Recognition of nucleosomes by protein factors that regulate chromatin structure and dynamics

Lee, Phoi Leng

2019

Lee, R. P. L. (2019). Recognition of nucleosomes by protein factors that regulate chromatin structure and dynamics. Doctoral thesis, Nanyang Technological University, Singapore.

<https://hdl.handle.net/10356/142395>

<https://doi.org/10.32657/10356/142395>

This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License (CC BY-NC 4.0).

Downloaded on 26 Apr 2023 14:02:26 SGT

**NANYANG
TECHNOLOGICAL
UNIVERSITY**

SINGAPORE

**Recognition of Nucleosomes by Protein Factors that Regulate
Chromatin Structure and Dynamics**

Lee Phoi Leng, Rachael

SCHOOL OF BIOLOGICAL SCIENCES

2019

**Recognition of Nucleosomes by Protein Factors that Regulate Chromatin
Structure and Dynamics**

Lee Phoi Leng, Rachael

SCHOOL OF BIOLOGICAL SCIENCES

A thesis submitted to the Nanyang Technological University
in partial fulfilment 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.

25th August 2019

.....
Date

.....
Lee Phoi Leng

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 and 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.

15th August 2019

.....
Date

.....
Associate Professor Curtis Alexander Davey

Authorship Attribution Statement

(A) This thesis does not contain any materials from papers published in peer-reviewed journals or from papers accepted at conferences in which I am listed as an author (referring to the time when I submitted this thesis).

25th August 2019

A handwritten signature in black ink, appearing to read 'Rachael', with a stylized flourish at the end.

.....
Date

.....
Lee Phoi Leng

Acknowledgements

I am deeply grateful to my supervisor, Prof. Curtis Alexander Davey, for giving me the opportunity to learn and flourish under his wings. He has provided invaluable advice and encouragements all along, which motivated me to stay passionate in science and research. His support and care empowers me to grow stronger as a fledging researcher.

I have crossed paths with many people during this PhD journey, where each of them adds values to make it a meaningful path. I want to express my heartfelt gratitude to all the present and past lab members for being like family members, whom I will often reminisce, - in no particular order- Dr. Zenita Adhireksan, Dr. Deepti Sharma, Dr. Gabriela Elizabeth Davey, Wayne Shum, Louis de Falco Jr, Dr. Sivaraman Padavattan, Dr. Sailatha Ravi, Dr. Eugene Tan, and Dr. Zhu Jun Ma. Those enriching discussions with them have helped me immensely in acquiring the scientific thinking required to carry out this endeavour. Not forgetting those friends that I made out of the lab, Dr. Wei Sheng Chia, Dr. Janet To, Dr. Ethan Lai, York Wiewo Chuah, Justin Ng, Tra Nguyen, and those who crossed my path, thanks for being the great companions in this journey.

Most importantly, I would like to express appreciation to my beloved parents, Lee Bon Huat and Koo Ai Ai, as well as my husband aka soulmate, Edmund Lai, for the unconditional love and support that they have provided me with to strive through all the difficulties and hard times. They are always my pillar of strength, without them, I will not be able to get to where I am today. Words are inadequate to express my love and gratitude towards them.

Graduate school is fraught with uncertainties, but this life experience has definitely inspired me to grow stronger and wiser. A self-reminder for my future research path: In pursuit of truth, be patient and open in accepting critiques and new scientific findings, which would only gain clarity along with the expansion of information and knowledge in the temporal dimension.

Table of contents

Statement of Originality.....	3
Supervisor Declaration Statement.....	4
Authorship Attribution Statement.....	5
Acknowledgements.....	6
Table of Contents.....	7
List of Figures	10
List of Tables	13
Abbreviations.....	14
Abstract.....	16
Chapter 1: Introduction.....	18
1.1 The discovery of DNA and chromatin	18
1.2 Chromatin organization and dynamics.....	20
1.3 Nucleosome, the fundamental unit of chromatin	24
1.4 Factors affecting nucleosome dynamics.....	26
1.5 Recognition of nucleosome by binding factors	27
1.6 Poly(ADP-Ribose) Polymerase 1 (PARP1)	29
1.7 Sirtuin 6 (SIRT6).....	38
1.8 Linker histones (H1).....	46
1.9 Project aims	54
Chapter 2: Materials and Methods.....	56
2.1 Materials.....	56
2.1.1 Chemicals	56
2.1.2 Buffers and media composition.....	57
2.1.3 Competent cells and Enzymes	60
2.1.4 Gel composition.....	61
2.1.5 Equipments	61
2.1.6 DNA and Protein Constructs	62
2.2 Methods.....	64

2.2.1 Large scale DNA production	65
2.1.1.1 Palindromic DNA treatment	67
2.1.1.2 Non-Palindromic DNA treatment	69
2.2.2 Human core histones production	70
2.2.3 PARP1 and SIRT6 protein expression and purification	72
2.2.4 Linker histones production	74
2.2.5 Human histone octamer refolding	76
2.2.6 Nucleosome reconstitution	76
2.2.7 Electrophoretic mobility shift assay (EMSA)	77
2.2.8 Bio-Layer Interferometry (BLI)	78
2.2.9 Electron Microscopy (EM)	78
2.2.10 Trace Fluorescence Labelling	79
2.2.11 Screening and optimization of the co-crystallization of the protein-nucleosome complex	80
2.2.12 Linker histone-Nucleosome Assembly Crystallization	81
2.2.13 Heavy metals crystallization soaking for SAD	81
2.2.14 X-ray Diffraction Data Collection	82
2.2.15 Data processing and Model building	82
Chapter 3: Results and Discussion Part 1 – Preparation of DNA, NCP and Proteins	84
3.1 DNA production	84
3.1.1 Palindromic cohesive ends DNA constructs	84
3.1.2 Non-palindromic cohesive ends DNA constructs	87
3.2 Protein production	89
3.2.1 Core histones preparation	89
3.2.2 PARP1 production	90
3.2.3 SIRT6 production	92
3.2.4 Linker histones production	93
3.3 Human Histone Octamer refolding	96
3.4 Nucleosome core particle (NCP), Nucleosomes (NUCs), and Di-Nucleosomes (diNUCs) reconstitution	97
Chapter 4: Results and Discussion Part 2 – Structural Recognition of Nucleosome by PARP1	100
4.1 Binding of PARP1 to nucleosome visualized by EMSA	100
4.2 Electron microscopy (EM) analysis of PARP1 protein assembled with nucleosome	101
4.2.1 Negative Staining Electron Microscopy	101
4.2.2 Cryogenic Electron Microscopy (Cryo-EM)	107

Chapter 5: Results and Discussion Part 3 – Structural Characterization of SIRT6 binding to nucleosomes	110
5.1 Binding of SIRT6 to nucleosome core particle by EMSA.....	110
5.2 Bio-Layer Interferometry (BLI)	110
5.3 Electron microscopy (EM) of SIRT6-NCP 601L assembly.....	115
5.3.1 Negative Stain EM.....	115
5.3.2 Cryo-EM.....	116
5.4 Co-crystallization attempts of the SIRT6-NCP complex	119
Chapter 6: Results and Discussion Part 4 – Linker histones binding in chromatin from the structural perspective	122
6.1 Nucleosome-Linker Histone Assembly (NLH) EMSA studies	122
6.2 X-Ray Crystallization and Data Collection.....	124
6.2.1 Nucleosome-Linker Histone Assembly (NLH).....	125
6.2.2 Assembly of Di-Nucleosome with Linker Histone H1.0 (NLH H1.0-diNUC)	128
6.3 Structural Refinement and Analysis.....	132
6.3.1 NLH H1.0-NUC structures	132
6.3.2 NLH H1.0-diNUC structures.....	142
Chapter 7: Current Perspectives and Future Directions.....	154
References.....	157
Appendix I	181
Appendix II	186
Appendix III.....	188

List of Figures

1. The observation of beads on a string structure of chromatin under electron microscope. Picture adapted from (D. E. Olins & Olins, 2003).
2. A model of hierarchical packaging of 10 nm chromatin fibers into the interphase chromosome. Figures adapted from (Hansen, et al., 2018).
3. The hierarchical view of nucleosome and chromatin organization. Picture adapted from (van Emmerik & van Ingen, 2019).
4. The nucleosome core particle (NCP) structure was obtained from pdb file 1KX5.
5. The PARPs family. They are categorized into four main classes depending on their functional domains. Figure is adapted from Schreiber et al., 2006.
6. The schematic domain organization of PARP1.
7. The structure of the interface between PARP1 and DNA double strand end (PDB code: 4DQY).
8. The structure of zinc finger 1 and 2 of PARP1 binding to the DNA single strand break (PDB code: 2N8A).
9. The schematic representation of the seven human SIRT family members.
10. Functions of SIRT6. Picture adapted from (Kugel & Mostoslavsky, 2014).
11. Structure of SIRT6 from residue 11 to 295 (PDB code: 3PKI).
12. Mammalian linker histone variants. Picture adapted from (Happel & Doenevke, 2008).
13. X-ray structure of Avian H5 (PDB code: 1HST).
14. Off-dyad and on-dyad binding mode of linker histone. The picture was adapted from Fyodorov et al., 2017.
15. Summary of nucleosome-linker histone structures in chronological order.
16. Overview flow chart of the methods and approaches that were attempted to achieve project aims.
17. The enzymatic treatment steps of the plasmid to produce the 601L DNA construct. For illustration purpose, only two repeats are shown here.
18. 601L DNA preparation.
19. DNA CMS 169s production.
20. Design of CMS 169s and CMS169sn.
21. Production of DNA CMS169sn.
22. DNA 343 production.
23. DNA 349 production.
24. Human core histone H2B preparation.
25. PARP1 purification steps.
26. SIRT6 purification steps.
27. H1.0 purification.
28. The size exclusion purification of human histone octamer.
29. Reconstitution of NCP and NUCs.
30. 6% native PAGE assessment of small scale diNUCs 343, 345, 349, 351, and 353 reconstitution to screen for the optimal DNA: HO ratio.

31. Final optimized DNA: HO ratio for large scale diNUCs 343, 345, 347, 349, 351, and 353 reconstitutions that were harvested for X-ray crystallization set ups.
32. Assessment of small scale diNUCs reconstitution for 361, 351z, 353z, and 355z constructs to screen for the optimal DNA: HO ratio via 6% native PAGE.
33. Final optimized DNA: HO ratio for large scale diNUCs reconstitution of 361, 351z, 353z, 355z, and 357z constructs that were utilized for X-ray crystallization set ups.
34. Increasing amount of PARP1 protein were titrated to the same amount of NCP 601L sample leading to increasing intense 1:1 and 1:2 complex bands as portrayed on the 6% native PAGE gel.
35. The EMSA of PARP1 binding to NCP 601* and NUC CMS-1m.
36. Electron microscopic negative stained image of NCP alone, as a control.
37. Negative stained sample of PARP1-NCP complex with T12 EM microscope.
38. Effect of NaCl on the density of the PARP1-NCP sample on the negative stained grid.
39. Comparison of commercial versus homemade carbon coated grids.
40. Effect of centrifuging the sample before loading onto the grid.
41. PARP1-NUC CMS-1m complex (1.2:1).
42. PARP1-NCP 601* complex (1.2:1).
43. Challenges in structural characterization of PARP1-NUC CMS-1m complex (1.2:1) using cryo EM technique.
44. Model of PARP1-induced conformational change in nucleosomal DNA after binding to the linker DNA (detected as DNA double-stranded break) as proposed by Sultanov et al., 2017.
45. EMSA binding of SIRT6 to 601L NCP.
46. Principles of Bio-Layer Interferometry.
47. Qualitative test of NCP601L-SIRT6 Interaction.
48. Buffer optimization of NCP601L-SIRT6 binding assay.
49. Quantitative binding kinetics assay of NCP601L-Sirt6 interaction.
50. The raw data was fitted with 1:1 Langmuir model.
51. The raw data was fitted with 2:1 heterogeneous ligand model.
52. The negative stained SIRT6-NCP 601L on the carbon-coated grid.
53. The negative stained SIRT6-NCP 601L on the carbon-coated grid with sample concentration 0.5 μ M.
54. The SIRT6-NCP 601L assembly with a concentration of 1.5 μ M.
55. The images taken on graphene oxide coated holey carbon grids.
56. Optimized SIRT6-NCP 601L assembly crystallization conditions.
57. 15% SDS-PAGE assessment to determine the content of the crystals.
58. Linker histone H1.0 was titrated against same amount of NUC 169s.
59. EMSA of H1.0 binding to diNUC 343 and diNUC 349.
60. NLH crystals.
61. Schematic representation for diNUC343 linker DNA.
62. Crystals of H10-diNUCs assembly.
63. X-ray diffraction images of crystals H10-diNUC 349 assembly soaked in different soaking buffer conditions.

64. NLH H1.0-NUC169s.
65. An $F_o - F_c$ omit electron density map (magenta; contoured at 3σ ; H1.0 atoms omitted from the model) corresponding to the H1.0 ($\alpha 1$ helix at the front) in the NLH H1.0-NUC169s structure.
66. Inter-nucleosomal contacts provide H1.0 binding sites.
67. The packing of the NLH H1.0-NUC169sn asymmetric unit in the crystal lattice.
68. The asymmetric unit of NLH H1.0-NUC169sn.
69. H1.0 is contacting 3 nucleosomes in the crystal packing of NLH 169sn: host nucleosome, paired nucleosome, and proximal nucleosome.
70. gH5-NUC167 (PDB: 4QLC) was superimposed onto H1.0 in the NLH H1.0-NUC169sn.
71. The two orientations that can be adopted by one base-paired NLH H1.0-NUC169sn nucleosome pseudo-dimer due to the two distinct restriction sites: KpnI and SacI.
72. The superimposition of NLH H1.0-NUC169s and its symmetry mate onto NLH H1.0-NUC169sn. Superimposition was done based on one nucleosome only to illustrate stacking geometry.
73. The arrangement of the NLH H1.0-diNUC343 in the crystal lattice.
74. The asymmetric unit of the NLH H1.0-diNUC343.
75. An $F_o - F_c$ omit map without (A) and with (B) H1.0 in the model. The side chain density of H1.0 was observed as portrayed.
76. H1.0 binding to NUC 1 and NUC 4 in the NLH H1.0-diNUC343 structure.
77. NUC 2, 3, and 4 are superimposed onto NUC 1 to compare the linker DNA distance from H1.0.
78. NLH H1.0-diNUC349 crystal lattice arrangement.
79. A) A top view and B) a side view of an asymmetric unit of NLH H1.0-diNUC349.
80. H1.0 binding in the NLH H1.0-diNUC349 crystal.
81. gH5-NUC167 (PDB: 4QLC) superimposed onto H1.0 in crystal lattice of NLH H1.0-diNUC349.
82. A) A side view and B) a top view of NLH H1.0-diNUC353 crystal lattice arrangement.
83. The asymmetric unit of NLH H1.0-diNUC353 and the order of nucleosome connectivity is shown as a guide.
84. The asymmetric unit of NLH H1.0-diNUC353 superimposed onto NLH H1.0-diNUC343.
85. Illustration of a potential fiber-like arrangement that could be achieved through sticky end base-pairing with the z series diNUC constructs.

List of Tables

1. The differences in the processing steps for preparation of DNA 601L and DNA 169s.
2. The concentration and strategies of heavy metals soaking for nucleosome-linker histone assembly crystals.
3. The purification steps and final yield as well as the purity for linker histone variants H1.0, H1.2, H1.3, and H1.5.
4. The types of Quantifoil holey carbon grids tested for SIRT6-NCP 601L assembly cryo-EM experiments.
5. The strategies and outcomes of heavy metals incorporation for NLH crystals phasing and labelling.
6. Data collection and refinement statistics tables of A) NLH H1.0-NUC169s and B) NLH H1.0-NUC169sn.
7. Summary of NLH H1.0-NUC169s interaction points.
8. Summary of NLH H1.0-NUC169sn interaction points.
9. Data collection and refinement statistics for the NLHs.
10. The detailed interaction points between H1.0 and NUC1, NUC4 in a NLH H1.0-diNUC343 asymmetric unit.
11. Distance between the H1.0 key interacting residues and the nearest linker DNA.
12. Summary of H1.0 binding mode across various NLH structures.

Abbreviations

ATP	Adenosine triphosphate
BLI	Bio-Layer Interferometry
BME	β -mercaptoethanol
CIA	Chloroform-isoamyl alcohol
CIAP	Calf Intestine Alkaline Phosphatase
CTD	Carboxy Terminal Domain
CVC	Chromatin Volume Concentration
DMSO	Dimethyl Sulfoxide
DNA	Deoxyribonucleic acid
DSB	Double Strand Break
DTT	Dithiothreitol
EDTA	Ethylenediaminetetraacetic Acid
EM	Electron Microscopy
EMSA	Electrophoretic Mobility Shift Assay
FDA	Food and Drug Administration
FPLC	Fast Performance Liquid Chromatography
FRAP	Fluorescence Recovery After Photobleaching
HCL	Hydrochloric Acid
HO	Histone Octamer
IMAC	Immobilized Metal Affinity Column
IPTG	Isopropyl- β -D-thiogalactopyranoside
LH	Linker Histone
LM-PGA	Low Molecular Weight Poly- γ --glutamic acid
MPD	2-methyl-2,4-pentanediol
NCP	Nucleosome Core Particle
NLH	Nucleosome Linker Histone assembly
NMR	Nuclear Magnetic Resonance
NUC	Nucleosome
NTA	Nitrilotriacetic Acid

NTD	N-Terminal Domain
PAGE	Polyacrylamide Gel Electrophoresis
PALM	Photoactivated Localization Microscopy
PEG	Polyethylene Glycol
PMSF	Phenylmethanesulfonylfluoride
PTM	Post-translational Modification
SAD	Single-wavelength Anomalous Diffraction
SAXS	Small Angle X-ray Scattering
SDS	Sodium Dodecyl Sulfate
SDS-PAGE	Sodium Dodecyl Sulfate - Polyacrylamide Gel Electrophoresis
SPR	Surface Plasmon Resonance
SSB	Single Strand Break
STORM	Stochastic Optical Reconstruction Microscopy
TAD	Topological Associated Domain
TEMED	Tetramethylethylenediamine
TEV	Tobacco Etch Virus

Abstract

One of the key players in genomic structure and dynamics is chromatin modification enzymes, which regulate epigenetic marks and alter chromatin activity. The first part of this thesis is focused on poly(ADP-ribose) polymerase 1 (PARP1) and sirtuin-6 (SIRT6), two chromatin modification enzymes that modulate distinct cellular pathways. Both proteins are involved in the DNA damage response and are crucial for upholding genome stability. PARP1 is a multipartite polymerase that modulates transcription, DNA damage repair and cell-death signaling and is an important cancer drug target since it plays a key role in maintaining genomic stability. The enzyme specifically recognizes single- and double-stranded DNA breaks, resulting in activation of the poly[ADP-ribos(e)]ylation (PARylation) activity towards self (PARP1) and other protein substrates. This automodification and PAR attachment to other nuclear factors evokes a cascade cellular response that leads to recruitment of additional DNA repair factors and chromatin remodeling machinery to the site of DNA damage. PARP1 is of further interest in chromatin studies because it is known to bind more strongly to nucleosomal substrates as compared to naked DNA. The SIRT6 protein is also involved in multiple signalling pathways, including DNA repair, telomere maintenance, glycolysis and inflammation. It is a histone deacetylase that is specific for K9 and K56 of H3.

