Structural and biochemical studies of protein poly(ADP-ribosyl)ation

Zhizhi Wang

A dissertation

submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

University of Washington

2012

Reading Committee:

Wenqing Xu, Chair

Ronald E. Stenkamp

Ning Zheng

Liguo Wang

Program Authorized to Offer Degree:

Biological Structure

University of Washington

Abstract

Structural and biochemical studies of protein poly(ADP-ribosyl)ation

Zhizhi Wang

Chair of the Supervisory Committee:

Professor Wenqing Xu

Department of Biological Structure

Protein poly(ADP-ribosyl)ation (PARylation) has been found to be involved in various cellular processes, such as DNA damage recognition and repair, chromatin structure dynamics, gene transcription and poly(ADP-ribose)-dependent cell death. Most recently, PARylation has been shown to control the polyubiquitination and degradation of Axin, a key regulator of the Wnt signaling pathway. RNF146, which contains a WWE domain and a RING domain, is the only proven E3 ubiquitin ligase to date that requires PARylation of the substrate for subsequent polyubiquitination. The RNF146 WWE domain has been shown to bind poly(ADP-ribose) (PAR). In this thesis, first I review the current knowledge on protein PARylation. Then I describe my doctoral research on the structural and biochemical studies on how WWE domain recognizes PAR polymer. My studies reveal the structural basis of the RNF146 WWE

domain/*iso*-ADP-ribose (*iso*-ADPR, the internal unit of PAR) interaction and, for the first time, define the PAR/*iso*-ADPR binding as a *bona fide* function of the WWE domain family. This suggests that PAR may be a signal for protein ubiquitination and this signal is decoded by WWE domain-containing E3 ubiquitin ligases. In Chapter 3, I describe my structural studies on poly(ADP-ribose) glycohydrolase (PARG), which is the principal enzyme responsible for the degradation of PAR polymers *in vivo*. By solving the structures of the mouse PARG catalytic domain and its complex with its inhibitor ADP-HPD, I reveal how the PAR polymer as a substrate is recognized by PARG. The structures also suggest how the N terminal flexible peptide preceding the PARG catalytic domain regulates the enzymatic activity of PARG. This study helps the understanding of PARG catalytic and regulatory mechanisms as well as the rational design of PARG inhibitors.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

I thank my advisor Dr. Wenqing Xu for his guidance and support, which influence every part of my graduate study. I thank Dr. Ronald Stenkamp, Dr. Ning Zheng, and Dr. Liguo Wang for the great help as the reading committee for this dissertation. I also thank Dr. Ning Zheng, Dr. Ronald Stenkamp, Dr. John Clark, and Dr. Liguo Wang for their precious advice as the committee members. I also appreciate the support from the Department of Biological Structure and Biological Physics, Structure and Design program. Also thanks to my collaborators, Dr. Gregory Michaud, Dr. Zhihong Cheng, Dr. Yue Zhang, Dr. Thomas R. Hinds, Dr. Erkang Fan and Dr. Feng Cong. This dissertation would never have been completed without the encouragement and devotion of my friends and my family, especially my wife.

DEDICATION

TO MY FAMILY

Chapter 1 Introduction to poly(ADP-ribosyl)ation

Protein functions and localizations inside the cell are usually regulated by post-translational modifications (PTMs). PTMs are covalent modifications to protein backbones and to amino acid side chains. Poly(ADP-ribosyl)ation (PARylation) is a reversible PTM that occurs mainly when cells are under stress. It is catalyzed by a family of poly(ADP-ribose) polymerases (PARPs), which modify the target protein side chains by transferring ADP-ribose (ADPR) units from nicotinamide adenine dinucleotide (NAD^+) molecules (Figure 1.1). PARylation is involved in various cellular processes, such as DNA damage recognition and repair, chromatin structure dynamics, gene transcription, cell apoptosis, protein degradation and post-transcriptional gene regulation (De Vos, Schreiber, & Dantzer, 2012; Hottiger, Hassa, Luescher, Schueler, & Koch-Nolte, 2010; Krishnakumar & Kraus, 2010; Luo & Kraus, 2012; Satoh & Lindahl, 1992).

1.1 The PARP family

PARPs have been identified in all eukaryotes except yeast (Hassa, Haenni, Elser, & Hottiger, 2006; Lautier, Lagueux, Thibodeau, Menard, & Poirier, 1993; Otto et al., 2005), and lately in some bacteria (Slade et al., 2011). Based on the homology to the PARP catalytic domain of PARP1, seventeen PARPs have been identified in humans (Schreiber, Dantzer, Ame, & de Murcia, 2006). Except for the catalytic domains, the PARPs have variable domain organizations, which determine their subcellular localizations, different biological activities and also different PARylation substrates. Six of them, PARP1-4, PARP5A (Tankyrase 1) and PARP5B (Tankyrase 2), have poly(ADP-ribose) transferase activity. They all possess the conserved Glu residue in the catalytic triad H-Y-E (Gibson & Kraus, 2012). Because of the lack of Glu in the catalytic triad,

the other members (PARP6-PARP16) are either enzymatically inactive, or only having mono(ADP-ribose) transferase activity (summarized in Table 1.1). Recently, a new nomenclature has been suggested for PARPs as diphtheria toxin-like ADP-ribosyl transferases (ARTDs) (Hottiger, et al., 2010) (Table 1.1).

1.2 The chemistry of PARylation

PARPs catalyze the transformation of NAD^+ into nicotinamide and poly(ADP-ribose) (PAR), which is negatively charged and mostly added to Glu, Asp and Lys residues of the target proteins. The ADP-ribose units in PAR are linked to each other by ribose-ribose glycosidic bonds, resulting in either linear or branched PAR polymers. There are at least two different steps for PARPs to synthesize PAR on the target residue. First, in the initiation step, PARP hydrolyses NAD^+ and transfers ADPR to the acceptor residues. Second, in the elongation step, PARP transfers additional ADPR units to the first ADPR unit subsequently through α(1''→2') ribose-O-ribose glycosidic bonds and forms linear PAR. Some PARPs like PARP1, can also transfer the ADPR unit through α(1'''→2'') glycosidic bonds and form branched PAR (Figure 1.1). PAR can reach over 200 units in length with new branches every 20-50 linear ADPR units (Kiehlbauch, Aboulela, Jacobson, Ringer, & Jacobson, 1993).

1.3 PARylation and DNA damage detection and repair

PARylation was firstly and extensively studied in the DNA damage response. PARylation by PARP1, the founding member of the PARP family, is one of the earliest responses to DNA

lesions, including single-strand breaks (SSBs) and double-strand breaks (DSBs) (Bouchard, Rouleau, & Poirier, 2003; Woodhouse & Dianov, 2008). The DNA-binding domain of PARP1 can recognize the SSBs and DSBs in a sequence-independent manner. Subsequently, PARP1 is catalytically activated and starts to hydrolyse NAD^+ and transfer ADPR units to target proteins. PARylation by PARP1 is mainly on itself (auto-PARylation), and also on histones and the DNA repair scaffold protein XRCC1. These PARylations promote the PAR-dependent recruitment of chromatin remodeling proteins, including CHD4, APLF, CHFR, macroH2A and ALC1. This provides a more favorable state of chromatin for DNA repair machinery to target the DNA damage site. Furthermore, the DNA repair enzymes, including DNA ligase III and ATM, are also recruited in a PAR-dependent manner (D. Ahel et al., 2009; I. Ahel et al., 2008; Chou et al., 2010; Li et al., 2010; Masson et al., 1998; Okano, Lan, Caldecott, Mori, & Yasui, 2003). PAR degradation is catalyzed by poly(ADP-ribose) glycohydrolase (PARG). PARG is also recruited to DNA damage site in a PAR-dependent manner, to rapidly degrade PAR polymer. PAR turnover concerted by PARP1 and PARG is crucial for dynamic recruitment and release of DNA repair machinery (Mortusewicz, Fouquerel, Ame, Leonhardt, & Schreiber, 2011).

1.4 PARylation and cell death

Under mild genotoxic stress, the activation of PARP1 by DNA damage will promote DNA repair and survival of the cell. However, in some pathological conditions, severe DNA damage hyper-activates PARP1, which subsequently promotes cell death by two different mechanisms in which the catalytic activity of PARG contributes. First, the rapid synthesis of PAR by PARP1 from NAD^+ and degradation of PAR by PARG to ADPR results in the depletion of the NAD^+ pool,

and ultimately ATP. This cellular energy depletion induced necrosis is called parthanatos (Andrabi, Dawson, & Dawson, 2008). Second, the endo-glycohydrolase activity of PARG degrades PAR polymer to oligo-ADPRs. They accumulate in the nucleus and translocate to the mitochondria, and further induce the release of AIF (apoptosis-inducing factor) from mitochondria to promote AIF-mediated apoptosis (Wang, Yang, Chai, Shi, & Xue, 2002; Yu et al., 2002).

1.5 PARylation and circadian clock

PARP1 activity (auto-PARylation) in liver oscillates with circadian rhythm (Asher et al., 2010). Furthermore, PARP1 interacts with transcription factors CLOCK and BMAL1 (brain and muscle aryl hydrocarbon receptor nuclear translocator [ARNT]-like) hetero-dimer, and PARylates CLOCK after nutrient intake. Since PAR polymers are bulky and negatively charged, PARylated CLOCK repulses the DNA elements that CLOCK binds. The reduced binding subsequently changes the transcription of its downstream genes. *Parp1*-knockout mice exhibit abnormal circadian rhythms in response to perturbed feeding times (Asher et al., 2010; Kumar & Takahashi, 2010).

1.6 PARylation and RNA interference

In cancers and neurodegenerative diseases, cells are under stress and form cytoplasmic stress granules, which are macrostructures comprised of aggregated RNAs and RNA-binding proteins. Recent studies showed that not only the stalled translation pre-initiation complexes are

aggregated in stress granules, but also microRNAs, microRNA-binding Argonaute family members, six members of the PARP family and two isoforms of PARG are also components of stress granules (Leung & Sharp, 2010; Leung, Todorova, Ando, & Chang, 2012; Leung et al., 2011). MicroRNA-binding Argonautes are PARylated, and together with microRNAs, form aggregates with PAR-binding domain-containing PARPs through PAR polymer as a part of the stress granules. Upon the forming of stress granules, microRNA-directed mRNA cleavage as well as microRNA mediated translational repression are inhibited. Overexpression of the PARPs and knockdown of PARG both can promote the formation of stress granules and further relief of mRNA cleavage and translation repression (Leung & Sharp, 2010; Leung, Todorova, Ando, & Chang, 2012; Leung et al., 2011).

1.7 PARylation and ubiquitination-dependent degradation

Programmed protein degradation is crucial for cell homeostasis. The ubiquitin-proteasome pathway is the most important mechanism for controlling protein turnover. In the past, protein phosphorylation was the most studied posttranslational modification to mark a protein for ubiquitination-dependent degradation (Bergink & Jentsch, 2009; Willems et al., 1999). In several cases, protein PARylation has also been observed to be associated with protein degradation (Aravind, 2001). However it was unclear whether PARylation is a specific degradation signal for some proteins. Recently, the first direct evidence for PARylation-dependent ubiquitination was demonstrated (Huang et al., 2009). It showed that Axin, a scaffold protein that regulates β-catenin degradation in Wnt signaling pathway, was ubiquitinated and degraded in a PARylation-

dependent manner. Inhibition of tankyrases, the PARPs that PARylate Axin, can completely abolish Axin ubiquitination and degradation (Huang, et al., 2009).

1.8 Recognition of PAR polymer

As mentioned above, PARylation plays crucial roles in many cellular processes. Most of them require interactions between PARylated proteins and other proteins in a PAR-dependent manner. Recent studies have shown that four different types of protein modules are involved in PAR recognition. They are PAR-binding motifs (PBMs), macrodomains, PAR-binding zinc-finger (PBZ) and WWE domains.

1.8.1 PAR-binding motifs (PBMs)

PBMs are the first discovered PAR-binding modules (Pleschke, Kleczkowska, Strohm, & Althaus, 2000). PBMs have eight amino acids with a consensus sequence $[HKR]_1$-X_2-X_3-$[AIQVY]_4$-$[KR]_5$-$[KR]_6$-$[AILV]_7$-$[FILPV]_8$ (Gagne et al., 2008; Pleschke, et al., 2000). PBMs are present in a large number of proteins, which are usually involved in DNA repair, chromatin remodeling and RNA metabolism. In the case of DNA repair, XRCC1, which is a PBM-containing protein, is rapidly recruited to the DNA damage sites after the activation of PARP1 in a PAR-binding manner. There is another example, AIF, which also contains a PBM. AIF is released from mitochondria after the recognition of PAR polymer in mitochondria, and induces cell apoptosis. In both cases, 3D structures of the PBM-containing proteins have been solved (Y. Wang et al., 2011). However, there is still no structural information about the interaction

between PAR polymer and PBM. The main reason is that PBMs interact with PAR polymer only, not ADPR or other derivatives. Meanwhile, it's still a big challenge to synthesize large amounts of homogenous PAR polymer of a specific size for crystallization with PBM-containing proteins. PBM has multiple positively charged residues (K and R) and also several hydrophobic residues. This feature probably enables PBM to recognize phosphate groups as well as the adenine rings on PAR (Figure 1.2).

1.8.2 Macrodomains

The evolutionally ancient macrodomains are highly conserved and have 130-190 amino acids. They are found in histone variant macroH2A chromatin remodeling proteins as well as some PARP family members. Macrodomains interact with ADPR, and the binding affinity is around micromolar to sub-micromolar (Karras et al., 2005; Kustatscher, Hothorn, Pugieux, Scheffzek, & Ladurner, 2005). There are several X-ray crystal structures of macrodomains from different proteins extensively interacting with ADPR (Figure 1.2). These structures show very high structure homology that all of them containing mixed α/β folds with a groove for ADPR binding. The 2'-OH of the distal ribose (adenine-bonded ribose) is blocked, and this only allows the terminal ADPR unit of PAR, instead of the internal ADPR units to be recognized by macrodomains.

After PARP1 is activated, macroH2A is found to interact with PAR at the DNA damage site during DNA repair. ALC1, macrodomain-containing chromatin remodeling protein, also rapidly binds to PAR in the DNA damage site and causes the chromatin structure to be more accessible

to DNA repair machinery. In the stress granules, macrodomain-containing PARPs (PARP14 and PARP15) aggregate with PARylated Argonautes through PAR.

Surprisingly, the latest solved structure of a bacterial PARG showed that the catalytic domain is macrodomain-like. Similarly, it also binds to ADPR tightly with the 2'-OH of distal ribose buried in the binding pocket. The bacterial PARG was also shown to have only exo-glycohydrolase activity (Slade, et al., 2011).

1.8.3 PAR binding zinc-finger (PBZ)

PBZ has a C2H2 zinc finger, with the consensus sequence [K/R]xxCx[F/Y]GxxCxbbxxxxHxxx[F/Y]xH. It was firstly identified in CHFR, which is a mitotic checkpoint protein (Scolnick & Halazonetis, 2000). Other proteins with PBZs, like APLF and SNM1, are all involved in DNA repair (I. Ahel, et al., 2008). There are several structural studies of PBZ domains of CHFR and APLF by NMR (nuclear magnetic resonance) or by X-ray crystallography (Eustermann et al., 2010; Isogai et al., 2010; Li, et al., 2010; Oberoi et al., 2010). The complex structures of APLF PBZs with RFA (2'-O-α-D-ribofuranosyladenosine) (Figure 1.2), which contains the characteristic α(1"→2') ribose-ribose glycosidic bond, showed that PBZ can recognize the internal structure of PAR polymer (Eustermann, et al., 2010). The tandem PBZ motifs found in APLF have much higher affinity with PAR polymer than individual PBZ motifs. The X-ray structure of CHFR with ADPR derivatives showed that the PBZ motif can recognize two adenine rings simultaneously from two ADPR units in the PAR polymer (Oberoi, et al., 2010).

During the DNA damage response, PARP1 recognizes the DNA damage site and is subsequently activated. Both CHFR and APLF are rapidly recruited to the damage site by the recognition of PAR polymer with PBZ domains.

1.8.4 WWE domains

WWE domains are named after their most conserved three residues (Trp, Trp and Glu). Based on *in silico* predictions, WWE domains are found either coupling with PARPs or ubiquitin E3 ligases (Aravind, 2001). Recent studies showed that RNF146, which is a WWE domain-containing ubiquitin E3 ligase, is responsible for the polyubiquitination of PARylated Axin and other substrates (Callow et al., 2011; Kang et al., 2011; Zhang et al., 2011). The WWE domain of RNF146 recognizes PAR polymer, and a single site mutation in the WWE domain (R163A) can abolish the binding of PAR and inhibit the turnover of Axin in the canonical Wnt signaling pathway (Zhang, et al., 2011).

However, there was no structural information about how the WWE domain recognizes PAR polymer. It was unknown whether it's a common function of all WWE domains, and whether PARylation-dependent ubiquitination is a general mechanism. My studies described here in Chapter 2 address these questions.

1.9 Degradation of PAR polymer

1.9.1 PARG

Like PARPs, PARG is present in all eukaryotic cell types, except yeast (Lautier, et al., 1993). A recent study showed that PARG is also present in several species of bacteria (Slade, et al., 2011). There are seventeen identified PARPs in humans that share homology to the PARP1 catalytic domain. However, there is only one known poly(ADP-ribose) glycohydrolase (PARG) which is responsible for most PAR degradation in humans (Figure 1.1). The *PARG* gene encodes for three different isoforms of PARG localizing in different cellular compartments. The 111 kDa full length PARG ($hPARG_{111}$) localizes in the nucleus. Both 99 kDa $hPARG_{99}$ and 102 kDa $hPARG_{102}$ isoforms localize in the cytoplasm. PARG possesses both exo-glycosidase and endo-glycosidase activities and therefore is able to hydrolyse ribose-ribose glycosidic bonds between ADP-ribose units at the terminus or within the PAR polymers (Ikejima & Gill, 1988; Miwa, Tanaka, Matsushi.T, & Sugimura, 1974).

In unstimulated cells, the basal concentration of PAR is very low. Once cells are exposed to genotoxic stress, the PAR level can increase by up to 500-fold because of activation of PARP1. However, the PAR polymer formed following DNA damage can be degraded by PARG only a few minutes after its synthesis. The function of PARG prevents the accumulation of PARylated proteins in the nucleus and also keeps PARP1 active during the DNA damage response by removing PAR polymer which results from the auto-inhibitory PARylation of PARP1 itself.

While the N-terminal region is absent in some PARG splicing forms and predicted to be disordered in mice, the conserved C-terminal 60kD catalytic domain is fully active (Meyer, Meyer-Ficca, Whatcott, Jacobson, & Jacobson, 2007). Despite the important biological functions, there was no available three-dimensional structure for any part of mammalian PARG. A recent study revealed the structure of bacterium *T. curvata* PARG, which is only about 30kDa (half the size of the catalytic domain of mammalian PARG). The core of *T. curvata* PARG has a

macrodomain fold, and can only bind to the terminal ADPR unit of PAR polymer, suggesting that it has exo-glycohydrolase activity, not endo-glycohydrolase activity (Slade, et al., 2011). The observed GGG-X_{6-8}-QEE catalytic loop, which is also present in mammalian PARGs (Patel, Koh, Jacobson, & Oliveira, 2005), catalyzes the hydrolysis of ADPR units from the distal end of the PAR polymer.

The mammalian PARG catalytic domain has a more complex structure, and both exo- and endo-glycohydrolase activities. The structure of mammalian PARG would shed light on the catalysis mechanism as well as the regulatory mechanism of this enzyme. It would also help inhibitor design for PARG. The studies described here in Chapter 3 provide a first look at the structures of active mouse PARG catalytic domain itself and in complex with inhibitors.

