

CHEMICAL GENETIC AND GENOMIC DISCOVERY OF PARP-1-DEPENDENT
MECHANISMS OF TRANSCRIPTIONAL REGULATION

APPROVED BY SUPERVISORY COMMITTEE

W. Lee Kraus, Ph.D.

Luke Rice, Ph.D.

Michael Rosen, Ph.D.

Bing Li, Ph.D.

DEDICATION

I would like to thank my family, friends, and especially Lindsey
for their patience, confidence, and patience.

CHEMICAL GENETIC AND GENOMIC DISCOVERY OF PARP-1-DEPENDENT
MECHANISMS OF TRANSCRIPTIONAL REGULATION

by

BRYAN ANDREW GIBSON

DISSERTATION

Presented to the Faculty of the Graduate School of Biomedical Sciences

The University of Texas Southwestern Medical Center at Dallas

In Partial Fulfillment of the Requirements

For the Degree of

DOCTOR OF PHILOSOPHY

The University of Texas Southwestern Medical Center at Dallas

Dallas, Texas

December 2015

Copyright

by

Bryan Andrew Gibson, 2015

All Rights Reserved

CHEMICAL GENETIC AND GENOMIC DISCOVERY OF PARP-1-DEPENDENT
MECHANISMS OF TRANSCRIPTIONAL REGULATION

Publication No. _____

Bryan Andrew Gibson, Ph.D.

The University of Texas Southwestern Medical Center at Dallas, 2015

W. Lee Kraus, Ph.D.

Poly(ADP-ribose) polymerases, or PARPs, are a family of enzymes that modulate diverse biological processes through covalent transfer of ADP-ribose from NAD^+ onto target proteins. These targets of post-translational modification as well as the genomic targets, or binding sites, to which the nuclear PARP family members localize reflects the molecular biology and cellular function of an individual PARP. Given that PARP proteins are implicated in the most devastating of human disease, including cancer, heart disease, stroke, and neuropathology, a deeper understanding of their proteomic and genomic targets may guide effective therapeutic intervention within this family of enzymes. Here, I report the development of new

methodologies in an effort to identify the targets of PARPs. To identify the sites of ADP-ribosylation, I've developed a simple and robust analog-sensitive approach for PARPs, which allows PARP-specific clickable ADP-ribosylation. Using this approach, I have mapped hundreds of protein targets and sites of ADP-ribosylation for PARPs 1, 2, and 3. I found that PARP-1 ADP-ribosylates and inhibits RNA-binding by NELF, a protein complex that regulates promoter-proximal pausing by RNA polymerase II. I have used this analog-sensitive approach to discover genomic sites of PARP-1-mediated ADP-ribosylation and their relationship to paused RNA Polymerase II. Furthermore, I have found that knockdown of PARP-1 resulted in the accumulation of paused RNA Polymerase II, implicating PARP-1 in RNA Polymerase II elongation through ADP-ribosylation and inhibition of NELF. PARP-1 is activated by both DNA lesions and nucleosomally wrapped DNA, the latter of which is a likely substrate for PARP-1 in its role as a regulator of transcription. In order to identify the nucleosomes that PARP-1 binds, I've developed an MNase ChIP-seq method using crosslinked cells called XL-MNase ChIP-seq. This technical advance has revealed that PARP-1 binds a nuclease-sensitive nucleosome that spans the "nucleosome free" region of regulatory elements of the genome, implicating PARP-1 as a modulator of genomic access at regulatory regions across the genome. The utility of these new methodologies is evident in their use in the discovery of new biological roles for PARP-1 in transcriptional regulation.

TABLE OF CONTENTS

PRIOR PUBLICATIONS	xii
LIST OF FIGURES	xiii
LIST OF TABLES	xvii
LIST OF DEFINITIONS	xviii
CHAPTER ONE. SMALL MOLECULES BIG EFFECTS: A ROLE FOR CHROMATIN-LOCALIZED METABOLITE BIOSYNTHESIS IN GENE REGULATION	1
CHAPTER TWO. NEW INSIGHTS INTO THE MOLECULAR AND CELLULAR FUNCTIONS OF POLY(ADP-RIBOSE) AND PARPS	8
PREFACE	8
INTRODUCTION	8
REGULATED FORMATION OF PAR POLYMERS	9
THE PARP FAMILY OF PROTEINS	11
CATALYSIS OF PARYLATION	14
DEGRADATION AND RECYCLING OF PAR	15
OTHER PAR-DEGRADING ENZYMES	17
MECHANISMS OF PARP ACTIVATION	17
MECHANISMS OF TARGET SELECTION	20
COOPERATION BETWEEN PARPS	22
CELLULAR FUNCTIONS OF PARYLATION	23

INHIBITION OF PROTEIN-PROTEIN AND PROTEIN-NUCLEIC ACID	
INTERACTIONS	24
PAR-DEPENDENT PROTEIN LOCALIZATION, INTERACTION SCAFFOLDS,	
AND PHASE SEPARATION	27
PAR-DEPENDENT UBIQUITYLATION	29
REGULATED FORMATION AND DEGRADATION OF PAR	31
ROLES OF PAR-BINDING MODULES	32
PAR-BINDING MOTIFS	32
PAR-BINDING ZINC FINGER	33
MACRODOMAIN FOLDS	35
WWE DOMAINS	36
PERSPECTIVES AND CONCLUSIONS	37
CHAPTER THREE. CHEMICAL GENETIC DISCOVERY OF PARP TARGETS REVEALS	
A ROLE FOR PARP-1 IN TRANSCRIPTIONAL ELONGATION	41
SUMMARY	41
INTRODUCTION	41
RESULTS	42
MATERIALS AND METHODS	68
ANTIBODIES	68
ALIGNMENT OF NAD ⁺ INTO THE ACTIVE SITE OF THE PARP-1 CATALYTIC	
DOMAIN	69
NAD ⁺ ANALOGS	69

SYNTHESIS OF NAD ⁺ ANALOGS	70
MOLECULAR BIOLOGY AND CLONING	73
EXPRESSION AND PURIFICATION OF RECOMBINANT PROTEINS	75
CELL CULTURE AND GENERATION OF KNOCKDOWN AND OVEREXPRESSION CELL LINES	79
PARP-1 ENZYME KINETICS ASSAYS	80
IN VITRO PARP AUTOMODIFICATION REACTIONS	82
IN VITRO NERF-E ADP-RIBOSYLATION REACTIONS	83
NERF-E/TAR RNA BINDING ASSAYS	83
PREPARATION AND LC-MS/MS ANALYSIS OF 8-BU(3-YNE)T-ADP- RIBOSYLATED HELA CELL NUCLEAR EXTRACT PROTEINS	84
ANALYSIS OF LC-MS/MS DATA	87
8-BU(3-YNE)T-ADP-RIBOSYLATION REACTIONS IN INTACT NUCLEI ...	88
CONFOCAL MICROSCOPY	89
CLICK-CHIP-SEQ	90
ANALYSIS OF CLICK-CHIP-SEQ AND CHIP-SEQ DATA	92
ANALYSIS OF CHIP-CHIP DATA	94
PREPARATION OF GRO-SEQ LIBRARIES	94
ANALYSIS OF GRO-SEQ DATA	95
GENOMIC DATASETS	97
OLIGONUCLEOTIDE SEQUENCES	97
CONCLUSIONS	98

CHAPTER FOUR. PARP-1 BINDS A NUCLEASE-SENSITIVE NUCLEOSOME

SPANNING REGULATORY REGIONS OF THE GENOME	101
SUMMARY	101
INTRODUCTION	101
RESULTS	106
PARP-1 BINDING AT ACTIVE REGIONS OF THE GENOME	106
DEVELOPMENT OF XL-MNASE CHIP-SEQ	107
CHARACTERIZATION OF PROMOTER NUCLEOSOME	
ARCHITECTURE	112
PARP-1 CO-LOCALIZATION WITH A NUCLEASE-SENSITIVE	
NUCLEOSOME	116
MATERIALS AND METHODS	121
ANTIBODIES	121
CELL CULTURE AND TNF α STIMULATION OF AC16 CELLS	122
PREPARATION OF CHIP-SEQ LIBRARIES	122
PREPARATION OF XL-MNASE CHIP-SEQ LIBRARIES	123
ANALYSIS OF CHIP-SEQ AND XL-MNASE CHIP-SEQ DATA	124
ANALYSIS OF GRO-SEQ DATA	126
GENOMIC DATASETS	127
CONCLUSIONS	128
CHAPTER FIVE. DESIGN AND USE OF SYNTHETIC ANTIBODY-LIKE ADP-	
RIBOSE DETECTION REAGENTS	131

SUMMARY	131
INTRODUCTION	131
RESULTS	132
MATERIALS AND METHODS	137
CLONING AND MOLECULAR BIOLOGY	140
PROTEIN EXPRESSION	141
PROTEIN PURIFICATION	142
IN VITRO PARYLATION OF MODEL SUBSTRATES	144
ADP-RIBOSE DETECTION USING ARBD-FC FUSION PROTEINS	144
CONCLUSIONS	145

PRIOR PUBLICATIONS

Gibson, B.A., Kraus, W.L. (2012) **New insights into the molecular and cellular functions of poly(ADP-ribose) and PARPs.** *Nature Reviews Molecular and Cellular Biology* 3:411-24.

Gibson, B.A., Kraus, W.L. (2011) **Small molecules, big effects: a role for chromatin-localized metabolite biosynthesis in gene regulation.** *Molecular Cell* 41:497-9.

Liu, J., Doty, T., Gibson, B., Heyer, W.D. (2010) **Human BRCA2 protein promotes RAD51 filament formation on RPA-covered single-stranded DNA.** *Nature Structural and Molecular Biology* 17:1260-2

LIST OF FIGURES

FIGURE ONE. Chromatin-localized biosynthesis of metabolites for chromatin- and transcription-modifying enzymes	150
FIGURE TWO. Structure of PAR and mechanisms of PAR catalysis and glycohydrolysis	150
FIGURE THREE. PARP-1 activation and target selection	150
FIGURE FOUR. PAR-dependent ubiquitylation and the human disease Cherubism	150
FIGURE FIVE. Cellular functions of PAR and the PAR-dependent recruitment of factors to sites of DNA damage	150
FIGURE SIX. Recognition strategies and functions of PAR-binding modules	150
FIGURE SEVEN. PARP inhibitors and their potential as therapeutic agents	150
FIGURE EIGHT. Structure-based engineering of an NAD ⁺ analog-sensitive PARP-1 (<i>as</i> PARP-1) mutant	150
FIGURE NINE. Structure-based alignment of NAD ⁺ in the catalytic domain of PARP-1	150
FIGURE TEN. Screening for an NAD ⁺ analog-sensitive PARP-1 mutant using an activity-based screen	150
FIGURE ELEVEN. Enzyme kinetics of <i>as</i> PARP-1 with 8-Bu(3-yne)T-NAD ⁺	150
FIGURE TWELVE. Transfer of NAD ⁺ analog sensitivity to other PARP family members	150
FIGURE THIRTEEN. Using analog-sensitive PARP-1 mutants to unambiguously identify the ADP-ribosylation targets of DNA-dependent PARPs	150

FIGURE FOURTEEN. Overlap of PARP targets with targets reported in the literature ..	150
FIGURE FIFTEEN. P-TEFb-dependent ADP-ribosylation of NELF by PARP-1	150
FIGURE SIXTEEN. Other post-translational modifications associated with PARP target modification sites	150
FIGURE SEVENTEEN. Functional links between PARP-1-catalyzed ADP- ribosylation, NELF binding, and RNA polymerase II pausing genome-wide	150
FIGURE EIGHTEEN. Development of Click-ChIP-seq, a nuclei-based assay using <i>as</i> PARP-1 technology to identify sites of PARP-1-mediated ADP-ribosylation genome-wide	150
FIGURE NINETEEN. Altered promoter-proximal RNA polymerase II pausing at specific genes in PARP-1-deficient MCF-7 cells	150
FIGURE TWENTY. PARP-1 modulates RNA polymerase II pausing release genome-wide	150
FIGURE TWENTY ONE. Model for the role of PARP-1 in the regulation of RNA polymerase pausing and release through NELF	150
FIGURE TWENTY TWO. PARP-1 binds open chromatin regions at promoters, enhancers, and insulators	150
FIGURE TWENTY THREE. Overlap of active histone marks with PARP-1 and CTCF binding	150
FIGURE TWENTY FOUR. Development of XL-MNase ChIP-seq	150
FIGURE TWENTY FIVE. NucPosSimulator-based quantification of nucleosome parameters from XL-MNase ChIP-seq data	150

FIGURE TWENTY SIX. A poorly positioned, lowly occupied, and MNase-sensitive nucleosome resides at the “nucleosome free” region of active promoters	150
FIGURE TWENTY SEVEN. PARP-1 binding is correlated with MNase-sensitive nucleosomes within the “nucleosome free” regions of active genes	150
FIGURE TWENTY EIGHT. PARP-1 binding and MNase-sensitive nucleosomes occupy NFκB p65 and CTCF binding site “nucleosome free” regions	150
FIGURE TWENTY NINE. Absence of PARP-1 binding and MNase sensitivity at NFκB p65 binding sites without TNFα treatment	150
FIGURE THIRTY. The chromatin landscape of MNase-sensitive nucleosomes at the “nucleosome free” region of NFκB p65 binding sites	150
FIGURE THIRTY ONE. Characterization of PARP-1 binding to MNase-sensitive nucleosomes at CTCF binding sites	150
FIGURE THIRTY TWO. Model for PARP-1 binding to an unstable nucleosome at regulatory regions throughout the genome	150
FIGURE THIRTY THREE. Schematic representation of recombinant proteins in which natural ADP-ribose binding domains (ARBDs) are fused to mammalian immunoglobulin G (IgG) Fc region	150
FIGURE THIRTY FOUR. Bacterial plasmids for expressing ADP-Ribose binding domain-Fc fusion proteins	150
FIGURE THIRTY FIVE. Analysis of purified ADP-ribose binding domain-Fc fusion proteins by polyacrylamide gel electrophoresis	150

FIGURE THIRTY SIX. Dot blot assays of mono, oligo, and poly(ADP-ribose) binding by ADP-ribose binding domain-Fc fusion proteins	150
FIGURE THIRTY SEVEN. Western blot assays of mono, oligo, and poly(ADP- ribose) binding by ADP-ribose binding domain-Fc fusion proteins	150

LIST OF TABLES

TABLE ONE. Some examples of metabolites, their synthases, and effectors involved in chromatin-dependent gene regulation	150
TABLE TWO. The organization and enzymatic activities of PARP family members	150

LIST OF DEFINITIONS

PARP – Acronym for poly(ADP-ribose) polymerase, an enzyme that uses NAD^+ to covalently transfers ADP-ribose onto target proteins using NAD^+

NAD^+ – Acronym for the coenzyme β -nicotinamide adenine dinucleotide

ADP-ribose – Metabolite, derived from NAD^+ , and the base unit of the PARylation post-translational modification

ADP-ribosylation – A type of post-translational modification involving the transfer of NAD^+ onto a target protein

PARylation – Written fully as poly(ADP-ribosyl)ation, this is a type of ADP-ribosylation reaction in which two or more units of ADP-ribose are linked onto a single site of post-translational modification through an $\alpha(1 \rightarrow 2)$ O-glycosidic bond

PAR – Acronym for poly(ADP-ribose)

PARP inhibitor – A small molecule which inhibits PARP proteins' ADP-ribosylation activity

PARG – Acronym for poly(ADP-ribose) glycohydrolase, a poly(ADP-ribose) degrading enzyme

***iso*-ADP-ribose** – The ADP-ribose unit surrounding the $\alpha(1 \rightarrow 2)$ O-glycosidic bond of poly(ADP-ribose), notably recognized by the WWE domain of RNF146

NAD^+ analog – A small molecule similar, or analogous, to NAD^+

Automodification – A type of post-translational modification where a post-translational modifying enzyme modifies itself

Post-translational modification – A modification to a protein's primary sequence not occurring during translation

Ubiquitylation – A type of post-translational modification catalyzed by an enzyme class called E3 ligases, where the protein ubiquitin is covalently added to another protein's lysine residue often triggering proteasome-mediated degradation of the protein that is modified

SUMOylation – A type of post-translational modification where the protein SUMO is covalently added to another protein.

Phosphorylation – A type of post-translational modification catalyzed by an enzyme class called kinases, where a phosphate group is added to serine, threonine, or tyrosine residues of a protein

Acetylation – A type of post-translational modification catalyzed by an enzyme class called acetyltransferases, where an amine group, most notably on lysine amino acid side chains, with an acetyl group from Acetyl-CoA

Methylation – A type of modification catalyzed by an enzyme class called methyltransferases, resulting in methylation of amino acids on proteins or cytosines on DNA

Proteasome – A Megadalton-size protein complex that degrades proteins following their damage or signaling for destruction

Stress Granule – A sub-organellar structure within the cytosol formed during conditions of stress composed of stalled translation pre-initiation complexes

Cajal Body – A sub-organellar structure within the nucleus

Chromatin – Term referring to the nucleoprotein structure of packaged DNA in eukaryotic cells, the basic unit of which is the nucleosome

Nucleosome – The basic unit of chromatin and consists of ~147 bp wrapped ~1.65 times around a histone octamer core composed of 2 copies each of histones H3, H4, H2A, and H2B

Histone – A highly basic protein and core chromatin component that includes histones H3, H4, H2A, H2B, their variants, and the linker histone H1 in human cells

Histone Variant – A histone protein incorporated into nucleosomes similarly to canonical histone H3, H4, H2A, and H2B, but differs in its sequence and structure

Genomics – A methodological approach aimed at studying the entirety of the genome

Proteomics – A methodological approach aimed at studying the entirety of the proteome

Click chemistry – A type of orthogonal chemistry that results in predictable joining of discrete small units, most notably the azide alkyne [3+2] Huisgen 1,3-dipolar cycloaddition

Analog sensitivity – A type of gain-of-function protein mutant where a large “gatekeeper” residue is mutated to a smaller amino acid to facilitate binding to an analog of either the natural substrate or non-specific protein inhibitor modified with a bulky R group to render it sterically unusable with the wild-type protein

Chemical genetics – A discipline combining the use of chemical and genetic approaches to screen for biological outcomes

LC-MS/MS – Acronym for liquid chromatography with mass spectrometry, with MS/MS referring to the use of tandem mass spectrometers

ChIP – Abbreviation for chromatin immunoprecipitation, a technique used for the identification of protein binding sites in the genome

ChIP-seq – Abbreviation for chromatin immunoprecipitation coupled with deep sequencing, a technique for the identification of protein binding sites throughout the genome

GRO-seq – Abbreviation for global run-on coupled with deep sequencing, a technique for the identification of engaged RNA polymerases throughout the genome

Click-ChIP-seq – Abbreviation for clicked chromatin interaction and precipitation coupled with deep sequencing, a technique for the identification of sites of PARP-specific ADP-ribosylation throughout the genome

XL-MNase ChIP-seq – Abbreviation for Micrococcal Nuclease and chromatin immunoprecipitation coupled with deep sequencing from crosslinked cells, a technique for the identification of protein binding sites throughout the genome at nucleosome resolution

ChIP-chip – Abbreviation for chromatin immunoprecipitation coupled with microarray chip, a technique for the identification of protein binding sites throughout the genome

PCR – Acronym for polymerase chain reaction, a technique for sequence specific DNA amplification

EMSA – Acronym for electrophoretic mobility shift assay, a technique for the identification and quantification of protein nucleic acid interactions

Deep sequencing – Also known as high throughput or next-gen sequencing, a technology of massively parallel sequencing capable of generating millions of individual sequence “reads” in tandem from libraries of prepared DNA material

Confocal microscopy – An optical imaging technique where use of sample detection is constrained to a confocal plane to eliminate out-of-focus light

Paired-end sequencing – A type of deep sequencing where both ends of a DNA fragment are sequenced

Monoclonal antibody – A clonal type of antibody produced to recognize a single epitope of an antigen

MNase – Abbreviation of micrococcal nuclease, a nuclease often used to digest chromatin to mononucleosomes

Gene ontology – A computational method for the identification of shared biology among a list of genes

Metagene – A plot of the average in from a genomic experiment from many individual genomic loci

Heatmap – A plot where numerical value, in genomics this is often read density, is indicated by intensity or type of color

ChIP-seq peak – A genomic feature calculated by statistical analysis of genomic signal accumulation at an individual genomic locus, which indicates protein binding

Sequence alignment – The matching of one sequence to another, in genomics this refers to matching sequence reads from a deep sequencing experiment to a reference genome

RNA Polymerase II pausing – A regulatory step in the RNA Polymerase II transcription cycle where RNA Polymerase is held ~30-60 base pairs into a gene following initiation

Cardiomyocytes – The muscle cells of the heart

Histone mark – A term referring broadly to post-translational modifications of histone proteins

Nucleosome position – Within a single cell this refers to the exact position on the genome where a nucleosome is located, and in genomic experiments from a population of cells this term refers to the consistency of positioning of at a given genomic locus within that population

Nucleosome occupancy – Within a single cell this refers to whether or not a genomic locus is wrapped around a nucleosome, and in genomic experiments from a population of cells this term refers to the overall signal of nucleosome density resulting from the population

DHS – Acronym for DNase I hypersensitivity site, which is a region of the genome or on a defined template that is especially sensitive to nuclease digestion by DNase I, which acts as a proxy for the relative accessibility of the DNA.

NFR – Acronym for nucleosome free region, which refers to a nucleosome depleted position spanning transcription factor binding sites and transcription start sites of active genes

Regulatory element – An element within the genome that regulates genomic activity

Promoter – The region surrounding a gene that promotes its transcription

Enhancer – An orientation-independent element within the genome that stimulates transcription

Insulator – An element of the genome that insulates one region of the genome from the other

TSS – Acronym for transcription start site

TFBS – Acronym for transcription factor binding site

Cell death – Generally refers to the cellular process of coordinated and regulated death

Transcription – The cellular process of single stranded RNA synthesis from a DNA template

DNA damage – A process most often extracellular which results in damage to the integrity of DNA within the genome, of which there are many types, including single- and double-strand breaks, abasic sites, base oxidation, and thymine dimer formation

DNA repair – The active process of repairing damaged DNA

Inflammation – A transcriptional response that arises from many types of cellular stress

Circadian rhythm – The cellular response to cyclic environmental inputs, chief among them a 24-hour cycle of light and dark

CHAPTER ONE

SMALL MOLECULES, BIG EFFECTS: A ROLE FOR CHROMATIN-LOCALIZED METABOLITE BIOSYNTHESIS IN GENE REGULATION

Regulated transcription of specific gene sets in response to cellular signals requires an extensive repertoire of highly coordinated molecular events, many of which are directed at the chromatin template. These molecular events tune the promoter chromatin environment to the required transcriptional outcome (e.g., permissive environment for enhanced transcription, or repressive environment for reduced transcription). These processes involve a wide variety of chromatin- and transcription-modulating enzymes that require small molecule cofactors or metabolites to support their catalytic activities. For example, (1) protein acetyltransferases, such as p300/CBP, PCAF/GCN5, and Tip60, require acetyl-CoA as a donor of acetyl groups in acetylation reactions; (2) poly(ADP-ribose) polymerases, such as PARP-1, require oxidized nicotinamide adenine dinucleotide (NAD^+), as a donor of ADP-ribose units in poly(ADP-ribosyl)ation reactions; and (3) sirtuin family protein deacetylases, such as Sirt1, require NAD^+ as a donor of ADP-ribose units to accept acetyl groups in deacetylation reactions ([Table 1](#)). The source of these metabolites in the nucleus and their use by chromatin-modifying enzymes has been receiving increased attention in the recent literature. In Katoh et al. (2011) the authors show how the production of *S*-adenosylmethionine (SAM) by promoter-bound MATII α , a methionine adenosyltransferase (MAT) isozyme, supports the activity of methyltransferases involved in transcriptional regulation (*1*).

Metabolite	Synthase	Effector	Outcome
Acetyl-CoA	Acetyl-CoA synthetase	Acetyltransferases	Histone acetylation
	ATP-citrate lyase	Acetyltransferases	Histone acetylation
ATP	ATP synthase	Kinases	Phosphorylation
		Chromatin remodeling complexes	Chromatin remodeling
FAD ⁺	FMN adenylyltransferase	LSD family demethylases	Histone demethylation
α -ketoglutarate	Isocitrate dehydrogenase	Jmj family demethylases	Histone demethylation
NAD ⁺ (and related cofactors)	NMNAT-1 ^a (and NAD ⁺ cofactor kinases and reductases)	Sirtuins (<i>NAD</i> ⁺)	Histone deacetylation
		PARPs (<i>NAD</i> ⁺)	ADP-ribosylation
		CtBP (<i>NAD</i> ⁺ / <i>NADH</i>)	Transcriptional corepression
		Clock:BMAL1 [<i>NAD(P)</i> ⁺ / <i>NAD(P)H</i>]	Modulation of DNA binding
		NPAS2:BMAL1 [<i>NAD(P)</i> ⁺ / <i>NAD(P)H</i>]	Modulation of DNA binding
O-acetyl-ADP-ribose	Sirtuins	Sir2/3/4	Regulation of enzyme activity
		MacroH2A1.1	Regulation of protein-protein interactions
S-adenosylmethionine	MatI α ^a	Protein methyltransferases	Histone methylation
		DNA methyltransferases	DNA methylation

Table 1. Some examples of Metabolites, Their Synthases, and Effectors Involved in Chromatin-Dependent Gene Regulation

Of the metabolite producers (“synthases”) listed here only MATI α (1) and NMNAT-1 (2) have been shown to be recruited to chromatin. The extent to which other metabolite-producing enzymes are recruited to target gene promoters has yet to be determined.

where chromatin-localized biosynthesis of metabolites can modulate the expression of nearby genes by providing metabolites for use by epigenetic effectors.

In their studies of gene regulation by MafK, a Maf family transcription factor, Katoh et al. (2011) used a proteomic approach to identify 289 MafK-interacting proteins. These interactors included a variety of chromatin- and transcription-related factors, such as the Maf corepressor Bach1, as well as the SAM synthetase MATII α , which catalyzes the production of SAM by joining methionine and ATP. Interestingly, a previous genetic screen in *Drosophila* identified a mutation in the gene encoding a fly MAT ortholog (*SamS*), called Su(z)5, as a suppressor of position effect variegation, implicating SAM biosynthesis in the maintenance of heterochromatin (3). Subsequent identification of 127 MATII α -interacting proteins revealed 39 proteins that overlapped with the set of MafK-interacting proteins, including components of the Swi/Snf, NuRD, PARP-1, and PcG complexes, which mediate nucleosome remodeling, histone modification, and the formation of heterochromatin. Based on these results, the authors investigated a possible MafK- and MatII α -dependent repressive chromatin network. To understand how these various factors might function together in gene regulation, the authors used the MafK-regulated and Maf recognition element (MARE)-containing heme oxygenase-1 (*HO-1*) gene as a model. Using biotinylation-coupled chromatin immunoprecipitation (ChIP), the authors found that both MafK and MatII α bound to the *HO-1* locus, with MafK binding more extensively to the MARE regions and MATII α binding more broadly, even into a neighboring gene. Upon robust stress-induced expression of the *HO-1* gene in response to diethyl maleate (DEM), loss of MATII α binding, but not MafK binding, from the *HO-1* locus was observed, suggesting that MATII α might play a critical role in repressing the expression of *HO-1* in the absence of an activating signal. Using RNAi-mediated knockdown of components of the individual complexes comprising this repressive module, the authors found that Swi/Snf, NuRD,

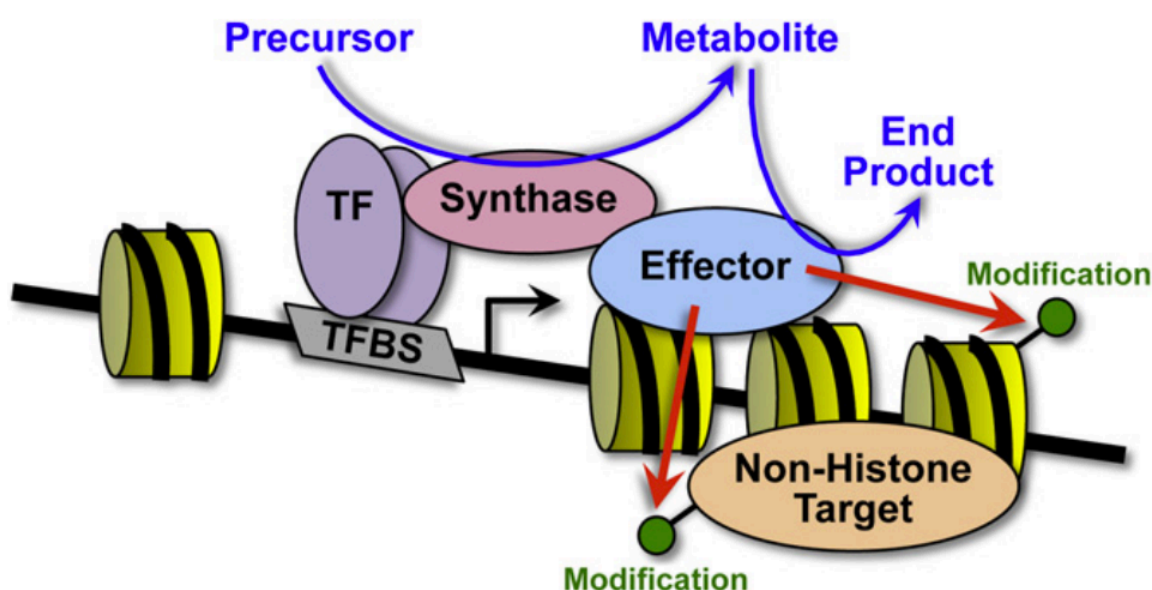


Figure 1. Chromatin-localized Biosynthesis of Metabolites Can Provide Chromatin- and Transcription-Modifying Enzymes with the Cofactors they Need to Catalyze Their Essential Chemical Reactions.

While bound to chromatin, metabolite producers (“synthases”; e.g., MATII α) can synthesize small molecule cofactors (e.g., SAM) that are used by chromatin- and transcription-modifying enzymes (“effectors”, e.g., methyltransferases) to chemically modify histones and nonhistone proteins to alter gene expression. The synthase may be recruited by a DNA-bound transcription factor (“TF”, e.g., MafK) or the effector. The specific examples noted here are from (1).

MATII α partner MATII β , and PARP-1 each play a role in *HO-1* gene repression. The critical observation made by the authors, however, was that introduction of a single point mutation in MATII α , which abolishes its SAM synthetase activity, further enhanced *HO-1* transcription upon DEM treatment. These results implicate MATII α enzymatic activity - its ability to produce SAM - in MATII α -dependent repression of the *HO-1* gene.

Since SAM is used by both histone and DNA methyltransferases, the authors examined which enzymes might be activated by SAM production to promote gene repression at the *HO-1* locus. By using MATII α knockdown in combination with bisulfite DNA sequencing or histone methylation ChIP, they excluded DNA methyltransferases but found evidence for histone methyltransferases as mediators of the SAM-dependent effects. In particular, H3K4me2 and H3K9me2 levels decreased upon MATII α knockdown. The authors did not identify the specific histone methyltransferases, but they did find a number of such enzymes in their MATII α -interacting protein set (e.g., G9a, Ehmt, and MLL1/KMT2A). Together, these results fit a model in which chromatin-localized biosynthesis of metabolites can provide chromatin-modifying enzymes with the cofactors they need to catalyze their essential chemical reactions (**Figure 1**).

This model represents an emerging paradigm in gene regulation (**Table 1**). For example, in mammalian cells the nuclear NAD⁺ synthase nicotinamide mononucleotide adenylyltransferase-1 (NMNAT-1) is recruited by Sirt1 to promoters, where it produces NAD⁺ to support deacetylation of H4K16 by Sirt1 (2). In yeast, the Sirt1 homolog Sir2 is allosterically regulated by O-acetyl-ADP-ribose, a byproduct that it produces during deacetylation reactions (4). O-acetyl-ADP-ribose allosterically regulates the quaternary association of Sir3 and Sir2/4, which accelerates Sir2 activity on acetylhistone substrates. More broadly, ribonucleotide

reductase, which synthesizes deoxyribonucleotides from ribonucleotides, is rapidly recruited by Tip60 to DNA damage foci to assist with the synthesis of new DNA, most likely due to low pools of dNTPs during quiescence (5). Beyond these specific cases of chromatin-bound biosynthetic enzymes, metabolite production in general has been shown to modulate the actions of chromatin-modifying proteins to alter gene expression. For example, histone acetyltransferases are sensitive to the availability of acetyl-CoA, produced by acetyl-CoA synthetase in yeast (6) and ATP-citrate lyase in mammals (7). Moreover, the corepressor function of carboxyl-terminal binding protein (CtBP) and the DNA binding activities of the circadian transcription factors Clock and BMAL are sensitive to the redox state of NAD^+ cofactors (8-10). Whether the metabolites in these latter cases are produced locally by chromatin-bound enzymes has yet to be determined.

The results of Katoh et al. (2011) and related results in the literature raise a number of questions. First, why should metabolite production occur by a chromatin-bound enzyme as opposed to a free enzyme in the nucleoplasm? Diffusion in the nucleus is rapid and unlikely to limit the availability of metabolites for use at specific loci. Thus, stable increases in the local free pools of metabolites are unlikely to be required or even occur. Perhaps there is a more intimate and functionally important association of the metabolite producers with the metabolite consumers. Such interactions could allow for (1) substrate channeling between the producer and consumer (11), which would shield the metabolite from other effectors that might use it for opposing outcomes (e.g., H3K4me3 versus H3K27me3) or (2) allosteric interactions, which can stimulate the activity of either, or both, enzymes. Also, a requirement for both the producer and consumer to be present in the same place at the same time adds precision to the regulatory

process. Second, how generally do metabolite producers and users associate across the genome? The answer to this question will require the application and integration of genomic, proteomic, and metabolomic approaches to the enzymes pairs described herein. Finally, what other metabolite-producing enzymes may localize to the nucleus to support the functions of transcription- and chromatin-related proteins? As more sensitive proteomic approaches yield more examples, the challenge will be to identify the functional relevance, including the specific proteins that use the metabolites, their localization across the genome, and the regulatory outcomes that they control.

CHAPTER TWO

Review of the Literature

NEW INSIGHTS INTO THE MOLECULAR AND CELLULAR FUNCTIONS OF POLY(ADP-RIBOSE) AND PARPS

Preface

Poly(ADP-ribose) polymerases (PARPs) are enzymes that transfer ADP-ribose groups to target proteins and thereby affect a wide variety of nuclear and cytoplasmic processes. The activity of PARP family members, such as PARP-1 and PARP-2, are tied to cellular signalling pathways and through PARylation they ultimately promote changes in: gene expression; RNA and protein abundance; and the location and activity of proteins mediating signalling responses. PARPs act in a complex response network driven by the cellular, molecular and chemical biology of poly(ADP-ribose) (PAR). This PAR-dependent response network is critical for a broad array of physiological and pathological responses and, thus, is a good target for chemical therapeutics for several diseases.

Introduction

Rapid and appropriate cellular responses to external stimuli are critical for adaptability and maintenance of cellular and organismal viability, especially in challenging environmental conditions. Post-translational modifications allow rapid transduction of extra- and intracellular signals into biological outcomes, through targeted control of protein localization and activity, RNA and protein levels and changes in gene expression. Poly(ADP-ribose) polymerase (PARP) proteins comprise a group of ADP-ribosyl transferase enzymes, which transfer negatively

charged ADP-ribose groups from donor NAD^+ molecules onto their target proteins post-translationally. Through poly(ADP-ribosyl)ation (PARylation) of target proteins, PARPs control a wide array of cellular processes, such as DNA repair, transcriptional regulation, RNA interference, mitochondrial function, formation of subnuclear bodies, and cell division. PARylation is particularly prevalent during stress responses that require rapid adaptation.

So far, the functions and mechanisms of PARylation have been best characterized for a few PARPs, including PARP-1, PARP-2, and the tankyrases. Recent reviews on PARPs have emphasized their: enzymology and organization into a family of related enzymes (12-14); nuclear functions (15); roles in cellular stress responses and inflammation (16, 17); mechanisms of chromatin modulation and transcriptional regulation (18, 19); roles in cellular signalling pathways (20, 21); and inhibition by small molecules and its therapeutic potential (22-25). In this review, I have summarized the molecular mechanisms of PAR formation, degradation, and recognition, as well as the consequences of PARylation for target protein function. In addition, I have highlighted the key mechanisms by which PARylation regulates many cellular responses, including: inhibition of protein-protein or protein-DNA interactions; formation of scaffolds that alter protein localization and promote protein interactions and regulation of other protein modifications, such as ubiquitylation. Finally, I have described the emerging molecular and cellular regulatory networks that depend on PARylation, and how these form the basis for the promising development of PARP inhibitor-based therapeutic approaches.

Regulated Formation of PAR Polymers

PARP family member	Alternative name	Transferase name*	Subclass	Size (aa) [‡]	Subcellular localization	Triad motif	Enzymatic activity [§]	Key functional motifs and domains
PARP1		ARTD1	DNA-dependent	1,014	Nuclear	H-Y-E	P and B	WGR, zinc-fingers and BRCT
PARP2		ARTD2	DNA-dependent	570	Nuclear	H-Y-E	P and B	WGR
PARP3		ARTD3	DNA-dependent	540	Nuclear	H-Y-E	M (P predicted)	WGR
PARP4	vPARP	ARTD4		1,724	Cytosolic (vault particle)	H-Y-E	P (predicted)	BRCT
PARP5A	Tankyrase 1	ARTD5	Tankyrase	1,327	Nuclear and cytosolic	H-Y-E	P and O	Ankyrin repeat
PARP5B	Tankyrase 2 and PARP6 [¶]	ARTD6	Tankyrase	1,166	Nuclear and cytosolic	H-Y-E	P and O	Ankyrin repeat
PARP6 [¶]		ARTD17		322	ND	H-Y-Y	M (predicted)	
PARP7	TIPARP and RM1	ARTD14	CCCH PARP	657	ND	H-Y-I	M (predicted)	Zinc-fingers and WWE
PARP8		ARTD16		854	ND	H-Y-I	M (predicted)	
PARP9	BAL1	ARTD9	macroPARP	854	ND	Q-Y-T	M (predicted)	Macrodomain
PARP10		ARTD10		1,025	Nuclear and cytosolic	H-Y-I	M	
PARP11		ARTD11		331	ND	H-Y-I	M (predicted)	WWE
PARP12	ZC3HDC1	ARTD12	CCCH PARP	701	Cytosolic (stress granules)	H-Y-I	M (predicted)	Zinc-fingers and WWE
PARP13	ZC3HAV1 and ZAP1	ARTD13	CCCH PARP	902	Cytosolic (stress granules)	H-Y-V	M (predicted)	Zinc-fingers and WWE
PARP14	BAL2 and COAST6	ARTD8	macroPARP	1,801	Cytosolic (stress granules)	H-Y-L	M	Macrodomain and WWE
PARP15	BAL3	ARTD7	macroPARP	444	Cytosolic (stress granules)	H-Y-L	M (predicted)	Macrodomain
PARP16		ARTD15		630	ND	H-Y-I	M (predicted)	

Table 2. The Organization and Enzymatic Activities of PARP Family Members.

aa, amino acid; ARTD, ADP-ribosyl transferase; BAL, B-aggressive lymphoma protein, COAST6, collaborator of signal transducer and activator of transcription 6; ND, not determined; PARP, poly(ADP-ribose) polymerase; vPARP, vault PARP; ZAP1, zinc-finger antiviral protein 1; ZC3HAV1, zinc-finger CCCH-type antiviral protein 1; ZC3HDC1, zinc-finger CCCH domain-containing protein 1. *Based on the revised nomenclature of Hottiger et al. ‡Size of the human protein in amino acids. §Known or predicted enzymatic activity: mono- (M), oligo- (O) or poly(ADP-ribosyl)ation (P), or branching (B). ||All PARP family members contain a PARP domain and the PARP signature motif. ¶PARP6, which refers to two different proteins, is an example of the degeneracy of the PARP nomenclature in the existing literature.

The PARP Family of Proteins. PARylation is a post-translational modification in which ADP-ribose units are added to glutamate, aspartate, and lysine residues of target (or acceptor) proteins by members of the PARP family. Seventeen PARP family members have been identified on the basis of homology to PARP-1, the founding member of the PARP family (13, 14) (Table 2). Not all PARP family members are enzymatically active, and some may function as mono(ADP-ribosyl) transferases, rather than PARPs; hence, a new nomenclature has been proposed that refers to PARPs more generally as ADP-ribosyl transferases, or ARTDs (26) (Table 2).

PARP family members comprise four subfamilies, which are largely based on their domain architectures (see references (13, 14) for an in-depth discussion). These include: the DNA-dependent PARPs, PARP-1, PARP-2 and PARP-3 (27-29) (reviewed in (13)); the tankyrases, including tankyrase-1 (PARP5a) and tankyrase-2 (PARP-5b); the CCCH (Cys-Cys-Cys-His) PARPs, including tiPARP (PARP-7), PARP-12, PARP13.1 and PARP13.2; and the macroPARPs, including BAL1 (PARP-9), BAL2 (PARP14, CoaSt6), and BAL3 (PARP-15) (Table 2). The DNA-dependent PARPs, PARP-1, PARP-2 and PARP-3, are activated by discontinuous DNA structures through their N-terminal DNA-binding domains. The tankyrases contain large ankyrin domain repeats that mediate protein-protein interactions facilitating target selection and activation (30). The CCCH PARPs contain Cys-Cys-Cys-His zinc fingers that bind to RNA, as well as WWE domains, which can exhibit PAR-binding activity. The macroPARPs contain macrodomain folds, which are ADP-ribose-binding modules that can facilitate localization of these PARPs to sites of poly- and possibly mono(ADP-ribosyl)ation. The other PARP family members have distinct domain structures (Table 2). These different PARP subfamilies vary in their enzymatic activities and some use non-enzymatic mechanisms of

[See the next page for the image for Figure 2]

Figure 2. Structure of PAR and mechanisms of PAR catalysis and glycohydrolysis..

a. The chemistry of PAR synthesis and the structure of PAR. *Top*, NAD⁺ and ADP from the terminal ADP-ribose of the PAR chain are shown in blue and purple, respectively, just prior to catalysis (oriented as shown in **b**). Upon catalysis of PAR synthesis, the nicotinamide moiety of NAD⁺ is released with $\alpha(1\rightarrow2)$ O-glycosidic bond formation. *Bottom*, During elongation, the adenine-proximal ribose from PAR chain termini are joined (purple), whereas during branching two nicotinamide-proximal ribose rings are joined (green). The terminal/newly added unit is shown in black.

b. The mechanism of elongation of PAR chains by PARP-1, as inferred by the position of carba-NAD⁺ (purple ADP) relative to NAD⁺ (blue) modelled into the active site of PARP-1 (31). The PARP fold is highlighted in the orange beta sheets and yellow alpha helices, with the conserved H-Y-E triad highlighted in green. The glutamate of this conserved triad is responsible for the catalysis of PAR elongation.

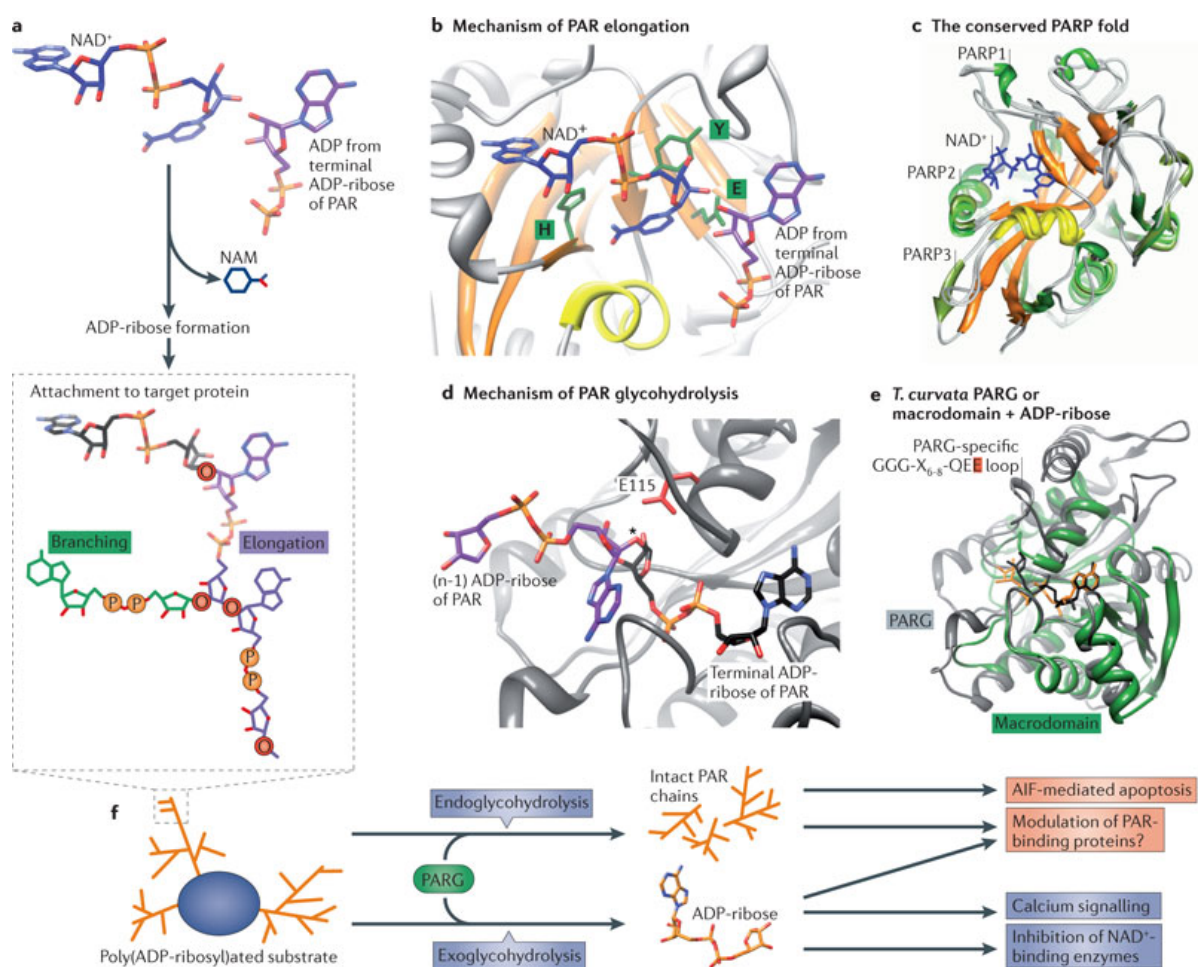
c. Structural comparison of the conserved PARP fold found in human PARP-1, PARP-2, and PARP-3 (PDB: 1A26, 3KCZ and 2PA9, respectively). The PARP fold is colored as in **b**, with PARP-1, PARP-2 and PARP-3 colored in green, light green, and olive green, respectively. The position of NAD⁺ in the PARP fold is based on Ruf *et al.*, 1998 (32).

d. Mechanism of PAR glycohydrolysis by PARG. The catalytic site of PARG is shown, with the catalytic glutamate shown in red. PARG is represented as in ref. (33), with the terminal ADP-ribose shown in black, and the n-1 ADP-ribose shown in purple. The site that is targeted by PARG, where the PARP-generated $\alpha(1\rightarrow2)$ O-glycosidic bond resides, is denoted with an asterisk (*).

e. Comparison of the structures of *T. curvata* PARG (gray; PDB: 3SIG) and the macrodomain of macroH2A1.1 (green; PDB: 3IID) bound to ADP-ribose (colored purple and orange, as indicated). The arrow highlights the characteristic PARG sequence GGG-X₆₋₈-QEE present in a non-homologous loop, which is not present in the macrodomain fold.

f. PAR degradation by PARG can yield different biological outcomes, depending on whether PARG-mediated degradation occurs by endoglycohydrolysis or exoglycohydrolysis. Intact PAR chains formed by endoglycohydrolysis can trigger AIF-mediate apoptosis (34), whereas free ADP-ribose formed by exoglycohydrolysis can trigger calcium signalling and inhibit NAD⁺-binding enzymes (35, 36). Both may modulate the function of macrodomain-containing proteins that can bind PAR.

