Molecular Basis for Ubiquitin and SUMO Recognition by RAP80, a DNA repair protein

by

Anamika Singh

A thesis submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Department of Biochemistry
University of Alberta

© Anamika Singh, 2017

Abstract

The information contained in our genome is essential for the proper functioning of different physiological processes in cells that are required for our growth and survival, making the maintenance of genome integrity a task of utmost priority. Our cells have evolved a highly elaborate, complex and well regulated process of DNA repair that accompanies cell cycle regulation, chromatin remodeling and transcription of repair proteins, collectively known as, DNA damage response. Amidst different kinds of DNA damage, double strand breaks pose a serious threat to the cell survival, however most of the time these breaks are repaired by Non Homologous End Joining (NHEJ) or Homologous Recombination (HR). HR involves the end resection of broken DNA strands followed by a homology search, DNA synthesis and the ligation of the broken strand. The end resection of the broken DNA strand is a special feature of the HR process, which is precisely controlled by many factors and performed by proteins such as CtIP and the MRN complex. RAP80, a DNA repair protein was initially believed to promote HR over NHEJ by recruiting proteins to DNA damage sites that favor end resection, such as BRCA1 that forms a multi protein complex with Abraxas, BRCC36, BARD1 and RAP80, known as BRCA1 A complex. BRCA1, a major player of the DNA repair pathway is a crucial tumor suppressor protein whose mutations show strong links to breast and ovarian cancer. Recent advances suggest that although RAP80 recruits BRCA1 to the DNA damage sites, eventually it represses HR by sequestering BRCA1 within the complex and restricting its association with other complexes that endorse end resection. This decisive role played by RAP80 in the repair pathway choice is due to its ability to recognize Lys-63 linked ubiquitin chains and SUMO moieties through its ubiquitin and SUMO interacting motifs present at the N-terminal region. A single deletion mutation in the

ubiquitin interacting motif of RAP80 has been linked with cases of familial breast cancer. Additionally, a mutation in the SUMO interacting motif of RAP80 has been linked to somatic mutations in cancer cells, highlighting the importance of ubiquitin and SUMO recognition in genome protection. We employed NMR spectroscopy, binding studies, stability studies and MD simulations to understand the molecular basis of ubiquitin and SUMO recognition of RAP80. We demonstrate that the ΔE81 mutation in the first ubiquitin interacting motif of RAP80 leads to a structural frameshift of the helix with causing modest changes in the stability. The prominent reason for the impaired binding of this mutant with ubiquitin is its loss of favourable electrostatic interaction, resulting in a loss of multivalent binding advantage.

We also determine the binding affinity of the SUMO interacting motif of RAP80 with SUMO-2 using NMR methods and show that the phosphorylation of RAP80 by casein kinase 2 enhances its affinity for SUMO-2 by ~25 fold due to increased electrostatic interactions between the phosphorylated serine residues and the basic loop of SUMO-2 present at the binding site. Understanding these binding interactions from a structural and thermodynamic point of view is fundamental in order to develop a holistic understanding of the ubiquitin and SUMO recognition process. Our research has provided a closer look at the critical interactions in ubiquitin and SUMO signaling in DNA repair and has opened up new avenues for future research in DNA repair field.

Preface

This thesis is organized in four chapters and two appendices. A part of the first chapter that forms the introduction of the thesis has been published before as a review article and the following two chapters have also been published before. Chapter 4 that concludes and throws light on the future directions for the research is original work. Both the appendices are original work and have not been published elsewhere.

Some parts of Chapter 1 are taken from the published review article. B.L. Lee, **Anamika**, J.N.M. Glover, M.J. Hendzel and Leo Spyracopouls. J. Mol. Biol.(2017). Molecular Basis for K-63 linked ubiquitination processes in double-strand DNA break repair : A focus on kinetics and dynamics. doi:10.1016/j.jmb.2017.05.029. I wrote manuscript for the introduction of the ubiquitination and SUMOylation in the DNA double strand break repair.

Chapter 2 has been published before as - **Anamika**, Craig J. Markin, Manoj K. Rout and Leo Spyracopoulos. (2014) Molecular Basis for Impaired DNA Damage Response Function Associated with the RAP80 ΔE81 Defect. *J. Biol. Chem*. 28:12852-12862. I purified the proteins with CJM, collected NMR data and performed the analysis with CJM and LS. I also contributed to the figure making and manuscript editing.

Chapter 3 has been published before as - **Anamika** and Leo Spyracopoulos. (2016) Molecular Basis for Phosphorylation-dependent SUMO Recognition by the DNA Repair Protein RAP80 *J. Biol. Chem*. **291**, 4417-4428. L. S. and I designed the study and wrote the paper. I labeled, expressed, and purified RAP80 and SUMO-2, performed NMR experiments, and analyzed NMR data with assistance from L. S. MD simulations with NMR restraints, and NMR lineshape analyses were conducted and analyzed by L. S. and myself. Aromatic HSQC NMR pulse sequences were written by L. S., set up and analyzed by me. All authors analyzed the results and approved the final version of the manuscript.

Chapter 4 is the overall discussion of the research carried out in Chapter 2 and 3 and forms the ground for the appendices.

Appendix A and B are original work and not published anywhere.

Dedicated to my father

You gave us wings to fly, Papa

And, you mean the world to us.

Acknowledgements

A span of six years pursuing my PhD has been the most meaningful experience of my life, so far. The people and experiences that made this journey worthwhile can never be thanked enough, yet here is a little attempt of mine.

First and foremost, I am grateful to the Department of Biochemistry at the University of Alberta for giving me an opportunity to pursue my higher education in a refreshingly beautiful country like Canada. I thank my supervisor Dr. Leo Spyracopoulos for his guidance, patience and excellent supervision during the entire period of my graduate studies. Thank you for teaching me the basic details of NMR and protein purification, all that mixed with a good dose of humor that definitely made my stay in the lab more pleasant and fun. I am also very happy and grateful to you for making me independent at every stage, polishing my writing and presentation skills, encouraging me to form collaborations and providing an environment conducive for my intellectual growth.

I would like to thank my supervisory committee members Dr. Brian Sykes and Dr. Mark Glover for their inputs, critical analysis of data, presentations, courses and assistance with the award applications. My initial years in the laboratory was mostly spent in learning the tools and techniques, and it would not have been possible without the assistance from senior lab members like Drs. Craig Markin, Manoj Rout and Brian Lee. Thank you Craig for helping me with almost everything from NMR to Mathematica, Sparky, protein purification and making the settling in lab easy and comfortable. Special thanks for teaching me how to play golf. I am thankful to Manoj for encouraging me to learn Illustrator, CARA and other softwares throughout my PhD program, that has saved me a lot of time and energy. Also, thank you Manoj for keeping me updated with the Indian politics, funny viral videos and always giving me the best advice. Thank you Brian for clearing all my doubts regarding NMR concepts, scripts and general science. You all have helped immensely in shaping my research career and I will always be grateful for that. Most part of my last year in the PhD program was spent in Dr. Michael Hendzel lab, with whom

Contents

List of Tables

List of Figures

Abbreviations

AIR- Abraxas interacting region

AMBER- Assisted model building with energy refinement

ATM- Ataxia-telangiectasia mutated

BACH1- BTB domain and CNC homolog 1

BARD1- BRCA1-associated RING domain protein 1

BLM- Bloom syndrome protein

BMI1- B lymphoma Mo-MLV insertion region 1 homolog

BMRB- Bio Magnetic Resonance Bank

BRCA1- Breast Cancer Type1 Susceptibility Protein

BRCC36- BRCA1 Containing Complex

CARA- Computer aided resonance assignment

CBX4- Chromobox 4

CHD3- Chromodomain helicase DNA binding protein 3

CHFR- Checkpoint with forkhead and ring finger domains

CK2- casein kinase 2

COSMIC- Catalogue of somatic mutations in cancer

CtiP- CtBP-interacting protein, CtBP- C terminal-binding protein

DAXX- Death domain-associated protein

DDR- DNA damage response

DMEM- Dulbecco's modified eagle's medium

DNA- Deoxyribonucleic acid

DSB- Double strand break

DSS-2,2-dimethyl-2-silapentane-5-sulfonate

DUB- Deubiquitinase

EXO1- Exonuclease 1

FBS- Fetal bovine serum

FHA- Forkhead-associated domain

FRAP- Fluorescence recovery after photobleaching

Gy- Gray

GST- Glutathione S-transferase

HECT- Homologous to the E6-AP carboxyl terminus

HERC2- HECT and RLD domain containing E3 ubiquitin protein ligase 2

HR- Homologous Recombination

HSQC- Heteronuclear Single Quantum Coherence Spectroscopy

IF- Immunofluorescence staining

IPTG- Isopropyl β –D-1-thiogalactopyranoside

KAP1- KRAP-associated protein 1

L3MBTL1- Lethal (3) malignant brain tumor-like protein 1

LRM- Ligand recognition motif

MALDI-TOF- Matrix Assisted Laser Desorption/Ionization – Time of Flight

MD- molecular dynamics

MDC1- Mediator of DNA damage checkpoint protein 1

MICS- Motif identification from chemical shifts

MIU- Motif interacting with ubiquitin

MRN- Mre11-Rad50-Nbs1

NBS1- Nijmegen breakage syndrome 1

NHEJ- Non homologous end joining

NMR- Nuclear magnetic resonance

NOE- Nuclear overhauser effect

NOESY- Nuclear overhauser effect spectroscopy

NZF- Np14 zinc finger

OTUB1- OTU domain-containing ubiquitin aldehyde-binding protein 1

PALB2- Partner and localizer of BRCA2

PAR- polyADP ribose

PARP- polyADP ribose polymerase

PBS- Phosphate buffer saline

PBZ- PAR-binding zinc finger

PCR- Polymerase chain reaction

PIAS- Protein inhibitor of activated STAT

PML- Promyelocytic leukaemia

polyUb- polyubiquitin

RAP80- Receptor associated protein 80

RIPA- Radioimmunoprecipitation assay

RNF- Ring finger protein

RPA- Replication protein A

SDS-PAGE- Sodium dodecyl suplhate-polyacrylamide gel

SIM-SUMO interacting motif

SENP- Sentrin specific protease

ssDNA- single strand DNA

STUBL- SUMO-targeting ubiquitin ligase

SUMO- small ubiquitin like modifier

TALOS- Torsion Angle Likeliness Obtained from Sequence and Shift Similarity

TEV-tobacco etch virus

TOCSY- Total correlation spectroscopy

TOPBP1- DNA topoisomerase 2-binding protein 1

TRAIP- TRAF interacting protein

TROSY- Transverse relaxation optimized spectroscopy

tUIM- tandem ubiquitin-interacting motif

Ub- ubiquitin

UBA- Ubiquitin associated domain

UBD- Ubiquitin binding domain

UIM- Ubiquitin interacting motif

USP- Ubiquitin specific protease

Vcp- Valosin containing protein

XRCC- X-ray repair cross-complementing protein

ZNF- Zinc finger

53BP1- p53 binding protein 1

Chapter 1

Introduction

DNA Damage and Repair: A Focus on DNA Double Strand Breaks

DNA contains the genetic information required for the survival, growth and reproduction of living organisms[1]. It serves as the blueprint of life. DNA is composed of four nucleotides, linked by phosphodiester bonds to form a polynucleotide chain; these double stranded polynucleotide chains are structurally organized in to basic units called nucleosomes, wherein 147 base pairs of DNA are wrapped around histone core that is composed of two copies of each histone, H2A, H2B, H3 and H4 and appear as beads on a string in electron micrographs[2]. The histone H1 binds the linker DNA connecting two nucleosome units and exists in one copy per nucleosome, unlike other core histones[2]. Multiple layers of extensive folding and packing of these basic units result in the condensation of ~3 billion base pairs of DNA that can be linearly stretched to 2 meters in length, into a nucleus of size 3 μm[4]. The condensed state of chromatin varies during a cell cycle to facilitate chromosomal segregation and cell division, wherein the interphase chromosome is in the most decondensed state and the metaphase stage represents the highly condensed state[5]. Since the DNA functions as the information center of the cell, the maintenance of genomic integrity is a well regulated process as any damage to DNA can cause changes or loss of genetic information that can give rise to various diseases and abnormalities, cancer being one of them[1]. On a daily basis our genome is endlessly challenged by various external and internal factors such as radiation, various chemicals and free radicals generated during metabolism that has the potential to hamper the physical integrity and alter the chemical composition of DNA causing DNA base mismatch, single strand DNA breaks and double strand DNA breaks[1]. Double strand DNA breaks pose a major threat to cells and it can lead to a loss of genetic information or initiate its rearrangement leading to chromosomal translocation which is an underlying cause for many diseases [1,7]. Double strand DNA breaks are repaired by one of the two mechanisms– Non Homologous End Joining (NHEJ) and Homologous Repair (HR), depending on the stage of the cell cycle. Both processes are accompanied with cell cycle checkpoint activation and transcription of repair proteins, collectively known as DNA damage response (DDR)[7,8]. NHEJ is active during most of the cell cycle and involves the direct ligation of the broken strands without being guided by the homologous

sequence of the sister chromatid which makes this process more likely to incorporate error in the repaired DNA strand[9]. In contrast to NHEJ, HR is active only in S and G2 phases of cell cycle and involves a search for a homologous region in the sister chromatid near the break site which renders it less prone to errors and assures the faithful restoration of genetic information[10]. Collectively, both repair pathways ensure that the genomic integrity of the cell is maintained. The choice between the two pathways is a crucial task and is well-regulated.

DNA Double Strand Break Repair by Homologous Recombination: A Brief Overview

DNA double strand break repair by homologous recombination is a highly elaborate and complex process that involves changes at multiple levels. At the molecular level, changes in nucleotides such as nucleotide excision and synthesis is accompanied with alterations at the chromatin level such as chromatin organization causing its decondensation for easy access of the enzymatic machinery and repair factors to the damage sites[11,8]. At a cellular level, it influences the nuclear architecture by changing the number, size and morphology of the nuclear bodies and regulates the cell cycle progression and apoptosis[12]. In order to accomplish these tasks efficiently, different proteins accumulate at the double strand break sites spatiotemporally, guided by different post translational modifications that lead to the sequential progression of events resulting in the repair of broken strand[13,8,11,14]. The broken DNA strands induce chromatin relaxation that is sensed by an unknown protein followed by concomitant or sequential sensing of DNA breaks by the MRN complex and activation of ataxia telangiectasia mutated (ATM) kinase[8]. ATM phosphorylates histone variant H2AX at serine 139, and the resulting phosphorylated histone (γ-H2AX), acts as a recruitment scaffold for downstream repair proteins, forming a foci like structure that can be visualized by fluorescence microscopy[8,11,15]. Hundreds to thousands copies of γ-H2AX span mega base pairs of DNA at one repair focus that also accommodates multiple copies of different DNA repair factors[15]. Immunofluorescence staining and detection of γ-H2AX

foci by fluorescence microscopy has dramatically expanded the understanding of recruitment and involvement of different proteins near damage sites[16]. MDC1 (mediator of DNA damage checkpoint 1) binds γ-H2AX and is phosphorylated by ATM kinase that acts as a binding platform for RNF8, a ubiquitin E3 Ub ligase[20,22]. A number of different Ub ligases, discussed in details below, lead to the conjugation of monoubiquitin and Ub chains of different topologies at damage sites, the most relevant one being Lys-63 linked Ub chains that enables the recruitment of key downstream repair factors[13,20]. These chains recruit repair proteins such as BRCA1, BRCA2, RAD51, and endonucleases such as EXO1 and CtIP, and play important roles in repair pathway choice between NHEJ and HR, chromatin reorganization, and repair pathway regulation. Mechanistically, repair of DSBs by HR involves extensive resection of DNA ends by nucleases such as EXO1, DNA2, CtIP, and Mre11 of the MRN complex, resulting in the formation of ssDNA, which then binds replication protein A (RPA). In order to perform a homology search within the sister chromatid, RPA is exchanged with RAD51 by the assistance of BRCA2, and the resulting RAD51-ssDNA strand undergoes a homology search, followed by DNA synthesis, ligation, and conclusion of repair, thereby regaining the integrity of the broken chromosome[20]. Figure 1.1 displays the sequential recruitment of important players of HR.
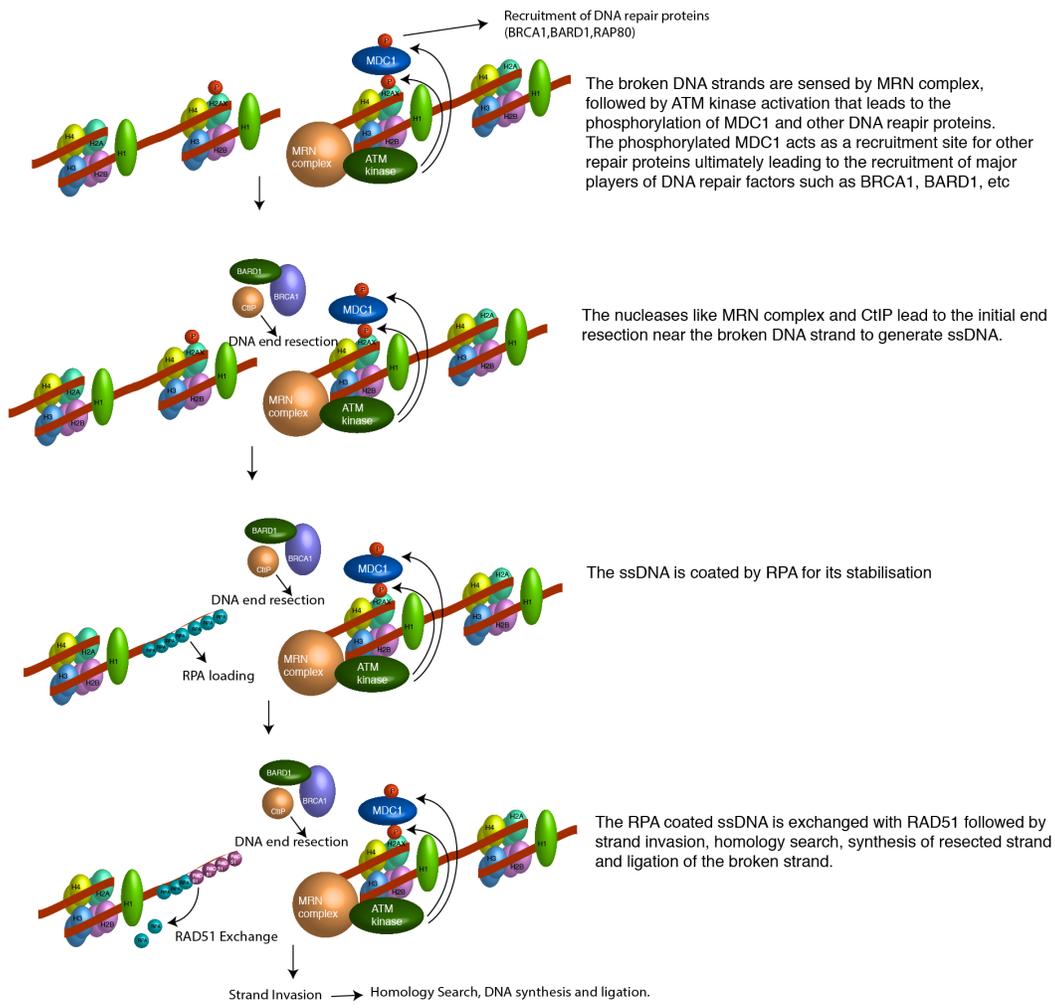
Figure 1.1: A brief overview of the DNA repair process by homologous recombination.

The progression of above mentioned steps in the DNA repair by homologous recombination that involves the recruitment of multiple repair factors to the damage sites is enabled by a multitude of post-translational modifications such as ubiquitination, SUMOylation, phosphorylation, acetylation and PARylation[8,14]. Amidst different kinds of post-translational modifications that play important roles in the seamless repair of double strand breaks by HR, the work presented here aims at understanding the molecular basis for the recognition of ubiquitin and SUMO modifications.

Ubiquitin - A Post Translational Modifier

The post-translational modification of proteins by ubiquitin is essential for the various tasks associated with the repair process such as timely recruitment, removal and retention of crucial proteins at the DSB sites, chromatin reorganization, the choice of repair pathway and ultimately the process of end resection and repair. Ubiquitin (Ub) is an 8.5 kDa, 76 residues globular protein, composed of a five-stranded β-sheet, an α-helix, a 3_{10} helix and a flexible C-terminal tail[21–23]. It can be conjugated to other proteins through an amide bond between its C-terminal glycine, and an acceptor lysine on substrate. Ub possesses seven surface lysine residues K6, K11, K27, K29, K33, K48 and K63, which can be enzymatically linked to the C-terminus of a sequential Ub and result in Ub chains of distinct links and topologies[24]. The conjugation of Ub to any protein equips it with an additional surface for interaction with other proteins possessing at least one Ub binding domain[23]. Ub conjugation to protein substrates is carried out by the sequential action of three enzymes. Initially, the Ub C terminal glycine is covalently attached to cysteine within an E1 enzyme through a thioester bond, followed by interaction of the E1~Ub complex with an E2 Ub conjugating enzyme, and subsequent Ub transfer to the E2 active site cysteine. E3 Ub ligases bind E2~Ub complexes and substrate proteins, leading to the formation of an amide bond between the ε-amino group of a substrate lysine, and the C-terminus of Ub[23,24]. Figure 1.2 is a schematic diagram of the ubiquitin conjugation mechanism. The ubiquitination is a reversible process where the deconjugation of Ub or Ub conjugates is carried out by a set of deubiquitinase enzymes[24,25]. Ub can itself be conjugated with one of the seven lysine residues on the

surface of another Ub and form Ub chains of different linkages and topologies which act as signals in various cellular pathways and broadens the repertoire of ubiquitin signaling[26]. K48-linked Ub chains typically signal for the proteasomal degradation of substrate, whereas K63-linked chains serve as a signaling platform for various pathways, one of them being DSB repair[27,28]. Chains of K6, K27 and K33 linkages have been recently reported to take part in the DNA repair process[29–31].
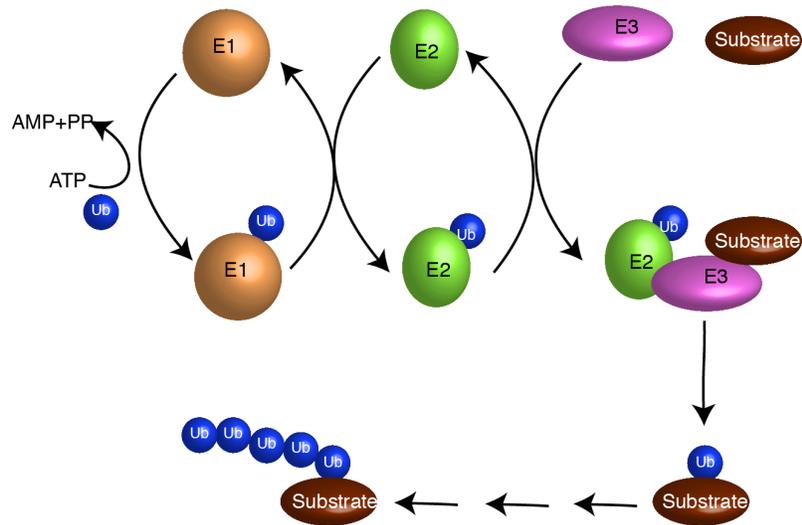
Figure 1.2: The Ubiquitin Enzyme Cascade- Ub activating enzyme, E1 is attached with an Ub to form E1~Ub complex, in the presence of ATP which then interacts with Ub conjugating enzyme, E2 and forms E2-Ub complex, that along with the enzyme Ub ligase E3 that has substrate specificity transfers the Ub to substrate. This enzymatic cascade is repeated multiple times to form Ub chains on the substrate.

Key Ubiquitination Events in the DSB Repair by Homologous Recombination

The ubiquitin E3 ligases responsible for different kinds of Ub modifications at the DSB sites are mentioned below:

RNF20-RNF40: The histone H2B monoubiquitination at K120 by this heterodimeric E3 Ub ligase regulates chromatin compaction; a reorganization step that is required for transcription elongation[32–35]. Histone H2B modification also occurs in response to DNA double strand breaks, in an ATM dependent manner with RNF20-RNF40, resulting in the accumulation of early repair proteins such as XRCC, RAD51 and BRCA1, for both NHEJ and HR repair pathways[32,35].

RNF2-BMI1: This heteromeric E3 Ub ligase belongs to the polycomb group proteins, an important group of epigenetic regulators. RNF2-BMI1 depleted cells show an impaired DNA damage response, and increased sensitivity to radiation[36]. RNF2-BMI1 is responsible for the monoubiquitination of the core histone H2A at K119/K120, to enable gene silencing, as well as to initiate DNA repair by facilitating recruitment of the downstream repair proteins RAP80, 53BP1, and BRCA1 to damage sites[37,38].

RNF8 and RNF168: The E3 Ub ligase RNF8 is recruited to damage sites through binding of phosphorylated MDC1 via its FHA domain[39–41]. RNF8 in conjunction with the heterodimeric E2/UEV complex Ubc13-Mms2, synthesizes K63 chains on linker histone H1[42]. A second E3 ligase, RNF168, then binds these K63 chains through an N-terminal Ub binding domain (UBD1), which leads to monoubiquitination of the core histone H2A at K13/K15, in conjunction with the E2 UbcH5[43–47]. It is also possible that RNF168, with Ubc13, can synthesize K63 chains at damage sites. Subsequently, K63 chain recognition by RAP80 in complex with BRCA1 and BRCC36 is obligatory for HR-mediated DNA repair[48–50]. The monoubiquitination of histone H2A at K13/15, and dimethylation of K20 at histone H4 are recognized by the Ub-dependent recruitment motif (UDR) and Tudor domain of 53BP1, respectively, resulting in initiation of

NHEJ[51–53]. In its constitutive state, histone H4 with dimethylated K20 is bound to the polycomb group protein L3MBTL1 and demethylases. These components are likely removed through attachment of K48 Ub chains by the E3/E2 pair RNF8/UbcH8, followed by removal of L3MBTL1 by the AAA-ATPase VCP/p97 segregase[54–58]. The ubiquitination activities of RNF8/RNF168 are thought to be involved in determining the balance between recruitment of the repair protein 53BP1 and the BRCA1-A complex to damage sites, which determines if repair will proceeds by NHEJ or HR.

HERC2: HERC2 is a 500 kDa protein belonging to the HECT family of E3 Ub ligases that is required for RNF8 and RNF168 dependent ubiquitination[59]. Upon DNA damage, it is phosphorylated by ATM and binds the FHA domain of RNF8, to facilitate accumulation of 53BP1 and BRCA1 at damage sites by promoting stabilization, or formation, of the RNF8- Ubc13 complex. HERC2 is not directly involved in K63 chain formation through the Ub ligase activity of its HECT domain[60,61].

CHFR: The E3 Ub ligase CHFR is recruited to sites of damage near the onset of repair, prior to recruitment of RNF8/RNF168[62]. It contains FHA, RING, and cysteine rich domains, as well as a polyADP-ribose (PAR) binding zinc finger (PBZ) motif[63]. It interacts with the E2s Ubc13 and UbcH5C, to synthesize K63- and K48-linked chains, respectively[63]. CHFR binds polyADP-ribose modifications at sites of DNA damage through its PBZ motif, and functions to ubiquitinate PAR polymerase 1, which leads to its early removal from the DSB repair process via proteasomal degradation[62].

RNF138: The E3 Ub ligase RNF138, is recruited to damage sites through DNA binding mediated by its zinc finger domains (ZNF); this promotes end resection and HR repair by ubiquitinating and displacing Ku80 from ssDNA, and by binding Mre11 of the MRN complex through its ZNF domains to recruit CtIP/EXO1 nuclease to damage sites[64,65].

BRCA1-BARD1: BRCA1 is a tumor suppressor protein, found to be mutated in numerous malignancies; this protein is composed of an N-terminal RING and C-terminal

BRCT domains which facilitate interactions with multiple proteins to function in the maintenance of genome integrity[66]. BRCA1 interacts with BARD1, forming a heteromeric E3 Ub ligase that modifies substrates with different Ub linkages, including BRCA2, CtIP and histone H2A[30,66–75]. Whilst BRCA1 has been the subject of intense research efforts, the molecular basis for its involvement in HR-driven DSB repair remains mysterious[76,77]. Figure 1.3 displays the diverse Ub linkages reported at the sites of double strand breaks. A continuous flux of ubiquitin signaling is maintained by the activity of deubiquitinases specific of ubiquitin linkages[13,78,79]. The DUB activity of BRCC36[73,74], OTUB1[75,80,81] and POH1[82,83] counteract RNF8/RNF168 mediated ubiquitination, USP3[84,85], a H2A and H2B specific DUB acts as a negative regulator of BMI1/RNF2 and RNF20-RNF40 mediated ubiquitination[20].
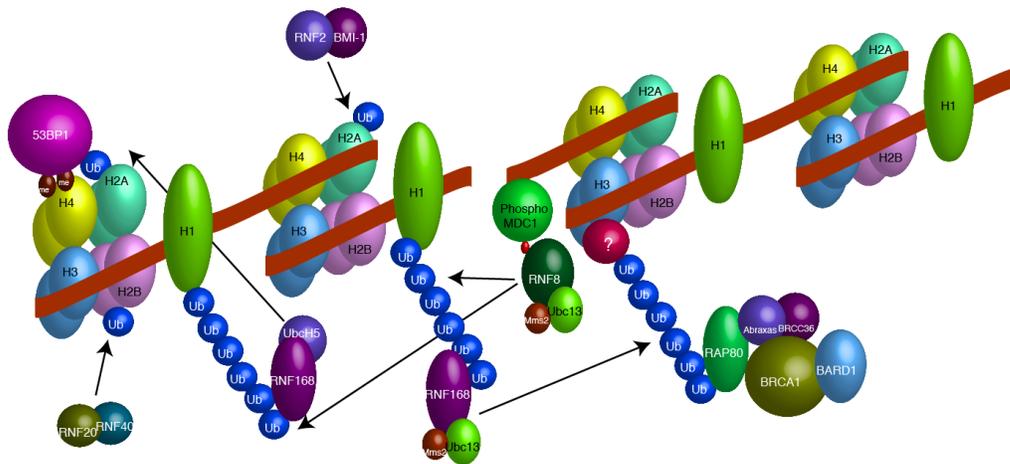
Figure 1.3: Diverse Ub modifications at the double strand break sites. RNF20-RNF40 leads to the monoubiquitination of histone H2B. RNF2-BMI1 is responsible for monoubiquitination of H2A. RNF8 binds to phosphorylated MDC1 and in conjunction with Ubc13 and Mms2 leads to the formation of K63 linked Ub chains on histone H1 that acts as a binding site for RNF168. RNF168 along with E2 UbcH5 leads to the monoubiquitination of H2A at K13/K15 that along with methylated histone H4 acts as a binding site for 53BP1. RNF168 can also partner with Ubc13 and Mms2 and catalyze Lys-63 linked Ub chains on an unknown protein. Lys-63 linked Ub chains are the binding sites for RAP80 that recruits BRCA1 A complex (RAP80, Abraxas, BRCC36, BRCA1, BARD1).

SUMO – A Post Translational Modifier

SUMO (Small Ubiquitin like MOdifier) belongs to the family of ubiquitin like modifiers and is expressed in five different isoforms: SUMO-1-5. They have been implicated in various biological functions like DNA repair, nuclear body formation and transcriptional regulation[26,86–88]. Although SUMO displays considerable similarities with Ub, it differs from Ub in the context of binding[26,87]. SUMO proteins are small, globular, possess a structural fold like ubiquitin (~10-12 kDa) and are conjugated to other proteins via an enzymatic cascade similar to ubiquitin involving SUMO specific E1, E2 and E3 enzymes[87]. SUMO-1 shows ~45% sequence identity with SUMO-2 and SUMO-3 whereas SUMO-2 and SUMO-3 share ~95% sequence identity[89], SUMO-4 awaits further research to expose its biological functions and SUMO-5 was recently discovered as a regulator of PML nuclear body with its expression being restricted to specific tissues[90]. SUMO is attached to the lysine residue of other proteins at a consensus sequence – ψKxE/D (ψ represents a large hydrophobic residue) via SUMO specific E1 activating enzyme that forms a thioester bond with the terminal glycine of SUMO that is generated upon maturation by proteases[90]. The resulting E1-SUMO complex then interacts with the only known SUMO E2, Ubc9 and the SUMO is transferred to the active site cysteine residue of the E2 enzyme. The conjugation of SUMO onto the substrate lysine can be carried out either by only E2 or a SUMO specific E3 enzyme can facilitate the reaction to achieve substrate specificity (See Figure 1.4)[87]. The isoforms SUMO-2 and SUMO-3 contain the consensus SUMOylation sequence and can be attached to form polySUMO chains of different or mixed isoforms. In contrast to SUMO-2 and SUMO-3, SUMO-1 does not contain a SUMOylation motif and is unable to form SUMO-1 chains thereby acting as a chain terminator, when conjugated. Like Ub signalling, the repertoire of SUMO signals is broadened by the presence of different isoforms and the ability of different lysines to be SUMOylated to form polySUMO chains of different linkages and isoforms.
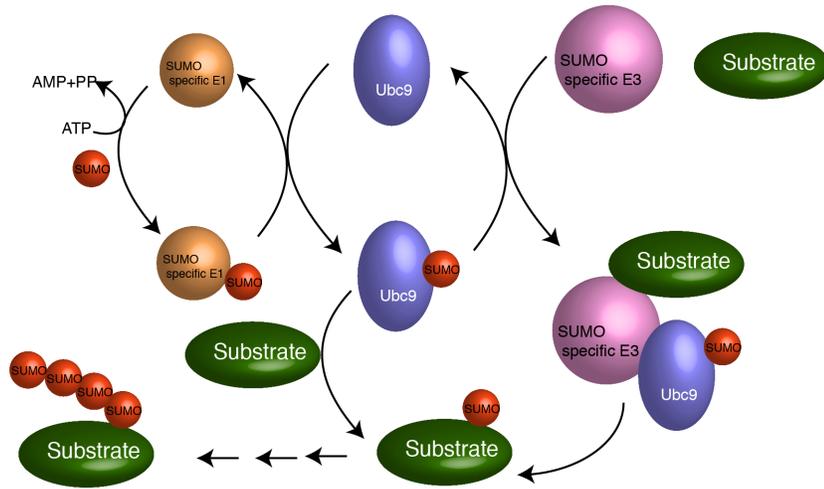
Figure 1.4: The SUMO conjugation machinery. SUMO is conjugated to SUMO specific E1 in an energy dependent manner followed by the interaction of E1~SUMO complex to E2 (Ubc9) and transfer of SUMO to E2. SUMO specific E3 with substrate specificity interact with E2-Ub and facilitate the transfer of SUMO to the substrate. Ubc9 (E2) can carry out the transfer of SUMO to substrate independently of E3. Subsequent rounds of this enzymatic cascade result in the formation of SUMO chains on the substrate. Note: SUMO-1 lacks a SUMOylation motif and hence incapable of forming chains.

Emerging Role of SUMO in DSB Repair by Homologous Recombination

Unlike ubiquitination, the role of SUMO in DSB repair is not well defined. However there is strong evidence indicating a deeper role for SUMO in the process. Different isoforms of SUMO have been shown to gather at sites of DSBs along with the SUMO E3 ligases PIAS1 and PIAS4[92,93], whose activity is required for proper accrual of Ub chains at DSBs. Although, the field is still in its nascent phase, these are the notable SUMOylation events affecting the DSB response, and involve timely recruitment and removal of repair factors from the sites of DNA damage.

- The polycomb group protein BMI1 has been shown to monoubiquitinate H2A, and its SUMOylation by CBX4 is a contributing factor for its accumulation at the DSB sites[94].

- The MDC1 protein is also SUMOylated upon DNA damage, a necessary step for its removal from sites of damage by the SUMO-targeted E3 Ub ligase (STUbL) RNF4[95].

- The RAP80 protein binds K63 chains at damage sites through tandem UIMs (Ub interacting motifs), and possesses a SIM (SUMO interacting motif) adjacent to the UIMs that is important for its recruitment to the damage sites.[48,96–99] RAP80 also interacts with the SUMO E2 Ubc9, and the ensuing SUMOylation is important for its functional role at DSBs[100].

- The BRCA1 is also modified by SUMO and co-localizes with SUMO1, SUMO2/3, and Ubc9 at DSB sites. SUMO modification of BRCA1 has been shown to enhance its Ub ligase activity in cells, which in turn, is important for DSB repair[92].

- The Mre11 protein is a constituent of the MRX complex in yeast (equivalent to the MRN complex in humans) and possesses two conserved SIMs. SIM1 is essential for the assembly of MRX complexes at DSB sites, and SIM2 is thought to recruit SUMOylated conjugates to Mre11 facilitating subsequent recruitment of SUMO E2 Ubc9 and the SUMO E3 Ub ligase Siz2, to enhance general SUMOylation of DNA repair proteins, presumably to assist repair[101].

- The exonuclease EXO1, that resects DNA ends during HR-mediated DSB repair, is constitutively SUMOylated by PIAS1/4-Ubc9, and this is a requirement for its Ub-dependent EXO1 degradation at stalled replication forks, to avoid excessive resection of free DNA ends. Moreover, it was found that the deSUMOylating enzyme SENP6 interacts with EXO1 to antagonize this process[102].

- The SUMO specific protease SENP7 assists DSB repair by cleaving SUMO chains on KAP1, a key transcriptional regulator; this releases CHD3, a chromatin remodelling complex, which then brings about chromatin decondensation, a physical requirement for repair proteins to gain access to the damage sites[103].

The above mentioned ubiquitination and SUMOylation events in the DSB repair process is just the tip of the iceberg, other kinds of post-translational modification such as phosphorylation, acetylation and methylation and their cross talk with each other contributes to the complexity of the process.

The Controversial Role of RAP80 in the Homologous Recombination

One of the special features of the homologous recombination repair is the DNA end resection, which is performed by the MRN complex and CtIP[11]. BRCA1 is also called "caretaker of genome" because of the multiple roles it plays at double strand break sites. It consists of N-terminal RING domain, required for its ubiquitin E3 ligase activity and a C-terminal BRCT domain through which it interacts with multiple proteins and accomplishes its several functions. The three complexes of BRCA1- A, B and C are named after its association with Abraxas, BACH1 and CtIP in a multiprotein complex. BRCA1 is recruited to the sites of DNA damage by its association with RAP80 in the A complex, promotes DNA end resection through its association with CtIP and assists in the process of RAD51 loading through its interaction with BRCA2[104]. Given its numerous functions in promoting homologous recombination it is not surprising to note that it is a tumor suppressor gene and women with BRCA1 mutations are at a high risk of developing breast and ovarian cancer in their lifetime[104,105]. Furthermore, approximately 20% of familial breast cancer has been linked to mutations in the BRCA1 and BRCA2 genes[105]. RAP80 due to its ability to recognize Lys-63 linked ubiquitin

chains and recruit BRCA1 in a multiprotein BRCA1 A complex at DNA damage sites, was believed to endorse HR over NHEJ[106-108]. However, recent advances suggest an opposite role for RAP80 in the repair process; RAP80 sequesters BRCA1 in the A complex and limits its association with other complexes that promote HR and ultimately acts as a negative regulator of HR[104,106].

RAP80 stands at the junction of Ub and SUMO signaling by recognizing both the moieties and leading the recruitment of crucial repair proteins at the damage sites[50,96,97,107]. The ability of RAP80 to recognize the Ub and SUMO linkages is crucial for the seamless progression of the repair process, underscored by the finding that a single deletion mutation in the first ubiquitin interacting motif is linked to cancer[105]. Moreover, a phosphorylation site mutant of the RAP80 SUMO interacting motif is recorded in the COSMIC (Catalog Of Somatic Mutation In Cancer); linking the failure to recognize these post-translational modifications at a molecular level to predisposition or progression of cancer[106]. Understanding the molecular details and energetics of the binding interaction between RAP80 and its partners Ub and SUMO is fundamental to gaining insight in to the Ub and SUMO signaling in the DNA repair pathway. This will also lead to a better understanding of other physiological processes in addition to DNA repair, that involve ubiquitin and SUMO modifications. Furthermore, it will aid the treatment of diseases like cancer by facilitating the design of rational drug therapies.

Molecular Basis for Recognition of Ubiquitin and SUMO Signals

Ubiquitin mostly interacts with its binding proteins through the hydrophobic patch centered around Ile44 and formed by the residues such as L8, I44, V70 and H68[26,27]. Another hydrophobic patch formed by I36, L71 and L73 of ubiquitin and the Phe4 patch formed by Q2, F4 and T12 serve as a site of interaction for different Ub interacting proteins[26]. All the surface of interaction on Ub along with the seven lysine residues are highlighted in Figure 1.5.

Ile-44 patch

Ile-36 patch

C-terminal   Phe 4 patch

Ubiquitin, PDB ID- 1UBQ
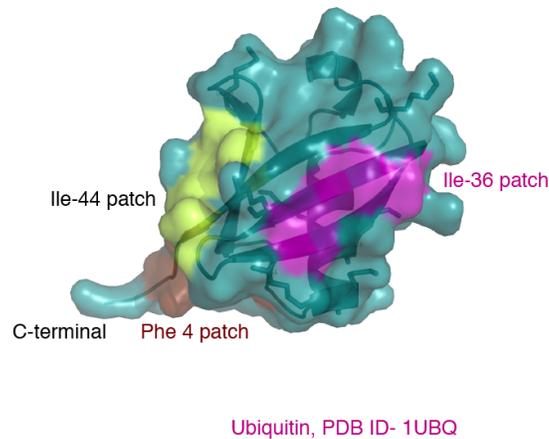
Figure 1.5: A surface representation of ubiquitin with Ile 44 patch highlighted in yellow, Ile 36 patch highlighted in pink and Phe4 patch highlighted in brown.

The proteins interacting with ubiquitin contain at least one of the ubiquitin binding domains (UBD): UBA (UBiquitin Associated domain), UIM (Ubiquitin Interacting Motif), NZF (Np14 Zinc Finger) or MIU (Motif Interacting with Ubiquitin)[109,110]. Although most of the UBDs bind Ub with a low affinity, they differ in their structure and site of interaction[109,110]. An extensively studied UBD is UIM, found in many trafficking proteins with a consensus motif of ~20 residues, centered around alanine (Ala) and flanked by acidic residues (Ac): *X*-Ac-Ac-Ac-Ac-Φ-*X*-*X*-Ala-*X*-*X*-*X*-Ser-*X*-*X*-Ac-*X*-*X*-*X*-*X* where X is less conserved residue and Φ stands for a large hydrophobic residue[111–113]. The UIMs interact with the Ile44 hydrophobic patch of Ub with a binding affinity ranging from 100 μM- 2 mM for monoubiquitin[110–113]. The DNA repair protein RAP80 that takes the centre stage of the thesis contains two tandem linked UIMs that binds K63 linked chains specifically at the sites of DNA damage[114,115]. As mentioned in the next chapter, Chapter 2, the NMR derived solution state structure of RAP80 reveals that the tandem UIMs adopt a helical structure, like other UIMs and the X-ray structure of tandem UIMs of RAP80 bound to K63 linked ubiquitin chain reveals

17

that the hydrophobic residues of the UIM contact the Ile 44 patch of Ub and the acidic residues flanking the hydrophobic residues of the UIM stabilize the interaction by contacting the positively charged flexible C-terminal tail of Ub, highlighting the impact of both the hydrophobic and acidic residues in the UIM-Ub interaction[114,116]. Since the UIM binds with a moderate to low affinity with Ub, most of the reported UIMs are placed close to each other in order to increase the affinity of the interaction to several fold, owing to the avidity effect[115]. The binding of RAP80 with Ub has been studied in detail, using NMR titrations and elaborate thermodynamic models and it was demonstrated that tandem UIMs of RAP80 binds Ub in a multivalent fashion to achieve a higher affinity, which is balanced by conformational selection between helix and coil states of the tandem UIMs[115]. The biological impact of this interaction was underscored when a single deletion mutation in the first UIM of RAP80 was linked with cases of familial breast and ovarian cancer[105]. The ΔE81 RAP80 mutant protein displayed impaired binding with Ub and multiple chromatid breaks were observed in the metaphase spread of cells expressing the ΔE81 RAP80 gene, displaying a highly compromised genomic integrity[116,105]. In Chapter 2, the structural and binding details of ΔE81 RAP80 were explored to understand the impaired binding with ubiquitin and using NMR titrations, stability studies and MD simulations we show that the deletion in the N-terminal of the first UIM of RAP80 leads to a structural frameshift of the N-cap motif and due to the loss of favorable electrostatic interaction, there is an abolishment of multivalent binding[116].

The field of SUMO research is in its incipient stage and there is much more to learn about the interaction with SUMO, which is evident in the finding that there are only two SUMO binding domains identified so far as opposed to numerous Ub binding domains reported. The SUMO interacting motif (SIM) and a novel SUMO binding zinc finger are the only two known SUMO binding domains, with extensive studies being done on SIM[87,117,118]. The SIMs have been reported in multitude of proteins involved in DNA repair, transcriptional repression, nuclear body formation and chromatin remodelling[87,119]. The SIMs studied so far are composed of a stretch of 4-5 hydrophobic residues flanked by acidic residues on either or both the sides of the

hydrophobic core[118]. In some cases, they also contain serine/threonine residues that can be phosphorylated juxtaposed to the hydrophobic core[49,120,121]. They adopt a β-strand conformation in isolation and upon binding to SUMO, they form an intermolecular β-sheet. The hydrophobic core of the SIM binds the hydrophobic groove formed between β_2 strand and α_1 helix of SUMO, highlighted in salmon pink and teal for SUMO-1 and SUMO-2 respectively, as shown in Figure 1.6. The flanking acidic residues or serines/threonine upon phosphorylation plays a major role in the SUMO-SIM binding. The negatively charged flanking region provides specificity to the orientation of the SIM binding to SUMO by forming favourable electrostatic interaction with the positively charged SUMO loop (connecting the β_2 strand and α_1 helix) and thereby stabilizing the interaction[49,91,120].
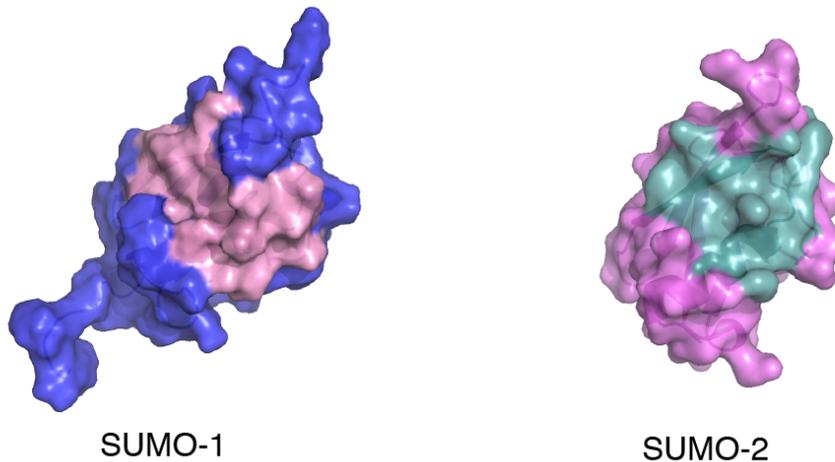
Figure 1.6: The surface representation of SUMO-1 and SUMO-2. The cleft between the β_2 strand and α_1 helix of SUMO-1 and SUMO-2 that serves as the binding site for proteins containing SUMO interacting motifs are coloured in salmon pink and teal, respectively.

In addition, these negatively charged residues also enhance the binding affinity by many fold[49,120]. RAP80 contains a SIM N-terminal to its tandem UIMs and have been reported to be required for the optimal recruitment of DNA repair complexes to the DNA damage sites[96,97]. The SIM of RAP80 also contains two serine residues C-terminal to the hydrophobic core that can be phosphorylated by CK2[49]. In Chapter 3, we explore the molecular basis of RAP80 SIM binding to SUMO-2 isoform using NMR techniques and MD simulations. We performed a detailed comparative analysis of RAP80 binding to SUMO-2 with and without phosphorylation and show that the binding affinity of RAP80 SIM for SUMO-2 increases by ~25 fold upon phosphorylation and the NMR lineshape analysis and chemical shift directed structure of phosphorylated RAP80 SIM bound to SUMO-2 reveals that the enhanced affinity is due to the increased electrostatic interactions between the phosphorylated residues of RAP80 SIM and the basic loop of SUMO connecting the β_2 strand and α_1 helix. In addition to deepening our understanding of SUMO: SIM interactions, our findings also indicate a role for CK2 enzyme in the recruitment of DNA repair complexes to the damage sites, mediated by RAP80.

Integration of Ubiquitin and SUMO Signals: The concept of SUMO-Ub Hybrid Chain

The complexity and diversity of Ub and SUMO signalling is further expanded by the discovery of STUbLs (SUMO Targeted Ubiquitin Ligases) such as RNF4 and RNF111. They conjugate Ub on to polySUMO chains and result in the formation of SUMO-Ub hybrid chains[122]. The STUbL RNF4 is recruited to the sites of DNA damage and also affects the recruitment of RAP80 and BRCA1[96,123]. RNF4 binds polySUMO chains through its multiple SIMs and ubiquitinates polySUMO chains through its RING domain[124]. The resulting SUMO-Ub hybrid chains (K63-Ub_2-SUMO-2) have been proposed as a highly specific signal in the DNA damage repair[96,97]. The hybrid chain bound the N-terminal SIM-UIMs of RAP80 with ~80 fold higher affinity as compared to RAP80 binding with Ub/SUMO alone[96]. The evidences for hybrid chain as a biologically relevant signal are exciting and compelling, however their physiological

presence is still in question. From a structural point of view, the interaction seems favourable due to the positioning of SUMO, Ub in hybrid chain and the relative distance between SIM and tandem UIMs of RAP80[20]. In terms of binding affinity, the multivalent binding can elevate the affinity substantially which can be expected to be enhanced further by several fold due to the phosphorylated SIM-SUMO binding[20]. Exploring the atomic details of the interaction will help us understand the recognition of such mixed signals in greater details.

References

[1]   Jackson, S. P., and Bartek, J. (2010) The DNA-damage response in human biology and disease. *Nature*. **461**, 1071–1078

[2]   Kornberg, R. D., and Lorch, Y. (1999) Twenty-Five Years of the Nucleosome , Fundamental Particle of the Eukaryote Chromosome. **98**, 285–294

[3]   Bednar, J., Horowitz, R. A., Grigoryev, S. A., Carruthers, L. M., Hansen, J. C., Koster, A. J., & Woodcock, C. L. (1998). Nucleosomes, linker DNA, and linker histone form a unique structural motif that directs the higher-order folding and compaction of chromatin. *PNAS*, **95**(24), 14173–14178.

[4]   Woodcock, C. L. (2005) A milestone in the odyssey of higher-order chromatin structure. *Nat. Str. Mol. Biol.*, **12**, 639–640

[5]   Manuelidis, L. (1990) View of Interphase Chromosomes. *Science,* **250**, 1533-1540.

[6]   Brown, J. S., Lukashchuk, N., Sczaniecka-Clift, M., Britton, S., le Sage, C., Calsou, P., Beli, P., Galanty, Y., and Jackson, S. P. (2015) Neddylation Promotes Ubiquitylation and Release of Ku from DNA-Damage Sites. *Cell Rep*. **11**, 704–714

[7]   Ciccia, A., and Elledge, S. J. (2010) The DNA Damage Response: Making It Safe to Play with Knives. *Mol. Cell*. **40**, 179–204

[8]   Thompson, L. H. (2012) Recognition, signaling, and repair of DNA double-strand breaks produced by ionizing radiation in mammalian cells: The molecular choreography. *Mutat. Res. - Rev. Mutat. Res.* **751**, 158–246

[9]   Lieber, M. R. (2010) The Mechanism of Double-Strand DNA Break Repair by the Nonhomologous DNA End-Joining Pathway. *Annu. Rev. Biochem.*, **79**(D)**,** 181-211.