The first part of this thesis work was aimed at structural and functional characterization of PARP-nucleosome and SIRT6-nucleosome recognition. The original objective was to determine the structure of a PARP1-nucleosome complex by X-ray crystallography or electron microscopy (EM) and to solve the structure of a SIRT6-nucleosome core particle complex. Full length PARP1 and SIRT6 proteins were overexpressed and purified. The enzymes were subsequently used to generate homogenous assemblies with a variety of different nucleosomal constructs, which were subjected to both crystallographic screening and EM analysis. In SIRT6 co-crystallizations with nucleosome core particle, crystals were obtained under several different conditions, but these did not contain SIRT6 incorporated into the lattice. Additionally a number of different preparative approaches were taken to try to optimize EM grids and material integrity. Alternative nucleosome constructs and conditions were also explored in an attempt to improve both crystallization behavior and EM imaging resolution. Nonetheless, a recent publication indicates that PARP1 binding to the nucleosome in fact induces (partial) unfolding of the

nucleosome core. This would mean that PARP1-nucleosome assemblies are likely highly dynamic and varied in structural nature, and as such would constitute highly challenging candidates for study by X-ray and EM approaches.

Condensing nucleosomes into highly compact states coincides with an epigenetic mechanism in chromatin, which is required for gene regulation, maintaining genomic integrity and generating mitotic chromosomes for cell division. Chromatin compaction is facilitated by nucleosome association of linker histones, which are essential proteins in higher eukaryotes. However, in spite of its importance for genomic function, the biological relevance of different types of condensed nucleosomal structures observed *in vitro* is unclear, and we lack a unified understanding of how linker histones bind to nucleosomes and promote compaction in the cell.

Chromatin's structural characterization is complicated by its heterogeneous composition in the cell and the conformational disorder of nucleosome assemblies *in vitro*. One approach is to explore methods that would allow one to elucidate crystal structures of assemblies with differing lengths of linker DNA connecting the nucleosome core regions. Yet, as with the nucleosome core particle, the challenge is finding constructs that yield well diffracting crystals. To overcome this obstacle, we developed an approach that promotes ordered lattice formation, which appears to have general utility for chromatin structural studies. In the second part of this thesis, X-ray structures of novel nucleosome-linker histone assemblies based on engineered nucleosomal constructs are described. The structures reveal different nucleosome packing configurations and linker histone binding modes and demonstrate that a slight alteration in the linker DNA length can result in a dramatically distinct configuration. We propose that different nucleosome packing configurations in crystals obtained by varying linker DNA length could serve as a platform for both rationalizing and understanding chromatin structural heterogeneity in the cell.

Chapter 1 Introduction

1.1 The discovery of DNA and chromatin

DNA was initially discovered by Sir Friedrich Miescher in 1871, while studying the leukocytes, he isolated a molecular substance containing high levels of phosphorus, which he coined “nuclein” (Dahm, 2005). The advancement in a field of microscopy and dyes subsequently allowed the investigation of chromosomes in cell division. The term “chromatin” was first introduced by Walther Flemming in 1879 when he observed by light microscope the easily-stainable threads during the process of mitosis in the cell nucleus, suggesting that chromatin might be nuclein.

Albrecht Kossel’s work discovering that nucleic acids, DNA and RNA, were composed of nitrogen bases: adenine, cytosine, guanine, thymine/uracil earned him a Nobel Prize in 1910 (Nobelprize.org). The chemical nature of the chromatin was elucidated as a complex of nucleic acids and proteins at the end of 19th century. The attribution of nuclein or chromatin as the hereditary substance was proposed by Theodor Boveri and Edmund B. Wilson (Boveri, 2008; Droscher, 2002). The Hershey–Chase experiment that was conducted in 1952 confirmed that the DNA from bacteriophage, but not the protein coat, was needed to confer its ability to produce progeny in bacterium (Hershey and Chase, 1952). After a century of studies, Oswald T. Avery, Colin MacLeod, and Maclyn McCarty also demonstrated that DNA, but not the protein, is the hereditary substance (Avery et al., 1979). This has laid the foundation that DNA was most likely the molecule of life which contains all the instruction for life and could pass the traits along to the next generation. This prevailing view continued for decades until the advent of epigenetics, which revealed that histone modifications can also be hereditary (Felsenfeld, 2014).

Strides were made not only in the function of genetic material, in 1953, James Watson and Francis Crick using a stick-and-ball models revealed the structure of DNA, which consisted of the Watson-Crick base-pairing, in the helical form, which was uncovered by Rosalind Franklin and Maurice Wilkins, through X-ray diffraction (Watson and Crick, 2003; Wilkins et al., 1953). In 1962, Francis Crick, James Watson, and Maurice Wilkins were awarded with Nobel Prize for their discovery of the double helix structure of the DNA. In the meantime, the

mechanism of DNA replication was proposed to be semi-conservative via radioisotopic labelling experiments by Meselson and Stahl (Meselson and Stahl, 1958).

Figure 1. The observation of beads on a string structure of chromatin under electron microscope. Picture adapted from (Olins and Olins, 2003).

The early clue for existence of nucleosome was made based by the observation of the beads on a string by electron microscopists in the 1970s (Figure 1) (Olins and Olins, 1974; Woodcock et al., 1976). A cross-linking experiment of chromatin illustrated that a distinct histone octamer was formed by two copies each of H2A, H2B, H3, and H4, with DNA wrapped around the histone octamers forming nucleosomes as the repeating units in chromatin (Kornberg, 1974; Richmond et al., 1984). Richmond's group unraveled the atomic details of nucleosome core particle (NCP) by solving its first X-ray structure at 2.8 Å in 1997, subsequently in 2002 and 2003, a high resolution structure was solved to 1.9 Å, revealing the solvent-mediated interactions and the details of the DNA bending (Davey et al., 2002; Luger et al., 1997). The

discovery of the nucleosome transformed the chromatin research field, from the higher-order packing of DNA to DNA-based processes, focusing on the biological significance of nucleosome in chromatin function (Kornberg and Lorch, 1999).

1.2 Chromatin organization and dynamics

In humans, most of the cells have a diploid genome, with 23 pairs of chromosomes, which totals in 6 billion of DNA base-pairs, spanning about 2 meters of DNA. The packaging of such enormous amount of genetic material into the nucleus, which is only several micrometers across is a challenging task, facilitated by wrapping of the DNA around the histone octamers, to form a primary structure of 10 nm chromatin fiber. The primary chromatin fiber is then further organized into a highly compact structure while maintaining the cell functionality through additional levels of fine-tuning. Deciphering the chromatin structural states are important for understanding of the functional impact of chromatin 3D architecture in regulating the accessibility to the DNA for genomic processes, such as DNA transcription, replication, and repair (Ausio, 2015; Li and Reinberg, 2011).

Chromatin exists in diverse conformational states. For non-dividing cells, the interphase chromatin is relatively decondensed and distributed throughout the nucleus. The compaction level of the chromatin fiber during interphase reflects the various functional states, heterochromatin, which is transcriptionally silenced, and euchromatin, which is transcriptionally accessible. These chromatin terms were derived from staining of DNA molecules using non-specific heavy metal as observed through transmission electron microscopy (TEM) (Belmont et al., 1989; Bohrmann and Kellenberger, 1994; Olins et al., 1989). Upon entering into mitosis, the chromatin condenses into discrete chromosomes, to prepare for cell division. As cell cycle progresses, the chromosome conformation changes drastically and condensin proteins are required for the processes (Gibcus et al., 2018).

As far as the higher order structure over the years, most studies supported the 30 nm fiber as the secondary structure (Felsenfeld and McGhee, 1986; Grigoryev and Woodcock, 2012; Robinson et al., 2006; Tremethick, 2007; Woodcock and Dimitrov, 2001; Zhu and Li, 2016), which was proposed to be the first step of hierarchical folding to form transcriptionally dormant

chromatin (Li et al., 2010). However, in spite of the 30 nm fiber readily assembling *in vitro*, its existence remained elusive due to a lack of supporting *in-vivo* evidence. Some groups have suggested that 30 nm fiber may only exist in certain cell types (Woodcock, 1994), or is absent from mitotic chromosomes *in-situ*, which are instead in a highly interdigitated and disordered state, comparable to a polymer melt (Eltsov et al., 2008).

In recent years, the paradigm of chromatin fiber has been gradually shifting, with the advancement of microscopic resolution, more and more evidence supports the 10 nm chromatin fiber to be the predominant structure (Hansen et al., 2018; Nozaki et al., 2017; Ou et al., 2017; Ricci et al., 2015). Chromatin fibers consist of regions with varying levels of nucleosome occupancy, which are interspaced by nucleosome-depleted stretches, especially at the promoter, enhancer, and terminator regions, this confers arguments against a hierarchical organization of chromatin fiber, where in 30 nm fiber there would be no ordered or periodic nucleosome positioning and occupancy along the DNA double helix (Struhl and Segal, 2013).

Utilizing quantitative super-resolution nanoscopy (Stochastic optical reconstruction microscopy, STORM) and computer simulations, nucleosomes were shown to assemble in discrete nanodomains, termed heterogeneous “clutches” with varying densities and sizes, interspersed with nucleosome-depleted regions. Despite heterogeneity in a given nucleus, larger and denser clutches were observed in average differentiated cells, as compared to stem cells, displaying a striking correlation between spatial distribution, size of compacted nucleosome clutches and cell pluripotency (Ricci et al., 2015).

A recent study utilized ChromEMT, which combined DNA labelling method and electron microscopy tomography, allowing large 3D nuclear volume ultrastructure to be observed *in-situ*. There was no hierarchical organization of chromatin fiber observed in interphase or mitotic cells, instead disordered chains with diameters between 5nm to 24nm, form distinct concentration densities and structural conformations. Chromatin volume concentrations (CVCs) *in situ* were calculated for quantification purposes, where interphase chromatin subvolumes ranged from 12% to 52%, while mitotic chromosome subvolumes were above 40% (Ou et al., 2017).

Mammalian interphase chromatin and mitotic chromosomes were depicted to comprise 10 nm zig-zag chromatin fiber assembled into a highly disordered and interdigitated state (Figure

2) without folding into 30 nm fiber, mimicking a ‘polymer-melt’ structure using SAXS, cryo-electron microscopy *in-situ*, synchrotron X-ray scattering, and electron spectroscopic imaging with tomography analyses (Fussner et al., 2012; Hansen et al., 2018; Joti et al., 2012; Nishino et al., 2012). Furthermore, combining super-resolution imaging Photo-Activated Localization Microscopy (PALM) and single-nucleosome tracking in mammalian live cell, chromatin dynamicity was demonstrated, where compact domains with a peak diameter of ~160 nm were moving coherently, and heterochromatin-rich regions, which showed higher number of domains with lesser movement. The similar chromatin domains were observed in mitotic chromosomes, which were organized by cohesin complexes (Nozaki et al., 2017). These estimated chromatin domain sizes are in good agreement with the contact domains of ~185 kb (Rao et al., 2014).

Figure 2. A model of hierarchical packaging of 10 nm chromatin fibers into the interphase chromosome. Figure adapted from (Hansen et al., 2018).

Nonetheless, correlation of transcriptional activity through visualization of chromatin fiber compaction using microscopy approaches might not precisely reflect the genomic transcriptional and regulatory mechanisms. Thus, chromosome conformation capturing

experiments (3C, 4C, 5C, Hi-C) were widely adopted, to provide comprehensive information on local and genome-wide chromatin interaction, allowing deciphering of long-range genome interaction and genome regulation, but little on the spatial perspective, as opposed to microscopy approach (Belmont, 2014).

Topologically-associated domains (TADs) comprises a three-dimension scaffold which represent the functional domains of gene regulation. The chromatin fiber is organized spatially as a functional unit of genome to facilitate long-range transcriptional regulation, however the underlying mechanism remains uncertain (Dixon et al., 2012). TADs in mammalian genomes ranges from tens of kb up to 1 or 2 Mb, with an average around 800 kb (Dekker and Heard, 2015; Nora et al., 2012). Architectural proteins, CTCF and cohesin, are important for intra-TAD interactions (Rao et al., 2017; Zuin et al., 2014). High-resolution Hi-C, 5C maps, and super-resolution microscopy have provided insights into finer details of TAD organization, smaller domains of 100-200 kb were visible, however their relationship with TADs are not clearly delineated (Nozaki et al., 2017; Rao et al., 2014). Although TADs are generally present, they are absent in some cases, for instance the inactive X chromosome appears to be depleted of TADs and no TADs were observed along the mitotic chromosomes (Deng et al., 2015; Naumova et al., 2013). Nonetheless, genetic rearrangement that disrupts TAD organization was shown to be implicated in pathogenic phenotypes by altering the long-range enhancer-promoter interactions (Groschel et al., 2014; Lupianez et al., 2015).

Condensation of nucleosomes into highly compact states coincides with an epigenetic mechanism in chromatin, which is required for gene regulation, maintenance of genomic integrity and generating mitotic chromosomes for cell division (Johnson and Dent, 2013; Tsompana and Buck, 2014). The dynamic nature of chromatin structure is finely regulated in order to confer the specificity of distinct DNA processes such as transcription, replication, and repair. Together with DNA methylation, histone modifications, histone variants, and various types of architectural proteins and remodelers, these factors are imposing additional levels of epigenetic control of chromatin structure and function.

1.3 Nucleosome, the fundamental unit of chromatin

As mentioned above, the fundamental repeating unit of the chromatin structure is the nucleosome, which is composed of around 200 base pairs (bp) of DNA wrapped around a core histone octamer (Figure 3) (Kornberg, 1974). Each nucleosome can be further subdivided into the nucleosome core and linker DNA regions. The nucleosome core contains ~147 bp wrapping around the histone octamer with two copies of each of the four core histones, H2A, H2B, H3 and H4 (Kornberg and Thomas, 1974; Richmond and Davey, 2003; Tan and Davey, 2011).

Linker DNA length varies from 10 to 90 bp and connects the nucleosome core region in a beads-on-string structure with a diameter of 11 nm (Oudet et al., 1975). Although the linker DNA length between nucleosomes *in vivo* is non-uniform, the nucleosome repeat lengths (NRLs) of 188 bp and 196 bp are the most commonly present in eukaryotes (Perisic et al., 2010; Widom, 1992). Such heterogeneity would play a role in determining the nucleosome-nucleosome interaction, and even the folding manner into chromatin fiber (van Holde and Zlatanova, 2007).

Nucleosomal DNA left-handedly wraps around the histone octamers in a total of 1.65 turns. Interaction is made between negatively charged DNA and positive charged histone proteins at a regular periodicity of 10 bp, which is also known as superhelical location (SHL) (Luger et al., 1997). The base pair that is centered on the nucleosome dyad, also known as the SHL 0, gives the NCP a pseudo-2-fold symmetry. Each SHL is referred to the location where the DNA major grooves facing toward the histone surface and the SHL ranges from -7 to 7, relative to the dyad location (Cutter and Hayes, 2015).

The nucleosome serves three basic functions which are genomic packaging, mediating higher-order chromatin compaction and serving as a scaffold for epigenetic marks to be displayed for chromatin factors to bind (McGinty and Tan, 2015). Higher order association between nucleosomes in chromatin can determine the gene activation status. For instance, the highly compacted state which is also known as heterochromatin generally indicates the silenced regions while the open chromatin state, the euchromatin is more transcriptionally active due to higher accessibility for chromatin protein factors involved in cellular processes (Campos and Reinberg, 2009; Kouzarides, 2007; Li et al., 2007). While at the nucleosomal level the nucleosomes are not rigid structures, in fact, they are dynamically remodelled to ensure a highly

regulated access to the packaged DNA (Koyama and Kurumizaka, 2018; Luger, 2006; Widom, 1998).

Figure 3. The hierarchical view of nucleosome and chromatin organization. Picture adapted from (van Emmerik and van Ingen, 2019).

1.4 Factors affecting nucleosome dynamics

The nucleosomal DNA wrapping and unwrapping is DNA-sequence and position-dependent. The DNA segment at the entry and exit site is more accessible from the either end of the nucleosome DNA and it is a sequence-directed nucleosome positioning (Gansen et al., 2018). The rate of DNA unwrapping and rewinding can be quantified by a site exposure equilibrium constant, K_{eq} , which ranges between 10^{-1} and 10^{-6} by taking into account the location on nucleosome (Anderson and Widom, 2000; Lowary and Widom, 1998a; Widom, 2001). This dynamic behaviour allows nucleosome to fine tune the binding of the DNA-binding proteins through differential accessibility (Kobayashi and Kurumizaka, 2019).

Other than DNA sequences, the factors that modulate nucleosome dynamics are post-translational modifications (PTMs) of histones, substitution of histone variants and the engagement of nucleosome recognition factors (Kale et al., 2019; Lequieu et al., 2017; Winogradoff and Aksimentiev, 2019). PTMs are covalent modifications that can be enzymatically added and removed, with different modifications for distinct but specific downstream functions (Bannister and Kouzarides, 2011; Kouzarides, 2007). Among the PTMs, the most extensively studied ones are acetylation, methylation, phosphorylation, ubiquitylation and ADP-ribosylation (Hassa et al., 2006; Nowak and Corces, 2004; Shilatifard, 2006; Sterner and Berger, 2000; Zhang and Reinberg, 2001). Nucleosome can undergo different types of structural fluctuations depending on the impact on distinct functioning regions such as histone tails, DNA sequence at the exit-entry region or dyad symmetry axis region, and interface between histone dimers (Simon et al., 2011). The functional consequences of these modifications are mainly regulation of chromatin state and recruitment of protein factors or complexes.

Moreover, canonical histones can be substituted by the histone variant that alters the chemical nature of the nucleosome core (Talbert and Henikoff, 2010; Weber and Henikoff, 2014). This substitution changes histone-DNA and histone-histone contacts which could lead to versatile nucleosome wrapping or vice versa (Bilokapic et al., 2018; Shaytan et al., 2016). PTMs can also be exerted on histone variants, thus adding another level of complexity in fine-tuning the dynamics of nucleosome (Loyola and Almouzni, 2007).