1.9.2 ARH3

In addition to PARG, ARH3 (ADP-ribosyl hydrolase 3) can also degrade PAR polymer *in vitro* (Oka, Kato, & Moss, 2006). It has a totally distinct structure from PARG and is shown to hydrolyze O-acetyl-ADP-ribose (Ono, Kasamatsu, Oka, & Moss, 2006). A recent study showed that the mitochondrial isoform of PARG is inactive and proposed that ARH3 is responsible for the PAR degradation in the mitochondria matrix (Niere et al., 2012). The function of ARH3 *in vivo* is still poorly studied.

1.10 PARPs and PARG as Drug targets

More and more evidence shows the importance of homeostasis of PAR metabolism concerted by PARPs and PARG. PARPs are involved in many important cellular processes, including DNA damage response and stress response. PARP inhibitors are potential drugs for cancer treatment as well as stress related disease including cardiovascular disease and stroke (Telli, 2011; Underhill, Toulmonde, & Bonnefoi, 2011). So far, there are many PARP inhibitors in clinical trial pipelines, most of which are for cancers.

Over decades, much effort has been made to delineate the biological functions of PARG and to explore the therapeutic potentials of PARG inhibition in pathophysiological conditions such as inflammation, ischemia, and stroke (Cuzzocrea & Wang, 2005; Davidovic, Vodenicharov, Affar, & Poirier, 2001; Koh, Dawson, & Dawson, 2005; W. Min & Wang, 2009). PARG has also been proposed as a cancer drug target because PARG deficiency enhances cytotoxic sensitivity induced by chemotherapy agents (Koh et al., 2004).

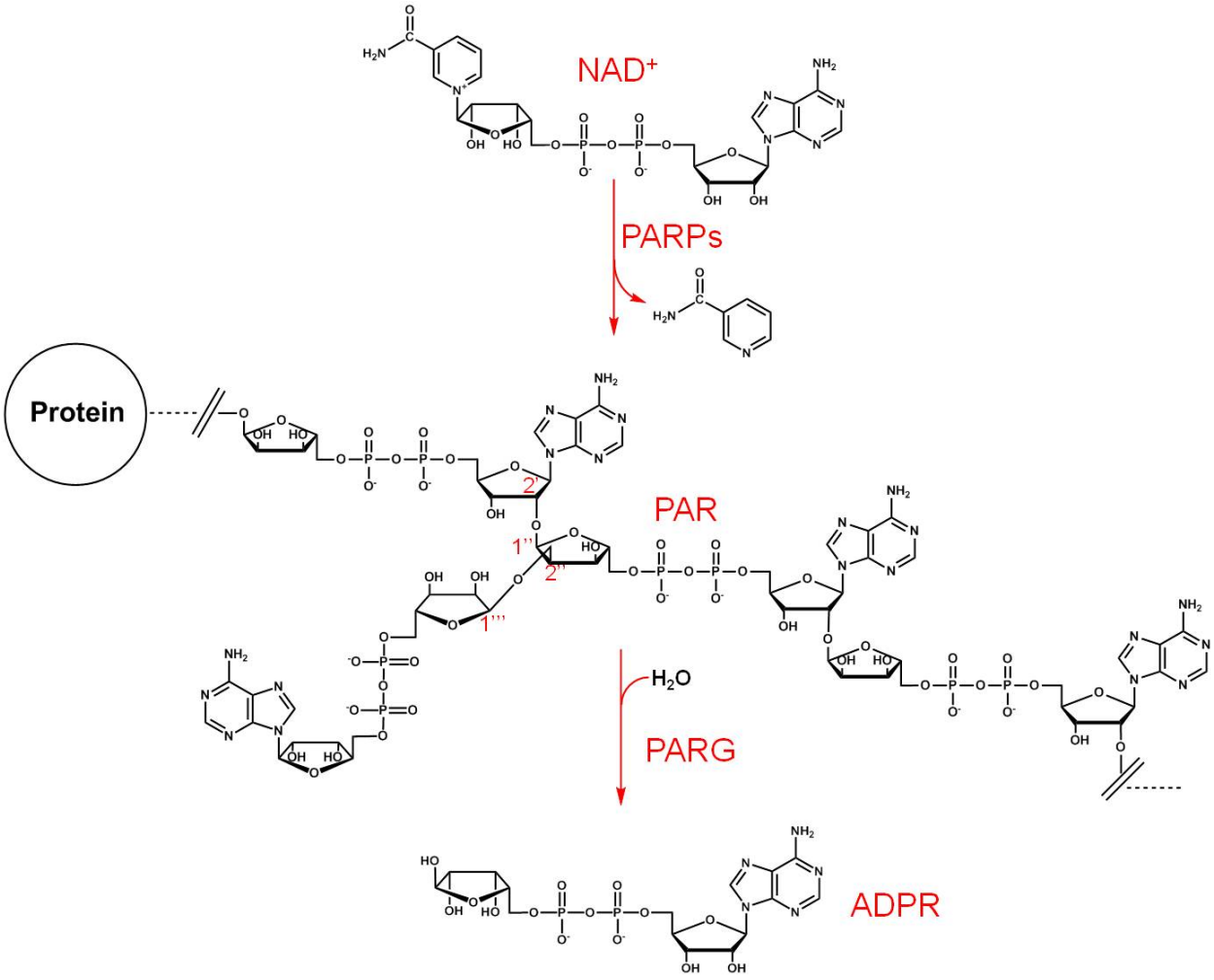

Figure 1.1 Poly(ADP-ribose) (PAR) metabolism by PARPs and PARG

The characteristic linkages of linear PAR [α(1''→2') ribose-O-ribose glycosidic bond] and branched PAR [α(1'''→2'') ribose-O-ribose glycosidic bond] are indicated.

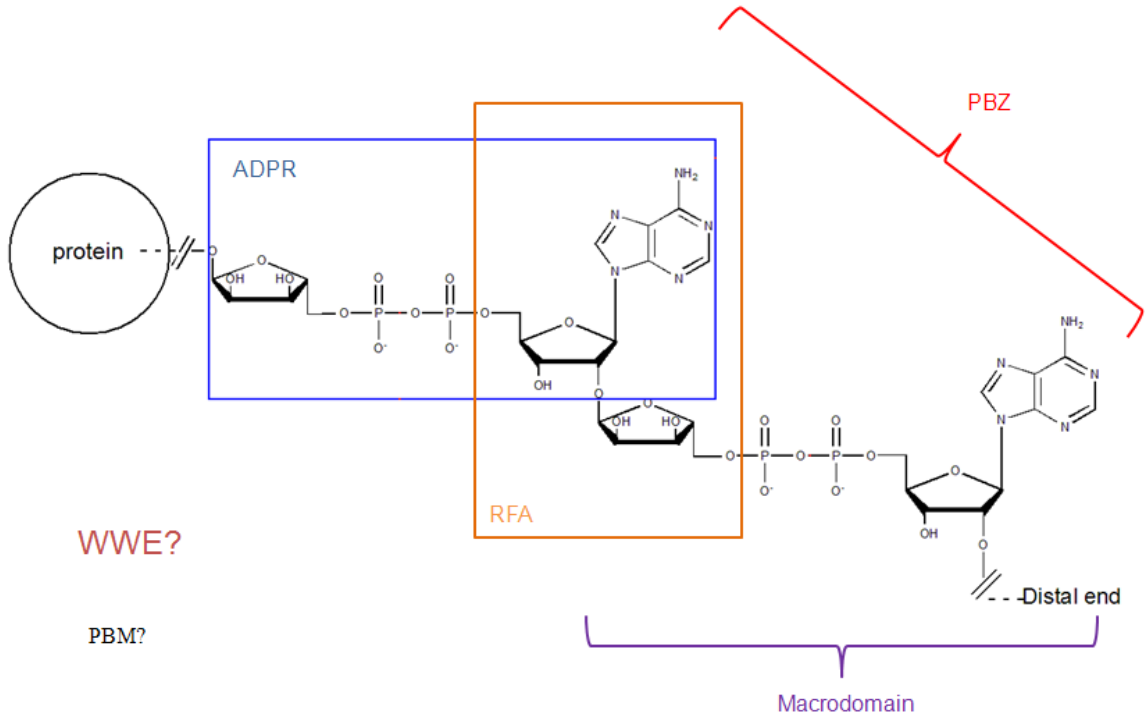

Figure 1.2 Structure of PAR and recognition of PAR by PAR-binding modules.

The brackets illustrate the region recognized by each PAR-binding module. The ADPR and its derivative RFA used in structural characterization of PAR-binding modules are framed. How the PBM recognizes PAR remains unknown; PBZ binds to RFA and can also recognize adenine rings on two successive ADPR units; Macrodomain only binds the ADPR unit on the terminus; How the WWE domain binds to PAR will be revealed by my studies described in Chapter 2.

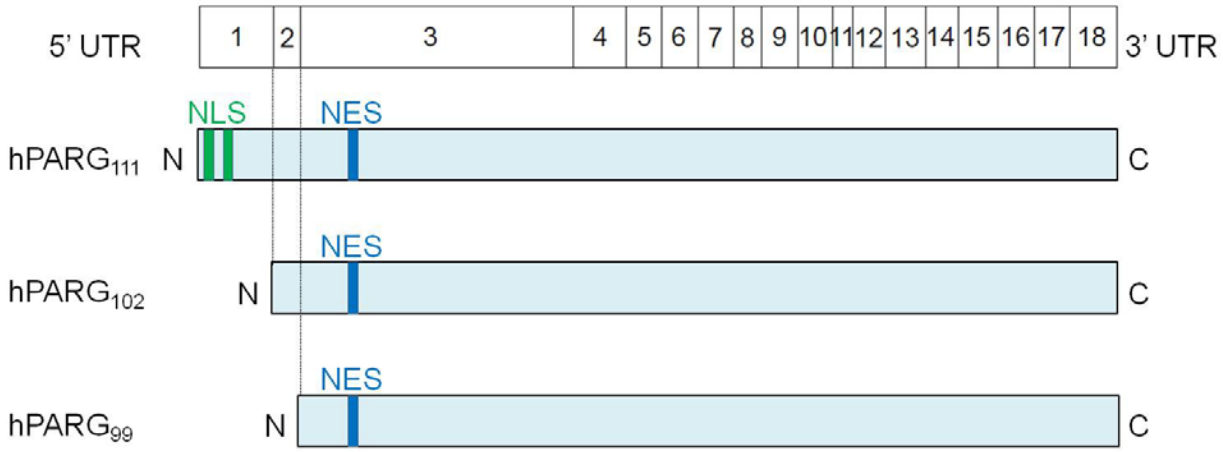

Figure 1.3 Domain structures of four human PARG isoforms.

NLS, nuclear localization signals; NES, nuclear export signals.

Table 1.1 PARP family members—enzymatic activities and functional domains

Member	Alternative name	Transferase name	Subcellular localization	Triad motif	Enzymatic activity	Functional domains
PARP1		ARTD1	nuclear	H-Y-E	P	Zinc-fingers, WGR, BRCT
PARP2		ARTD2	nuclear	H-Y-E	P	WGR
PARP3		ARTD3	nuclear	H-Y-E	P	WGR
PARP4	vPARP	ARTD4	cytosolic	H-Y-E	P	BRCT
PARP5A	Tankyrase 1	ARTD5	Nuclear cytosolic	H-Y-E	P	Ankyrin repeat, SAM
PARP5B	Tankyrase 2	ARTD6	Nuclear cytosolic	H-Y-E	P	Ankyrin repeat, SAM
PARP6		ARTD17	ND	H-Y-Y	M	
PARP7		ARTD14	ND	H-Y-I	M	Zinc-fingers, WWE
PARP8		ARTD16	ND	H-Y-I	M	
PARP9	BAL1	ARTD9	ND	Q-Y-T	M	Macrodomain
PARP10		ARTD10	Nuclear cytosolic	H-Y-I	M	
PARP11		ARTD11	ND	H-Y-I	M	WWE
PARP12		ARTD12	cytosolic	H-Y-I	M	Zinc-fingers
PARP13	ZAP1	ARTD13	cytosolic	H-Y-V	M	Zinc-fingers
PARP14	BAL2	ARTD8	cytosolic	H-Y-L	M	Macrodomian, WWE
PARP15	BAL3	ARTD7	cytosolic	H-Y-L	M	Macrodomian
PARP16		ARTD15	ND	H-Y-I	M	

ARTD, ADP-ribosyl transferase; BAL, B-aggressive lymphoma protein; ND, not determined; vPARP, vault PARP; ZAP1, zinc-finger antiviral protein 1. Known or predicted enzymatic activity: mono- (M), or poly(ADP-ribosyl)ation (P). [Adapted from table 1 of (Gibson & Kraus, 2012)]

Chapter 2 PAR signal recognition by WWE domain

The material of this chapter has been published in:

Wang Z, Michaud GA, Cheng Z, Zhang Y, Hinds TR, Fan E, Cong F, Xu W.

(2012) *Genes & Development* Feb 1; 26(3):235-40.

Recognition of the *iso*-ADP-ribose moiety in poly(ADP-ribose) by WWE domains suggests a general mechanism for poly(ADP-ribosyl)ation-dependent ubiquitination.

2.1 Introduction

Protein ubiquitination regulates diverse biological processes; however, the mechanism by which proteins are earmarked for ubiquitination remains incompletely understood. Other than phosphorylation, which is a general mechanism for many cases, hydroxylation of a substrate (i.e., HIF1-a) and the binding of small molecules (e.g., the plant hormone auxin) to E3 ligases have been shown to control protein ubiquitination in sporadic cases (Bergink & Jentsch, 2009; J. H. Min et al., 2002; Tan & Zheng, 2009; Willems, et al., 1999). Protein poly(ADP-ribosyl)ation (PARylation), catalyzed by PAR polymerases (PARPs), also regulates a myriad of biological processes, including DNA damage responses, transcriptional regulation, intracellular trafficking, energy metabolism, circadian rhythm, and cell survival and cell death programs, among others (Curtin, 2005; Jagtap & Szabo, 2005; Kim, Zhang, & Kraus, 2005; Krishnakumar & Kraus, 2010; Schreiber, et al., 2006). How PARylation affects so many biological functions remains largely mysterious.

In many cases, such as PARylation of histones in transcriptional regulation, PARylation is considered to control activities of the substrate proteins via the negative charges in the PAR polymer (Krishnakumar & Kraus, 2010; Schreiber, et al., 2006). In other cases, PAR polymers

have been implicated as signaling molecules that can induce cell death, especially in the brain (Andrabi et al., 2011; Andrabi et al., 2006). Most recently, PARylation has been shown to control the polyubiquitination and degradation of Axin, a key regulator of the Wnt signaling pathway (Callow, et al., 2011; Huang, et al., 2009; Kang, et al., 2011; Zhang, et al., 2011). In all of these reports, RNF146 (aka Iduna), which contains a WWE domain and a RING domain (Figure 2.1), is the only known E3 ubiquitin ligase to date that requires PARylation of the substrate for subsequent polyubiquitination (Callow, et al., 2011; Kang, et al., 2011; Zhang, et al., 2011). The RNF146WWEdomain has been shown to bind PAR (Callow, et al., 2011; Zhang, et al., 2011), and it was reported that a short PAR-binding motif (PBM) within the domain retains this binding activity (Andrabi, et al., 2011). The PBM was originally found in histones and several other proteins (Gagne, et al., 2008). However, the PBM identified in RNF146 is not conserved in other WWE domains, so it remains unclear whether the WWE domain represents a novel PAR-binding domain. Here we reveal the structural basis of the RNF146 WWE domain/*iso*-ADPR interaction and, for the first time, define PAR/*iso*-ADPR binding as a *bona fide* function of the WWE domain family. Importantly, the structural coupling of WWE domains and E3 ligase domains in many WWE domain-containing proteins suggests a functional coupling of protein PARylation and ubiquitination.

2.2 Results and discussion

2.2.1 The RNF146 WWE domain recognizes *iso*-ADPR, but not ADPR

I sought to clarify the requirement of the entire RNF146WWE domain structure for PAR binding through structural analysis. PAR polymers display high chemical heterogeneity (in both lengths

and branching patterns) and are not suitable for quantitative and structural analysis. Thus, I first examined its interaction with ADPR, the building unit added to PAR during PAR synthesis (Figure 2.2). Isothermal titration calorimetry (ITC) analysis demonstrated that the RNF146 WWE domain does not interact with ADPR, even at high concentrations (>0.1 mM) (Figure 2.3). Then I turned to *iso*-ADPR, which is the smallest PAR structural unit containing the ribose-ribose glycosidic bond unique to PAR, formed during PAR synthesis by PARPs (Figure 2.2). It remains a major challenge to obtain PAR of a specific length in sufficient quantities for biochemical analysis, and *iso*-ADPR is not commercially available, nor has it been used previously in structural and biochemical studies. I therefore developed a protocol to biosynthetically generate *iso*-ADPR (Figure 2.4). I synthesized PAR using histone PARylation by PARP1, following previously published protocols (Fahrer, Kranaster, Altmeyer, Marx, & Buerkle, 2007; Kiehlbauch, et al., 1993). After removing small molecules by size exclusion chromatography (SEC), I digested PAR polymers with a phosphodiesterase to generate *iso*-ADPR. Finally, I purified *iso*-ADPR by ion exchange and a second SEC step to remove all remaining large molecules. The purity and identity of purified *iso*-ADPR were confirmed by reverse-phase high-performance liquid chromatography (RP-HPLC) and mass spectrometry. ITC analysis demonstrated that, in contrast to the poor interaction with ADPR, the RNF146 WWE domain interacted with *iso*-ADPR avidly, with a dissociation constant of 0.37 μM (Figure 2.5). Selectivity of the RNF146 WWE domain for *iso*-ADPR was also demonstrated by its co-migration with *iso*-ADPR, but not ADPR, in SEC (Figure 2.6). Thus, the RNF146 WWE domain interacts with the PAR polymer, but not mono-ADPR. This is likely to be important for RNF146 function, since ADP-ribosylation and PARylation are performed by different enzymes and have distinct biological functions (Corda & Di Girolamo, 2003; Curtin, 2005; Jagtap & Szabo, 2005;

Krishnakumar & Kraus, 2010; Okazaki & Moss, 1996; Schreiber, et al., 2006). The ability to generate purified *iso*-ADPR provides a unique reagent for biochemical and structural analysis of PAR-binding proteins and mechanistic analysis of PAR-metabolizing enzymes, such as PARG, that cleave the glycosidic bond.

2.2.2 Crystal structure of the RNF146 WWE domain in complex with *iso*-ADPR

To understand how the RNF146 WWE domain interacts with *iso*-ADPR, I determined the crystal structure of the RNF146WWE domain in complex with *iso*-ADPR at 1.63 Å resolution (Figures 2.7, 2.8; Table 2.1). The WWE domain contains six β strands, forming half of an α/β barrel, with the other side of this half β barrel covered by an α helix. The high resolution of our structure allowed me to define unambiguously that the ribose-ribose linkage in *iso*-ADPR is an α(1→2) glycosidic bond (Figure 2.8). This confirmed that PAR synthesis catalyzed by PARP1 is α(1→2)-specific. The adenine ring of *iso*-ADPR inserts into the pocket formed by the half β barrel and α helix (Figure 2.7). The two ribose-phosphate moieties on both sides of *iso*-ADPR sit on the edge of the half β barrel, which is highly positively charged (Figure 2.9). Both sides of *iso*-ADPR, especially the two separated phosphate groups, are involved in extensive interactions with WWE domain residues (Figure 2.10), providing an explanation for why the RNF146 WWE domain specifically binds to *iso*-ADPR and thus to PAR, but not ADPR, which has the two phosphate groups on the same side (Figure 2.2). Compared with a previously determined NMR structure of an unliganded RNF146 WWE domain (Protein Data Bank [PDB] code 1UJR), it appears that *iso*-ADPR binding induces significant conformational changes in the WWE domain, particularly in the C-terminal tail region (residues 169-183), which folds back to support the

distal ribose–phosphate groups of *iso*-ADPR (Figure 2.11). The mode of binding displayed in my crystal structure, in which the two phosphate groups lie on an open surface of the RNF146 WWE domain, should allow internal *iso*-ADPR units in PAR polymers to bind the WWE domain in the same manner.