[10]   Krejci, L., Altmannova, V., Spirek, M., and Zhao, X. (2012) Homologous recombination and its regulation. *Nucleic Acid Research,* **40**, 5795-5818

[11]   Schwertman, P., Bekker-Jensen, S., and Mailand, N. (2016) Regulation of DNA double-strand break repair by ubiquitin and ubiquitin-like modifiers. *Nat Rev Mol Cell Biol*. **17**, 379–394

[12]   Dellaire, G., and Bazett-Jones, D. P. (2004) PML nuclear bodies: Dynamic sensors of DNA damage and cellular stress. *BioEssays*. **26**, 963–977

[13]   Smeenk, G., and Mailand, N. (2016) Writers, readers, and erasers of histone ubiquitylation in DNA double-strand break repair. *Front. Genet*. **7**, 1–14

[14]   Polo, S., and Jackson, S. (2011) Dynamics of DNA damage response proteins at DNA breaks: a focus on protein modifications. *Genes Dev*. **25**, 409–433

[15]   Bekker-Jensen, S., and Mailand, N. (2010) Assembly and function of DNA double-strand break repair foci in mammalian cells. *DNA Repair (Amst)*. **9**, 1219–1228

[16]   Kuo, L. J., and Yang, L.-X. (2008) Gamma-H2AX - a novel biomarker for DNA double-strand breaks. *In Vivo*. **22**, 305–309

[17]   Stewart, G. S., Wang, B., Bignell, C. R., Taylor, A. M. R., and Elledge, S. J. (2003) MDC1 is a mediator of the mammalian DNA damage checkpoint. *Nature*. **421**, 961–966

[18]   Goldberg, M., Stucki, M., Falck, J., D'Amours, D., Rahman, D., Pappin, D., Bartek, J., and Jackson, S. P. (2003) MDC1 is required for the intra-S-phase DNA damage checkpoint. *Nature*. **421**, 952–956

[19]   Lou, Z., Silva Chini, C. C., Minter-Dykhouse, K., and Chen, J. (2003) Mediator of DNA damage checkpoint protein 1 regulates BRCA1 localization and phosphorylation in DNA damage checkpoint control. *J. Biol. Chem*. **278**, 13599–13602

[20]   Lee, B. L., Singh A., Glover, J. N. M.,Hendzel, M. J and Spyracopoulos, Leo. (2017) Molecular basis for K63-linked ubiquitination processes in double-strand

DNA break repair: A focus on kinetics and dynamics. *J. Mol. Biol.* http://dx.doi.org/10.1016/j.jmb.2017.05.029

[21] Hershko, A., and Ciechanover, A. (1998) The ubiquitin system. *Annu. Rev. Biochem.* **67**, 425–479

[22] Varshavsky, A. (2006) The early history of the ubiquitin field. *Protein Sci.* **15**, 647–654

[23] Pickart, C. M., and Eddins, M. J. (2004) Ubiquitin: Structures, functions, mechanisms. *Biochim. Biophys. Acta - Mol. Cell Res.* **1695**, 55–72

[24] Pickart, C. M. (2001) Echanisms nderlying ubiquitination. *Annu. Rev. Biochem.* **70**, 503–33

[25] Komander, D., Clague, M. J., and Urbé, S. (2009) Breaking the chains: structure and function of the deubiquitinases. *Nat. Rev. Mol. Cell Biol.* **10**, 550–563

[26] Komander, D., and Rape, M. (2012) The Ubiquitin Code. *Annual Review of Biochemistry,* **81,** 203-229.

[27] Swatek, K. N., and Komander, D. (2016) Ubiquitin modifications. *Cell Res.* **26**, 399–422

[28] Ikeda, F., and Dikic, I. (2008) Atypical Ubiquitin Chains: New Molecular Signals. 'Protein Modifications: Beyond the Usual Suspects' Review Series. *EMBO Rep.* **9**, 536–542

[29] Gatti, M., Pinato, S., Maiolica, A., Rocchio, F., Prato, M., Aebersold, R., and Penengo, L. (2015) RNF168 promotes noncanonical K27 ubiquitination to signal DNA damage. *Cell Rep.* **10**, 226–238

[30] Morris, J. R., and Solomon, E. (2004) BRCA1: BARD1 induces the formation of conjugated ubiquitin structures, dependent on K6 of ubiquitin, in cells during DNA replication and repair. *Hum. Mol. Genet.* **13**, 807–817

[31] Elia, A. E. H., Boardman, A. P., Wang, D. C., Huttlin, E. L., Everley, R. A., Dephoure, N., Zhou, C., Koren, I., Gygi, S. P., and Elledge, S. J. (2015) Quantitative Proteomic Atlas of Ubiquitination and Acetylation in the DNA Damage Response. *Mol. Cell.* **59**, 867–881

[32] Nakamura, K., Kato, A., Kobayashi, J., Yanagihara, H., Sakamoto, S., Oliveira, D. V. N. P., Shimada, M., Tauchi, H., Suzuki, H., Tashiro, S., Zou, L., and Komatsu,

K. (2011) Regulation of Homologous Recombination by RNF20-Dependent H2B Ubiquitination. *Mol. Cell.* **41**, 515–528

[33] Zhu, B., Zheng, Y., Pham, A. D., Mandal, S. S., Erdjument-Bromage, H., Tempst, P., and Reinberg, D. (2005) Monoubiquitination of human histone H2B: The factors involved and their roles in HOX gene regulation. *Mol. Cell.* **20**, 601–611

[34] Kim, J., Hake, S. B., and Roeder, R. G. (2005) The human homolog of yeast BRE1 functions as a transcriptional coactivator through direct activator interactions. *Mol. Cell.* **20**, 759–770

[35] Moyal, L., Lerenthal, Y., Gana-Weisz, M., Mass, G., So, S., Wang, S. Y., Eppink, B., Chung, Y. M., Shalev, G., Shema, E., Shkedy, D., Smorodinsky, N. I., van Vliet, N., Kuster, B., Mann, M., Ciechanover, A., Dahm-Daphi, J., Kanaar, R., Hu, M. C. T., Chen, D. J., Oren, M., and Shiloh, Y. (2011) Requirement of ATM-Dependent Monoubiquitylation of Histone H2B for Timely Repair of DNA Double-Strand Breaks. *Mol. Cell.* **41**, 529–542

[36] Ismail, H., Andrin, C., McDonald, D., and Hendzel, M. J. (2010) BMI1-mediated histone ubiquitylation promotes DNA double-strand break repair. *J. Cell Biol.* **191**, 45–60

[37] Ginjala, V., Nacerddine, K., Kulkarni, A., Oza, J., Hill, S. J., Yao, M., Citterio, E., van Lohuizen, M., and Ganesan, S. (2011) BMI1 is recruited to DNA breaks and contributes to DNA damage-induced H2A ubiquitination and repair. *Mol. Cell. Biol.* **31**, 1972–82

[38] Pan, M. R., Peng, G., Hungs, W. C., and Lin, S. Y. (2011) Monoubiquitination of H2AX protein regulates DNA damage response signaling. *J. Biol. Chem.* **286**, 28599–28607

[39] Huen, M. S. Y., Grant, R., Manke, I., Minn, K., Yu, X., Yaffe, M. B., and Chen, J. (2007) RNF8 Transduces the DNA-Damage Signal via Histone Ubiquitylation and Checkpoint Protein Assembly. *Cell.* **131**, 901–914

[40] Mailand, N., Bekker-Jensen, S., Faustrup, H., Melander, F., Bartek, J., Lukas, C., and Lukas, J. (2007) RNF8 Ubiquitylates Histones at DNA Double-Strand Breaks and Promotes Assembly of Repair Proteins. *Cell.* **131**, 887–900

[41] Kolas, N. K., Chapman, J. R., Nakada, S., Ylanko, J., Chahwan, R., Sweeney, F.

D., Panier, S., Mendez, M., Wildenhain, J., Thomson, T. M., Pelletier, L., Jackson, S. P., and Durocher, D. (2007) Orchestration of the DNA-damage response by the RNF8 ubiquitin ligase. *Science*. **318**, 1637–1640

[42] Thorslund, T., Ripplinger, A., Hoffmann, S., Wild, T., Uckelmann, M., Villumsen, B., Narita, T., Sixma, T. K., Choudhary, C., Bekker-Jensen, S., and Mailand, N. (2015) Histone H1 couples initiation and amplification of ubiquitin signalling after DNA damage. *Nature*. **527**, 389–93

[43] Doil, C., Mailand, N., Bekker-Jensen, S., Menard, P., Larsen, D. H., Pepperkok, R., Ellenberg, J., Panier, S., Durocher, D., Bartek, J., Lukas, J., and Lukas, C. (2009) RNF168 Binds and Amplifies Ubiquitin Conjugates on Damaged Chromosomes to Allow Accumulation of Repair Proteins. *Cell*. **136**, 435–446

[44] Stewart, G. S., Panier, S., Townsend, K., Al-Hakim, A. K., Kolas, N. K., Miller, E. S., Nakada, S., Ylanko, J., Olivarius, S., Mendez, M., Oldreive, C., Wildenhain, J., Tagliaferro, A., Pelletier, L., Taubenheim, N., Durandy, A., Byrd, P. J., Stankovic, T., Taylor, A. M. R., and Durocher, D. (2009) The RIDDLE Syndrome Protein Mediates a Ubiquitin-Dependent Signaling Cascade at Sites of DNA Damage. *Cell*. **136**, 420–434

[45] Wang, B., and Elledge, S. J. (2007) Ubc13/Rnf8 ubiquitin ligases control foci formation of the Rap80/Abraxas/Brca1/Brcc36 complex in response to DNA damage. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 20759–63

[46] Mattiroli, F., Vissers, J. H. A., Van Dijk, W. J., Ikpa, P., Citterio, E., Vermeulen, W., Marteijn, J. A., and Sixma, T. K. (2012) RNF168 ubiquitinates K13-15 on H2A/H2AX to drive DNA damage signaling. *Cell*. **150**, 1182–1195

[47] Gatti, M., Pinato, S., Maspero, E., Soffientini, P., Polo, S., and Penengo, L. (2012) A novel ubiquitin mark at the N-terminal tail of histone H2As targeted by RNF168 ubiquitin ligase. *Cell Cycle*. **11**, 2538–2544

[48] Kim, H., Chen, J., and Yu, X. (2007) Ubiquitin-Binding Protein RAP80 Mediates BRCA1-Dependent DNA Damage Response. *Science*, **316**, 1202-1205

[49] Anamika, and Spyracopoulos, L. (2016) Molecular basis for phosphorylation-dependent SUMO recognition by the DNA repair protein RAP80. *J. Biol. Chem.* **291**, 4417–4428

[50] Sobhian, B., Shao, G., Lilli, D. R., Culhane, A. C., Moreau, L. A., Xia, B., Livingston, D. M., and Greenberg, R. A. (2007) RAP80 Targets BRCA1 to Specific Ubiquitin Structures at DNA Damage Sites. *Science*, **316**, 1198-1202

[51] Botuyan, M. V., Lee, J., Ward, I. M., Kim, J. E., Thompson, J. R., Chen, J., and Mer, G. (2006) Structural Basis for the Methylation State-Specific Recognition of Histone H4-K20 by 53BP1 and Crb2 in DNA Repair. *Cell*. **127**, 1361–1373

[52] Wilson, M. D., Benlekbir, S., Fradet-Turcotte, A., Sherker, A., Julien, J.-P., McEwan, A., Noordermeer, S. M., Sicheri, F., Rubinstein, J. L., and Durocher, D. (2016) The structural basis of modified nucleosome recognition by 53BP1. *Nature*. **536**, 100–103

[53] Fradet-Turcotte, A., Canny, M. D., Escribano-Díaz, C., Orthwein, A., Leung, C. C. Y., Huang, H., Landry, M.-C., Kitevski-LeBlanc, J., Noordermeer, S. M., Sicheri, F., and Durocher, D. (2013) 53BP1 is a reader of the DNA-damage-induced H2A Lys 15 ubiquitin mark. *Nature*. **499**, 50–54

[54] Min, J., Allali-Hassani, A., Nady, N., Qi, C., Ouyang, H., Liu, Y., MacKenzie, F., Vedadi, M., and Arrowsmith, C. H. (2007) L3MBTL1 recognition of mono- and dimethylated histones. *Nat. Struct. Mol. Biol.* **14**, 1229–1230

[55] Acs, K., Luijsterburg, M. S., Ackermann, L., Salomons, F. a, Hoppe, T., and Dantuma, N. P. (2011) The AAA-ATPase VCP/p97 promotes 53BP1 recruitment by removing L3MBTL1 from DNA double-strand breaks. *Nat. Struct. Mol. Biol.* **18**, 1345–1350

[56] Mallette, F. A., Mattiroli, F., Cui, G., Young, L. C., Hendzel, M. J., Mer, G., Sixma, T. K., and Richard, S. (2012) RNF8- and RNF168-dependent degradation of KDM4A/JMJD2A triggers 53BP1 recruitment to DNA damage sites. *EMBO J.* **31**, 1865–1878

[57] Nakada, S. (2016) Opposing roles of RNF8/RNF168 and deubiquitinating enzymes in ubiquitination-dependent DNA double-strand break response signaling and DNA-repair pathway choice. *J. Radiat. Res.* **57**, i33–i40

[58] Lok, G. T.-M., Sy, S. M.-H., Dong, S.-S., Ching, Y.-P., Tsao, S. W., Thomson, T. M., and Huen, M. S. Y. (2012) Differential regulation of RNF8-mediated Lys48- and Lys63-based poly-ubiquitylation. *Nucleic Acids Res.* **40**, 196–205

[59] Bekker-Jensen, S., Rendtlew Danielsen, J., Fugger, K., Gromova, I., Nerstedt, A., Lukas, C., Bartek, J., Lukas, J., and Mailand, N. (2010) HERC2 coordinates ubiquitin-dependent assembly of DNA repair factors on damaged chromosomes. *Nat. Cell Biol.* **12**, 12–80

[60] Mohiuddin, Kobayashi, S., Keka, I. S., Guilbaud, G., Sale, J., Narita, T., Abdel-Aziz, H. I., Wang, X., Ogawa, S., Sasanuma, H., Chiu, R., Oestergaard, V. H., Lisby, M., and Takeda, S. (2016) The role of HERC2 and RNF8 ubiquitin E3 ligases in the promotion of translesion DNA synthesis in the chicken DT40 cell line. *DNA Repair (Amst).* **40**, 67–76

[61] Oestergaard, V. H., Pentzold, C., Pedersen, R. T., Iosif, S., Alpi, A., Bekker-Jensen, S., Mailand, N., and Lisby, M. (2012) RNF8 and RNF168 but not HERC2 are required for DNA damage-induced ubiquitylation in chicken DT40 cells. *DNA Repair (Amst).* **11**, 892–905

[62] Liu, C., Wu, J., Paudyal, S. C., You, Z., and Yu, X. (2013) CHFR is important for the first wave of ubiquitination at DNA damage sites. *Nucleic Acids Res.* **41**, 1698–1710

[63] Bothos, J., Summers, M. K., Venere, M., Scolnick, D. M., and Halazonetis, T. D. (2003) The Chfr mitotic checkpoint protein functions with Ubc13-Mms2 to form Lys63-linked polyubiquitin chains. *Oncogene.* **22**, 7101–7107

[64] Ismail, I. H., Gagné, J.-P., Genois, M.-M., Strickfaden, H., McDonald, D., Xu, Z., Poirier, G. G., Masson, J.-Y., and Hendzel, M. J. (2015) The RNF138 E3 ligase displaces Ku to promote DNA end resection and regulate DNA repair pathway choice. *Nat. Cell Biol.* **17**, 1446–1457

[65] Schmidt, C. K., Galanty, Y., Sczaniecka-Clift, M., Coates, J., Jhujh, S., Demir, M., Cornwell, M., Beli, P., and Jackson, S. P. (2015) Systematic E2 screening reveals a UBE2D–RNF138–CtIP axis promoting DNA repair. *Nat. Cell Biol.* **17**, 1458–1470

[66] Prakash, R., Zhang, Y., Feng, W., and Jasin, M. (2015) Homologous Recombination and Human Health. *Perspect. Biol.* **7(4)**; a016600

[67] Wu-Baer, F., Lagrazon, K., Yuan, W., and Baer, R. (2003) The BRCA1/BARD1 heterodimer assembles polyubiquitin chains through an unconventional linkage involving lysine residue K6 of ubiquitin. *J. Biol. Chem.* **278**, 34743–34746

[68] Baer, R., and Ludwig, T. (2002) The BRCA1/BARD1 heterodimer, a tumor suppressor complex with ubiquitin E3 ligase activity. *Curr. Opin. Genet. Dev.* **12**, 86–91

[69] Kalb, R., Mallery, D. L., Larkin, C., Huang, J. T. J., and Hiom, K. (2014) BRCA1 is a histone-H2A-specific ubiquitin ligase. *Cell Rep.* **8**, 999–1005

[70] Yu, X., Fu, S., Lai, M., Baer, R., Chen, J., Breast, B., and Susceptibility, C. (2006) BRCA1 ubiquitinates its binding partner CtIP. *Genes Dev.* **1**, 1721–1726

[71] Drost, R., Bouwman, P., Rottenberg, S., Boon, U., Schut, E., Klarenbeek, S., Klijn, C., van der Heijden, I., van der Gulden, H., Wientjens, E., Pieterse, M., Catteau, A., Green, P., Solomon, E., Morris, J. R., and Jonkers, J. (2011) BRCA1 RING function is essential for tumor suppression but dispensable for therapy resistance. *Cancer Cell.* **20**, 797–809

[72] Shakya, R., Reid, L. J., Reczek, C. R., Cole, F., Egli, D., Lin, C.-S., deRooij, D. G., Hirsch, S., Ravi, K., Hicks, J. B., Szabolcs, M., Jasin, M., Baer, R., and Ludwig, T. (2011) BRCA1 Tumor Suppression Depends on BRCT Phosphoprotein Binding, But Not Its E3 Ligase Activity. *Science.* **334**, 525–528

[73] Chen, X., Arciero, C. A., Wang, C., Broccoli, D., and Godwin, A. K. (2006) BRCC36 is essential for ionizing radiation-induced BRCA1 phosphorylation and nuclear foci formation. *Cancer Res.* **66**, 5039–5046

[74] Shao, G., Lilli, D. R., Patterson-Fortin, J., Coleman, K. a, Morrissey, D. E., and Greenberg, R. a (2009) The Rap80-BRCC36 de-ubiquitinating enzyme complex antagonizes RNF8-Ubc13-dependent ubiquitination events at DNA double strand breaks. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 3166–3171

[75] Nakada, S., Tai, I., Panier, S., Al-Hakim, A., Iemura, S.-I., Juang, Y.-C., O'Donnell, L., Kumakubo, A., Munro, M., Sicheri, F., Gingras, A.-C., Natsume, T., Suda, T., and Durocher, D. (2010) Non-canonical inhibition of DNA damage-dependent ubiquitination by OTUB1. *Nature.* **466**, 941–946

[76] Yun, M. H., and Hiom, K. (2009) Understanding the functions of BRCA1 in the DNA-damage response. *Biochem. Soc. Trans.* **37**, 597–604

[77] Wang, B. (2012) BRCA1 tumor suppressor network: focusing on its tail. *Cell Biosci.* **2**, 6

[78]   Jacq, X., Kemp, M., Martin, N. M. B., and Jackson, S. P. (2013) Deubiquitylating Enzymes and DNA Damage Response Pathways. *Cell Biochem. Biophys.* **67**, 25–43

[79]   Messick, T. E., and Greenberg, R. A. (2009) The ubiquitin landscape at DNA double-strand breaks. *J. Cell Biol.* **187**, 319–326

[80]   Wiener, R., Zhang, X., Wang, T., and Wolberger, C. (2012) The mechanism of OTUB1-mediated inhibition of ubiquitination. *Nature.* **483**, 618–22

[81]   Juang, Y. C., Landry, M. C., Sanches, M., Vittal, V., Leung, C. C. Y., Ceccarelli, D. F., Mateo, A. R. F., Pruneda, J. N., Mao, D. Y. L., Szilard, R. K., Orlicky, S., Munro, M., Brzovic, P. S., Klevit, R. E., Sicheri, F., and Durocher, D. (2012) OTUB1 Co-opts Lys48-Linked Ubiquitin Recognition to Suppress E2 Enzyme Function. *Mol. Cell.* **45**, 384–397

[82]   Butler, L. R., Densham, R. M., Jia, J., Garvin, A. J., Stone, H. R., Shah, V., Weekes, D., Festy, F., Beesley, J., and Morris, J. R. (2012) The proteasomal de-ubiquitinating enzyme POH1 promotes the double-strand DNA break response. *EMBO J.* **31**, 3918–3934

[83]   Morris, J. R. (2012) Attenuation of the ubiquitin conjugate DNA damage signal by the proteasomal DUB POH1. *Cell Cycle.* **11**, 4103–4104

[84]   Lancini, C., van den Berk, P. C. M., Vissers, J. H. a, Gargiulo, G., Song, J.-Y., Hulsman, D., Serresi, M., Tanger, E., Blom, M., Vens, C., van Lohuizen, M., Jacobs, H., and Citterio, E. (2014) Tight regulation of ubiquitin-mediated DNA damage response by USP3 preserves the functional integrity of hematopoietic stem cells. *J. Exp. Med.* **211**, 1759–1777

[85]   Chen, D., Dai, C., and Jiang, Y. (2015) Histone H2A and H2B Deubiquitinase in Developmental Disease and Cancer. *Cancer Transl. Med.* **1**, 170

[86]   Müller, S., Hoege, C., Pyrowolakis, G., and Jentsch, S. (2001) SUMO, ubiquitin's mysterious cousin. *Nat. Rev. Mol. Cell Biol.* **2**, 202–210

[87]   Gareau, J. R., and Lima, C. D. (2010) The SUMO pathway: emerging mechanisms that shape specificity, conjugation and recognition. *Nat. Rev. Mol. Cell Biol.* **11**, 861–871

[88]   Morris, J. R., and Garvin, A. J. (2017) SUMO in the DNA Double-Stranded Break

Response : Similarities , Differences , and Cooperation with Ubiquitin. *J. Mol. Biol.* 10.1016/j.jmb.2017.05.012

[89]   Matic, I., van Hagen, M., Schimmel, J., Macek, B., Ogg, S. C., Tatham, M. H., Hay, R. T., Lamond, A. I., Mann, M., and Vertegaal, A. C. O. (2007) In Vivo Identification of Human Small Ubiquitin-like Modifier Polymerization Sites by High Accuracy Mass Spectrometry and an in Vitro to in Vivo Strategy. *Mol. Cell. Proteomics.* **7**, 132–144

[90]   Liang, Y., Lee, C., Yao, Y., Lai, C., and Schmitz, M. L. (2016) SUMO5 , a Novel Poly-SUMO Isoform , Regulates PML Nuclear Bodies. *Nat. Publ. Gr.* 10.1038/srep26509

[91]   Kerscher, O. (2007) SUMO junction-what's your function? New insights through SUMO-interacting motifs. *EMBO Rep.* **8**, 550–555

[92]   Morris, J. R., Boutell, C., Keppler, M., Densham, R., Weekes, D., Alamshah, A., Butler, L., Galanty, Y., Pangon, L., Kiuchi, T., Ng, T., and Solomon, E. (2009) The SUMO modification pathway is involved in the BRCA1 response to genotoxic stress. *Nature.* **462**, 886–890

[93]   Galanty, Y., Belotserkovskaya, R., Coates, J., Polo, S., Miller, K. M., and Jackson, S. P. (2009) Mammalian SUMO E3-ligases PIAS1 and PIAS4 promote responses to DNA double-strand breaks. *Nature.* **462**, 935–939

[94]   Ismail, I. H., Gagné, J. P., Caron, M. C., McDonald, D., Xu, Z., Masson, J. Y., Poirier, G. G., and Hendzel, M. J. (2012) CBX4-mediated SUMO modification regulates BMI1 recruitment at sites of DNA damage. *Nucleic Acids Res.* **40**, 5497–5510

[95]   Luo, K., Zhang, H., Wang, L., Yuan, J., and Lou, Z. (2012) Sumoylation of MDC1 is important for proper DNA damage response. *EMBO J.* **31**, 3008–3019

[96]   Guzzo, C. M., Berndsen, C. E., Zhu, J., Gupta, V., Greenberg, R. A., Wolberger, C., Matunis, M. J., Guzzo, C. M., Berndsen, C. E., Zhu, J., Gupta, V., and Datta, A. (2013) RNF4-Dependent Hybrid SUMO-Ubiquitin Chains Are Signals for RAP80 and Thereby Mediate the Recruitment of BRCA1 to Sites of DNA Damage. *Sci. Signal.* **5**, ra88

[97]   Hu, X., Paul, A., and Wangs, B. (2012) Rap80 protein recruitment to DNA double-

strand breaks requires binding to both small ubiquitin-like modifier (SUMO) and ubiquitin conjugates. *J. Biol. Chem.* **287**, 25510–25519

[98] Yan, J., Kim, Y., Yang, X., Albers, M., Koegl, M., and Jetten, A. M. (2007) Ubiquitin-interaction motifs of RAP80 are critical in its regulation of estrogen receptor α. *Nucleic Acids Research*, **35**, 1673–1686

[99] Wang, B., Matsuoka, S., Ballif, B., Zhang, D., Smogorzewska, A., Giyi, S., and Elledge, S. J. (2007) Abraxas and RAP80 form a BRCA1 Protein Complex required for the DNA damage response. *Science*, **316**, 1194–1198

[100] Yan, J., Yang, X. P., Kim, Y. S., Joo, J. H., and Jetten, A. M. (2007) RAP80 interacts with the SUMO-conjugating enzyme UBC9 and is a novel target for sumoylation. *Biochem. Biophys. Res. Commun.* **362**, 132–138

[101] Chen, Y. J., Chuang, Y. C., Chuang, C. N., Cheng, Y. H., Chang, C. R., Leng, C. H., and Wang, T. F. (2016) S. cerevisiae Mre11 recruits conjugated SUMO moieties to facilitate the assembly and function of the Mre11-Rad50-Xrs2 complex. *Nucleic Acids Res.* **44**, 2199–2213

[102] Bologna, S., Altmannova, V., Valtorta, E., Koenig, C., Liberali, P., Gentili, C., Anrather, D., Ammerer, G., Pelkmans, L., Krejci, L., and Ferrari, S. (2015) Sumoylation regulates EXO1 stability and processing of DNA damage. *Cell Cycle*. **14**, 2439–2450

[103] Garvin, A. J., Densham, R. M., Blair-Reid, S. A., Pratt, K. M., Stone, H. R., Weekes, D., Lawrence, K. J., and Morris, J. R. (2013) The deSUMOylase SENP7 promotes chromatin relaxation for homologous recombination DNA repair. *EMBO Rep.* **14**, 975–983

[104] Her, J., Soo Lee, N., Kim, Y., and Kim, H. (2016) Factors forming the BRCA1-A complex orchestrate BRCA1 recruitment to the sites of DNA damage. *Acta Biochim. Biophys. Sin. (Shanghai)*. **48**, 658–664

[105] Nikkilä, J., Coleman, K. A., Morrissey, D., Pylkäs, K., Erkko, H., Messick, T. E., Karppinen, S.-M., Amelina, A., Winqvist, R., and Greenberg, R. A. (2009) Familial breast cancer screening reveals an alteration in the RAP80 UIM domain that impairs DNA damage response function. *Oncogene*. **28**, 1843–1852

[106] Lombardi, P. M., Matunis, M. J., and Wolberger, C. (2017) RAP80, ubiquitin and

SUMO in the DNA damage response. *J. Mol. Med.* **95**, 799–807

[107] Kim, H., Chen, J., and Yu, X. (2007) Ubiquitin-binding protein RAP80 mediates BRCA1-dependent DNA damage response. *Science.* **316**, 1202–1205

[108] Yan, J., Yang, X. P., Kim, Y. S., and Jetten, A. M. (2008) RAP80 responds to DNA damage induced by both ionizing radiation and UV irradiation and is phosphorylated at Ser205. *Cancer Res.* **68**, 4269–4276

[109] Hurley, J. H., Lee, S., and Prag, G. (2006) Ubiquitin-binding domains. *Biochem. J.,***399**, 361–372

[110] Dikic, I., Wakatsuki, S., and Walters, K. J. (2009) Ubiquitin-binding domains — from structures to functions. *Nat. Rev. Mol. Cell Biol.* **10**, 659–671

[111] Fisher, R. D., Wang, B., Alam, S. L., Higginson, D. S., Robinson, H., Sundquist, W. I., and Hill, C. P. (2003) Structure and Ubiquitin Binding of the Ubiquitin-interacting Motif *.*J. Biol. Chem.*, **278**, 28976–28984

[112] Hofmann, K., and Falquet, L. (2001) A ubiquitin-interacting motif conserved in components of the proteasomal and lysosomal protein degradation systems. *Trends in Biochemical Sciences* **26**, 347–350

[113] Young, P., Deveraux, Q., Beal, R. E., Pickart, C. M., and Rechsteiner, M. (1998) Characterization of Two Polyubiquitin Binding Sites in the 26 S Protease Subunit 5a *. *J. Biol. Chem.*, **273**, 5461–5467

[114] Sato, Y., Yoshikawa, A., Mimura, H., Yamashita, M., Yamagata, A., and Fukai, S. (2009) Structural basis for specific recognition of Lys 63-linked polyubiquitin chains by tandem UIMs of RAP80. *EMBO J.* **28**, 2461–2468

[115] Markin, C. J., Xiao, W., and Spyracopoulos, L. (2010) Mechanism for recognition of polyubiquitin chains: Balancing affinity through interplay between multivalent binding and dynamics. *J. Am. Chem. Soc.* **132**, 11247–11258

[116] Anamika, Markin, C. J., Rout, M. K., and Spyracopoulos, L. (2014) Molecular basis for impaired DNA damage response function associated with the RAP80 ΔE81 defect. *J. Biol. Chem.* **289**, 12852–12862

[117] Minty, A., Dumont, X., Kaghad, M., and Caput, D. (2000) Covalent Modification of p73α by SUMO-1. *J. Biol. Chem.,***275**, 36316–36323

[118] Hecker, C. M., Rabiller, M., Haglund, K., Bayer, P., and Dikic, I. (2006)

Specification of SUMO1- and SUMO2-interacting motifs. *J. Biol. Chem.* **281**, 16117–16127

[119] Heun, P. (2007) SUMOrganization of the nucleus. *Curr. Opin. Cell Biol.* **19**, 350–355

[120] Cappadocia, L., Mascle, X. H., Bourdeau, V., Tremblay-Belzile, S., Chaker-Margot, M., Lussier-Price, M., Wada, J., Sakaguchi, K., Aubry, M., Ferbeyre, G., and Omichinski, J. G. (2015) Structural and Functional Characterization of the Phosphorylation-Dependent Interaction between PML and SUMO1. *Structure.* **23**, 126–138

[121] Escobar-cabrera, E., Okon, M., Lau, K. W., Dart, C. F., Alexandre, M., Bonvin, J. J., and Mcintosh, L. P. (2011) Characterizing the N- and C-terminal Small Ubiquitin-like Modifier (SUMO) Interacting Motifs of the Scaffold Protein. *J. Biol. Chem.,* **286**, 19816-19829

[122] Perry, J. J. P., Tainer, J. A., and Boddy, M. N. (2008) A SIM-ultaneous role for SUMO and ubiquitin. *Trends Biochem. Sci.* **33**, 201–208

[123] Galanty, Y., Belotserkovskaya, R., Coates, J., and Jackson, S. P. (2012) RNF4, a SUMO-rageted ubiquitin E3 ligase , promotes DNA double-strand break repair. *Genes Dev.* **26**, 1179–1195

[124] Plechanovová, A., Jaffray, E. G., McMahon, S. A., Johnson, K. A., Navrátilová, I., Naismith, J. H., and Hay, R. T. (2011) Mechanism of ubiquitylation by dimeric RING ligase RNF4. *Nat. Struct. Mol. Biol.* **18**, 1052–1059

Chapter 2

Molecular Basis for Impaired DNA Damage Response

Function Associated with the RAP80 ΔE81 Defect

Introduction

The recognition of K63-linked polyubiquitin (polyUb) chains plays a central role in the DNA damage response (DDR) by recruiting repair protein complexes to sites of double stranded DNA breaks (DSB), or DNA damage foci[1]. The protein RAP80 is involved in multivalent recognition of polyUb chains through N-terminal tandem ubiquitin interacting motifs (tUIMs), and is fundamental for protein recruitment in the DDR[1,2]. Defects within mechanisms underlying the binding of repair proteins to DNA damage foci may give rise to tumorigenesis and the progression of disease[3,4]. This highlights intense efforts to identify binding partners for BRCA1 and BRCA2, tumor suppressors whose recruitment to sites of DNA damage is essential for repair[5,6]. RAP80 is a large, 80 kDa protein comprised of a number of domains; the polyUb binding function is localized to a pair of closely spaced, or tandem UIMs, short α-helices that bind the canonical hydrophobic interface centered on Ile44 of Ub[7-9]. The individual UIMs of RAP80 have low affinity for Ub monomers (K_D ~500 mM). In tandem, the UIMs employ a multivalent binding mechanism to substantially increase the affinity for K63-linked or linear polyUb chains with two to four linked monomers, to K_D values ranging from 2–75 mM[7,8]. Approximately 20% of familial breast cancers arise from germline mutations in the *BRCA1* and *BRCA2* genes[3]. The genetic basis for the remaining cancers is not known, but is suspected to involve mutations in genes with lower penetrance than *BRCA1* and *BRCA2*, potentially in combination with environmental or other factors[3]. Given that BRCA1 and BRCA2 proteins function within protein complexes that localize to sites of DNA damage, mutations in genes for proteins that interact with these tumor suppressors may also be involved in cancer susceptibility[3]. Recently, an alteration within the highly conserved N-terminal UIM region from the *RAP80* gene was discovered in a group of 112 *BRCA1/BRCA2* mutation negative Finnish breast cancer families[3]. The alteration, ΔE81, is an in-frame deletion of the first of three sequential Glu residues that occur at the N-terminus of the first UIM α-helix. The RAP80 ΔE81 mutant demonstrated reduced ubiquitin binding in GST pull down assays, and reduced localization to DNA damage

foci in confocal microscopy studies with U2OS cells. Furthermore, cells expressing ΔE81 RAP80 show increases in chromosomal abnormalities that are consistent with defective DSB repair[3]; these results point to the biological relevance of the ΔE81 mutation, and a potential familial link to cancer. We investigated the molecular basis for impaired recognition of Ub by ΔE81 RAP80, through the use of solution state NMR spectroscopy in conjunction with detailed thermodynamic binding models. In addition, we used the temperature dependence of the $^1H_\alpha$ chemical shifts to measure the stability of the α-helices for the tandem UIMs in ΔE81. The results indicate that the multivalent binding advantage for tandem UIMs in RAP80 is lost in the ΔE81 mutant, and provide a molecular description for a primarily structural defect that may be involved in the pathogenesis of cancer.

Experimental Procedures

Mutagenesis, protein expression and purification

The ΔE81-RAP80-tUIM mutant was made using PCR mutagenesis with recombinant pGEX-6P-1 RAP80-tUIM plasmid[8] as a template. Inserts harboring the deletion mutant were re-ligated into pGEX-6P-1 and verified by sequencing (ClearView BioStructures, Inc., Edmonton, Alberta). Overexpression and purification of $[U-^{15}N]$ and $[U-^{13}C,^{15}N]$-wild type-RAP80-tUIM was carried out as previously described [8,10]. Overexpression and purification of $[U-^{15}N]$ and $[U-^{13}C,^{15}N]$- ΔE81-RAP80-tUIM was also carried out as previously described for wild type RAP0-tUIM[8,10], with the exception that 10 M urea was added to the purified protein solution, and removed by dialysis before the final size-exclusion chromatography step (Superdex 30). This step was included in order to denature any remaining protease following cleavage of the GST affinity tag. Tandem Ub_2 was expressed and purified as previously described[8], whereas recombinant ubiquitin was purchased from Boston Biochem (Cambridge, MA).

NMR spectroscopy, main chain relaxation measurements and structure determination

NMR spectra for titrations of ΔE81-RAP80-tUIM with Ub/Ub$_2$ were collected at 25 °C using a Varian Unity INOVA 600 MHz spectrometer. Samples (350 *m*L) were placed in 5 mm SHIGEMI microcell NMR tubes, with buffer containing 50 mM TRIS, 150 mM NaCl, 2 mM DTT pH 7.3 and 10 % D$_2$O. Chemical shift assignment for ΔE81-RAP80-tUIM was conducted at 5°C, the NMR sample contained 0.1 mM DSS as an internal chemical shift reference, and the protein concentration was 0.6 mM, with buffer conditions identical to those at 25°C. Backbone ^{1}H^N, ^{1}H$_\alpha$, ^{13}C$_\alpha$, ^{13}C$_\beta$, and ^{15}N chemical shifts were assigned using 3D HNCACB[11,12], CBCA(CO)NNH[12,13], and HBHA(CO)NH[14] spectra, and the automatic assignment feature within the program CARA[15]. These assignments were verified manually using the Sparky program[16]. Main chain dynamics for wild type and ΔE81-RAP80-tUIM were analyzed through measurement of ^{15}N-R_1, R_2, and {^{1}H}–^{15}N nuclear Overhauser effect (NOE) relaxation data[17] at 5 °C and 600 MHz. NMR samples (350 ml) were placed in a SHIGEMI microcell NMR tubes in buffer consisting of 50 mM TRIS, 150 mM NaCl, 2 mM DTT pH 7.3 and 10 % D$_2$O. Protein concentrations were 0.5 and 1.3 mM for [*U*-^{15}N]-wild type and [*U*-^{15}N]- ΔE81-RAP80-tUIM, respectively. The relaxation data were analyzed according to the model free approach as previously described [18,19]. For NMR structure determination, NMR spectra for wild type and ΔE81-RAP80-tUIM were collected at 5°C using a Varian Unity INOVA 600 MHz spectrometer. NMR samples (350 ml) were prepared in 5 mm SHIGEMI microcell NMR tubes in buffer containing 50 mM TRIS, 150 mM NaCl, 2 mM DTT pH 7.3 and 10 % D$_2$O, with 1.2 and 0.6 mM for [*U*-^{13}C,^{15}N]-wild type and [*U*-^{13}C,^{15}N]- ΔE81-RAP80-tUIM, respectively. Backbone ^{1}H^N, ^{1}H$_\alpha$, ^{13}C$_\alpha$, ^{13}C$_\beta$, ^{13}CO and ^{15}N chemical shifts at 5°C for wild type RAP80-tUIM were assigned using 3D HNCO[20], HNCACB[11,12], CBCA(CO)NNH[12,13], and HBHA(CO)NH[14] spectra, and the automatic assignment feature within the program CARA[15]; assignments were verified manually using the Sparky program[16]. Side chain ^{1}H chemical shifts were assigned using 3D ^{15}N TOCSY-HSQC spectra[21] with the aid of assignments at 25°C. Interproton distance restraints were derived from 3D ^{15}N NOESY-

HSQC spectra[21] with a mixing time of 200 ms. For these NMR experiments, sample conditions were similar to those used for chemical shift assignment at 5°C, with the exception that protein concentrations were 0.3 and 1.3 mM for $[U\text{-}^{13}C,^{15}N]$-wild type and $[U\text{-}^{15}N]$- ΔE81-RAP80-tUIM, respectively. Quantitative main chain ϕ and ψ dihedral angle restraints as well as rotamer probabilities for the side chain χ^1 dihedral angles, were derived from NMR chemical shifts using the TALOS-N program[22]. In addition, helix N-cap box motifs were identified from NMR chemical shifts using the MICS program [23]. Structures for wild type and ΔE81-RAP80 were calculated using torsion angle dynamics and the simulated annealing protocol in Xplor-NIH[24]. NOE restraints were sorted into strong (1.8–2.9 Å), medium (1.8–3.5 Å), and weak (1.8–5.0 Å) distance ranges. Backbone torsion angle restraints were obtained from the values calculated by the TALOS-N program. χ^1 angle restraints were included for residues for which TALOS-N indicated that a single rotamer was favored. Twenty structures were calculated in this manner for wild type RAP80-tUIM and ΔE81-RAP80-tUIM. Structures and restraints have been deposited in the Protein Data Bank under accession codes 2MKG and 2MKF, and BMRB codes 19774 and 19773, respectively.

RAP80 UIM a-helix stability measurements

The temperature dependencies of $^1H_{\alpha}$ chemical shifts for ΔE81-RAP80-tUIM were measured by collecting 2D $^1H\text{-}^{13}C$ HSQC spectra from 5–50°C, in steps of 5°C. The NMR sample was prepared identically to that used for main chain chemical shift assignment at 5°C, as described above. Due to extensive overlap for $^1H_{\alpha}$ resonances, per residue chemical shifts were obtained by deconvolution of the spectra through fitting to a lineshape function defined as the sum of one to four Gaussian curves, depending on the degree of overlap, using *Mathematica*[25]. Spectra were converted to tabular format using the *pipe2txt.tcl* program implemented in NMRPipe[26], and this text file was subsequently imported into *Mathematica* for deconvolution. Regions surrounding peaks of interest were defined, and chemical shifts and intensities within these regions were extracted. 1D ^1H projections were obtained by summing $^{13}C_{\alpha}$ planes together; the projection was subsequently fit to the deconvolution function. Fitted parameters included

linewidths, intensity scaling factors, and chemical shifts for the Gaussian peaks. In order to obtain an estimate of the chemical shift error, the dimensions of the initial selection box were varied within reasonable bounds and the procedure repeated; the chemical shifts thus obtained were averaged, and the standard deviation was taken to be the error. The resonances analyzed in this manner were F85, L89, and A113. Chemical shift overlap for most other peaks was too severe for effective deconvolution. In the case of F85 at 50°C, an overlapping water artifact precluded chemical shift determination, thus, this point was not included in the subsequent analyses. In order to facilitate direct comparison of the helical stabilities of wild type and ΔE81-RAP80-tUIM, we reanalyzed our previously measured $^1H_\alpha$ data[8] using this deconvolution methodology. The chemical shift changes were fit to a two state cooperative helix-coil transition, as previously described [8,27]. Errors in the fitted values of T_m and ΔT, the midpoint and width of transition respectively, were obtained from Monte Carlo simulations using the errors determined during the chemical shift deconvolutions.

Chemical shift titrations for ΔE81-RAP80-tUIM with Ub and tandem Ub$_2$

2D ^1H-^{15}N HSQC NMR spectra[28] for titrations of [U-^{15}N]- ΔE81-RAP80-tUIM with unlabeled human recombinant ubiquitin (Boston Biochem) and of [U-^{13}C,^{15}N]- ΔE81-RAP80-tUIM with tandem Ub$_2$ were collected at 25°C and 600 MHz. Protein concentrations for the stock solutions of ΔE81-RAP80-tUIM, Ub, and Ub$_2$ used in NMR-monitored titrations were determined using amino acid analyses. Titrations were conducted by diluting labeled ΔE81-RAP80-tUIM with aliquots from stock solutions of Ub or Ub$_2$, and following chemical shift changes, as previously described[8]. The concentrations of [U-^{15}N]- ΔE81-RAP80-tUIM were 215, 209, 201, 188, 174, 160, 146, 133, 119, 105, and 53 μM; the concentrations of Ub were 0, 90, 221, 443, 666, 889, 1112, 1334, 1557, 1780, and 890 μM, corresponding to RAP80:Ub ratios of 0, 0.4, 1.1, 2.4, 3.8, 5.5, 7.6, 10.0, 13.1, 16.9, and 16.9. The decrease in protein concentration during the titration allows for more extensive sampling of the binding isotherm resulting in an increase for the precision of the fitted K_D value, as previously described[29]. For the Ub$_2$ titration, concentrations of [U-^{13}C,^{15}N]- ΔE81-RAP80-tUIM were 206, 203, 196, 190,

182, 160, 136, and 121 μM, and the corresponding concentrations of Ub_2 were 0, 40, 131, 215, 310, 596, 908, and 1096 μM, corresponding to RAP80:Ub_2 ratios of 0, 0.2, 0.7, 1.1, 1.7, 3.7, 6.7, and 9.1. For some residues, overlap prevented chemical shift determination for certain titration points; these were not included in the fitting procedure to extract binding constants. Determination of the dissociation constants for binding of Ub to the N and C-terminal UIMs, $K_{D,N}$ and $K_{D,C}$, respectively, was achieved using a global fit of the chemical shift changes of all significantly shifting residues to a binding model describing the interaction of Ub with the N-terminal UIM (N-UIM), the C-terminal UIM (C-UIM), or two Ub molecules simultaneously bound to both UIMs (8):

$$K_{D,N} = \frac{[P][L]}{[PL_N]} \quad K_{D,C} = \frac{[P][L]}{[PL_C]}$$

$$[PL_{N,C}] = \frac{[PL_N][L]}{(K_{D,N} + K_{D,C})} + \frac{[PL_C][L]}{(K_{D,N} + K_{D,C})}$$

[1]

where,

$[P]$ is the concentration of free RAP80

$[L]$ is the concentration of free Ub

$[PL_N]$ is the concentration of RAP80 with Ub bound to the N-UIM

$[PL_C]$ is the concentration of RAP with Ub bound to the C-UIM

$[PL_{N,C}]$ is the concentration of RAP80 with Ub bound to both the N- and C-UIM, with $K_{D,N}$ and $K_{D,C}$ as described above.

The total protein concentration is given by $[P] + [PL_N] + [PL_C] + [PL_{N,C}]$

The total ligand concentration is given by $[L] + [PL_N] + [PL_C] + 2[PL_{N,C}]$.

Chemical shift changes for residues from the N-UIM were taken to reflect either Ub binding to this UIM alone, or with a second Ub bound to the neighboring UIM, but not Ub bound only to the C-UIM; a corollary approach was also used to assess Ub binding for residues belonging to the C-UIM. Unlike binding of Ub to wild-type RAP80-tUIM as characterized in our previous work[8], the deletion of E81 abrogates Ub binding to the N-UIM. This is manifested as drastically reduced $\Delta\delta_{obs}$ values for residues of the N-UIM, the result of greatly reduced levels of saturation. In order to achieve an estimate of this

weakened affinity, it was necessary to restrain $\Delta\delta_{max}$ values for some residues of the N-UIM during the global fit as follows: the free chemical shifts and their direction of movement upon titration with Ub matched those observed for wild type RAP80-tUIM, thus, $\Delta\delta_{max}$ values for K90 and S92 were used as the corresponding $\Delta\delta_{max}$ values for ΔE81-RAP80-tUIM. The $K_{D,N}$ and $K_{D,C}$ values from fits using these constraints individually were then averaged to give the reported values. The dissociation constants ($K_{D,N}$ and $K_{D,C}$) for binding of Ub_2 to ΔE81-RAP80-tUIM were obtained in an analogous manner, except with the substitution of a binding model for Ub_2 rather than Ub, as follows:

$$K_{D,N} = \frac{[P][L]}{[P_N L_N]} \quad K_{D,N} = \frac{[P][L]}{[P_N L_C]}$$

$$K_{D,C} = \frac{[P][L]}{[P_C L_C]} \quad K_{D,C} = \frac{[P][L]}{[P_C L_N]}$$

[2]

$$[P_{N,C}L_N L_C] = \frac{[P_N L_N][L]}{(K_{D,N} + K_{D,C})} + \frac{[P_C L_C][L]}{(K_{D,N} + K_{D,C})}$$

$$[P_N P_C L_{N,C}] = \frac{[P_N L_N][P]}{(K_{D,N} + K_{D,C})} + \frac{[P_C L_C][P]}{(K_{D,N} + K_{D,C})}$$

where,

$[P]$ is the concentration of free RAP80

$[L]$ is the concentration of free Ub_2

$[P_N L_N]$ is the concentration of RAP80 with the N-Ub from Ub_2 bound to the N-UIM,

$[P_N L_C]$ is the concentration of RAP80 with the C-Ub from Ub_2 bound to the N-UIM,

$[P_C L_C]$ is the concentration of RAP80 with the C-Ub from Ub_2 bound to the C-UIM,

$[P_C L_N]$ is the concentration of RAP80 with the N-Ub of Ub_2 bound to the C-UIM,

$[P_N P_C L_{N,C}]$ is the concentration of one Ub_2 molecule with the N-UIM from one RAP80 molecule bound to the N-Ub and the C-UIM from a second RAP80 bound to the C-Ub from Ub_2

41

$[P_{N,C}L_NL_C]$ is the concentration of one RAP80 molecule with the N-UIM bound to the N-Ub from one Ub_2 and the C-UIM bound to the C-Ub from another Ub_2. $K_{D,N}$ and $K_{D,C}$ are as described above.

The total protein concentration is given by $[P] + [P_NL_N] + [P_NL_C] + [P_CL_N] + [P_CL_C] + [P_{N,C}L_NL_C] + 2[P_NP_CL_{N,C}]$

The total ligand concentration is given by $[L] + [P_NL_N] + [P_CL_C] + [P_CL_N] + [P_NL_C] + 2[P_{N,C}L_NL_C] + [P_NP_CL_{N,C}]$.

Δd_{max} values for residues (K90, and S92) of the N-UIM were likewise individually constrained at their wild type RAP80-tUIM values in order to achieve a reasonable fit, and the fitted $K_{D,N}$ and $K_{D,C}$ values were again averaged. Chemical shift errors were determined using Monte Carlo methods; where possible, isolated peaks used in the K_D determination were extracted from the spectra, noise was added to each point in a random fashion (based on the overall level of spectral noise, obtained from NMRDraw [26], and the chemical shift was determined using a parabolic fit, as described in reference[29]. This procedure was repeated 10000 times, and the standard deviation of the chemical shift was determined for the ensemble of values. When it was not possible to extract an isolated peak, the error was estimated as the median of the other chemical shift errors calculated for the spectrum. However, uncertainty in the stock protein concentrations is typically much larger than that for chemical shifts, and is the main source of error in the determination of K_D values from NMR titrations. The stock protein concentration errors were taken to be 10%[29,30]. Subsequently, both chemical shift and protein concentration errors were used to estimate errors for the fitted dissociation constants, using 250 Monte Carlo trials.

Stability calculations for the Ub–RAP80 N-UIM interaction

MD simulations were conducted to estimate the free energy of interaction for wild type or $\Delta E81$ RAP80 N-UIM with Ub using the AMBER 11/12 suites of biomolecular simulation programs[31]. A starting model for the interaction between the N-UIM from RAP80 and Ub was generated from the structure of K63-Ub_2 bound to wild type RAP80 (pdb id: 3A1Q)[9]. Given that the crystallographic structure lacks RAP80 residues N-

terminal to E81, and contains a GPLGS cloning artifact instead, these five residues were removed and replaced with the residues M79–T80, whose main chain and side chain dihedral angles were set at standard N-cap values. This wild type model was then used to create a starting model for the interaction between ΔE81-RAP80 N-UIM and Ub. However, M79 in the ΔE81 mutant cannot be accommodated in the N-cap conformation due to steric clashes with Ub. Thus, the main chain and dihedral angles for M79 and T80 were set at standard a-helical values, the side chain of T80 was set at $\chi^1 \sim \pm 180°$ to form a favorable electrostatic interaction with R42 from Ub, and the side chain from M79 was set at $\chi^1 \sim +60°$ to form favorable van der Waals interactions with L73 from Ub. The starting models were then energy minimized using a generalized Born implicit solvent model[32]. Using these initial models, MD simulations were conducted with the *ff*99SBNMR forcefield[33], the TIP3P water model, and a particle mesh Ewald approach (PMEMD) with default parameters was used for calculating long-range electrostatics, as implemented in CUDA[34]. SHAKE was used to restrain covalent bonds to hydrogen, temperature was regulated using a Berensden thermostat, and pairwise non-bonded and electrostatic interactions were cutoff at 8 Å. The initial structural models were solvated in a truncated octahedral water box with a distance of 24 Å between protein atoms and their images in adjacent unit cells, and the systems were neutralized with Na^+ ions. The systems were heated over 50 ps to 298 K with 2 kcal/mol restraints on solute atoms, and equilibrated to 1 atm pressure for an additional 50 ps. Production dynamics were conducted for ~16 ns. After discarding the first ~4 ns, the free energies of binding were determined from the average over 12 snapshots (1 per ns), using the MMPBSA python scripts in AMBER 11 with the *ff*99SBNMR forcefield[33], a Poisson-Boltzmann implicit solvent model, and entropy refinement using the normal mode approximation[35].