In addition, nucleosome stability and dynamics can also be regulated by protein factors such as linker histones, histone chaperones, ATP-dependent chromatin remodelers and DNA-

binding proteins. Linker histones have been established as an obligatory chromatin architectural protein for their contribution towards higher order chromatin structure formation (Vignali and Workman, 1998; Woodcock and Ghosh, 2010). Other than facilitating the replacement of histone variants and stabilization of histones in the absence of DNA, histone chaperones are known to have a role in chromatin assembling and disassembling for DNA repair, replication and transcription purposes as well (Das et al., 2010; Loyola and Almouzni, 2004). Besides, histone chaperones were revealed to have some functional crosstalk with epigenetic marks in modulating chromatin dynamics (Avvakumov et al., 2011). On the other hand, chromatin remodelers are essential for genomic function, as they utilize the coupling of ATP hydrolysis to catalyse diverse structural transformations of nucleosomes (Clapier et al., 2017). A growing evidence for the activities of nucleosome remodeling correlate with the PTM machineries (Petty and Pillus, 2013).

Epigenetic mark or the PTM is undeniably one of the crucial characteristics in determining the chromatin states in conjunction with other factors. In short, there are many elements contributing to nucleosome dynamics and to the higher order structure, as well as to chromatin accessibility. All these factors can crosstalk and work on par to strictly fine-tune the regulation processes.

1.5 Recognition of nucleosome by protein factors

Nucleosome is a disk-shaped particle of around $100\text{\AA}$ in diameter. This 200 kDa complex provides extensive surface heterogeneity for interactions with chromatin factors. The nucleosome displays about $74\,000\text{\AA}^2$ of multifaceted, accessible solvent surface. The thickness of the nucleosome ranges from around $25\text{\AA}$ at the dyad region to about $60\text{\AA}$ at the H2B α C helices protrusion area. Primarily, chromatin factors can interact with one or more of the three nucleosome parts which are the histone tails that splay out from the core histones, the histone octamer disk surface and the nucleosomal DNA (Figure 4).

Figure 4. The nucleosome core particle (NCP) structure was obtained from pdb file 1KX5. H2A is shown in yellow; H2B was shown in red; and H3 and H4 are shown in blue and green respectively. A) The stick model of the NCP. B) The mesh surface representation of the NCP. The figures were created using Pymol (DeLano, 2002).

The extended histone tails from the histone core make up about 25% of the histone mass and range from 15 to 36 amino acids. These flexible tails protrude from the nucleosome surface or between the DNA gyres. They can bind intranucleosomal DNA or DNA and histones from the neighbouring nucleosomes (du Preez and Patterson, 2013). Histone tails are abundant in PTMs which are involved in epigenetic regulation. The presence of many arginine and lysine residues render the histone tails strong net positive charge which is the basis underlying the recognition by modification factors (Ng and Cheung, 2016).

Recent advances in co-crystallization of chromatin complexes have well-depicted the molecular recognition of the histone octamer disk surface by chromatin factors (Armache et al., 2011; Arnaudo et al., 2013; Barbera et al., 2006; Kato et al., 2013; Makde et al., 2010; McGinty et al., 2014; Yang et al., 2013a). Among all the structures solved, a common binding motif which is the arginine anchor motif was revealed to bind to the H2A/H2B acidic patch. The acidic patch refers to an extensive negatively charged surface situated in a groove on the H2A/H2B dimer surface. The charge is contributed by eight residues, E56, E61, E64, D90, E91, and E92 from H2A as well as E102 and E110 from H2B. This acidic patch is topologically poised for the positive side chain residues like arginine from the chromatin factors to anchor.

The DNA phosphodiester backbone displays a highly negative electrostatic charge surface. Most transcription factors bind to DNA sequences only in nucleosome free regions.

Nonetheless, nucleosomal DNA confers distinct type of interactions with chromatin factors due to the unique curvature and the alignment of the DNA gyres on nucleosome (McGinty and Tan, 2015). Free DNA has a different conformation compared to the nucleosomal DNA; this unique feature can become the recognition mark for some chromatin factors (Makowski et al., 2018; Zhu et al., 2018).

1.6 Poly(ADP-Ribose) Polymerase 1 (PARP1)

PARP family

The PARP family of enzymes catalyze the transfer of an ADP-ribose moiety from a NAD^+ substrate onto the Lys, Glu and Asp residues of the target proteins (Nishizuka et al., 1967). PARP activity is dependent on NAD^+ molecule to be the substrate for the synthesis of poly(ADP-ribose) (PAR). There are at least 17 members in the mammalian PARP family in which all contain a PARP domain and a PARP signature motif (Ame et al., 2004). A new nomenclature of ADP-ribosyltransferases (ARTs) has been proposed since some of the PARP family members are only able to catalyze mono(ADP-ribosyl)ation instead of poly(ADP-ribosyl)ation (Hottiger et al., 2010).

The PARP family members can be categorized into four subfamilies depending on their putative functional domains such as DNA-dependent PARPs (PARP1, PARP2, PARP3), tankyrases (PARP5a, PARP5b), CCCH PARPs (PARP7, PARP12, PARP13.1, PARP13.2) and macroPARPs (PARP9, PARP14, PARP15) (Figure 5) (Schreiber et al., 2006). DNA-dependent PARPs act as a sensor for DNA-strand breaks. Despite the similarities of PARPs 1, 2 and 3 having Trp-Gly-Arg (WGR) motif and catalytic domains, their N-termini are quite distinct. These unique N-termini yield differences in subcellular localization, target specificity and function. Tankyrases consist of large Ankyrin domain repeats that mediate protein-protein interactions. Tankyrase-1 and 2 have some overlapping functions as they associate with each other and share most of the protein interaction partners (Sbodio et al., 2002). CCCH PARPs are named as such due to the presence of Cys-Cys-Cys-His zinc fingers that bind to RNA and the WWE domains that can bind to PAR. MacroPARPs contain macrodomain folds that can bind ADPR and derivatives.

Figure 5. The PARPs family. They are categorized into four main classes depending on their functional domains. Figure is adapted from Schreiber et al., 2006.

PARP family members display a broad variety of subcellular localizations, activities and roles as well as target partners. Impact of PARPs can be understood from the molecular, cellular, and biological levels. At the molecular level, poly(ADP-ribosyl)ation is one of the post-

translational modifications that can alter the properties of the protein acceptor, markedly depending on the site of ribosylation, as well as the length and branching of the poly(ADP-ribose) polymer (Schreiber et al., 2006). The chemical biology of PAR has been associated with different types of regulatory mechanisms such as serving as an interaction scaffold, inhibition of protein-protein interaction and regulation of ubiquitylation (Gibson and Kraus, 2012).

Furthermore, PARPs participate in a wide range of cellular processes, for instance, transcription regulation, DNA repair, chromatin modification, cell death, tumorigenesis, aging and homeostatic mechanism as well as inflammation (Swindall et al., 2013; Vyas et al., 2014).

Functional role of PARP1

PARP1 is one of the most studied members of the PARP family. It is a large, multifunctional enzyme modulated via DNA and protein interactions as well as post-translational modifications (Ame et al., 2004). PARP1 has been well-characterized for its role in DNA damage detection and repair machinery (Ko and Ren, 2012; Tallis et al., 2014). Its catalytic activity is NAD⁺-dependent and is stimulated by more than 500-fold upon binding to DNA breaks, leading to extensive PARylation of the partner proteins such as histones, repair factors, transcription factors and itself (Petrucchio and Percudani, 2008). The negatively charged PAR polymers production on the PARP1 itself and on target proteins results in several outcomes, such as recruitment of repair factors, modulation of chromatin structure, progression of cell death and disruption of the binding with the interacting partners (Schreiber et al., 2006).

Activities of PARP1 have been observed in distinct DNA repair pathways such as base excision repair (BER), nucleotide excision repair (NER), mismatch repair (MMR), single-strand break (SSB) repair as well as double-strand break (DSB) repair through both direct and indirect approaches (Bouchard et al., 2003; Dantzer et al., 2000; Woodhouse and Dianov, 2008). PARP1 detects DNA breaks efficiently through N-terminal zinc finger domains binding (Ikejima et al., 1990) and activates the DNA repair machinery through poly(ADP-ribose)ylation signaling (Dantzer et al., 2006). The immediate PARP1 mediated PARylation response targets a wide range of nuclear proteins and depending on the extent of stress, the response can be transient or sustained. PARP1 knockout mice are highly susceptible to DNA damaging agents and exhibit

considerable genomic instability; suggesting that PARP1 is indispensable in surveillance and maintenance of the genome integrity (de Murcia et al., 1997; Wang et al., 1997).

Besides safeguarding genome integrity (Jeggo, 1998), PARP1 is also involved in transcriptional regulation through modulating chromatin plasticity (Poirier et al., 1982b; Tallis et al., 2014). PARP1 remodels transcription states through both DNA sequence-dependent and independent manners. Polynucleosomes which are nick-free have been depicted to be a more potent stimulator for PARP1 enzymatic activity when compared to other forms of free DNA (Kim et al., 2004). When activated, the automodified PARP1 confers a strong anionic charge through the PAR network to decompact chromatin and increase accessibility towards transcription factors (Kim et al., 2004; Rouleau et al., 2004).

Moreover, PARP1 mediates PARylation on core histones and chromatin modifying proteins such as linker histone H1, chromatin remodelers and transcription factors in order to alter chromatin and regulate gene transcription levels (D'Amours et al., 1999; Krishnakumar and Kraus, 2010). Association of PARP1 with the transcription factors can be non-enzymatic as well, as seen on its co-stimulatory effect on NF- κ B gene expression (Hassa et al., 2003). In fact, PARP1 can serve as a transcription activator or repressor through direct binding to the sequence-specific octamer motif on the promoter region (Ko and Ren, 2011; Simon et al., 2011; Zhang et al., 2002). The recognition motif of PARP1 was found within the promoter region of many DNA repair genes, suggesting that PARP1 also mediates DNA repair through transcription regulation of the specific repair genes (Ko and Ren, 2012).

Additionally, PARP1 engages in modulating cell death and acts on inflammatory pathways mainly through co-activating the NF- κ B factor. Studies using PARP1 knockout mice and cell lines showed impaired expression of NF- κ B-dependent pro-inflammatory genes (Hassa et al., 2003; Hassa and Hottiger, 2002). The role of PARP1 in DNA damage response is to mediate the cell death pathways in the case the DNA damage is irreparable. Prolonged PARylation response which is corresponding to high level of DNA damage promotes cell death instead of repair (Wang et al., 2016). PARP1 is cleaved by caspase into 89 kDa and 24 kDa fragments, with the functional impact still unclear but it is considered as a marker for apoptosis (Soldani et al., 2001). The extensive PARylation leads to PARP1-dependent cell death (parthanatos) via the apoptosis-inducing factor (AIF)-mediated mechanism in a caspase-

independent manner (David et al., 2009; Wang et al., 2011; Yu et al., 2002). On the other hand, over-activation of PARP1 can result in cell death by necrosis due to energy failure by depleting the cellular NAD⁺ pool (Heeres and Hergenrother, 2007; Koh et al., 2005).

Structural features of PARP1 binding to DNA

PARP1 is a 116 kDa protein with its functional domains well-characterized, it contains an amino-terminal with two zinc fingers for DNA break sensing, a third zinc finger that contributes to DNA binding and the activation mechanism, an automodification domain which comprises a BRCT motif, a WGR motif and last but not least, a carboxyl-terminal catalytic domain which consisting of the highly conserved PARP signature motif for NAD⁺ binding (Figure 6).

Figure 6. The schematic domain organization of PARP1. DNA binding domain comprises three zinc fingers and a nuclear localization signal from residues 1 to 372, it is followed by the auto-modification domain from residues 373 to 524. Residues 524 to 650 constitute the WGR domain. The catalytic domain which spans residues 650 to 1014, consists of a PARP signature motif which is involved in the catalysis of NAD⁺. Abbreviations are as follows: Zn = zinc binding domain, NLS= nuclear localization signal, BRCT= BRCA1 C-terminal motif, WGR= tryptophan-glycine-arginine-containing motif.

DNA repair proteins recognize the weakened base stacking and flexible hinge at DNA lesion sites (Yang, 2006). The structure of PARP1 binding to the DNA strand break is essential in giving an insight in the molecular basis of its response to genomic DNA damage. PARP1 comprises six domains with flexible linkers connecting each of them, and in the unbound state, it is in an open conformation structure like “beads on a string”. Structures of PARP1 domains in complex with SSB and DSB models revealed that upon binding to DNA breaks, there is a

conformational change to form a network of inter-domain contacts which lead to cis-activation (Langelier et al., 2012; Lilyestrom et al., 2010). Accordingly, a cooperative and allosteric activation mechanism was proposed for PARP1 mediated DNA damage response.

The binding interface involves the Zn1, Zn3 and WGR domains, which are collectively in contact with the DNA double-strand break (Figure 7A). The interaction is DNA sequence-independent as PARP1 domains mainly come in contact with the DNA backbone. Hydrophobic contacts with the exposed nucleotide bases allow PARP1 to specifically interact with damaged DNA (Figure 7B). Full length PARP1 was modelled by Langeliar et al., which suggested that a network of inter-domain contacts, which are crucial for cis modification and activation of PARP1 upon binding to DNA (Langelier et al., 2012).

Figure 7. The structure of the interface between PARP1 and DNA double strand end (PDB code: 4DQY). A) Zinc-finger 1 is shown in purple blue, zinc-finger 3 in cyan, WGR-CAT domain in magenta, and the zinc ions in red. The zinc-finger 2 and the automodification domain of PARP1 were absent in the structure. This figure was created using Pymol (DeLano, 2002). B) The structure details at the binding interface, the zinc finger 1 and WGR domain are interacting to the DNA in a sequence-independent manner through hydrophobic and charge-based interaction as observed from the figures.

Eustermann et al. revealed the structure of PARP1 binding to the single-strand DNA break by using integrated nuclear magnetic resonance (NMR)/ X-ray approach. Only the zinc

fingers 1 and 2 of PARP1 were used to obtain the crystal structure with the single-strand nicked DNA. The two flexibly linked zinc binding domains were shown to bind cooperatively and adopted single directionality on the DNA with the second zinc finger recognizing the 3' stem before the first zinc finger could bind to the 5' stem (Figure 8). Next, the structure solved by Langeliar et al. was used to combine with this structure to obtain a full-length PARP1 model assembled on the DNA SSB. Complemented by NMR studies, the structure model of full-length PARP1 led to the hypothesis of multi-domain folding pathway upon binding to SSB and resulting in cis-automodification through PAR catalysis (Eustermann et al., 2015).

Figure 8. The structure of zinc finger 1 and 2 of PARP1 binding to the DNA single strand break (PDB code: 2N8A). A) Zinc-finger 1 is shown in green, zinc-finger 2 in cyan and the zinc ions in grey. This figure was created using Pymol (DeLano, 2002). B) The structure details at the binding interface of 5' and 3' stems, the zinc finger 1 (in green) is interacting with the 5' DNA through hydrophobic and charge-based interaction while the zinc finger 2 (in cyan) mostly interacting with charge-based residues as seen from the figures.

PARP1 binding in the context of chromatin

PARP1 can bind to both nucleosomes and various types of DNA. The binding of PARP1 to DNA has been extensively studied, but very little is known about its interaction with chromatin. Other than serving as the first line of sensing of DNA damage, PARP1 has been implicated to play a role in many different cellular functions that the underlying mechanisms were not well-delineated. Inactive PARP1 was demonstrated to associate with nucleosomes and produce a compact and transcriptionally silenced chromatin (Ji and Tulin, 2010; Wacker et al., 2007). On the other hand, earlier studies showed that when PARP1 was automodified and activated, it regulated chromatin plasticity and promoted chromatin decompaction through its enzymatic effect on core histones and other chromatin associated proteins (Huletsky et al., 1989; Kraus and Lis, 2003; Poirier et al., 1982b). Nonetheless, the mode of PARP1 recognition and PARylation on its protein targets has not been fully characterized, especially from the structural perspective.

The molecular basis of PARP1 recognition in DNA and nucleosomes was illustrated by Clarke et al. biophysically. Nucleosomes with two DNA linkers were depicted to be the most preferred substrate for PARP1 binding when compared to free DNA and other types of nucleosomes. A full length PARP1 was required to bind with high affinity to nucleosome, while only the N-terminal fragment of PARP1 was sufficient in DNA binding. Thus, an alternative mode was proposed for PARP1 nucleosome vs. free DNA binding (Clark et al., 2012).

Most studies were done with truncated fragments of PARP1 providing only limited insight into the activation process. More research involving full length PARP1 binding in the context of chromatin could shed light on its potential as a therapeutic target for cancer, aging, and other degenerative diseases.

Therapeutic potential of PARP1

The role of PARP1 in maintaining genomic stability elevated it to being a promising target in cancer therapeutics. PARP1 is exploited by cancer cells to safeguard against cytotoxic mutations and DNA lesions through upregulating its expression and enzymatic activity (Shimizu

et al., 2004). Thus, disruption of DNA repair in cancer cells through developing PARP1 inhibitors (PARPi) has been actively researched for application in chemotherapy (Javle and Curtin, 2011; Pommier et al., 2016). PARPi were designed to structurally mimic the catalytic by-product of PARP1, nicotinamide for acting as an antagonist of the NAD⁺ cofactor in binding to the catalytic domain. 3-aminobenzamide (3-AB) is one of the earliest PARP inhibitors (De Blasio et al., 2003; Thraves et al., 1985). Two separate strategies have been exploited for the development of PARPi for chemotherapy: 1) synthetic lethality and 2) the combination of DNA-damaging therapy.

Synthetic lethality can be achieved through targeting PARP1 in cancer cells with defects in DNA damage response (DDR). Cells with defects in responding to DNA damage such as BRCA-deficient cells are more dependent on PARP1 to maintain the genomic stability, thus PARPi treatment would induce specific synthetic lethality only towards the DDR defective cancer cells but not normal cells (Cerrato et al., 2016). Studies of PARP1 inhibitors in BRCA1 and BRCA2 deficient preclinical models have validated this specific approach (Bryant et al., 2005; Farmer et al., 2005). Consistently, Olaparib (AZD2281) was identified as the first single anti-tumoral agent showing promising activity against cancers associated with BRCA1 or BRCA2 mutations (Fong et al., 2009).