2.2.3 Mutagenesis analysis of the RNF146 WWE domain

Seven RNF146 WWE domain surface residues are involved in *iso*-ADPR binding (Figure 2.10). Among them, the Tyr 107 phenol group stacks on the side of the *iso*-ADPR adenine ring within the pocket, and Gln 153 near the bottom of the adenine-binding pocket forms two hydrogen bonds with the adenine ring and appears to confer binding specificity as well as binding affinity. The Arg 163 and Tyr 144 side chain groups interact with the proximal phosphate group, and Trp 114, Arg 110, and Lys 175 interact with the ribose-phosphate groups in the distal side of *iso*-ADPR (Figure 2.10). To validate our structure and define key interactions between the RNF146 WWE domain and *iso*-ADPR, I analyzed the *iso*-ADPR binding of seven RNF146 WWE domain mutants by ITC analysis. Mutants Y107A, Y144A, and R163A lack a detectable interaction, and mutants W114A and Q153A have much lower affinity than the wild-type RNF146 WWE domain. The mutations R110A and K175A had only a minor effect on the interaction (Figure 2.12; summarized in Table 2.2). Then my collaborators tested the binding of RNF146 mutants to PAR in the context of a full-length protein. Consistent with the ITC analysis using the WWE domain, the RNF146 mutants Y107A, Y144A, R163A, and Q153A all lost the ability to interact with PAR in a coimmunoprecipitation assay (data not shown).

To demonstrate the importance of the RNF146 WWE domain residues involved in PAR binding, my collaborators examined the effect of their mutation in full-length RNF146 protein on Axin turnover in a cellular context. Expression of siRNA resistant RNF146 completely rescued the effect of RNF146 siRNA on Axin1, whereas expression of RNF146 mutants failed to do so (data not shown). These experiments confirm that the specific interactions observed in the crystal structure are important for the *in vivo* functions of RNF146. The previously proposed PBM in RNF146 (residues 144-167) (Andrabi, et al., 2011) only accounts for three β strands in the WWE domain structure (Figure 2.7). Residues outside the PBM region interact extensively with *iso*-ADPR (Figure 2.10), and missense mutations of the RNF146 WWE domain outside the PBM, such as Y107A, abolish the PAR-binding activity (Figure 2.12, Table 2.2). Therefore, we conclude that the RNF146 WWE domain is a *bona fide* PAR-binding domain, which specifically recognizes the *iso*-ADPR moiety of PAR.

2.2.4 PAR/*iso*-ADPR binding is a common function of WWE domains

The WWE domain family exhibits a low degree of sequence homology (Figure 2.13). The most conserved residues include the two Trp and one Glu residues that give rise to the name ‘‘WWE’’ domain. These residues are involved in stabilizing the WWE domain fold (Figure 2.14). Based on the crystal structure of the RNF146 WWE domain/*iso*- ADPR complex and our mutagenesis analysis, we note that critical residues for *iso*-ADPR and PAR binding are conserved in most WWE domains, suggesting that PAR binding may be a common function of the WWE domain family (Figure 2.13). Surface plasma resonance (SPR) analysis done by my collaborators demonstrated that GST-tagged WWE domains from HUWE1, ULF, Deltex1, and PARP11 bind

to the PAR polymers, whereas the DDHD2 WWE domain did not interact with PAR (data not shown). This is in perfect concordance with our structure and sequence alignment. While the four crucial residues involved in *iso*-ADPR binding in the RNF146 WWE domain are conserved among most WWE domains in the human genome, including these from RNF146, HUWE1, ULF, Deltex1, Deltex2, Deltex4, and PARP11, residues corresponding to RNF146 Q153 and R163 are not conserved in DDHD2 and PARP14 (Figure 2.13).

Given the sequence conservation, we predict that other PAR-binding WWE domains also recognize *iso*-ADPR in the same way as the RNF146 WWE domain. While many of the WWE domains we tested were not suitable for ITC analysis due to aggregation, our ITC analysis with the HUWE1 WWE domain showed that this domain interacts with *iso*-ADPR with a K_d of 13 μM (Figure 2.15). The interactions between WWE domains and PAR are further enhanced by the high local concentrations of *iso*-ADPR around PAR, as there can be as many as 200 units in each PAR chain. Deltex1/2/4 proteins contain two tandem WWE domains. Structural modeling based on the unliganded *Drosophila* Deltex WWE domain structure (Zweifel, Leahy, & Barrick, 2005) also suggests that these two WWE domains may recognize two neighboring *iso*-ADPR units, with each interacting in a manner similar to the RNF146 WWE domain (Figure 2.16). Therefore, with the exception of DDHD2 and PARP14, PAR binding through the recognition of *iso*-ADPR is a common function for most WWE domains.

2.2.5 Structural and functional coupling of protein PARylation and ubiquitination

It has been previously noticed that the majority of WWE domains are structurally coupled with E3 ubiquitin ligases (Figure 2.1) (Aravind, 2001). Here we show that the WWE domains from all

four E3 domain-containing proteins tested (RNF146, HUWE1, ULF, and Deltex1) interact specifically with PAR. HUWE1 (aka Mule, ARFBP1, LASU1, and HectH9) is a HECT-type E3 ligase critical for cell regulation; it ubiquitinates the tumor suppressor p53, core histones, and the Bcl-2 family member Mcl-1 (Bernassola, Karin, Ciechanover, & Melino, 2008; D. Chen, Brooks, & Gu, 2006). ULF/TRIP12 is an E3 ubiquitin ligase of ARF, a key activator of p53 (Delin Chen, Shan, Zhu, Qin, & Gu, 2010). The Deltex family, which plays an important role in Notch signaling, contains a RING domain and two tandem WWE domains (Katoh & Katoh, 2007; Zweifel, et al., 2005). The definition of WWE domains as a novel PAR-binding domain and the structural coupling of the WWE domain with E3 domains suggest that PAR may be a signal for protein ubiquitination — either polyubiquitination that may lead to protein degradation by proteasomes, or divergent monoubiquitination that controls protein activities in the cell. The WWE domains may therefore be a key link between protein PARylation and ubiquitination. Furthermore, since the small WWE domains recognize only one (or two) internal *iso*-ADPR unit, numerous WWE domain-containing proteins may potentially cluster around a PAR polymer, a property that may have functional importance. The role of PARPs and PARylation is well established in many biological processes, including DNA repair. The insights provided by the studies reported here will facilitate design of specific mutations in WWE domain-containing E3 ligases that can be used to unravel the role and the molecular mechanisms of PARylation in biological systems. In summary, protein PARylation may be another general mechanism to label proteins for ubiquitination other than protein phosphorylation, and many of the protein PARylation events may function through ubiquitination. Our studies have not only identified *iso*-ADPR as the critical signaling unit in PAR, but also suggested a role for the WWE domains in a superfamily of ubiquitin ligases in decoding the protein PARylation signal.

2.3 Materials and methods

2.3.1 Generation and purification of *iso*-ADPR

Based on the methods reported previously (Kiehlbauch, et al., 1993) (Fahrer, et al., 2007), PAR polymer was synthesized *in vitro* with some modifications. The histone PARylation reaction was set up in a 15-mL incubation mixture comprising 50 mM Tris-HCl (pH 8.0), 20 mM $MgCl_2$, 50 mM NaCl, 10mMDTT, 1.5mM β-NAD^+, 0.1 mg/mL histone H1 (Millipore), 0.1 mg/mL histone type IIA (Sigma-Aldrich), 67 mg/mL activator oligonucleotide GGAATTCC, and 5000 U of human PARP1 (Trevigen). The reaction was stopped after 1.5 h at room temperature by ice-cold trichloroacetic acid. The white pellet was washed with ice-cold ethanol and was further dissolved in buffer containing 0.2 mg/mL Protease K (Invitrogen) and 0.1% SDS. PAR polymer was detached using NaOH and was purified by phenol/chloroform extraction and isopropanol precipitation. After removing small molecules by Superdex 75 on fast protein liquid chromatography (FPLC) (GE healthcare), we digested purified PAR polymer by 50 U of snake venom phosphodiesterase (Worthington) with 15 mM $MgCl_2$ overnight at room temperature. The product of the phosphodiesterase digestion, *iso*-ADPR, was further purified by anion exchange chromatography and Superdex 75 on FPLC (GE Healthcare). Purified *iso*-ADPR was air-dried and dissolved by ddH_2O to 40 mM final concentration and stored at -20°C. LC-MS detected m/e 558.1 $(M-H)^-$ with high purity (the calculated mass of *iso*-ADPR in acidic form is 559.1).

2.3.2 Protein expression and purification

Human RNF146 WWE domain (99-183) gene was cloned into pGEX-4T1 with an N-terminal GST tag and a TEV cleavage site in between. Native GST fusion protein was over-expressed in

E. coli BL21 (DE3) cells (Novagen), grown in Luria broth media. Se-Met substituted GST-RNF146 WWE domain was over-expressed in auto-induction media. Bacteria cell pellets were lysed by sonication. Both native and Se-Met GST fusion proteins were eluted from Glutathione Sepharose 4B beads. GST tag was removed by TEV at 4 °C overnight. Then the proteins were further purified by a cation exchange column, and finally purified by a Superdex 200 column on FPLC (GE Healthcare). The peak fractions were pooled, and concentrated to ~10 mg/ml in a buffer containing 20 mM Tris·HCl pH 7.5, 100 mM NaCl, 2 mM DTT. The RNF146 WWE domain (residues 99-183) mutants were cloned by site-directed mutagenesis. All mutant proteins were expressed and purified as described above. Untagged mutant proteins were further purified by SEC Superdex 75 for ITC analysis.

The HUWE1 WWE domain was cloned into pGEX-6P1. Other WWEs (ULF, Deltex1, PARP11 and DDHD2) were cloned into pGEX-4T1. All GST-WWE proteins were expressed and purified as described above. GST-tagged fusion proteins were concentrated to ~1 mg/ml after anion exchange chromatography and stored at -80 °C for SPR analysis. The GST-tag on HUWE1 WWE domain was removed by human rhinovirus 3C protease, and further purified by a Superdex 75 column for ITC analysis.

2.3.3 Crystallization and structure determination

Purified Se-Met human RNF146 WWE domain (residues 99-183) at 10 mg/mL was mixed at 1:1.5 molar ratios with purified *iso*-ADPR and incubated for 30 min on ice prior to co-crystallization. The hanging-drop method was used to prepare crystals of the Se-Met RNF146 WWE domain in complex with *iso*-ADPR. 1μL of protein–ligand mixture solution was mixed

with 1μL of well solution containing 45% PEG 400, 100 mM Tris·HCl (pH 8.0), and 10 mM DTT. Ship-shaped crystals usually appeared in 1 day at 22°C and grew to full sizes in three days. The crystals were directly flash-frozen in liquid nitrogen. Screening and data collection were performed at the Advanced Light Source (ALS), beamline 8.2.1 All diffraction data were processed by HKL2000 (Otwinowski & Minor, 1997). The structure was determined by single-wavelength anomalous dispersion (SAD) using one data set collected at wavelength 0.9793 Å, which was also used for refinement (Table 2.1). The selenium sites and the initial phases were determined by PHENIX (Adams et al., 2010). Four selenium sites were found in one asymmetric unit, and the experimental electron density map clearly showed the presence of two WWE molecules with two ligands in one asymmetric unit. The complex model was improved using iterative cycles of manual rebuilding with the program COOT (Emsley, Lohkamp, Scott, & Cowtan, 2010) and refinement with Refmac5 of the CCP4 6.1.2 program suite (Collaborative Computational Project, 1994). There is no Ramachandran outlier (98.1% most favored, 1.9% allowed). The electrostatic potential surfaces shown were generated by the APBS tool in Pymol (Delano & Brunger, 1994).

2.3.4 ITC

ITC analyses were carried out using a VP-ITC Microcal calorimeter (MicroCal) at 30°C for the RNF146 WWE domain and its mutants and at 16°C for the HUWE1 WWE domain. All proteins underwent buffer exchange to 20 mM HEPES (pH 7.5), 150 mM NaCl, and 1 mM DTT by a Superdex 75 column, with a final concentration of ~20 mM. Ligands (ADPR and *iso*-ADPR) were also diluted by the same buffer to ~500 mM. A typical titration consisted of injecting 30-40

5 mL aliquots of the ligand into the protein sample (1.4218 mL chamber) at time intervals of 4min, to ensure that the titration peak reached the baseline. The ITC data were analyzed using the software Origin 7.0 provided by the manufacturer. Data were fit by a one-site model.

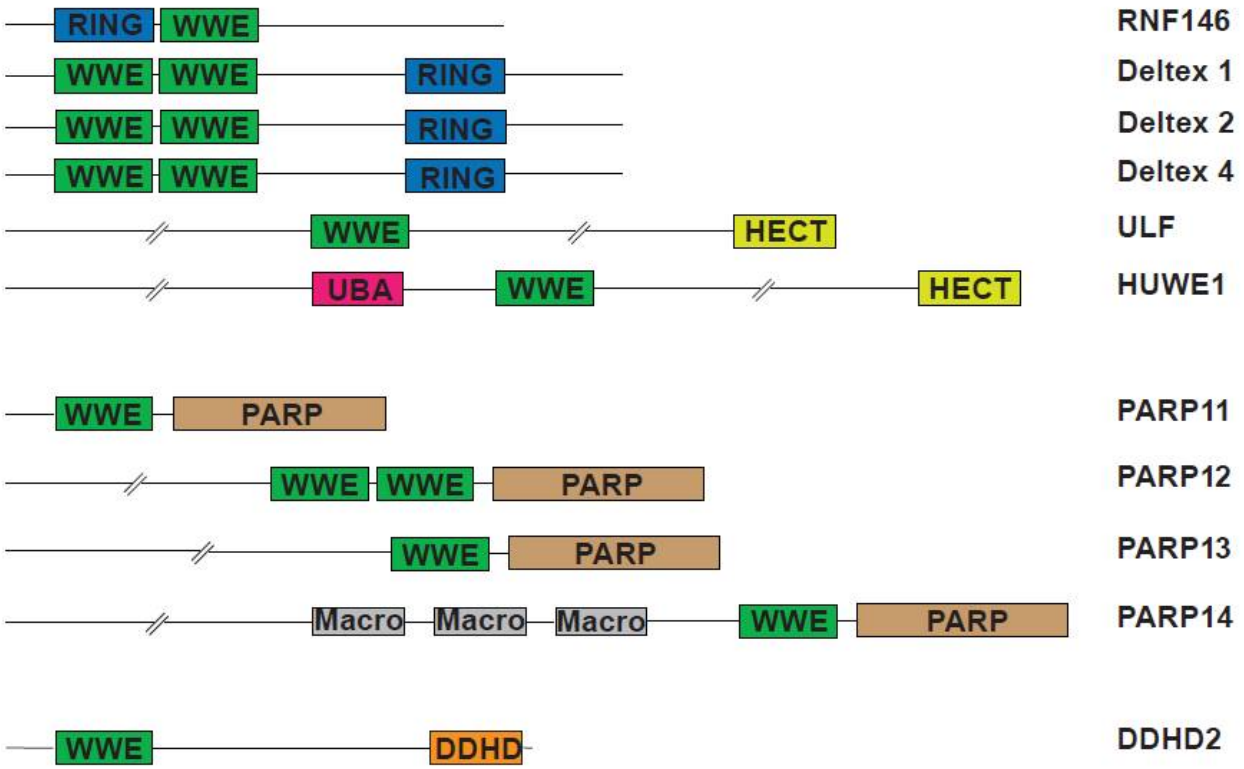

Figure 2.1 Domain structures of WWE domain containing proteins.

Domain structures of representative human WWE domain-containing proteins are shown. The majority of WWE domains are associated with E3 ubiquitin ligases. More domain structures of WWE domain-containing proteins can be found at the following pfam link (WWE PF02825). http://pfam.sanger.ac.uk/family/PF02825#tabview=tab1

Abbreviations: RING: Really Interesting New Gene finger domain; HECT: Homologous to E6-AP Carboxyl Terminus domain; UBA: Ubiquitin-Associated domain; PARP: Poly(ADP-ribose) polymerase domain. ULF: ubiquitin ligase for ARF; HUWE1: HECT, UBA, and WWE domain containing 1; PARP11: Poly(ADP-ribose) polymerase 11; PARP12: Poly(ADP-ribose) polymerase 12; PARP13: Poly(ADP-ribose) polymerase 13; PARP14: Poly(ADP-ribose) polymerase 14; DDHD2: phospholipase DDHD domain containing 2.

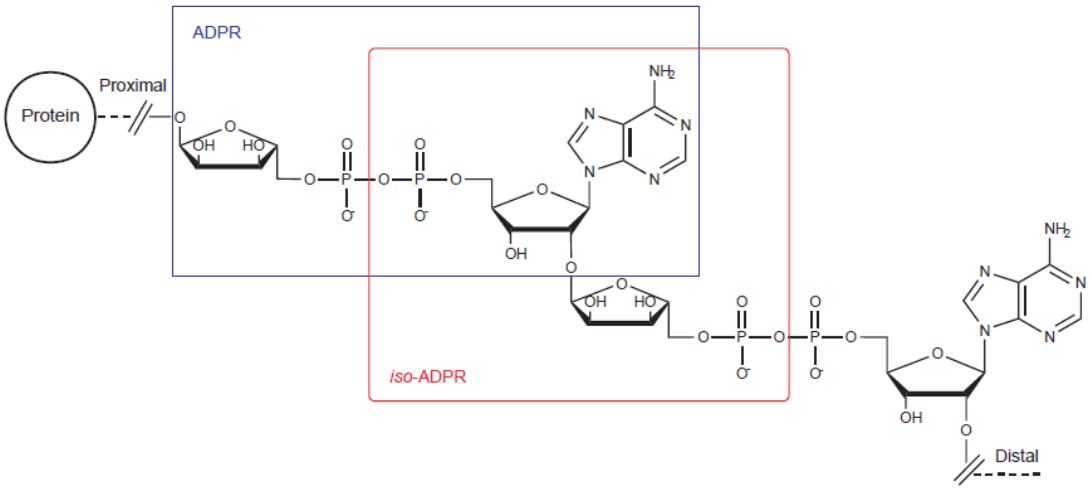

Figure 2.2 Structures of PAR and its structural units.

ADPR is the unit within the blue frame, and iso-ADPR is within the red frame.