Results

NMR spectra for ΔE81-RAP80-tUIM

The superposition of ^1H–^{15}N HSQC NMR spectra for wild type and ΔE81-RAP80-tUIM is shown in Fig. 2.1. The spectra reveal large chemical shift changes near the site of the

deletion mutation, with smaller or insignificant chemical shift changes in the remaining N- and C-terminal UIM domains.
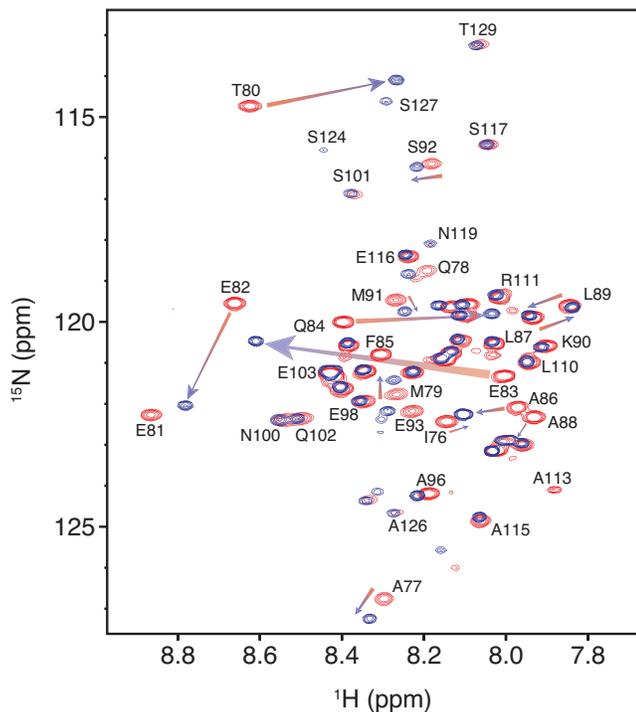
Figure 2.1: 2D ^{1}H–^{15}N HSQC NMR spectra for wild type (red) and ΔE81 (blue) RAP80-tUIM.

NMR monitored titrations for the interaction of ΔE81-RAP80-tUIM with Ub and tandem Ub$_2$

The interaction between ΔE81-RAP80-tUIM and mono or tandem Ub2 chains was analyzed using NMR-monitored titrations. [U-15N]- ΔE81-RAP80-tUIM was titrated with unlabeled Ub (Fig. 2.2). For a Ub: ΔE81-RAP80-tUIM ratio of 17:1, substantial chemical shift changes are localized to the C-UIM (Fig. 2.2a). Representative chemical shift changes for residue E82 during titration with monoUb are shown in Fig. 2.2b. The

changes are linear, small in magnitude, and not accompanied by substantial line broadening, indicative of weak 1:1 binding with fast kinetics ($k_{off} > 10000$ s^{-1}). For the C-UIM, representative chemical shift changes for A115 upon titration with Ub are linear and large in magnitude, indicating a 1:1 binding interaction that is substantially stronger than that for the N-UIM (Fig. 2.2c). As in the case of the N-UIM, lack of substantial line broadening is indicative of fast binding kinetics. The NMR-monitored titration was analyzed using a binding model wherein the N- and C-UIMs bind Ub independently with separate dissociation constants, $K_{D,N}$ and $K_{D,C}$, respectively (Fig. 2.2d and eq. 1). Representative fits of the chemical shift changes for residue S117 from the C-UIM are shown in Fig. 2.2e, and remain similar to wild type, with $K_{D,C} = 590 \pm 80$ μM. In contrast, Ub binding to the N-UIM is significantly impaired with $K_{D,N}$ increasing ~20 fold to 24 ± 8 mM (Fig. 2.2f). The binding of ΔE81-RAP80-tUIM to tandem Ub$_2$ chains was assessed by monitoring chemical shift changes for [U-^{13}C,^{15}N]- ΔE81-RAP80-tUIM upon titration with unlabeled Ub$_2$ (Fig. 2.3). As in the case of binding to mono Ub, large chemical shift changes are confined to the C-UIM for ΔE81-RAP80-tUIM:Ub$_2$ ratios of 1 to 9 (Fig. 2.3a). Representative spectral changes for both UIMs (Figs. 2.3b,c) are similar to the results for titration with mono Ub. The chemical shift changes within the N-UIM are linear, small, and consistent with fast kinetics, whereas those for the C-UIM are linear, but large in magnitude, and indicative of fast binding kinetics. The titration was analyzed using a binding model lacking multivalent effects, wherein the different Ub molecules in tandem Ub$_2$ bound independently to the individual UIMs in ΔE81-RAP80-tUIM with separate dissociation constants, $K_{D,N}$ and $K_{D,C}$ (Fig. 2.3d and eq. 2). Representative fits of the chemical shift changes upon addition of Ub$_2$ to [U-^{13}C,^{15}N -ΔE81-RAP80-tUIM are shown in Fig. 2.3e,f. For Ub$_2$ binding, $K_{D,N}$ is 8 ± 2 mM, modestly stronger than binding of Ub, indicative of severe impairment in Ub recognition in comparison to wild type. The value for $K_{D,C}$ remains the same as that observed for binding of Ub, 700 ± 122 μM. The slightly enhanced affinity observed for $K_{D,N}$ for Ub$_2$ in comparison to Ub binding likely reflects the presence of some residual multivalency when the C-terminal UIM is bound.
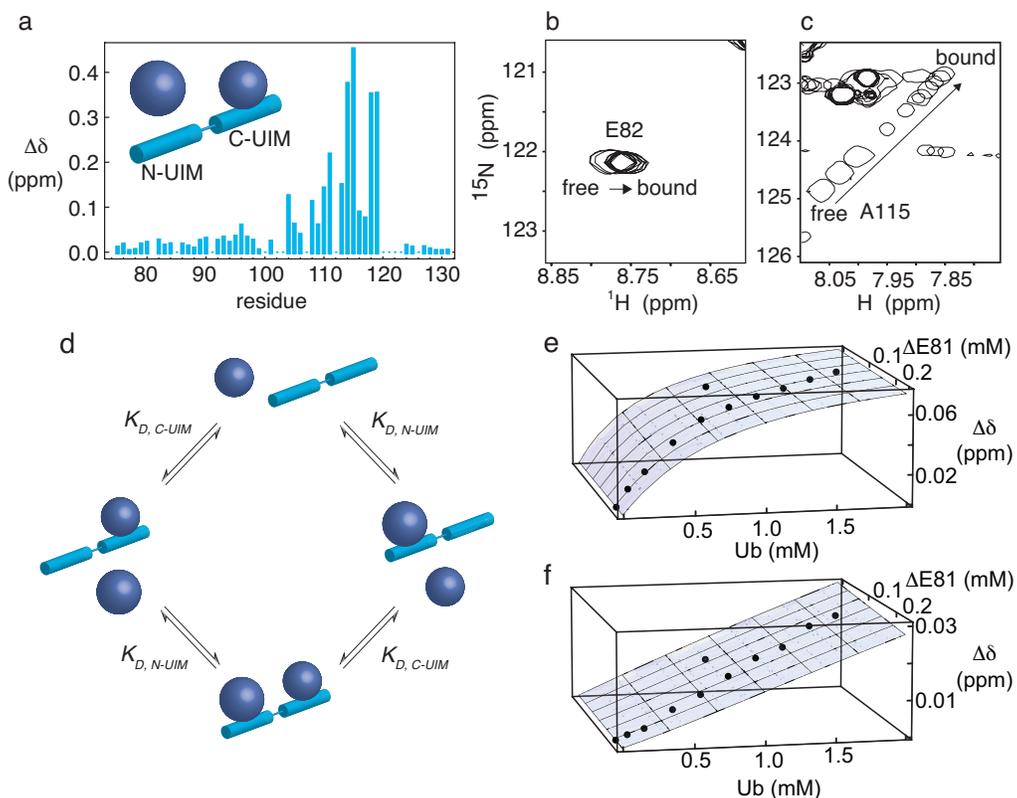
Figure 2.2: Titration of ΔE81-RAP80-tUIM with Ub. **a),** Maximum chemical shift changes for the interaction of Ub with ΔE81-RAP80-tUIM. Representative regions from 2D ^{1}H–^{15}N HSQC NMR spectra from the N-UIM **b)** and C-UIM **c),** from ΔE81-RAP80-tUIM titrated with unlabeled Ub. **d),** Binding model used to analyze the interaction of ΔE81-RAP80 with Ub (eq. 1). **e),** Chemical shift changes for S117 within the C-UIM are indicated on the vertical axis, and concentrations of ΔE81-RAP80-tUIM and unlabeled Ub titrant are indicated on the horizontal axes. Experimentally determined chemical shift changes are shown as points, and the best fits to the binding isotherms are shown as surfaces. **f),** Chemical shift changes for K90 within the N-UIM, additional details as in (e).
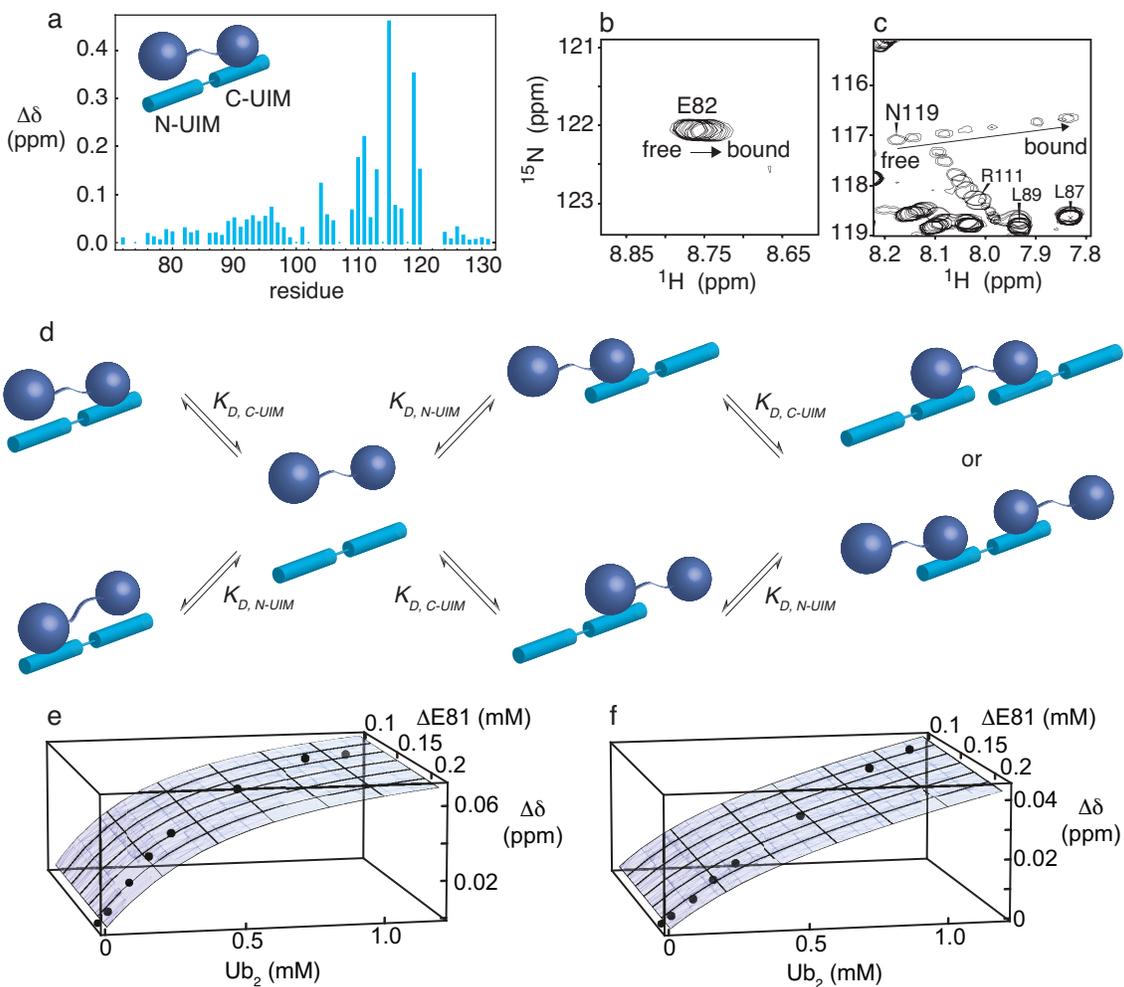
Figure 2.3: Titration of ΔE81-RAP80-tUIM with tandem Ub₂. **a),** Maximum chemical shift changes for the interaction of Ub₂ with ΔE81-RAP80-tUIM. Representative regions from 2D ¹H–¹⁵N HSQC NMR spectra from the N-UIM **b),** and C-UIM **c),** from ΔE81-RAP80-tUIM titrated with unlabeled Ub₂. **d),** Binding model used to analyze the interaction of ΔE81-RAP80 with Ub₂ (eq. 2). **e),** Fits of chemical shift perturbation data to binding models for RAP80-tUIM binding to tandem Ub₂. Chemical shift changes for S117 within the C-UIM are indicated on the vertical axis, and concentrations of ΔE81-RAP80-tUIM and unlabeled Ub titrant are indicated on the horizontal axes. Experimentally determined chemical shift changes are shown as points, and the best fits to the binding isotherms are shown as surfaces. **f),** Chemical shift changes for K90 within the N-UIM, additional details as in (e).

NMR chemical shifts indicate that the N-UIM helix N-cap is maintained in ΔE81-RAP80-tUIM

The N-cap is a common structural motif that stabilizes the N-termini of α-helices. The N-cap motif possesses an NMR signature that is defined by an upfield shift for the random coil $^{13}C_\alpha$ value of the capping residue, and a structural signature that includes main chain ϕ and ψ dihedral angles of ~ –94° and +167°, respectively [36,37]. The $^{13}C_\alpha$ chemical shift values for the N-terminal residues from the N-UIM show the characteristic signature for a helix N-cap for both wild type and ΔE81-RAP80-tUIM (Fig. 2.4a). Using $^1H^N$, $^1H_\alpha$, $^{13}C_\alpha$, $^{13}C_\beta$, ^{13}CO (wild type), and ^{15}N chemical shifts, quantitative values for the main chain ϕ and ψ angles were derived from the TALOS-N program, and helix capping motifs were predicted from the MICS program. The calculated main chain dihedral angles indicate the presence of an N-UIM helix N-cap for both wild type and ΔE81-RAP80-tUIM (Table 2.1 and Supplemental Tables 2.4 and 2.5). In addition, residue T80 is predicted to be in an N-cap conformation with 90% and 80% probability for wild type and ΔE81-RAP80-tUIM, respectively (Supplemental Tables 2.6 and 2.7).

Residue	wild type		ΔE81	
	ϕ (°)	ψ (°)	ϕ (°)	ψ (°)
M79	–77 ± 11	141 ± 10	–71 ± 8	141 ± 10
T80	–72 ± 5	162 ± 4	–74 ± 6	161 ± 6
E81	–58 ± 4	–40 ± 4	–	–
E82	–67 ± 4	–40 ± 6	–59 ± 4	–37 ± 6

Table 2.1: Main chain dihedral angles for the N-cap motif in wild type and ΔE81-RAP80 N-UIM α-helices from NMR chemical shifts.

NMR structures for wild type and ΔE81-RAP80-tUIM

In addition to identifying a helix N-cap for the N-UIM, the quantitative values for the main chain ϕ and ψ angles from the TALOS-N program indicate that the structure of the N and C-UIMs from wild type and ΔE81-RAP80-tUIM are α-helical. Additionally, the TALOS-N program predicts that a number of side chains within the helices adopt the expected χ^1 angles (typically –60° or ±180° for Val, Supplemental Tables 2.8 and 2.9) that avoid steric clashes between side chain and main chain atoms. Structural statistics for the wild type and ΔE81 mutant structures calculated on the basis of NOE and TALOS-derived dihedral angle restraints are given in Table 2.2. Importantly, the χ^1 angle for T80 for wild type and ΔE81-RAP80-tUIM predominantly adopts the expected +60° rotamer for the N-cap conformation, allowing the side chain hydroxyl to hydrogen bond with the main chain amide of the N+3 residue, or E83 and Q84 for wild type and ΔE81, respectively. The TALOS main chain dihedral angles, and distance restraints derived from ^{15}N NOESY HSQC NMR spectra, were used to calculate structures for the N-terminal UIM for wild type and ΔE81-RAP80 (Fig. 2.4b-d). The NMR-derived structures indicate that the N-terminus from the N-UIM for ΔE81-RAP80 undergoes a structural frameshift, wherein E81 in wild type is replaced by T80 in the ΔE81 mutant.

	wild type RAP80-tUIM	ΔE81 RAP80-tUIM
distance restraints		
total	65	120
intraresidue	24	56
sequential ($|i–j| = 1$)	40	57
medium ($2 \leq |i–j| \leq 4$)	1	7
long ($|i–j| \geq 5$)	0	0
dihedral restraints	57 ϕ, 57 ψ, 15 χ^1	56 ϕ, 56 ψ, 14 χ^1
restraints violation		
distance of > 0.5 Å	0	5
dihedral of > 5 °	8	8
ϕ/ψ in the most-favored region (%)	94.2	93.7

Table 2.2 - Structural statistics for twenty NMR-derived structures for wild type and ΔE81-RAP80-tUIM.
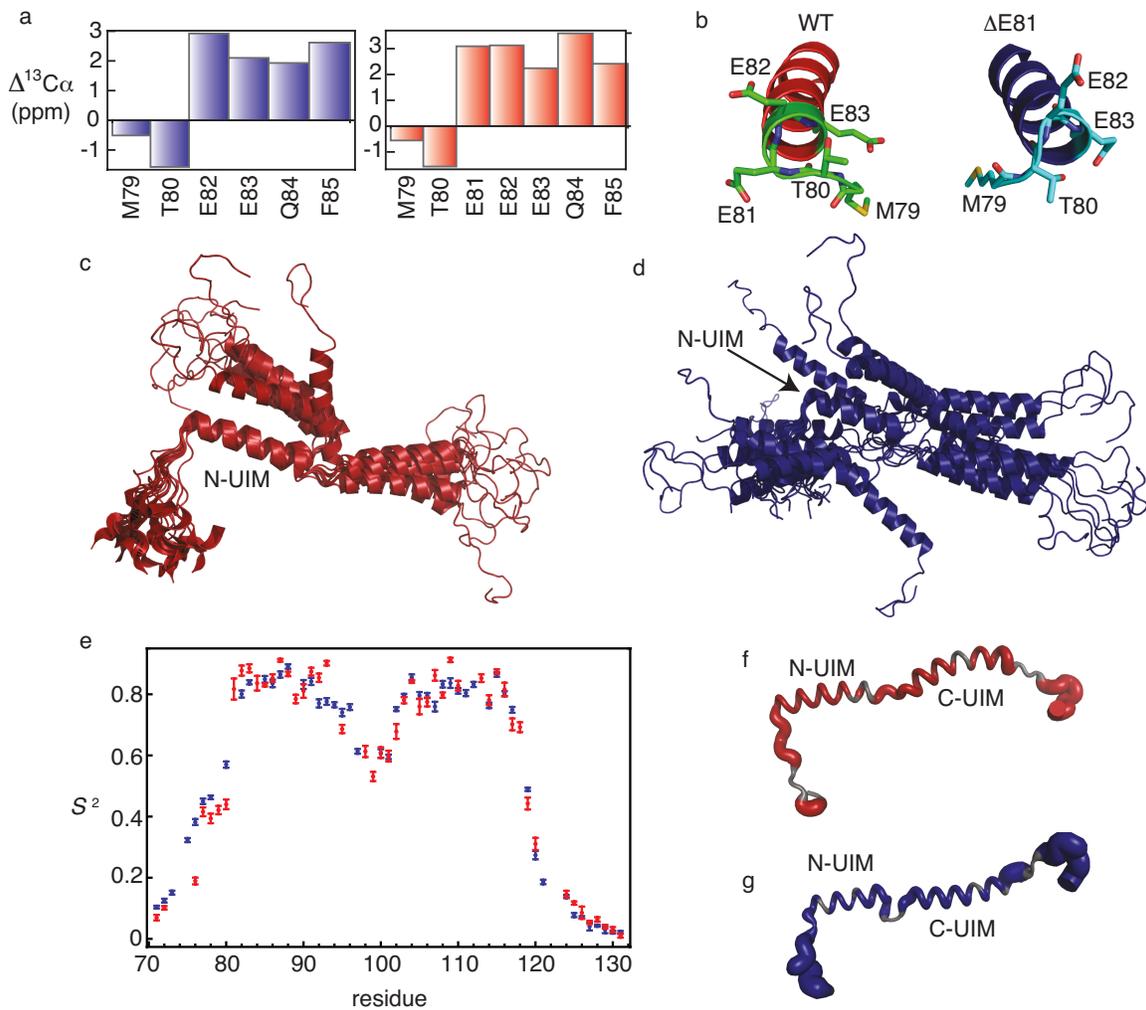
Figure 2.4: (a), Characteristic upfield $^{13}C_a$ chemical shift changes for the helix N-cap in ΔE81-RAP80-tUIM (blue, left panel) and wild type (red, right panel). **(b)**, Representative NMR structures for the N-UIM from wild type and ΔE81-RAP80, wild type is shown with main chain atoms in the cartoon representation and colored red, and ΔE81 is shown in blue, key residues near the N-cap are shown in the stick representation. Ensemble of twenty NMR structures for the N- and C-UIMs from **(c)**, wild type and **(d)**, ΔE81-RAP80. The structures are superimposed on the a-helix of the N-UIM, wild type is shown with main chain atoms in the cartoon representation colored red, ΔE81 is shown in blue. **(e)**, Main chain order parameters (S^2) for wild type (red) and ΔE81-RAP80-tUIM (blue) from ^{15}N NMR relaxation data at 600 MHz and 5° C. The relative flexibility is indicated by $1-S^2$, and shown as increasing width of the main chain for wild type **(f)**, and ΔE81-RAP80-tUIM **(g)**. Residues lacking relaxation data are colored grey, and the structures correspond to those closest to the average structure from the respective ensembles.

Main chain dynamics for wild type and ΔE81-RAP80-tUIM from ^{15}N relaxation measurements

$^{15}N–R_1$, $–R_2$, and {^1H}–^{15}N NOE NMR relaxation measurements at 5°C (data not shown) indicate that in general, the UIM α-helices (N-UIM residues 81–95; C-UIM residues ~108–117) for wild type and ΔE81-RAP80-tUIM are as rigid as well-defined regions of secondary structure in globular proteins with order parameter (S^2) values slightly in excess of 0.8 (Fig. 2.4e-g). The extreme N- and C-termini of wild type and ΔE81-RAP80-tUIM, residues 72–79 and 120–131, respectively, are highly dynamic with S^2 values lower than ~0.5 (Fig. 2.4e-g). Additionally, the linker region (residues 96 – 102) between the N- and C-UIMs is also flexible, though not to the same extent as the termini, with S^2 values between ~0.5 – 0.7. Similarly, residue T80 that forms the N-cap within the N-UIM for wild type and ΔE81-RAP80 is also flexible, with S^2 values of ~0.5 (Fig. 2.4e). Commensurate with the increased flexibility evident in S^2 values, the extreme termini and linker region also show slow timescale motions, characterized by the time constant t_S, in the nanosecond range (data not shown). Residues E81 (wild type) or E82 (ΔE81) immediately adjacent to the capping threonine residue are rigid with S^2 values in excess of 0.8 (Fig. 2.4e).

RAP80 UIM α-helix stability

The temperature dependence of $^1H_α$ chemical shifts indicate that, in general, the N-UIM and C-UIM α-helices retain stability for ΔE81 in comparison to wild type RAP80, with ~70% and >90% helical content at the centers of the α-helices at 25 and 5°C, respectively (residue 89, N-UIM; residue 113, C-UIM) for wild type and ΔE81 (Table 2.3 and Fig. 2.5). However, near the site of the deletion mutation (residue 85), there is a decrease in stability, with $\Delta T_m = –13\pm6$ K, which corresponds to ~10 and 20% less helical content at 5 and 25°C, respectively, for residue F85 of ΔE81 near the N-terminus of the N-UIM in comparison to wild type (Fig. 2.5). This modest destabilization at the N-terminus is consistent with results from quantitative chemical shift analysis at 5°C using the TALOS-N program that indicates T80 has ~10% lower probability of adopting the N-cap conformation for ΔE81 in comparison to wild type. Assuming that Ub binds only the α-

helical state of the UIM (conformational selection), the experimentally determined dissociation constants are larger than expected. Using eqs. 3 and 4 in reference (8), the dissociation constant for binding to a purely a-helical state can be derived from the experimental values of $K_{D,N}$ and $K_{D,C}$ using:

$$K_D = K_D^{exp} \times f_H \qquad [3]$$

where K_D^{exp} is the experimental, or apparent, K_D value for the N- or C-UIM ($K_{D,N}$ or $K_{D,C}$), and f_H is the fraction of α-helix.

Residue	wild type		ΔE81	
	T_m (K)	ΔT (K)	T_m (K)	ΔT (K)
85	309 ± 4	46 ± 10	296 ± 4	41 ± 9
89	302 ± 2	15 ± 3	305 ± 3	20 ± 4
113	316 ± 6	21 ± 4	312 ± 5	17 ± 4

Table 2.3: Stability of wild type and ΔE81-RAP80 α-helices from $^1H_\alpha$ NMR monitored temperature titrations.
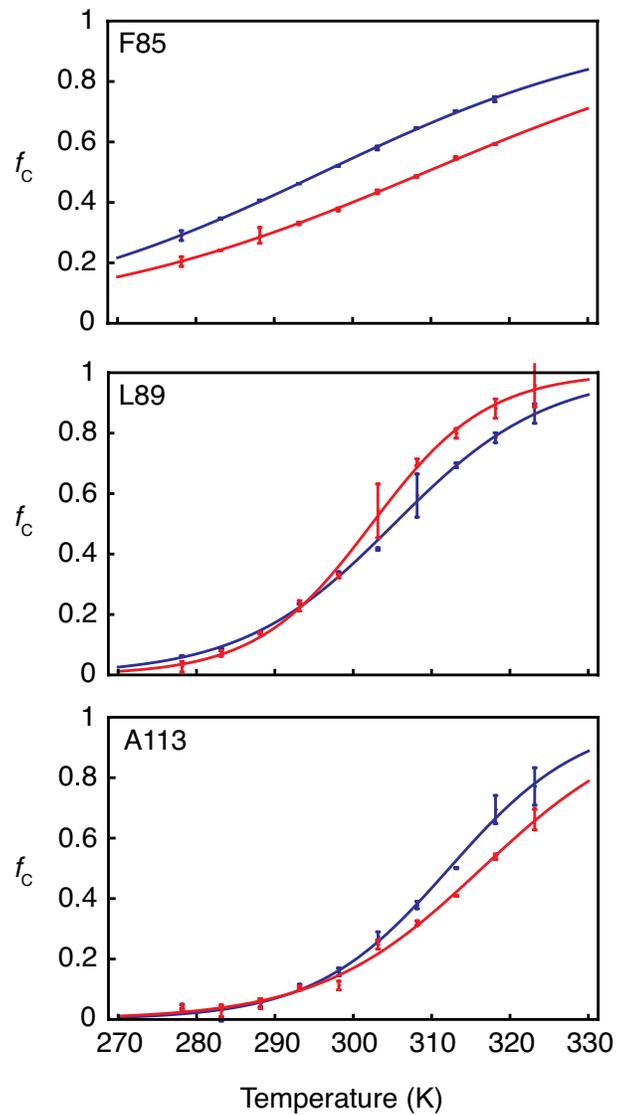
Figure 2.5: Temperature dependence for $^1H_\alpha$ chemical shifts for residues F85, L89, and A113 in wild type (red) and ΔE81-RAP80-tUIM (blue). f_c indicates fraction in the random coil conformation.

Stability calculations for the Ub–RAP80 N-UIM interaction

From the MD simulations for the complex between wild type RAP80 and the N-UIM and Ub (Fig. 2.6a), MMPBSA calculations using AMBER 11 give a favorable free energy of interaction for wild type RAP80 of -19 ± 4 kcal/mol. The structural frameshift for the ΔE81-RAP80 N-UIM causes residues M79 and T80 to rotate 90° about the helix axis. Thus, for the N-UIM of ΔE81 in the N-cap conformation, the side chain from M79 will sterically clash with residues 70–73 from Ub at the N-UIM-Ub interface. However, when residues M79 and T80 adopt a-helical conformations, the side chain from M79 at the N-terminus of the N-UIM can be accommodated. Using the MD simulations as reasonable models for the ΔE81 N-UIM–Ub interaction gives a free energy of interaction of -9 ± 4 kcal/mol, leading to an unfavorable change in the free energy of interaction for the mutant with $\Delta\Delta$G (mutant–wildtype) of $+10\pm5$ kcal/mol. The pairwise energy decomposition with respect to van der Waals, electrostatic, and polar solvation terms is shown in Fig. 2.6b, and demonstrates the unfavorable changes in the energetics of binding between the C-terminus of Ub and the N-terminus of the N-UIM as a result of the ΔE81 mutation. To facilitate comparison to the experimental values for $K_{D,N}$, the experimental free energy of binding was calculated according to:

$$\Delta G = RT \ln\left(K_D\right) \qquad [4]$$

where R is the gas constant (1.987×10^{-3} kcal K^{-1} mol^{-1}), T is temperature (K), and K_D (mol L^{-1}) is given by eq. 4, thus including the effects of conformational selection.
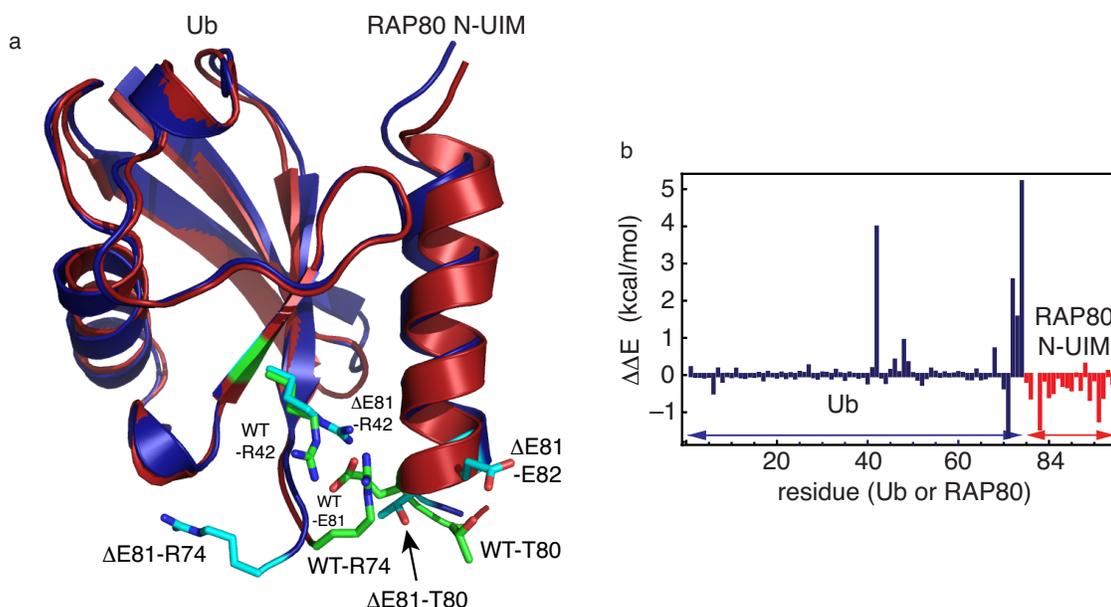
Figure 2.6: (a) Structural consequences of the ΔE81-RAP80 N-UIM N-cap frameshift for the interaction of the N-UIM with Ub. A representative snapshot from the MD simulation for the wild type RAP80 N-UIM–Ub interaction is shown with main chain atoms in the cartoon representation colored red, and one for the ΔE81-RAP80 N-UIM–Ub interaction is shown in blue; labeled residues colored green and in the stick representation correspond to wild type, those colored cyan and in the stick representation correspond to ΔE81. (b) Differences in the per residue pairwise energy decomposition for the interaction of ΔE81 and wild type RAP80-tUIM with Ub (ΔΔE = ΔE81– wild type).

Discussion

2D ^{1}H–^{15}N NMR spectra for wild type and ΔE81-RAP80-tUIM indicate that the overall α-helical structure of the UIM domains remains intact (Fig. 2.1). As expected, the crosspeak for E81 vanishes, and is essentially replaced by E82 in the deletion mutant. Interestingly, the crosspeaks for E83 and Q84 for ΔE81 RAP80 shift to the approximate locations for residues E82 and E83, respectively, in the wild type protein. Residue T80, adjacent to the deletion site, also experiences a large chemical shift change. These main chain amide chemical shift changes indicate that the structure of the C-UIM remains unchanged, whereas the N-UIM undergoes a structural frameshift, maintaining helical secondary structure. Consistent with these results, main chain and side chain dihedral

56

angles derived from NMR chemical shifts, show that the N-cap structure is maintained in ΔE81-RAP80-tUIM through the structural frameshift, with E82 in ΔE81-RAP80-tUIM taking the role of E81 in wild type RAP80-tUIM. The structural implications for the frameshift are evident in the NMR structures for the N-cap in ΔE81 and wild type RAP80-tUIM (Fig. 2.4b). For ΔE81-RAP80-tUIM, residues M79 and T80 translate ~1.5 Å along the N-UIM helix axis towards the C-terminus, and rotate about the helix axis ~90°. In the context of the interaction between Ub and the N-UIM, the structural frameshift causes E81 in wild type RAP80 to be replaced by T80 in the ΔE81 mutant, and the N-cap can no longer be accommodated at the Ub-UIM interface due to steric clashes with the side chain of M79. However, if residues M79 and T80 adopt an α-helical conformation, the side chain of M79 moves away from the interface, which can facilitate binding, but favorable electrostatic interactions between E81 from RAP80 and residues R42, R72, L73, and R74 from Ub are disrupted, and replaced by interactions between the shorter side chain of T80 and R42, R72, L73, and R74 from Ub (Fig. 2.6a). It is likely that the α-helical conformation at residues M79 and T80 can be adopted to some extent, given the flexibility at these residues observed through ^{15}N NMR S^2 values (Fig. 2.4e-g) as well as random coil index chemical shift S^2 values[38], determined through the TALOS-N program (Supplemental Tables 2.4 and 2.5). NMR-monitored ^{1}H–^{15}N chemical shift titrations of Ub into wild type and ΔE81 RAP80-tUIM reveal that loss of a single N-terminal glutamic acid residue results in near abolishment of Ub binding for the N-UIM. The twenty-fold increase in $K_{D,N}$ for binding of Ub to ΔE81 RAP80 N-UIM corresponds to a 2.2 kcal/mol loss in binding affinity (eqs. 3 and 4). Similarly, the NMR titration results for binding of Ub$_2$ to ΔE81-RAP80 show a substantial impairment in Ub recognition for the N-UIM, with a twelve-fold increase in $K_{D,N}$, or a loss of 1.5 kcal/mol in binding affinity. Taken together, the results for Ub and Ub$_2$ binding to ΔE81-RAP80-tUIM indicate that this single residue mutation leads to abolishment of multivalent binding in comparison to wild type RAP80-tUIM. Interestingly, stability calculations from MD simulations for the free energy of interaction between wild type or ΔE81-RAP80 N-UIM and Ub are in agreement with the experimental values determined from NMR, with the calculated ΔΔG of +10±5 kcal/mol comparing favorably with the experimental ΔΔG of +2 kcal/mol. Importantly, these calculations suggest that disruption

of the interaction between RAP80 and polyubiquitin is mainly due to the loss of key electrostatic interactions between E81 within the N-UIM and residues R42, and 72–74 from the C-terminus of Ub (Fig. 2.6). In further support of a predominantly structural basis for the defective function of ΔE81-RAP80, the stabilities of the α-helices for the N- and C-UIMs of the deletion mutant remain similar to wild type, with ~70% α-helical content at 25°C. A modest destabilization near the N-terminus of the N-UIM is observed as a 20% loss in helical structure for residue F85 (Table 2.3). Using eqs. 3 and 4, this difference in helical content indicates that the destabilization results in a modest loss of 0.2 kcal/mol in binding energy. In other words, taking into account this loss in inherent helicity, and assuming $K_{D,N} = K_{D,C}$ [8], the $K_{D,N}$ for Ub binding to ΔE81 would be expected to increase from 590 to 767 mM, rather than to the experimentally measured value of 24 mM.

In light of intense efforts to investigate the putative association between defects within binding partners for BRCA1/BRAC2 tumor suppressors and a hereditary disposition to breast cancer, the results provide a compelling molecular basis for genomic instability resulting from a single amino acid deletion in RAP80. This defect impedes polyUb recognition, is likely responsible for aberrant targeting of BRCA1/BRCA2 to DNA damage foci, and is potentially involved in the pathogenesis of disease. Considering the broad, vital role that multivalent interactions play in human biology[39], and the variety of different UIM–Ub pairs[40,41], it is of interest that near complete abolishment of multivalent binding is implicated in the development of cancer. That is, binding of polyUb by ΔE81-RAP80 is not abolished; the C-UIM binds Ub normally, whereas abrogated binding of Ub by the N-UIM is responsible for the near total loss of multivalent recognition.

Supplemental Information

Supplemental Table 2.4: TALOS output for main chain ϕ and ψ dihedral angles for wild type RAP80-tUIM.

```
REMARK TALOS-N Protein Backbone Torsion Angle Prediction Table
REMARK  Prediction Summary for Chemical Shift Input wtRAP80UIM_5C_TALOS_withGLUstretch.tab
REMARK
REMARK  PHI is the predicted torsion angle C(i-1) N(i)  CA(i) C(i)   (degrees).
REMARK  PSI is the predicted torsion angle N(i)   CA(i) C(i)  N(i+1) (degrees).
REMARK
REMARK  DPHI and DPSI are the estimated standard deviations of the
REMARK  prediction errors in PHI and PSI (degrees).
REMARK
REMARK  DIST is the TALOS-N database matching score.
REMARK
REMARK  S2 is the Wishart RCI chemical shift order parameter,
REMARK  JACS, 127(43), 14970-14971.
REMARK
REMARK  COUNT is the number of database triplets used to form
REMARK  the torsion angle predictions.
REMARK
REMARK  CLASS is the classification of the prediction result:
REMARK    None: no torsion prediction was made.
REMARK
REMARK    Strong/Generous: majority consensus in database matches;
REMARK    prediction is likely to be good.
REMARK
REMARK    Warn: no consensus in database matches, do not use prediction.
REMARK
REMARK    Dyn:  RCI-S2 value indicates that residue has dynamic conformation.
REMARK
REMARK Reference:
REMARK  Y. Shen, and A. Bax:
REMARK  Protein backbone and sidechain torsion angles predicted from
REMARK  NMR chemical shifts using artificial neural networks
REMARK  J. Biomol. NMR (in press).
REMARK
REMARK  TALOS-N Version 4.01 Rev 2013.148.15.55 TALOSN_INFO
DATA FIRST_RESID 1
DATA SEQUENCE LGSRKIAQMTEEEQFALALKMSEQEAREVNSQEEEEEELLRKAIAESLNSCRPSD
DATA SEQUENCE ASATRS
VARS    RESID RESNAME PHI PSI DPHI DPSI DIST S2 COUNT CS_COUNT CLASS
FORMAT %4d %s %8.3f %8.3f %8.3f %8.3f %8.3f %5.3f %2d %2d %s
 -2 L 9999.000 9999.000   0.000    0.000   0.000 0.000  0  9 None
 -1 G   97.551   22.591   29.405   76.924   0.826 0.104  8  9 Dyn
  0 S  -67.246  149.052    5.702    9.140   0.428 0.132 25  6 Dyn
 74 R  -70.139  140.509    8.786   11.966   0.347 0.185 25  9 Dyn
 75 K  -68.049  144.031    8.803    8.234   0.302 0.226 25 15 Dyn
 76 I  -76.810  136.103   14.427    9.684   0.342 0.290 10 18 Dyn
 77 A  -78.189  151.141   18.929   22.450   0.424 0.335  8 18 Dyn
 78 Q  -70.269  138.680   13.834   12.566   0.344 0.367 25 18 Dyn
 79 M  -76.614  141.217   10.694    9.563   0.171 0.479 25 18 Dyn
 80 T  -72.125  161.755    5.201    4.298   0.089 0.639 25 18 Strong
 81 E  -58.191  -39.675    3.782    4.400   0.057 0.841 25 18 Strong
 82 E  -67.475  -40.324    4.319    5.546   0.058 0.893 25 18 Strong
 83 E  -67.412  -40.454    3.415    3.116   0.057 0.900 25 18 Strong
 84 Q  -64.776  -38.577    3.323    4.308   0.062 0.891 25 18 Strong
 85 F  -66.223  -40.569    3.666    5.043   0.057 0.872 25 18 Strong
 86 A  -64.275  -38.715    5.057    4.062   0.062 0.857 25 18 Strong
 87 L  -67.352  -41.999    5.328    5.927   0.063 0.832 25 18 Strong
 88 A  -66.822  -38.958    5.383    4.358   0.070 0.824 25 18 Strong
 89 L  -67.356  -38.841    4.283    4.617   0.080 0.809 25 18 Strong
 90 K  -68.934  -36.486    3.419    3.928   0.101 0.800 25 18 Strong
 91 M  -68.669  -36.115    4.449    6.149   0.126 0.782 25 18 Strong
 92 S  -69.330  -34.294    6.866    5.802   0.161 0.761 25 17 Strong
 93 E  -68.702  -37.908    3.187    5.204   0.203 0.748 25 14 Strong
 94 Q  -66.626  -36.623    4.116    6.898   0.246 0.722 25 14 Strong
 95 E  -68.260  -34.478    6.111    8.456   0.312 0.671 25 15 Strong
 96 A  -69.236  -29.380    4.472    8.239   0.366 0.600 25 18 Dyn
 97 R  -73.373  -14.808    9.298    9.728   0.441 0.529 25 18 Dyn
 98 E  -87.447   -3.738   11.367   10.529   0.582 0.485  8 18 Dyn
 99 V  -72.538  141.241    9.018   11.385   0.519 0.403  9 18 Dyn
100 N  -69.177  142.263   14.532   19.048   0.448 0.396  9 18 Dyn
101 S  -64.420  -33.434   10.028    9.440   0.375 0.450  8 18 Dyn
102 Q  -66.923  -35.355    4.651    4.858   0.269 0.616 25 18 Strong
103 E  -67.910  -36.326    4.124    4.363   0.190 0.768 25 18 Strong
```

```
104 E  -68.817  -37.863    4.786    4.373    0.133 0.829 25 18 Strong
105 E  -67.264  -39.553    3.739    5.462    0.104 0.858 25 18 Strong
106 E  -65.635  -39.778    3.736    4.090    0.079 0.873 25 18 Strong
107 E  -65.629  -37.715    3.584    4.581    0.078 0.879 25 17 Strong
108 E  -67.275  -41.857    3.727    4.317    0.097 0.870 25 17 Strong
109 L  -66.706  -36.591    4.656    3.989    0.117 0.836 25 16 Strong
110 L  -68.563  -36.216    5.115    5.725    0.160 0.815 25 11 Strong
111 R  -66.608  -35.896    4.368    3.721    0.202 0.791 25  8 Strong
112 K  -65.916  -38.693    6.030    4.692    0.160 0.826 25  9 Strong
113 A  -65.415  -35.566    5.394    5.040    0.124 0.837 25 15 Strong
114 I  -67.805  -40.509    4.268    4.972    0.107 0.855 25 18 Strong
115 A  -66.288  -36.605    3.856    4.380    0.097 0.839 25 18 Strong

116 E  -67.538  -36.661    3.850    6.461    0.124 0.791 25 18 Strong
117 S  -71.251  -28.744    6.500    9.766    0.174 0.693 25 18 Strong
118 L  -75.617  -18.846    6.091   13.251    0.256 0.549 25 18 Dyn
119 N  -82.906   -7.893   14.303   16.523    0.358 0.448  4 17 Dyn
120 S  -91.674   -3.762   13.204   10.832    0.550 0.395  4 11 Dyn
121 C 9999.000 9999.000    0.000    0.000    0.000 0.000  0  5 None
122 R 9999.000 9999.000    0.000    0.000    0.000 0.000  0  3 None
123 P  -61.414  146.442    5.455    6.242    0.622 0.639 25  9 Strong
124 S  -70.200  155.084    8.144    9.080    0.507 0.645 25 15 Strong
125 D  -66.590  142.312   12.109   20.220    0.556 0.623  8 18 Warn
126 A  -66.589  -21.370    8.284    9.476    0.602 0.417 10 18 Dyn
127 S  -85.055   -8.892    9.262   11.723    0.606 0.329  8 18 Dyn
128 A   92.447    0.572   12.514   48.191    0.797 0.240  7 18 Dyn
129 T  -79.268  140.403   14.356   17.498    0.894 0.210  9 18 Dyn
130 R  -89.179  127.129   12.130   11.924    1.532 0.144 25 17 Dyn
131 S 9999.000 9999.000    0.000    0.000    0.000 0.000  0 11 None
```

Supplemental Table 2.5: TALOS output for main chain ϕ and ψ dihedral angles for DE81-RAP80-tUIM.

```
REMARK TALOS-N Protein Backbone Torsion Angle Prediction Table
REMARK  Prediction Summary for Chemical Shift Input delE81RAP80UIM_talos_input_Manoj.tab
REMARK
REMARK  PHI is the predicted torsion angle C(i-1) N(i)  CA(i) C(i)   (degrees).
REMARK  PSI is the predicted torsion angle N(i)   CA(i) C(i)  N(i+1) (degrees).
REMARK
REMARK  DPHI and DPSI are the estimated standard deviations of the
REMARK  prediction errors in PHI and PSI (degrees).
REMARK
REMARK  DIST is the TALOS-N database matching score.
REMARK
REMARK  S2 is the Wishart RCI chemical shift order parameter,
REMARK  JACS, 127(43), 14970-14971.
REMARK
REMARK  COUNT is the number of database triplets used to form
REMARK  the torsion angle predictions.
REMARK
REMARK  CLASS is the classification of the prediction result:
REMARK    None: no torsion prediction was made.
REMARK
REMARK    Strong/Generous: majority consensus in database matches;
REMARK    prediction is likely to be good.
REMARK
REMARK    Warn: no consensus in database matches, do not use prediction.
REMARK
REMARK    Dyn:  RCI-S2 value indicates that residue has dynamic conformation.
REMARK
REMARK Reference:
REMARK  Y. Shen, and A. Bax:
REMARK  Protein backbone and sidechain torsion angles predicted from
REMARK  NMR chemical shifts using artificial neural networks
REMARK  J. Biomol. NMR (in press).
REMARK
REMARK  TALOS-N Version 4.01 Rev 2013.148.15.55 TALOSN_INFO
DATA FIRST_RESID 0
DATA SEQUENCE GPLGSRKIAQ MTEEQFALAL KMSEQEAREV NSQEEEEEL LRKAIAESLN
DATA SEQUENCE SCRPSDASAT RS
VARS   RESID RESNAME PHI PSI DPHI DPSI DIST S2 COUNT CS_COUNT CLASS
FORMAT %4d %s %8.3f %8.3f %8.3f %8.3f %8.3f %5.3f %2d %2d %s
-2 L 9999.000 9999.000   0.000    0.000   0.000 0.000  0  7 None
-1 G  122.684  168.458   70.797   96.295   1.741 0.403  6 11 Dyn
 0 S  -77.712  -13.020   15.135   12.055   1.142 0.499  6 11 Dyn
74 R  -68.593  -19.410    5.292    8.921   1.050 0.598 25 12 Dyn
75 K  -98.133    4.140   12.035   11.015   1.156 0.480  8 12 Dyn
76 I  -76.737  133.513    8.929    8.872   0.859 0.403 25 12 Dyn
77 A  -80.734  137.157   16.166   31.133   0.868 0.324 10 12 Dyn
78 Q  -70.257  136.225   14.188   19.411   0.522 0.332  8 12 Dyn
79 M  -70.897  141.494    7.600   10.133   0.278 0.411 25 12 Dyn
80 T  -74.067  161.229    6.332    5.991   0.217 0.559 25 12 Dyn
82 E  -59.005  -37.156    4.474    6.308   0.206 0.771 25 12 Strong
83 E  -67.038  -40.567    4.978    4.698   0.202 0.819 25 12 Strong
84 Q  -67.202  -36.536    3.288    5.664   0.201 0.808 25 12 Strong
85 F  -66.058  -43.085    2.790    5.732   0.188 0.785 25 12 Strong
86 A  -63.471  -37.591    3.604    4.178   0.196 0.758 25 12 Strong
87 L  -66.817  -41.142    4.977    7.128   0.205 0.719 25 12 Strong
88 A  -66.547  -37.536    6.150    5.709   0.230 0.679 25 12 Strong
89 L  -67.854  -39.372    3.562    5.200   0.245 0.640 25 12 Strong
90 K  -68.778  -35.935    4.504    5.257   0.289 0.623 25 11 Strong
91 M  -68.831  -35.740    4.280    6.419   0.312 0.626 25 11 Strong
92 S  -69.438  -33.286    7.086    7.421   0.346 0.647 25 11 Strong
93 E  -69.017  -30.665    5.712    6.656   0.402 0.669 25 12 Strong
94 Q  -74.184  -25.542    8.799   10.260   0.457 0.652 25 12 Strong
95 E  -87.880   -5.686   12.836   16.278   0.579 0.594  6 12 Dyn
96 A   83.006   -7.487   12.146   42.808   0.712 0.557  8  8 Dyn
97 R 9999.000 9999.000    0.000    0.000   0.000 0.000  0  4 None
```

```
98 E 9999.000 9999.000   0.000   0.000   0.000 0.000  0  4 None

99 V  -77.098 140.754  11.152  10.762   0.505 0.305 25  8 Dyn
100 N -68.429 144.289   6.850  15.647   0.419 0.316 10 12 Dyn
101 S -63.958 -31.733  11.294  12.136   0.466 0.380  5 12 Dyn
102 Q -66.182 -35.230   4.010   6.299   0.383 0.523 25 12 Dyn
103 E -67.800 -35.506   5.116   4.431   0.321 0.700 25 12 Strong
104 E -69.047 -37.635   4.156   4.709   0.271 0.775 25 11 Strong
105 E -66.506 -39.639   3.738   5.141   0.238 0.812 25 11 Strong
106 E -66.308 -38.163   4.389   6.156   0.204 0.832 25 11 Strong
107 E -64.162 -40.219   3.422   3.822   0.191 0.837 25 12 Strong
108 E -67.155 -39.546   2.951   4.685   0.180 0.817 25 12 Strong
109 L -66.581 -41.084   3.890   5.251   0.165 0.800 25 12 Strong
110 L -67.100 -39.654   4.428   4.642   0.171 0.798 25 12 Strong
111 R -64.748 -38.840   4.497   4.704   0.185 0.806 25 12 Strong
112 K -66.216 -38.705   3.597   4.896   0.183 0.802 25 12 Strong
113 A -65.219 -38.883   3.263   3.247   0.181 0.788 25 12 Strong
114 I -68.298 -41.524   4.134   4.597   0.189 0.786 25 12 Strong
115 A -65.269 -37.413   3.800   4.705   0.206 0.782 25 12 Strong
116 E -68.755 -36.654   3.953   5.071   0.245 0.743 25 12 Strong
117 S -66.588 -32.330   4.727   7.744   0.257 0.639 25 12 Strong
118 L -69.059 -25.303   5.004   9.427   0.358 0.520 25 12 Dyn
119 N -90.477  -3.228  14.081  12.978   0.386 0.517 25 12 Dyn
120 S -75.484 144.098  10.818  13.208   0.555 0.587  7 12 Dyn
121 C -98.809 153.711  21.968  11.927   0.893 0.740 10 12 Generous
122 R -65.846 139.706   7.277   9.837   0.946 0.787 25  8 Strong
123 P -62.475 149.365   7.658   9.184   0.621 0.701 25  8 Strong
124 S -65.293 149.486   8.174  11.348   0.479 0.587 25  8 Dyn
125 D -71.623 141.188  11.622  20.029   0.612 0.569  7 12 Dyn
126 A -75.437 -18.664  15.549  14.825   0.641 0.434  6 12 Dyn
127 S -91.251  -1.794  12.360  15.141   0.726 0.375  3 12 Dyn
128 A  87.046   4.489   8.019  11.049   0.755 0.305  7 12 Dyn
129 T -77.362 142.438  10.895  11.501   0.696 0.281 25 12 Dyn
130 R -81.343 137.581  10.191  11.651   1.162 0.229 25 12 Dyn
131 S 9999.000 9999.000   0.000   0.000   0.000 0.000  0  8 None
```

Supplemental Table 2.6: MICS output for structural motifs within wild type RAP80-tUIM.

