Structural investigation of the ubiquitylation of nucleosome core particles by the E3 ligase BRCA1/BARD1

Sivaraman, Hariharan

2019

Sivaraman, H. (2019). Structural investigation of the ubiquitylation of nucleosome core particles by the E3 ligase BRCA1/BARD1. Doctoral thesis, Nanyang Technological University, Singapore.

<https://hdl.handle.net/10356/84338>

<https://doi.org/10.32657/10220/49322>

**NANYANG
TECHNOLOGICAL
UNIVERSITY**

SINGAPORE

**STRUCTURAL INVESTIGATION OF THE
UBIQUITYLATION OF NUCLEOSOME CORE PARTICLES**

BY THE E3 LIGASE BRCA1/BARD1

SIVARAMAN HARIHARAN

SCHOOL OF BIOLOGICAL SCIENCES

2019

**STRUCTURAL INVESTIGATION OF THE
UBIQUITYLATION OF NUCLEOSOME CORE PARTICLES
BY THE E3 LIGASE BRCA1/BARD1**

SIVARAMAN HARIHARAN

SCHOOL OF BIOLOGICAL SCIENCES

A thesis submitted to the Nanyang Technological
University in partial fulfillment of the requirement for
the degree of Doctor of Philosophy

2019

Statement of Originality

I hereby certify that the work embodied in this thesis is the result of original research done by me except where otherwise stated in this thesis. The thesis work has not been submitted for a degree or professional qualification to any other university or institution. I declare that this thesis is written by myself and is free of plagiarism and of sufficient grammatical clarity to be examined. I confirm that the investigations were conducted in accord with the ethics policies and integrity standards of Nanyang Technological University and that the research data are presented honestly and without prejudice.

15 January 2019

Date

A handwritten signature in blue ink, consisting of a stylized circular mark followed by the name 'Sivaraman Hariharan' written in a cursive script.

Sivaraman Hariharan

Supervisor Declaration Statement

I have reviewed the content and presentation style of this thesis and declare it of sufficient grammatical clarity to be examined. To the best of my knowledge, the thesis is free of plagiarism. The research and writing are those of the candidate's except as acknowledged in the Author Attribution Statement. I confirm that the investigations were conducted in accord with the ethics policies and integrity standards of Nanyang Technological University and that the research data are presented honestly and without prejudice.

15 January 2019

Date

A handwritten signature in black ink, appearing to read 'Daniela Rhodes', written in a cursive style.

Professor Daniela Rhodes

“Science, my lad, is made up of mistakes, but they are mistakes which it is useful to make, because they lead little by little to the truth.”

Jules Verne, A journey to the centre of the earth

Acknowledgments

The course of Ph.D. has put me through a lot of different experiences and emotions which has changed my intellectual ability overall. I want to take this opportunity to express my gratitude to all those who supported me and stood by me with all the hope and optimism.

First and foremost, I would like to heartfully thank my supervisor Professor Daniela Rhodes for giving me the opportunity to work with her, and on this project, it has been a great experience. I want to thank Professor Daniela Rhodes for all her support and encouragement. It has been a great experience to work with a renowned scientist like Professor Daniela Rhodes (student and colleague of Sir Aaron Klug), who has always shown her burning desire to address central questions about Chromatin and DNA-protein interactions. The opportunity to work with a scientist of that caliber has shaped me professionally and personally. I want to thank my co-supervisor Associate Professor Curt Alexander Davey for his support and the opportunity he gave to work in his lab.

I thank all my friends and colleagues at NTU. I thank Dr. Andrew Wong, Dr. Ufuk Borucu, Dr. Bilal Ahsan, Mr. Liu Zhehui, Dr. Abbas El Sahili, Dr. Vasuki Ranjani, and Ms. Srujana Kishore for all the scientific discussions, guidance, emotional support and help they provided, which enhanced my ability to cruise through the difficult times.

Above all, I thank my family for their immense love, care, and emotional support. I am forever grateful to my dad, mom, brother, my darling wife Dr. Usha Sreekumar Lekshmy Kunjamma, my father-in-law, mother-in-law and sister-in-law, who stood by me in all my ups and downs. I thank Singapore and NTU for the financial support they provided to fulfill this beautiful journey.

Abstract

Based on the reports from the American Cancer Society, cancer is the leading cause of death in the world, followed by heart diseases. Scientific methods and technology for early detection and intervention would save not only millions of lives but also billions of dollars. Among the types of cancers, according to the World Health Organization, breast cancer is the most common type of cancer in women worldwide. Statistics show that one in eight women is diagnosed with breast cancer, and eight out of 10 women do not have a family history of the disease. Approximately 10% of the breast cancers can be traced back to inherited mutations in the BRCA1 and BRCA2 genes, most of the mutations occurring in BRCA1.

Though BRCA1 was found to be the dominant gene involved in breast cancer, only recently insights into how BRCA1 protein-dependent tumor suppression occurs, has started to emerge. A potential tumour suppressor that heterodimerizes with BRCA1 *in vitro and in vivo*, increasing the stability of BRCA1 is BARD1.¹ The heterodimerization of the BRCA1 and BARD1 tumour suppressor proteins are mediated by the N-terminal RING domains of BRCA1 and BARD1.² Ample information is available regarding the diverse functions and characteristics of BRCA1 and BARD1 tumor suppressor proteins, but only in the context of cell biology and biochemistry. Even though structural information on the RING domains and BRCT domains of BRCA1 and BARD1 proteins themselves is available, there is no structural information on the mechanism by which BRCA1-BARD1 heterodimeric complex functions as a chromatin remodeler. Results over the last two decades have provided evidence that histone post-translational modifications (e.g., acetylation, methylation, and ubiquitylation) play a vital role in regulating numerous processes within the eukaryotic

cell, such as transcriptional regulation, DNA replication, and repair, leading to genomic instability and tumor progression.

In recent years, H2A specific ubiquitin ligases have been discovered and link H2A ubiquitylation (H2AUb) to gene silencing and DNA damage repair. A critical E3 ligase involved in DNA break repair is the tumor suppressor BRCA1. It is well known that mutations in BRCA1 are associated with the occurrence of breast and ovarian cancer. Prominent amongst the many *in vivo* identified interaction partners of BRCA1 is the interaction of BRCA1 with the BRCA1-associated RING domain protein 1 (BARD1). BRCA1 and BARD1 interact with each other to form a heterodimeric RING-RING complex that ubiquitylates K125, K127, and K129 explicitly in the C-terminal tail of histone H2A in the human nucleosome.³

Complete understanding of the mechanism by which the chromatin remodelers function and bring about a specific biological outcome is yet to be accomplished. Structurally determining the highly dynamic interactions of E3 ubiquitin ligases with related proteins and their substrates has also proved to be a challenge for a mechanistic investigation into these complex systems, which is further complicated by the technical difficulty in isolating homogenous samples. The latest advancements in the field of cryo-electron microscopy (cryo-EM) are proving helpful.

I have used cryo-electron microscopy and biochemical techniques, to understand the mechanism of interaction between BRCA1-BARD1 heterodimer and nucleosome core particle (NCP) to specifically modify K125, K127, and K129 in the C-terminal tail of H2A. Determining the mechanism of interaction at the molecular level would shed light on mechanisms involved to recognize, alter chromatin and how mutations in these proteins are involved in the onset of diseases, cancer. The structural information along

with biochemical data and data from functional studies would help in developing novel therapeutic compounds that could target diseases like cancer. In this dissertation, I will describe approaches taken towards the expression and purification of various BRCA1 and BARD1 constructs, characterization of the BRCA1-BARD1 heterodimer in the nucleosomal context, recognition of the NCP, and preliminary structural information using cryo-EM on how the heterodimeric complex binds to the NCP is presented.

Table of Contents

Acknowledgments	6
Abstract	7
Table of Contents	10
List of Tables	14
List of Figures	15
Abbreviations	19
Chapter 1: Introduction	23
1.1. Chromatin	23
1.2. The Nucleosome Core Particle	26
1.3. Chromatin higher-order structure	30
1.4. Nucleosomal DNA	33
1.5. The H2A-H2B acidic patch	34
1.6. Histone modifications and modifiers	36
1.6.1. Histone modifications involved in carcinogenesis.....	39
1.7. Ubiquitylation	40
1.7.1. Types of ubiquitylation.....	44
1.7.2. Ubiquitylation and E3 ligases in the development of diseases	47
1.7.3. Ubiquitin	48
1.7.4. E2 conjugating enzymes	49
1.7.5. E3 ubiquitin ligase.....	52
1.7.5.1. BRCA1 and BARD1	54
1.7.5.2. BRCA1-BARD1 associated genome regulation complex	60
1.8. Objective of the study	65
Chapter 2: Materials and methods	66
2.1. General Materials and Methods	66
2.1.1. Plasmids	66
2.1.2. Ligation Independent Cloning	68
2.1.3. Preparation of proteins and protein complexes	68

2.2. Nucleosome core particles	71
2.2.1. Expression and purification of human histones	71
2.2.2. Large-scale production of Widom 601 DNA.....	73
2.2.3. Assembly of the histone octamer	74
2.2.4. Preparation of NCPs	76
2.3. Assays	77
2.3.1. Electrophoretic Mobility Shift Assay	77
2.3.2. Ubiquitylation assay	78
2.4. Insect cell expression system.....	79
2.4.1. MultiBac system.....	79
2.4.2. Cre-loxP recombination reaction	81
2.4.3. Blue-white screening and bacmid isolation	81
2.5. Electron microscopy.....	82
2.5.1. Preparation of Cryo-EM grids.....	82
2.5.2. Data collection	83
2.5.3. Data processing	83
Chapter 3: Biochemical analysis of the BRCA1-BARD1 interaction with NCP ..	85
3.1. Overview	85
3.2. NCPs for biochemical characterization and structural studies	88
3.2.1. Production of Widom 601 nucleosome positioning sequence	88
3.2.2. Production of the core histone proteins H2A, H2B, H3, and H4.....	89
3.2.3. Assembly of the histone octamer	89
3.2.4. Reconstitution of NCPs	91
3.3. Purification of BRCA1-BARD1 E3 ligase and UbcH5c conjugating enzyme..	92
3.3.1. Co-expression and co-purification of BRBA-S from <i>E. coli</i>	92
3.3.2. Expression and purification of UbcH5c from <i>E. coli</i>	98

3.3.3. Co-expression of BRCA1 and BARD1 constructs in SF9 cells	100
3.3.4. Ni-NTA affinity purification of BRBA-L from Insect cells	103
3.4. Analysis of histone H2A ubiquitylation by BRCA1-BARD1	105
3.4.1. Multiple mono-ubiquitylation of histone H2A	105
3.4.2. Native gel analysis of the NCP modified by ubiquitin molecules	108
3.4.3. Self-ubiquitylation of BRBA-S	110
3.4.4. Analysis of BRCA1-BARD1 mediated H2A ubiquitylation by wt-Ub	112
3.4.5. Analysis of BRCA1-BARD1 mediated H2A ubiquitylation by K0-Ub	115
3.5. Conclusions.....	119
Chapter 4: Stabilization of complexes for cryo-EM analysis.....	122
4.1. Overview	122
4.2. Results	123
4.2.1. Cross-linking of histones at increasing concentrations of glutaraldehyde	123
4.2.2. Analysis of cross-linking of histones at different time points	125
4.2.3. Analysis of cross-linking of BRBA-S at increasing concentrations of glutaraldehyde.....	127
4.2.4. Cross-linking of BRBA-S to NCP	129
4.2.5. Cross-linking analysis of NCP _{K0-Ub} -UbcH5c-BRBA-S with glutaraldehyde.....	131
4.3. Conclusions.....	133
Chapter 5: Structural analysis by single-particle cryo-EM	135
5.1. Introduction to electron microscopy	135
5.1.1. Background	135
5.1.2. Direct electron detectors	137

5.1.3. Phase plate	138
5.1.4. Electron microscopy for NCPs and NCP-chromatin remodeler complexes	139
5.2. Results	139
5.2.1. Cryo-EM analysis of non-cross-linked NCP-UbcH5c-BRBA-S complex	139
5.2.2. Structural analysis of non-cross-linked NCP-UbcH5c-BRBA-S complex	142
5.2.3. Cryo-EM analysis of cross-linked NCP _{KO-Ub} -UbcH5c-BRBA-S complex	145
5.2.4. Data collection and 2D classification of cross-linked complex.....	146
5.3. Conclusions.....	154
Chapter 6: Discussion	156
6.1. Cellular functions and physiological importance of BRCA1-BARD1	156
6.2. BRCA1-BARD1 ubiquitylation and interaction with NCP	156
6.3. Structure of NCP-BRBA-S complex obtained by cryo-EM and NCP-RING1B- BMI1 complex	157
6.4. Sequence comparison of BRCA1 and RING1B	163
6.5. Interaction of chromatin remodelers with NCP	164
6.6. Improving the resolution of EM structures.....	168
6.7. Concluding remarks	169
References	170
Supplementary Data	187

List of Tables

Table 1: Name of plasmids with corresponding tagged inserts, and encoded

antibiotic resistance 66

Table 2: Antibiotics, media, and reagents..... 67

Table 3: Data collection and processing parameters for non-cross-linked NCP-

UbcH5c-BRBA-S..... 141

Table 4: Data collection and processing parameters for cross-linked NCP_{K0-Ub}-

UbcH5c-BRBA-S complex 147

List of Figures

Figure 1: History of chromatin.	24
Figure 2: Different levels of DNA compaction: DNA-nucleosome-chromatin-chromosomes.⁵	25
Figure 3: Electron microscopic visualization of chromatin.¹¹	26
Figure 4: Nucleosome core particle in different orientations (PDBID: 1KX5).²¹ ..	28
Figure 5: Two proposed models of the higher-order chromatin structure.⁴⁵	33
Figure 6: TA steps in nucleosome reconstituted with Widom 601L sequence.⁵⁵ ..	34
Figure 7: Location of the acidic patch on the surface of the NCP (PDBID: 1KX5).	35
Figure 8: The covalent histone modifications.¹¹	38
Figure 9: Pictorial representation of the ubiquitylation cascade.	42
Figure 10: Picture illustrating the biological consequences from ubiquitylation of substrates.¹²⁸	43
Figure 11: Structure of ubiquitin with lysines.	44
Figure 12: Schematic representation of the ubiquitylation types.	45
Figure 13: Structure of the ubiquitin chain, alongside a schematic representation.¹³⁵	46
Figure 14: Crystal structure of ubiquitin (PDBID: 1UBQ).¹⁶²	49
Figure 15: Solution NMR structure of UbCH5c (PDBID: 2FUH).¹⁷⁰	50
Figure 16: Placement of ubiquitin in E2 (light orange)-Ub (blue) complex.	52
Figure 17: Domain architecture of BRCA1 and BARD1.	56
Figure 18: NMR structure of BRCA1-BARD1 heterodimer (PDBID: 1JM7).² ...	57
Figure 19: Mutations in the RING domain of BRCA1-BARD1 heterodimer (PDBID: 1JM7).²	58

Figure 20: BRCA1 interacting residues in helix 1 and loop 2 of UbcH5c.....	59
Figure 21: Diagram illustrating the subunits of the BRCA1-A complex.....	61
Figure 22: Diagram illustrating the subunits of the BRCA1-B complex.....	61
Figure 23: Diagram illustrating the subunits of the BRCA1-C complex.....	61
Figure 24: Schematic illustration of the process by which the histone octamer is assembled.	71
Figure 25: Schematic illustration of the process by which the Widom 601 145bp DNA is produced.....	74
Figure 26: Schematic diagram of the process by which the NCPs are produced.	77
Figure 27: A MultiBac cloning strategy for co-expression of proteins.	79
Figure 28: Production of Widom 601 145bp DNA.	88
Figure 29: Representative chromatogram and SDS-PAGE analysis of histone H2B.....	89
Figure 30: Production of the histone octamer.	90
Figure 31: Production of the NCP.....	92
Figure 32: Purification of the BRBA-S heterodimer.....	95
Figure 33: Purification of BRBA-L from E. Coli.	97
Figure 34: Western blot analysis of the purified BRCA1¹⁻²⁵⁰ & BARD1¹⁴⁻¹⁸⁶ expressed in E. coli.....	98
<i>Figure 35: Purification of UbcH5c (E2 conjugating enzyme).</i>	99
Figure 36: Screening for co-expression of BRCA1-BARD1 complexes.....	102
Figure 37: An example of monitoring expression using YFP as a reporter.	103
Figure 38: Purification of BRBA-L from SF9 cells.	104
Figure 39: Poly-ubiquitylation and multi-mono-ubiquitylation of histone H2A.	107

Figure 40: The conjugation of ubiquitin molecules promotes retardation of the NCP.	109
Figure 41: BRBA-S is prone to self-ubiquitylation.	110
Figure 42: H2A ubiquitylation with wt-Ub.	113
Figure 43: H2A ubiquitylation with K0-Ub.	117
Figure 44: Self-association of nucleosomal histones when treated with different concentrations of glutaraldehyde.	125
Figure 45: Time course of NCPs crosslinked with glutaraldehyde.	126
Figure 46: Cross-linking pattern of BRBA-S heterodimer when treated with increasing concentrations of glutaraldehyde.	128
Figure 47: Cross-linking of BRBA-S to NCP with 0.01% glutaraldehyde.	130
Figure 48: Cross-linking of BRBA-S to NCP with 0.025% glutaraldehyde.	131
Figure 49: Cross-linking of NCP_{K0-Ub}-UbcH5c-BRBA-S with glutaraldehyde.	133
Figure 50: Workflow in single particle Cryo-EM.	136
Figure 51: Cryo-EM screening of non-cross-linked NCP-UbcH5c-BRBA-S.	141
Figure 52: 2D projections of the non-cross-linked NCP-UbcH5c-BRBA-S sample categorized into 35 classes.	143
Figure 53: 7 Å resolution representation of the 3D class.	144
Figure 54: The FSC curve and the final 3.9 Å structure of the NCP.	145
Figure 55: Comparison between micrographs showing non-cross-linked and cross-linked samples.	146
Figure 56: Representative micrographs of the cross-linked NCP_{K0-Ub}-UbcH5c-BRBA-S complex.	148
Figure 57: Gallery of 2D class averages from 390555 particles.	148
Figure 58: Gallery of good 2D class averages.	149

Figure 59: The ‘gold-standard’ Fourier Shell Correlation (FSC) plot for the reconstructed EM density map.....	150
Figure 60: Representation of the final 3D map fitted with the NCP structure obtained using phase plate.....	151
Figure 61: Fitting of BRCA1-BARD1 heterodimeric RING-RING complex into EM surface map.....	153
Figure 62: Crystal structure of RING1B-BMI-NCP complex.⁶⁴	160
Figure 63: A comparison between the final 3D model of NCP_{K0-Ub}-BRBA-S and published NCP-RING1B-BMI1 structure.....	162
Figure 64: Side view comparison between the final 3D model of BRBA-S-NCP and published RING1B-BMI1-NCP structure.	163
Figure 65: BRCA1 and RING1B require similar loop region for NCP recognition and binding.	164
Figure 66: Nucleosome ‘acidic patch’ recognition by the binding factors.⁵⁶	167

Abbreviations

μl	microliter
μm	micrometre
μM	micromolar
Acetyl CoA	Acetyl Coenzyme A
A. I	After induction
Amp	Ampicillin
APC	Anaphase-promoting complex
ATP	Adenosine triphosphate
bp	base pair
BARD1	BRCA1 associated RING domain
BAH	Bromo-associated homology
Bbd	Bar body deficient
BER	Base excision repair
B.I	Before induction
BRCT	BRCA1 C-terminus
BRCA1	Breast cancer associated protein 1
BV	Baculovirus
Cam	Chloramphenicol
CCD	Charge coupled device
CD	Circular dichroism
CMOS	Complementary metal oxide semiconductor
Cryo-EM	Cryo-Electron Microscopy
CTF	Contrast transfer function
d	day
DNA	Deoxyribonucleic acid
DDR	DNA damage response
DPC	Defocus phase contrast
DQE	Detector quantum efficiency
DSB	Double strand breaks
DUB	Deubiquitinases
e.g.	for example
EDTA	Ethylenediaminetetraacetic acid

EMSA	Electrophoretic mobility shift assay
EM	Electron microscopy
EPDE	End point dilution experiment
ERG	Estrogen-responsive genes
EtOH	Ethanol
FAD	Flavin adenine dinucleotide
FCS	Foetal calf serum
FSC	Fourier shell correlation
FT	Flow-through
g	gravity of earth
Gen	Gentamycin
YFP	Yellow fluorescent protein
GuHCl	Guanidinium hydrochloride
h	hour
HBOC	Hereditary breast and ovarian cancer
HDAC	Histone deacetylases
HDM	Histone demethylases
HDR	Homology directed repair
HECT	Homologous to the E6-AP Carboxyl Terminus
HER2	Human epidermal growth factor receptor 2
His6	Hexa-histidine
His10	Deca-histidine
HMGN	High mobility group nucleosome binding
HMT	Histone methyltransferases
HR	Homologous recombination
HSP	Heat shock protein
H2A, H2B, H3, H4	Histone proteins
IBR	In-between-RINGs
IPTG	Isopropyl β -D-1-thiogalactopyranoside
Kan	Kanamycin
kDa	Kilodalton
KOAc	Potassium acetate
K0-Ub	Lysine free ubiquitin
KSHV	Kaposi's sarcoma-associated herpes virus

LANA	Latency-associated nuclear antigen
LIC	Ligation independent cloning
LiCl	Lithium chloride
LUBAC	Linear ubiquitin chain assembly complex
MDM2	Mouse double minute 2
MOI	Multiplicity of Infection
M	molar
mg	milligram
min	minute
mL	millilitre
mM	millimolar
MMR	Mismatch repair
MS	Mass spectroscopy
MW	Molecular weight
MWCO	Molecular Weight Cut-Off
NaCl	Sodium chloride
NCP	Nucleosome core particle
NER	Nucleotide excision repair
NHEJ	Non-homologous end joining
ng	nanogram
nM	nanomolar
NADH	Nicotinamide adenine dinucleotide hydrogen
NMR	Nuclear magnetic resonance
OD ₆₀₀	Optical Density at 600 nm wavelength
PAGE	Polyacrylamide gel electrophoresis
PCR	Polymerase chain reaction
PDB	Protein data bank
PEG	Polyethylene glycol
PML	Promyelocytic leukaemia
PPP	Protein production platform
PRC1	Polycomb repressive complex 1
PTM	Post-translational modification
RCC1	Regulator of chromatin condensation

RBR	RING between RING
RELION	Regularized likelihood optimization
RFP	ret proto-oncoprotein finger protein
RPM	Revolutions per minute
RT	Room temperature
SCF	Skp1-CUL1-F-box protein
SDS	Sodium dodecyl sulfate
SELEX	Systematic evolution of ligands by exponential enrichment
SHL	Super helical location
SIR	Silent information regulator
SUMO	Small ubiquitin like modifier
Sup.	Supernatant
RING	Really interesting new gene
T18	Transforming factor 18
TB	Tris borate
TE	Tris-EDTA
u	Unit
UIM	Ubiquitin interacting motif
USP	Ubiquitin-specific protease
VPP	Volta phase plate
wt-Ub	Ubiquitin
X-Gal	5-bromo-4-chloro-3-indolyl- β -D-galactopyranoside
YFP	Yellow fluorescent protein
ZPP	Zernike phase plate
ZnCl ₂	Zinc chloride

Chapter 1: Introduction

“At lunch Francis [Crick] winged into the Eagle public bar to tell everyone within hearing distance that we had found the secret of life.”

- James Watson

1.1. Chromatin

In 1953, James D. Watson, Francis Crick, with the scientific contribution from Rosalind Franklin and Maurice Wilkins, discovered the structure of deoxyribonucleic acid (DNA) that forms the basis of heredity and contains the code for the synthesis of proteins. DNA comprises of two strands twined with each other (double helix), of alternating sugar and phosphates, twined strands stabilized by hydrogen bonds amid the combination of adenine (A) with thymine (T), and guanine (G) with cytosine (C), nucleotide bases.

The revelation of the DNA and its structure was built on many scientific advancements like the invention of X-ray crystallography, findings by Rollin D. Hotchkiss, Erwin Chargaff and Linus Pauling. Erwin Chargaff came up with two basic rules that paved a way to the discovery of DNA, 1) the number of A equals the number of T and G equals C in DNA. 2) relative amounts of guanine, cytosine, adenine, and thymine bases differ from one species to another, hint that DNA could be the genetic material, and not protein. Discovery of the DNA structure was the foundation for all technological, diagnostic, and therapeutic advancements in modern biotechnology. The chronology of seminal discoveries allowed researchers to establish the structure of chromatin and nucleosomes. Figure 1 illustrates the chronology of developments.

1871	• F. Miescher discovered nucleic acids (Miescher, 1871)
1880	• The term 'Chromatin' was coined by W. Flemming
1883	• A. Weismann related chromatin with heredity
1884	• A. Kossel discovered histones (Kossel, 1911)
1888	• Sutton and Boveri proposes the continuity theory of chromatin
1889	• The term 'chromosome' was coined by W. V. Waldemeyer
1928	• The term 'euchromatin' and 'heterochromatin' was coined by E. Heitz
1942	• C. Waddington proposed 'epigenetic landscapes'
1944	• O. Avery, C. MacLeod and M. McCarty identified DNA as the 'transforming principle' (Avery et al, 2000)
1953	• Structure of DNA was determined by J. Watson, F. Crick, M. Wilkins, and R. Franklin (Franklin & Gosling, 1953; Watson & Crick, 1953)
1961	• M. Lyon proposed the concept of 'X-Chromosome inactivation'
1967	• Histones were fractionated by E. W. Johns (Johns, 1969)
1973	• Visualization of chromatin using electron microscopy by A. Olins & D. Olins and C. Woodcock (Woodcock, 1973)
1974	• A. Olins & D. Olins and R. Kornberg & J. Thomas proposed the chromatin subunit model (Kornberg, 1974; Olins & Olins, 1974; Woodcock, 1973)
1975	• P. Oudet, M. Bross-Bellard and P. Chambon coined the term 'nucleosome' (Oudet, Bellard, & Chambon, 1975)
1984	• 7.0 Å structure of nucleosome solved by A. Klug and group (Richmond, Finch, Rushton, Rhodes, & Klug, 1984)
1993	• B. Turner proposed that epigenetic information resides in histone tail modifications (Turner, 1993)
1997	• 2.8 Å structure of nucleosome determined by T. Richmond and group (K. Luger, Mäder, Richmond, Sargent, & Richmond, 1997)
2001	• T. Jeruwein and C. Allis proposed the concept of 'Histone code' (Jenuwein & Allis, 2001)

Figure 1: History of chromatin.

The human genome comprises of approximately 3 billion bp of DNA with a length of about 2 meters (6.56 foot) in each cell and trillion cells in a human together, there would be a total of 100 billion miles of DNA in a human.⁴ Sun is at 92.9 million miles from the earth. How is the DNA that is so long stored in a cell that is just 10-20 microns in diameter? How is it dynamic to be accessible for proteins, enzymes, and other cellular components to obtain information for use in regulating complex cellular processes?

To overcome these challenges, 'nature' has evolved a set of nuclear proteins, the histones, around which the linear DNA is wrapped, thus condensing the DNA to fit within the nucleus of the cell. Wrapping of DNA around the histone proteins demands work and work requires energy. Therefore, histones cater energy in the mode of electrostatic interactions to bend DNA. The resulting compacted protein-DNA complex is termed chromatin. The term 'Chromatin' was coined by W. Flemming. Work in the last several decades has revealed that chromatin is a dynamic genome organizer, a

scaffold that regulates most of the processes whose substrate is DNA, such as transcription, replications, repair, and recombination. Chromatin is further compacted and condenses to form chromosomes. The different levels at which the DNA is compacted to form higher-order structures is depicted in Figure 2.

Figure 2: Different levels of DNA compaction: DNA-nucleosome-chromatin-chromosomes.⁵

Even though images of natural chromatin fibers were available before the ones published by Oudet et al., those images were challenging to decode. Oudet et al. were able to accomplish it by optimizing sample preparation techniques and by the removal of linker histone that enhanced the quality of the picture to exhibit what was called to be ‘beads on the string.’^{6–10} 30 nm thick chromatin fiber and decondensed chromatin are shown in Figure 3.

Figure 3: Electron microscopic visualization of chromatin.¹¹

A) 30 nm chromatin fiber isolated from the nucleus at the interphase step of the cell cycle (Courtesy of Barbara Hamkalo). B) Electron micrograph of decondensed chromatin ('beads on a string') where the nucleosomes are shown (Courtesy of Victoria Foe).

1.2. The Nucleosome Core Particle

The nucleosomes are the basic building blocks of chromatin. The NCP consists a histone octamer constituting two copies each of the four positively charged proteins, histones H2A, H2B, H3, H4, surrounding which 145-147 bp of negatively charged DNA is wound in almost 1.7 left-handed tight turns, as shown in Figure 4.^{6,12,13}

The NCPs are highly conserved amongst eukaryotes and forms a structure that is disc-shaped measuring a height of 5.5 nm and diameter of 11 nm.¹³ NCPs serve as a scaffold for binding and modification by chromatin remodelers and histone modifying enzymes. In general, the coiling of DNA surrounding the histone octamer obstruct the interaction of non-histone, transcription factors and chromatin remodelers. It is necessary for the DNA to dissociate partially, enabling access, the property termed as 'site accessibility,' for the modifications to occur.

Initial experiments with the aim of characterizing chromatin and determining the structure of the nucleosome, required samples from native sources like beef kidney, rat liver, calf thymus, myeloma cells, and chicken blood, resulting in mixed genomic sequences.^{13,14} To produce native NCP, chromatin from difficult sources (rat liver, chicken blood, beef kidney) was digested extensively with micrococcal nuclease resulting in individual nucleosomes containing about DNA of 147 ± 2 bp in length, hence the term nucleosome core particle (NCP).^{15,16} In this process the linker histone was lost, and the first crystal structure of the native NCP was determined at 7 Å resolution.¹³

After the discovery of 7 Å resolution structure, in order to produce more homogenous NCPs, specific sequence DNA were used to reconstitute nucleosomes.^{17,18} Such crystals gave a better diffraction and the 2.8 Å crystal structure of the NCP was determined using human α -satellite centromeric DNA repeat sequence in 1997.¹⁹ Subsequently the Widom 601 positioning sequence chosen by SELEX method has been the most popular nucleosome positioning sequence to produce ordered nucleosome and chromatin.

Histone-DNA direct interactions occur in 10 bp TA step periodicity, and there are 14 sites at which direct interaction between the nucleosomal DNA and histones happen, termed Superhelical location (SHL).²⁰ Electrostatic and van der Waals forces govern direct interactions between the DNA and the histones. The electrostatic interactions between the side chains of S, T, R, or K and DNA phosphate backbone being more predominant. The NCP is shown in Figure 4.

Figure 4: Nucleosome core particle in different orientations (PDBID: 1KX5).²¹

DNA and histones depicted in cartoon representations. H2A, H2B, H3, and H4 are colored in yellow, red, blue, and green, respectively.

Core histones H2A, H2B, H3 and H4 in eukaryotes display a high level of sequence conservation and are globular with long N-terminal tails that are unstructured and highly charged. Histone tails are rich in R and K residues, conferring a steady net positive charge. Each histone core confers a common structural organization called the ‘histone fold’ that promotes the physical interaction within the histone proteins to assemble the octamer protein disc. This serves as a spool onto which the DNA is wrapped. Histone-fold motif is formed by three central alpha-helices linked by loops $\alpha 1$ -L1- $\alpha 2$ -L2- $\alpha 3$, with N and C-terminal flanking sequences.^{19,22,23} Histone-fold of one histone (H2A) team with histone-fold of another histone (H2B) to assemble a histone-fold heterodimer, similar heterodimer form between histone H3 and H4. The antiparallel combination of the histone folds from complimentary histones results in a crescent-shaped heterodimeric structure. The H3-H4 heterodimers interact with each other via the four-helix bundle to assemble the (H3-H4)₂ tetramer. The H2A-H2B dimers directly associate with the (H3-H4)₂ through the H2B and H4 histone folds, together forming a complete octamer. C-terminal of H2A and H2B add to the surface of the nucleosome

core. The histone octamer assembly serves as spool for the wrapping of nucleosomal DNA. Site-specific hydroxyl radical footprinting experiments and X-ray crystallographic studies confirmed the nucleosome dyad aligns with a central TA base pair, having DNA length of 73 bp on each side of the nucleosome core.²⁴

It is not appropriate to consider histone octamer just as a structural organizing scaffold as many findings associate post-translational modifications on histones to diseases like cancer, other epigenetic disorders. Also, nucleosomal stability and diverse cellular processes are further regulated by modifications of histones and histone variants like H2AX and H2AZ.²⁵

Nature has created a set of histone variants like H2A.Z and H2A Bbd which are instrumental in modifying the surface of the acidic patch, by replacing histones. Replacement of H2A with H2A.Z increases the acidic nature of the patch, and H2A Bbd decreases that property. This change in the acidic property has been shown to affect the dynamics of the higher-order chromatin structure.^{26,27} In addition to the core histone proteins, an additional histone, the histone H1 to the entry and exit of the DNA prevents it from unwrapping. H1 is vital for the assembly of the higher-order structure in chromatin.²⁸

In addition to histone variants, post-translational modification of histones can also bring changes to the surface properties of the nucleosome, histone-nucleosomal DNA interactions, and the ability of the nucleosome to recruit chromatin remodelers. Altogether, these local changes in the NCP impact the higher-order chromatin structure and their functions in cells.²⁹⁻³³

1.3. Chromatin higher-order structure

The NCPs are linked together by short stretches of linker DNA with lengths ranging between 0 and 80 bp, assembling higher order chromatin structures.^{34–36} The compaction of nucleosomal arrays into higher order chromatin structures is governed by the internucleosomal synergy mediated by the tail regions of the core histones, histone H4 tail in particular.³⁷ Hybrid trypsinization experiments has shown that the H3/H4 tails favour compaction of the nucleosomal arrays into higher order structure, in comparison with H2A/H2B tail, implying the importance of H3/H4 tails in chromatin compaction.³⁸ Analysis showed that the formation of compacted chromatin fibre was achieved with deletions in the tail regions of histones H2A or H2B or H3, but not with H4 tail deletions, suggesting the vital role of histone H4 tail in chromatin compaction and stabilization.³⁹ The H4 tail of a nucleosome interacts with the acidic patch of the neighbouring nucleosome, eventually leading to compaction of the chromatin fibre. In 1997, Richmond's group saw the interaction between H4^{16–25} N-terminal tail and the acidic patch of the neighboring nucleosome in the high-resolution crystal structure of the NCP.¹⁹ The H4^{16–25} region formed hydrogen bonds and salt bridges with the acidic chains of E56, E61, E64, D90, E91, E92 of H2A and E110 of H2B. The modelling simulations presented by Nordenskiöld's group showed that there could be a possible interaction between the H4K16 and the H2B^{96–99}.⁴⁰

The results from computational modelling simulations demonstrated that the H4^{16–22} shows tendency to form α -helix and fits well in the groove formed by the acidic patch.⁴¹ The N-terminal region of H4 forms an extended structure when interacting with the acidic patch, validated by the results showing H4V21 in close proximity to H2A E64 in the compacted nucleosomal arrays.^{21,42}

Analytical ultracentrifugation sedimentation velocity investigations has shown that the H4¹⁴⁻¹⁹ region plays an important role in mediating chromatin condensation.³⁹ Nordenskiöld's group through sedimentation velocity experiments showed that, mono-, di-, and tri-valent cations promote internucleosomal interactions, leading to condensation of arrays.⁴⁰ Although it is evident that the H4 tail is vital for mediating the chromatin compaction and stabilization, it is also important to know that the acetylation of H4K5, H4K8, H4K12, and H4K16 leads to partial unfolding of the chromatin and destabilization.^{19,41,43}

Overall, the chromatin remodelers that act on DNA such as polymerases can gain accessibility to the genome by fine-tuning the level of chromatin compaction. The crucial cellular processes like repair, recombination, and replication are affected by packing of DNA into chromatin and the changes in their structure. The changes in the chromatin structure are due to the post-translational modifications of histone residues that in turn has an impact on the interactions between nucleosomes affecting the packing of nucleosomal arrays into higher-order chromatin structure. Upon formation of chromatin fibers, fiber-fiber interactions are thought to increase compaction as in condensed chromosomes.⁴⁴

Rigorous biochemical, biophysical, and structural investigations by electron microscopy led to two different models of its structure. Daniela Rhodes and group reconstituted and studied long chromatin fibers with varying nucleosomal repeat lengths of 178, 188, 197, 207, 217, 227, and 237 bp incorporated with histone H5 from chicken erythrocytes and contained between 47 and 80 nucleosomes in length, respectively.^{45,46} Analysis of the nucleosomal arrays reconstituted with 177-207 bp nucleosomal DNA gave rise to chromatin fibres which is 33 nm in diameter and 11 nucleosomes per 11

nm. Nucleosomal arrays reconstituted with 217-237 bp of nucleosomal DNA produced chromatin fibres which is 44 nm in diameter and 15 nucleosomes per 11 nm. Comparing the fibers, the diameter of the fiber does not increase linearly with increase in length of the linker DNA.