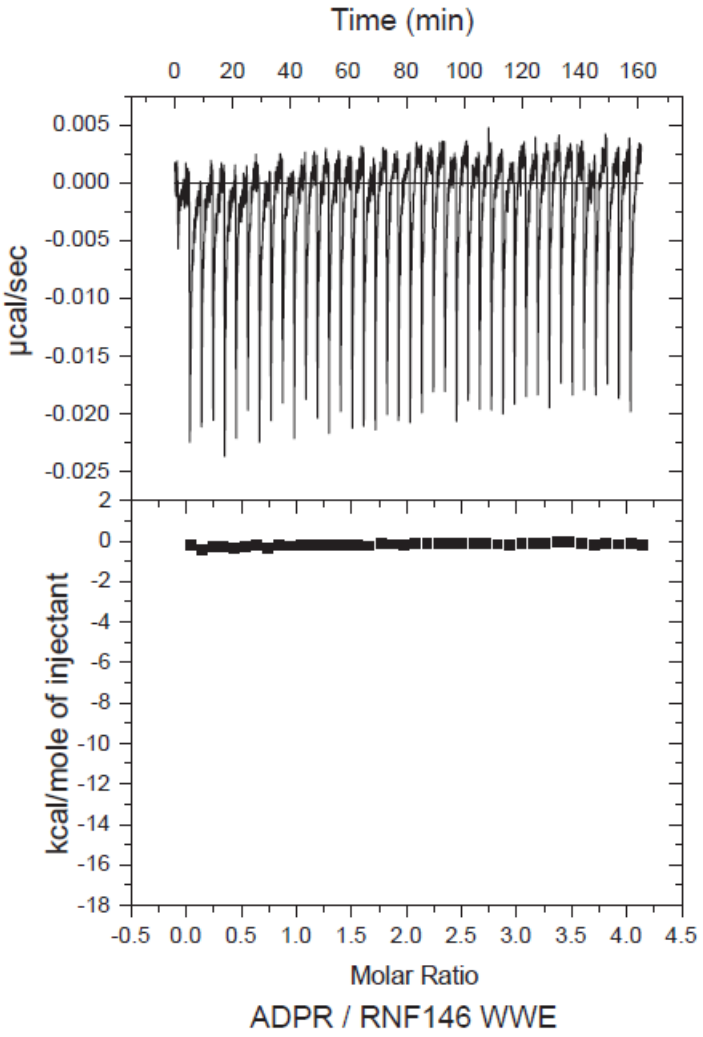

Figure 2.3 ITC analysis showed no binding of the RNF146 WWE domain with ADPR.

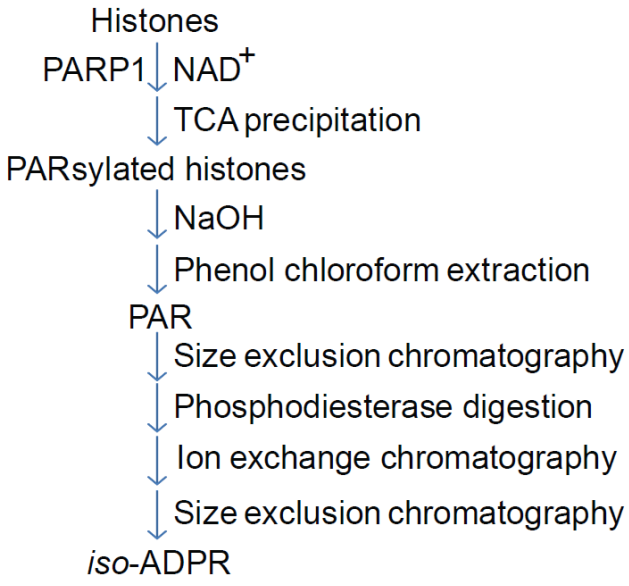

Figure 2.4 Procedural outline of *iso*-ADPR *in vitro* biosynthesis and purification.

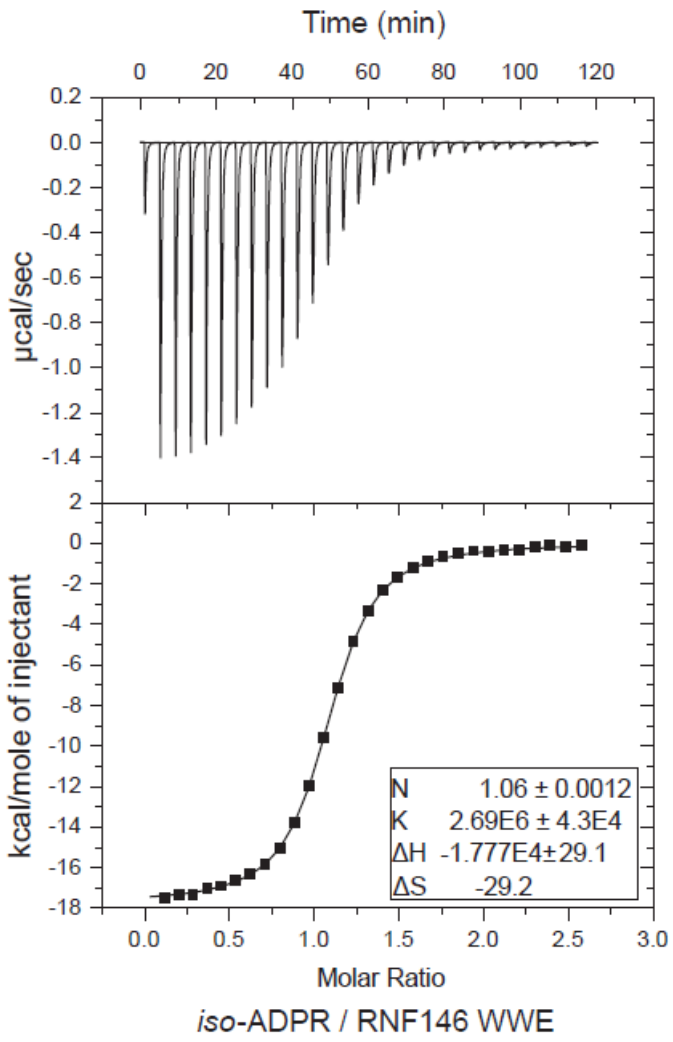

Figure 2.5 ITC analysis showed the binding of the RNF146 WWE domain with *iso*-ADPR (K_d ~ 370 nM).

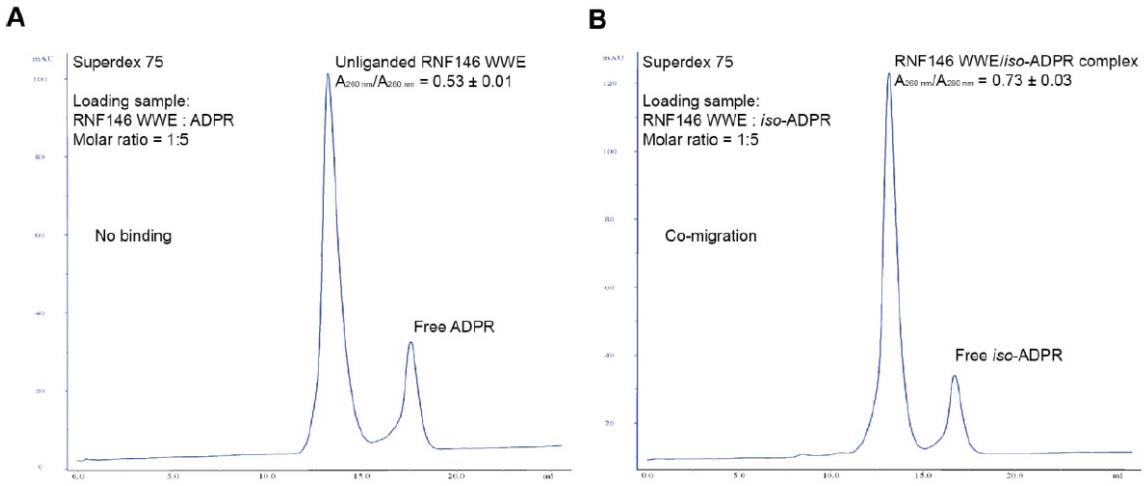

Figure 2.6 Size exclusion chromatography of (A) RNF146WWE + ADPR mixture and (B) RNF146 WWE + *iso*-ADPR mixture.

UV_{280nm} absorbance is shown in chromatogram. UV_{260nm}/UV_{280nm} values of peak fraction, as an indicator of ADPR/*iso*-ADPR binding, were determined separately. The value with ADPR (A) is the same as that of the unliganded RNF146 WWE domain. The value with *iso*-ADPR (B) indicates roughly stoichiometric co-migration of the RNF146 WWE domain and *iso*-ADPR during SEC. This experiment was repeated three times.

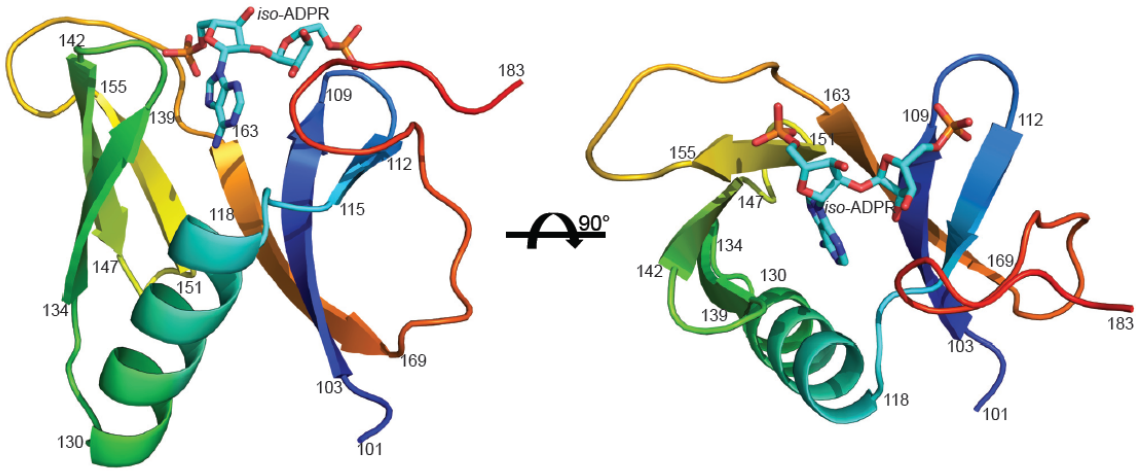

Figure 2.7 Overall structure of the complex. The protein is shown in rainbow, and *iso*-ADPR is shown in cyan.

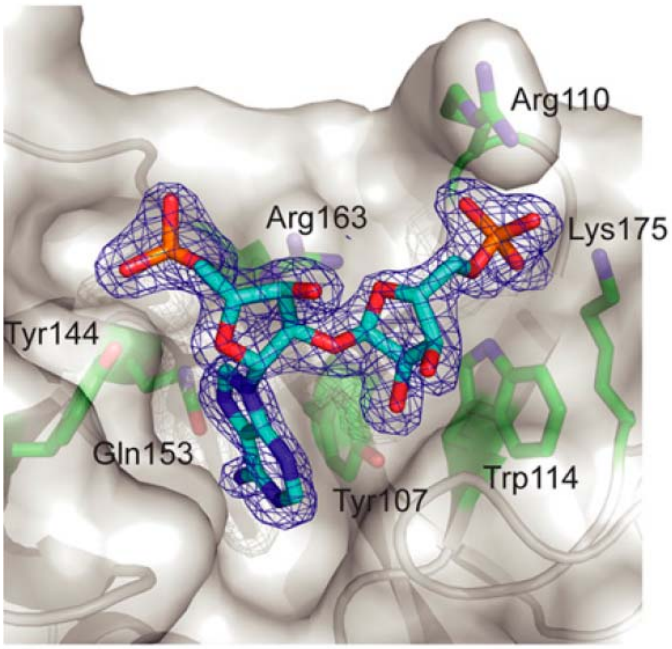

Figure 2.8 F_o - F_c difference density (blue mesh) calculated when *iso*-ADPR is omitted (contoured at 2.5σ). Protein is shown in gray, and key binding residues are highlighted in sticks.

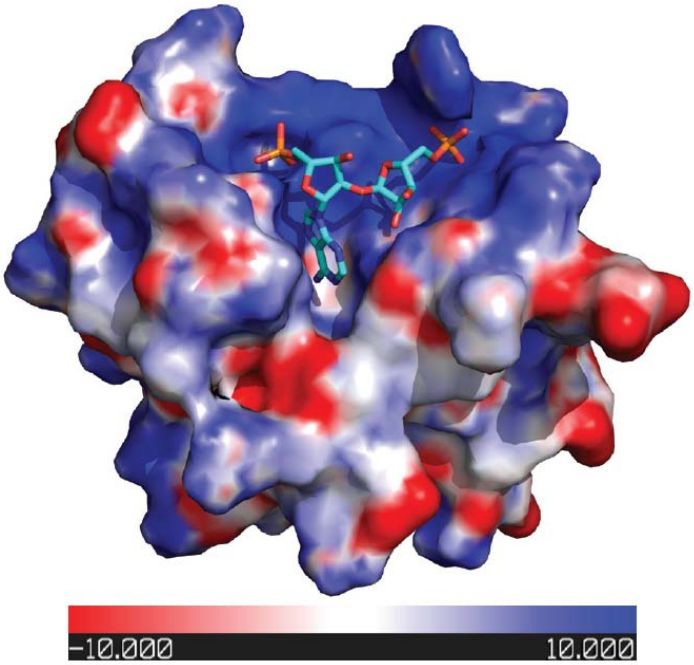

Figure 2.9 The electrostatic surface potentials of the *iso*-ADPR-binding region, calculated using Pymol. Negative potential patches are shown in red, and positive potential patches are shown in blue.

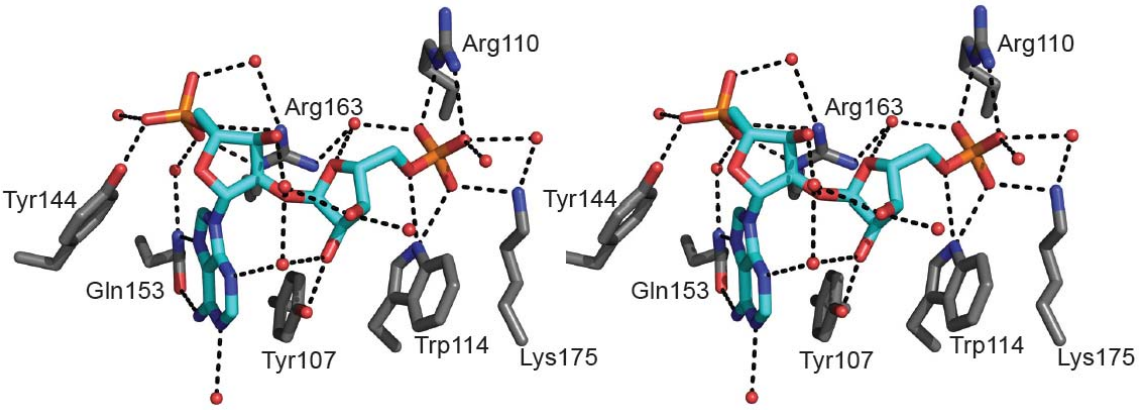

Figure 2.10 Stereo view of key interactions (in black dash lines) involved in *iso*-ADPR binding with the RNF146 WWE domain.

Key binding residues are highlighted in sticks, and water molecules are shown as pink spheres.

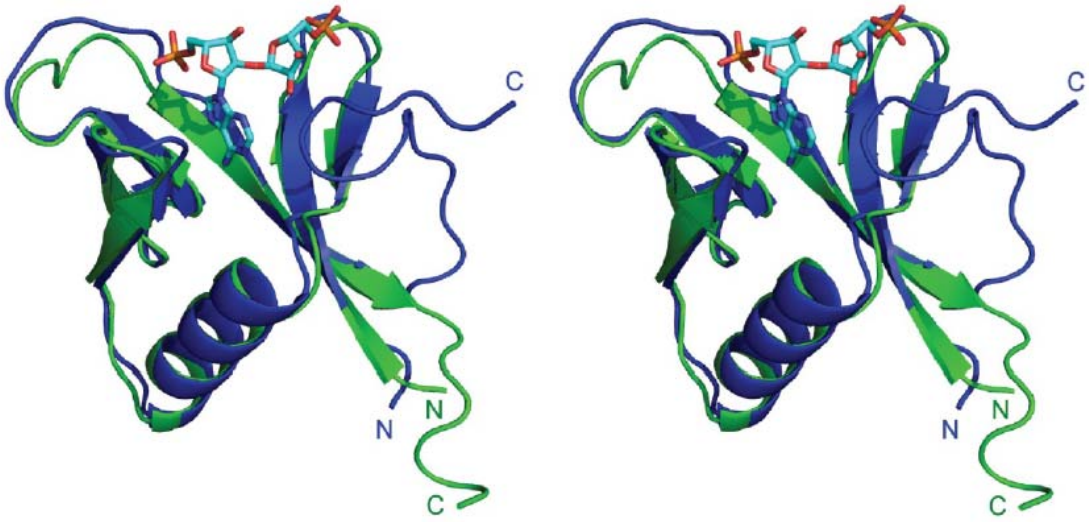

Figure 2.11 Superposition of the crystal structure of RNF146 WWE domain (in blue) in complex with *iso*-ADPR (in sticks), and the NMR structure of unliganded RNF146 WWE domain (in green, PDB 1UJR)

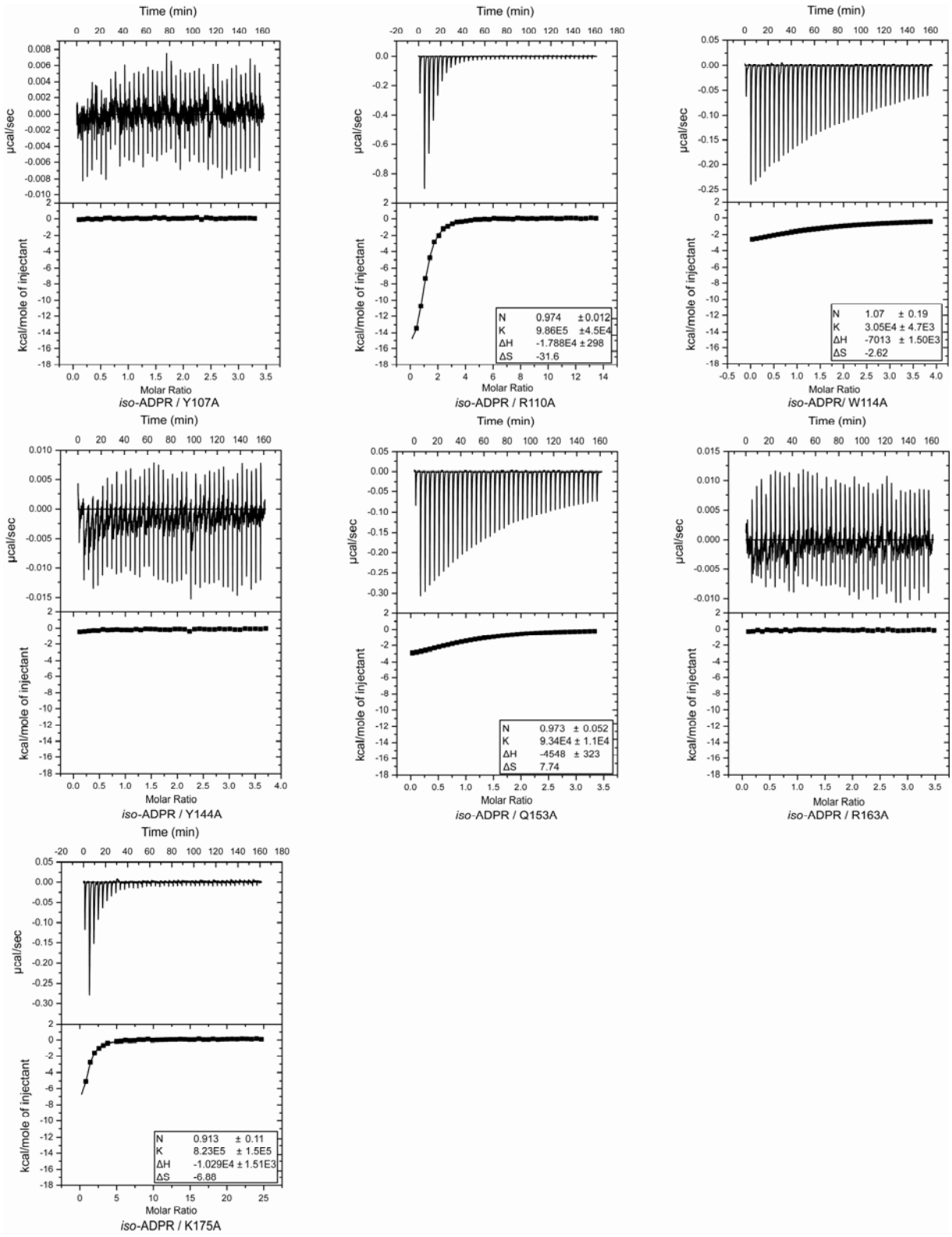

Figure 2.12 ITC measurements of the interactions between *iso*-ADPR and RNF146 WWE domain mutants.