REMARK MICS Protein Structural Motif Prediction Table
REMARK Prediction Summary for Chemical Shift Input WT-5C_07jan14-update-from-21Jan14_talos.tab

REMARK CS_COUNT is the number of chemical shifts for given residue
REMARK Q_H is the predicted probability to be a residue in a helix
REMARK Q_E is the predicted probability to be a residue in a strand
REMARK Q_L is the predicted probability to be a residue in a loop
REMARK Q_NCAP is the predicted probability to be a Ncap residue in a Ncap motif
REMARK Q_CCAP is the predicted probability to be a Ccap residue in a Ccap motif
REMARK Q_T1@2 is the predicted probability to be the 2nd residue in a type I beta-turn
REMARK Q_T2@2 is the predicted probability to be the 2nd residue in a type II beta-turn
REMARK Q_T1p@2 is the predicted probability to be the 2nd residue in a type I' beta-turn
REMARK Q_T2p@2 is the predicted probability to be the 2nd residue in a type II' beta-turn
REMARK Q_T8@2 is the predicted probability to be the 2nd residue in a type VIII beta-turn
REMARK S2 is the Wishart RCI chemical shift order parameter [JACS, 127, 14970-14971]

DATA FIRST_RESID 69
DATA SEQUENCE   GPLGSRKIAQ MTEEEQFALA LKMSEQEARE VNSQEEEEEE LLRKAIAESL
DATA SEQUENCE   NSCRPSDASA TRS

VARS RESID RESNAME CS_CNT SS_CLASS Q_H Q_E Q_L   Q_NCAP Q_CCAP   Q_T1@2 Q_T2@2
Q_T1p@2 Q_T2p@2 Q_T8@2 S2
FORMAT %4d %1s %2d %1s %8.3f %8.3f %8.3f %8.3f %8.3f %8.3f %8.3f %8.3f %8.3f %8.3f %8.3f

RESID	RESNAME	CS_CNT	SS_CLASS	Q_H	Q_E	Q_L	Q_NCAP	Q_CCAP	Q_T1@2	Q_T2@2	Q_T1p@2	Q_T2p@2	Q_T8@2	S2
-1	G	3	L	0.114	0.252	0.579	0.000	0.000	0.041	0.000	0.000	0.001	0.014	0.101
0	S	0	L	0.042	0.345	0.514	0.003	0.000	0.041	0.002	0.023	0.014	0.016	0.140
74	R	3	L	0.128	0.149	0.518	0.000	0.000	0.060	0.017	0.003	0.005	0.121	0.206
75	K	6	L	0.040	0.185	0.706	0.012	0.000	0.021	0.000	0.000	0.002	0.034	0.241
76	I	6	L	0.189	0.211	0.545	0.000	0.000	0.000	0.017	0.006	0.000	0.032	0.289
77	A	6	L	0.123	0.173	0.610	0.000	0.009	0.000	0.000	0.019	0.015	0.051	0.332
78	Q	6	L	0.032	0.439	0.514	0.000	0.004	0.000	0.004	0.007	0.000	0.000	0.364
79	M	6	L	0.014	0.270	0.687	0.000	0.002	0.007	0.010	0.010	0.000	0.000	0.476
80	T	6	L	0.005	0.018	0.081	0.871	0.008	0.000	0.011	0.006	0.000	0.000	0.636
81	E	6	H	0.919	0.003	0.049	0.000	0.013	0.000	0.000	0.006	0.000	0.009	0.838
82	E	6	H	0.973	0.000	0.005	0.000	0.017	0.000	0.002	0.001	0.000	0.001	0.890
83	E	6	H	0.963	0.000	0.003	0.001	0.028	0.000	0.002	0.001	0.000	0.001	0.897
84	Q	6	H	0.955	0.000	0.003	0.001	0.037	0.000	0.001	0.001	0.000	0.002	0.888
85	F	6	H	0.951	0.001	0.005	0.001	0.037	0.000	0.003	0.002	0.000	0.001	0.869
86	A	6	H	0.950	0.001	0.003	0.000	0.037	0.004	0.000	0.000	0.005	0.000	0.854
87	L	6	H	0.941	0.000	0.014	0.001	0.036	0.001	0.003	0.001	0.000	0.003	0.829
88	A	6	H	0.952	0.000	0.004	0.002	0.037	0.000	0.001	0.001	0.000	0.003	0.821
89	L	6	H	0.942	0.001	0.014	0.003	0.036	0.003	0.000	0.000	0.000	0.000	0.806
90	K	6	H	0.944	0.001	0.010	0.005	0.037	0.003	0.000	0.000	0.000	0.001	0.797
91	M	6	H	0.922	0.001	0.022	0.011	0.033	0.004	0.003	0.002	0.000	0.003	0.785
92	S	6	H	0.767	0.088	0.106	0.008	0.028	0.000	0.000	0.000	0.005	0.000	0.771
93	E	5	H	0.837	0.007	0.057	0.018	0.040	0.013	0.014	0.011	0.000	0.003	0.765
94	Q	3	H	0.831	0.010	0.071	0.020	0.038	0.014	0.000	0.006	0.010	0.000	0.730
95	E	6	H	0.816	0.007	0.090	0.021	0.024	0.026	0.003	0.007	0.005	0.000	0.672
96	A	6	H	0.695	0.078	0.148	0.006	0.021	0.029	0.000	0.000	0.013	0.010	0.596
97	R	6	H	0.694	0.018	0.213	0.002	0.020	0.015	0.007	0.002	0.000	0.029	0.526
98	E	6	L	0.306	0.083	0.549	0.009	0.034	0.000	0.000	0.000	0.011	0.008	0.482
99	V	6	L	0.215	0.141	0.540	0.002	0.031	0.000	0.040	0.012	0.000	0.020	0.400
100	N	6	L	0.059	0.048	0.830	0.003	0.043	0.002	0.000	0.000	0.012	0.005	0.393
101	S	6	L	0.050	0.156	0.675	0.004	0.004	0.101	0.007	0.000	0.000	0.003	0.451
102	Q	6	H	0.563	0.036	0.276	0.002	0.019	0.069	0.000	0.000	0.000	0.037	0.600
103	E	5	H	0.716	0.051	0.188	0.007	0.016	0.007	0.000	0.000	0.000	0.014	0.709
104	E	0	H	0.621	0.061	0.288	0.005	0.016	0.000	0.000	0.000	0.008	0.000	0.709
105	E	0	H	0.394	0.240	0.298	0.019	0.018	0.000	0.000	0.001	0.000	0.031	-9998.000
106	E	0	L	0.372	0.085	0.333	0.078	0.015	0.000	0.072	0.010	0.000	0.034	-9998.000
107	E	0	H	0.649	0.026	0.045	0.246	0.020	0.000	0.000	0.001	0.000	0.011	0.812
108	E	3	H	0.947	0.001	0.012	0.000	0.029	0.005	0.002	0.003	0.000	0.001	0.832
109	L	6	H	0.913	0.000	0.020	0.012	0.034	0.000	0.001	0.001	0.017	0.002	0.841
110	L	5	H	0.884	0.002	0.060	0.018	0.032	0.000	0.001	0.001	0.000	0.000	0.856
111	R	0	H	0.940	0.001	0.006	0.006	0.036	0.000	0.000	0.001	0.008	0.002	0.860
112	K	3	H	0.944	0.002	0.005	0.001	0.037	0.005	0.003	0.000	0.000	0.003	0.860
113	A	6	H	0.930	0.001	0.005	0.002	0.036	0.008	0.000	0.001	0.016	0.002	0.859
114	I	6	H	0.930	0.001	0.028	0.002	0.036	0.000	0.000	0.002	0.000	0.000	0.854
115	A	6	H	0.885	0.048	0.016	0.000	0.033	0.001	0.005	0.003	0.008	0.001	0.836
116	E	6	H	0.878	0.008	0.024	0.001	0.031	0.021	0.016	0.015	0.004	0.000	0.788

```
117 S  6 H   0.762  0.115  0.044  0.000  0.033  0.027  0.017  0.000  0.001  0.000  0.690
118 L  6 H   0.731  0.005  0.171  0.002  0.026  0.060  0.000  0.001  0.003  0.001  0.551
119 N  6 L   0.260  0.043  0.574  0.004  0.055  0.003  0.000  0.000  0.011  0.051  0.459
120 S  5 L   0.056  0.325  0.545  0.002  0.002  0.004  0.033  0.007  0.005  0.020  0.411
121 C  0 L   0.028  0.395  0.475  0.005  0.002  0.001  0.030  0.003  0.000  0.061  0.409
122 R  0 L   0.038  0.087  0.818  0.001  0.014  0.000  0.007  0.000  0.000  0.034  0.661
123 P  3 L   0.126  0.042  0.543  0.004  0.009  0.197  0.024  0.009  0.003  0.043  0.656
```

Supplemental Table 2.7: MICS output for structural motifs within ΔE81-RAP80-tUIM.

REMARK MICS Protein Structural Motif Prediction Table
REMARK Prediction Summary for Chemical Shift Input ConvertedSequence_cara_mics.tab

REMARK CS_COUNT is the number of chemical shifts for given residue
REMARK Q_H is the predicted probability to be a residue in a helix
REMARK Q_E is the predicted probability to be a residue in a strand
REMARK Q_L is the predicted probability to be a residue in a loop
REMARK Q_NCAP is the predicted probability to be a Ncap residue in a Ncap motif
REMARK Q_CCAP is the predicted probability to be a Ccap residue in a Ccap motif
REMARK Q_T1@2 is the predicted probability to be the 2nd residue in a type I beta-turn
REMARK Q_T2@2 is the predicted probability to be the 2nd residue in a type II beta-turn
REMARK Q_T1p@2 is the predicted probability to be the 2nd residue in a type I' beta-turn
REMARK Q_T2p@2 is the predicted probability to be the 2nd residue in a type II' beta-turn
REMARK Q_T8@2 is the predicted probability to be the 2nd residue in a type VIII beta-turn
REMARK S2 is the Wishart RCI chemical shift order parameter [JACS, 127, 14970-14971]

DATA SEQUENCE   LGSRKIAQMT EEQFALALKM SEQEAREVNS QEEEEEELLR KAIAESLNSC
DATA SEQUENCE   RPSDASATRS

VARS RESID RESNAME CS_CNT SS_CLASS Q_H Q_E Q_L Q_NCAP Q_CCAP Q_T1@2 Q_T2@2 Q_T1p@2 Q_T2p@2 Q_T8@2 S2
FORMAT %4d %1s %2d %1s %8.3f %8.3f %8.3f %8.3f %8.3f %8.3f %8.3f %8.3f %8.3f %8.3f %8.3f

RESID	RESNAME	CS_CNT	SS_CLASS	Q_H	Q_E	Q_L	Q_NCAP	Q_CCAP	Q_T1@2	Q_T2@2	Q_T1p@2	Q_T2p@2	Q_T8@2	S2
-1	G	3	L	0.067	0.058	0.862	0.008	0.000	0.000	0.000	0.000	0.000	0.005	0.359
0	S	5	L	0.030	0.247	0.591	0.025	0.000	0.048	0.023	0.014	0.008	0.014	0.472
74	R	4	L	0.290	0.044	0.544	0.006	0.000	0.051	0.021	0.001	0.013	0.030	0.629
75	K	5	L	0.361	0.085	0.430	0.016	0.000	0.050	0.003	0.000	0.014	0.041	0.622
76	I	5	L	0.374	0.092	0.487	0.010	0.017	0.000	0.000	0.002	0.000	0.018	0.605
77	A	5	L	0.164	0.083	0.694	0.000	0.011	0.000	0.000	0.004	0.007	0.037	0.548
78	Q	5	L	0.018	0.135	0.812	0.000	0.014	0.000	0.005	0.015	0.000	0.000	0.523
79	M	5	L	0.006	0.159	0.801	0.006	0.009	0.000	0.012	0.007	0.000	0.000	0.560
80	T	5	L	0.006	0.031	0.148	0.790	0.004	0.001	0.013	0.006	0.000	0.000	0.659
82	E	5	H	0.898	0.004	0.073	0.000	0.010	0.000	0.001	0.005	0.000	0.009	0.813
83	E	5	H	0.974	0.001	0.007	0.000	0.014	0.000	0.002	0.001	0.000	0.001	0.866
84	Q	5	H	0.967	0.001	0.005	0.001	0.023	0.000	0.001	0.001	0.000	0.002	0.873
85	F	5	H	0.958	0.002	0.005	0.000	0.028	0.000	0.002	0.002	0.000	0.002	0.859
86	A	5	H	0.932	0.004	0.017	0.000	0.037	0.003	0.000	0.000	0.005	0.001	0.840
87	L	5	H	0.945	0.001	0.006	0.002	0.036	0.005	0.004	0.000	0.000	0.001	0.815
88	A	5	H	0.935	0.003	0.018	0.003	0.036	0.003	0.000	0.000	0.000	0.002	0.793
89	L	5	H	0.936	0.001	0.014	0.005	0.036	0.003	0.001	0.001	0.000	0.003	0.784
90	K	5	H	0.924	0.004	0.027	0.006	0.036	0.003	0.000	0.000	0.000	0.001	0.784
91	M	4	H	0.903	0.001	0.032	0.016	0.032	0.009	0.003	0.002	0.000	0.002	0.793
92	S	5	H	0.756	0.116	0.072	0.026	0.030	0.000	0.000	0.000	0.000	0.000	0.787
93	E	5	H	0.832	0.006	0.068	0.011	0.043	0.028	0.004	0.007	0.000	0.000	0.783
94	Q	5	H	0.832	0.009	0.074	0.009	0.030	0.030	0.003	0.008	0.005	0.000	0.754
95	E	5	H	0.823	0.007	0.088	0.011	0.022	0.046	0.000	0.000	0.002	0.000	0.705
96	A	5	H	0.590	0.120	0.171	0.004	0.021	0.070	0.000	0.000	0.004	0.019	0.630
97	R	5	L	0.445	0.059	0.394	0.008	0.045	0.000	0.014	0.014	0.005	0.016	0.571
98	E	5	L	0.072	0.065	0.801	0.011	0.028	0.000	0.000	0.003	0.013	0.007	0.545
99	V	5	L	0.113	0.105	0.728	0.018	0.009	0.000	0.013	0.014	0.000	0.000	0.522
100	N	5	L	0.010	0.042	0.865	0.046	0.005	0.000	0.009	0.010	0.013	0.000	0.519
101	S	5	L	0.099	0.152	0.568	0.101	0.005	0.046	0.015	0.009	0.003	0.002	0.563
102	Q	5	H	0.666	0.012	0.225	0.086	0.007	0.000	0.000	0.000	0.000	0.003	0.665
103	E	5	H	0.932	0.003	0.039	0.013	0.012	0.000	0.000	0.000	0.000	0.001	0.789
104	E	5	H	0.959	0.001	0.018	0.002	0.016	0.000	0.001	0.000	0.001	0.001	0.842
105	E	4	H	0.957	0.001	0.012	0.002	0.026	0.000	0.001	0.001	0.000	0.001	0.864
106	E	5	H	0.955	0.000	0.007	0.001	0.031	0.000	0.002	0.001	0.001	0.001	0.873
107	E	5	H	0.950	0.000	0.005	0.001	0.035	0.000	0.002	0.001	0.004	0.001	0.875
108	E	5	H	0.951	0.000	0.003	0.001	0.037	0.000	0.002	0.001	0.004	0.002	0.864
109	L	5	H	0.947	0.001	0.002	0.001	0.036	0.000	0.002	0.002	0.007	0.001	0.853
110	L	5	H	0.955	0.002	0.002	0.001	0.037	0.001	0.000	0.000	0.000	0.001	0.848
111	R	5	H	0.952	0.000	0.005	0.001	0.037	0.001	0.000	0.001	0.000	0.003	0.852
112	K	5	H	0.940	0.002	0.007	0.001	0.036	0.006	0.004	0.003	0.000	0.001	0.854
113	A	5	H	0.915	0.004	0.037	0.002	0.036	0.007	0.000	0.000	0.000	0.000	0.849
114	I	5	H	0.886	0.002	0.027	0.003	0.035	0.000	0.017	0.012	0.013	0.005	0.843
115	A	5	H	0.864	0.066	0.022	0.001	0.035	0.000	0.000	0.000	0.012	0.000	0.827
116	E	5	H	0.834	0.007	0.038	0.002	0.029	0.032	0.030	0.013	0.017	0.000	0.798
117	S	5	H	0.736	0.102	0.071	0.003	0.026	0.057	0.003	0.000	0.003	0.000	0.732
118	L	5	H	0.627	0.013	0.182	0.000	0.049	0.092	0.014	0.011	0.004	0.009	0.659
119	N	5	L	0.086	0.087	0.753	0.009	0.009	0.000	0.007	0.000	0.004	0.046	0.651
120	S	5	L	0.013	0.261	0.662	0.000	0.005	0.012	0.021	0.003	0.000	0.024	0.692
121	C	4	E	0.008	0.555	0.405	0.005	0.009	0.005	0.010	0.001	0.000	0.004	0.757
122	R	0	L	0.024	0.067	0.849	0.010	0.007	0.000	0.022	0.000	0.000	0.021	0.791

```
123 P  0 L   0.154   0.035   0.434   0.003   0.008   0.232   0.050   0.012   0.001   0.071   0.609
124 S  5 L   0.091   0.168   0.632   0.009   0.007   0.054   0.003   0.000   0.000   0.036   0.628
125 D  5 L   0.095   0.205   0.576   0.015   0.009   0.063   0.007   0.000   0.000   0.030   0.644


126 A  5 L   0.037   0.285   0.595   0.006   0.007   0.000   0.006   0.000   0.000   0.064   0.601
127 S  5 L   0.360   0.083   0.487   0.021   0.009   0.021   0.004   0.000   0.000   0.015   0.554
128 A  5 L   0.395   0.118   0.440   0.000   0.007   0.022   0.004   0.000   0.000   0.015   0.498
129 T  5 L   0.115   0.122   0.755   0.000   0.008   0.000   0.000   0.000   0.000   0.000   0.459
130 R  5 L   0.175   0.404   0.418   0.000   0.003   0.000   0.000   0.000   0.000   0.000   0.411
```

Supplemental Table 2.8: TALOS output for side chain χ^1 angles for wild type RAP80-tUIM.

```
REMARK TALOS-N Protein Chi1 Conformation Prediction Table
REMARK  Prediction Summary for Chemical Shift Input WT-5C_07jan14-update-from-21Jan14_talos.tab
REMARK
REMARK  Q_Gm is the probability to be with a g- (gauche-) rotamer.
REMARK  Q_Gp is the probability to be with a g+ (gauche+)rotamer.
REMARK  Q_t  is the probability to be with a t (trans) rotamer.
REMARK
REMARK  CS_COUNT is the number of chemical shifts in the query triplets
REMARK  used to predict chi1 conformation.
REMARK
REMARK  CLASS is the classification of the prediction result:
REMARK    na: no chi1 prediction was made.
REMARK    g-/g+/t: 3-state chi1 prediction
REMARK
REMARK Reference:
REMARK  Y. Shen, and A. Bax:
REMARK  Protein backbone and sidechain torsion angles predicted from
REMARK  NMR chemical shifts using artificial neural networks
REMARK  J. Biomol. NMR (in press).
REMARK
REMARK  TALOS-N Version 4.01 Rev 2013.148.15.55 TALOSN_INFO

VARS   RESID RESNAME CS_COUNT Q_Gm Q_Gp Q_T CLASS
FORMAT %4d %s %2d %5.3f %5.3f %5.3f %s

  -2  L  9 0.333 0.333 0.333 na
   0  S  6 0.351 0.486 0.163 na
  74  R  9 0.580 0.131 0.289 na
  75  K 15 0.517 0.078 0.405 na
  76  I 18 0.886 0.064 0.050 g-
  78  Q 18 0.651 0.099 0.250 g-
  79  M 18 0.529 0.102 0.369 na
  80  T 18 0.050 0.901 0.050 g+
  81  E 18 0.462 0.194 0.344 na
  82  E 18 0.497 0.208 0.295 na
  83  E 18 0.567 0.061 0.372 na
  84  Q 18 0.533 0.102 0.366 na
  85  F 18 0.189 0.026 0.785 t
  87  L 18 0.562 0.040 0.399 na
  89  L 18 0.633 0.041 0.326 g-
  90  K 18 0.493 0.082 0.425 na
  91  M 18 0.627 0.083 0.290 g-
  92  S 17 0.289 0.553 0.158 na
  93  E 14 0.536 0.065 0.399 na
  94  Q 14 0.629 0.110 0.261 g-
  95  E 15 0.613 0.066 0.321 g-
  97  R 18 0.583 0.083 0.335 na
  98  E 18 0.823 0.088 0.088 g-
  99  V 18 0.158 0.100 0.741 t
 100  N 18 0.388 0.225 0.388 na
 101  S 18 0.241 0.519 0.241 na
 102  Q 17 0.634 0.103 0.263 g-
 103  E 11 0.595 0.093 0.312 na
 104  E  5 0.333 0.333 0.333 na
 105  E  0 0.333 0.333 0.333 na
 106  E  0 0.333 0.333 0.333 na
 107  E  3 0.333 0.333 0.333 na
 108  E  9 0.567 0.061 0.372 na
 109  L 14 0.692 0.045 0.263 g-
 110  L 11 0.562 0.040 0.399 na
 111  R  8 0.539 0.079 0.383 na
 112  K  9 0.500 0.096 0.403 na
 114  I 18 0.898 0.051 0.051 g-
 116  E 18 0.613 0.066 0.321 g-
 117  S 18 0.289 0.553 0.158 na
 118  L 18 0.705 0.043 0.252 g-
 119  N 17 0.578 0.211 0.211 na
 120  S 11 0.273 0.497 0.230 na
 121  C  5 0.333 0.333 0.333 na
 122  R  3 0.333 0.333 0.333 na
 124  S 15 0.289 0.467 0.244 na
 125  D 18 0.340 0.193 0.466 na
```

127 S 18 0.289 0.467 0.244 na
129 T 18 0.444 0.444 0.112 na
130 R 17 0.422 0.086 0.492 na
131 S 11 0.333 0.333 0.333 na

Supplemental Table 2.9: TALOS output for side chain χ^1 angles for DE81-RAP80-tUIM.

REMARK TALOS-N Protein Chi1 Conformation Prediction Table
REMARK Prediction Summary for Chemical Shift Input delE81RAP80UIM_talos_input_Manoj.tab
REMARK
REMARK Q_Gm is the probability to be with a g- (gauche-) rotamer.
REMARK Q_Gp is the probability to be with a g+ (gauche+)rotamer.
REMARK Q_t is the probability to be with a t (trans) rotamer.
REMARK
REMARK CS_COUNT is the number of chemical shifts in the query triplets
REMARK used to predict chi1 conformation.
REMARK
REMARK CLASS is the classification of the prediction result:
REMARK   na: no chi1 prediction was made.
REMARK   g-/g+/t: 3-state chi1 prediction
REMARK
REMARK Reference:
REMARK Y. Shen, and A. Bax:
REMARK Protein backbone and sidechain torsion angles predicted from
REMARK NMR chemical shifts using artificial neural networks
REMARK J. Biomol. NMR (in press).
REMARK
REMARK TALOS-N Version 4.01 Rev 2013.148.15.55 TALOSN_INFO
VARS   RESID RESNAME CS_COUNT Q_Gm Q_Gp Q_T CLASS
FORMAT %4d %s %2d %5.3f %5.3f %5.3f %s
-2 L 7 0.333 0.333 0.333 na
0 S 11 0.148 0.582 0.270 na
74 R 12 0.562 0.115 0.323 na
75 K 12 0.561 0.085 0.354 na
76 I 12 0.843 0.061 0.096 g-
78 Q 12 0.651 0.099 0.250 g-
79 M 12 0.564 0.138 0.298 na
80 T 12 0.050 0.901 0.050 g+
82 E 12 0.497 0.133 0.370 na
83 E 12 0.537 0.143 0.319 na
84 Q 12 0.570 0.069 0.361 na
85 F 12 0.189 0.026 0.785 t
87 L 12 0.562 0.040 0.399 na
89 L 12 0.633 0.041 0.326 g-
90 K 11 0.538 0.089 0.373 na
91 M 11 0.595 0.090 0.314 na
92 S 11 0.314 0.514 0.172 na
93 E 12 0.613 0.066 0.321 g-
94 Q 12 0.629 0.110 0.261 g-
95 E 12 0.613 0.066 0.321 g-
97 R 4 0.333 0.333 0.333 na
98 E 4 0.333 0.333 0.333 na
99 V 8 0.165 0.104 0.731 t
100 N 12 0.356 0.206 0.438 na
101 S 12 0.158 0.553 0.289 na
102 Q 12 0.537 0.095 0.368 na
103 E 12 0.568 0.134 0.298 na
104 E 11 0.613 0.066 0.321 g-
105 E 11 0.536 0.065 0.399 na
106 E 11 0.565 0.100 0.335 na
107 E 12 0.536 0.065 0.399 na
108 E 12 0.536 0.065 0.399 na
109 L 12 0.633 0.041 0.326 g-
110 L 12 0.480 0.039 0.480 na
111 R 12 0.436 0.128 0.436 na
112 K 12 0.456 0.088 0.456 na
114 I 12 0.847 0.051 0.102 g-
116 E 12 0.565 0.100 0.335 na
117 S 12 0.314 0.514 0.172 na
118 L 12 0.692 0.045 0.263 g-
119 N 12 0.676 0.050 0.274 g-
120 S 12 0.310 0.429 0.262 na
121 C 12 0.305 0.475 0.220 na
122 R 8 0.514 0.151 0.335 na
124 S 8 0.289 0.467 0.244 na
125 D 12 0.360 0.232 0.409 na
127 S 12 0.289 0.467 0.244 na
129 T 12 0.351 0.551 0.098 na
130 R 12 0.444 0.112 0.444 na
131 S 8 0.333 0.333 0.333 na

References

[1]   Sobhian, B., Shao, G. Z., Lilli, D. R., Culhane, A. C., Moreau, L. A., Xia, B., Livingston, D. M., and Greenberg, R. A. (2007) RAP80 targets BRCA1 to specific ubiquitin structures at DNA damage sites. *Science* **316**, 1198-1202

[2]   Wang, B., Matsuoka, S., Ballif, B. A., Zhang, D., Smogorzewska, A., Gygi, S. P., and Elledge, S. J. (2007) Abraxas and RAP80 form a BRCA1 protein complex required for the DNA damage response. *Science* **316**, 1194-1198

[3]   Nikkila, J., Coleman, K. A., Morrissey, D., Pylkas, K., Erkko, H., Messick, T. E., Karppinen, S. M., Amelina, A., Winqvist, R., and Greenberg, R. A. (2009) Familial breast cancer screening reveals an alteration in the RAP80 UIM domain that impairs DNA damage response function. *Oncogene* **28**, 1843-1852

[4]   Yin, Z., Menendez, D., Resnick, M. A., French, J. E., Janardhan, K. S., and Jetten, A. M. (2012) RAP80 is critical in maintaining genomic stability and suppressing tumor development. *Cancer Res.* **72**, 5080-5090

[5]   Osorio, A., Barroso, A., Garcia, M. J., Martinez-Delgado, B., Urioste, M., and Benitez, J. (2009) Evaluation of the BRCA1 interacting genes RAP80 and CCDC98 in familial breast cancer susceptibility. *Breast Cancer Res. Tr.* **113**, 371-376

[6]   Rebbeck, T. R., Mitra, N., Domchek, S. M., Wan, F., Friebel, T. M., Tran, T. V., Singer, C. F., Tea, M. K., Blum, J. L., Tung, N., Olopade, O. I., Weitzel, J. N., Lynch, H. T., Snyder, C. L., Garber, J. E., Antoniou, A. C., Peock, S., Evans, D. G., Paterson, J., Kennedy, M. J., Donaldson, A., Dorkins, H., Easton, D. F., Epidemiological Study of, B., Carriers, B. M., Rubinstein, W. S., Daly, M. B., Isaacs, C., Nevanlinna, H., Couch, F. J., Andrulis, I. L., Freidman, E., Laitman, Y., Ganz, P. A., Tomlinson, G. E., Neuhausen, S. L., Narod, S. A., Phelan, C. M., Greenberg, R., and Nathanson, K. L. (2011) Modification of BRCA1-associated breast and ovarian cancer risk by BRCA1-interacting genes. *Cancer Res.* **71**, 5792-5805

[7]   Sims, J. J., and Cohen, R. E. (2009) Linkage-specific avidity defines the lysine 63-linked polyubiquitin-binding preference of RAP80. *Mol. Cell* **33**, 775-783

[8]   Markin, C. J., Wei, X. A., and Spyracopoulos, L. (2010) Mechanism for recognition of polyubiquitin chains: Balancing affinity through interplay between multivalent binding and dynamics. *J. Am. Chem. Soc.* **132**, 11247-11258

[9]   Sato, Y., Yoshikawa, A., Mimura, H., Yamashita, M., Yamagata, A., and Fukai, S. (2009) Structural basis for specific recognition of Lys 63-linked polyubiquitin chains by tandem UIMs of RAP80. *EMBO J.* **28**, 2461-2468

[10]   Marley, J., Lu, M., and Bracken, C. (2001) A method for efficient isotopic labeling of recombinant proteins. *J Biomol. NMR* **20**, 71-75

[11]   Wittekind, M., and Mueller, L. (1993) HNCACB, a high-sensitivity 3D NMR experiment to correlate amide proton and nitrogen resonances with the a- and a-carbon resonances in proteins. *J. Magn. Reson.* **B 101**, 201-205

[12]   Muhandiram, D. R., and Kay, L. E. (1994) Gradient-enhanced triple-resonance 3-dimensional NMR experiments with improved sensitivity. *J. Magn. Reson.* **B 103**, 203-216

[13]   Grzesiek, S., and Bax, A. (1992) Correlating backbone amide and side chain resonances in larger proteins by multiple relayed triple resonance NMR. *J. Am. Chem. Soc.* **114**, 6291-6293

[14]   Grzesiek, S., and Bax, A. (1993) Amino acid type determination in the sequential assignment procedure of uniformly ^{13}C/^{15}N enriched proteins. *J. Biomol. NMR* **3**, 185-204

[15]   Keller, R. L. J. (2004) The computer aided resonance assignment tutorial. Cantina Verlag, Zürich

[16]   Goddard, T. D., and Kneller, D. G. SPARKY 3. University of California, San Francisco

[17]   Farrow, N. A., Muhandiram, R., Singer, A. U., Pascal, S. M., Kay, C. M., Gish, G., Shoelson, S. E., Pawson, T., Forman-Kay, J. D., and Kay, L. E. (1994) Backbone dynamics of a free and a phosphopeptide-complexed Src homology-2 domain studied by ^{15}N NMR relaxation. *Biochemistry* **33**, 5984-6003

[18]   Mandel, A. M., Akke, M., and Palmer, A. G. (1995) Backbone dynamics of *Escherichia coli* ribonuclease HI - Correlations with structure and function in an active enzyme. *J. Mol. Biol.* **246**, 144-163

[19]   Spyracopoulos, L. (2006) A suite of Mathematica notebooks for the analysis of protein main chain ^{15}N NMR relaxation data. *J Biomol. NMR* **36**, 215-224

[20]   Kay, L. E., Xu, G. Y., and Yamazaki, T. (1994) Enhanced sensitivity triple resonance spectroscopy with minimal H_2O saturation. *J. Magn. Reson.* **A 109**, 129-133

[21]   Zhang, O., Kay, L. E., Olivier, J. P., and Forman-Kay, J. D. (1994) Backbone ^{1}H and ^{15}N resonance assignments of the N-terminal SH3 domain of drk in folded and unfolded states using enhanced-sensitivity pulsed field gradient NMR techniques. *J Biomol. NMR* **4**, 845-858

[22]   Shen, Y., and Bax, A. (2013) Protein backbone and sidechain torsion angles predicted from NMR chemical shifts using artificial neural networks. *J Biomol. NMR* **56**, 227-241

[23]   Shen, Y., and Bax, A. (2012) Identification of helix capping and b-turn motifs from NMR chemical shifts. *J Biomol. NMR* **52**, 211-232

[24]   Schwieters, C. D., Kuszewski, J. J., Tjandra, N., and Clore, G. M. (2003) The Xplor-NIH NMR molecular structure determination package. *J. Magn. Reson.* **160**, 65-73

[25]   Wolfram, S. (2003) *The Mathematica book*, 5th ed., Wolfram Media, Champaign, IL

[26]   Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J Biomol. NMR* **6**, 277-293

[27]   Nesmelova, I., Krushelnitsky, A., Idiyatullin, D., Blanco, F., Ramirez-Alvarado, M., Daragan, V. A., Serrano, L., and Mayo, K. H. (2001) Conformational exchange on the microsecond time scale in a-helix and b-hairpin peptides measured by ^{13}C NMR transverse relaxation. *Biochemistry* **40**, 2844-2853

[28]   Kay, L. E., Keifer, P., and Saarinen, T. (1992) Pure absorption gradient enhanced heteronuclear single quantum correlation spectroscopy with improved sensitivity. *J. Am. Chem. Soc.* **114**, 10663-10665

[29]   Markin, C. J., and Spyracopoulos, L. (2012) Increased precision for analysis of protein-ligand dissociation constants determined from chemical shift titrations. *J Biomol. NMR* **53**, 125-138

[30]   Markin, C. J., and Spyracopoulos, L. (2012) Accuracy and precision of protein-ligand interaction kinetics determined from chemical shift titrations. *J Biomol. NMR* **54**, 355-376

[31]   Case, D. A., Darden, T. A., Cheatham, T. E., Simmerling, C. L., Wang, J., Duke, R. E., Luo, R., Walker, R. C., Zhang, W., Merz, K. M., Roberts, B., Wang, B., Hayik, S., Roitberg, A., Seabra, G., Kolossváry, I., Wong, K. F., Paesani, F., Vanicek, J., Liu, J., Wu, X., Brozell, S. R., Steinbrecher, T., Gohlke, H., Cai, Q., Ye, X., Wang, J., Hsieh, M.-J., Cui, G., Roe, D. R., Mathews, D. H., Seetin, M. G., Sagui, C., Babin, V., Luchko, T., Gusarov, S., Kovalenko, A., and Kollman, P. A. (2010) AMBER 11.

[32]   Mongan, J., Simmerling, C., McCammon, J. A., Case, D. A., and Onufriev, A. (2007) Generalized Born model with a simple, robust molecular volume correction. *J. Chem. Theory Comput.* **3**, 156-169

[33]   Li, D. W., and Brüschweiler, R. (2010) NMR-based protein potentials. *Angew. Chem. Int. Edit.* **49**, 6778-6780

[34]   Gotz, A. W., Williamson, M. J., Xu, D., Poole, D., Le Grand, S., and Walker, R. C. (2012) Routine microsecond molecular dynamics simulations with AMBER on GPUs. 1. Generalized born. *J. Chem. Theory Comput.* **8**, 1542-1555

[35]   Kollman, P. A., Massova, I., Reyes, C., Kuhn, B., Huo, S. H., Chong, L., Lee, M., Lee, T., Duan, Y., Wang, W., Donini, O., Cieplak, P., Srinivasan, J., Case, D. A., and Cheatham, T. E. (2000) Calculating structures and free energies of complex molecules: Combining molecular mechanics and continuum models. *Accounts Chem. Res.* **33**, 889-897

[36]   Gronenborn, A. M., and Clore, G. M. (1994) Identification of N-terminal helix capping boxes by means of ^{13}C chemical shifts. *J Biomol. NMR* **4**, 455-458

[37]   Seale, J. W., Srinivasan, R., and Rose, G. D. (1994) Sequence determinants of the capping box, a stabilizing motif at the N-termini of a-helices. *Protein Sci.* **3**, 1741-1745

[38] Berjanskii, M. V., and Wishart, D. S. (2005) A simple method to predict protein flexibility using secondary chemical shifts. *J. Am. Chem. Soc.* **127**, 14970-14971

[39] Mammen, M., Choi, S. K., and Whitesides, G. M. (1998) Polyvalent interactions in biological systems: Implications for design and use of multivalent ligands and inhibitors. *Angew. Chem. Int. Edit.* **37**, 2755-2794

[40] Fisher, R. D., Wang, B., Alam, S. L., Higginson, D. S., Robinson, H., Sundquist, W. I., and Hill, C. P. (2003) Structure and ubiquitin binding of the ubiquitin-interacting motif. *J. Biol. Chem.* **278**, 28976-28984

[41] Safadi, S. S., and Shaw, G. S. (2010) Differential interaction of the E3 ligase parkin with the proteasomal subunit S5a and the endocytic protein Eps15. *J. Biol. Chem.* **285**, 1424-1434

Chapter 3

Molecular Basis for Phosphorylation Dependent SUMO

Recognition by the DNA Repair Protein RAP80

Introduction

The DNA repair process in eukaryotic cells is an indispensible life process responsible for maintaining the fidelity of the genome[1,2]. The genomic information encoded within the molecular structure of DNA is relentlessly compromised as a result of factors that include free radicals arising from metabolic processes, radiation, and replication errors[1]. By virtue of highly regulated and efficient DNA repair mechanisms, most DNA damage does not progress to viable malignant tumors[1]. Amongst the different kinds of damage that alter DNA structure, double strand breaks are the most deleterious[1]. Depending the nature of DNA damage, checkpoint activation and cell cycle arrest accompany a number of repair pathways, including homologous recombination, non-homologous end joining, or alternative non-homologous end joining repair, function to combat the damage[3]. Similar to many life processes, homologous recombination is governed by the hierarchical and synergistic action of different post-translational modifications, such as phosphorylation, ubiquitination, and SUMOylation[4,5]. The severed ends of damaged DNA are sensed by the MRE11/Rad50/NBS1 (MRN) protein complex, followed by recruitment of ATM kinase and its concomitant activation[3,5,6]. This results in the phosphorylation of nearby histones, which serves as a marker for initiation of repair[3,5]. Phosphorylated histones comprise the binding site for MDC1, which is also phosphorylated by ATM kinase; subsequently, phosphorylated MDC1 recruits the ubiquitination enzyme RNF8 which acts with the Ubc13/Mms2 heterodimer to attach K63-linked Ub chains at damage sites, in combination with the Ub ligase RNF168[3,7,8]. One of the biological functions for K63-linked polyUb chains is to serve as a signal for BRCA1 recruitment, a key protein that is obligatory for repair of DNA damage, and cell cycle checkpoint activation. RAP80, an 80 kDa nuclear protein, is responsible for recruitment of the BRCA1 A complex (BRCA1, BARD1, BRCC36, Abraxas and RAP80) to sites of DNA damage by binding K63-linked Ub chains through tandem a-helical UIMs[9,10]. In addition to ubiquitination, SUMOylation of different DNA repair proteins by PIAS4, a SUMO specific E3 ligase, is involved in BRCA1 recruitment[11,12]. The well established role for K63-linked polyUb recognition by

RAP80 in BRCA1 recruitment was recently modified by the finding that RAP80 possesses a SIM, N-terminal to the tandem UIMs, which is partly responsible for BRCA1 A complex recruitment to DNA damage sites[13,14]. Optimal recruitment of BRCA1 to damage sites depends on the combined action of the SIM and UIM of RAP80; this implies that there are two possibilities for SUMO and polyUb binding: independent recognition of the individual modifier proteins, or recognition of SUMO–Ub hybrid chains[13,14]. Hybrid chain recognition is appealing in comparison to independent modifier binding, as a result of an ~80-fold higher affinity for RAP80 as compared to SUMO and Ub binding alone[14]. In addition, the requirement of RNF4, a SUMO-binding Ub ligase, for BRCA1 recruitment by RAP80, suggests that hybrid chains are the preferred candidates for RAP80 binding[14].

There are four SUMO isoforms in mammalian cells- SUMO-1, 2, 3 and 4. SUMO-1 shows 45% sequence identity to SUMO-2 and SUMO-3, whereas SUMO-2 and SUMO-3 share 95% sequence identity, and can form polySUMO chains. The function of SUMO-4 is currently unknown. Although their sequences and chain forming capabilities vary, all SUMO isoforms assume a Ub-like fold[15]. From a structural perspective, the binding of SUMO to its cognate partners typically involves electrostatic and hydrophobic interactions, unlike Ub interactions, which typically involve a hydrophobic patch centered on Ile44[16-18]. The SIM is the most extensively studied SUMO binding motif, with a hydrophobic module (V/I)X(V/I)(V/I), bordered by N and C-terminal acidic modules[18]. The three isoforms of SUMO bind SIMs within a hydrophobic groove between the α_1 helix and the β_2 strand, typically forming an intermolecular β-sheet at the interface. The orientation of the β strand has been observed in both parallel and antiparallel conformations depending on the specific SIM sequence and SUMO isoform[19]. This is believed to result from the distribution of negatively charged residues adjacent to the hydrophobic SUMO-interacting module from the SIM. This region also possesses serine residues that are typically phosphorylation sites, and play a role in determining SUMO isoform preference[17,18]. Phosphorylation of the SIM serine residues provides enhanced electrostatic interactions, that generally result in a substantial increase in the affinity of the SUMO–SIM interaction[20,21]. A number of structures for

76

phosphorylated SIMs bound to SUMO-1 are have been reported[19,21-25]. However, the molecular basis for the interaction between SUMO-2 and its cognate phosphorylated SIM is unknown. The DNA repair protein RAP80 has been shown to preferentially interact with SUMO-2[14], and possesses a canonical CK2 phosphorylation site within its SIM. In this study, the structure of the N-terminal UIM and SIM domains from RAP80, as well as the molecular basis for binding of SUMO-2 to the SIM were investigated using NMR spectroscopy. We also determined the first structure of SUMO-2 bound to a phosphorylated SIM, which in conjunction with measurement of the thermodynamics and kinetics of SUMO-2 binding for the phosphorylated and non-phosphorylated states of RAP80, provide insight into the molecular determinants underlying the SUMO-2 specificity of this critical DNA repair protein.

Experimental Procedures

Cloning, protein expression and purification of RAP80

RAP80 is a 719 residue, multi-domain protein consisting of N-terminal nuclear localization signals (~residues 3–35), two N-terminal tandem UIMs (~residues 80 – 120), an N-terminal SIM (~residues 35–50), and in the C-terminal half, an Abraxas interacting region (AIR) and two putative zinc fingers[26]. The central AIR domain binds phosphorylated Abraxas within the BRCA1 complex[9,10,27]. This function, combined with the independent SUMO and Ub binding properties of the N-terminal region (~residues 30–120)[13,14,28,29], facilitates DNA damage recognition and repair by the BRCA1 complex. To study the SUMO binding properties of human RAP80, residues 33-131 were cloned into the EcoRI and BamHI sites of pHis-P1, and the insert sequence was verified by sequencing. Expression of the His$_6$-tagged fusion constructs results in an N-terminal GAMDP cloning artifact following cleavage with TEV protease. For expression of unlabeled proteins, 100μL of electrocompetent *Escherichia coli* strain BL21(DE3)-RIPL cells were transformed with 300 ng plasmid, and allowed to grow overnight on agar plates containing ampicillin and chloramphenicol at 37 °C. A single colony was picked and used to inoculate 50 mL of LB starter culture, which was incubated at 37 °C

overnight. LB containing ampicillin and chloramphenicol (500 ml) was inoculated with 5 mL of starter culture and incubated at 37 °C with shaking at 250 rpm. Upon optimal growth to OD_{600} ~0.6-0.8, cells were induced with 0.4 mM IPTG. Post induction, cells were grown overnight at 25 °C and subsequently harvested. Cells were suspended in 100 mL of lysis buffer containing 20 mM imidazole, 500 mM NaCl, 20 mM sodium phosphate, 2 mM DTT, 10 mM $MgSO_4$, 5 μg/mL DNase I and 0.5% protease inhibitor cocktail II (Calbiochem catalog no. 538132), pH 7.3, and subjected to sonication. Following cell rupture, lysate was centrifuged at 25000 rpm in a Beckman JA-25.5 rotor for 30 min at 4 °C. The supernatant was filtered using a Millipore steriflip 0.45 μm vacuum filtration unit. The filtrate was affinity-purified using a His-prep FF 16/10 column equilibrated with buffer containing 20 mM imidazole, 500 mM NaCl, 20 mM sodium phosphate and 2 mM DTT at pH 7.3. Bound protein was eluted using a gradient of increasing imidazole concentration ranging from 20 mM to 500 mM. Fractions containing protein, as detected by UV absorbance, were pooled and the His_6 tag was cleaved by addition of 100 μL of 210 μM TEV protease with incubation at 4 °C overnight. The cleaved His_6 affinity tag was removed by passing the sample over a His-prep FF 16/10 column; unbound protein was collected, and exchanged using a dialysis membrane with a 3.5 kDa cutoff, into buffer containing 50 mM Tris, 150 mM NaCl, 2 mM DTT, at pH 7.3. Final purification was carried out by size exclusion with a HiLoad 26/60 Superdex 75 column equilibrated with 50 mM Tris, 150 mM NaCl, and 2 mM DTT at pH 7.3. Fractions containing protein were pooled and concentrated using an Amicon Ultra 15 centrifugal membrane filtration device with a cutoff of 3 kDa. For expression of $[U$-$^{15}N]$ and $[U$-$^{13}C,^{15}N]$-labeled protein for NMR studies, cells were grown to OD_{600} ~0.6-0.8 in 2 L of LB media, and pelleted by centrifugation using a Beckman JA-10.5 rotor for 30 minutes at 5000 rpm. Cells were washed in M9 medium, and suspended in 250 mL of M9 media containing ^{15}N labeled ammonium sulphate as the sole nitrogen source or both ^{15}N labeled ammonium sulphate and ^{13}C labeled glucose as the sole carbon source. Cells were acclimatized to the change in media conditions for ~2 hours through incubation at 25 °C and shaking at 250 rpm. Protein expression was induced using 0.4 mM IPTG. Following induction, the incubation temperature was reduced to 18 °C for overnight growth. Purification was achieved as described for unlabeled proteins.

Protein expression and purification of SUMO-2

pET28a plasmid harboring residues 1-93 of human SUMO-2 was a gift from Dr. Lawrence McIntosh, University of British Columbia. Expressed SUMO-2 contained an N-terminal His_6 tag, which results in a GSH cloning artifact following cleavage by thrombin. Protein expression and purification strategies for unlabeled, [U-^{15}N] and [U-^{13}C,^{15}N] labeled SUMO-2 were similar to those for RAP80 constructs, except where noted. Thrombin cleavage of the SUMO-2 His_6 tag was carried out using a Thrombin CleanCleave kit (Sigma-Aldrich). Cleaved SUMO-2 was dialyzed against thrombin cleavage buffer containing 50 mM Tris HCl, and 2.5 mM $CaCl_2$ at pH 7.9, followed by incubation with thrombin immobilized agarose beads at room temperature for 48 hours. Post cleavage, thrombin beads were removed by centrifugation and the protein was dialysed in nickel column binding buffer (20 mM imidazole, 20 mM sodium phosphate, 150 mM NaCl, 2 mM DTT, pH 7.3). Purification by size exclusion chromatography was similar to that for RAP80 constructs. For all protein samples, purity and molecular weight were confirmed using SDS-PAGE and MALDI-TOF mass spectrometry.

RAP80 SIM peptide synthesis and phosphorylation

Peptide containing residues 35-50 (RLEDAFIVISDSDGEE) from RAP80, with an acetylated N-terminus and amidated C-terminus, was chemically synthesized (Biomatik, 99% pure). Doubly phosphorylated peptide was synthesized through incubation of 0.18 mM peptide with 1000 units of CK2 (Casein Kinase 2) (New England Biolabs), in reaction buffer containing 0.63 mM ATP, 50 mM Tris, 10 mM $MgCl_2$ 0.1 mM EDTA, 2 mM DTT, 0.01% Brij35, pH 7.5. The reaction was carried out at 30 °C for 4 hours. Complete phosphorylation at residues S44 and S46 was confirmed by MALDI-TOF mass spectrometry. For NMR-monitored binding studies and structure determination of SUMO-2 with phosphorylated RAP80 peptide complex, phosphopeptide containing RAP80 residues 37-49 (EDAFIVIpSDpSDGE) was chemically synthesized (Biomatik, 95.96% pure).

NMR chemical shift assignment

All NMR experiments were carried out on a Varian Unity INOVA 600 MHz spectrometer, except where noted. All samples were prepared in SHIGEMI microcell NMR tubes. Main chain resonance assignments for RAP80$_{33-131}$ were accomplished at 25° C using a [U-^{13}C,^{15}N] NMR sample in 50 mM Tris, 150 mM NaCl, 45 mM TCEP, and pH 6.1, with a protein concentration of 0.3 mM, 0.1 mM DSS as an internal reference, and 10% D$_2$O, and the 3D HNCACB[30,31], CBCA(CO)NNH[31,32], HNN(CA,CO)[33,34], HNCO[35], and HN(CA)CO[36] experiments. Backbone ^{1}H^N, ^{15}N, ^{13}C$_\alpha$, ^{13}CO and ^{13}C$_\beta$ chemical shifts were assigned manually using the SPARKY program [37] and verified using the automatic assignment feature of CARA[38]. These main chain chemical shift assignments were readily translated to pH 7.3 for subsequent NMR experiments, as pH-dependent chemical shift changes were minimal. Main chain resonance assignments for SUMO-2 were obtained in a similar fashion using a [U-^{13}C/^{15}N] NMR sample containing 0.4 mM SUMO-2 in 50 mM Tris, 150 mM NaCl, 2 mM DTT, 10% D$_2$O, and 0.1 mM DSS as an internal standard at pH 7.3. Chemical shifts were verified against those deposited in the BMRB (accession number 6801).

Main chain ^{1}H^N, ^{1}H$_\alpha$, and side chain proton chemical shifts for chemically synthesized RAP80$_{35-50}$ were assigned using 2D TOCSY and NOESY experiments[39] at 5 and 25 °C, for a sample containing 0.7 mM peptide in 50 mM Tris, 150 mM NaCl, 2 mM DTT, and 0.1 mM DSS, and 10 % D$_2$O, at pH 7.3. Chemical shifts for synthesized, doubly phosphorylated pRAP80$_{37-49}$, were assigned similarly using a sample containing 0.5 mM peptide in 50 mM Tris, 150 mM NaCl, 2 mM DTT, 0.1 mM DSS, and 10 % D$_2$O at pH 7.3.

Chemical shift perturbation mapping for [U-^{15}N]-RAP80$_{33-131}$ upon SUMO-2 addition

In addition to following SUMO-2 chemical shift changes upon titration with RAP80 peptides, we determined RAP80$_{33-131}$ main chain ^{1}H^N and ^{15}N chemical shift changes upon interaction with SUMO-2, by recording 2D ^{1}H-^{15}N HSQC NMR spectra for 0.3 mM [U-^{15}N]-RAP80$_{33-131}$ in 50 mM Tris 150 mM NaCl, 2 mM DTT, and 10% D$_2$O at pH 6.1 until there were no significant changes in RAP80 resonances upon SUMO-2 addition (>10-fold excess of SUMO-2).