Investigation of the reconstituted sample using negative stain electron microscopy suggested that the 30 nm fiber confers an interdigitated solenoid structure (Figure 5) comparable to the structure proposed by Antonio Bermudez and group.⁴⁷ In the one-start helix structure, the interaction between the successive nucleosomes resulted in a helical trajectory with interdigitated pattern.⁴⁵ In the second model, named as a two-start helical crossed-linker model or two-start zigzag structure (Figure 5), nucleosomes in two rows assemble two-start helix, resulting in the interaction between alternating nucleosomes with the DNA linker following a straight trajectory over the center of the fiber.⁴⁸ Be it one-start helix model or two-start helical cross-linker model, linker histone is essential to achieve higher level of compaction.⁴⁶

Consistent with what was observed earlier by Rhodes group on the crystal structure of NCPs, further investigation after deletion of H4 tail residues and mutation experiments highlighted the importance of H4¹⁴⁻¹⁹ in chromatin compaction, especially K16 residue that is found acetylated in transcribed genes.⁴² Acetylation of K16 is an indication of transcriptionally functional chromatin, and such modification has a mechanistic effect in the modulation of higher-order structure of chromatin.⁴⁹ The work done by David Tremethick and group highlighted the importance of the acidic patch in chromatin dynamics.²⁷

A)

B)

Figure 5: Two proposed models of the higher-order chromatin structure.⁴⁵

A) Interdigitated solenoid one-start model is showing the interaction between the successive nucleosomes. B) Zigzag two-start model is showing the interaction between the alternating nucleosomes.

1.4. Nucleosomal DNA

Wrapping of DNA around the histone octamer requires energy. The DNA sequence to an extent influences the binding affinity and positioning of the histone octamer. Analysis of nucleosomal DNA from native sources and nucleosome positioning across the genome showed A-T steps prefer to position inwards and G-C steps position outwards due to their narrow and wide minor grooves, respectively.^{50,51} The propensity shown by A-T rich sequences in the minor groove towards the octamer influences bending of the DNA molecule.^{24,52–54}

In the NCP structure, the phosphate groups of the DNA are gripped by the $\alpha 2$ helices of the four histones and N-terminal $\alpha 1$ helix of the histone H3, H2B, and H4, via the positive charge caused by the helix dipole. The loops L1-L2, N-terminal helices $\alpha 1$ - $\alpha 1$ and loop2-loop1 of (H3-H4)₂ tetramer anchor phosphate backbone of the DNA where

the minor groove faces the octamer (Figure 6). TA base steps occur 5 bp away from the dyad, then in 10 bp intervals.

Figure 6: TA steps in nucleosome reconstituted with Widom 601L sequence.⁵⁵

A) Nucleosome core particle showing TA base pairs interacting with H2A-H2B (orange) and H3-H4 (red). H2A, H2B, H3, and H4 are colored yellow, red, blue, and green, respectively. B) Magnified view of H3-H4 heterodimer interacting with DNA, phosphate backbone bound to H3-H4 is shown in spheres (red).

1.5. The H2A-H2B acidic patch

In the histone octamer, the histones interact in an ordered fashion to give rise to a broad, asymmetrically charged protein surface with a deep groove. The formation of a negatively charged acidic surface on the NCP by histones H2A-H2B (residues E56, E61, E64, D90, E91, E92 of H2A and E102, E110 of H2B), remarkably has been found to serve as a platform for a variety of nucleosome binding factors. The negatively charged area on the NCP disk termed as the ‘acidic patch’.¹⁹ The acidic patch contains a hydrophobic pocket formed by residues Y50, V54 and Y57 of H2A. The acidic patch has a deep grooved binding interface with a solid negative charged surface.²¹ Many nucleosome binding factors recognize the H2A-H2B acidic patch and surrounding

residues in close proximity. The deep groove of the acidic patch is the binding site for most chromatin remodelers and histone-modifying enzymes containing a wide range of binding motifs, including RCC1/Sir3 (loops), LANA (hairpin peptide), HMGN2/CENP-C (extended conformations), also attracts helix and β -sheet structural elements.^{56–64}

RCC1, Sir3, LANA peptide, and HMGN2 all have an arginine residue that forms hydrogen bonds with E61, D90, and D92 of H2A on the acidic surface, which suggests that at any given time, only one protein can bind the surface of any one specific nucleosome. This exclusive property of nucleosome to accommodate ‘one at a time’ implies that the nucleosome binding factors need to compete to do their job. Till date, all crystal structures with modifying factors bound to the NCP contribute to a common interaction motif, an arginine residue bound to the H2A-H2B acidic surface.⁵⁵ More detailed view is presented in discussion.

Figure 7: Location of the acidic patch on the surface of the NCP (PDBID: 1KX5).

The H2A-H2B dependent binding is not only common to chromatin remodelers, but also to E3 ubiquitin ligases such as RNF168 and BRCA1. RNF168 E3 ligase interact with the acidic patch and ubiquitylates H2A tail residues K13, K14, and K15. The importance of the H2A-H2B acidic patch in the NCP, for activity and binding of E3 ubiquitin ligases, was shown by mutation of critical residues that are involved in the generation of an acidic patch.^{62,64}

1.6. Histone modifications and modifiers

In the early 1960s after the discovery that histones had a role in suppressing gene activity and regulation, histone modifications were identified.^{65–70} Through mass spectrometry or modification-specific antibodies, there are many residues on histone proteins where modifications have been identified. There are three ways by which chromatin remodelers or histone-modifying enzymes can bind to nucleosomes: (1) interaction via the N-terminal and C-terminal histone tails (2) direct interaction with the surface of the octamer disc (3) through the DNA of a nucleosome.⁷¹ There are two steps in histone modifications; the initial phase is the breaking of nucleosome-nucleosome contacts to loosen chromatin structure, the final step is the recruitment of modifiers to make changes.⁷²

Nucleosome binding factors and chromatin remodelers are recruited to regulate gene associated activities and chromatin architecture.⁷³

Biochemical, biophysical, and structural studies have given us insights about how the acidic patch of the NCP, distinct marks on the nucleosome, nucleosomal DNA, and histone tails can serve as interacting surfaces to direct binding of modifying enzymes to nucleosome and introduce specific modifications. The essential enzymes involved in the histone modifications are histone deacetylases (HDACs), histone acetyl

transferases (HATs), histone demethylases (HDMs), histone methyltransferases (HMTs), histone phosphatases, histone kinases, and E3 ubiquitin ligases.⁷⁴ On a global scale, these modifications could change cell fate and functions. Histones are prone to diverse modifications, including acetylation (e.g., HATs with acetyl CoA as cofactor mediate the transfer of acetyl group to the ϵ -amino group of lysine side chains nullifying the positive charge of the lysines, that in turn devitalize the interaction between histones and DNA), SUMOylation of lysine residues, phosphorylation of serine and tyrosine residues (e.g., histone kinases with ATP mediate the transfer of a phosphate group to the targeted side chains increasing the histone's negative charge, that in turn affects the electrostatic interactions between histones and DNA), methylation of arginine and lysine residues (e.g., HMTs in the presence of S-adenosylmethionine transfers a methyl group to the ϵ -amino group of lysine side chain maintaining the positive charge of the lysine, that in turn maintains the interaction between histones and DNA), ubiquitylation of lysine residues (e.g., unlike other modifications, ubiquitylation attaches a large molecule to a histone that might alter the overall conformation of the nucleosome, that can in turn affect interactions between the nucleosomes).⁷⁵⁻⁸⁰ Apart from many residues that are modified, its more complex due to different forms of each modification: mono, di, or trimethylation, mono, multi-mono, or poly-ubiquitylation.

All histone modifications are written, read, or erased by a variety of enzymes that are at times a part of large multi-protein complexes that act in concert to bring about a final DNA dependent biological outcome (e.g., DNA transcription, replication, repair, regulation of cell differentiation, and cell apoptosis) to maintain genomic stability, resulting in healthy living.^{81,82}

Figure 8: The covalent histone modifications.¹¹

Core histone proteins are shown with their respective modifications. Methylation, phosphorylation, acetylation and ubiquitylation are indicated in blue, yellow, pink, and turquoise blue, respectively.

Chromatin remodelers are key players in accurately decoding the information given by histones & DNA. Many chromatin remodelers need zinc, iron or co-substrates (ATP, acetyl-CoA, S-adenosyl-methionine) or organic coenzymes (FAD, NADH, acetyl-CoA) during enzymatic activity.⁷¹ These enzymes might perform modification independently, or enzyme complexes work together to modify histones. An example is Polycomb group complexes.⁸³

Chromatin remodelers affect histones either locally, via precise recruitment by sequence-specific transcription factors or can affect larger regions of chromatin like coding areas and non-promoter sequences.^{84,85} Such marks affect chromatin dynamics and are programmed for different functions. An example in *Saccharomyces cerevisiae*, Snf1 kinase promotes phosphorylation of H3S10. This modification, in turn, facilitates

the Gcn5 acetyl transferase promoted acetylation of H3K14.⁸⁶ Another example is the acetylation mark on H3K18 and H3K23 that facilitates methylation of R17 by CARM1 methyltransferase, leading to estrogen-responsive gene (ERG) activation.⁸⁷ The RNA polymerase II-associated Set2 methyltransferase promotes the methylation of K36 on H3, facilitating deacetylation of histone H3 and H4 by the Rpd3S complex.⁸⁸ An example of histone modification that is shown to alter the physical properties of the nucleosomes is H3K56 acetylation. Acetylation of H3K56 increases the unwrapping of the nucleosomal DNA from the nucleosome at the entry-exit region, the property termed as 'DNA breathing'. H3K56 acetylation has been implicated in the nucleosome assembly during the process of DNA replication and repair, and in disassembly of nucleosomes during transcription.⁸⁹

1.6.1. Histone modifications involved in carcinogenesis

Histone marks vary between normal cells and cancer cells. Correct patterns of histone modifications are crucial for cell maintenance and functioning, a disorder of which is associated with carcinogenesis.⁹⁰ Deficiency in H4K16 acetylation, along with deficiency in H4K20 tri-methylation is related to hypomethylation of DNA at repeated DNA sequences in cancer cells.⁹¹ Loss of H4K16 acetylation is associated with loosely packed chromatin, leading to genomic instability.⁹¹⁻⁹³

Decrease in the levels of di- or tri-methylation of H3K4 has been identified in a variety of abnormal tissue mass (neoplasms). H3K9 tri-methylation is related to transcriptional repression.⁹⁴ Elevated or reduced levels in H3K9 acetylation has been linked to hepatocellular carcinoma and ovarian cancer, respectively.⁹⁵ Elevated levels of H3K9 methylation has been related to many types of cancer.⁹⁶

H3K27 trimethylation is involved in the repression of genes by the assembly of compact chromatin structure.^{97,98} It is implicated in silencing of genes during differentiation. In humans, H3K27 trimethylation is implicated in breast, oesophageal, ovarian, pancreatic, and prostate cancers.^{99,100} In oesophageal cancers, the expression of H3K27 tri-methylation is elevated and is related to poor prognosis in candidates with cancer. Reduced levels in H3K27 trimethylation is associated with shorter survival time for candidates with breast, ovarian, pancreatic, and prostate cancers.^{101,102} As elevated or reduced levels of H3K27 trimethylation can affect different cancer types, a balance in the levels of H3K27 trimethylation is crucial for healthy cell growth and maintenance.¹⁰³

H3K56 acetylation is found genome-wide, influencing genes associated with many different pathways involved in DNA damage signalling, DNA repair, and apoptosis. Decontrolled expression of these genes leads to tumorigenesis.¹⁰⁴ H3K56 acetylation is also associated with transcription, DNA repair and chromatin remodeling, contributing to genomic stability.^{105–107} H3K56 is located at the entry and exit region of a nucleosome, acetylation of which weakens the interactions between the histone and the DNA near the entry-exit points.¹⁰⁸ A standard marker for active transcription is acetylation of H4K16 and H3K18. It has been found that patients with breast, kidney, lung and pancreatic cancers have decreased H3K18 acetylation that correlates with poor prognosis.^{109,110}

1.7. Ubiquitylation

Until 1980s, it was believed that the proteins are long-lived within the cells. Hershko's group, along with his student Aaron Ciechanover, and collaborator Irwin Rose found that few proteins when added to extracts from immature red blood cells

(reticulocytes) were degraded by an ATP mediated protease present in the reticulocyte extract.^{111,112} This important finding led to the discovery and characterization of the enzymes involved in the covalent attachment of ubiquitin to the proteins, namely, E1, E2 and E3.^{113–115} The ATP mediated protease was later found to be the 26S proteasome.¹¹⁶ Although the enzymes involved were discovered, the diverse functions of the ubiquitylation system was not understood. Subsequent studies by Varshavsky's group through molecular biology investigations and biochemical characterization of *S. cerevisiae* and mammalian cells discovered the ubiquitylation system's role in apoptosis, cell cycle regulation, DNA repair and regulation of transcription.^{117,118}

The ubiquitylation system consists of many different proteins with different functions, namely, ubiquitin-activating enzyme (E1), a ubiquitin-conjugating enzyme (E2), ubiquitin-ligase (E3), and ubiquitin. Ubiquitylation system is complex and depends on the successive enzymatic activity of the E1, E2, and E3 ligase, that results in the attachment of ubiquitin to the ϵ -amino group of a lysine residue. The initial step in ubiquitylation cascade is the activation of ubiquitin catalyzed by the E1 activating enzyme. First, E1 activating enzyme binds to ubiquitin in the presence of ATP-Mg²⁺ and catalyzes adenylation of the ubiquitin C-terminal, next is the formation of a thioester bond between the E1 activating enzyme and the ubiquitin moiety, followed by the release of the AMP group. The next step in the ubiquitylation cascade is the transfer of the ubiquitin moiety from the cysteine residue of E1 to a cysteine residue on the E2 via transthioesterification reaction. The ubiquitylation cycle is complete when an E3 ubiquitin ligase binds to both the substrate and an E2-Ub complex, catalysing the transfer of ubiquitin from E2 to the substrate. The ϵ -amino group of the substrate lysine attacks the thioester leading to the formation of an isopeptide bond between the ubiquitin moiety and the substrate.¹¹⁹ In the process of ubiquitylation, ubiquitin molecule is

covalently attached to the substrate via an isopeptide bond between the lysine of the substrate and the C-terminal glycine residue of ubiquitin. Ubiquitin itself can be acetylated on K residues, or phosphorylated on S, T, or Y residues. Each modification has the potential to alter the signaling outcome dramatically. Ubiquitylation is regulated by deubiquitinases (DUBs). Deubiquitinases such as carboxyl-terminal hydrolase BAP1, CYLD, and A20 are identified as tumor suppressors whose mutations are associated with malignancies.^{120–122}

Figure 9: Pictorial representation of the ubiquitylation cascade.

In terms of kinetics, difference in kinetics across varied ligase systems is reported. Considering Cdc34 (E2) in SCF, or Ube2C and Ube2S in APC, a lackadaisical chain initiation is reported. In SCF substrate Beta-catenin, chain initiation rate ($k_{u,0}$) is slower than the dissociation rate constant (k_{off}) by an order of magnitude. A lackadaisical chain initiation step protects nonsubstrates that has weak affinity for E3 ligases from being ubiquitylated, minimizing energy cost. After the chain initiation step, rapid short chain elongation provides signalling and offers proofreading to sieve nonsubstrates that

interact with E3 ligase with higher affinity.¹²³ Anaphase-promoting complex substrates have k_{off} of 0.01 s^{-1} and k_{u} of 0.1 s^{-1} .¹²⁴ SCF substrates have a k_{off} of 0.3 to 0.4 s^{-1} and k_{u} of an estimated 5 s^{-1} , due to which ubiquitylation happens in seconds.¹²⁵ This speed in ubiquitylation is achieved due to swift binding and dissociation of Cdc34 E2 conjugating enzyme to the SCF E3 ubiquitin ligases.¹²⁶ DUBs activity beyond a threshold switches on the proofreading mode that avoids nonsubstrates for polyubiquitylation. It has also been reported, DUBs activity much higher than the threshold could potentially prevent substrate ubiquitylation. Based on the mechanism proposed, DUBs activity should be well regulated within the cells either through phosphorylation or degradation of DUBs. In the case of APC, DUBs stabilize low-processive substrates and let on high-processive substrates for poly-ubiquitylation for proteasome mediated degradation.¹²⁷

Figure 10: Picture illustrating the biological consequences from ubiquitylation of substrates.¹²⁸

The lysines in ubiquitin are prone to ubiquitylation linkages. Polyubiquitin chains mediated by the K48 residue on the ubiquitin is the most common type and is shown to be responsible for proteasome-dependent degradation of the substrate.¹²⁹ K11 linked chains have been shown to be associated with signaling for proteasomal degradation particularly during cell cycle regulation.¹³⁰ Polyubiquitylation or

monoubiquitylation of the substrate via K63 of ubiquitin serves as a signal for cellular processes like DNA damage response or trafficking of proteins within cells. K6 linked polyubiquitylation is associated with DNA repair.^{131,132} It was discovered that a RING E3 ligase, RNF 168, assembles non-canonical K27 linked chains on histone H2A in cells. Together, the data from various studies imply that K27 associated ubiquitylation signals proteins involved in DNA damage response (DDR).¹³³

Figure 11: Structure of ubiquitin with lysines.

The biological outcome related to the linkage, and the relative abundance (red numbers in parentheses) of specific biological outcome is shown.^{134,135}

1.7.1. Types of ubiquitylation

There are mainly three types of ubiquitylation, 1) modification of a substrate protein on a single residue, with a single ubiquitin molecule (mono-ubiquitylation), 2) modification of a substrate protein on multiple residues, with a single ubiquitin molecule on each residue (multi-mono-ubiquitylation), and 3) modification of substrate residue with a chain of ubiquitin molecules (poly-ubiquitylation). Mono-ubiquitylation has been reported to be a critical ‘priming’ step for the sequential assembly of polyubiquitin chains, and usually different E2 conjugating enzymes determine the mono and/or poly-ubiquitylation of the substrate.^{136–138} Mono-ubiquitylation of histones has been

associated with response to DNA damage.¹³⁹ Polyubiquitylation has been recognized in both degradation of the substrates and non-degradative functions such as protein trafficking.¹⁴⁰ The destiny of the substrates that are ubiquitylated is determined by the number of ubiquitin molecules attached and the topology of the ubiquitin chain.¹⁴¹

Figure 12: Schematic representation of the ubiquitylation types.

The layouts of mono-ubiquitylation, multi-mono-ubiquitylation, and poly-ubiquitylation. Homotypic poly-ubiquitylation, where ubiquitin chain has one linkage type. Heterotypic poly-ubiquitylation, where mixed and branched linkages are depicted.

Chains formed by ubiquitylation either prefer “open conformations,” where no binding surfaces are available except for the site at which the linkage happens, or “compact” conformations, where adjacent ubiquitin molecules interact with each other. The ‘open’ and ‘compact’ conformations are shown in Figure 13. In the standard K48 linked polyubiquitylation, ubiquitin chain embraces compact conformations in which ubiquitin molecules cooperate through their I44 patches, or in minor cases, an I36 patch of a ubiquitin interacts with an I44 patch of the other. Like K48 linkages, K6, and K11 ubiquitin chain conformations are compact as ubiquitin molecules interact with each

other through I36 hydrophobic patches and I44 is exposed to the solvent available to interact with binding partners.^{142–145}

Figure 13: Structure of the ubiquitin chain, alongside a schematic representation.¹³⁵ Hydrophobic patches associated with the interaction between the ubiquitin molecules are shown, I44 (blue) and I36 (green). Proximal moiety is shown in orange and distal moiety is shown in yellow. K6, K11, and K48 are examples for ‘compact’ conformations. Met 1, and K63 are examples of ‘open’ conformations.

Ubiquitin binding proteins recognize ubiquitin via their ubiquitin binding domains (UBDs). Most of the UBDs identify ubiquitin’s I44 hydrophobic patch, with a few of them having the ability to recognize additional surfaces such as I36, L8 on the ubiquitin. The affinity of UBDs for ubiquitin differs dependent on whether it is monoubiquitin or polyubiquitin chain, with UBDs showing higher affinity for polyubiquitin chain compared to monoubiquitin. Mechanisms proteins employ to increase the affinity between UBDs and ubiquitin are dimerization of ubiquitin binding

proteins, tandem UBDs, and compartmentalization. Some examples of UBD containing ubiquitin binding proteins are UbcH5c, UBC1, USP5, RAP80, ataxin-3, and HDAC6.¹⁴⁶

1.7.2. Ubiquitylation and E3 ligases in the development of diseases

Ubiquitylation regulates the activity of proteins associated with DNA repair, tumor suppressors, and oncogenes. Impaired ubiquitylation leads to deregulated cellular pathways that are implicated in improper assembly and functioning of protein networks, building up misfolded proteins, mislocalization of proteins, and oncogenesis. Data generated over the decades has highlighted the role of ubiquitylation machinery and E3 ligases in tumorigenesis and tumor suppression.^{147–150}

There are three types of E3 ligases: RING, HECT and RING-between-RING (RBR) E3s. BRCA1 and BARD1 are RING type E3 ubiquitin ligases. Mutations in the E3 ligase c-Cbl are associated with colorectal cancer, myeloid neoplasms, and myelodysplastic syndromes.¹⁵¹ Another E3 ligase VHL has been linked to stabilization of HIF-1 α , that facilitates tumor vascularization. In usual conditions, HIF-1 α is degraded by the proteasome, but in oxygen-deprived conditions, HIF-1 α cannot be degraded and promotes angiogenesis, providing tumor cells with nutrients and oxygen.^{152,153} NF-kB pathway is associated to cancer promotion by hampering programmed cell death. One example of how NF-kB signaling pathway inhibits apoptosis and promotes inflammation of gastric tracts and lungs, is its involvement in mucosa-associated lymphoid tissue lymphoma (MALT). The protease activity of MALT regulates NF-kB signalling and the protease activity of MALT is dependent on its mono-ubiquitylation. Ubiquitylation preventing mutations in the MALT protein results in improper NF-kB signalling and decreases survival of cells.^{154,155}

Protein aggregation and accumulation of misfolded proteins are believed to be the cause of neurodegenerative diseases, such as Alzheimer's and Parkinson's. The unique characteristic of Parkinson's is the aggregation of α -synuclein protein, also called the Lewy bodies, due to mono-ubiquitylation or di-ubiquitylation at various lysines. Based on the lysine residue modified, ubiquitylation of α -synuclein could either promote or inhibit the formation of α -synuclein fibrils.¹⁵⁶ The aggregation of tau proteins is due to deficient poly-ubiquitylation, proper poly-ubiquitylation of the tau proteins believed to be associated with Alzheimer's disease would lead to efficient degradation of the misfolded and aggregated proteins by the proteasome.^{157,158} A very interesting and medically important complex, the BRCA1-BARD1 is a heterodimer conferring E3 ligase activity.¹⁵⁹ BRCA1 proteins are essential for the maintenance of chromosomal integrity and they function at the site of DNA damage.¹⁶⁰

1.7.3. Ubiquitin

The conjugation of a ubiquitin moiety to the substrate is termed 'Ubiquitylation'. Ubiquitin is a small molecule of 76 amino acids, highly conserved stable protein with β -grasp fold and six residues at the C-terminus that are highly flexible. Its remarkable degree of conservation from plants to yeast to humans confers multiple functionalities.^{161,162} Ubiquitin's seven lysines and its N-terminus are the linkage sites for ubiquitylation. The seven lysine residues are spread throughout the molecule and point in different directions. Ubiquitin interacting surface is made of sixteen to seventeen hydrophobic residues, the first loop comprising L8 is prone to multiple conformational changes enabling association with diverse ubiquitin-binding proteins. I44 hydrophobic patch consisting of residues I44, L8, V70, and H68, interacts with the proteasome and ubiquitin-binding proteins. Third hydrophobic patch consisting of I36, L71, and L73 in the tail region of ubiquitin can modulate interactions between ubiquitin

moieties and is identified by E3 ligases, deubiquitinases, and ubiquitin-binding proteins. There is another critical hydrophobic patch termed as F4 patch comprising of Q2, F4, and T12 that has been reported to interact with ubiquitin-specific protease (USP) domain in deubiquitinases.^{163,164} Ubiquitin itself can be modified by small modifications such as phosphorylation and acetylation.^{165,166} Modifications of ubiquitin are reported to contribute to additional regulatory mechanism in the ubiquitylation system.¹²⁸ Structure of ubiquitin determined by X-ray crystallography is shown in Figure 14.

Figure 14: Crystal structure of ubiquitin (PDBID: 1UBQ).¹⁶²

A) Structure of ubiquitin with lysines (orange) associated with chain linkages. B) Structure of ubiquitin with residues (green) interacting with ubiquitin-binding proteins and other partners.

1.7.4. E2 conjugating enzymes

Humans have approximately 35 E2 conjugating enzymes that coordinate the transfer of ubiquitin or ubiquitin-like moieties (e.g., SUMO) in cells. In cells, E2 conjugating enzymes predominantly exist as E2-Ub complex.^{167,168} In HEK293 cells, Ub is estimated to be in the molar concentration of ~85 μ M.¹⁶⁹ Most E2 conjugating enzymes have a single domain called the ubiquitin-conjugating (UBC) domain; few

other E2s have short extra sequence either at the N-terminus or C-terminus or both, imparting additional functionality. UBC domain comprises four α -helices and four β -sheets with some of the loop region possessing the E3 ligase binding site. The short extra sequences at the N-terminus and C-terminus are mostly disordered.

Figure 15: Solution NMR structure of UbchH5c (PDBID: 2FUH).¹⁷⁰

α and β sheets forming the UbchH5c is labeled. The binding site for E1 activating enzyme and E3 ubiquitin ligase are shown.

E2 conjugating enzymes show a preference for free lysines to transfer ubiquitin and has been reported to function with different types of E3 ligases.^{167,171} Crystallography and solution NMR studies identified a canonical E3 binding surface, formed by the residues in $\alpha 1$, and the loop connecting $\alpha 1$ and $\beta 4$. The characterized binding surface has been shown to collaborate with both RING and HECT E3 ligases, and some residues in the binding surface serve as a platform for E1 enzyme binding. Almost all E2-E3 interactions are characterized to be weak or moderate.^{172–174}

When a ubiquitin moiety is covalently conjugated to the E2 conjugating enzyme, the remarkable flexibility of the residues at the C-terminus of ubiquitin (residues 72-76) permits multiple conformations concerning the E2 structured domain. Structures

determined by solution NMR experiments showed open and closed conformations of the E2-Ub complexes, the interactions involving the $\alpha 2$ helix residues in the E2 and I44 residue in the ubiquitin during the closed conformation stabilize the complex.¹⁷⁵ To facilitate the transfer of ubiquitin and further stabilization of the complex in the closed conformation, the backside of the E2 conjugating enzyme interacts with the ubiquitin with low affinity. When E2 is conjugated to a ubiquitin moiety, binding of another ubiquitin moiety to the backside of E2 enhances the affinity of the E2-Ub conjugate E3s.^{170,176} E2s like Ube2N, Ube2R1, and Ube2S can add ubiquitin moieties only to substrates to which a ubiquitin moiety is already conjugated. This is because they can transfer their conjugated ubiquitin to another ubiquitin molecule only (i.e.,) one E2 initiates the process by adding a ubiquitin molecule and a second E2 extends the ubiquitin chains (e.g., APC/C E3 complex during the regulation of cell cycle).¹³⁶

A)

B)

C)

Figure 16: Placement of ubiquitin in E2 (light orange)-Ub (blue) complex.

A) Cartoon showing open and closed conformations of the complex. B) Crystal structure of an E2-Ub complex in its open conformation with the I44 surface exposed to the solvent (PDBID: 3A33). C) Crystal structure of an E2-Ub complex in its closed conformation upon binding of the RING domain containing E3 ubiquitin ligase (PDBID: 4AUQ). All the three representations are the same structures in different orientations.

1.7.5. E3 ubiquitin ligase

E3 ubiquitin ligases recruit E2 conjugating enzymes that are conjugated with ubiquitin, bind to a target and coordinate the transfer of ubiquitin from the E2 enzyme to target substrate. E3 ligases can bind to both E2 conjugating enzyme and the substrates at a given time and are the prominent origin of substrate specificity in the process of ubiquitylation.¹⁷⁷⁻¹⁷⁹ E3s and E2s work hand in hand when determining the linkage specificity and processivity of polyubiquitin chains.

Studies over decades led to the understanding that the RING domain-containing proteins are associated with ubiquitin ligase activity.^{180,181} There are estimated over 300 genes in humans that encode for RING domain containing E3 ligases, and are implicated

in many DNA dependent cellular processes and diseases.^{182,183} RING domain E3 ligases constitute the clear majority of E3 ligases and do not participate chemically in the transfer of ubiquitin. RING domain E3s bind to the substrate and an E2-Ub complex, coordinating the transfer of ubiquitin directly from the E2 active site to the substrate. Examples of RING E3s are Anaphase-promoting complex (APC), BRCA1 and Skp1-CUL1-F-box protein (SCF) complex.^{3,184,185}

Unlike RING domain E3s, Homologous to the E6-AP Carboxyl Terminus (HECT) and RING between RING (RBR) E3 ligases assemble E3-Ub intermediates before transferring the ubiquitin to the substrate. HECT E3s have a bilobed (N-lobe and C-lobe) catalytic domain at the C-terminus connected by a short hinge. Unlike RING E3-E2 conjugates, HECT C-terminus determines the linkage specificity and facilitates poly-ubiquitylation of the substrate with specific poly-ubiquitin chain linkages.¹⁸⁶ (e.g.,) Rsp5 catalyzes K63 linked polyubiquitin chains whereas E6AP catalyzes K48 linked polyubiquitylation, by replacing the 62 amino acids of the Rsp5 protein with the corresponding amino acid sequence in E6AP, the difference in the linkage specificity was abolished, and Rsp5 was able to promote K48 polyubiquitylation.¹⁸⁶ RBR E3 ligases are characterized by two RING domains (RING1 and RING2), and a central zinc-binding domain in-between-RINGs (IBR). RBR E3 ligases were discovered when they were found to form a complex with UbcH7. But RBR E3 ligases could not directly transfer ubiquitin molecules to their substrates, instead formed RBR-Ub intermediates. Subsequently, Klevit and group discovered that cysteines in the conserved RING2 domain of RBR E3s had the potential to attack and discharge UbcH7 (E2)-Ub conjugates, and formed a reversible thioester intermediate with ubiquitin.^{187–189} Examples of RBR E3 ligases are PARKIN and Linear ubiquitin chain assembly complex (LUBAC). To avoid inappropriate auto-ubiquitylation and non-specific ubiquitylation

of the substrates, post-translational modifications (PTMs) of E3 ligases activate them only when required, at other times, they are maintained in a dormant state. Post-translational modifications result in wide range of conformational changes of E3 ligases themselves and their functions in cells. Post-translational modification of E3 ligases includes phosphorylation, ubiquitylation and PARylation.^{190–192}

Phosphorylation is the most common post-translational modification involved in the regulation of enzymatic activity. Examples of phosphorylation activated E3 ligases are the c-Cbl RING E3 ligase, mouse double minute 2 (MDM2) and NEDD4 HECT E3s. Post-translational modifications of E3 ligases facilitate their catalytic efficiency by several hundred folds.^{193–195} Autoubiquitylation of E3 ligases can signal for self-destruction through proteasomal degradation or can moderate the ligase activity directed towards its substrates.^{196–198}

1.7.5.1. BRCA1 and BARD1

Breast cancer is the most common disease and the leading cause of cancer-associated deaths in women, around the globe.¹⁹⁹ The BRCA1 gene mapped to chromosome arm 17q and BRCA2 mapped to chromosome arm 13q, are both associated with early onset of breast cancer.^{200,201} Malfunctioning of the BRCA1 or reduction in the expression of BRCA1 protein increases the risk of breast, ovarian, and different cancer types. One of the unique characteristics of BRCA1-mutated breast cancer is that 15% of such cancers are ‘triple negative’ (i.e.,) the tumor cells are negative for estrogen, progesterone, and human epidermal growth factor receptor 2 (HER2). This unusual characteristics of the breast cancer tumor cell make them less viable for hormonal therapies and drugs that target estrogen, progesterone, and HER2 receptors.²⁰²

In 1990, Mary-Claire King's group published a seminal paper, reporting the presence of BRCA1 gene on chromosome 17.²⁰¹ Mary-Claire King's finding instituted the competition to clone and sequence the BRCA1 gene, eventually Mark Skolnick's group and an American molecular diagnostic company named "Myriad genetics" winning the race in 1994.²⁰³ The BRCA1 gene consists of 24 exons, out of which 22 are coding, and 2 are non-coding.²⁰³ BRCA1 protein predominantly exists as a heterodimer with a RING domain-containing protein, BRCA1-Associated RING Domain 1 (BARD1).²⁰⁴ BARD1 shares sequence homology with the RING domain and BRCT domains of BRCA1. In 1996 Richard Baer's group performed mammalian two-hybrid assay and discovered BARD1 interacts with BRCA1 *in vivo*.²⁰⁵ Although not as prevalent as BRCA1 mutations, germline mutations in BARD1 have also been reported in breast, ovarian, and other types of cancer.^{206,207} BRCA1-BARD1 heterodimer formation is associated with DNA damage signaling and DNA repair by homologous recombination.^{132,208–210}

The BRCA1 protein is 1863 amino acids in length, molecular weight of 207.72 kDa. The BRCA1 protein consists of different domains: a RING domain (Zinc finger, C3HC4, amino acids 20-64) at the N-terminus, a Serine-rich domain in the middle, nuclear localization signaling region, and two BRCT repeats at the C-terminus. BARD1 protein is 777 amino acids in length, having a molecular weight of 86.64 kDa. The BARD1 protein consists of a RING domain at the N-terminus, ankyrin repeats in the middle and two BRCT repeats at the C-terminus. BRCA1 are phosphoprotein binding regions that share structural homology with proteins in the DNA repair and cell cycle regulation pathway.²¹¹ The domain architecture of BRCA1 and BARD1 proteins are illustrated in Figure 17.

Figure 17: Domain architecture of BRCA1 and BARD1.

BRCA1 and BARD1 proteins share a similarity in the domains as they both comprise of RING domains at the N-terminus and two BRCT domains at the C-terminus. RING represents zinc RING finger domain; NLS represents nucleic acid localization signal; BRCT indicates BRCA1 associated C-terminus; ANK represents ankyrin motif.

The RING domain is a Cysteine and Histidine-rich motif that binds two zinc atoms (Figure 18). RING domains of BRCA1 and BARD1 alone are not sufficient to form stable heterodimeric structure, instead larger regions containing residues 1-109 of BRCA1 and 26-119 of BARD1 is essential.²¹² RING domain of BRCA1 and BARD1 contain mutations involved in diseases and E3 ubiquitin ligase activity.²¹³⁻²¹⁵ The structure of the RING domain of BRCA1 consists of three-stranded short β -sheet, zinc binding loop, and an α -helix. RING domain of BARD1 is structurally homologous to the RING domain of BRCA1, except it misses the central α -helix. Stretch of amino acids (8-22) before the RING domain and the stretch of amino acids (81-96) after the RING domain in BRCA1 form helices that are antiparallel and are involved in the dimerization with BARD1 (Figure 18). Similarly, in the structure of the RING domain in BARD1, stretch of amino acids (36-48) before the RING domain and amino acids (101-116) after the RING domain form antiparallel helices that are involved in the interaction with BRCA1. These four antiparallel helices form the hydrophobic interface for heterodimerization between BRCA1 and BARD1. The amino acid sequence flanking the RING domain of BRCA1 is necessary for recognizing and binding other proteins in

cells.² Solution NMR structure of minimal RING domain containing BRCA1-BARD1 complex is shown in Figure 18.

Figure 18: NMR structure of BRCA1-BARD1 heterodimer (PDBID: 1JM7).²

The structure of the N-terminal region (1-103 amino acids) of BRCA1 (highlighted in 'green') and the N-terminal region (26-122) of BARD1 (highlighted in 'blue'). Coordinated zinc atoms are shown in red (BRCA1) and yellow (BARD1).