Figure 2.13 Multiple sequence alignment by ClustalW of known human WWE domain sequences.

Most conserved residues important for WWE domain folding are shown in green. Highly conserved residues potentially involved in *iso*-ADPR binding (corresponding to Tyr 107, Tyr 144, Gln 153, and Arg 163 in RNF146) are highlighted in red, and two other non-conserved *iso*-ADPR-binding residues in RNF146 (Arg 110 and Trp 114) are in purple. The sequence surrounding RNF146 Lys 175 in the C-terminal tail is not conserved and is not shown here. In accordance with the mutagenesis results, conserved residues are important for *iso*-ADPR binding, whereas non-conserved residues are not. Deltex proteins contain tandem WWE domains (A and B represent the N-terminal and C-terminal WWE domains, respectively).

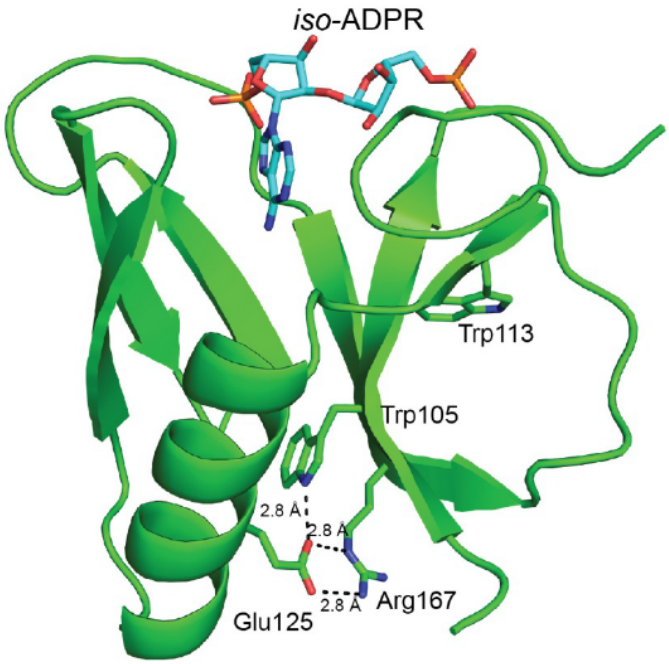

Figure 2.14 Critical residues involved in stabilizing the WWE domain structure

W105, W113, E125 and R167, as shown here, are conserved in most WWE domains (Figure 2.13). Hydrogen bonds and salt bridges are labeled as black dash lines. W105 and W113 are also involved in hydrophobic interactions stabilizing the WWE domain structure.

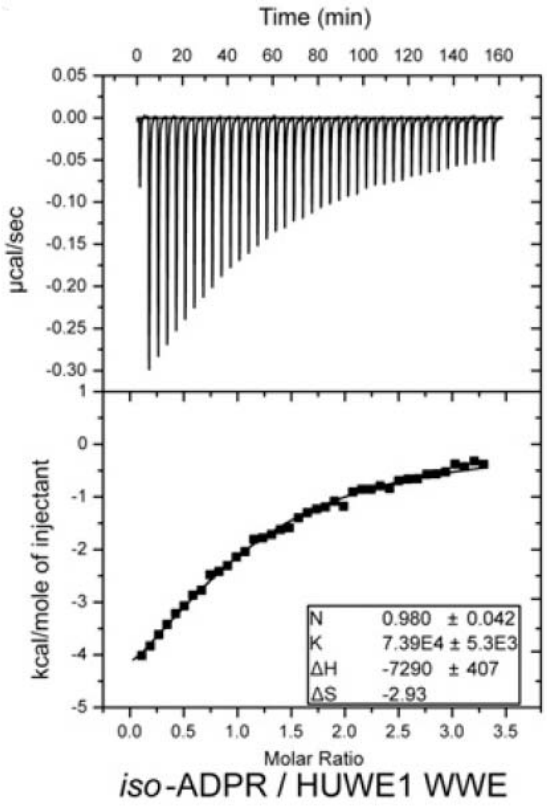

Figure 2.15 ITC analysis of the interaction between the HUWE1 WWE domain and *iso*-ADPR.

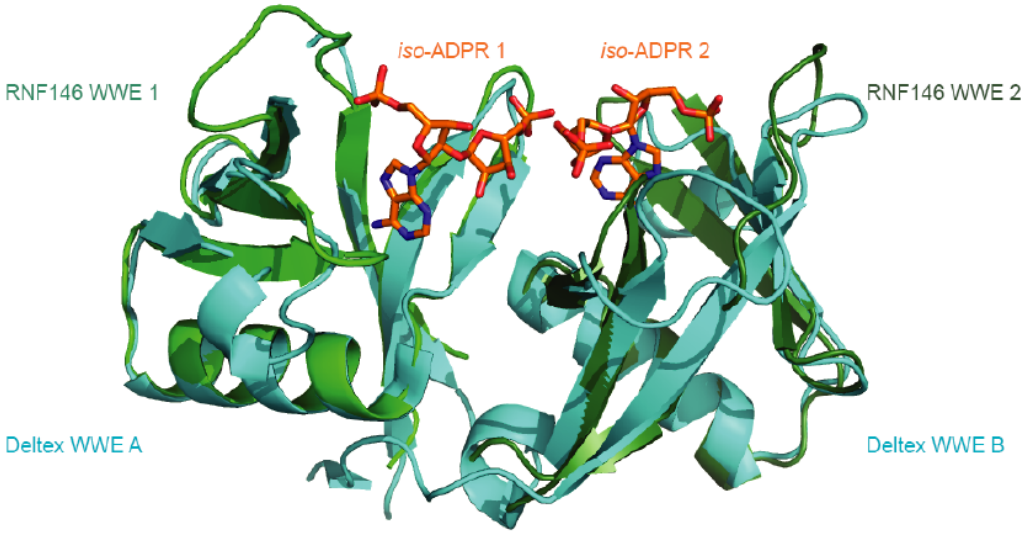

Figure 2.16 Superposition of the structures of RNF146 WWE domain (in green), in complex with *iso*-ADPR (in orange sticks), and the structure of tandem WWE domains of *Drosophila* Deltex (in cyan, PDB 2A90).

The N-terminal WWE domain of Deltex is labeled as 'WWE A', and the C-terminal one is 'WWE B'. Two RNF146 WWE domain structures are superimposed with the tandem WWE domains of Deltex separately. The two predicted *iso*-ADPR binding sites in Deltex are on the same side and close to each other.

Table 2.1 Data collection and refinement statistics

	SeMet SAD
Data collection	
Space group	$P3_112$
Cell dimensions	
a, *b*, *c* (Å)	55.34, 55.34, 146.47
α, β, γ (°)	90, 90, 120
Resolution (Å)	20.0 - 1.63 (1.68 - 1.63)*
R_{sym}(%)	7.1 (58.2)
$I/\sigma I$	60.3 (1.9)
Completeness (%)	99.6 (96.0)
Redundancy	10.9 (6.0)
Refinement	
Resolution (Å)	1.63
No. reflections	29606
R_{work}/ R_{free}	15.1 / 18.1
No. atoms	
Protein	1526
Ligand/ion	72
Water	125
B-factors	
Protein	38.6
Ligand/ion	27.5
Water	52.4
R.m.s deviations	
Bond lengths (Å)	0.012
Bond angles (°)	1.57

One crystal was used for data collection for this structure.

*Highest resolution shell is shown in parenthesis.

Table 2.2 Dissociation constants of RNF146 WWE domain mutants with *iso*-ADPR as measured by ITC analysis.

RNF146 WWE domain	Kd measured by ITC for *iso*-ADPR(μM)
Wild type	0.37 ± 0.02
Y107A	No binding
R110A	1.01 ± 0.05
W114A	32.79 ± 0.15
Y144A	No binding
Q153A	10.71 ± 0.12
R163A	No binding
K175A	1.22 ± 0.18

Chapter 3 Structural studies of PARG

3.1 Introduction

PARylation is involved in various cellular processes including DNA repair, chromatin remodeling, circadian clocks, RNA interference and cell death (De Vos, et al., 2012; Hottiger, et al., 2010; Krishnakumar & Kraus, 2010; Luo & Kraus, 2012; Satoh & Lindahl, 1992). PAR metabolism in cells is regulated by PARPs and PARG. PARG is the predominant protein responsible for PAR polymer hydrolysis *in vivo*. Loss of PARG function in *Drosophila melanogaster* results in progressive neurodegeneration with a reduced lifespan due to the excessive production of PAR in the central nervous system (Hanai et al., 2004). The PARG null mutation in mouse causes the lethal phenotype in early embryos (Koh, et al., 2004). The hypomorphic mutation of PARG ($PARG_{110}^{-/-}$) in mouse showed impaired DNA repair response with high genomic instability, including chromosome aberrations and a high frequency of sister chromatid exchange (Cortes et al., 2004; W. Min, Cortes, Herceg, Tong, & Wang, 2010).

PARG has both exo-glycosidase and endo-glycosidase activities, and is responsible for the hydrolysis of ribose-ribose glycosidic bonds between ADPR units located at the terminus of PAR and within the polymer. PARG degrades shorter and linear PAR much faster than branched, longer PAR polymers (Hatakeyama, Nemoto, Ueda, & Hayaishi, 1986; Malanga & Althaus, 1994). The PAR formed following the activation of PARP1 by DNA damage has a very short half-life (Alvarezgonzalez & Althaus, 1989). It's degraded by PARG only a few minutes after its synthesis. This function of PARG prevents the accumulation of highly PARylated proteins with very long PAR modification in the nucleus and may also keep PARP1 active by removing PAR polymer which results from the auto-inhibitory PARylation of PARP1 itself.

There is one known PARG inhibitor, adenosine diphosphate-(hydroxymethyl)-pyrrolidinediol (ADP-HPD), an analogue of ADPR. It has an IC50 (Concentration needed for 50% inhibition) of about 120 nM. ADP-HDP has been used for *in vitro* studies for PARG inhibition. However, it is not cell permeable and can be hydrolyzed by phosphodiesterases in the cell. So it's not feasible for cell based studies. The lack of an ideal small compound inhibitor for PARG is still a major hurdle for the scientists who want to further study the biological functions of PARG.

Recently, inhibitors of PARG have been proposed as drug targets in pathophysiological conditions such as inflammation, ischemia, and stroke (Cuzzocrea & Wang, 2005; Davidovic, et al., 2001; Koh, et al., 2005; W. Min & Wang, 2009). In addition, because PARG deficiency enhances cytotoxic sensitivity induced by chemotherapy agents (Koh, et al., 2004), PARG inhibitors are potential anti-cancer drug sensitizers.

The first PARG gene was cloned from bovine in the 1990s (Lin, Ame, AboulEla, Jacobson, & Jacobson, 1997), however there has been no structural data available for any mammalian PARG until now. The recently reported bacterial PARG structure revealed it has a macrodomain fold. However, this bacterial PARG is only half the size of the catalytic domain of mammalian PARG. Based on sequence alignments, there is very low similarity between the bacterial and mammalian PARGs in regions except for the catalytic loops. The bacterial PARG can only bind to the terminal ADPR unit of PAR polymer, and therefore only has exo-glycohydrolase activity (Slade, et al., 2011), whereas the mammalian PARG catalytic domain has both exo- and endo-glycohydrolase activities (Ikejima & Gill, 1988; Miwa, et al., 1974). The much more complex structure of mammalian PARG will shed light on the catalysis and regulatory mechanism of this enzyme. This work will also be important for understanding cell regulation by protein PARylation and help with rational drug design.

3.2 Results and discussion

3.2.1 Overall structure of mouse PARG catalytic domain

PARG comprises an N-terminal regulatory domain and C-terminal catalytic domain. Based on the metaPrDOS prediction (Figure 3.1) (Ishida & Kinoshita, 2008), the N-terminal region (1-438), which is absent in some PARG splicing forms, is predicted to be disordered in mice, while the conserved C-terminal 60kD catalytic domain is well-folded and fully active for PARG activity (Botta & Jacobson, 2010; Meyer, et al., 2007). I was able to purify and crystallize the recombinant mPARG (439-959) proteins from *E. coli*. The unliganded structure of mPARG (439-959) was determined with Se-Met SAD data, and refined at 1.91Å resolution (Table 3.1).

mPARG (439-959) has a kidney-shaped structure with a deep cleft in the middle on the abdominal side, which is the active site. A nine-strand mixed β sheet is sandwiched by one α helix bundle at top and another at the bottom. The top helix bundle (N-terminal half) has nine α helices, while the bottom (C-terminal half) one has five. The N-terminal half of the β sheet is made up of anti-parallel β strands, and the strands in the C-terminal half are parallel. They together merge into one twisted β sheet. The very N-terminal segment, containing the sixteen-residue mitochondrial targeting sequence (MTS, residues 454-469, MRKMPRCGIHLPSLRP), wraps around from top to bottom on the back side (Figure 3.2).

This MTS is proposed to be the signal peptide to direct the import of PARG into mitochondria. Previous studies showed that MTS plays a crucial role in PARG activity, and the deletion or mutations of MTS result in the total or partial loss of PARG enzymatic activity (Botta & Jacobson, 2010). In my structure, this MTS together with residues preceding it, has an extended conformation and wraps along the back side of the PARG catalytic domain (Figure 3.2). The

MTS docks in a hydrophobic groove on the back side of the β sheet, which is the opposite side from the active site (Figure 3.3). Hydrophobic residues Met454, Met457, Leu464 and Leu467 on the MTS pack tightly with this hydrophobic groove. This explains why the mutants of these leucine residues have no detectable enzymatic activity (Botta & Jacobson, 2010). Once the MTS is cleaved during the import into mitochondria, PARG would have a large hydrophobic surface exposed to solvent. PARG may then adopt a totally different conformation or even aggregate to accommodate this, and subsequently lose its enzymatic activity.

The core structure of the mPARG catalytic domain also has an ADPR-binding macrodomain fold, despite missing the first β strand of the macrodomain (Figure 3.4). Although the primary sequence homology is very low among PARGs from bacteria to mammalians, they all possess a conserved macrodomain fold.

There are three loops in the PARG catalytic cleft: the GGG-X_{6-8}-QEE PARG signature catalytic loop, the di-phosphate binding loop that is highly conserved between PARGs and other macrodomain structures (Slade, et al., 2011), and the third loop from a β hairpin that is unique in the mPARG structure not found in bacterial PARG (Figure 3.5). A previously biochemical study showed that Tyr788, the tip residue in the β hairpin pointing into the cleft, plays an important role for the recognition of the PARG inhibitor ADP-HPD (Koh et al., 2003).

3.2.2 Structure of mouse PARG with inhibitor

To explore how mPARG recognizes ADP-HPD, I tried to co-crystallize this inhibitor with mPARG and to soak it into mPARG crystals. With my first crystal form, both methods failed, because the cleft is blocked by another copy of the mPARG molecule in the crystal lattice. Under

a new crystallization condition, which gave me a new crystal form, I successfully soaked in ADP-HPD, and solved the mPARG/ADP-HPD complex structure (Table 3.2).

The overall structure of ADP-HPD bound to PARG is similar to the unliganded structure. ADP-HPD sits in the deep cleft of mPARG (Figure 3.6). The signature catalytic loop (GGGVTGAGLVQEE) interacts with the pyrrolidine ring of the ADP-HPD. Residue Glu748 forms hydrogen bond with the 2''-OH, while the key catalytic residue Glu749 is very close to the C1''. Glu749 may work as a general acid to protonate the 2'-OH of adenine bound ribose on the (n-1) ADPR. Residue Asn733 forms another hydrogen bond with 3''-OH to recognize the ligand. The glycine rich region may interact with the diphosphate group in the (n-1) ADPR. The second conserved glycine rich loop (GCGAFGGD) interacts with the diphosphate group of the ADP-HPD. Residue Phe868 side chain is also in close contact with the pyrrolidine ring (Figure 3.7).

On the adenine bound ribose side, the adenine ring interacts extensively with mPARG. Aromatic rings of Tyr788 and Phe895 form perpendicular and parallel π stacking interactions with the adenine ring of ADP-HPD, respectively. Tyr785 and Glu720 both form hydrogen bonds with the NH_2 group of the adenine ring. Thr718 and Ile719 are in close contact with the N1 of the adenine ring. In addition, Asn862 forms a hydrogen bond to 2'-OH of the adenine bound ribose. Tyr788 also forms a hydrogen bond with one of the phosphates (Figure 3.7). All these interactions position the PAR polymer in the right orientation to be hydrolyzed by PARG.

Upon binding, the second conserved glycine rich loop (GCGAFGGD) undergoes a major conformational change to tightly interact with the ADP-HPD (Figure 3.8). The side chain dihedral angle of Phe868 rotates about 90° to form a close contact with the pyrrolidine ring. The

side chain dihedral angle of Phe895 rotates about 120° to form parallel π stacking interactions with the adenine ring and subsequently close the deep pocket for the adenine ring (Figure 3.8).

In contrast to bacterial PARG in which the 2'-OH of the adenine bound ribose is buried, the same 2'OH group in my mouse PARG structure is exposed to the solvent, which might allow mPARG to bind (n+1) ADPR (Figure 3.9). This allows the binding and hydrolysis of the internal ADPR unit by mPARG and explains the fact that mammalian PARG has both endo- and exo-glycohydrolase enzymatic activities.

3.2.3 Structure of mouse PARG inactive mutants

In order to study the enzymatic mechanism of mPARG, I tried to soak the catalytic residue Glu749 mutant E749Q with the *iso*-ADPR, which is the "minimal" substrate for PARG because it has the α(1→2) ribose-ribose glycosidic bond to be cleaved by PARG. Unfortunately, I could only observe the structure for E749Q itself without *iso*-ADPR in the electron density map (Table 3.3). This may result from either the weak binding between PARG and *iso*-ADPR, or the minimal enzymatic activity of the E749Q mutant. Compared with the wild type mPARG structure, the E749Q mutant has a slightly different conformation in the catalytic loop (Figure 3.10).

3.3 Implications for PARG catalytic and regulatory mechanisms

3.3.1 Catalytic mechanism

Based on the bacterial PARG structure, the mechanism of PARG catalysis was proposed that binding of the terminal ADPR unit positions the ribose-ribose O-glycosidic bond in direct hydrogen bonding contact with the last Glu residue of the signature catalytic loop (GGG-X_{6-8}-QEE). A putative oxocarbenium intermediate is formed by the protonation of the (n-1) ADPR adenine bound ribose 2'-OH leaving group through Glu. This positively charged oxocarbenium is stabilized by the proximal diphosphate group of bound ADPR. A water molecule is positioned to attack the oxocarbenium intermediate, which is activated through deprotonation by Glu. This results in the release of ADP-β-ribose and (n-1) PAR (Slade, et al., 2011).