NMR monitored titrations for [U-^{15}N]-SUMO-2 with RAP80$_{33-131}$, RAP80$_{35-49}$, and pRAP80$_{37-49}$

Generally, changes in main chain amide chemical shifts for SUMO-2 upon addition of various RAP80 peptides were followed at different titration points using 2D ^{1}H-^{15}N HSQC NMR spectra [40,41]. Titration of 225 μM [U-^{15}N] SUMO-2 in 50 mM Tris, 150 mM NaCl, 10 mM DTT, and 10% D$_2$O, at pH 7.3 was carried out by adding increasing amounts of stock solution containing unlabeled RAP80$_{33-131}$ in the same buffer. During titration, the increase in RAP80$_{33-131}$ is accompanied by dilution of SUMO-2 with concentrations of 225, 211, 198, 184, 171, 157, 144, 117, 91, 64 and 11 μM, and RAP80$_{33-131}$ concentrations of 0, 35, 70, 104, 139, 174, 208, 278, 347, 416 and 555 μM. For these titration points, the RAP80$_{33-131}$/SUMO-2 concentration ratios were 0, 0.16, 0.35, 0.56, 0.81, 1.12, 1.44, 2.36, 3.82, 6.49 and 51.98. For this, and all subsequent titrations, the combined ^{15}N and ^{1}H^N chemical shift change for each residue was calculated according to $\Delta\delta = [(\Delta\delta^{15}N/5)^2 + (\Delta\delta^1H^N)^2]^{1/2}$, where $\Delta\delta^{15}N$ and $\Delta\delta^1H^N$ are the respective ^{15}N and ^{1}H^N chemical shift changes in ppm. These combined chemical shift changes were numerically fit to a 1:1 binding isotherm to determine the dissociation constant (K_D), as previously described [41]. For all titrations, protein concentrations were measured using the BCA assay and verified with amino acid analyses. For NMR-monitored titrations of SUMO-2 with RAP80$_{35-50}$, a stock solution of unlabeled RAP80$_{35-50}$ was titrated into a 380 μM [U-^{15}N] SUMO-2 sample in 50 mM Tris, 150 mM NaCl, 2

mM DTT, and 10% D_2O, at pH 7.3. The concentrations of SUMO-2 and RAP80$_{35-50}$ at the various titration points were 376, 373, 368, 359, 342, 324, 303, 288, 253, 217, 181, 145, 108, 72, 36 μM, and 0, 5, 15, 31, 64, 98, 136, 163, 229, 296, 364, 431, 498, 565, and 633 μM, respectively, giving concentration ratios for RAP80$_{35-50}$/SUMO-2 of 0, 0.013, 0.040, 0.08, 0.18, 0.30, 0.44, 0.56, 0.90, 1.36, 2.01, 2.97, 4.6, 7.84 and 17.5(41).

Titration of SUMO-2 with synthetic pRAP80$_{37-49}$ peptide was carried out by adding increasing concentrations of peptide to a 190 μM [U-^{15}N] SUMO-2 sample in 50 mM Tris, 150 mM NaCl, 2 mM DTT, and 10% D_2O, at pH 7.3. Concentrations of SUMO-2 and pRAP80$_{37-49}$ for the different titration points were 190, 188, 185, 179, 172, 159, 147, 136, 124, 113, 102, 91, 80, 69, 68, 61, 51, 46 μM, and 0, 1, 4, 7, 12, 22, 30, 38, 46, 54, 62, 70, 78, 86, 88, 100, 120, 135 μM, respectively, with corresponding concentration ratios for pRAP80$_{37-49}$/SUMO-2 of 0, 0.006, 0.02, 0.04, 0.07, 0.14, 0.20, 0.28, 0.37, 0.47, 0.60, 0.77, 0.97, 1.24, 1.29, 1.63, 2.35 and 2.93.[41].

Chemical shift differences were calculated for the first and last points from the titrations of SUMO-2 with various RAP80 peptides, and mapped onto the structure of free SUMO-2 (PDB ID: 1WM2). For these chemical shift maps, common missing residues included N- and C-terminal residues and prolines: M1, A2, D3, N15, D16, Q90-G93, P6, P39, P66, P73, and broadened residues: K33, K35. For the titration with RAP80$_{33-131}$, missing residues in the ^{1}H-^{15}N NMR spectra due to low SUMO-2 concentration include E4, Q25, R36, R50, M55, and E81.

NMR structure determination for RAP80$_{33-131}$ and the SUMO-2/pRAP80$_{37-49}$ complex

The secondary structure for RAP80$_{33-131}$ was determined using protein backbone (ϕ, ψ) and side chain (χ^1) torsion angles determined with the TALOS-N program], using quantitative analysis of backbone ^{1}H^N, ^{1}H$_\alpha$, ^{13}CO, ^{13}C$_\alpha$ and side chain ^{13}C$_\beta$ chemical shifts.

To determine a structure for pRAP80$_{37-49}$ in complex with SUMO-2, a sample containing 0.4 mM [U-^{13}C/^{15}N]-SUMO-2 and 1.2 mM pRAP80$_{37-49}$ was prepared in buffer containing

25 mM Tris, 2 mM DTT and 10% D_2O, pH 7.3. For assignment of bound SUMO-2 backbone $^1H^N$, ^{15}N, $^{13}C_\alpha$, ^{13}CO, and side chain $^{13}C_\beta$ chemical shifts, 3D HNCACB, CBCA(CO)NNH, HNCO, HN(CA)CO, and 2D ^1H-^{15}N HSQC NMR experiments were collected. Side chain carbon and proton chemical shifts were obtained using the 3D (H)CCTOCSY(CO)NNH and H(CC)TOCSY(CO)NNH NMR experiments[43-46]. Chemical shift assignment was accomplished manually using the SPARKY NMR software, and automatically using CARA. The six aromatic side chains of SUMO-2 H17, F32, H37, Y47, F60, F62 and F87, were assigned using an in-house modified 2D ^1H-^{13}C aromatic TROSY HSQC[47], an in-house 2D ^1H-^{13}C aromatic NOESY TROSY HSQC pulse sequence, as well as the 2D (HB)CB(CGCD)HD and (HB)CB(CGCDCE)HE aromatic side chain correlation experiments[48]. To assign pRAP80$_{37-49}$ side chain and backbone proton chemical shifts for peptide bound to SUMO-2, 2D TOCSY and NOESY experiments[49] with suppression of SUMO-2 signals from protons bound to ^{13}C and ^{15}N were carried out on a Unity INOVA 800 MHz NMR spectrometer using the same sample. To determine interacting residues between [U-^{13}C/^{15}N]-SUMO-2 and unlabeled pRAP80$_{37-49}$, 2D ^{13}C/^{15}N F_1-filtered, F_3-edited NOESY experiments for aliphatic and aromatic protons were carried out[49]. Intermolecular NOEs were assigned using side chain proton assignments for bound SUMO-2, and side chain proton assignments for bound pRAP80$_{37-49}$. A total of fifteen intermolecular NOEs between SUMO-2 and pRAP80$_{37-49}$ were assigned, as well as 28 intramolecular pRAP80$_{37-49}$ NOEs. Restraints for NOEs were determined on the basis of peak intensities, and assigned to distance ranges using the auxiliary programs from the Amber 14 biomolecular simulation suite of programs[50]. Main chain torsion angle restraints for bound SUMO-2 were determined with the TALOS-N program using main chain $^1H^N$, $^1H_\alpha$, ^{13}CO, $^{13}C_\alpha$ and side chain $^{13}C_\beta$ chemical shifts. Torsion angles for bound peptide were determined from $^1H^N$ and $^1H_\alpha$ chemical shifts using the PREDITOR program[51]. Torsion angle restraints were used in structure calculations for residues with a PREDITOR confidence score higher than 0.7.

For structure calculations, the starting model for SUMO-2 was derived from the high-resolution crystal structure (PDB ID: 1WM2)[52]. N-terminally acetylated and C-terminally amidated pRAP80$_{37-49}$ peptide was manually docked to strand b$_2$ from SUMO-

2, in the parallel b-strand conformation, to generate a starting model for the complex. The choice for this initial model was based on the following experimental observations, that suggested a parallel b-strand conformation for peptide: intermolecular NOEs between SUMO-2 and RAP80 residues I41/I43, combined with a lack of intermolecular NOEs between F40 and V42 from RAP80 and SUMO-2, significant main chain amide chemical shift perturbations for a network of positively charged residues including H17, K35, H37, and K42 for SUMO-2, and corresponding large ^1H_a and $^1H^N$ chemical shift changes for pS44, D45, and pS46 for SIM peptide, upon SUMO-SIM interaction. This initial model was solvated in a truncated octahedral TIP3P water box, with a distance 24 Å between protein atoms from images in adjacent unit cells. The starting model was energy minimized using the *ff*14SB force field and the sander program within the Amber 14 suite of biomolecular simulation programs, with pairwise long range electrostatics and van der Waals interactions cut off at 8 Å. In addition, default parameters for phosphoserine bearing a –2 charge (S2P residue) were employed[53]. The simulation system was heated for 50 ps to a temperature of 298 K, solute atoms were subjected to 2 kcal/mol restraints, and allowed to equilibrate to 1 atm pressure for a further 50 ps. The system was then subjected to production dynamics for 40 ps with the inclusion of NMR-derived distance and dihedral restraints. Structural statistics for twenty snapshots from the last 40 ps of the simulations with NMR restraints are given in Table 3.1.

	pRAP80$_{37\text{-}49}$	SUMO-2
distance restraints		–
total	30	–
intraresidue	3	–
sequential ($|i\text{–}j| = 1$)	21	–
medium ($2 \leq |i\text{–}j| \leq 4$)	6	–
long ($|i\text{–}j| \geq 5$)	0	–
Intermolecular	15	15
dihedral restraints	11 ϕ, 11 ψ, 0 χ^1	68 ϕ, 68 ψ, 30 χ^1
restraint violations		
distance > 0.5 Å	0	0
dihedral > 5 °	4	8
ϕ/ψ in the most-favoured region (%)	88.8	

Table 3.1: Structural statistics for twenty NMR-derived structures for SUMO-2/pRAP80$_{37\text{-}49}$.

NMR lineshape analyses

For titrations of SUMO-2 with RAP80$_{35-50}$ and pRAP80$_{37-49}$, lineshape analyses were carried out using the Bloch-McConnell equations for two-site chemical exchange, as previously described[54], to yield the kinetics (k_{on} and k_{off}) of the RAP80 SIM/SUMO-2 interaction, as well as changes in kinetics upon phosphorylation.

Results

Secondary structure for RAP80$_{33-131}$ from quantitative chemical shift analysis

The Ub and SUMO binding properties of RAP80 are confined approximately to residues 30–120[13,14,28,29], with the tandem UIMs (residues 80–120) forming α-helical structure[55]. In order to determine the molecular basis of SUMO-2 recognition by residues 30–50 from RAP80, we explored the structure of free RAP80$_{33-131}$ using NMR spectroscopy. 2D ^{1}H-^{15}N HSQC NMR spectra for RAP80$_{33-131}$ are shown in Fig. 3.1a. The ^{1}H^N chemical shifts range from ~7.7-8.9 ppm; this relatively narrow dispersion suggests that for RAP80$_{33-131}$, other than the tandem α-helical UIM domains, a specific global fold is not adopted. Quantitative chemical shift analysis of main chain torsion angles using the TALOS-N program indicates that residues 33-39, N-terminal to the SIM, adopt a random coil conformation (Fig. 3.1b), and are flexible, as indicated by random coil index derived S^2 values of ~0.50 [56]. TALOS-N chemical shift analysis indicates that residues 40-47 that form the SIM, adopt a β strand conformation with a high probability of ~0.95 (Fig. 3.1b). Residues 48-78 that connect the SIM to the tandem UIMs also adopt a primarily random coil conformation, with higher flexibility, as indicated by an average random coil derived S^2 value of ~0.3. Residues belonging to the tandem UIMs adopt α-helical conformations, as previously described (28,29). The LR motif (residues 60-78), believed to assist in recruitment of BRCA1 to damage sites [57], does not adopt a specific secondary structure and appears flexible, at least in uncomplexed states.
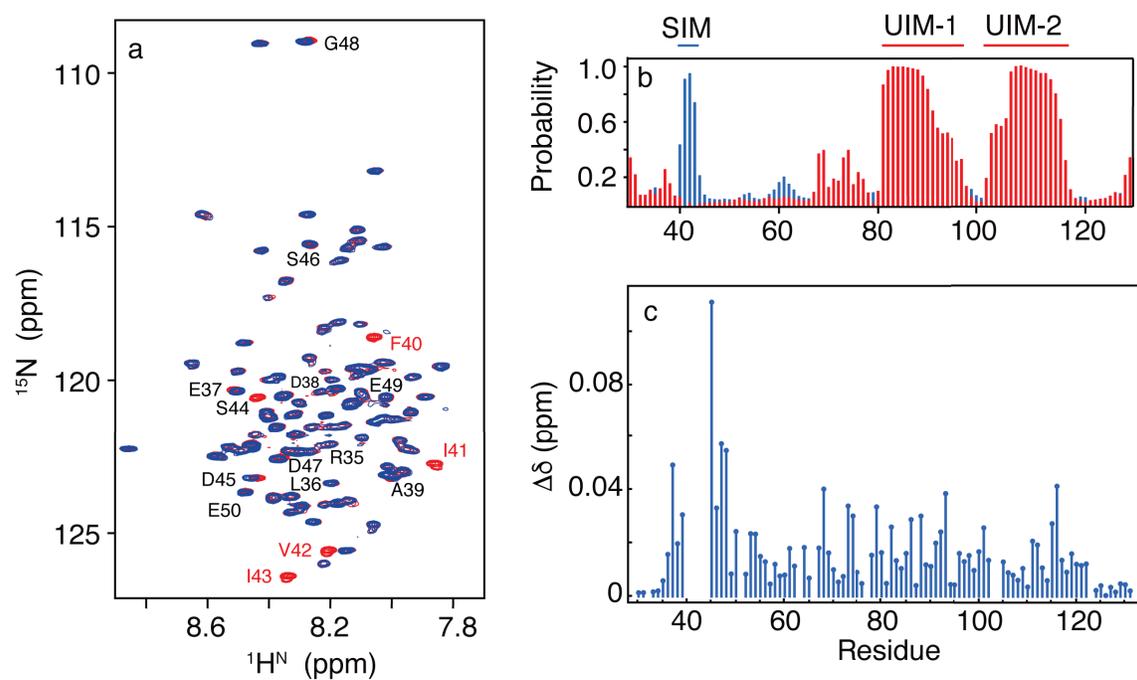
Figure 3.1: (a), 2D ^{1}H–^{15}N HSQC NMR spectra for free RAP80$_{33\text{-}131}$ (*red*) and bound to SUMO-2 (*blue*). Residues 40–44 are broadened beyond detection upon SUMO-2 binding. (b), secondary structure for free RAP80$_{33\text{-}131}$ derived from quantitative chemical shift analysis, with a-helix shown in *red*, and b-strand shown in *blue*. (c), chemical shift perturbations for RAP80$_{33\text{-}131}$ bound to an eleven-fold excess of SUMO-2.

Chemical shift perturbation mapping for [U-^{15}N]-RAP80$_{33-131}$ upon interaction with SUMO-2

To delineate the N-terminal residues of RAP80 that interact with SUMO-2, we followed changes in main chain amide resonances for RAP80$_{30-131}$ upon SUMO-2 binding using 2D ^{1}H-^{15}N HSQC NMR spectroscopy (Fig. 3.1a). In general, a significant chemical shift change accompanying a protein-protein interaction can be expected to have a threshold value of ~0.1 ppm, up to a maximum of ~0.5 ppm [55]. There are few residue-specific, significant chemical shift changes for RAP80$_{33-131}$ within the typical range (Fig. 3.1c). However, the main chain ^{1}H^N and ^{15}N resonances for residues expected to be directly involved in the interaction with SUMO-2, F40, I41, V42 and I43, could not be detected upon interaction with SUMO-2 due to extensive line broadening. The resonances for residues adjacent to the hydrophobic SIM residues, on the other hand, S44, D45, S46, and D47, shifted linearly with increasing SUMO-2, without severe line broadening, consistent with fast chemical exchange. These results suggest that RAP80 interacts with SUMO-2 exclusively *via* the SIM, similar to the interaction of RAP80 with Ub, wherein only UIM residues are directly involved in binding [29].

NMR-monitored titrations for the [U-^{15}N]-SUMO-2/RAP80$_{33-131}$ interaction

To gain insight into the molecular basis underlying specificity of the SUMO-2/RAP80 interaction, we employed NMR spectroscopy to determine the RAP80 binding site on SUMO-2 using chemical shift mapping, and to determine a quantitative dissociation constant for the interaction. 2D ^{1}H-^{15}N HSQC NMR spectra of SUMO-2 in the absence and presence of RAP80$_{33-131}$ are shown in Fig. 3.2a with per residue combined chemical shift changes shown in Fig. 3.2b. SUMO-2 main chain amide resonances showing significant changes upon interaction with RAP80$_{33-131}$ are located within strand β_2, helix α_1, the loop connecting them, as well as loop residues near the 3_{10} helix in free SUMO-2 (Fig. 2c). These regions form the typical SIM interaction site for the various SUMO isoforms[18]. Residues K33 and K35 are undetectable as a result of line broadening, and remained unobservable upon saturation of the binding site. Other residues in the binding

cleft, including V30 (Fig. 3.2d,e), Q31, F32, I34, R36, T38, L40, S41, K42, and L43, showed linear chemical shift changes which could be followed throughout the titration. The main chain amide resonances for F32 and L40 were broadened at high RAP80$_{33-131}$/SUMO-2 ratios. Fitting chemical shift changes for twenty SUMO-2 residues that could be followed during the titration to a 1:1 binding isotherm gives an average dissociation constant (K_D) of 195 ± 33 μM.

Figure 3.2: (a), 2D ^{1}H–^{15}N HSQC NMR spectra for free SUMO-2 (*red*) and bound to RAP80$_{33-131}$ (*blue*). (b), chemical shift perturbations for SUMO-2 bound to a 52-fold excess of RAP80$_{33-131}$. (c), Residues experiencing chemical shift changes greater than one standard deviation from the mean (*red*) and those broadened beyond detection (*orange*) are mapped on the surface of SUMO-2, between α$_1$ helix and β$_2$ strand. (PDB ID 1WM2). (d), fits of chemical shift perturbation data to 1:1 binding isotherms for the SUMO-2 interaction with RAP80$_{33-131}$. (e), Chemical shift changes for V30 are indicated on the vertical axis, and concentrations for SUMO-2 and unlabeled RAP80$_{33-131}$ are indicated on the horizontal axes. Experimentally determined chemical shift changes are shown as points, and the best fit to a 1:1 binding isotherm is shown as a surface. (f), expanded region from the 2D ^{1}H–^{15}N HSQC NMR spectra taken during titration of SUMO-2 with RAP80$_{33-131}$, showing chemical shift changes for V30.

Chemical shift perturbation mapping, NMR monitored titrations, and lineshape analysis for interaction of [U-^{15}N]-SUMO-2 with RAP80$_{35-50}$

To facilitate NMR and structural studies, we synthesized peptides encompassing the minimal binding motif from RAP80 (residues 35-50 or 37-49), given the lack of structure adjacent to these regions, and that no significant chemical shift changes are observed beyond these regions in NMR binding studies employing the longer RAP80$_{33-131}$ construct. The synthetic peptides were N-terminally amidated and C-terminally acetylated to maintain neutrality for the terminal residues, thereby avoiding introduction of non-physiologically relevant electrostatic interactions. To ensure that the peptides retained the SUMO interactions of the longer RAP80$_{33-131}$ construct, we conducted chemical shift mapping and NMR-monitored titrations to determine the binding site of the RAP80$_{35-50}$ peptide on SUMO-2, as well as the K_D and kinetic constants for the interaction, as described in detail below.

Main chain amide ^{1}H^N chemical shifts from SUMO-2 residues showing significant changes upon interaction with RAP80$_{35-50}$ are evident in 2D ^{1}H-^{15}N HSQC NMR spectra (Fig. 3.3a,b), and the results recapitulate those for titrations with the longer RAP80$_{33-131}$ construct. Similar to the RAP80$_{33-131}$–SUMO-2 interaction, residues K33 and K35 could not be detected due to extensive line broadening. Residues in the vicinity of K33 and K35, such as F32, I34, R36, H37, and T38 showed linear main chain amide chemical shift changes upon interaction with RAP80$_{35-50}$. Fitting chemical shift changes for seventeen residues to a 1:1 binding isotherm yields an average K_D of 239 ± 54 μM (Fig. 3.3c), with corresponding linear chemical shift changes for V30 shown in Fig. 3d. The K_D for the SUMO-2 interaction with RAP80$_{35-50}$ is the same within error for interaction of the longer RAP80$_{33-131}$ construct with SUMO-2 (195 ± 33 mM). We conducted lineshape analysis for the chemical shift changes of V30 during the titration to determine the kinetics of interaction (Fig. 3.3e). Fits of the lineshape changes to two site chemical exchange using the Bloch-McConnell equations yields a k_{on} value of 5.9 x 10^6 M^{-1} s^{-1}, and a k_{off} value of 1511 s^{-1}.
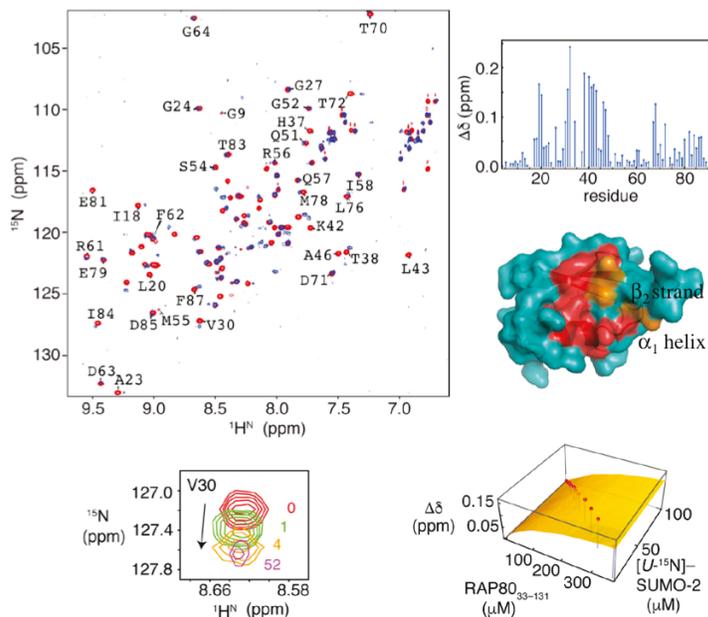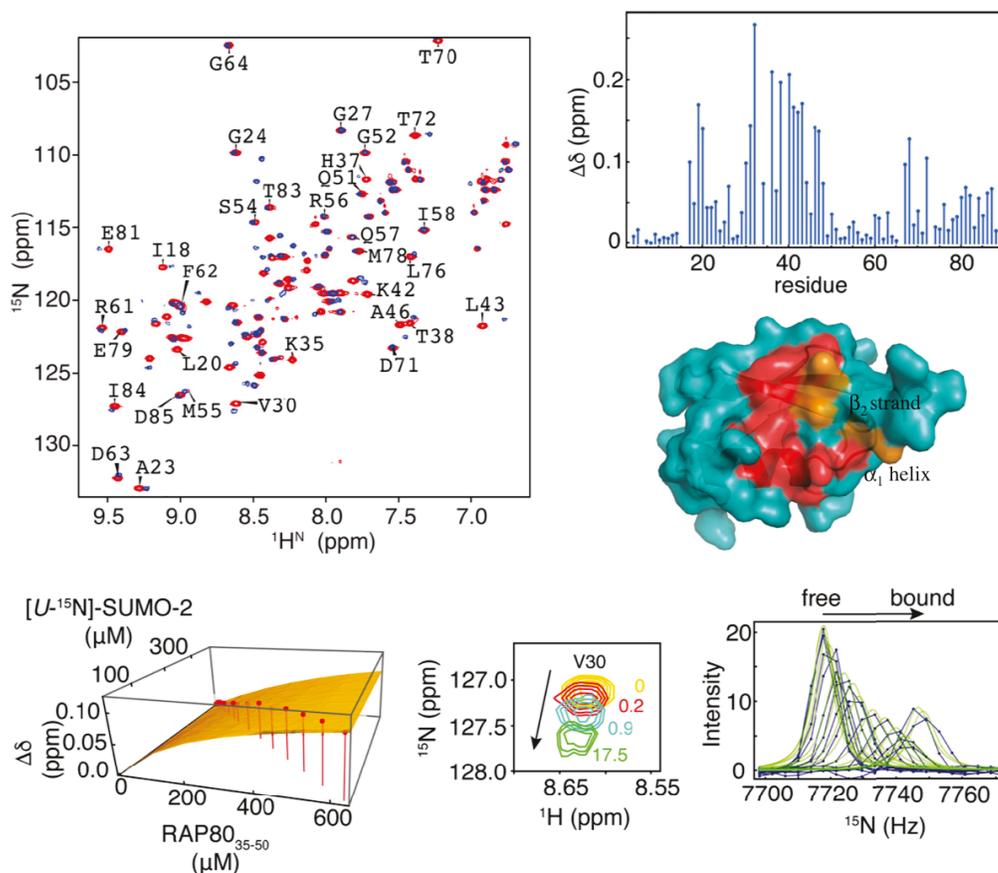
Figure 3.3: (a), 2D ^{1}H–^{15}N HSQC NMR spectra for free SUMO-2 (*red*) and bound to RAP80$_{35\text{-}50}$ (*blue*). (b), Chemical shift perturbations for SUMO-2 bound to a seventeen-fold excess of RAP80$_{35\text{-}50}$. (c), Residues experiencing chemical shift changes greater than one standard deviation from the mean (*red*) and those broadened beyond detection (*orange*) are mapped on the surface of SUMO-2, between α_1-helix and β_2-strand. (PDB ID: 1WM2). (d), fits of chemical shift changes to 1:1 binding isotherms for the SUMO-2 interaction with RAP80$_{35\text{-}50}$. Chemical shift changes for V30 are indicated on the vertical axis, and concentrations for SUMO-2 and unlabeled RAP80$_{35\text{-}50}$ are indicated on the horizontal axes. Experimentally determined chemical shift changes are shown as points, and the best fit to a 1:1 binding isotherm is shown as a surface. (e), expanded region from 2D ^{1}H–^{15}N HSQC NMR spectra taken during titration of SUMO-2 with RAP80$_{35\text{-}50}$, with ligand/protein ratios indicated. (f), lineshape analysis for V30 ^{15}N chemical shift changes taken from 2D ^{1}H–^{15}N HSQC NMR spectra taken during titration of SUMO-2 with RAP80$_{35\text{-}50}$, experimental data are shown as *blue* dots connected by lines, and the best fits are shown as *green* lines.

Chemical shift perturbation mapping for RAP80$_{35-50}$ upon phosphorylation

RAP80 contains a canonical CK2 recognition motif [S/T]-X-X-[D/E][58] adjacent to hydrophobic module of the SIM, that possesses the sequence: S$_{44}$D$_{45}$S$_{46}$D$_{47}$. We assessed the phosphorylation of RAP80 by CK2 using ^{1}H NMR spectroscopy. N-terminally acetylated, C-terminally amidated RAP80$_{35-50}$ was doubly phosphorylated at Ser44 and Ser46 with CK2. Phosphorylation of the two serines was confirmed by mass spectrometry, and the observation of significant chemical shift changes for the backbone and side chain protons from Ser44/Ser46 (Fig. 3.4). In addition, a significant change for the ^{1}H^N chemical shift for D47 is also observed. The ~0.3 ppm downfield shift for the ^{1}H^N of serine upon phosphorylation is suggestive of a hydrogen bond between the amide proton and phosphoryl group, as previously observed [59].
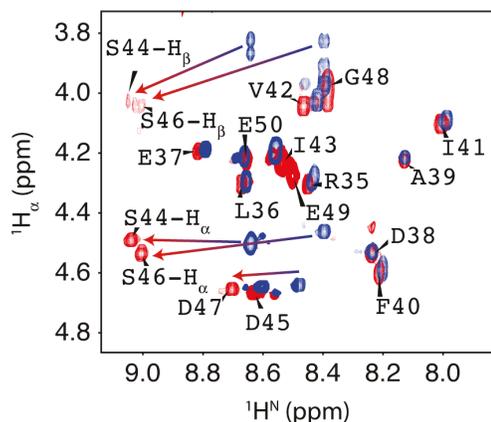
Figure 3.4: ^{1}H$_a$–^{1}H^N fingerprint region from the NMR spectra for RAP80$_{35-50}$ (*blue*) and phosphorylated RAP80$_{35-50}$ (*red*).

Chemical shift perturbation mapping, NMR monitored titrations, and lineshape analysis for the SUMO-2 interaction with pRAP80$_{37-49}$

In order to assess the impact of RAP80 phosphorylation on the specificity of the RAP80/SUMO-2 interaction, we determined the RAP80 binding site on SUMO-2, as well as the dissociation and kinetic rate constants using NMR spectroscopy. The interaction of [U-^{15}N] SUMO-2 with SIM peptide phosphorylated at S44 and S46 (pRAP80$_{37-49}$) was studied using chemical shift titrations of SUMO-2 with pRAP80$_{37-49}$. The 2D ^{1}H-^{15}N HSQC NMR spectra, and associated chemical shift maps for SUMO-2 in the absence and presence of pRAP80$_{37-49}$ are shown in Figs. 3.5a,b. For the interaction between pRAP80$_{37-49}$ and SUMO-2, the binding site is similar to that for unphosphorylated RAP80$_{35-50}$. However, the NMR resonances for residues F32, I34, R36, T38, L40, and L43 showed extensive line broadening and concomitant signal loss for [pRAP80$_{37-49}$]:[SUMO-2] ratios<1. For peptide/protein ratios>1, the resonances appear less broad, and upon saturation, become more intense in comparison to titrations with unphosphorylated peptide. This behaviour indicates that the SUMO-2 interaction with RAP80 SIM peptide becomes more specific upon phosphorylation. However, the main chain amide chemical shifts for residues K33 and K35 remain broad, even at saturating concentrations of phosphorylated SIM peptide. For ten residues that could be followed during the titration, the average K$_D$ value was determined to be 9±3 μM (Fig. 3.5c,d), a ~25 fold increase in binding affinity upon phosphorylation. From lineshape analysis (Fig. 5e), the values of k$_{on}$ and k$_{off}$ were determined to be 1.1 x 10^8 M^{-1} s^{-1} and 1020 s^{-1}, respectively. Compared to the kinetics of binding for unphosphorylated RAP80 SIM, k$_{on}$ increases by ~100 fold, whereas k$_{off}$ decreases by 1.5 fold.
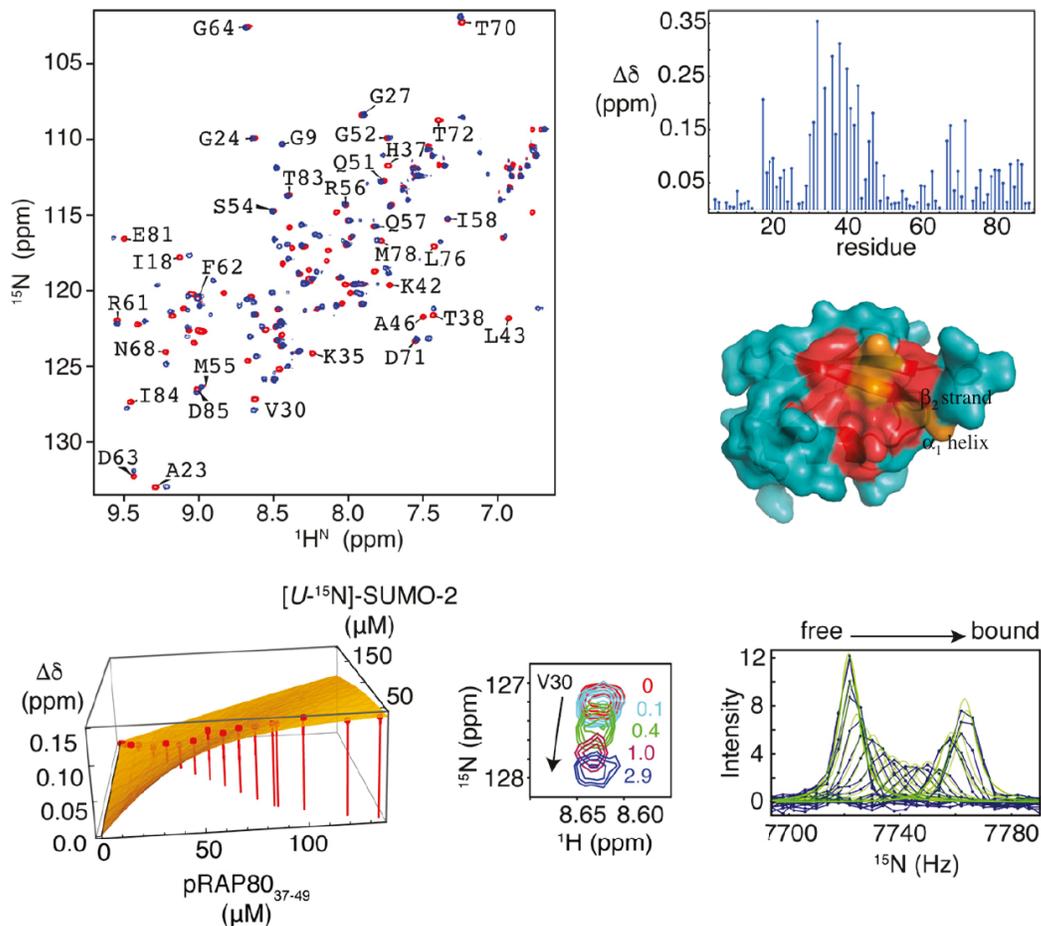
Figure 3.5: (a), 2D ^{1}H–^{15}N HSQC NMR spectra for free SUMO-2 (*red*) and bound to doubly phosphorylated RAP80$_{37-49}$ (*blue*). (b), chemical shift perturbations for SUMO-2 bound to a three-fold excess of pRAP80$_{37-49}$. (c), residues experiencing chemical shift changes greater than one standard deviation from the mean (*red*) and those broadened beyond detection (*orange*) are mapped on the surface of SUMO-2, between α_1-helix and β_2-strand. (PDB ID: 1WM2). (d), fits of chemical shift changes to 1:1 binding isotherms for the SUMO-2 interaction with pRAP80$_{37-49}$. Chemical shift changes for V30 are indicated on the vertical axis, and concentrations for SUMO-2 and unlabeled pRAP80$_{37-49}$ are indicated on the horizontal axes. Experimentally determined chemical shift changes are shown as points, and the best fit to a 1:1 binding isotherm is shown as a surface. (e), expanded region from 2D ^{1}H–^{15}N HSQC NMR spectra taken during titration of SUMO-2 with pRAP80$_{37-49}$, with ligand/protein ratios indicated. (f), lineshape analysis for V30 ^{15}N chemical shift changes taken from 2D ^{1}H–^{15}N HSQC NMR spectra taken during titration of SUMO-2 with pRAP80$_{37-49}$, experimental data are shown as *blue* dots connected by lines, and the best fits are shown as *green* lines.

NMR structure for the SUMO-2/pRAP80$_{37-49}$ complex

In general, the timescale for exchange between the free and bound forms of SUMO/SIM complexes results in significant NMR resonance broadening, rendering structural studies of the complex difficult[24]. For the SUMO-2/pRAP80$_{37-49}$ complex, intermolecular NOEs between the hydrophobic region of the SIM ($F_{40}I_{41}V_{42}I_{43}$), and NMR experiments selective for aromatic residues on SUMO-2, particularly F32 at the interface, help define the structure of the complex for this region. In contrast, for the electrostatic binding region of the interface, extensive line broadening for SUMO-2 residues, and the inability to obtain NOE-based NMR distance restraints between SIM phosphate groups and SUMO-2, present a significant challenge to NMR structure determination. To overcome these difficulties, we adopted an NMR-guided structural approach wherein sparse NMR data are combined with molecular dynamics simulations using a modern force field, in explicit solvent, to determine a model for the structure. To that end, quantitative main chain ϕ and ψ dihedral angles for both SUMO-2 and doubly phosphorylated RAP80 SIM peptide, intramolecular peptide NOEs, and intermolecular SIM–SUMO-2 NOEs, in combination with the highly developed AMBER *ff*14SB molecular dynamics forcefield[50], facilitated the calculation of NMR structures for the complex. A representative SUMO-2–pRAP80$_{37-49}$ structure (Fig. 3.6a) highlights the parallel SIM orientation with phosphorylated serine residues interacting with a basic region within the SIM binding cleft from SUMO-2. The parallel β-strand from pRAP80$_{37-49}$ completes a β sheet in the intermolecular complex with strand β_2 from SUMO-2, consistent with large, positive chemical shift changes ($\Delta\delta$ = bound–free) in the $^1H^N$–^1H_a fingerprint region of the bound peptide (Fig. 3.6b,c)[60]. The fingerprint region also indicates that the N and C termini of the peptide are flexible and not involved in binding, as supported by a lack of significant changes for the main chain proton chemical shifts of E37, D38, A39, G48 and E49 upon binding SUMO-2 (Fig. 3.6b,c). At the N-terminus, A39 and F40 side chain atoms do not directly contact SUMO-2. The side chain from SIM residue I41 forms multiple intermolecular contacts with the side chains of V30, F32, A46, and R50, from SUMO-2. In addition, SIM residue I43 contacts the side chains from SUMO-2 residues K42 and L43. The phosphate group from S44 is involved in electrostatic interactions with

the side chains from K35 and H17 from SUMO-2. This close interaction is supported by an ~0.5 ppm downfield shift for the histidine side chain $H^{\delta 2}$ protons, as shown in Fig. 3.6d. D45 from pRAP80$_{37-49}$ is involved in favourable electrostatic interactions with K42 and the T38 hydroxyl group from SUMO-2. The pS46 phosphate group from the SIM peptide shows favourable electrostatic interactions with the side chains of H37 and K35 from SUMO-2. The involvement of H37 in the interaction is supported by a downfield shift for the histidine $H^{\delta 2}$ protons by ~0.5 ppm (Fig. 3.6d). At the C-terminus of the SIM peptide, residues D47, G48, and E49 appear flexible, consistent with narrower $^{1}H^{N}$ resonances in $^{13}C/^{15}N$-filtered experiments for the protein peptide complex, in comparison to directly interacting residues. The atomic coordinates (code 2N9E) have been deposited in the Protein Data Bank (http://wwpdb.org/). NMR chemical shifts have been deposited in the BioMagRes Bank, www.bmrb.wisc.edu (accession no. 104587).
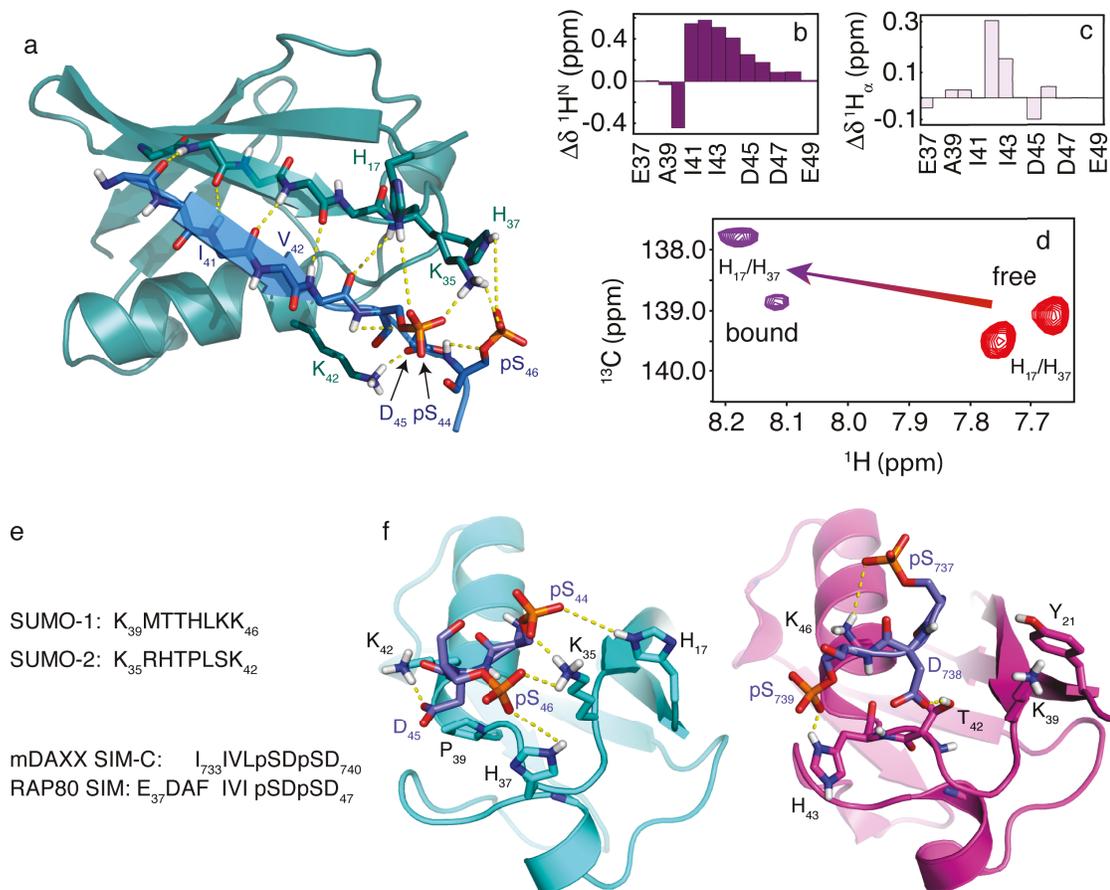
Figure 3.6: (a), NMR structure for the SUMO-2/pRAP80$_{37-49}$ complex. The main chain atoms are shown in the schematic representation (RAP80 SIM *blue*, SUMO-2 *teal*). Key electrostatic interactions are indicated by yellow dashed lines, and include hydrogen bonds between the main chain atoms of the SIM hydrophobic module, and interactions between the SUMO-2 specificity module and the negatively charged side chains of the SIM. (b), $^1H^N$ main chain amide chemical shift changes (bound–free) for pRAP80$_{37-49}$ upon binding of SUMO-2. (c), ^1H_a main chain chemical shift changes (bound-free) for pRAP80$_{37-49}$ upon binding of SUMO-2. (d), 2D ^1H–^{13}C HSQC NMR spectra showing side chain chemical shifts for the SUMO-2 H^{d2} atoms from residues H17 and H37 in the free state (*red*) and upon interaction with pRAP80$_{37-49}$ (*purple*). (e), sequence alignment for the SIM binding loop from SUMO-1 and -2, and the alignment for the SIM from RAP80 and that from "modified" DAXX which contains residues from the PML SIM N-terminal to I$_{733}$ to facilitate crystallization. (f), Comparison of the structure of the SUMO-2/pRAP80$_{37-49}$ complex with the SUMO-1/pDAXX$_{733-740}$ complex (PDB ID: 4WJP). SUMO-1 and -2 are shown in *purple* and *cyan*, respectively, the electrostatic SUMO recognition modules from the pSIMs are shown in *blue*. Dashed *yellow* lines highlight key intermolecular electrostatic interactions. The hydrophobic SUMO recognition modules from the pSIMs are not shown for clarity.

Discussion

For RAP80$_{33-131}$, 2D ^{1}H-^{15}N NMR spectra, and quantitative chemical shift analyses were used to derive per residue flexibility and main chain secondary structure, and indicate that the N-terminal region of RAP80 consists of three independent domains. These domains include a partially structured SIM, and two partially structured, tandem UIMs, surrounded by flexible regions. From chemical shift perturbation mapping and NMR-monitored titrations for various RAP80 constructs, we observe that the minimal region sufficient for interaction with SUMO-2 encompasses residues 40-47 from the RAP80 SIM. Similar to Ub binding to the tandem UIMs [29], the RAP80 SIM binds SUMO-2 independently. Furthermore, we determined that residues 50–78 between the SIM and tandem UIMs are not involved in Ub binding. The affinity of the RAP80–SUMO-2 interaction is weak, with a K_D of ~200 mM, and interestingly, double phosphorylation of the SIM at S44 and S46 gives rise to a ~25-fold increase in the affinity of the interaction. NMR lineshape analysis indicates that this is due to a substantial increase in k_{on} for SUMO-2/RAP80 SIM association; a result of enhanced electrostatic interactions between the phosphate groups of S44/S46, and a positively charged SUMO-2 region at one end of the SIM interaction site. The key role for electrostatics in the modulation of SUMO-SIM interactions is underscored by a prominent dependence on salt concentration[24]. The binding affinity of unphosphorylated RAP80 SIM for SUMO-2 is comparable to the affinities of SIM-C from DAXX for both SUMO-1 and SUMO-2 with K_D values of ~140 mM at 200 mM KCl, but weaker in comparison to the ~40 mM K_D for SIM-N binding to both SUMO-1 and SUMO-2 at 200 mM KCl[24]. It should be noted that for these affinity comparisons, SIM-C from DAXX shares the SDSD electrostatic module sequence with RAP80, whereas SIM-N from DAXX does not (DDDD). Upon double phosphorylation of RAP80, the SUMO-2–SIM interaction affinity increases, with a K_D ~9 mM; this is comparable to a K_D of ~1 mM for the tetraphosphorylated PML SIM–SUMO-1 interaction, though the greater affinity for the phosphoPML/SUMO-1 interaction, in comparison to that for RAP80/SUMO-2, is likely a result of lower salt concentration for the former interaction[21,23].

The NMR-derived solution state structure of SUMO-2 in complex with $pRAP80_{37-49}$ is the first structure of SUMO-2 bound to a phosphorylated SIM. The structure of the complex shows a rich variety of electrostatic interactions, as well as key hydrophobic interactions, generally separated into two distinct, but adjacent regions, or modules of the binding interface, as generally observed for SUMO-SIM interactions. The negatively charged region of RAP80 [pS_{44}-D_{45}-pS_{46}-D_{47}], adjacent to hydrophobic SIM module [F_{40}-I_{41}-V_{42}-I_{43}] is involved in extensive electrostatic interactions with the positively charged residues of the SIM binding interface on SUMO-2. Specifically, the observed chemical shift changes for the SUMO-2 H17 and H37 side chain H^{d2} protons upon phosphoSIM binding, suggest electrostatic interactions with a distance range of 3–5 Å between the histidine N^{e2} and SIM phosphate oxygen atoms for the histidine–phosphoserine pairs. This is consistent with the known relationship between crystallographic structures and chemical shift changes for histidine side chain H^{d2} protons from RNase A upon interaction with nucleotide phosphate groups[61,62]. Other key electrostatic interactions include the intermolecular hydrogen bonds across the peptide planes from RAP80 and SUMO-2 that form the intermolecular b-sheet, as well as intramolecular SIM hydrogen bonds between the phosphoserine side chain g-oxygens and their respective main chain amide protons.

The RAP80 SIM-SUMO-2 structure is similar to the mDAXX C-terminal SIM/SUMO-1 interaction (PDB ID: 4WJP)[21], though there are important differences in the electrostatic module of the binding interface that form the basis of specificity determinants (Fig. 3.6e, f). The residues from the SIM hydrophobic modules, FIVI for RAP80 and IIVL for DAXX, insert into the hydrophobic SUMO binding cleft similarly, with the side chain of the second and fourth residues buried, and in direct contact with SUMO. As with other SUMO-SIM interactions, key structural differences are found in the electrostatic binding module adjacent to the hydrophobic module. Specifically, there are key sequence differences for positively charged residues in the electrostatic SIM binding module of SUMO, that includes a SIM binding loop flanking the interaction site (Fig. 6*e*, residues M40 to K46, and R36 to K42 in SUMO-1 and SUMO-2, respectively).

In the case of the electrostatic phosphorylated SDSD modules from both RAP80 and the C-SIM from DAXX (PDB ID: 4WJP), the first phosphoserine from RAP80, pS44, forms favourable electrostatic interactions with K35 from strand b2 adjacent to the N-terminus of the binding loop, as well as H17 from strand b1. In contrast, the first phosphoserine from DAXX, interacts with K46 from helix a1 of SUMO-1, which lies on the opposite, or C-terminal side of the SIM binding loop. In addition, SUMO-1 lacks an equivalent to the positively charged H17 SUMO-2 residue at the N-terminal side of the binding loop, and possesses Y21 instead. Furthermore, a key SIM binding loop histidine, H37 from SUMO-2 and H43 from SUMO-1 occurs at opposite ends of the SIM binding loop. These differences result in the second phosphate from DAXX interacting with the C-terminal portion of the SIM binding loop, whereas the second phosphate from the RAP80 SIM interacts with the N-terminal portion of the SIM binding loop (Fig. 3.6f). Thus, the change in orientation for various SIM phosphoserine side chains upon interacting with cognate SUMO isoforms can be attributed to different charge distributions in the electrostatic binding modules from SUMO.

From a broader biological perspective, the N-terminal region from RAP80 connects numerous signalling pathways, including SUMOylation, ubiquitination, phosphorylation, and lysine and acetylation, with DNA damage repair[63,64]. For example, it has recently been suggested that phosphorylated SIMs function as a node in a network that connects CK2 signalling with SUMOylation[20]. CK2 has also been shown to be involved in double strand break repair through homologous recombination by assisting the association of the DNA repair protein Rad51 to the MRN complex through phosphorylation[63]. Double phosphorylation of the canonical CK2 site in RAP80 substantially elevates the binding affinity for SUMO-2, and suggests a deeper role for CK2 in the repair of double stranded DNA breaks, at least within the context of RAP80 mediated recruitment of BRCA1 to DNA damage sites. Interestingly, the phosphorylation of SIMs may be cross-regulated through acetylation of K33 in SUMO-2 and K37 in SUMO-1, although additional lysines near the SIM binding cleft may be involved [65-67]. From a structural perspective, K33 and K35 are critical for electrostatic interactions with SIM or phosphorylated SIM; their acetylation abolishes SUMO-SIM binding,

although not in all cases. In the case of RAP80, it will be of interest to determine if the interaction with SUMO-2 is regulated by acetylation of K33 and K35, linking regulation of SUMO acetylation with DNA repair by homologous recombination.

The molecular basis underlying the role of ubiquitination and SUMOylation in the RAP80-mediated function of the DNA damage response is not yet fully understood. It is unclear if these post-translational modifications function independently, or if hybrid SUMO-Ub chain recognition is necessary for RAP80 recruitment to DNA damage sites. The latter model is alluring, given that the polySUMO binding Ub ligase RNF4 has been shown to be necessary for BRCA1 recruitment[14], and can catalyze the covalentl attachment of Ub to SUMO chains[68]. This study provides the first molecular details underlying the SUMO/RAP80 interaction, and its regulation by phosphorylation. It will be of interest to gain a deeper understanding of the interplay between various regulatory and signalling processes in RAP80-mediated BRCA1 recruitment to the DNA damage sites.