Figure 2. See the previous page for the figure legend.



action: PARPs 1 and 2, vPARP, and tankyrase 1 and 2 catalyse poly(ADP-ribosyl)ation, whereas PARPs 3, 10, 14 and 15 catalyse mono(ADP-ribosyl)ation; the remaining PARP family members are either catalytically inactive or require specific and, as yet undetermined, cofactors to trigger ADP-ribosylation (37-39) (**Table 2**).

Catalysis of PARylation. The enzymatically active PARPs use NAD^+ as a donor of ADP-ribose units during catalysis, generating one ADP-ribose and one nicotinamide per molecule of NAD^+ metabolized (**Fig. 2a**). ADP-ribose can then be attached to target proteins via lysine residues and on the **carboxyl** group of glutamate and aspartate. PAR polymers can be generated through two different types of bond formation, promoting elongation or branching. In the elongation reaction, the adenine-proximal ribose (A-ribose) from PAR chain termini are joined in an $\alpha(1 \rightarrow 2)$ *O*-glycosidic bond, whereas in the branching reaction the ADP-ribose junction occurs between two nicotinamide-proximal ribose (N-ribose) rings (**Fig. 2a, b**). PAR polymers can grow to over 200 units in size, with branching reactions occurring on average once per every 50 elongation reactions (40). This elongation bias is thought to be achieved through preferential binding of PAR ADP-ribose units that are orientated in an elongation-competent manner with the surface of the PARP catalytic domain, consistent with analysis of where the terminal ADP-ribose site resides in the co-crystal structure of carba- NAD^+ with the PARP-1 catalytic domain (31) (**Fig. 2b**).

The PARP catalytic domain binds NAD^+ using a protein fold that shares homology with mono(ADP-ribosyl)ating bacterial exotoxins, such as *Diphtheria* toxin and exotoxin A (**Fig. 2b, c**). The PARP domain consists of a unique fold, set apart from the Rossman fold of other NAD^+ -

binding enzymes (for example, lactate dehydrogenase), which coordinates NAD^+ in a distinct orientation at nicotinamide and N-ribose moieties and is conserved from bacterial exotoxins to eukaryotic PARPs (32). The positioning of adenine in the catalytic pocket is more poorly understood, as the structure of a PARP and its NAD^+ substrate have yet to be determined (NAD^+ has been modelled into the structures shown in [Fig. 2b, c](#)). On the basis of both structural homology with the contacts made between *Diphtheria* toxin and NAD^+ and *in silico* characterization of PARPs, it has been inferred that a histidine and a tryptophan that are highly conserved in enzymatically active ADP-ribosyltransferases are critical in positioning, respectively, the A-ribose and N-ribose moieties of NAD^+ in the correct orientation for (ADP-ribosyl)ation (41, 42) ([Fig. 2b](#)). In poly(ADP-ribosyl)ating PARPs, a catalytic glutamate is responsible for catalysis of additional ADP-ribose units onto the acceptor residues. These three atoms form the ‘HYE’ (that is, histidine, tryptophan, glutamate) triad that appear in bona fide poly(ADP-ribosyl)ating PARP proteins, as well as a number of homologous mono(ADP-ribosyl)ating exotoxins in bacteria ([Fig. 2b](#)). The loss of the catalytic glutamate residue from the PARP catalytic domain is responsible for altering the activity of PARPs from poly- to mono(ADP-ribosyl)ating enzymes, with at least a subset of mono(ADP-ribosyl)ating PARPs achieving catalysis through a substrate-assisted mechanism (39).

Degradation and Recycling of PAR. PARP proteins are responsible for the synthesis of PAR and, as might be expected, a number of proteins with PAR-degrading activities promote the rapid catabolic destruction of PAR almost immediately after synthesis. Such turnover of PAR polymers allows temporal control over the recruitment and release of PAR-binding proteins from

specific cellular locations, as well as transient signalling and the formation of transient sub-organellar structures in the cytoplasm and nucleus. The best studied of the PAR-degrading enzymes is the endo- and exoglycohydrolase poly(ADP-ribose) glycohydrolase (PARG). PARG is present in mammalian cells in three different isoforms: PARG₉₉ and PARG₁₀₂ (99 and 102 kDal, respectively), which localize in the cytoplasm; and PARG₁₁₀ (110 kDal), which localizes predominantly to the nucleus (43). As discussed below, in spite of their seemingly opposing activities, PARG and PARPs function in a coordinated manner to regulate a wide range of cellular processes.

The catalytic domains of known PARGs are structurally related to the macrodomain fold, which is an ADP-ribose binding module (33) (**Fig. 2d, e**). Structural analysis of a PARG enzyme from the bacterium *T. curvata* has revealed a PARG-specific GGG-X₆₋₈-QEE loop that has evolved near the N-ribose moiety of the ADP-ribose-binding site in the macrodomain fold (**Fig. 2d, e**). The integrity of the glutamates in this loop, in which the C-terminal glutamate is the catalytic residue, is critical for PARG glycohydrolysis of the $\alpha(1\rightarrow2)$ *O*-glycosidic bond that covalently links individual ADP-ribose monomers together in PAR polymers (**Fig. 2d**). The crystal structure of *T. curvata* PARG in complex with relevant ligands suggests that it has only exoglycohydrolytic activity, which degrades from the terminal ADP-ribose unit in PAR rather than internally between ribose units of the polymer, whereas PARGs from higher eukaryotes have both exo- and endoglycohydrolytic activities (**Fig. 2f**). The endoglycohydrolytic activity may have co-evolved with PARP enzymes that catalyse more complex branching, as the bacterial PARP homologue does not appear to PARylate to the same extent as human PARP-1,

and the hydrolysis of branched ADP-ribose units by bovine PARG is markedly slower than that of elongated ADP-ribose units (44).

Other PAR-degrading Enzymes. In addition to the better-studied PARG enzymes, other PAR-degrading enzymes have been identified and characterized. These include the ADP-ribosyl hydrolase (ARH) and NUDIX (nucleoside diphosphate linked to another moiety, X) families of proteins. ARH3 is a PAR-degrading hydrolase that, similarly to PARG, can catalyse the removal of PAR, but not mono(ADP-ribose) (45). ARH3 is a structurally distinct enzyme from PARG and, in addition to PAR hydrolysis, can also remove the *O*-acetyl group from the NAD⁺ metabolite *O*-acetyl-ADP-ribose (46). ARH3, an enzyme conserved in all vertebrates, has been implicated in the degradation of PAR that is associated with the mitochondrial matrix, but a complete understanding of its function in vivo remains poorly understood (47). The NUDIX class of hydrolases (reviewed in ref. (48)) are highly conserved proteins that regulate the catabolism of diverse nucleoside-diphosphate-containing molecules, which includes ADP-ribose covalently linked in PAR chains. There is little evidence so far showing that NUDIXs affect PAR catabolism in vivo, but their ability to rapidly metabolize ADP-ribose suggests that they indeed may have such a role.

Mechanisms of PARP Activation

PAR production is a tightly controlled process, as PARylation of DNA repair proteins, transcription factors, and metabolic regulators, to name a few, can have dramatic consequences for cellular physiology. Thus, regulated activation of PARPs is central to their function. Of all

the PARP family members, the activation of PARP-1 is the best understood. PARP-1 has a DNA-binding domain containing three zinc finger motifs, which are important for recognition of various features in DNA structure, such as single- and double-strand breaks, cruciform structures, and nucleosome linker DNA (49, 50). As revealed in the structure of PARP-1 bound to a DNA double-strand break, a conformational change occurs in PARP-1 upon DNA binding, resulting in modulation of the active site through WGR-domain-mediated remodeling of an inhibitory helical domain near the ADP-ribosyl transferase active site (51). While the structural work by *Langelier et al.* has clarified how information is translated from the DNA binding domain to the catalytic domain of PARP-1, important questions still remain. The structure by Langelier et al does not include the second zinc finger of the DNA binding domain, as it is dispensable for double-strand break recognition, yet this zinc finger is most potent in single strand break recognition (51, 52). Moreover, it remains unclear how recognition of intact nucleosome linker DNA occurs, suggesting the possibility that different PARP-1 substrates may alter DNA-induced allostery to connect inputs to different post-translational targets, as has been shown for DNA binding transcription factors such as the glucocorticoid receptor (53, 54).

DNA is not the only stimulatory factor for PARP-1; nucleosomes, as well as a variety of protein interaction partners, also stimulate PARP-1 activity (55-57) (**Fig. 3**). For example, the enzymatic activity of the fly (*Drosophila*) homologue of mammalian PARP-1, dPARP, is activated or inhibited, respectively, by histones H4 or H2B and H2A (58). Interestingly, the fly H2A variant, H2Av, which localizes to gene promoters, is less inhibitory than canonical H2A (59). This may allow localization and activation of dPARP near the promoters of expressed genes, as has been shown to occur at the *HSP70* locus (60). The activation of dPARP at the

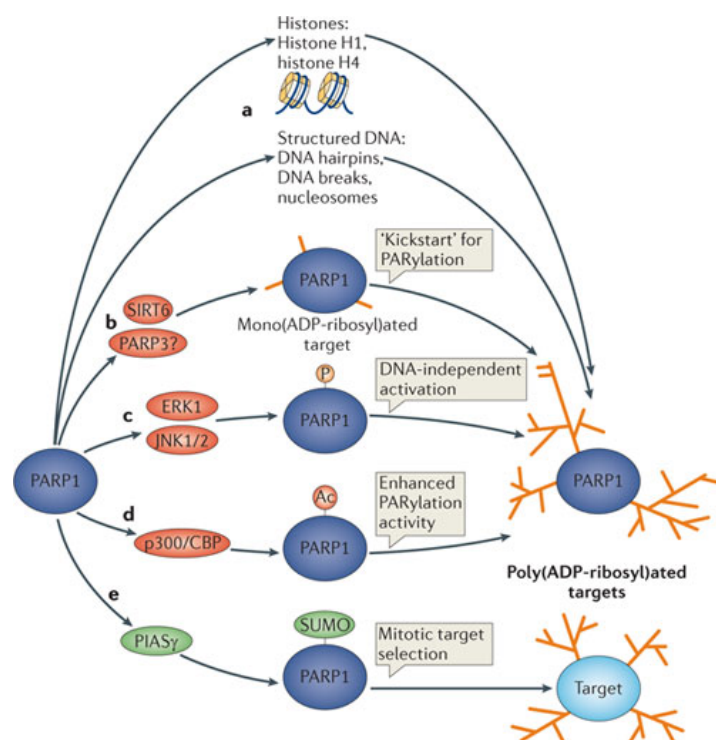


Figure 3. PARP-1 activation and target selection.

Poly(ADP-ribose) polymerase 1 (PARP-1) activity and substrate specificity is modulated by association with substrates and post-translational modifications during both PARP-1 auto-poly(ADP-ribosylation) (PARylation) and PARylation of other targets.

a. PARP-1 association with histones or structured DNA can directly stimulate its activity.

b. Mono(ADP-ribosyl)ation of PARP-1 or its targets may provide a 'kickstart' mechanism for PAR synthesis. The mono(ADP-ribosyl)ating enzymes SIRT6 and PARP-3 both stimulate PAR synthesis by PARP-1, which has kinetically impaired mono(ADP-ribosyl)ation activity relative to poly(ADP-ribosyl)ation activity. The 'pre-addition' of a mono(ADP-ribose) unit to PARP-1 by SIRT6 or PARP-3 may promote subsequent poly(ADP-ribosyl)ation of target proteins by PARP-1. This 'kickstart' mechanism may provide a means whereby PARPs and mono(ADP-ribosyl)transferases can cooperate in response to cellular stimuli.

c. Extracellular signal-related kinase 2 (ERK2), JUN N-terminal kinase 1 (JNK1) and JNK2 phosphorylate (P) PARP-1 to promote activation of its catalytic activity. Phosphorylation by ERK2 can stimulate PARP-1 autoPARylation independently of DNA.

d. Acetylation (Ac) by the acetyltransferase p300/CBP can stimulate PARP-1 activity. This acetylation can be reversed by the NAD^+ -dependent protein deacetylase SIRT1, inactivating PARP-1.

e. During mitosis, SUMOylation (SUMO) of PARP-1 by PIASy alters PARP-1 target selection, by restricting PARylation activity to a subset of mitotic targets.

HSP70 promoter is mediated in part by the histone H2A acetyltransferase dTIP60, as well as globally by phosphorylation of histone H2Av, implying that the chromatin state can affect PARP-1 activity (61). PARP-2 is also activated by breaks in DNA, which results in its automodification and activation, much like PARP-1 (62).

Post-translational modification of PARPs can also alter PARP activation. PARP family members, as well as their targets, are subjected to a wide range of post-translational modifications, including phosphorylation, acetylation, methylation, SUMOylation and ubiquitylation (reviewed in references (15, 16)). In many cases, the PARPs function as hubs that link various post-translational modifications in a signal-dependent manner. For example, PARP-1 and the tankyrases are targets for kinases, and phosphorylation by ERK2 and MAPK, respectively, leads to activation of these PARPs, driving specific cellular outcomes (63, 64) (Fig. 3c). Thus, post-translational modifications of PARPs can integrate signalling inputs and outputs in a wide variety of cellular processes.

Mechanisms of Target Selection

The rules that govern how PARPs interact with their targets and are activated in response to target binding are not well understood in most PARPs, although progress is being made. The specific domain architecture of each PARP family member, as well as the target proteins, is probably important. For example, in PARP-1, the BRCA1 C-terminus (BRCT) motif forms a phosphoprotein-binding module that is thought to drive protein-protein interactions between PARP-1 and its substrates. Furthermore, a tankyrase interaction motif found in tankyrase target proteins helps determine their binding to the ankyrin repeats that are characteristic of tankyrases,

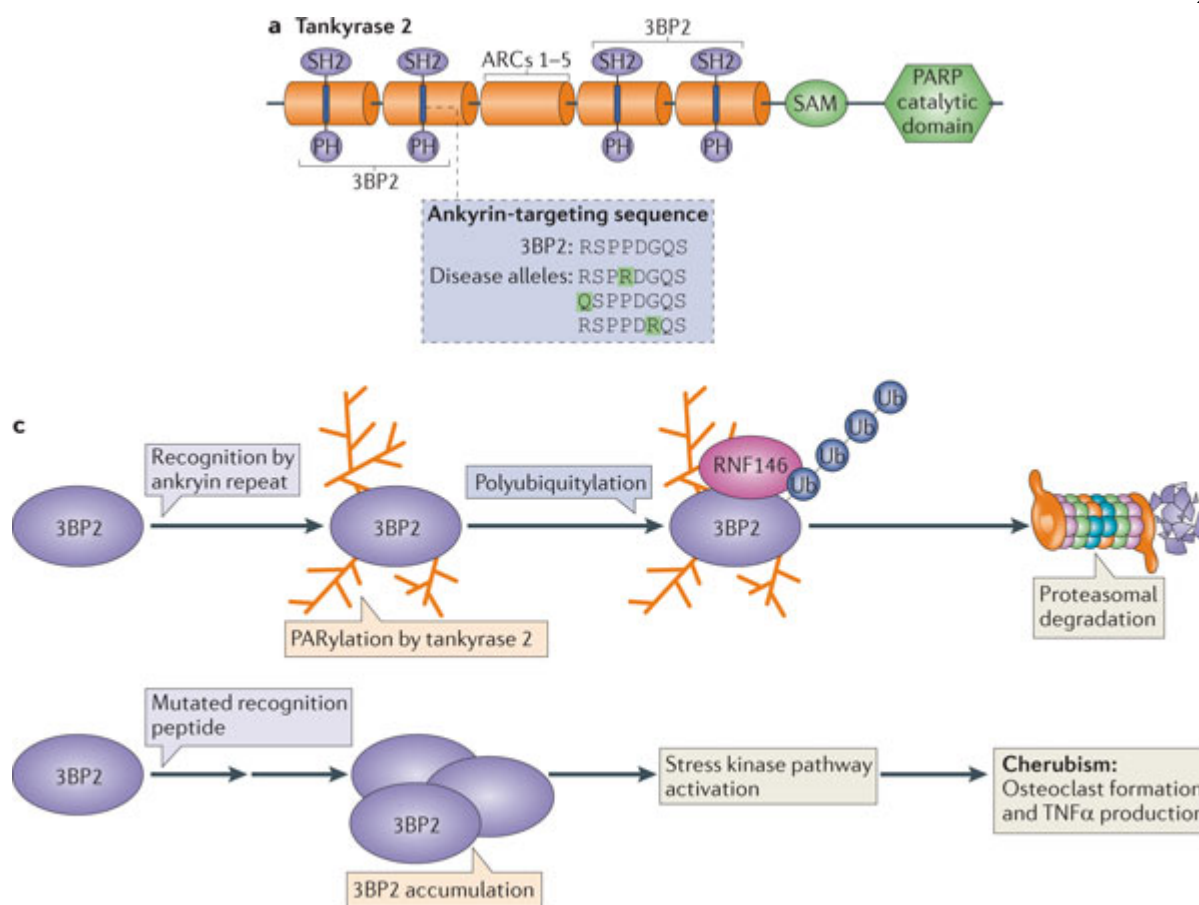


Figure 4. PAR-dependent ubiquitylation and the human disease Cherubism.

PAR-dependent ubiquitylation is a broad phenomenon used in a diverse set of biological processes and its misregulation results in the human disease Cherubism.

a. Schematic representation of the domain architecture of tankyrase-2 and its interaction with 3BP2, through the ankyrin-targeting sequence, RSPPDGQS, and ankyrin repeat clusters 1, 2, 4, and 5. The changes in the ankyrin-targeting sequences in 3BP2 that result in the human disease Cherubism are shown in the inset. **b.** PAR-mediated destruction of 3BP2 underlies the etiology of Cherubism. 3BP2 is directed through its ankyrin-targeting sequence for PARylation by tankyrase-2 and subsequent ubiquitylation by the E3 ligase RNF146 (also known as Iduna), which leads to proteasome-mediated degradation of 3BP2. Specific mutation of the ankyrin-targeting sequence in 3BP2 can prevent its recognition by tankyrase-2 and thereby prevent its PARylation and ubiquitylation; this results in abnormal accumulation of 3BP2, causing stress kinase pathway activation, systemic inflammation through TNF α signalling and the disease phenotypes of Cherubism. SAM, sterile alpha motif; SH2, Src homology; PH, pleckstrin homology.

facilitating target selection and activation (30) (**Fig. 4a, b**). This tankyrase interaction motif, historically defined as RXXPDG, has more recently been quantitatively determined to have an optimal sequence motif of REAGDGEE (30, 65). Amino-acid substitutions in this motif in the tankyrase target protein 3BP2 have shown that tankyrase recognition of this motif is critical for 3BP2 recognition and PARylation (**Fig. 4b**). Mutation of the motif in 3BP2 disrupts PAR-directed, ubiquitin-mediated degradation of 3BP2 and causes Cherubism, a rare autosomal dominant human disease of the maxilla and mandible in which the bone is replaced with fibrous tissue (65, 66) (**Fig. 4c**). Other diverse structural and functional modules are found in in PARP family members (reviewed in reference (13)) (**Table 2**; for example, WGR, WWE, zinc finger, macrodomain), indicating that they are key drivers of the different substrate specificities among PARPs.

Post-translational modification of PARPs can also control how they recognize specific target proteins. For example, during cell division, PARP-1 recruits and PARylates PIAS γ , an E3 SUMO ligase; PIAS γ in turn SUMOylates PARP-1, and thereby promotes association with a specific subset of PARP-1 targets *in vivo* (67) in a cell cycle-specific manner (**Fig. 3**). SUMOylated PARP-1 is also bound by and ubiquitylated by the ubiquitin E3 ligase RNF4, which promotes the degradation of PARP-1 (68). How post-translational modifications of PARP targets affects their interactions with, and modification by, PARPs *in vivo* remains poorly understood, but studies with PARP-1 indicate that they are likely to also be important.

Cooperation Between PARPs. Recent work has highlighted the interplay between mono(ADP-ribosyl)ation and the activation of PARPs. Notably, both PARP-3 and SIRT6 are mono(ADP-

ribosyl)ating proteins that can activate PARP-1 in vitro and in vivo (37, 69). Mono(ADP-ribosyl)ation of PARP-1 may function as a ‘kickstart’ to PARP-1-dependent poly(ADP-ribosyl)ation and acts as a control mechanism for PARP-1 activation (**Fig. 3b**). In this regard, biochemical studies have shown that mono(ADP-ribosyl)ation in the first cycles of PARylation are the most kinetically impaired step of the PARylation reaction (37, 69, 70). Consistent with this, SIRT6 functions as a critical component of the PARP-1-mediated response to DNA damage in both the non-homologous end joining and homologous recombination pathways. Although this ‘kickstart’ mechanism has yet to be demonstrated for other PARP proteins, the widespread colocalization of poly(ADP-ribosyl) transferases with mono(ADP-ribosyl) transferases suggests that this interplay is more general and may be an underappreciated and critical mechanism for controlling PARP activity (71).

Cellular Functions of PARylation

PARPs regulate a wide array of cellular processes, from transcription and DNA repair, to mitochondrial function and the formation of cytoplasmic and nuclear sub-organellar bodies. Although not exclusively so, the PARP family of proteins is a key component of stress responses and mediates these responses in diverse ways (16). The use of a post-translational signal, such as PAR, in pathways that require dramatic and immediate responses is perhaps advantageous due to the nature of PAR itself: bulky, charged and flexible. A broad view of the literature reveals three major types of regulatory mechanisms that are mediated by PAR: inhibition of protein-protein and protein-nucleic acid interactions; formation of scaffolds that alter protein localization and promote protein interactions and regulation of ubiquitylation events (**Fig. 5**). The latter two of

these mechanisms are driven by distinct PAR-binding modules that are present in key regulatory proteins.

Inhibition of Protein-Protein and Protein-Nucleic Acid Interactions. The best-studied mechanism of ADP-ribose function, either as a mono- or poly-modification, is its ability to inhibit acceptor proteins from interacting with their protein-binding partners or substrates (Fig. 5a). The logical simplicity of attaching the large, charged ADP-ribose modification to inhibit target protein function is illustrated by its adoption by a diverse array of proteins over large periods of evolutionary time. Bacterial toxins, such as exotoxin A and *Diphtheria* toxin, are closely related structurally, and enzymatically in their catalytic fold, to the PARP fold in the catalytic domain of PARP-1 and other PARPs (32, 42). For example, the extraordinarily potent *Diphtheria* toxin inhibits RNA translation after infection by *C. diphtheriae* through the addition of a single ADP-ribose unit onto a diphthamide residue in eEF-2, which perturbs its interaction with the eukaryotic translation machinery (72).

Similar mechanisms of ADP-ribose-dependent inhibition have been widely adopted in eukaryotic cells for the control of protein-protein and protein-nucleic acid interactions. PAR has properties that are similar to RNA and DNA and, as such, can repulse nucleic acid targets, as well as attract nucleic-acid-binding proteins. Such PAR-dependent control is used by a variety of regulatory proteins in diverse biological systems, as exemplified by regulation of the transcription factor CLOCK, which is crucial for metabolic entrainment of circadian rhythms in mice (73). PARP-1 knockout mice show aberrant circadian rhythms when normal feeding cycles are perturbed. Molecular analyses have shown that PARP-1 PARylates CLOCK in response to

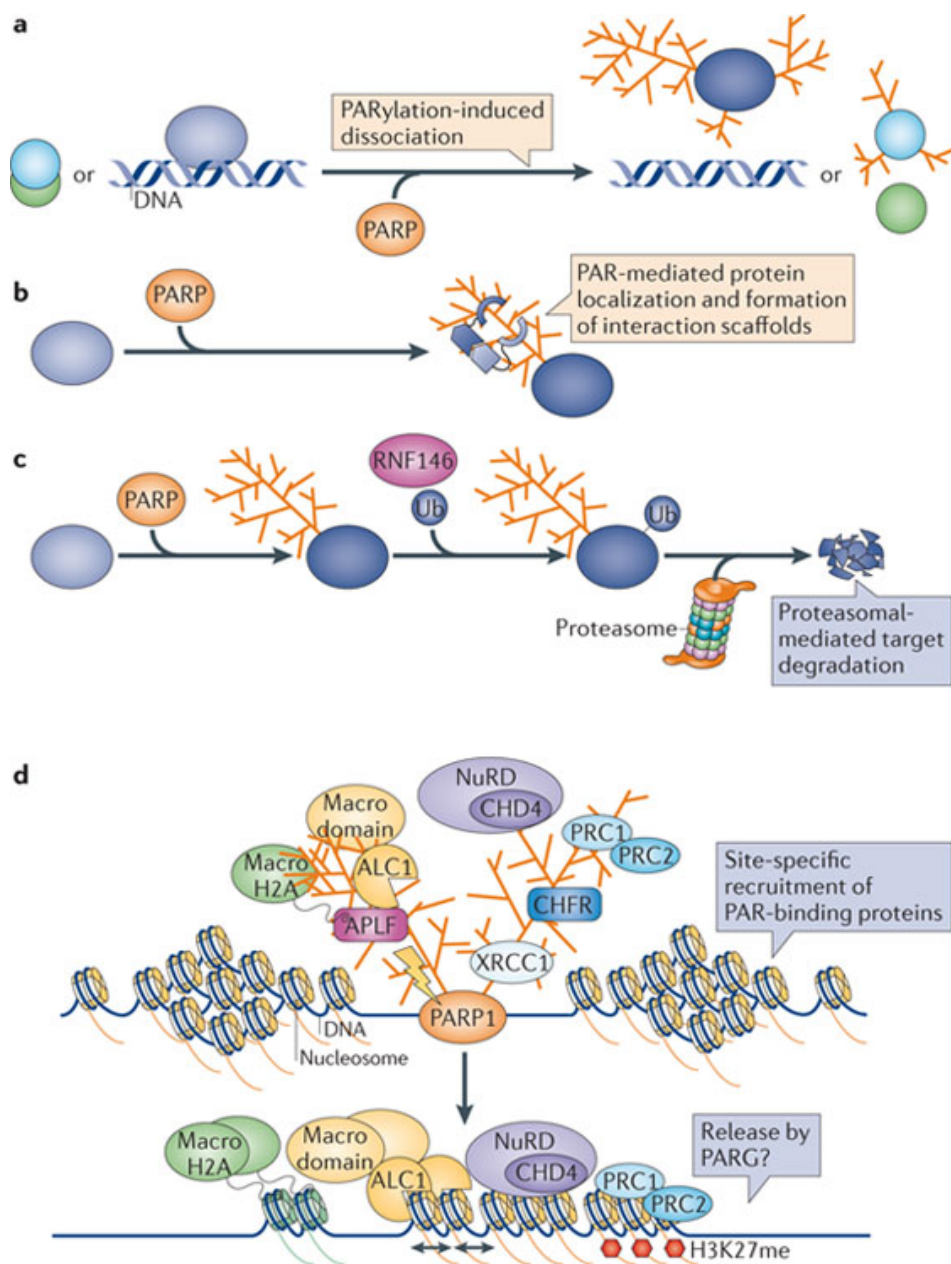
[See the next page for the image for Figure 5]

Figure 5. Cellular functions of PAR and the PAR-dependent recruitment of factors to sites of DNA damage.

PAR can modulate cellular functions in several ways, dictated by its bulky and charged nature.

- a.** PAR modification of a protein often inhibits protein-protein or protein-nucleic acid interactions.
- b.** Site-specific PARylation of target proteins can mediate the recruitment of PAR-binding proteins and facilitate protein-protein interactions, resulting in a PAR-based interaction scaffold.
- c.** PARylation of a target protein can trigger recruitment of E3 ligases that contain PAR-binding modules and thereby subsequent ubiquitylation and proteasome-mediated degradation of the target protein.
- d.** During the DNA damage response, poly(ADP-ribose) polymerase 1 (PARP-1) binds at the site of DNA damage, undergoes automodification, and PARylates other factors including histone H1 and XRCC1. This promotes rapid recruitment of PAR-binding proteins to site of DNA damage, which is important for efficient damage repair. These include: the PBM-containing proteins XRCC1 (a DNA damage scaffold protein) and CHD4 (a component of the NuRD chromatin remodelling complex); PBZ-containing proteins APLF and CHFR; and macro domain-containing proteins macroH2A and ALC1. Deficiencies in the PAR-binding abilities of these proteins can impair the kinetics of repair or promote chromosomal aberrations (74-79). The recruitment, and action of (black bi-directional arrows) chromatin remodelers (ALC1, NuRD) and chromatin modifying (H3K27me) and silencing complexes (PRC1 and PRC2) are important in pre- and post-repair remodeling of the nucleosome landscape during DNA repair. Also, the recruitment of In addition, PAR turnover by PARG may also be important the release of recruited factors and the procession of DNA repair.

Figure 5. See the previous page for the figure legend.



feeding stimuli, disrupting CLOCK binding to its DNA response element. Disruption of this PAR-dependent signalling pathway in PARP-1-deficient mice causes dramatic changes in their sleep schedule. Although questions still remain as to exactly how PARP-1 is activated by food intake, this exciting discovery highlights the variety of ways in which PARylation can support complex regulatory networks. Indeed, PARylation regulates the activities of various proteins, impairing: chromatin remodelling by ISW1; histone demethylation by KDM5B; telomere loop binding by TRF-1; nucleosome-nucleosome interactions to facilitate chromatin decondensation; and nuclear to cytoplasmic transport of NF κ B and p53 transcription factors, to name just a few examples (80-84).

PAR-dependent Protein Localization, Interaction Scaffolds, and Phase Separation.

PARylation of a protein target can promote the formation of a molecular scaffold that recruits other proteins required for a particular process (Fig. 5b). This is exemplified by the role of PARylation at sites of DNA damage, where it recruits DNA repair factors, as well as other chromatin-binding proteins (Fig. 5d). In response to DNA damage, PAR polymers are generated by PARP-1 and this recruits the early DNA damage response factor ATM, as well as the SUMO E3 ligase PIAS γ , both of which contain PAR-binding motifs. ATM and PIAS γ interact with, and modify, I κ B kinase γ (IKK γ), triggering its phosphorylation and SUMOylation, respectively. This activates IKK γ , and thereby triggers phosphorylation and degradation of the inhibitory I κ B protein to allow an NF κ B-mediated transcriptional response to damage (85). This PAR-dependent ‘signalosome’ is a mechanism that highlights how the unique scaffolding properties of PAR allow it to integrate diverse cellular signals.

The most dramatic demonstration of how PAR-binding proteins form a functional scaffold during DNA damage responses has arisen from the use of laser micro-irradiation coupled with fluorescence microscopy (**Fig. 5d**). In these studies, genomic lesions are induced by a 355 nm laser beam that is directed across the nucleus (75). Fluorescence microscopy is then used to track localization of damage response factors in real-time. Within seconds of irradiation, markers of DNA damage, such as γ H2AX and PARP-1-dependent PAR accumulation, can be readily observed within laser-demarcated sites of DNA damage. PAR-binding proteins are also rapidly recruited to these sites within seconds of irradiation in a PAR- and PAR-binding domain dependent manner, including: the DNA damage scaffold protein XRCC1 and a component of the NuRD chromatin remodelling complex, CHD4, both of which contain PAR-binding modules; PAR-binding zinc finger (PBZ)-containing proteins APLF and CHFR; and macro-domain-containing proteins macroH2A and ALC1; as well as RNA-binding proteins such as FUS, which create a phase separated liquid at the site of DNA damage, the function of which has yet to be determined (**Fig. 5d**). Deficiencies in the PAR-binding abilities of these proteins can impair the kinetics of repair or promote chromosomal aberrations (74-79).

PAR-dependent scaffolding also affects the formation and function of stress granules, which are cytosolic phase-separated liquids comprised of RNAs and RNA-binding proteins (including stalled translation pre-initiation complexes) that form in response to cellular stress partly to protect mRNAs against degradation (71). Upon activation of the stress response, RNA-binding PARPs 12, 13.1, and 13.2 cooperate with tankyrase-1 and PAR-binding, macrodomain-containing PARPs 14 and 15 to form stress granules. Importantly, stress granule formation can be generated de novo by overexpression of any of these PARP family members, and stress-

dependent stress granule formation can be inhibited by overexpression of the cytoplasmic PARG isoforms PARG₉₉ and PARG₁₀₂. The phase-separated stress granule probably results from multivalent cooperation between the PARylation activity of tankyrase, RNA binding by PARPs 12, 13.1 and 13.2, and the PAR-binding properties of macro domains within macroPARPs 14 and 15. These stress granules function to insulate mRNA from micro RNA (miRNA)-mediated degradation, in cooperation with PARylation of Ago2 and other components of the miRNA machinery, which inhibits their function (71).

PAR also functions as a protein interaction scaffold to promote formation of other sub-organellar structures. For example, spindle pole proteins are recruited by PAR during mitosis (86-88) and Cajal bodies in the nucleus are enriched in PAR and nucleic-acid-binding proteins (89). Both the Cajal body component Coilin and PARP-1 itself are targets of PARylation at Cajal Bodies. Future studies in this area will benefit from a more detailed understanding of PAR polymer dynamics and how phase separation by PARylation assists in the biology of these diverse PAR-driven structures (90).

PAR-dependent Ubiquitylation. PAR-dependent ubiquitylation resulting in subsequent proteolysis of the target protein is a broad phenomenon that can be mediated by multiple PARP proteins (91-94). These results are changing our understanding of how PARylation can alter cell signalling and the fate of proteins that are targets of PARP family members. For example, RNF146 (also known as Iduna) is a PAR-binding E3 ligase, which binds PAR through its WWE domain (95, 96) (**Fig. 5c**). RNF146-mediated, PAR-dependent ubiquitylation is an important part of a tankyrase-dependent protein degradation pathway, which regulates cellular signalling in

several biological contexts, as well as proteasome-directed degradation of PARP-1 targets during glutamate excitotoxicity in the brain (97, 98). The WWE domain in RNF146 binds specifically to PAR and subsequently ubiquitylates lysines on PARylated proteins through its RING E3 ligase domain (95, 96). Ubiquitylation increases with the extent of PARylation, but studies have yet to reveal how PAR polymer length affects the efficiency of ubiquitylation, what the optimal length of the PAR polymer is and how PAR turnover might affect the efficiency of E3 ligase activity towards PARylated proteins.

Although RNF146-mediated PAR-dependent ubiquitylation has thus far been studied most extensively, there is increasing evidence that crosstalk between PARylation and ubiquitylation is not restricted to the case of RNF146. For example, the ubiquitin ligase CHFR contains a PAR-binding zinc finger (PBZ), as well as a RING E3 ligase domain (77). Although initial studies have showed that auto-ubiquitylation of CHFR does not require PAR binding, the same is true for RNF146 (ref. (77)). CHFR can ubiquitylate PARP-1 both *in vitro* and *in vivo* in a PAR-dependent manner, similarly to RNF146 (ref. (99)). Upon mitotic stress, CHFR interacts with PARylated PARP-1 through its PBZ domain, triggering ubiquitylation and degradation of PARP-1, which results in cell cycle arrest. Altering this PAR-dependent crosstalk prevents the cell cycle arrest during mitotic stress, resulting in genomic aberrations and mitotic catastrophe (99). Moreover, several ubiquitin ligases contain PAR-binding WWE domains (94), suggesting that the examples of PAR-dependent ubiquitylation identified so far may only represent the beginning of a common theme of connections between PARylation and ubiquitylation (93, 100). More broadly, ubiquitylation of PARylation targets may be part of a larger negative feedback mechanism that restrains overactivation of PARylation-induced cell signalling.

Regulated Formation and Degradation of PAR Polymers. PAR catabolism by PARG is a crucial component of PARylation-mediated control. As such, PARG has several important cellular and physiological roles in higher eukaryotes, particularly during acute signalling or stress responses. Under conditions of cell stress, PARP-1 becomes hyper-activated, modifies target proteins in stress response pathways, and with continued activation may promote cell death. PARG catalytic activity can contribute to cell death during PARP-1 hyperactivation via two different mechanisms: by rapidly degrading PARP-1-synthesized PAR polymers, contributing to the depletion of NAD^+ and, ultimately, ATP in a necrotic cell death process called parthanatos (101); and by endoglycohydrolysis of PAR chains, whose oligomeric products are exported to the mitochondria where they are necessary and sufficient to promote an AIF-mediated cell death programme (Fig. 1f)(34, 102). Moreover, PARG also generates free ADP-ribose during exoglycohydrolysis, which can inhibit the activity of NAD^+ -binding enzymes, alter calcium signalling, and modulate the function of macrodomain-containing proteins, implicating PARG in cellular signalling events downstream of its effects on PAR chains (Fig. 1f) (35, 36, 103). Commensurate with the role of PARG in cell death decisions, PARG-deficient mice are embryonic lethal at E3.5, which is probably mediated through the accumulation of PAR and uncontrolled PAR-dependent signaling (104).

PARG is also important for the response to DNA damage. PARG is rapidly recruited to sites of DNA damage in a PAR-dependent manner, and works together with PARP to ensure turnover of PAR and dynamic recruitment and release of factors at sites of DNA damage (105). Thus, PARG is an integral and crucial part of the cellular responses mediated by PARP family

members, and the coordination of PARG and PARP activities reveals much about the nature and regulation of PAR function (106).

Roles of PAR-binding Modules

During the formation of PAR-dependent protein interaction scaffolds, PAR recognition and association occurs through four distinct protein modules: PAR-binding motifs (PBMs); PBZ domains; macrodomain folds; and WWE domains. Some of these are found in the PARPs themselves. Several of these, which I discuss below, have been best characterized in the context of protein recruitment to PAR during DNA repair.

PAR-binding Motifs. PBMs are present in a strikingly large variety of proteins, including DNA damage response proteins and enzymes involved in chromatin remodelling and RNA processing (107) (reviewed in refs (13, 15)). PBMs, which comprise a short sequence of amino acids, were the first PAR-interacting modules discovered (108). The sequence, [HKR]₁-X₂-X₃-[AIQVY]₄-[KR]₅-[KR]₆-[AILV]₇-[FILPV]₈, has a number of Lys-Arg clusters interspersed with more hydrophobic amino acids (108) (**Fig. 6**). Although it is likely that the hydrophobic and basic residues of this motif cooperate to recognize the unique features contained within PAR, little detailed structural information is available on the PAR-PBM interaction. The variations that are observed in the sequence specificity of the PBM motif may reflect more promiscuous binding by proteins containing PBMs, in which PBMs may in fact direct binding to various nucleic acid species, but then be ‘tuned’ by their specific sequence for PAR, RNA or DNA binding (109).

One example of a PBM-containing protein is dMi-2, a *Drosophila* homologue of the mammalian nucleosome remodelling enzyme CHD4 (110). It is recruited to heat shock ‘puffs’ (that is, regions of decondensed chromatin in polytene chromosomes that include the gene loci encoding heat shock proteins, such as HSP70) in a PAR-dependent manner, which is facilitated through PAR-binding K/R-rich regions in its amino-terminus, each of which vary in their relatedness to the consensus PBM. The dMi-2 PBM also binds DNA and RNA, possibly allowing a ‘recruitment and handoff’ mechanism of PAR- and DNA-binding, followed by RNA-binding, which would allow dMi-2 to participate in multiple steps of gene regulation at the *HSP70* locus.

PAR-binding Zinc Finger. The PAR-binding zinc finger (PBZ) domain is another relatively small and unstructured PAR-binding module found in various proteins, including the DNA damage response proteins APLF and CHFR (77) (Fig. 6). PBZ domains are related to ssRNA-binding zinc fingers of the C3H1 class; both contain a similarly unordered fold punctuated by a rigid zinc-coordinated backbone that stacks with the adenine base. The specificity of the PBZ domain for PAR polymers is dictated by: a zinc-coordinated fold that promotes stacking interactions with either one or two adenine moieties of a poly(ADP-ribose) unit; and the unique $\alpha(1\rightarrow2)$ *O*-glycosidic bond that occurs at ADP-ribose-ADP-ribose junctions (76, 111, 112). PBZ domains generally exhibit nanomolar affinities for ADP-ribose; when found in tandem, the effective affinities can be sub-nanomolar. The high affinity of the PBZ domain for PAR allows DNA damage response proteins such as CHFR and APLF to recognize and bind to sites of

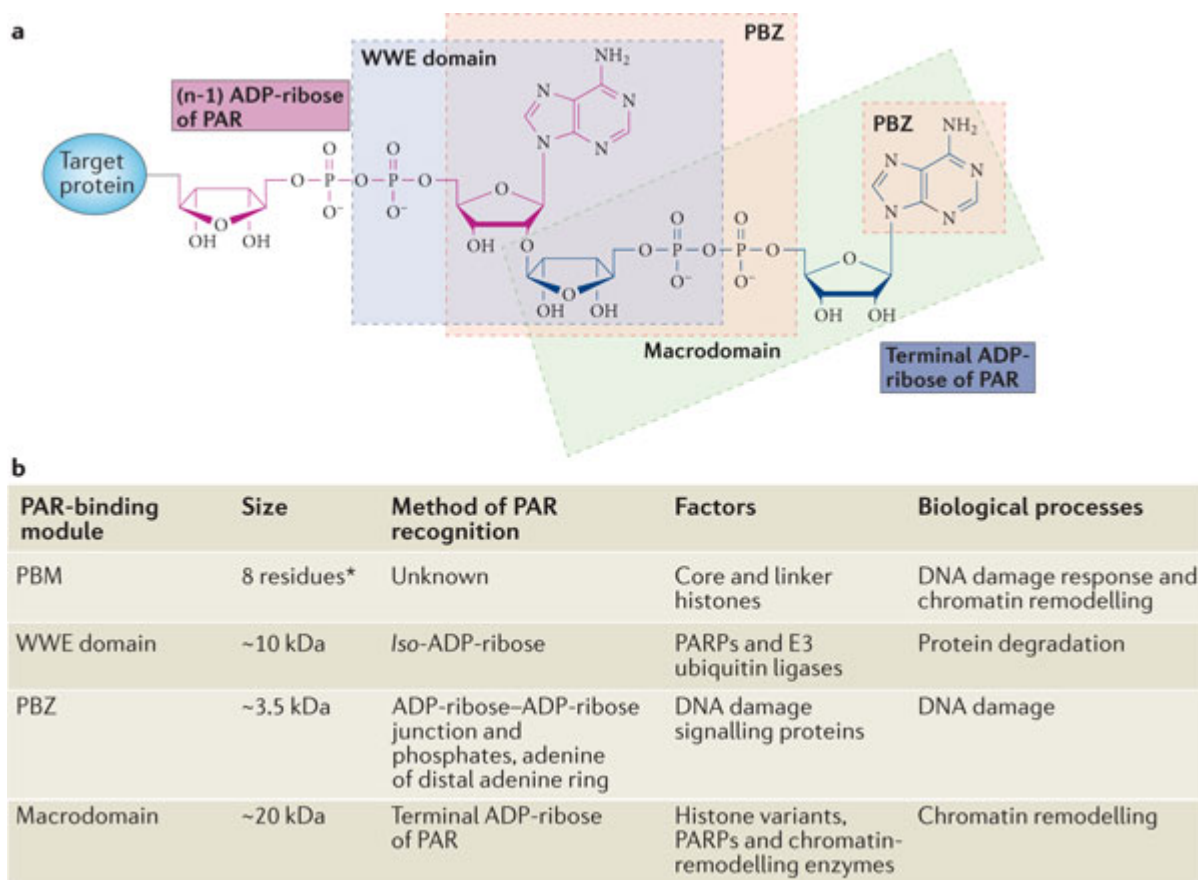


Figure 6. Recognition strategies and functions of PAR-binding modules.

a. PAR-binding domains in PAR-binding proteins bind to different structural features of PAR. The elongated form of PAR is shown here, with a terminal ADP-ribose (black) joined to an (n-1) ADP-ribose (purple). WWE domains recognize the iso-ADP-ribose residue that is unique to poly(ADP-ribose) residing at the $\alpha(1 \rightarrow 2)$ O-glycosidic bond(94). PAR-binding zinc fingers (PBZs) use a zinc-coordinated fold that recognizes the $\alpha(1 \rightarrow 2)$ O-glycosidic bond much like the WWE domain, but in some cases can also bind the phosphates and adenine ring of the more distal ADP-ribose residue (111, 112). Macrodomains bind to the terminal ADP-ribose residue of PAR, making extensive contacts throughout the ADP-ribose molecule (43, 103).

b. Different PAR-binding modules that are found in PAR-binding proteins, and their relevant structural and functional features. *Asterisk*, The consensus sequence of a PAR-binding motif defined by Pleschke *et al.*, (108).

PARylation in vivo, and thereby rapidly mediate the local response to PAR-inducing signals, such as DNA damage.

Macrodomain Folds. The evolutionarily ancient macrodomain fold is a PAR-binding module found in several proteins, including the histone variant macroH2A, the macroPARP family members, chromatin-remodelling enzymes, and the oestrogen-regulated MACROD1 protein (reviewed in ref. (113)). Much like the PBZ domain, the macrodomain recognizes PAR polymers through extensive interactions with ADP-ribose, resulting in micro- to sub-micromolar affinities (**Fig. 6 b**) (103, 114). As with PBZ domains, macro domains can be organized in tandem arrays, which increases their effective affinity for PAR in a cooperative manner. The macrodomain fold is found in ten proteins in humans, with homologues in all three taxonomic domains, as well as viruses. Although binding to ADP-ribose is the most ancient and often observed function for the macrodomain and, thus, probably its original function, it has since evolved to support poly(ADP-ribose) glycohydrolase activity, as well as catabolism of ADP-ribose-like molecules that are byproducts of NAD^+ metabolic pathways (115, 116).

Recent studies have shown that macrodomain-containing proteins are rapidly recruited to sites of PARylation (**Fig. 5c, d**) (75, 76, 117, 118). The interactions between PAR and PAR-binding proteins, as well as among the PAR-binding proteins themselves, are complex and cooperative. For example, the PBZ-containing protein APLF is a component of the macrodomain-containing ALC1 complex. APLF is a histone chaperone that promotes the binding of macroH2A to PAR at sites of DNA damage (117) and ALC1 binds rapidly to chromatin and remodels nucleosomes in order to facilitate access of DNA repair factors (75).