Our PARG structure in complex with ADPR-like ADP-HPD suggests that eukaryotic and bacterial PARGs use a very similar mechanism. mPARG perfectly positions PAR to the cleft, aligning the ribose-O-ribose glycosidic bond in direct hydrogen bonding contact with Glu749, which is the also the last Glu in the PARG signature catalytic loop (GGGVTGAGLVQEE). The bound ADPR unit in mammalian PARG can be either the terminal unit or an internal unit on PAR polymer. Therefore mammalian PARGs have both endo- and exo-glycohydrolase enzymatic activities.

3.3.2 Regulation of PARG activity

Previous work has revealed the N-terminal MTS segment, which is encoded by the PARG exon 4, as a regulatory component of PARG activity. In addition, it was shown that deletion of the PARG segment encoded by exon 5 can also abolish PARG activity (Niere, et al., 2012). In my crystal structure, the PARG exon 4 encoded segment (residues 439-479), and exon 5 encoded segment (residues 480-519) together form a extended loop, and wrap around from top to bottom

on the back side of PARG catalytic domain (Figure 3.2). Most hydrophobic residues of exon 4 and 5 are buried towards the internal side of PARG structure in my crystal structure. Therefore, it is plausible that proteins in the cell that can bind specifically to hydrophobic residues of the exon 4 and/or 5 segments can "pull" the extended exon 4/5 segments away from the PARG β-sheet and the N-terminal helical domain and lead to the inhibition of PARG activity. Determining what kinds of proteins may bind to the PARG exon 4/5 segments and how this binding leads to PARG inhibition are of crucial importance for understanding protein PARylation and await future investigation.

3.4 Materials and methods

3.4.1 Protein expression and purification

The mouse PARG catalytic domain (439-959) gene was cloned into pGEX-4T1 with an N-terminal GST tag and a TEV cleavage site in-between. Native GST fusion protein was over-expressed in *E. coli* BL21 (DE3) cells (Novagen), grown in Luria broth media. Se-Met substituted GST-mPARG (439-959) was over-expressed in auto-induction media. Bacteria cell pellets were lysed by sonication. Both native and Se-Met GST fusion proteins were eluted from Glutathione Sepharose 4B beads. GST tag was removed by TEV at 4 °C overnight. Then the proteins were further purified by an anion exchange column, and finally purified by a Superdex 200 column on FPLC (GE Healthcare). The peak fractions were pooled, and concentrated to ~5 mg/ml in a buffer containing 10 mM Tris·HCl pH 8.5, 100 mM NaCl, 2 mM DTT. The catalytic residue mutant E749Q of mPARG (439-959) was cloned by site-directed mutagenesis. The mutant protein was expressed and purified using the same methods as for the wild type protein.

3.4.2 Crystallization and structure determination

The hanging-drop vapor diffusion method for crystallization was used to prepare crystals of the Se-Met mPARG(439-959). For the crystal form for unliganded structure and mutant E749Q, 1μL of protein sample (5mg/mL) was mixed with 1μL of well solution containing 20% PEG3350, 0.2M $(NH_4)_2SO_4$ at 4°C. The best crystals were obtained by further micro-seeding in 14% PEG3350, 0.2M $(NH_4)_2SO_4$. Thick plate-shaped crystals usually appeared in one day at 4°C after seeding and grew to their full sizes in three days. The crystals were frozen by liquid nitrogen in the cryo solution containing 10% glycerol and 20% PEG3350. For the crystals of ADP-HPD bound structure, 1μL of protein sample with 25% glycerol was mixed with 1μL of well solution containing 0.22M KI, 20% PEG3350 10 mM DTT at room temperature. The crystals grew to full size in 2-3 days, and were soaked with 1mM ADP-HPD overnight at room temperature. Then they were frozen by liquid nitrogen in cryo solution containing 10% glycerol and 20% PEG3350.

Screening and data collection were performed at the Advanced Light Source (ALS), beamlines 8.2.1 and 8.2.2 at wavelength 0.9793 Å. All diffraction data were processed by HKL2000 (Otwinowski & Minor, 1997). The unliganded structure was determined by single-wavelength anomalous dispersion (SAD) using one data set collected at wavelength 0.9793 Å, which was also used for refinement. The selenium sites and the initial phases were determined by PHENIX (Adams, et al., 2010). Thirty-six selenium sites were found in one asymmetric unit, and the experimental electron density map clearly showed the presence of four molecules of mPARG (439-959) in one asymmetric unit. The initial phases for ADP-HPD bound mPARG (439-959) and E749Q mutant were determined by molecular replacement in Phaser (McCoy et al., 2007). All models were improved using iterative cycles of manual rebuilding with the program COOT

(Emsley, et al., 2010) and refinement with Refmac5 of the CCP4 6.1.2 program suite (Collaborative Computational Project, 1994).

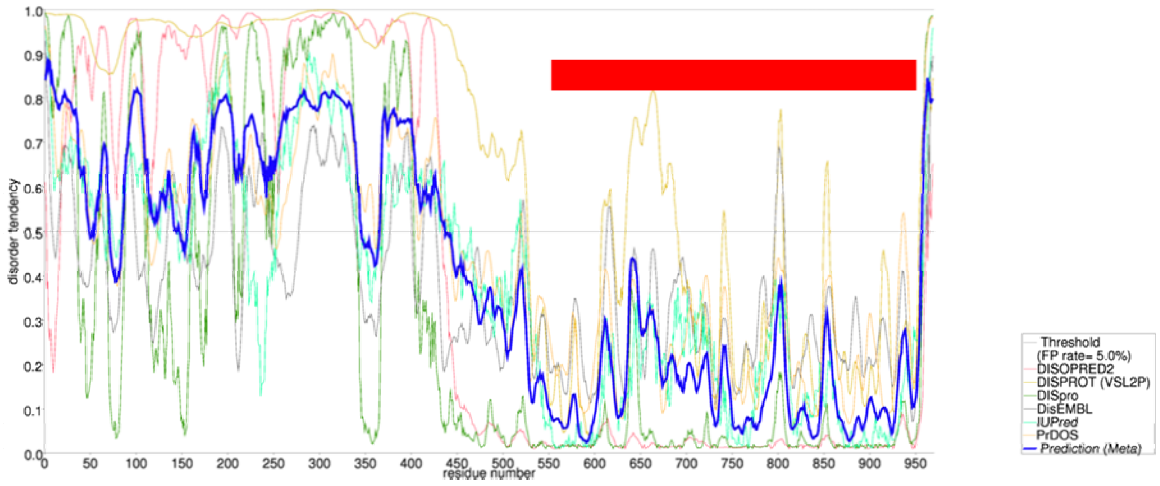

Figure 3.1 The disorder prediction for mouse PARG from metaPrDOS server.

It indicates the N-terminal regulatory domain of mPARG (1-438) is disordered, whereas the mPARG (439-959) protein (highlighted in red solid bar) that was used for crystallization trials was predicted to be well-folded.

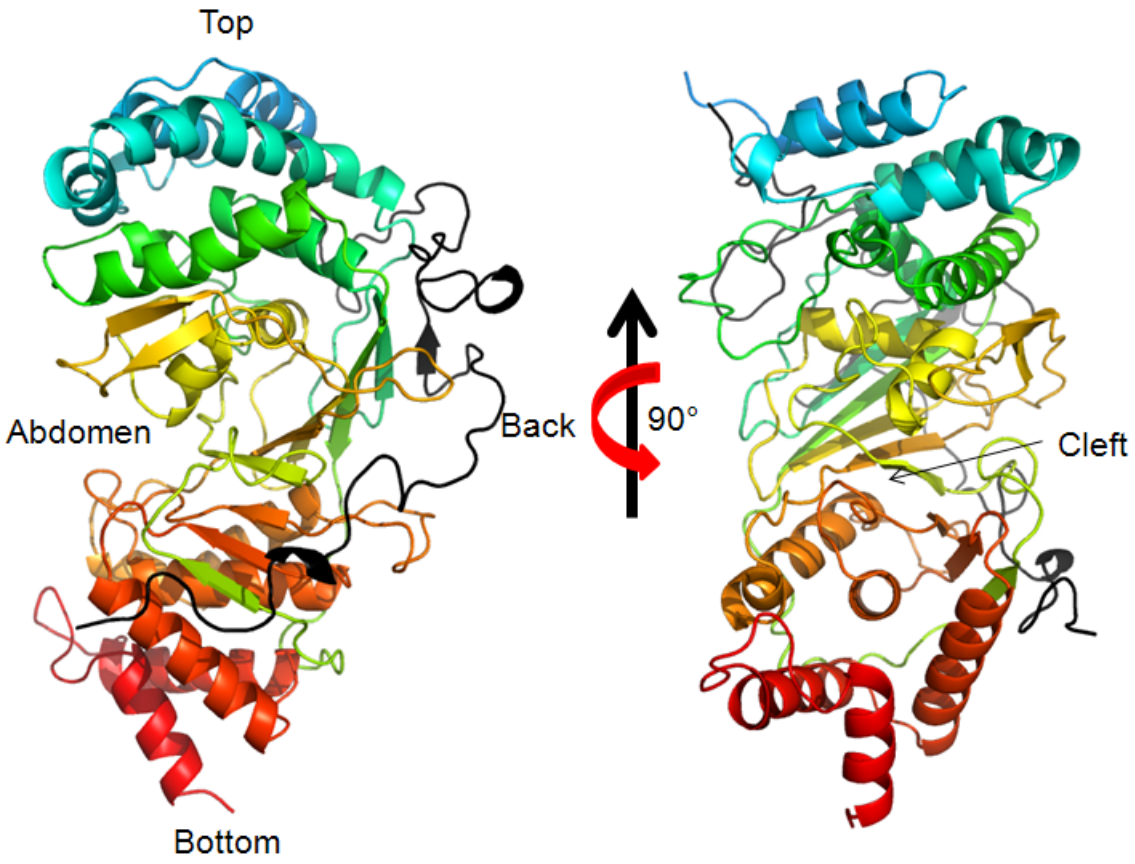

Figure 3.2 Overall structure of unliganded mPARG(439-959).

The N terminal MTS containing loop is in black. The cleft is right in the middle of the abdominal side.

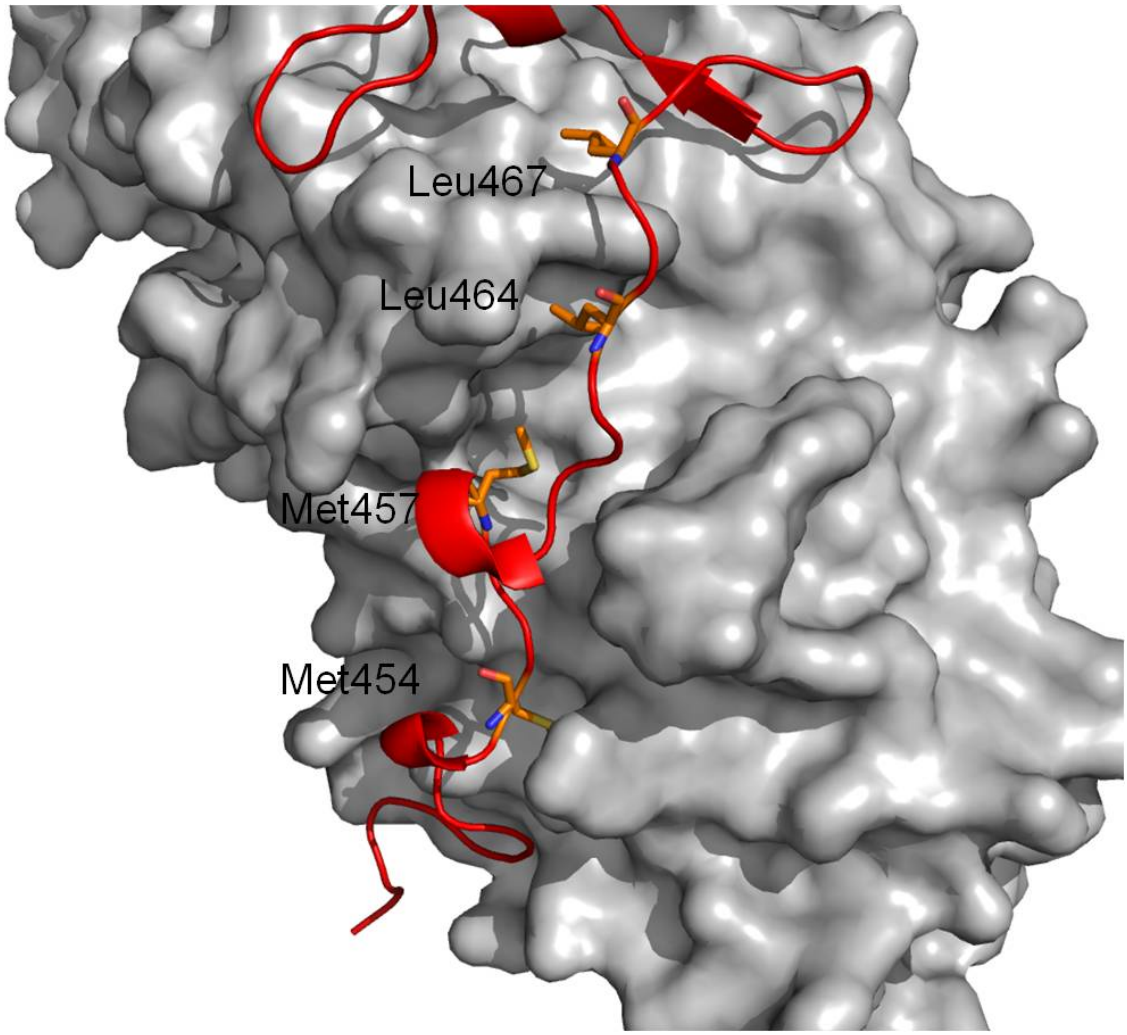

Figure 3.3 MTS docks on hydrophobic groove of the back side of mPARG catalytic domain.

Met454, Met457, Leu464 and Leu467 are highlighted in orange sticks.

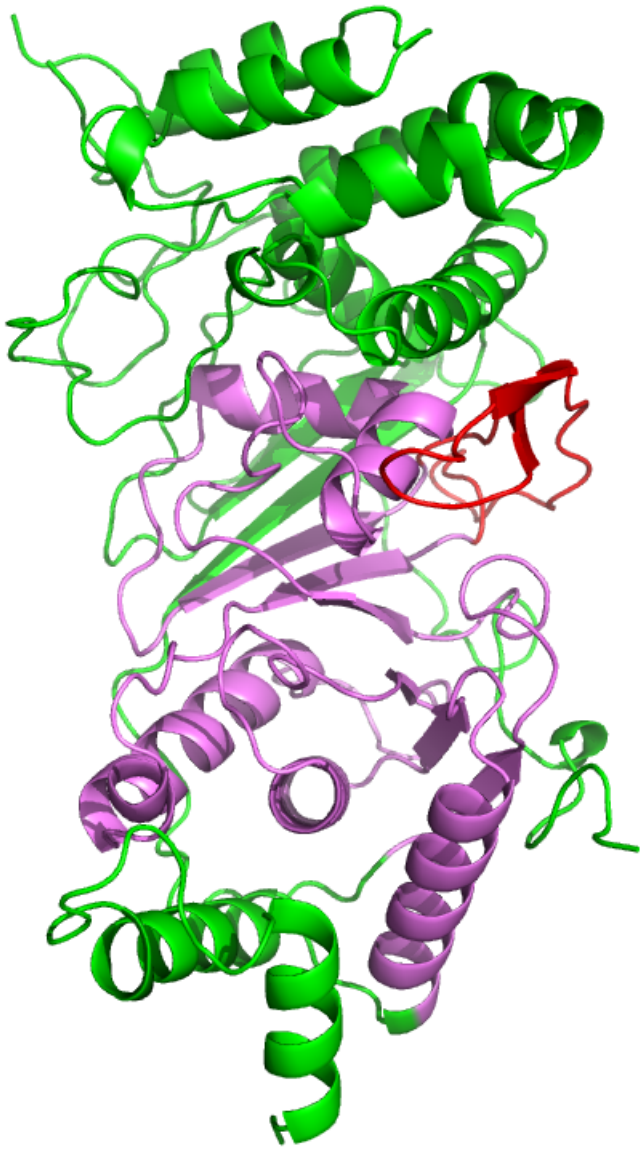

Figure 3.4 The core structure of mPARG has a macrodomain-like fold.

The macrodomain-like region is highlighted in pink. mPARG has more delicate structure than macrodomain, including the N-terminal extended loop, seven more helices in the N-terminal helix bundle, two more helices in the C-terminal helix bundle, and three more N-terminal β strands (all highlighted in green). In addition, mPARG has one β hairpin insertion in the macrodomain-like region (highlighted in red).

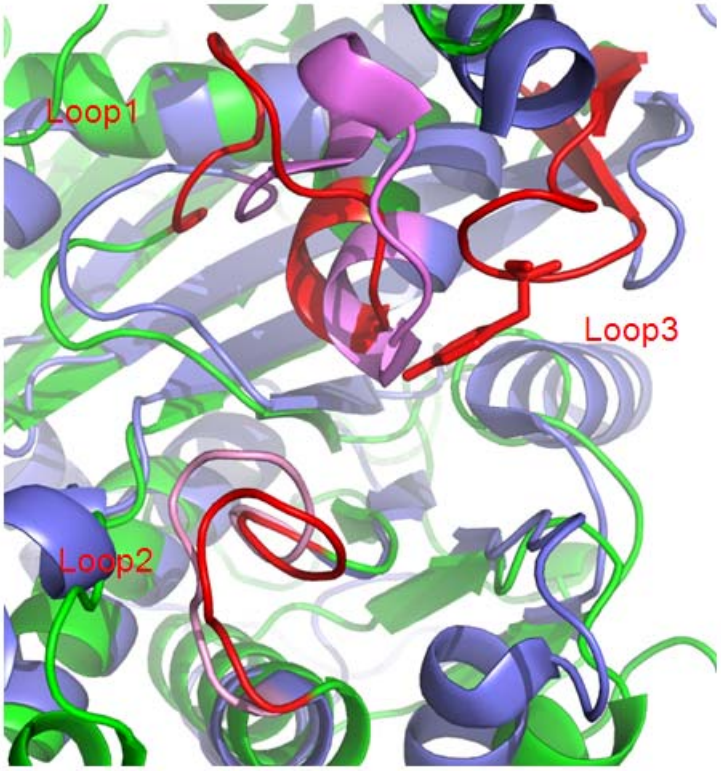

Figure 3.5 Superposition of mPARG and bacterial PARG at the cleft.

mPARG is shown in green, and bacterial PARG in light blue. The three loops (1-3) of mPARG at the cleft are highlighted in red, including the Tyr788 on the β hairpin. Bacterial PARG only has the conserved loop1 and loop2 (highlighted in pink).