References

[1]   Jackson, S. P., and Bartek, J. (2010) The DNA damage response in human biology and disease. *Nature* **461**, 1071-1078

[2]   Giglia-Mari, G., Zotter, A., and Vermeulen, W. (2011) DNA damage response. *Cold Spring Harb. Perspect. Biol.* **3**, 1-19

[3]   Bekker-Jensen, S., and Mailand, N. (2010) Assembly and function of DNA double-strand break repair foci in mammalian cells. *DNA Repair* **9**, 1219-1228

[4]   Bologna, S., and Ferrari, S. (2013) It takes two to tango: Ubiquitin and SUMO in the DNA damage response. *Front. Genet.* **4**, 1-18

[5]   Huen, M. S., and Chen, J. (2008) The DNA damage response pathways: at the crossroad of protein modifications. *Cell Res.* **18**, 8-16

[6]   Matsuoka, S., Ballif, B. A., Smogorzewska, A., McDonald, E. R., Hurov, K. E., Luo, J., Bakalarski, C. E., Zhao, Z. M., Solimini, N., Lerenthal, Y., Shiloh, Y., Gygi, S. P., and Elledge, S. J. (2007) ATM and ATR substrate analysis reveals extensive protein networks responsive to DNA damage. *Science* **316**, 1160-1166

[7]   Kolas, N. K., Chapman, J. R., Nakada, S., Ylanko, J., Chahwan, R., Sweeney, F. D., Panier, S., Mendez, M., Wildenhain, J., Thomson, T. M., Pelletier, L., Jackson, S. P., and Durocher, D. (2007) Orchestration of the DNA-damage response by the RNF8 ubiquitin ligase. *Science* **318**, 1637-1640

[8]   Mailand, N., Bekker-Jensen, S., Faustrup, H., Melander, F., Bartek, J., Lukas, C., and Lukas, J. (2007) RNF8 ubiquitylates histones at DNA double-strand breaks and promotes assembly of repair proteins. *Cell* **131**, 887-900

[9]   Kim, H., Chen, J. J., and Yu, X. H. (2007) Ubiquitin-binding protein RAP80 mediates BRCA1-dependent DNA damage response. *Science* **316**, 1202-1205

[10]   Sobhian, B., Shao, G. Z., Lilli, D. R., Culhane, A. C., Moreau, L. A., Xia, B., Livingston, D. M., and Greenberg, R. A. (2007) RAP80 targets BRCA1 to specific ubiquitin structures at DNA damage sites. *Science* **316**, 1198-1202

[11]   Galanty, Y., Belotserkovskaya, R., Coates, J., Polo, S., Miller, K. M., and Jackson, S. P. (2009) Mammalian SUMO E3-ligases PIAS1 and PIAS4 promote responses to DNA double-strand breaks. *Nature* **462**, 935-939

[12]   Morris, J. R., Boutell, C., Keppler, M., Densham, R., Weekes, D., Alamshah, A., Butler, L., Galanty, Y., Pangon, L., Kiuchi, T., Ng, T., and Solomon, E. (2009) The SUMO modification pathway is involved in the BRCA1 response to genotoxic stress. *Nature* **462**, 886-890

[13]   Hu, X., Paul, A., and Wang, B. (2012) RAP80 protein recruitment to DNA double-strand breaks requires binding to both small ubiquitin-like modifier (SUMO) and ubiquitin conjugates. *J. Biol. Chem.* **287**, 25510-25519

[14]   Guzzo, C. M., Berndsen, C. E., Zhu, J., Gupta, V., Datta, A., Greenberg, R. A., Wolberger, C., and Matunis, M. J. (2012) RNF4-dependent hybrid SUMO-ubiquitin chains are signals for RAP80 and thereby mediate the recruitment of BRCA1 to sites of DNA damage. *Sci. Signal.* **5**, ra88

[15]   Kerscher, O., and William, C. (2007) SUMO junction—what's your function? New insights through SUMO-interacting motifs. **8**, 550-555

[16]   Muller, S., Hoege, C., Pyrowolakis, G., and Jentsch, S. (2001) SUMO, ubiquitin's mysterious cousin. *Nat. Rev. Mol. Cell Biol.* **2**, 202-210

[17]   Hecker, C.-M., Rabiller, M., Haglund, K., Bayer, P., and Dikic, I. (2006) Specification of SUMO1- and SUMO2-interacting motifs. *J. Biol. Chem.* **281**, 16117-16127

[18]   Gareau, J. R., and Lima, C. D. (2010) The SUMO pathway: emerging mechanisms that shape specificity, conjugation and recognition. *Nat. Rev. Mol. Cell Biol.* **11**, 861-871

[19]   Song, J., Zhang, Z., Hu, W., and Chen, Y. (2005) Small ubiquitin-like modifier (SUMO) recognition of a SUMO binding motif: a reversal of the bound orientation. *J. Biol. Chem.* **280**, 40122-40129

[20]   Stehmeier, P., and Muller, S. (2009) Phospho-regulated SUMO interaction modules connect the SUMO system to CK2 signaling. *Mol. Cell* **33**, 400-409

[21]   Cappadocia, L., Mascle, X. H., Bourdeau, V., Tremblay-Belzile, S., Chaker-Margot, M., Lussier-Price, M., Wada, J., Sakaguchi, K., Aubry, M., Ferbeyre, G., and Omichinski, J. G. (2015) Structural and functional characterization of the phosphorylation-dependent interaction between PML and SUMO1. *Structure* **23**, 126-138

[22]   Sekiyama, N., Ikegami, T., Yamane, T., Ikeguchi, M., Uchimura, Y., Baba, D., Ariyoshi, M., Tochio, H., Saitoh, H., and Shirakawa, M. (2008) Structure of the

small ubiquitin-like modifier (SUMO)-interacting motif of MBD1-containing chromatin-associated factor 1 bound to SUMO-3. *J. Biol. Chem.* **283**, 35966-35975

[23] Chang, C. C., Naik, M. T., Huang, Y. S., Jeng, J. C., Liao, P. H., Kuo, H. Y., Ho, C. C., Hsieh, Y. L., Lin, C. H., Huang, N. J., Naik, N. M., Kung, C. C. H., Lin, S. Y., Chen, R. H., Chang, K. S., Huang, T. H., and Shih, H. M. (2011) Structural and functional roles of Daxx SIM phosphorylation in SUMO paralog-selective binding and apoptosis modulation. *Mol. Cell* **42**, 62-74

[24] Escobar-Cabrera, E., Okon, M., Lau, D. K., Dart, C. F., Bonvin, A. M., and McIntosh, L. P. (2011) Characterizing the N- and C-terminal small ubiquitin-like modifier (SUMO)-interacting motifs of the scaffold protein DAXX. *J. Biol. Chem.* **286**, 19816-19829

[25] Reverter, D., and Lima, C. D. (2005) Insights into E3 ligase activity revealed by a SUMO-RanGAP1-Ubc9-Nup358 complex. *Nature* **435**, 687-692

[26] Yan, Z., Kim, Y. S., and Jetten, A. M. (2002) RAP80, a novel nuclear protein that interacts with the retinoid-related testis-associated receptor. *J. Biol. Chem.* **277**, 32379-32388

[27] Wang, B., Matsuoka, S., Ballif, B. A., Zhang, D., Smogorzewska, A., Gygi, S. P., and Elledge, S. J. (2007) Abraxas and RAP80 form a BRCA1 protein complex required for the DNA damage response. *Science* **316**, 1194-1198

[28] Sims, J. J., and Cohen, R. E. (2009) Linkage-specific avidity defines the lysine 63-linked polyubiquitin-binding preference of RAP80. *Mol. Cell* **33**, 775-783

[29] Markin, C. J., Wei, X. A., and Spyracopoulos, L. (2010) Mechanism for recognition of polyubiquitin chains: Balancing affinity through interplay between multivalent binding and dynamics. *J. Am. Chem. Soc.* **132**, 11247-11258

[30] Wittekind, M., and Mueller, L. (1993) HNCACB, a high-sensitivity 3D NMR experiment to correlate amide proton and nitrogen resonances with the alpha- and beta-carbon resonances in proteins. *J. Magn. Reson. B* **101**, 201-205

[31]   Muhandiram, D. R., and Kay, L. E. (1994) Gradient-enhanced triple-resonance three-dimensional NMR experiments with improved sensitivity. *J. Magn. Reson. B* **103**, 203-216

[32]   Grzesiek, S., and Bax, A. (1992) Correlating backbone amide and side chain resonances in larger proteins by multiple relayed triple resonance NMR. *J. Am. Chem. Soc.* **114**, 6291-6293

[33]   Szyperski, T., Braun, D., Fernandez, C., Bartels, C., and Wüthrich, K. (1995) A novel reduced-dimensionality triple-resonance experiment for efficient polypeptide backbone assignment, 3D CO HN N CA. *J. Magn. Reson. B* **108**, 197-203

[34]   Panchal, S. C., Bhavesh, N. S., and Hosur, R. V. (2001) Improved 3D triple resonance experiments, HNN and HN(C)N, for H^N and ^{15}N sequential correlations in (^{13}C, ^{15}N) labeled proteins: Application to unfolded proteins. *J. Biomol. NMR* **20**, 135-147

[35]   Kay, L. E., Xu, G. Y., and Yamazaki, T. (1994) Enhanced-sensitivity triple-resonance spectroscopy with minimal H_2O saturation. *J. Magn. Reson. A* **109**, 129-133

[36]   Yamazaki, T., Lee, W., Arrowsmith, C. H., Muhandiram, D. R., and Kay, L. E. (1994) A suite of triple resonance experiments for the backbone assignment of ^{15}N, ^{13}C, ^2H labeled proteins with high sensitivity. *J. Am. Chem. Soc.* **116**, 11655-11666

[37]   Goddard, T. D., and Kneller, D. G. SPARKY 3. University of California, San Francisco

[38]   Keller, R. L. J. (2004) The Computer-aided resonance assignment tutorial. Cantina Verlag, Zürich

[39]   Wüthrich, K. (1986) *NMR of proteins and nucleic acids.*, John Wiley & Sons, New York

[40] Kay, L. E., Keifer, P., and Saarinen, T. (1992) Pure absorption gradient enhanced heteronuclear single quantum correlation spectroscopy with improved sensitivity. *J. Am. Chem. Soc.* **114**, 10663-10665

[41] Markin, C. J., and Spyracopoulos, L. (2012) Increased precision for analysis of protein-ligand dissociation constants determined from chemical shift titrations. *J. Biomol. NMR* **53**, 125-138

[42] Shen, Y., and Bax, A. (2013) Protein backbone and sidechain torsion angles predicted from NMR chemical shifts using artificial neural networks. *J. Biomol. NMR* **56**, 227-241

[43] Grzesiek, S., Anglister, J., and Bax, A. (1993) Correlation of backbone amide and aliphatic side chain resonances in ^{13}C/^{15}N-enriched proteins by isotropic mixing of ^{13}C magnetization. *J. Magn. Reson. B* **101**, 114-119

[44] Logan, T. M., Olejniczak, E. T., Xu, R. X., and Fesik, S. W. (1993) A general method for assigning NMR spectra of denatured proteins using 3D HC(CO)NH-TOCSY triple resonance experiments. *J. Biomol. NMR* **3**, 225-231

[45] Lyons, B. A., and Montelione, G. T. (1993) An HGCCNH triple-resonance experiment using C-13 isotropic mixing for correlating backbone amide and side-chain aliphatic resonances in isotopically enriched proteins. *J. Magn. Reson. B* **101**, 206-209

[46] Gardner, K. H., Konrat, R., Rosen, M. K., and Kay, L. E. (1996) An (H)C(CO)NH-TOCSY pulse scheme for sequential assignment of protonated methyl groups in otherwise deuterated ^{15}N, ^{13}C-labeled proteins. *J. Biomol. NMR* **8**, 351-356

[47] Pervushin, K., Riek, R., Wider, G., and Wüthrich, K. (1998) Transverse relaxation-optimized spectroscopy (TROSY) for NMR studies of aromatic spin systems in ^{13}C-labeled proteins. *J. Am. Chem. Soc.* **120**, 6394-6400

[48]   Yamazaki, T., Forman-Kay, J. D., and Kay, L. E. (1993) Two-dimensional NMR experiments for correlating $^{13}C_b$ and $^{1}H_{d/e}$ chemical shifts of aromatic residues in ^{13}C-labeled proteins via scalar couplings. *J. Am. Chem. Soc.* **115**, 11054-11055

[49]   Otting, G., and Wüthrich, K. (1989) Extended heteronuclear editing of 2D ^{1}H NMR spectra of isotope-labeled proteins, using the X(w$_1$, w$_2$) double half filter. *J. Magn. Reson.* **85**, 586-594

[50]   Case, D. A., Berryman, J. T., Betz, R. M., Cerutti, D. S., Cheatham, T. E., Darden, T. A., Duke, R. E., Giese, T. J., Gohlke, H., Goetz, A. W., Homeyer, N., Izadi, S., Janowski, P., Kaus, J., Kovalenko, A., Lee, T. S., Le Grand, S., Li, P., Luchko, T., Luo, R., Madej, B., Merz, K. M., Monard, G., Needham, P., Nguyen, H., Nguyen, H. T., Omelyan, I., Onufriev, A., Roe, D. R., Roitberg, A., Salomon-Ferrer, R., Simmerling, C., Smith, W., Swails, J., Walker, R. C., Wang, J., Wolf, R. M., Wu, X., York, D. M., and Kollman, P. A. (2015) AMBER 2015.

[51]   Berjanskii, M. V., Neal, S., and Wishart, D. S. (2006) PREDITOR: A web server for predicting protein torsion angle restraints. *Nucleic Acids Res.* **34**, 63-69

[52]   Huang, W. C., Ko, T. P., Li, S. S., and Wang, A. H. (2004) Crystal structures of the human SUMO-2 protein at 1.6 Å and 1.2 Å resolution: Implication on the functional differences of SUMO proteins. *Eur. J. Biochem.* **271**, 4114-4122

[53]   Homeyer, N., Horn, A. H., Lanig, H., and Sticht, H. (2006) AMBER force-field parameters for phosphorylated amino acids in different protonation states: Phosphoserine, phosphothreonine, phosphotyrosine, and phosphohistidine. *J. Mol. Model.* **12**, 281-289

[54]   Markin, C. J., and Spyracopoulos, L. (2012) Accuracy and precision of protein-ligand interaction kinetics determined from chemical shift titrations. *J. Biomol. NMR* **54**, 355-376

[55]   Anamika, Markin, C. J., Rout, M. K., and Spyracopoulos, L. (2014) Molecular basis for impaired DNA damage response function associated with the RAP80 ΔE81 defect. *J. Biol. Chem.* **289**, 12852-12862

[56]   Berjanskii, M. V., and Wishart, D. S. (2005) A simple method to predict protein flexibility using secondary chemical shifts. *J. Am. Chem. Soc.* **127**, 14970-14971

[57]   Panier, S., Ichijima, Y., Fradet-Turcotte, A., Leung, Charles C. Y., Kaustov, L., Arrowsmith, C. H., and Durocher, D. (2012) Tandem protein interaction modules organize the ubiquitin-dependent response to DNA double-strand breaks. *Mol. Cell* **47**, 383-395

[58]   Olsten, M. E., and Litchfield, D. W. (2004) Order or chaos? An evaluation of the regulation of protein kinase CK2. *Biochem. Cell Biol.* **82**, 681-693

[59]   Tholey, A., Lindemann, A., Kinzel, V., and Reed, J. (1999) Direct effects of phosphorylation on the preferred backbone conformation of peptides: A nuclear magnetic resonance study. *Biophys. J.* **76**, 76-87

[60]   Wishart, D. S., Sykes, B. D., and Richards, F. M. (1991) Relationship between nuclear magnetic resonance chemical shift and protein secondary structure. *J. Mol. Biol.* **222**, 311-333

[61]   Rüterjans, H, and Witzel, H. (1969) NMR studies on the structure of the active site of pancreatic ribonuclease A. *Eur. J. Biochem.* **9**, 118-127

[62]   Leonidas, D. D., Chavali, G. B., Oikonomakos, N. G., Chrysina, E. D., Kosmopoulou, M. N., Vlassi, M., Frankling, C., and Acharya, K. R. (2003) High-resolution crystal structures of ribonuclease A complexed with adenylic and uridylic nucleotide inhibitors. Implications for structure-based design of ribonucleolytic inhibitors. *Protein Sci.* **12**, 2559-2574

[63]   Yata, K., Lloyd, J., Maslen, S., Bleuyard, J. Y., Skehel, M., Smerdon, S. J., and Esashi, F. (2012) Plk1 and CK2 act in concert to regulate Rad51 during DNA double strand break repair. *Mol. Cell* **45**, 371-383

[64]   Krejci, L., Altmannova, V., Spirek, M., and Zhao, X. (2012) Homologous recombination and its regulation. *Nucleic Acids Res.* **40**, 5795-5818

[65]   Choudhary, C., Kumar, C., Gnad, F., Nielsen, M. L., Rehman, M., Walther, T. C., Olsen, J. V., and Mann, M. (2009) Lysine acetylation targets protein complexes and co-regulates major cellular functions. *Science* **325**, 834-840

[66]   Cheema, A., Knights, C. D., Rao, M., Catania, J., Perez, R., Simons, B., Dakshanamurthy, S., Kolukula, V. K., Tilli, M., Furth, P. A., Albanese, C., and Avantaggiati, M. L. (2010) Functional mimicry of the acetylated C-terminal tail of p53 by a SUMO-1 acetylated domain, SAD. *J. Cell. Physio*. **225**, 371-384

[67]   Ullmann, R., Chien, C. D., Avantaggiati, M. L., and Muller, S. (2012) An acetylation switch regulates SUMO-dependent protein interaction networks. *Mol. Cell* **46**, 759-770

[68]   Tatham, M. H., Plechanovova, A., Jaffray, E. G., Salmen, H., and Hay, R. T. (2013) Ube2W conjugates ubiquitin to alpha-amino groups of protein N-termini. *Biochem. J*. **453**, 137-145

Chapter 4

Conclusion and Future Directions

In Chapter 2 we quantified the impaired binding of the RAP80 ΔE81 mutant with ubiquitin and diubiquitin and studied the structural and stability details of the mutant protein in comparison with the wild type protein. We demonstrated using NMR titrations, molecular dynamics simulations and structural studies that due to the deletion of E81 at the N-terminal region of UIM1 of RAP80 there is a structural frameshift of the α-helix. Although the N-cap motif and the structural integrity of the protein is maintained, the binding is impaired due to the loss of favourable electrostatic interaction between E81 and R74 and the surrounding residues of ubiquitin that leads to a loss of the multivalent binding and results in the diminished recruitment of repair factors to the sites of DNA damage subsequently leading to a possible compromise in the genomic integrity[1][2]. Chapter 3 focuses on the binding details of RAP80 SIM in both its free and phosphorylated form with SUMO-2 isoform. We also explored the structural details of phosphorylated RAP80 SIM binding with SUMO-2. It expanded the molecular understanding of SUMO:SIM binding and its intimate link with phosphorylation to the DNA repair field. We demonstrate using NMR chemical shift titrations and lineshape analysis that upon phosphorylation of RAP80 SIM by the CK2 there is a substantial increase in the binding affinity due to ~100 fold increase in the k_{on} of the interaction. The structural details of the complex indicate that this is due to the presence of additional electrostatic interaction between the phosphorylated SIM and the basic loop of the SUMO-2 at the binding interface. The structural aspects of phosphorylated RAP80 SIM binding with SUMO-2 and its comparison with previously reported SUMO-1:phosphorylated SIM structure has helped us gain invaluable insight into the SUMO isoform specificity[3]. In addition to contributing new knowledge to the ubiquitin and SUMO signalling in the DNA repair, we provide the molecular level understanding of an impaired binding process in the DNA repair that can potentially lead to the pathogenesis of deadly disease like cancer. However, there are many open questions that need attention in order to gain a deeper insight of the ubiquitin and SUMO signalling in isolation, and their implications on a broader scale.

Critical Mutations in BRCA1 Interacting Proteins

As mentioned in Nikkila *et. al*[2], approximately 20% of familial breast cancer are linked to mutations in the BRCA1 and BRCA2 genes whose prominence in the maintenance of genome integrity cannot be emphasized enough. The genetic cause for the rest 80% is still unknown and may involve mutations in BRCA1 BRCA2 interacting proteins or any other unidentified high penetrance genes[2]. In addition to mutations in RAP80, BACH1, PALB2, RAD50, NBS1 and TopBP1, a screening for other interacting proteins with critical mutations will be beneficial in identifying the mutations that can be a threat to genomic integrity and a potential cause for the predisposition to cancer[4-6].

The Role of RAP80 LRM in the Recognition Process

The N-terminal region of RAP80 connects numerous signalling pathways such as ubiquitination, SUMOylation and phosphorylation, owing to the presence of SIM, CK2 phosphorylation sites and tandem linked UIMs. As mentioned in Panier *et. al*[7], the optimal recruitment of RAP80 to the DNA damage sites is dependent on its LRM motif in addition to its UIMs, which is 8 residues N-terminal to the UIM1. We showed that the residues N-terminal to RAP80 do not adopt the same alpha helical conformation and do not show any change in the chemical shift upon ubiquitin addition, indicating that they do not contribute directly to the ubiquitin binding at the damage sites. Nevertheless, these residues might be critical in recognizing the ubiquitinated substrate or contribute to the binding of N-terminal region of RAP80 to the high affinity ligand - SUMO-Ub$_2$ hybrid chain, synthesized at the DNA damage sites by RNF4, a SUMO targeted ubiquitin ligase[8]

The Interaction of RAP80 with Different SUMO Isoforms

The binding of different isoforms of SUMO with the SUMO interacting motif of PML and Daxx have been studied in both in its free and phosphorylated form[9][10]. The high degree of similarity between RAP80 SIM ($_{40}$FIVISDSD$_{47}$) and Daxx SIM ($_{733}$IIVLSDSD$_{740}$) sequences inspired us to compare their affinities. As reported in Chang *et. al*[10], upon phosphorylation of both the serines of Daxx SIM by CK2 its affinity for SUMO-1 is ~10 fold higher than SUMO-3 which is 11.9 ± 2.7 μM (measured by ITC) and similar to our reported affinity of 9 ± 3 μM (measured by NMR titration). We restricted our study to SUMO-2 due to its preferred binding to SUMO-2 over SUMO-1 as indicated in invitro pull down assays[8,11]. Due to a similar charge distribution of SUMO-1 and SUMO-2 in the SIM binding groove and the loop connecting the α_1 helix and β_2 strand, we expected pRAP80 to bind SUMO-1 with a similar affinity and later carried out the NMR titration of phosphorylated RAP80 SIM with SUMO-1, reported in Appendix A. We determine the binding affinity to be ~8 ± 5 μM, similar to the phosphorylated RAP80 binding with SUMO-2, but ~4 fold less than Daxx SIM binding with SUMO-1 which is ~2 ± 0.4 μM which might be due to the method used for determination of binding affinity. The structure of pRAP80:SUMO-1 complex will shed more light on the atomic details of isoform specificity.

SUMO Acetylation and its Interaction with SIM

Our work also opens new avenues of research, such as the involvement of SUMO lysine acetylation in the SUMO signalling. The residues K33 and K35 of SUMO-2 play a key role in the SIM recognition by providing an electrostatically positive site for interaction with the negatively charged or phosphorylated serine residues of SIM[3]. The residues K33 and K35 have been reported be acetylated in cells and demonstrated to abolish its interaction with SIM containing proteins. This might lead to an additional layer of regulation in the SUMO:SIM signalling[12–14]. The structure of pRAP80:SUMO-2 and the NMR titration experiments establish the significant role played by K33 and K35 in its interaction with RAP80 SIM and its acetylation may be an important mode of regulating

113

the SUMO:SIM signalling in the DNA repair pathway. More studies need to be carried out to understand the intricate signalling details of the SUMO:SIM interaction in DNA repair pathway.

The Role of CK2 in the RAP80 Mediated BRCA1 Recruitment to the DNA Damage Sites

We also hinted at the role of CK2 in the RAP80 mediated BRCA1 recruitment to the DNA damage sites inspired by the dramatic increase in the affinity of RAP80 SIM for SUMO-2 upon phosphorylation. The inhibition of CK2 has been reported to delay the foci removal suggesting its role in the DNA repair, however in order to investigate the direct role of CK2 in the recruitment of RAP80 to the DNA damage sites, we compared the foci forming ability of a RAP80-S44A,S46A mutant with the wild type RAP80. We did not observe any evident change in the number of foci formed by both the CK2 site mutant and the wild type RAP80 (unpublished data). Although the recruitment was unaffected, there might be an effect on the retention of RAP80 at the damage sites. Consistent with our predicted role of CK2 in DNA repair, S44F mutation has been listed in the COSMIC (Catalogue Of Somatic Mutation In Cancer), indicating a deeper role for CK2 mediated phosphorylation in the regulation of the function of RAP80 in genome protection[15]. We also studied the cellular expression of cells expressing a gfp-rap80-S44A,S46A mutation and we observe a reduced puncta formation when compared to the wild type protein in the absence of DNA damage, implications and details of which are mentioned in Appendix B. From a structural point of view, the substitution of phenylalanine in place of serine will not only lead to the loss of a potential phosphorylation site but the mutation also has the ability to impede or obstruct the phosphorylation of S46, the phosphorylated residue at -2 to +2 position serves as a positive specificity determinant[16]. The presence of a bulky hydrophobic residue near the site of phosphorylation might also affect the specificity for the site, negatively[16]. The effect of mutation at one site might have more deleterious consequences that can lead to a total loss of phosphorylation, resulting in the loss of binding affinity. Apart from the

114

ability to be phosphorylated, the interaction of RAP80S44F to SUMO-2 also poses few concerns from a structural standpoint. The incorporation of phenylalanine in place of serine might cause steric hindrance due to its bulky side chain and the interaction between the hydrophobic side chain of phenylalanine and the positively charged residues such as K35, R36 and H37 of SUMO-2 will not be a favoured, leading to an impaired binding[3]. However, an experimental analysis of structure and binding will provide a firm understanding of the interaction.

The Role and Recognition of Hybrid Chain by RAP80

The chemical shift derived secondary structure of N-terminal RAP80 shows that it contains a SIM in a β-strand conformation at the near N-terminus, a 32 residue flexible linker connecting the SIM and the tandem UIMs that are α-helical[3]. The RAP80 N-terminal SIM-tUIMs bind the SUMO-2-Ub_2 chains with a significantly higher affinity as compared to SUMO and Ub alone, in vitro[8]. Although structural studies need to be carried out to understand the molecular level of the binding, it can be safely assumed that the binding of hybrid chain to RAP80 will have additional multivalent advantage and alongwith the SIM phosphorylation, the increase in affinity can be significant, as reported in Guzzo *et*. *al*[8][17]. Besides the role of N-terminal SIM-tUIMs of RAP80 in DNA repair, we also discovered its role in the nuclear puncta formation independent of DNA damage, explained in Appendix B. Using fluorescence microscopy we show that the SIM-tUIMs of RAP80 have the ability to form nuclear puncta in the absence of induced DNA damage and their deletion leads to a complete absence of these puncta. These puncta partially co-localize with PML nuclear bodies and their co-localization may be through its binding to hybrid chain, as both the SIM and UIMs collectively contribute towards the puncta formation. PML protein is modified by hybrid chain mediated by the STUbL, RNF4 which is necessary for PML degradation and RNF168 has been reported to localize to PML bodies thorough its ability to bind hybrid chains[18][19]. The hybrid chain may be a biologically relevant signal in DNA repair and PML regulation offering additional specificity over SUMO and Ub chains of specific linkage.

Our work presented here has contributed to the knowledge base of ubiquitin and SUMO signalling and opened new avenues of research for a holistic understanding of the signalling process in DNA repair and other processes for the benefit the humankind.

References

[1]   Anamika, Markin, C. J., Rout, M. K., and Spyracopoulos, L. (2014) Molecular basis for impaired DNA damage response function associated with the RAP80 ΔE81 defect. *J. Biol. Chem.* **289**, 12852–62

[2]   Nikkilä, J., Coleman, K. A., Morrissey, D., Pylkäs, K., Erkko, H., Messick, T. E., Karppinen, S.-M., Amelina, A., Winqvist, R., and Greenberg, R. A. (2009) Familial breast cancer screening reveals an alteration in the RAP80 UIM domain that impairs DNA damage response function. *Oncogene.* **28**, 1843–1852

[3]   Anamika, and Spyracopoulos, L. (2016) Molecular basis for phosphorylation-dependent SUMO recognition by the DNA repair protein RAP80. *J. Biol. Chem.* **291**, 4417–4428

[4]   Heikkinen, K., Rapakko, K., Karppinen, S.-M., Erkko, H., Knuutila, S., Lundán, T., Mannermaa, A., Børresen-Dale, A.-L., Borg, Å., Barkardottir, R. B., Petrini, J., and Winqvist, R. (2006) RAD50 and NBS1 are breast cancer susceptibility genes associated with genomic instability . *Carcinogenesis.* **27**, 1593–1599

[5]   Karppinen, S.-M., Erkko, H., Reini, K., Pospiech, H., Heikkinen, K., Rapakko, K., Syväoja, J. E., and Winqvist, R. (2006) Identification of a common polymorphism in the TopBP1 gene associated with hereditary susceptibility to breast and ovarian cancer. *Eur. J. Cancer.* **42**, 2647–2652

[6]   Stratton, M. R., and Rahman, N. (2008) The emerging landscape of breast cancer susceptibility. *Nat Genet.* **40**, 17–22

[7]   Panier, S., Ichijima, Y., Fradet-Turcotte, A., Leung, C. C. Y., Kaustov, L., Arrowsmith, C. H., and Durocher, D. (2012) Tandem Protein Interaction Modules Organize the Ubiquitin-Dependent Response to DNA Double-Strand Breaks. *Mol.*

Cell. **47**, 383–395

[8]   Guzzo, C. M., Berndsen, C. E., Zhu, J., Gupta, V., Greenberg, R. A., Wolberger, C., Matunis, M. J., Guzzo, C. M., Berndsen, C. E., Zhu, J., Gupta, V., and Datta, A. (2013) RNF4-Dependent Hybrid SUMO-Ubiquitin Chains Are Signals for RAP80 and Thereby Mediate the Recruitment of BRCA1 to Sites of DNA Damage. *Sci*. *Signal*. **5**, ra88

[9]   Escobar-cabrera, E., Okon, M., Lau, K. W., Dart, C. F., Alexandre, M., Bonvin, J. J., and Mcintosh, L. P. (2011) Characterizing the N- and C-terminal Small Ubiquitin-like Modifier (SUMO) interacting Motifs of the Scaffold Protein. *J*. *Biol*. *Chem*. **286**, 19816-19829

[10]   Chang, C. C., Naik, M. T., Huang, Y. S., Jeng, J. C., Liao, P. H., Kuo, H. Y., Ho, C. C., Hsieh, Y. L., Lin, C. H., Huang, N. J., Naik, N. M., Kung, C. C. H., Lin, S. Y., Chen, R. H., Chang, K. S., Huang, T. H., and Shih, H. M. (2011) Structural and functional roles of Daxx SIM phosphorylation in SUMO paralog-selective binding and apoptosis modulation. *Mol*. *Cell* **42**, 62-74

[11]   Hu, X., Paul, A., and Wangs, B. (2012) Rap80 protein recruitment to DNA double-strand breaks requires binding to both small ubiquitin-like modifier (SUMO) and ubiquitin conjugates. *J*. *Biol*. *Chem*. **287**, 25510–25519

[12]   Ullmann, R., Chien, C. D., Avantaggiati, M. L., and Muller, S. (2012) An acetylation switch regulates SUMO-dependent protein interaction networks. *Mol*. *Cell* **46**, 759-770

[13]   Cheema, A., Knights, C. D., Rao, M., Catania, J., Perez, R., Simons, B., Dakshanamurthy, S., Kolukula, V. K., Tilli, M., Furth, P. A., Albanese, C., and Avantaggiati, M. L. (2010) Functional mimicry of the acetylated C-terminal tail of p53 by a SUMO-1 acetylated domain, SAD. *J*. *Cell*. *Physio*. **225**, 371-384

[14]   Choudhary, C., Kumar, C., Gnad, F., Nielsen, M. L., Rehman, M., Walther, T. C., Olsen, J. V., and Mann, M. (2009) Lysine acetylation targets protein complexes and co-regulates major cellular functions. *Science* **325**, 834-840

[15]   Lombardi, P. M., Matunis, M. J., and Wolberger, C. (2017) RAP80, ubiquitin and SUMO in the DNA damage response. *J*. *Mol*. *Med*. **95**, 799–807

[16]   Meggio, F., Marin, O., and Pinna, L. A. (1994) Substrate specificity of protein kinase CK2. *Cell. Mol. Biol. Res.* **40(5-6)**, 401-409

[17]   Lee, B. L., Singh A., Glover, J. N. M., Hendzel, M. J and Spyracopoulos Leo.(2017). Molecular basis for K63-linked ubiquitination processes in double strand DNA breaks repair: A focus on kinetics and dynamics. *J. Mol. Biol.* 10.1016/j.jmb.2017.05.029

[18]   Tatham, M. H., Geoffroy, M.-C., Shen, L., Plechanovova, A., Hattersley, N., Jaffray, E. G., Palvimo, J. J., and Hay, R. T. (2008) RNF4 is a poly-SUMO-specific E3 ubiquitin ligase required for arsenic-induced degradation. *Nat. Cell Biol.* **10**, 538–546

[19]   Shire, K., Wong, A. I., Tatham, M. H., Anderson, O. F., Ripsman, D., Gulstene, S., Moffat, J., Hay, R. T., and Frappier, L. (2016) Identification of RNF168 as a PML nuclear body regulator. *J. Cell Sci.* **129**, 580–591

Bibliography

Jackson, S. P., and Bartek, J. (2010) The DNA-damage response in human biology and disease. *Nature*. **461**, 1071–1078

Kornberg, R. D., and Lorch, Y. (1999) Twenty-Five Years of the Nucleosome , Fundamental Particle of the Eukaryote Chromosome. **98**, 285–294

Bednar, J., Horowitz, R. A., Grigoryev, S. A., Carruthers, L. M., Hansen, J. C., Koster, A. J., & Woodcock, C. L. (1998). Nucleosomes, linker DNA, and linker histone form a unique structural motif that directs the higher-order folding and compaction of chromatin. *PNAS*, **95**(24), 14173–14178.

Woodcock, C. L. (2005) A milestone in the odyssey of higher-order chromatin structure. *Nat. Str. Mol. Biol.*, **12**, 639–640

Manuelidis, L. (1990) View of Interphase Chromosomes. *Science,* **250**, 1533-1540.

Brown, J. S., Lukashchuk, N., Sczaniecka-Clift, M., Britton, S., le Sage, C., Calsou, P., Beli, P., Galanty, Y., and Jackson, S. P. (2015) Neddylation Promotes Ubiquitylation and Release of Ku from DNA-Damage Sites. *Cell Rep.* **11**, 704–714

Ciccia, A., and Elledge, S. J. (2010) The DNA Damage Response: Making It Safe to Play with Knives. *Mol. Cell*. **40**, 179–204

Thompson, L. H. (2012) Recognition, signaling, and repair of DNA double-strand breaks produced by ionizing radiation in mammalian cells: The molecular choreography. *Mutat. Res. - Rev. Mutat. Res.* **751**, 158–246

Lieber, M. R. (2010) The Mechanism of Double-Strand DNA Break Repair by the Nonhomologous DNA End-Joining Pathway. *Annual Review of Biochemistry*, **79**(D)**,** 181-211.

Krejci, L., Altmannova, V., Spirek, M., and Zhao, X. (2012) Homologous recombination and its regulation. *Nucleic Acid Research*, **40**, 5795-5818

Schwertman, P., Bekker-Jensen, S., and Mailand, N. (2016) Regulation of DNA double-strand break repair by ubiquitin and ubiquitin-like modifiers. *Nat Rev Mol Cell Biol*. **17**, 379–394

Dellaire, G., and Bazett-Jones, D. P. (2004) PML nuclear bodies: Dynamic sensors of DNA

damage and cellular stress. *BioEssays*. **26**, 963–977

Smeenk, G., and Mailand, N. (2016) Writers, readers, and erasers of histone ubiquitylation in DNA double-strand break repair. *Front. Genet.* **7**, 1–14

Polo, S., and Jackson, S. (2011) Dynamics of DNA damage response proteins at DNA breaks: a focus on protein modifications. *Genes Dev.* **25**, 409–433

Bekker-Jensen, S., and Mailand, N. (2010) Assembly and function of DNA double-strand break repair foci in mammalian cells. *DNA Repair (Amst).* **9**, 1219–1228

Kuo, L. J., and Yang, L.-X. (2008) Gamma-H2AX - a novel biomarker for DNA double-strand breaks. *In Vivo.* **22**, 305–309

Stewart, G. S., Wang, B., Bignell, C. R., Taylor, A. M. R., and Elledge, S. J. (2003) MDC1 is a mediator of the mammalian DNA damage checkpoint. *Nature.* **421**, 961–966

Goldberg, M., Stucki, M., Falck, J., D'Amours, D., Rahman, D., Pappin, D., Bartek, J., and Jackson, S. P. (2003) MDC1 is required for the intra-S-phase DNA damage checkpoint. *Nature.* **421**, 952–956

Lou, Z., Silva Chini, C. C., Minter-Dykhouse, K., and Chen, J. (2003) Mediator of DNA damage checkpoint protein 1 regulates BRCA1 localization and phosphorylation in DNA damage checkpoint control. *J. Biol. Chem.* **278**, 13599–13602

Lee, B. L., Singh A., Glover, J. N. M., Hendzel, M. J. and Spyracopoulos, Leo. (2017) Molecular basis for K63-linked ubiquitination processes in double-strand DNA break repair: A focus on kinetics and dynamics. *J. Mol. Biol.* http://dx.doi.org/10.1016/j.jmb.2017.05.029

Hershko, A., and Ciechanover, A. (1998) The ubiquitin system. *Annu. Rev. Biochem.* **67**, 425–479

Varshavsky, A. (2006) The early history of the ubiquitin field. *Protein Sci.* **15**, 647–54

Pickart, C. M., and Eddins, M. J. (2004) Ubiquitin: Structures, functions, mechanisms. *Biochim. Biophys. Acta - Mol. Cell Res.* **1695**, 55–72

Pickart, C. M. (2001) Echanisms nderlying ubiquitination. *Annu. Rev. Biochem.* **70**, 503–33

Komander, D., Clague, M. J., and Urbé, S. (2009) Breaking the chains: structure and function of the deubiquitinases. *Nat. Rev. Mol. Cell Biol.* **10**, 550–563

Komander, D., and Rape, M. (2012) The Ubiquitin Code. *Annual Review of Biochemistry,* **81,** 203-229.

Swatek, K. N., and Komander, D. (2016) Ubiquitin modifications. *Cell Res.* **26**, 399–422

Ikeda, F., and Dikic, I. (2008) Atypical Ubiquitin Chains: New Molecular Signals. 'Protein Modifications: Beyond the Usual Suspects' Review Series. *EMBO Rep*. **9**, 536–542

Gatti, M., Pinato, S., Maiolica, A., Rocchio, F., Prato, M., Aebersold, R., and Penengo, L. (2015) RNF168 promotes noncanonical K27ubiquitination to signal DNA damage. *Cell Rep*. **10**, 226–238

Morris, J. R., and Solomon, E. (2004) BRCA1: BARD1 induces the formation of conjugated ubiquitin structures, dependent on K6 of ubiquitin, in cells during DNA replication and repair. *Hum. Mol. Genet*. **13**, 807–817

Elia, A. E. H., Boardman, A. P., Wang, D. C., Huttlin, E. L., Everley, R. A., Dephoure, N., Zhou, C., Koren, I., Gygi, S. P., and Elledge, S. J. (2015) Quantitative Proteomic Atlas of Ubiquitination and Acetylation in the DNA Damage Response. *Mol. Cell*. **59**, 867–881

Nakamura, K., Kato, A., Kobayashi, J., Yanagihara, H., Sakamoto, S., Oliveira, D. V. N. P., Shimada, M., Tauchi, H., Suzuki, H., Tashiro, S., Zou, L., and Komatsu, K. (2011) Regulation of Homologous Recombination by RNF20-Dependent H2B Ubiquitination. *Mol. Cell*. **41**, 515–528

Zhu, B., Zheng, Y., Pham, A. D., Mandal, S. S., Erdjument-Bromage, H., Tempst, P., and Reinberg, D. (2005) Monoubiquitination of human histone H2B: The factors involved and their roles in HOX gene regulation. *Mol. Cell*. **20**, 601–611

Kim, J., Hake, S. B., and Roeder, R. G. (2005) The human homolog of yeast BRE1 functions as a transcriptional coactivator through direct activator interactions. *Mol. Cell*. **20**, 759–770

Moyal, L., Lerenthal, Y., Gana-Weisz, M., Mass, G., So, S., Wang, S. Y., Eppink, B., Chung, Y. M., Shalev, G., Shema, E., Shkedy, D., Smorodinsky, N. I., van Vliet, N., Kuster, B., Mann, M., Ciechanover, A., Dahm-Daphi, J., Kanaar, R., Hu, M. C. T., Chen, D. J., Oren, M., and Shiloh, Y. (2011) Requirement of ATM-Dependent Monoubiquitylation of Histone H2B for Timely Repair of DNA Double-Strand Breaks. *Mol. Cell*. **41**, 529–542

Ismail, H., Andrin, C., McDonald, D., and Hendzel, M. J. (2010) BMI1-mediated histone ubiquitylation promotes DNA double-strand break repair. *J. Cell Biol*. **191**, 45–60

Ginjala, V., Nacerddine, K., Kulkarni, A., Oza, J., Hill, S. J., Yao, M., Citterio, E., van Lohuizen, M., and Ganesan, S. (2011) BMI1 is recruited to DNA breaks and contributes to DNA damage-induced H2A ubiquitination and repair. *Mol. Cell.,* **31**, 1972-1982

Pan, M. R., Peng, G., Hungs, W. C., and Lin, S. Y. (2011) Monoubiquitination of H2AX protein regulates DNA damage response signaling. *J. Biol. Chem.* **286**, 28599–28607

Huen, M. S. Y., Grant, R., Manke, I., Minn, K., Yu, X., Yaffe, M. B., and Chen, J. (2007) RNF8 Transduces the DNA-Damage Signal via Histone Ubiquitylation and Checkpoint Protein Assembly. *Cell.* **131**, 901–914

Mailand, N., Bekker-Jensen, S., Faustrup, H., Melander, F., Bartek, J., Lukas, C., and Lukas, J. (2007) RNF8 Ubiquitylates Histones at DNA Double-Strand Breaks and Promotes Assembly of Repair Proteins. *Cell.* **131**, 887–900

Kolas, N. K., Chapman, J. R., Nakada, S., Ylanko, J., Chahwan, R., Sweeney, F. D., Panier, S., Mendez, M., Wildenhain, J., Thomson, T. M., Pelletier, L., Jackson, S. P., and Durocher, D. (2007) Orchestration of the DNA-damage response by the RNF8 ubiquitin ligase. *Science.* **318**, 1637–1640

Thorslund, T., Ripplinger, A., Hoffmann, S., Wild, T., Uckelmann, M., Villumsen, B., Narita, T., Sixma, T. K., Choudhary, C., Bekker-Jensen, S., and Mailand, N. (2015) Histone H1 couples initiation and amplification of ubiquitin signalling after DNA damage. *Nature.* **527**, 389–393

Doil, C., Mailand, N., Bekker-Jensen, S., Menard, P., Larsen, D. H., Pepperkok, R., Ellenberg, J., Panier, S., Durocher, D., Bartek, J., Lukas, J., and Lukas, C. (2009) RNF168 Binds and Amplifies Ubiquitin Conjugates on Damaged Chromosomes to Allow Accumulation of Repair Proteins. *Cell.* **136**, 435–446

Stewart, G. S., Panier, S., Townsend, K., Al-Hakim, A. K., Kolas, N. K., Miller, E. S., Nakada, S., Ylanko, J., Olivarius, S., Mendez, M., Oldreive, C., Wildenhain, J., Tagliaferro, A., Pelletier, L., Taubenheim, N., Durandy, A., Byrd, P. J., Stankovic, T., Taylor, A. M. R., and Durocher, D. (2009) The RIDDLE Syndrome Protein Mediates a Ubiquitin-Dependent Signaling Cascade at Sites of DNA Damage. *Cell.* **136**, 420–434

Wang, B., and Elledge, S. J. (2007) Ubc13/Rnf8 ubiquitin ligases control foci formation of the Rap80/Abraxas/Brca1/Brcc36 complex in response to DNA damage. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 20759–20763

Mattiroli, F., Vissers, J. H. A., Van Dijk, W. J., Ikpa, P., Citterio, E., Vermeulen, W., Marteijn, J. A., and Sixma, T. K. (2012) RNF168 ubiquitinates K13-15 on H2A/H2AX to drive DNA damage signaling. *Cell.* **150**, 1182–1195

Gatti, M., Pinato, S., Maspero, E., Soffientini, P., Polo, S., and Penengo, L. (2012) A novel ubiquitin mark at the N-terminal tail of histone H2As targeted by RNF168 ubiquitin ligase. *Cell Cycle*. **11**, 2538–2544

Kim, H., Chen, J., and Yu, X. (2007) Ubiquitin-Binding Protein RAP80 Mediates BRCA1-Dependent DNA Damage Response. *Science*, **316**, 1202-1205

Anamika, and Spyracopoulos, L. (2016) Molecular basis for phosphorylation-dependent SUMO recognition by the DNA repair protein RAP80. *J. Biol. Chem*. **291**, 4417–4428

Sobhian, B., Shao, G., Lilli, D. R., Culhane, A. C., Moreau, L. A., Xia, B., Livingston, D. M., and Greenberg, R. A. (2007) RAP80 Targets BRCA1 to Specific Ubiquitin Structures at DNA Damage Sites. *Science*, **316**, 1198-1202

Botuyan, M. V., Lee, J., Ward, I. M., Kim, J. E., Thompson, J. R., Chen, J., and Mer, G. (2006) Structural Basis for the Methylation State-Specific Recognition of Histone H4-K20 by 53BP1 and Crb2 in DNA Repair. *Cell*. **127**, 1361–1373

Wilson, M. D., Benlekbir, S., Fradet-Turcotte, A., Sherker, A., Julien, J.-P., McEwan, A., Noordermeer, S. M., Sicheri, F., Rubinstein, J. L., and Durocher, D. (2016) The structural basis of modified nucleosome recognition by 53BP1. *Nature*. **536**, 100–103

Fradet-Turcotte, A., Canny, M. D., Escribano-Díaz, C., Orthwein, A., Leung, C. C. Y., Huang, H., Landry, M.-C., Kitevski-LeBlanc, J., Noordermeer, S. M., Sicheri, F., and Durocher, D. (2013) 53BP1 is a reader of the DNA-damage-induced H2A Lys 15 ubiquitin mark. *Nature*. **499**, 50–54

Min, J., Allali-Hassani, A., Nady, N., Qi, C., Ouyang, H., Liu, Y., MacKenzie, F., Vedadi, M., and Arrowsmith, C. H. (2007) L3MBTL1 recognition of mono- and dimethylated histones. *Nat. Struct. Mol. Biol*. **14**, 1229–1230

Acs, K., Luijsterburg, M. S., Ackermann, L., Salomons, F. a, Hoppe, T., and Dantuma, N. P. (2011) The AAA-ATPase VCP/p97 promotes 53BP1 recruitment by removing L3MBTL1 from DNA double-strand breaks. *Nat. Struct. Mol. Biol*. **18**, 1345–1350

Mallette, F. A., Mattiroli, F., Cui, G., Young, L. C., Hendzel, M. J., Mer, G., Sixma, T. K., and Richard, S. (2012) RNF8- and RNF168-dependent degradation of KDM4A/JMJD2A triggers 53BP1 recruitment to DNA damage sites. *EMBO J*. **31**, 1865–1878

Nakada, S. (2016) Opposing roles of RNF8/RNF168 and deubiquitinating enzymes in ubiquitination-dependent DNA double-strand break response signaling and DNA-repair

pathway choice. *J. Radiat. Res.* **57**, i33–i40

Lok, G. T.-M., Sy, S. M.-H., Dong, S.-S., Ching, Y.-P., Tsao, S. W., Thomson, T. M., and Huen, M. S. Y. (2012) Differential regulation of RNF8-mediated Lys48- and Lys63-based poly-ubiquitylation. *Nucleic Acids Res.* **40**, 196–205

Bekker-Jensen, S., Rendtlew Danielsen, J., Fugger, K., Gromova, I., Nerstedt, A., Lukas, C., Bartek, J., Lukas, J., and Mailand, N. (2010) HERC2 coordinates ubiquitin-dependent assembly of DNA repair factors on damaged chromosomes. *Nat. Cell Biol.* **12**, 12–80

Mohiuddin, Kobayashi, S., Keka, I. S., Guilbaud, G., Sale, J., Narita, T., Abdel-Aziz, H. I., Wang, X., Ogawa, S., Sasanuma, H., Chiu, R., Oestergaard, V. H., Lisby, M., and Takeda, S. (2016) The role of HERC2 and RNF8 ubiquitin E3 ligases in the promotion of translesion DNA synthesis in the chicken DT40 cell line. *DNA Repair (Amst).* **40**, 67–76

Oestergaard, V. H., Pentzold, C., Pedersen, R. T., Iosif, S., Alpi, A., Bekker-Jensen, S., Mailand, N., and Lisby, M. (2012) RNF8 and RNF168 but not HERC2 are required for DNA damage-induced ubiquitylation in chicken DT40 cells. *DNA Repair (Amst).* **11**, 892–905

Liu, C., Wu, J., Paudyal, S. C., You, Z., and Yu, X. (2013) CHFR is important for the first wave of ubiquitination at DNA damage sites. *Nucleic Acids Res.* **41**, 1698–1710

Bothos, J., Summers, M. K., Venere, M., Scolnick, D. M., and Halazonetis, T. D. (2003) The Chfr mitotic checkpoint protein functions with Ubc13-Mms2 to form Lys63-linked polyubiquitin chains. *Oncogene.* **22**, 7101–7107

Ismail, I. H., Gagné, J.-P., Genois, M.-M., Strickfaden, H., McDonald, D., Xu, Z., Poirier, G. G., Masson, J.-Y., and Hendzel, M. J. (2015) The RNF138 E3 ligase displaces Ku to promote DNA end resection and regulate DNA repair pathway choice. *Nat. Cell Biol.* **17**, 1446–57

Schmidt, C. K., Galanty, Y., Sczaniecka-Clift, M., Coates, J., Jhujh, S., Demir, M., Cornwell, M., Beli, P., and Jackson, S. P. (2015) Systematic E2 screening reveals a UBE2D–RNF138–CtIP axis promoting DNA repair. *Nat. Cell Biol.* **17**, 1458–1470

Prakash, R., Zhang, Y., Feng, W., and Jasin, M. (2015) Homologous Recombination and Human Health. *Perspect. Biol.* 7(**4**); a016600

Wu-Baer, F., Lagrazon, K., Yuan, W., and Baer, R. (2003) The BRCA1/BARD1 heterodimer assembles polyubiquitin chains through an unconventional linkage involving

lysine residue K6 of ubiquitin. *J. Biol. Chem.* **278**, 34743–34746

Baer, R., and Ludwig, T. (2002) The BRCA1/BARD1 heterodimer, a tumor suppressor complex with ubiquitin E3 ligase activity. *Curr. Opin. Genet. Dev.* **12**, 86–91

Kalb, R., Mallery, D. L., Larkin, C., Huang, J. T. J., and Hiom, K. (2014) BRCA1 is a histone-H2A-specific ubiquitin ligase. *Cell Rep.* **8**, 999–1005

Yu, X., Fu, S., Lai, M., Baer, R., Chen, J., Breast, B., and Susceptibility, C. (2006) BRCA1 ubiquitinates its binding partner CtIP. *Genes Dev.* **1**, 1721–1726

Drost, R., Bouwman, P., Rottenberg, S., Boon, U., Schut, E., Klarenbeek, S., Klijn, C., van der Heijden, I., van der Gulden, H., Wientjens, E., Pieterse, M., Catteau, A., Green, P., Solomon, E., Morris, J. R., and Jonkers, J. (2011) BRCA1 RING function is essential for tumor suppression but dispensable for therapy resistance. *Cancer Cell.* **20**, 797–809

Shakya, R., Reid, L. J., Reczek, C. R., Cole, F., Egli, D., Lin, C.-S., deRooij, D. G., Hirsch, S., Ravi, K., Hicks, J. B., Szabolcs, M., Jasin, M., Baer, R., and Ludwig, T. (2011) BRCA1 Tumor Suppression Depends on BRCT Phosphoprotein Binding, But Not Its E3 Ligase Activity. *Science.* **334**, 525–528

Chen, X., Arciero, C. A., Wang, C., Broccoli, D., and Godwin, A. K. (2006) BRCC36 is essential for ionizing radiation-induced BRCA1 phosphorylation and nuclear foci formation. *Cancer Res.* **66**, 5039–5046

Shao, G., Lilli, D. R., Patterson-Fortin, J., Coleman, K. a, Morrissey, D. E., and Greenberg, R. a (2009) The Rap80-BRCC36 de-ubiquitinating enzyme complex antagonizes RNF8-Ubc13-dependent ubiquitination events at DNA double strand breaks. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 3166–3171

Nakada, S., Tai, I., Panier, S., Al-Hakim, A., Iemura, S.-I., Juang, Y.-C., O'Donnell, L., Kumakubo, A., Munro, M., Sicheri, F., Gingras, A.-C., Natsume, T., Suda, T., and Durocher, D. (2010) Non-canonical inhibition of DNA damage-dependent ubiquitination by OTUB1. *Nature.* **466**, 941–946

Yun, M. H., and Hiom, K. (2009) Understanding the functions of BRCA1 in the DNA-damage response. *Biochem. Soc. Trans.* **37**, 597–604