Other examples of such functional interplay between proteins with PAR-binding modules are likely to exist.

WWE Domains. The PAR-binding module discovered most recently is the WWE domain (Fig. 6). Although the functional link between WWE domains, PARPs and ubiquitylation has been recognized for some time, it is only recently that several WWE domains have been shown to be PAR-binding modules (91, 94, 96). WWE domains that bind the PAR model substrate *iso*-ADP-ribose *in vitro* are present in PARP-11, as well as a host of ubiquitin ligases, including RNF146, Deltex1, ULF and HUWE1, suggesting that these WWE-domain-containing ligases may be functional PAR-binding modules *in vivo*. This would further support the idea that the PARylation and ubiquitylation pathways are functionally linked. In the case of RNF146, the PAR-binding role of its WWE domain was first incorrectly ascribed to the PBM of RNF146 (97). Subsequent crystallographic and biochemical analyses have demonstrated that PAR binding of RNF146 is not through its PBM, but rather through a novel WWE domain that specifically recognizes the *iso*-ADP-ribose moiety surrounding $\alpha(1\rightarrow2)$ *O*-glycosidic bonds that are unique to, and characteristic of, PARylation (94). Importantly, other proteins containing WWE domains that show conservation in structurally important residues for *iso*-ADP-ribose interaction and do not contain PAR-binding motifs, unlike RNF146, bind with similar affinity to the *iso*-ADP-ribose moiety of PAR. Although the importance of RNF146-dependent ubiquitylation of PARylated proteins is clear, the role of other WWE domain-containing ubiquitylases has not been determined.

Perspectives and Conclusions

PARPs are a unique class of enzymes whose activity is a key component of cellular stress responses leading to physiological or pathological outcomes. As such, PARPs have recently been touted as pharmacological targets in several human disease states, especially those involving genotoxic and pro-inflammatory stress responses (for example, infection, obesity and cardiovascular disease) (**Fig. 7**). The activation of PARPs results in the PARylation of PARP target proteins, leading to the modulation of target protein activity through: inhibition of protein or nucleic acid interactions; creation of new protein interaction scaffolds; and, in some cases, ubiquitin-mediated proteolysis of PARP targets. Moreover, the dynamic production and degradation of PAR polymers can promote cell death or the recruitment of PAR-binding proteins to specific cellular loci during transcription and DNA damage. Great strides have been made over the past few years in the characterization of these regulatory processes and the underlying molecular mechanisms that control them. But many questions that are central to our understanding of PARP functions remain.

We need a better characterization of the subcellular localization, function and activity of the less well-characterized PARP family members. In addition, we need a clearer definition of the specific targets of ADP-ribosylation in cells and how they vary under different conditions. Finally, we need a broader and more integrative view of the diverse roles of PARP family members in physiology and disease, which will require the generation of cell-type-specific knockout mice for particular PARPs. The development of new molecular and chemical tools will be required to answer these questions; these include cell-permeable PARG and ARH inhibitors, PARP inhibitors that are specific to individual PARPs, and ways to analyse PAR

[See the next page for the image for Fig. 7]

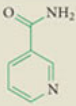
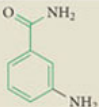
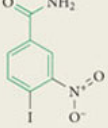
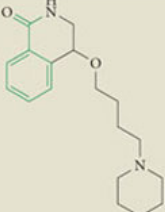
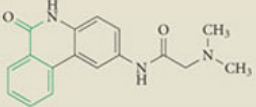
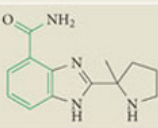
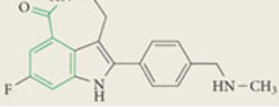
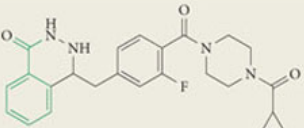
Figure 7. PARP inhibitors and their potential as therapeutic agents.

The inhibition of PARP proteins has become a promising therapeutic approach for several human diseases (119, 120). There are 89 clinical trials either recruiting, ongoing, or recently completed in the United States. The involvement of PARP proteins in the DNA damage response, telomere maintenance, and stress responses gives hope that PARP inhibition might be used in selective ‘next generation’ therapies for cancer, as well as stress-related diseases that exhibit pro-inflammatory signatures, such as cardiovascular diseases, stroke, metabolic disorders, diabetes and autoimmunity (24, 25, 121-124).

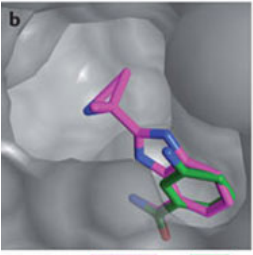
So far, most clinical trials of PARP inhibitors have focused on cancer therapy. PARP-1 can contribute to carcinogenesis by promoting cancer cell survival in the face of genotoxic insults, which may allow cells to survive and accumulate DNA damage. The rationale for PARP-directed therapeutics in cancer treatment is to selectively induce synthetic lethality of specific homologous repair-deficient (for example, BRCA1- or BRCA2-deficient) tumours by blocking PARP-1-dependent base excision repair pathways (125-127). However, the therapeutic benefits of PARP inhibitors are also evident in homologous repair-sufficient tumours, where they may decrease the apoptotic threshold in co-treatments with chemo- and radiotherapies, regulating the expression of tumour-related genes, suppressing angiogenesis, and altering gene expression programmes (22, 24, 25). Difficult to treat triple-negative (that is, ER-, PR-, HER2-negative) breast cancers may respond favorably to PARP inhibitors alone or in combination with other chemotherapeutics (128, 129). Recent setbacks with Iniparib, a supposed PARP inhibitor that had advanced the furthest in clinical trials (129, 130), are not as problematic as they may have seemed. Iniparib is not a particularly potent PARP inhibitor, and its mode of action on PARPs is not fully understood (130). In fact, recent studies suggest that it is not a *bona fide* PARP inhibitor (131, 132). Other PARP inhibitors with greater specificity and better defined mechanisms still hold great promise and are likely to be more successful.

All of the current classes of PARP inhibitors are based on the natural inhibitor nicotinamide (panel **a**), and nearly all PARP inhibitors use the competitive binding strategy first observed with nicotinamide (133). They target the nicotinamide binding pocket (panel **b**), adding other chemical groups to increase specificity and affinity (32, 134). Early PARP inhibitors (for example, nicotinamide, 3-aminobenzamide) have little specificity for individual PARP proteins, with IC₅₀ values in the hundreds of micromolar, and elicit significant off-target effects and toxicity (135). Newer PARP inhibitors have improved efficacy and specificity, with IC₅₀ values reaching the low nanomolar range and even PARP family member selectivity in some cases (for example, ABT-888 for PARP-1 and PARP-2; XAV939 for tankyrases (98, 133, 134)).

Figure 7. See the previous page for the figure legend.

a	Common or drug name*	Chemical structure
<i>Natural inhibitor (NAD⁺ metabolite)</i>		
	Nicotinamide	
<i>Benzamide</i>		
	INO-1001 (3-AB)	
	Iniparib (BSI-201)	
	DPQ	
<i>Phenanthridinone</i>		
	PJ34	
<i>Benzamidazole</i>		
	Veliparib (ABT-888)	
	Rucaparib (AG-014699)	
<i>Phthalazinone</i>		
	Olaparib (AZD-2281)	

b



PARP2 with **ABT-888** and **3-AB**
PDB ID: **3KJD** and **3KCZ**, respectively.

*Chemical identifier shown in brackets

structure and dynamics in vivo. The promise of PARP inhibitors should be better realized with a more complete understanding of the roles of each individual PARP family member and its functional interplay with the other PARP family members (**Fig. 7**). Adapting technologies from other fields, such as genomics and proteomics, as well as analogue-sensitive mutant approaches from chemical biology (*136*), will be vital in teasing out the molecular models of PARP biology. The continued study of PARPs holds great promise for answering fundamental questions in biology, as well as revealing new ways to target them for the treatment of disease.

CHAPTER THREE

CHEMICAL GENETIC DISCOVERY OF PARP TARGETS REVEALS A ROLE FOR PARP-1 IN TRANSCRIPTIONAL ELONGATION

Summary

Poly(ADP-ribose) polymerases (PARPs) are a family of enzymes that modulate diverse biological processes through covalent transfer of ADP-ribose from NAD^+ onto target proteins. Here, I report a simple and robust analog-sensitive approach for PARPs, which allows PARP-specific clickable ADP-ribosylation. Using this approach, I've mapped hundreds of protein targets and sites of ADP-ribosylation for PARPs 1, 2, and 3. I found that PARP-1 ADP-ribosylates and inhibits RNA-binding by NELF, a protein complex that regulates promoter-proximal pausing by RNA polymerase II (Pol II). I also used this analog-sensitive approach to discover genomic sites of PARP-1-mediated ADP-ribosylation and their relationship to paused Pol II. Furthermore, knockdown of PARP-1 resulted in the accumulation of paused Pol II, implicating PARP-1 in Pol II elongation through ADP-ribosylation and inhibition of NELF. The robustness and utility of this analog-sensitive approach should accelerate the discovery of new functions for ADP-ribosylation across the PARP family.

Introduction

ADP-ribosylation of proteins is an important regulator of cellular processes, from the regulation of chromatin and transcription to protein translation and stability (137). Most of the

17 PARP family members encoded in the human genome are enzymes with either mono- or poly(ADP-ribosyl) transferase activities, which covalently link ADP-ribose derived from NAD^+ to their target proteins, primarily at glutamate, aspartate, and lysine residues (138). ADP-ribose, especially when polymerized, is a large and negatively charged moiety, which can alter the activities and intermolecular interactions of PARP target proteins. PARPs 1, 2, and 3, collectively referred to as the DNA-dependent PARPs, are a group of nuclear proteins with DNA-dependent mono (PARP-3) or poly (PARPs 1 and 2) ADP-ribosyl transferase activities (138). They have been implicated in a number of nuclear processes, from chromosome maintenance and DNA repair to chromatin regulation and gene expression (15, 139). Previous studies aimed at identifying the targets of PARP family members have historically relied on immune-affinity enrichment using a monoclonal antibody against poly(ADP-ribose) (PAR) comprising >10 ADP-ribose units (140-142), with more recent approaches using a variety of affinity resins (141, 142), protein microarrays (143, 144), or a chemical-genetics approach using a surface-facing mutation in the PARP catalytic domain, which ablates poly(ADP-ribosyl) transferase activity (145). While these studies have provided a first glance of the ADP-ribosylated proteome, the lack of a single, robust, and transferable methodology to identify the sites of ADP-ribosylation on target proteins modified by a specific PARP family member has hampered our understanding of the biology of the PARP enzymes.

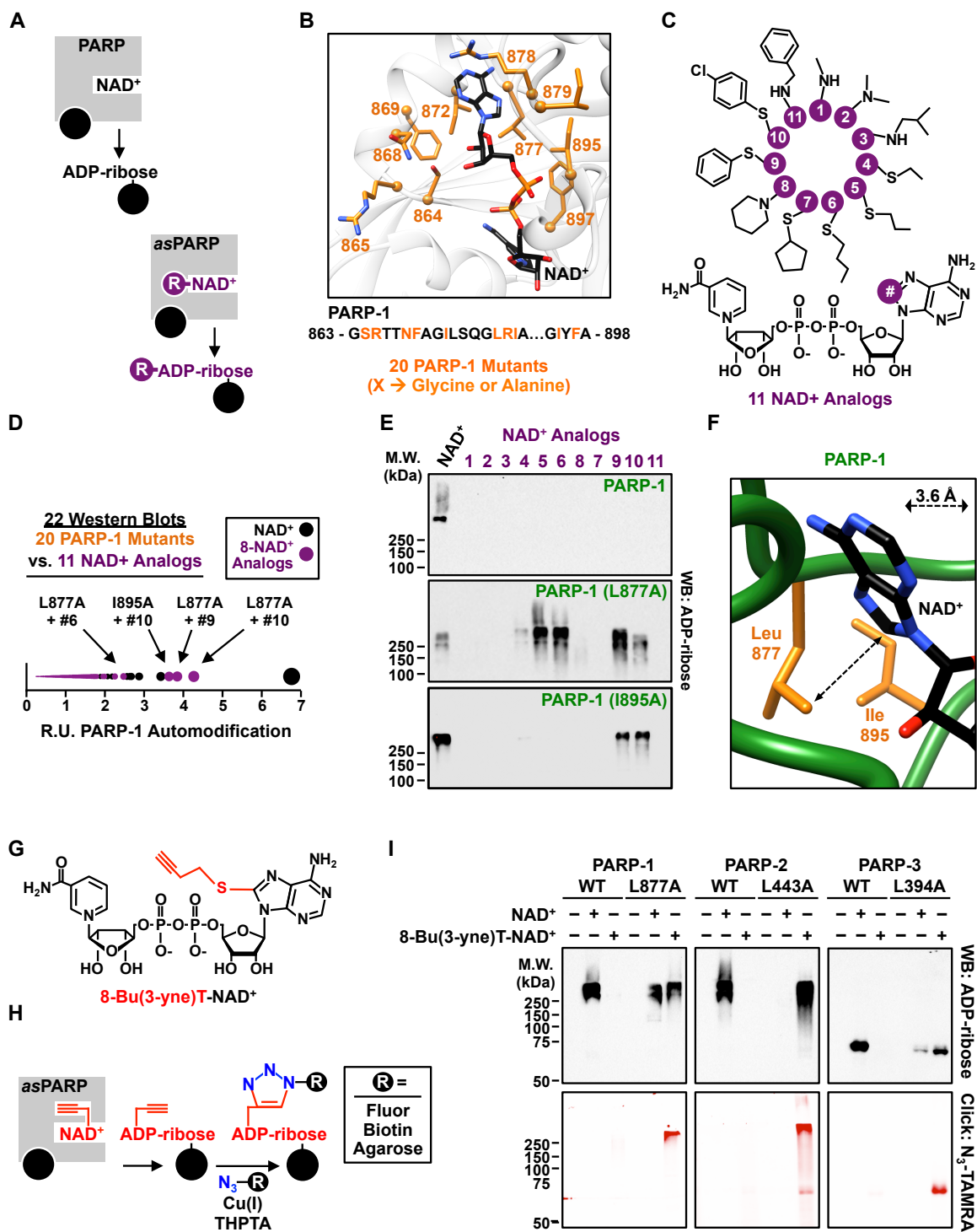
Results

In this regard, I have developed an NAD^+ analog-sensitive approach for PARPs that preserves the natural mono- or poly(ADP-ribosyl) transferase activities of the enzymes (Fig. 8A) and is

[See the next page for the image for Figure 8]

Figure 8. Structure-based engineering of an NAD⁺ analog-sensitive PARP-1 (*as*PARP-1) mutant.

- (A) Schematic illustrating the logic of engineering NAD⁺ analog-sensitivity in PARP proteins.
- (B) Depiction of residues in PARP-1, in orange, within the PARP-1 active site (*top*) and in the context of the PARP-1 amino acid sequence (*bottom*) selected for mutation to glycine or alanine for discovery of a gatekeeper position which might confer NAD⁺ analog-sensitivity.
- (C) Chemical structures of the 11 NAD⁺ analogs used in screening for analog-sensitive PARP-1 activity.
- (D) Comparison of relative automodification activity of purified wild-type and mutant PARP-1 proteins with NAD⁺ and NAD⁺ analogs (in black and purple, respectively). The PARP-1 automodification signals (in relative units; R.U.) were determined by densitometry of ADP-ribosylation-induced shifts of PARP-1 mobility in Western blot assays.
- (E) Western blot for ADP-ribose from automodification reactions containing PARP-1 or PARP-1 mutants (L877A and I895A) and NAD⁺ or NAD⁺ analogs. WB = Western blot.
- (F) Depiction of the spatial relationship between position 8 of the adenine ring in NAD⁺ and the gatekeeper residues from a structural model of the PARP-1 catalytic domain with NAD⁺ aligned in its active site (see details in Fig. S1). The dashed arrow indicates the distance, in angstroms (Å), between the position 8 and the indicated C γ of Leucine 877.
- (G) Chemical structure of the bi-functional NAD⁺ analog 8-Bu(3-yne)T-NAD⁺ with the clickable analog sensitivity-inducing, alkyne-containing R group highlighted in red.
- (H) Schematic illustrating *as*PARP activity-dependent, click chemistry-mediated covalent attachment of fluorophores, biotin, or agarose resin to 8-Bu(3-yne)T-ADP-ribosylated proteins.
- (I) Automodification reactions with wild-type or analog-sensitive PARP-1, PARP-2, and PARP-3 analyzed by Western blotting for ADP-ribose (*top*) or in-gel fluorescence (excitation: 532 nm, emission: 605 nm) following copper-catalyzed cycloaddition to azido-rhodamine (*bottom*).

Figure 8. See the previous page for the figure legend.

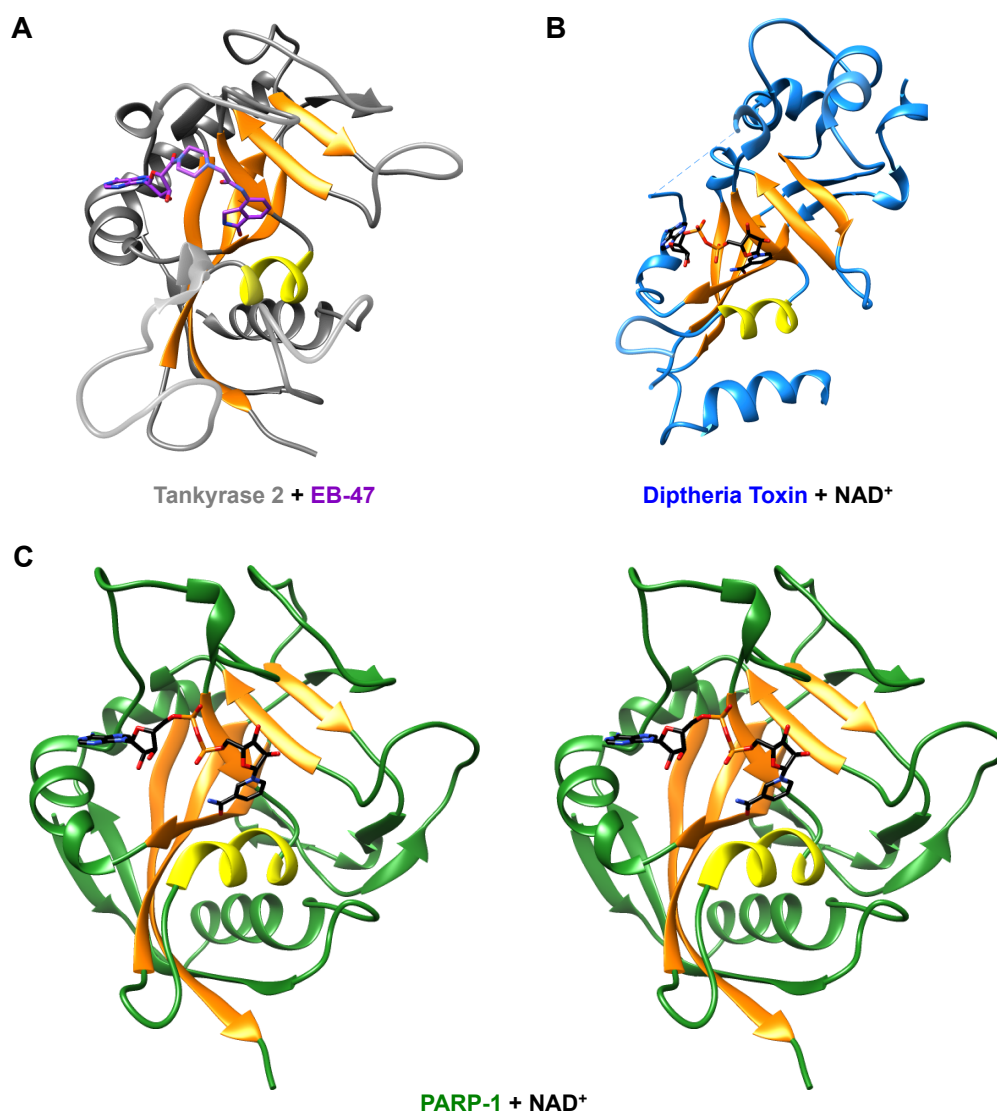


Figure 9. Structure-based alignment of NAD⁺ in the catalytic domain of PARP-1.

(A) Structure of the Tankyrase 2 catalytic domain with EB-47 (PDBID:4BJ9), a PARP inhibitor that has a chemical structure related to NAD⁺. Tankyrase 2 and EB-47 are in grey and purple, respectively. The beta-sheets and alpha-helical folds in the conserved ADP-ribosyltransferase secondary structure are colored in orange and yellow, respectively.

(B) Structure of Diphtheria toxin with NAD⁺ (PDBID:1TOX) in blue and black, respectively. The conserved ADP-ribosyltransferase secondary structure colored as in panel A.

(C) Alignment of NAD⁺ into the catalytic domain of PARP-1, based on 1TOX and 4BJ9, as well as a structure of PARP-1 with the inhibitor 3-methoxybenzamide (PDBID:3PAX), shown in a stereoscopic view. The NAD⁺ and PARP-1 are colored in black and green, respectively, with the conserved ADP-ribosyltransferase secondary structure colored as in panel A.

capable of identifying the specific targets of individual PARP family members. Analog-sensitivity is achieved by mutation of a large “gatekeeper” amino acid in the active site of a protein to a smaller residue, creating a pocket that fits a bulky R group on an engineered substrate whose interactions with the wild-type enzyme are sterically blocked (136). The analog-sensitive strategy, which creates substrate specificity within a class of enzymes that all use the same substrate and generate the same post-translational modification, remains one of the most definitive approaches for identifying direct targets of protein-modifying enzymes. In

order to engineer analog-sensitivity in PARPs, I initially focused on PARP-1, an abundant and ubiquitously expressed PARP protein. To identify a “gatekeeper” residue in PARP-1, I changed 10 large residues buried within the active site and facing the adenine ring of NAD^+ to both glycine and alanine based on a molecular model (Figs. 8B and 9). I selected the 8 position of the adenine ring of NAD^+ as the site for R group addition since its modification precludes ADP-ribosylation with wild-type PARP-1 (*wt*PARP-1) or other PARPs (146), a feature critical to the analog-sensitive approach. I then synthesized in collaboration or purchased a library of 11 NAD^+ analogs, each with a different R group at position 8, from 8-methylamino- NAD^+ to 8-benzylamino- NAD^+ (Fig. 8C). In a screen of the 20 PARP-1 mutants versus the 11 NAD^+ analogs (Figs. 8D, E and 10), I identified two different gatekeeper residues, leucine 877 and isoleucine 895, whose mutation to alanine results in analog-sensitive activity in a PARP-1 automodification assay. PARP-1 (L877A) (i.e., analog-sensitive PARP-1 or *as*PARP-1) was active with five NAD^+ analogs, whereas PARP-1 (I895A) was active with two (Fig. 8E). While L877 and I895 are 18 amino acids from one another in the PARP-1 linear sequence, they are adjacent to one another and proximal to the 8 position of the adenine ring in three dimensional

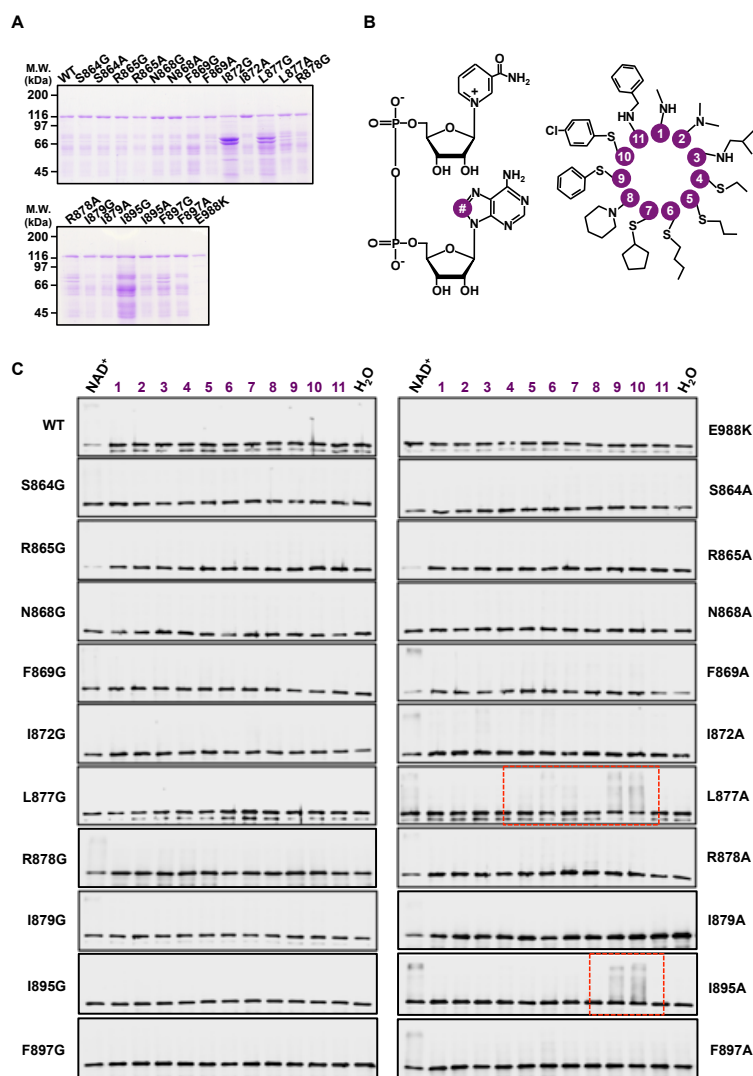


Figure 10. Screening for an NAD⁺ analog-sensitive PARP-1 mutant using an activity-based screen.

(A) Recombinant PARP-1 proteins used in the analog-sensitive PARP-1 activity screen. Expression of His₁₀-tagged wild-type PARP-1 and 21 PARP-1 site-specific mutants in *E. coli* and purification using nickel-NTA affinity chromatography. The mutants were designed around the adenine ring of NAD⁺. Coomassie stained SDS-PAGE gel of the purified proteins. Molecular weight (M.W.) markers in kilodaltons (kDa) are shown.

(B) Chemical structures of the 11 NAD⁺ analogs used in screening for analog-sensitive PARP-1 (*as*PARP-1) activity. The R groups, which are shown in the “wheel” on the right, are linked at position 8 of the adenine ring of NAD⁺ (#).

(C) Western blot analyses of the *as*PARP-1 mutant automodification reactions performed with NAD⁺ or the NAD⁺ analogs shown in panel B using an anti-PARP-1 antibody. Dashed red boxes highlight positive hits from the screen as indicated by an auto PARylation-induced migration shift of PARP-1 with the NAD⁺ analog.

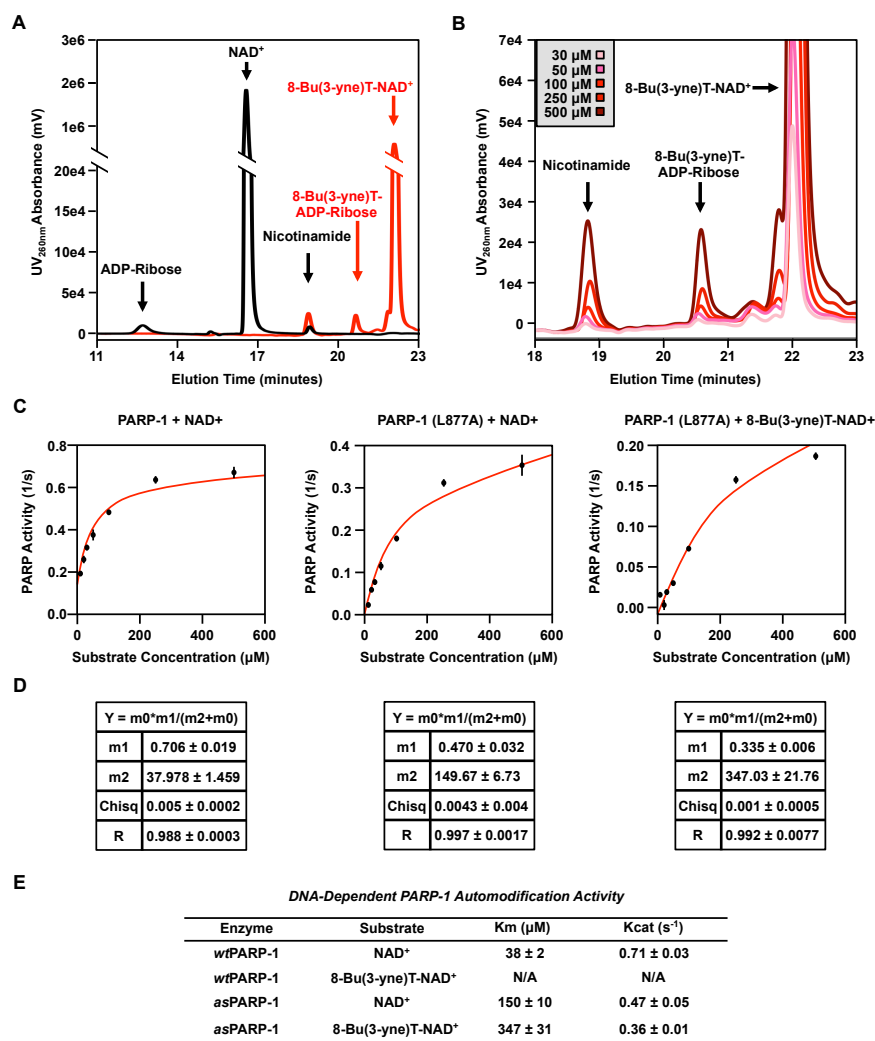


Figure 11. Enzyme kinetics of *as*PARP-1 with 8-Bu(3-yne)T-NAD⁺.

(A) LC-HPLC traces of NAD⁺- or 8-Bu(3-yne)T-NAD⁺-derived small molecules from automodification reactions containing PARP-1 with 500 μM NAD⁺ (black traces) or PARP-1 (L877A) with 100 μM 8-Bu(3-yne)T-NAD⁺ (red traces).

(B) LC-HPLC traces showing 8-Bu(3-yne)T-NAD⁺ concentration-dependent production of small molecules produced from an automodification reaction containing PARP-1 (L877A) with the NAD⁺ analog.

(C) Saturation curves for PARylation activity with PARP-1 and NAD⁺ (left), PARP-1 (L877A) and NAD⁺ (middle), and PARP-1 (L877A) and 8-Bu(3-yne)T-NAD⁺ (right). No saturation was detectable for PARP-1 and 8-Bu(3-yne)T-NAD⁺ (data not shown).

(D) Quantification of Michaelis-Menten enzyme kinetics parameters from the saturation curves shown in panel C for PARP-1 and NAD⁺ (left), PARP-1 (L877A) and NAD⁺ (middle), and PARP-1 (L877A) and 8-Bu(3-yne)T-NAD⁺ (right).

(E) Summary of Michaelis-Menten enzyme kinetics parameters for PARP-1 and PARP-1 (L877A) with NAD⁺ and 8-Bu(3-yne)T-NAD⁺.

space (Fig. 8F). These results support our molecular model of PARP-1 interaction with NAD⁺, as well as the structural basis for our *as*PARP approach.

To extend the utility of the *as*PARP approach, I functionalized the R group of NAD⁺ analog 6 (Fig. 8C), 8-Butylthio-NAD⁺, with an alkyne to generate 8-Bu(3-yne)T-NAD⁺ (Fig. 8G). 8-Bu(3-yne)T-NAD⁺ is a “clickable” NAD⁺ analog with a single bi-functional R group at position 8, facilitating *as*PARP-selective ADP-ribosylation, as well as alkyne incorporation into the post-translationally modified target for subsequent use in azide-alkyne cycloaddition reactions to label or purify the PARP targets (Fig. 8H). PARP-1 (L877A) with 8-Bu(3-yne)T-NAD⁺ yields similar activity as previously screened analogs, nearing wild-type enzyme kinetics when compared to PARP-1 and NAD⁺ (Fig. 11). Critically, this clickable NAD⁺ analog also exhibits activity with *as*PARP-2 and *as*PARP-3 mutants (L443A and L394A, respectively), which contain alanine substitutions at residues homologous to L877 of PARP-1 (Figs 8I and 12 A, B). The ability to transfer this analog-sensitive activity with 8-Bu(3-yne)T-NAD⁺ by homology at the conserved gatekeeper residue (Figs 8I and 12 D-G) suggests broad utility of this approach across the PARP family, for both mono- and poly(ADP-ribosyl) transferases. In subsequent studies, described below, I used the *as*PARP approach to identify the protein targets of specific PARPs, as well as the sites of PARP-1-mediated ADP-ribosylation across the genome.

We used the *as*PARP approach to identify site-specific nuclear targets of PARPs 1, 2, and 3 with an approach that focuses on glutamate and aspartate residues. I incubated purified recombinant *as*PARPs 1, 2, or 3 with HeLa cell nuclear extract in the presence of 8-Bu(3-yne)T-NAD⁺, which resulted in PARP-specific labeling of extract proteins (Fig. 13A). I then clicked

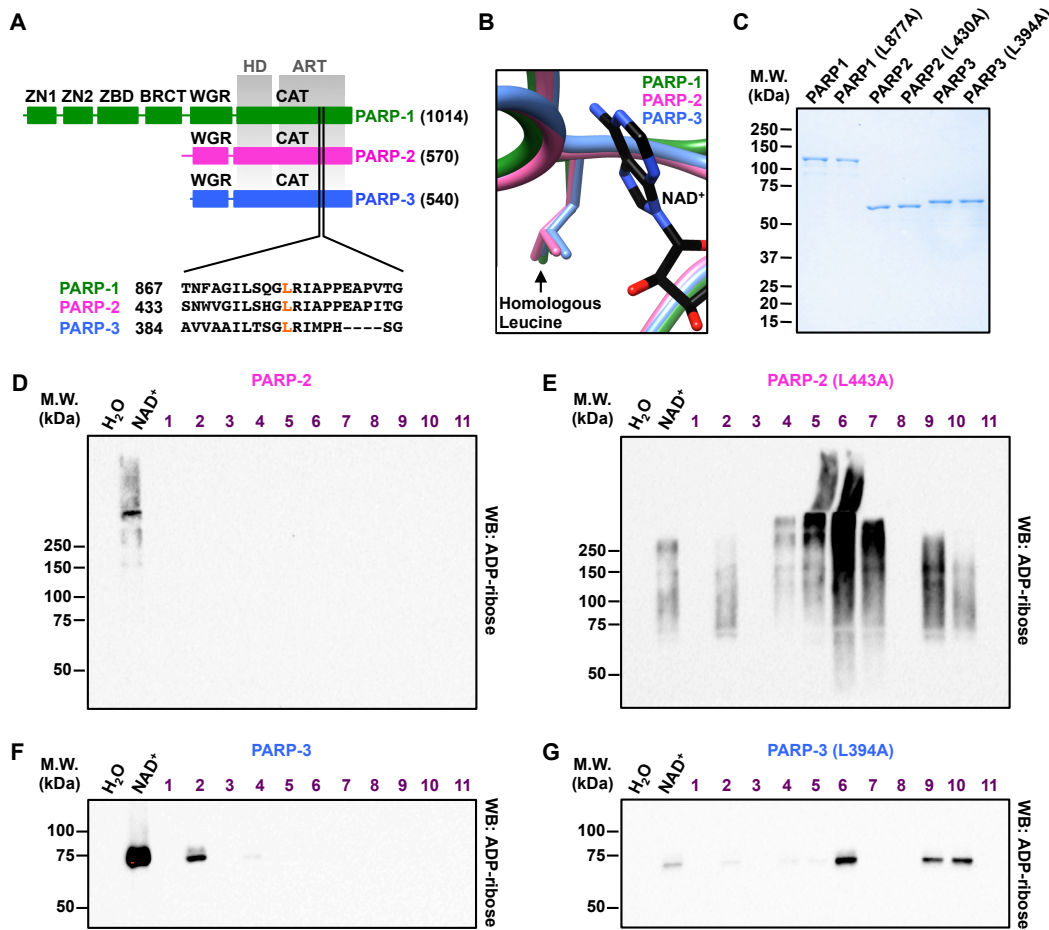


Figure 12. Transfer of NAD⁺ analog sensitivity to other PARP family members.

(A) Schematic of the domain structures of PARP-1, PARP-2, and PARP-3 aligned based on the catalytic domain (CAT) (top). Amino acid sequences flanking the homologous gatekeeper residue (L, in orange) in PARP-1, PARP-2, and PARP-3 (bottom). PARP-1, PARP-2, and PARP-3 are colored in green, pink, and blue, respectively. ZN1, zinc finger 1; ZN2, zinc finger 2; ZBD, zinc binding domain; WGR, WGR repeat domain; CAT, catalytic domain; HD, helical domain; ART, ADP-ribosyltransferase domain.

(B) Depiction of the binding site for NAD⁺ (black), which was modeled into PARP-1 (green), PARP-2 (pink), and PARP-3 (blue). The homologous leucine gatekeeper residues within the PARP-1, PARP-2, and PARP-3 catalytic domains are indicated.

(C) SDS-PAGE analysis, with subsequent staining using Coomassie blue, of purified FLAG-tagged PARP-1, PARP-1 (L877A), PARP-2, PARP-2 (L443A), PARP-3, and PARP-3 (L394A) expressed in Sf9 insect cells. Molecular weight (M.W.) markers in kilodaltons (kDa) are shown.

(D through G) Western blot analyses of wild-type and analog-sensitive PARP-2 and PARP-3 automodification reactions performed with NAD⁺ or the NAD⁺ analogs shown in Fig. S2B using an ADP-ribose detection reagent. (D) PARP-2, (E) PARP-2 (L443A), (F) PARP-3, and (G) PARP-3 (L394A).

[See the next page for the image for Figure 13]

Figure 13. Using analog-sensitive PARP-1 mutants to unambiguously identify the ADP-ribosylation targets of DNA-dependent PARPs.

(A) In-gel fluorescence (ex: 532nm, em: 605nm) of HeLa cell nuclear extract proteins conjugated to azido-TAMRA using copper-catalyzed cycloaddition following 8-Bu(3-yne)T-ADP-ribosylation reactions with 8-Bu(3-yne)T-NAD⁺ in the presence of wild-type (*wt*) or analog-sensitive (*as*) PARP-1, PARP-2, or PARP-3.

(B) Depiction of the strategy for LC-MS/MS detection of PARP-specific ADP-ribosylation sites. HeLa cell nuclear extract (N.E.) is incubated with a single purified recombinant analog sensitive PARP (*as*PARP) in the presence of 8-Bu(3-yne)T-NAD⁺. Following in vitro modification, the extract proteins are covalently linked to azido-agarose beads via copper-catalyzed cycloaddition. The conjugated beads are washed, trypsinized to release peptides for protein identification, and then washed again. The remaining peptides containing ADP-ribosylation sites are eluted from the resin using hydroxylamine (NH₂OH). The cleaved modification produces a 15.0109 Dalton increase in mass yielding a signature *m/z* change during LC-MS/MS identifying the specific site of glutamate or aspartate modification. Both the tryptic digest and hydroxylamine eluate are subjected to LC-MS/MS analysis.

(C) The number of unique ADP-ribosylation sites and protein targets identified for PARP-1, PARP-2, and PARP-3 using the LC-MS/MS approach illustrated in (B).

(D) Venn diagram depicting the overlap of the protein targets of PARP-1, PARP-2 and PARP-3.

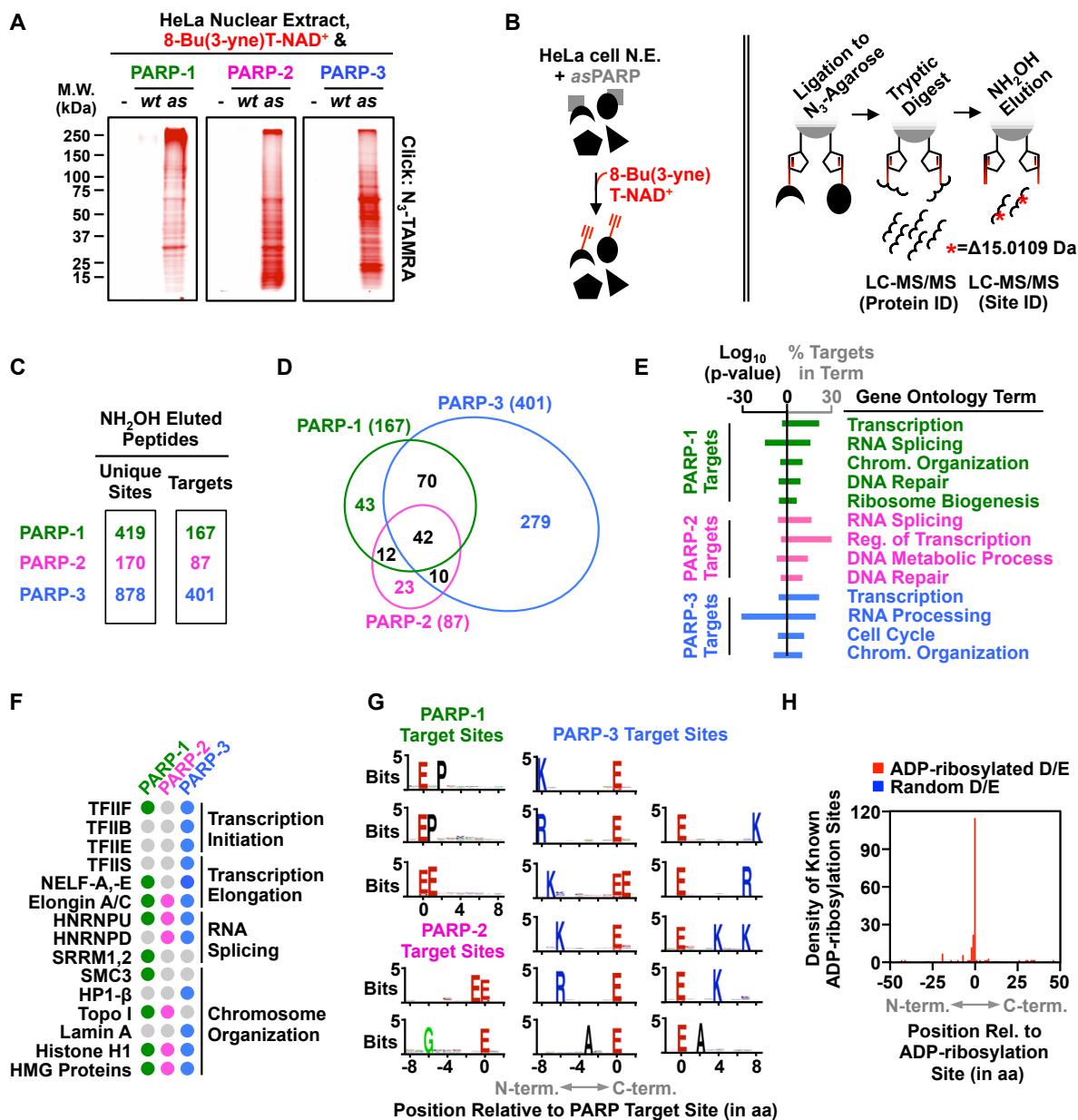
(E) Gene ontology terms enriched for the sets of PARP-1, PARP-2 and PARP-3 targets, with both the *p*-value and percent of total targets included for selected enriched GO terms.

(F) Selected targets of PARP-specific ADP-ribosylation identified using *as*PARP-1, *as*PARP-2, and *as*PARP-3.

(G) Enriched amino acid sequences \pm 8 residues on either side of identified PARP-1, PARP-2, and PARP-3 ADP-ribosylation sites. aa = amino acids.

(H) Histogram of the two-dimensional relationship between previously identified PARylation sites (142) and ADP-ribosylation sites identified using the PARP-specific LC-MS/MS approach described herein.

Figure 13. See the previous page for the figure legend.



the 8-Bu(3-yne)T-ADP-ribose-labeled proteins to azide-agarose resulting in their covalent attachment to the agarose resin, allowing extensive washing with denaturants, strong detergents, and organic solvents. After isolation of the PARP-specific ADP-ribosylated proteins, I performed trypsin-based peptide identification by LC-MS/MS (Peptide ID), washed extensively again, and eluted the ADP-ribosylated peptides using hydroxylamine to identify the sites of ADP-ribosylation by LC-MS/MS (Site ID), as previously described (*142*) ([Fig. 13B](#)). This approach revealed unique, as well as overlapping, sites of PARP-1-, 2-, and 3-mediated ADP-ribosylation ([Fig. 13C, D](#)). Ontological analyses of the target proteins revealed enrichment of terms related to transcription and DNA-repair, consistent with the known biology of PARPs 1, 2, and 3 ([Fig. 13E, F](#)). In addition, I observed an enrichment of terms and specific target proteins related to RNA splicing and processing (all three PARPs), DNA metabolism (PARP-2), and cell cycle regulation (PARP-3) ([Fig. 13E, F](#)), suggesting new functions for these PARPs.

Motif analyses at the sites of PARP-1-, 2-, and 3-mediated ADP-ribosylation indicate some similarities in sequence preference among the three PARPs (e.g., glutamate proximal to the site of modification, labeled as position “0” in [Fig. 13G](#)), but differences as well (e.g., lysine or arginine 6-8 amino acids N-terminal or 4-8 amino acids C-terminal to the site of modification for PARP-3 only; [Fig. 13G](#)). PARP-1 modification sites are enriched for proline 1-2 amino acids C-terminal to the ADP-ribosylated glutamates ([Fig. 13G](#)). Interestingly, this same preference for proline was observed when profiling bulk ADP-ribosylation sites during a response to oxidative DNA damage (*142*), which primarily stimulates PARP-1 activity. The sites of PARP-1-, 2-, and 3-mediated ADP-ribosylation that I identified herein partially overlapped and were more numerous than sites of ADP-ribosylation identified using other approaches ([Fig. 14](#)), although

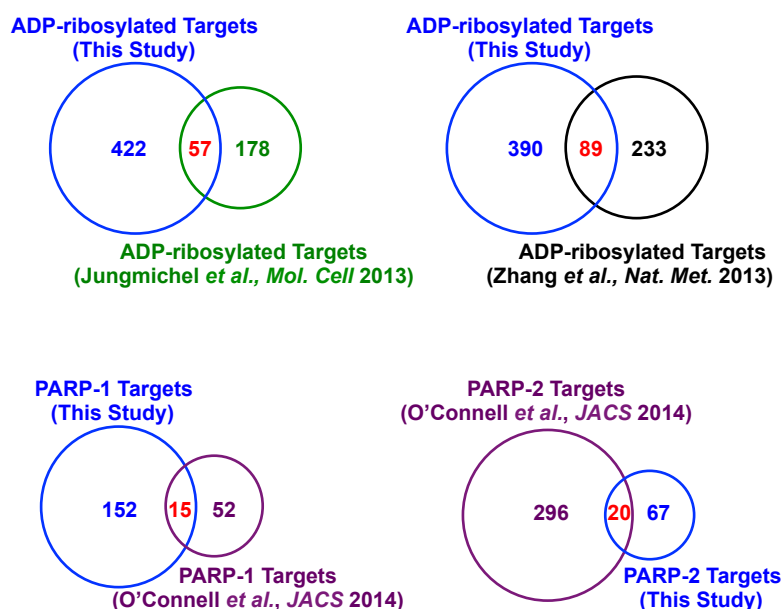


Figure 14. Overlap of PARP targets with targets reported in the literature.