Many Cysteine residues (C24R, C39S/Y, C44F, C47G/F, C61G and C64G/R/Y) are mutated in the RING domain of BRCA1 that expose individuals to breast and ovarian cancer.^{2,212} In addition to the mutations in the RING domain, another set of mutations are seen throughout the length of N-terminal region of BRCA1 (1-103) that are identified in patients with breast or ovarian cancer patients. Several of the mutations mentioned here are placed on the surface of the protein involved in the heterodimerization of BRCA1 and BARD1.² The BRCA1-BARD1 interaction and their stability as a heterodimer is vital for their role in homologous recombination, tumor suppression and E3 ubiquitin ligase activity.^{2,159} The BRCA1-BARD1 RING-RING

heterodimer is very stable, present at concentrations lower than 10^{-7} M.^{2,216} Mutation of the residues in the RING domain of BRCA1-BARD1 heterodimer are shown in Figure 19.

Figure 19: Mutations in the RING domain of BRCA1-BARD1 heterodimer (PDBID: 1JM7).²

A) Structure showing RING domains of BRCA1 and BARD1, RING domains are highlighted in 'red'. B) Structure illustrating the mutations that are present in the RING domain of BRCA1. C) Structure depicting the mutations in the N-terminal region of BRCA1.

The RING domain in BRCA1 confer the E3 ligase activity, interacting with E2 conjugating enzymes, facilitating ubiquitylation. The disease-associated mutations disrupt the interaction between the BRCA1 and the E2 conjugating enzyme, resulting in inefficent ubiquitylation of the substrates.²¹⁷ Though BARD1 contains a RING domain, it possesses very low E3 ligase activity by itself. However, BARD1 when forming a stable heterodimeric complex with BRCA1, shows high activity.¹⁵⁹ BRCA1-BARD1 can moderate both mono- and poly-ubiquitylation depending on the E2 conjugating enzyme.^{218,131,219} The discovery that the BRCA1-BARD1 complex functions in ubiquitylation of the histones opened up a new aspect of post-translational signalling.^{3,220} BRCA1-BARD1 exhibits E3 ligase activity to other putative substrates, namely ER α , TFIIE, RNA polymerase II (RPB1), and Ct1P.^{221–225}

In order to determine which lysine residue in ubiquitin is involved in the BRCA1-BARD1 mediated poly-ubiquitylation of substrates, one of the seven lysines present in ubiquitin was mutated with arginine one at a time. It was found that the poly-ubiquitylation was abrogated in the presence of Ub-K6R. Along with this finding, it was also found that BRCA1-BARD1 dependent poly-ubiquitylation was reduced to an extent when K48 and K63 were mutated with arginine.¹³¹ Experiments confirmed that though BRCA1-BARD1 preferred catalyzing K6 linked poly-ubiquitylation, it could also promote poly-ubiquitylation through other lysines.¹³¹ When BRCA1-BARD1 RING heterodimer was paired with different E2 conjugating enzymes, BRCA1-BARD1-E2 pairs resulted in distinct ubiquitylation products, that included mono-ubiquitylation and specific poly-ubiquitin chains.²¹⁸ It was shown by site-directed mutagenesis and solution NMR that, RING domain of BRCA1 heterodimer directly binds to the E2 conjugating enzyme (UbcH5c), without BARD1 physically participating in the interaction. Most of the E2s directly bind to the RING domain of BRCA1, but not all (e.g., Ubc9).²¹⁸ Though BARD1 does not physically interact with E2, BRCA1-BARD1 heterodimeric assembly is crucial for E3 ligase activity.¹⁷⁴

Figure 20: BRCA1 interacting residues in helix 1 and loop 2 of UbH5c.

Surface residues that directly interact with BRCA1 are labeled with an asterisk and highlighted in bold. Amino acids that are conserved among E2s and directly interact with BRCA1 are highlighted in grey.

1.7.5.2. BRCA1-BARD1 associated genome regulation complex

Many external and internal factors contribute to cellular DNA damage, the DNA damage accumulate over time leading to genomic instability and tumor progression.²²⁶ Genomic instability is a common trait in cancer cells.²²⁷ The integrity of chromatin during cell cycle progression within the cells are monitored by cell cycle regulators and, upon DNA damage, the cell cycle regulators activate the checkpoint to permit the cells with sufficient time to repair damaged DNA. Following DNA damage, activation of cell cycle checkpoint is critical for maintaining genomic stability as it protects the duplication of damaged DNA and transfer of lesioned DNA to the dividing cells. Defective proteins involved in the cell cycle regulation leads to developmental abnormality, genomic instability, and disease prognosis.^{228,229}

Different DNA repair mechanisms include nucleotide excision repair (NER), base excision repair (BER), mismatch repair (MMR), non-homologous end joining (NHEJ), and homology-directed repair (HDR).^{230–235} The first clue that BRCA1 was associated with DNA damage repair came from localization data of BRCA1 protein present at DNA damage sites.²³⁶ It is established that BRCA1 plays an essential role in genome integrity and cancer.^{237,238}

BRCA1 is a multi-functional protein that is associated with diverse cellular processes, including transcription, DNA repair, cell cycle regulation, protein ubiquitylation, chromatin remodeling, and apoptosis.^{3,239–244} There are different phases in a cell cycle, and BRCA1 is one of the proteins that play an essential role in all these phases.^{245,246} BRCA1 forms and functions as a part of the sizeable multi-protein complex (BARD1, BRCC36, BRCC45, NDA1, and more) that targets the sites of DNA

damage to maintain genome integrity. Illustrations of the BRCA1 associated multi-protein complexes are shown in Figure 21, Figure 22, & Figure 23.

Figure 21: Diagram illustrating the subunits of the BRCA1-A complex.

Figure 22: Diagram illustrating the subunits of the BRCA1-B complex.

Figure 23: Diagram illustrating the subunits of the BRCA1-C complex.

A study conducted by Inder Verma's group, lead to an exciting discovery showing that BRCA1 E3 ligase activity contributes to the maintenance of the genome by ubiquitylating histone H2A at the pericentric heterochromatin, consequently inhibiting transcription of pericentric α -satellite DNA repeats.²²⁰ Absence or mutated BRCA1 is associated with the de-repression of α -satellite DNA repeats, in turn resulting in chromatin decompaction and decrease in the levels of histone H2A ubiquitylation.²²⁰ Transcription of α -satellite DNA repeats leads to deficient homologous recombination, growth arrest, and malfunctioning of other DNA dependent cellular processes, altogether leading to genomic instability. It is confirmed that the BRCA1 function as E3 ubiquitin ligase is essential for its functioning as caretaker in the DNA repair machinery and maintaining genomic stability, but its mechanism in tumor suppression is still an interesting topic for debate. Like BRCA1, Polycomb repressor complexes (PRC1) are E3 ubiquitin ligases, whose characterization revealed that the loss or depletion of the ligase in cells lead to the reduction in ubiquitylated H2A. In the perspective of histone ubiquitylation, histone ubiquitylation is implicated in the transcriptional repression by modulating chromatin structures²⁴⁷, regulating proteins involved in the transcriptional machinery²⁴⁸, and acting as a signal for other histone modifications.²⁴⁹ PRC1 mediated ubiquitylation of histone H2A is involved in X-chromosome inactivation, Polycomb silencing, and silencing of HOX genes.^{248,250} H2A ubiquitylation is implicated in two other important functions, repression of BCL6 target genes and inhibition of MLL3-orchestrated di- and tri-methylation mediated transcriptional repression.^{249,251} In the proximity of DNA lesion, RING2 mediated mono-ubiquitylation of histone H2A is obligatory for nucleotide excision repair.²⁵²

In order to map the exact residues ubiquitylated by BRCA1-BARD1 E3 ligase, Kevin Hiom's group conducted *In vitro* and *In vivo* experiments focusing on the

interaction of BRCA1-BARD1 with chromatin. Ubiquitylation assay performed in the presence of BRCA1-BARD1 and recombinant *Xenopus laevis* individual histones, nucleosomes, and chromatin extracted from HeLa cells, confirmed that BRCA1-BARD1 specificity to histone H2A is only in the context of nucleosomes and chromatin.³ However, it is important to remember that BRCA1-BARD1 ubiquitylates individual histones H2A, H2B, H3 and H4 with no probable propensity.²⁴²

In order to confirm the specificity for H2A *In vivo*, the mCherry fluorescent reporter protein fused BARD1²⁶⁻¹²⁶-GGSGG-BRCA1²⁻¹⁰⁹ (BDfBC) product was expressed in U2OS 2-6-3 cells and, it resulted in the exhibition of mCherry colocalization with ubiquitylated H2A (H2AUb). In addition, increase in the accumulation of H2AUb was seen. This data proved that H2A is the substrate. To determine the exact residues that are ubiquitylated in nucleosomal H2A *in vitro*, mutation experiments were performed in the presence of *Xenopus laevis* nucleosomes. There was a decrease in the ubiquitylation when K124, K127 and K129 of the H2A C-terminal tail were mutated, suggesting K124, K127 and K129 are the potential sites for ubiquitylation by BRCA1-BARD1.³

DT40 cells stably expressing wild type or mutant forms of FLAG-H2A were used to identify the specific lysines that are ubiquitylated by BRCA1-BARD1 *in vivo*. Immunoprecipitation with anti-FLAG and western blot analysis targeting H2AUb suggested K125, K127, and K129 are involved in the ubiquitylation by BRCA1-BARD1. However, *in vivo* experiments did not establish if one or more of the three lysines (K125, K127 and K129) are ubiquitylated. Isolation of H2AUb from HEK293 cells expressing BDfBC, pepsin digest, analysis by mass spectrometry and above mentioned *in vivo* experimental data combined, the results suggest K127 and K129 are

the potential sites of ubiquitylation *in vivo*. However, *in vitro* results suggest K124, K127 and K129 (*Xenopus laevis*) are the ubiquitylation sites.³

1.8. Objective of the study

Approximately 5-15% of all histone H2A in eukaryotic cells is modified with ubiquitin. In cells H2AUb coincides with highly compacted heterochromatin and has been associated with gene silencing. In recent years, many H2A-specific ubiquitin ligases (E3) have been identified that link H2AUb to gene silencing, X-chromosome inactivation and DNA repair. A key E3 ubiquitin ligase involved in DNA repair is the tumour suppressor BRCA1. In cells, BRCA1 forms a heterodimeric complex with BARD1 that specifically modify K127, and K129 in the C-terminal tail of histone H2A.

The seminal findings by Kevin Hiom's group in collaboration with Professor Daniela Rhodes, while they were colleagues at the MRC Laboratory of Molecular Biology in Cambridge motivated us to investigate the structural mechanism involved in the functioning of the BRCA1-BARD1 E3 ubiquitin ligase to ubiquitylate histone H2A. In recent years, many high-resolution structures of chromatin remodeler bound NCP complexes have been solved. However, there is only one structure of E3 ligase (RING1B-BMI1) in complex with NCP that has been solved and published in 2014. BRCA1-BARD1 has two most important functions: 1) as an E3 ubiquitin ligase, and 2) as a chromatin remodeler. The main aim is to reveal through biochemical and structural analysis how BRCA1-BARD1 recognizes and binds to the nucleosome core particle to specifically ubiquitylate K125, K127, and K129 in the C-terminal tail of histone H2A. I am interested in determining the structure of the NCP-BRCA1-BARD1 complex. Elucidating the structure of NCP-BRCA1-BARD1 complex will provide a structural and mechanistic view of how a histone specific ubiquitin ligase act on the NCP substrate at the molecular level. The understanding of the E3 ubiquitin ligase functioning to install histone modifications at the molecular level has therapeutic implications.

Chapter 2: Materials and methods

“Nothing in life is to be feared, it is only to be understood. Now is the time to understand more, so that we may fear less.”

- Marie Curie

2.1. General Materials and Methods

2.1.1. Plasmids

Plasmids containing gene inserts are listed below, all the plasmids were produced following Ligation Independent Cloning protocol.

Table 1: Name of plasmids with corresponding tagged inserts, and encoded antibiotic resistance

Plasmid	Gene Insert	Tag	Resistance	Expression system
pST50Trc1	BRCA1 (1-103)	N' Strep II-His ₁₀	Ampicillin	<i>E. coli</i>
pST50Trc2	BARD1 (26-130)	Untagged	Ampicillin	<i>E. coli</i>
pST50Trc1	BRCA1 (1-250)	N' Strep II-His ₁₀	Ampicillin	<i>E. coli</i>
pST50Trc2	BARD1 (14-186)	Untagged	Ampicillin	<i>E. coli</i>
pET30a	BARD1 (27-186)	N' His ₆	Kanamycin	<i>E. coli</i>
pET15b	BARD1 (1-186)	N' His ₆	Ampicillin	<i>E. coli</i>
pST50Trc3	UbcH5c	His ₁₀	Ampicillin	<i>E. coli</i>
pST44	BRCA1 (1-103), BARD1 (26-130)	N' Strep II-His ₁₀ (BRCA1), Un- tagged (BARD1)	Ampicillin	<i>E. coli</i>

pST44	BRCA1 (1-250), BARD1 (14-186)	N' Strep II-His ₁₀ (BRCA1), Un- tagged (BARD)	Ampicillin	<i>E. coli</i>
pFL	BRCA1 (1-250)	N' Strep II	Gentamicin	SF9
PUCDM	BARD1 (14-186)	N' His ₁₀	Chloramphenicol	SF9
pFL	BRCA1 (1-263)	N' Strep II	Gentamicin	SF9
pFL	BRCA1 (1-1580)	N' Strep II	Gentamicin	SF9
pFL	BRCA1 (1-1745)	N' Strep II	Gentamicin	SF9
pFL	BRCA1 (1-1863)	N' Strep II	Gentamicin	SF9
pUCDM	BARD1 (1-777)	N' His ₁₀	Chloramphenicol	SF9

Table 2: Antibiotics, media, and reagents

Antibiotics/media/reagents	Stock composition	Final composition
Lysogeny Broth	-	1% tryptone, 0.5% yeast extract, 0.5% NaCl, pH 7.5
2xYT	-	1.6% tryptone, 1% yeast extract, 0.5% NaCl, pH 7.3
SF900-III	see https://www.thermofisher.com	-
Grace's media	see https://www.thermofisher.com	
Ampicillin	100 mg/mL	100 µg/mL
Chloramphenicol	34 mg/mL	34 µg/mL
Kanamycin	50 mg/mL	50 µg/mL
Gentamicin	7 mg/ml	7 µg/mL

Tetracycline	10 mg/mL	10 µg/mL
Antibiotic-Antimycotic	100x	1x
IPTG	1 M	0.1 - 0.5 mM
FBS	100%	2%
X-Gal	40 mg/mL	40 g/mL

2.1.2. Ligation Independent Cloning

A double-strand-specific Exonuclease III-based ligation independent cloning strategy was employed to generate plasmids for expressing protein constructs of interest.²⁵³ First, primers were designed in such a way that both 5' and 3' end of the primer comprises of a 20-nucleotide sequence complementary to the ends of the linearized vector followed by a sequence that is complementary to the gene of interest. Primers were designed to have a melting temperature of 60°C. Classically, the gene of interest was amplified by PCR with a pair of forward and reverse oligonucleotides and extracted from the gel using the *GeneJet Gel Extraction Kit* (Thermo Fischer Scientific). 250 ng of linearized vector and insert were prepared in a NEB1 (New England Biolabs) buffered solution in a final volume of 13.5 µl and was stored at 4°C. To the solution containing linearized vector and the insert, 10 U/µL Exonuclease III (NEB) was added, and the reaction was incubated at 14°C for 60 sec, quenched by the addition of 1.5 µL 0.5 M EDTA. The sample was snap frozen in liquid nitrogen. Annealing of the linearized vector and insert was accomplished by running an annealing program in a thermal cycler (Applied Biosystems).

2.1.3. Preparation of proteins and protein complexes

Co-expression of proteins in *E. coli* was accomplished by using pST44 polycistronic expression system.²⁵⁴ Co-purification of the N-terminal STR-His₁₀-

BRCA1¹⁻¹⁰³ with untagged-BARD1²⁶⁻¹³⁰ (hereafter called BRBA-S) was based on the protocol described by Song Tan's group,⁶⁴ with changes mentioned below.

N-terminal STR-His₁₀-BRCA1¹⁻¹⁰³, untagged-BARD1²⁶⁻¹³⁰, STR-His₁₀-BRCA1¹⁻²⁵⁰, and untagged-BARD1¹⁴⁻¹⁸⁶ were cloned into pST50Trc transfer vectors and subsequently subcloned into the pST44 polycistronic plasmid for co-expression. The BRBA-S heterodimeric complex and STR-His₁₀-BRCA1¹⁻²⁵⁰ with untagged-BARD1¹⁴⁻¹⁸⁶ complex were co-expressed in BL21 (DE3) RIPL *E. coli* strain at 18°C. Cells expressing the complex were resuspended in lysis buffer (100 mM Hepes pH 8.0, 500 mM NaCl, 10% glycerol, 20 mM imidazole, 10 mM β-mercaptoethanol, 10 μM ZnCl₂, Benzonase, Protease Inhibitor Cocktail), and lysed using LM20 (*Microfluidics Corporation*). The lysate was clarified by centrifugation at 44, 000 g for 30 min. Clarified lysate was passed through Ni-NTA affinity (GE Healthcare) column, column was washed with wash buffer 1 (50 mM Hepes pH 8.0, 500 mM NaCl, 10% glycerol, 40 mM imidazole, 0.1% Triton, 10 mM β-mercaptoethanol), followed by wash buffer 2 (50 mM Hepes pH 8.0, 10% glycerol, 60 mM Imidazole, 10 mM β-mercaptoethanol). After the wash step, the sample was eluted by a linear gradient in elution buffer (50 mM Hepes pH 8.0, 150 mM NaCl, 10% glycerol, 750 mM imidazole, 10 mM β-mercaptoethanol). Fractions were analyzed in 12% Bis-Tris gel. Pure fractions were pooled together, concentrated and was loaded onto Superdex 75 16/600 or Superdex 200 16/600 (GE Healthcare) size-exclusion chromatography column. Protein complex was collected in elution buffer (50 mM Hepes pH 8.0, 150 mM NaCl, 10% glycerol, 2 mM TCEP) and was analyzed in 12% Bis-Tris gel. Chosen samples were pooled together, concentrated, and stored at -80°C. Protocol for purification of STR-BRCA1¹⁻²⁵⁰-His₁₀-BARD1¹⁴⁻¹⁸⁶ was like that of BRBA-S, except an intermediate heparin purification step. All the cloning, co-expression and purification of BRCA1 and BARD1 constructs were

done by me in our lab, except cloning of Strep-BRCA1¹⁻²⁵⁰-His-BARD1¹⁴⁻¹⁸⁶ (hereafter called BRBA-L) for insect cell expression. BRBA-L was cloned into the MultiBac system by Dr. Alessandra Beltrami, who was a post-doctoral fellow in Daniela Rhodes' lab. Though I cloned the N-terminal Strep-II tagged full-length BRCA1 and N-terminal His₁₀ tagged full-length BARD1 and co-expressed in Sf9 cells, I was aware that the full-length BRCA1 and BARD1 are prone to degradation due to the presence of disordered core. Since N-terminal tagged full-length BRCA1 and BARD1 are challenging to purify and was prone to degradation, I attempted to co-purify C-terminal tagged full-length BRCA1 and C-terminal tagged full-length BARD1, obtained from Professor Kevin Hiom.

wt-Ubch5c was expressed in BL21(DE3) RIPL at 18°C overnight. Small-scale expression test for wt-Ubch5c showed good expression. Briefly, cells from large-scale expression were lysed in lysis buffer (30 mM MES pH 6.0, 50 mM NaCl, 1 mM EDTA, 2 mM DTT), followed by centrifugation to clarify the supernatant. The clarified supernatant was filtered through 0.45 µm membrane, and as mentioned before, since small-scale expression tests showed fat bands, the clarified lysate was directly passed through SP sepharose (GE Healthcare) cation exchange chromatography column and the bound proteins were eluted in elution buffer (30 mM MES pH 6.0, 1 M NaCl, 1 mM EDTA, 2 mM DTT). Quality of the eluted fractions was analyzed in 12% Bis-Tris gel. Pure fractions were pooled together, concentrated, and subjected to size-exclusion chromatography on Superdex 75 16/600 (GE Healthcare) column. Buffer exchange was achieved during size-exclusion chromatography, and the eluted fractions were analyzed in 12% Bis-Tris gel. Sample was stored in storage buffer (50 mM Hepes pH 7.0, 150 mM NaCl, 2 mM DTT) at -80°C. Milligram quantities of Ubch5c was obtained, and the enzyme was used for conducting biochemical and structural investigations.

2.2. Nucleosome core particles

2.2.1. Expression and purification of human histones

Figure 24: Schematic illustration of the process by which the histone octamer is assembled.

Expression and purification of histones were based on the protocol devised by Lynda Chapman (Medical Research Council Laboratory of Molecular Biology, Cambridge, UK) and previously established protocols.²⁵⁵ Plasmids encoding histone H2A, H2B, H3, and H4 genes were a generous gift from Curt Alexander Davey (Nanyang Technological University, SBS).

Individual histones were transformed into Rosetta (DE3) pLysS *E. coli* strain and plated on LB-agar plates. Plates containing transformed cells were incubated at 37°C overnight. Pre-inoculum was cultivated overnight with one single colony from the LB-agar plate, and a specific volume of the pre-inoculum was used for initiating a large-scale culture. Large-scale cultures were grown at 37°C in 2xTY media with 100 µg/mL ampicillin and 34 µg/mL chloramphenicol. At OD₆₀₀ of 0.6-0.8, the culture was induced by adding IPTG to a concentration of 0.4 mM and left in a shaker for approximately 4 hours. Cells were harvested by centrifugation at 4,000 rpm for 30 minutes at 4°C.

The cell pellet from 6-liter culture was resuspended in lysis buffer (50 mM Tris-HCl pH 7.6, 100 mM NaCl, 1 mM EDTA, 5 mM β-Mercaptoethanol, 1% Triton, and protease inhibitor) and the cells were lysed using LM20 (*Microfluidics Corporation*). The lysate was clarified by centrifugation at 44,000 g for 20 minutes at 4°C. Inclusion body pellet containing histones were washed thoroughly. To solubilize and soften the inclusion bodies, DMSO was added and incubated for 30 minutes at RT. The softened pellet was resuspended in unfolding buffer (20 mM Tris-HCl pH 7.6, 6 M Guanidinium hydrochloride, and 2 mM DTT) and was kept rolling at RT for 1 hour so that the pellet dissolves completely. Cell debris was removed by centrifugation, and the unfolded histones present in the supernatant was isolated.

To further purify the individual histones, the highly positively charged histones were purified by ion-exchange chromatography. Histones were buffer exchanged from unfolding buffer to dialysis buffer (10 mM Tris-HCl pH 8.0, 0.1 M NaCl, 7 M urea, 1 mM EDTA, 2 mM DTT, and protease inhibitors). Urea solution used for purifications were subjected to deionizing mixed bed resin (*Sigma Aldrich*). Dialysed sample was passed through a 0.45 µm filter membrane to isolate large contaminants from the

solution containing histones. Solution containing histones was passed through tandemly connected HiPrep 16/10 Q Fast Flow that excluded DNA and negatively charged contaminants, and HiPrep 16/10 SP Fast Flow columns that have negatively charged matrix to capture histones (GE healthcare). Immediately before tandem purification, both the columns were equilibrated with 10% buffer B (histone H2A and H2B) or 20% buffer B (histone H3 and H4) depending on the histones.

Linear gradient purification was performed using buffer A (10 mM Tris-HCl pH 8.0, 7 M urea, 1 mM EDTA, 2 mM DTT, and protease inhibitors) and buffer B (10 mM Tris-HCl pH 8.0, 1 M NaCl, 7 M urea, 1 mM EDTA, 2 mM DTT, and protease inhibitors). Eluted fractions were analyzed by 18% SDS-PAGE. Fractions assessed to be pure were pooled and dialyzed against water containing β -Mercaptoethanol and protease inhibitors. The concentration of histones was determined and lyophilized. Lyophilized samples were stored at -80°C.

2.2.2. Large-scale production of Widom 601 DNA

The Widom 601 145 bp DNA template was a gift from Associate Professor Curt A. Davey in Nanyang Technological University, Singapore. The Widom 601 145 bp DNA was purified using the methods described previously.²⁵⁶ Widom 601 145 bp DNA was used to reconstitute NCPs for biochemical characterization and structural studies. A plasmid containing the Widom 601 DNA was amplified, isolated, and purified from *E. coli*. To remove RNA in the solution, RNase A was added and left to incubate at RT for a few hours or overnight. To isolate the plasmid DNA from fragmented RNA, PEG fractionation was carried out. Plasmid DNA in TE buffer was treated twice with Phenol/Chloroform to remove PEG.

The ethanol pre-incubated in -20°C freezer was added to the buffer containing plasmid DNA and was left to incubate at -80°C for 30 min. Ethanol precipitation step was performed twice, followed by drying of the plasmid DNA. Once dried, plasmid DNA was suspended in molecular biology grade water or buffer of choice. Plasmid DNA was ready for enzyme digestion to release the Widom 601 DNA.

Figure 25: Schematic illustration of the process by which the Widom 601 145bp DNA is produced.

2.2.3. Assembly of the histone octamer

The histone octamer was reconstituted following the protocol described by Lynda Chapman (Medical Research Council Laboratory of Molecular Biology, Cambridge, UK) and Luger.²⁵⁷ The protocol was used for the refolding of histone

octamers from lyophilized individual histones. Each of the four histones was resuspended in unfolding buffer (20 Tris-HCl pH 7.6, 6 M GuHCl, 5 mM DTT) to a concentration of 3 mg/mL and entirely dissolved by pipetting up and down without using vortexer. Tubes were left for rolling at RT for 30-40 minutes for the unfolding of the histones. Individual histones were mixed in equimolar ratio, except a 25% excess of both H2A, H2B, and the final protein concentration was adjusted to 1 mg/mL in the unfolding buffer. H2A and H2B were added in excess to prevent or minimize formation of H3-H4 (53.02 kDa) tetramers and facilitate the formation of the octamers. Surplus H2A-H2B dimer that has not been involved in the formation of the octamers could be removed during size exclusion chromatography.

To promote refolding of the histones, a sample containing the histone mix was dialyzed against 1 liter of refolding buffer (10 mM Tris-HCl pH 7.6, 2 M NaCl, 1 mM EDTA, and 5 mM β -Mercaptoethanol) with four changes including an overnight change in the third dialysis step. All the dialysis steps were performed at 4°C to prevent dissociation of the complex. Precipitated sample was removed by centrifugation at 44,000 g for 20 minutes and the supernatant containing histone octamer was concentrated to purify further using size exclusion chromatography.

A solution containing histone octamer was loaded onto a gel filtration column (HiLoad 16/600 Superdex 200 pg, *GE Healthcare*). Higher molecular weight aggregates eluted close to the void volume, histone octamer eluted between 65 to 70 ml, followed by the H2A/H2B dimer around 85 ml. It is recommended to have symmetric histone octamer peak that does not contain shoulders. Eluted fractions were analyzed on 18% SDS-PAGE gels. It was necessary to vigilantly visualize and confirm if the individual histones were stoichiometric. The sample was also analyzed on in-house prepared native

gels to establish the presence of one single band and no other. Fractions containing stoichiometric amounts of four histone proteins were pooled, and the sample was concentrated until the desired concentration was reached. A solution containing histone octamer was mixed with an equal volume of 100% glycerol to obtain the final concentration of 50% glycerol. The final concentration was determined spectrophotometrically (for a solution of 1 mg/mL, A_{276} is 0.45) and was stored at -20°C.

2.2.4. Preparation of NCPs

Protocols established by Lynda Chapman (Medical Research Council Laboratory of Molecular Biology, Cambridge, UK) and Luger²⁵⁷ were referred to reconstitute NCPs. For the preparation of NCPs, histone octamer, Widom 601 DNA, and other components necessary for the reconstitution of NCPs were added with maximum accuracy and in sequential order (H_2O , KCl/NaCl, Dithiothreitol, histone octamer, and 601 DNA). Inaccurate mixing of histones and DNA results in unusable, insoluble aggregates. Histone octamer to Widom 601 DNA ratio was determined by performing small-scale reconstitutions, and the ones that gave the best results were chosen to scale up. The concentration of histone octamer needed to saturate the complex was determined by keeping the Widom 601 DNA constant and titrating increasing concentrations of the histone octamer. Small-scale salt dialysis was performed in 1 kDa mini dialysis kit (GE Healthcare). Though small-scale reconstitutions were performed every single time, NCP samples were prepared, in general, the molar ratio of 1.05:1 provided quality samples. Briefly, Widom 601 DNA and the histone octamer were mixed in buffer containing 10 mM Hepes, pH 7.4, 2 M NaCl, 1 mM DTT and dialysis was performed at 4°C in steps. Salt concentration was reduced in steps from 2 M to 0 M to facilitate proper positioning of DNA to the histone octamer, resulting in the formation of nucleosome core. After salt removal by dialysis, the solution sample was

centrifuged at 18,000 g, 4°C for 20 minutes. Samples were analyzed on 5.5% native gel in 0.4X TB buffer. Gels were stained using SYBR safe or SYBR gold depending on the concentration of the sample loaded onto the gel. Only saturated NCPs were qualified for biochemical characterizations and structural studies.

Figure 26: Schematic diagram of the process by which the NCPs are produced.

2.3. Assays

2.3.1. Electrophoretic Mobility Shift Assay

In the classical way, a solution containing a mix of proteins-nucleic acids or protein-protein are subjected to electrophoresis by running the sample through polyacrylamide or agarose gel in native conditions. In common, protein-nucleic acid or protein-protein complexes migrate slowly in comparison with their individual counterparts due to change in the size of the molecule and intrinsic properties. In work presented in this thesis, interactions of BRCA1-BARD1 (protein-protein) heterodimeric

complex with NCPs (Protein-DNA complex), and analysis of ubiquitin interaction with NCPs are of particular interest. When protein bind to NCP, there is an increase in the molecular weight and the net charge is modified resulting in a change of migration dynamics in the gel. Interaction of molecules or formation of the complex is detected by analyzing the shift in the bands. In the assay, the concentration of NCP was kept constant, and the protein complex was titrated with increasing concentrations. In general, samples for EMSA were diluted in binding buffer (50 mM Hepes pH 7.5, 50 mM NaCl). Before sample loading, gels were pre-run for 30 min at 120 V at 4°C. Samples were separated for 1 h 20 min at 100 V. Gels were post-stained with 1x SYBR Gold (*Thermo Fisher Scientific*, S-11494) for 5 min with/without shaking and visualized using a Bio-Rad imager. Having mentioned the popularity of the assay and one of its uses, it also has limitations. One of the significant limitations of this assay is that samples are not at equilibrium during the electrophoresis run. Rapid dissociation of the complexes would inhibit visualizing band shifts (i.e.,) detection of complexes. Therefore, unable to visualize band shifts does not imply there is no interaction between the molecules.²⁵⁸

2.3.2. Ubiquitylation assay

Unless stated otherwise, ubiquitylation assays were performed by mixing 2 μ M NCP, 5 μ M His₆-wt-Ub (*Sigma*, U5507), 200 μ M K0-Ub (*Boston Biochem*, UM-NOK), 0.03 μ M Uba1 (*Boston Biochem*, E-305), 0.375 μ M UbcH5c, 0.375 μ M BRBA-S or BRBA-L in 50 mM Hepes pH 7.5, 75 mM NaCl, 1 mM DTT, 10 μ M ZnCl₂, 2.5 mM MgCl₂, 3 mM ATP. Assay samples were incubated in an incubator for 60 min at 30°C after which samples were quenched with SDS-PAGE gel loading dye and boiled for 5 minutes before separating in 12% Bis-Tris gels. Coomassie-stained gels and western blots were scanned with a Bio-Rad imager.

2.4. Insect cell expression system

2.4.1. MultiBac system

Figure 27: A MultiBac cloning strategy for co-expression of proteins.

A) The figure illustrates the Cre-loxP based recombineering of acceptor and donor plasmids, resulting in a single expression vector with all the four expression cassettes encompassed by Tn7 transposon elements. B) The genes of interest integrated into the baculoviral genome by Tn7 transposition.

In cells, most of the proteins are a component of large multi-protein complexes; there is an increase in interest for isolating and purifying multi-protein complexes to study their functions and structure. Conventional strategies to express sufficient quantities of multi-protein complexes have been demanding resources, and time. MultiBac expression system, a non-sequential modular design, and its detailed protocol,

comparatively fast and robust strategy developed by Imre Berger and his group, was used to co-express protein complexes.^{259,260}

Briefly, the MultiBac system comprises of acceptor and donor vectors; each contains two expression cassettes and a multiplication module in between the expression cassettes. One cassette includes a polH late promoter, and another cassette consists of a p10 late promoter. Genes of interest were inserted into the cassette of choice. Both acceptor and donor plasmids contain loxP sites, and for large multi-protein complexes, it is required to fuse acceptor and donor vectors via Cre-loxP fusion (Figure 27), as only the acceptor vectors contain Tn7 transposon elements that enable the integration of the gene cargo into the baculoviral genome. The loxP sites on the vectors are placed in a way that the gene cargo is flanked by the Tn7R and Tn7L transposon sequences when the recombineering is complete. When the Cre-loxP recombineering is complete, acceptor-donor fusions were integrated into the MultiBac genome by Tn7 transposition in DH10EmBacY cells combined with a YFP gene, which could be used as a reporter. YFP reporter was invaluable in simplifying the feedback mechanism, as it gave direct input regarding the health of the cells and the effect of the virus on the cells. Endpoint dilution experiments were performed to determine the viral titer.

In work presented in this thesis, pFL (acceptor) and pUCDM (donor) vectors were used. BRCA1 constructs were cloned into the pFL vector cassette containing a polH promoter, and BARD1 constructs into the pUCDM vector. pFL vector containing the gene of interest was propagated in OmniMax2 cells, whereas pUCDM vector containing gene of interest was propagated in DH10 α pir strain as donor vectors include an origin of replication derived from R6K γ . Genes encoding for proteins were inserted by ligation independent cloning. In vitro Cre-lox recombineering of pFL and pUCDM vectors were performed. Resultant recombined vector was then selected and amplified

in DH10 α pir competent strain in the presence of ampicillin (pFL) and chloramphenicol (pUCDM). Subsequently, a transfer vector containing both proteins of interest were transformed into DH10EmBacY cells, where the gene cargo was integrated into the LacZ α mini-attenuator sequence disrupting the LacZ α gene, through Tn7 transposition *in vivo*. The integrity of the LacZ α gene was confirmed by blue-white screening performed on IPTG-X-gal plates containing the right combination of antibiotics. To completion, positive clones were cultured, and the bacmid DNA was isolated.

2.4.2. Cre-loxP recombination reaction

For Cre-Lox recombination of acceptor and donor vectors, 50 mM Tris-HCl pH 7.5, 33 mM NaCl, 10 mM MgCl₂, 1 μ L Cre recombinase enzyme, 1 μ g of pFL acceptor vector, 1 μ g of pUCDM donor vector, and molecular biology grade H₂O to final volume of 20 μ L, were mixed together and incubated at 30°C for 1 h. Vector fusions were streaked on an LB-agar plate containing the appropriate combination of antibiotics, and the positive clones were isolated. It is recommended not to use very high concentrations of Cre recombinase in the reaction mix to drive the reaction as it can promote the formation of Cre-DNA aggregates.

2.4.3. Blue-white screening and bacmid isolation

Vector fusions were transformed into competent strains as per the standard molecular biology procedure. 150 μ L of 1:10, 1:100, 1:1000 dilutions were streaked on LB agar plates containing 100 μ g/mL ampicillin, 37 μ g/mL chloramphenicol, 40 μ g/mL IPTG, and 500 μ g/mL X-gal. Post-incubation for three days at 37°C, single white colonies were picked and restreaked for confirmation and incubated for 24 h at 37°C. Confirmed positive clones were grown in 2xYT medium and isolated through the alkaline lysis method. Multigene transfer vector was integrated into the bacmid using

the Tn7 trans positioning. Bacmid containing the YFP reporter gene were propagated and purified from *E. coli*. Purified bacmids with the gene of interest were transfected into SF9 cells. The 1st generation (V₀) was carefully isolated from the plates and was amplified to further two generations (V₁ & V₂) to have enough healthy virus stock for infecting large scale SF9 cell cultures.