Wang, B. (2012) BRCA1 tumor suppressor network: focusing on its tail. *Cell Biosci.* **2**, 6

Jacq, X., Kemp, M., Martin, N. M. B., and Jackson, S. P. (2013) Deubiquitylating Enzymes and DNA Damage Response Pathways. *Cell Biochem. Biophys.* **67**, 25–43

Messick, T. E., and Greenberg, R. A. (2009) The ubiquitin landscape at DNA double-strand breaks. *J. Cell Biol.* **187**, 319–326

Wiener, R., Zhang, X., Wang, T., and Wolberger, C. (2012) The mechanism of OTUB1-mediated inhibition of ubiquitination. *Nature.* **483**, 618–22

Juang, Y. C., Landry, M. C., Sanches, M., Vittal, V., Leung, C. C. Y., Ceccarelli, D. F., Mateo, A. R. F., Pruneda, J. N., Mao, D. Y. L., Szilard, R. K., Orlicky, S., Munro, M., Brzovic, P. S., Klevit, R. E., Sicheri, F., and Durocher, D. (2012) OTUB1 Co-opts Lys48-Linked Ubiquitin Recognition to Suppress E2 Enzyme Function. *Mol. Cell.* **45**, 384–397

Butler, L. R., Densham, R. M., Jia, J., Garvin, A. J., Stone, H. R., Shah, V., Weekes, D., Festy, F., Beesley, J., and Morris, J. R. (2012) The proteasomal de-ubiquitinating enzyme POH1 promotes the double-strand DNA break response. *EMBO J.* **31**, 3918–34

Morris, J. R. (2012) Attenuation of the ubiquitin conjugate DNA damage signal by the proteasomal DUB POH1. *Cell Cycle.* **11**, 4103–4104

Lancini, C., van den Berk, P. C. M., Vissers, J. H. a, Gargiulo, G., Song, J.-Y., Hulsman, D., Serresi, M., Tanger, E., Blom, M., Vens, C., van Lohuizen, M., Jacobs, H., and Citterio, E. (2014) Tight regulation of ubiquitin-mediated DNA damage response by USP3 preserves the functional integrity of hematopoietic stem cells. *J. Exp. Med.* **211**, 1759–77

Chen, D., Dai, C., and Jiang, Y. (2015) Histone H2A and H2B Deubiquitinase in Developmental Disease and Cancer. *Cancer Transl. Med.* **1**, 170

Müller, S., Hoege, C., Pyrowolakis, G., and Jentsch, S. (2001) SUMO, ubiquitin's mysterious cousin. *Nat. Rev. Mol. Cell Biol.* **2**, 202–210

Gareau, J. R., and Lima, C. D. (2010) The SUMO pathway: emerging mechanisms that shape specificity, conjugation and recognition. *Nat. Rev. Mol. Cell Biol.* **11**, 861–871

Morris, J. R., and Garvin, A. J. (2017) SUMO in the DNA Double-Stranded Break Response : Similarities , Differences , and Cooperation with Ubiquitin. *J. Mol. Biol.* 10.1016/j.jmb.2017.05.012

Matic, I., van Hagen, M., Schimmel, J., Macek, B., Ogg, S. C., Tatham, M. H., Hay, R. T., Lamond, A. I., Mann, M., and Vertegaal, A. C. O. (2007) In Vivo Identification of Human Small Ubiquitin-like Modifier Polymerization Sites by High Accuracy Mass Spectrometry and an in Vitro to in Vivo Strategy. *Mol. Cell. Proteomics.* **7**, 132–144

Liang, Y., Lee, C., Yao, Y., Lai, C., and Schmitz, M. L. (2016) SUMO5 , a Novel Poly-

SUMO Isoform , Regulates PML Nuclear Bodies. *Nat. Publ. Gr.* 10.1038/srep26509

Kerscher, O. (2007) SUMO junction-what's your function? New insights through SUMO-interacting motifs. *EMBO Rep.* **8**, 550–555

Morris, J. R., Boutell, C., Keppler, M., Densham, R., Weekes, D., Alamshah, A., Butler, L., Galanty, Y., Pangon, L., Kiuchi, T., Ng, T., and Solomon, E. (2009) The SUMO modification pathway is involved in the BRCA1 response to genotoxic stress. *Nature.* **462**, 886–890

Galanty, Y., Belotserkovskaya, R., Coates, J., Polo, S., Miller, K. M., and Jackson, S. P. (2009) Mammalian SUMO E3-ligases PIAS1 and PIAS4 promote responses to DNA double-strand breaks. *Nature.* **462**, 935–939

Ismail, I. H., Gagné, J. P., Caron, M. C., McDonald, D., Xu, Z., Masson, J. Y., Poirier, G. G., and Hendzel, M. J. (2012) CBX4-mediated SUMO modification regulates BMI1 recruitment at sites of DNA damage. *Nucleic Acids Res.* **40**, 5497–5510

Luo, K., Zhang, H., Wang, L., Yuan, J., and Lou, Z. (2012) Sumoylation of MDC1 is important for proper DNA damage response. *EMBO J.* **31**, 3008–3019

Guzzo, C. M., Berndsen, C. E., Zhu, J., Gupta, V., Greenberg, R. A., Wolberger, C., Matunis, M. J., Guzzo, C. M., Berndsen, C. E., Zhu, J., Gupta, V., and Datta, A. (2013) RNF4-Dependent Hybrid SUMO-Ubiquitin Chains Are Signals for RAP80 and Thereby Mediate the Recruitment of BRCA1 to Sites of DNA Damage. *Sci. Signal.* **5**, ra88

Hu, X., Paul, A., and Wangs, B. (2012) Rap80 protein recruitment to DNA double-strand breaks requires binding to both small ubiquitin-like modifier (SUMO) and ubiquitin conjugates. *J. Biol. Chem.* **287**, 25510–25519

Yan, J., Kim, Y., Yang, X., Albers, M., Koegl, M., and Jetten, A. M. (2007) Ubiquitin-interaction motifs of RAP80 are critical in its regulation of estrogen receptor a. **35**, 1673–1686

Wang, B., Matsuoka, S., Ballif, B., Zhang, D., Smogorzewska, A., Giyi, S., and Elledge, S. J. (2007) Abraxas and RAP80 form a BRCA1 Protein Complex required for the DNA damage response. *Science*, **316**, 1194–1198

Yan, J., Yang, X. P., Kim, Y. S., Joo, J. H., and Jetten, A. M. (2007) RAP80 interacts with the SUMO-conjugating enzyme UBC9 and is a novel target for sumoylation. *Biochem. Biophys. Res. Commun.* **362**, 132–138

Chen, Y. J., Chuang, Y. C., Chuang, C. N., Cheng, Y. H., Chang, C. R., Leng, C. H., and Wang, T. F. (2016) S. cerevisiae Mre11 recruits conjugated SUMO moieties to facilitate the assembly and function of the Mre11-Rad50-Xrs2 complex. *Nucleic Acids Res*. **44**, 2199–2213

Bologna, S., Altmannova, V., Valtorta, E., Koenig, C., Liberali, P., Gentili, C., Anrather, D., Ammerer, G., Pelkmans, L., Krejci, L., and Ferrari, S. (2015) Sumoylation regulates EXO1 stability and processing of DNA damage. *Cell Cycle*. **14**, 2439–2450

Garvin, A. J., Densham, R. M., Blair-Reid, S. A., Pratt, K. M., Stone, H. R., Weekes, D., Lawrence, K. J., and Morris, J. R. (2013) The deSUMOylase SENP7 promotes chromatin relaxation for homologous recombination DNA repair. *EMBO Rep*. **14**, 975–983

Her, J., Soo Lee, N., Kim, Y., and Kim, H. (2016) Factors forming the BRCA1-A complex orchestrate BRCA1 recruitment to the sites of DNA damage. *Acta Biochim. Biophys. Sin. (Shanghai)*. **48**, 658–664

Nikkilä, J., Coleman, K. A., Morrissey, D., Pylkäs, K., Erkko, H., Messick, T. E., Karppinen, S.-M., Amelina, A., Winqvist, R., and Greenberg, R. A. (2009) Familial breast cancer screening reveals an alteration in the RAP80 UIM domain that impairs DNA damage response function. *Oncogene*. **28**, 1843–1852

Lombardi, P. M., Matunis, M. J., and Wolberger, C. (2017) RAP80, ubiquitin and SUMO in the DNA damage response. *J. Mol. Med*. **95**, 799–807

Kim, H., Chen, J., and Yu, X. (2007) Ubiquitin-binding protein RAP80 mediates BRCA1-dependent DNA damage response. *Science*. **316**, 1202–1205

Yan, J., Yang, X. P., Kim, Y. S., and Jetten, A. M. (2008) RAP80 responds to DNA damage induced by both ionizing radiation and UV irradiation and is phosphorylated at Ser205. *Cancer Res*. **68**, 4269–4276

Hurley, J. H., Lee, S., and Prag, G. (2006) Ubiquitin-binding domains. **372**, 361–372

Dikic, I., Wakatsuki, S., and Walters, K. J. (2009) Ubiquitin-binding domains — from structures to functions. *Nat. Rev. Mol. Cell Biol*. **10**, 659–671

Fisher, R. D., Wang, B., Alam, S. L., Higginson, D. S., Robinson, H., Sundquist, W. I., and Hill, C. P. (2003) Structure and Ubiquitin Binding of the Ubiquitin-interacting Motif *. **278**, 28976–28984

Hofmann, K., and Falquet, L. (2001) A ubiquitin-interacting motif conserved in components

of the proteasomal and lysosomal protein degradation systems. **26**, 347–350

Young, P., Deveraux, Q., Beal, R. E., Pickart, C. M., and Rechsteiner, M. (1998) Characterization of Two Polyubiquitin Binding Sites in the 26 S Protease Subunit 5a *. **273**, 5461–5467

Sato, Y., Yoshikawa, A., Mimura, H., Yamashita, M., Yamagata, A., and Fukai, S. (2009) Structural basis for specific recognition of Lys 63-linked polyubiquitin chains by tandem UIMs of RAP80. *EMBO J.* **28**, 2461–8

Markin, C. J., Xiao, W., and Spyracopoulos, L. (2010) Mechanism for recognition of polyubiquitin chains: Balancing affinity through interplay between multivalent binding and dynamics. *J. Am. Chem. Soc.* **132**, 11247–11258

Anamika, Markin, C. J., Rout, M. K., and Spyracopoulos, L. (2014) Molecular basis for impaired DNA damage response function associated with the RAP80 ΔE81 defect. *J. Biol. Chem.* **289**, 12852–12862

Minty, A., Dumont, X., Kaghad, M., and Caput, D. (2000) Covalent Modification of p73 ␣ by SUMO-1. **275**, 36316–36323

Hecker, C. M., Rabiller, M., Haglund, K., Bayer, P., and Dikic, I. (2006) Specification of SUMO1- and SUMO2-interacting motifs. *J. Biol. Chem.* **281**, 16117–16127

Heun, P. (2007) SUMOrganization of the nucleus. *Curr. Opin. Cell Biol.* **19**, 350–355

Cappadocia, L., Mascle, X. H., Bourdeau, V., Tremblay-Belzile, S., Chaker-Margot, M., Lussier-Price, M., Wada, J., Sakaguchi, K., Aubry, M., Ferbeyre, G., and Omichinski, J. G. (2015) Structural and Functional Characterization of the Phosphorylation-Dependent Interaction between PML and SUMO1. *Structure.* **23**, 126–138

Escobar-cabrera, E., Okon, M., Lau, K. W., Dart, C. F., Alexandre, M., Bonvin, J. J., and Mcintosh, L. P. (2011) Characterizing the N- and C-terminal Small Ubiquitin-like Modifier (SUMO) -interacting Motifs of the Scaffold Protein. 10.1074/jbc.M111.231647

Perry, J. J. P., Tainer, J. A., and Boddy, M. N. (2008) A SIM-ultaneous role for SUMO and ubiquitin. *Trends Biochem. Sci.* **33**, 201–208

Galanty, Y., Belotserkovskaya, R., Coates, J., and Jackson, S. P. (2012) RNF4, a SUMO-rageted ubiquitin E3 ligase , promotes DNA double-strand break repair. *Genes Dev.* **26**, 1179–1195

Plechanovová, A., Jaffray, E. G., McMahon, S. A., Johnson, K. A., Navrátilová, I., Naismith, J. H., and Hay, R. T. (2011) Mechanism of ubiquitylation by dimeric RING ligase RNF4. *Nat. Struct. Mol. Biol.* **18**, 1052–1059

Sobhian, B., Shao, G. Z., Lilli, D. R., Culhane, A. C., Moreau, L. A., Xia, B., Livingston, D. M., and Greenberg, R. A. (2007) RAP80 targets BRCA1 to specific ubiquitin structures at DNA damage sites. *Science* **316**, 1198-1202

Wang, B., Matsuoka, S., Ballif, B. A., Zhang, D., Smogorzewska, A., Gygi, S. P., and Elledge, S. J. (2007) Abraxas and RAP80 form a BRCA1 protein complex required for the DNA damage response. *Science* **316**, 1194-1198

Nikkila, J., Coleman, K. A., Morrissey, D., Pylkas, K., Erkko, H., Messick, T. E., Karppinen, S. M., Amelina, A., Winqvist, R., and Greenberg, R. A. (2009) Familial breast cancer screening reveals an alteration in the RAP80 UIM domain that impairs DNA damage response function. *Oncogene* **28**, 1843-1852

Yin, Z., Menendez, D., Resnick, M. A., French, J. E., Janardhan, K. S., and Jetten, A. M. (2012) RAP80 is critical in maintaining genomic stability and suppressing tumor development. *Cancer Res.* **72**, 5080-5090

Osorio, A., Barroso, A., Garcia, M. J., Martinez-Delgado, B., Urioste, M., and Benitez, J. (2009) Evaluation of the BRCA1 interacting genes RAP80 and CCDC98 in familial breast cancer susceptibility. *Breast Cancer Res. Tr.* **113**, 371-376

Rebbeck, T. R., Mitra, N., Domchek, S. M., Wan, F., Friebel, T. M., Tran, T. V., Singer, C. F., Tea, M. K., Blum, J. L., Tung, N., Olopade, O. I., Weitzel, J. N., Lynch, H. T., Snyder, C. L., Garber, J. E., Antoniou, A. C., Peock, S., Evans, D. G., Paterson, J., Kennedy, M. J., Donaldson, A., Dorkins, H., Easton, D. F., Epidemiological Study of, B., Carriers, B. M., Rubinstein, W. S., Daly, M. B., Isaacs, C., Nevanlinna, H., Couch, F. J., Andrulis, I. L., Freidman, E., Laitman, Y., Ganz, P. A., Tomlinson, G. E., Neuhausen, S. L., Narod, S. A., Phelan, C. M., Greenberg, R., and Nathanson, K. L. (2011) Modification of BRCA1-associated breast and ovarian cancer risk by BRCA1-interacting genes. *Cancer Res.* **71**, 5792-5805

Sims, J. J., and Cohen, R. E. (2009) Linkage-specific avidity defines the lysine 63-linked polyubiquitin-binding preference of RAP80. *Mol. Cell* **33**, 775-783

Markin, C. J., Wei, X. A., and Spyracopoulos, L. (2010) Mechanism for recognition of polyubiquitin chains: Balancing affinity through interplay between multivalent binding and dynamics. *J. Am. Chem. Soc.* **132**, 11247-11258

Sato, Y., Yoshikawa, A., Mimura, H., Yamashita, M., Yamagata, A., and Fukai, S. (2009) Structural basis for specific recognition of Lys 63-linked polyubiquitin chains by tandem UIMs of RAP80. *EMBO J.* **28**, 2461-2468

Marley, J., Lu, M., and Bracken, C. (2001) A method for efficient isotopic labeling of recombinant proteins. *J Biomol. NMR* **20**, 71-75

Wittekind, M., and Mueller, L. (1993) HNCACB, a high-sensitivity 3D NMR experiment to correlate amide proton and nitrogen resonances with the a- and a-carbon resonances in proteins. *J. Magn. Reson.* **B 101**, 201-205

Muhandiram, D. R., and Kay, L. E. (1994) Gradient-enhanced triple-resonance 3-dimensional NMR experiments with improved sensitivity. *J. Magn. Reson.* **B 103**, 203-216

Grzesiek, S., and Bax, A. (1992) Correlating backbone amide and side chain resonances in larger proteins by multiple relayed triple resonance NMR. *J. Am. Chem. Soc.* **114**, 6291-6293

Grzesiek, S., and Bax, A. (1993) Amino acid type determination in the sequential assignment procedure of uniformly ^{13}C/^{15}N enriched proteins. *J. Biomol. NMR* **3**, 185-204

Keller, R. L. J. (2004) The computer aided resonance assignment tutorial. Cantina Verlag, Zürich

Goddard, T. D., and Kneller, D. G. SPARKY 3. University of California, San Francisco

Farrow, N. A., Muhandiram, R., Singer, A. U., Pascal, S. M., Kay, C. M., Gish, G., Shoelson, S. E., Pawson, T., Forman-Kay, J. D., and Kay, L. E. (1994) Backbone dynamics of a free and a phosphopeptide-complexed Src homology-2 domain studied by ^{15}N NMR relaxation. *Biochemistry* **33**, 5984-6003

Mandel, A. M., Akke, M., and Palmer, A. G. (1995) Backbone dynamics of *Escherichia coli* ribonuclease HI - Correlations with structure and function in an active enzyme. *J. Mol. Biol.* **246**, 144-163

Spyracopoulos, L. (2006) A suite of Mathematica notebooks for the analysis of protein main chain ^{15}N NMR relaxation data. *J Biomol. NMR* **36**, 215-224

Kay, L. E., Xu, G. Y., and Yamazaki, T. (1994) Enhanced sensitivity triple resonance spectroscopy with minimal H_2O saturation. *J. Magn. Reson.* **A 109**, 129-133

Zhang, O., Kay, L. E., Olivier, J. P., and Forman-Kay, J. D. (1994) Backbone ^1H and ^{15}N resonance assignments of the N-terminal SH3 domain of drk in folded and unfolded states using enhanced-sensitivity pulsed field gradient NMR techniques. *J Biomol. NMR* **4**, 845-858

Shen, Y., and Bax, A. (2013) Protein backbone and sidechain torsion angles predicted from NMR chemical shifts using artificial neural networks. *J Biomol. NMR* **56**, 227-241

Shen, Y., and Bax, A. (2012) Identification of helix capping and b-turn motifs from NMR chemical shifts. *J Biomol. NMR* **52**, 211-232

Schwieters, C. D., Kuszewski, J. J., Tjandra, N., and Clore, G. M. (2003) The Xplor-NIH NMR molecular structure determination package. *J. Magn. Reson.* **160**, 65-73

Wolfram, S. (2003) *The Mathematica book*, 5th ed., Wolfram Media, Champaign, IL

Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J Biomol. NMR* **6**, 277-293

Nesmelova, I., Krushelnitsky, A., Idiyatullin, D., Blanco, F., Ramirez-Alvarado, M., Daragan, V. A., Serrano, L., and Mayo, K. H. (2001) Conformational exchange on the microsecond time scale in a-helix and b-hairpin peptides measured by ^{13}C NMR transverse relaxation. *Biochemistry* **40**, 2844-2853

Kay, L. E., Keifer, P., and Saarinen, T. (1992) Pure absorption gradient enhanced heteronuclear single quantum correlation spectroscopy with improved sensitivity. *J. Am. Chem. Soc.* **114**, 10663-10665

Markin, C. J., and Spyracopoulos, L. (2012) Increased precision for analysis of protein-ligand dissociation constants determined from chemical shift titrations. *J Biomol. NMR* **53**, 125-138

Markin, C. J., and Spyracopoulos, L. (2012) Accuracy and precision of protein-ligand interaction kinetics determined from chemical shift titrations. *J Biomol. NMR* **54**, 355-376

Case, D. A., Darden, T. A., Cheatham, T. E., Simmerling, C. L., Wang, J., Duke, R. E., Luo, R., Walker, R. C., Zhang, W., Merz, K. M., Roberts, B., Wang, B., Hayik, S., Roitberg, A., Seabra, G., Kolossváry, I., Wong, K. F., Paesani, F., Vanicek, J., Liu, J., Wu, X., Brozell, S.

R., Steinbrecher, T., Gohlke, H., Cai, Q., Ye, X., Wang, J., Hsieh, M.-J., Cui, G., Roe, D. R., Mathews, D. H., Seetin, M. G., Sagui, C., Babin, V., Luchko, T., Gusarov, S., Kovalenko, A., and Kollman, P. A. (2010) AMBER 11.

Mongan, J., Simmerling, C., McCammon, J. A., Case, D. A., and Onufriev, A. (2007) Generalized Born model with a simple, robust molecular volume correction. *J. Chem. Theory Comput.* **3**, 156-169

Li, D. W., and Brüschweiler, R. (2010) NMR-based protein potentials. *Angew. Chem. Int. Edit.* **49**, 6778-6780

Gotz, A. W., Williamson, M. J., Xu, D., Poole, D., Le Grand, S., and Walker, R. C. (2012) Routine microsecond molecular dynamics simulations with AMBER on GPUs. 1. Generalized born. *J. Chem. Theory Comput.* **8**, 1542-1555

Kollman, P. A., Massova, I., Reyes, C., Kuhn, B., Huo, S. H., Chong, L., Lee, M., Lee, T., Duan, Y., Wang, W., Donini, O., Cieplak, P., Srinivasan, J., Case, D. A., and Cheatham, T. E. (2000) Calculating structures and free energies of complex molecules: Combining molecular mechanics and continuum models. *Accounts Chem. Res.* **33**, 889-897

Gronenborn, A. M., and Clore, G. M. (1994) Identification of N-terminal helix capping boxes by means of ^{13}C chemical shifts. *J Biomol. NMR* **4**, 455-458

Seale, J. W., Srinivasan, R., and Rose, G. D. (1994) Sequence determinants of the capping box, a stabilizing motif at the N-termini of a-helices. *Protein Sci.* **3**, 1741-1745

Berjanskii, M. V., and Wishart, D. S. (2005) A simple method to predict protein flexibility using secondary chemical shifts. *J. Am. Chem. Soc.* **127**, 14970-14971

Mammen, M., Choi, S. K., and Whitesides, G. M. (1998) Polyvalent interactions in biological systems: Implications for design and use of multivalent ligands and inhibitors. *Angew. Chem. Int. Edit.* **37**, 2755-2794

Fisher, R. D., Wang, B., Alam, S. L., Higginson, D. S., Robinson, H., Sundquist, W. I., and Hill, C. P. (2003) Structure and ubiquitin binding of the ubiquitin-interacting motif. *J. Biol. Chem.* **278**, 28976-28984

Safadi, S. S., and Shaw, G. S. (2010) Differential interaction of the E3 ligase parkin with the proteasomal subunit S5a and the endocytic protein Eps15. *J. Biol. Chem.* **285**, 1424-1434

Anamika, Markin, C. J., Rout, M. K., and Spyracopoulos, L. (2014) Molecular basis for impaired DNA damage response function associated with the RAP80 ΔE81 defect. *J. Biol.*

Chem. **289**, 12852–62

Nikkilä, J., Coleman, K. A., Morrissey, D., Pylkäs, K., Erkko, H., Messick, T. E., Karppinen, S.-M., Amelina, A., Winqvist, R., and Greenberg, R. A. (2009) Familial breast cancer screening reveals an alteration in the RAP80 UIM domain that impairs DNA damage response function. *Oncogene*. **28**, 1843–1852

Anamika, and Spyracopoulos, L. (2016) Molecular basis for phosphorylation-dependent SUMO recognition by the DNA repair protein RAP80. *J. Biol. Chem*. **291**, 4417–4428

Heikkinen, K., Rapakko, K., Karppinen, S.-M., Erkko, H., Knuutila, S., Lundán, T., Mannermaa, A., Børresen-Dale, A.-L., Borg, Å., Barkardottir, R. B., Petrini, J., and Winqvist, R. (2006) RAD50 and NBS1 are breast cancer susceptibility genes associated with genomic instability . *Carcinogenesis*. **27**, 1593–1599

Karppinen, S.-M., Erkko, H., Reini, K., Pospiech, H., Heikkinen, K., Rapakko, K., Syväoja, J. E., and Winqvist, R. (2006) Identification of a common polymorphism in the TopBP1 gene associated with hereditary susceptibility to breast and ovarian cancer. *Eur. J. Cancer*. **42**, 2647–2652

Stratton, M. R., and Rahman, N. (2008) The emerging landscape of breast cancer susceptibility. *Nat Genet*. **40**, 17–22

Panier, S., Ichijima, Y., Fradet-Turcotte, A., Leung, C. C. Y., Kaustov, L., Arrowsmith, C. H., and Durocher, D. (2012) Tandem Protein Interaction Modules Organize the Ubiquitin-Dependent Response to DNA Double-Strand Breaks. *Mol. Cell*. **47**, 383–395

Guzzo, C. M., Berndsen, C. E., Zhu, J., Gupta, V., Greenberg, R. A., Wolberger, C., Matunis, M. J., Guzzo, C. M., Berndsen, C. E., Zhu, J., Gupta, V., and Datta, A. (2013) RNF4-Dependent Hybrid SUMO-Ubiquitin Chains Are Signals for RAP80 and Thereby Mediate the Recruitment of BRCA1 to Sites of DNA Damage. *Sci. Signal*. **5**, ra88

Escobar-cabrera, E., Okon, M., Lau, K. W., Dart, C. F., Alexandre, M., Bonvin, J. J., and Mcintosh, L. P. (2011) Characterizing the N- and C-terminal Small Ubiquitin-like Modifier (SUMO) interacting Motifs of the Scaffold Protein. *J. Biol. Chem*. **286**, 19816-19829

Chang, C. C., Naik, M. T., Huang, Y. S., Jeng, J. C., Liao, P. H., Kuo, H. Y., Ho, C. C., Hsieh, Y. L., Lin, C. H., Huang, N. J., Naik, N. M., Kung, C. C. H., Lin, S. Y., Chen, R. H., Chang, K. S., Huang, T. H., and Shih, H. M. (2011) Structural and functional roles of Daxx

SIM phosphorylation in SUMO paralog-selective binding and apoptosis modulation. *Mol. Cell* **42**, 62-74

Hu, X., Paul, A., and Wangs, B. (2012) Rap80 protein recruitment to DNA double-strand breaks requires binding to both small ubiquitin-like modifier (SUMO) and ubiquitin conjugates. *J. Biol. Chem.* **287**, 25510–25519

Ullmann, R., Chien, C. D., Avantaggiati, M. L., and Muller, S. (2012) An acetylation switch regulates SUMO-dependent protein interaction networks. *Mol. Cell* **46**, 759-770

Cheema, A., Knights, C. D., Rao, M., Catania, J., Perez, R., Simons, B., Dakshanamurthy, S., Kolukula, V. K., Tilli, M., Furth, P. A., Albanese, C., and Avantaggiati, M. L. (2010) Functional mimicry of the acetylated C-terminal tail of p53 by a SUMO-1 acetylated domain, SAD. *J. Cell. Physio.* **225**, 371-384

Choudhary, C., Kumar, C., Gnad, F., Nielsen, M. L., Rehman, M., Walther, T. C., Olsen, J. V., and Mann, M. (2009) Lysine acetylation targets protein complexes and co-regulates major cellular functions. *Science* **325**, 834-840

Lombardi, P. M., Matunis, M. J., and Wolberger, C. (2017) RAP80, ubiquitin and SUMO in the DNA damage response. *J. Mol. Med.* **95**, 799–807

Meggio, F., Marin, O., and Pinna, L. A. (1994) Substrate specificity of protein kinase CK2. *Cell. Mol. Biol. Res.* **40(5-6)**, 401-409

Lee, B. L., Singh A., Glover, J. N. M., Hendzel, M. J. and Spyracopoulos, Leo. (2017) NU SC. *J. Mol. Biol.* 10.1016/j.jmb.2017.05.029

Tatham, M. H., Geoffroy, M.-C., Shen, L., Plechanovova, A., Hattersley, N., Jaffray, E. G., Palvimo, J. J., and Hay, R. T. (2008) RNF4 is a poly-SUMO-specific E3 ubiquitin ligase required for arsenic-induced PML degradation. *Nat. Cell Biol.* **10**, 538–546

Shire, K., Wong, A. I., Tatham, M. H., Anderson, O. F., Ripsman, D., Gulstene, S., Moffat, J., Hay, R. T., and Frappier, L. (2016) Identification of RNF168 as a PML nuclear body regulator. *J. Cell Sci.* **129**, 580–591

Guzzo, C. M., Berndsen, C. E., Zhu, J., Gupta, V., Greenberg, R. A., Wolberger, C., Matunis, M. J., Guzzo, C. M., Berndsen, C. E., Zhu, J., Gupta, V., and Datta, A. (2013) RNF4-Dependent Hybrid SUMO-Ubiquitin Chains Are Signals for RAP80 and Thereby Mediate the Recruitment of BRCA1 to Sites of DNA Damage. *Sci. Signal.* **5**, ra88

Hu, X., Paul, A., and Wangs, B. (2012) Rap80 protein recruitment to DNA double-strand

breaks requires binding to both small ubiquitin-like modifier (SUMO) and ubiquitin conjugates. *J. Biol. Chem.* **287**, 25510–25519

Escobar-cabrera, E., Okon, M., Lau, K. W., Dart, C. F., Alexandre, M., Bonvin, J. J., and Mcintosh, L. P. (2011) Characterizing the N- and C-terminal Small Ubiquitin-like Modifier (SUMO) interacting Motifs of the Scaffold Protein. *J. Biol. Chem.* **286**, 19816-19829

Chang, C. C., Naik, M. T., Huang, Y. S., Jeng, J. C., Liao, P. H., Kuo, H. Y., Ho, C. C., Hsieh, Y. L., Lin, C. H., Huang, N. J., Naik, N. M., Kung, C. C. H., Lin, S. Y., Chen, R. H., Chang, K. S., Huang, T. H., and Shih, H. M. (2011) Structural and functional roles of Daxx SIM phosphorylation in SUMO paralog-selective binding and apoptosis modulation. *Mol. Cell* **42**, 62-74

Anamika, and Spyracopoulos, L. (2016) Molecular basis for phosphorylation-dependent SUMO recognition by the DNA repair protein RAP80. *J. Biol. Chem.* **291**, 4417–4428

Tikoo, S., Madhavan, V., Hussain, M., Miller, E. S., Arora, P., Zlatanou, A., Modi, P., Townsend, K., Stewart, G. S., and Sengupta, S. (2013) Ubiquitin-dependent recruitment of the Bloom Syndrome helicase upon replication stress is required to suppress homologous recombination. *EMBO J.* **32**, 1–15

Soo Lee, N., Jin Chung, H., Kim, H.-J., Yun Lee, S., Ji, J., Seo, Y., Hun Han, S., Choi, M., Yun, M., Lee, S., Myung, K., Kim, Y., Chul Kang, H., and Kim, H. (2016) TRAIP/RNF206 is required for recruitment of RAP80 to sites of DNA damage. *Nat. Commun.* **7**, 10463

Müller, S., Matunis, M. J., Dejean(1998) Conjugation with the ubiquitin-related modifier SUMO-1 regulates the partitioning of PML within the nucleus. *EMBO J.* **17**, 61–70

Kamitani, T., Kito, K., Nguyen, H. P., Wada, H., Fukuda-kamitani, T., and Yeh, E. T. H. (1998) Identification of Three Major Sentrinization Sites in PML. *J. Biol. Chem.* **273**, 26675–26682

Appendix A

The interaction of phosphorylated RAP80$_{37\text{-}49}$ with SUMO-1

Introduction

The SUMO interacting motif of RAP80 has been shown to preferentially interact with the SUMO-2 isoform in pull down assays[1-2]. However, the Daxx SIM, with a highly similar sequence as RAP80 SIM has been shown to interact with both SUMO-1 and SUMO-2 isoforms with a similar affinity in its unphosphorylated state, at a physiological salt concentration[3]. Furthermore, the Daxx SIM upon phosphorylation binds SUMO-1 preferentially over SUMO-2, with a very modest increase in affinity[4]. Please see the table (Table A1) below for the experimentally determined K_D values. To investigate the RAP80 SIM interaction with SUMO-1 isoform, we performed a NMR chemical shift titration and our results indicate that the RAP80 SIM binds SUMO-1 with a similar affinity as SUMO-2.

Peptide	SUMO-1	SUMO-2/3
RAP80 SIM (RLEDAFIVISDSDGEE)	-	$K_D = 239 \pm 54$ μM
RAP80 pSIM (EDAFIVIpSDpSDGE)	-	$K_D = 9 \pm 3$ μM
Daxx SIM (IYKTSVATQCDPEEIIVLSDSD)	$K_D = 110 \pm 14$ μM	$K_D = 170 \pm 24$ μM
Daxx pSIM (KTSVATQCDPEEIIVLpSDpSD)	$K_D = 1.6 \pm 0.4$ μM	$K_D = 11.9 \pm 2.7$ μM

Table A.1: The experimentally determined K_D values for RAP80 and Daxx SIMs in free and phosphorylated states for SUMO-1 and SUMO-2.

Experimental Procedures

Protein and peptide expression and purification

pET28a plasmid encoding 1-97 of human SUMO-1 (P63165) was provided by Dr. Lawrence McIntosh, University of British Columbia. The expression and purification

protocol for [U-^{15}N] labeled SUMO-1 was similar to that for SUMO-2, as previously described[5]. Phosphorylated RAP80 SIM peptide encompassing residues 37-49 (EDAFIVIpSDpSDGE) was chemically synthesized Biomatik[5].

NMR monitored titration of SUMO-1 with phosphorylated RAP80$_{37\text{-}49}$

An NMR-monitored chemical shift titration experiment was carried out using [U-^{15}N] SUMO-1 in buffer containing 50 mM Tris, 150 mM NaCl, 2 mM DTT, and 10% D$_2$O at pH 7.3, with a protein concentration of 200 μM. The sample was prepared in a SHIGEMI microcell NMR tube. Phosphorylated RAP80 peptide was prepared with a concentration of 500 mM in the same buffer as SUMO-1. 2D ^{1}H-^{15}N HSQC spectra were acquired at each addition of stock peptide solution. To ensure that the protein-peptide interaction was saturated, peptide was titrated into SUMO-1 was conducted until no further changes were evident in the NMR spectrum of SUMO-1. The concentrations of [U-^{15}N] SUMO-1 and unlabeled peptide for the 19 titration points were 196, 194, 191, 186, 178, 164, 153, 141, 129, 117, 106, 95, 83, 66, 65, 59, 49, 44, 40 and 0, 1, 4, 7, 13, 22, 30, 39, 47, 55, 63, 71, 79, 82, 86, 109, 147, 165, 250, respectively. For the titration, the pRAP80/SUMO-1 ligand/protein ratios were 0, 0.006, 0.02, 0.04, 0.07, 0.13, 0.2, 0.27, 0.36, 0.46, 0.59, 0.74, 0.94, 1.23, 1.31, 1.84, 3.0, 3.7, and 6.2. The concentrations of the stock protein and peptide solutions were determined using amino acid analyses. The combined chemical shift changes, $\Delta\delta$, for backbone amide ^{1}H and ^{15}N chemical shifts were calculated according to $\Delta\delta = [(\Delta\delta^{15}\text{N}/5)^2 + (\Delta\delta^1\text{H}^\text{N})^2]^{1/2}$, where $\Delta\delta^{15}$N and $\Delta\delta^1$H$^\text{N}$ are the respective ^{15}N and ^{1}H$^\text{N}$ chemical shift changes in ppm. The combined chemical shift changes for each residue were fit to a 1:1 binding isotherm to determine the binding dissociation constant (K_D), as described previously[5].

Results

CK2 phosphorylated RAP80 SIM binds SUMO-1 with a similar binding affinity as SUMO-2.

In our previous work we determined the binding affinity of SUMO-2 for CK2 phosphorylated $RAP80_{37-49}$ and we reported the dissociation constant, K_D to be ~9±3 µM[5]. Here, we used CK2 phosphorylated RAP80 SIM to quantify its binding with SUMO-1 as phosphorylation of SIM has been reported to increase the affinity and specificity for its substrate. As indicated in Figure A.1 a, ^{1}H-^{15}N HSQC spectrum of SUMO-1 (free-red, bound-blue) many residues of SUMO-1 displayed significant change in the chemical shift upon addition of $pRAP80_{37-49}$ and residue wise change in combined chemical shift change is displayed in Figure A.1 b. The residues displaying largest chemical shift change are from 35-47, which is highlighted in pink in the cartoon representation of SUMO-1, Figure A.1 c (PDB ID- 1A5R). The residues showing significant chemical shift changes belonged to the SIM binding cleft of SUMO-1 located between α_1 helix and β_2 strand. The fitting of chemical shift change to protein and peptide concentration for L47 is shown in Figure A.1 d. The chemical shift change of residue L47 is shown in Figure A.1 e and the average K_D values for 25 residues showing significant change in chemical shift was determined to be 8 ± 5 µM.
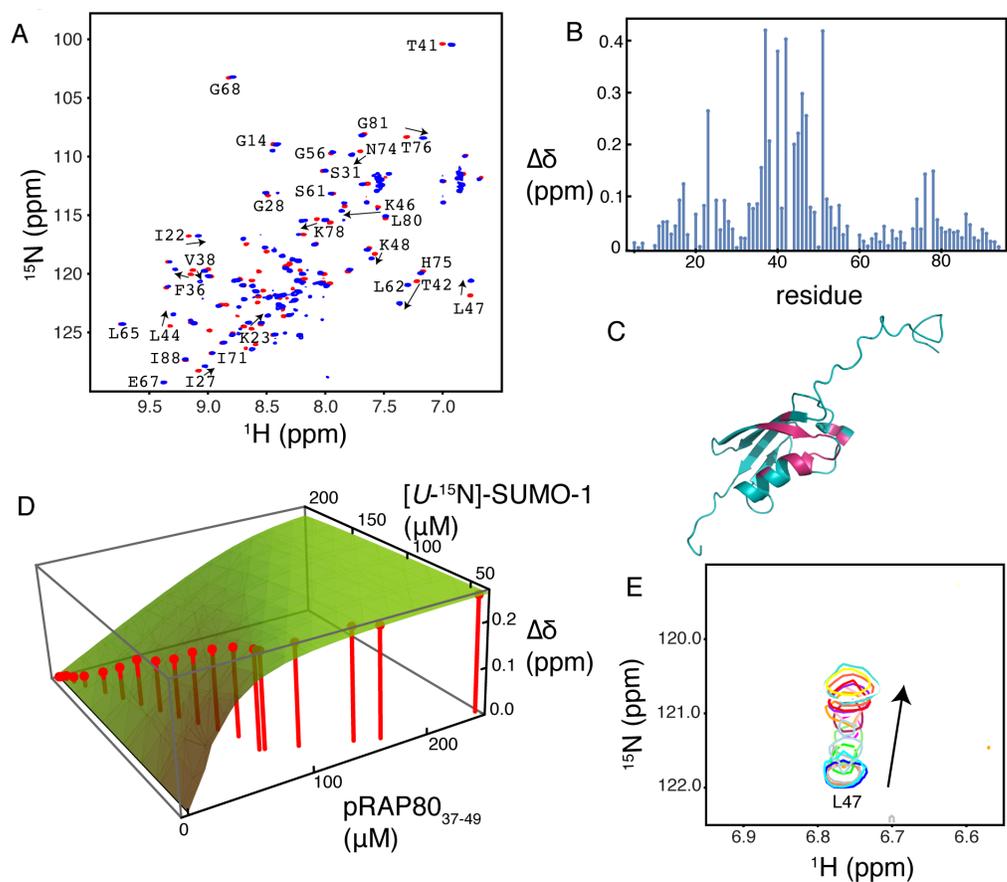
140

Figure A.1: (a), Two dimensional ¹H-¹⁵N HSQC NMR spectra for free SUMO-1(red) and bound to pRAP80₃₇₋₄₉ (blue). The arrows indicate the direction of chemical shift movement from free to bound upon addition of pRAP80₃₇₋₄₉ **(b),** Chemical shift perturbation map of SUMO-1 upon addition of ~6 fold excess of pRAP80₃₇₋₄₉ **(c),** Residues showing significant change in combined chemical shift ($\Delta\delta > 0.15$) and the ones showing broadening beyond detection are colored in pink on SUMO-1 structure (PDB ID- 1A5R) **(d),** The chemical shift changes of residue L47 upon addition of pRAP80₃₇₋₄₉ is displayed on vertical axis and the concentrations of SUMO-1 and pRAP80₃₇₋₄₉ are shown on horizontal axis. Experimental values are denoted in points and the surface represents the best fit to a 1:1 binding isotherm **(e),** ¹H-¹⁵N HSQC spectrum of SUMO-1 with the residue L47 is expanded and displayed at different concentrations of pRAP80₃₇₋₄₉.

Discussion

The RAP80 SIM upon phosphorylation by CK2 binds SUMO-1 with a similar affinity as SUMO-2. The phosphorylated RAP80 SIM binding with SUMO-2 might be relevant in other physiological processes such as its interaction with PML body. PML body is modified by all the SUMO isoforms and our immunofluorescence experiments in Appendix B, and others show that the RAP80 puncta localizes with PML body, which might be dependent on the interaction of RAP80 SIM with SUMO-1 isoform[6-9].

References

[1]   Guzzo, C. M., Berndsen, C. E., Zhu, J., Gupta, V., Greenberg, R. A., Wolberger, C., Matunis, M. J., Guzzo, C. M., Berndsen, C. E., Zhu, J., Gupta, V., and Datta, A. (2013) RNF4-Dependent Hybrid SUMO-Ubiquitin Chains Are Signals for RAP80 and Thereby Mediate the Recruitment of BRCA1 to Sites of DNA Damage. *Sci. Signal.* **5**, ra88

[2]   Hu, X., Paul, A., and Wangs, B. (2012) Rap80 protein recruitment to DNA double-strand breaks requires binding to both small ubiquitin-like modifier (SUMO) and ubiquitin conjugates. *J. Biol. Chem.* **287**, 25510–25519

[3]   Escobar-cabrera, E., Okon, M., Lau, K. W., Dart, C. F., Alexandre, M., Bonvin, J. J., and Mcintosh, L. P. (2011) Characterizing the N- and C-terminal Small Ubiquitin-like Modifier (SUMO) interacting Motifs of the Scaffold Protein. *J. Biol. Chem.* **286**, 19816-19829

[4]   Chang, C. C., Naik, M. T., Huang, Y. S., Jeng, J. C., Liao, P. H., Kuo, H. Y., Ho, C. C., Hsieh, Y. L., Lin, C. H., Huang, N. J., Naik, N. M., Kung, C. C. H., Lin, S. Y., Chen, R. H., Chang, K. S., Huang, T. H., and Shih, H. M. (2011) Structural and functional roles of Daxx SIM phosphorylation in SUMO paralog-selective binding and apoptosis modulation. *Mol. Cell* **42**, 62-74

[5]   Anamika, and Spyracopoulos, L. (2016) Molecular basis for phosphorylation-dependent SUMO recognition by the DNA repair protein RAP80. *J. Biol.*

Chem. **291**, 4417–4428

[6]   Tikoo, S., Madhavan, V., Hussain, M., Miller, E. S., Arora, P., Zlatanou, A., Modi, P., Townsend, K., Stewart, G. S., and Sengupta, S. (2013) Ubiquitin-dependent recruitment of the Bloom Syndrome helicase upon replication stress is required to suppress homologous recombination. *EMBO J.* **32**, 1–15

[7]   Soo Lee, N., Jin Chung, H., Kim, H.-J., Yun Lee, S., Ji, J., Seo, Y., Hun Han, S., Choi, M., Yun, M., Lee, S., Myung, K., Kim, Y., Chul Kang, H., and Kim, H. (2016) TRAIP/RNF206 is required for recruitment of RAP80 to sites of DNA damage. *Nat. Commun.* **7**, 10463

[8]   Müller, S., Matunis, M. J., Dejean(1998) Conjugation with the ubiquitin-related modifier SUMO-1 regulates the partitioning of PML within the nucleus. *EMBO J.* **17**, 61–70

[9]   Kamitani, T., Kito, K., Nguyen, H. P., Wada, H., Fukuda-kamitani, T., and Yeh, E. T. H. (1998) Identification of Three Major Sentrinization Sites in PML. *J. Biol. Chem.* **273**, 26675–26682

Appendix B

The SUMO and Ubiquitin binding drives the DNA damage independent puncta formation by RAP80.

Introduction

RAP80 has been extensively studied for its role in DNA repair, where it binds SUMO and ubiquitin modifications at the DNA double strand break sites through its N-terminal SUMO and Ubiquitin binding motifs[1-4]. Through its central region it interacts with Abraxas, a component of BRCA1 A complex and leads to the recruitment of crucial DNA repair proteins such as BRCA1, BRCC36, BARD1 and Abraxas to the sites of DNA damage[5-7]. In addition to its well-defined role in DNA repair, it has also been shown to interact with different nuclear receptors such as estrogen receptor α and retinoid related testis-associated receptor and modulate their transcriptional activities[2-8]. RAP80 (Receptor Associated Protein) is a 79 kD protein consisting of a SIM followed by tandem UIMs at its N-terminal(2,8). It also contains two zinc fingers at the C-terminus that interacts with the histone H2B of nucleosome[9]. The SIM of RAP80 binds SUMO-2 with a modest affinity of ~200 μM which is enhanced by ~20 fold upon phosphorylation of two serines adjacent to the hydrophobic core of SIM by CK2[10]. The UIMs of RAP80 specifically recognize Lys-63 linked ubiquitin chains; each UIM binds monoubiquitin with an affinity of ~500 μM. They work together in a multivalent fashion to enhance the affinity of RAP80 for Ub chain by ~20 fold[11-12]. The multivalent recognition of Ub chains by RAP80 and the affinity enhancement through increased electrostatic interaction between phosphorylated SIM of RAP80 and SUMO underlies the specificity of N-terminal RAP80 for Ub chains and SUMO modifications. The complete structural features of RAP80 still needs to be elucidated, however the N-terminal of RAP80 (33-131) contains a beta strand SIM joined by a flexible linker to the two tandem linked alpha helices held together by a flexible linker, giving it an overall characteristic of a disordered segment of a protein[10]. The crosstalk between SUMO and Ub modifiers often play a major role in different cellular processes such as protein degradation and DNA repair[13-15]. The ubiquitin E3 ligase, RNF4 is a polySIM containing protein that binds polySUMO chains and ubiquitinates them resulting in the synthesis of SUMO-Ub hybrid chains is required for the optimal recruitment of BRCA1 and RAP80 to the sites of DNA damage(8). One such hybrid chain, SUMO-K63(Ub)$_2$ has been implicated in DNA repair and has been proposed as a binding partner of N-terminal RAP80[3,4]. The

N-terminal region of RAP80 binds the hybrid chains SUMO-2-K63(Ub)$_2$ linked Ub$_2$ with an ~80 fold higher affinity as compared to its binding affinity for SUMO-2 or Ub alone, in vitro(8). The mechanism of hybrid chains binding with RAP80 N-terminal RAP80 has not been studied in atomic details, however the presence of a long ~32 residues flexible linker between the SIM and the tandem UIMs of RAP80 can facilitate a multivalent interaction between the two and along with the ~25 fold increased affinity of RAP80 for SUMO-2 upon phosphorylation by CK2 can be a plausible cause for the ~80 fold enhanced affinity of N-terminal RAP80 for SUMO-K63(Ub)$_2$ hybrid chains[8,16].

Multiple, low affinity interactions of proteins with their binding partners guide their phase separation in to protein-RNA or protein dense cellular compartments[16,17]. Intrinsically disordered regions of proteins containing multivalent binding motifs drive such interactions that are crucial for the phase separation of proteins. Post-translational modifications such as phosphorylation also play a crucial role in the phase separation of proteins. The valency and affinity of the binding sites dictate the separation or likeliness of proteins into these protein dene compartments[17-19]. Our fluorescence images of RAP80 show that in the absence of DNA damage (under normal conditions) it has a propensity to form nuclear puncta. Despite the striking resemblance of these puncta with repair foci, our colocalisation studies indicate that not all the RAP80 nuclear puncta are DNA repair foci. By abolishing the binding of RAP80 with ubiquitin and SUMO, the numbers of nuclear puncta are reduced, probably due to the diminished phase separation of RAP80 into protein dense compartments. Similar results were observed after mutating the two CK2 phosphorylation sites of RAP80, adjacent to its SIM. The functional relevance of RAP80 puncta formation was shown by its colocalization with PML bodies and histone, H2AK119Ub. By forming puncta and being sequestered in PML nuclear body RAP80 is able to respond to the DNA damage repair pathway quicker and is protected from the proteasomal degradation. The localization of RAP80 puncta with histone, H2AK119Ub suggests its role in transcriptional repression. In summary, our findings indicate that in addition to the well-characterized role of SIM and UIMs of RAP80 at the sites of DNA repair, both the motifs and associated post-translational modification of RAP80 SIM by CK2 are crucial for the formation of nuclear puncta, probably by phase separation which is required for the accomplishment of its roles in

DNA repair and transcriptional repression.

Experimental Procedures

Cell line, plasmids and transfection

Human U2OS cells were used for all the experiments, and were cultured in DMEM media supplemented with 10% FBS (Fetal Bovine Serum). The DNA sequence for plasmids encoding for wild type GFP-RAP80, GFP-RAP80 ΔSIM, GFP-RAP80 ΔUIM, GFP-RAP80 ΔSIM ΔUIM, GFP-RAP80 SIM-UIM and GFP-RAP80-S44AS46A are listed in Supplemental Data Table 1. The plasmid encoding for wild type GFP-RAP80 was obtained from Dr. Anton Jetten (National Institute of Environmental Health Sciences, NIH, USA) and the GFP-RAP80 ΔSIM mutant was created by deleting the residues 40-47 using site directed mutagenesis (GenScript). For immunofluorescence staining, U2OS cells were grown on coverslips and transfected with the 0.5-1 μg of plasmid according to the manufacturer's instructions on Effectene, a transfection kit from QIAGEN. For FRAP experiments U2OS cells were grown on MaTek dishes and transfected using the similar conditions to those for growth on coverslips.

Immunofluorescence staining

Cells grown on coverslips were exposed to an ionizing radiation (2Gy) using a Mark I 68A irradiator (JL Shepherd & Associates), with a ^{137}Cs source and incubated for 1 hour prior to IF staining, where indicated. To carry out the IF staining, cells were pre-extracted by incubating in RIPA buffer, or 0.5% Triton X-100 for 3-5 minutes, washed with 1X PBS (Phosphate Buffer Saline); cells were subsequently fixed by incubating in 4% paraformaldehyde for 10 minutes. Cells were then washed in 1X PBS and coverslips were inverted for incubation with primary antibody for 60 minutes at 25° C or overnight at 4° C, then washed in 0.1% TritonX-100 and 1X PBS. Cells were then incubated with specific secondary antibodies for 30 minutes at 25° C followed by washing with 0.1%

TritonX-100 twice and once with 1X PBS. Coverslips were then mounted using mounting media containing DAPI on the microscope slide.

Antibodies

The following antibodies were used for IF staining. Anti-RAP80 (Cell Signalling), anti-γ-H2AX (Active motif, Cell Signalling), anti-PML (Santa Cruz Biotechnology), anti-BLM(Bethyl Laboratories), anti-Ubiquityl Histone H2A (Upstate) and anti-RING-1B (MBL Life Science). Anti mouse and anti-rabbit Alexa fluor 488, Cy3 and Cy5 conjugated secondary antibodies from Invitrogen were used as applicable.

Fluorescence Microscopy

IF stained cells were observed and imaged using an Axiovert 200 M microscope (Carl Zeiss Inc.), equipped with Metamorph software (Molecular Devices), a cooled CCD camera and different fluorescence optics. Images were analyzed using the Image J software to study nuclear foci as well as co-localization of proteins. The brightness and contrast levels for few images were adjusted to reduce the background intensity and were converted to 8-bit for the preparation of figures using Adobe Illustrator CC.