Venn diagram representations of overlap between PARP-1, PARP-2, and PARP-3 targets identified in this study with targets identified using other approaches reported in the literature: ADP-ribosylated targets - Jungmichel *et al.* (2013) (141) and Zhang *et al.* (2013) (142); PARP-1 targets and PARP-2 targets - O'Connell *et al.* (2014) (145).

[See the next page for the image for Figure 15]

Figure 15. P-TEFb-dependent ADP-ribosylation of NELF by PARP-1.

(A) Cumulative distribution of enriched 7-mer amino acid sequences ± 50 residues from PARP-1, 2, and 3 ADP-ribosylation sites. The PARP-1 target-predictive sequence RSRSRDR is highlighted.

(B) Schematic showing the distribution of PARP-1 ADP-ribosylation sites, P-TEFb phosphorylation sites, and the PARP target-enriched 7-mer RSRSRDR on proteins in the NELF complex.

(C) Histogram of the two-dimensional relationship between ADP-ribosylation sites identified herein and the nearest incidence of known phosphorylation modifications on PARP target proteins.

(D) Western blot analysis of immunoprecipitated FLAG-tagged NELF-E or GFP from 293T cells.

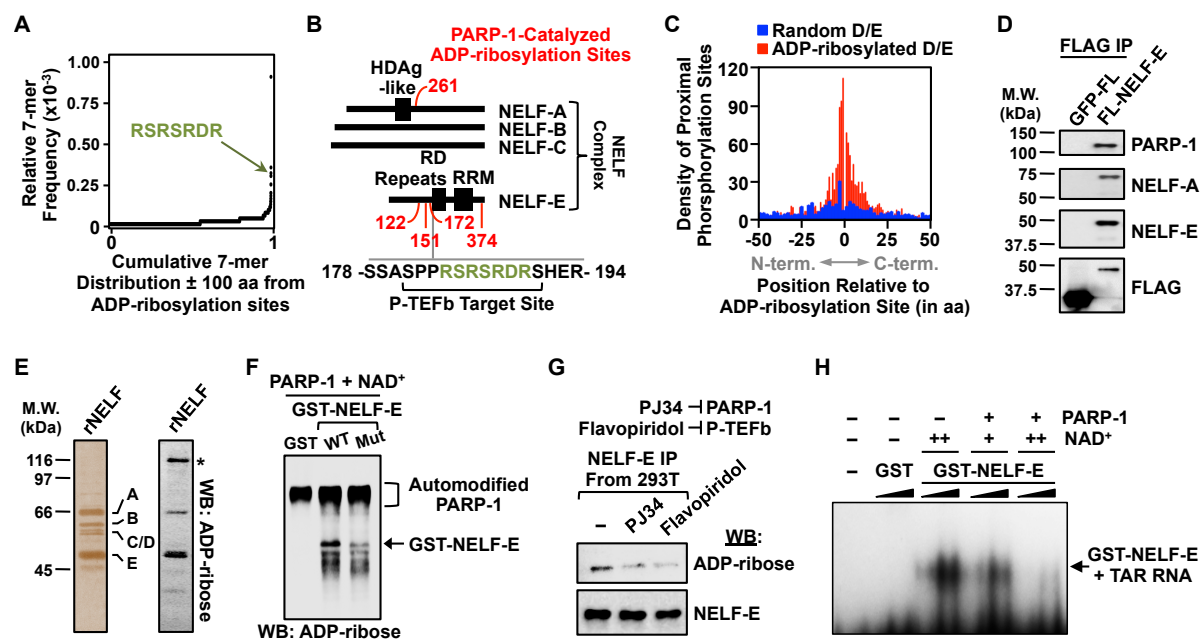
(E) Silver stained SDS-PAGE gel (*left*) and ADP-ribose Western blot (*right*) of immunopurified NELF complex showing ADP-ribosylated proteins migrating at the expected molecular weights of NELF-E and NELF-A. The asterisk indicates an ADP-ribosylated protein migrating at the expected molecular weight of PARP-1.

(F) Western blot for ADP-ribose of in vitro modification reactions containing GST, GST-tagged wild-type NELF-E, or GST-tagged ADP-ribosylation site point mutant NELF-E, PARP-1, and NAD^+ as indicated.

(G) Western blot analysis of immunoprecipitated FLAG-tagged NELF-E from 293T cells treated with vehicle, the PARP inhibitor PJ34, or the P-TEFb/CDK9 inhibitor flavopiridol.

(H) NELF-E/TAR RNA electrophoretic mobility shift assay with or without PARP-1-mediated ADP-ribosylation. GST or GST-NELF-E was titrated between 0.1 to 1.0 μM and NAD^+ was added at 25 μM (+) or 100 μM (++) during the ADP-ribosylation reaction.

Figure 15. See the previous page for the figure legend.



good agreement for the specific sites of ADP-ribosylation was observed between the results of our *as*PARP biochemical approach and a previous cell-based bulk ADP-ribosylation assay for common targets (142) (Fig. 13H). Collectively, these results demonstrate the power of our *as*PARP approach to robustly and faithfully identify sites of ADP-ribosylation mediated by a specific PARP family member.

To explore in more detail potential sequence-based “rules” for ADP-ribosylation, I determined the enrichment of 7-mer amino acid sequences near sites of PARP-1-, 2-, and 3-mediated ADP-ribosylation. I observed enrichment of a number of different sequences, including RSRSRDR (Fig. 15A). Interestingly, over half of the occurrences of this 7-mer within the human genome are targets of PARP-1 (not shown), including NELF-E (Figs. 13F and 15B), a subunit of the NELF complex whose function is to restrict transcriptional elongation and stimulate promoter-proximal pausing by RNA polymerase II (Pol II) (147). In NELF-E, the motif is located near a cluster of PARP-1- release of paused Pol II into productive elongation through phosphorylation of Pol II, DSIF, and NELF (148, 149) (Fig. 15B). Interestingly, phosphorylation sites (Fig. 15C) and, to a lesser mediated ADP- ribosylation sites and is within the previously identified phosphorylation target site for P-TEFb, a complex containing CDK9 and cyclin T1, which coordinates the extent, sites of other post-translational modifications (Fig. 16), are frequently found at or near ADP-ribosylation sites across the proteome. This suggests a broad role for ADP-ribosylation as a modulator at hubs of regulatory activity, as well as a more specific regulatory role for ADP-ribosylation (and PARPs) in cooperation with phosphorylation (and kinases) across the proteome.

Given previous reports implicating the *D. melanogaster* homolog of PARP-1 as a key

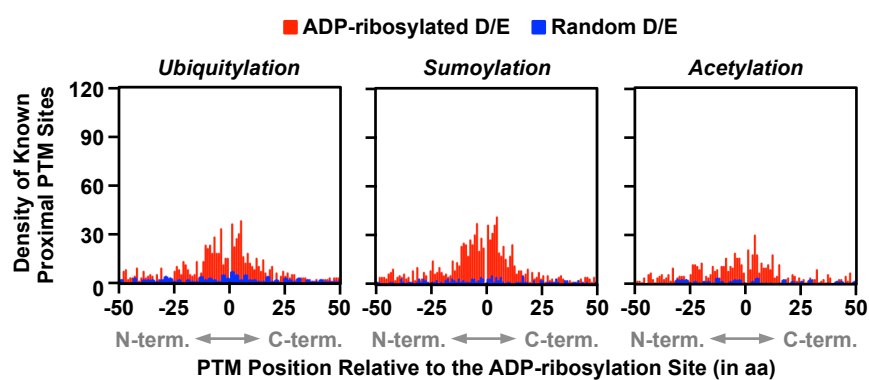


Figure 16. Other post-translational modifications associated with PARP target modification sites.

Histograms of the two-dimensional relationship between ADP-ribosylation sites identified herein and the nearest incidence of any known sites of ubiquitylation (*left*), sumoylation (*middle*), and acetylation (*right*) on PARP target proteins.

[See the next page for the image for Figure 17]

Figure 17. Functional links between PARP-1-catalyzed ADP-ribosylation, NELF binding, and RNA polymerase II pausing genome-wide.

(A) Schematic representation of Click-ChIP-seq, an *as*PARP-1-based method for identifying the genome-wide distribution of PARP-1-catalyzed ADP-ribosylation.

(B) Genome browser view of a multi-gene locus of the mouse genome showing PARP-1-catalyzed ADP-ribosylation (from Click-ChIP-seq), NELF-E and H3K4me3 enrichment (from ChIP-seq), and transcription (from GRO-seq).

(C) Genome-wide correlations between the enrichment of chromatin- and transcription-related proteins, histone modifications, and PARP-1-catalyzed ADP-ribosylation. Asterisks indicate proteins with previously reported physical or genetic interactions with PARP-1 (150-152).

(D) Heatmap representations showing PARP-1-catalyzed ADP-ribosylation (from Click-ChIP-seq), NELF-E, PARP-1, and H3K4me3 enrichment (from ChIP-seq), and transcription (from GRO-seq) at the promoters of all RefSeq genes [± 5 kb on either side of the annotated transcription start sites (TSSs)]. The data, which are from MEF cells, are ordered bottom to top based on increasing transcription.

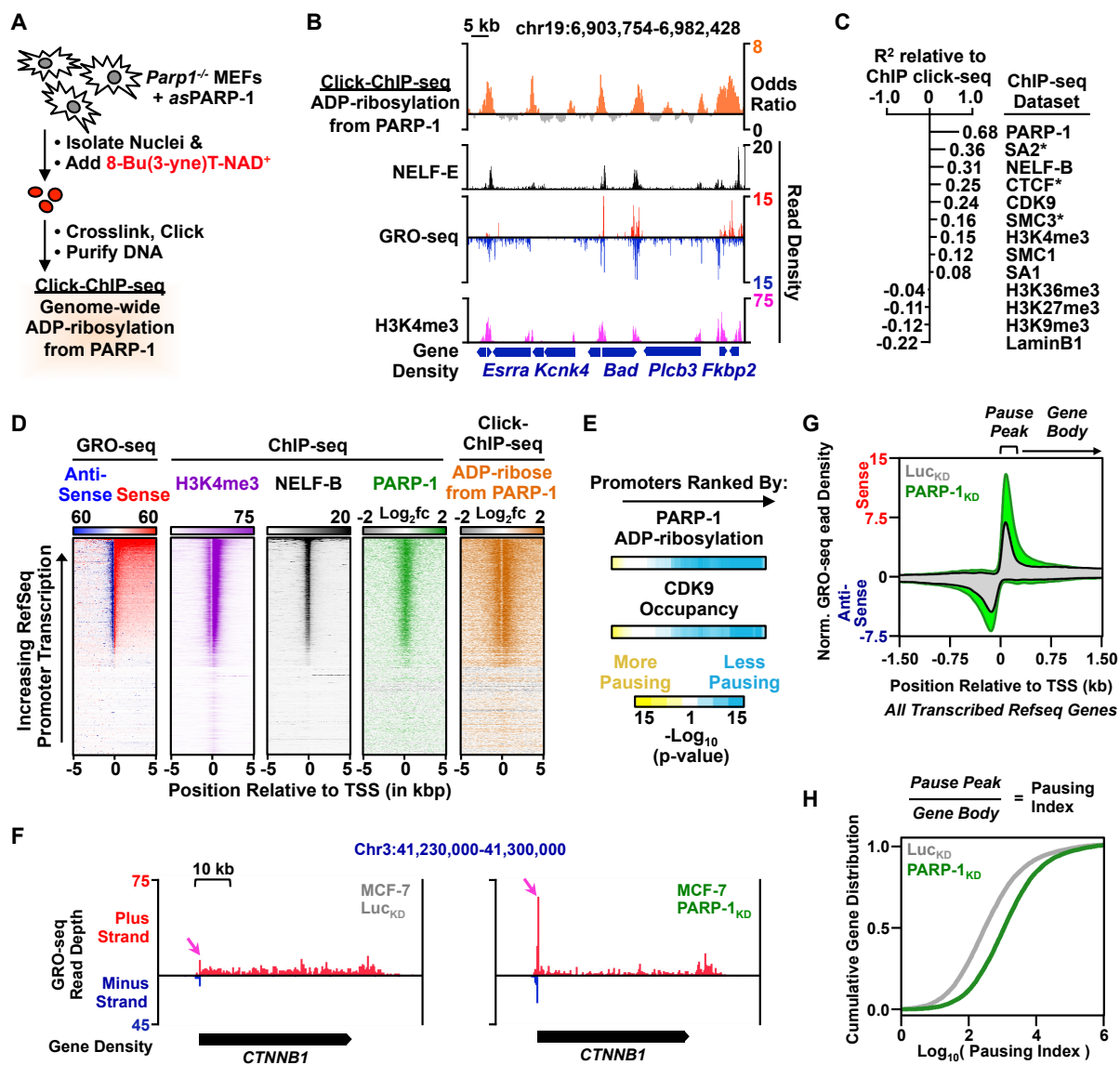
(E) Heatmap representation of RNA polymerase II pausing indices at RefSeq promoters ranked by PARP-1 activity or CDK9 occupancy.

(F) Genome browser view of a representative locus of the human genome containing the gene *CTNNB1*, showing normalized GRO-seq read density from MCF-7 cells subjected to shRNA-mediated knockdown with either control/luciferase (Luc_{KD}, *left*) or PARP-1 (PARP-1_{KD}, *right*) shRNAs. Pink arrows indicate the location of the altered peak of paused Pol II.

(G) Metagene of GRO-seq read density at the promoters of all expressed RefSeq genes from MCF-7 cells subjected to shRNA-mediated knockdown with either control/luciferase or PARP-1 shRNAs.

(H) RNA polymerase II pausing indices at the promoters of all transcribed RefSeq genes from MCF-7 cells subjected to shRNA-mediated knockdown with either control/luciferase or PARP-1 shRNAs.

Figure 17. See the previous page for the figure legend.



modulator of Pol II pause release at heat shock loci (60, 153), the identification of NELF-E as an ADP-ribosylated PARP-1 target protein led us to explore the functional interplay between these two proteins in more detail. Immunoaffinity purification of NELF from mammalian cells expressing FLAG epitope-tagged NELF-E indicates that PARP-1 interacts with the NELF complex (Fig. 15D), and that NELF-E and NELF-A are ADP-ribosylated in mammalian cells (Fig. 15E). I confirmed that the four glutamate residues that I identified in our proteomic screen (E122, E151, E172, and E374; Fig. 15B) are indeed major sites of ADP-ribosylation by PARP-1 using an *in vitro* modification assay with purified PARP-1 and ADP-ribosylation site mutant GST-tagged NELF-E (Fig. 15F). Mutation of these glutamates to glutamines, a structurally similar residue refractory to ADP-ribosylation, resulted in a substantial reduction in NELF-E modification by PARP-1 (Fig. 15F). Using a PARP inhibitor (i.e., PJ34) and a CDK9 inhibitor (i.e., flavopiridol), I observed that ADP-ribosylation of NELF-E in mammalian cells is dependent on phosphorylation by CDK9/P-TEFb (Fig. 15G), a kinase that phosphorylates Pol II, DSIF, and NELF-E. Finally, using an electrophoretic mobility shift assay with a model NELF-E-interacting RNA (i.e., HIV TAR), I found that ADP-ribosylation of NELF-E ablates its ability to bind RNA, a function of NELF-E necessary for the establishment paused Pol II (154) (Fig. 15H).

PARP-1 is a key regulator of gene expression outcomes in a variety of biological systems, modulating chromatin structure through its nucleosome-binding activity, and regulating components of chromatin and the transcriptional machinery through its catalytic activity (155). To define sites of PARP-1-mediated ADP-ribosylation across the genome, I developed an assay, which I call “Click-ChIP-seq” (click chemistry-based chromatin isolation and

precipitation with deep sequencing), using the *as*PARP-1 approach in nuclei. I re-expressed GFP (as a control), *wt*PARP-1, or *as*PARP-1 in *Parp1*^{-/-} mouse embryo fibroblasts (MEFs) (Figs. 17A and 18A). ADP-ribosylation following addition of 8-Bu(3-yne)T-NAD⁺ was clearly evident in the nuclei of *Parp1*^{-/-} MEFs expressing *as*PARP-1, but not *wt*PARP-1 (Fig. 18B). I then (1) collected 8-Bu(3-yne)T-NAD⁺-treated nuclei, (2) crosslinked them with formaldehyde, (3) and (6) purified the genomic DNA from those complexes (Figs. 17A and 18C). A qPCR-based assay of the enriched genomic DNA revealed *as*PARP-1-specific ADP-ribosylation at gene clicked the 8-Bu(3-yne)T-ADP-ribose to biotin, (4) sheared the chromatin by sonication, (5) affinity purified the 8-Bu(3-yne)T-ADP-ribose-chromatin complexes using streptavidin-agarose, promoters in nuclei isolated from MEFs (Fig. 18D). To explore chromatin-associated ADP-ribosylation genome-wide, I subjected the enriched genomic DNA to deep sequencing (Fig. 17A).

Click-ChIP-seq revealed robust enrichment of PARP-1-mediated ADP-ribosylation at the promoters of transcriptionally active genes, which were defined by an enrichment of histone H3 lysine 4 trimethylation (H3K4me3, a mark of active promoters, from ChIP-seq) and actively transcribing Pol II (from GRO-seq) (Fig. 17B). Genome-wide correlation analyses between PARP-1-mediated ADP-ribosylation and a variety of other histone modifications and chromatin/transcription-related factors revealed positive correlations with PARP-1, NELF-B, and CDK9, as well as components of a CTCF-cohesion (SMC1 and SA1/SA2) complex thought to be regulated by PARP-1-mediated ADP-ribosylation (Fig. 17C) (156). Heatmap representations of the genomic data highlight the striking relationships at gene promoters among PARP-1-mediated ADP-ribosylation, Pol II accumulation, and H3K4me3, NELF-B, and PARP-1 enrichment (Fig. 17D). Interestingly, PARP-1-mediated ADP-ribosylation and CDK9

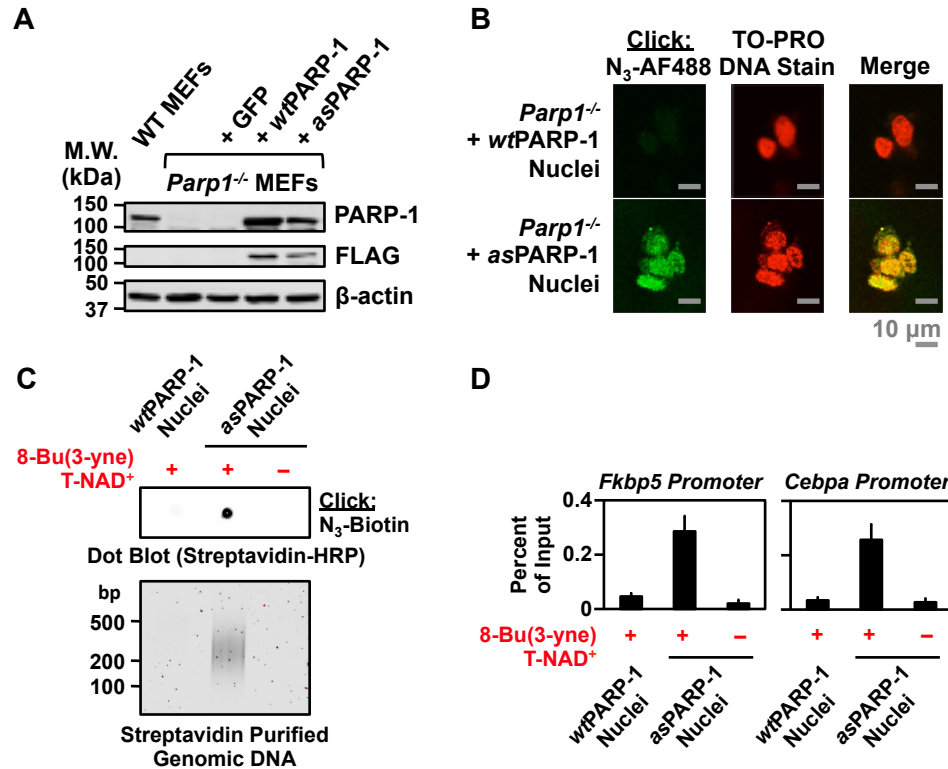


Figure 18. Development of Click-ChIP-seq, a nuclei-based assay using *as*PARP-1 technology to identify sites of PARP-1-mediated ADP-ribosylation genome-wide.

(A) Re-expression of PARP-1 in *Parp1*^{-/-} MEFs. Western blot analyses of PARP-1 expression in PARP-1 knockout MEFs expressing a GFP control, wild-type PARP-1, or analog-sensitive PARP-1.

(B) Confocal fluorescence microscopy of PARP-1-mediated ADP-ribosylation in intact nuclei from the *Parp1*^{-/-} MEFs expressing *wt*PARP-1 or *as*PARP-1 described in panel A following addition of 8-Bu(3-yne)T-NAD⁺.

(C) Streptavidin-HRP-based dot blot of nuclear proteins (*top*) and ethidium bromide-stained agarose gel of streptavidin-purified genomic DNA (*bottom*) from azido-biotin clicked nuclear samples following a reaction with 8-Bu(3-yne)T-NAD⁺ in *Parp1*^{-/-} MEFs expressing *wt*PARP-1 or *as*PARP-1.

(D) Quantitative PCR (qPCR) analysis of streptavidin-purified genomic DNA from Click-ChIP assaying for the proximal promoter regions of two expressed genes (*Fkbp5* and *Cebpa*) in *Parp1*^{-/-} MEFs expressing *wt*PARP-1 or *as*PARP-1.

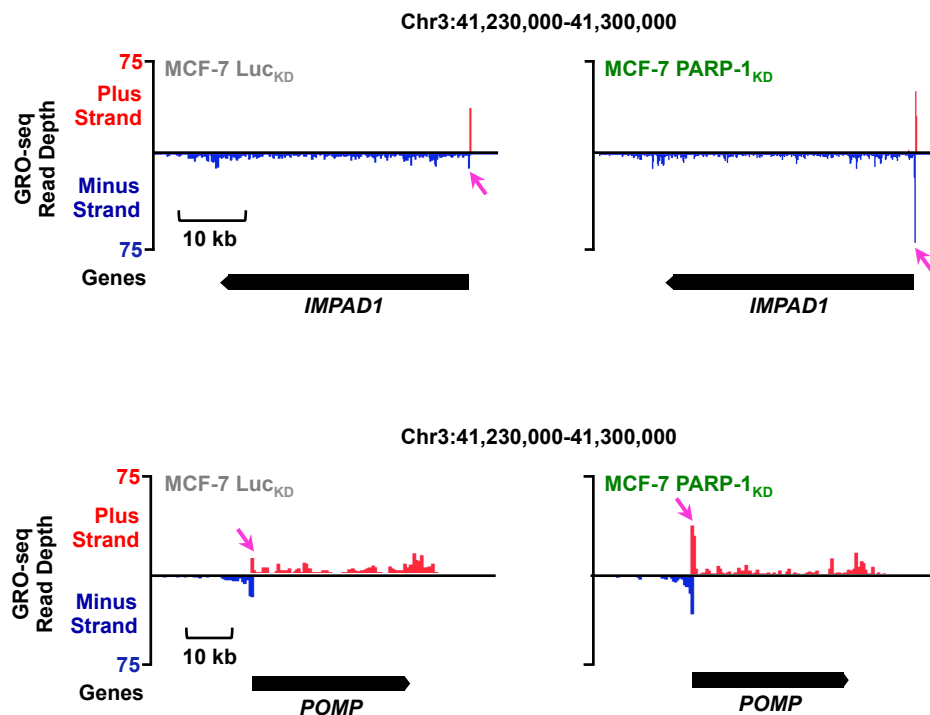


Figure 19. Altered promoter-proximal RNA polymerase II pausing at specific genes in PARP-1-deficient MCF-7 cells.

MCF-7 human breast cancer cells were subjected to knockdown (KD) with control (luciferase; Luc_{KD}) or PARP-1 (PARP-1_{KD}) shRNAs, followed by global run-on sequencing (GRO-seq). Genome browser view of two representative loci of the human genome, containing the genes *IMPAD1* (top) and *POMP* (bottom), showing normalized GRO-seq read densities from MCF-7 cells subjected to control (Luc_{KD}, left) or PARP-1 (PARP-1_{KD}, right) knockdown. Pink arrows indicate the location of the altered peak of paused Pol II.

occupancy at promoters strongly correlated with low levels of Pol II pausing (Fig. 17E). These results suggested to us that PARP-1-mediated ADP-ribosylation may act similarly to CDK9/P-TEFb-mediated phosphorylation to promote the release of paused Pol II into productive elongation.

To test this hypothesis, GRO-seq was performed in MCF-7 breast cancer cells to monitor the effects of shRNA-mediated PARP-1 knockdown on Pol II pausing. I observed a dramatic accumulation of reads in the peaks of paused Pol II upon PARP-1 knockdown (compared to a control luciferase knockdown; Luc) at gene promoters, as determined by GRO-seq (Figs. 17F and 19). This effect was evident genome-wide (Fig. 17G), with a clear increase in global Pol II pausing indices upon PARP-1 knockdown (Fig. 17H). At active promoters with a significant accumulation of GRO-seq reads in the paused Pol II peak upon PARP-1 knockdown, I observed decreased GRO-seq reads in the gene bodies, lower NELF-E occupancy at the promoters, and lower Pol II pausing prior to PARP-1 depletion (Fig. 20). These results suggest that PARP-1 modulates NELF activity at these promoters to achieve an efficient release of Pol II into productive elongation. Collectively, our data point to a clear functional link between CDK9-mediated phosphorylation, PARP-1-mediated ADP-ribosylation, and NELF-mediated Pol II pausing (Fig. 21).

Materials and Methods

Antibodies. The custom rabbit polyclonal antiserum against PARP-1 used for Western blotting and ChIP assays was generated by using an antigen comprising the amino-terminal half of PARP-1 (55) (now available from Active Motif; cat. no. 39559). The custom recombinant

[See the next page for the image for Figure 20]

Figure 20. PARP-1 modulates the levels of RNA polymerase II pausing levels genome-wide in cells.

(A) MA plot highlighting statistically significant changes (p -value < 0.001 ; red dots) in normalized pause peak read counts (first 250 bp of gene) of RNA Polymerase II transcribed genes between replicates of control or PARP-1 knockdown MCF-7 cells as determined from GRO-seq data using edgeR (157).

(B) Log_{10} (fold change) of normalized GRO-seq gene body read counts (for a region spanning 1 - 13 kb downstream of the TSS) between control and PARP-1 knockdown MCF-7 cells for genes that show (red; “Yes”) or do not show (grey; “No”) significant (p -value < 0.001) differential read counts in the Pol II pause peak region of a transcribed gene.

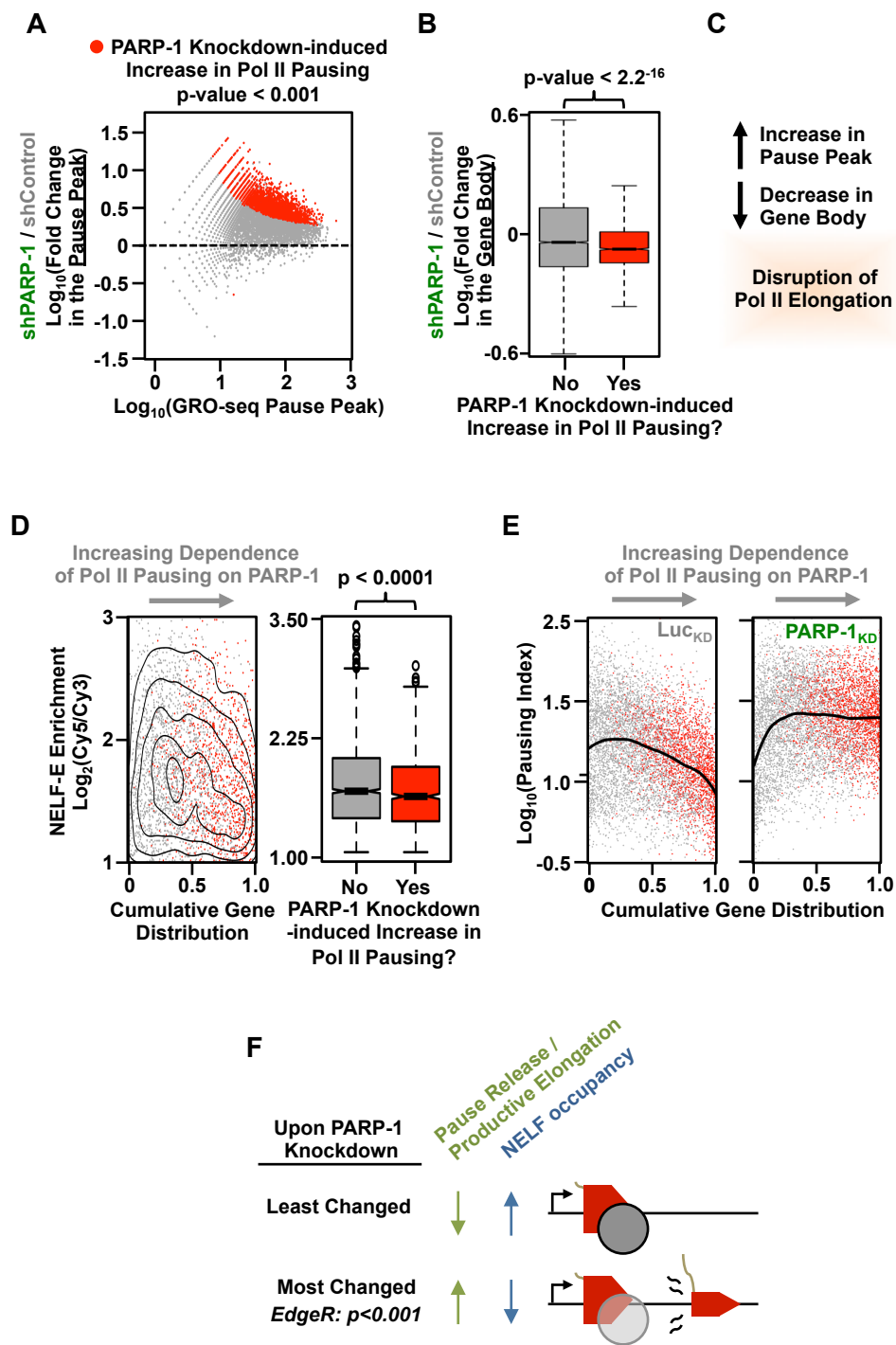
(C) Graphical representation of the relationship between Pol II pause peak and gene body read counts between control and PARP-1 knockdown cells.

(D) Intensity of NELF-E enrichment at NELF-E ChIP peaks from a ChIP-chip microarray experiment in MCF-7 cells (158) across promoters (1) ordered by the extent of dependence of RNA polymerase II pausing on PARP-1 (red, genes that show increased Pol II pausing upon knockdown of PARP-1, p -value < 0.001 ; grey, genes that do not show an increase in Pol II pausing upon knockdown of PARP-1) (*left*) or (2) segregated for genes that show (red; “Yes”) or do not show (grey; “No”) significant (p -value < 0.001) differences in RNA polymerase II pausing (*right*).

(E) Absolute value of the pausing indices for promoters ordered by the extent of dependence of RNA polymerase II pausing on PARP-1 for control (Luc_{KD} , *left*) or PARP-1 ($\text{PARP-1}_{\text{KD}}$, *right*) knockdown MCF-7 cells. Red, genes that show increased Pol II pausing upon knockdown of PARP-1, p -value < 0.001 ; grey, genes that do not show an increase in Pol II pausing upon knockdown of PARP-1.

(F) Schematic diagram summarizing the relationship between the extent of dependence of RNA polymerase II pausing on PARP-1, pausing index values, and NELF occupancy upon PARP-1 knockdown in MCF-7 cells.

Figure 20. See the previous page for the figure legend.



antibody-like anti-poly-ADP-ribose binding reagent (anti-PAR) and anti-pan-ADP-ribose binding reagent (anti-panADPR) were generated and purified in-house (now available from EMD Millipore; cat. nos. MABE1031 and MABE1016, respectively). Additional antibodies were purchased from commercial sources: M2 anti-FLAG monoclonal (Sigma; F1804), NELF-E (Santa Cruz; H-140), NELF-A (Bethyl; A301), and Pol II (Santa Cruz; N-20). For Western blotting, the primary antibodies were used at a 1:4000 dilution in 5% non-fat milk made in TBST, with subsequent detection using an appropriate HRP-conjugated secondary antibody (Pierce) used at a 1:5000 dilution in 5% non-fat milk made in TBST. NELF-E occupancy at the promoters, and lower Pol II pausing prior to PARP-1 depletion ([Fig. 20](#)). These results suggest that PARP-1 modulates NELF activity at these promoters to achieve an efficient release of Pol II into productive elongation. Collectively, our data point to a clear functional link between CDK9-mediated phosphorylation, PARP-1-mediated ADP-ribosylation, and NELF-mediated Pol II pausing ([Fig. 21](#)).

Materials and Methods

Antibodies. The custom rabbit polyclonal antiserum against PARP-1 used for Western blotting and ChIP assays was generated by using an antigen comprising the amino-terminal half of PARP-1 (55) (now available from Active Motif; cat. no. 39559). The custom recombinant antibody-like anti-poly-ADP-ribose binding reagent (anti-PAR) and anti-pan-ADP-ribose binding reagent (anti-panADPR) were generated and purified in-house (now available from EMD Millipore; cat. nos. MABE1031 and MABE1016, respectively). Additional antibodies were purchased from commercial sources: M2 anti-FLAG monoclonal (Sigma; F1804), NELF-E

(Santa Cruz; H-140), NELF-A (Bethyl; A301), and Pol II (Santa Cruz; N-20). For Western blotting, the primary antibodies were used at a 1:4000 dilution in 5% non-fat milk made in TBST, with subsequent detection using an appropriate HRP-conjugated secondary antibody (Pierce) used at a 1:5000 dilution in 5% non-fat milk made in TBST.

Alignment of NAD⁺ into the Active Site of the PARP-1 Catalytic Domain. PDB files from co-crystal structures of Diphtheria toxin with NAD⁺ (PDBID:1TOX), Tankyrase 2 with EB-47 (PDBID:4BJ9), and PARP-1 with 3-methoxybenzamide (PDBID:3PAX) were downloaded from the RCSB Protein Data Bank for analysis. Using matchmaker in the UCSF Chimera program, the structures of Diphtheria toxin, Tankyrase 2, and PARP-1 were aligned to one another. Subsequently, an estimate of NAD⁺ binding within the PARP-1 active site was obtained in two steps: First, by aligning the nicotinamide-ribose portion of NAD⁺ from 1TOX with the nicotinamide-based inhibitor, 3-MB. Second, by positioning the adenine-ribose portion of NAD⁺ according to the atomic coordinates of adenine-ribose from the NAD⁺-like inhibitor EB-47 from 4BJ9.

NAD⁺ Analogs. The NAD⁺ analogs used herein were either purchased from, or synthesized collaboratively with, the BIOLOG Life Science Institute (LSI), Bremen, Germany. The following NAD⁺ analogs used for initial screening were purchased from the BIOLOG LSI catalog: (1) β -nicotinamide-8-methylaminoadenine dinucleotide (8-MA-NAD⁺), (2) β -nicotinamide-8-dimethylaminoadenine dinucleotide (8-DMA-NAD⁺), (3) β -nicotinamide-8-ethylthioadenine dinucleotide (8-ET-NAD⁺), (4) β -nicotinamide-8-propylthioadenine

dinucleotide (8-PrT-NAD⁺), (5) β -nicotinamide-8-butylthioadenine dinucleotide (8-BuT-NAD⁺), (6) β -nicotinamide-8-cyclopentylthioadenine dinucleotide (8-cPeT-NAD⁺), (7) β -nicotinamide-8-piperidinoadenine dinucleotide (8-PIP-NAD⁺), (8) β -nicotinamide-8-phenylthioaminoadenine dinucleotide (8-PT-NAD⁺), and (9) β -Nicotinamide-8-(4-chlorophenylthio)adenine dinucleotide (8-pCPT-NAD⁺).

Synthesis of NAD⁺ Analogs

General Reagents and Synthetic Methods. All reagents were analytical grade or the best grade available from commercial suppliers. Dimethyl sulfoxide (DMSO) was stored over activated molecular sieves (3 Å) for at least two weeks before use. Nucleotides were quantified and aliquoted using their extinction coefficient at λ_{max} at pH 7.0. Extinction coefficients were determined using a cyanide assay as described (159), with modifications. 8-Br-NAD⁺ and 8-SH-NAD⁺ (160) were synthesized as described and provided by BIOLOG LSI.

Synthesis of 8-Bu(3-yne)T-NAD⁺. Sixty μmoles of 8-SH-NAD⁺, triethylammonium salt, were dissolved in 1200 μL of dried DMSO in a 3 mL polypropylene reaction tube with a screw cap. After addition of 468.8 μmoles (44 μL , 7.8 equivalents) of 4-bromo-1-butyne and 120 μmoles (20.4 μL , 2 eq.) of diisopropylethylamine, the reaction mixture was shaken under argon atmosphere in a MHL 20 thermomixer (HLC Biotech), set at 25°C and 500 rpm. After 8 hours, no further reaction progress was detected by analytical HPLC performed with a LaChrom Elite instrument using an L-2130 pump, a L-2420 variable wavelength UV/Vis detector, a L-2350 column oven (set at 30 °C), and EZChrom software version 3.3.1 SP1 (all VWR-Hitachi) with YMC ODS-A 12 nm, S-11 μm (YMC) resin packed in a 250 x 4.6 mm stainless steel column.

HPLC was run with a mobile phase of 25 mM sodium dihydrogen phosphate, pH 6.8, 10% acetonitrile, and 4 mM tributylammonium hydrogensulfate. The reaction was quenched by the addition of 1,200 μ L 1 M sodium dihydrogen phosphate buffer, pH 4.7, and frozen at -70°C for 1 hour. All volatile components of the reaction mixture were evaporated overnight in a SpeedVac concentrator under reduced pressure with oil pump vacuum. The residue was dissolved in ~ 30 mL water, passed through a $0.45\text{ }\mu\text{m}$ regenerated cellulose filter, and purified by preparative MPLC at room temperature using a C-605 pump (Büchi), a preparative K 2001 UV-detector (Knauer), and a L200E analog recorder (Linseis) with Merck LiChroprep RP-18 6 nm, $15 - 25\text{ }\mu\text{m}$ (Merck-Hitachi) resin packed in a $435 \times 25\text{ mm}$ glass column (Kronlab) equilibrated with 20 mM triethylammonium formate (TEAF), pH 7.0. The reaction mixture was applied to the column, washed with the same buffer, and then eluted with 100 % water. The product-containing fractions with a purity of $>98\%$ (by HPLC) were exchanged to sodium salt via cation exchange MPLC with Toyopearl SP-650M, $65\text{ }\mu\text{m}$, sodium form (Tosoh Bioscience) resin packed in a $125 \times 35\text{ mm}$ glass column (Kronlab). All purified fractions were carefully evaporated under vacuum to yield $36.6\text{ }\mu\text{moles}$ 8-Bu(3-yne)T-NAD⁺, sodium salt.

Synthesis of 8-BuT-NAD⁺. Synthesis was performed in parallel reactions with $3 \times 100\text{ }\mu\text{moles}$ 8-Br-NAD⁺, sodium salt. One hundred μmoles of the starting material were dissolved in $1,280\text{ }\mu\text{L}$ of H_2O in a 10 mL polypropylene vial with a screw cap. After addition of $2,560\text{ }\mu\text{L}$ of 100 mM sodium tetraborate buffer, pH 9.0, $1,280\text{ }\mu\text{L}$ CH_3CN , $580\text{ }\mu\text{L}$ butyl mercaptan (5 mmol , 50 eq.) and $720\text{ }\mu\text{L}$ of 2.5 N NaOH (18 eq.), the reaction tube was closed tightly and placed in a thermomixer, set at 50°C and 500 rpm . After completion of the reactions ($\sim 35\text{ min}$), the solutions were transferred to a 500 mL separation funnel, 200 mL of water were added, and the

pH was rapidly titrated to pH 7.0 with 6 N HCl. The aqueous phase was extracted 3 times with 50 mL dichloromethane, followed by 3 x 30 mL ethyl acetate. Purification with preparative MPLC (product elution with 0.5 % isopropanol) and salt exchange to sodium were performed as described above for 8-Bu(3-yne)T-NAD⁺. The final product was 180.6 μmoles of 8-BuT-NAD⁺, sodium salt with a purity of 99.8 % (by HPLC) and a yield of 60 %.

Synthesis of 8-PrT-NAD⁺. Synthesis and workup was performed with propyl mercaptan as described for 8-BuT-NAD⁺ in parallel reactions with 3 x 100 μmoles of 8-Br-NAD⁺, sodium salt. Purification with preparative MPLC (product elution with 0.5 % isopropanol) and salt exchange to sodium were performed as described above for 8-Bu(3-yne)T-NAD⁺. The final product was 200.6 μmoles of 8-PrT-NAD⁺, sodium salt with a purity of 99.7 % (by HPLC) and a yield of 67 %.

Synthesis of 8-BnA-NAD⁺. Fifty-nine μmoles of 8-bromo-adenosine-5-monophosphate were mixed in 206 μL of absolute ethanol while under nitrogen, followed by the addition of 117 μL benzylamine (20 eq) and 11.8 mg of calcium carbonate (2.0 eq). The reaction was heated to reflux under nitrogen and carried out overnight. Product formation was detected by LC-MS at 260 nm. The reaction was diluted with ethanol and the insoluble calcium salt was separated by centrifugation. The supernatant was evaporated to obtain dry 8-benzylamine-adenosine-5-monophosphate, which was then purified by preparative HPLC and coupled to nicotinamide mononucleotide as previously described (146) to obtain β-nicotinamide-8-benzylaminoadenine dinucleotide.

Synthesis of 8-iBuA-NAD⁺. Fifty-nine μmoles of 8-bromo-adenosine-5-monophosphate was mixed in 206 μL of absolute ethanol while under nitrogen, followed by the addition of 117

μ L isobutylamine (20 eq) and 11.8 mg of calcium carbonate (2.0 eq). The reaction was heated to reflux while under nitrogen and carried out overnight. Product formation was detected by LC-MS at 260 nm. The reaction was diluted with ethanol and the insoluble calcium salt was separated by centrifugation. The supernatant was evaporated to obtain dry 8-isobutylamine-adenosine-5-monophosphate, which was then purified by preparative HPLC and coupled to nicotinamide mononucleotide as previously described (146) to obtain β -nicotinamide-8-isobutylaminoadenine dinucleotide.

Molecular Biology and Cloning

Human and Mouse cDNAs. cDNA pools from 293T cells (human) or 3T3-L1 cells (mouse) were prepared by extraction of total RNA from the cells using Trizol (Life Technologies), followed by reverse transcription using superscript III reverse transcriptase (Promega) and an oligo(dT) primer according to manufacturer's instructions.

Construction of Bacterial Expression Vectors. The following expression vectors were constructed for expression of PARP-1 and NELF-E in bacteria.

PARP-1: Human PARP-1 cDNA in pET19b (Novagen), described previously (55), was subjected to PCR-based site-directed mutagenesis at codons corresponding to amino acid residues 864, 865, 868, 869, 872, 877, 878, 879, 895, and 897 to alter the coding to either glycine or alanine.

NELF-E: Human NELF-E cDNA was isolated by PCR from the cDNA pools described above and cloned into the pGEX2T bacterial expression vector (GE Life Sciences), allowing expression of GST-tagged NELF-E. An ADP-ribosylation site point mutant human NELF-E

cDNA was generated by mutating codons for glutamate at amino acid positions 121, 151, 152, 171, 172, and 374 to codons for glutamine (the mutant includes two glutamate residues adjacent to two the four identified, for a total of six sites). The mutant cDNA was amplified by PCR and cloned into pGEX2T for expression of GST-tagged NELF-E with mutated PARP-1 ADP-ribosylation sites.

Construction of Insect Expression Vectors. Human PARP-2, mouse PARP-1, and mouse PARP-3 cDNAs were isolated by PCR from the cDNA pools described above, adding sequences encoding an N-terminal FLAG epitope to the cDNAs via the primers during amplification. Leucine to alanine-encoding base pair alterations at “gatekeeper” positions were introduced into the cDNAs using PCR-based site-directed mutagenesis. All constructs were sequenced ensure fidelity of the sequences. Recombinant bacmids were then prepared for Sf9 transfection via transformation into the DH10BAC *E. coli* strain with subsequent blue/white colony screening using the Bac-to-Bac system (Invitrogen) according to the manufacturer’s instructions.

Construction of Mammalian Expression Vectors. The following expression vectors were constructed for expression of PARP-1, NELF subunits, and GFP in mammalian cells.

PARP-1: PCR products for mouse PARP-1 and PARP-1 (L877A) were amplified from their cognate cDNAs, adding a sequence encoding an N-terminal FLAG epitope during the amplification process. The PCR products were then cloned into the pTY-U6 plasmid (provided by Yi Zhang, UT Southwestern Medical Center) after removal of an ORF encoding GFP-3xFLAG cDNA.

NELF subunits: Constructs for the expression of NELF subunits were generated using two different mammalian expression vectors: (1) pCDNA3 - A cDNA encoding an N-terminal HA epitope-tagged NELF-A protein followed by an IRES2 sequence was amplified using two-step PCR from 293T cDNA and the pIRES2dsRED plasmid (Clontech), respectively, and ligated into a cloning vector. A cDNA encoding N-terminal FLAG epitope-tagged NELF-E was amplified from cDNA and cloned into the NELFA-IRES2 plasmid described above. The resulting HA-NELF-A-IRES2-FLAG-NELF-E construct was then sub-cloned into pCDNA3. (2) pINDUCER - A cDNA encoding wild-type NELF-E or ADP-ribosylation site point mutant NELF-E with an N-terminal FLAG tag was amplified by PCR and cloned into the pINDUCER plasmid (Addgene).

GFP: Constructs for the expression of GFP were generated using two different mammalian expression vectors: (1) pCDNA3 - A cDNA encoding GFP with a C-terminal 3xFLAG epitope tag was sub-cloned from pTY-U6-GFP-3xFLAG into pCDNA3; (2) pINDUCER - A cDNA encoding GFP was amplified by PCR and cloned into the pINDUCER plasmid.