Although protocols recommend six months to a year of storage of virus at 4°C, repeated cooling and thawing cycles during use has resulted in a 10 to 20-fold reduction in the viral titer. Therefore, fresh transfections and propagation of the virus were performed for stocks older than three months to have the best potent virus and optimal viral titer during infection. Once infected, the SF9 culture was incubated at 27°C for 48 to 72 hours depending on the construct. Time course was performed to check for optimal expression. YFP signal was monitored for infection efficiency.

2.5. Electron microscopy

2.5.1. Preparation of Cryo-EM grids

The techniques used in the preparation of cryo-EM grids is crucial in obtaining good quality cryo-EM data helpful in obtaining high-resolution information. The Quantifoil R1.2/1.3 holey carbon grids were glow discharged for 60 secs (*Harrick plasma*, PDC-32G). In general, for cryo-EM, individual proteins or complexes in micromolar concentrations provide well-distributed particles within the holes present in a carbon film. Afterward, 4 µL of a solution containing NCP-UbcH5c-BRBA-S or NCP_{K0-Ub}-UbcH5c-BRBA-S complexes were applied on the grid and blotted by Vitrobot Mark IV (FEI) for 2 sec with blot force of 1 at 4°C and 100% humidity. The grids were plunged into liquid ethane surrounded by liquid nitrogen. Prepared grids were stored in dewar holding liquid nitrogen.

2.5.2. Data collection

Data was collected using a Titan Krios microscope from *Thermo-Fisher Scientific* (formerly FEI), equipped with a phase-plate and direct electron detector, operating at 300 kV. The advantage of using Krios is that, the higher accelerating voltage decreases the extent of inelastic electron scattering. The use of phase plates introduces a phase shift between the unscattered and scattered waves during data collection, eventually increasing contrast. Though phase plate provides an advantage in terms of contrast, it is important to remember that inconsistencies in the amount of phase shift is possible. A magnification and dose rate of 130000 x and 5 e⁻/Å²/sec was used, respectively. Defocus values varying between -1.0 to -1.5 µm was used during data collection. Automated data acquisition was possible using EPU single particle acquisition software.

2.5.3. Data processing

2.5.3.1. CTF estimation, particle extraction, 2D and 3D classification

I thank Liu Zhehui and Bilal Ahsan for guiding and helping me in the processing of the cryo-EM data. Micrographs were converted to “.mrc” files for processing with **RE**gularized **L**ikelihood **O**ptimization (**RELION**),²⁶¹ a Bayesian approach dependent program developed in the group of Sjors Scheres at the MRC Laboratory of Molecular Biology. It is believed that the beam-induced motion is caused by the reaction of the sample to the high-energy electron beam, eventually leading to the increase in the accumulation of positive charge on the specimen. Beam-induced motion blurs the captured images, therefore beam-induced motion needs to be corrected. In the absence of CTF estimation, the obtained map from a reconstruction could possibly contain local density displacements. Therefore, estimation or correction of CTF is a necessity. CTFFIND is a commonly employed software for estimating the defocus parameters of

the objective lens from the obtained micrographs. In a broader sense, the defocus parameters are predicted by fitting the microscope's CTF model to a micrograph's amplitude spectrum. Motion correction and contrast transfer function (CTF) estimation were performed using Motioncor 2-1.1 and CTFFIND 4.1, respectively.^{262,263} Particles were auto-picked using EMAN2,^{264,265} and all the picked particles were fed as input for 2D classification with 25 iterations. The classes were ordered in a way that the ones holding the highest number of particles came first in the order, followed by classes with comparatively fewer particles

Class averages with good particles were selected and subjected to 3D classification with 25 iterations. The 3D model with an extra density attached to the NCP was chosen for 3D refinement. The validating metric for appraising a reconstructed 3D EM map is the gold-standard Fourier Shell Correlation (FSC) curve. An FSC curve is a 3D extension of the 2D Fourier ring correlation curve, that defines the two “half-maps,” each reconstructed from randomly selected $\frac{1}{2}$ subsets of the data. A low pass filtered NCP model was given as reference. Post-refinement, the FSC curves were plotted and the 0.143 FSC resolution criterion were noted. The crystal structure of the NCP (PDBID: 1KX5) and BRCA1-BARD1 (PDBID: 1JM7) was docked into the cryo-electron microscopy map using CHIMERA.

Chapter 3: Biochemical analysis of the BRCA1-BARD1 interaction with NCP

“Success is not final, failure is not fatal: it is the courage to continue that counts.”

- Winston Churchill

3.1. Overview

Ubiquitylation of histones is a vital epigenetic imprint that serves as a signal for many cellular processes and acts as a mark for the activation or suppression of other histone modifications.²⁶⁶ Ubiquitylation of histones is associated with transcriptional regulation.^{267,268} In eukaryotes, many E3 ubiquitin ligases are involved in the ubiquitylation reaction and determines the specificity of ubiquitylation.^{269,270} The BRCA1-BARD1 heterodimer belongs to the family of E3 ubiquitin ligases.^{217,159} BRCA1 is a large protein with a molecular weight of 207 kDa containing two functionally important domains: a RING domain at the N-terminus and two copies of BRCT domains at the C-terminus.^{203,271} BARD1 has a molecular weight of 87 kDa resembling BRCA1 in the sense that BARD1 also contains a RING domain at the N-terminus and two tandem copies of BRCT domains at the C-terminus. The NMR structure of the RING domain-containing BRCA1 and BARD1 showed that the interaction between BRCA1 and BARD1 is mediated by a four-helix bundle and sequences adjacent to their respective RING domains.²⁰⁵

For many diverse biological functions to be performed, nature has produced many protein sequences and interaction domains. Among these domains are the RING domains which require zinc for folding and DNA binding activity.^{272–275} RING domains are highly conserved zinc-coordinating Cysteine-rich motifs that are present in many

proto-oncoproteins, such as BRCA1, promyelocytic leukemia (PML), transforming factor 18 (TF 18) and ret proto-oncoprotein finger protein (RFP).^{203,276,277}

The RING domain consists of eight zinc-binding residues that coordinate two zinc atoms, in which either four cysteines (C₄) or three cysteines and a histidine (C₃H) coordinate each zinc atom tetrahedrally.^{278,279} The eight zinc-binding residues can be classed into four pairs, where first pair (cysteine 1 and cysteine 2) partner with the third pair (cysteine 4 and cysteine 5) to coordinate the first zinc atom and the second pair (cysteine 3 and histidine 1) partner with the fourth pair (cysteine 6 and cysteine 7) to coordinate the second zinc atom. The pattern formed by this pairing leads to what is called the ‘cross-brace motif’.^{2,280,281} Mutational analysis have provided evidence that the integrity of the RING domains is crucial for the protein’s enzymatic activity.^{282,283} The chelation of zinc results in loss of ubiquitylation by E3 ubiquitin ligases, implying proper folding of the RING domain by coordinating with zinc is essential for its function.²¹⁷

Biochemical characterization and structure determination of proteins and protein complexes require purified samples and in sufficient quantities. Over the years, methods to express recombinant proteins have considerably reduced the need to obtain protein samples from their natural sources or study proteins that are not abundant in cells. Many studies have established that the stability, enzymatic activity, binding affinity, and tendency to crystallize changes between different protein constructs. The rate of success in purifying a full-length protein varies from the ability to express and purify a single functional domain of a multi-domain protein, especially protein of large sizes. The approach of designing several constructs to select the good ones for biochemical characterization and structural studies has been proven effective.

Kevin Hiom and group reported that the RING domain-containing BRCA1¹⁻¹⁰⁹ and BARD1²⁶⁻¹¹⁹ were insufficient for E3 ligase activity. However, longer constructs of BRCA1¹⁻²⁵⁰ and BARD1¹⁴⁻¹⁸⁶ (abbreviated as BRBA-L) containing N and C terminal amino acid sequences, was able to ubiquitylate histone H2A. Therefore, it was concluded that BRCA1¹⁻¹⁰⁹ and BARD1²⁶⁻¹¹⁹ containing RING domains are sufficient to form a stable heterodimer, but sequences flanking the RING domain are necessary for E3 ligase activity and might contain important residues essential for the transfer of ubiquitin.²⁴²

Sequence and structural comparison of another E3 ligase RING1B¹⁻¹¹⁶, used in the crystal structure of RING1B¹⁻¹¹⁶-BMI1¹⁻¹⁰⁹-UbcH5c in complex with NCP, shows a high degree of conservation between the NCP binding loop of RING1B¹⁻¹¹⁶ and the corresponding loop in the BRCA1¹⁻¹⁰³.⁶⁴ This observation suggested, BRCA1 might use the conserved lysine and arginine residues to bind to the NCP. Therefore, it was essential to establish if BRCA1¹⁻¹⁰³-BARD1²⁶⁻¹³⁰ (called as BRBA-S) that is comparable to RING1B¹⁻¹¹⁶-BMI1¹⁻¹⁰⁹ can bind stably to NCP to conduct a structural analysis of the BRCA1¹⁻¹⁰³-BARD1²⁶⁻¹³⁰ heterodimer in complex with NCP.

Based on the primary sequence, secondary structure predictions, available structural information and detailed analysis of the previous experimental data, I designed and cloned constructs of different lengths of both BRCA1 and BARD1, co-expressed them and purified proteins from *E. coli* and SF9 cells. In order to study the interaction and complex formation of BRCA1-BARD1 with NCPs, NCPs were assembled from purified Widom 601 145 bp nucleosome positioning sequence and recombinant human histones, in sufficient quantities to carry out both biochemical and structural analysis.

3.2. NCPs for biochemical characterization and structural studies

3.2.1. Production of Widom 601 nucleosome positioning sequence

In order to obtain stable and properly packed NCPs, NCPs were reconstituted using the Widom 601 DNA sequence. This sequence was identified from SELEX experiments to have the highest affinity for the histone octamer and positioning potential.²⁰ Widom 601 DNA sequence has a single nucleosome positioning sequence at the centre of the DNA matching the dyad of the nucleosome.²⁰ To perform biochemical and structural investigations, large quantities of NCPs were reconstituted. Figure 28 shows the samples analysed by gel electrophoresis. Lane 5 shows the released Widom 601 fragment. The uncut plasmid (lane 1) and the digested plasmid (lane 2, 3 and 4) are shown in the agarose gel. The analysis of the released Widom 601 fragment shows very high purity suitable for reconstitution of the NCPs.

Figure 28: Production of Widom 601 145bp DNA.

Uncut plasmids containing the multimers of Widom 601 DNA insert (lane 1), 500 ng of the sample after treating with EcoR V for 3 h (lane 2), 500 ng of the sample after cutting with EcoR V overnight (lane 3), 1 µg of the sample after treating with EcoR V overnight (lane 4), Widom 601 145bp DNA fragments after purification (lane 5).

3.2.2. Production of the core histone proteins H2A, H2B, H3, and H4

The protocol detailed in materials and methods, results in the production of human histones H2A, H2B, H3 and H4 in sufficient quantities for biochemical and structural studies. Typical yields for recombinant human histones H2A, H2B, H3, are three to four times higher from *E.coli* than that of histone H4. As histone purification profiles are similar for the individual histones, Figure 29 shows a representative Ion exchange chromatogram and SDS-PAGE analysis of histone H2B. From the analysis of the samples by 18% SDS-PAGE denaturing gel electrophoresis, the purified histones were 90-95% pure suitable for assembly of the NCPs.

Figure 29: Representative chromatogram and SDS-PAGE analysis of histone H2B.

The slopes in red and green represent conductivity and concentration, respectively.

3.2.3. Assembly of the histone octamer

The assembly of the histone octamer involves the co-folding of the histones. The strategy to successful histone octamer production is to add 25% excess of histones H2A and H2B to facilitate the separation of the histone octamers from dimers in size-exclusion chromatography. Histone octamers, tetramers and dimers have a molecular weight of 108 kDa, 54 kDa, and 17 kDa, respectively. Histone octamers and dimers can be separated efficiently using a Superdex 200 16/600 column. Therefore, the human

histone octamer was assembled by combining histones H3, H4, H2A and H2B in a 1:1:1.25:1.25 stoichiometry.

By dialyzing the unfolded histone mix against a buffer containing a salt concentration of 2M NaCl, the unfolded histones slowly refolded to assemble into a histone octamer. The salt concentration of 2M NaCl (higher ionic strength) keeps the histone octamer intact, the histone octamer dissociates into H2A-H2B dimers and an H3-H4 tetramer at a concentration below 2M NaCl.²⁸⁴

The size-exclusion profile in Figure 30 shows three peaks, a high sharp peak migrating at the molecular weight of a histone octamer that is well separated from two other peaks, a peak close to the void volume, representing the fraction containing aggregates and a peak representing the H2A-H2B dimer.

Figure 30: Production of the histone octamer.

The chromatogram from the Superdex 200 column shows three peaks, the first peak represents aggregates, the second peak represents the refolded histone octamer, and the third peak represents the H2A-H2B dimer. The samples analysed on 18% SDS-PAGE and 12% Bis-Tris gels.

The samples were analysed by two different gels, a 18% SDS-PAGE denaturing gel and a 12% Bis-Tris denaturing gel. The samples were analysed on two different gels as histones H2B and H3 are difficult to separate on 12% Bis-Tris gels. The analysis of the fractions corresponding to the histone octamer peak by SDS-PAGE (18%) shows stoichiometric amounts of histones H2A, H2B, H3, and H4 (Figure 30, left). The histone octamer fractions analyzed in a 12% Bis-Tris gel shows human histones H2A, H4 and a comparatively thick band represent overlapping H2B and H3, shown in Figure 30 (right).

3.2.4. Reconstitution of NCPs

The histone octamer and Widom 601 DNA were combined to produce NCPs. A challenge in NCP reconstitution is that the concentration of the histone octamer is difficult to be determined accurately from standard OD₂₇₆ readings, due to the lack of aromatic amino acids in histones. To find the optimal concentration of histone octamer to reconstitute an NCP, titrations were carried out. A small-scale reconstitution with increasing histone octamer to DNA ratios were performed to establish the optimal ratio. Widom 601 DNA sequence were titrated with increasing concentrations of histone octamer reconstituted from purified recombinant human histones. The ratio at which the nucleosome positioning sequence is saturated with histone octamer and the presence of very minimal free DNA was selected as the best ratio for large-scale reconstitution. As mentioned earlier, at salt concentrations below 2M NaCl, the H2A-H2B exist as dimers and H3-H4 exist as a tetramer.²⁸⁴ The NCPs were assembled by gradually reducing the salt concentration from 2M NaCl to 0M NaCl, by step dialysis to obtain saturated and compact NCPs. During the reconstitution of an NCP, at 1M NaCl, the H3-H4 tetramer gets installed onto the Widom 601 DNA first. At 0.8M NaCl, H2A-H2B dimers settle on the DNA. Further reduction in the salt concentration and incubation overnight was

to ensure the proper assembly of the NCPs. Figure 31 shows the analysis of assembled NCPs at increasing histone octamer: DNA ratio. As the histone octamer concentration increased, saturated NCPs were assembled with a decrease in the concentration of free Widom 601 145 bp DNA.

Figure 31: Production of the NCP.

Nucleosomes reconstituted with histone octamer and Widom 601 145bp DNA at the molar ratio of 0.85, 0.95, 1.05, 1.15. Native PAGE analysis of assembled NCPs.

3.3. Purification of BRCA1-BARD1 E3 ligase and UbchH5c conjugating enzyme

3.3.1. Co-expression and co-purification of BRBA-S from *E. coli*

It is evident that the functions of a protein are dependent on its physical and chemical properties. Co-expression of proteins with their *in vivo* partners maintain or improves stability, activity, and helps in correct folding within the cell. Therefore, co-expression of BRCA1 constructs with BARD1 was a reasonable strategy. Previously the group of Rachel Klevit obtained BRCA1 and BARD1 construct by refolding of the proteins in inclusion bodies.^{212,285} To investigate if BRBA-S has efficient E3 ligase activity and

binding to the NCP, I cloned, co-expressed, and co-purified the heterodimeric complex as per the steps detailed in materials and methods. By co-expression, soluble BRBA-S heterodimeric complex was obtained as reported.⁶⁴ The growth media used for expression and buffers used for purification of BRBA-S were supplemented with zinc to facilitate efficient folding of the proteins.

In order to facilitate the purification of the BRBA-S heterodimeric complex, BRCA¹⁻¹⁰³ was tagged with tandem Strep II-His₁₀ tag at the N-terminus, BARD1²⁶⁻¹³⁰ was untagged. Since BARD1 has a high affinity for BRCA1, the aim was to pull-down the untagged BARD1²⁶⁻¹³⁰ when enriching the tagged BRCA1¹⁻¹⁰³ by affinity purification and size-exclusion chromatography. Strep-II was chosen for its higher specificity for Strep-Tactin beads, and His₁₀ was chosen for its higher affinity for the Ni-NTA resins. Strep-His₁₀-BRCA1¹⁻¹⁰³ has a molecular weight of 16.08 kDa, and non-tagged BARD1²⁶⁻¹³⁰ has a molecular weight of 11.86 kDa.

The tagging of BRCA1 with two different affinity tags provided the option of enriching the complex with two affinity purification steps, if necessary. Strep-II tag consists of eight amino acids (WSHPQFEK) whose interaction is highly selective to the Strep-Tactin (an engineered form of Streptavidin), whose binding is reversible upon treatment with Desthiobiotin, a biotin analog. Instead of the classical His₆, His₁₀ was used to increase the affinity to Ni-NTA resin during the purification. Though two affinity tags were incorporated at the N-terminus of BRCA1, the attempt to purify the complex using His₁₀ tag alone was successful. Increasing the affinity of the histidine tag to Ni-NTA resin allowed the imidazole concentration to be increased in the wash buffers helping to eliminate almost all the contaminants and in obtaining pure homogenous BRBA-S heterodimeric complex after the first Ni-NTA affinity purification step. The

analysis of the complex in a denaturing gel shown in Figure 32 B shows the purity of the sample containing only two strong bands. From the denaturing gel, the purity of the complex was estimated to be above 95%.

The Ni-NTA purified fractions were pooled, concentrated and subjected to size-exclusion chromatography. The comparison of the purified sample with the known molecular weight standards confirmed that the molecular weight of the heterodimeric complex is 28 kDa, as expected. The presence of BRCA1 and BARD1 in the complex, was confirmed by both western blot analysis and mass spectrometry. Figure 32 G shows the homogeneity of BRCA1¹⁻¹⁰³ and BARD1²⁶⁻¹³⁰ in the purified heterodimeric complex and the sample of this purity was used for biochemical studies and cryo-EM analysis. A total of 8-10 milligrams of the pure complex was obtained from four litres of *E. coli* culture.

Figure 32: Purification of the BRBA-S heterodimer.

A) Ni-NTA affinity chromatogram of the BRBA-S heterodimeric complex. The green slope represents increasing concentration of imidazole during elution. B) Denaturing gel (12% Bis-Tris) and Coomassie stain analysis of Ni-NTA eluted samples. C and D) Superdex 10/300 GL chromatogram of the purified complex (C) and the standards (D). E and F) Superdex 200 PG chromatogram of the purified complex (E) and the standards (F). G) Denaturing gel (12% Bis-Tris) and Coomassie stain analysis of the samples from size-exclusion chromatography.

Co-expression and purification of BRBA-L from *E. coli*

BRBA-L is longer than BRBA-S. BRBA-L contains more extended sequence at the N- and C-terminus of BRCA1 and BARD1, compared to BRBA-S. To check for activity and compare if BRBA-L is more active than BRBA-S, due to the presence of additional sequence at the N- and C-terminus, BRBA-L was co-expressed and co-purified. BRBA-L was cloned for co-expression in *E. coli* in pST44 polycistronic

expression vector. Standard optimization procedures (temperature, time, stabilizing agents like glycerol and detergents) were carried out to obtain optimal expression and a pure protein sample. Zinc was added to *E. coli* growth media and protein purification buffers to ensure proper folding of the proteins. Slow growth and expression at 18°C were necessary for optimum yield. Ni-NTA affinity-based enrichment gave high yields of BRCA1¹⁻²⁵⁰, but a band corresponding to the molecular weight of BARD1¹⁴⁻¹⁸⁶ (19.38 kDa) was very faint (Figure 33 B). This suggest either that co-expression is not working or degradation of BARD1. When Ni-NTA purified fractions were pooled, concentrated and subjected to heparin column, the analysis of the fractions eluted from heparin column showed increased purity of the sample. Figure 33 B and D shows analysis of the fractions obtained from Ni-NTA affinity column and heparin column, respectively. A prominent band representing BRCA1¹⁻²⁵⁰ is seen. A faint band corresponding to the molecular weight of BARD1¹⁴⁻¹⁸⁶ is indicated with a red arrow. Analysis of the sample by size-exclusion chromatography (Figure 33 E) shows a single sharp peak close to the void volume, suggesting the sample is a soluble aggregate. Analysis of the eluted fraction from the size-exclusion column on a denaturing gel showed a band comparable to the 20 kDa molecular weight marker (Figure 33 F).

Figure 33: Purification of BRBA-L from *E. Coli*.

A and B) Ni-NTA affinity chromatogram of BRBA-L complex purification (A) and denaturing gel (12% Bis-Tris) analysis of Ni-NTA affinity eluted samples (B). C and D) Heparin chromatogram of the purified sample (C) and denaturing gel (12% Bis-Tris) analysis of the heparin eluted samples (D). E and F) Superdex 75 10/300 chromatogram of the purified sample (E) and denaturing gel (12% Bis-Tris) analysis of the sample eluted from size-exclusion chromatography (F).

In order to investigate if BARD1¹⁴⁻¹⁸⁶ is expressed and to confirm the expression of BRCA1¹⁻²⁵⁰, the samples were probed with anti-His and anti-BARD1 antibodies. Figure 34 A shows the expression of BRCA1¹⁻²⁵⁰, confirmed by immunoblotting using anti-His antibody. The sample was taken before induction (B.I) and analysis by anti-His western blot did not exhibit a signal, confirming the BRCA1¹⁻²⁵⁰ was not expressed before inducing protein expression. The western blot analysis performed with anti-

BARD1 antibody show several bands for BARD1 and hence suggest that BARD1 is expressed but degraded (Figure 34 B).

Figure 34: Western blot analysis of the purified *BRCA1*¹⁻²⁵⁰ & *BARD1*¹⁴⁻¹⁸⁶ expressed in *E. coli*

A) Anti-His western blot image showing expression of *BRCA1*¹⁻²⁵⁰. B) Anti-BARD1 western blot image showing expression and degradation of *BARD1*¹⁴⁻¹⁸⁶.

3.3.2. Expression and purification of Ubch5c from *E. coli*

Ubch5c belongs to the family of E2 conjugating enzymes. E2 conjugating enzymes interact with ubiquitin and E3 ligase via their conserved α/β core domain to catalyze ubiquitylation.²⁶⁹ Ubch5c was expressed and purified from *E. coli*. The aim was to produce sufficient quantities of Ubch5c for performing ubiquitylation assays and determine the structure of NCP-BRCA1-BARD1-Ubch5c complex. Small-scale expression tests showed abundant expression of Ubch5c. As the expression was abundant, based on the method described by Rachel Klevit and group, cation exchange chromatography was used in the initial purification step.¹⁷⁰ Figure 35 A shows a peak with high absorbance representing the eluted fractions from the cation exchange column (SP column). The expected molecular weight of Ubch5c is 16.687 kDa. Analysis of the fractions obtained from the cation exchange column on a 12% Bis-Tris denaturing gel showed distinct fat bands migrating little above the 15 kDa molecular weight marker

(Figure 35 B). The comparison of the expected molecular weight with the molecular weight of the migrating band in the denaturing gel implied that the migrating band represents UbcH5c. The UbcH5c containing fractions were pooled, concentrated and subjected to size-exclusion chromatography for further purification and simultaneous buffer exchange. The chromatogram in Figure 35 C shows a single sharp symmetrical peak, indicating a homogenous population. The gel electrophoretic analysis in Figure 35 D shows the sample is 95% pure. A total quantity of 11 milligrams of the protein was obtained from six litres of *E. coli* culture. The purified UbcH5c was used for conducting ubiquitylation assays and in the cryo-EM analysis.

Figure 35: Purification of UbcH5c (E2 conjugating enzyme).

A and B) SP column chromatogram of UbcH5c purification (A) and denaturing gel analysis of the samples eluted from cation exchange chromatography (B). The slopes in red and green represent conductivity and concentration. C and D) Superdex 75 16/600 chromatogram of the

purified sample (C) and denaturing gel analysis of the samples eluted from size-exclusion chromatography (D).

3.3.3. Co-expression of BRCA1 and BARD1 constructs in SF9 cells

Because of the problems encountered when expressing BRCA1¹⁻²⁵⁰-BARD1¹⁴⁻¹⁸⁶ in *E. coli* and BRCA1¹⁻²⁵⁰-BARD1¹⁴⁻¹⁸⁶ being a eukaryotic protein, the complex was expressed in insect cells using Baculovirus. Insect cells are often used to express mammalian proteins as the proteins produced carry post-translational modifications similar to mammalian cells. Furthermore, many large eukaryotic proteins and/or complexes do not fold correctly in *E. coli* forming inclusion bodies because it lacks the appropriate chaperones for folding. The full-length BRCA1 and BARD1 proteins have been reported to be recombinantly expressed in insect cells. Based on the primary sequence elements, secondary structure predictions, and the previously published data, the constructs were designed encompassing the critical domains of interest. The designed constructs were submitted to Protein Production Platform (PPP) based in School of Biological Sciences at Nanyang Technological University (Singapore) for cloning and individual expression screening in insect cells. Attempts by PPP to individually express BRCA1 and BARD1 constructs of different lengths failed (data not shown). Therefore, I expressed BRCA1¹⁻²⁵⁰-BARD1¹⁴⁻¹⁸⁶ and different lengths of BRCA1 with full-length BARD1 in insect cells.

The strategy employed to co-express different lengths of BRCA1 with full-length BARD1 was to clone the constructs into the pFL and pUCDM vectors, respectively. BRCA1 constructs were cloned with a Strep-II tag at the N-terminus, and full-length BARD1 was cloned with a His₁₀ at the N-terminus. To co-express the proteins, pFL vectors containing different lengths of BRCA1 and pUCDM vector

containing full-length BARD1 was combined by Cre recombinase-dependent fusion of the plasmids at the loxP sites present in the vectors.

To establish if there was expression, cell lysates were analysed by western blot. The western blot images in Figure 36 A & B shows insect cell expressed BRCA1 constructs co-expresses with full-length BARD1. The two images show the same samples electrophoretically separated on two different denaturing gels and probed with anti-Strep and anti-BARD1 antibody to confirm co-expression.

Careful analysis of the anti-Strep western blot (Figure 36 A) of samples taken three days post-infection of SF9 cells indicated higher expression level of BRCA1¹⁻²⁶³ (lane 2), compared to the longer constructs BRCA1¹⁻¹⁵⁸⁰ (lane 3), BRCA1¹⁻¹⁷⁴⁵ (lane 4), and full-length BRCA1 (lane 5). Anti-BARD1 western blot of the same samples shows little to no difference, except for lane 2 where the expression is comparatively high (Figure 36 B). The separation of a protein by denaturing gel electrophoresis is influenced by its pI. The migration of the BRCA1 constructs in the denaturing gel is higher than that expected from the molecular weight in the denaturing gel.

Attempt to co-express different constructs of BRCA1 and BARD1 was successful, but subsequent purification attempts resulted in extensive degradation of longer BRCA1 constructs. The degradation of the BRCA1 constructs is substantiated by the data presented by Cheryl Arrowsmith and group.²⁸⁶ NMR studies, circular dichroism (CD) spectrometric analysis, and limited proteolysis experiments showed that the 1500 amino acids in the middle region of the full-length protein spanning 1863 amino acid residues are intrinsically disordered that serves as a platform for the binding of other native interacting partners.²⁸⁷

Figure 36: Screening for co-expression of BRCA1-BARD1 complexes.

A) Anti-strep western blot of samples taken 72 hours post infection, uninfected cells (lane 1), *BRCA1*¹⁻²⁶³ (31.90 kDa, lane 2), *BRCA1*¹⁻¹⁵⁸⁰ (178.02 kDa, lane 2), *BRCA1*¹⁻¹⁷⁴⁵ (196.49 kDa, lane 3), Full-length *BRCA1* (207.72 kDa, lane 5). B) Anti-BARD1 western blot of samples taken 72 hours post infection, uninfected cells (lane 1), Full-length *BARD1* (86.64 kDa, lanes 2,3,4, and 5).

Although the expression of proteins in insect cells is advantageous for the reasons discussed earlier, the process itself is time-consuming and expensive. Although the protein expression could be confirmed by analyzing the samples on a denaturing gel and subsequent western blot analysis each time, monitoring the effect of the virus on the cells and deciding the appropriate time to harvest the cells expressing the protein would be difficult. The use of yellow fluorescent protein (YFP) gene integrated into the bacmid and hence generating fluorescent proteins was very advantageous in obtaining real-time information on the performance of the virus and protein expression in insect cells. The use of YFP fluorescence signal allowed me to monitor the condition of the cells and protein expression, by simply imaging cells and looking for a fluorescence signal.

An example of how YFP was used in monitoring the effect of the virus on cells is shown in Figure 37. The images in Figure 37 shows bright field representation and fluorescence representation post infection for 65 hours with virus encoding BRCA1¹⁻²⁵⁰-BARD1¹⁴⁻¹⁸⁶. Bright field representation shows cells that are healthy (circled in 'green') and unhealthy/dead (circled in 'red'). The healthy cells circled in green are shiny with well-defined membranes. The unhealthy or dead cells highlighted in red are opaque and dark (Figure 37 A). The image in Figure 37 B display cell density and percentage of cells (YFP %). The image shows total cell density of 2.66 million/mL at the time of harvest. 76% of the cells (2.01 million/mL) show fluorescence exhibited by YFP reporter indicating the cells are infected with the virus and the proteins of interest are expressed.

Figure 37: An example of monitoring expression using YFP as a reporter.

A) Bright field image of cells three days post-infection with Baculovirus encoding for the BRBA-L complex. B) Fluorescence image of BRBA-L expressing cells at the time of harvest.

3.3.4. Ni-NTA affinity purification of BRBA-L from Insect cells

BRCA1¹⁻²⁵⁰-BARD1¹⁴⁻¹⁸⁶ (BRBA-L) co-expressed in SF9 cells was first purified using Ni-NTA affinity chromatography. BRBA-L was purified to perform the biochemical investigation on the ability of the heterodimer to ubiquitylate histone H2A in the NCP. The Ni-NTA eluate analyzed by denaturing gel and staining shows BRCA1¹⁻²⁵⁰ and BARD1¹⁴⁻¹⁸⁶ bands as well as two higher molecular weight bands, one

protein which runs at approximately 70 kDa and the other protein migrating at approximately 85 kDa (Figure 38 C). Since the attempts to remove the contaminants using heparin column and ATP-MgCl₂ wash was not successful, mass spectrometry analysis of the bands was carried out and confirmed the proteins were molecular chaperones. Binding to chaperones is reasonable as the same construct when purified from *E. coli* showed degradation (Figure 34), establishing the presence of unstructured regions. Chaperones are associated with the stabilization of the proteins and help in protein folding. In agreement with the obtained results, binding of molecular chaperone HSP70 to BRCA1 core was reported by Ashok Varma and group.²⁸⁸ Though the chaperones were bound, the BRCA¹⁻²⁵⁰-BARD1¹⁴⁻¹⁸⁶ heterodimeric complex was active.

Figure 38: Purification of BRBA-L from SF9 cells.

A and B) Ni-NTA affinity column chromatogram of BRBA-L purification (A) and denaturing gel (12% Bis-Tris) analysis of the samples eluted from Ni-NTA affinity column chromatography (B). C) Sample after pooling and concentrating.

3.4. Analysis of histone H2A ubiquitylation by BRCA1-BARD1

3.4.1. Multiple mono-ubiquitylation of histone H2A

To test the activity and specificity of the purified complexes BRBA-S and BRBA-L to histone H2A, ubiquitylation assays were performed using NCP as the substrate. As the conjugation of a ubiquitin molecule to the substrate is through the isopeptide bond, commonly used denaturing agent like SDS or denaturing conditions such as boiling do not break the peptide bond allowing us to visualize the ubiquitylated products on a denaturing gel. The strategy used to check and confirm the number of ubiquitylated lysine residues in histone H2A is to substitute wt-Ub with K0-Ub, as K0-Ub is incapable of conjugating to another ubiquitin group due to the absence of a ϵ -amino group of lysine and hence leads to mono-ubiquitylation. Figure 39 A illustrate the conjugation of wt-Ub forming a poly-ubiquitin chain on the substrate (vs) conjugation of a single K0-Ub to the substrate due to the absence of lysine. Figure 39 B illustrate the poly-ubiquitylation of the substrate lysine by wt-Ub and multi-mono-ubiquitylation of the substrate lysines by K0-Ub.

The NCPs were incubated with UBE1 (E1 activating enzyme), UbcH5c (E2 conjugating enzyme), BRCA1-BARD1 (E3 ubiquitin ligase), and wt-Ub or K0-Ub, in a ubiquitylation reaction compatible buffer containing ATP-MgCl₂. The ubiquitylation reaction samples were incubated for 60 min at 30°C. The expectation was that when histone H2A is ubiquitylated, the molecular weight of histone H2A is increased due to the conjugation of ubiquitin, resulting in higher molecular weight bands representing ubiquitylated histone. The denaturing gel in Figure 39 C show products from the ubiquitylation reactions performed with wt-Ub and lysine free Ubiquitin (K0-Ub - all lysines mutated to arginine). The ubiquitin group has a molecular weight of 8.6 kDa. In

the carried out ubiquitylation assays, I used wt-ubiquitin tagged with His at the N-terminus and untagged K0-Ub. The molecular weight of the wt-ubiquitin increased to 10.7 kDa (lanes 2,3 and 5) due to the additional His tag residues at the N-terminus. The molecular weight of the untagged K0-Ub is 8.6 kDa (lanes 4 and 6). The molecular weight of recombinant human histone H2A is 14 kDa. Lanes 3 and 4 show ubiquitylation products from the reactions performed with BRBA-S and lanes 5 and 6 show ubiquitylation products from the reactions carried with BRBA-L.

The result of the experiment show poly-ubiquitylation and multi-mono-ubiquitylation of histone H2A. Ubiquitylation reactions performed with wt-Ub, and the subsequent analysis of the samples by denaturing gel electrophoresis show more than three distinct bands (Figure 39, lanes 3 and 5) suggesting poly-ubiquitylation of one or more of the three residues in the histone H2A tail (residues K125, K127, and K129). The four prominent bands run at approximately 25 kDa, close to 37 kDa, little lower than 50 kDa, between 50 to 75 kDa. The expected increase in the molecular weights of histone H2A after conjugation of ubiquitin molecules are 24.7 kDa, 35.4 kDa, 46.1 kDa & 56.8 kDa. Therefore, the migration of the products are comparable to the expected molecular weights. The result confirmed poly-ubiquitylation of histone H2A *in vitro*.

The diagram illustrating the difference between poly-ubiquitylation and multi-mono-ubiquitylation is depicted in Figure 39 A and B. Figure 39 C, lanes 4 and 6 show the ubiquitylation products from the reactions conducted with K0-Ub. The assumption was that if the histone H2A tail is ubiquitylated at amino acid residues K125, K127, and K129, three bands representing ubiquitylated H2A should appear at 22.6 kDa, 31.2 kDa, and 39.8 kDa.

Figure 39: Poly-ubiquitylation and multi-mono-ubiquitylation of histone H2A.

Denaturing gel electrophoretic analysis of the ubiquitylation reaction samples are shown. A) Diagram illustrating the difference between poly-ubiquitylation and multi-mono-ubiquitylation of substrates. B) Diagram illustrating ubiquitylation by wt-Ub and K0-Ub. C) Coomassie-stained gel (12% Bis-Tris) showing poly-ubiquitylation (marked in red) and multi-mono-

ubiquitylation (marked in yellow) of histone H2A, BRBA-S mediated polyubiquitylation of histone H2A (lane 3), BRBA-S mediated multi-mono-ubiquitylation of histone H2A (lane 4), BRBA-L dependent poly-ubiquitylation of histone H2A (lane 5), BRBA-L dependent multi-mono-ubiquitylation of histone H2A.

The gel electrophoretic analysis of the reaction samples post ubiquitylation show, three noticeable bands. The first band run below 25 kDa, the second band is seen above 25 kDa, and the third band is noticed above 37 kDa. The comparison between the expected and apparent molecular weight confirms ubiquitylation of histone H2A at three sites, implying ubiquitylated K125, K127, and K129 residues. The data implies multi-mono-ubiquitylation of histone H2A *in vitro*. The NCP and the sample from the ubiquitylation reaction carried without E3 ligase acted as the control. Ubiquitylation products from wt-Ub mediated ubiquitylation reactions were further confirmed by western blot and mass spectrometry (data not shown). Altogether, the results confirm that BRBA-S and BRBA-L can both multi-mono-ubiquitylate and poly-ubiquitylate human histone H2A.