Alternatively, PARPi can be paired with DNA-damaging agents to induce a synergistic cytotoxicity. Extensive preclinical studies showed that PARPi enhance the sensitivity of the cancer cells towards methylating agents, topoisomerase I inhibitors, and ionising radiation (Curtin, 2005; Zaremba and Curtin, 2007). Combination of PARPi with the methylating agent temozolomide (TMZ) displayed a remarkable anti-tumor effect *in vivo* in diverse cell types, including breast, ovarian, prostate, colorectal, pancreatic, lung and B-cell lymphoma cancers (Calabrese et al., 2004; Donawho et al., 2007). Moreover, PARPi Olaparib increased the sensitivity of breast cancer towards platinum drugs in a mouse model (Rottenberg et al., 2008). To date, there are nine PARPi which are under clinical development. A few PARPi are currently undergoing clinical trials in phase I and II, with the exception of BSI-201 that is in phase III trial (Benafif and Hall, 2015). Olaparib was the first PARP inhibitor to obtain accelerated approval by the US Food and Drug Administration (FDA) for BRCA1/2 mutant ovarian cancer therapy

(Brown et al., 2016). Nonetheless, the detailed mechanisms of action of these PARP1 inhibitors are not well understood.

In addition to PARP1's role in DNA damage repair pathway, it has been implicated in other cellular pathways such as inflammation and cell death, making it a potential target in various disorders, such as diabetes, sepsis, stroke and myocardial infarction (Jagtap and Szabo, 2005; Masutani et al., 2005; Peralta-Leal et al., 2009). Mouse models with streptozotocin-induced diabetes and liposaccharide-induced septic shock demonstrated the protective effects of PARP inhibition, which suggest potential for clinical application in the future (Burkart et al., 1999; Mabley et al., 2001).

Furthermore, hyperactivation of PARP1 as a result of severe DNA damage could lead to necrotic cell death through diminishing the NAD^+ and consequently the energetic pool. Under these circumstances, the inhibition of PARP1 or the absence of PARP1 would allow the oxidative stressed cells to die by apoptosis instead of the less desirable necrosis through preserving the cellular NAD^+ pool (Virag and Szabo, 2002). The PARP has been associated with acute neuronal and myocardial cell death following ischaemia reperfusion injury, which was demonstrated through improving the functional outcomes by using treatment with PARP inhibitors to preserve the cellular energetic pools (Skaper, 2003; Szabo et al., 2004). The dramatic consequences of PARP1 hyperactivation are also observed in human diseases such as Cockayne Syndrome and Xeroderma Pigmentosum A, where sustained PARP1 activation induces NAD^+ depletion and leads to mitochondrial defects through inactivation of SIRT1 (Fang et al., 2014; Scheibye-Knudsen et al., 2014). In summary, these findings indicate therapeutic benefits of PARP inhibitors not only in treatment of cancer, but also inflammatory, neurological and cardiac disorders.

1.7 Sirtuin 6 (SIRT6)

Background on SIRT family

Sirtuins, Silent Information Regulator Two proteins, are a family of proteins which are conserved evolutionarily across prokaryotes and eukaryotes (North and Verdin, 2004). The

founding member of this family is the yeast Sir2 which is capable of extending replicative lifespan through its NAD⁺ dependent histone deacetylation activity in genomic silencing (Imai et al., 2000).

The mammalian Sirtuins comprise seven members, SIRT 1 to 7. Each of them is characterized by the highly conserved core catalytic domain and variable N-terminal and C-terminal sequences (Figure 9) (Frye, 2000). Despite the similarity of the core, they exhibit distinct cellular localization, expression pattern, enzymatic activities as well as binding partners. SIRT6 and SIRT7 are predominantly found in the nucleus while SIRT3, SIRT4 and SIRT5 localize in mitochondria. SIRT1 and SIRT2 can shuttle between nucleus and cytoplasm but SIRT1 is mostly distributed in the nucleus while SIRT2 in the cytoplasm. In the nucleus compartment, SIRT1 is mainly associated with euchromatin, SIRT6 predominantly with heterochromatin, while SIRT7 is localized inside the nucleolus (Houtkooper et al., 2012).

Most Sirtuins catalyze NAD⁺-dependent deacetylation and ADP-ribosylation (Jackson and Denu, 2002; Liszt et al., 2005). Subsequent studies revealed the deacylation mechanism by Sirtuin enzymes such as desuccinylation, demalonylation, deglutarylation and long chain fatty acid deacylation (Bheda et al., 2016; Du et al., 2011; Peng et al., 2011). Sirtuins are involved in post-translational modification of both histones and non-histone proteins. Due to the NAD⁺-dependency, Sirtuins are classified as class III histone deacetylases. The requirement of NAD⁺ for their catalytic activity renders Sirtuins to be sensors of the cellular NAD⁺ levels which are corresponding to metabolism and nutritional state of the cell (Haigis and Guarente, 2006). On the other hand, the modifications mediated by Sirtuins allow them to play an indispensable role in diverse metabolic and cellular pathways.

Figure 9. The schematic representation of the seven human SIRT family members. N-terminal is represented in red while the core catalytic domain and C-terminal in purple and pink respectively.

As the founding member of this family, SIRT1 has been extensively studied over the past decade (Blander and Guarente, 2004; Kumar and Chauhan, 2016). Initial attention was drawn to SIRT1 due to its possible role in aging and longevity (Hubbard and Sinclair, 2014). Nonetheless, SIRT6 has emerged to the forefront due to its broad functionalities ranging from DNA damage repair, glucose metabolism, telomere maintenance, aging and cancer.

Functions of the SIRT6

SIRT6 displays both NAD^+ -dependent deacetylation and ADP-ribosyltransferase activities. Recent work characterized the mechanism of deacylation mediated by SIRT6 (Jiang et al., 2013). H3K9 and H3K56, which are involved in regulating chromatin state, were reported to be the deacetylation targets of SIRT6 (Michishita et al., 2008; Michishita et al., 2009). During deacetylation, the acetyl group from the lysine side chain of the target protein is transferred to the cofactor NAD^+ , thus generating nicotinamide, 2'-O-acetyl-ADP-ribose and deacetylated protein. Deacetylation of the histone by SIRT6 is shown to be cell cycle-dependent, it promotes the compaction of chromatin and is mostly associated with gene silencing.

SIRT6 is critical in safeguarding genomic integrity by acting on the DNA damage response pathway. The mechanisms of action for SIRT6 are the modification of protein factors involved in DNA damage repair, altering chromatin accessibility to repair factors, or modulation of the reactive oxygen species from metabolic processes (Figure 10) (Mostoslavsky et al., 2006). SIRT6 is proposed to deacetylate various non-histone protein factors such as CtIP, a process which promotes DNA end resection that facilitates homologous recombination (Kaidi et al., 2010). SIRT6 also enhances double strand break (DSB) repair under oxidative stress by physically associating with PARP1 and mono-ADP-ribosylating it on K521 to activate its function (Mao et al., 2011). SIRT6 modulates DNA repair through deacetylation of nucleosomes in the vicinity of DSBs, which stabilizes DNA-PKcs association at the break (McCord et al., 2009). SIRT6 whole body knock-out mice die from severe hypoglycemia at 4 weeks age. SIRT6 liver-specific knock-out mice display elevated glycolysis, triglyceride synthesis and fatty liver formation. Conversely, the mice overexpressing SIRT6 are protected from diet-induced obesity by showing increased glucose tolerance and reduced fat accumulation.

Furthermore, SIRT6 maintains chromosome stability through acting on the chromosome ends, which are known as telomeres. Deacetylation of the histone marks on the telomeric chromatin enables SIRT6 to modulate the chromatin state to condense the telomere and lock the genome ends into a stable state. SIRT6 association with telomere is required for WRN to bind stably. Consistently, SIRT6 depletion cause end-to-end chromosomal fusions which in turn lead to telomere dysfunction and premature cellular senescence. The resulting abnormal telomeric structures resemble the telomeric defects of Werner Syndrome which is caused by mutation in the WRN factor (Multani and Chang, 2007). To summarize, SIRT6 deficiency or inactivation of SIRT6 in mice give rise to shorter lifespan and severe metabolic defects while SIRT6 overexpression results in increased life span and favorable metabolic outcomes (Kanfi et al., 2010; Mostoslavsky et al., 2006). Thus, SIRT6 links the regulation of telomeric chromatin state to mammalian life-span modulation and the abnormality in its expression to premature aging syndrome (Michishita et al., 2008).

Besides, SIRT6 plays a crucial role in glucose homeostasis through regulation of the promoter region of multiple glycolytic genes expression as well as in the direct deacetylation of factors involved in glycolytic pathway (Ghosh et al., 2010). SIRT6 modulates glycolytic gene

expression indirectly through a Hif1 α -dependent manner. It functions as the corepressor of Hif1 α to suppress the glycolytic pathway and upregulate mitochondrial respiration (Zhong et al., 2010). SIRT6 acts on the suppression of hepatic glucose production through the regulation of PGC-1 α by deacetylation reaction on the acetyltransferase GCN5 (Dominy et al., 2012). Moreover, SIRT6 plays a role in lipid metabolism through negatively regulating the PPAR γ -mediated pathway to prevent triglyceride formation (Kanfi et al., 2010).

An additional key target of SIRT6 is the NF- κ B. SIRT6 functions as a corepressor by binding to the NF- κ B subunit RelA and deacetylating H3K9 at the promoter of the target genes governing proinflammatory cytokines and anti-apoptotic genes, thus resulting in attenuation of NF- κ B signaling (Kawahara et al., 2009). Similarly, SIRT6 directly inhibits the expression of a subset of NF- κ B target genes, especially those involved in stress response and aging (Kawahara et al., 2011). Furthermore, SIRT6 promotes TNF α secretion through hydrolysis of long chain fatty acyl lysine (Jiang et al., 2013).

Figure 10. Functions of SIRT6. Picture adapted from (Kugel and Mostoslavsky, 2014).

Structure of SIRT6

SIRT6 NAD⁺ deacetylase activity was assessed by *in vitro* charcoal-binding, HPLC and continuous assays to monitor the formation of [³H]Acetate, [³H]OAADPr and nicotinamide respectively. The three different assays were in line in demonstrating SIRT6 weak deacetylation activity rate in the range from 3.2 to 4.6 $\mu\text{M/h}$ with a specificity of 0.6-0.8 h^{-1} (Pan et al., 2011). Molecular basis of SIRT6 activity was further explored from the structural perspective to explain the low deacetylase activity. X-ray structure of partial SIRT6 due to *in-situ* proteolysis (from 11-295 residues with ADPr bound) was solved. The N-terminus of SIRT6 is indispensable for chromatin association and intrinsic catalytic activity while the C-terminus is important for proper nuclear localization (Tennen et al., 2010).

SIRT6 comprises two globular domains, a large Rossmann fold for NAD⁺ binding and a smaller zinc binding domain with a total of eight α -helices and nine β -strands (Figure 11). The Rossmann fold domain is made up of six-stranded parallel β -sheet sandwiched between two and four helices at each side. The small zinc binding domain consists of three-stranded β -sheet connecting to the large domain through two extending loops β 3 and α 6. In comparison with other human Sirtuins, SIRT6 has a more open conformation where the small zinc-binding domain is splayed from the large Rossmann fold domain. This conformation may be contributed by several structural features of SIRT6 including the lack of the conserved salt bridge and formation of a new stabilizing hydrogen bond between the zinc-binding motif and the substrate binding loop. Aside from that, SIRT6 lacks the cofactor-binding loop and it is replaced by the ordered helix α 3, indicating that the substrate binding pocket of SIRT6 does not change its conformation significantly when distinct ligands are bound. These structural features were speculated to contribute to its 1000 times lower deacetylase activity and suggesting a distinct binding mechanism where no substrate was required prior to NAD⁺ binding (Pan et al., 2011).

Figure 11. Structure of SIRT6 from residue 11 to 295 (PDB code: 3PKI). A) The chains are coloured from N-terminal to C-terminal by using rainbow spectrum, from blue to red. B) The catalytic site where ADPr bound in SIRT6. The residues sitting within 4 Å from ADPr are labelled as these are the residues that interact with ADPr to facilitate its binding.

Understanding SIRT6 interactions with chromatin

SIRT6 exhibits low deacetylase activity *in vitro* towards single acetylated lysines or acetylated H3K9Ac peptide (Pan et al., 2011). Physiological relevance of SIRT6 was investigated by Gil et al., who revealed a significant deacetylase activity *in vivo*. This discrepancy was further explored and an alternative mode of binding of SIRT6 to nucleosome was postulated. SIRT6 was demonstrated to deacetylate H3K9Ac and H3K56Ac at much higher efficiency when bound to nucleosome by using semi-quantitative methods (Gil et al., 2013). Besides, SIRT6 has a higher deacylation activity on long chain fatty acyl groups than deacetylation (Jiang et al., 2013). Thus, more structural studies are required to characterize the SIRT6 interactions with nucleosome, which could be exploited for therapeutics.

Potential therapeutic properties of SIRT6

Role of SIRT6 in the maintenance of genomic DNA integrity, telomere stability, inflammation response regulation and metabolic homeostasis confers its ability to impact

tumorigenesis, metabolism and aging. Modulation of SIRT6 functions at the molecular level could lead to development of new therapeutic interventions. Current research on SIRT6 modulators has been revolving around Sirtuin activators and inhibitors depending on the pathophysiological state. Identification of Sirtuin modulators is generally achieved through chemical library screening and catalytic mechanism-based design approaches combined with the studies of structure-activity relationship (Carafa et al., 2016; Li et al., 2015).

The search for sirtuin-activating compounds (STACs) has begun more than a decade ago to promote its activities in enhancing DNA repair, prevent malignancy, treatment of metabolic diseases, and to extend male mice lifespan which might hold potential for therapeutics in humans (Sanchez-Fidalgo et al., 2012; Villalba and Alcain, 2012). Allosteric activators for SIRT1 which exploit its unique N-terminus have been studied widely, however activators specific for SIRT2-7 have been lacking. A recent study has revealed the first synthetic activator for SIRT6 deacetylation activity through substrate-independent binding to the catalytic core (You et al., 2016). This has provided a structural basis for development of SIRT6 activators as a potent therapeutic tool.

On the other hand, functions of SIRT6 in inflammation responses through promoting the release of TNF- α and its role in DNA damage response through activating PARP1 have placed SIRT6 as a druggable target. Inhibition of SIRT6 holds potential for therapeutic application. A few SIRT6 inhibitors ranging from plant extracts, small molecules and peptides have been scrutinized, but most of them are weak binders and not specific to SIRT6 (Heinke et al., 2011). Novel selective SIRT6 inhibitors derived via structure-based drug design technique could be developed into future therapeutics (Parenti et al., 2014).

Nonetheless, the pleiotropic effects of SIRT6 on distinct cellular pathways have to be further explored in order to specifically target the particular enzymatic activity that leads to the pathological state.

1.8 Linker histones (H1)

Mammalian linker histone variants

Linker histones (H1) are lysine- and proline-rich, and, like core histones, highly basic. In contrast to core histones, linker histones are more variable and much less studied. In addition to variation between species, many organisms are known to possess multiple linker histone variants. There have been 11 mammalian linker histone variants identified so far, which are sub-categorized into seven somatic and four are germ-cell specific (Figure 12). The presence of multiple variants in mammalian cells adds an additional level of complexity when studying H1 function due to their functional redundancy (Izzo et al., 2008).

In addition, 11 H1 variants can be categorized into distinct groups according to their cell-type and cell-cycle expression pattern. Accordingly, H1.1 to H1.5, are replication-dependent, while H1x is ubiquitously expressed. By contrast, the H1.0 is replication-independent and accumulates in terminally differentiated cells. On the other hand, the germ-line specific variants are H1oo (oocytes), HILS (testis), H1.t (testis), and H1.T2 (testis). The replication-dependent variants, H1.1 to H1.5, are synthesized only during the S-phase with varying amount of proteins, indicating that there is a functional heterogeneity, which is regulated by the cell-cycle stages (Happel et al., 2009). In agreement with this, H1 variants also portray distinct residency time and localization pattern *in vivo* (Th'ng et al., 2005).

H1 protein (and synonyms) ^a	H1 gene (and synonyms)	Gene locus ^b (chromosome)	Length (amino acids) ^c	Expression (cell, tissue)
H1.1 (H1a)	<i>HIST1HA (H1F1)</i>	6p21.3-22	214	Ubiquitous ^d
H1.2 (H1c)	<i>HIST1HC (H1F2)</i>	6p21.3-22	212	Ubiquitous
H1.3 (H1d)	<i>HIST1HD (H1F3)</i>	6p21.3-22	220	Ubiquitous
H1.4 (H1e)	<i>HIST1HE (H1F4)</i>	6p21.3-22	218	Ubiquitous
H1.5 (H1b)	<i>HIST1HB (H1F5)</i>	6p21.3-22	225	Ubiquitous
H1t	<i>HIST1H1T (H1FT)</i>	6p21.3-22	206	Spermatocytes
H1T2	<i>H1FNT (HANP1)</i>	12q13.11	233	Spermatids
H1oo (H1foo)	<i>H1FOO</i>	3q21.3 ^f	345	Oocytes
HILS1 (Hils1)	<i>HILS1</i>	17q21.33	230	Spermatids
H1x (H1X, H1.X)	<i>H1FX</i>	3q21.3 ^f	212	Ubiquitous
H1.0 (H1°)	<i>H1FV (H1F0)</i>	22q13.1	193	Differentiated cells ^h

Figure 12. Mammalian linker histone variants. Picture adapted from (Happel & Doenecke, 2008).

Structure and functional roles of linker histones

The structure of linker histone in solution, an avian H5 (homolog of H1.0 in human) was determined by Clore et al. in 1987, using the nuclear magnetic resonance (NMR) technique (Clore et al., 1987). In 1993, an X-ray crystal structure of the avian H5 was solved, revealing the near-atomic details of the H5 globular domain. Linker histone typically comprises a tripartite structure: a conserved globular domain of ~75 residues, a short 20–35 residues N-terminal domain (NTD), and a long, intrinsically disordered C-terminal domain (CTD) of ~100 residues. The globular domain adopts a winged helix fold, consisting of three alpha helices and a beta hairpin motif connected with loops (Figure 13) (Ramakrishnan et al., 1993).

Figure 13. X-ray structure of Avian H5 (PDB code: 1HST).

Linker histones extend the length of the core nucleosomal DNA by 20 bp by protecting it from micrococcal nuclease digestion (Bednar et al., 1998). The full structure of linker histone binding to nucleosome is sometimes referred as chromatosome in literature. The H1 variants have a differential affinity binding to DNA and chromatin, influencing the nucleosome repeat length and nucleosomal spacing *in vivo* (Oberg et al., 2012; Woodcock et al., 2006). The H1/core nucleosome ratio varies between different cell types, from 0.5 in embryonic stem (ES) cells to

0.8 in many somatic cells, and up to 1.3 in chicken erythrocytes (Bates and Thomas, 1981; Woodcock et al., 2006). The abundance of linker histones in the nucleus is about equal to that of the nucleosomes in higher eukaryotes.