Expression and Purification of Recombinant Proteins

Purification of PARP-1 Expressed in E. coli. BL21(DE3) *E. coli*, transformed with a pET19b plasmid encoding wild-type PARP-1 or mutant PARP1, were grown in LB with ampicillin at 37°C to a density (OD_{600 nm}) of 0.4. Recombinant protein expression was then induced by addition of IPTG to 1 mM for 2 hours at 37°C. The cells were collected by centrifugation, and the bacterial cell pellets were flash frozen in liquid N₂ and stored at -80°C.

The cell pellets were thawed on wet ice and lysed by sonication in IMAC Lysis Buffer (10 mM Tris•HCl, pH 7.5, 0.5 M NaCl, 0.1 mM EDTA, 0.1% NP-40, 10% glycerol, 10 mM imidazole, 1 mM PMSF, 1 mM β -mercaptoethanol). The lysates were clarified by centrifugation. Recombinant PARP-1 was purified from the clarified lysate using Ni-NTA affinity chromatography, with washing of the resin using IMAC Wash Buffer (10 mM Tris•HCl, pH 7.5, 1 M NaCl, 0.2% NP-40, 10% glycerol, 10 mM imidazole, 1 mM PMSF) and elution of the PARP-1 protein using IMAC Elution Buffer (10 mM Tris•HCl, pH 7.5, 0.2 M NaCl, 0.1% NP-40, 10% glycerol, 500 mM imidazole, 1 mM PMSF, 1 mM β -mercaptoethanol). The eluates were dialyzed in IMAC Dialysis Buffer (10 mM Tris•HCl, pH 7.5, 0.2 M NaCl, 10% glycerol, 1 mM PMSF, 1 mM β -mercaptoethanol) and concentrated as needed with a centrifugal concentrator. The purified PARP-1 protein was flash frozen in liquid N₂ in single-use aliquots, at a final concentration of 1 mg/mL and stored at -80°C.

Purification of PARPs Expressed in Sf9 Insect Cells. Sf9 insect cells, cultured in SF-II 900 medium (Invitrogen), were transfected with 1 μ g of bacmid driving expression of wild-type PARP-1, PARP-1(L877A), wild-type PARP-2, PARP-2(L443A), wild-type PARP-3, or PARP-3(L394A) using Cellfectin transfection reagent as described by manufacturer (Invitrogen). After three days, the medium was supplemented with 10% FBS, penicillin and streptomycin, and collected as a baculovirus stock. After multiple rounds of amplification of the stock, the resulting high titer baculovirus was used to infect fresh Sf9 cells to induce expression of PARP protein for two days. The PARP-expressing Sf9 cells were then collected by centrifugation, flash frozen in liquid N₂, and stored at -80°C.

PARP-1-, PARP-2- or PARP-3-containing Sf9 cell pellets were thawed on wet ice. The cells were resuspended in FLAG Lysis Buffer [20 mM HEPES, pH 7.9, 0.5 M NaCl, 4 mM MgCl₂, 0.4 mM EDTA, 20% glycerol, 250 mM nicotinamide, 2 mM β-mercaptoethanol, 2x protease inhibitor cocktail (Roche)] and lysed by Dounce homogenization (Wheaton). The lysate was clarified by centrifugation, mixed with an equal volume of FLAG Dilution Buffer (20 mM HEPES, pH 7.9, 10% glycerol, 0.02% NP-40), sonicated, and then clarified by centrifugation again. The clarified lysate was mixed with anti-FLAG M2 agarose resin (Sigma), washed twice with FLAG Wash Buffer #1 (20 mM HEPES, pH 7.9, 150 mM NaCl, 2 mM MgCl₂, 0.2 mM EDTA, 15 % glycerol, 0.01% NP-40, 100 mM nicotinamide, 0.2 mM β-mercaptoethanol, 1 mM PMSF, 1 μM aprotinin, 100 μM leupeptin), twice with FLAG Wash Buffer #2 (20 mM HEPES, pH 7.9, 1 M NaCl, 2 mM MgCl₂, 0.2 mM EDTA, 15% glycerol, 0.01% NP-40, 100 mM nicotinamide, 0.2 mM β-mercaptoethanol, 1 mM PMSF, 1 μM aprotinin, 100 μM leupeptin), and twice with FLAG Wash Buffer #3 (20 mM HEPES, pH 7.9, 150 mM NaCl, 2 mM MgCl₂, 0.2 mM EDTA, 15% glycerol, 0.01% NP-40, 0.2 mM β-mercaptoethanol, 1 mM PMSF). The FLAG-tagged PARP proteins were eluted from the anti-FLAG M2 agarose resin with FLAG Wash Buffer #3 containing with 0.2 mg/mL FLAG peptide (Sigma). The eluted proteins (~0.5 mg/mL) were distributed into single use aliquots, flash frozen in liquid N₂, and stored at -80°C until use.

Purification of NELF-E Expressed in E. coli. BL21(DE3) Rosetta2 pLysS *E. coli*, transformed with a pGEX2T plasmid encoding GST-tagged wild-type or ADP-ribosylation site point mutant NELF-E, were grown in LB with ampicillin and chloramphenicol at 37°C to a density (OD_{600 nm}) of 0.4. The growth temperature was lowered to 18°C and the expression of

the recombinant protein was induced at a density ($OD_{600\text{ nm}}$) of 0.7 by the addition of IPTG to 0.5 mM for 18 hours at 18°C. The cells were collected by centrifugation, and the bacterial cell pellets were flash frozen in liquid N₂ and stored at -80°C.

The cell pellets were thawed on wet ice and lysed by sonication in GST Purification Buffer (50 mM Tris•HCl, pH 7.5, 0.15 M NaCl, 1 mM EDTA, 0.1% NP-40, 1 mM PMSF, 5 mM β-mercaptoethanol). The lysates were clarified by centrifugation. Recombinant GST-NELF-E was purified from the clarified lysate using glutathione-agarose (Sigma) affinity chromatography, with washing of the resin using GST Purification Buffer and elution of the purified proteins in GST Purification Buffer supplemented with 15 mg/mL reduced glutathione. The eluates were dialyzed in GST Purification Buffer and flash frozen in liquid N₂ in single-use aliquots and stored at -80°C.

Purification of NELF Proteins Expressed in Mammalian Cells. pCDNA3-based plasmid encoding a bicistronic expression cassette for N-terminal HA epitope-tagged NELF-A and N-terminal FLAG epitope-tagged NELF-E were transfected into 293T cells using GeneJuice transfection reagent (Novagen). Seventy-two hours post-transfection and two hours after vehicle (DMSO), 300 nM Flavopiridol (Sigma), or 20 μM Olaparib (ApexBio) treatment, the cells were collected in ice cold PBS and pelleted by centrifugation. The cells were then resuspended in ice cold Nuclei Isolation Buffer (10 mM HEPES, pH 8.0, 2 mM MgCl₂, 3 mM CaCl₂, 300 mM sucrose) with freshly added protease inhibitors and 250 nM APD-HPD (Millipore) and allowed to swell for 10 minutes on ice. The nuclei were released from the cells by the addition of NP-40 to 0.65% v/v with vortexing (medium speed) and immediately collected by centrifugation. The supernatant was removed and proteins were extracted from the nuclei by resuspension on ice for

30 minutes in Immunoprecipitation Buffer (25 mM Tris•HCl, pH 7.5, 150 mM NaCl, 1% NP-40, 1 mM EDTA) with freshly added protease inhibitors and 250 nM APD-HPD. The resulting nuclear extract was clarified by two rounds of centrifugation and incubated with equilibrated anti-FLAG M2 agarose resin for 2.5 hours at 4°C with gentle mixing. The agarose beads were washed thoroughly with IP Wash Buffer (25 mM Tris•HCl, pH 7.5, 250 mM NaCl, 1% NP-40, 1 mM EDTA) with freshly added protease inhibitors and 250 nM APD-HPD and then equilibrated in IP Extraction Buffer (25 mM Tris•HCl, pH 7.5, 25 mM NaCl, 0.1% NP-40, 1 mM EDTA) with freshly added protease inhibitors and 250 nM APD-HPD. The immunoprecipitated NELF complex was eluted from the agarose resin by addition of Immunoprecipitation Buffer containing 0.2 mg/mL FLAG peptide. The eluted proteins were distributed in single use aliquots, flash frozen in liquid N₂, and stored at -80°C until use.

Cell Culture and Generation of Knockdown and Overexpression Cell Lines. MCF-7, HeLa S3, 3T3-L1, and 293T cell lines were obtained from the ATCC and used for extract preparation, cDNA library generation, and the variety of assays described herein. Wild-type and *Parp1*^{-/-} mouse embryonic fibroblasts (MEF) cells were a gift from Zhao-Qi Wang, Leibniz Institute for Age Research. MCF-7 cells were maintained in Minimum Essential Medium Eagle supplemented with 5% calf serum. Prior to all experiments, MCF-7 cells were grown for 3 days in phenol red-free MEM Eagle medium supplemented with 5% charcoal-dextran-treated calf serum. MEF, HeLa S3, and 293T cells were cultured in DMEM supplemented with 10% fetal bovine serum, penicillin, and streptomycin.

shRNA-Mediated Knockdown in MCF-7 Cells. Retroviruses were generated by transfection of pSUPER.retro vectors, each expressing a different shRNA sequence directed against the cognate target, with an expression vector for the VSV-G envelope protein into Phoenix Ampho cells using the GeneJuice transfection reagent (Novagen) according to the manufacturer's protocol. The resulting viruses were collected, filtered, and used to infect the parental MCF-7 cell line. Stably transduced cells were isolated under appropriate drug selection with 0.5 µg/mL puromycin or 800 µg/mL G418, expanded, and frozen in aliquots for future use.

Ectopic Protein Expression in MEFs, MCF-7 Cells, and 293T Cells. Lentiviral particles were generated essentially as previously described (161).

MEFs: *Parp1*^{-/-} MEFs were infected with lentivirus, with subsequent selection using 1 µg/mL puromycin to enrich for cells expressing GFP, wild-type PARP-1, or PARP-1(L877A). Ectopic expression of the cognate proteins was confirmed by Western blotting.

MCF-7 cells: MCF-7 cells were infected with lentivirus, with subsequent selection using 1 mg/mL Geneticin (Life Technologies) to enrich for cells with doxycycline-inducible expression of GFP, NELF-E, or ADP-ribosylation site point mutant NELF-E.

293T cells: pCDNA3-based plasmids encoding GFP with a c-terminal 3xFLAG epitope tag or a bicistronic expression cassette for N-terminal HA epitope-tagged NELF-A and N-terminal FLAG epitope-tagged NELF-E were transfected into 293T cells using GeneJuice transfection reagent (Novagen).

PARP-1 Enzyme Kinetics Assays. The enzyme kinetics of wild-type and analog-sensitive mutants of human PARP-1 were determined essentially as previously described (146). Purified

PARP-1 proteins were incubated for 5 minutes in PARP Enzyme Kinetics Buffer (50 mM Tris•HCl, pH 7.9, 4 mM MgCl₂) containing 0.5 mg/mL sonicated salmon sperm DNA (Stratagene) at 25°C with 10, 20, 30, 50, 100, 250, and 500 μ M NAD⁺ or 8-Bu(3-yne)T-NAD⁺. The concentrations of the purified PARP-1 proteins were (1) 25 nM for PARP-1 with NAD⁺, (2) 50 nM for PARP-1(L877A) with NAD⁺, and (3) 250 nM for PARP-1 and PARP-1(L877A) with 8-bu(3-yne)T-NAD⁺. The reactions were brought to 450 mM perchloric acid (HClO₄) to stop them, incubated for 5 minutes on ice, and then brought to 26.5 mM potassium carbonate (K₂CO₃) to quench them. The quenched reactions were centrifuged at maximum speed in a microcentrifuge at 4°C, and the supernatant containing small molecules from the automodification reaction was collected.

The supernatants from the extracted ADP-ribosylation reactions were analyzed using a Shimadzu LCMS-QP8000 α with a Phenomenex Kinetex C18 column (150 x 4.6 mm, 5 μ m XB-C18 100A) monitored at 260 nm. The solvents were 50 mM ammonium acetate pH 5.4 (solvent A) and 50% methanol in water (solvent B). The compounds were eluted at a flow rate of 0.3 mL per minute with 0% solvent B for 1 minute, a linear gradient of 0% to 1% solvent B over 5 minutes, a linear gradient of 1% to 50% solvent B for 5 minutes, an equilibration in 50% solvent B for 1 minute, and finishing with an equilibration back to 0% solvent B for 2 minutes. The retention times for ADP-ribose, NAD⁺, nicotinamide, 8-Bu(3-yne)T-ADP-ribose, and 8-Bu(3-yne)T-NAD⁺ were 12.7, 16.6, 18.9, 20.6, and 22.0 minutes, respectively. Quantitative values for the small molecules and metabolites were obtained by comparison to a standard curve. The k_{cat} and K_m values were obtained by curve-fitting the $V/[E] \sim [S]$ plot using KaleidaGraph. ADP-

ribosylation activity was derived using the formation of nicotinamide after subtraction of the signals for ADP-ribose or 8-Bu(3-yne)T-ADP-ribose.

In Vitro PARP Automodification Reactions

Automodification Reactions. Two hundred nanograms of purified recombinant PARP protein (PARP-1, 2, or 3) were incubated in Automodification Buffer [30 mM HEPES, pH 8.0, 5 mM MgCl₂, 5 mM CaCl₂, 0.01% NP-40, 1 mM DTT, 100 ng/μL sonicated salmon sperm DNA (Stratagene), 100 ng/μL BSA (Sigma)] with 250 μM NAD⁺ or NAD⁺ analog at 25°C for 5 min. for PARP-1 and PARP-2 purified from Sf9 cells, 30 min. for PARP-3 purified from Sf9 cells, or 45 min. for PARP-1 purified from *E. coli*.

Detection of Automodification by Western Blotting. The automodification reactions were stopped by the addition of one third of a reaction volume of 4x SDS-PAGE Loading Buffer (200 mM Tris•HCl, pH 6.8, 8% SDS, 40% glycerol, 4% β-mercaptoethanol, 50 mM EDTA, 0.08% bromophenol blue) followed by heating to 100°C for 5 min. The reaction products were then resolved on a 10% PAGE-SDS gel, transferred to a nitrocellulose membrane, and blotted with a polyclonal antibody against PARP-1 or an ADP-ribose detection reagent (MABE1016, EMD Millipore). Relative quantification of PARP-1 automodification was performed using densitometry and was expressed as the ratio of the PARylation-induced decrease in PARP-1 mobility relative to non-shifted PARP-1.

Detection of Automodification by In-gel Fluorescence. The automodification reactions were stopped by methanol:chloroform precipitation (162), with subsequent collection of the precipitates by centrifugation. The protein pellets were redissolved clicked to azido-rhodamine

(Click Chemistry Tools) in Denaturing CC Buffer [100 mM HEPES, pH 8.0, 4 M urea, 0.5 M NaCl, 2% CHAPS, 100 μ M azido-rhodamine, 5 mM THPTA (Click Chemistry Tools), 1 mM CuSO₄, 5 mM sodium ascorbate] following the step-wise addition of azido-rhodamine, THPTA:CuSO₄ complex, and sodium ascorbate. After a 2 hour reaction in the dark at room temperature, the clicked proteins were collected by a methanol:chloroform precipitation with centrifugation, redissolved in 1x SDS Loading Buffer (50 mM Tris•HCl, pH 6.8, 2% SDS, 10% glycerol, 1% β -mercaptoethanol, 12.5 mM EDTA, 0.02% bromophenol blue), and heated at 100°C for 5 min. in the dark. The clicked proteins were resolved on a 10% PAGE-SDS gel. The gel was then washed quickly with MilliQ H₂O, washed twice with 10% methanol, washed twice with water, and visualized on a Bio-Rad Pharos FX Plus Molecular Imager (excitation: 532 nm, emission: 605 nm).

In Vitro NELF-E ADP-ribosylation Reactions. One hundred nM of recombinant PARP-1 protein purified from Sf9 insect cells was (1) incubated in Automodification Buffer, followed by (2) the addition of 1 μ M GST, GST-NELF-E, or GST-NELF-E mutant, and (3) the addition of 25 μ M NAD⁺ at 25°C for 20 min. Detection of ADP-ribosylated GST-tagged NELF-E was performed as described above for the automodification of PARP-1 using ADP-ribose detection reagent (MABE1016; EMD Millipore).

NELF-E/TAR RNA Binding Assays

TAR RNA folding and end-labeling. HIV TAR RNA was ordered from Integrated DNA Technologies (Coralville, Iowa) and resuspended in DEPC-treated 1x TE to a concentration of

100 μ M. TAR RNA was folded at a concentration of 10 μ M in RNA Folding Buffer (10 mM Tris•HCl, pH 7.5, 100 mM KCl, 10 mM MgCl₂) by incubation for 2 minutes at 90°C, followed by rapid cooling on wet ice for 2 minutes, and incubation at room temperature for 30 minutes to equilibrate. The RNA was end-labeled at a concentration of 1 μ M in T4 PNK Buffer (70 mM Tris•HCl, pH 7.5, 10 mM MgCl₂, 5 mM DTT) using T4 PNK (NEB) and 835 nM γ -³²P-ATP (3000 Ci/mmol; Perkin-Elmer) for 1 hour at room temperature. TAR RNA was then desalted into 10 mM Tris•HCl, pH 7.5, using an RNase free Micro Bio-Spin P-30 Gel Column (Bio-Rad) according to the manufacturer's instructions prior to use in electrophoretic mobility shift assays.

NELF-E ADP-ribosylation and Electrophoretic Mobility Shift Assay. One μ M of GST or GST-NELF-E was incubated with or without 100 nM PARP-1 in Gel Shift Buffer (20 mM Tris•HCl, pH 7.5, 37.5 mM NaCl, 0.025% NP-40, 5 mM MgCl₂, 5 mM CaCl₂, 0.1 mg/mL BSA, 0.1 mg/ml sonicated salmon sperm DNA, 5 % Glycerol, 20 nM end-labeled TAR RNA). Then 25 or 100 μ M of NAD⁺ was added to stimulate NELF-E ADP-ribosylation. The reaction was stopped after incubation for 20 minutes at room temperature by the addition of 100 μ M Olaparib (Apex Bio). Half of each reaction was then diluted 10-fold in Gel Shift Buffer, followed by a 1 hour incubation at room temperature to equilibrate. The samples were run on 5% polyacrylamide gels in 0.5x TBE for 2 hours at 4°C. The gels were dried and exposed to autoradiographic film overnight.

Preparation and LC-MS/MS Analysis of 8-Bu(3-yne)T-ADP-ribosylated HeLa Cell Nuclear Extract Proteins

The following protocols were used for small-scale and large-scale labeling, with subsequent in-gel fluorescence or mass spectrometry analyses, respectively.

Small-scale Modification of HeLa Cell Nuclear Extract Proteins and Detection by In-gel Fluorescence. HeLa cell nuclear extract was prepared as previously described (163). One microgram of PARP protein (PARP-1, 2, or 3; wild-type or analog sensitive) was incubated in separate reactions as follows: (1) 5 min. in Automodification Buffer, (2) 5 min. with 50 μ g of HeLa cell nuclear extract, followed by 15 min. with 250 μ M NAD⁺, or (3) 5 min. with 50 μ g of HeLa cell nuclear extract, followed by 15 min. with 250 μ M 8-Bu(3-yne)T-NAD⁺. The reactions were stopped by methanol:chloroform precipitation and the precipitates were collected by centrifugation. The protein pellets were clicked to azido-rhodamine (Click Chemistry Tools) in Denaturing CC Buffer, run on an SDS-PAGE gel, and visualized using a Bio-Rad Pharos FX Plus Molecular Imager, as described above.

Large-scale 8-Bu(3-yne)T-ADP-ribosylation of HeLa Cell Nuclear Extract Proteins and LC-MS/MS. Twenty μ g of asPARP protein were incubated sequentially as follows: (1) 5 minutes in Automodification Buffer, (2) 5 minutes upon addition of 1000 μ g of HeLa nuclear extract, and then (3) 15 minutes with 250 μ M 8-Bu(3-yne)T-NAD⁺. The reactions were stopped by methanol:chloroform precipitation. The protein pellets were resuspended in 1 mL of Urea Solubilization Buffer (200 mM HEPES, pH 8.0, 8 M urea, 1 M NaCl, 4% CHAPS) and the insoluble proteins were pelleted by centrifugation for 1 minute at maximum speed in a microcentrifuge. Soluble proteins in Urea Solubilization Buffer were combined sequentially in a 2 mL tube in the following order with mixing: 100 μ L azido-agarose Beads (Click Chemistry Tools), 820 μ L water, 40 μ L of a 50:250 mM CuSO₄:THPTA pre-formed catalytic complex, 20

μL of 500 mM aminoguanidine hydrochloride, and 20 μL of 500 mM sodium ascorbate. After 18 hours of reaction time in the dark with slow mixing in a rotating mixer, the beads were collected by centrifugation at room temperature for 1 minute at 1000 RCF in a microcentrifuge and the reaction supernatant was aspirated. The beads were resuspended in 1.8 mL MilliQ H_2O and were collected by centrifugation at room temperature for 1 minute at 1000 RCF. The beads were then resuspended in 1 mL of SDS Wash Buffer (100 mM Tris•HCl, pH 8.0, 1% SDS, 250 mM NaCl, 5 mM EDTA) supplemented with freshly made 1 mM DTT, heated to 70°C for 15 minutes, and then allowed to cool to room temperature. The resin was collected by centrifugation at room temperature for 5 minutes at 1000 RCF in a microcentrifuge and the supernatant was aspirated. The resin was then resuspended in 1 mL of SDS Wash Buffer containing 40 mM iodoacetamide and incubated at room temperature for 30 minutes in the dark to alkylate the cysteine residues. The resin was then transferred to a 2 mL single use column (Bio-Rad) and washed as follows: 10 washes of 2 mL each with SDS Wash Buffer, 10 washes of 2 mL each with Urea Wash Buffer (100 mM Tris•HCl, pH 8.0, 8M urea), and 10 washes of 2 mL each with 20% acetonitrile. Following these extensive washes, the resin was resuspended in 500 μL of Trypsin Digestion Buffer (100 mM Tris•HCl, pH 8.0, 2 mM CaCl_2 , 10% acetonitrile).

Trypsin digestion of bead-bound 8-Bu(3-yne)T-ADP-ribosylated HeLa cell nuclear extract proteins was performed by adding of 1 μg of trypsin (Promega) to the Trypsin Digestion Buffer, with incubation at room temperature overnight with slow mixing on a rotating mixer. The peptides from the tryptic digest were prepared for LC-MS/MS by desalting on a C18 stage tip (Thermo) according to the manufacturer's protocol and lyophilized for storage at -20°C prior to the LC-MS/MS runs for peptide ID. Post-tryptic digest resin, containing peptides covalently

linked to the agarose resin through the azide-clicked 8-Bu(3-yne)T-ADP-ribosylation site, were transferred to a fresh 2 mL single use column (Bio-Rad) and washed as follows: 10 washes of 2 mL each with SDS Wash Buffer, 10 washes of 2 mL each with Urea Wash Buffer, 10 washes of 2 mL each with 20% acetonitrile, and 5 washes of 2 mL each with Peptide Elution Buffer (100 mM HEPES, pH 8.5). The resin was transferred to a microcentrifuge tube and hydroxylamine (Sigma) was added to 0.5 M to elute the glutamate- and aspartate-modified 8-Bu(3-yne)T-ADP-ribosylated peptides from the resin, using a modification of an approach previously described (142). The eluted peptides were prepared for LC-MS/MS by desalting on a C18 stage tip (Thermo) according to the manufacturer's protocol and then lyophilized for storage at -20°C prior to LC-MS/MS analysis. Samples were prepared for LC-MS/MS exactly as described previously (142).

Analysis of LC-MS/MS Data

LC-MS/MS Peptide and Site Identification. The sites of ADP-ribosylation were obtained from LC-MS/MS analysis as described previously (142). All ADP-ribosylation sites identified from both replicates were used in the data analysis. The software, scripts, and other information about the analyses can be obtained by contacting the corresponding author (W.L.K.).

Gene Ontology Analyses. Gene ontology analyses were performed using the DAVID (Database for Annotation, Visualization, and Integrated Discovery) tool (164, 165). The input was the PARP target proteins found to be ADP-ribosylated in the LC-MS/MS analysis.

Location of Sites of ADP-ribosylation Relative to Other Post-translational Modifications. ADP-ribosylation sites were mapped from IPI accession numbers to Uniprot IDs

giving preference to polypeptides matching gene names to IPI database nomenclature, followed by polypeptides with the longest amino acid sequence length. A knowledge base of known post-translational modifications (PTMs) for comparison versus the ADP-ribosylation sites identified in this study was obtained from the PhosphoSite Plus database (166). Additional content for sumoylation (167) and ADP-ribosylation (142), which was downloaded and processed from the publically available data, was added. The sites of ADP-ribosylation sites identified in this study, as well as an aspartate/glutamate ratio-normalized random control, were analyzed for amino acid-to-amino acid relationships to other PTMs within the assembled PTM knowledge base. On any given polypeptide, the PTM with the closest spatial relationship to an ADP-ribosylation site, or within the random aspartate or glutamate control, were retained for analysis and visualization.

Determination of ADP-ribosylation Site Motifs. Sequences ± 8 amino acids from all unique ADP-ribosylation sites for PARP-1, PARP-2, and PARP-3 were analyzed for statistically significant enrichment of amino acid sequences using the Motif-X server (168, 169), with a significance threshold of 0.005.

Determination Enriched 7-mer Amino Acid Sequences Proximal to ADP-ribosylation Sites. Amino acid sequences ± 100 residues from all unique ADP-ribosylation sites for PARP-1, PARP-2, and PARP-3 were analyzed for the relative frequency of all overlapping 7-mer sequences using a custom script.

8-Bu(3-yne)T-ADP-ribosylation Reactions in Intact Nuclei

MEFs were harvested in cold PBS and collected by centrifugation. The cells were swollen in Nuclei Isolation Buffer (10 mM HEPES, pH 8.0, 2 mM MgCl_2 , 3 mM CaCl_2 , 300

mM sucrose, with freshly added 1 mM DTT, protease inhibitors, and phosphatase inhibitors) and the nuclei were released by the addition of 0.65% NP-40 with moderate vortexing. Following collection by centrifugation, the nuclei were resuspended in PARP Reaction Buffer (30 mM Tris•HCl, pH 7.5, 10 mM KCl, 5 mM MgCl₂, 5 mM CaCl₂, 0.01% NP-40, 0.05 mM EDTA, 20% glycerol, with freshly added 1 mM DTT, protease inhibitors, and phosphatase inhibitors) containing 250 μM 8-Bu(3-yne)T-NAD⁺ for 30 minutes at 25°C with occasional gentle mixing to allow ADP-ribosylation to occur in the isolated nuclei.

Confocal Microscopy. Following 8-Bu(3-yne)T-ADP-ribosylation in intact nuclei from MEFs as described above, the nuclei were washed once with Nuclei Click Reaction Buffer (10 mM HEPES, pH 8.0, 10 mM NaCl, 10 mM spermidine, 0.1% NP-40, with freshly added protease inhibitors) and then clicked twice to Alexa Fluor 488 by the sequential addition of 100 μM azido-Alexa Fluor 488 (Life Technologies), a pre-formed complex of CuSO₄:THPTA (1.0 mM:5.0 mM), and 5 mM sodium ascorbate. The nuclei were then washed with Nuclei Clicking Buffer to remove unclicked azido-Alexa Fluor 488. The nuclear DNA was stained by incubating the nuclei with 1 μM To-Pro DNA Stain (Life Technologies) in Nuclei Click Reaction Buffer for 2 minutes, with subsequent washing using Nuclei Click Reaction Buffer to remove unincorporated DNA stain. The nuclei were pipetted onto a glass cover slip, wicked dry with Whatman 3MM filter paper, covered with VectaShield (Vector Laboratories), and adhered to a glass slide. Fluorescence was visualized via confocal microscopy using a Leica TCS SP2 Confocal Microscope.

Click-ChIP-seq. I developed a methods based on the analog-sensitive PARP technology for localizing ADP-ribosylation events mediated by a specific nuclear PARP (e.g., PARP-1) to specific sites in the genome. I call this click chemistry-based technology “Click-ChIP-seq,” for clicked chromatin interaction and precipitation coupled with deep sequencing, which is similar in many respects to standard ChIP-seq protocols.

Formaldehyde Crosslinking in Intact Nuclei. Following 8-Bu(3-yne)T-ADP-ribosylation in intact nuclei from MEFs as described above, the nuclei were collected by centrifugation and resuspended to homogeneity in Nuclei Crosslinking Buffer (10 mM HEPES, pH 8.0, 10 mM NaCl, 10 mM MgCl₂, 0.1% NP-40, with freshly added protease inhibitors) and crosslinked by addition of 0.5% methanol-free formaldehyde (Pierce) for 10 minutes while mixing slowly on a rotary mixer. Crosslinking was quenched by the addition of glycine to a final concentration of 275 mM with incubation on ice for 10 minutes. The nuclei were centrifuged at 500 RCF for 10 minutes before the 8-Bu(3-yne)T-ADP-ribosylated proteins were clicked to biotin for subsequent analysis.

Enrichment of 8-Bu(3-yne)T-ADP-ribose-associated Regions of the Genome. Following crosslinking as described above, the nuclei were washed thoroughly with Nuclei Click Reaction Buffer. ADP-ribosylated proteins in the nuclei were then clicked twice to biotin by resuspension of the nuclei in Nuclei Click Reaction Buffer with sequential addition of 100 μ M azido-biotin (Click Chemistry Tools), a pre-formed complex of CuSO₄:THPTA (1.0 mM:5.0 mM), and 5 mM sodium ascorbate. The nuclei were then washed thoroughly in Nuclei Click Reaction Buffer to remove unclicked azido-biotin and resuspended in ChIP Lysis Buffer [50 mM Tris•HCl, pH 7.9, 1% SDS, 10 mM EDTA, 1x protease inhibitor cocktail (Roche)]. The

resuspended nuclei were sonicated in a water bath sonicator (Diagenode) to a DNA fragment size distribution of 200 to 500 bp. The biotinylated and sonicated chromatin was diluted ten-fold in ChIP Dilution Buffer (20 mM Tris•HCl, pH 7.9, 2 mM EDTA, 150 mM NaCl, 0.5% Triton X-100, 1 mM DTT, 1x protease inhibitor cocktail, 1% BSA) and bound to MyOne T1 streptavidin-conjugated magnetic beads (Invitrogen) by incubation overnight at 4°C with gentle mixing.

After the incubation, the beads were collected in a magnetic field and washed by resuspension as specified in the following buffers: (1) twice in with 2% SDS, (2) once with Low Salt Precipitated Complex Wash Buffer (20 mM Tris•HCl, pH 8.1, 2 mM EDTA, 125 mM NaCl, 0.05% SDS, 1% Triton X-100, 1 μ M aprotinin, and 100 μ M leupeptin), (3) three times with High Salt Immune Complex Wash Buffer (20 mM Tris•HCl, pH 8.1, 2 mM EDTA, 500 mM NaCl, 0.05% SDS, 1% Triton X-100, 1 μ M aprotinin, and 100 μ M leupeptin), (4) once with LiCl Wash Buffer (10 mM Tris•HCl, pH 8.1, 1 mM EDTA, 250 mM LiCl, 1% NP-40, 1% sodium deoxycholate, 1 μ M aprotinin, and 100 μ M leupeptin), and (5) twice with 1x TE (10 mM Tris•HCl, 1 mM EDTA). The beads were collected in a magnetic field after the final wash and resuspended in 200 μ L of De-crosslinking Buffer (100 mM sodium bicarbonate, 1% SDS, 100 mM NaCl). The 8-Bu(3-yne)T-ADP-ribose-associated DNA was released from beads by incubation in the De-crosslinking Buffer overnight at 65°C and isolated away from the beads by magnetic separation. The beads were washed with once with 1x TE, which was combined with the initial eluate. The eluted DNA was treated with RNase and then protease, and purified by phenol:chloroform:isoamyl alcohol extraction (25:24:1). The resulting DNA was ethanol precipitated, washed with 70% ethanol, dried, and resuspended in 10 mM Tris•HCl in a volume determined by the subsequent assays. For quantitative PCR, the 8-Bu(3-yne)T-ADP-ribose-

precipitated DNA was analyzed using a Roche LightCycler 480 (170) and compared to a retained input fraction (de-crosslinked and purified as above).

Preparation of Click-ChIP-Seq and ChIP-seq Libraries for Deep Sequencing. Click-ChIP-seq libraries for deep sequencing were prepared from the 8-Bu(3-yne)T-ADP-ribose-precipitated DNA essentially as previously described (171). For ChIP-seq libraries prepared from intact nuclei, the isolated nuclei were subjected to mock ADP-ribosylation and azido-biotin click chemistry reactions, sonicated, and diluted in ChIP Dilution Buffer as described above. PARP-1 bound to chromatin was immunoprecipitated from the sonicated chromatin essentially as described previously (171) using anti-FLAG M2 agarose (Sigma). The ChIP'ed DNA was prepared for high throughput sequencing as previously described (171). The Click-ChIP-seq and ChIP-seq libraries were sequenced using an Illumina HiSeq instrument at the McDermott Center Sequencing Core Facility at UT Southwestern Medical Center.

Analysis of Click-ChIP-seq and ChIP-seq Data. Software, scripts, and other information about the analyses can be obtained by contacting the corresponding author (W.L.K.).

Quality Control. Quality control for the Click-ChIP-seq and ChIP-seq data was performed using the FastQC tool (<http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>).

Read Alignment and Visualization. The deep sequencing reads were aligned to the human genome (hg19) using default parameters in bowtie (172). The Click-ChIP-seq data were prepared for visualization by calculating the odds ratio of Click-ChIP-seq relative to input across the genome in 2500 bp windows with 250 bp steps. The Click-ChIP-seq odds ratios were then converted into bigwig files using BEDTools (173) and visualized using the IGV genome browser

(174, 175) with a 2-fold cutoff for color intensity. The H3K4me3 and NELF ChIP-seq data were from published sources (176, 177). Uniquely mappable ChIP-seq read densities were converted into bigWig files using BEDTools and visualized using the IGV genome browser.

Peak Calling and Genome-wide Dataset Correlations. Genomic sites of enrichment for transcription or chromatin proteins, or histone modifications were calculated using SICER (178) with a false discovery rate of 1×10^{-2} for all data sets evaluated. Enrichment of NELF-E, SA1, SA2, SMC1, SMC3, CTCF, CDK9, and H3K4me3 was determined using a window and gap size of 200 bp. Enrichment of PARP-1, H3K36me3, H3K27me3, and H3K9me3 was determined using a window size of 200 bp and a gap size of 600 bp. Enrichment of lamins was determined using a window size of 1,000 bp and a gap size of 3000 bp. To determine the correlation of PARP-1-dependent ADP-ribosylation with a transcription or chromatin protein, or a histone modification, a Pearson's correlation coefficient was calculated between the normalized read depth for each factor underneath their requisite peaks and the input-normalized Click-ChIP-seq signals.

Heatmaps. Read densities in a 10 kb window surrounding the RefSeq TSS (± 5 kb) were determined using a custom script in the programming language R and visualized as heatmaps using Java TreeView (179).

Analysis of CDK9 Enrichment, PARP-1-Mediated ADP-ribosylation, and RNA Polymerase Pausing. For every actively transcribed RefSeq mouse promoter, CDK9 occupancy and ADP-ribosylation enrichment were calculated in a 1 kb window surrounding the TSS (± 500 bp). RNA polymerase II pausing indices were then calculated for these promoters as described below. Promoters were binned according to increasing occupancy of CDK9 or enrichment of

PARP-1-mediated ADP-ribosylation, and an increase or decrease in RNA polymerase II pausing was calculated for the binned promoters versus all RefSeq promoters. The results of these analyses were plotted according to the absolute value of the increase or decrease in RNA polymerase II pausing and the associated t-test-calculated p-value for this determination.

Analysis of ChIP-chip Data. NELF-E ChIP-chip data from MCF-7 cells (158) were downloaded from the GEO database. Probe sequences for the custom chip array, each associated with a log2 odds ratio value from the NELF-E ChIP-chip experiment, were aligned to the human reference genome (hg19) using the bowtie aligner(172). Statistically significant NELF-E peaks were called as described previously (180) with a 2-fold odds ratio cut-off. NELF-E binding sites were calculated from the called peaks and assigned to a promoter when occurring within the first 500 bp of a RefSeq promoter. Software, scripts, and other information about the analyses can be obtained by contacting the corresponding author (W.L.K.).

Preparation of GRO-Seq Libraries

Preparation of Nuclei. MCF-7 cells with shRNA-mediated knockdown of luciferase (Luc; as a control) or PARP-1 were described previously (106). Nuclei were isolated from the Luc and PARP-1 knockdown cell lines were subjected to GRO-seq as described previously (170). Briefly, MCF-7 cells were washed three times with ice-cold PBS and then resuspended for swelling in ice-cold Hypotonic Lysis Buffer [10 mM Tris•HCl, pH 7.4, 0.5% NP-40, 10% Glycerol, 3 mM CaCl₂, 2 mM MgCl₂, and 1 mM DTT containing 1x protease inhibitor cocktail (Sigma-Aldrich) and 4 units/mL SUPERase-In (Ambion)]. The swollen cells were collected by

centrifugation at 1000 RCF for 10 min at 4°C and then resuspended in 1.5 ml of lysis buffer and pipetted up and down through a narrow opening tip 30 to 50 times to lyse the cells and release the nuclei. The nuclei were collected by centrifugation and washed once with 1 mL of Hypotonic Lysis Buffer. After a final collection by centrifugation, the resulting pellets of nuclei were resuspended in 500 μ L of Freezing Buffer (50 mM Tris•HCl, pH 8.3, 40% glycerol, 5 mM MgCl₂, 0.1 mM EDTA, and 4 units/mL of SUPERase-In per mL), counted, frozen in liquid nitrogen in 100 μ L aliquots containing 5×10^6 nuclei, and stored at -80°C until use.

Nuclear Run-On and Library Preparation. Nuclear run-on and GRO-seq library preparation were performed as previously described (170, 181). Briefly, nuclear run-on reactions were performed for ~100 bases in the presence of sarkosyl (to prevent reengagement of RNA polymerases), rNTPs, $\alpha^{32}\text{P}$ -CTP, and 5-bromo-UTP. The nascent RNAs were isolated, hydrolyzed to ~100 bases, and enriched using α -BrdUTP antibody-conjugated agarose beads (Santa Cruz). The bound RNAs were washed several times and eluted. The 5' RNA cap was removed and the ends were repaired in preparation for adapter ligation. Small RNA adapters were ligated to the 5' end, followed by another bead binding enrichment using α -BrdUTP antibody-conjugated agarose beads. These steps were repeated using a 3' adapter. The resulting RNAs were reverse transcribed, amplified using PCR, and analyzed by high throughput sequencing using an Illumina 1G Genome Analyzer.

Analysis of GRO-seq Data. The GRO-seq data were analyzed using software described previously (170) and the approaches described below. Software, scripts, and other information about the analyses can be obtained by contacting the corresponding author (W.L.K.).

Quality Control. Quality control for the GRO-seq data was performed using the FastQC tool (<http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>). The GRO-seq reads were trimmed to remove adapter contamination using the default parameters of Cutadapt software (182). Reads >32 bp long were retained for alignment.

Read Alignment and Gene Annotations. Trimmed human GRO-seq reads were aligned to the human reference genome (hg19) using the bwa aligner (172) with default settings (uniquely aligned, 2 mismatches allowed, and 19 bp seed sequence). The 5'-most base pair from each read was used in all analyses, with no more than 2 duplicates allowed at any genomic location. Mouse GRO-seq reads were aligned to the mouse reference genome (mm9) using the bowtie aligner (172) with default settings (uniquely aligned, 2 mismatches allowed, and 19 bp seed sequence). As above, the 5'-most base pair from each read was used in all analyses. For genes with multiple TSSs, I used the TSS with the most GRO-seq reads within the first 150 bp in the genic sense direction.

Determination of Transcription Levels. Transcription levels were calculated by counting the total GRO-seq reads across the entire transcript and dividing by the length of the transcript in base pairs. Overlaps and redundancies were removed from the combined gene lists to eliminate the possibility of double counting genes. .

Analysis of Pausing Indices. Pausing indices representing the base pair normalized difference in read depth between the promoter proximal region (1- 250 bp) and the gene body region (1-13 kb) were calculated as described (183). The effect of PARP-1 knockdown on the level of paused RNA polymerase was determined using edgeR (157). Pausing indices was calculated for every RefSeq gene as described above from GRO-seq replicates normalized for

read depth by random read subtraction. EdgeR was run without a library read depth adjustment and with a p-value cutoff of 0.001 to determine significant changes in RNA polymerase pausing.

Heatmaps and Metagenes. The read densities of sense and anti-sense reads were calculated on adjacent lines for a 10 kb window surrounding each RefSeq TSS (± 5 kb) of using a custom script in the R programming language. The data were visualized as heatmaps using Java TreeView (179), with sense and anti-sense reads for each RefSeq promoter and colored red and blue, respectively. Metagenes were generated as previously described (170).

Genomic Datasets. The new genomic data sets generated for these studies are as follows: (1) MCF-7 GRO-seq (Luc knockdown and PARP-1 knockdown), and (2) MEF PARP-1 ChIP- and Click-ChIP-seq in intact nuclei. They are available from the NCBI's GEO database using accession numbers GSE74141 and GSE74142 respectively. The following publically available deep sequencing data sets (including their cognate controls) were downloaded from NCBI's GEO archive using the following accession numbers: H3K36me3 ChIP-seq (GSE12241); Lamin B1-DamID (GSE17051); NELF-B ChIP-seq (GSE24113); SA1, SA2, SMC1, and SMC3 ChIP-seq (GSE32319); CTCF and H3K4me3 ChIP-seq (GSE29218); H3K9me3 and H3K27me3 (GSE22268); CDK9 ChIP-seq (GSE45517).

Oligonucleotide Sequences

• ***shRNAs*** (listed 5' to 3')

Luc #1	GGAAUCCAGUGUGUGAAGA[dT][dT]
Luc #2	GAGAGAAAAAAUCAACAGC[dT][dT]

PARP-1 #1 GUGUAGACAUCCUCCGUAU[dT][dT]

PARP-1 #2 CAUACUCUAUCCGAGUAU[dT][dT]

• ***Primers for Click-ChIP-qPCR (listed 5' to 3')***

FKBP5 Fwd: GTCCAGCCAGACCAAACAGT

FKBP5 Rev: AAGGGACACAGGGTGTGAAG

CEBPA Fwd: CTGGAAGTGGGTGACTTAGAGG

CEBPA Rev: GAGTGGGGAGCATAGTGCTAG

Conclusions

Herein, I have described the development of an NAD^+ analog-sensitive approach for PARP proteins that preserves the natural catalytic activities of both mono- and poly(ADP-ribosyl) transferases, in contrast to a previous approach (145). Our *as*PARP approach uses a single point mutation buried within the PARP active site in concert with a single alkyne-containing R group on NAD^+ to achieve PARP-specific clickable ADP-ribosylation. I have demonstrated the robustness and transferability of this *as*PARP approach by using it to (1) identify the PARP-1, 2, and 3 target proteins, as well as the specific sites of Glu and Asp ADP-ribosylation on those proteins, by using mass spectrometry and (2) identify the PARP-1-mediated sites of ADP-ribosylation across the mammalian genome using deep sequencing. Importantly, this combination of chemical genetics, proteomics, and genomics has allowed us to develop and test new hypotheses about the biology of ADP-ribosylation in gene regulation. This has led to a model, supported by our data, in which PARP-1 (via ADP-ribosylation) and P-TEFb (via

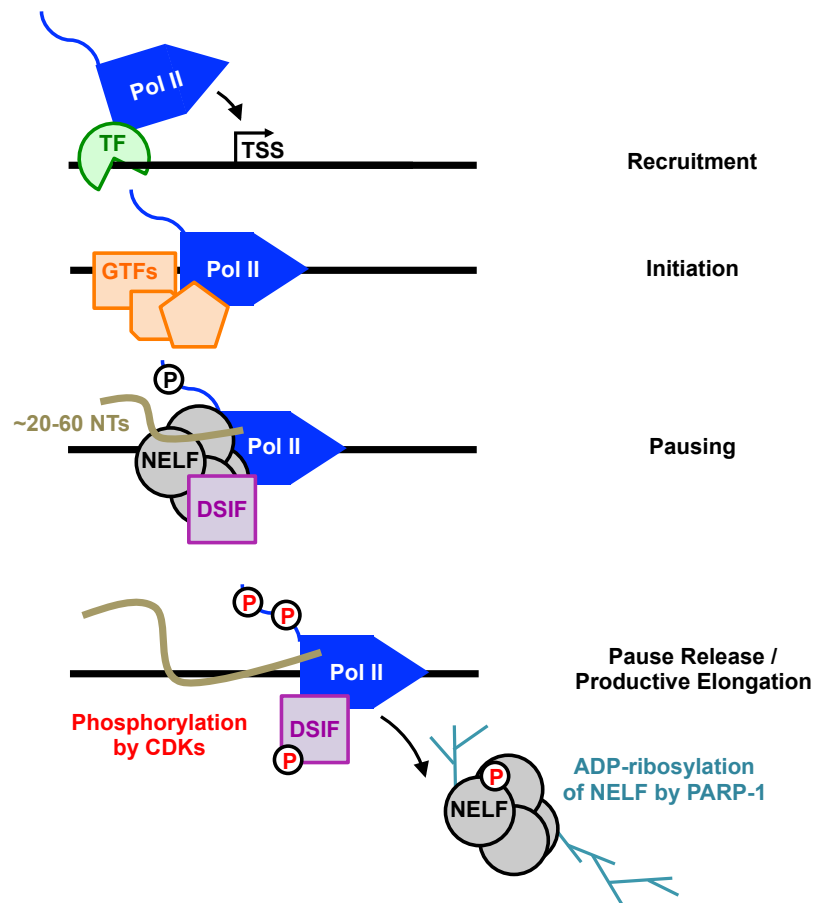


Figure 21. Model for the role of PARP-1 in the regulation of RNA polymerase pausing and release through NELF.

Transcription factors (TF) help to recruit RNA polymerase II (RNA Pol II) to promoters (“Recruitment”), where it engages the promoter, forms a preinitiation complex containing general transcription factors (GTFs), and initiates transcription (“Initiation”). After transcribing ~20 to 60 nucleotides (NTs), RNA Pol II pauses downstream of the transcription start site (TSS) upon the binding of NELF and DSIF (“Pausing”). Phosphorylation of the carboxyl-terminal domain (CTD) of RPB1, the largest subunit of RNA Pol II, DSIF, and NELF by cyclin-dependent kinases (CDKs), such as CDK9, a subunit of the positive transcription elongation factor P-TEFb, as well as ADP-ribosylation of NELF by PARP-1, promote the dissociation or inactivation of NELF, allowing productive elongation by RNA Pol II (“Pause Release / Productive Elongation”). Phosphorylation by cyclin-dependent kinases (CDKs), such as CDK9, may promote subsequent ADP-ribosylation of NELF by PARP-1

phosphorylation) act together to control Pol II pausing and release through the negative elongation factor NELF ([Fig. 21](#)). This integrated approach should have great utility across the family of PARPs, accelerating the discovery of previously unknown biological functions for ADP-ribosylation.