3.4.2. Native gel analysis of the NCP modified by ubiquitin molecules

Analysis of the ubiquitylation reaction samples by denaturing gel electrophoresis does not provide information on the integrity of the NCPs after ubiquitylation. In order to investigate the migration pattern of the ubiquitylated NCPs and to investigate if the ubiquitylation is mediated only by the E3 ligase BRCA1-BARD1 and not any other enzyme participating in ubiquitylation reaction, the samples were analyzed by native gel electrophoresis. The ubiquitylated NCPs were expected to migrate slowly compared to the unmodified NCPs. Figure 40 shows that the ubiquitylated NCP migrates slower in the native gel due to the ubiquitylation of the histone H2A by wt-Ub, mediated by BRBA-S (lane 12) and BRBA-L (lane 13) E3 ligases. Lane 1 show unmodified NCPs.

BRBA-S (lane 2) and BRBA-L (lane 3) complexes are not visible as the nucleic acid specific SYBR gold dye was used to stain the native gel, post-electrophoresis. The electrophoretic analysis shows a retracted band for the ubiquitylated NCP. The broader bands seen in lane 12 and 13 suggests conjugation of multiple wt-Ub molecules to the NCPs, resulting in differently modified NCPs. The modified NCPs that migrate faster in the native gel indicate a lesser number of wt-Ub molecules conjugated to NCP. The bands representing slowly migrating modified NCPs indicate more number of wt-Ub molecules attached. The variation in length of the ubiquitin chain formed does contribute to the sample retardation of different degrees. In the band shift, the lower the number of wt-Ub molecules attached and the minimum length of the ubiquitin chain yields the most modest shift. The disappearance of the band representing unmodified NCPs suggests that all the NCPs in the reaction are modified.

Figure 40: The conjugation of ubiquitin molecules promotes retardation of the NCP. Illustration of the nucleosomal shift caused by the attachment of ubiquitin molecules. The samples were resolved on a pre-electrophoresed 5.5% native polyacrylamide gel, at 4°C. Samples were incubated for 60 min @ 30°C. The reaction was performed with NCP, UBE1 (0.03 μ M), UbcH5c (1 μ M), BRBA-S (1 μ M) or BRBA-L (1 μ M), wt-Ub (20 μ M), in the presence of ATP-MgCl₂. 0.4X TB (without EDTA) was used as running buffer.

3.4.3. Self-ubiquitylation of BRBA-S

Many E3 ligases can ubiquitylate itself, referred to as self-ubiquitylation or more commonly, auto-ubiquitylation. Self-ubiquitylation of E3s promotes conformational changes and the formation of higher order structures (e.g., trimers, oligomers), to enhance or suppress their ligase activity. An example of the regulation of the ligase activity by self-ubiquitylation is Nedd4 ligase in humans.²⁸⁹ Previous studies have suggested that full-length BRCA1-BARD1 and BRCA¹⁻²⁵⁰-BARD1¹⁴⁻¹⁸⁶ self-ubiquitylates.²⁴²

Figure 41: BRBA-S is prone to self-ubiquitylation.

Coomassie-stained gel (12% Bis-Tris) showing self-ubiquitylation of BRBA-S. Samples were incubated for 60 min at 30°C. Reactions were performed with UBE1 (0.03 μ M), UbcH5c (2 μ M), BRBA-S (2 μ M), wt-Ub (20 μ M), in the presence of ATP-MgCl₂.

I wanted to investigate if the minimal RING domain-containing heterodimeric complex BRBA-S can self-ubiquitylate, in the absence of the additional N- and C-terminal sequences present in BRBA-L. The ubiquitylation reactions were performed in the presence of E1, E2, E3, wt-Ub, and ATP-MgCl₂ in a reaction compatible buffer, without the NCP (substrate). The samples from the reaction were analyzed by denaturing gel electrophoresis. The appearance of a higher molecular weight product suggests self-ubiquitylation of BRBA-S.

UbcH5c has a molecular weight of 16.68 kDa, and non-ubiquitylated BRCA1 monomer in BRBA-S heterodimer has a molecular weight of 16.08 kDa. As UbcH5c and non-ubiquitylated BRCA1 show very minimal difference in the migration, when analyzed individually, UbcH5c and non-ubiquitylated BRCA1 run parallel on a denaturing gel (Figure 41, lane 1 & 2). UbcH5c and non-ubiquitylated BRCA1 overlap when the samples from the reaction mix are analyzed on a denaturing gel (lane 4,7 & 8). The control samples did not show any signs of ubiquitylation. When the reaction mix containing UBE1, UbcH5c, BRBA-S, wt-Ub, and ATP-MgCl₂ were incubated, ubiquitylated BRBA-S was seen at a molecular weight of approximately 250 kDa (indicated by a red arrow), signaling self-ubiquitylation of BRBA-S (Figure 41, lane 8). Self-ubiquitylation of BRBA-S is further corroborated by the decrease in the intensity of the band representing wt-Ub (Figure 41, lane 8). BRCA1 self-ubiquitylates by forming K6 linked poly-ubiquitin chains, resulting in an enhanced activity to ubiquitylate histones *in vitro*.^{131,242} BARD1 increases the self-ubiquitylating activity of BRCA1.¹⁵⁹ The fact that the RING domain containing BRBA-S can self-ubiquitylate endorses the understanding that the RING domain is crucial for the ubiquitylation activity.

3.4.4. Analysis of BRCA1-BARD1 mediated H2A ubiquitylation by wt-Ub

In order to compare the ubiquitylation activity of BRBA-S to that of BRBA-L heterodimeric complex, time courses were carried out in the presence of wt-Ub. The concentrations of the BRBA-S, BRBA-L, and NCP were kept constant. The concentrations of BRBA-S, BRBA-L, and NCP used in the reaction were 1 μ M, 1 μ M, and 2 μ M, respectively. As mentioned earlier in the previous sections, wt-Ub is capable of building poly-ubiquitin chains. Therefore, there is no control over the number of wt-Ub groups that can attach to residues on the substrate. NCPs were incubated with UBE1, UbcH5c, wt-Ub, and same concentrations of BRBA-S and BRBA-L complexes in a ubiquitylation reaction compatible buffer containing ATP-MgCl₂. The reactions were performed and analyzed on a denaturing as well as a native gel (Figure 42 A & B). To establish the difference in the activity between the BRBA-S and BRBA-L heterodimers, consumed fraction of unmodified H2A vs time and ubiquitylated H2A vs time, curves were plotted (Figure 42 C & D). The denaturing gel in Figure 42 A show ubiquitylated H2As. Lanes 4-8 and lanes 10-14 represent ubiquitylation reactions at 0 min, 8 min, 16 min, 32 min, 64 min, mediated by BRBA-S and BRBA-L, respectively.

At time points 8 min, 16 min, 32 min, and 64 min, ubiquitylated H2As are noticeable in both BRBA-S and BRBA-L mediated ubiquitylation reactions (Figure 42 A, lanes 5-8 & lanes 11-14). In lanes 4-8 and 10-14, a decrease in the intensity of bands representing unmodified H2A and wt-Ub are evident. In both BRBA-S and BRBA-L mediated ubiquitylation reactions, a decrease in the intensity of the band representing unmodified H2A is seen at 16 min indicating a higher degree of ubiquitylation (Figure 42, lane 6 & 12). Lane 6 and lane 12 show BRBA-S and BRBA-L mediated ubiquitylation reactions, respectively. In Figure 42 A, lane 7, BRBA-S mediated

ubiquitylation reaction at 32 min show almost all the unmodified H2A is consumed, whereas in BRBA-L mediated ubiquitylation reaction at 16 min show almost all the unmodified H2A is consumed (Figure 42 A, lane 12). The ability to ubiquitylate H2A in a shorter time suggests BRBA-L is more active.

Figure 42: H2A ubiquitylation with wt-Ub.

A) Denaturing gel showing poly-ubiquitylation of histone H2A, mono-, di-, tri-, tetra-, and penta-ubiquitylated H2A are labeled correspondingly. B) Native gel showing sample retardation due to conjugation of wt-Ub to the substrate. C) The curve showing consumed amount of histone H2A vs. time. D) The bar graph showing ubiquitylated H2A vs. time. The ubiquitylation assay was performed at the specified time points at 30°C in a total reaction volume of 30 μ L buffer containing 3 mM ATP (50 mM Hepes pH 7.5, 75 mM NaCl, 1 mM DTT, 10 μ M ZnSO₄, 3 mM MgCl₂). The protein concentrations were 0.03 μ M UBE1, 1 μ M UbcH5c,

1 μ M BRBA-S, 1 μ M BRBA-L, 35 μ M wt-Ub, and 2 μ M NCP). Subsequently, samples were denatured by heating at 95°C, followed by analysis on a denaturing gel.

Since wt-Ub has a molecular weight of 10.7 kDa, and recombinant human histone H2A has a molecular weight of 14 kDa, the expected molecular weights of mono-, di-, tri-, tetra-, and penta-, ubiquitylated H2As are 24.7 kDa, 35.4 kDa, 46.1 kDa, 56.8 kDa, and 67.5 kDa, respectively. Careful analysis of the denaturing gel in Figure 42 A show bands at ~25, ~37, ~50, above 50, below 75 kDa, signifying ubiquitylated H2A. Each ubiquitylated H2A band that appeared on the denaturing gel was a match to the expected molecular weights - the denaturing gel in Figure 42 A show the mono-, di-, tri-, tetra- and penta-ubiquitylated H2As as Ub₁, Ub₂, Ub₃, Ub₄, and Ub₅, respectively. Ub₁, Ub₂, Ub₃, Ub₄, and Ub₅ bands were evident in both BRBA-S and BRBA-L mediated ubiquitylation reactions.

The gel in Figure 42 A show that, at 32 min and 64 min, Ub₃ (tri-ubiquitylated H2A) was more intense with a decrease in the intensity of Ub₁ and Ub₂ (mono- and di-ubiquitylated H2A), signifying the preference shown by wt-Ub to engage free ubiquitylation sites. Together the results of the analysis suggest that BRBA-L is more active compared to BRBA-S.

Having analysed the reaction products on a denaturing gel to visualize the ubiquitylation of histone H2A, the same set of samples were electrophoretically separated on a native gel to compare the migration or shift in the modified NCPs. The native gel in Figure 42 B shows the shift in the NCP due to ubiquitylation of histone H2A. Lanes 1-3 represent control samples showing no shift, indicating NCPs are unmodified. Lanes 4-8 and 10-14 represent ubiquitylation reactions at 0 min, 8 min, 16 min, 32 min, 64 min, mediated by BRBA-S and BRBA-L, respectively. BRBA-L modified NCPs show higher band shift in comparison to the shift induced by BRBA-S

mediated ubiquitylation (Figure 42 B). From the comparison of the levels of H2A modification and the migration of the modified NCP it is concluded that BRBA-L has higher activity compared to BRBA-S.

In order to quantify the difference in activity between the BRBA-S and BRBA-L complexes, the bands in the denaturing gel were quantified using Bio-Rad Image Lab software, and curves were plotted using Prism (*GraphPad Software Inc*). Figure 42 C shows the curves of the quantification of the unmodified H2A, against time (min). The consumed fraction of unmodified H2A (vs) time show that the BRBA-L catalyzed ubiquitylation reactions depleted more unmodified H2A compared to BRBA-S catalyzed ubiquitylation reactions, confirming BRBA-L is more active. The bands Ub₁, Ub₂, Ub₃, Ub₄, and Ub₅, were quantified individually and integrated for each time point to plot a curve against time (min). The curve plotted with values obtained from integrating ubiquitylated H2A quantifications (vs) time also pronounced that BRBA-L has higher activity (Figure 42 D).

3.4.5. Analysis of BRCA1-BARD1 mediated H2A ubiquitylation by K0-Ub

The ubiquitylation time course experiments to measure the E3 ligase activity between BRBA-S and BRBA-L was also performed in the presence of K0-Ub. In K0-Ub, the arginines substituting the lysines present in wt-Ub inhibits the formation of poly-ubiquitin chains and hence interpreting the pattern of ubiquitylation should be simpler. In the ubiquitylation reactions, K0-Ub was in large excess. The expectation was that if any other lysines in the histone H2A are ubiquitylated, the concentration of K0-Ub should not be limiting.

NCPs were incubated with UBE1, UbcH5c, K0-Ub, and same concentrations of BRBA-S and BRBA-L complexes in a ubiquitylation reaction compatible buffer

containing ATP-MgCl₂. The same concentrations of BRBA-S and BRBA-L complexes were used in the reactions. The reactions were performed and analyzed on a denaturing as well as a native gel (Figure 43 A & B). To quantify the activity between the BRBA-S and BRBA-L heterodimers, consumed fractions of unmodified H2A vs time and ubiquitylated H2A vs time, curves were plotted (Figure 43 C & D). The denaturing gel in Figure 43 A show ubiquitylated H2As. Lanes 4-8 and lanes 10-14 represent ubiquitylation reactions at 0 min, 8 min, 16 min, 32 min, 64 min, mediated by BRBA-S and BRBA-L, respectively. At 0 min, ubiquitylated H2A is not visible in both BRBA-S and BRBA-L mediated reactions. At time points 8 min, 16 min, 32 min, and 64 min, ubiquitylated H2As are visible in both BRBA-S and BRBA-L mediated ubiquitylation reactions (Figure 43 A, lanes 5-8 & lanes 11-14). In Figure 43 A, lane 8, BRBA-S mediated ubiquitylation reaction at 64 min show almost all the unmodified H2A is utilized, whereas in BRBA-L mediated ubiquitylation reaction at 32 min show almost all the unmodified H2A is utilized (Figure 43 A, lane 13). The potential to fully modify H2A in a shorter time again indicates BRBA-L is a more active complex.

K0-Ub has a molecular weight of 8.6 kDa, and recombinant human histone H2A has a molecular weight of 14 kDa. The expected molecular weights of mono-, di-, and tri-ubiquitylated H2As are 22.6 kDa, 31.2 kDa, and 39.8 kDa, respectively. The analysis of the modification in Figure 43 A shows bands migrating at approximately 25 kDa, above 25 kDa, and little above 37 kDa, indicating ubiquitylated H2A. Each ubiquitylated H2A band that appeared on the denaturing gel was comparable to the expected molecular weights. The denaturing gel in Figure 43 A show the mono-, di-, and tri-ubiquitylated H2As indicated as Ub₁, Ub₂, and Ub₃, respectively. Ub₁, Ub₂, and Ub₃ bands were noticeable in both BRBA-S, and BRBA-L mediated ubiquitylation reactions.

Figure 43: H2A ubiquitylation with K0-Ub.

A) Denaturing gel showing multiple-mono-ubiquitylation of histone H2A, mono-, di-, and tri-ubiquitylated H2A are labeled as Ub₁, Ub₂, and Ub₃, respectively. B) Native gel showing nucleosome retardation due to conjugation of K0-Ub to the H2A. C) The curve showing consumed fraction of histone H2A vs. time. D) The histogram showing ubiquitylated H2A vs. time. The ubiquitylation assay was performed at the specified time points at 30°C in a total reaction volume of 30 μ L buffer containing 3 mM ATP, 50 mM Hepes pH 7.5, 75 mM NaCl, 1 mM DTT, 10 μ M ZnSO₄, and 3 mM MgCl₂. The protein concentrations were 0.03 μ M UBE1, 1 μ M UbcH5c, 1 μ M BRBA-S, 1 μ M BRBA-L, 100 μ M K0-Ub, and 2 μ M NCP.

Lane 14 in Figure 43 A show BRBA-L mediated ubiquitylation reaction samples taken at 64 min. At 64 min, Ub₃ (tri-ubiquitylated H2A) became more abundant with a decrease in the intensity of Ub₁ and Ub₂ (mono- and di-ubiquitylated H2A), implying the preference shown by K0-Ub to engage unoccupied ubiquitylation sites. The activity comparison is shown in Figure 43 A and advocates that the BRBA-L is more active compared to BRBA-S.

Having investigated the reaction time courses on a denaturing gel and confirming that BRBA-L is more active than BRBA-S, the same samples were analyzed by native gel electrophoresis to compare the shift of the modified NCPs. In Figure 43 B, NCPs are retarded due to ubiquitylation of histone H2A. Lanes 1-3 represent control samples and show no apparent shift, indicating the NCPs are unmodified. Lanes 4-8 and 10-14 represent ubiquitylation reactions at 0 min, 8 min, 16 min, 32 min, 64 min, mediated by BRBA-S and BRBA-L, respectively. BRBA-L modified NCPs show higher band shift in comparison to the shift promoted by BRBA-S mediated ubiquitylation (Figure 43 B). The excess K0-Ub is seen in the wells. In summary, the interpretations made based on two different analysis (denaturing gel electrophoresis and native gel electrophoresis) advocates that BRBA-L has higher ubiquitylation activity compared to BRBA-S.

In order to quantify the modification levels between the BRBA-S and BRBA-L, the bands in the denaturing gel were quantified using Bio-Rad Image Lab software, and curves plotted using Prism (*GraphPad Software Inc*). The Figure 43 C shows the plot derived from the quantification of the unmodified H2A, against time (min). The consumed fraction of unmodified H2A (vs) time curve show that the BRBA-L catalyzed ubiquitylation is more rapid compared to BRBA-S catalyzed ubiquitylation reactions,

again confirming BRBA-L is the more active construct. The bands Ub₁, Ub₂, and Ub₃ were quantified individually and integrated for each time point to be represented in a histogram. The histogram with values obtained from integrating ubiquitylated H2A quantifications (vs) time also show that BRBA-L has higher activity (Figure 43 D). The conclusions drawn based on the data obtained from denaturing gel electrophoresis, native gel electrophoresis, the curve plotted, and the histogram uphold one another and confirms BRBA-L has higher E3 ubiquitin ligase activity in comparison to shorter heterodimeric complex BRBA-S. Although K0-Ub was used in large excess compared to that of wt-Ub, band shift was less in K0-Ub modified NCPs compared to wt-Ub modified NCPs, a characteristic difference caused due to mono and poly-ubiquitylation of substrates.

3.5. Conclusions

Widom 601 nucleosome positioning sequence, recombinant human histones, histone octamer, and the NCPs were purified and assembled in milligram quantities (using protocols established by Rhodes laboratory at MRC and the Richmond's laboratory) to characterize E3 ligase function of BRCA1-BARD1 heterodimer biochemically and for cryo-EM studies. In order to compare the E3 ligase activity, two different length constructs of the BRCA1-BARD1 complex were expressed. BRBA-S and BRBA-L were co-expressed and co-purified from *E. coli* and SF9 cells, respectively. BRBA-S was obtained pure and homogenous, but purification of BRBA-L from insect cells yielded chaperone bound heterodimer. For biochemical experiments and structural studies, UbCH5c was purified from *E. coli*.

In order to check if the purified BRBA-S and BRBA-L are active, ubiquitylation assays were performed, and the activity established. To check if BRBA-S can self-

ubiquitylate, ubiquitylation assay was carried out with E1, E2, E3, NaCl, and ATP-MgCl₂, without the NCP. The ubiquitylation assay confirmed self-poly-ubiquitylation of BRBA-S. The time course experiments to compare E3 ligase activity, and their quantification suggested BRBA-L is more active than BRBA-S. However, due to the difference in the expression systems used to express the complexes and the variation in the homogeneity of the BRBA-S and BRBA-L complexes, the findings can be substantiated when both the complexes are purified from the same expression system, the homogeneity of the BRBA-L is improved, and by assessing the PTMs by mass spectrometry.

The ubiquitylation assays and time course experiments performed with wt-Ub and K0-Ub showed poly-ubiquitylation and multi-mono-ubiquitylation of the histone H2A, respectively. The UbcH5c-BRCA1 interactions govern the diversity or dual activity exhibited by BRCA1-BARD1 in both poly-ubiquitylating and mono-ubiquitylating substrates.²¹⁸ Although two different length constructs were tested for the activity and the activity compared, the BRBA-S complex was about 95% pure, whereas the BRBA-L was bound by chaperones. BRBA-S was more suitable for the subsequent studies and cryo-EM analysis. It is logical to choose BRBA-S as it is more stable, and it contains the NCP binding loop.

The critical step for ubiquitylation of histone H2A is the binding of BRCA-BARD1 to the NCP, directing the transfer of ubiquitin. However, for structural determination of the NCP-BRCA1-BARD1 complex, it is essential that the complex is stable. To investigate if BRBA-S stably bind to the NCP and to determine their binding affinity, EMSAs (Electro-Mobility Shift Assay) were performed. The concentration of the NCP was kept constant with increasing concentration of BRBA-S. Analysis of the

samples by native gel electrophoresis showed no visible shift (data not shown). Since it was known that UbcH5c increases the binding affinity of RING1B-BMI1 to NCP,⁶⁴ I investigated if adding UbcH5c gave a stable NCP-UbcH5c-BRBA-S for cryo-EM studies. The EMSAs performed in the presence of UbcH5c did not show any apparent shift of the NCP. It is possible that BRBA-S interact with the NCP transiently and in turn, it is challenging to capture a stable NCP-BRBA-S complex. It is reasonable to think that the weak or transient interaction of E3 ligase to the substrate is due to the coordinated functioning of E3 ligases and deubiquitinases in conjugating or detaching ubiquitin, respectively. Indeed it has been reported that the nature of the interaction between E3 ligases and substrates are transient.²⁹⁰ In order to determine the cryo-EM structure of NCP-BRBA-S complex, I chose to stabilize the NCP-BRBA-S complex by crosslinking with glutaraldehyde, which is a commonly used strategy to stabilize protein complexes.²⁹¹ Crosslinking is a technique by which covalent coupling of two functional groups between two different proteins can be accomplished, that helps in stabilizing low affinity interactions. Though crosslinking using glutaraldehyde is a common technique, it is important to note that increasing specificity by fine-tuning parameters such as protein concentration, incubation time, temperature, pH, and buffer composition are important to avoid unspecific and/or intramolecular crosslinking.

Chapter 4: Stabilization of complexes for cryo-EM analysis

“Somewhere, something incredible is waiting to be known.”

- Carl Sagan

4.1. Overview

Although protein-protein interactions have different stability, weak or transient interactions present a challenge when determining structure. In general, interactions associated with the enzymatic reactions are short lived and transient. It is possible to covalently trap the proteins that physically interact with one another when the interacting partners are proximate, using a chemical crosslinker. The length of the cross-linker determines how far apart the two interactive sites can be in space. In order to avoid changes in the native protein structure and prevent aggregation, it is essential to determine the minimal concentration of cross-linker that stabilizes the complex without perturbations. In order to stabilize the NCP-BRBA-S complex to determine the structure, glutaraldehyde was the chosen cross-linker. Glutaraldehyde is a 5-carbon dialdehyde (bifunctional) chemical crosslinker that reacts with amine groups, it crosslinks only proteins, unlike formaldehyde (crosslinks protein and DNA). When cross-linked, covalent bonds are formed between the functional groups of glutaraldehyde and the proteins, stabilizing the macromolecular complex.

Glutaraldehyde was initially used in fixing tissues for microscopic analysis, in the preparation of stable crystals for structure determination by X-ray crystallography, and to produce antibody-enzyme conjugates.^{292–294} Presently, it is frequently being used in stabilizing complexes for negative stain and single-particle cryo-electron microscopy.^{291,295} Glutaraldehyde targets lysine residues in proteins and the reaction buffer should be free from amines that have the potential to react with cross-linkers.

This property of amines is used to quench the cross-linking reaction with amine-containing solutions like Glycine or Tris. Cross-linking is also used as a technique to stabilize complexes that has a high affinity but falls apart during the process of vitrification.

There are three essential criterions to be taken into consideration when performing cross-linking studies, 1) the temperature at which the cross-linking reactions are performed, 2) duration of incubation of the proteins with the chemical cross-linker and 3) the final concentration of the cross-linker used for cross-linking the proteins. The efficiency of cross-linker in cross-linking and stabilizing complexes is based on the size of the protein, protein molecular weight and the number of reactive groups. To determine the optimal glutaraldehyde concentration that could stabilize NCP-BRBA-S complex and at the same time prevent cross-links within NCPs or BRBA-S, I first conducted concentration and time-dependent cross-linking assays with glutaraldehyde as a crosslinker for NCPs and BRBA-S, individually and then based on this, the concentration of glutaraldehyde to stabilize the interaction of BRBA-S with NCPs was determined. The optimal glutaraldehyde concentration was determined to stabilize BRBA-S with NCP. The aim was to find the glutaraldehyde concentration that is suitable to cross-link and stabilize BRBA-S to NCP with negligible or no NCP-NCP and BRBA-S-BRBA-S interparticle cross-linking. Cross-linking reactions were performed at 4°C to stabilize complexes.

4.2. Results

4.2.1. Cross-linking of histones at increasing concentrations of glutaraldehyde

In order to analyze the effect of glutaraldehyde concentration on histones in the NCP, cross-linking analysis was conducted for NCPs at 1 μ M and 3 μ M that are suitable

concentrations for cryo-EM studies. The NCPs were incubated with increasing concentrations of glutaraldehyde (0.05%, 0.1%, 0.2%, 0.4%, 0.8%, 1%) for 10 min at 4°C. Figure 44 A & B show a concentration of 1 μ M and 3 μ M of NCPs treated with increasing concentrations of glutaraldehyde, respectively. Due to the addition of glutaraldehyde to NCPs, the amount of non-crosslinked histone bands started to decrease, and cross-linked higher molecular weight bands started to appear. NCPs without glutaraldehyde served as the controls.

At 0.05% glutaraldehyde (lane 3), higher molecular weight bands between 25 kDa and 37 kDa (lane 3) started to show up, implying cross-linking of histones at 0.05% glutaraldehyde. As the concentration of glutaraldehyde increased in reactions, the appearance of cross-linked histone products (dimers, tetramers, and octamers) became more and more apparent (lanes 4-6). At glutaraldehyde concentrations above 0.6% (lanes 7-9), aggregated higher molecular weight cross-linked products are visible in the wells. At 0.6%, 0.8% and 1% of glutaraldehyde, the bands representing non-cross-linked histones have disappeared, implying almost all the histones are cross-linked.

The denaturing gel in Figure 44 B show two faint bands between 25 and 37 kDa at a glutaraldehyde concentration of 0.05%, representing cross-linked histones. Careful analysis of the denaturing gels shows that 0.05% of glutaraldehyde is sufficient to promote cross-linking of histones in 1 μ M and 3 μ M of NCP when incubated for 10 min at 4°C. Therefore, I decided to lower the concentration of glutaraldehyde and investigate the effects of lower concentrations of glutaraldehyde at different time points.

Figure 44: Self-association of nucleosomal histones when treated with different concentrations of glutaraldehyde.

A) Denaturing gel electrophoresis showing cross-linking of histones when 1 μ M of NCP was titrated with increasing concentrations of glutaraldehyde. B) Denaturing gel showing cross-linking of histones when 3 μ M of NCP was titrated with increasing concentrations of glutaraldehyde. The cross-linking reaction was performed in a buffer containing 50 mM Hepes, pH 7.5, 50 mM NaCl and 1 mM DTT. The reactions were quenched with 1 M Tris-HCl pH 7.5. The samples were analyzed by 4-12% gradient Bis-Tris gel.

4.2.2. Analysis of cross-linking of histones at different time points

It was observed that treatment of NCPs for 10 min with a glutaraldehyde concentration of 0.05% and above, cross-linked histones in the NCP. Therefore, it was essential to perform cross-linking assays with glutaraldehyde concentrations lower than 0.05% at increasing time points to determine the glutaraldehyde concentration at which the histones do not cross-link. The images in Figure 45 A, B, C, and D show denaturing gel analysis of NCPs treated with 0.005%, 0.01%, 0.025%, 0.05% of glutaraldehyde for 10 min, 30 min, 60 min, and 120 min, respectively.

Figure 45: Time course of NCPs crosslinked with glutaraldehyde.

A) Denaturing gel showing cross-linking of histones when 3 μ M of NCP was treated with glutaraldehyde for 10 min. B) 12% denaturing gel showing cross-linking of histones when 3 μ M of NCP was treated with glutaraldehyde for 30 min. C) 12% denaturing gel showing cross-linking of histones when 3 μ M of NCP was treated with glutaraldehyde for 60 min. D) 12% denaturing gel showing cross-linking of histones when 3 μ M of NCP was treated with glutaraldehyde for 120 min. The cross-linking reaction was performed in buffer containing 50

mM Hepes pH 7.5, 50 mM NaCl and 1 mM DTT. The reactions were quenched with 1 M Tris-HCl pH 7.5. Samples were analysed in 12% Bis-Tris denaturing gels.

The denaturing gel in Figure 45 A show cross-linked histones at a glutaraldehyde concentration of 0.05%. There are no visible cross-linked products in the samples treated with concentrations of 0.005%, 0.01%, and 0.025% glutaraldehyde. At 30 min, 60 min and 120 min, cross-linked histones are seen at 0.025%, 0.025% and 0.01%, respectively. Figure 45 B (sample at 30 min) & C (samples at 60 min) show no major difference in the rate at which the histones cross-link, up to a glutaraldehyde concentration of 0.025%. Detailed analysis exhibit cross-linked higher molecular weight products in samples treated with 0.05% glutaraldehyde, at 60 min and 120 min. Overall, the cross-linking assays suggest a glutaraldehyde concentration of 0.01% for 30 min would be a reasonable starting point to stabilize NCP-BRBA-S complex. Having determined the glutaraldehyde concentration and time at which the histones starts to cross-link, cross-linking assays were performed to determine the glutaraldehyde concentration at which BRBA-S does not show cross-linked products.

4.2.3. Analysis of cross-linking of BRBA-S at increasing concentrations of glutaraldehyde

Because the non-tagged BRCA1¹⁻¹⁰³ has a molecular weight of 11.61 kDa, and non-tagged BARD1 has a molecular weight of 11.86 kDa, gel electrophoretic fractionation of the heterodimeric complex on denaturing gels leads to overlapping of protein bands. Therefore, to facilitate the interpretation of cross-linking experiments, the Strep-His₁₀ tag at the N-terminus of BRCA1¹⁻¹⁰³ was not cleaved off, resulting in a difference in the migration of BRCA1¹⁻¹⁰³ and BARD1²⁶⁻¹³⁰ in denaturing gels. BRBA-S was treated with 0.05%, 0.1%, 0.2%, 0.4%, 0.6%, 0.8%, and 1% of glutaraldehyde for a period of 10 min at two different BRBA-S concentrations at 4°C.

Figure 46: Cross-linking pattern of BRBA-S heterodimer when treated with increasing concentrations of glutaraldehyde.

A) Denaturing gel showing cross-linking of BRBA-S when 4 μ M of BRBA-S was treated with increasing concentrations of glutaraldehyde for 10 min. B) 4-12% gradient denaturing gel showing cross-linking of BRBA-S when 12 μ M of BRBA-S was treated with increasing concentrations of glutaraldehyde for 10 min. The denaturing gels were stained by Coomassie blue dye. Analysis was carried out in a 4-12% denaturing gradient gel.

The denaturing gel in Figure 46 A shows BRBA-S treated with increasing concentrations of glutaraldehyde. BRBA-S at a concentration of 12 μ M treated with increasing concentrations of glutaraldehyde is shown in Figure 46 B. Figure 46 A shows, at 0.05% glutaraldehyde, cross-linked products appear at approximately 25 kDa, 37 kDa, 50 kDa. At glutaraldehyde concentrations between 0.1% and 0.8%, 25 kDa, 37 kDa, 50 kDa bands representing cross-linked products appear along with the smear above 50 kDa (Figure 46 A). In Figure 46 B, there is no appearance of cross-linked products at a glutaraldehyde concentration of 0.05%. Glutaraldehyde concentrations between 0.1% to 0.8% shows a cross-linked product running at 25 kDa (Figure 46 B). Wells in the denaturing gels do not show any signs of aggregation. Careful analysis of the gels suggests BRCA1 shows greater propensity to get cross-linked, indicated by the

rate of disappearance of the band representing BRCA1. BRCA1 is a zinc binding protein and contains many cysteines. Cysteines have a very reactive thiol group that form disulfide bridges between cysteines that are in proximity, under oxidizing conditions. For cryo-EM analysis, 3 μ M: 12 μ M of NCP: BRBA-S was chosen.

4.2.4. Cross-linking of BRBA-S to NCP

Based on the results obtained from glutaraldehyde concentration titration and time course experiments performed for NCP and BRBA-S, individually, mild cross-linking condition for 30 min were chosen. A glutaraldehyde concentration of 0.01% for 30 min at a temperature of 4°C was tested. The concentration of NCP was kept constant with increasing concentrations of BRBA-S. The analysis of the crosslinking reaction was both by denaturing gel electrophoresis and by EMSA.

NCPs were incubated with increasing concentrations of BRBA-S, and each point in the titration was treated with 0.01% glutaraldehyde for 30 min. Figure 47 shows the analysis by denaturing gel electrophoresis (A) and by EMSA (B). Lane 1 in Figure 47 A and Figure 47 B represent the NCPs. No cross-linked products were observed when the NCPs were treated with 0.01% of glutaraldehyde for 30 minutes, in the absence of BRBA-S (Figure 47 A, lane 2). This control was added to check if internal cross-linking occurs within the histones at a glutaraldehyde concentration of 0.01% and at a time point of 30 min. As the concentration of BRBA-S increases, the BRCA1 band became thicker overlapping H2A, and thick BARD1 bands overlap H4 (Figure 47 A, lane 4-7). The overlapping of the BRCA1 with H2A, and BARD1 with H4 made interpretation of the results from the denaturing gel electrophoresis difficult and inconclusive. Analyzing the samples by native gel electrophoresis, the signal intensities of the bands representing non-cross-linked NCPs were not significantly reduced (Figure 47 B, lanes 3-7),

implying no apparent band shift due to cross-linking. Although a minimal difference in the band intensity was noticed in lanes 5-7 shown in Figure 47 B, when compared to the control lanes (Figure 47 B, lanes 1-2), the difference was not sufficiently substantial to take sample for cryo-EM. Altogether, attempt to stabilize BRBA-S to NCP with a glutaraldehyde concentration of 0.01% for 30 min did not provide a stable complex for cryo-EM analysis.

Figure 47: Cross-linking of BRBA-S to NCP with 0.01% glutaraldehyde.

A) Analysis of samples treated with 0.01% glutaraldehyde on a 12% denaturing gel and post-stained with Coomassie blue. B) Samples were analyzed by native gel electrophoresis on 5.5% polyacrylamide and post-stained with Coomassie dye.

As cross-linking of proteins with 0.01% glutaraldehyde did not give a stable complex, the glutaraldehyde concentration was increased to 0.025% for 30 min. The native gel in Figure 48 B, lane 2 shows no products from cross-linking when NCPs were treated with 0.025% of glutaraldehyde for 30 minutes, suggesting 0.025% of glutaraldehyde for 30 min does not target the histones in the NCPs. There was a reduction in the intensity of the bands representing non-cross-linked NCPs (Figure 48 B, lanes 3-7) suggesting the

complex is cross-linked, but the cross-linking was not efficient as the full-band shift was not noticed. The data obtained from cross-linking assays performed with 0.01% and 0.025% glutaraldehyde did not provide a stable complex for cryo-EM analysis.

Figure 48: Cross-linking of BRBA-S to NCP with 0.025% glutaraldehyde.

A) Analysis of samples treated with 0.025% glutaraldehyde on a 12% Bis-Tris denaturing gel and post-stained with Coomassie blue. B) Samples were analyzed by native gel electrophoresis on 5.5% polyacrylamide and post-stained with Coomassie dye.

4.2.5. Cross-linking analysis of NCP_{K0-Ub}-UbcH5c-BRBA-S with glutaraldehyde

As mild cross-linking concentrations with a more extended incubation period did not provide stable complexes for structural studies, cross-linking assays were performed with higher concentrations of glutaraldehyde for a shorter time, in the buffer containing ubiquitylated NCP (hereafter called as NCP_{K0-Ub}), BRBA-S, and UbcH5c, since UbcH5c has been shown to increase the affinity of E3 ligase to NCPs by a factor of ten.⁶⁴ NMR studies have shown that E3 ligases, such as RNF169 and RAD18 have a higher affinity for ubiquitylated NCPs.²⁹⁶ The NCPs were ubiquitylated with K0-Ub as it results in

mono-ubiquitylation of the NCPs. wt-Ub were not used for ubiquitylating NCPs as it would result in a more heterogeneous population of NCPs.