FRAP Experiments

U2OS cells were cultured on MaTek dishes and transfected with plasmids for GFP-tagged protein according to the experiment performed. FRAP was performed by photobleaching a strip within the nucleus as described previously [20-21]. Images were collected until complete fluorescence recovery was achieved. The total recovery time for different proteins was determined based on the time taken for the slope of the recovery curve to flatten. Wild type GFP-RAP80 and the various mutants were imaged for 40 seconds. The fluorescence was normalized to the background and photobleaching during the imaging as mentioned in the references. The first image collected for all the different proteins before photobleaching was used to represent the live cell image from the GFP-tagged wild type and mutants of RAP80 and RYBP. A minimum of 15 cells were imaged

for each experiment, and each FRAP experiment was performed in triplicates in most cases, except where indicated. The mean recovery data were calculated from each cell to represent the FRAP curve for each protein.

Results

RAP80 forms nuclear puncta independent of DNA damage

The visual analysis of live cell images of ~50 U2OS cells transiently expressing GFP-WT-RAP80 displayed the expression of RAP80 in three patterns, under normal conditions: a population evenly distributed in the nucleoplasm, small foci like puncta expressed in >20 number in few cells and medium sized (0.1-1 μM) puncta ranging from 1-15 in numbers in most of the cells, as shown in Figure B.1a. To rule out this observation as a case of protein overexpression, we IF stained U2OS cells for endogenous RAP80 and a similar pattern of expression was observed which is shown in Figure B.1b. We focused our study to medium sized puncta. The foci like puncta may be the sites of DNA repair caused due to incipient DNA damage or replication associated repair(18,19). In order to investigate that, we IF stained U2OS cells expressing GFP-WT-RAP80 with antibody against endogenous γ-H$_2$AX, a reliable marker of DNA repair(24). As shown in Figure B.1c, a subset of the mid-sized RAP80 puncta colocalized with γ-H2AX and a population of these puncta did not co-localize with γ-H2AX. A visual analysis of the mid-sized puncta in ~50 GFP-WT-RAP80 expressing U20S cells fixed and stained with γ-H$_2$AX showed that nearly half of the mid sized RAP80 nuclear puncta co-localized with γ-H$_2$AX whereas the rest half of them did not show any colocalization. The abovementioned results indicate that RAP80 forms mid sized nuclear puncta independent of DNA damage. In addition to the co-localization studies we also performed photobleaching experiments and monitored the recovery curve for GFP-WT-RAP80 nuclear puncta under normal conditions and after exposing the cells to ionizing radiation of 2 Gy and incubating for an hour before preforming the experiment. The data shown here is from one experiment performed with an approximate of 10 nuclear puncta was photobleached and analyzed. The arithmetic mean of the normalized FRAP curve of

the cells under normal conditions was calculated and compared to the average of normalized FRAP curve intensity of similar number of puncta in cells exposed to IR. As shown in Figure B.1d, the half recovery of fluorescence was observed within first 2-3 seconds suggesting the interior of the puncta to be dynamic. The nuclear puncta under normal conditions also exhibited a fraction of immobile or stable population in comparison with RAP80 puncta in the repair foci. Upon visual analysis of the images captured during FRAP experiment, one third of the WT-RAP80 nuclear puncta showed delayed recovery, and the complete recovery was not achieved on the FRAP timescale (40 seconds), indicating that a certain population of the GFP-WT-RAP80 puncta forms a more stable structure as compared to repair foci. We also performed a similar comparison of FRAP recovery curves for the nucleoplasm of U2OS cells expressing GFP-WT-RAP80 under normal conditions with the nucleoplasm of U2OS cells expressing GFP-WT-RAP80 and exposed to ionizing radiation. As indicated in Figure B.1e, the nucleoplasmic RAP80 recovery curve under normal conditions is similar to when cells were exposed to ionizing radiation. The mobile and immobile population remains same under both the conditions along with the rate of recovery. The normalized recovery intensity was of both puncta and nucleoplasm was fit to the equation mentioned below:

$$F(T) = C1(1-e^{(-k1*T)}) + C2(1-e^{(-k2*T)})$$

 where F is the normalized fluorescence intensity, C1 and C2 are the constants and k1 and k2 are the rates for two rates of reaction and t is the time interval[25]. The above mentioned equation was used to detect the presence of any binding interaction in addition to diffusion. A shown in Figure B.2a and B.2b and Table B.1, in both puncta and nucleoplasm there is a binding interaction of RAP80 in addition to diffusion, reflected in the k1 and k2 values, where k2 is approximately 4-6 times slower than k1, indicating a binding interaction of RAP80 in both its puncta and nucleoplasm. A similar pattern is observed for the GFP-RAP80-WT puncta and nucleoplasm population under the conditions of DNA repair, shown in Figure B.2g and B.2h. We show RAP80 forms medium sized (0.1-1 μM) nuclear puncta independent of DNA damage; under normal

conditions a subset of these puncta forms stable complex. We also show that RAP80 in both the nucleoplasm and puncta is involved in binding interaction.
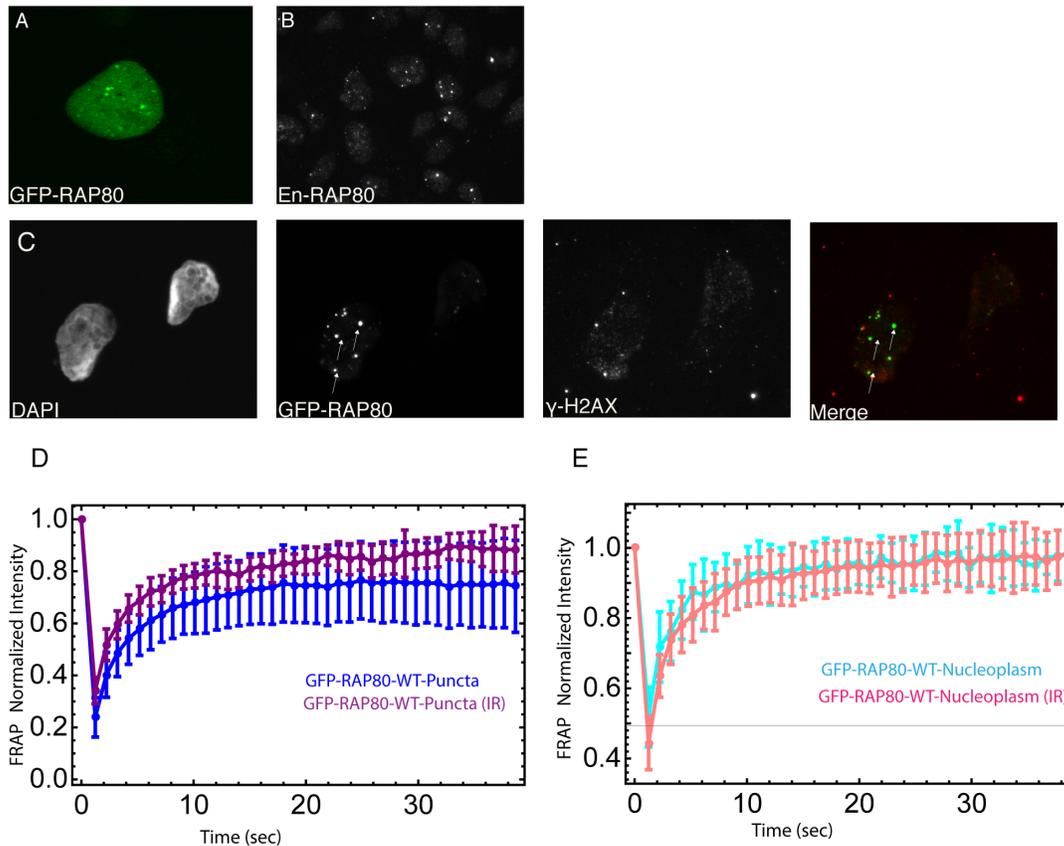

Figure B.1: (a), U2OS cells expressing wild type-GFP-RAP80. **(b),** U2OS cells expressing endogenous RAP80. **(c),** U2OS cells expressing wild type GFP-RAP80, IF stained with γ-H2AX, without exposing the cells to radiation. Cells were analyzed for co-localization of GFP-RAP80 foci with γ-H2AX foci. The GFP-RAP80 foci marked with arrow do not co-localize with γ-H2AX foci. Also shown are the images for nuclear staining by DAPI and the merged image for RAP80 stained in green and γ-H2AX in red. **(d),** FRAP curves for the puncta of wild type GFP-RAP80 (Blue) without any exposure to radiation and after exposure to radiation (dark pink). The fluorescence recovery curves for puncta under both the conditions (normal and radiation exposed) are displayed together. **(e),** FRAP curves for the nucleoplasm of wild type GFP-RAP80 (cyan) without any exposure to radiation and after exposure to radiation (light pink). The fluorescence recovery curves for puncta under both the conditions (normal and radiation exposed) are displayed together.
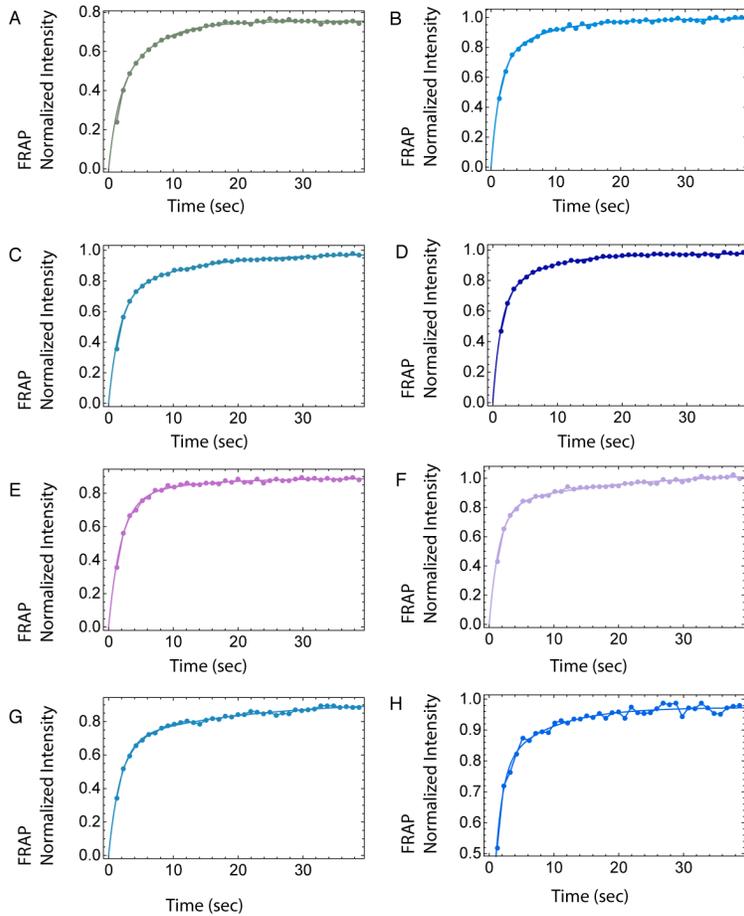
Table 1

FRAP Region	K1	K2
WT-RAP80 Puncta	0.773 ± 0.14	0.183 ± 0.01
WT-RAP80 Nucleoplasm	0.754 ± 0.07	0.135 ± 0.01
RAP80-ΔSIM Puncta	0.585 ± 0.02	0.089 ± 0.007
RAP80-ΔSIM Nucleoplasm	0.801 ± 0.05	0.148 ± 0.01
RAP80-ΔUIM Puncta	0.529 ± 0.03	0.105 ± 0.03
RAP80-ΔUIM Nucleoplasm	0.647 ± 0.03	0.051 ± 0.01
WT-RAP80 Puncta (IR)	0.545 ± 0.03	0.049 ± 0.01
WT-RAP80 Nucleoplasm (IR)	0.899 ± 0.10	0.129 ± 0.02

Figure B.2: The normalized FRAP intensity values were fit to a two rate equation with the fits for GFP-WT-RAP80 puncta shown in **(a),** GFP-WT-RAP80 nucleoplasm region shown in **(b),** GFP-ΔSIM-RAP80 puncta shown in **(c)**, GFP-ΔSIM-RAP80 nucleoplasm region shown in **(d),** GFP-ΔUIM-RAP80 puncta shown in **(e),** GFP-ΔUIM-RAP80 nucleoplasm region shown in **(f),** The puncta formed by U2OS cells expressing GFP-WT-RAP80 and exposed to a radiation of 2 Gy in **(g),** The nucleoplasm of U2OS cells expressing GFP-WT-RAP80 and exposed to a radiation of 2 Gy in **(h)**.

Table B.1: The values of the rates, k1 and k2 obtained after fitting the normalized FRAP intensity to the two rate equation for all the mutants in puncta and nucleoplasm is shown here. The errors associated with both the values are also shown

The SUMO interacting motif of RAP80 is crucial for its puncta formation

To investigate further in to the RAP80 domains responsible for the DNA damage independent puncta formation, we deleted the SUMO interacting motif of the RAP80 and examined the expression pattern and mobility of the GFP-RAP80-ΔSIM using fluorescence microscopy and photobleaching experiments. The mid sized nuclear puncta were counted for ~50 U2OS cells expressing GFP-RAP80-ΔSIM and compared with U2OS cells expressing GFP-WT-RAP80. As shown in Figure B.3a, B.3b, B.3b and Table B.2, the number of puncta is almost reduced by half upon the deletion of SIM motif of RAP80. Whilst the ability to form puncta was retained, a significant reduction in the number of puncta suggest that the SIM of RAP80 contributes partially to the formation and maintenance of the structural integrity of these mid-sized nuclear puncta. The reduction in the number of these mid-sized puncta may be partially due to the impairment in the ability of the RAP80 ΔSIM to form repair foci, however not entirely due to the residual ability of RAP80 ΔSIM protein to form repair foci by recognizing the ubiquitin modifications at the DNA damage sites through its intact tandem Ub interacting motifs. We performed the photobleaching experiment of U2OS cells expressing GFP-RAP80-ΔSIM mutant to study the molecular dynamics of its mid-sized puncta in comparison to WT-RAP80. The mean values of normalised fluorescence intensity of ~10 puncta from three independent experiments were plotted against time as shown in Figure B.4a, and compared with GFP-WT-RAP80 puncta. The FRAP curve of GFP-RAP80 ΔSIM mutant displays a dynamic interior like WT-RAP80 nuclear puncta. In addition a decrease in the immobile or stable population of puncta was also observed confirming that SIM of RAP80 has a role in the immobile or stable fraction of RAP80 puncta. The dynamic characteristics of the nucleoplasmic population of GFP-RAP80 ΔSIM and GFP-WT - RAP80 did not show any difference in the mobile and immobile population, as well as the recovery curve, as shown in the Figure B.4d. The FRAP recovery curve of both the puncta and nucleoplasm of GFP-RAP80-ΔSIM were fit to the biexponential equation, shown in Figure B.2c and B.2d, respectively. The estimated values of k1 and k2 for the puncta and nucleoplasm is shown in Table B.2. Like, GFP-WT-RAP80, GFP-RAP80 ΔSIM population in both puncta and nucleoplasm displays a binding interaction.
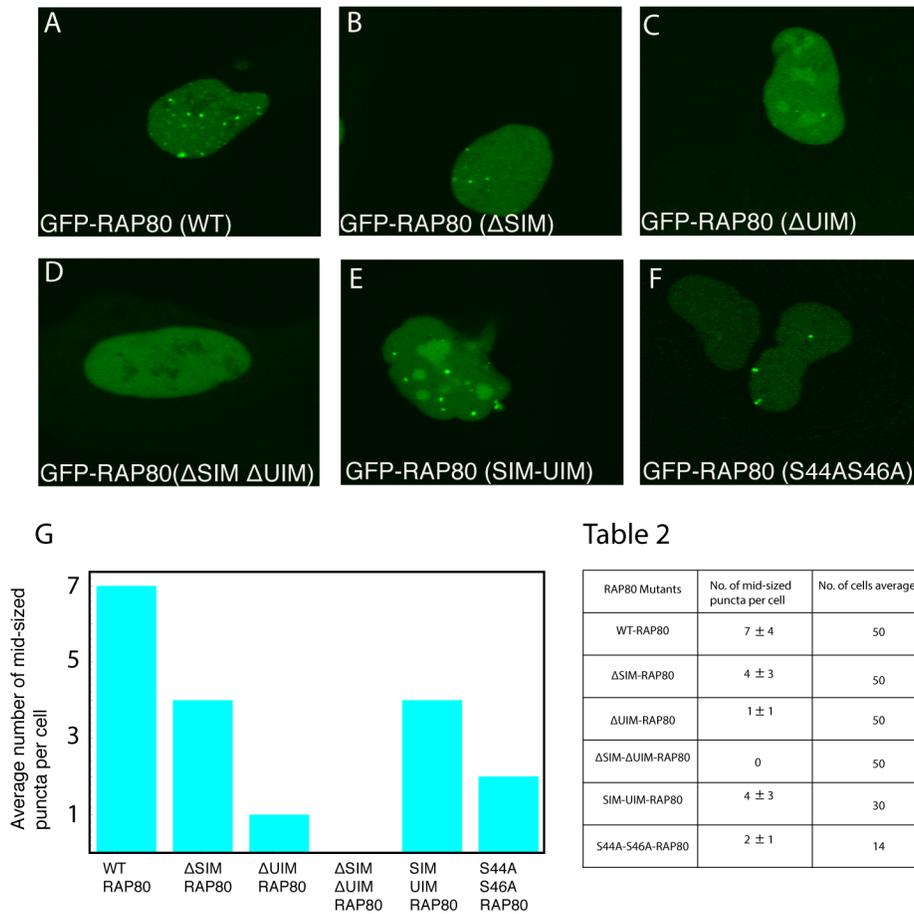
A

GFP-RAP80 (WT)

B

GFP-RAP80 (ΔSIM)

C

GFP-RAP80 (ΔUIM)

D

GFP-RAP80(ΔSIM ΔUIM)

E

GFP-RAP80 (SIM-UIM)

F

GFP-RAP80 (S44AS46A)

G

Average number of mid-sized puncta per cell (y-axis: 1, 3, 5, 7)

Bar graph categories: WT RAP80, ΔSIM RAP80, ΔUIM RAP80, ΔSIM ΔUIM RAP80, SIM UIM RAP80, S44A S46A RAP80

Table 2

RAP80 Mutants	No. of mid-sized puncta per cell	No. of cells averaged
WT-RAP80	7 ± 4	50
ΔSIM-RAP80	4 ± 3	50
ΔUIM-RAP80	1 ± 1	50
ΔSIM-ΔUIM-RAP80	0	50
SIM-UIM-RAP80	4 ± 3	30
S44A-S46A-RAP80	2 ± 1	14

Figure B.3: (a), U2OS cells were transfected with plasmid encoding wild type-GFP-RAP80, grown and imaged live using 2-photon microscope **(b),** U2OS cells were transfected with plasmid encoding GFP-tagged SIM deletion mutant of RAP80- GFP-RAP80ΔSIM, grown and imaged live using 2-photon microscope **(c),** U2OS cells were transfected with plasmid encoding GFP-tagged UIM deletion mutant of RAP80- GFP-RAP80ΔUIM, grown and imaged live using 2-photon microscope **(d),** U2OS cells were transfected with plasmid encoding GFP-tagged SIM and UIM deletion mutant of RAP80- GFP-RAP80ΔSIM ΔUIM, grown and imaged live using 2-photon microscope **(e),** U2OS cells were transfected with plasmid encoding only the N-terminal of RAP80 consisting of SIM and UIM motifs, GFP-RAP80-SIMUIM, grown and imaged live using 2-photon microscope **(f),** U2OS cells were transfected with plasmid encoding GFP-tagged CK2 phosphorylation mutant of RAP80- GFP-RAP80S44AS46A, grown and imaged live using 2-photon microscope **(g),** The average number of mid-sized puncta of RAP80 expressed by different constructs of RAP80 is shown as a bar graph. **Table B.2:** The average number of mid sized puncta formed by different mutants of GFP-tagged-RAP80 with the number of cells averaged are indicated.

The ubiquitin interacting motifs of RAP80 is required for its puncta formation

 The tandem linked ubiquitin interacting motifs of RAP80 is responsible for its multivalent binding to Ub chains and increasing the affinity of the RAP80 towards the ubiquitin moieties. To study the effect and contribution of UIMs in the formation and maintenance of mid-sized RAP80 nuclear puncta, in the absence of DNA damage we transfected U2OS cells with GFP-RAP80 ΔUIM construct and visually analysed the live images of ~ 50 cells for the appearance of nuclear puncta. As shown in Figure B.3c, 3g and Table B.2, there was a drastic reduction in the number of mid sized puncta formed by the U2OS cells expressing GFP-RAP80 ΔUIM construct. More than 60% of the cells completely lost the ability to form mid-sized nuclear puncta. Similar to GFP-RAP80 ΔSIM expressing cells, a subset of the nuclear puncta formed in the cells expressing GFP-RAP80 ΔUIM may be the DNA repair foci with RAP80 recruited to the damage by the retained ability of SIM to recognize SUMO modifications, as demonstrated before. Upon analysis of the FRAP recovery curves of GFP-RAP80 ΔUIM cells and its comparison to GFP-RAP80-WT, we observed that the deletion of UIMs of the RAP80 resulted in a loss of the stable/immobile population of the protein, shown in Figure 4B. Like GFP-RAP80 ΔSIM and GFP-RAP80-WT, the interior of the puncta was highly dynamic due to the fast recovery of the fluorescence. Due to a significant loss in the ability of GFP-RAP80 ΔUIM mutant to form puncta (as indicated in Table B.2), the FRAP curves of only 15 puncta were analyzed from three independent experiments. The photobleaching curve obtained from the nucleoplasm region displayed similar characteristics to GFP-RAP80 ΔSIM mutant and GFP-RAP80-WT. The rates of diffusion and binding obtained by fitting the FRAP curve to a two rate equation are shown in Figure B.2e and B.2f, suggesting diffusion and binding interaction of the puncta and nucleoplasmic population of RAP80.

The SIM and UIM of RAP80 collectively contribute towards its puncta formation

To explore the combined contribution of the SIM and UIM of RAP80 towards its mid-sized puncta formation, we transfected U2OS cells with GFP-RAP80-ΔSIMΔUIM

mutants and examined the live cell images for its puncta formation. Out of ~50 cells examined, only 1 cell showed a single puncta. The ability to form the puncta was completely abolished by the deletion of the SIM and UIM motifs of RAP80, shown in Figure B.3d, B.3g and Table B.2. Surprisingly, the ability to form puncta was recovered by the U2OS cells expressing GFP-RAP80-SIM-UIM construct, shown in Figure B.3e, 3g and Table B.2. The above results indicate that the N-terminal SIM and UIM of RAP80 possess the ability to form nuclear puncta, independent of DNA damage.

The RAP80 nuclear puncta formation is dependent on its post-translational modification by Casein Kinase 2

The SIM of RAP80 has been demonstrated to enhance the affinity and specificity for its binding partner upon phosphorylation by CK2. To study the cellular expression of CK2 site mutant of RAP80 (S44AS46A), U2OS cells were transfected with GFP-RAP80-S44AS46A and the live cell images were examined for puncta formation. Out of 15 cells examined, there was a significant reduction in the number of puncta formed after the CK2 phosphorylation site mutation, as shown in Figure B.3f, B.3g and Table B.2. The molecular dynamics of this mutant was studied by photobleaching 8 puncta formed by the GFP-RAP80-S44AS46A mutant expressing cells. As shown in Figure 4° C, the recovery curve of both puncta and nucleoplasm were similar to WT-RAP80. However, there was no loss or reduction in the stable or immobile puncta population of RAP80, as seen with GFP-RAP80 UIM and SIM deletion mutants.
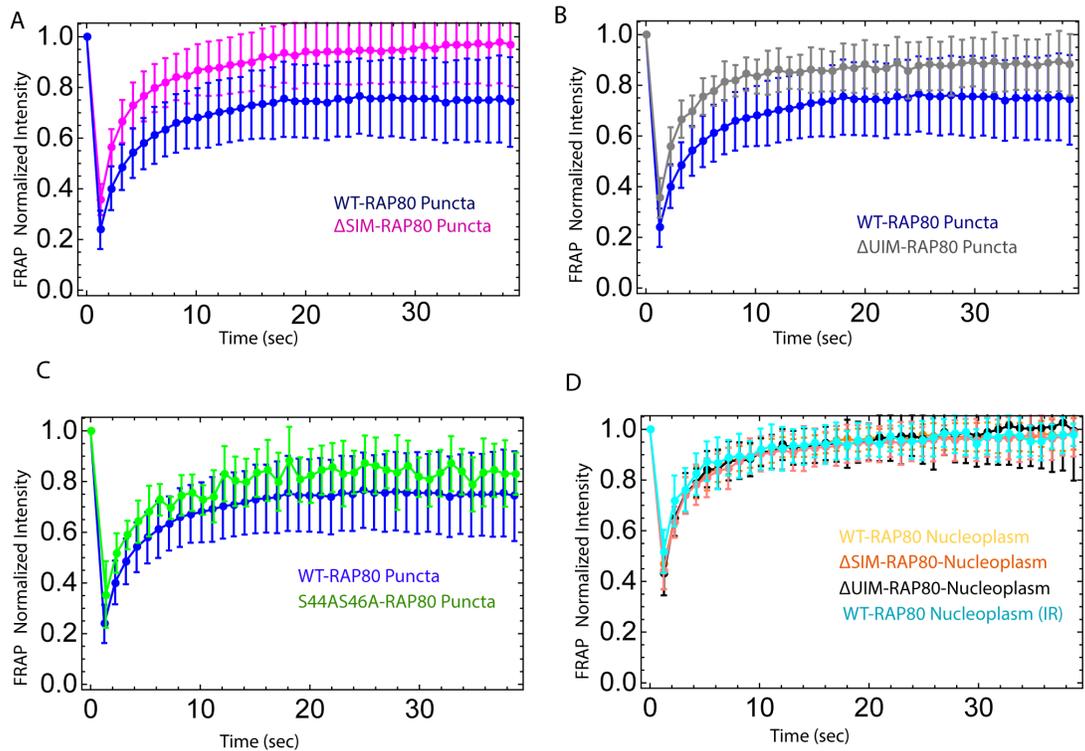
Figure B.4: (a), The normalized FRAP intensity of the puncta formed by U2OS cells expressing GFP-RAP80-WT (blue) and GFP-RAP80-ΔSIM (pink) are shown **(b),** The normalized FRAP intensity of the puncta formed by U2OS cells expressing GFP-RAP80-WT (blue) and GFP-RAP80-ΔUIM (grey) are shown **(c),** The normalized FRAP intensity of the puncta formed by U2OS cells expressing GFP-RAP80-WT (blue) and GFP-RAP80-S44AS46A (green) are shown **(d),** The normalized FRAP intensity of the nucleoplasm of U2OS cells expressing GFP-RAP80-WT (yellow), GFP-RAP80-ΔSIM (orange), GFP-RAP80-ΔUIM (black), GFP-RAP80-WT and exposed to IR (cyan) are shown.

The RAP80 nuclear puncta colocalizes with PML protein

To explore the functional significance and identity of RAP80 nuclear puncta, co-localization of RAP80 with proteins belonging to different well-studied nuclear bodies was screened, due to the striking resemblance of RAP80 puncta with the nuclear bodies. According to previous published work by Tikoo *et. al*[26] and Kim *et. al*[27], RAP80 is believed to reside in PML nuclear bodies through its interaction with BLM and TRAIP. RAP80 co-localizes with BLM and BLM localized to PML body, in the absence of damage suggesting the localization of RAP80 to PML body. In Kim et. al[27], TRAIP

was shown to interact with RAP80 and it localized to PML nuclear body. Upon overexpression of TRAIP, RAP80 was shown to form nuclear body and its co-localization with TRAIP implied its localization to PML body. This intrigued us to screen for co-localization of endogenous RAP80 with endogenous PML body in the absence of DNA repair and our results do not indicate a strong and direct co-localization of RAP80 with PML protein in U2OS cells, however we only found few RAP80 foci that localized with PML in U2OS cells, pointed with arrows in the Figure B.5a. We also screened for GFP-RAP80 foci localization with endogenous BLM helicase and endogenous PML body and as showed in Figure B.5b, almost all the BLM helicase foci display co-localization with PML but only a fraction of RAP80 foci localized with BLM and PML body, marked with arrows.
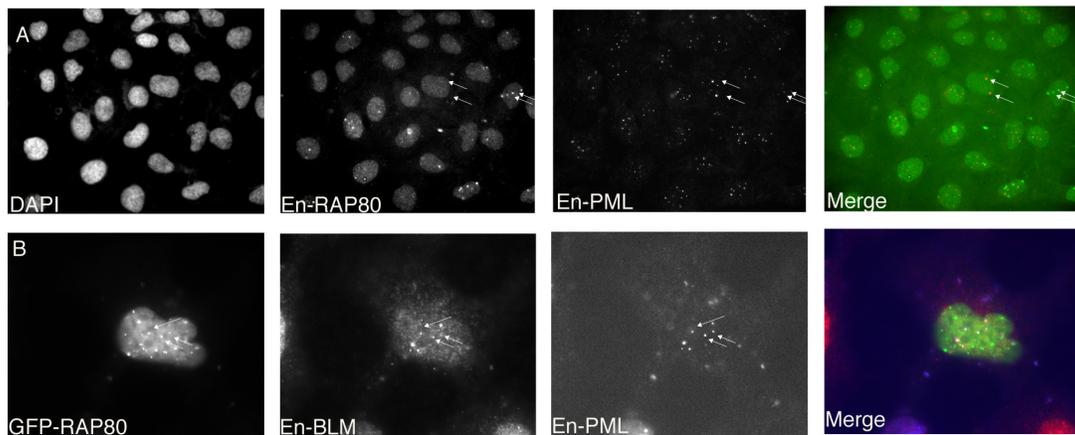

Figure B.5: (a), U2OS cells were grown, fixed, stained using fluorescent-tagged antibodies against endogenous RAP80 and endogenous PML protein. Cells were imaged using fluorescence microscopy and RAP80 puncta were analyzed for co-localization with PML bodies. The RAP80 puncta colocalizing with PML protein are marked with arrows. **(b),** GFP-RAP80 expressing U2OS cells were immune-fluorescence stained for endogenous BLM helicase protein and endogenous PML body and co-localized BLM protein and PML body are marked with arrow.

RAP80 nuclear puncta is associated with transcriptional repression marker of chromatin, H2AK119Ub

The RAP80 protein was initially identified as a transcriptional modulator of estrogen and RTR mediated signaling[6,27]. To delve deeper in identifying the functional significance of these RAP80 nuclear puncta we performed its co-localization with different markers of PcG body such as BMI-1, RING-1B and H2AK119Ub[29]. As shown in Figure B.6a and B.6b we did not observe any co-localization of RAP80 with BMI-1 or RING-1B. We further investigated the co-localization of RAP80 foci with a well-known modification of genetic repression by PcG body- H2AK119Ub. Transcriptional repression is coupled with the DNA repair process and due to the well-established role of RAP80 in DNA repair[30]; we expected a co-localization of RAP80 puncta with H2AK119Ub. In order to segregate the puncta formed due to spontaneous DNA damage, we performed the co-localization of RAP80 puncta with H2AK119Ub and γ-H2AX. As shown in Figure B.6c, some of the RAP80 foci (not marked) localized with γ-H2AX and the RAP80 foci marked with arrow did not co-localize with γ-H2AX but co-localized with H2AK119Ub marker suggesting RAP80 is associated with transcriptional repression, independent of DNA damage. This co-localization experiment was also performed after treating U2OS cells with PRC inhibitor to validate the specificity of the antibody in-house. As shown in Figure B.6d-e, the GFP-RAP80 expressing U2OS cells treated with DMSO and PRC inhibitor, stained for H2AK119Ub and γ-H2AX. In the cells treated with inhibitor, the polycomb repressive complexes did not maintain their foci and the staining was diffused throughout the nucleus as seen by the anti-H2AK119Ub staining and did not co-localize with RAP80 foci (Figure B.6d-e), thereby confirming the specificity of the antibody H2AK119Ub.
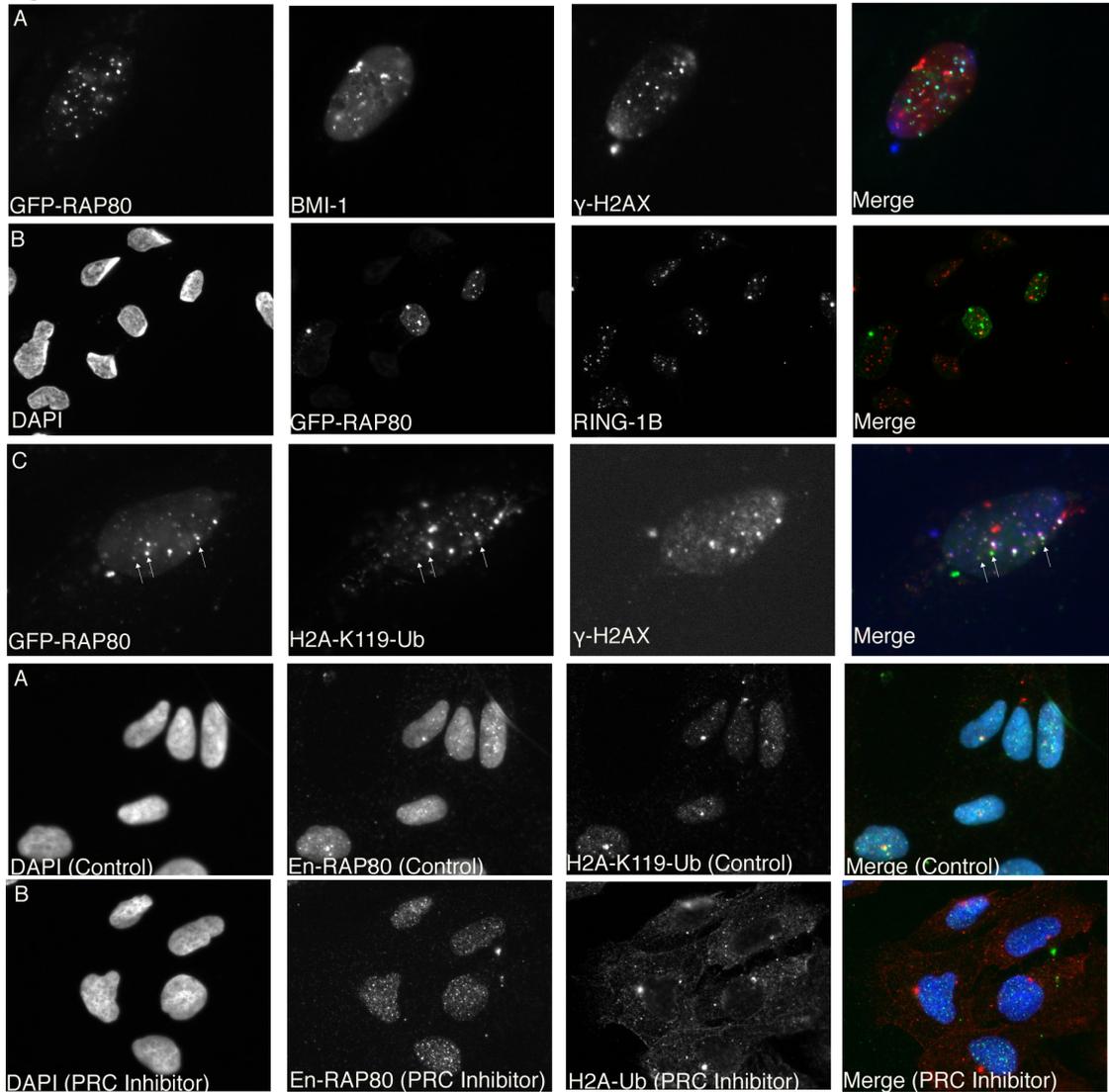
Figure B.6: (a), U2OS cells expressing wild type GFP-RAP80 were IF stained with endogenous BMI1 and endogenous γ-H2AX, without any exposure to radiation **(b),** U2OS cells expressing GFP-RAP80 were IF stained with RING-1B without exposing cells to radiation **(c),** U2OS cells expressing wild type GFP-RAP80 were IF stained with PRC marker H2AK119Ub and endogenous γ-H2AX, without any exposure to radiation.. **(d),** U2OS cells incubated in DMSO for 2 hours and were grown, fixed, immunofluorescence stained for endogenous RAP80 and H2AK119Ub modification to act as a control for PRC inhibitor experiment. Cells were imaged by fluorescence microscopy and analyzed for co-localization of the above-mentioned proteins **(e),** U2OS cells incubated in of PRC inhibitor, PRC 4165 (40 μM) for 2 hours and were grown, fixed, immunofluorescence stained for endogenous RAP80 and H2AK119Ub modification.

Discussion

The immunofluorescence staining of γ-H$_2$AX by specific antibody and the appearance of DSB sites as a foci, resolvable by fluorescence microscopy has revolutionized the identification of proteins involved in the DNA repair process[23,30]. We adopted a similar strategy to separate the puncta formed by RAP80 into DNA damage dependent foci and DNA damage independent puncta. The comparison of FRAP recovery curves of both DNA damage dependent and independent puncta revealed that in the absence of DNA damage, a fraction of RAP80 puncta formed are more stable or immobile over FRAP timescale. This ability of RAP80 to form stable puncta is diminished upon deletion of its SUMO and Ubiquitin interacting motifs whereas it remains unchanged upon mutation of CK2 phosphorylation sites of RAP80 SIM. Overall, our findings suggest that the stability of a fraction of DNA damage independent RAP80 puncta is dependent on its SUMO and Ubiquitin interacting motifs. The significant decrease in the number of puncta by impairing the RAP80 binding to Ubiquitin or SUMO binding either by deleting the SIM and UIM or by mutating the CK2 phosphorylation site of RAP80 SIM strongly suggests that the integrity of DNA damage independent foci is dependent on its Ub and SUMO interacting motifs. A complete absence of the RAP80 puncta upon simultaneous deletion of both the SIM and UIM of RAP80 and their recovery of to a sub-optimal level upon expression of only the N-terminal SIM and UIM motifs of RAP80 confirms that the DNA damage independent RAP80 puncta are highly dependent on its ability to recognize Ub and SUMO motifs. The FRAP curves of the different RAP80 mutants suggest that the interior of the puncta are highly dynamic and in addition to diffusion the recovery of the fluorescence in the photobleached region is dependent on the RAP80 unbinding event with a nuclear component, in both the puncta and the nucleoplasm. In the absence of DNA damage, RAP80 has been reported to bind nuclear receptors, ubiquitinated histone H2A through its UIM domain and H2B through its C-terminal zinc fingers[2,8,9] ; one or all of these might be responsible for its FRAP recovery curve in the nucleoplasm. In order to determine the true values of the binding parameters and compare it upon the deletion of RAP80 motifs, we need to fit the FRAP curve to advanced models[25].

The phosphorylation of two serine residues adjacent to RAP80 SIM hydrophobic sequence is crucial for its binding to SUMO and the mutation S44F is linked to cancerous mutation[32]. The cellular expression of the GFP-RAP80S4AS46A mutant show reduced puncta as compared to GFP-RAP80-WT but similar to GFP-RAP80-ΔSIM mutant suggesting a role of these CK2 phosphorylation sites under normal conditions, through the enhancement of its SUMO binding.

The dependence of RAP80 puncta formation on its closely placed SIM and UIMs at its N-terminal disordered region indicate that the formation of these puncta may be guided by the phase separation of RAP80 in protein dense or protein-RNA dense compartments, which has been demonstrated to be driven by multivalent interactions, in addition to post-translational modifications[16-17]. Although we demonstrate the regulation of puncta formation by SIM and UIM of RAP80 in a cellular environment, biophysical experiments examining the puncta formation and their dependence of multivalent binding of RAP80 and post-translational modification will help us gain deeper insights in to the mechanism of the puncta formation. Regarding the functional significance, the colocalization experiment of RAP80 with different nuclear body marker such as PML protein and PcG body marker, H2AK119Ub suggest that through puncta formation and localization to these structures RAP80 is able to respond quicker to the DNA repair process by being sequestered in PML body and protected from proteasomal degradation[31-32]. PML nuclear bodies are closely associated with the DNA repair process and stress response and they have been shown to change their number, size and morphology upon DNA damage induction and exposure to different forms of stress[33-35]. They respond to such conditions probably by sequestering and protecting crucial proteins from proteasomal degradation and providing a quick supply of these proteins, when needed[36]. Many DNA repair proteins such as BLM, p53 and ATM have been reported to reside in the PML nuclear body[22,33-36]. PML protein contains a SIM and is SUMOylated at three sites[39]. It recruits and sequester proteins such as Daxx and Sp100 within the PML body through SUMO SIM interactions. The hybrid chains have also been implicated in degradation of PML body and the dependence of both SIM and UIM of RAP80 for its DNA damage independent puncta formation hence localization with PML body raises the possibility that RAP80 may be involved in its interaction with PML body through the

hybrid chains[14], however studies needs to be carried out to investigate this. The association of RAP80 with chromatin through its co-localization with histone H2AK119Ub may be required to fulfill its role as a transcriptional modulator[40]. Nucleosome, the basic unit of chromatin is composed of two copies of each histone, H2A, H2B, H3 and H4, with each histone consisting of a flexible tail that is modified chemically and acts as sites for recruitment and retention of different factors required for the progression of replication, transcription, silencing or repair of broken DNA strands. The multivalent recognition of different kinds of modifications on nucleosomes by a single effector has been shown to regulate the chromatin related processes[41]. The association of RAP80 with chromatin through ubiquitinated histone H2A and H2B, in the absence of DNA damage and its SIM and UIM dependent partitioning in the puncta suggest that RAP80 may be involved in recognizing different modifications at the chromatin, simultaneously through its SIM and UIM, leading to the convergence of different signals at the chromatin[9].

In summary, we show that the N-terminal disordered region of RAP80 containing Ub and SUMO binding motif is possess the ability to form nuclear puncta. We also show that the interior of these puncta is highly dynamic and RAP80 population in both puncta and nucleoplasm is involved in binding interaction. A fraction of these puncta is more stable and is dependent on the Ub and SUMO interacting motif of RAP80. The principle behind the formation of puncta may be caused by the phase separation due to the presence of low affinity, multivalent binding components in the N-terminal of RAP80. In vitro studies will throw light on the mechanism of the formation and assembly of these puncta help us gain more insight into their biological relevance and role in DNA repair.

References

[1]   Kim, H., Chen, J., and Yu, X. (2007) Ubiquitin-binding protein RAP80 mediates BRCA1-dependent DNA damage response. *Science*. **316**, 1202–1205

[2]   Hu, X., Paul, A., and Wangs, B. (2012) Rap80 protein recruitment to DNA double-strand breaks requires binding to both small ubiquitin-like modifier (SUMO) and ubiquitin conjugates. *J. Biol. Chem.* **287**, 25510–25519

[3]   Savage, K. I., and Harkin, D. P. (2015) BRCA1, a "complex" protein involved in the maintenance of genomic stability. *FEBS J.* **282**, 630–646

[4]   Liu, Z., Wu, J., and Yu, X. (2007) CCDC98 targets BRCA1 to DNA damage sites. *Nat. Struct. Mol. Biol.* **14**, 716–720

[5]   Wang, B., Matsuoka, S., Ballif, B. A., Zhang, D., Smogorzewska, A., Gygi, S. P., and Elledge, S. J. (2007) Abraxas and RAP80 form a BRCA1 protein complex required for the DNA damage response. *Science*. **316**, 1194–1198

[6]   Yan, J., Kim, Y. S., Yang, X. P., Albers, M., Koegl, M., and Jetten, A. M. (2007) Ubiquitin-interaction motifs of RAP80 are critical in its regulation of estrogen receptor α. *Nucleic Acids Res.* **35**, 1673–1686

[7]   Yan, Z., and Jetten, A. M. (2000) Characterization of the repressor function of the nuclear orphan receptor retinoid receptor-related testis-associated receptor/germ cell nuclear factor. *J. Biol. Chem.* **275**, 35077–35085

[8]   Guzzo, C. M., Berndsen, C. E., Zhu, J., Gupta, V., Datta, A., Greenberg, R. A., Wolberger, C., and Matunis, M. J. (2012) RNF4-dependent hybrid SUMO-ubiquitin chains are signals for RAP80 and thereby mediate the recruitment of BRCA1 to sites of DNA damage. *Sci Signal*. **5**, ra88

[9]   Wu, J., Huen, M. S. Y., Lu, L.-Y., Ye, L., Dou, Y., Ljungman, M., Chen, J., and Yu, X. (2009) Histone ubiquitination associates with BRCA1-dependent DNA damage response. *Mol. Cell. Biol.* **29**, 849–60

[10]   Anamika, and Spyracopoulos, L. (2016) Molecular basis for phosphorylation-dependent SUMO recognition by the DNA repair protein RAP80. *J. Biol. Chem.*

291, 4417–4428

[11]   Sims, J. J., and Cohen, R. E. (2009) Linkage-Specific Avidity Defines the Lysine 63-Linked Polyubiquitin-Binding Preference of Rap80. *Mol. Cell*. **33**, 775–783

[12]   Markin, C. J., Xiao, W., and Spyracopoulos, L. (2010) Mechanism for recognition of polyubiquitin chains: Balancing affinity through interplay between multivalent binding and dynamics. *J. Am. Chem. Soc.* **132**, 11247–11258

[13]   Schwertman, P., Bekker-Jensen, S., and Mailand, N. (2016) Regulation of DNA double-strand break repair by ubiquitin and ubiquitin-like modifiers. *Nat Rev Mol Cell Biol*. **17**, 379–394

[14]   Tatham, M. H., Geoffroy, M.-C., Shen, L., Plechanovova, A., Hattersley, N., Jaffray, E. G., Palvimo, J. J., and Hay, R. T. (2008) RNF4 is a poly-SUMO-specific E3 ubiquitin ligase required for arsenic-induced PML degradation. *Nat. Cell Biol*. **10**, 538–546

[15]   Galanty, Y., Belotserkovskaya, R., Coates, J., and Jackson, S. P. (2012) RNF4, a SUMO-rageted ubiquitin E3 ligase , promotes DNA double-strand break repair. *Genes Dev*. **26**, 1179–1195

[16]   Lee, B. L., Singh A., Glover, J. N. M.,Hendzel, M. J. and Spyracopoulos, Leo. (2017). *J. Mol. Biol*. 10.1016/j.jmb.2017.05.029

[17]   Li, P., Banjade, S., Cheng, H.-C., Kim, S., Chen, B., Guo, L., Llaguno, M., Hollingsworth, J. V, King, D. S., Banani, S. F., Russo, P. S., Jiang, Q.-X., Nixon, B. T., and Rosen, M. K. (2012) Phase transitions in the assembly of multivalent signalling proteins. *Nature*. **483**, 336–340

[18]   Banjade, S., and Rosen, M. K. (2014) Phase transitions of multivalent proteins can promote clustering of membrane receptors. *Elife*. **3**, 1–24

[19]   Bah, A., and Forman-Kay, J. D. (2016) Modulation of intrinsically disordered protein function by post-translational modifications. *J. Biol. Chem*. **291**, 6696–6705

[20]   Lever, M. A., Th, J. P. H., Sun, X., and Hendzel, M. J. (2000) Rapid exchange of histone H1 . 1 on chromatin in living human cells. *Nature*, **8916**, 873–876

[21]   McDonald, D., Carrero, G., Andrin, C., De Vries, G., and Hendzel, M. J. (2006)

Nucleoplasmic beta-actin exists in a dynamic equilibrium between low-mobility polymeric species and rapidly diffusing populations. *J. Cell Biol*. **172**, 541–552

[22] Ismail, I. H., and Hendzel, M. J. (2008) The γ-H2A.X: is it just a surrogate marker of double-strand breaks or much more? *Environ. Mol. Mutagen*., **49**, 73-82

[23] Jackson, S., and Bartek, J. (2009) The DNA-damage response in human biology and disease. *Nature*. **461**, 1071–1078

[24] Kuo, L. J., and Yang, L.-X. (2008) γ-H2AX - a novel biomarker for DNA double-strand breaks. *In Vivo*. **22**, 305–9

[25] Sadegh Zadeh, K., Montas, H. J., and Shirmohammadi, A. (2006) Identification of biomolecule mass transport and binding rate parameters in living cells by inverse modeling. *Theor. Biol. Med. Model*. **3**, 36

[26] Tikoo, S., Madhavan, V., Hussain, M., Miller, E. S., Arora, P., Zlatanou, A., Modi, P., Townsend, K., Stewart, G. S., and Sengupta, S. (2013) Ubiquitin-dependent recruitment of the Bloom Syndrome helicase upon replication stress is required to suppress homologous recombination. *EMBO J*. **32**, 1–15

[27] Soo Lee, N., Jin Chung, H., Kim, H.-J., Yun Lee, S., Ji, J., Seo, Y., Hun Han, S., Choi, M., Yun, M., Lee, S., Myung, K., Kim, Y., Chul Kang, H., and Kim, H. (2016) TRAIP/RNF206 is required for recruitment of RAP80 to sites of DNA damage. *Nat. Commun*. **7**, 10463

[28] Yan, Z., Kim, Y. S., and Jetten, A. M. (2002) RAP80, a novel nuclear protein that interacts with the retinoid-related testis-associated receptor. *J. Biol. Chem*. **277**, 32379–32388

[29] Di Croce, L., and Helin, K. (2013) Transcriptional regulation by Polycomb group proteins. *Nat. Struct. Mol. Biol*. **20**, 1147–1155

[30] Shanbhag, N. M., Rafalska-Metcalf, I. U., Balane-Bolivar, C., Janicki, S. M., and Greenberg, R. A. (2010) ATM-Dependent chromatin changes silence transcription in cis to dna double-strand breaks. *Cell*. **141**, 970–981

[31] Kinner, A., Wu, W., Staudt, C., and Iliakis, G. (2008) Gamma-H2AX in recognition and signaling of DNA double-strand breaks in the context of chromatin. *Nucleic Acids Res*. **36**, 5678–5694

[32] Lombardi, P. M., Matunis, M. J., and Wolberger, C. (2017) RAP80, ubiquitin and SUMO in the DNA damage response. *J. Mol. Med.* **95**, 799–807

[33] Eskiw, C. H. (2003) Size, position and dynamic behavior of PML nuclear bodies following cell stress as a paradigm for supramolecular trafficking and assembly. *J. Cell Sci.* **116**, 4455–4466

[34] Breitenbach, V. L., H. De (2010) PML Nuclear Bodies. *Cold Spring Harb. Perspect. Biol.,***2**:a000661

[35] Ryung, H., Munkhjargal, A., Kim, M., and Young, S. (2017) Mutat Res Fund Mol Mech Mutagen The functional roles of PML nuclear bodies in genome maintenance. *Mutat   Res   Fund   Mol   Mech   Mutagen.* http://dx.doi.org/10.1016/j.mrfmmm.2017.05.002

[36] Dellaire, G., and Bazett-Jones, D. P. (2004) PML nuclear bodies: Dynamic sensors of DNA damage and cellular stress. *BioEssays.* **26**, 963–977

[37] D'Orazi, G., Cecchinelli, B., Bruno, T., Manni, I., Higashimoto, Y., Saito, S., Gostissa, M., Coen, S., Marchetti, A., Del Sal, G., Piaggio, G., Fanciulli, M., Appella, E., and Soddu, S. (2002) Homeodomain-interacting protein kinase-2 phosphorylates p53 at Ser 46 and mediates apoptosis. *Nat. Cell Biol.* **4**, 11–19

[38] Hofmann, T. G., Möller, A., Sirma, H., Zentgraf, H., Taya, Y., Dröge, W., Will, H., and Schmitz, M. L. (2002) Regulation of p53 activity by its interaction with homeodomain-interacting protein kinase-2. *Nat. Cell Biol.* **4**, 1–10

[39] Kamitani, T., Kito, K., Nguyen, H. P., Wada, H., Fukuda-kamitani, T., and Yeh, E. T. H. (1998) Identification of Three Major Sentrinization Sites in PML. *J. Biol. Chem.,***273**, 26675–26682

[40] Uckelmann, M., and Sixma, T. K. (2017) Histone ubiquitination in the DNA damage   response.   *DNA   Repair   (Amst).* http://dx.doi.org/10.1016/j.dnarep.2017.06.011

[41] Ruthenburg, A. J., Li, H., Patel, D. J., and Allis, C. D. (2007) Multivalent engagement of chromatin modifications by linked binding modules. *Nat. Rev. Mol. Cell Biol.* **8**, 983–94