CHAPTER FOUR

PARP-1 BINDS A NUCLEASE-SENSITIVE NUCLEOSOME SPANNING REGULATORY ELEMENTS ACROSS THE GENOME

Summary

PARP-1 is a chromatin-binding protein that plays key roles in transcriptional regulation. To gain insight into the mechanism by which PARP-1 affect transcriptional regulation, I have developed crosslinked MNase ChIP-seq (XL-MNase ChIP-seq) methodology to identify the nucleosomes that PARP-1 binds *in vivo*. I have found that PARP-1 binds a nucleosome spanning the “nucleosome free” region of regulatory elements across the genome. This nucleosome shows poor positioning, low occupancy, and exceptional nuclease sensitivity, indicates that it is structurally altered from a canonical nucleosome. This nuclease-sensitive nucleosome and its binding by PARP-1 may affect access of regulatory elements to transcriptional machinery thereby modulating gene expression and cellular response to stimuli.

Introduction

In eukaryotes the genome is packaged into a nucleoprotein structure known as chromatin, the basic unit of which is the nucleosome. Nucleosomes contain ~147 bp of DNA wrapped about 1.65 times around a protein octamer core, composed of 2 copies each of histones H3, H4, H2A, and H2B (184, 185). These nucleosome units are separated by linker DNA, with each nucleosome spaced on average every 30-60 base pairs from one another in mammalian cells

(186). Nucleosome position and occupancy throughout the genome is a critical point of genome regulation and is determined by the underlying DNA sequence, ATP-dependent nucleosome remodeling enzymes, and transcription factor binding (187-190). Also regulating genome access is the composition of chromatin, which alters through both use of histone variants (i.e. H2A.Z, H3.3) within the nucleosome octamer, as well as histone post-translational modifications or “marks” (i.e. Histone H3K4me3 and H3K27me3) (191, 192). Histone variants, such as Histone H2A.Z, can affect chromatin structure by resulting in a more “open” or “closed” chromatin depending on their incorporation (193). Histone marks most often function through their ability to recruit chromatin-binding proteins known as “readers”, which then facilitate diverse biological changes from transcription initiation, DNA replication, and the DNA damage response (194-196).

Identification of the positions of nucleosomes across the genome has been made possible by the advent of genome-wide approaches (197). In nucleosome-determining experiments Micrococcal Nuclease (MNase), which preferentially digests linker DNA, is used to liberate nucleosomes from chromatin, after which nucleosomal DNA is subjected to massively parallel sequencing to generate a map of nucleosome density across the genome. This technique, known as MNase-seq, has revealed that the promoter regions of transcribed genes as well as transcription factor binding sites often exhibit well-positioned nucleosomes adjacent to transcription start sites (TSS) and transcription factor binding sites (TFBS) as well as a region of nucleosome depletion known as the nucleosome free region (NFR), which spans the TSS and TFBS (186, 197). These nucleosome features are created, in part, by the concerted action of transcription factors and a class of chromatin-binding proteins known as nucleosome remodeling

enzymes (190). Chromatin binding proteins can affect nucleosome architecture at active regions of the genome, affecting transcriptional outcomes (198).

One of these chromatin-binding proteins is Poly(ADP-ribose) Polymerase 1 (PARP-1). PARP-1 is a nuclear protein that plays diverse roles in genome maintenance, transcriptional regulation, and modulation of chromatin structure (137). PARP-1 binds to nucleosomes at the their dyad axis, the point at which DNA begins and ends it's wrapping around the histone octamer, and modulates chromatin structure, access, and transcriptional activation (55, 60, 84, 153). PARP-1 binds DNA at promoters, reciprocally to the linker Histone H1, modulating transcriptional through a mechanism that is not yet fully understood (180). PARP-1 is an integral player in the transcriptional response to inflammatory stimulus, and it's genetic depletion or use of PARP inhibitors results in diminution of the pro-inflammatory response, resulting in cardioprotective effects *in vivo* (199). How PARP-1's binding to nucleosomes relates to fine-scale nucleosome architecture throughout the genome and how that chromatin-binding activity modulates the inflammatory response in cardiomyocytes remains unknown.

In order to understand how and where PARP-1 binds to nucleosomes across the genome I have developed XL-MNase-ChIP-seq, a technique capable of determining nucleosome resolution profiles of chromatin architecture, histone marks, and chromatin-binding proteins across the genome. I have used this technique to develop a deeper understanding of PARP-1 binding in human cardiomyocyte cells, where PARP-1 has an established and important role in transcriptional regulation. I have found that PARP-1 binds to a specific and unique nucleosome at regulatory elements, providing new insights into how PARP-1 might regulate genome access and transcriptional regulation.

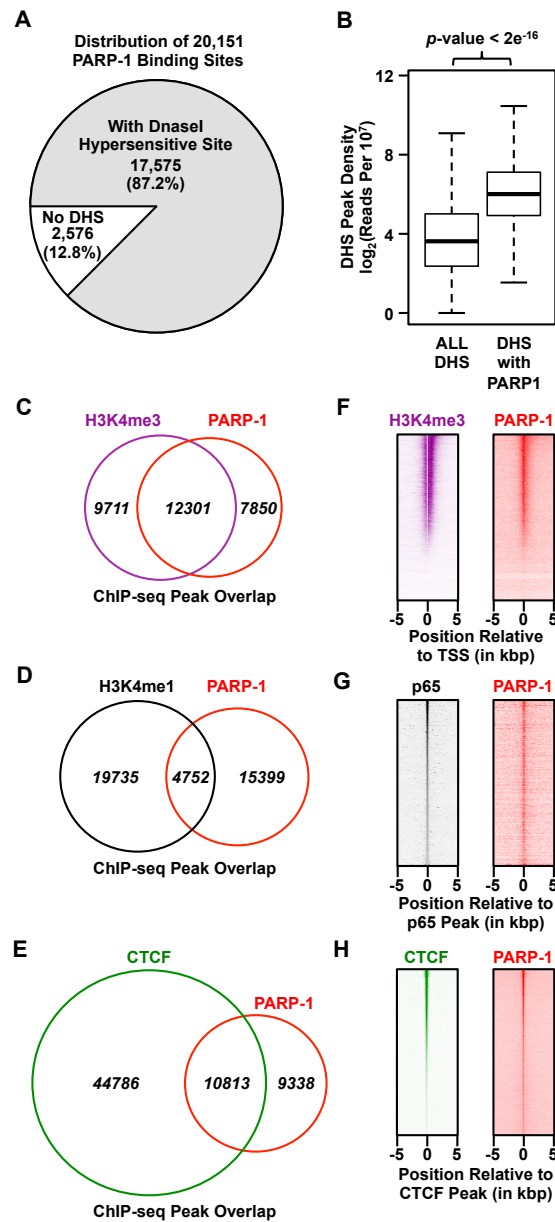


Figure 22. PARP-1 binds open chromatin regions at promoters, enhancers, and insulators.

(A) Pie chart illustrating the fraction of PARP-1 binding sites, as determined by ChIP-seq in AC16 cells, that are also Dnase I hypersensitivity sites in human cardiomyocytes.

(B) Boxplot of normalized DNase I hypersensitivity site read depth at all sites or those that overlap with PARP-1 binding sites.

Venn Diagram depiction of PARP-1 binding site overlap with ChIP-seq peaks of (C) Histone H3K4me3, (D) Histone H3K4me1, and (E) CTCF.

ChIP-seq peak intensity ranked heatmaps of the (F) Histone H3K4me3 modification at transcription start sites, (G) NF κ B p65 binding sites, and (H) CTCF binding sites.

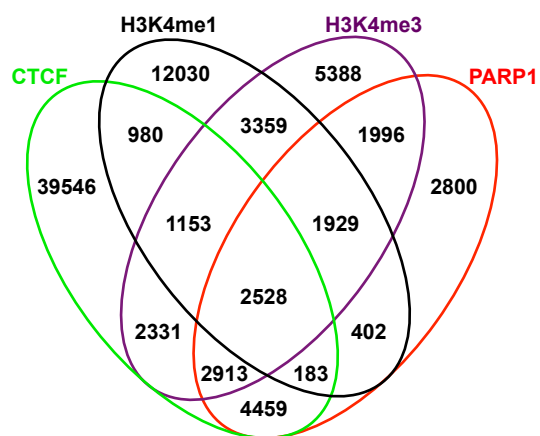


Figure 23. Overlap of Active Histone Marks, PARP-1, and CTCF binding.

Four-way Venn diagram quantifying the ChIP-seq peak overlap between PARP-1, CTCF, and active histone marks Histone H3K4me1 and Histone H3K4me3.

Results

PARP-1 Binding at Active Regions of the Genome. In order to better understand where PARP-1 binds to the genome, a standard ChIP-seq experiment was first performed for PARP-1, and regions of statistically significant PARP-1 binding were compared to digital DNase I hypersensitivity site I (DHS) profiling and transcription factor ChIP-seq data. I found that the majority of PARP-1 binding occurs at active regions of the genome as indicated by PARP-1 ChIP-seq peak overlap with adult human cardiomyocyte DHSs, a method that reflects genomic accessibility (Fig. 22A). DHSs that overlap with PARP-1 binding show significantly increased read density, suggesting that among the full category of accessible regions of the genome, PARP-1 binds to those that are most “open” and active (Fig. 22B). In order to identify which specific types of regulatory regions PARP-1 bound, I used cardiomyocyte-specific Histone H3K4me3, Histone H3K4me1, and CTCF ChIP-seq peaks as genomic waypoints indicating the genomic locations of active promoters, enhancers and insulators, respectively. I compared these ChIP-seq peaks to sites of PARP-1 binding and found substantial overlap with each of these diverse types of regulatory elements (Figs. 22C-E and 23) suggesting that genomic action by PARP-1 might extend beyond promoter proximal regions, as previously described (180). I then ranked gene promoters, NFκB binding sites (a specific enhancer), and CTCF binding sites (or insulators) by Histone H3K4me3, NFκB p65, and CTCF ChIP-seq read depth and found that PARP-1 reads were similarly ranked indicating that PARP-1 binding reflects and mirrors the level of binding by these histone modifications or transcription factors (Fig. 22F-H). Interestingly, the breadth of PARP-1 binding at the TSS or TFBSs was unexpectedly small surrounding these regulatory elements, with PARP-1 enrichment spanning less than distance

from the TSS as compared to the Histone H3K4me3 mark and more similar to that of either NFκB p65 and CTCF enrichment. These observations suggested that PARP-1 binding and action at these diverse regulatory elements might be constrained and directed toward a shared and specific nucleosomal feature.

Development of XL-MNase ChIP-seq. In order to define the substrate that PARP-1 binds I focused first on gaining deeper understanding of the exact chromatin features at regulatory elements to which PARP-1 binds. Toward this goal, I developed a modification of ChIP-seq I call XL-MNase ChIP-seq, where chromatin in formaldehyde crosslinked (XL) cells are fragmented using MNase instead of sonication and subjected to ChIP and sequencing of nucleosomal fragments. MNase assisted ChIP-seq has been performed before from uncrosslinked isolated nuclei, followed by extraction of small nucleosomal fragments under native conditions (200). Although informative, the MNase material generated using native conditions have been shown to differ from nucleosome positions *in vivo*, due to processing issues during nuclei isolation, genome fragmentation, and nucleosome extraction (201). Moreover, this technique limits nucleosome-level analysis to histones and their post-translational modifications, excluding chromatin-binding proteins such as PARP-1.

The largest obstacle to the melding of traditional MNase approaches with traditional crosslinking-based ChIP techniques is that crosslinking cells causes them to become refractory to nuclei or chromatin isolation, which is most typically performed in MNase assisted genome-wide analyses. In order to surpass this technical limitation, I instead permeabilized crosslinked cells using a non-ionic detergent and performed the MNase digestion within these intact cells (Fig.

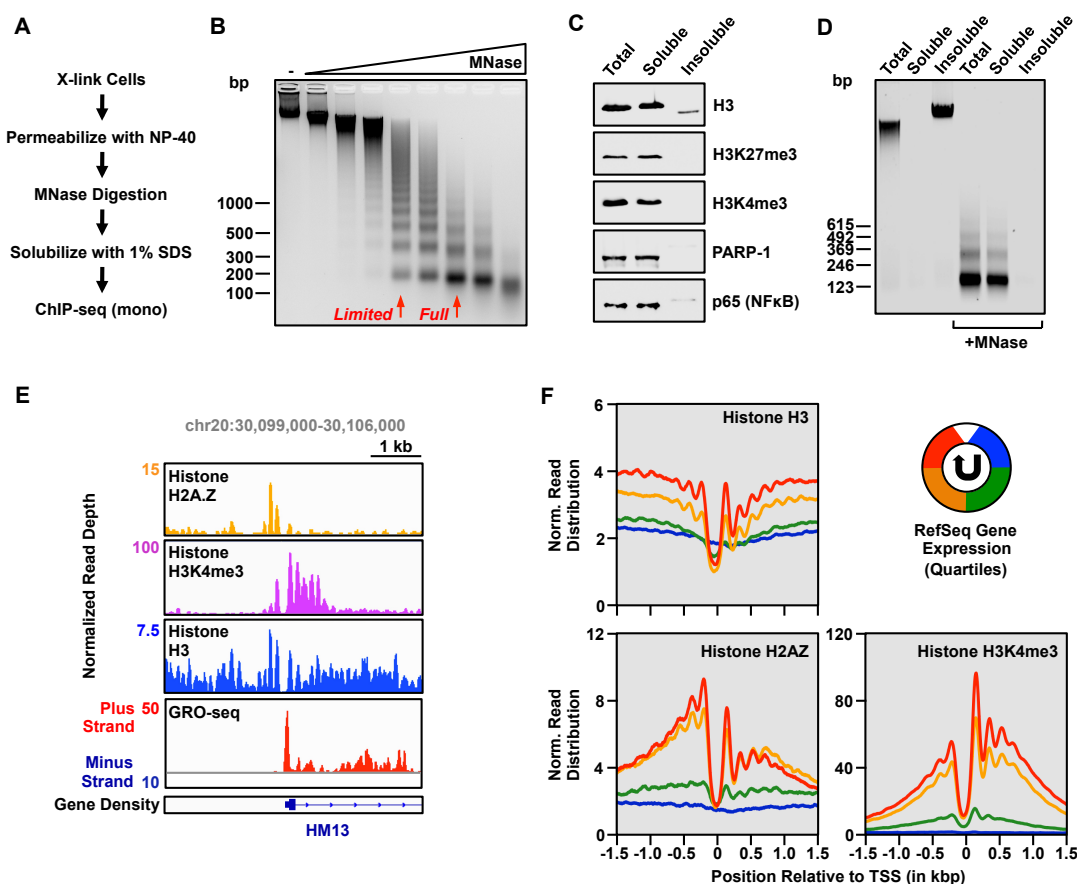


Figure 24. Development of XL-MNase ChIP-seq from crosslinked AC16 cells.

(A) Schematic depicting the overall workflow of XL-MNase ChIP-seq, wherein crosslinked cells are permeabilized to allow MNase digestion, immunoprecipitation, and deep sequencing of histones, histone marks, or chromatin binding proteins to produce nucleosome resolution maps of chromatin protein occupancy.

(B) Representative MNase digestion using crosslinked and permeabilized cells with a titration of MNase revealing a nucleosomal ladder, highlighting the “limited” and “full” conditions used in this study.

(C) Western blot analysis of total, 1% SDS solubilized, and insoluble material following full MNase digestion.

(D) Agarose gel electrophoresis of AC16 cells with and without MNase addition, showing the MNase-assisted generation of soluble chromatin for immunoprecipitation.

(E) Genome browser view of the HM13 gene promoter in human cells showing GRO-seq read density as well as a representative map of histone variant H2A.Z, histone mark H3K4me3, and Histone H3 read densities from an XL-MNase ChIP-seq experiment in AC16 cells.

(F) Metagenome analysis of average XL-MNase ChIP-seq read densities of Histone H3 (*upper left*), Histone H2A.Z (*bottom left*), and Histone H3K4me3 (*bottom right*), ranked by quartiles of expression as determined by GRO-seq.

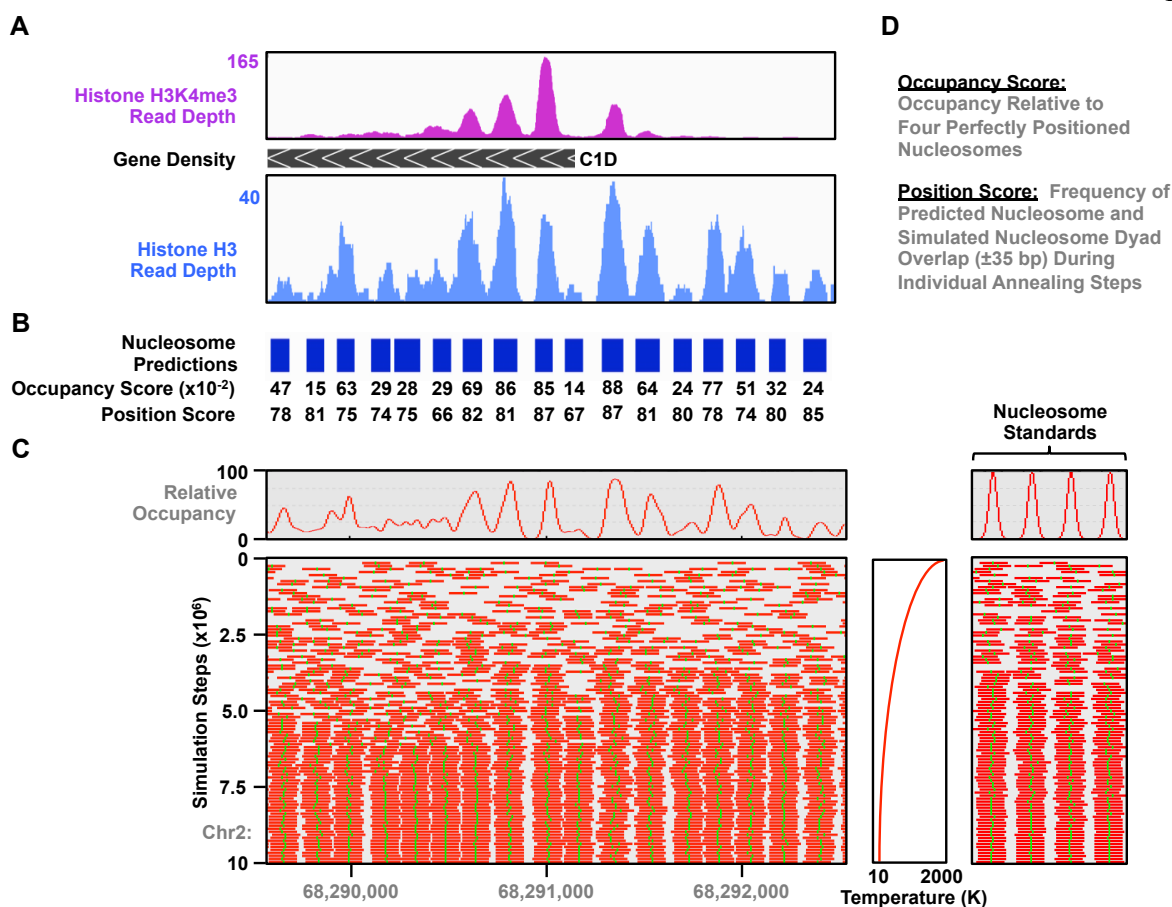


Figure 25. NucPosSimulator-based Quantification of Nucleosome Parameters from XL-MNase ChIP-seq Data.

(A) Genome browser view of the C1D gene promoter in human cells with a map of Histone H3K4me3 (*top*) and Histone H3 (*bottom*) XL-MNase ChIP-seq read densities.

(B) NucPosSimulator nucleosome predictions from Histone H3 XL-MNase ChIP-seq data in blue boxes, with relative occupancy and Position Score metric calculated for each nucleosome underneath.

(C) Plot of occupancy and predicted nucleosome positions during a NucPosSimulator simulated annealing run at the C1D gene promoter for Histone H3 XL-MNase ChIP-seq reads relative to four perfectly positioned artificial nucleosomes heuristically tuned for accurate occupancy calculation. The relative occupancy metric was calculated by NucPosSimulator relative to the artificial nucleosomes and with a sigma smoothing factor of 20. Simulations were run from 2000K to 10K and annealed nucleosome positions during individual simulation steps, depicted in red with dyad axis (nucleosome center) indicated in green, were collected every 100 steps.

(D) Definitions of relative occupancy and Position Score metrics.

24A). After either full digestion (~70% mononucleosomes) or a more limited digestion (less than 5% mononucleosomes) using MNase (Fig. 24B) soluble chromatin can be isolated from these crosslinked cells by lysis with 1% SDS (Fig. 24C-D). The SDS used for lysis can then be neutralized by dilution with non-ionic detergent as is typical with traditional sonication-based strategies for chromatin fragmentation, yielding immunoprecipitable chromatin capable of being prepared for deep sequencing.

We then used this experimental approach to perform MNase-seq and MNase ChIP-seq for Histone H3, Histone H3K4me3, Histone H3K4me1, Histone H2A.Z, and PARP-1 in AC16 cells with and without 30 minutes of TNF α stimulation, with paired-end sequencing of ChIPped mononucleosomal material. This new approach to MNase digestion and chromatin immunoprecipitation yields extraordinarily high quality maps of nucleosome density (Fig. 24E) and exceptional retention of specific nucleosomes, such as the first nucleosome after the transcription start site known as the +1 nucleosome, which often exhibit processing loss during native mononucleosome extraction protocols (Fig. 24E,F) (186, 202). I then ranked genes by quartiles of expression using GRO-seq data from AC16 cells (171) and performed metagene analyses for Histone H3, Histone H3K4me3, and histone variant H2A.Z MNase ChIP-seq, revealing an overall similar and highly resolved pattern of transcription-dependent ordering of promoter nucleosomes (Histone H3), histone variant incorporation (Histone H2A.Z) and histone mark deposition (Histone H3K4me3) as has previously been reported (Fig. 24F) (203-205). Notable in these metagene analyses is the lack in a need for smoothing during processing to generate these average profiles of nucleosome density, suggesting that not only is mammalian nucleosome architecture much less disordered than previously thought, but also that the “fuzzy”

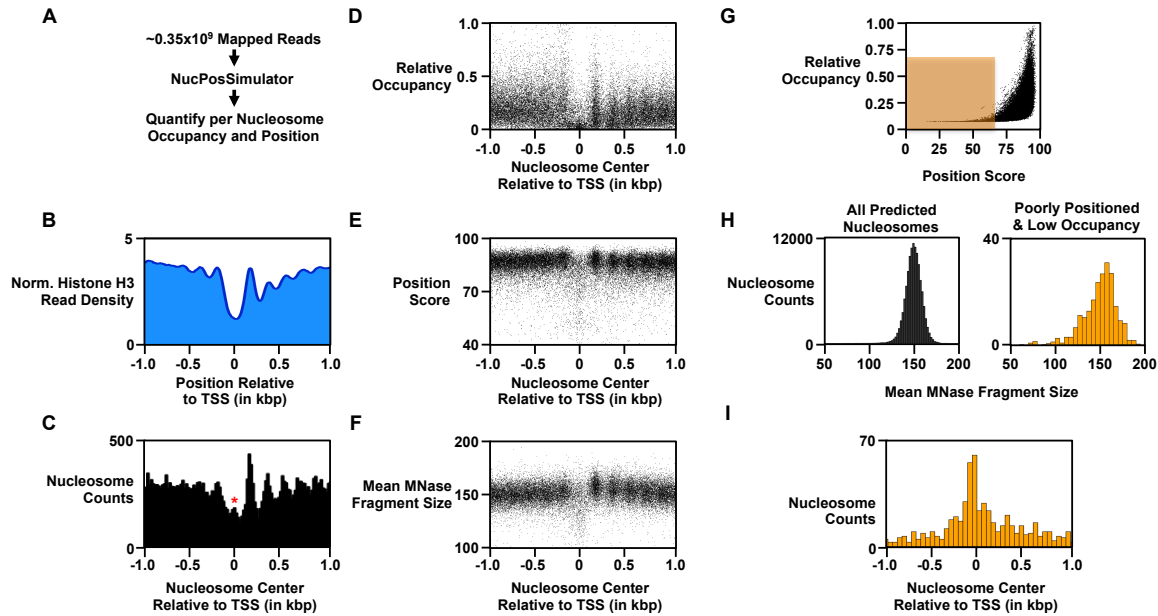


Figure 26. A poorly positioned, lowly occupied, and MNase-sensitive nucleosome resides at the “nucleosome free region” of active promoters.

(A) Schematic depicting the computational pipeline used to identify the position and occupancy of nucleosomes from fully digested Histone H3 XL-MNase ChIP-seq reads using NucPosSimulator.

(B) Metagene of normalized Histone H3 XL-MNase ChIP-seq read density from the top quartile of transcribed RefSeq genes, as determined by GRO-seq read density.

(C) Histogram of nucleosome centers called by NucPosSimulator for the top quartile of transcribed genes relative to transcription start sites, with * indicating an unexpected density within the “nucleosome free region”.

Plot of NucPosSimulator called nucleosome centers for the top quartile of transcribed genes relative to transcription start sites and their (D) relative occupancy, (E) Position Score metric, and (F) mean MNase fragment size as determined from paired-end sequencing reads.

(G) Plot of the Position Score metric and relative occupancy of all predicted nucleosomes with a population of poorly positioned (<65) and lowly occupied (<0.65) nucleosomes highlighted with a transparent orange box.

(H) Histogram of the mean MNase fragment size of all predicted nucleosomes (*left*) and the highlighted poorly positioned and lowly occupied nucleosomes (*right*) as described in figure 3G.

(I) Histogram depiction of nucleosome centers of nucleosomes highlighted in orange in figure 3G relative to the TSS.

nucleosome profiles from previous attempts at MNase-seq and MNase ChIP-seq experiments are likely the result of experimental manipulations that shroud the nucleosomal order within cells *in vivo* (200).

Characterization of Promoter Nucleosome Architecture. Given the high quality and resolution of the chromatin maps generated by our XL-MNase ChIP-seq approach, I used the NucPosSimulator analysis tool (206) to more thoroughly describe the chromatin landscape at promoters in an effort to gain deeper understanding of the subtleties of nucleosome structure to which PARP-1 might bind (Fig. 25A,B). NucPosSimulator uses a simulated annealing approach to calculate nucleosome position from raw MNase digested reads (Fig. 25C) (206). Simulated annealing is a type of metaheuristic that is uniquely capable at finding a global optimum within complex energy landscapes, such as so-called “fuzzy” regions within nucleosome maps that result from cell-to-cell heterogeneity of nucleosome positioning. This type of algorithm excels as a computational solution in the identification of a consensus nucleosome position even in the context of noisy and complex MNase chromatin maps. I iteratively ran the NucPosSimulator algorithm across all RefSeq genes to calculate the position and relative occupancy of nucleosomes relative to transcription start sites (Fig. 26A). I also incorporated into NucPosSimulator’s “off the shelf” functions a set of nucleosome standards to account for the iterative nature of the nucleosome calculations as well as an additional analysis metric I call the Position Score for each nucleosome. The Position Score metric uses the frequency of “correct” nucleosome prediction during the NucPosSimulator’s simulated annealing run for each nucleosome as a proxy for how well positioned that given nucleosome is within the cellular

population from which the MNase material was generated (Fig. 25D).

Comparison of a histogram of computationally determined individual nucleosome centers and a metagene analysis of Histone H3 MNase ChIP-seq reads from the top quartile of expressed gene promoters reveals accurate determination of individual nucleosome positions as indicated by the high degree of similarity between these two very different representations of the data (Fig. 26B,C). A surprising finding in this analysis was an unexpected density of nucleosomes found within the previously described “nucleosome free” region spanning the TSS. While previous reports show definitively that transcription factors and other chromatin binding proteins are capable of generating sub-nucleosomal footprints during MNase digestions (207), the Histone H3 ChIP in this XL-MNase CHIP-seq approach would suggest that these identified nucleosomes arise from nucleosomes or nucleosome-like particles with histone content.

Plotting NucPosSimulator-identified nucleosome centers against their assigned relative occupancy, Position Score, and mean MNase fragment size from the paired-end sequencing reads (Fig. 26D-F) revealed that this NFR-spanning nucleosome is categorically lowly occupied, poorly positioned and exceptionally nuclease sensitive. I next plotted the Position Score and relative occupancy of all identified nucleosomes, revealing the interrelated nature of these two definitions of the nucleosomal landscape, as well as a group of poorly positioned and lowly occupied nucleosomes (Fig. 26G). While the majority of nucleosomes have an essentially normally distributed mean MNase paired-end fragment distribution consistent with their being mostly composed of canonical nucleosomes, the poorly positioned and lowly occupied nucleosomes show an altered nuclease sensitivity (Fig. 26H). Mapping of the centers of these outlier group of nucleosomes reveal that they localize to the “nucleosome free” region of

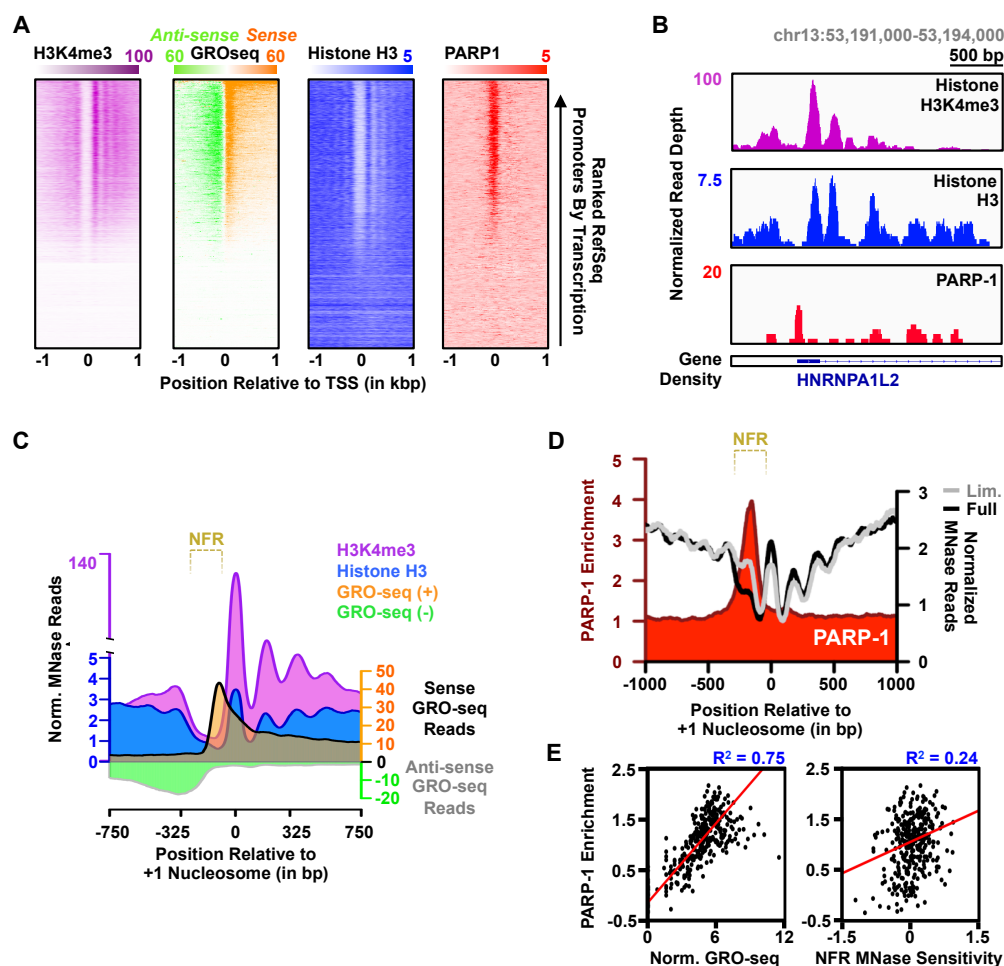


Figure 27. PARP-1 binding is correlated with MNase-sensitive nucleosomes within the “nucleosome free” regions of active genes.

(A) Heatmap representation of normalized read density for Histone H3K4me3 XL-MNase ChIP-seq (far left), GRO-seq (middle left), Histone H3 XL-MNase ChIP-seq (middle right), and PARP-1 XL-MNase ChIP-seq enrichment (far right) for all RefSeq promoters ranked by transcription (genic sense direction GRO-seq read density).

(B) Genome browser view of the HNRNPA1L2 gene promoter showing normalized XL-MNase ChIP-seq read density for Histone H3K4me3, Histone H3, and PARP-1.

Metagenome analyses of the chromatin and transcription landscape for the top quartile of transcribed genes relative to the first detectable nucleosome (+1) depicting the relationship between average read densities for (C) H3K4me3 XL-MNase ChIP-seq, Histone H3 XL-MNase ChIP-seq, and transcription (GRO-seq) as well as (D) PARP-1 XL-MNase ChIP-seq enrichment, and MNase sensitivity (limited versus full normalized MNase-seq read density).

(E) Genome-wide correlation of PARP-1 enrichment at the “nucleosome free” regions of promoters with normalized GRO-seq read density (left) or MNase sensitivity (right, limited over full normalized MNase-seq reads) for groups of 500 genes clustered by their transcription level.

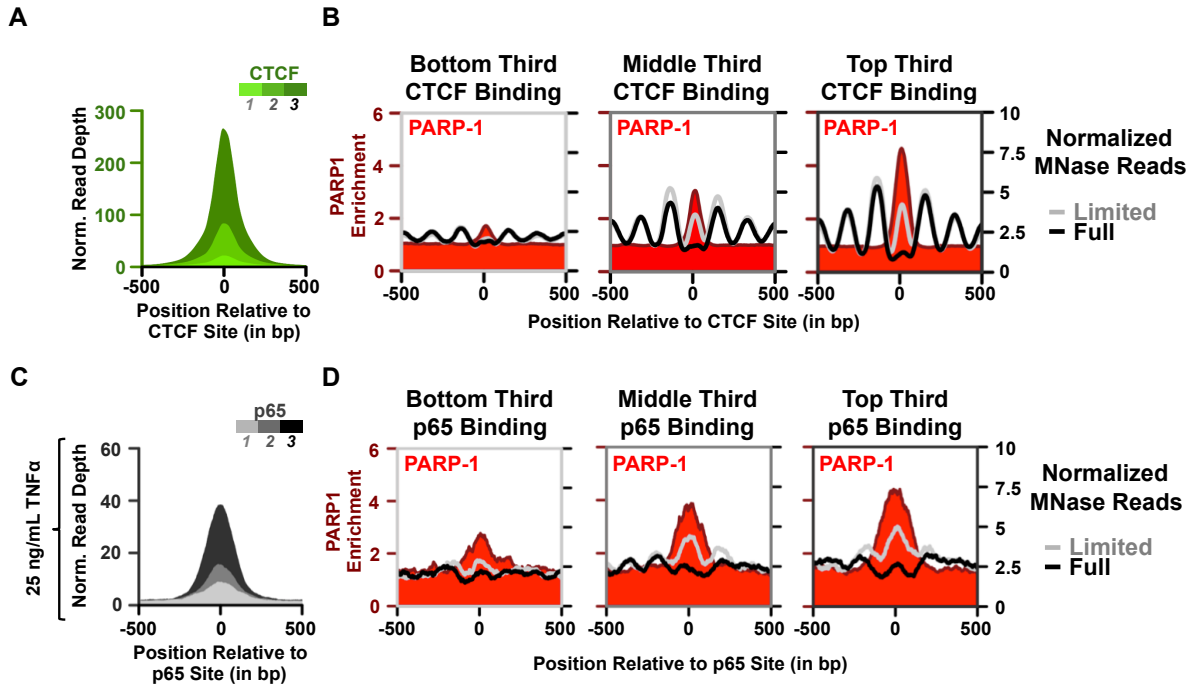


Figure 28. PARP-1 binding and MNase-sensitive nucleosomes occupy NFκB p65 and CTCF binding site “nucleosome free” regions.

(A) Metagene analysis of the bottom, middle and top thirds of CTCF ChIP-seq read density at CTCF binding sites colored by increasing intensities of green.

(B) Metagene analysis of PARP-1 XL-MNase ChIP-seq enrichment and MNase sensitivity (limited and full normalized MNase-seq read density) within the bottom (*left*), middle (*center*) and top (*right*) thirds of CTCF ChIP-seq read density at CTCF binding sites.

(C) Metagene analysis of the bottom, middle and top thirds of NFκB p65 ChIP-seq read density from TNFα-treated AC16 cells at NFκB p65 binding sites colored by increasing intensities of grayscale.

(D) Metagene analysis of PARP-1 XL-MNase ChIP-seq enrichment and MNase sensitivity (limited and full normalized MNase-seq read density) from TNFα-treated AC16 cells within the bottom (*left*), middle (*center*) and top (*right*) thirds of NFκB p65 ChIP-seq read density at NFκB p65 binding sites.

promoters spanning the transcription start site (Fig. 26I). These data are consistent with previous reports that a poorly positioned and unstable double-variant nucleosome composed of H2A.Z and H3.3 are localized to this region (208), and suggest that XL-MNase ChIP-seq is capable of identifying nucleosome structures most typically lost during previous iterations of MNase-seq.

PARP-1 Co-localization with a Nuclease-sensitive Nucleosome. Following thorough characterization of nucleosome architecture at promoters, I analyzed PARP-1 binding relative to these nucleosome features and found PARP-1 enrichment spanning the “nucleosome free” region of transcribed genes, in a pattern similar to that of the nuclease sensitive nucleosome described above (Fig 27A, B). Metagene analysis of transcription, nucleosome density and sensitivity, and PARP-1 enrichment within the top quartile of expressed genes relative to the +1 nucleosome revealed finite enrichment of PARP-1 and MNase sensitive signal within the NFR spanning the TSS (Fig. 27C, D). PARP-1 enrichment at the NFR was well-correlated with GRO-seq read density for expressed genes, consistent with previous work on PARP-1 binding in MCF-7 breast cancer cells (180), as well as with the extent of nucleosome free region nuclease sensitivity (Fig. 27E), consistent with PARP-1 binding to an MNase-sensitive nucleosome at these regions.

Since PARP-1 binding occurs not just at promoters, but also at insulators and enhancers as indicated by ChIP-seq peak overlap of PARP-1 with CTCF and Histone H3K4me1 (Figs. 22D, E and 23) I wondered whether insulator and enhancer elements might also show evidence of an MNase-sensitive nucleosome and finite PARP-1 enrichment spanning a “nucleosome free” region. Indeed at both CTCF binding sites and NFκB p65 binding sites, PARP-1 enrichment mirrored transcription factor binding intensity specifically at regions typically considered

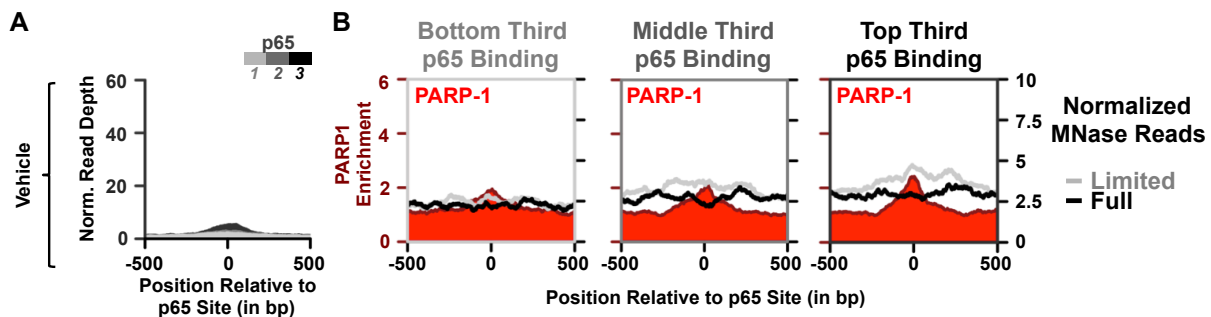


Figure 29. Absence of PARP-1 binding and MNase sensitivity at NFκB p65 binding sites without TNFα treatment.

(A) Metagene analysis of the bottom, middle and top thirds of NFκB p65 Chip-seq read density from untreated AC16 cells at NFκB p65 binding sites called in TNFα-treated AC16 cells and colored by increasing intensities of grayscale.

(B) Metagene analysis of PARP-1 enrichment and MNase sensitivity (limited and full normalized MNase-seq read density) from untreated AC16 cells within the bottom (*left*), middle (*center*) and top (*right*) thirds of NFκB p65 Chip-seq read density at NFκB p65 binding sites called in TNFα-treated AC16 cells.

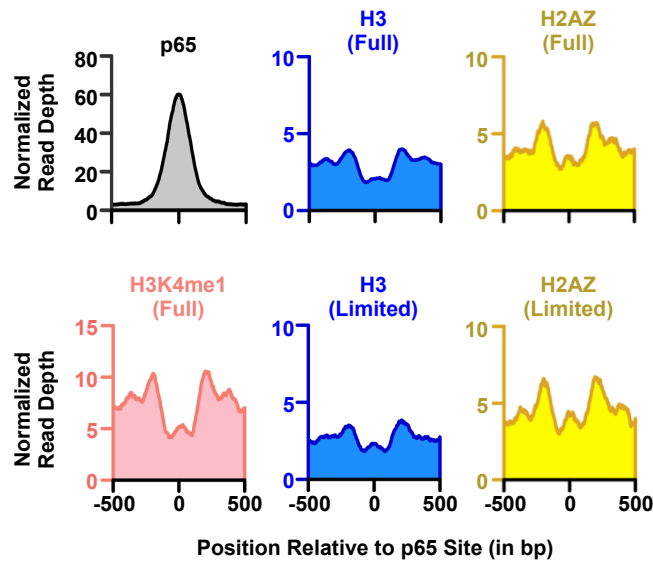


Figure 30. The chromatin landscape of MNase-sensitive nucleosomes at the “nucleosome free” region of NFκB p65 binding sites.

Metagene analysis at NFκB p65 binding sites of normalized ChIP-seq read densities for NFκB p65 (*upper left*), and XL-MNase ChIP-seq read densities with full MNase digestion for Histone H3 (*upper center*), Histone H2A.Z (*upper right*), Histone H3K4me1 (*bottom left*), and limited MNase digestion for Histone H3 (*bottom center*), and Histone H2A.Z (*bottom right*).

“nucleosome free” (Fig. 28A-D). Moreover, the NFR of these regulatory elements showed evidence of MNase sensitivity similar to expressed gene promoters as indicated by an enrichment of MNase-seq read density in a limited versus full MNase digestion condition. Notably, the presence of this nuclease-sensitive nucleosome, as well as PARP-1 enrichment were dependent on NFκB binding as indicated by the absence of these chromatin features without TNFα stimulation (Fig. 29). I then performed a metagene analysis of NFκB p65 binding sites for the histone variant H2A.Z as well as known histone enhancer mark H3K4me1 in order to gain a more full understanding of the epigenomic landscape of this NFR-spanning nucleosome (Fig. 30). I found that the NFR-spanning nucleosome contains the H3K4me1 mark and contains the histone variant H2A.Z at a level commensurate with the adjacent -1 and +1 nucleosomes (Fig. 30), suggesting that this nucleosome might be a H2A.Z/H3.3 double-variant salt sensitive nucleosome that has previously described (208).

We then looked to more fully characterize the nucleosomal landscape of CTCF sites, which are known to have a less heterogeneous nucleosomal landscape as compared to promoters (209). As previously reported, CTCF sites indeed contain very ordered nucleosomal arrays as shown in metagene analyses (Fig. 31A). At the NFR of CTCF binding sites, I found increased read density in a more limited MNase condition as well as a clearly identified nucleosome population after performing nucleosome detection using NucPosSimulator as above (Fig. 31A, B). Similar to promoters, this NFR-spanning nucleosome was poorly positioned, had low occupancy, and showed sensitivity to nuclease as indicated by a reduced mean paired-end MNase fragment size relative to adjacent nucleosomes (Fig. 31C). I analyzed whether PARP-1 enrichment occurring the NFR region of CTCF Binding sites (Fig. 31D) is correlated with the

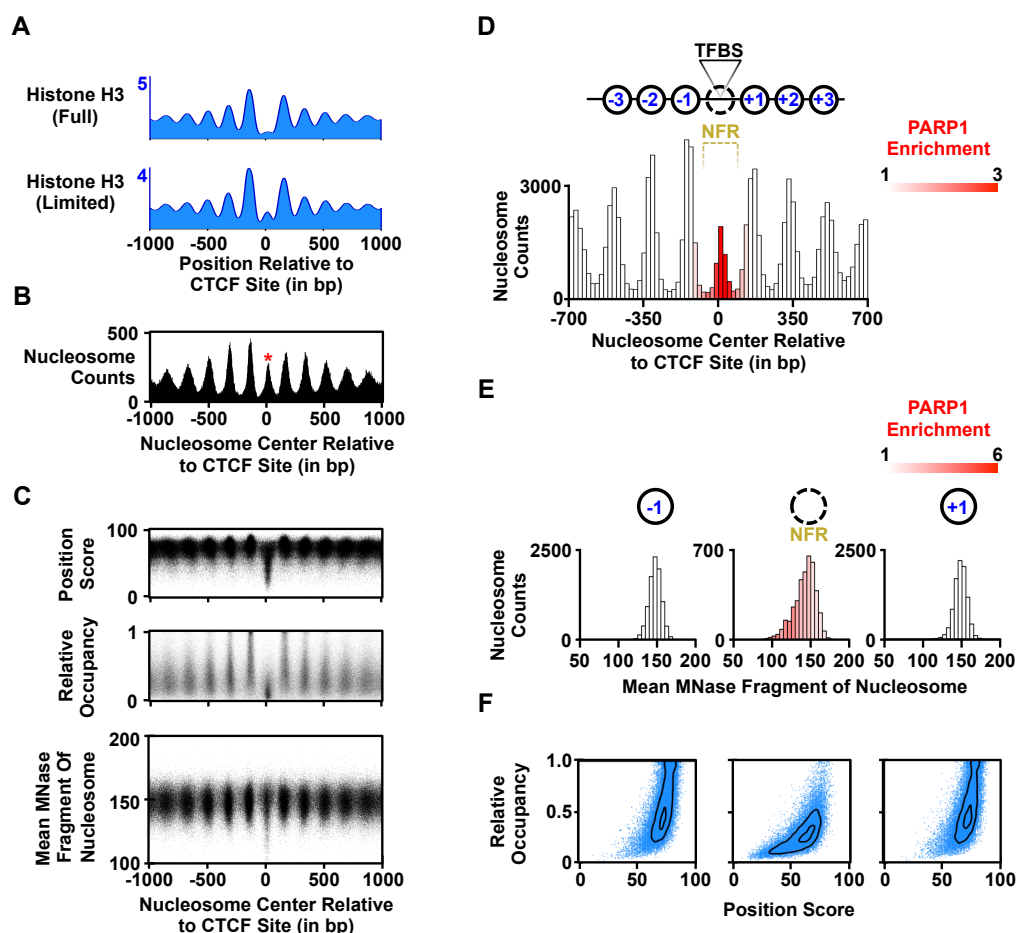


Figure 31. Characterization of PARP-1 binding to MNase-sensitive nucleosomes at CTCF binding sites.