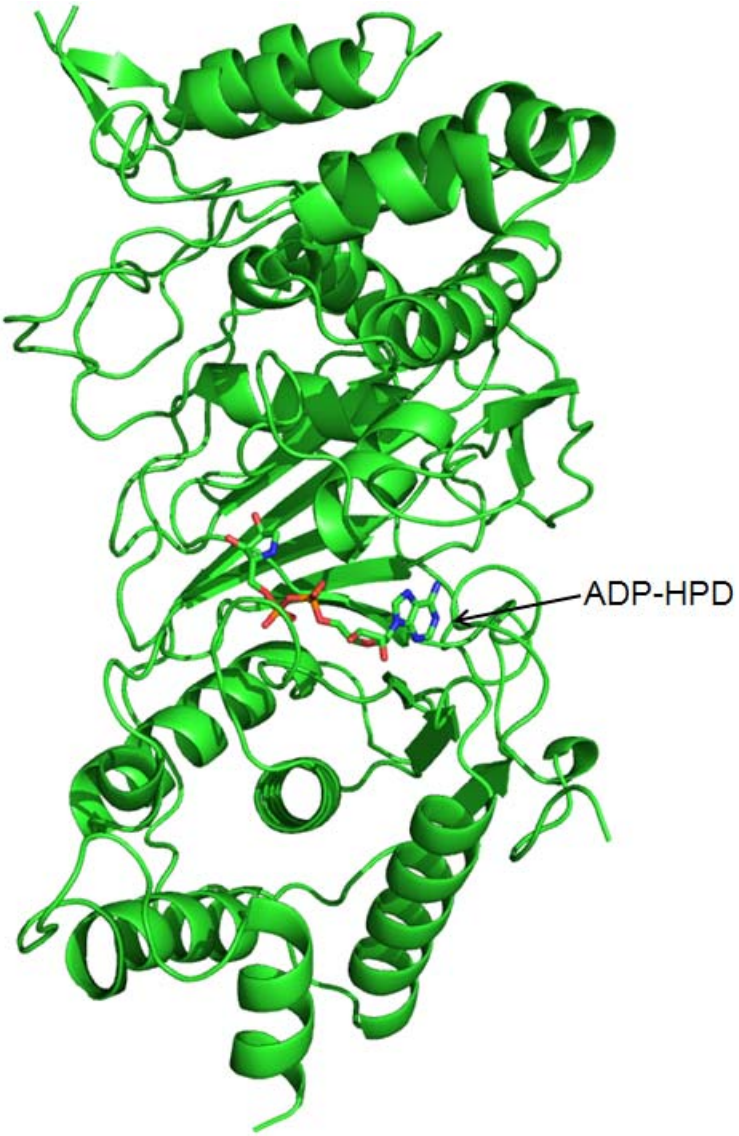

Figure 3.6 Structure of ADP-HPD bound mPARG.

ADPR analog ADP-HPD molecule (in stick) sits in the deep cleft of mPARG.

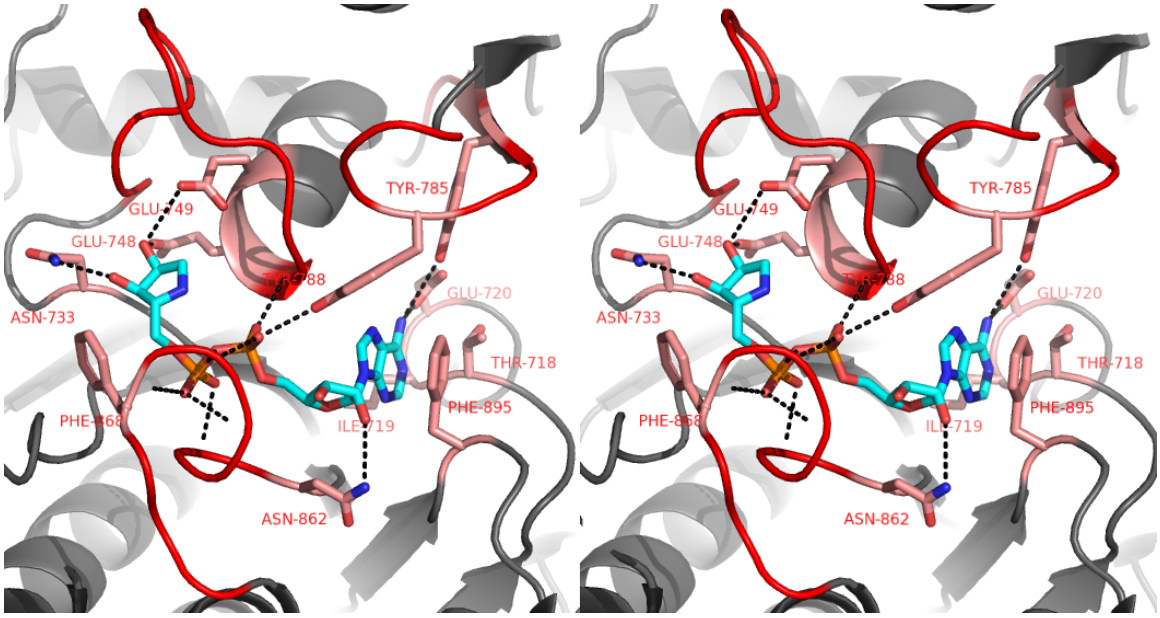

Figure 3.7 Stereo view of the ADP-HPD binding in the cleft.

Residues involved in ADP-HPD binding are highlighted in pink sticks. Hydrogen bonds are labeled as black dash lines. ADP-HPD is highlighted as stick. Three loops are highlighted in red.

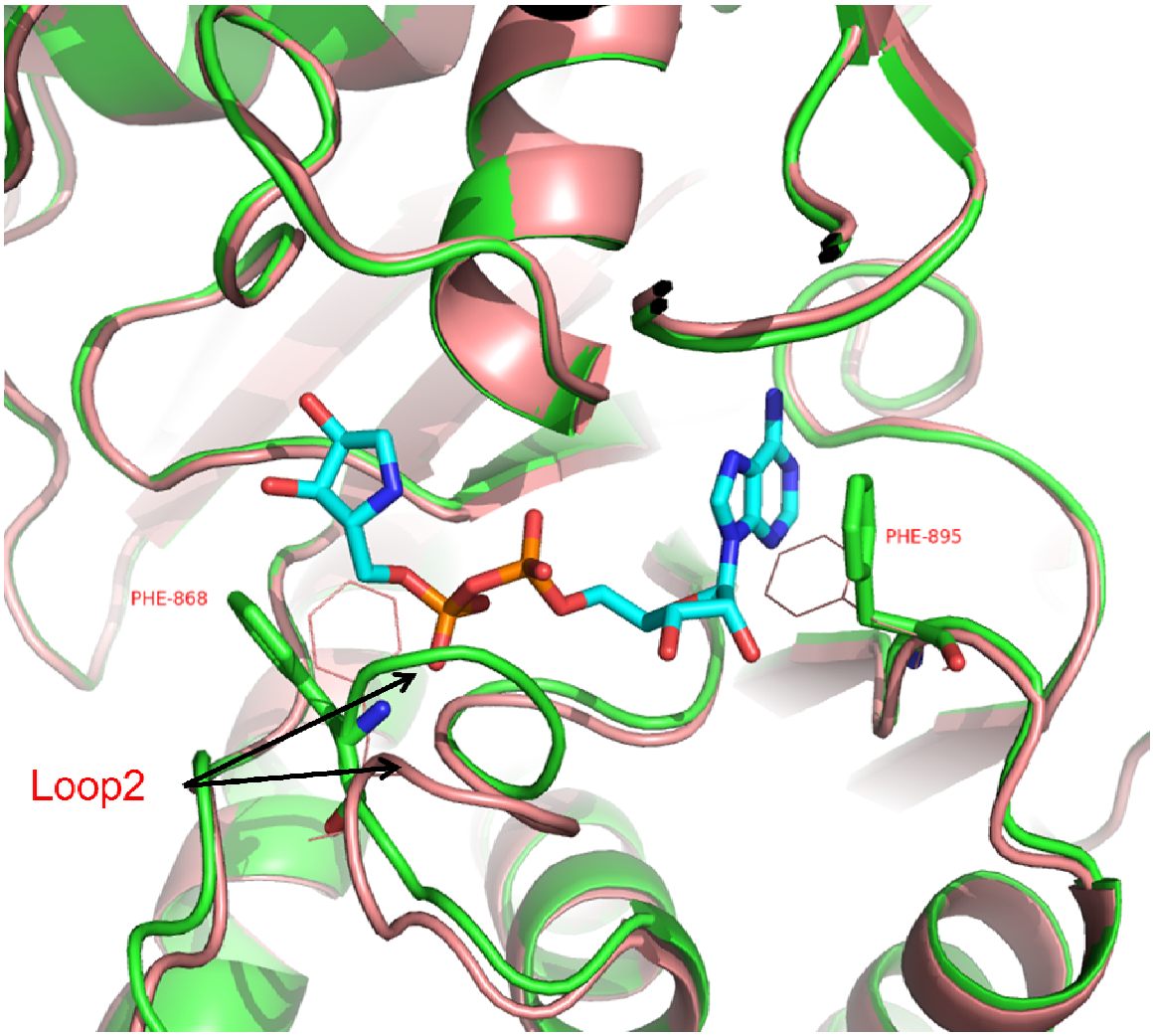

Figure 3.8 Superposition of unliganded mPARG (light pink) and ADP-HPD bound mPARG (green) structures.

Loop2 undergoes conformational change to tightly pack the ADP-HPD. Both side chains of Phe868 and Phe895 (highlighted in green sticks) rotate to strongly interact with ADP-HPD.

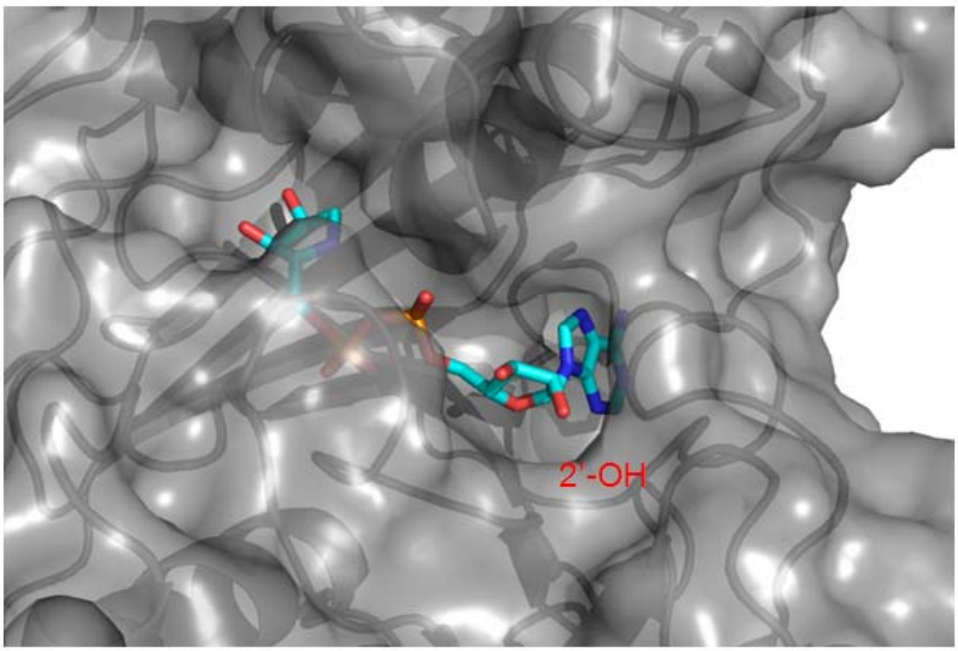

Figure 3.9 2’-OH of the adenine bound ribose is exposed to solvent.

The surface of the mPARG is shown as gray. ADPR analog ADP-HPD is highlighted as stick. Unlike bacterial PARG, mPARG doesn’t block the 2’-OH of the adenine bound ribose. This allows the binding of (n+1) ADPR unit. This structure feature supports that mPARG has both exo- and endo- glycohydrolase activity.

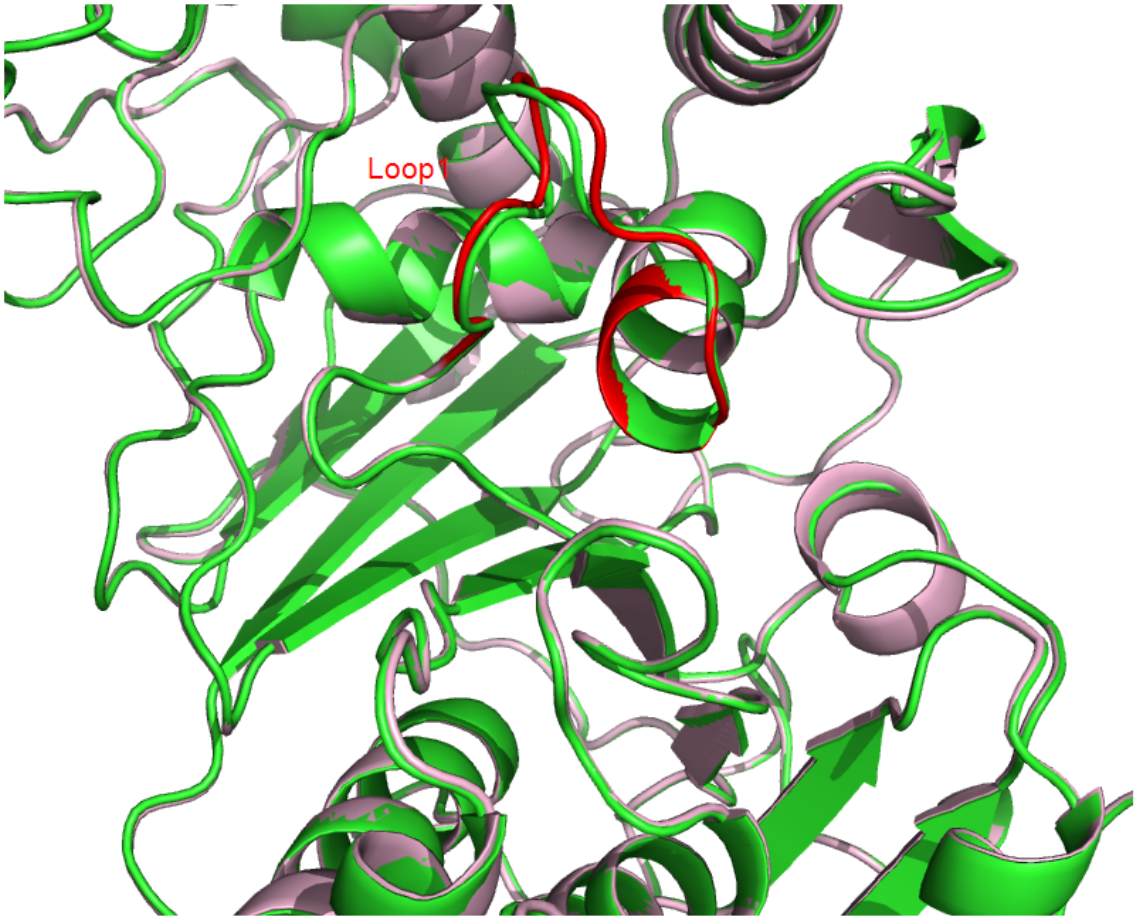

Figure 3.10 Superposition of mutant E749Q and wild type mPARG structures.

E749Q is in light pink, and wild type mPARG is in green. There is minor conformational change in the signature catalytic loop (highlighted in red) in E749Q.

Table 3.1 Data collection and refinement statistics for unliganded mPARG(439-959)

	SeMet SAD
Data collection	
Space group	*P*1
Cell dimensions	
a, *b*, *c* (Å)	67.14, 90.4, 104.7
α, β, γ (°)	81.64, 88.41, 89.36
Resolution (Å)	50.0-1.91 (1.98-1.91)*
R_{sym}(%)	12.8 (58.5)
$I/\sigma I$	16.7 (1.5)
Completeness (%)	92.1 (54.8)
Redundancy	3.4 (2.6)
Refinement	
Resolution (Å)	1.91
No. reflections	173052
R_{work}/ R_{free}	17.4 / 21.4
No. atoms	
Protein	16481
Ligand/ion	140
Water	1461
B-factors	
Protein	37.9
Ligand/ion	74.0
Water	42.6
R.m.s deviations	
Bond lengths (Å)	0.009
Bond angles (°)	1.12

One crystal was used for data collection for this structure.

*Highest resolution shell is shown in parenthesis.

Table 3.2 Data collection and refinement statistics for ADP-HPD/mPARG(439-959)

	ADP-HPD bound
Data collection	
Space group	$P2_12_12$
Cell dimensions	
a, *b*, *c* (Å)	188.95, 55.57, 165.99
α, β, γ (°)	90, 90, 90
Resolution (Å)	40.0-2.50 (2.57-2.50)*
R_{sym}(%)	9.0 (83.4)
$I/\sigma I$	25.7 (1.5)
Completeness (%)	99.6 (97.7)
Redundancy	5.0 (4.2)
Refinement	
Resolution (Å)	2.50
No. reflections	58701
R_{work}/ R_{free}	24.3/28.2
No. atoms	
Protein	12219
Ligand/ion	110
Water	93
B-factors	
Protein	21.2
Ligand/ion	46.6
Water	18.5
R.m.s deviations	
Bond lengths (Å)	0.007
Bond angles (°)	1.015

One crystal was used for data collection for this structure.

*Highest resolution shell is shown in parenthesis.

Table 3.3 Data collection and refinement statistics for mPARG(439-959) E749Q

	E749Q
Data collection	
Space group	$P2_1$
Cell dimensions	
a, *b*, *c* (Å)	66.92, 91.28, 102.52
α, β, γ (°)	90, 91.20, 90
Resolution (Å)	30.0-1.90 (1.96-1.90)*
R_{sym}(%)	11.0 (33.4)
$I/\sigma I$	13.4 (3.9)
Completeness (%)	99.1 (97.8)
Redundancy	4.5 (4.3)
Refinement	
Resolution (Å)	1.90
No. reflections	90220
R_{work}/ R_{free}	18.6/22.8
No. atoms	
Protein	8251
Ligand/ion	122
Water	967
B-factors	
Protein	18.1
Ligand/ion	42.3
Water	23.9
R.m.s deviations	
Bond lengths (Å)	0.011
Bond angles (°)	1.287

One crystal was used for data collection for this structure.

*Highest resolution shell is shown in parenthesis.

Bibliography

Adams, P. D., Afonine, P. V., Bunkoczi, G., Chen, V. B., Davis, I. W., Echols, N., et al. (2010). PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallographica Section D-Biological Crystallography, 66*.

Ahel, D., Horejsi, Z., Wiechens, N., Polo, S. E., Garcia-Wilson, E., Ahel, I., et al. (2009). Poly(ADP-ribose)-Dependent Regulation of DNA Repair by the Chromatin Remodeling Enzyme ALC1. *Science, 325*(5945).

Ahel, I., Ahel, D., Matsusaka, T., Clark, A. J., Pines, J., Boulton, S. J., et al. (2008). Poly(ADP-ribose)-binding zinc finger motifs in DNA repair/checkpoint proteins. *Nature, 451*(7174).

Alvarezgonzalez, R., & Althaus, F. R. (1989). POLY(ADP-RIBOSE) CATABOLISM IN MAMMALIAN-CELLS EXPOSED TO DNA-DAMAGING AGENTS. *Mutation Research, 218*(2).

Andrabi, S. A., Dawson, T. M., & Dawson, V. L. (2008). Mitochondrial and Nuclear Cross Talk in Cell Death Parthanatos. *Mitochondria and Oxidative Stress in Neurodegenerative Disorders, 1147*.

Andrabi, S. A., Kang, H. C., Haince, J.-F., Lee, Y.-I., Zhang, J., Chi, Z., et al. (2011). Iduna protects the brain from glutamate excitotoxicity and stroke by interfering with poly(ADP-ribose) polymer-induced cell death. *Nature Medicine, 17*(6).

Andrabi, S. A., Kim, N. S., Yu, S.-W., Wang, H., Koh, D. W., Sasaki, M., et al. (2006). Poly(ADP-ribose) (PAR) polymer is a death signal. *Proceedings of the National Academy of Sciences of the United States of America, 103*(48).

Aravind, L. (2001). The WWE domain: a common interaction module in protein ubiquitination and ADP ribosylation. *Trends in Biochemical Sciences, 26*(5).