The globular domain has high sequence homology among the H1 variants and some studies have shown that the globular domain is sufficient for structure-specific nucleosome recognition (Allan et al., 1980b). Nonetheless, a study based on human H1.4 (hH1.4) demonstrated that the NTD is required for determining the affinity and specificity of hH1.4 binding to chromatin (Oberg and Belikov, 2012). On the other hand, the sequence of the CTD is quite variable among H1 species and isoforms, and amino acid composition required for linker stem formation (Syed et al., 2010) and stabilizing secondary chromatin structures. The importance of CTD for high-affinity binding to chromatin *in vivo* was confirmed using FRAP and mutation studies (Hendzel et al., 2004). The sequences and length in the N-terminal and C-terminal tails vary substantially within variants. For instance, the oocyte-specific linker histone (H1oo) consisting of more than 200 residues at the CTD, is more mobile and likely contributes to a more dynamic chromatin state (Perez-Montero et al., 2016; Tanaka et al., 2003).

Chromatin compaction has been long known to be facilitated by nucleosome association of linker histones, which are essential proteins in higher eukaryotes (Kowalski and Palyga, 2016). Linker histone, which is considered a transcriptional repressor due to its role in facilitating chromatin condensation (Zlatanova et al., 2000), regulates gene transcription through stabilizing the chromatin positioning, maintaining nucleosome spacing, and altering chromatin structure (Hu et al., 2018). H1 occupancy is lower in highly-expressed gene region and vice versa. Despite that, H1 mediates transcription process at a more specific level, in complex with transcriptional activators or repressors, which is variant-specific and regulated by post-translational modification (PTM) (Daujat et al., 2005; Hergeth and Schneider, 2015; Kim et al., 2008; Kim et al., 2013; Trojer et al., 2007).

Recent study revealed that linker histones also engaged in epigenetic regulation, DNA regulation and repair (Fyodorov et al., 2018). H1 was proposed to have a role in DNA methylation after most of the genes regulated by DNA methylation were misregulated in H1 triple-knockout mice appeared to be (Fan et al., 2005). Subsequently, H1 was found to

specifically interact with DNA methyltransferase and facilitate the recruitment to those specific sites, however, the effect is H1 variant-specific (Yang et al., 2013b). The yeast H1 homolog was shown to have a role in the homologous recombination pathway of DNA repair (Downs et al., 2003). Furthermore, H1 has been linked to DNA repair in higher eukaryotes, where it altered sensitivity to DNA damaging agents, as observed in the chicken H1 variant knockout and mouse triple-knockout (Hashimoto et al., 2007; Murga et al., 2007).

Linker histones are post-translationally modified through phosphorylation, methylation, acetylation, ADP-ribosylation (mono- and poly-), ubiquitination, formylation, and citrullination on residues in both the tail regions and globular domain (Bonet-Costa et al., 2012; Lu et al., 2009; Sarg et al., 2015; Wisniewski et al., 2007; Wisniewski et al., 2008). Among all the post-translational modifications, phosphorylation of H1 has been most extensively studied. The presence of high number of serine and threonine residues in H1, especially the CTD region, allows considerable phosphorylation to take place, in a cell-cycle regulated manner (Talaszy et al., 2009). Differential phosphorylation levels observed among the H1 variants, are likely to confer the functional heterogeneity (Garcia et al., 2006; Villar-Garea and Imhof, 2008).

The H1 variant-specific modifications could be exploited as biomarkers for clinical prognosis and therapeutic development. Notwithstanding, highly specific antibodies would be required for that purpose (Izzo and Schneider, 2016). The versatility of H1 functional roles in regulating nucleosome spacing, chromatin organization, and diverse cellular functions implies more work would be needed to decipher the role of H1 epigenetics.

Structural basis of nucleosome-linker histone interaction

The nucleosome and chromatosome were first described in the 1970s (Kornberg, 1974; Olins and Olins, 1974; Simpson, 1978). Over the years, various structural models have been proposed (Allan et al., 1980a; Brown et al., 2006; Pruss et al., 1996; Zhou et al., 1998). In 2013, a NMR structure of *Drosophila* H1 chromatosome was solved, in which the globular domain binds at an off-dyad position and interacts with only one linker DNA of the nucleosome (Figure 14) (Zhou et al., 2013). In the subsequent year, an 11 Å low resolution cryo-EM structure of human H1.4 binding to nucleosome array was modelled (Song et al., 2014), where the globular

domain of H1.4 is adopting off-dyad binding mode too. Nonetheless, the observed binding orientation of H1.4 is distinct from the *Drosophila* H1. Although a full length H1.4 was utilized, the tail regions were not visible, suggesting that they are flexible.

Despite three decades of efforts, the first high resolution structure of linker histone-nucleosome complex was only solved in 2015 (Zhou et al., 2015). The structure revealed that the interaction was mainly mediated by positively-charged residues with the phosphates of the DNA backbone in nucleosome. Consistent with the earlier models (Fan and Roberts, 2006; Zhou et al., 1998), the globular domain of avian H5 binds at the on-dyad position, interacting with both linker DNAs. Bednar et al. utilized a nucleosome with longer linker DNA (197 bp) to investigate *Xenopus* H1.0 binding using cryo-EM and X-ray crystallography, where xH1.0 was depicted to interact with both linker DNAs and nucleosome DNA at the dyad axis, with the CTD localizing to only single linker DNA, conferring an asymmetric positioning (Figure 15) (Bednar et al., 2017). By far, linker histone interactions are electrostatic in nature and show little DNA sequence specificity, it is most likely that the binding of linker histone to chromatin is mediated by negatively-charged DNA backbone and positively-charged residues of H1 (Clark and Kimura, 1990). Consensus from all the solved structures is that the mechanism of action of linker histones entails holding onto the linker DNAs and altering the DNA trajectories facilitating chromatin folding.

Figure 14. Off-dyad and on-dyad binding mode of linker histone. The picture was adapted from Fyodorov et al., 2017. A) The off-dyad binding mode observed in NMR structure of *Drosophila* H1 bound to nucleosome. The on-dyad binding mode observed in the crystal structure of Avian H5 bound to nucleosome (PDB code: 4QLC).

Figure 15. Summary of nucleosome-linker histone structures in chronological order.

Implications of linker histone placement dynamics on chromatin

Linker histones bind dynamically to chromatin in living cells via transient interaction with the DNA packaged into nucleosomes (Flanagan and Brown, 2016). Dynamicity of H1 binding is measured by using *in vivo* FRAP technique, which demonstrated residence time of one or two minutes for the rapid dissociation and translocation event of H1 in both euchromatin and heterochromatin (Brown et al., 2006; Lever et al., 2000; Misteli et al., 2000). The process is ATP-independent and does not require fiber-fiber interactions. As far as time scale, it is considered to be a fast exchange event comparing to core histones, which do not exchange over the same time period; while a slow event when comparing to the exchange rate of transcription factors, which takes seconds (Voss and Hager, 2014).

The quantitative FRAP analysis combined with mutational studies of the 16 basic residues on mouse H1.0 globular domain revealed that 9 residues are critical, and molecular modeling exposed two binding sites with nucleosome, consistent with the asymmetric model (Brown et al., 2006). However, the first caveat of this study would be a lack of utilization of the

full-length H1.0, as the tail regions are required to contribute to the high affinity binding of linker histone. Secondly, an assumption is made that H1 binds at or near the nucleosome dyad, which might create bias on the linker histone positioning. Additional less intrusive approaches to study the linker histone binding in the native chromatin would be required to cross-validate these findings.

Taking into consideration that *in vivo* positioning of nucleosomes and linker histones is a dynamic process, it becomes clear that the *in vitro* structural studies are limited to portraying the static contacts between the nucleosome and linker histones. With the volume of *in vivo* data validating the notion that chromatin is highly polymorphic with loose, zig-zag 10 nm fiber (Hansen et al., 2018), an alternative view of chromatin compaction emerges, involving an intriguing speculation that linker histones are preferentially recognizing a specific DNA geometry in a highly interdigitated polymer-melt and different H1 variants might accommodate distinct forms of chromatin structure in facilitating chromatin folding (Kowalski and Palyga, 2016). This can be further facilitated through distinct types of PTM in regulating linker histone dynamicity, thus imposing additional level of regulation on chromatin functions.

Linker histone and diseases

H1 composition and expression levels have been found to be altered in several diseases and cancer types, for example a characteristic expression profile is exhibited for specific H1 variants in ovarian cancer (Medrzycki et al., 2012). These findings illustrate that the H1 expression is finely regulated. Another study reveals that the expression level and reactivity of H1.5 are positively correlated with the aggressiveness of the cancer cells, thus the level of H1.5 could be potentially used as a diagnostic biomarker to distinguish between the benign and malignant prostate tumors (Khachaturov et al., 2014). Surprisingly, the opposite expression pattern is observed for H1.0, which generally accumulates in terminally-differentiated cells. The quantity of H1.0 in human gliomas correlate inversely with the histopathological grade of the tumor and the patient's survival rates (Gabrovsky et al., 2013).

Other than expression level, recurrent mutations in linker histone genes are identified in follicular lymphoma, suggesting that altered H1 functions would results in functional

consequences and lead to diseases (Li et al., 2014; Okosun et al., 2014). Most of which are missense mutations, accumulated in the globular domain and CTD of H1, known to impact the H1 binding properties. However, the mechanism underlying the functional consequences of these mutations are still poorly understood. In addition, mutations of linker histone genes have also been detected in chronic lymphocytic leukemia, colorectal cancer, and diffuse large B-cell lymphoma, highlighting possible implications in precision medicine (Landau and Wu, 2013; Lohr et al., 2012; Sjoblom et al., 2006).

Furthermore, linker histones are depicted to play a role in cancer and embryogenesis via mediating interactions with tumor-suppressors in various context. For instance, phosphatase and tensin homolog (PTEN) interacts with linker histones and promotes their binding to chromatin in order to indirectly condense chromatin and suppress the oncogenes expression; however, the loss of PTEN in cancer cells lead to upregulation of oncogenes transcriptional level (Chen et al., 2014). In contrast, during embryogenesis linker histones interact with p53, which is mediated by chromodomain helicase DNA-binding protein CHD8 in order to inhibit p53-dependent transactivation and apoptosis (Nishiyama et al., 2009).

Linker histones emerge in association with several other diseases such as Alzheimer's disease, male infertility and autoimmune disease (Nayernia et al., 2005; Roque et al., 2015; Schett et al., 2002; Ye et al., 2017). Despite that, the connection between linker histones and these diseases require further investigation to delineate the underlying mechanism of H1. Nonetheless, these findings are appealing in bringing attention to the potential of linker histone as a novel therapeutic intervention.

1.9 Project aims

The initial project of my thesis work involved structural, biochemical and biophysical study of two nuclear factors with the aim to characterize their roles in DNA damage recognition/repair and chromatin epigenetics. These included poly(ADP-ribose) polymerase (PARP) and sirtuin-6 (SIRT6). PARP1 and PARP2 are multipartite polymerases that modulate transcription, DNA damage repair and cell-death signaling and are important cancer drug targets since they play a key role in maintaining genomic stability. These enzymes specifically recognize single- and double-stranded DNA breaks, resulting in activation of the poly[ADP-ribos(e)]ylation (PARylation) activity towards self (PARP1/2) and other protein substrates. This automodification and PAR attachment to other nuclear factors evokes a cascade cellular response that leads to recruitment of additional DNA repair factors and chromatin remodeling machinery to the site of DNA damage. PARP1 (and possibly also PARP2) is especially interesting in chromatin exploration because it is known to bind more strongly to nucleosomal substrates as compared to naked DNA. While the SIRT6 protein is involved in multiple signaling pathways, including DNA repair, telomere maintenance, glycolysis and inflammation, it has multiple modification capabilities, such as deacetylation, deacylation, and ADP-ribosylation, thus understanding of the underlying basis of its recognition to nucleosomal substrates is critical.

My ultimate objectives were: 1) determining the structure of a PARP1-nucleosome complex by X-ray crystallography or electron microscopy (EM) and 2) solving the structure of a SIRT6-nucleosome core particle complex. Given the significant challenges encountered in both of the PARP1 and SIRT6 undertakings, the project was re-directed toward obtaining crystal structures of nucleosome-linker histone assemblies (NLHs).

Chromatin's structural characterization is complicated by its heterogeneous composition in the cell and the conformational disorder of nucleosome assemblies *in vitro*. One approach would be to explore methods allowing one to elucidate crystal structures of assemblies with differing lengths of linker DNA connecting the nucleosome core regions. However, as with the nucleosome core particle, the challenge is finding constructs that yield well diffracting crystals. In order to overcome this obstacle, we developed a screening approach to test the nucleosomal constructs with varying lengths and proceeded with those that could form ordered lattice.

From the insights gained from the NLH packing in the crystal lattice, we then undertook the design of a family of di-nucleosome constructs with varying length of linker DNA aiming to obtain continuous nucleosomal packing which mimics the higher order structure of chromatin. The crystal structures of di-nucleosome with differing linker DNA lengths revealed that a dramatically distinct configuration can be achieved just by a subtle alteration of the linker DNA length. This may help to rationalize observations of chromatin structural heterogeneity in the cell. At the current phase, sustained efforts on the rational design of nucleosomal DNA are being carried out for chromatin structural studies with the goal of understanding linker histone binding to compact heterogeneous chromatin structures.

Chapter 2 Materials and methods

2.1 Materials

2.1.1 Chemicals

Chemical	Source
2-methyl-2,4-pentadiol (MPD)	Sigma-Aldrich, USA
B-mercaptoethanol (BME)	Sigma-Aldrich, USA
Acetic acid	Merck, USA
Acrylamide	Bio-Rad, USA
Acrylamide: Bis (29:1)	Bio-Rad, USA
Agarose	Affimetrix, USA
Ammonium Persulfate (APS)	Acros Organics, Belgium
Ampicillin	Sigma-Aldrich, USA
Bacto agar	BD Biosciences, USA
Bacto tryptone	BD Biosciences, USA
Benzamidine	Sigma-Aldrich, USA
Bis-acrylamide	Bio-Rad, USA
Boric acid	Merck, USA
Bromophenol blue	Sigma-Aldrich, USA
Cacodylic acid	Sigma-Aldrich, USA
Chloramphenicol	Acros Organics, Belgium
Chloroform	Fisher Scientific UK Ltd., UK
D-(+)-trehalose	Sigma-Aldrich, USA
Dipotassium hydrogen phosphate	Merck, USA
Dithiothreitol (DTT)	Gold Biotechnology, USA
DNA ladders and 6x loading dye	Promega, USA
Ethanol	Merck, USA
Ethylenediaminetetraacetic Acid (EDTA)	Bio-Rad, USA
Ethidium bromide	Bio-Rad, USA
Formic acid	Sigma-Aldrich, USA
Glucose	Alfa Aesar, USA
Glycerol	USB Corporation, USA
Glycine	1 st Base Pte. Ltd., Singapore
Guanidine hydrochloride	Sigma-Aldrich, USA
Hydrochloric acid (HCl)	Merck, USA
Imidazole	Sigma-Aldrich, USA
Instant blue	Expdeon, UK
Isoamyl alcohol	Acros Organics, Belgium
Isopropanol	Merck, USA
Isopropyl- β -D-thiogalactopyranoside	Gold Biotechnology, USA

Lithium chloride (LiCl)	Sigma-Aldrich, USA
Loading dye, Blue/Orange, 6x	Promega, USA
Magnesium chloride (MgCl₂)	Sigma-Aldrich, USA
Magnesium sulfate (MgSO₄)	Sigma-Aldrich, USA
Manganese chloride (MnCl₂)	Sigma-Aldrich, USA
Phenol	Sigma-Aldrich, USA
Phenylmethanesulfonylfluoride (PMSF)	Sigma-Aldrich, USA
Potassium dihydrogen phosphate	Merck, USA
Potassium acetate	Alfa Aesar, USA
Potassium chloride (KCl)	Merck, USA
Potassium Cacodylate (K Caco)	Sigma-Aldrich, USA
Potassium hydroxide (KOH)	Merck, USA
Polyethylene glycol (PEG) 6000	Sigma-Aldrich, USA
Protease Inhibitor Cocktail Tablet	Roche, USA
Protein ladder	Bio-Rad, USA
Riboadenosine triphosphate (rATP)	Sigma-Aldrich, USA
Sodium acetate	Sigma-Aldrich, USA
Sodium chloride (NaCl)	Merck, USA
Sodium Dodecyl Sulfate (SDS)	Hoefer Inc., USA
Sodium hydroxide (NaOH)	Merck, USA
Spermine tetrahydrochloride	Sigma-Aldrich, USA
Sucrose	USB Corporation, USA
Tetramethylethylenediamine (TEMED)	Bio-Rad, USA
Triton X-100	Bio-Rad, USA
Tris	USB Corporation, USA
Urea	USB Corporation, USA
Yeast Extract	USB Corporation, USA

2.1.2 Buffers and media composition

Buffers/ Media	Composition
Chloroform-isoamyl alcohol (CIA)	24:1 (v/v) chloroform: isoamyl alcohol
H1 Lysis Buffer	50 mM Tris-HCl pH 8.0 2% (v/v) Triton X-100 0.05% (w/v) protease inhibitor cocktail tablet
Heparin Buffer A	50 mM Tris-HCl pH 8.0 150 mM NaCl
Heparin Buffer B	50 mM Tris-HCl pH 8.0 1 M NaCl
Histone Sonication Buffer	50 mM Tris-HCl pH 8.0 500 mM NaCl 5% (v/v) glycerol 5 mM BME

	0.2% (v/v) PMSF 2 mM Benzamidine
Histone Unfolding Buffer	7 M guanidine hydrochloride 50 mM Tris-Cl pH 8.0 5% (v/v) glycerol 5 mM BME
Histone His Buffer A	50 mM Tris-Cl pH 8.0 500 mM NaCl 5% (v/v glycerol) 6 M urea 5 mM BME (for hH3 only)
Histone His Buffer B	50 mM Tris-Cl pH 8.0 500 mM NaCl 5% (v/v glycerol) 6 M urea 1 M imidazole 5 mM BME (for hH3 only)
HisTrap Buffer A	50 mM Tris-Cl pH 8.8 500 mM NaCl 20 mM imidazole
HisTrap Buffer B	50 mM Tris-Cl pH 8.0 500 mM NaCl 500 mM imidazole
IMAC Buffer A	50 mM Tris-Cl pH 8.0 500 mM NaCl 20 mM imidazole
IMAC Buffer B	50 mM Tris-Cl pH 8.0 500 mM NaCl 500 mM imidazole
Lysis solution I	50 mM glucose 10 mM EDTA 25 mM Tris-Cl pH 8.0
Lysis solution II	0.2 M NaOH 1% SDS
Lysis solution III	4 M K-Acetate 2 M acetic acid
Mono/ Resource Q Buffer A	20 mM Tris-Cl pH 7.5 0.1 mM EDTA 0.3 M NaCl
Mono/ Resource Q Buffer B	20 mM Tris-Cl pH 7.5 0.1 mM EDTA 1 M NaCl
Mono/ Resource S Buffer A	50 mM Tris-Cl pH 8.0 5 mM BME 6 M urea
Mono/Resource S Buffer B	50 mM Tris-Cl pH 8.0