(A) Metagenome analysis of Normalized XL-MNase Chip-seq read density at CTCF binding sites for Histone H3 with either limited (*top*) or full (*bottom*) MNase digestion.

(B) Histogram of nucleosome centers called by NucPosSimulator relative to CTCF binding sites, with * indicating called nucleosomes at the “nucleosome free region”.

Plot of NucPosSimulator called nucleosome centers relative to CTCF binding sites and their (D) Position Score metric (top), relative occupancy (middle), and mean MNase fragment size (bottom) as determined from paired-end sequencing reads.

(D) Histogram of nucleosome centers called by NucPosSimulator relative to CTCF binding sites, colored from white to red by PARP-1 XL-MNase ChIP-seq enrichment.

(E) Histogram of nucleosome counts relative to their mean MNase fragment size, as determined from paired-end sequencing reads, colored from white to red by PARP-1 XL-MNase ChIP-seq enrichment for the -1 (*left*), “NFR” (*middle*), and +1 (*right*) nucleosomes from CTCF binding sites.

(F) Plot of the Position Score and relative occupancy of -1 (*left*), “NFR” (*middle*), and +1 (*right*) nucleosomes from CTCF binding sites in blue, with a kernel density estimation of plotted points in black.

level of MNase sensitivity. I compared the mean paired-end MNase fragment size distributions from the -1, NFR, or +1 nucleosomes of CTCF sites and found that the MNase fragment size of the +1 and -1 nucleosomes were well-positioned and occupied with essentially normally distributed around the canonical nucleosome size, whereas NFR nucleosomes were poorly positioned, lowly occupied, and substantially enriched for sub-nucleosomal fragments (Fig. 31E). PARP-1 enrichment relative to mean paired-end MNase fragment size at CTCF binding site NFR

nucleosomes show that PARP-1 binding correlates with the extent of a nucleosome's sensitivity to nuclease, suggesting that the structural difference of this unique nucleosome or nucleosome-like particle are reflective of or drive PARP-1 binding at regulatory regions across the genome.

Materials and Methods

Antibodies. The custom rabbit polyclonal antiserum against PARP-1 used for Western blotting and ChIP assays was generated by using an antigen comprising the amino-terminal half of PARP-1 (55)(now available from Active Motif; cat. no. 39559). Additional antibodies were purchased from commercial sources: Histone H3 polyclonal (Abcam; ab1791), Histone H2A.Z polyclonal (Abcam; ab4171), Histone H3K4me3 polyclonal (Active Motif; cat. no. 39159), Histone H3K4me1 polyclonal (Abcam; ab8895). For Western blotting, the primary antibodies were used at a 1:4000 dilution in 5% non-fat milk made in TBST, with subsequent detection using an appropriate HRP-conjugated secondary antibody (Pierce) used at a 1:5000 dilution in 5% non-fat milk made in TBST.

Cell Culture and TNF α Stimulation of AC16 Cells. AC16 Cells were cultured and treated as previously described (171). Briefly, AC16 cells were maintained in cell culture in DMEM F-12 supplemented with 12.5% FBS, Penicillin, and Streptomycin. For stimulation of AC16 cells with TNF α , cells were (1) grown to 75% confluence, (2) switched to serum free media for 24 hours, and (3) treated with either vehicle (H₂O) or 25 ng/mL TNF α for 30 minutes.

Preparation of ChIP-seq Libraries

Formaldehyde Crosslinking and Harvest of AC16 Cells. Following either vehicle (H₂O) or 25 ng/mL TNF α treatment for 30 minutes, AC16 serum free cell culture medium was replaced with PBS containing 1% methanol-free formaldehyde to form crosslinks. After 10 minutes crosslinking was quenched by addition of glycine to 125 mM and cells were then harvested in ice cold PBS.

Sonication of Crosslinked AC16 Cells. Crosslinked AC16 cell pellets were thawed on wet ice. Cells were washed once with Farnham Lysis Buffer [5 mM PIPES, pH 8.0, 85 mM KCl, 0.5% NP-40, 1 mM DTT, 1x Protease Inhibitor Cocktail (Roche)] and lysed in ChIP Lysis Buffer (50 mM Tris•HCl, pH 7.9, 1% SDS, 10 mM EDTA, 1 mM DTT, 1x Protease Inhibitor Cocktail). Chromatin was then fragmented using a Bioruptor water bath sonicator (Diagenode) to an average of 300 bp in length.

Chromatin Immunoprecipitation. Chromatin Immunoprecipitation was then performed as previously described (171), using 8 μ L of anti-PARP-1 antisera and 80 μ L Protein A agarose resin (Millipore) for each PARP-1 ChIP.

Preparation of ChIP-seq Libraries for Deep Sequencing. Input DNA and immunoprecipitated DNA were prepared for deep sequenced as previously described (171).

Preparation of XL-MNase ChIP-seq Libraries

MNase Digestion of Permeabilized AC16 Cells. Crosslinked AC16 cell pellets consisting of cells from 2x15 cm² plates (~3x10⁶ cells), obtained as described above, were resuspended in 1 mL Buffer TM2 [10 mM Tris•HCl, pH 7.5, 2 mM MgCl₂, 1 Protease Inhibitor Cocktail (Roche), 1 mM DTT] and equilibrated on wet ice for 10 minutes. NP-40 was then added drop by drop while under gentle vortex until a final concentration of 1.5 % was reached. Permeabilized AC16 cells were then incubated on ice 10 minutes and then spun 10 minutes at 2,000 RPM at 4°C in a microcentrifuge. The permeabilized cell pellets were then resuspended in Buffer TM2 containing 2 mM CaCl₂ and 20 units (limited digestion) or 200 units (full digestion) of Micrococcal Nuclease (MNase) (Worthington Biochemical) was added for 10 minutes at 37°C and agitation by flicking every 2 minutes. MNase reactions were stopped by addition of EGTA and SDS to 10 mM and 1%, respectively. Chromatin digested using the limited digestion condition was subjected to 5 seconds of sonication in a water bath sonicator, which showed no additional fragmentation, but dramatically increased the solubility of chromatin fragments. MNase fragmented chromatin was then run on a 1% agarose gel in 1xTAE by electrophoresis to confirm the quality of the digestion and subsequently diluted and used as input for immunoprecipitation in an identical fashion to standard sonicated chromatin, as previously described (171).

Chromatin Immunoprecipitation of Nucleosomal DNA. Chromatin

immunoprecipitation was performed as previously described (171), using 1-2 micrograms of antibody per immunoprecipitation per 40 μ L of Protein A agarose resin (Millipore).

Preparation of XL-MNase ChIP-seq Libraries for Deep Sequencing. Sequencing library preparation was carried out as previously described, with the only modification being the a library size selection step corresponding to a broad ~50-300 bp using a gel extraction kit.

Analysis of ChIP-seq and XL-MNase ChIP-seq Data. Software, scripts, and other information about the analyses can be obtained by contacting the corresponding author (W.L.K.).

Quality Control. Quality control for the ChIP-seq, MNase-seq, and XL-MNase ChIP-seq data was performed using the FastQC tool (<http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>).

Read Alignment and Visualization. Both ChIP-seq single-end and XL-MNase ChIP-seq paired-end (2x50) deep sequencing reads were aligned to the human genome (hg19) using default single-end or paired-end alignment parameters in bowtie (172). The XL-MNase ChIP-seq data were prepared for visualization by calculating read depth of the middle halves of paired-end MNase fragments. The ChIP-seq data were prepared for visualization by extending reads from their 5' end to a genomic coordinate representing the middle half of the average fragment size as calculated by MACS during peak calling (see below). The ChIP-seq and XL-MNase ChIP-seq read depths were then converted into bigwig files using BEDTools (173) and visualized using the IGV genome browser (174, 175).

Peak Calling and Genome-wide Dataset Correlations. Genomic sites of enrichment for transcription or chromatin proteins, or histone modifications were calculated using MACS (178) with p-value cutoff of 1×10^{-5} for all data sets evaluated using input DNA as a control. Overlap of ChIP-seq peaks was based on overlap of 1 bp or more between MACS-called peak size and position. Correlations between GRO-seq read depth, MNase Sensitivity (normalized NFR read depth of limited over full MNase-seq) and PARP-1 Enrichment (PARP-1 over H3 MNase ChIP-seq) were performed in groups of 500 genes clustered by transcription (per base GRO-seq reads across the entire RefSeq gene annotation).

Identification of NFκB and CTCF Binding Sites Under ChIP-seq Peaks. DNA sequence underneath MACS-called ChIP-seq peaks for either NFκB p65 or CTCF were extracted from the UCSC genome browser. The consensus CTCF or RelA transcription factor motifs underneath each CTCF or p65 peaks, respectively, were found in a directed approach using FIMO. In cases where multiple CTCF or RelA motifs were found within a single MACS-called ChIP-seq peak, the site closest to the center of the ChIP-seq peak was retained. CTCF or p65 binding sites were then ordered for MetaGenes or Heatmaps by Normalized ChIP-seq read depth ± 100 bp from the 5'-most base pair of the consensus motif.

Heatmaps and MetaGenes. Read densities were calculated in a 10 kb window (± 5 kbp) surrounding the RefSeq transcription start sites, CTCF binding sites, or p65 binding sites in 25 base pair windows using a custom script in the R programming language and visualized as heatmaps using Java TreeView (179). All Metagene analyses were calculated as previously described (170) and calculated as an average of per base read depth, without smoothing of any kind.

Analysis of Nucleosome Positioning, Occupancy, and MNase Sensitivity Using NucPosSimulator. For 10 kbp surrounding (± 5 kbp) every annotated human Refseq promoter and CTCF binding site, nucleosome locations were calculated using NucPosSimulator as well as those nucleosome's relative occupancy, strength of positioning, and mean MNase fragment size (i.e. MNase Sensitivity). Nucleosome positioning and relative occupancy were calculated in relation to four perfectly positioned nucleosomes, the occupancy of which was heuristically tuned to the read depth of the Histone H3 XL-MNase ChIP-seq dataset used, in order to reduce variability from segmented calculation of position score and relative occupancy for nucleosomes at single promoters at a time with individual simulated annealing experiments. NucPosSimulator simulated annealing runs were performed with 10,000,000 steps saved every 100 steps from 2000K to 10K with a smoothing sigma of 20, a nucleosome length of 150, and a binding energy of -8. Positioning Score was defined as the percent nucleosome predictions during a simulated annealing run that fall within ± 35 bp of the final nucleosome center. Relative occupancy of each nucleosome was as calculated by NucPosSimulator. The mean MNase fragment length was calculated from Histone H3 XL-MNase ChIP-seq paired-end sequencing reads overlapping ± 35 bp of final the final nucleosome center.

Analysis of GRO-seq Data. The GRO-seq data were analyzed using software described previously (170) and the approaches described below. Software, scripts, and other information about the analyses can be obtained by contacting the corresponding author (W.L.K.).

Quality Control. Quality control for the GRO-seq data was performed using the FastQC tool (<http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>). The GRO-seq reads were

trimmed to remove adapter contamination using the default parameters of Cutadapt software (182). Reads >32 bp long were retained for alignment.

Read Alignment and Gene Annotations. Trimmed human GRO-seq reads were aligned to the human reference genome (hg19) using the bwa aligner (172) with default settings (uniquely aligned, 2 mismatches allowed, and 19 bp seed sequence). The 5'-most base pair from each read was used in all analyses. For genes with multiple TSSs, I used the TSS with the most GRO-seq reads within the first 150 bp in the genic sense direction.

Determination of Transcription Levels. Transcription levels were calculated by counting the total GRO-seq reads across the entire transcript and dividing by the length of the transcript in base pairs.

Heatmaps and Metagenes. The read densities of sense and anti-sense reads were calculated on adjacent lines for a 10 kb window surrounding each RefSeq TSS (± 5 kb) in 25 bp windows using a custom script in the R programming language. The data were visualized as heatmaps using Java TreeView (179), with sense and anti-sense reads for each RefSeq promoter and colored orange and green, respectively. Metagenes were generated as previously described (170).

Genomic Datasets. The new genomic data sets generated for these studies are as follows: (1) AC16 MNase-seq (\pm TNF α treatment) with both full and limited digestion, (2) AC16 MNase ChIP-seq (\pm TNF α treatment) with full MNase digestion and immunoprecipitation against Histone H3, Histone H3K4me3, Histone H3K4me1, Histone H2A.Z, and PARP-1, (3) AC16 MNase ChIP-seq (\pm TNF α treatment) with limited MNase digestion and immunoprecipitation

against Histone H3, Histone H2A.Z, and PARP-1, and (4) sonication fragmentation-based AC16 ChIP-seq (\pm TNF α treatment) with immunoprecipitation against PARP-1. They are available from the NCBI's GEO database using accession numbers [GSEXXXXX](#) and [GSEXXXXX](#) (in submission).

The following publically available deep sequencing data sets (including their cognate controls) were downloaded from NCBI's GEO archive using the following accession numbers: AC16 GRO-seq and NF κ B p65 ChIP-seq (GSE51225); H3K4me1 ChIP-seq (GSM706848); CTCF ChIP-seq (GSM1022657 and GSM 1022677); DNase I Hypersensitivity (GSM1014515).

Conclusions

ChIP-seq and native MNase profiling experiments are useful techniques capable of defining where proteins and nucleosomes localize across the genome. In order to identify the high-resolution nucleosomal landscape and nucleosome binding site locations of a chromatin-binding protein I developed a combination of standard ChIP-seq and MNase digestion called XL-MNase ChIP-seq. This technique uses MNase to digest chromatin to mononucleosomes inside of permeabilized and crosslinked cells, generating soluble ChIPpable material that results in chromatin maps of exceptional quality.

By adapting the simulated annealing approach of the NucPosSimulator algorithm, I were able to characterize the nucleosomal landscape at promoters and CTCF binding sites, finding at each type of regulatory element a nucleosome spanning the “nucleosome free” region. This nucleosome was generally poorly positioned, lowly occupied and exhibited an unusual sensitivity to nuclease digestion. This NFR-spanning nucleosome may be a H2A.Z/H3.3 double-

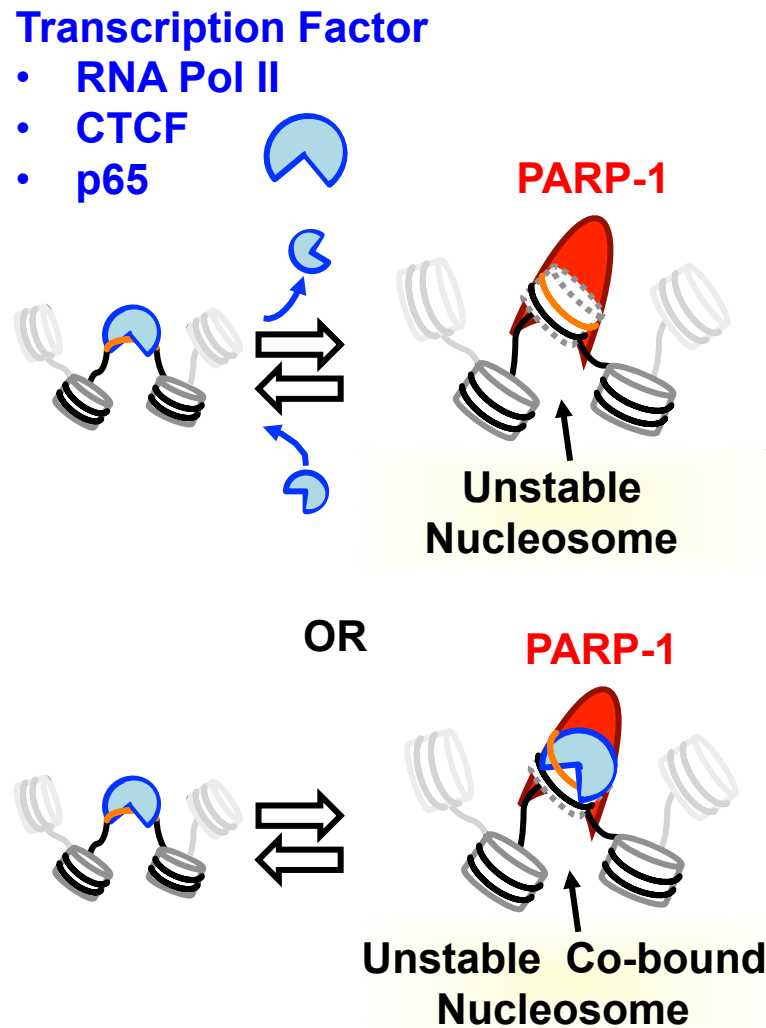


Figure 32. Model for PARP-1 binding to an unstable nucleosome at regulatory regions throughout the genome.

Depiction of a model for PARP-1 binding to an unstable NFR-localized nucleosome at regulatory regions of the genome. Transcription factors, such as RNA Polymerase II, CTCF, and NFκB, bind to target regulatory elements throughout the genome, in orange, by recognizing “open” regions with their target DNA exposed. While typically an accessible and bound regulatory element is free of nucleosomes, which are thought to sterically occlude transcription factor binding, I propose that a detectable fraction of regulatory elements contain a PARP-1-bound unstable nucleosome, which is unusually sensitive to nuclease. These unstable PARP-1-bound nucleosomes might occur at regulatory elements following loss of transcription factor binding, or transcription factors might co-bind with PARP-1 and this unstable nucleosome. This model indicates that the nucleosome landscape, PARP-1 binding, and transcription factor access to regulatory DNA is more dynamic than previously appreciated.

variant nucleosome as it shares the genomic location, H2A.Z incorporation, instability (nuclease sensitivity/salt sensitivity), and poor positioning characteristics of this previously reported nucleosome (208).

We found that PARP-1 binds to the most accessible regions of the genome, namely promoters, enhancers, and insulators. Using XL-MNase ChIP-seq I found that PARP-1 specifically bound to the nuclease-sensitive nucleosome at NFR regions shared between these three types of regulatory elements. The amount of sensitivity that nucleosomes localized to this region correlated with the intensity of PARP-1 binding, suggesting that PARP-1 may specifically recognize structural elements of this nucleosome which cause its instability or otherwise induce this phenomenon upon binding. Overall these data suggest that PARP-1 binds an unstable nucleosome at the NFR region of regulatory elements, implicating PARP-1 functioning as a regulator of access to transcription start sites and transcription factor binding sites across the genome (Fig. 32). Moreover, the ability to define a specific nucleosomal substrate for PARP-1 in vivo validates the utility of the XL-MNase ChIP-seq approach as a new functional genomics tool capable of identifying the nucleosome substrates of chromatin binding proteins.

CHAPTER FIVE

DESIGN AND USE OF SYNTHETIC ANTIBODY-LIKE ADP-RIBOSE DETECTION REAGENTS

Summary

One limitation to more accurately studying the variety of ADP-ribosylation events that occur *in vivo*, is the current state of the art in ADP-ribosylation post-translational modification detection, namely the 10H mouse monoclonal anti-poly(ADP-ribose) antibody. This antibody recognizes only extensive poly(ADP-ribosyl)ation (>10 units), incapable of recognizing short oligo or mono(ADP-ribosyl)ated proteins. Herein, I describe a suite of reagents that are able to detect and distinguish monomers, oligomers, and polymers of ADP-ribose. These reagents are constructed for recombinant expression with the advantageous properties of antibodies and deterministic specificity, selectivity, and affinity superior to the previously reported anti-ADP-ribose antibodies (*e.g.*, the 10H monoclonal antibody).

Introduction

Mono-ADP-ribose (MAR) and poly-ADP-ribose (PAR) transferase enzymes, known as PARPs, catalyse the transfer (and in the case of PAR transferase enzymes, polymerization) of ADP-ribose units from NAD^+ , which can be covalently linked glutamate, aspartate, and lysine residues of acceptor proteins (137). DNA-strand breakage has been considered the main trigger of MAR and PAR synthesis, leading either to repair of the damaged site and cell survival, or cell death, depending on the cellular context and on the intensity of the DNA insult (13). However,

other cellular components (*e.g.*, interacting proteins, nucleosomes, posttranslational modifications, *etc.*) may also stimulate MAR and PAR synthesis and the size and branching of PAR synthesized under normal conditions is still unclear (15).

There are at present 17 PARP family members, and most all PARP family members can effectively synthesize MAR and/or PAR, yet whether they produce structures comparable to that synthesized by the founding member PARP1 is still unknown (138). Some PARP family members lack conserved residues crucial for polymer elongation and may instead be mono(ADP-ribose) transferases. A detailed biochemical characterization of each PARP family member is necessary to answer the numerous questions that remain regarding PAR synthesis, transfer, function and degradation.

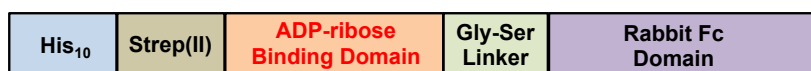
At present, the most widely used (and possibly only available) reagent to investigate ADP-ribose utilization is a monoclonal antibody against poly(ADP-ribose) (PAR). A major drawback of this reagent is that it only detects ADP-ribose in chains of more than ~10 units. In many biological systems, monomers or oligomers of ADP-ribose are the most relevant form.

Results

I have cloned, expressed, and purified a number of antibody-like ADP-ribose binding proteins comprising one of a number of different ADP-ribose binding domains (ARBDs) fused the Fc region of an immunoglobulin (Fig. 33). The recombinant fusion proteins can be purification in large (milligram) quantities of nearly homogenous and well-characterized protein from plasmid vectors (Figs. 34 and 35).

The ADP-ribose binding domains chosen to exemplify the utility of this approach are the

ADP-Ribose Binding Domain-Fc Fusion Proteins



Natural ADP-ribose binding domains used:

- 1) **WWE(RNF146)** = RNF146 WWE domain from *Homo sapiens* (Human).
- 2) **Macro(AF1521)** = AF1521 macrodomain from *Archaeoglobus fulgidus* (Archeabacterium).
- 3) **Macro(mH2A1.1)** = MacroH2A1.1 macrodomain from *Homo sapiens* (Human).
- 4) **Macro3X(PARP14)** = Macrodomains 1 through 3 of PARP14 from *Homo sapiens* (Human).

Figure 33. Schematic representation of recombinant proteins in which natural ADP-ribose binding domains (ARBDs) are fused to mammalian immunoglobulin G (IgG) Fc region. The ARBD-Fc fusion proteins contain the following components fused N-terminally to C-terminally: (1) a His₁₀ tag, which can be used for purification using nickel-affinity chromatography (or related purification methods), (2) a Strep-tag (synthetic peptide consisting of eight amino acids Trp-Ser-His-Pro-Gln-Phe-Glu-Lys), which can be used for detection or purification, (3) a natural ADP-ribose binding domain (*e.g.*, WWE(RNF146), Macro(AF1521), Macro(mH2A1.1), Macro3X(PARP14)), (4) a short linker to accommodate a restriction enzyme site in the plasmide vector, and (5) an Fc regions from a mammalian IgG (*e.g.*, rabbit, mouse, or human).

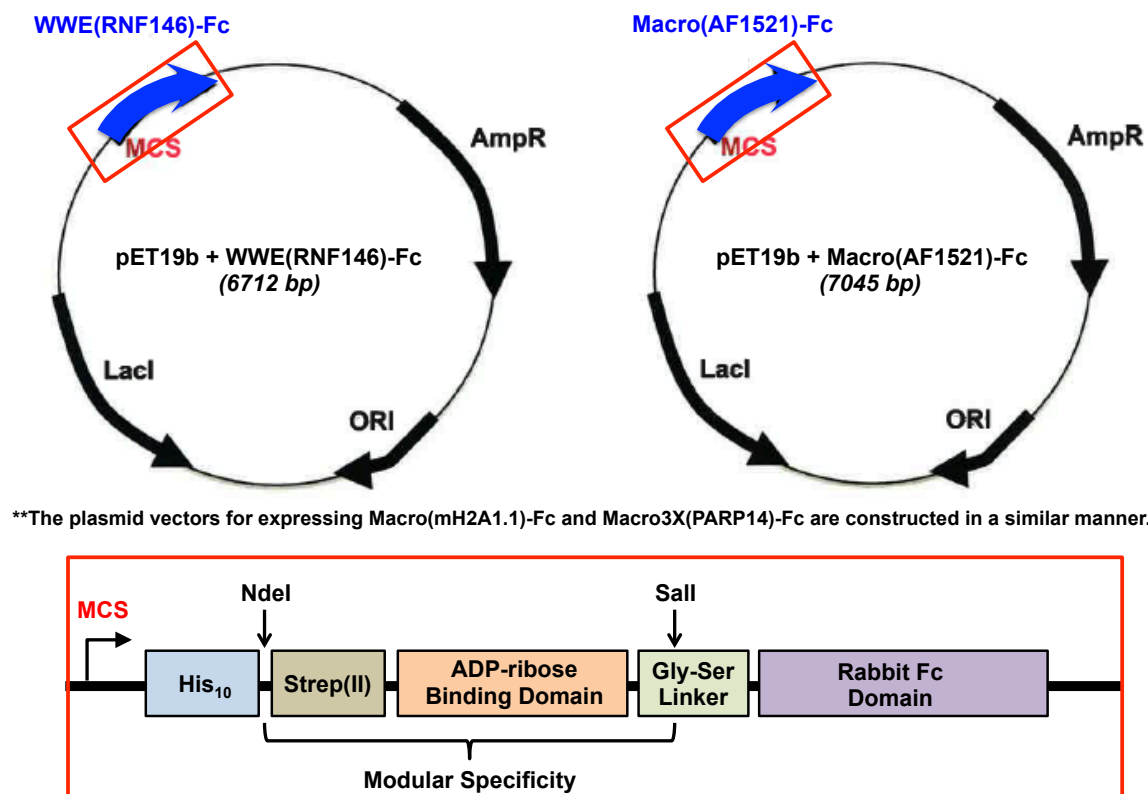


Figure 34. Bacterial plasmids for expressing ADP-Ribose binding domain-Fc fusion proteins. A standard bacterial vector for expressing ADP-Ribose binding domain-Fc fusion proteins containing a DNA sequence cassette with a sequence encoding the key components listed in Figure 33 (blue arrow in the schematics at top; red boxed inset at bottom) inserted into a multi-cloning site (MCS) of the vector. The vector also contains promoter driving transcription of the cassette (bent arrow in inset), an antibiotic selectable marker gene (*e.g.*, AmpR), and a bacterial origin of DNA replication (*e.g.*, ORI). One example of a vector containing a MCS, promoter, antibiotic selectable marker gene, and bacterial origin of DNA replication is the pET19b vector, although any vector with such elements could be used as a host for the DNA sequence cassette encoding the ADP-Ribose binding domain-Fc fusion protein.

H. sapiens WWE domain from protein RNF146, the *A. fulgidus* AF1521 macro domain, the *H. sapiens* macrodomain from histone variant macroH2A1.1, the *H. sapiens* macrodomain triplet from PARP14 (Fig. 33). The WWE domain was chosen as an ADP-ribose binding domain due to its unique mode of binding to poly(ADP-ribose) (94). The WWE domain from the RNF146 protein binds to the connection between ADP-ribose moieties in oligo and poly(ADP-ribose) chains, allowing its use as a oligo and poly(ADP-ribose) specific reagent (Figs. 36 and 37). The AF1521 macrodomain was chosen as a general ADP-ribose binding module (Figs. 36 and 37), since it has the highest reported affinity to ADP-ribose of any form for any macrodomain yet reported (103). Similarly, the macroH2A1.1 macrodomain was chosen for its ability to bind mono- and poly-ADP-ribose (Figs. 36 and 37). The PARP14 triple macrodomain was chosen due to its reported preference for mono(ADP-ribose) as a target (Figs. 36 and 37) (210).

The sequence content of ADP-ribose binding domains and the Fc portion of immunoglobulin were guided by and largely limited to reported functional crystallizable fragments of chosen domains or the homology to reported crystallizable fragments (211). The construct encodes for a protein with an N-terminal decahistidine tag (His₁₀), an optional Strep (II) tag, the ADP-ribose binding domain, a flexible glycine and serine rich linker, and the Fc region of an immunoglobulin (Fig. 1.1-1.5). The basic outline of the ARBD-Fc fusions were designed such that they contain a modular construction, wherein any ARBD could be easily placed into the construct by way of traditional cloning methods using the convenient NdeI and SalI sites flanking the ADP-ribose binding domain, which is responsible for substrate specificity (Fig. 34). These ADP-ribose binding domains were cloned into the pET19b *E. coli* expression

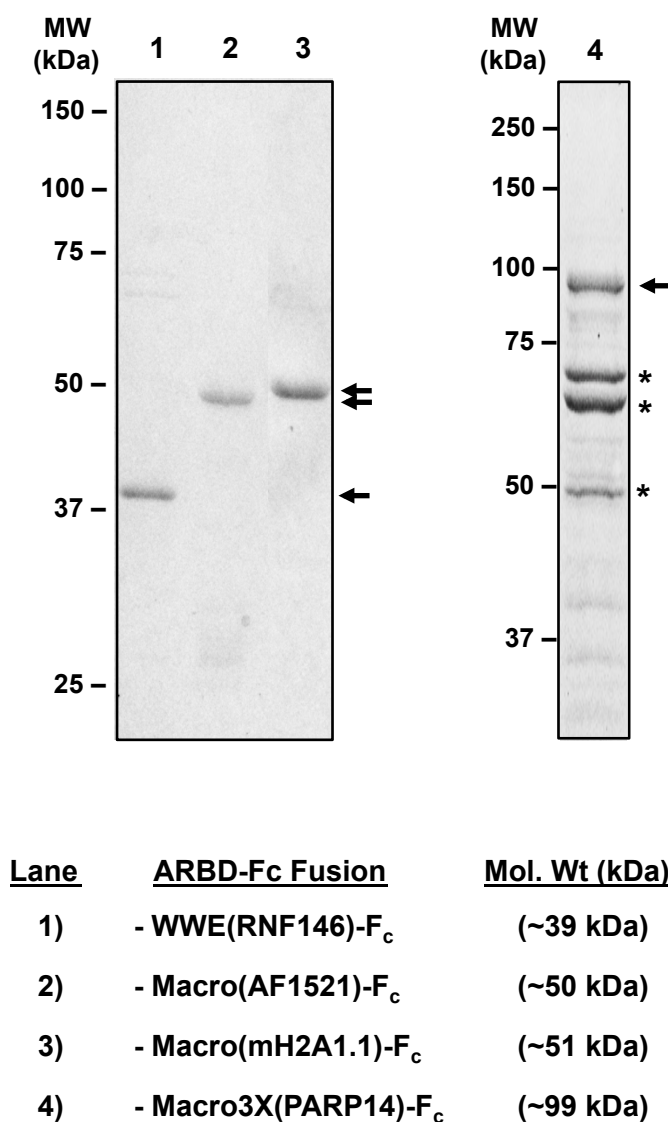


Figure 35. Analysis of purified ADP-ribose binding domain-Fc fusion proteins by polyacrylamide gel electrophoresis. ARBD-Fc fusion proteins containing the following nature ADP-ribose-binding domains: WWE(RNF146), Macro(AF1521), Macro(mH2A1.1), Macro3X(PARP14), were expressed in *E. coli*, purified under native/denaturing conditions using nickel-NTA affinity chromatography and then analyzed on a 10% PAGE-SDS. The sizes of the marker protein in kilodaltons (kDa) are indicated.

system, (Fig. 34) where the constructs expressed well yielding many milligrams of soluble purifiable protein per liter of culture (Fig. 35). Any relate vector with similar features could be used as well.

In order to test the specificity and affinity of these ARBD-Fc fusion constructs, they were purified and screened against prepared model substrates of mono-ADP-ribose, a short version of the polymer called oligo-ADP-ribose (<10 ADP-ribose units), and a lengthy version of the polymer called poly-ADP-ribose. Using both a dot blotting and Western blotting approaches, the commercially available 10H antibody along with the different ARBD-Fc fusion proteins show distinctly different patterns of recognition, demonstrably related to their known biophysical modes of recognition of ADP-ribosylation. As reported in the literature, the 10H antibody binds to extensive poly(ADP-ribosyl)ation (> ~10 ADP-ribose units), but does not bind to mono- or oligo-ADP-ribose (Figs. 36 and 37). The WWE domain-containing fusion construct, which should bind to 2 or more units of poly(ADP-ribose) since it biochemically identifies junctions between ADP-ribose units, binds both oligo and poly(ADP-ribose) with great specificity and affinity (Figs. 36 and 37). The AF1521 macro domain-containing fusion binds well to all prepared substrates with a high affinity, likely owed to its nanomolar affinity for ADP-ribose (Figs. 36 and 37). The macroH2A1.1 macro domain fusion protein binds to both mono- and poly-ADP-ribose (Figs. 36 and 37). The PARP14 triple macro domain fusion protein binds preferentially to mono-ADP-ribose, as has been reported previously (Figs. 36 and 37) (210).

Materials and Methods

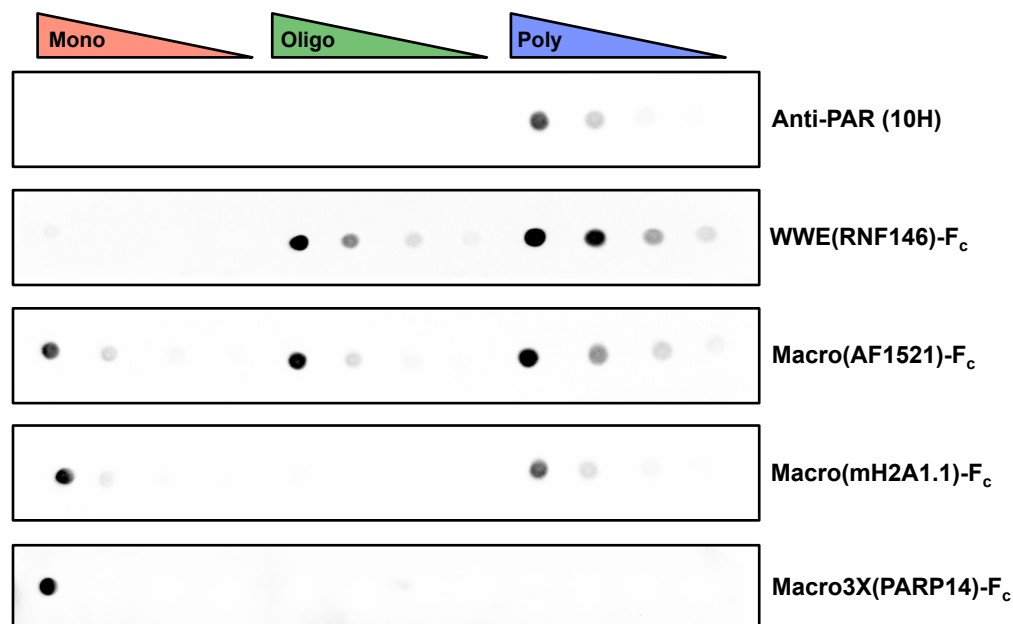


Figure 36. Dot blot assays of mono, oligo, and poly-ADP-ribose binding by ADP-ribose binding domain-Fc fusion proteins. Mono-ADP-ribose was generated by incubating purified recombinant mouse PARP3 with NAD⁺ in the presence of DNA, resulting in auto-mono-ADP-ribosylated PARP3 protein. Oligo and poly-ADP-ribose was generated by incubating purified recombinant mouse PARP1 with NAD⁺ (modulating the amount of PARP1, concentration of NAD⁺, and length of time to control the length of polymer production) in the presence of DNA, resulting in auto-oligo-ADP-ribosylated (less than 10 ADP-ribose monomers) or auto-poly-ADP-ribosylated PARP1 protein. The ADP-ribosylated proteins were spotted onto a nitrocellulose membrane in decreasing amounts as shown, then blotted using the indicated ARBD-Fc fusion proteins with a goat anti-rabbit IgG HRP detection system.

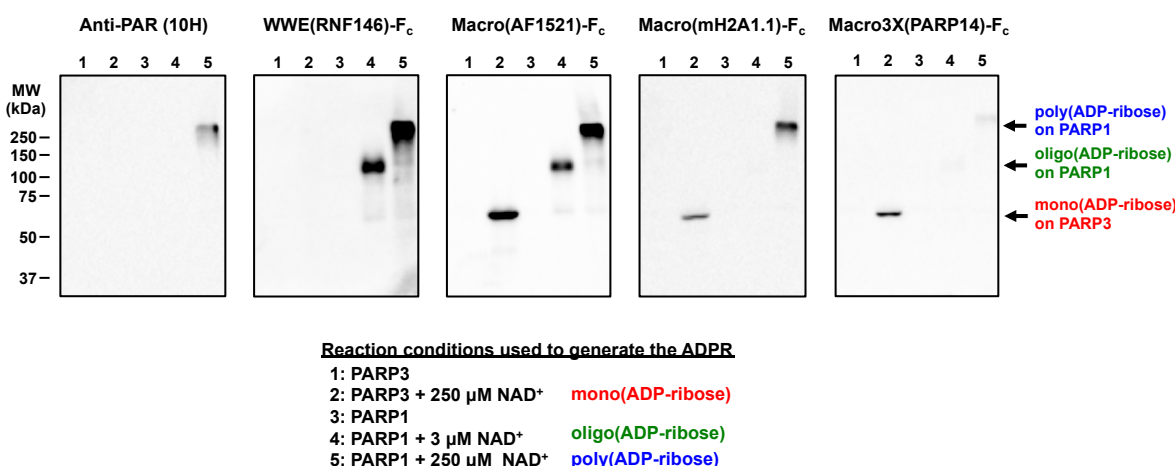


Figure 37. Western blot assays of mono, oligo, and poly-ADP-ribose binding by ADP-ribose binding domain-Fc fusion proteins. Mono-ADP-ribose on PARP3, and oligo-ADP-ribose and poly-ADP-ribose PARP1 were generated as described in FIG. 3. The ADP-ribosylated proteins, as well as non-ADP-ribosylated versions of PARP1 and PARP3 (as controls) were run on a 10% PAGE-SDS resolving gel and transferred onto a nitrocellulose membrane, then blotted using the indicated ARBD-Fc fusion proteins with a goat anti-rabbit IgG HRP detection system. The locations of the mono-ADP-ribosylated PARP3, oligo-ADP-ribosylated PARP1, and poly-ADP-ribosylated PARP1 bands are indicated by the labeled arrows. The molecular weights (MW) in kilodaltons (kDa) of marker proteins run on the same gels are shown. Note that the signal for the blot performed with Macro3X(PARP14)-Fc is from a longer exposure than the others.

Cloning and molecular biology. cDNA was generated from Trizol extracted cellular RNA originating from 293T cells or 3T3-L1 cells by using a commercially available hi-fidelity reverse transcriptase, Superscript III (Invitrogen), as described by the manufacturer.

ARBD-Fc Fusion Proteins. The DNA sequences encoding RNF146 WWE, and AF1521, MacroH2A1.1, and PARP14 macrodomains were amplified from the cDNA by Polymerase Chain Reaction (PCR) with a 5'-terminal NdeI site, 3'-terminal SalI site. The Rabbit Fc region of Immunoglobulin G (IgG) was synthesized by IDT as three DNA sequence blocks. One of the above ARBD amplicons and the three DNA sequence blocks encoding the rabbit Fc region were combined at equimolar concentrations and amplified into a contiguous linear DNA fragment using PCR with 5' ARBD forward primers and 3' Fc reverse primers. The amplified DNA cassette was cloned into pET19b (Novagen) using NdeI and BamHI sites within the MCS of the plasmid. All constructs were confirmed by sequencing.

PARP1 and PARP3. PARP1 (human; 293T as a source) and PARP3 (mouse; 3T3-L1 as a source) were amplified from high quality cDNA, made as described above, using Polymerase Chain Reaction (PCR) with Phusion DNA polymerase (NEB). A start codon and FLAG tag coding sequence were integrated 5' of the ORFs of PARP1 and PARP3 through primer design and cloned into the multiple cloning site of the pFastBac1 plasmid (Invitrogen). pFastBac1 with flag tagged PARP1 or PARP3 were recombined into a bacmid in *Escherichia coli* strain DH10bac using the Bac-to-Bac protocol, as described by manufacturer (Invitrogen). Clones passing two independent white/blue screens were inoculated into LB supplemented with 7 µg per milliliter Gentamicin, 50 µg per milliliter Kanamycin, and 10 µg per milliliter Tetracycline with

overnight growth at 37 °C. Bacmids were isolated from bacteria for transfection into Sf9 culture using sodium hydroxide/potassium acetate lysis and isopropanol precipitation of nucleic acids.

Protein Expression. The recombinant ARBD-Fc fusions, PARP1, and PARP3 were expressed as described below.

ARBD-Fc Fusion Proteins. The pET19b-ARBD-Fc expression vectors were individually transformed into separate chemically-competent *Escherichia coli* strain BL21(DE3) cells using a heat shock approach. The transformed cells were inoculated into 5 ml of Luria Broth supplemented with 100 µg/ml of Ampicillin and grown overnight with shaking at 37 °C. After overnight growth, separate 1 L cultures of LB supplemented with 100 µg/ml of Ampicillin were inoculated with 5 ml of the individual cultures and incubated at 37 °C with shaking until reaching an optical density of 0.4 OD per ml at a wavelength of 595 nm. At an optical density of 0.4 OD per milliliter, these cultures were induced for protein overexpression by addition of Isopropyl β-D-1 thiogalactopyranoside (IPTG) at a final concentration of 1 mM. The cells were grown for 2 hours at 37 °C post-induction followed by harvesting via centrifugation. Harvested bacterial cell pellets were flash frozen in liquid nitrogen and stored at -80 °C.

PARP proteins. Uninfected Serum Free Sf9 cells, cultured in SF-II 900 media (Invitrogen) were plated onto 10 cm diameter cell culture plates at a density of 1×10^6 cells per milliliter of culture. The cells were transfected with 1 µg of bacmid for expression of PARP1 or PARP3 using Cellfectin transfection reagent, as described by manufacturer (Invitrogen). After 5 hours, filter sterilized FBS, Penicillin, and Streptomycin were added to final concentrations of 10%, 100 I.U. per milliliter, and 100 µg/ml, respectively. After three days, the cell culture

medium was collected, which contained the desired baculovirus. Following two subsequent rounds of infection and amplification, a 150 ml culture of Sf9 cells in SF-II 900 media with 10% FBS in a sterile spinner flask was infected for overexpression of either PARP1 or PARP3. Two days after infection, the Sf9 cells expressing either PARP1 or PARP3 were collected by centrifugation, washed once with cold PBS, pelleted in a 50 mL plastic tube, flash frozen with liquid N₂, and stored at -80 °C for future use.

Protein Purification. The recombinant ARBD-Fc fusions, PARP1, and PARP3 were expressed as described below.

ARBD-Fc Fusion Proteins. Induced frozen bacterial cell pellets were thawed on wet ice until liquid and lysed by sonication in IMAC Lysis Buffer (10 mM Tris pH 7.5, 0.5 M NaCl, 0.1 mM EDTA, 0.1% NP-40, 10% Glycerol, 10 mM Imidazole, 1 mM PMSF, 1 mM β-Mercaptoethanol). The lysate was clarified by high speed centrifugation at 15,000 RPM for 30 minutes at 4 °C in an SS34 Rotor. The clarified supernatant was applied to 1 mL bed volume of Ni-NTA Resin (QIAGEN) equilibrated in IMAC Equilibration Buffer (10 mM Tris pH 7.5, 0.5 M NaCl, 0.1% NP-40, 10% Glycerol, 10 mM Imidazole) and incubated at 4 °C for 2 hours on a Nutator. IMAC beads were washed 4 times in 30 mL of IMAC Wash Buffer (10 mM Tris pH 7.5, 1 M NaCl, 0.2% NP-40, 10% Glycerol, 10 mM Imidazole, 1 mM PMSF) with collection by centrifugation. The purified ARBD-Fc fusion proteins were eluted by three consecutive washes of IMAC elution buffer (10 mM Tris pH 7.5, 0.2 M NaCl, 0.1% NP-40, 10% Glycerol, 500 mM Imidazole, 1 mM PMSF, 1 mM β-Mercaptoethanol). Eluates were combined and dialyzed overnight in IMAC Dialysis Buffer (10 mM Tris pH 7.5, 0.2 M NaCl, 10% Glycerol, 1 mM

PMSF, 1 mM β -Mercaptoethanol). Dialysate containing the purified fusion proteins was collected, spun at maximum speed in a microcentrifuge to remove any debris, and quantified using a Bradford Protein Assay. The fusion proteins were aliquoted in 100 μ g amounts for future use, flash frozen with liquid N₂, and stored at -80 degrees Celsius.

PARP1 and PARP3. PARP1- or PARP3-expressed frozen Sf9 cell pellets from 150 mL of cell culture was thawed on wet ice. The cells were resuspended in 7 ml of FLAG PARP Lysis Buffer [20 mM HEPES pH 7.9, 0.5 M NaCl, 4 mM MgCl₂, 0.4 mM EDTA, 20% Glycerol, 250 mM Nicotinamide, 2 mM β -Mercaptoethanol, 2x Protease Inhibitor cocktail (Roche)] and dounced 10 times on ice with tight pestle in a Wheaton dounce homogenizer. The lysate was centrifugated 30 min. at 15,000 RPM at 4 °C in an SS34 Rotor. The clarified supernatant was removed, mixed with an equal volume of FLAG Dilution Buffer (20 mM HEPES pH 7.9, 10% Glycerol, 0.02% NP-40), and sonicated with a Branson Digital Sonifier for 15 seconds at 65% amplitude in a salt ice bath. The lysate was centrifugated again as described above and the clarified lysate was removed. Clarified lysate was applied to 200 μ l of FLAG resin equilibrated in FLAG PARP Wash Buffer #1 (20 mM HEPES pH 7.9, 150 mM NaCl, 2 mM MgCl₂, 0.2 mM EDTA, 15% Glycerol, 0.01% NP-40, 100 mM Nicotinamide, 0.2 mM β -Mercaptoethanol, 1 mM PMSF, 1 μ M Aprotinin, 100 μ M Leupeptin) and incubated at 4 °C for 3 hours on a Nutator. The resin was washed once with 100 volumes of FLAG PARP Wash Buffer #1, twice with FLAG PARP Wash Buffer #2 (20 mM HEPES pH 7.9, 1 M NaCl, 2 mM MgCl₂, 0.2 mM EDTA, 15% Glycerol, 0.01% NP-40, 100 mM Nicotinamide, 0.2 mM β -Mercaptoethanol, 1 mM PMSF, 1 μ M Aprotinin, 100 μ M Leupeptin), and twice with FLAG PARP Wash Buffer #3 (20 mM HEPES pH 7.9, 150 mM NaCl, 2 mM MgCl₂, 0.2 mM EDTA, 15% Glycerol, 0.01% NP-40, 0.2 mM β -

Mercaptoethanol, 1 mM PMSF). Purified PARP1 and PARP3 were eluted from the resin using FLAG PARP Wash Buffer #3 containing 0.2 mg/mL FLAG peptide. Eluted proteins were distributed in 10 μ l aliquots, flash frozen with liquid N₂, and stored at -80 °C until use.