NCP_{K0-Ub}, Ubch5c, and BRBA-S in the ratio of 1:4:4 ratio was incubated with 0.05%, 0.1%, 0.2%, 0.4%, 0.8%, and 1% of glutaraldehyde for 5 min at 4°C. The gel in Figure 49 shows samples cross-linked with increasing concentrations of glutaraldehyde. Lane 1, lane 2, lane 3, lane 4 and lane 5 shows a molecular weight marker, unmodified NCPs, BRBA-S, Ubch5c, and K0-Ub, respectively. Molecular weights of unmodified NCP, BRBA-S, Ubch5c, and K0-Ub are 198.947 kDa, 28 kDa, 16.68 kDa and 8.6 kDa, respectively. When separated by gradient denaturing gel electrophoresis, BRCA1, Ubch5c, H2A, and Ubch5c overlap with one another, as shown by a prominent band running a little above 15 kDa (Figure 49, lane 6). Presuming all the three lysines are ubiquitylated in an H2A, an NCP_{K0-Ub} will have a molecular weight of 250.54 kDa. BRBA-S and Ubch5c bound to interfaces on both sides of the NCP_{K0-Ub} will increase the molecular weight of the entire complex to 339.9 kDa. It is also possible that the NCP_{K0-Ub} could be bound to one side if the complex is unsaturated, having a molecular weight of 295.22 kDa. The gels in Figure 49 A & B show, at glutaraldehyde concentrations of 0.4% and 0.6%, a higher molecular weight species run above 250 kDa marker band, comparable to the molecular weight of BRBA-S complexes bound to NCPs. The cross-linking assays were performed twice to confirm if the results obtained are consistent. The higher molecular weight product was consistent in both the assays, shown in Figure 49 A, lane 9 and Figure 49 B, lane 4, with no aggregates in the well. Although the analysis by gel suggests the sample could be heterogeneous, it was decided to go ahead with cryo-EM analysis as recent developments in data processing enable analysis of heterogenous samples.

Figure 49: Cross-linking of NCP_{K0-Ub}-UbcH5c-BRBA-S with glutaraldehyde.

Analysis of samples treated with increasing concentrations of glutaraldehyde on a 4-12% gradient denaturing gel and post-stained with Coomassie blue. A) Marker (lane 1), NCP (lane 2), UbcH5c (lane 3), BRBA-S (lane 4), K0-Ub (lane 5), samples cross-linked with 0.05%, 0.1%, 0.2%, 0.4%, 0.6%, 0.8%, and 1% of glutaraldehyde (lane 6-12). B) Samples cross-linked with 0.05%, 0.1%, 0.2%, 0.4%, 0.6%, 0.8%, and 1% (lanes 1-7), and molecular weight marker (lane 8).

4.3. Conclusions

The experiments conducted, and results presented in this chapter lead to the determination of the optimal concentrations for glutaraldehyde cross-linking of histones in the histone octamer and the cross-linking of BRBA-S complex. Based on the obtained results from individual cross-linking experiments, cross-linking experiments were performed to determine the optimal glutaraldehyde concentration at which NCP-BRBA-S is stabilized. However, the mild glutaraldehyde concentrations of 0.01% and 0.025% indicated by the individual cross-linking experiments did not result in the stabilized complex for cryo-EM analysis. However, cross-linking of NCP_{K0-Ub}, UbcH5c and BRBA-S with 0.4% glutaraldehyde for a shorter period provided samples worth

analyzing using cryo-EM. The results obtained from cryo-EM studies are presented in the next chapter.

Chapter 5: Structural analysis by single-particle cryo-EM

“Learning never exhausts the mind.”

- Leonardo da Vinci

5.1. Introduction to electron microscopy

5.1.1. Background

For almost five decades X-ray crystallography was the preferred tool to determine high-resolution three-dimensional structures of biological macromolecules, but with limitations: (1) Many large multi-protein complexes refused to crystallize, (2) truncation of flexible sequences was used as a strategy to promote crystallization, (3) even after crystallization, it could hinder the investigation of molecular interactions as crystal packing favors some modes of interactions whereas disfavours other types, and (4) for crystallization, we need relatively large volumes of very homogenous samples at high protein concentrations.

To overcome these challenges, cryo-electron microscopy has emerged as a spectacularly successful structure determination method. Single particle cryo-electron microscopy was initially used as a tool to obtain low-resolution information about the morphology of higher molecular weight protein complexes that refused to crystallize. Although the steps involved have not significantly changed, recent developments in instrumentation, direct electron detection and software revolutionized cryo-EM resulting in a process that gives near atomic structures of multi-component complexes.²⁹⁷ The protein complexes of interest can be investigated in all the functional conformations. A critical pre-requisite to obtaining high-resolution structures is the characterization and optimizing the biochemistry to obtain homogenous samples,

choosing the right grids and grid preparation to get well-distributed molecules in the ice. Single particle cryo-EM is successful in determining the high-resolution structure of a wide range of molecules from remarkably symmetrical virus particles to ribosomes to nucleosomes to challenging membrane proteins. Figure 50 shows the entire workflow involved in cryo-electron microscopy.

Figure 50: Workflow in single particle Cryo-EM.

Taylor and Glaeser showed that crystals at a temperature below -120°C could be used to collect data with more prolonged exposure as it would reduce damage due to radiation.²⁹⁸ Richard Henderson at the MRC laboratory of molecular biology in Cambridge and colleagues elucidated the structure of bacteriorhodopsin at a resolution of $7\text{ \AA}$ using electron crystallography from images of unstained 2D crystals.²⁹⁹ Another seminal discovery came from the comprehensive work done by Jacques Dubochet and colleagues at European Molecular Biology Laboratory (EMBL). They found a way to freeze thin films of water (vitreous ice) fast enough to avoid the formation of ice crystals

known to damage biological samples and developed a manual plunge-freezing technique for grid preparation.³⁰⁰ The drawback with cryo-EM is high background and low contrast. Particles in a micrograph contain different sources of noise: (1) heterogeneity in the particles caused by preparation artifacts and beam induced damage, (2) irregularities in the structure of the carbon film, and (3) low contrast. To overcome these problems, Joachim Frank and his colleagues developed processor driven image averaging methods resulting in improved signal-to-noise ratio.³⁰¹ For their contribution to the development of cryo-EM, Richard Henderson, Jacques Dubochet, and Joachim Frank shared the 2017 Nobel prize in chemistry.

5.1.2. Direct electron detectors

A leap of change came with the use of the complementary metal oxide semiconductor (CMOS) based direct electron detectors, instead of films.^{302,303} In the direct electron detectors, electrons directly hit the less doped silicon layer followed by the supporting highly doped silicon base, with each frame being read continuously, commonly called the ‘movie’ mode. ‘Dope’ is the fabrication of detectors with desired chemical impurities to regulate its properties. CMOS detectors enabled bigger arrays, quicker readout, reduced noise, and reduced sensitivity to radiation damage, compared to charge coupled devices (CCDs). Increased speed and detector quantum efficiency (DQE) of CMOS detectors were critical in the ‘resolution revolution.’³⁰⁴ The ability to continuously read out images (movie mode) results in the collection of data with higher frame rates, eventually decreasing radiation damage.

‘Beam-induced motion’ and blurring of images has been a long-standing problem in electron microscopy. The use of direct electron detectors has the advantage of reducing beam-induced motion, improved signal-to-noise ratio by computational

alignment, correction of individual frames, and minimizes the backscattering of electrons from the substrate, leading to reduced noise.³⁰⁵ The precise identification of the position and orientation of each particle in the micrograph is enabled by motion correction, which forms the basis for obtaining high-resolution 3D structures of a molecule. The method of single particle cryo-electron microscopy for determining the high-resolution 3D structures of biological samples is still evolving with monthly improvements being reported.

5.1.3. Phase plate

One of the challenges in single particle cryo-electron microscopy is radiation damage caused to the biological samples by the beam of electrons. To overcome the problem of radiation damage, the biomolecules are vitrified by being embedded in a thin layer of vitreous ice. Frozen-hydrated biological samples have low contrast. Therefore, phase contrast has been applied to extract the maximum amount of information with a minimized electron dose. A common way of achieving phase contrast is by applying ‘defocus,’ termed defocus phase contrast (DPC). Another tool is ‘phase plate.’ Years ago, an effective phase plate was developed, called Zernike phase plate (ZPP), consisting of a thin carbon film (approximately 25 nm in thickness) and a hole in the center (about 500 nm). However the ZPP had many practical difficulties, as it has a lifespan of just a few days to a week and the precision needed to centre the hole of the phase plate in the beam path limiting automated data collection.^{306,307} The next significant development was the Volta phase plate (VPP).³⁰⁸ It is similar to that of ZPP but without a hole in the center. In VPP, the phase shift is generated when the electron beam strikes the carbon thin film, the phase shift is temporary and returns to its native state after a few days. Upon striking the carbon thin film with the electron beam, a negative potential is generated resulting in a $\pi/2$ phase shift. The primary advantage of

VPP is its reusability, long lifespan of up to few years and compatibility with automated data collection.

5.1.4. Electron microscopy for NCPs and NCP-chromatin remodeler complexes

The structure known from both crystallography²¹ and cryo-EM³⁰⁹ shows NCP is compact and produces increased contrast when visualized in the electron microscope due to the presence of DNA wrapping around the histone octamer. NCP is compact but relatively small for EM with a diameter of 10 nm and height of 5 nm. This relative increase in the contrast when compared to other protein is due to the higher electron scattering provided by the DNA. Sara Sandin's group at NTU used NCP as a model to validate the use of VPP to increase the contrast of biological macromolecules. The results obtained from the use of VPP showed it is possible to unambiguously visualize DNA and NCPs in thin ice, subsequently solving the structure of NCP to a resolution of 3.9 Å.³⁰⁹ It was shown that using VPP in cryo-EM increases the contrast and allows structural elucidation of challenging and relatively small biological macromolecules to near atomic resolution.³⁰⁹

5.2. Results

5.2.1. Cryo-EM analysis of non-cross-linked NCP-UbcH5c-BRBA-S complex

The NCPs, UbcH5c, and BRBA-S used for cryo-EM studies were pure and homogenous as shown by size-exclusion chromatography and analysis in denaturing gels. Investigation of the purified samples by ubiquitylation assays, analysis of the ubiquitylation reactions on a denaturing gel and native gel electrophoresis showed BRBA-S is enzymatically active. After purification of the proteins, biochemical characterization of the complex and optimization of grid conditions, EM micrographs were recorded. The NCPs, UbcH5c, and BRBA-S were mixed at a ratio of 1:4:4 to have

a saturated complex, without cross-linking for cryo-EM analysis. Quantifoil R1.2/1.3 holey carbon grids were used to prepare cryo-specimens with the NCP-UbcH5c-BRBA-S complex. Subsequently, the prepared grids were screened using Tecnai Arctica (200 kV) instrument with the phase plate. Cryo-EM analysis of the plunge-frozen, thin layers of a vitrified solution containing the sample on holey carbon grids showed a good distribution of particles (Figure 51).

Alignment and clustering of images in EM is challenging as a data set contains images having protein or protein complexes projected from different angles, resulting in a variety of arrangements. Image alignment and clustering is achieved by subjecting the micrographs through computational analysis. Figure 51 A & B show distribution of NCPs with the disk (indicated by a blue circle) and side (indicated by a red circle) orientations. The side orientation of the NCP was prevalent around the edges of the holes in the EM grid, many disk orientations were noticed in the center of the micrograph (Figure 51 C). The particles in the micrograph did not show aggregation or rafting. The use of VPP enhanced the contrast significantly, enabling us to identify the particles and orientations unambiguously. As the distribution of particles and the overall quality of the micrographs were satisfactory, we decided to proceed with the collection of a dataset for processing. The data was collected on the FEI Titan Krios (300 kV) using the phase plate. Parameters used during data collection are summarized in Table 3.

Figure 51: Cryo-EM screening of non-cross-linked NCP-UbcH5c-BRBA-S.

A and B) Orientations of particles when visualized with VPP. Disk orientation (blue) and side view (red) are indicated. C) Micrograph showing more side orientations towards the edge (black arrow) and disk orientations towards the center of the hole. Scale bars are 100 nm.

Table 3: Data collection and processing parameters for non-cross-linked NCP-UbcH5c-BRBA-S

Sample support	Quantifoil R1.2/1.3
Microscope	Titan Krios
Pixel size (Å)	1.1
Phase plate	VPP
Voltage (kV)	300
Detector	K2
Defocus (µm)	-1.5
Electron exposure (e⁻/Å²)	60
Exposure time (s)	12
Magnification	1,30,000
Micrographs recorded	1794
Micrographs used	1568
CTF	Corrected

Particles picked	2,57,992
Symmetry	C1
Box size	200
Map resolution (Å)	3.9
Fourier shell correlation (FSC) threshold	0.143
B factor	-160.6

5.2.2. Structural analysis of non-cross-linked NCP-UbcH5c-BRBA-S complex

The spatial resolution of a 3D model depends on the quality of the images recorded, number of images and structural homogeneity of the complex. A total of 1794 micrograph were recorded. The defocus parameters were estimated for the recorded micrographs to select the best micrographs from which the particles were picked. A total of 257792 particles were picked, and 35 classes of two-dimensional (2D) projections were obtained from 25 iteration steps (Figure 52). The classified 2D averages showed disk view, side view, and tilted views of the NCPs with clearly distinguishable features such as the DNA gyres, but the projections did not show distinctive extra density on either side of the NCP for the BRBA-S. 2D classes showing distinct features are shown in the first three rows of the classification matrix (Figure 52). Careful analysis of the 2D class averages, did not reveal any extra density, suggesting the absence of BRBA-S heterodimeric complex and UbcH5c. It was decided to move forward with the reconstruction of the 3D map to ensure the conclusion of lack of an additional density is unambiguous.

Figure 52: 2D projections of the non-cross-linked NCP-UbcH5c-BRBA-S sample categorized into 35 classes.

Thirty-five 2D class averages to which, 257792 particles contributed. Disk, side and tilted views are shown.

In order to reconstruct a 3D map and validate the obtained model, a low-pass filtered 60 Å NCP structure resembling just a disk was used as the reference model, commonly termed ‘initial model.’ A total of 257792 particles that contributed to 35 classes of 2D averages were used to generate five 3D classes. The number of particles that contributed to each 3D class was almost equal. Among the five 3D classes, the 3rd 3D class containing 48504 particles that were well refined exhibiting discernible features was chosen for further processing (Figure 53). After the reconstruction of the 3D volume, NCPs were observed with no extra density (BRBA-S heterodimeric complex and UbcH5c).

After post-processing and refinement, a resolution of 3.9 Å was estimated by plotting the gold standard FSC curve, at a threshold of 0.143 (Figure 54 A). Comparable to the already existing crystal structures of NCP^{13,19} and the phase-plate cryo-EM

structure of the NCP³⁰⁹ published by Sara Sandin's group, the refined 3D structure presented here clearly shows the 145 bp DNA wrapping around the histone octamer and at this resolution (3.9 Å), the DNA grooves and the alpha helices in the histone proteins are clearly discernible (Figure 54 B).

Figure 53: 7 Å resolution representation of the 3D class.

A and B) Disk and tilted view of the class selected for final 3D auto-refinement and post-processing.

B)

Figure 54: The FSC curve and the final 3.9 Å structure of the NCP.

A) The 'gold-standard' Fourier Shell Correlation (FSC) plot for the EM density map. B) Disk view, side view and tilted view of the 3.9 Å structure of NCP.

5.2.3. Cryo-EM analysis of cross-linked NCP_{K0-Ub}-UbcH5c-BRBA-S complex

There are three critical steps in grid preparation that decides the quality of the information obtained, 1) blotting to obtain a very thin film of the sample solution, 2) rapid plunge-freezing in liquid ethane to minimize ice formation and 3) choice of the grid. Cryo-EM analysis of the non-cross-linked sample containing NCP-UbcH5c-BRBA-S confirmed that the BRBA-S nor UbcH5c were bound to the NCP. As the non-cross-linked sample did not provide a stable NCP-UbcH5c-BRBA-S complex, with the intention of capturing the UbcH5c-BRBA-S complex bound to NCP, ubiquitylation reactions were initiated followed by cross-linking with different concentrations of glutaraldehyde at a time point. The cross-linked NCP_{K0-Ub}-UbcH5c-BRBA-S complex were screened using Tecnai Arctica (200 kV). The cryo-specimens containing cross-linked sample was prepared by plunge freezing.

The micrograph in Figure 55 B shows particle distribution of the cross-linked sample. Comparing the micrographs containing non-cross-linked and cross-linked sample shows a clear difference in the particle features captured in the ice (Figure 55 A

& B). Unlike the non-cross-linked sample, cross-linked complex show orientation bias, as side views and tilted views are not observed.

Figure 55: Comparison between micrographs showing non-cross-linked and cross-linked samples.

A) Micrograph showing non-cross-linked sample. B) Micrograph exhibiting cross-linked complex.

It has been shown by electron cryo-tomography (cryo-ET) that almost all the proteins or protein complexes are adsorbed in the air-water interface.^{310,311} The preferred orientation could be due to the properties of the cross-linked sample or due to the adsorption of the cross-linked complex in the air-water interface.³¹¹ Although the cause for the orientation bias is not very clear, it is indeed different from the non-cross-linked sample, visually. Therefore, it is logical to collect a dataset and process the data to analyze the structure of the cross-linked complex at the 2D and 3D level.

5.2.4. Data collection and 2D classification of cross-linked complex

Data collection was performed using Titan Krios (300 kV) transmission electron microscope (TEM). The parameters used during data collection and processing of the dataset are given in Table 4. A total of 2468 micrographs were recorded for processing.

The recorded micrographs were subjected to CTF estimation to deduce the defocus and astigmatism parameters before particle selection and picking for 2D class averaging. Post CTF estimation, 1985 micrographs were selected from which the particles were auto picked for further processing. The micrographs in Figure 56 show particles from the cross-linked complex. From 1985 micrographs, a total of 390555 particles were picked.

Table 4: Data collection and processing parameters for cross-linked $NCP_{K0-Ub-UbcH5c-BRBA-S}$ complex

Sample support	Quantifoil R1.2/1.3
Microscope	Titan Krios
Pixel size (Å)	1.1
Phase plate	VPP
Voltage (kV)	300
Detector	K2
Defocus (µm)	-1.5
Electron exposure ($e^-/\text{Å}^2$)	60
Exposure time (s)	12
Magnification	1,30,000
Micrographs recorded	2468
Micrographs used	1985
CTF	Corrected
Particles picked	3,90,555
Symmetry	C1
Box size	200
Map resolution (Å)	6.3
Fourier shell correlation (FSC) threshold	0.143
B factor	-217.2

Figure 56: Representative micrographs of the cross-linked NCP_{K0-Ub} -UbcH5c-BRBA-S complex.

390555 particles picked were used to generate 70 2D class averages in 25 iteration steps. Identification of homogenous subsets, increasing the quality of group averages and elimination of bias from 2D image clustering in EM is a crucial step in single-particle cryo-EM analysis. This is achieved by performing classification in 25-50 iterations. The 2D classes show a mix of good, bad and noisy classes (Figure 57).

Figure 57: Gallery of 2D class averages from 390555 particles.
Tile images showing a mix of good, bad and noisy class averages.

Careful analysis of the 2D class averages showed a blurry extra density on one side of the NCP in the good class averages, suggesting that the density seen could be due to the binding of BRBA-S and/or UbcH5c to the NCP_{K0-Ub} (Figure 58). Although this presumption gave hope, it is also possible that the seen blurry density could be noise. This ambiguity of whether the BRBA-S heterodimer and/or UbcH5c is bound to NCP could be resolved by processing the data further to obtain the final 3D volume. Therefore, the particles that contributed to the good 2D class averages were selected for 3D classification of the complex. From the total class averages, 115000 particles that contributed to the good 2D class averages were chosen for 3D classification of the complex (Figure 58).

Figure 58: Gallery of good 2D class averages.

Class averages are showing particles with blurry extra density (specified by a white star) and DNA unwrapping (specified by the yellow star).

The selected 115000 particles produced five 3D classes, using low pass filtered NCP as the reference model. Each class showed a 3D EM map, but the 4th class among the five classes resolved well and showed clear extra density on one side of the NCP. A total of 37600 particles contributed to the generation of the 4th class. The 4th 3D class

had a higher resolution and showed clear extra density on the disk surface of the NCP with discernible features. Three arguments convinced us that the extra density is valid, 1) the extra density is observed on the disk surface of the NCP, 2) the extra density showed features resembling protein secondary structures, and 3) the extra density showed similarity to the structurally related RING1B-BMI1 complex bound to the NCP. After post-processing and refinement, the resolution of the map representing the NCP was determined to be around 7 Å, and the density corresponding to the bound volume was estimated to be around 8-12 Å. The FSC curve was plotted with 0.143 as the cutoff value (Figure 59).

Figure 59: The ‘gold-standard’ Fourier Shell Correlation (FSC) plot for the reconstructed EM density map.

Post-reconstruction of a 3D electron density map, it is essential to validate the reconstructed model with a known structure, since it was NCP and BRBA-S, the reconstructed 3D volume was validated by fitting the 3.9 Å structure of the NCP determined by phase-plate cryo-EM (accession code: EMDB-8140). Subsequently, the published NMR structure of the RING domain-containing BRCA1¹⁻¹⁰³-BARD1²⁶⁻¹²²,

hereafter called as BR^{RING}-BA^{RING} (accession code: PDB-1JM7)² was fitted to electron density 3D map to validate the obtained 3D model. The fit is shown in Figure 60 and Figure 61. An extra density at the side of the NCP is observed. Since glutaraldehyde reacts with lysines, the lysine residues that primarily hold the DNA on the surface of the histone octamer are affected. Upon breathing of the DNA at the edges of the nucleosome, lysines become modified that are unable to rebind the DNA. It could also be that the unwrapping of the DNA could be the result of BRBA-S binding to the surface of the NCP, as the side at which the DNA unwraps and the face of the NCP that binds the BRBA-S correlates with each other. The face of the NCP that does not show extra density has the DNA intact on its side. Figure 60 A, B, and C, show disk view, side view, and tilted view of the obtained 3D model fitted with the NCP. The extra density is clearly seen. The missing density of the nucleosomal DNA as seen in the 3D map (Figure 60 A) correlates with the unwrapping of the DNA visualized in the 2D class averages (Figure 58).

Figure 60: Representation of the final 3D map fitted with the NCP structure obtained using phase plate.

Illustration showing the disk, side and tilted views of the reconstructed 3D map. The reconstructed map is shown in 'grey,' and the fitted NCP is shown in 'blue.' The region where

the nucleosomal DNA starts to dissociate is indicated in 'dashed' red and the region where the nucleosomal DNA is missing in the reconstructed 3D volume is marked in 'arced' red line.

Based on the location and size of the extra density on one face of the NCP and the features of that density resembling protein secondary structure, we believe that the extra density on one side of the NCP represent BRBA-S. However, there was challenge in fitting the available BR^{RING}-BA^{RING} within the visible electron density. The challenge in fitting the available BR^{RING}-BA^{RING} structure into the obtained electron density is due to lower resolution and structural heterogeneity caused by flexibility. The obtained surface map did not show any distinctive features that could be compared to the UbcH5c E2 conjugating enzyme. Although the extra density on one side of the NCP was validated, at this point, it is not possible to describe the cause for the absence of density on the other face of the NCP that also contains a binding site for BRBA-S as the NCP has two-fold symmetry.

Figure 61: Fitting of BRCA1-BARD1 heterodimeric RING-RING complex into EM surface map.

A) Illustration of the surface and ribbon representation of the BRCA1-BARD1 heterodimeric RING-RING complex fitted in the EM map, shown in disk view. B) The depiction of the surface and ribbon representation of the BRCA1-BARD1 heterodimeric RING-RING complex fitted in the EM map, shown in side view. C) Image of the surface and ribbon representation of the BRCA1-BARD1 heterodimeric RING-RING complex fitted in the NCP EM map, shown in the tilted view. Reconstructed EM map is shown in 'grey,' NCP structure obtained by cryo-EM is highlighted in 'red,' BRCA1 is marked in 'green' and BARD1 is indicated in 'blue.'

However, some visual cues helped in assessing the credibility of the model. NCP-UbcH5c-RING1B-BMI1 complex structure published by Song Tan's group showed RING1B-BMI1 (PRC1) binding to the acidic patch of the NCP.⁶⁴ Based on this structure and sequence comparison between the acidic patch binding loop in RING1B (a subunit of the PRC1 complex) and the comparable loop in BRCA1 expected to bind to the acidic patch of the NCP, in the 3D model obtained, the extra density is positioned on the acidic patch of the NCP and hence is consistent with the interpretation presented (Figure 61 A). The sequence comparison between the NCP binding loop in RING1B and BRCA1 is shown in Chapter 6. Figure 61 B & C shows the obtained 3D map fitted with NCP and BR^{RING}-BA^{RING}. The fitting confirms that the size of the extra density on the NCP acidic patch matches the size of the BR^{RING}-BA^{RING}, which implies the fitted representation is reliable.

5.3. Conclusions

Non-cross-linked and cross-linked samples containing NCP-UbcH5c-BRBA-S was screened using Cryo-EM. The non-cross-linked sample showed a good distribution of particles in the grid with features resembling only NCPs. The cross-linked sample showed particle features differing from non-cross-linked particles, along with preferential orientation. Processing of the recorded cryo-EM datasets for both non-cross-linked and cross-linked samples confirmed that the non-cross-linked sample did not have BRBA-S or UbcH5c bound to the NCP. The 3D volume reconstructed from the data recorded for the non-cross-linked sample resulted in a 3.9 Å NCP structure.

Processing of data collected by cryo-EM for the cross-linked sample resulted in a 3D map with extra density resembling features of the protein secondary structure. The final cryo-EM 3D map had the resolution of the NCP estimated at 7 Å and resolution

representing the extra density approximated at 8-12 Å. Fitting of the published solution NMR structure of BR^{RING}-BA^{RING} into the obtained density map, though fitted to a degree, was not perfect. The imperfect fitting is due to the low-resolution and flexibility. Structural information elucidated and presented in this thesis will be crucial in conducting more biochemical experiments, collecting large volumes of cryo-EM datasets to obtain a high-resolution structure of the complex.

Chapter 6: Discussion

“Human curiosity, the urge to know, is a powerful force and is perhaps the best secret weapon of all the struggle to unravel the workings of the natural world.”

- Aaron Klug

6.1. Cellular functions and physiological importance of BRCA1-BARD1

The importance of epigenetics in regulating cellular processes whose substrate is DNA is unquestionable. One bulky histone modification that is involved in the DNA damage repair and genomic stability is ubiquitylation of histones.^{3,312,313} The last step in the ubiquitylation cascade, the transfer of ubiquitin from E2 conjugating enzyme to the substrate is mediated by enzymes called ‘E3 ubiquitin ligases’ that bind to E2 conjugating enzyme and the substrate to catalyse the transfer of ubiquitin to the substrate. RING domain containing E3 ligases are the largest class of ubiquitin ligases.³¹⁴ BRCA1-BARD1 are particularly interesting because of their functional diversity as a clinically significant tumour suppressor protein conferring E3 ligase activity and belongs to the RING family of E3 ligases. The biological functions of BRCA1 depend on its ability to form a heterodimer with its associated proteins BARD1, another RING domain-containing protein.^{1,315,316} BRCA1 E3 ligase-mediated ubiquitylation of histone H2A is crucial for DNA damage repair and heterochromatin-mediated silencing.^{3,220,317} The thesis focuses on understanding how BRCA1-BARD1 recognizes the NCP.

6.2. BRCA1-BARD1 ubiquitylation and interaction with NCP

BRBA-S and BRBA-L complexes were co-expressed and purified from *E.coli* and Sf9 cells, respectively. Ubiquitylation assays performed in the absence of a substrate

showed BRBA-S itself acts as a substrate and self-ubiquitylates. Ubiquitylation time courses to compare the E3 ligase activity between the minimal RING domain complex BRBA-S, and the additional N- and C-terminal sequence containing complex BRBA-L performed with both wt-Ub and K0-Ub confirmed that the additional sequences adjacent to the RING domains contribute to the increase in the E3 ligase activity. Having known the additional sequence increases the activity, the next right thing to do is to perform mutation experiments to identify the specific amino acids that contribute to the enhanced activity.

It was established that BRBA-S and BRBA-L in the presence of UbcH5c and wt-Ub, poly-ubiquitylated histone H2A in the nucleosomal context. BRBA-S and BRBA-L dependent ubiquitylation assay showed mono-ubiquitylation of histone H2A in the presence of UbcH5c and K0-Ub. The three bands that represented mono-ubiquitylated H2A in my time course experiments with K0-Ub, substantiate the *in vitro* mutation experiment results reported by Kevin Hiom's group showing K125, K127 and K129 in H2A are the sites of ubiquitylation *in vitro*.

6.3. Structure of NCP-BRBA-S complex obtained by cryo-EM and NCP-RING1B-BMI1 complex

The final and most important part of the work presented in this thesis concludes in determining the low-resolution structure of the RING domain-containing N-terminal sequence of BRCA-BARD1 in complex with NCP using cryo-EM. I was able to stabilize the complex and demonstrated a way to solve the low-resolution structure of the complex.

Although a thorough comparison between the NCP-BRBA-S complex and NCP-RING1B-BMI complex would not be possible with the structure currently available,

comparison focusing on the sequence similarity between the NCP binding loop in RING1B and comparable loop in BRCA1, an insight about the region in which BRBA-S binds to the NCP provides validity for the obtained 3D structure. Sequence comparison of the two E3 ligases, BRCA1 and RING1B, are presented in this chapter.

Although the obtained 3D EM model was validated by fitting, there are two observations that needs to be addressed, 1) the extra density seen on only one face of the NCP, and 2) the sample flexibility. The NCPs are dynamic particles varying in their dynamic and structural properties, resulting in distinct functionalities. The visualization of the extra density on one face of the NCP and not on the other could possibly be due to the change in the structural conformation of the NCP, a property noticed even during modification of histones.⁷⁹ When BRBA-S binds to one face of the NCP, the binding might induce a change in the overall structure of the NCP preventing another molecule from binding to the other surface of the NCP. The effect could be a result of increased degree of unwrapping or breathing of NCPs.^{79,318} Another possibility is that the BRBA-S binding to the NCP does not result in a saturated complex, resulting in extra density on only one side of the NCP. Interestingly, the surface of the NCP which shows the extra density corresponding to the BRBA-S and the unwrapping of the DNA appears to be on the same side of the NCP, implying correlation between the binding and DNA unwrapping only on the side where the BRBA-S binds, among the available two DNA ends. This correlation needs to be validated by conducting further biochemical investigations.

Sample heterogeneity induced by flexibility, that is autonomous of mechanical drift of the microscope stage and beam induced motion in the course of data acquisition curtails the quality of the recorded micrographs.^{319,320} Flexibility lead to movement of

regions of molecules, with the molecule being unfixed in one single position. During data recording, the final micrograph obtained is a result of the sum of all these flexible movements, that compromises the quality of the data recorded. The quality of the recorded micrographs is an important criterion that decides the final resolution. Accounting for the missing density on one face of the NCP and the flexibility, the final 3D structure provides an insight into the BRBA-S binding to the acidic patch region of the NCP, validated by fitting the BR^{RING}-BA^{RING} (PDBID: 1JM7) solution structure elucidated by NMR.

The first high-resolution structure of an E3 ligase showing the recognition of the acidic patch on the NCP disk surface was NCP-UbcH5c-RING1B-BMI1 complex at a resolution of 3.3 Å.⁶⁴ RING1B-BMI1 is an E3 ligase complex that belongs to the Polycomb family (PcG) first implicated in the *Hox* gene silencing in *Drosophila*.³²¹ RING1B-BMI1 is a histone-modifying enzyme associated with transcriptional repression by mono-ubiquitylation of histone H2A at K119.^{248,250} PRC1 is associated with the regulation of transcriptional repression by monoubiquitylation of residue K119 of histone H2A and chromatin compaction.⁹⁷ RING1B-BMI1 E3 ligase can pair with one of many E2 conjugating enzymes, including UbcH5c.³²² In the NCP-UbcH5c-RING1B-BMI1 complex, in order to overcome the challenge of weak binding affinity and salt-sensitive E2-E3 interaction, the C-terminus of the RING1B minimal RING construct was fused to the N-terminus of UbcH5c.^{64,323} The final 3D EM model of NCP_{K0-Ub}-UbcH5c-BRBA-S complex did not show extra density comparable to UbcH5c, the reason could be due to the low-affinity interaction, as it was not fused to BRCA1 or BARD1.

RING1B-BMI1 bind to the histone surfaces, along with the H2A-H2B acidic patch.^{63,64} K65, K85, K93, V95, R81, R98, K97 and K93 residues of the RING1B govern the interaction with the NCP disk. R98 of RING1B interacts with E61, D90, E92 of H2A and E105, H109 of H2B.⁶⁴ R81 of RING1B hydrogen bonds with D72 and N89.⁶⁴ The vital information elucidated from the structure of NCP-RING1B-BMI1 complex is that the interactions between the RING1B-BMI1 and NCP is distant from the catalytic site.⁶⁴ Structure of NCP-RING1B-BMI1 complex is shown in Figure 62.

Figure 62: Crystal structure of RING1B-BMI1-NCP complex.⁶⁴

A) Orthogonal view of the complex with RING1B indicated in 'cyan' and BMI1 indicated in 'yellow.' B) Disk view of the complex showing RING1B-BMI1 heterodimer interaction with the histone surface. Crystal structure incorporates the minimal RING domain-containing heterodimeric complex RING1B²⁻¹¹⁶-BMI1²⁻¹⁰⁹.

Due to the symmetrical nature of the NCPs, one would expect that the binding would be identical on both sides of the NCP. Song Tan's group has reported variations between the binding of RING1B-BMI1 to one side of the NCP over the other.⁶⁴ Song Tan's group suspect that the variation in the two halves of the structure result from

different environments of each RING1B-BMI1 complex within the crystal lattice.⁶⁴ Because of this variation in binding, the UbcH5c is placed too far from the surface of the NCP on one side compared to the other. In the NCP-BRBA-S structure we reconstructed, the extra density is seen only on one side of the NCP and not on the other. In the NCP-RING1B-BMI1 complex, the RING1B-BMI1 is shown to bind on both sides of NCP, for comparison with the NCP-BRBA-S, the binding of RING1B-BMI1 to one side of the NCP is only shown in the following depictions.

Figure 63 & Figure 64 shows the structural and binding similarity between the NCP-RING1B-BMI1 and NCP-BRBA-S complexes. Figure 63 A exhibits the obtained structure of the NCP-BRBA-S complex in surface representation, and Figure 63 B displays the published structure of the NCP-RING1B-BMI1 complex in surface representation. Figure 63 C shows the disk view of the obtained 3D map fitted with NCP and BR^{RING}-BA^{RING} in cartoon representation. Top view or the disk view of the NCP bound by RING1B-BMI1 E3 ligase is shown in Figure 63 D. A magnified disk view of the obtained 3D map fitted with NCP and BR^{RING}-BA^{RING} in cartoon representation is shown in Figure 63 E.

Figure 63: A comparison between the final 3D model of NCP_{K0-Ub}-BRBA-S and published NCP-RING1B-BMI1 structure.

A) Low-resolution structure of NCP-BRBA-S complex, BRCA1 shown in 'green' and BARD1 shown in 'blue.' B) Published structure of NCP-RING1B-BMI1 complex, RING1B shown in 'cyan' and BMI1 shown in 'yellow. NCP is shown in 'red.' C) Cartoon representation of the BRBA-S heterodimer fitted in the extra density of the obtained 3D EM map. D) Cartoon representation of the NCP-RING1B-BMI1 complex. E) Higher-magnification cartoon representation of the BRBA-S heterodimer fitted in the extra density of the obtained 3D EM map, acidic patch is marked in 'yellow' and a surface representation of the NCP with the acidic patch marked in 'yellow' is presented for visual reference.

Similar to what is shown in the published high-resolution structure of NCP-RING1B-BMI1 complex, BRBA-S heterodimer also binds to the central histone surface of the NCP. Based on the comparison of the structures, the histone surface on which the extra density is seen validates our final 3D model.

Figure 64: Side view comparison between the final 3D model of BRBA-S-NCP and published RING1B-BMI1-NCP structure.

E) Final 3D model of NCP-BRBA-S complex, BRCA1 shown in 'green' and BARD1 shown in 'blue.' F) Published structure of NCP-RING1B-BMI1 complex, RING1B shown in 'cyan' and BMI1 shown in 'yellow. NCP is shown in 'red.'