	5 mM BME 6 M urea 1 M NaCl
Octamer Unfolding buffer	7 M guanidine hydrochloride 10 mM Tris-Cl pH 7.5 10 mM DTT
Octamer Refolding Buffer	2 M NaCl 10 mM Tris-Cl pH 7.5 1 mM EDTA 10 mM BME
TB	12 mg/ml Bacto Tryptone 24 mg/ml yeast extract 72 mM K ₂ HPO ₄ 17 mM KH ₂ PO ₄ 0.4% (v/v) glycerol
TBE (1X)	89 mM Tris 89 mM boric acid 2 mM EDTA
TCS-0	20 mM Tris-HCl pH 7.5 1 mM EDTA 1 mM DTT
TCS-2	20 mM Tris-HCl pH 7.5 1 mM EDTA 1 mM DTT 2 M KCL
TE (10,50)	10 mM Tris-Cl pH 8.0 50 mM EDTA
TE (10,0.1)	10 mM Tris-Cl pH 8.0 0.1 mM EDTA
TE(10,1)	10 mM Tris-Cl pH 8.0 1 mM EDTA
TGS	25 mM Tris 200 mM Glycine 0.1% SDS
TYE agar	10 mg/ml Bacto Tryptone 5 mg/ml yeast extract 8 mg/ml NaCl 15 mg/ml bacto agar
2xTY	16 mg/ml Bacto Tryptone 10 mg/ml yeast extract 5 mg/ml NaCl 0.1% (w/v) glucose

ZYP-0.8G	1 μ M MgSO ₄ 0.8% (w/v) glucose 3 mg/ml (NH ₄) ₂ SO ₄ 7 mg/ml KH ₂ PO ₄ 7 mg/ml Na ₂ HPO ₄
ZYP-5052	1 μ M MgSO ₄ 5 mg/ml glycerol 0.5 mg/ml glucose 2 mg/ml lactose 3 mg/ml (NH ₄) ₂ SO ₄ 7 mg/ml KH ₂ PO ₄ 7 mg/ml Na ₂ HPO ₄

2.1.3 Competent cells and Enzymes

Competent cell	Usage	Source
BL21 (DE3)	Protein expression	Novagen, USA
HB 101	Plasmid amplification for DNA constructs	Promega, USA
DH5α	Plasmid amplification for DNA constructs	New England Biolabs, USA

Enzyme	Source
Calf intestine alkaline phosphatase (CIAP)	New England Biolabs, USA
EcoRV	New England Biolabs, USA
HinfI	New England Biolabs, USA
HRV-3C Protease	New England Biolabs, USA
KpnI	New England Biolabs, USA
Lysozyme (from chicken egg white)	Sigma-Aldrich, USA
RNaseA	Roche, USA
SacI	New England Biolabs, USA
PstI	New England Biolabs, USA
T4 DNA ligase	New England Biolabs, USA
Thrombin	GE Healthcare, USA
TEV	Purified in lab

2.1.4 Gels Composition

Gel	Composition
6% PAGE	6% (v/v) acrylamide 0.15% (v/v) bis-acrylamide 0.25x TBE buffer 0.1% (w/v) APS 0.1% (v/v) TEMED
10% PAGE	10% (v/v) acrylamide 0.17% (v/v) bis-acrylamide 1x TBE buffer 0.1% (w/v) APS 0.1% (v/v) TEMED
SDS-PAGE 15% resolving gel	15% (v/v) acrylamide 0.25% (v/v) bis-acrylamide 375 mM Tris-Cl pH 8.8 0.1% (w/v) SDS 0.1% (w/v) APS 0.1% (v/v) TEMED
SDS-PAGE 18% resolving gel	18% (v/v) acrylamide 0.3% (v/v) bis-acrylamide 375 mM Tris-Cl pH 8.8 0.1% (w/v) SDS 0.1% (w/v) APS 0.1% (v/v) TEMED
SDS-PAGE 5% stacking gel	5% (v/v) acrylamide 0.25% (v/v) bis-acrylamide 125 mM Tris-Cl pH 6.8 0.1% (w/v) SDS 0.1% (w/v) APS 0.1% (v/v) TEMED

2.1.5 Equipment

Instrument	Source
AKTA FPLC	GE Healthcare, USA
Eppendorf Benchtop Centrifuge	Thermo Fisher Scientific, USA
G:Box machine	Syngene, UK
HiLoad Superdex 200 column	GE Healthcare, USA
HiTrap Heparin HP (5 ml) column	GE Healthcare, USA
HiTrap IMAC HP (5 ml) column	GE Healthcare, USA
HisTrap HP (5 ml) column	GE Healthcare, USA

JA-10 Fixed-Angle Aluminum Rotor	Beckman Coulter Life Sciences, USA
JA-25.5 Fixed-Angle Aluminum Rotor	Beckman Coulter Life Sciences, USA
MasterFlex L/S Peristaltic Pump	Cole Parmer, USA
Mono Q 5/50 GL column	GE Healthcare, USA
Mono S 5/50 GL column	GE Healthcare, USA
Nanodrop 200 Spectrophotometer	Thermo Fisher Scientific, USA
Pilatus3 R 300K detector	DECTRIS, Switzerland
Resource S (6 ml) column	GE Healthcare, USA
Resource Q (6 ml) column	GE Healthcare, USA
Rigaku-FRX	Rigaku, Japan
Sonics Vibra Cell Sonicator	Sonics & Materials, USA

2.1.6 DNA and Protein Constructs

All the sequence details of DNA and protein constructs is listed under the Appendix section.

Constructs	Type	Vector	Remarks
601L	DNA	pUC57	Palindromic DNA, 145 bp, 601-based sequence
CMS169s	DNA	pUC19	Palindromic DNA, 169 bp with 12 bp linker DNAs, 601-based sequence
CMS169sn	DNA	pUC19	Non-palindromic DNA, 169 bp with 12 bp linker DNAs, 601-based sequence
diNUC343	DNA	pUC19	Non-palindromic DNA, 343 bp with 29 bp linker between the di-nucleosomes and 12 bp at each linker DNA end
diNUC349	DNA	pUC19	Non-palindromic DNA, 349 bp with 29 bp linker between the di-nucleosomes and 15 bp at each linker DNA end
diNUC351	DNA	pUC19	Non-palindromic DNA, 351 bp with 29 bp linker between the di-nucleosomes and 16 bp at each linker DNA end
diNUC353	DNA	pUC19	Non-palindromic DNA, 353 bp with 29 bp linker between the di-nucleosomes and 17 bp at each linker DNA end
diNUC351z	DNA	pUC19	Non-palindromic DNA, 351 bp with 31 bp linker between the di-nucleosomes and 15 bp at each linker DNA end
diNUC353z	DNA	pUC19	Non-palindromic DNA, 353 bp with 31 bp linker between the di-nucleosomes and 16 bp at each linker DNA end
diNUC355z	DNA	pUC19	Non-palindromic DNA, 355 bp with 31 bp linker between the di-nucleosomes and 17 bp at each linker DNA end

diNUC357z	DNA	pUC19	Non-palindromic DNA, 357 bp with 31 bp linker between the di-nucleosomes and 18 bp at each linker DNA end
H2A	Protein	pET-28a	<i>Homo sapiens</i> histone H2A
H2B	Protein	pET-28a	<i>Homo sapiens</i> histone H2B
H3	Protein	pET-28a	<i>Homo sapiens</i> histone H3
H4	Protein	pET-15b	<i>Homo sapiens</i> histone H4
H1.0	Protein	pET-15b	<i>Homo sapiens</i> full length linker histone H1.0
H1.2	Protein	pET-15b	<i>Homo sapiens</i> full length linker histone H1.2
H1.3	Protein	pNIC28-	<i>Homo sapiens</i> full length linker histone H1.3
H1.5	Protein	pET-15b	<i>Homo sapiens</i> full length linker histone H1.5
PARP1	Protein	pET-28a	<i>Homo sapiens</i> full length PARP1
SIRT6	Protein	pET-28a	<i>Homo sapiens</i> full length SIRT6

2.2 Methods

Below is the overview flow chart for the methods adopted in order to achieve the project aims (Figure 16).

Figure 16. Overview flow chart of the methods and approaches that were attempted to achieve project aims.

2.2.1 Large scale DNA production

The protocol for DNA purification differed depending on the palindromicity of the DNA, however, identical over-expression and plasmid isolation steps were shared. The palindromic DNA constructs used were 601L and CMS169s, while the non-palindromic constructs were CMS169sn, diNUC343, diNUC349, diNUC351, and diNUC353 (see Appendix for construct details). The 601L construct was synthesized and subcloned into the pUC57 plasmid by EZBiolab, USA. The CMS169s was cloned into pUC19 vector as multiple repeats of 96 bp 601L-based half sites, by Dr. Bao Qiuye. All the non-palindromic constructs were synthesized in pUC19 vector by Genscript, China.

Transformation

Competent HB101 cells (Promega, USA) were transformed with plasmid comprising the gene of interest by mixing 100 ng of plasmid with 25 μ l of cells and incubated on ice for 10 min. The mixture was then heat-shocked at 42°C for 45 s and kept on ice for 2 min before adding 120 μ l of 2xTY media into the Eppendorf tube for recovery that lasted for at least 45 min at 37°C. 70 μ l of cells were then plated onto TYE agar, made with 100 μ g ml⁻¹ ampicillin and incubated at 37°C for 15 hrs.

Cell culture

Four large, well-separated colonies were used to inoculate media in 4 separate tubes each containing 5 ml of 2xTY media with 100 μ g/ml ampicillin. The tubes were incubated at 37°C, shaking at 180 rpm for 3 to 4 hrs. The turbid sub-cultures were then transferred into 100 ml TB media flask with 100 μ g/ml ampicillin, and incubated for another 1 to 2 hrs until the OD₆₀₀ reached 0.8. Then, the culture was split into 10 two-liter flasks containing 500 ml TB media with 100 μ g ml⁻¹ ampicillin, and incubated at 37°C, shaking at 170 rpm for 20 hrs.

Cell lysis and plasmid isolation

The cells were harvested by centrifugation at 7,000 rpm for 7 min at 20°C (JA 10 rotor, Beckman). The pelleted cells in each centrifugation bottle were resuspended in 50 ml alkaline lysis buffer I (refer to the top of the Material and Methods section for composition of all the buffers), followed by addition of 100 ml of alkaline lysis buffer II into each tube and shaken vigorously for 2 min to lyse the cells. The bottles were incubated on ice for a total of 20 min, with vigorous shaking in every 5 min interval. Subsequently, 175 ml of alkaline lysis buffer III (pre-chilled at 4°C) was added to each bottle, and inverted gently for 2 min to renature plasmid DNA and precipitate chromosomal DNA. The bottles were incubated on ice for a total of 20 min, with gentle mixing every 5 min.

The samples were then centrifuged at 10,000 g for 20 min at 4°C. The supernatant was filtered through four layers of sterile gauze, followed by the addition of 0.52 volumes of isopropanol. The mixture was stirred and incubated at room temperature for 30 min to precipitate the plasmid DNA. Then, centrifugation was carried out at 16,000 g for 30 min at 20°C. The supernatant was discarded and the pellet was air-dried before being dissolved in a total of 40 ml TE(10/50). A total of 240 µl RNase A was added to the dissolved pellet and the digestion was allowed to proceed overnight at 37°C.

Insoluble matter was removed by centrifugation at 15,000 g for 15 min at 20°C. The supernatant was transferred to two clean 50 ml Teflon tubes, and contaminant protein extraction was carried out by adding 0.5x of buffered phenol. The mixture was centrifuged at 27,000 g for 20 min at 20°C. The aqueous phase was removed and the phenol extraction steps were repeated 1 – 2 times until the interphase was clear. The aqueous phase was then extracted with 10 ml CIA (chloroform: isoamyl alcohol, 24:1) followed by centrifugation at 12,000 g for 10 min at 20°C. To the aqueous phase, 0.2 volumes of 4 M NaCl and 0.4 volumes of 40% (w/v) PEG-6000 was added to precipitate the plasmid DNA. After 30 min incubation on ice, the sample was centrifuged at 27,000 g for 10 min at 4°C.

Subsequently, the supernatant was discarded and the pellet was dissolved in 15 ml TE(10/0.1). CIA extraction was performed twice more as described above to remove the PEG residue. Final ethanol precipitation was done by addition of 0.1 volumes of 3 M sodium acetate

and 2.5 volumes of 100% ethanol, followed by incubation on ice for 30 min and centrifugation at 15,000 g for 10 min at 4°C. The pellet containing plasmid DNA was dissolved in TE (10/0.1) and concentration measurements were done on Nanodrop 200 spectrophotometer.

2.2.1.1 Palindromic DNA treatment

Purified plasmid containing inserts for palindromic DNA constructs were subjected to a series of enzymatic treatments (Figure 17). 601L DNA was digested with EcoRV enzyme, while CMS169s DNA was digested using PstI enzyme (Table 1). The completion of the digestion was verified by 10% polyacrylamide gel electrophoresis (PAGE) in 1x TBE buffer. Upon completion, the sample was then subjected to the first round of purification by using Resource Q column to remove the plasmid backbone. The purified half sites were then pooled and dephosphorylated.

Figure 17. The enzymatic treatment steps of the plasmid to produce the 601L DNA construct. For illustration purpose, only two repeats are shown here.

Differences in steps	DNA 601L	DNA 169s
Digestion of half sites out from plasmid	EcoRV	PstI
Purification of half sites from empty vector	PEG precipitation	Resource Q chromatography

Table 1. The differences in the processing steps for preparation of DNA 601L and DNA 169s.

CIAP dephosphorylation

Dephosphorylation of the 5' ends of the half-sites was required to prevent the re-ligation of the repeats into the original conformation that was present in the plasmid. Dephosphorylation was conducted by using Calf-Intestine Alkaline Phosphatase (CIAP) in a final reaction mixture of 2 mg ml⁻¹ DNA, 50 mM Tris-HCl (pH 8.0), 10 mM MgCl₂, 0.1 mM EDTA, and 38.4 U CIAP per mg of DNA, incubated at 37°C for 24 hrs. In order to assess the completeness of dephosphorylation, a ligation test was performed, which included a positive control containing the non-dephosphorylated DNA sample. The reactions were analyzed using 10% PAGE.

CIAP was removed from the reaction mixture by adding 0.25 volumes of phenol and CIA, followed by centrifugation at 22,000 *g* for 20 min at 20°C, and these steps were repeated until the interphase was clear. Dephosphorylated half-sites were then ethanol-precipitated from the aqueous phase, pelleted by centrifugation, and dissolved in TE (10, 0.1) to the desired concentration.

HinfI digestion

HinfI digestion was used to create 3 bases overhang ligation sites for forming palindromic constructs. The digestion was conducted in a final reaction mixture of 2 mg ml⁻¹ DNA, 50 mM Tris-HCl (pH 7.5), 50 mM NaCl, 10 mM MgCl₂, 1 mM DTT, and 276 U HinfI per mg of DNA. The reaction was incubated at 37°C for 24 hrs, and the completeness of digestion was assessed with 10% PAGE.

The desired DNA fragments with 3 bases overhang were purified from the digested 10 bp oligonucleotides and from the undigested fragments using a Mono Q 5/50 GL column on an AKTA FPLC system. The sample was injected into the column, which was pre-equilibrated with the Mono Q Buffer A. A gradual salt gradient up to 100% was applied using the Buffer B (Buffer A with 1 M NaCl) to separate the DNA fragments with distinct ionic binding properties. All the peak fractions were examined on 10% PAGE and those which contained the pure DNA fragments were pooled, ethanol-precipitated, and dissolved in TE (10, 0.1).

Ligation

Two half-sites with compatible overhang sites were ligated with T4 DNA ligase to give the final palindromic DNA construct. Ligation was conducted in a reaction mixture of 2 mg ml⁻¹ DNA, 100 mM Tris-HCl pH 7.5, 20 mM MgCl₂, 10 mM DTT, 2 mM rATP, and 500 U T4 DNA ligase per mg of DNA. The reaction was incubated at 18°C for 24 hrs, and its completeness was assessed using 10% native PAGE. The ligated palindromic DNA constructs were again purified by using the Mono Q 5/50 GL column, as described above. Fractions containing the desired DNA fragments were pooled, ethanol-precipitated, dissolved in TE (10/0.1), lastly stored at -20°C.

2.2.1.2 Non-Palindromic DNA production

Subsequent to plasmid isolation, non-palindromic DNA sequences were subjected to their respective enzyme treatments. Upon cleavage, the full construct was obtained without the need for subsequent dephosphorylation, digestion or ligation. CMS169sn plasmid DNA was double digested using 200U/mg of KpnI and SacI, while other DNA constructs such as 343, 349, 351 and 353, as well as the z series, were cleaved with KpnI enzyme only.

The digestion progress was assessed by a 10% PAGE. The complete digested sample was subjected to Resource Q column purification and the fractions containing pure fragments were

pooled for storage at -20°C. Fractions with contamination of upper bands were kept aside for re-run through the Resource Q column.

2.2.2 Human core histones production

The *Homo sapiens* core histones were kindly provided by Prof Hitoshi Kurumizaka. Histones hH2a, hH2b and hH3 were subcloned into pET-28a vector while hH4 was subcloned into pET-15b with hexa-histidine tag upstream of the gene of interest by Dr. Deepti Sharma in our laboratory. The expression of the recombinant histone proteins was carried out by using auto-induction media as described by F. William Studier (Studier, 2005).

Transformation

Competent cells strain BL21 (DE3) were transformed with plasmids containing genes of histone proteins hH2A, hH2B or hH3, while JM109 strain was transformed with hH4 plasmid. 100 ng of plasmid of interest was added into the Eppendorf tube with 50 µl of respective cells, and then the mixture was incubated on ice for 10 mins. The cells were heat-shocked at 42°C for 45 seconds and returned on ice for another 2 mins. Then, 150 µl of cold 2xTY media was added into the sample and incubated at 37°C for 1 hour. Subsequently, 70 µl of the transformed cell mixture was plated onto LB agar plates containing 50 µg/ml of kanamycin for hH2A, hH2B and hH3 plasmids or 100 µg/ml of ampicillin was used for hH4 plasmid containing cells.

Insoluble histone expression and inclusion body preparation

A 100 ml ZYP-0.8G no induction media with respective antibiotics was used as a starter culture. 7 ml of this media was added to the agar plate and all the colonies were well suspended before slowly adding the bacterial suspension into the starter culture until the OD reaches 0.1. Next, the starter culture was incubated in the shaker at 170 rpm for around 3 hrs. When the OD of the starter culture reached 0.8, the culture was split evenly into 4 of 2 L flasks with 500 ml of ZYP-5052 auto-induction media. The large cultures were then incubated for at least 22 hrs in 37°C shaker at 170 rpm.