In Vitro PARylation of Model Substrates. PARP1 and PARP3 were incubated with either reaction buffer alone (20 mM HEPES, 5 mM MgCl₂, 5 mM CaCl₂, 0.01% NP-40, 25 mM KCl, 1mM DTT, 0.1 mg/mL BSA, 0.1 mg/mL sssDNA), as a negative control or with NAD⁺. Mono(ADP-ribose) was generated by incubation of 2 μ M PARP3 with 250 μ M NAD⁺ for 30 minutes at room temperature. Oligo(ADP-ribose) was generated by incubation of 1 μ M PARP1 with 3 μ M NAD⁺ for 30 minutes at room temperature. Poly(ADP-ribose) was generated by incubation of 100 nM PARP1 with 250 μ M NAD⁺ for 5 minutes. Reactions were stopped by addition of on third reaction volume 4x SDS Loading Buffer (8% SDS, 0.4% Bromophenol Blue, 20% Glycerol, 700 mM β -Mercaptoethanol).

ADP-ribose Detection Using ARBD-Fc Fusion Proteins. Two hundred and fifty ng of PARP3 with mono(ADP-ribose), 85 ng of PARP1 with oligo(ADP-ribose), or 8.5 ng of PARP1 with poly(ADP-ribose) or three-fold serial dilutions of each were spotted onto dry nitrocellulose membrane (dot blots). Alternatively these samples (in quantities as noted above, or similar quantified of unmodified PARP1 or PARP3 controls) were resolved by 10% PAGE-SDS and transferred to a nitrocellulose membrane (Western blots). Both the dot blots and Western blots were incubated for 1 hour at room temperature in 5% (w/v) non-fat milk in TBST (Tris buffered saline with Tween-20) with shaking. The blots were washed thoroughly with TBST and then

incubated with a specific ARBD-Fc fusion protein at 10 ng/ μ l in 1% (w/v) non-fat milk in TBST for 1 hour at room temperature with shaking. Membranes were again washed thoroughly with TBST and incubated for 1 hour with 1:8000 dilution Goat anti-Rabbit antibody conjugated to horse radish peroxidase (Thermo) in 1% (w/v) non-fat milk powder in TBST for 1 hour at room temperature with shaking. The membranes were washed thoroughly with TBST, developed with SuperSignal West Chemiluminescent reagents (Pierce), and detected with a Chemi-Doc system (Bio-Rad).

Conclusions

Here, I've described the design, creation, and use of a novel and superiorly sensitive approach to detect ADP-ribosylation modifications. Where before, detection of ADP-ribosylation was only possible on extensively post-translationally modified proteins using a monoclonal antibody, I have now generated a series of protein fusions capable of detecting mono, oligo, and poly(ADP-ribosyl)ated substrates with great sensitivity and specificity. These immune-based ADP-ribose detection reagents use a modular design, whereby one or more ADP-ribose-binding domains are fused to the Fc region of an immunoglobulin, which facilitates their use in the all biotechnological applications currently designed for antibodies. This spectrum of reagents allows more sensitive detection of all forms of the ADP-ribose modification than previously possible with available reagents. Moreover these reagents can detect mono(ADP-ribose), and oligo(ADP-ribose), where no reagent capable of this feat existed before.

This collection of fusion proteins allows detection of multiple different forms of ADP-ribose (mono, oligo, poly) with affinities and specificities that are not observed with available

reagents (e.g., the 10H monoclonal antibody) with the functionality of an immune-related reagent. An advantage of this approach is that the domains conferring ADP-ribose specificity use well characterized biophysical mechanisms for mono, oligo, or poly-ADP-ribose recognition, allowing predictably identification of the target modifications. In this way the reagent can be made to act in a predictable and rational manner related to the domain of choice, whereas an antibody, such as the 10H anti-PAR monoclonal antibody, binds its target in an uncharacterized and biophysically unknown manner. In addition, the recombinant nature of the fusion proteins allows purification of large quantities of nearly homogenous and well-characterized protein from plasmid vectors.

Current understanding of cellular ADP-ribosylation is that its use as a post-translational modification is often in the context of rapid signal transduction, often being a marker of various types of cellular stress. Where previously only the most egregious DNA lesion-induced PARylation events were routinely detectable using current technology, the sensitive and substrate-specific reagents described herein constitute a new method for understanding how ADP-ribosylation might play a role in more subtly important biological events. Even more importantly, the combination of these reagents on a given set of biological samples now allows parsing of the type of modification contributing to, or correlated with, a particular biological process, including disease states such as cancer. By using the ARBD-Fc Fusion proteins which can detect either all ADP-ribosylated substrates, mono(ADP-ribosyl)ated substrates, or oligo and poly(ADP-ribosyl)ated substrates, the type of ADP-ribosylation content of a biological sample can be deduced using these reagents. This type of specificity, heretofor unachievable using current technology, may also be used to shed light on clinical trials using PARP protein inhibitors in

order to screen patient samples and determine whether their ADP-ribosylation content reflects study outcomes.

BIBLIOGRAPHY

1. Y. Katoh *et al.*, Methionine Adenosyltransferase II serves as a transcriptional corepressor of Maf Oncoprotein. *Mol Cell*, (2011).
2. T. Zhang *et al.*, Enzymes in the NAD⁺ salvage pathway regulate SIRT1 activity at target gene promoters. *J Biol Chem* **284**, 20408-20417 (2009).
3. J. Larsson, J. Zhang, A. Rasmuson-Lestander, Mutations in the *Drosophila melanogaster* gene encoding S-adenosylmethionine synthetase [corrected] suppress position-effect variegation. *Genetics* **143**, 887-896 (1996).
4. G. G. Liou, J. C. Tanny, R. G. Kruger, T. Walz, D. Moazed, Assembly of the SIR complex and its regulation by O-acetyl-ADP-ribose, a product of NAD-dependent histone deacetylation. *Cell* **121**, 515-527 (2005).
5. H. Niida *et al.*, Essential role of Tip60-dependent recruitment of ribonucleotide reductase at DNA damage sites in DNA repair during G1 phase. *Genes Dev* **24**, 333-338 (2010).
6. H. Takahashi, J. M. McCaffery, R. A. Irizarry, J. D. Boeke, Nucleocytosolic acetyl-coenzyme a synthetase is required for histone acetylation and global transcription. *Mol Cell* **23**, 207-217 (2006).
7. K. E. Wellen *et al.*, ATP-citrate lyase links cellular metabolism to histone acetylation. *Science* **324**, 1076-1080 (2009).
8. V. Kumar *et al.*, Transcription corepressor CtBP is an NAD(+)-regulated dehydrogenase. *Mol Cell* **10**, 857-869 (2002).
9. Q. Zhang, D. W. Piston, R. H. Goodman, Regulation of corepressor function by nuclear NADH. *Science* **295**, 1895-1897 (2002).
10. J. Rutter, M. Reick, L. C. Wu, S. L. McKnight, Regulation of clock and NPAS2 DNA binding by the redox state of NAD cofactors. *Science* **293**, 510-514 (2001).
11. P. A. Srere, Complexes of sequential metabolic enzymes. *Annu Rev Biochem* **56**, 89-124 (1987).
12. P. O. Hassa, S. S. Haenni, M. Elser, M. O. Hottiger, Nuclear ADP-ribosylation reactions in mammalian cells: where are we today and where are we going? *Microbiol Mol Biol Rev* **70**, 789-829 (2006).
13. P. O. Hassa, M. O. Hottiger, The diverse biological roles of mammalian PARPS, a small but powerful family of poly-ADP-ribose polymerases. *Front Biosci* **13**, 3046-3082 (2008).
14. V. Schreiber, F. Dantzer, J. C. Ame, G. de Murcia, Poly(ADP-ribose): novel functions for an old molecule. *Nat Rev Mol Cell Biol* **7**, 517-528 (2006).
15. R. Krishnakumar, W. L. Kraus, The PARP side of the nucleus: molecular actions, physiological outcomes, and clinical targets. *Mol Cell* **39**, 8-24 (2010).
16. X. Luo, W. L. Kraus, On PAR with PARP: cellular stress signaling through poly(ADP-ribose) and PARP-1. *Genes Dev* **26**, 417-432 (2012).
17. P. O. Hassa, M. O. Hottiger, The functional role of poly(ADP-ribose)polymerase 1 as novel coactivator of NF-kappaB in inflammatory disorders. *Cell Mol Life Sci* **59**, 1534-1553 (2002).

18. W. L. Kraus, Transcriptional control by PARP-1: chromatin modulation, enhancer-binding, coregulation, and insulation. *Curr Opin Cell Biol* **20**, 294-302 (2008).
19. Y. Ji, A. V. Tulin, The roles of PARP1 in gene control and cell differentiation. *Curr Opin Genet Dev* **20**, 512-518 (2010).
20. M. Y. Kim, T. Zhang, W. L. Kraus, Poly(ADP-ribosyl)ation by PARP-1: 'PAR-laying' NAD⁺ into a nuclear signal. *Genes Dev* **19**, 1951-1967 (2005).
21. M. Cohen-Armon, PARP-1 activation in the ERK signaling pathway. *Trends Pharmacol Sci* **28**, 556-560 (2007).
22. M. Rouleau, A. Patel, M. J. Hendzel, S. H. Kaufmann, G. G. Poirier, PARP inhibition: PARP1 and beyond. *Nat Rev Cancer* **10**, 293-301 (2010).
23. R. K. Sodhi, N. Singh, A. S. Jaggi, Poly(ADP-ribose) polymerase-1 (PARP-1) and its therapeutic implications. *Vascul Pharmacol* **53**, 77-87 (2010).
24. C. Underhill, M. Toulmonde, H. Bonnefoi, A review of PARP inhibitors: from bench to bedside. *Ann Oncol* **22**, 268-279 (2011).
25. M. L. Telli, PARP inhibitors in cancer: moving beyond BRCA. *Lancet Oncol* **12**, 827-828 (2011).
26. M. O. Hottiger, P. O. Hassa, B. Luscher, H. Schuler, F. Koch-Nolte, Toward a unified nomenclature for mammalian ADP-ribosyltransferases. *Trends Biochem Sci* **35**, 208-219 (2010).
27. J. C. Ame *et al.*, PARP-2, A novel mammalian DNA damage-dependent poly(ADP-ribose) polymerase. *J Biol Chem* **274**, 17860-17868 (1999).
28. C. Boehler *et al.*, Poly(ADP-ribose) polymerase 3 (PARP3), a newcomer in cellular response to DNA damage and mitotic progression. *Proc Natl Acad Sci U S A* **108**, 2783-2788 (2011).
29. S. L. Rulten *et al.*, PARP-3 and APLF function together to accelerate nonhomologous end-joining. *Mol Cell* **41**, 33-45 (2011).
30. J. I. Sbodio, N. W. Chi, Identification of a tankyrase-binding motif shared by IRAP, TAB182, and human TRF1 but not mouse TRF1. NuMA contains this RXXPDG motif and is a novel tankyrase partner. *J Biol Chem* **277**, 31887-31892 (2002).
31. A. Ruf, V. Rolli, G. de Murcia, G. E. Schulz, The mechanism of the elongation and branching reaction of poly(ADP-ribose) polymerase as derived from crystal structures and mutagenesis. *J Mol Biol* **278**, 57-65 (1998).
32. A. Ruf, G. de Murcia, G. E. Schulz, Inhibitor and NAD⁺ binding to poly(ADP-ribose) polymerase as derived from crystal structures and homology modeling. *Biochemistry* **37**, 3893-3900 (1998).
33. D. Slade *et al.*, The structure and catalytic mechanism of a poly(ADP-ribose) glycohydrolase. *Nature* **477**, 616-620 (2011).
34. S. W. Yu *et al.*, Mediation of poly(ADP-ribose) polymerase-1-dependent cell death by apoptosis-inducing factor. *Science* **297**, 259-263 (2002).
35. M. Kolisek, A. Beck, A. Fleig, R. Penner, Cyclic ADP-ribose and hydrogen peroxide synergize with ADP-ribose in the activation of TRPM2 channels. *Mol Cell* **18**, 61-69 (2005).
36. B. C. Smith, W. C. Hallows, J. M. Denu, A continuous microplate assay for sirtuins and nicotinamide-producing enzymes. *Anal Biochem* **394**, 101-109 (2009).

37. O. Loseva *et al.*, PARP-3 is a mono-ADP-ribosylase that activates PARP-1 in the absence of DNA. *J Biol Chem* **285**, 8054-8060 (2010).
38. R. C. Aguiar, K. Takeyama, C. He, K. Kreinbrink, M. A. Shipp, B-aggressive lymphoma family proteins have unique domains that modulate transcription and exhibit poly(ADP-ribose) polymerase activity. *J Biol Chem* **280**, 33756-33765 (2005).
39. H. Kleine *et al.*, Substrate-assisted catalysis by PARP10 limits its activity to mono-ADP-ribosylation. *Mol Cell* **32**, 57-69 (2008).
40. C. C. Kiehlbauch, N. Aboul-Ela, E. L. Jacobson, D. P. Ringer, M. K. Jacobson, High resolution fractionation and characterization of ADP-ribose polymers. *Anal Biochem* **208**, 26-34 (1993).
41. H. Otto *et al.*, In silico characterization of the family of PARP-like poly(ADP-ribosyl)transferases (pARTs). *BMC Genomics* **6**, 139 (2005).
42. C. E. Bell, D. Eisenberg, Crystal structure of diphtheria toxin bound to nicotinamide adenine dinucleotide. *Biochemistry* **35**, 1137-1149 (1996).
43. M. L. Meyer-Ficca, R. G. Meyer, D. L. Coyle, E. L. Jacobson, M. K. Jacobson, Human poly(ADP-ribose) glycohydrolase is expressed in alternative splice variants yielding isoforms that localize to different cell compartments. *Exp Cell Res* **297**, 521-532 (2004).
44. M. Miwa, M. Tanaka, T. Matsushima, T. Sugimura, Purification and properties of glycohydrolase from calf thymus splitting ribose-ribose linkages of poly(adenosine diphosphate ribose). *J Biol Chem* **249**, 3475-3482 (1974).
45. S. Oka, J. Kato, J. Moss, Identification and characterization of a mammalian 39-kDa poly(ADP-ribose) glycohydrolase. *J Biol Chem* **281**, 705-713 (2006).
46. T. Ono, A. Kasamatsu, S. Oka, J. Moss, The 39-kDa poly(ADP-ribose) glycohydrolase ARH3 hydrolyzes O-acetyl-ADP-ribose, a product of the Sir2 family of acetyl-histone deacetylases. *Proc Natl Acad Sci U S A* **103**, 16687-16691 (2006).
47. M. Niere *et al.*, ADP-Ribosylhydrolase 3 (ARH3), not Poly-ADP-Ribose Glycohydrolase (PARG) Isoforms, are Responsible for Degradation of Mitochondrial Matrix-Associated Poly-ADP-Ribose. *J Biol Chem*, (2012).
48. A. G. McLennan, The Nudix hydrolase superfamily. *Cell Mol Life Sci* **63**, 123-143 (2006).
49. M. F. Langelier, J. L. Planck, S. Roy, J. M. Pascal, Crystal structures of poly(ADP-ribose) polymerase-1 (PARP-1) zinc fingers bound to DNA: structural and functional insights into DNA-dependent PARP-1 activity. *J Biol Chem* **286**, 10690-10701 (2011).
50. K. W. Caldecott, S. Aoufouchi, P. Johnson, S. Shall, XRCC1 polypeptide interacts with DNA polymerase beta and possibly poly (ADP-ribose) polymerase, and DNA ligase III is a novel molecular 'nick-sensor' in vitro. *Nucleic Acids Res* **24**, 4387-4394 (1996).
51. M. F. Langelier, J. L. Planck, S. Roy, J. M. Pascal, Structural basis for DNA damage-dependent poly(ADP-ribosyl)ation by human PARP-1. *Science* **336**, 728-732 (2012).
52. W. Lilyestrom, M. J. van der Woerd, N. Clark, K. Luger, Structural and biophysical studies of human PARP-1 in complex with damaged DNA. *J Mol Biol* **395**, 983-994 (2010).

53. M. Altmeyer, S. Messner, P. O. Hassa, M. Fey, M. O. Hottiger, Molecular mechanism of poly(ADP-ribosyl)ation by PARP1 and identification of lysine residues as ADP-ribose acceptor sites. *Nucleic Acids Res* **37**, 3723-3738 (2009).
54. L. C. Watson *et al.*, The glucocorticoid receptor dimer interface allosterically transmits sequence-specific DNA signals. *Nat Struct Mol Biol* **20**, 876-883 (2013).
55. M. Y. Kim, S. Mauro, N. Gevry, J. T. Lis, W. L. Kraus, NAD⁺-dependent modulation of chromatin structure and transcription by nucleosome binding properties of PARP-1. *Cell* **119**, 803-814 (2004).
56. D. D'Amours, S. Desnoyers, I. D'Silva, G. G. Poirier, Poly(ADP-ribosyl)ation reactions in the regulation of nuclear functions. *Biochem J* **342** (Pt 2), 249-268 (1999).
57. W. L. Kraus, J. T. Lis, PARP goes transcription. *Cell* **113**, 677-683 (2003).
58. A. Pinnola, N. Naumova, M. Shah, A. V. Tulin, Nucleosomal core histones mediate dynamic regulation of poly(ADP-ribose) polymerase 1 protein binding to chromatin and induction of its enzymatic activity. *J Biol Chem* **282**, 32511-32519 (2007).
59. E. Kotova *et al.*, Drosophila histone H2A variant (H2Av) controls poly(ADP-ribose) polymerase 1 (PARP1) activation in chromatin. *Proc Natl Acad Sci U S A* **108**, 6205-6210 (2011).
60. S. J. Petesch, J. T. Lis, Rapid, transcription-independent loss of nucleosomes over a large chromatin domain at Hsp70 loci. *Cell* **134**, 74-84 (2008).
61. S. J. Petesch, J. T. Lis, Activator-Induced Spread of Poly(ADP-Ribose) Polymerase Promotes Nucleosome Loss at Hsp70. *Mol Cell* **45**, 64-74 (2012).
62. V. Schreiber *et al.*, Poly(ADP-ribose) polymerase-2 (PARP-2) is required for efficient base excision DNA repair in association with PARP-1 and XRCC1. *J Biol Chem* **277**, 23028-23036 (2002).
63. M. Cohen-Armon *et al.*, DNA-independent PARP-1 activation by phosphorylated ERK2 increases Elk1 activity: a link to histone acetylation. *Mol Cell* **25**, 297-308 (2007).
64. N. W. Chi, H. F. Lodish, Tankyrase is a golgi-associated mitogen-activated protein kinase substrate that interacts with IRAP in GLUT4 vesicles. *J Biol Chem* **275**, 38437-38444 (2000).
65. S. Guettler *et al.*, Structural basis and sequence rules for substrate recognition by Tankyrase explain the basis for cherubism disease. *Cell* **147**, 1340-1354 (2011).
66. N. Levaot *et al.*, Loss of Tankyrase-mediated destruction of 3BP2 is the underlying pathogenic mechanism of cherubism. *Cell* **147**, 1324-1339 (2011).
67. H. Ryu *et al.*, PIASy mediates SUMO-2/3 conjugation of poly(ADP-ribose) polymerase 1 (PARP1) on mitotic chromosomes. *J Biol Chem* **285**, 14415-14423 (2010).
68. N. Martin *et al.*, PARP-1 transcriptional activity is regulated by sumoylation upon heat shock. *EMBO J* **28**, 3534-3548 (2009).
69. Z. Mao *et al.*, SIRT6 promotes DNA repair under stress by activating PARP1. *Science* **332**, 1443-1446 (2011).
70. H. Mendoza-Alvarez, R. Alvarez-Gonzalez, Poly(ADP-ribose) polymerase is a catalytic dimer and the automodification reaction is intermolecular. *J Biol Chem* **268**, 22575-22580 (1993).

71. A. K. Leung *et al.*, Poly(ADP-Ribose) Regulates Stress Responses and MicroRNA Activity in the Cytoplasm. *Mol Cell* **42**, 489-499 (2011).
72. R. J. Collier, Understanding the mode of action of diphtheria toxin: a perspective on progress during the 20th century. *Toxicon* **39**, 1793-1803 (2001).
73. G. Asher *et al.*, Poly(ADP-ribose) polymerase 1 participates in the phase entrainment of circadian clocks to feeding. *Cell* **142**, 943-953 (2010).
74. D. M. Chou *et al.*, A chromatin localization screen reveals poly (ADP ribose)-regulated recruitment of the repressive polycomb and NuRD complexes to sites of DNA damage. *Proc Natl Acad Sci U S A* **107**, 18475-18480 (2010).
75. D. Ahel *et al.*, Poly(ADP-ribose)-dependent regulation of DNA repair by the chromatin remodeling enzyme ALC1. *Science* **325**, 1240-1243 (2009).
76. G. Y. Li *et al.*, Structure and identification of ADP-ribose recognition motifs of APLF and role in the DNA damage response. *Proc Natl Acad Sci U S A* **107**, 9129-9134 (2010).
77. I. Ahel *et al.*, Poly(ADP-ribose)-binding zinc finger motifs in DNA repair/checkpoint proteins. *Nature* **451**, 81-85 (2008).
78. M. Masson *et al.*, XRCC1 is specifically associated with poly(ADP-ribose) polymerase and negatively regulates its activity following DNA damage. *Mol Cell Biol* **18**, 3563-3571 (1998).
79. S. Okano, L. Lan, K. W. Caldecott, T. Mori, A. Yasui, Spatial and temporal cellular responses to single-strand breaks in human cells. *Mol Cell Biol* **23**, 3974-3981 (2003).
80. R. Krishnakumar, W. L. Kraus, PARP-1 regulates chromatin structure and transcription through a KDM5B-dependent pathway. *Mol Cell* **39**, 736-749 (2010).
81. Z. Y. Abd Elmageed, A. S. Naura, Y. Errami, M. Zerfaoui, The Poly(ADP-ribose) polymerases (PARPs): New roles in intracellular transport. *Cell Signal* **24**, 1-8 (2012).
82. A. Sala *et al.*, The nucleosome-remodeling ATPase ISWI is regulated by poly-ADP-ribosylation. *PLoS Biol* **6**, e252 (2008).
83. S. Smith, I. Gariat, A. Schmitt, T. de Lange, Tankyrase, a poly(ADP-ribose) polymerase at human telomeres. *Science* **282**, 1484-1487 (1998).
84. D. A. Wacker *et al.*, The DNA binding and catalytic domains of poly(ADP-ribose) polymerase 1 cooperate in the regulation of chromatin structure and transcription. *Mol Cell Biol* **27**, 7475-7485 (2007).
85. M. Stilmann *et al.*, A nuclear poly(ADP-ribose)-dependent signalosome confers DNA damage-induced I κ B kinase activation. *Mol Cell* **36**, 365-378 (2009).
86. P. Chang, M. Coughlin, T. J. Mitchison, Tankyrase-1 polymerization of poly(ADP-ribose) is required for spindle structure and function. *Nat Cell Biol* **7**, 1133-1139 (2005).
87. P. Chang, M. Coughlin, T. J. Mitchison, Interaction between Poly(ADP-ribose) and NuMA contributes to mitotic spindle pole assembly. *Mol Biol Cell* **20**, 4575-4585 (2009).
88. P. Chang, M. K. Jacobson, T. J. Mitchison, Poly(ADP-ribose) is required for spindle assembly and structure. *Nature* **432**, 645-649 (2004).

89. E. Kotova, M. Jarnik, A. V. Tulin, Poly (ADP-ribose) polymerase 1 is required for protein localization to Cajal body. *PLoS Genet* **5**, e1000387 (2009).
90. M. Altmeyer *et al.*, Liquid demixing of intrinsically disordered proteins is seeded by poly(ADP-ribose). *Nature communications* **6**, 8088 (2015).
91. L. Aravind, The WWE domain: a common interaction module in protein ubiquitination and ADP ribosylation. *Trends Biochem Sci* **26**, 273-275 (2001).
92. T. Wang, C. M. Simbulan-Rosenthal, M. E. Smulson, P. B. Chock, D. C. Yang, Polyubiquitylation of PARP-1 through ubiquitin K48 is modulated by activated DNA, NAD⁺, and dipeptides. *J Cell Biochem* **104**, 318-328 (2008).
93. W. Chang, J. N. Dynek, S. Smith, TRF1 is degraded by ubiquitin-mediated proteolysis after release from telomeres. *Genes Dev* **17**, 1328-1333 (2003).
94. Z. Wang *et al.*, Recognition of the iso-ADP-ribose moiety in poly(ADP-ribose) by WWE domains suggests a general mechanism for poly(ADP-ribosyl)ation-dependent ubiquitination. *Genes Dev* **26**, 235-240 (2012).
95. H. C. Kang *et al.*, Iduna is a poly(ADP-ribose) (PAR)-dependent E3 ubiquitin ligase that regulates DNA damage. *Proc Natl Acad Sci U S A* **108**, 14103-14108 (2011).
96. Y. Zhang *et al.*, RNF146 is a poly(ADP-ribose)-directed E3 ligase that regulates axin degradation and Wnt signalling. *Nat Cell Biol* **13**, 623-629 (2011).
97. S. A. Andrabi *et al.*, Iduna protects the brain from glutamate excitotoxicity and stroke by interfering with poly(ADP-ribose) polymer-induced cell death. *Nat Med* **17**, 692-699 (2011).
98. S. M. Huang *et al.*, Tankyrase inhibition stabilizes axin and antagonizes Wnt signalling. *Nature* **461**, 614-620 (2009).
99. L. Kashima *et al.*, CHFR regulates the mitotic checkpoint by targeting PARP-1 for ubiquitination and degradation. *J Biol Chem*, (2012).
100. M. G. Bacalini *et al.*, Poly(ADP-ribosyl)ation affects stabilization of Che-1 protein in response to DNA damage. *DNA Repair (Amst)* **10**, 380-389 (2011).
101. S. A. Andrabi, T. M. Dawson, V. L. Dawson, Mitochondrial and nuclear cross talk in cell death: parthanatos. *Ann N Y Acad Sci* **1147**, 233-241 (2008).
102. X. Wang, C. Yang, J. Chai, Y. Shi, D. Xue, Mechanisms of AIF-mediated apoptotic DNA degradation in *Caenorhabditis elegans*. *Science* **298**, 1587-1592 (2002).
103. G. I. Karras *et al.*, The macro domain is an ADP-ribose binding module. *EMBO J* **24**, 1911-1920 (2005).
104. D. W. Koh *et al.*, Failure to degrade poly(ADP-ribose) causes increased sensitivity to cytotoxicity and early embryonic lethality. *Proc Natl Acad Sci U S A* **101**, 17699-17704 (2004).
105. O. Mortusewicz, E. Fouquerel, J. C. Ame, H. Leonhardt, V. Schreiber, PARG is recruited to DNA damage sites through poly(ADP-ribose)- and PCNA-dependent mechanisms. *Nucleic Acids Res* **39**, 5045-5056 (2011).
106. K. M. Frizzell *et al.*, Global analysis of transcriptional regulation by poly(ADP-ribose) polymerase-1 and poly(ADP-ribose) glycohydrolase in MCF-7 human breast cancer cells. *J Biol Chem* **284**, 33926-33938 (2009).

107. J. P. Gagne *et al.*, Proteome-wide identification of poly(ADP-ribose) binding proteins and poly(ADP-ribose)-associated protein complexes. *Nucleic Acids Res* **36**, 6959-6976 (2008).
108. J. M. Pleschke, H. E. Kleczkowska, M. Strohm, F. R. Althaus, Poly(ADP-ribose) binds to specific domains in DNA damage checkpoint proteins. *J Biol Chem* **275**, 40974-40980 (2000).
109. O. Huambachano, F. Herrera, A. Rancourt, M. S. Satoh, Double-stranded DNA binding domain of poly(ADP-ribose) polymerase-1 and molecular insight into the regulation of its activity. *J Biol Chem* **286**, 7149-7160 (2011).
110. M. Murawska, M. Hassler, R. Renkawitz-Pohl, A. Ladurner, A. Brehm, Stress-induced PARP activation mediates recruitment of Drosophila Mi-2 to promote heat shock gene expression. *PLoS Genet* **7**, e1002206 (2011).
111. S. Eustermann *et al.*, Solution structures of the two PBZ domains from human APLF and their interaction with poly(ADP-ribose). *Nat Struct Mol Biol* **17**, 241-243 (2010).
112. J. Oberoi *et al.*, Structural basis of poly(ADP-ribose) recognition by the multizinc binding domain of checkpoint with forkhead-associated and RING Domains (CHFR). *J Biol Chem* **285**, 39348-39358 (2010).
113. W. Han, X. Li, X. Fu, The macro domain protein family: structure, functions, and their potential therapeutic implications. *Mutat Res* **727**, 86-103 (2011).
114. G. Kustatscher, M. Hothorn, C. Pugieux, K. Scheffzek, A. G. Ladurner, Splicing regulates NAD metabolite binding to histone macroH2A. *Nat Struct Mol Biol* **12**, 624-625 (2005).
115. F. C. Peterson *et al.*, Orphan macrodomain protein (human C6orf130) is an O-acyl-ADP-ribose deacylase: solution structure and catalytic properties. *J Biol Chem* **286**, 35955-35965 (2011).
116. D. Chen *et al.*, Identification of macrodomain proteins as novel O-acetyl-ADP-ribose deacetylases. *J Biol Chem* **286**, 13261-13271 (2011).
117. P. V. Mehrotra *et al.*, DNA repair factor APLF is a histone chaperone. *Mol Cell* **41**, 46-55 (2011).
118. A. J. Gottschalk *et al.*, Poly(ADP-ribosyl)ation directs recruitment and activation of an ATP-dependent chromatin remodeler. *Proc Natl Acad Sci U S A* **106**, 13770-13774 (2009).
119. C. K. Anders *et al.*, Poly(ADP-Ribose) polymerase inhibition: "targeted" therapy for triple-negative breast cancer. *Clin Cancer Res* **16**, 4702-4710 (2010).
120. G. Papeo *et al.*, Poly(ADP-ribose) polymerase inhibition in cancer therapy: are we close to maturity? *Expert Opin Ther Pat* **19**, 1377-1400 (2009).
121. P. Pacher, C. Szabo, Role of poly(ADP-ribose) polymerase 1 (PARP-1) in cardiovascular diseases: the therapeutic potential of PARP inhibitors. *Cardiovasc Drug Rev* **25**, 235-260 (2007).
122. H. Shevalye *et al.*, Poly(ADP-ribose) polymerase (PARP) inhibition counteracts multiple manifestations of kidney disease in long-term streptozotocin-diabetic rat model. *Biochem Pharmacol* **79**, 1007-1014 (2010).
123. M. Masutani, H. Nakagama, T. Sugimura, Poly(ADP-ribosyl)ation in relation to cancer and autoimmune disease. *Cell Mol Life Sci* **62**, 769-783 (2005).

124. R. A. Mota *et al.*, Inhibition of poly(ADP-ribose) polymerase attenuates the severity of acute pancreatitis and associated lung injury. *Lab Invest* **85**, 1250-1262 (2005).
125. H. E. Bryant *et al.*, Specific killing of BRCA2-deficient tumours with inhibitors of poly(ADP-ribose) polymerase. *Nature* **434**, 913-917 (2005).
126. H. Farmer *et al.*, Targeting the DNA repair defect in BRCA mutant cells as a therapeutic strategy. *Nature* **434**, 917-921 (2005).
127. P. C. Fong *et al.*, Inhibition of poly(ADP-ribose) polymerase in tumors from BRCA mutation carriers. *N Engl J Med* **361**, 123-134 (2009).
128. K. A. Gelmon *et al.*, Olaparib in patients with recurrent high-grade serous or poorly differentiated ovarian carcinoma or triple-negative breast cancer: a phase 2, multicentre, open-label, non-randomised study. *Lancet Oncol* **12**, 852-861 (2011).
129. J. O'Shaughnessy *et al.*, Iniparib plus chemotherapy in metastatic triple-negative breast cancer. *N Engl J Med* **364**, 205-214 (2011).
130. M. Guha, PARP inhibitors stumble in breast cancer. *Nat Biotechnol* **29**, 373-374 (2011).
131. A. G. Patel, S. B. De Lorenzo, K. S. Flatten, G. G. Poirier, S. H. Kaufmann, Failure of iniparib to inhibit poly(ADP-Ribose) polymerase in vitro. *Clin Cancer Res* **18**, 1655-1662 (2012).
132. X. Liu *et al.*, Iniparib nonselectively modifies cysteine-containing proteins in tumor cells and is not a Bona Fide PARP inhibitor. *Clin Cancer Res* **18**, 510-523 (2012).
133. E. Wahlberg *et al.*, Family-wide chemical profiling and structural analysis of PARP and tankyrase inhibitors. *Nat Biotechnol* **30**, 283-288 (2012).
134. M. Narwal, H. Venkannagari, L. Lehtio, Structural Basis of Selective Inhibition of Human Tankyrases. *J Med Chem*, (2012).
135. G. J. Southan, C. Szabo, Poly(ADP-ribose) polymerase inhibitors. *Curr Med Chem* **10**, 321-340 (2003).
136. K. M. Specht, K. M. Shokat, The emerging power of chemical genetics. *Current opinion in cell biology* **14**, 155-159 (2002).
137. B. A. Gibson, W. L. Kraus, New insights into the molecular and cellular functions of poly(ADP-ribose) and PARPs. *Nat Rev Mol Cell Biol* **13**, 411-424 (2012).
138. S. Vyas *et al.*, Family-wide analysis of poly(ADP-ribose) polymerase activity. *Nature communications* **5**, 4426 (2014).
139. M. O. Hottiger, Nuclear ADP-Ribosylation and Its Role in Chromatin Plasticity, Cell Differentiation, and Epigenetics. *Annual review of biochemistry* **84**, 227-263 (2015).
140. J. P. Gagne *et al.*, Quantitative proteomics profiling of the poly(ADP-ribose)-related response to genotoxic stress. *Nucleic Acids Res* **40**, 7788-7805 (2012).
141. S. Jungmichel *et al.*, Proteome-wide identification of poly(ADP-Ribosyl)ation targets in different genotoxic stress responses. *Mol Cell* **52**, 272-285 (2013).
142. Y. Zhang, J. Wang, M. Ding, Y. Yu, Site-specific characterization of the Asp- and Glu-ADP-ribosylated proteome. *Nature methods* **10**, 981-984 (2013).
143. K. L. Feijs *et al.*, ARTD10 substrate identification on protein microarrays: regulation of GSK3beta by mono-ADP-ribosylation. *Cell communication and signaling : CCS* **11**, 5 (2013).

144. S. Troiani *et al.*, Identification of candidate substrates for poly(ADP-ribose) polymerase-2 (PARP2) in the absence of DNA damage using high-density protein microarrays. *FEBS J* **278**, 3676-3687 (2011).
145. I. Carter-O'Connell, H. Jin, R. K. Morgan, L. L. David, M. S. Cohen, Engineering the substrate specificity of ADP-ribosyltransferases for identifying direct protein targets. *Journal of the American Chemical Society* **136**, 5201-5204 (2014).
146. H. Jiang, J. H. Kim, K. M. Frizzell, W. L. Kraus, H. Lin, Clickable NAD analogues for labeling substrate proteins of poly(ADP-ribose) polymerases. *Journal of the American Chemical Society* **132**, 9363-9372 (2010).
147. K. Adelman, J. T. Lis, Promoter-proximal pausing of RNA polymerase II: emerging roles in metazoans. *Nature reviews. Genetics* **13**, 720-731 (2012).
148. K. Fujinaga *et al.*, Dynamics of human immunodeficiency virus transcription: P-TEFb phosphorylates RD and dissociates negative effectors from the transactivation response element. *Mol Cell Biol* **24**, 787-795 (2004).
149. H. Kwak, J. T. Lis, Control of transcriptional elongation. *Annual review of genetics* **47**, 483-508 (2013).
150. W. Yu *et al.*, Poly(ADP-ribosyl)ation regulates CTCF-dependent chromatin insulation. *Nature genetics* **36**, 1105-1110 (2004).
151. N. J. O'Neil, D. M. van Pel, P. Hieter, Synthetic lethality and cancer: cohesin and PARP at the replication fork. *Trends in genetics : TIG* **29**, 290-297 (2013).
152. M. L. Bailey *et al.*, Glioblastoma cells containing mutations in the cohesin component STAG2 are sensitive to PARP inhibition. *Molecular cancer therapeutics* **13**, 724-732 (2014).
153. A. Tulin, A. Spradling, Chromatin loosening by poly(ADP)-ribose polymerase (PARP) at *Drosophila* puff loci. *Science* **299**, 560-562 (2003).
154. Y. Yamaguchi, N. Inukai, T. Narita, T. Wada, H. Handa, Evidence that negative elongation factor represses transcription elongation through binding to a DRB sensitivity-inducing factor/RNA polymerase II complex and RNA. *Mol Cell Biol* **22**, 2918-2927 (2002).
155. K. W. Ryu, D. S. Kim, W. L. Kraus, New facets in the regulation of gene expression by ADP-ribosylation and poly(ADP-ribose) polymerases. *Chemical reviews* **115**, 2453-2481 (2015).
156. C. T. Ong, K. Van Bortle, E. Ramos, V. G. Corces, Poly(ADP-ribosyl)ation regulates insulator function and intrachromosomal interactions in *Drosophila*. *Cell* **155**, 148-159 (2013).
157. M. D. Robinson, D. J. McCarthy, G. K. Smyth, edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139-140 (2010).
158. M. Kininis, G. D. Isaacs, L. J. Core, N. Hah, W. L. Kraus, Postrecruitment regulation of RNA polymerase II directs rapid signaling responses at the promoters of estrogen target genes. *Mol Cell Biol* **29**, 1123-1133 (2009).
159. S. P. Colowick, N. O. Kaplan, M. M. Ciotti, The reaction of pyridine nucleotide with cyanide and its analytical use. *J Biol Chem* **191**, 447-459 (1951).

160. A. Kumar, R. F. Colman, 8-(4-Bromo-2,3-dioxobutylthio)NAD: a new affinity label for NAD-specific isocitrate dehydrogenase. *Archives of biochemistry and biophysics* **308**, 357-366 (1994).
161. G. Liang, J. He, Y. Zhang, Kdm2b promotes induced pluripotent stem cell generation by facilitating gene activation early in reprogramming. *Nat Cell Biol* **14**, 457-466 (2012).
162. D. Wessel, U. I. Flugge, A method for the quantitative recovery of protein in dilute solution in the presence of detergents and lipids. *Anal Biochem* **138**, 141-143 (1984).
163. J. D. Dignam, R. M. Lebovitz, R. G. Roeder, Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated mammalian nuclei. *Nucleic Acids Res* **11**, 1475-1489 (1983).
164. D. W. Huang *et al.*, The DAVID Gene Functional Classification Tool: a novel biological module-centric algorithm to functionally analyze large gene lists. *Genome biology* **8**, R183 (2007).
165. D. W. Huang *et al.*, DAVID Bioinformatics Resources: expanded annotation database and novel algorithms to better extract biology from large gene lists. *Nucleic Acids Res* **35**, W169-175 (2007).
166. P. V. Hornbeck *et al.*, PhosphoSitePlus: a comprehensive resource for investigating the structure and function of experimentally determined post-translational modifications in man and mouse. *Nucleic Acids Res* **40**, D261-270 (2012).
167. I. A. Hendriks *et al.*, Uncovering global SUMOylation signaling networks in a site-specific manner. *Nat Struct Mol Biol* **21**, 927-936 (2014).
168. M. F. Chou, D. Schwartz, Biological sequence motif discovery using motif-x. *Current protocols in bioinformatics / editorial board, Andreas D. Baxevanis ... [et al.]* **Chapter 13**, Unit 13 15-24 (2011).
169. D. Schwartz, S. P. Gygi, An iterative statistical approach to the identification of protein phosphorylation motifs from large-scale data sets. *Nat Biotechnol* **23**, 1391-1398 (2005).
170. N. Hah *et al.*, A rapid, extensive, and transient transcriptional response to estrogen signaling in breast cancer cells. *Cell* **145**, 622-634 (2011).
171. X. Luo, M. Chae, R. Krishnakumar, C. G. Danko, W. L. Kraus, Dynamic reorganization of the AC16 cardiomyocyte transcriptome in response to TNFalpha signaling revealed by integrated genomic analyses. *BMC Genomics* **15**, 155 (2014).
172. B. Langmead, C. Trapnell, M. Pop, S. L. Salzberg, Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome biology* **10**, R25 (2009).
173. A. R. Quinlan, I. M. Hall, BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841-842 (2010).
174. J. T. Robinson *et al.*, Integrative genomics viewer. *Nat Biotechnol* **29**, 24-26 (2011).
175. H. Thorvaldsdottir, J. T. Robinson, J. P. Mesirov, Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Briefings in bioinformatics* **14**, 178-192 (2013).

176. Y. Shen *et al.*, A map of the cis-regulatory sequences in the mouse genome. *Nature* **488**, 116-120 (2012).
177. J. Sun *et al.*, Genetic and genomic analyses of RNA polymerase II-pausing factor in regulation of mammalian transcription and cell growth. *J Biol Chem* **286**, 36248-36257 (2011).
178. C. Zang *et al.*, A clustering approach for identification of enriched domains from histone modification ChIP-Seq data. *Bioinformatics* **25**, 1952-1958 (2009).
179. A. J. Saldanha, Java Treeview--extensible visualization of microarray data. *Bioinformatics* **20**, 3246-3248 (2004).
180. R. Krishnakumar *et al.*, Reciprocal binding of PARP-1 and histone H1 at promoters specifies transcriptional outcomes. *Science* **319**, 819-821 (2008).
181. L. J. Core, J. J. Waterfall, J. T. Lis, Nascent RNA sequencing reveals widespread pausing and divergent initiation at human promoters. *Science* **322**, 1845-1848 (2008).
182. M. Martin, Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnetjournal* **17**, (2011).
183. C. G. Danko *et al.*, Signaling pathways differentially affect RNA polymerase II initiation, pausing, and elongation rate in cells. *Mol Cell* **50**, 212-222 (2013).
184. R. D. Kornberg, Chromatin structure: a repeating unit of histones and DNA. *Science* **184**, 868-871 (1974).
185. K. Luger, A. W. Mader, R. K. Richmond, D. F. Sargent, T. J. Richmond, Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature* **389**, 251-260 (1997).
186. A. Valouev *et al.*, Determinants of nucleosome organization in primary human cells. *Nature* **474**, 516-520 (2011).
187. E. Segal *et al.*, A genomic code for nucleosome positioning. *Nature* **442**, 772-778 (2006).
188. K. Struhl, E. Segal, Determinants of nucleosome positioning. *Nat Struct Mol Biol* **20**, 267-273 (2013).
189. M. J. Pazin, P. Bhargava, E. P. Geiduschek, J. T. Kadonaga, Nucleosome mobility and the maintenance of nucleosome positioning. *Science* **276**, 809-812 (1997).
190. Z. Zhang *et al.*, A packing mechanism for nucleosome organization reconstituted across a eukaryotic genome. *Science* **332**, 977-980 (2011).
191. I. Maze, K. M. Noh, A. A. Soshnev, C. D. Allis, Every amino acid matters: essential contributions of histone variants to mammalian development and disease. *Nature reviews. Genetics* **15**, 259-271 (2014).
192. A. D. Goldberg, C. D. Allis, E. Bernstein, Epigenetics: a landscape takes shape. *Cell* **128**, 635-638 (2007).
193. J. Y. Fan, F. Gordon, K. Luger, J. C. Hansen, D. J. Tremethick, The essential histone variant H2A.Z regulates the equilibrium between different chromatin conformational states. *Nature structural biology* **9**, 172-176 (2002).
194. M. Vermeulen *et al.*, Selective anchoring of TFIID to nucleosomes by trimethylation of histone H3 lysine 4. *Cell* **131**, 58-69 (2007).
195. J. Zhou, C. Chau, Z. Deng, W. Stedman, P. M. Lieberman, Epigenetic control of replication origins. *Cell cycle* **4**, 889-892 (2005).

196. E. R. Foster, J. A. Downs, Histone H2A phosphorylation in DNA double-strand break repair. *FEBS J* **272**, 3231-3240 (2005).
197. Z. Zhang, B. F. Pugh, High-resolution genome-wide mapping of the primary structure of chromatin. *Cell* **144**, 175-186 (2011).
198. R. T. Koerber, H. S. Rhee, C. Jiang, B. F. Pugh, Interaction of transcriptional regulators with specific nucleosomes across the *Saccharomyces* genome. *Mol Cell* **35**, 889-902 (2009).
199. N. J. Curtin, C. Szabo, Therapeutic applications of PARP inhibitors: anticancer therapy and beyond. *Molecular aspects of medicine* **34**, 1217-1256 (2013).
200. A. Barski *et al.*, High-resolution profiling of histone methylations in the human genome. *Cell* **129**, 823-837 (2007).
201. C. A. Meyer, X. S. Liu, Identifying and mitigating bias in next-generation sequencing methods for chromatin biology. *Nature reviews. Genetics* **15**, 709-721 (2014).
202. Z. Li *et al.*, Foxa2 and H2A.Z mediate nucleosome depletion during embryonic stem cell differentiation. *Cell* **151**, 1608-1616 (2012).
203. I. Albert *et al.*, Translational and rotational settings of H2A.Z nucleosomes across the *Saccharomyces cerevisiae* genome. *Nature* **446**, 572-576 (2007).
204. B. E. Bernstein *et al.*, A bivalent chromatin structure marks key developmental genes in embryonic stem cells. *Cell* **125**, 315-326 (2006).
205. D. E. Schones *et al.*, Dynamic regulation of nucleosome positioning in the human genome. *Cell* **132**, 887-898 (2008).
206. R. Schopflin *et al.*, Modeling nucleosome position distributions from experimental nucleosome positioning maps. *Bioinformatics* **29**, 2380-2386 (2013).
207. J. G. Henikoff, J. A. Belsky, K. Krassovsky, D. M. MacAlpine, S. Henikoff, Epigenome characterization at single base-pair resolution. *Proc Natl Acad Sci U S A* **108**, 18318-18323 (2011).
208. C. Jin *et al.*, H3.3/H2A.Z double variant-containing nucleosomes mark 'nucleosome-free regions' of active promoters and other regulatory regions. *Nature genetics* **41**, 941-945 (2009).
209. Y. Fu, M. Sinha, C. L. Peterson, Z. Weng, The insulator binding protein CTCF positions 20 nucleosomes around its binding sites across the human genome. *PLoS Genet* **4**, e1000138 (2008).
210. A. H. Forst *et al.*, Recognition of mono-ADP-ribosylated ARTD10 substrates by ARTD8 macrodomains. *Structure* **21**, 462-475 (2013).
211. L. J. Harris, S. B. Larson, K. W. Hasel, A. McPherson, Refined structure of an intact IgG2a monoclonal antibody. *Biochemistry* **36**, 1581-1597 (1997).