6.4. Sequence comparison of BRCA1 and RING1B

Sequence comparison between E3 ubiquitin ligases RING1B and BRCA1 exhibit high degree of conservation between the NCP binding loop of RING1B and the comparable loop in the BRCA1 tumor suppressor E3 ligase protein.⁶⁴ This gives a hint that similar to RING1B, BRCA1 might use the similar lysine and arginine residues that are conserved to bind to the NCP. However, unlike H2A K119 which is ubiquitylated by RING1B, BRCA1 ubiquitylates K125, K127, and K129 *in vitro*. Due to the difference in the lysine residues that are ubiquitylated, there is a possibility that BRCA1-

BARD1 interaction with the NCP might be slightly different. In order to have a thorough understanding of how BRCA1-BARD1 interacts and binds to the NCP, obtaining the high-resolution molecular model is crucial. Figure 65 A shows NCP binding loop in BRCA1, highlighted by colored spheres. Sequence comparison between RING1B⁸⁷⁻¹⁰¹ and BRCA1⁶¹⁻⁷⁴ is shown in Figure 65 B.

Figure 65: BRCA1 and RING1B require similar loop region for NCP recognition and binding.

A) NCP binding loop in BRCA1. BRCA1 is highlighted in 'green' and BARD1 is highlighted in 'blue.' B) Alignment of sequences between BRCA1 and RING1B shows similar binding loop required for interacting with the nucleosome and for zinc binding. The figure representing sequence comparison between RING1B⁸⁷⁻¹⁰¹ and BRCA1⁶¹⁻⁷⁴ was based on the comparison presented in McGinty et al.⁶⁴

6.5. Interaction of chromatin remodelers with NCP

Like what is expected in the interaction between NCP and BRBA-S complex, where the R71 in BRCA1 is probable to bind to the NCP, already published chromatin remodelers interact with the NCP via their critical arginine residues. The published

structures of the chromatin remodelers in complex with the NCP is discussed in this section. In 2006, the crystal structure of Latency-associated nuclear antigen (LANA) peptide-NCP complex was solved at 2.9 Å resolution. The N-terminal residues (1-23) of the LANA protein, associated with Kaposi's sarcoma-associated herpesvirus binds to the H2A-H2B acidic patch.⁵⁸ The specific interaction consists of a hairpin-like structure that enters in the deep groove of the acidic patch, contributing to hydrophobic and charged interactions with acidic patch residues. LANA peptide arginine-anchor residues, R7, and R9, interact with residues of H2B and H2A, respectively. R7 interacts with E110 of H2B, and R9 interacts with E61, D90, & D92 of H2A, interactions are via charged interactions. Mutation of R9 eliminates binding of the LANA peptide to nucleosome. Though arginine residues are the critical residues for binding to nucleosomes, hydrophobic residues M6 and L8 interact with residues Y50, V54, Y57 of H2A, the residues that form a hydrophobic patch next to the nucleosome acidic surface.

The 2.9 Å crystal structure of the NCP-RCC1 complex, obtained by co-crystallization showed RCC1 binding to the acidic patch and the nucleosomal DNA via the RCC1 N-terminal tail and the two loops of the β propeller sheets in RCC.⁵⁷ The loop containing arginine anchor R223 binds to residues E61, D90, E92 of H2A through ionic interactions and hydrogen bonds, like similar to that of arginine anchor residue R9 of LANA peptide. Another critical residue in RCC1 is the R216 that binds to E61 and E64 of H2A. The side chains of K241, R239, and Q259 of the second loop interact with the phosphate backbone of the DNA.³²⁴ Apart from the critical residues involved in binding that are mentioned above, other residues including S214 and S227 interact with residues on H2A to provide additional anchoring points.

Silent information regulator (SIR) proteins by *Saccharomyces cerevisiae* maintain chromatin in a transcriptionally repressive form specifically at telomeres.³²⁵ In Sir3-nucleosome interaction, the surface of the nucleosome participating in recognition of $\alpha 3$ helix of H2B, L1 loop and $\alpha 1$ helix of H3 as well as the L2 loop and $\alpha 2$ helix of H4. The 3.0 Å crystal structure of Sir3-NCP complex revealed the loop 1 of the Sir3 BAH domain also interacts with the H2A-H2B acidic patch.³²⁶ Arginine residues R28, R29, R30, R32, and R34, align on the groove of the acidic surface to establish charge dependent interactions with the nucleosome surface. Interestingly, R32 engages the same site as R9 of the LANA peptide and R223 of RCC1. In addition to the acidic surface, BAH interacts with the surface of H3-H4, H4 N-terminal tail, and H2B C-terminal helix through residues from A1 helix, B5 β strand, loops 2 and 4. The negatively charged Sir3 BAH domain surface and positively nature of the histone H4 tail makes the interaction between them mostly electrostatic. An L3 loop of the Sir3 BAH domain gets structured upon binding to NCP resulting in a 30-fold increase in the binding affinity to the NCP, and the acetylation of Sir3 N-terminus stabilizes interaction with the NCP further.⁶⁰ Two commonly noticed chromatin marks, H4 K16 acetylation, and H3K79 methylation disengages Sir3, canonical histone marks implicated in transcriptional activation and silencing.^{327–329} Similar to the LANA peptide, RCC1, Sir3, CENP-C (a histone H3 variant in centromeric chromatin) also use R717 and R719 arginine-anchor residues to interact with the acidic patch.³³⁰

High mobility group nucleosome binding (HMGN2) proteins are thought to decompact chromatin and are involved in DNA repair and histone post-translational modifications.^{331–333} Structural characteristics of the HMGN2 protein binding to NCP was carried out using NMR. HMGN2 bind to the acidic patch and the DNA in the NCP. The evolutionarily conserved N-terminal sequence RRSARLSA of HMGN2 protein

interacts with residues L65 of H2A and V45, L103 of H2B.^{59,334} In conclusion, the cryo-EM model presented in this thesis confirms that the BRCA1-BARD1 binds to the NCP acidic patch, like most of the chromatin remodeling enzymes that interact with the NCP disk. The details of the interaction between the BRBA-S and NCP will require a higher resolution structure of the NCP-BRBA-S complex.

A)

B)

Figure 66: Nucleosome ‘acidic patch’ recognition by the binding factors.⁵⁶

A) Nucleosome-chromatin remodeler complexes showing the interaction between the arginine anchor motif and acidic patch on the nucleosome B) Magnified view of the arginine residues from RCC1, Sir3, RING1B, LANA, and CENP-C shown with H2A side chains (represented in the sticks).

6.6. Improving the resolution of EM structures

The key step in obtaining high resolution EM structures is in the preparation of EM grids to provide images where the particles are visible in ice. In the field it was generally thought that the optimal method was to use holey carbon grids without any support. However, it is often necessary to have a support, but amorphous carbon is unideal as a support for cryo-EM as it increases the background and reduces contrast.^{335,336} More recently graphene and graphene oxide coated grids are being used as they provide a very homogenously thin layer and better contrast.³³⁷ In addition and surprisingly, it has recently emerged that protein samples get rapidly denatured in the air-water interface and denaturing of the proteins can be prevented by having a surface.³¹¹ Until recently, the use of holey carbon or grids with carbon support have been extensive in cryo-EM studies. Currently, the use of graphene or graphene oxide grids for single-particle cryo-EM analysis is rising due to the advantages they offer.³³⁸ Graphene is the thinnest, strongest material, a very good conductor, and is very stable in an electron beam.^{338–340}

The period between sample application onto the grid and vitrification provides ample time for the proteins in the buffer to form an insoluble layer at the air-water interface, as the time taken for the proteins to diffuse to the air-water interface is in milliseconds.³⁴¹ It has been reported that the air-water interface is an inhospitable environment for the proteins, resulting in denaturation of proteins and facilitating particle orientation bias.^{310,342,343} Adsorption and denaturation of proteins at the air-water interface eventually result in lesser number of particles that are used in the reconstruction of the 3D structure.

A method to prevent protein denaturation and orientation bias at the air-water interface is by using hydrophilized graphene grids.³¹¹ In graphene grids, it has been shown that the adsorption of proteins at the air-water interface can be adjusted by increasing or decreasing the concentration of the chemical compounds used in non-covalent chemical doping, by capitalizing on the property of the π - π stacking interaction between graphene and chemical compounds like 1-pyrenecarboxylic acid.³⁴⁴ Having understood the advantage of using hydrophilized graphene grids for cryo-EM, it would be reasonable to prepare hydrophilized graphene grids with cross-linked NCP_{K0-Ub}-UbcH5c-BRBA-S complex, collect a larger dataset to have more particles for reconstruction of a higher resolution 3D model and investigate the complex. Countering sample flexibility can be achieved at the 2D and the 3D level. At the level of 3D reconstruction and refinement, masking of definite regions of interest can enhance the resolution of the individual domains, as achieved in γ -secretase.^{345,346}

6.7. Concluding remarks

Based on the low-resolution structural information I have obtained, It would be reasonable to say with confidence that BRCA1-BARD1 and RING1B-BMI1, similarly bind to the disk surface of the NCP. Though BRCA1-BARD1 and RING1B-BMI1 are expected to bind similarly, additional contacts may be expected and the orientation of BRCA1-BARD1 could possibly be different when compared to RING1B-BMI1, due to the positioning of the K125, K127, K129 residues on the histone H2A tail. A mechanistic understanding at the molecular level demands a high-resolution structure. Based on this primary structural information obtained by cryo-EM, it is motivating to center the focus on biochemical experiments, trying alternative sample support (e.g., graphene), and collecting more datasets using cryo-EM to solve the high-resolution structure of the complex that can shed light at the molecular level.

References

1. Baer, R. & Ludwig, T. The BRCA1/BARD1 heterodimer, a tumor suppressor complex with ubiquitin E3 ligase activity. *Curr. Opin. Genet. Dev.* **12**, 86–91 (2002).
2. Brzovic, P. S. *et al.* Structure of a BRCA1-BARD1 heterodimeric RING-RING complex. *Nat. Struct. Biol.* **8**, 833-837 (2001).
3. Kalb, R. *et al.* BRCA1 is a histone-H2A-specific ubiquitin ligase. *Cell Rep.* **8**, 999-1005 (2014).
4. Annunziato, A. T. DNA packaging: nucleosomes and chromatin. *Nat. Educ.* **1**, 1 (2008).
5. Pierce, B. A. Genetics: a conceptual approach. (W. H. Freeman, 2013). doi:10.1016/0002-9343 (86) 90442-0
6. Kornberg, R. D. Chromatin structure: a repeating unit of histones and DNA. *Science*. **184**, 868-871 (1974).
7. Hewish, D. R. & Burgoyne, L. A. Chromatin sub-structure. The digestion of chromatin DNA at regularly spaced sites by a nuclear deoxyribonuclease. *Biochem. Biophys. Res. Commun.* **52**, 504-510 (1973).
8. Dekker, J. *et al.* Electron microscopic and biochemical evidence that chromatin structure is a repeating Unit. *Cell*. **4**, 281-300 (1975).
9. Heitz, E. D. Heterochromatin der Moose. *Jahrbücher für Wissenschaftliche Bot.* **69**, 762-818 (1928).
10. Passarge, E. Emil Heitz and the concept of heterochromatin: longitudinal chromosome differentiation was recognized fifty years ago. *J. Hum. Genet.* **31**, 106-115 (1979).
11. Alberts, B. *et al.* Molecular biology of the cell. *Garl. Sci.* 608 (2008).
12. Kornberg, R. D. & Thomas, J. O. Chromatin structure; oligomers of the histones. *Science*. **80**, 865-868 (1975).
13. Richmond, T. J. *et al.* Structure of the nucleosome core particle at 7 Å resolution. *Nature* **311**, 532-537 (1984).
14. Finch, J. T. *et al.* X-ray diffraction study of a new crystal form of the nucleosome core showing higher resolution. *J. Mol. Biol.* **145**, 757-769 (1981).
15. Holde, K. E. *Chromatin*. (Springer Verlag, 1989).
16. Richmond, T. J., Finch, J. T. & Klug, A. Studies of nucleosome structure. *Cold Spring Harb. Symp. Quant. Biol.* **47**, 493-501 (1982).
17. Richmond, T. J., Searles, M. A. & Simpson, R. T. Crystals of a nucleosome core particle containing defined sequence DNA. *J. Mol. Biol.* **199**, 161-170. (1988).
18. Harp, J. M. *et al.* Preparative separation of nucleosome core particles containing defined-sequence DNA in multiple translational phases. *Electrophoresis*. **16**, 1861-1864 (1995).
19. Luger, K. *et al.* Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature*. **389**, 251-260 (1997).
20. Lowary, P. T. & Widom, J. New DNA sequence rules for high affinity binding to histone octamer and sequence-directed nucleosome positioning. *J. Mol. Biol.* **276**, 19-42 (1998).
21. Richmond, R. K. *et al.* Crystal structure of the nucleosome resolution core particle at 2.8 Å. *Nature*. **389**, 251-260 (1997).

22. Arents, G. & Moudrianakis, E. N. The histone fold: a ubiquitous architectural motif utilized in DNA compaction and protein dimerization. *Proc. Natl. Acad. Sci.* **92**, 11170-11174 (1995).
23. Arents, G. *et al.* The nucleosomal core histone octamer at 3.1 Å resolution: a tripartite protein assembly and a left-handed superhelix. *Proc. Natl. Acad. Sci.* **88**, 10148-10152 (1991).
24. Flaus, A. *et al.* Mapping nucleosome position at single base-pair resolution by using site-directed hydroxyl radicals. *Proc. Natl. Acad. Sci.* **93**, 1370-1375 (1996).
25. Redon, C. *et al.* Histone H2A variants H2AX and H2AZ. *Curr. Opin. Genet. Dev.* **12**, 162-169 (2002).
26. Chadwick, B. P. & Willard, H. F. A novel chromatin protein, distantly related to histone H2A, is largely excluded from the inactive X chromosome. *J. Cell Biol.* **152**, 375-384 (2001).
27. Fan, J. Y. *et al.* H2A.Z alters the nucleosome surface to promote HP1 α -mediated chromatin fiber folding. *Mol. Cell.* **16**, 655-661 (2004).
28. Thoma & Koller. Influence of histone H1 on chromatin structure. *Cell* **12**, 101-107 (1977).
29. Simon, M. *et al.* Histone fold modifications control nucleosome unwrapping and disassembly. *Proc. Natl. Acad. Sci.* **108**, 12711-12716 (2011).
30. North, J. A. *et al.* Histone H3 phosphorylation near the nucleosome dyad alters chromatin structure. *Nucleic Acids Res.* **42**, 4922-4933 (2014).
31. North, J. A. *et al.* ATP-dependent nucleosome unwrapping catalyzed by human RAD51. *Nucleic Acids Res.* **41**, 7302-7312 (2013).
32. Manohar, M. *et al.* Acetylation of histone H3 at the nucleosome dyad alters DNA-histone binding. *J. Biol. Chem.* **284**, 23312-23321 (2009).
33. Lu, X. *et al.* The effect of H3K79 dimethylation and H4K20 trimethylation on nucleosome and chromatin structure. *Nat. Struct. Mol. Biol.* **15**, 1122-1124 (2008).
34. Thoma, F., Koller, T. H. & Klug, A. Involvement of histone H1 in the organization of the nucleosome and of the salt-dependent superstructures of chromatin. *J. Cell Biol.* **83**, 403-427 (1979).
35. Kornberg, R. D. Structure of Chromatin. *Annu. Rev. Biochem.* **46**, 931-954 (1977).
36. Widom, J. A relationship between the helical twist of DNA and the ordered positioning of nucleosomes in all eukaryotic cells. *Proc. Natl. Acad. Sci.* **89**, 1095-1099 (1992).
37. Kan, P. Y., Caterino, T. L. & Hayes, J. J. The H4 tail domain participates in intra- and internucleosome interactions with protein and DNA during folding and oligomerization of nucleosome arrays. *Mol. Cell. Biol.* **29**, 538-546 (2009).
38. Tse, C. & Hansen, J. C. Hybrid trypsinized nucleosomal arrays: Identification of multiple functional roles of the H2A/H2B and H3/H4 N-Termini in chromatin fiber compaction. *Biochemistry* **36**, 11381-11388 (1997).
39. Dorigo, B. *et al.* Chromatin fiber folding: Requirement for the histone H4 N-terminal tail. *J. Mol. Biol.* **327**, 85-96 (2003).
40. Allahverdi, A. *et al.* The effects of histone H4 tail acetylations on cation-induced chromatin folding and self-association. *Nucleic Acids Res.* **39**, 1680-1691 (2011).

41. Yang, D. & Arya, G. Structure and binding of the H4 histone tail and the effects of lysine 16 acetylation. *Phys. Chem. Chem. Phys.* **13**, 2911-2921 (2011).
42. Dorigo, B. *et al.* Nucleosome arrays reveal the two-start organization of the chromatin fiber. *Science* **306**, 1571-1573 (2004).
43. Kalashnikova, A. A. *et al.* The role of the nucleosome acidic patch in modulating higher order chromatin structure. *J. R. Soc. Interface* **10**, (2013).
44. Tremethick, D. J. Higher-order structures of chromatin: The elusive 30 nm fiber. *Cell* **128**, 651-654 (2007).
45. Robinson, P. J. *et al.* EM measurements define the dimensions of the '30-nm' chromatin fiber: Evidence for a compact, interdigitated structure. *Proc. Natl. Acad. Sci.* **103**, 6506-6511 (2006).
46. Routh, A., Sandin, S. & Rhodes, D. Nucleosome repeat length and linker histone stoichiometry determine chromatin fiber structure. *Proc. Natl. Acad. Sci.* **105**, 8872-8877 (2008).
47. Daban, J. R. & Bermúdez, A. Interdigitated solenoid model for compact chromatin fibers. *Biochemistry* **37**, 4299-4304 (1998).
48. Schalch, T. *et al.* X-ray structure of a tetranucleosome and its implications for the chromatin fibre. *Nature* **436**, 138-141 (2005).
49. Shogren-Knaak, M. *et al.* Histone H4-K16 acetylation controls chromatin structure and protein interactions. *Science* **311**, 844-847 (2006).
50. Chua, E. Y. D. *et al.* The mechanics behind DNA sequence-dependent properties of the nucleosome. *Nucleic Acids Res.* **40**, 6338-6352 (2012).
51. Drew, H. R. & Travers, A. A. DNA bending and its relation to nucleosome positioning. *J. Mol. Biol.* **186**, 773-790 (1985).
52. Travers, A. A. DNA bending and nucleosome positioning. *Trends Biochem. Sci.* **12**, 108-112 (1987).
53. Jiang, C. & Pugh, B. F. Nucleosome positioning and gene regulation: Advances through genomics. *Nat. Rev. Genet.* **10**, 161-172 (2009).
54. Zhang, Y. *et al.* Intrinsic histone-DNA interactions are not the major determinant of nucleosome positions in vivo. *Nat. Struct. Mol. Biol.* **16**, 847-852 (2009).
55. McGinty, R. K. & Tan, S. Nucleosome Structure and Function. *Chem. Rev.* **115**, 2255-2273 (2015).
56. McGinty, R. K. & Tan, S. Recognition of the nucleosome by chromatin factors and enzymes. *Curr. Opin. Struct. Biol.* **37**, 54-61 (2016).
57. Makde, R. D. *et al.* Structure of RCC1 chromatin factor bound to the nucleosome core particle. *Nature* **467**, 562-566 (2010).
58. Barbera, A. J. *et al.* The nucleosomal surface as a docking station for Kaposi's Sarcoma Herpesvirus LANA. *Science* **311**, 856-861 (2006).
59. Kato, H. *et al.* Architecture of the high mobility group nucleosomal protein 2-nucleosome complex as revealed by methyl-based NMR. *Proc. Natl. Acad. Sci.* **108**, 12283-12288 (2011).
60. Arnaudo, N. *et al.* The N-terminal acetylation of Sir3 stabilizes its binding to the nucleosome core particle. *Nat. Struct. Mol. Biol.* **20**, 1-8 (2013).
61. Yang, D. *et al.* N α -acetylated Sir3 stabilizes the conformation of a nucleosome-binding loop in the BAH domain. *Nat. Struct. Mol. Biol.* **20**, 1116-1118 (2013).
62. Mattioli, F. *et al.* The nucleosome acidic patch plays a critical role in RNF168-dependent ubiquitination of histone H2A. *Nat. Commun.* **5**, 3291 (2014).

63. Leung, J. W. *et al.* Nucleosome acidic patch promotes RNF168 and RING1B/BMI1-dependent H2AX and H2A ubiquitination and DNA damage signaling. *PLoS Genet.* **10**, (2014).
64. McGinty, R. K., Henrici, R. C. & Tan, S. Crystal structure of the PRC1 ubiquitylation module bound to the nucleosome. *Nature* **514**, 591-596 (2014).
65. Huang, R. C. & Bonner, J. Histone, a suppressor of chromosomal RNA synthesis. *Proc. Natl. Acad. Sci.* **48**, 1216-1222 (1962).
66. Allfrey, V. G., Littau, V. C. & Mirsky, A. E. On the role of histones in regulating ribonucleic acid synthesis in the cell nucleus. *Rockefeller Inst.* **49**, 414-421 (1962).
67. Allfrey, G., Faulkner, R. & Mirsky, A. E. Acetylation and methylation of histones and their possible role in the regulation of RNA synthesis. *Rockefeller Inst.* **315**, 786-794 (1964).
68. Valouev, A. *et al.* Determinants of nucleosome organization in primary human cells. *Nature* **474**, 516-522 (2011).
69. Stedman, E. & Stedman, E. The basic proteins of cell nuclei. *Philos. Trans. R. Soc. B: Biol. Sci.* **235**, 565-595 (1951).
70. Müller, M. M. & Muir, T. W. Histones: at the crossroads of peptide and protein chemistry. *Chem. Rev.* **115**, 2296-2349 (2015).
71. Speranzini, V. *et al.* Touch, act and go: landing and operating on nucleosomes. *EMBO J.* **35**, 376-388 (2016).
72. Kouzarides, T. Chromatin modifications and their function. *Cell* (2007).
73. Suganuma, T. & Workman, J. L. Crosstalk among histone modifications. *Cell* **135**, 604-607 (2008).
74. Allis, C. D. *et al.* New nomenclature for enzymes. *Cell* **131**, 633-636 (2007).
75. Kaimori, J.Y. *et al.* Histone H4 lysine 20 acetylation is associated with gene repression in human cells. *Sci. Rep.* **6**, 1-10 (2016).
76. Dhall, A. *et al.* Sumoylated human histone H4 prevents chromatin compaction by inhibiting long-range internucleosomal interactions. *J. Biol. Chem.* **289**, 33827-33837 (2014).
77. Brehove, M. *et al.* Histone core phosphorylation regulates DNA accessibility. *J. Biol. Chem.* **290**, 22612-22621 (2015).
78. Greer, E. L. & Shi, Y. Histone methylation: a dynamic mark in health, disease and inheritance. *Nat. Rev. Genet.* **13**, 343-357 (2014).
79. Krajewski, W. A., Li, J. & Dou, Y. Effects of histone H2B ubiquitylation on the nucleosome structure and dynamics. *Nucleic Acids Res.* **46**, 7631-7642 (2018).
80. Bannister, A. J. & Kouzarides, T. Regulation of chromatin by histone modifications. *Cell Research* **21**, 381-395 (2011).
81. Berger, S. L. The complex language of chromatin regulation during transcription. *Nature* **447**, 407-412 (2007).
82. Jenuwein, T. & Allis, C. D. Translating the histone code. *Science* **293**, 1074-1080 (2001).
83. Schwartz, Y. B. & Pirrotta, V. Polycomb silencing mechanisms and the management of genomic programmes. *Nat. Rev. Genet.* **8**, 9-22 (2007).
84. Vogelauer, M. *et al.* Global histone acetylation and deacetylation in yeast *Nature* **408**, 495-498 (2000).
85. Rundlett, S. E. *et al.* Transcriptional repression by UME6 involves deacetylation

- of lysine 5 of histone H4 by RPD3. *Nature* **392**, 831-835 (1998).
86. Walter, W. *et al.* Interaction with histone H3 involves a dual modification pattern of phosphoacetylation. *Mol. Cell. Biol.* **28**, 2840-2849 (2008).
 87. Daujat, S. *et al.* Crosstalk between CARM1 methylation and CBP acetylation on histone H3. *Curr. Biol.* **12**, 2090-2097 (2002).
 88. Lee, J. S. & Shilatifard, A. A site to remember: H3K36 methylation a mark for histone deacetylation. *Mutat. Res. Mol. Mech. Mutagen.* **618**, 130-134 (2007).
 89. Neumann, H. *et al.* A method for genetically installing site-specific acetylation in recombinant histones defines the effects of H3 K56 acetylation. *Mol. Cell* **36**, 153-163 (2009).
 90. Füllgrabe, J., Kavanagh, E. & Joseph, B. Histone onco-modifications. *Oncogene* **30**, 3391-3403 (2011).
 91. Fraga, M. F. *et al.* Loss of acetylation at Lys16 and trimethylation at Lys20 of histone H4 is a common hallmark of human cancer. *Nat. Genet.* **37**, 391-400 (2005).
 92. Elsheikh, S. E. *et al.* Global histone modifications in breast cancer correlate with tumor phenotypes, prognostic factors, and patient outcome. *Cancer Res.* **69**, 3802-3809 (2009).
 93. Broeck, A. *et al.* Loss of histone H4K20 trimethylation occurs in preneoplasia and influences prognosis of non-small cell lung cancer. *Clin. Cancer Res.* **14**, 7237-7245 (2008).
 94. Kallestad, L. *et al.* Transcriptional repression is epigenetically marked by H3K9 methylation during SV40 replication. *Clin. Epigenetics* **6**, 1-9 (2014).
 95. Bai, X. *et al.* Overexpression of myocyte enhancer factor 2 and histone hyperacetylation in hepatocellular carcinoma. *J. Cancer Res. Clin. Oncol.* **134**, 83-91 (2008).
 96. Nguyen, C. T. *et al.* Histone H3 Lysine 9 methylation is associated with aberrant gene silencing in cancer cells and is rapidly reversed by 5-Aza-2'-deoxycytidine. *Cancer Res.* **62**, 6456-6461 (2002).
 97. Francis, N. J. Chromatin compaction by a polycomb group protein complex. *Science* **306**, 1574-1577 (2004).
 98. Ringrose, L., Ehret, H. & Paro, R. Distinct contributions of histone H3 lysine 9 and 27 methylation to locus-specific stability of polycomb complexes. *Mol. Cell* **16**, 641-653 (2004).
 99. Yu, J. *et al.* A polycomb repression signature in metastatic prostate cancer predicts cancer outcome. *Cancer Res.* **67**, 10657-10663 (2007).
 100. Tzao, C. *et al.* Prognostic significance of global histone modifications in resected squamous cell carcinoma of the esophagus. *Mod. Pathol.* **22**, 252-260 (2009).
 101. He, L. R. *et al.* Prognostic impact of H3K27me3 expression on locoregional progression after chemoradiotherapy in esophageal squamous cell carcinoma. *BMC Cancer* **9**, 461 (2009).
 102. Wei, Y. *et al.* Loss of trimethylation at lysine 27 of Histone H3 is a predictor of poor outcome in breast, ovarian, and pancreatic cancers. *Mol. Carcinog.* **47**, 701-706 (2008).
 103. Martinez G. E. & Licht, J. D. Dereglulation of H3K27 methylation in cancer. *Nat. Genet.* **42**, 100-101 (2010).
 104. Vempati, R. K. *et al.* p300-mediated acetylation of Histone H3 lysine 56

- functions in DNA damage response in mammals. *J. Biol. Chem.* **285**, 28553-28564 (2010).
105. Xu, F., Zhang, K. & Grunstein, M. Acetylation in histone H3 globular domain regulates gene expression in yeast. *Cell* **121**, 375-385 (2005).
 106. Tjeertes, J. V., Miller, K. M. & Jackson, S. P. Screen for DNA-damage-responsive histone modifications identifies H3K9Ac and H3K56Ac in human cells. *EMBO J.* **28**, 1878-1889 (2009).
 107. Rufiange, A. *et al.* Genome-wide replication-independent histone H3 exchange occurs predominantly at promoters and implicates H3 K56 acetylation and Asf1. *Mol. Cell* **27**, 393-405 (2007).
 108. Masumoto, H. *et al.* A role for cell-cycle-regulated histone H3 lysine 56 acetylation in the DNA damage response. *Nature* **436**, 294-298 (2005).
 109. Seligson, D. B. *et al.* Global levels of histone modifications predict prognosis in different cancers. *Am. J. Pathol.* **174**, 1619-1628 (2009).
 110. Seligson, D. B. *et al.* Global histone modification patterns predict risk of prostate cancer recurrence. *Nature* **435**, 1262-1266 (2005).
 111. Ciechanover, A., Hod, Y. & Hershko, A. A heat-stable polypeptide component of an ATP-dependent proteolytic system from reticulocytes. *Biochem. Biophys. Res. Commun.* **81**, 1100-1105 (1978).
 112. Hershko, A. *et al.* Proposed role of ATP in protein breakdown: Conjugation of proteins with multiple chains of the polypeptide of ATP-dependent proteolysis. *Proc. Natl. Acad. Sci.* **77**, 1783-1786 (1980).
 113. Ciechanover, A. *et al.* "Covalent affinity" purification of Ubiquitin-activating enzyme. *J. Biol. Chem.* **257**, 2537-2542 (1982).
 114. Hershko, A. *et al.* Components of Ubiquitin-protein ligase system. *J. Biol. Chem.* **258**, 8206-8214 (1983).
 115. Hershko, A., Ciechanover, A. & Varshavsky, A. The ubiquitin system. *Nat. Med.* **6**, 1073-1081 (2000).
 116. Baumeister, W. *et al.* The proteasome: Paradigm of a self-compartmentalizing protease. *Cell* **92**, 367-380 (1998).
 117. Ciechanover, A., Finley, D. & Varshavsky, A. Ubiquitin dependence of selective protein degradation demonstrated in the mammalian cell cycle mutant ts85. *Cell* **37**, 57-66 (1984).
 118. Varshavsky, A. The early history of the ubiquitin field. *Protein Sci.* **15**, 647-654 (2006).
 119. Pickart, C. M. & Eddins, M. J. Ubiquitin: Structures, functions, mechanisms. *Biochim. Biophys. Acta. Mol. Cell Res.* **1695**, 55-72 (2004).
 120. Dey, A. *et al.* Loss of the tumor suppressor BAP1 causes myeloid transformation. *Science* **337**, 1541-1546 (2012).
 121. Bignell, G. R. *et al.* Identification of the familial cylindromatosis tumour-suppressor gene. *Nat. Genet.* **25**, 160-165 (2000).
 122. Hymowitz, S. G. & Wertz, I. E. A20: From ubiquitin editing to tumour suppression. *Nat. Rev. Cancer* **10**, 332-340 (2010).
 123. Yin, Y. & Yang, J. Signal and specificity of protein ubiquitination for proteasomal degradation. *bioRxiv* 038737 (2016).
 124. Lu, Y., Wang, W. & Kirschner, M. W. Specificity of the anaphase-promoting complex: A single-molecule study. *Science* **348**, 12487371-12487379 (2015).

125. Pierce, N. W. *et al.* Detection of sequential polyubiquitylation on a millisecond timescale. *Nature* **462**, 615-619 (2009).
126. Kleiger, G. *et al.* Rapid E2-E3 assembly and disassembly enable processive ubiquitylation of Cullin-RING ubiquitin ligase substrates. *Cell* **139**, 957-968 (2009).
127. Rape, M., Reddy, S. K. & Kirschner, M. W. The processivity of multiubiquitination by the APC determines the order of substrate degradation. *Cell* **124**, 89-103 (2006).
128. Swatek, K. N. & Komander, D. Ubiquitin modifications. *Cell Res.* **26**, 399-422 (2016).
129. Thrower, J. S. Recognition of the polyubiquitin proteolytic signal. *EMBO J.* **19**, 94-102 (2000).
130. Bremm, A. & Komander, D. Emerging roles for Lys11-linked polyubiquitin in cellular regulation. *Trends Biochem. Sci.* **36**, 355-363 (2011).
131. Wu-Baer, F. *et al.* The BRCA1/BARD1 heterodimer assembles polyubiquitin chains through an unconventional linkage involving lysine residue K6 of ubiquitin. *J. Biol. Chem.* **278**, 34743-34746 (2003).
132. Morris, J. R. & Solomon, E. 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 (2004).
133. Gatti, M. *et al.* RNF168 promotes noncanonical K27 ubiquitination to signal DNA damage. *Cell Rep.* **10**, 226-238 (2015).
134. Xu, P. *et al.* Quantitative proteomics reveals the function of unconventional ubiquitin chains in proteasomal degradation. *Cell* **137**, 133-145 (2010).
135. Komander, D. & Rape, M. The ubiquitin code. *Annu. Rev. Biochem.* **81**, 203-229 (2012).
136. Windheim, M., Pegg, M. & Cohen, P. Two different classes of E2 ubiquitin-conjugating enzymes are required for the mono-ubiquitination of proteins and elongation by polyubiquitin chains with a specific topology. *Biochem. J.* 723-729 (2008).
137. Rodrigo-Brenni, M. C. & Morgan, D. O. Sequential E2s drive polyubiquitin chain assembly on APC targets. *Cell* **130**, 127-139 (2007).
138. Petroski, M. D. *et al.* Substrate modification with lysine 63-linked ubiquitin chains through the UBC13-UEV1A ubiquitin-conjugating enzyme. *J. Biol. Chem.* **282**, 29936-29945 (2007).
139. Meas, R. & Mao, P. Histone ubiquitylation and its roles in transcription and DNA damage response. *DNA Repair* **36**, 36-42 (2015).
140. Komander, D. The emerging complexity of protein ubiquitination. *Biochem. Soc. Trans.* **37**, 937-953 (2009).
141. Pickart, C. M. Ubiquitin in chains. *Trends Biochem. Sci.* **25**, 544-548 (2000).
142. Varadan, R. *et al.* Structural properties of polyubiquitin chains in solution. *J. Mol. Biol.* **324**, 637-647 (2002).
143. Cook, W. J. *et al.* Structure of a diubiquitin conjugate and a model for interaction with ubiquitin conjugating enzyme (E2). *J. Biol. Chem.* **267**, 16467-16471 (1992).
144. Tenno, T. *et al.* Structural basis for distinct roles of lys63- and lys48-linked polyubiquitin chains. *Genes to Cells* **9**, 865-875 (2004).