Asher, G., Reinke, H., Altmeyer, M., Gutierrez-Arcelus, M., Hottiger, M. O., & Schibler, U. (2010). Poly(ADP-Ribose) Polymerase 1 Participates in the Phase Entrainment of Circadian Clocks to Feeding. *Cell, 142*(6).

Bergink, S., & Jentsch, S. (2009). Principles of ubiquitin and SUMO modifications in DNA repair. *Nature, 458*(7237).

Bernassola, F., Karin, M., Ciechanover, A., & Melino, G. (2008). The HECT family of E3 ubiquitin ligases: Multiple players in cancer development. *Cancer Cell, 14*(1).

Botta, D., & Jacobson, M. K. (2010). Identification of a Regulatory Segment of Poly(ADP-ribose) Glycohydrolase. *Biochemistry, 49*(35).

Bouchard, V. J., Rouleau, M., & Poirier, G. G. (2003). PARP1, a determinant of cell survival in response to DNA damage. *Experimental Hematology (New York), 31*(6).

Callow, M. G., Tran, H., Phu, L., Lau, T., Lee, J., Sandoval, W. N., et al. (2011). Ubiquitin Ligase RNF146 Regulates Tankyrase and Axin to Promote Wnt Signaling. *Plos One, 6*(7).

Chen, D., Brooks, C. L., & Gu, W. (2006). ARF-BP1 as a potential therapeutic target. *British Journal of Cancer, 94*(11).

Chen, D., Shan, J., Zhu, W.-G., Qin, J., & Gu, W. (2010). Transcription-independent ARF regulation in oncogenic stress-mediated p53 responses. *Nature, 464*(7288).

Chou, D. M., Adamson, B., Dephoure, N. E., Tan, X., Nottke, A. C., Hurov, K. E., et al. (2010). A chromatin localization screen reveals poly (ADP ribose)-regulated recruitment of the repressive polycomb and NuRD complexes to sites of DNA damage. *Proceedings of the National Academy of Sciences of the United States of America, 107*(43).

Collaborative Computational Project, N. (1994). The CCP4 Suite: Programs for protein crystallography. *Acta Crystallographica Section D Biological Crystallography, 50*(5).

Corda, D., & Di Girolamo, M. (2003). Functional aspects of protein mono-ADP-ribosylation. *Embo Journal, 22*(9).

Cortes, U., Tong, W. M., Coyle, D. L., Meyer-Ficca, M. L., Meyer, R. G., Petrilli, V., et al. (2004). Depletion of the 110-kilodalton isoform of poly(ADP-ribose) glycohydrolase increases sensitivity to genotoxic and endotoxic stress in mice. *Molecular and Cellular Biology, 24*(16).

Curtin, N. J. (2005). PARP inhibitors for cancer therapy. *Expert Reviews in Molecular Medicine, 7*(4).

Cuzzocrea, S., & Wang, Z. Q. (2005). Role of poly(ADP-ribose) glycohydrolase (PARG) in shock, ischemia and reperfusion. *Pharmacological Research, 52*(1).

Davidovic, L., Vodenicharov, M., Affar, E. B., & Poirier, G. G. (2001). Importance of poly(ADP-ribose) glycohydrolase in the control of poly(ADP-ribose) metabolism. *Experimental Cell Research, 268*(1).

De Vos, M., Schreiber, V., & Dantzer, F. (2012). The diverse roles and clinical relevance of PARPs in DNA damage repair: Current state of the art. *Biochemical Pharmacology, 84*(2).

Delano, W. L., & Brunger, A. T. (1994). HELIX PACKING IN PROTEINS - PREDICTION AND ENERGETIC ANALYSIS OF DIMERIC, TRIMERIC, AND TETRAMERIC GCN4 COILED-COIL STRUCTURES. *Proteins-Structure Function and Genetics, 20*(2).

Emsley, P., Lohkamp, B., Scott, W. G., & Cowtan, K. (2010). Features and development of Coot. *Acta Crystallographica Section D-Biological Crystallography, 66*.

Eustermann, S., Brockmann, C., Mehrotra, P. V., Yang, J.-C., Loakes, D., West, S. C., et al. (2010). Solution structures of the two PBZ domains from human APLF and their interaction with poly(ADP-ribose). *Nature Structural & Molecular Biology, 17*(2).

Fahrer, J., Kranaster, R., Altmeyer, M., Marx, A., & Buerkle, A. (2007). Quantitative analysis of the binding affinity of poly(ADP-ribose) to specific binding proteins as a function of chain length. *Nucleic Acids Research, 35*(21).

Gagne, J.-P., Isabelle, M., Lo, K. S., Bourassa, S., Hendzel, M. J., Dawson, V. L., et al. (2008). Proteome-wide identification of poly(ADP-ribose) binding proteins and poly(ADP-ribose)-associated protein complexes. *Nucleic Acids Research, 36*(22).

Gibson, B. A., & Kraus, W. L. (2012). New insights into the molecular and cellular functions of poly(ADP-ribose) and PARPs. *Nature Reviews Molecular Cell Biology, 13*(7).

Hanai, S., Kanai, M., Ohashi, S., Okamoto, K., Yamada, M., Takahashi, H., et al. (2004). Loss of poly(ADP-ribose) glycohydrolase causes progressive neurodegeneration in Drosophila melanogaster. *Proceedings of the National Academy of Sciences of the United States of America, 101*(1).

Hassa, P. O., Haenni, S. S., Elser, M., & Hottiger, M. O. (2006). Nuclear ADP-ribosylation reactions in mammalian cells: Where are we today and where are we going? *Microbiology and Molecular Biology Reviews, 70*(3).

Hatakeyama, K., Nemoto, Y., Ueda, K., & Hayaishi, O. (1986). PURIFICATION AND CHARACTERIZATION OF POLY(ADP-RIBOSE) GLYCOHYDROLASE - DIFFERENT MODES OF ACTION ON LARGE AND SMALL POLY(ADP-RIBOSE). *Journal of Biological Chemistry, 261*(32).

Hottiger, M. O., Hassa, P. O., Luescher, B., Schueler, H., & Koch-Nolte, F. (2010). Toward a unified nomenclature for mammalian ADP-ribosyltransferases. *Trends in Biochemical Sciences, 35*(4).

Huang, S.-M. A., Mishina, Y. M., Liu, S., Cheung, A., Stegmeier, F., Michaud, G. A., et al. (2009). Tankyrase inhibition stabilizes axin and antagonizes Wnt signalling. *Nature, 461*(7264).

Ikejima, M., & Gill, D. M. (1988). POLY(ADP-RIBOSE) DEGRADATION BY GLYCOHYDROLASE STARTS WITH AN ENDONUCLEOLYTIC INCISION. *Journal of Biological Chemistry, 263*(23).

Ishida, T., & Kinoshita, K. (2008). Prediction of disordered regions in proteins based on the meta approach. *Bioinformatics, 24*(11).

Isogai, S., Kanno, S.-I., Ariyoshi, M., Tochio, H., Ito, Y., Yasui, A., et al. (2010). Solution structure of a zinc-finger domain that binds to poly-ADP-ribose. *Genes to Cells, 15*(2).

Jagtap, P., & Szabo, C. (2005). Poly(ADP-ribose) polymerase and the therapeutic effects of its inhibitors. *Nature Reviews Drug Discovery, 4*(5).

Kang, H. C., Lee, Y.-I., Shin, J.-H., Andrabi, S. A., Chi, Z., Gagne, J.-P., et al. (2011). Iduna is a poly(ADP-ribose) (PAR)-dependent E3 ubiquitin ligase that regulates DNA damage. *Proceedings of the National Academy of Sciences of the United States of America, 108*(34).

Karras, G. I., Kustatscher, G., Buhecha, H. R., Allen, M. D., Pugieux, C., Sait, F., et al. (2005). The macro domain is an ADP-ribose binding module. *Embo Journal, 24*(11).

Katoh, M., & Katoh, M. (2007). Notch signaling in gastrointestinal tract (Review). *International Journal of Oncology, 30*(1).

Kiehlbauch, C. C., Aboulela, N., Jacobson, E. L., Ringer, D. P., & Jacobson, M. K. (1993). HIGH-RESOLUTION FRACTIONATION AND CHARACTERIZATION OF ADP-RIBOSE POLYMERS. *Analytical Biochemistry, 208*(1).

Kim, M. Y., Zhang, T., & Kraus, W. L. (2005). Poly(ADP-ribosyl)ation by PARP-1: 'PAR-laying' NAD(+) into a nuclear signal. *Genes & Development, 19*(17).

Koh, D. W., Dawson, V. L., & Dawson, T. M. (2005). The road to survival goes through PARG. *Cell Cycle, 4*(3).

Koh, D. W., Lawler, A. M., Poitras, M. F., Sasaki, M., Wattler, S., Nehls, M. C., et al. (2004). Failure to degrade poly(ADP-ribose) causes increased sensitivity to cytotoxicity and early embryonic lethality. *Proceedings of the National Academy of Sciences of the United States of America, 101*(51).

Koh, D. W., Patel, C. N., Ramsinghani, S., Slama, J. T., Oliveira, M. A., & Jacobson, M. K. (2003). Identification of an inhibitor binding site of poly(ADP-ribose) glycohydrolase. *Biochemistry, 42*(17).

Krishnakumar, R., & Kraus, W. L. (2010). The PARP Side of the Nucleus: Molecular Actions, Physiological Outcomes, and Clinical Targets. *Molecular Cell, 39*(1).

Kumar, V., & Takahashi, J. S. (2010). PARP around the Clock. *Cell, 142*(6).

Kustatscher, G., Hothorn, M., Pugieux, C., Scheffzek, K., & Ladurner, A. G. (2005). Splicing regulates NAD metabolite binding to histone macroH2A. *Nature Structural & Molecular Biology, 12*(7).

Lautier, D., Lagueux, J., Thibodeau, J., Menard, L., & Poirier, G. G. (1993). MOLECULAR AND BIOCHEMICAL FEATURES OF POLY (ADP-RIBOSE) METABOLISM. *Molecular and Cellular Biochemistry, 122*(2).

Leung, A. K. L., & Sharp, P. A. (2010). MicroRNA Functions in Stress Responses. *Molecular Cell, 40*(2).

Leung, A. K. L., Todorova, T., Ando, Y., & Chang, P. (2012). Poly(ADP-ribose) regulates post-transcriptional gene regulation in the cytoplasm. *Rna Biology, 9*(5).

Leung, A. K. L., Vyas, S., Rood, J. E., Bhutkar, A., Sharp, P. A., & Chang, P. (2011). Poly(ADP-Ribose) Regulates Stress Responses and MicroRNA Activity in the Cytoplasm. *Molecular Cell, 42*(4).

Li, G.-Y., McCulloch, R. D., Fenton, A. L., Cheung, M., Meng, L., Ikura, M., et al. (2010). Structure and identification of ADP-ribose recognition motifs of APLF and role in the DNA damage response. *Proceedings of the National Academy of Sciences of the United States of America, 107*(20).

Lin, W. S., Ame, J. C., AboulEla, N., Jacobson, E. L., & Jacobson, M. K. (1997). Isolation and characterization of the cDNA encoding bovine poly(ADP-ribose) glycohydrolase. *Journal of Biological Chemistry, 272*(18).

Luo, X., & Kraus, W. L. (2012). On PAR with PARP: cellular stress signaling through poly(ADP-ribose) and PARP-1. *Genes & Development, 26*(5).

Malanga, M., & Althaus, F. R. (1994). POLY(ADP-RIBOSE) MOLECULES FORMED DURING DNA-REPAIR IN-VIVO. *Journal of Biological Chemistry, 269*(26).

Masson, M., Niedergang, C., Schreiber, V., Muller, S., Menissier-de Murcia, J., & de Murcia, G. (1998). XRCC1 is specifically associated with poly(ADP-ribose) polymerase and negatively regulates its activity following DNA damage. *Molecular and Cellular Biology, 18*(6).

McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C., & Read, R. J. (2007). Phaser crystallographic software. *Journal of Applied Crystallography, 40*.

Meyer, R. G., Meyer-Ficca, M. L., Whatcott, C. J., Jacobson, E. L., & Jacobson, M. K. (2007). Two small enzyme isoforms mediate mammalian mitochondrial poly(ADP-ribose) glycohydrolase (PARG) activity. *Experimental Cell Research, 313*(13).

Min, J. H., Yang, H. F., Ivan, M., Gertler, F., Kaelin, W. G., & Pavletich, N. P. (2002). Structure of an HIF-1 alpha-pVHL complex: Hydroxyproline recognition in signaling. *Science, 296*(5574).

Min, W., Cortes, U., Herceg, Z., Tong, W.-M., & Wang, Z.-Q. (2010). Deletion of the nuclear isoform of poly(ADP-ribose) glycohydrolase (PARG) reveals its function in DNA repair, genomic stability and tumorigenesis. *Carcinogenesis, 31*(12).

Min, W., & Wang, Z.-Q. (2009). Poly (ADP-ribose) glycohydrolase (PARG) and its therapeutic potential. *Frontiers in Bioscience, 14*.

Miwa, M., Tanaka, M., Matsushi.T, & Sugimura, T. (1974). PURIFICATION AND PROPERTIES OF A GLYCOHYDROLASE FROM CALF THYMUS SPLITTING RIBOSE-RIBOSE LINKAGES OF POLY(ADENOSINE DIPHOSPHATE RIBOSE). *Journal of Biological Chemistry, 249*(11).

Mortusewicz, O., Fouquerel, E., Ame, J.-C., Leonhardt, H., & Schreiber, V. (2011). PARG is recruited to DNA damage sites through poly(ADP-ribose)- and PCNA-dependent mechanisms. *Nucleic Acids Research, 39*(12).

Niere, M., Mashimo, M., Agledal, L., Dolle, C., Kasamatsu, A., Kato, J., et al. (2012). ADP-ribosylhydrolase 3 (ARH3), Not Poly(ADP-ribose) Glycohydrolase (PARG) Isoforms, Is Responsible for Degradation of Mitochondrial Matrix-associated Poly(ADP-ribose). *Journal of Biological Chemistry, 287*(20).

Oberoi, J., Richards, M. W., Crumpler, S., Brown, N., Blagg, J., & Bayliss, R. (2010). Structural Basis of Poly(ADP-ribose) Recognition by the Multizinc Binding Domain of Checkpoint with Forkhead-associated and RING Domains (CHFR). *Journal of Biological Chemistry, 285*(50).

Oka, S., Kato, J., & Moss, J. (2006). Identification and characterization of a mammalian 39-kDa poly(ADP-ribose) glycohydrolase. *Journal of Biological Chemistry, 281*(2).

Okano, S., Lan, L., Caldecott, K. W., Mori, T., & Yasui, R. (2003). Spatial and temporal cellular responses to single-strand breaks in human cells. *Molecular and Cellular Biology, 23*(11).

Okazaki, I. J., & Moss, J. (1996). Mono-ADP-ribosylation: A reversible posttranslational modification of proteins. *Advances in Pharmacology, 35*.

Ono, T., Kasamatsu, A., Oka, S., & Moss, J. (2006). The 39-kDa poly(ADP-ribose) glycohydrolase ARH3 hydrolyzes O-acetyl-ADP-ribose, a product of the Sir2 family of acetyl-histone deacetylases. *Proceedings of the National Academy of Sciences of the United States of America, 103*(45).

Otto, H., Reche, P. A., Bazan, F., Dittmar, K., Haag, F., & Koch-Nolte, F. (2005). In silico characterization of the family of PARP-like poly(ADP-ribosyl) transferases (pARTs). *Bmc Genomics, 6*.

Otwinowski, Z., & Minor, W. (1997). Processing of X-ray diffraction data collected in oscillation mode. *Macromolecular Crystallography, Pt A, 276*.

Patel, C. N., Koh, D. W., Jacobson, M. K., & Oliveira, M. A. (2005). Identification of three critical acidic residues of poly(ADP-ribose) glycohydrolase involved in catalysis: determining the PARG catalytic domain. *Biochemical Journal, 388*.

Pleschke, J. M., Kleczkowska, H. E., Strohm, M., & Althaus, F. R. (2000). Poly(ADP-ribose) binds to specific domains in DNA damage checkpoint proteins. *Journal of Biological Chemistry, 275*(52).

Satoh, M. S., & Lindahl, T. (1992). ROLE OF POLY(ADP-RIBOSE) FORMATION IN DNA-REPAIR. *Nature, 356*(6367).

Schreiber, V., Dantzer, F., Ame, J.-C., & de Murcia, G. (2006). Poly(ADP-ribose): novel functions for an old molecule. *Nature Reviews Molecular Cell Biology, 7*(7).

Scolnick, D. M., & Halazonetis, T. D. (2000). Chfr defines a mitotic stress checkpoint that delays entry into metaphase. *Nature, 406*(6794).

Slade, D., Dunstan, M. S., Barkauskaite, E., Weston, R., Lafite, P., Dixon, N., et al. (2011). The structure and catalytic mechanism of a poly(ADP-ribose) glycohydrolase. *Nature, 477*(7366).

Tan, X., & Zheng, N. (2009). Hormone signaling through protein destruction: a lesson from plants. *American Journal of Physiology-Endocrinology and Metabolism, 296*(2).

Telli, M. L. (2011). PARP inhibitors in cancer: moving beyond BRCA. *Lancet Oncology, 12*(9).

Underhill, C., Toulmonde, M., & Bonnefoi, H. (2011). A review of PARP inhibitors: from bench to bedside. *Annals of Oncology, 22*(2).

Wang, X. C., Yang, C. L., Chai, J. J., Shi, Y. G., & Xue, D. (2002). Mechanisms of AIF-mediated apoptotic DNA degradation in Caenorhabditis elegans. *Science, 298*(5598).

Wang, Y., Kim, N. S., Haince, J.-F., Kang, H. C., David, K. K., Andrabi, S. A., et al. (2011). Poly(ADP-Ribose) (PAR) Binding to Apoptosis-Inducing Factor Is Critical for PAR Polymerase-1-Dependent Cell Death (Parthanatos). *Science Signaling, 4*(167).

Willems, A. R., Goh, T., Taylor, L., Chernushevich, I., Shevchenko, A., & Tyers, M. (1999). SCF ubiquitin protein ligases and phosphorylation-dependent proteolysis. *Philosophical Transactions of the Royal Society of London Series B-Biological Sciences, 354*(1389).

Woodhouse, B. C., & Dianov, G. L. (2008). Poly ADP-ribose polymerase-1: An international molecule of mystery. *DNA Repair, 7*(7).

Yu, S. W., Wang, H. M., Poitras, M. F., Coombs, C., Bowers, W. J., Federoff, H. J., et al. (2002). Mediation of poly(ADP-ribose) polymerase-1-dependent cell death by apoptosis-inducing factor. *Science, 297*(5579).

Zhang, Y., Liu, S., Mickanin, C., Feng, Y., Charlat, O., Michaud, G. A., et al. (2011). RNF146 is a poly(ADP-ribose)-directed E3 ligase that regulates axin degradation and Wnt signalling. *Nature Cell Biology, 13*(5).

Zweifel, M. E., Leahy, D. J., & Barrick, D. (2005). Structure and Notch receptor binding of the tandem WWE domain of Deltex. *Structure, 13*(11).

VITA

Zhizhi Wang was born in Fenghua, Zhejiang, China in Sept. 1986 as the only child in his family. He received his Bachelor of Science degree in Biological Science at the Peking University, Beijing, China in 2007. Afterwards, he joined Biological Physics, Structure and Design Program and Department of Biological Structure at the University of Washington, where he earned a Doctor of Philosophy degree in Biological Structure in 2012.