The bacterial cells were harvested by centrifugation at 7000 rpm (JA 10 rotor, Beckman) for 10 mins at 20°C. The cells were then resuspended in 100 ml of pre-chilled sonication buffer. Sonication step was carried out at 35% amplitude for 20s on and 30s off for 4 mins on ice, this step was repeated for at least 3 times with stirring at each interval until the sample become homogenous. After the sample was lysed, centrifugation at 20000 rpm (JA 25.5 rotor, Beckman) for 15 mins at 4°C was carried out to remove the contaminant proteins in the soluble fraction.

The supernatant was discarded and the pellet was resuspended in another 100 ml of sonication buffer. The sonication cycle was then repeated with same amplitude and duration. Second centrifugation was carried out at 20000 rpm (JA 25.5 rotor, Beckman) for 15 mins at 4°C. The pellet collected this time was dissolved in unfolding buffer (7 M guanidinium chloride), followed by sonication and constant stirring at room temperature for an hour. The final centrifugation was done at 50000 g (JA 25.5 rotor, Beckman) for 20 mins at 4°C. The supernatant was collected carefully to avoid any debris that would clog the purification column.

Histone protein purification

The collected supernatant was loaded into the pre-equilibrated immobilized metal ion affinity chromatography (IMAC) column charged with nickel ions at 2 ml/min. The loaded column was then washed with IMAC buffer A with 2 M NaCl to remove the bulk DNA contamination. The histone proteins bound were eluted by using five buffer A with step gradients of 5 mM, 25 mM, 50 mM, 250 mM and 500 mM imidazole.

The eluted proteins were fractionated into 10 ml each and the fractions were assessed by 15% sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE, refer to section 2.1.4 for composition details). The pure fractions were then pooled and dialyzed against milli-Q water at 4°C with three changes overnight. The dialyzed sample was then subjected to thrombin (1.5 U/mg) digestion at room temperature for 3 h with stirring to cleave off the Histidine tag. Another SDS-PAGE gel was utilized to assess the digestion progress after two hours, using undigested aliquots as a negative control.

The cleaved histone sample was then loaded onto Resource S column (for hH2A and hH2B) or Mono S column (for hH3 and hH4). The flow rate was set at 1 ml/min. A NaCl gradient from 100 mM to 600 mM over 100 ml was used to elute the histone proteins. The fractions eluted were evaluated by 15% SDS-PAGE and the pure fractions were pooled and dialysed against dH₂O at 4°C overnight. The concentration and the absorbance value (A_{260/280}) of the final protein sample were measured by using the Nanodrop 2000 Spectrophotometer. The sample was aliquoted and lyophilized before storage at -20°C.

2.2.3 PARP1 and SIRT6 protein expression and purification

Transformation

Competent BL21 (DE3) plysS (Promega, USA) cells were transformed with plasmid pET-28a containing SIRT6, while BL21 (DE3) RIPL (Agilent technologies, USA) cells were transformed with plasmid pET28a containing PARP1. One hour recovery was allowed before plating the cells on agar plate with kanamycin antibiotics.

Soluble protein expression

The transformed cells were grown in 3 L- 6 L of 2xTY media containing the final concentration of 0.1% (w/v) glucose and 50 µg/ml of kanamycin. Prior to that, the colonies on the plates were scraped and dissolved uniformly in 7 ml media and 300-400 µl were added to each 2 L-flask containing 500 ml media. 100 µM of ZnSO₄ was added into each flask when the OD of the cell culture reached 0.3. The flasks were incubated at 37°C and shaken at 180 rpm until the OD reached 0.6, and then induced with 0.5 mM of IPTG. SIRT6 protein expression was allowed to take place for 3-4 hours. For PARP1, the protein expression after adding the IPTG was carried out at 16°C overnight.

The bacterial cells were harvested by centrifuging at 7,000 rpm for 7 min at 4°C. The cell pellet was re-suspended in 50 ml of lysis buffer comprising of 50 mM Tris-HCl pH7.5, 500 mM NaCl, 5% glycerol and 5 mM βME with protease inhibitor cocktail tablet (Roche, USA), 200 µl

of saturated PMSF as well as 200 μ l of 1M benzamidine added. The dissolved cell suspension was frozen using liquid nitrogen and could be stored at -80°C.

Protein Purification

The frozen cells were thawed at room temperature and the cell sample was sonicated on ice at 35% amplitude, with 20 sec pulse and 30 sec rest for 10 min, until the sample became homogenous. The sample was then centrifuged at 18,000 rpm for 15 min at 4°C and the supernatant was collected. The supernatant was filtered using 0.2 μ m filter (Sartorius, Germany) before injecting it into the superloop connected to the AKTA purifier.

First round of purification was done using immobilized metal ion affinity chromatography (IMAC) column charged with nickel ions. After loading the sample, the column was washed with 50 ml of binding buffer (IMAC buffer A) containing 2 M salt and additional 20 ml of binding buffer before starting sample elution. Elution was carried out with a gradient of elution buffer (binding buffer with 0.5 M imidazole). The flow through and fractions eluted were assessed with SDS-PAGE.

The fractions with pure protein were pooled and dialyzed against buffer containing 25 mM Tris pH 7.5, 100 mM NaCl, 5% glycerol and 2 mM β ME at 4°C. After 2 changes of dialysis buffer, overnight thrombin (1.5 U/mg) digestion was set up while the third dialysis ongoing at 4°C. SDS-PAGE was run to ensure the digestion was complete. The digested protein sample was passed through 5ml heparin column.

The heparin binding buffer used was 20 mM Tris pH 7.5, 5% glycerol, 5 mM β ME and PMSF. The protein bound was eluted using the elution buffer (binding buffer with 1M NaCl). A SDS-PAGE was used to assess the purity of the protein sample. The pure fractions of the protein were pooled. Then, the concentration and the absorbance value (A_{260/280}) of the protein sample were measured by using the Nanodrop 2000 Spectrophotometer.

For SIRT6 protein, two steps of purification were sufficient to obtain a pure pool of the protein. Thus, the pure protein sample was then concentrated to around 5-10 mg/ml and centrifuged to remove any precipitate. The protein was then aliquoted and stored at -20°C.

For PARP1 protein, another size-exclusion chromatography purification step was included to remove the remaining protein contaminants. The sample pooled from heparin step was concentrated down to less than 5 ml and injected into the Hiloal Superdex 200 column. The protein fractions eluted were analyzed by SDS-PAGE, and the pure fractions containing PARP1 were pooled and concentrated down to 5-10 mg/ml before aliquoting and storing at -20°C.

2.2.4 Linker histones production

The plasmid consisting *Homo sapiens* full length linker histone genes were obtained as follows, H1.0 gene was cloned into pET-15b vector by EZBiolab Inc., USA, H1.3 gene was cloned into pNIC28-Bsa4 vector by Protein Production Platform (PPP), NTU, Singapore, while H1.2 and H1.5 genes were cloned into pET-15b vector by Genscript, China.

Transformation

Competent BL21 (DE3) (Promega, USA) cells (50 µl) were transformed with 100 ng of plasmid containing linker histone, the mixture was incubated on ice for 10 mins. The cells were heat-shocked at 42°C for 45 seconds and returned on ice for another 2 mins. Then, 150 µl of cold 2xTY media was added into the sample and incubated at 37°C for 1 hour recovery. Subsequently, around 70 µl of the bacteria cell mixture was spread plated onto LB agar plates containing 100 µg/ml ampicillin (for pET-15b plasmid) or 50 µg/ml of kanamycin (for pNIC-Bsa4 plasmid).

Cell culture and lysis

The colonies on the agar plate were scraped off by adding 7 ml of 2xTY media. From this media, 500 µl was added into the 100 ml 2xTY starter culture with 0.1% (w/v) glucose and respective antibiotics. The starter culture was grown at 37°C for 2 to 3 hrs until the turbidity of the media visible. Then, the starter culture was split evenly into 10 flasks of 500 ml 2xTY media with 0.1% (w/v) glucose and respective antibiotics. The cells were grown at 37°C for 2 to 3

hours and the OD was monitored regularly. IPTG induction at 18°C was carried out overnight (16-18 hrs) once the OD reached around 1.0.

The induced bacteria cells were pelleted down at 9000 rpm for 7 mins at 4°C and re-suspended in 100 ml of linker histone lysis buffer (50 mM Tris-Cl pH 8.0, 2% (v/v) Triton X-100, protease inhibitor cocktail tablet (Roche, USA), 2 mM Benzamidine and 2 mM PMSF) with 4 mg/ml of lysozyme. The sample was left to stir for 1 hour at 4°C and subsequently was centrifuged at 20,000 rpm for 20 mins at 4°C. The supernatant which contained DNA and soluble contaminant proteins was discarded while the pellet was re-dissolved in linker histone lysis buffer with 1 M NaCl. The sample was sonicated for 3 mins at 35% amplitude with the pulse of 30 sec on and 30 sec off, repeated for 1-2 times, then left to stir overnight at 4°C.

Linker histones purification

The suspension was spun at 20,000 rpm for 20 mins at 4°C and the supernatant was collected and subjected to affinity purification by using 5 ml His Trap High Performance column. After loading the sample, the column was washed with 50 ml of binding buffer (50 mM Tris-HCl pH 8.0, 500 mM NaCl, and 20 mM imidazole) containing 2 M salt to reduce the DNA contamination followed by additional 20 ml of binding buffer before starting sample elution.

The linker histone proteins were eluted by using the Buffer B (50 mM Tris-HCl pH 8.0, 500 mM NaCl, 500 mM imidazole) gradient from 0% to 50% over 100 ml. The peak fractions were assessed with 15% SDS-PAGE gel and the fractions containing the proteins were pooled and loaded onto Heparin column for second round of affinity purification. Heparin Buffer A (50 mM Tris pH8.0) and Buffer B (50 mM Tris pH8.0, 1 M NaCl) were used to elute the protein with baseline set at 50% and increased gradually to 100% over 50 ml. The eluted fractions were analyzed by 15% SDS-PAGE.

Relatively pure fractions were pooled and enzyme digestion was set up overnight at 4°C to cleave the hexa-histidine tag at the N-terminal. TEV enzyme was used for all the linker histone proteins at a molar ratio of 1:75, except for the H1.0 protein which was instead treated with HRV 3C protease (Novagen, USA) at 10 units per mg. Small aliquots of proteins were left undigested as controls for assessing digestion completion.

After digestion, the sample was loaded on HisTrap HP column, this round the flow through and the wash fractions were analyzed by SDS-PAGE to confirm whether they contained the protein of interest which had no His-tag. Those fractions were pooled and subjected to the last purification step using Heparin column. A salt gradient from 50% to 100% in 100 ml was used to better separate the protein contaminants. Highly pure fractions were concentrated using Amicon spin concentrator (Millipore, USA) with 3 kD cut-off and then buffer exchanged to a buffer with 20 mM potassium cacodylate pH 6.0. The linker histone protein was aliquoted at 300 µg/ tube and stored at -80°C.

2.2.5 Human histone octamer refolding

The lyophilized histones were dissolved individually by using unfolding buffer to give a concentration of 4 to 5 mg/ml. The unfolding process was assisted by thorough mixing and incubating for 30 mins. The four histones were mixed in the ratio of H2A/H2B: H3/H4 = 1:1.1 by taking into consideration of the extinction coefficient and molarity into the calculation. Dialysis reaction was set up overnight against refolding buffer with two changes in 3 hrs intervals at 4°C.

The refolded octamer was centrifuged twice for 10 mins at 18000 g at 4°C to remove any precipitate. The supernatant was loaded into pre-equilibrated HiLoad 16/60 Superdex 200. The peak fractions were collected and assessed on 18% SDS-PAGE. The fractions containing the pure histone octamer were pooled and concentrated down to around 8 mg/ml using an Amicon concentrator device (Millipore, USA) with a molecular weight cut-off of 50 kDa and subsequently stored in 50% glycerol at -20°C.

2.2.6 Nucleosome reconstitution

The protocol for nucleosome core particle (NCP), nucleosome (NUC) and di-nucleosome (diNUC) in vitro reconstitution was modified from Luger et al., 1997a (Luger et al., 1999). The nucleosome was reconstituted by using reducing salt gradient stepwise in order to position the DNA on the histone octamer gradually. The reconstitution reaction with the final concentration

of 2 M KCl, 10 mM DTT, 6 μ M DNA and histone octamer was setup by mixing water, KCL, DTT, histone octamer and DNA in a sequential manner. The mixture was dialyzed against TCS-0.85, 0.65, 0.45, 0.25 and 0 at 2 ½ h for each interval before changing to the next buffer at room temperature. Small scale of 100 μ l reactions were set up with titrating octamer: DNA molar ratios of 0.9, 0.95, 1.0, 1.05, and 1.1 to optimize the yield and quality before proceeding to large scale.

For diNUC small scale reconstitution, the titrating octamer: DNA ratios screened were 1.6, 1.7, 1.8, 1.9, and 2.0. The DNA concentration was set at 2 μ M as higher concentration would lead to heavy precipitation. The large reconstitution was scaled up to 1 ml and similar gradient dialysis process was carried out. The reconstituted NCP or nucleosome was harvested and spun down at 14500 rpm for 10 mins at 4°C to remove any precipitate. The quality of the nucleosome and di-nucleosome were then assessed by 6% native PAGE before storing the sample at 4°C. For crystallization purposes, the nucleosome and di-nucleosome sample were buffer exchanged to 20 mM potassium cacodylate pH 6.0 and concentrated down to 7 mg/ml.

2.2.7 Electrophoretic Mobility Shift Assay (EMSA)

EMSA was utilized for determining the molar ratio of protein bound to nucleosome via the visible shift of the nucleosome band on native 6% PAGE gel (Refer to section 2.1.4 for composition details). The reaction mixture of the protein: NCP molar ratios were set up in a range and the nucleosome concentration was fixed at 1 μ M. After incubation on ice for 15 mins, the respective reaction mixture in 5 % sucrose was loaded onto a 6% native gel and the electrophoresis was conducted for 90 mins at 100 V at 4°C. The gel was then stained by using 1 μ g/ml ethidium bromide solution and visualized with gel documentation machine (Syngene, UK). The best ratio of protein: NCP was determined through evaluation of the gel band with the most completely shifted nucleosome bands with the least precipitant in the well.

2.2.8 Bio-Layer Interferometry (BLI)

Bio-layer interferometry (BLI) is a label-free, real-time technology to study the biomolecular interactions by using simple Dip and Read method. It analyzes the interference pattern of the reflection white light from two different surfaces which are the internal reference layer and the biosensor tip in real-time. A ligand would be immobilized on the biosensor tip and the binding of analyte would result in the increase of the optical thickness on the tip which in turn causes a wavelength shift. The wavelength shift would be calculated and converted to the kinetic and affinity information of the binding interaction.

The streptavidin biosensor was used to capture the biotinylated nucleosome sample (NCP 601L) which was the ligand. The purified protein, SIRT6 was utilized as the analyte. After the test binding and buffer optimization of the interaction were established, quantitative tests were carried out by simultaneously dipping of the loaded biosensors into the wells containing different serial dilutions of the analyte. The raw data was imported into the Octet Data Analysis Software and fitted with the 1:1 Langmuir model to calculate the kinetic and affinity values.

2.2.9 Electron Microscopy (EM)

Negative stain and cryo techniques were used for direct visualization of PARP1-NUC and SIRT6-NCP complexes under the electron microscope. The EM studies were done in collaboration with Prof. Shashi Bhusan's and Prof. Sara Sandin's groups.

Negative-stained EM

Negative stain technique involves staining of the sample by using 2% uranyl acetate. Carbon-coated copper grids (homemade by NTU Cryo EM lab and from Nanoserv, Singapore) were glow discharged for 1 min. 4 µl of sample were pipetted onto the grid, incubated for 1 min and blotted with Whatman paper. 4 µl of stain solution of 2% uranyl acetate was placed on the grid, incubated for 1 min and blotted. The grid was air-dried for 10 minutes. Electron microscope

Icor T12 microscope was operated at 120 kV. Negative stained grids were imaged at calibrated magnification of 49,000X with -1.0 to -2.0 μm defocus applied.

Cryo EM

The grids used for cryo-EM were Quantifoil R2/2 (200 mesh) copper grids with the carbon coated only at the background (not in the holes). Cryo technique requires plunging of the sample loaded-grid into liquid ethane and the whole process has to be in liquid nitrogen. Similar glow discharge process was conducted and the grid was placed in the Vitrobot machine (Thermo Fisher Scientific, USA) for the sample loading, blotting, plunging into liquid ethane steps. The frozen grid was clipped and inserted into the cartridge for loading into cryo-EM Tecnai Arctica operated at 200 kV.

2.2.10 Trace Fluorescence Labelling

Trace fluorescence labeling for protein crystallization was utilized to aid in facilitating the process of protein crystallization (Pusey et al., 2015). Protein factor was labeled with fluorescence probe in less than 1% of the final protein concentration for setting up of the crystallization. The screening process was then done under fluorescent illumination to identify the lead conditions of which the trace labeled proteins were incorporated into the crystal lattice.

The protein sample was buffer exchanged to 25 mM HEPES pH 7.5 to completely remove the Tris buffer which contains tertiary amine and could interfere with the labeling process. The protein in HEPES buffer was concentrated down to 5-10 mg/ml. The total amount of protein was calculated and a 1/5 molar equivalent quantity of carboxyrhodamine-succinimidyl ester (CR-SE, Sigma Aldrich) was added. The mixture was mixed well and the reaction was allowed to proceed in the dark at room temperature for an hour. To obtain 0.5% derivatization level,

No. of moles of protein $\times$ 0.005% $\times$ 5 (25% yield) = moles of dye

10 ml of ethanolamine was added to quench the reaction and the sample was passed through gel filtration column to separate the unbound dye from the labeled protein. The labeling

percentage and the concentration of the protein were determined by measuring the absorbance readings at 280 nm and 532 nm.

$$1. [\text{Protein}] = (A_{280\text{total}} - A_{280\text{fluor}}) / \epsilon_{\text{protein}}$$

Given $\epsilon_{\text{protein}} = 38,350$ and $A_{532\text{fluor}} / A_{280\text{fluor}} = \text{Correction factor} = 2.5$

$$2. \text{Percent of labeling} = (A_{532} / \epsilon_{\text{fluor}}) / [\text{Protein}]$$

Given $\epsilon_{\text{fluor}} = 92,000$

2.2.11 Screening and optimization of the co-crystallization of the protein-nucleosome complex

The best ratio of protein-NCP complex formed as observed from the EMSA was used to set up the 96 wells crystallization screening across different buffer conditions from both the commercial and homemade kits. Sitting drop approach was adopted for the screening, which utilizes the 96 wells Intelli-plates (Art Robbins Instruments, Canada) and the Mosquito Crystal machine (TTP Labtech, UK). The setup crystallization plate was incubated at 18°C and monitored daily to detect crystal formation across all the different conditions.