145. Eddins, M. J. *et al.* Crystal structure and solution NMR studies of lys48-linked tetraubiquitin at neutral pH. *J. Mol. Biol.* **367**, 204-211 (2007).
146. Husnjak, K. & Dikic, I. Ubiquitin-binding proteins: decoders of ubiquitin-mediated cellular functions. *Annu. Rev. Biochem.* **81**, 291-322 (2012).
147. Qi, J. & Ronai, Z. A. Dysregulation of ubiquitin ligases in cancer. *Drug Resist Updat.* **23**, 1-11 (2015).
148. Davis, R. J., Welcker, M. & Clurman, B. E. Tumor suppression by the Fbw7 ubiquitin ligase: mechanisms and opportunities. *Cancer Cell* **26**, 455-464 (2014).
149. Shi, D. & Grossman, S. R. Ubiquitin becomes ubiquitous in cancer: Emerging roles of ubiquitin ligases and deubiquitinases in tumorigenesis and as therapeutic targets. *Cancer Biol. Ther.* **10**, 737-747 (2010).
150. Liu, L. *et al.* Functional significance and therapeutic implication of ring-type E3 ligases in colorectal cancer. *Oncogene* **37**, 148-159 (2018).
151. Katzav, S. & Schmitz, M. L. Mutations of c-Cbl in myeloid malignancies. *Oncotarget* **6**, 10689-10696 (2015).
152. Lipkowitz, S. & Weissman, A. M. RINGs of good and evil: RING finger ubiquitin ligases at the crossroads of tumour suppression and oncogenesis. *Nat. Rev. Cancer* **11**, 629-643 (2011).
153. Shi, Y. H. & Fang, W. G. Hypoxia-inducible factor-1 in tumour angiogenesis. *World J. Gastroenterol.* **10**, 1082 (2004).
154. Pelzer, C. *et al.* The protease activity of the paracaspase MALT1 is controlled by monoubiquitination. *Nat. Immunol.* **14**, 337-345 (2013).
155. Popovic, D., Vucic, D. & Dikic, I. Ubiquitination in disease pathogenesis and treatment. *Nat. Med.* **20**, 1242-1253 (2014).
156. Tofaris, G. K. *et al.* Ubiquitination of α -Synuclein in lewy bodies is a pathological event not associated with impairment of proteasome function. *J. Biol. Chem.* **278**, 44405-44411 (2003).
157. Schaeffer, V. & Goedert, M. Stimulation of autophagy is neuroprotective in a mouse model of human tauopathy. *Autophagy* **8**, 1686-1687 (2012).
158. Seibenhener, M. & Babu, J. Sequestosome 1/p62 is a polyubiquitin chain binding protein involved in ubiquitin proteasome degradation. *Mol. Cell. Biol.* **24**, 8055-8068 (2004).
159. Hashizume, R. *et al.* The RING heterodimer BRCA1-BARD1 is a ubiquitin ligase inactivated by a breast cancer-derived mutation. *J. Biol. Chem.* **276**, 14537-14540 (2001).
160. Venkitaraman, A. R. Functions of BRCA1 and BRCA2 in the biological response to DNA damage. *J. Cell Sci.* **114**, 3591-3598 (2001).
161. Goldstein, G. *et al.* Isolation of a polypeptide that has lymphocyte-differentiating properties and is probably represented universally in living cells. *Proc. Natl. Acad. Sci.* **72**, 11-5 (1975).
162. Vijay-Kumar, S., Bugg, C. E. & Cook, W. J. Structure of ubiquitin refined at 1.8 Å resolution. *J. Mol. Biol.* **194**, 531-544 (1987).
163. Dikic, I., Wakatsuki, S. & Walters, K. J. Ubiquitin-binding domains from structures to functions. *Nat. Rev. Mol. Cell Biol.* **10**, 659-671 (2009).
164. Sloper-mould, K. E. *et al.* Distinct functional surface regions on ubiquitin. *J. Biol. Chem.* **276**, 30483-30489 (2001).
165. Swaney, D. L., Rodriguez-Mias, R. A. & Villen, J. Phosphorylation of ubiquitin

- at Ser65 affects its polymerization, targets, and proteome-wide turnover. *EMBO Rep.* **16**, 1131-1144 (2015).
166. Ohtake, F. *et al.* Ubiquitin acetylation inhibits polyubiquitin chain elongation. *EMBO Rep.* **16**, 192-201 (2015).
 167. Stewart, M. D. *et al.* E2 enzymes: More than just middle men. *Cell Res.* **26**, 423-440 (2016).
 168. Siepmann, T. J. *et al.* Protein interactions within the N-end rule ubiquitin ligation pathway. *J. Biol. Chem.* **278**, 9448-9457 (2003).
 169. Kaiser, S. E. *et al.* Protein standard absolute quantification (PSAQ) method for the measurement of cellular ubiquitin pools. *Nat. Methods* **8**, 691-696 (2011).
 170. Brzovic, P. S. *et al.* A UbcH5c/Ubiquitin noncovalent complex is required for processive BRCA1-directed ubiquitination. *Mol. Cell* **21**, 873-880 (2006).
 171. Cohen, I. *et al.* Distinct activation of an E2 ubiquitin-conjugating enzyme by its cognate E3 ligases. *Proc. Natl. Acad. Sci.* (2015).
 172. Zheng, N. *et al.* Structure of a c-Cbl-UbcH7 complex: RING domain function in a ubiquitin-protein ligases. *Cell* **102**, 533-539 (2000).
 173. Huang, L. *et al.* Structure of an E6AP-UbcH7 Complex: Insights into ubiquitination by the E2-E3 enzyme cascade. **286**, 1321-1327 (1999).
 174. Brzovic, P. S. *et al.* Binding and recognition in the assembly of an active BRCA1/BARD1 ubiquitin-ligase complex. *Proc. Natl. Acad. Sci.* **100**, 5646-5651 (2003).
 175. Pruneda, J. N. *et al.* Ubiquitin in motion: Structural studies of the ubiquitin-conjugating enzyme-ubiquitin conjugate. *Biochemistry* **50**, 1624-1633 (2011).
 176. Buetow, L. *et al.* Activation of a primed RING E3-E2-Ubiquitin complex by non-covalent ubiquitin. *Mol. Cell* **58**, 297-310 (2015).
 177. David, Y. *et al.* E3 ligases determine ubiquitination site and conjugate type by enforcing specificity on E2 enzymes. *J. Biol. Chem.* **286**, 44104-44115 (2011).
 178. Harper, J. W. & Tan, M.K. Understanding Cullin-RING E3 biology through proteomics-based substrate identification. *Mol. Cell. Proteomics* **11**, 1541-1550 (2012).
 179. Persaud, A. *et al.* Comparison of substrate specificity of the ubiquitin ligases Nedd4 and Nedd4-2 using proteome arrays. *Mol. Syst. Biol.* **5**, 1-18 (2009).
 180. Bailly, V. *et al.* Yeast DNA repair proteins Rad6 and Rad18 form a heterodimer that has ubiquitin conjugating, DNA binding, and ATP hydrolytic activities. *J. Biol. Chem.* **272**, 23360-23365 (1997).
 181. Bordallo, J. *et al.* Der3p/Hrd1p is required for endoplasmic reticulum-associated degradation of misfolded luminal and integral membrane proteins. *Mol. Biol. Cell* **9**, 209-222 (1998).
 182. Li, W. *et al.* Genome-wide and functional annotation of human E3 ubiquitin ligases identifies MULAN, a mitochondrial E3 that regulates the organelle's dynamics and signalling. *PLoS One* **e1487**, 1-14 (2008).
 183. Deshaies, R. J. & Joazeiro, C. A. P. RING domain E3 ubiquitin ligases. *Annu. Rev. Biochem.* **78**, 399-434 (2009).
 184. Zachariae, W. *et al.* Mass spectrometric analysis of the anaphase-promoting complex from yeast: Identification of a subunit related to cullins. *Science* **279**, 1216-1219 (1998).
 185. Jia, L. & Sun, Y. SCF E3 ubiquitin ligases as anticancer targets. *Curr. Cancer*

- Drug Targets* **11**, 347-356 (2011).
186. Kim, H. C. & Huibregtse, J. M. Polyubiquitination by HECT E3s and the determinants of chain type specificity. *Mol. Cell. Biol.* **29**, 3307-3318 (2009).
 187. Wenzel, D. M. *et al.* UBCH7 reactivity profile reveals parkin and HHARI to be RING/HECT hybrids. *Nature* **474**, 105-108 (2011).
 188. Smit, J. J. *et al.* The E3 ligase HOIP specifies linear ubiquitin chain assembly through its RING-IBR-RING domain and the unique LDD extension. *EMBO J.* **31**, 3833-3844 (2012).
 189. Stieglitz, B. *et al.* LUBAC synthesizes linear ubiquitin chains via a thioester intermediate. *Nat. Publ. Gr.* **13**, 840-846 (2012).
 190. Lewandowski, K. T. & Piwnica-Worms, H. Phosphorylation of the E3 ubiquitin ligase RNF41 by the kinase Par-1b is required for epithelial cell polarity. *J. Cell Sci.* **127**, 315-327 (2014).
 191. De Bie, P. & Ciechanover, A. Ubiquitination of E3 ligases: Self-regulation of the ubiquitin system via proteolytic and non-proteolytic mechanisms. *Cell Death Differ.* **18**, 1393-1402 (2011).
 192. DaRosa, P. A. *et al.* Allosteric activation of the RNF146 ubiquitin ligase by a Poly(ADP-ribosyl)ation signal. *Nature* **517**, 223-226 (2015).
 193. Dou, H. *et al.* Structural basis for autoinhibition and phosphorylation-dependent activation of c-Cbl. *Nat. Struct. Mol. Biol.* **19**, 1-22 (2014).
 194. Khosravi, R. *et al.* Rapid ATM-dependent phosphorylation of MDM2 precedes p53 accumulation in response to DNA damage. *Proc. Natl. Acad. Sci.* **96**, 14973-14977 (1999).
 195. Debonneville, C. *et al.* Phosphorylation of Nedd4-2 by Sgk1 regulates epithelial Na⁺ channel cell surface expression. *EMBO J.* **20**, 7052-7059 (2001).
 196. Zheng, N. *et al.* Ubiquitin ligases : structure, function, and regulation. *Annu. Rev. Biochem.* **86**, 1-29 (2017).
 197. Ben-saadon, R. *et al.* The polycomb protein RING1B generates self atypical mixed ubiquitin chains required for its In vitro histone H2A ligase activity. *Mol. Cell* **24**, 701-711 (2006).
 198. Ranaweera, R. S. & Yang, X. Auto-ubiquitination of Mdm2 enhances its substrate ubiquitin ligase activity. *J. Biol. Chem.* **288**, 18939-18946 (2013).
 199. Torre, L. A. *et al.* *Global cancer in women: Burden and trends.* *Cancer Epidemiology Biomarkers and Prevention* **26**, (2017).
 200. Miki, Y. *et al.* Candidate ovarian and breast susceptibility gene for the cancer. *Science* **266**, 5-11 (1994).
 201. Hall, J. M. *et al.* Linkage of early-onset familial breast cancer to chromosome 17q21. *Science* **250**, 1684-1689 (1990).
 202. Foulkes, W. D. Germline BRCA1 mutations and a basal epithelial phenotype in breast cancer. *J. Natl. Cancer Inst.* **95**, 1482-1485 (2003).
 203. Miki, Y. *et al.* A strong candidate for the breast and ovarian cancer susceptibility gene BRCA1. *Science* **266**, 66-71 (1994).
 204. Yu, X. & Baer, R. Nuclear localization and cell cycle-specific expression of CtIP, a protein that associates with the BRCA1 tumor suppressor. *J. Biol. Chem.* **275**, 18541-18549 (2000).
 205. Wu *et al.* Identification of a RING protein that can interact in vitro with the BRCA1 gene product. *Nat Genet.* **14**, 430-440 (1996).

206. Ramus, S. J. *et al.* Germline mutations in the BRIP1, BARD1, PALB2, and NBN genes in women with ovarian cancer. *J. Natl. Cancer Inst.* **107**, 1-8 (2015).
207. Ghimenti, C. *et al.* Germline mutations of the BRCA1-associated ring domain (BARD1) gene in breast and breast/ovarian families negative for BRCA1 and BRCA2 alterations. *Genes Chromosom. Cancer* **33**, 235-242 (2002).
208. McCarthy, E. E. *et al.* Loss of Bard1, the heterodimeric partner of the BRCA1 tumor suppressor, results in early embryonic lethality and chromosomal instability. *Mol. Cell. Biol.* **23**, 5056-5063 (2003).
209. Westermarck, U. K. *et al.* BARD1 participates with BRCA1 in homology-directed repair of chromosome breaks. *Mol. Cell. Biol.* **23**, 7926-936 (2003).
210. Ye, Q. *et al.* BRCA1-induced large-scale chromatin unfolding and allele-specific effects of cancer-predisposing mutations. *J. Cell Biol.* **155**, 911-921 (2001).
211. Bork, P. *et al.* A superfamily of conserved domains in DNA damage-responsive cell cycle checkpoint proteins. *FASEB J.* **11**, 68-76 (1997).
212. Meza, J. E. *et al.* Mapping the functional domains of BRCA1. Interaction of the ring finger domains of BRCA1 and BARD1. *J. Biol. Chem.* **274**, 5659-5665 (1999).
213. Ruffner, H. *et al.* Cancer-predisposing mutations within the RING domain of BRCA1: Loss of ubiquitin protein ligase activity and protection from radiation hypersensitivity. *Proc. Natl. Acad. Sci.* **98**, 5134-5139 (2001).
214. Brzovic, P. S. *et al.* BRCA1 RING domain cancer-predisposing mutations. Structural consequences and effects on protein-protein interactions. *J. Biol. Chem.* **276**, 41399-41406 (2001).
215. Linger, R. J. & Kruk, P. A. BRCA1 16 years later: Risk-associated BRCA1 mutations and their functional implications. *FEBS J.* **277**, 3086-3096 (2010).
216. Saurin, A. J. *et al.* Does this have a familiar RING? *Trends Biochem. Sci.* **21**, 208-214 (1996).
217. Lorick, K. L. *et al.* RING fingers mediate ubiquitin-conjugating enzyme (E2)-dependent ubiquitination. *Proc. Natl. Acad. Sci.* **96**, 11364-11369 (1999).
218. Christensen, D. E., Brzovic, P. S. & Klevit, R. E. E2-BRCA1 RING interactions dictate synthesis of mono- or specific polyubiquitin chain linkages. *Nat. Struct. Mol. Biol.* **14**, 941-948 (2007).
219. Nishikawa, H. *et al.* Mass spectrometric and mutational analyses reveal lys-6-linked polyubiquitin chains catalyzed by BRCA1-BARD1 ubiquitin ligase. *J. Biol. Chem.* **279**, 3916-3924 (2004).
220. Zhu, Q. *et al.* BRCA1 tumour suppression occurs via heterochromatin-mediated silencing. *Nature* **477**, 179-184 (2011).
221. Yu, X. *et al.* BRCA1 ubiquitinates its phosphorylation-dependent binding partner CtIP. *Genes Dev.* **20**, 1721-1726 (2006).
222. Kleiman, F. E. *et al.* BRCA1/BARD1 inhibition of mRNA 3' processing involves targeted degradation of RNA polymerase II. *Genes Dev.* **19**, 1227-1237 (2005).
223. Starita, L. M. *et al.* BRCA1/BARD1 ubiquitinate phosphorylated RNA polymerase II. *J. Biol. Chem.* **280**, 24498-24505 (2005).
224. Eakin, C. M. *et al.* Estrogen receptor is a putative substrate for the BRCA1 ubiquitin ligase. *Proc. Natl. Acad. Sci.* **104**, 1-6 (2007).
225. Wu, W. *et al.* The ubiquitin E3 ligase activity of BRCA1 and its biological functions. *Cell Div.* **3**, 1-10 (2008).

226. Tubbs, A. & Nussenzweig, A. Endogenous DNA damage as a source of genomic instability in cancer. *Cell* **168**, 644-656 (2017).
227. Jeggo, P. A., Pearl, L. H. & Carr, A. M. DNA repair, genome stability and cancer: A historical perspective. *Nat. Rev. Cancer* **16**, 35-42 (2016).
228. Beamish, H. *et al.* Defect in multiple cell cycle checkpoints in Ataxia-Telangiectasia post-irradiation. *J. Biol. Chem.* **271**, 20486-20493 (1996).
229. Collins, K., Jacks, T. & Pavletich, N. P. The cell cycle and cancer. *Proc. Natl. Acad. Sci.* **94**, 2776-2778 (1997).
230. Marteijn, J. A. *et al.* Understanding nucleotide excision repair and its roles in cancer and ageing. *Nat. Rev. Mol. Cell Biol.* **15**, 465-481 (2014).
231. De Boer, J. & Hoeijmakers, J. H. J. Nucleotide excision repair and human syndromes. *Carcinogenesis* **21**, 453-460 (2000).
232. Carter, R. J. & Parsons, J. L. Base excision repair, a pathway regulated by posttranslational modifications. *Mol. Cell. Biol.* **36**, 1426-1437 (2016).
233. Li, G. M. Mechanisms and functions of DNA mismatch repair. *Cell Res.* **18**, 85-98 (2008).
234. Lieber, M. R. The mechanism of human nonhomologous DNA end joining. *J. Biol. Chem.* **283**, 1-5 (2008).
235. Wright, W. D., Shah, S. S. & Heyer, W. D. Homologous recombination and the repair of DNA double-strand breaks. *J. Biol. Chem.* **293**, 10524-10535 (2018).
236. Scully, R. *et al.* Genetic analysis of BRCA1 function in a defined tumor cell line. *Mol. Cell* **4**, 1093-1099 (1999).
237. Li, M. L. & Greenberg, R. A. Links between genome integrity and BRCA1 tumor suppression. *Trends Biochem. Sci.* **37**, 418-424 (2012).
238. Chen, C.C. *et al.* Homology-directed repair and the role of BRCA1, BRCA2, and related proteins in genome integrity and cancer. *Annu. Rev. Cancer Biol.* **2**, 313-336 (2018).
239. Wu, J., Lu, L. Y. & Yu, X. The role of BRCA1 in DNA damage response. *Protein Cell* **1**, 117-123 (2010).
240. Mullan, P. B., Quinn, J. E. & Harkin, D. P. The role of BRCA1 in transcriptional regulation and cell cycle control. *Oncogene* **25**, 5854-5863 (2006).
241. Di, L. J. *et al.* Transcriptional regulation of BRCA1 expression by a metabolic switch. *Nat. Struct. Mol. Biol.* **17**, 1406-1413 (2010).
242. Mallery, D. L., Vandenberg, C. J. & Hiom, K. Activation of the E3 ligase function of the BRCA1/BARD1 complex by polyubiquitin chains. *EMBO J.* **21**, 6755-6762 (2002).
243. Thangaraju, M., Kaufmann, S. H. & Couch, F. J. BRCA1 facilitates stress-induced apoptosis in breast and ovarian cancer cell lines. *J. Biol. Chem.* **275**, 33487-33496 (2000).
244. Scully, R. *et al.* Association of BRCA1 with Rad51 in mitotic and meiotic cells. *Cell* **88**, 265-275 (1997).
245. MacLachlan, T. K. *et al.* BRCA1 effects on the cell cycle and the DNA damage response are linked to altered gene expression. *J. Biol. Chem.* **275**, 2777-2785 (2000).
246. Ruffner, H. & Verma, I. M. BRCA1 is a cell cycle-regulated nuclear phosphoprotein. *Proc. Natl. Acad. Sci.* **94**, 7138-7143 (1997).
247. Fierz, B. *et al.* Histone H2B ubiquitylation disrupts local and higher-order

- chromatin compaction. *Nat. Chem. Biol.* **7**, 113-119 (2011).
248. Cao, R., Tsukada, Y.I. & Zhang, Y. Role of BMI1 and RING1A in H2A ubiquitylation and Hox gene silencing. *Mol. Cell* **20**, 845-854 (2005).
 249. Nakagawa, T. *et al.* Deubiquitylation of histone H2A activates transcriptional initiation via trans-histone cross-talk with H3K4 di- and trimethylation. *Genes Dev.* **22**, 37-49 (2008).
 250. Hengbin, W. *et al.* Role of histone H2A ubiquitination in polycomb silencing. *Nature* **431**, 873-878 (2004).
 251. Gearhart, M. D. *et al.* Polycomb group and SCF ubiquitin ligases are found in a novel BCOR complex that is recruited to BCL6 targets. *Mol. Cell. Biol.* **26**, 6880-6889 (2006).
 252. Bergink, S. *et al.* DNA damage triggers nucleotide excision repair-dependent monoubiquitylation of histone H2A. *Genes Dev.* **20**, 1343-1352 (2006).
 253. Hsiao, K. Exonuclease III induced ligase-free directional subcloning of PCR products. **21**, 5528-5529 (1993).
 254. Tan, S., Kern, R. C. & Selleck, W. The pST44 polycistronic expression system for producing protein complexes in Escherichia coli. *Protein Expr. Purif.* **40**, 385-395 (2005).
 255. Luger, K., Rechsteiner, T. J. & Richmond, T. J. Expression and purification of recombinant histones and nucleosome reconstitution. *Chromatin Protoc.* **119**, 1-16
 256. Birnboim, H. C. & Doly, J. A rapid alkaline extraction procedure for screening recombinant plasmid DNA. **7**, 1513-1523 (1979).
 257. Luger, K. *et al.* Preparation of nucleosome core particle from recombinant histones. *Methods Enzymol.* **304**, (1999).
 258. Hellman, L. M. & Fried, M. G. Electrophoretic mobility shift assay (EMSA) for detecting protein–nucleic acid interactions. *Nat. Protoc.* **2**, 1849-1861 (2007).
 259. Berger, I., Fitzgerald, D. J. & Richmond, T. J. Baculovirus expression system for heterologous multiprotein complexes. *Nat. Biotechnol.* **22**, 1583-1587 (2004).
 260. Fitzgerald, D. J. *et al.* Protein complex expression by using multigene baculoviral vectors. *Nat. Methods* **3**, 1021-1032 (2006).
 261. Scheres, S. H. W. RELION: Implementation of a Bayesian approach to cryo-EM structure determination. *J. Struct. Biol.* **180**, 519-530 (2012).
 262. Mindell, J. A. & Grigorieff, N. Accurate determination of local defocus and specimen tilt in electron microscopy. *J. Struct. Biol.* **142**, 334-347 (2003).
 263. Rohou, A. & Grigorieff, N. CTFFIND4: Fast and accurate defocus estimation from electron micrographs. *J. Struct. Biol.* **192**, 216-221 (2015).
 264. Tang, G. *et al.* EMAN2: An extensible image processing suite for electron microscopy. *J. Struct. Biol.* **157**, 38-46 (2007).
 265. Ludtke, S. J., Baldwin, P. R. & Chiu, W. EMAN: Semiautomated software for high-resolution single-particle reconstructions. *J. Struct. Biol.* **128**, 82-97 (1999).
 266. Lee, J. S., Smith, E. & Shilatifard, A. The language of histone crosstalk. *Cell* **142**, 682-685 (2010).
 267. Xiao, T. *et al.* Histone H2B ubiquitylation is associated with elongating RNA polymerase II. *Mol. Cell. Biol.* **25**, 637-651 (2005).
 268. Minsky, N. *et al.* Monoubiquitinated H2B is associated with the transcribed region of highly expressed genes in human cells. *Nat. Cell Biol.* **10**, 483-488

- (2008).
269. Pickart, C. M. Mechanisms underlying ubiquitination. *Annu. Rev. Biochem.* **70**, 503-533 (2001).
 270. Weissman, A. M. Themes and variations on ubiquitylation. *Nat. Rev. Mol. Cell Biol.* **2**, (2001).
 271. Koonin, E. V., Altschul, S. F. & Bork, P. BRCA1 protein products: functional motifs. *Nat. Genet.* **13**, 266-267 (1996).
 272. Harrison, S. C. A structural taxonomy of DNA-binding domains. *Nature* **353**, 715-719 (1991).
 273. Freemont, P. S., Lane, A. N. & Sanderson, M. R. Structural aspects of protein-DNA recognition. *Biochem. J* **278**, 1-23 (1991).
 274. Klug, A. & Rhodes, D. Zinc Fingers: A novel protein fold for nucleic acid recognition. *Cold Spring Harb. Symp. Quant. Biol.* **52**, 473-482 (1987).
 275. Berg, J. M. Zinc fingers and other metal-binding domains. *J. Biol. Chem.* **265**, 6513-6516 (1990).
 276. Tagawa, M. *et al.* Expression of novel DNA-binding protein with zinc finger structure in various tumor cells. *J. Biol. Chem.* **265**, 20021-20026 (1990).
 277. Freemont, P. S. The RING finger: A novel protein sequence motif related to the zinc finger. *Ann. NY Acad. Sci* **684**, 174-192 (1993).
 278. Everett, R. D. *et al.* A novel arrangement of Zinc-binding residues and secondary structure in the C3HC4 motif of an alpha herpes virus protein family. *J. Mol. Biol.* **234**, 1038-1047 (1993).
 279. Lovering, R. *et al.* Identification and preliminary characterization of a protein motif related to the zinc finger. *Proc. Natl. Acad. Sci.* **90**, 2112-2116 (1993).
 280. Barlow, P. N. *et al.* Structure of the C3HC4 Domain by 1H-nuclear magnetic resonance spectroscopy. *J. Mol. Biol* **237**, 201-211 (1994).
 281. Borden, K. L. *et al.* The solution structure of the RING finger domain from the acute promyelocytic leukaemia proto-oncoprotein PML. *EMBO J.* **14**, 1532-1541 (1995).
 282. Joazeiro, C. A. P. & Weissman, A. M. RING Finger Proteins : Mediators of Ubiquitin Ligase Activity. **102**, 549-552 (2000).
 283. Jackson, P. K. *et al.* The lore of the RINGS: substrate recognition and catalysis by ubiquitin ligases. *Trends Cell Biol.* **10**, (2000).
 284. Eickbush, T. H. & Moudrianakis, E. N. The histone core complex: an octamer assembled by two sets of protein-protein interactions. *Biochemistry* **17**, 4955-4964 (1978).
 285. Brzovic, P. S. *et al.* The cancer-predisposing mutation C61G disrupts homodimer formation in the NH2-terminal BRCA1 RING finger domain. *J. Biol. Chem.* **273**, 7795-7799 (1998).
 286. Mark, W. Y. *et al.* Characterization of segments from the central region of BRCA1: An intrinsically disordered scaffold for multiple protein-protein and protein-DNA interactions *J. Mol. Biol.* **345**, 275-287 (2005).
 287. Sturdy, A., Naseem, R. & Webb, M. Purification and characterisation of a soluble N-terminal fragment of the breast cancer susceptibility protein BRCA1. *J. Mol. Biol.* **340**, 469-475 (2004).
 288. Yadav, L. R. *et al.* Functional assessment of intrinsic disorder central domains of BRCA1. *J. Biomol. Struct. Dyn.* **33**, 2469-2478 (2015).

289. Attali, I. *et al.* Ubiquitylation-dependent oligomerization regulates activity of Nedd4 ligases. *EMBO J.* **36**, 425-440 (2017).
290. Iconomou, M. & Saunders, D. N. Systematic approaches to identify E3 ligase substrates. *Biochem. J.* **473**, 4083-4101 (2016).
291. Chang, L. *et al.* Atomic structure of the APC/C and its mechanism of protein ubiquitination. *Nature* **522**, 450-454 (2015).
292. Quiocho, F. A. & Richards, F. M. Intermolecular cross linking of a protein in the crystalline state: carboxypeptidase-A. *Proc. Natl. Acad. Sci.* **52**, 833-839 (1964).
293. Reeke, G. & Hartsuck, J. The structure of carboxypeptidase A, some results at 2.0 Å resolution, and the complex with glycyl-tyrosine at 2.8 Å resolution. *Proc. Natl. Acad. Sci.* **58**, 2220-2226 (1967).
294. Richards, F. M. & Knowles, J. R. Glutaraldehyde as a protein cross-linking reagent. *J. Mol. Biol.* **37**, 231-233 (1968).
295. Machida, S. *et al.* Structural basis of heterochromatin formation by human HP1. *Mol. Cell* **69**, 385-397 (2018).
296. Hu, Q. *et al.* Mechanisms of ubiquitin-nucleosome recognition and regulation of 53BP1 chromatin recruitment by RNF168/169 and RAD18. *Mol Cell* **66**, 473-487 (2017).
297. Doerr, A. Single-particle cryo-electron microscopy. *Nat. Methods* **13**, 23-23 (2016).
298. Taylor, K. A. & Glaeser, R. M. Electron diffraction of frozen, hydrated protein crystals. *Science* **186**, 1036-1037 (1974).
299. Unwin, P. N. T. & Henderson, R. Molecular structure determination by electron microscopy of unstained crystalline specimens. *J. Mol. Biol.* **94**, 425-440 (1975).
300. Dubochet, J. *et al.* Electron microscopy of frozen water and aqueous solutions. *J. Microsc.* **128**, 219-237 (1982).
301. Frank, J. *et al.* Reconstruction of glutamine synthetase using computer averaging. *Ultramicroscopy* **3**, 283-290 (1978).
302. Faruqi, A. R. *et al.* Direct single electron detection with a CMOS detector for electron microscopy. *Nucl. Instruments Methods Phys. Res.* **546**, 170-175 (2005).
303. Milazzo, A. C. *et al.* Active pixel sensor array as a detector for electron microscopy. *Ultramicroscopy* **104**, 152-159 (2005).
304. Li, X. *et al.* Electron counting and beam-induced motion correction enable near-atomic-resolution single-particle cryo-EM. *Nat. Methods* **10**, 584-590 (2013).
305. Scheres, S. H. W. Beam-induced motion correction for sub-megadalton cryo-EM particles. *Elife* **3**, e03665 (2014).
306. Danev, R. & Nagayama, K. Transmission electron microscopy with Zernike phase plate. *Ultramicroscopy* **88**, 243-252 (2001).
307. Nagayama, K. Another 60 years in electron microscopy: Development of phase-plate electron microscopy and biological applications. *J. Electron Microsc.* (Tokyo). **60**, 43-62 (2011).
308. Danev, R. *et al.* Volta potential phase plate for in-focus phase contrast transmission electron microscopy. *Proc. Natl. Acad. Sci.* **111**, 15635-15640 (2014).
309. Chua, E. Y. D. *et al.* 3.9 Å structure of the nucleosome core particle determined by phase-plate cryo-EM. *Nucleic Acids Res.* **44**, 8013-8019 (2016).
310. Noble, A. J. *et al.* Routine single particle CryoEM sample and grid

- characterization by tomography. *Elife* **7**, 1-42 (2018).
311. D'Imprima, E. *et al.* The deadly touch: protein denaturation at the water-air interface and how to prevent it. *bioRxiv* 400432 (2018).
 312. Northam, M. R. & Trujillo, K. M. Histone H2B mono-ubiquitylation maintains genomic integrity at stalled replication forks. *Nucleic Acids Res.* **44**, 9245-9255 (2016).
 313. Mailand, N. *et al.* RNF8 ubiquitylates histones at DNA double-strand breaks and promotes assembly of repair proteins. *Cell* **131**, 887-900 (2007).
 314. Metzger, M. B. *et al.* RING-type E3 ligases: Master manipulators of E2 ubiquitin-conjugating enzymes and ubiquitination. *Biochimica et Biophysica Acta* **1843**, 47-60 (2014).
 315. Fabbro, M. *et al.* BRCA1-BARD1 complexes are required for p53 Ser-15 phosphorylation and a G 1/S arrest following ionizing radiation-induced DNA damage. *J. Biol. Chem.* **279**, 31251-31258 (2004).
 316. Fabbro, M. *et al.* BARD1 regulates BRCA1 apoptotic function by a mechanism involving nuclear retention. *Exp. Cell Res.* **298**, 661-673 (2004).
 317. Densham, R. M. *et al.* Human BRCA1-BARD1 ubiquitin ligase activity counteracts chromatin barriers to DNA resection. *Nat. Struct. Mol. Biol.* **23**, 647-655 (2016).
 318. Zentner, G. E. & Henikoff, S. Regulation of nucleosome dynamics by histone modifications. *Nat. Struct. Mol. Biol.* **20**, 259-266 (2013).
 319. Campbell, M. G. *et al.* Movies of ice-embedded particles enhance resolution in electron cryo-microscopy. *Structure* **20**, 1823-1828 (2012).
 320. Bai, X. C. *et al.* Ribosome structures to near-atomic resolution from thirty thousand cryo-EM particles. *Elife* **2013**, 2-13 (2013).
 321. Lewis, E. B. A gene complex controlling segmentation in drosophila. *Nature* **276**, 565-570 (1978).
 322. Buchwald, G. *et al.* Structure and E3-ligase activity of the RING-RING complex of polycomb proteins BMI1 and RING1B. *EMBO J.* **25**, 2465-2474 (2006).
 323. Bentley, M. L. *et al.* Recognition of UbcH5c and the nucleosome by the BMI1/RING1B ubiquitin ligase complex. *EMBO J.* **30**, 3285-3297 (2011).
 324. Nemergut, M. E. *et al.* Chromatin docking and exchange activity enhancement of RCC1 by histones H2A and H2B. *Science* **292**, 1540-1543 (2001).
 325. Rusche, L. N., Kirchmaier, A. L. & Rine, J. The establishment, inheritance, and function of silenced chromatin in *Saccharomyces cerevisiae*. *Annu. Rev. Biochem.* **72**, 481-516 (2003).
 326. Connelly, J. J. *et al.* Structure and function of the *Saccharomyces cerevisiae* Sir3 BAH domain. *Society* **26**, 3256-3265 (2006).
 327. Singer, M. S. *et al.* Identification of high-copy disruptors of telomeric silencing in *Saccharomyces cerevisiae*. *Genetics* **150**, 613-632 (1998).
 328. Ng, H. H. *et al.* Lysine-79 of histone H3 is hypomethylated at silenced loci in yeast and mammalian cells: A potential mechanism for position-effect variegation. *Proc. Natl. Acad. Sci.* **100**, 1820-1825 (2003).
 329. Van Leeuwen, F., Gafken, P. R. & Gottschling, D. E. Dot1p modulates silencing in yeast by methylation of the nucleosome core. *Cell* **109**, 745-756 (2002).
 330. Kato, H. *et al.* A conserve mechanism for centromere nucleosome recognition by centromeric protein CENP-C. *Science* **725**, 1110-1114 (2013).

331. Lim, J. H. *et al.* Chromosomal protein HMGN1 modulates histone H3 phosphorylation. *Mol. Cell* **15**, 573-584 (2004).
332. Rattner, B. P., Yusufzai, T. & Kadonaga, J. T. HMGN proteins act in opposition to ATP-dependent chromatin remodeling factors to restrict nucleosome mobility. *Mol. Cell* **34**, 620-626 (2009).
333. Lim, J. H. *et al.* Chromosomal protein HMGN1 enhances the acetylation of lysine 14 in histone H3. *EMBO J.* **24**, 3038-3048 (2005).
334. Postnikov, Y. & Bustin, M. Regulation of chromatin structure and function by HMGN proteins. *Biochim. Biophys. Acta* **1799**, 62-68 (2010).
335. Brink, J. *et al.* Evaluation of charging on biological specimens in an electron cryomicroscope. *Ultramicroscopy* **72**, 41-52 (1998).
336. Larson, D. M., Downing, K. H. & Glaeser, R. M. The surface of evaporated carbon films is an insulating, high-bandgap material. *J. Struct. Biol.* **174**, 420-423 (2011).
337. Passmore, L. A. & Russo, C. J. Specimen preparation for high-resolution cryo-EM. *Methods Enzymol.* **579**, 51-86 (2016).
338. Geim, A. K. & Novoselov, K. S. The rise of graphene. *Nat. Mater.* **6**, 183-191 (2007).
339. Sader, K. *et al.* Cryomicroscopy of radiation sensitive specimens on unmodified graphene sheets: Reduction of electron-optical effects of charging. *J. Struct. Biol.* **183**, 531-536 (2013).
340. Pantelic, R. S. *et al.* The application of graphene as a sample support in transmission electron microscopy. *Solid State Commun.* **152**, 1375-1382 (2012).
341. Naydenova, K. & Russo, C. J. Measuring the effects of particle orientation to improve the efficiency of electron cryomicroscopy. *Nat. Commun.* **8**, 629 (2017).
342. Ramsden, J. J. Experimental methods for investigating protein adsorption kinetics at surfaces. *Q. Rev. Biophys.* **27**, 41 (1994).
343. Trurnit, H. A theory and method for the spreading of protein monolayers. *J. Colloid Sci.* **15**, 1-13 (1960).
344. Pantelic, R. S. *et al.* Rendering graphene supports hydrophilic with non-covalent aromatic functionalization for transmission electron microscopy. *Appl. Phys. Lett.* **104**, 1341031-1341035 (2014).
345. Lu, P. *et al.* Three-dimensional structure of human γ -secretase. *Nature* **512**, 166-170 (2014).
346. Bai, X. *et al.* An atomic structure of human γ -secretase. *Nature* **525**, 212-217 (2015).

Supplementary Data

Extended Data Figure 1: H2A ubiquitylation with wt-Ub concentration of 20 μ M.

A) Denaturing gel showing BRBA-S mediated poly-ubiquitylation of histone H2A, mono-, di-, tri-, tetra-, and penta-ubiquitylated H2A are labeled correspondingly. B) Denaturing gel showing BRBA-L mediated poly-ubiquitylation of histone H2A, mono-, di-, tri-, tetra-, and penta-ubiquitylated H2A are labeled correspondingly. The ubiquitylation assay was performed at the specified time points at 30°C in a buffer containing 3 mM ATP (50 mM Hepes pH 7.5, 75 mM NaCl, 1 mM DTT, 10 μ M ZnSO₄, 2.5 mM MgCl₂). The protein concentrations were 0.03 μ M UBE1, 1 μ M UbcH5c, 1 μ M BRBA-S, 1 μ M BRBA-L, 20 μ M wt-Ub, and 2 μ M NCP). Subsequently, samples were denatured by heating at 95°C, followed by analysis on a denaturing gel.

Extended Data Figure 2: H2A ubiquitylation with K0-Ub concentration of 20 μ M.

A) Denaturing gel showing multiple-mono-ubiquitylation of histone H2A, mono-, di-, and tri-ubiquitylated H2A are labeled as Ub₁, Ub₂, and Ub₃, respectively. The ubiquitylation assay was performed at the specified time points at 30°C in a buffer containing 3 mM ATP (50 mM Hepes pH 7.5, 75 mM NaCl, 1 mM DTT, 10 μ M ZnSO₄, and 2.5 mM MgCl₂). The protein concentrations were 0.03 μ M UBE1, 1 μ M UbcH5c, 1 μ M BRBA-S, 1 μ M BRBA-L, 20 μ M K0-Ub, and 2 μ M NCP.