The conditions, which yielded crystals were used for crystallization setups with a bigger drop size. Hanging drop vapor diffusion setup and salting out approach was first employed to determine if the similar crystal could be reproduced. The sitting drop setup with both salting in and salting out methods were tested as well to optimize the crystal size and morphology.

In the next round of optimization, the concentration of the respective buffer components were varied. When the sizes of the crystals were large enough, some crystals were harvested and analyzed by 12% SDS-PAGE to verify the presence of protein and NCP. Some of the crystals were then harvested and treated with post-crystallization soaking buffer, directly before mounting them in the cryo-stream of the X-ray machine (Rigaku Microfocus Micromax 007 X-ray generator with R-Axis IV++ image plate detector) to assess their diffraction quality.

2.2.12 Linker Histone-Nucleosome Assembly Crystallization

Crystallizations were set up in hanging drop vapor diffusion method with the sample concentration of 4 mg/ml in 4-8 μ l drop volume with 500 μ l reservoir buffer at 18°C. Salting in approach was adopted by setting 2x crystallization buffer (20 mM sodium acetate pH 4.5, 50 mM potassium chloride, and 88-102 mM calcium chloride) in the drop and equilibrated against the reservoir of 1x crystallization buffer. Linker histones were mixed with nucleosome at ratio 1.2:1, which was determined from the EMSA results.

Crystals were harvested in the stabilization buffer (10 mM sodium acetate pH 4.5, 12.5 mM potassium chloride, and 20 mM calcium chloride). The crystals were then gradually buffer exchanged in a step-wise manner by using the stabilization buffer with up to 2% trehalose and 65% 2-methyl-2, 4-pentadiol (MPD) for post-crystallization dehydration effect to enhance the resolution.

2.2.13 Heavy metals crystallization soaking for SAD

Heavy metal soaking was tested at the post-crystallization stage in order to attempt to improve electron density maps associated with disordered linker histone chains for lower resolution dataset by maximizing the signal to noise. Several heavy metals were selected to provide good coverage on various classes which suit our protein as well as the crystallization condition. Solubility tests were done beforehand to ensure that the pH and salt composition in the stabilization buffer would be well-suited for the soaking experiment. The working concentration and strategy for the crystallization were detailed in the table below.

Heavy atoms	Working concentration	Strategy
1. Strontium chloride	20 mM	Replace CaCl ₂ during the post-crystallization soaking step
2. Lead acetate	1 mM	Add into the crystals after post-crystallization soaking step for overnight
3. Mercury acetate	1 mM	Add into the crystals after post-crystallization soaking step for 6 hours
4. Ethyl mercury phosphate	1 mM	Add into the crystals after post-crystallization soaking step for 6 hours
5. Tantalum cluster	1 mM	Add into the crystals after post-crystallization soaking step for 6 hours

Table 2. The concentration and strategies of heavy metals soaking for nucleosome-linker histone assembly crystals.

2.2.14 X-ray Diffraction Data Collection

Preliminary X-ray diffraction test was carried out at in-house collection facility set up by NTU Institute of Structural Biology (NISB), the Rigaku-FRX machine with PILATUS3R 300K detector manufactured by DECTRIS. Long exposure data collection was carried out to compensate for the lower energy of the in-house X-ray beam, 360 images were collected with 0.5° oscillation and 30 sec exposure.

Collection of high resolution X-ray diffraction data was undertaken at beam line X06DA of the Swiss Light Source (SLS) (Paul Scherrer Institute, Villigen, Switzerland) with the beamline wavelengths range from 0.71-2.25 Å and the PILATUS 2M-F detector. X-ray wavelength was set at 1 Å for linker histone-nucleosome assembly crystals and full data set consisting 1800 images were collected with 0.2° oscillation and 0.3 sec exposures.

2.2.15 Data Processing and Model Building

For home source collected data sets, data processing was done by using the HKL3000 software at the NISB facility. At synchrotron, data processing was done with go.com which was an automated processing pipeline developed at SLS. iMOSFLM or XDS (Battye et al., 2011;

Kabsch, 2010) were also used to integrate the X-ray data sets for output statistics comparison purposes. Integrated output files were further phased (PHASER or MOLREP) and refined (REFMAC5) by using CCP4 package version 7.0.024 (Collaborative Computational Project, 1994; Winn et al., 2011). Model building was carried out with WinCoot 0.8.7.1 (Emsley and Cowtan, 2004) and figures were made with PyMOL or CCP4MG (McNicholas et al., 2011).

Chapter 3 Results and Discussion Part 1: Preparation of DNA, Nucleosomes, and Chromatin proteins

3.1 DNA production

3.1.1 Palindromic cohesive ends DNA constructs

Preparation of palindromic DNA constructs required cloning of the tandem repeat half sites into the plasmid due to recombination activities in bacteria cells. With this approach, the two half sites are ultimately ligated to generate the full construct.

DNA 601L

The palindromic 601L DNA sequence is based on the left half of the high affinity, strong positioning Widom 601 sequence (Lowary and Widom, 1998b), and it yields nucleosome with exceptionally high salt stability (Chua et al., 2012). The plasmid containing 8 repeats of 601L was extracted via several alkaline lysis steps (as described above), then subjected to EcoRV digestion to release the half-sites from the vector (Figure 18A) and differential PEG precipitation was used to separate the excised half-sites from the empty vector (Figure 18B). The half-sites DNA was de-phosphorylated using CIAP (Figure 18C) to prevent re-ligation, after the HinfI digestion (Figure 18D). The ligated DNA would give a full 145 bp, which had to be purified through anionic exchange chromatography to get the final product (Figure 18E).

Figure 18. 601L DNA preparation. A) Before and after the EcoRV digestion. Red arrowhead is indicating the half sites. B) Before and after PEG fractionation to remove the vector from the half sites. C) The test ligation confirming the completion of CIAP dephosphorylation step. D) Hinfi digestion check of the half sites. The red arrowhead here is referring to the digested product while blue arrow refers to half sites before digestion. The excised small fragment was separated from the 71bp+3 half sites through anionic exchange purification as showed on chromatogram. E) Ligation product was shown as indicated by red arrowhead. F) The pure ligated product was obtained from purification through the MonoQ column.

DNA CMS169s

The palindromic DNA CMS 169s was designed by utilizing the 601L sequence as the core of the half-site A and B. In contrast to the DNA 601L, half-sites of the DNA CMS 169s had to be purified via anionic exchange chromatography instead of the differential PEG precipitation due to the presence of the cohesive ends. The use of cohesive DNA ends was intended to promote base-pairing interactions between nucleosomes in the lattice, which would promote improved order in the crystals.

The production of CMS 169s required three rounds of purification which followed after PstI digestion, Hinfi digestion and ligation. The PstI step did not proceed to complete digestion, which further aggravated purification in a Resource Q purification step (Figure 19A & B). The restriction sites were then switched to KpnI or SacI for the subsequent construct designs.

Multiple rounds of Resource Q purification were performed to ensure pure half-sites were obtained before proceeding to dephosphorylation and *HinfI* digestion. A second round of anionic exchange chromatography, the Mono Q column was utilized to remove the small digested DNA fragment as shown in Figure 19E. The final product of full-length CMS 169s was produced after ligation and the last Mono Q purification step (Figure 19G).

Figure 19. DNA CMS 169s production. Red arrowhead points to the desired product. M designates DNA marker. A) *PstI* digestion to release the half sites from plasmid vector. B) Resource Q purification to separate the pure half sites from the incomplete digested products and vector backbone as seen on chromatogram at the right. C) Test ligation to confirm the completion of dephosphorylation by CIAP. D) *HinfI* digestion. Blue arrowhead refers to the before digestion, while red arrowhead to the after digestion-band. E) The 10bp+3 was removed by using Mono Q purification as observed on chromatogram (right) and PAGE gel (left). F) Ligation assessment. The ligated product migrates close to 200 bp band of DNA ladder. G) The final Mono Q step to separate the ligated from the unligated product.

3.1.2 Non-palindromic cohesive end DNA constructs.

Preparation and purification of non-palindromic DNA constructs were simpler as the full construct repeats were cloned in the plasmid. Production of the non-palindromic DNA fragment only required one-step digestion to release the full repeat and one purification step to remove the empty vector and non-fully digested product, which substantially reduced the preparation time.

CMS169sn

The CMS 169sn (non-palindromic) construct was designed after facing persistent challenges in the PstI digestion step of DNA CMS 169s production. The sequence of the CMS 169sn construct was also based on that of the Widom 601 (Lowary and Widom, 1998b), but in this case made non-palindromic in order to reduce purification steps and preparation time. Both the 169s and 169sn consist of 169 bp + 4 nucleotide 3'overhangs (Figure 20). The 169sn construct however was designed to have two distinct restriction sites for differential DNA end base-pairing.

Figure 20. Design of CMS 169s and CMS169sn.

The extracted plasmid containing 4 repeats of DNA CMS 169sn was double digested using KpnI and SacI enzymes before purifying it through anionic exchange chromatography to yield the final DNA product (Figure 21).

Figure 21. Production of DNA CMS169sn. M designates DNA marker. A) KpnI and SacI double digestion-assessment. Red arrowhead points to the digested product. B) Products of Resource Q purification, which removes the vector backbone as shown by PAGE gel and on chromatogram.

Di-Nucleosomal DNA constructs

All DNA constructs for di-nucleosomes (diNUCs) were designed based on the Widom 601 (Lowary and Widom, 1998b) DNA sequence as the core element to achieve near maximal nucleosome positioning strength and histone octamer affinity. A series of diNUC DNA construct lengths ranging from 343, 345, 347, 349, 351, 353, 357, 359, to 361 bp were constructed by adding one bp sequentially to each of the terminal linker DNA arms.

On the other hand, the z series of the diNUC constructs (351z, 353z, 355z, and 357z) were designed by adding 2 bp to the continuous (shared) linker DNA region of the 349, 351, 353, and 355 fragments. KpnI enzyme restriction sites were adopted for both the linker DNA arms, and the sequences were made non-palindromic to facilitate testing purpose.

All the diNUC DNA constructs were designed in-lab and submitted to the Genscript (China) for synthesis. Four repeats of the construct were inserted into pUC19 plasmid and expressed using HB 101 cells. The extracted plasmid was subjected to overnight KpnI digestion and then purified using one round of anionic exchange chromatography with Resource Q column (Figure 22 & 23). The final products migrated in between the 300 bp and 400 bp bands of DNA marker.

Figure 22. DNA 343 production. M designates DNA marker. The digested sample was subjected to Resource Q purification to separate the pure product (Peak 1) from vector backbone (Peak 2).

Figure 23. DNA 349 production. M designates DNA marker. The digested sample was subjected to Resource Q purification to separate the pure product (Peak 1) and smaller fragments (initial elution of Peak 1).

3.2 Proteins production

3.2.1 Core histones preparation

Human core histones, H2A, H2B, H3, and H4, were each individually expressed and purified using the same protocol as detailed in Materials and Methods. Expression of histone proteins into the inclusion bodies required unfolding step in guanidinium hydrochloride to solubilize the histone proteins, while purification steps were done in the 6 M urea condition.

Hexa-histidine tag at the N-termini of the protein construct was utilized to facilitate the first round of affinity-based purification by using the nickel-charged chromatography column (Figure 24A). Elution was carried out through gradual increments of the imidazole concentration to outcompete the column binding by histidine tag. In order to remove the hexa-histidine tag, refolding of histone proteins was done before thrombin digestion (Figure 24B). A final round of purification using the cationic exchange chromatography allowed the removal of the digested His-tag as well as the rest of the protein contaminants (Figure 24C).

Figure 24. Human core histone H2B preparation. The red arrowhead points to the protein of interest. M designates protein marker. A) Ni-NTA affinity purification for His-tagged hH2B. B) The thrombin digestion assessment. Lane 2 and 3 refers to the before and after digestion respectively. C) The SDS-PAGE and chromatogram from the cationic exchange purification for hH2B.

3.2.2 PARP1 production

His-tagged PARP1 protein construct was expressed using BL21 RIPL cells in the soluble fraction of the cell lysate which was then applied on the nickel-charged affinity chromatography

to remove the bulk protein and DNA contaminants. The UV measurement started to spike at 12.5 mM imidazole concentration, nonetheless, PARP1 protein only eluted at the later part of the peak as seen on SDS PAGE gel (Figure 25A).

A second affinity chromatography step using a heparin column (Figure 25B) was carried out to further purify the PARP1 protein sample pooled from the first elution. With the high DNA binding affinity of the PARP1 protein, it was eluted only at the second UV peak which was at 400 mM to 700 mM NaCl buffer condition.

Due to the degradation issue of PARP1 protein, a last step of purification was required by utilizing the size-exclusion chromatography. The His-tagged full-length PARP1 was eluted at around 60 ml and the degraded protein would be eluted closely after as seen on SDS-PAGE gel (Figure 25C). The final yield of PARP1 protein was ~2 mg out of the 5 L cell culture and was stored at -20°C.

Figure 25. PARP1 purification steps. The red arrowheads point to the full length PARP1 protein with His-tag and the red bracket indicates the fractions pooled and proceeded to the next step. The chromatogram of the purification profiles were portrayed alongside on the right with the SDS-PAGE gels of each purification step. M–Protein marker. A) PARP1 Ni-NTA purification. B) PARP1 heparin purification. C) PARP1 gel filtration purification step.

3.2.3 SIRT6 production

SIRT6 protein was also expressed in the soluble cell lysate. The purification of SIRT6 protein was a two-step chromatography process which involved nickel-charged IMAC column

and heparin column. For the first affinity purification step, SIRT6 protein was eluted from 15 mM to 165 mM imidazole buffer concentration, with the second peak containing more pure SIRT6 protein (Figure 26A).

His-tag was cleaved off from SIRT6 protein by using thrombin enzyme before proceeding to the heparin chromatography (Figure 26B). SIRT6 protein was eluted from 300 mM to 500 mM NaCl buffer concentration (Figure 26C), the purity was sufficient and the yield of SIRT6 protein was 12 mg out of a 5 L cell culture.

Figure 26. SIRT6 purification steps. The red arrowheads point to the SIRT6 protein and the red brackets indicate the fractions pooled and subjected to next step. M–Protein marker. A) SIRT6 Ni-NTA purification. B) The thrombin digestion of SIRT6. The lane 1 and 2 refer to the protein before and after the digestion. The red arrow points to the digested product. C) The heparin purification of SIRT6.

3.2.4 Linker histones production

There is total of 11 variants within human linker histone family, with H1.x, H1.0, H1.1, H1.2, H1.3, H1.4, and H1.5 found in the somatic cells while H1.oo, H1S1, H1.T, and H1.T2 are expressed in the germ cells.

All the full-length human linker histone (LH) proteins were attempted to be expressed and purified. The 7 somatic LHs were successfully produced in joint effort with Dr. Zenita Adhireksan. The expression of the germ cell-specific linker histone variants in the *E.coli* bacteria system were so far unsuccessful, with the H1T2 and HILS1 proteins ultimately identified as the bacterial ribosomal protein after mass spectrometric analysis.

The purification steps were customized according to the purity of the specific linker histone variant after each step of chromatography. Since all the linker histone variants were designed with the Hexa-Histidine tag, the first purification step involved HiTrap IMAC HP column charged with nickel ions to remove the bulk contaminants (Figure 27A). A high salt buffer wash step was incorporated before the elution to remove the DNA contaminants that bound strongly to the linker histone proteins.

The eluted fractions containing the linker histone protein were pooled and subjected to a second round of heparin affinity-based chromatography if a too high level of contaminants was present in the sample (Figure 27B). Linker histone proteins were eluted at the latter part of the peak with a higher salt concentration. Subsequently, overnight digestion was carried out to cleave off the hexa-histidine tag before proceeding to the next chromatography step (Figure 27C). Purification steps, which followed were determined based on the purity of the eluted fractions containing linker histone proteins as detailed on the table below (Table 3).

LH variants	MW (kDa)	Purification steps	Final purity (A260/280)	Final Yield (5 L)
H1.0	20.86	His-trap, Heparin, HRV 3C protease digestion, Heparin	0.56	13.7 mg
H1.2	22.58	His-trap, Heparin, TEV digestion, His-trap+Heparin	0.65	7.3 mg
H1.3	22.35	His-trap, TEV digestion, His-trap+Heparin	0.64	18.3 mg
H1.5	22.58	His-trap, TEV digestion, His-trap+Heparin	0.66	12.6 mg

Table 3. The purification steps and final yield as well as the purity for linker histone variants H1.0, H1.2, H1.3, and H1.5.

Figure 27. H1.0 purification. The red arrowheads point to the H1.0 protein and the red brackets indicate the fractions pooled and subjected to next step. M—Protein marker. A) His-trap purification of H1.0, with the chromatogram profile at the right. B) SDS-PAGE gel (left) and chromatogram (right) of H1.0 heparin chromatography. C) His-tag digestion assessment. Red arrowhead points to the digested product. D) Final heparin purification to obtain the pure H1.0 protein.

3.3 Human Histone Octamer (HO) refolding

In order to obtain nucleosome substrate *in vitro*, the four core histones, H2A, H2B, H3, and H4 have to be refolded to form the octamer complex before undergoing reconstitution step to assemble with the DNA fragment.

All four human core histones, which were purified separately were unfolded and mixed together in stoichiometry ratio of dimer: tetramer at 1.0: 1.1. The refolding process was carried out through dialysis against the refolding buffer with 2 M NaCl for 20 hours. The refolded histone octamer was then passed through the gel filtration column to remove the excess histone dimer. The peak fractions, which eluted around 60 ml containing the octamer complex, were analyzed by 15% SDS-PAGE to confirm the presence of the four histone proteins. Those fractions eluted at the tail of the peak were uncollected to avoid tetramer contamination. Higher complex or the aggregated materials were eluted at 40-50 ml while the H2A-H2B dimer peak eluted at 80-90 ml (Figure 28).

From the SDS-PAGE assessment, H4 with a molecular weight of 11.6 kD was the fastest running band while the H3 protein of 15.7 kD was the slowest, which corresponded to the highest band. The two middle bands represented the H2A (14.4 kD) and H2B (14.2 kD) proteins. All the core histone proteins migrated slightly slower than the corresponding sizes of the protein marker, with the H2A migrating closely to H3.

Figure 28. The size exclusion purification of human histone octamer. The eluted peak fractions were assessed by 18% SDS-PAGE.

3.4 Nucleosome core particle (NCP), Nucleosomes (NUCs) and Di-Nucleosomes (diNUCs)

Reconstitution

DNA and histone octamer were mixed in high salt (2 M KCl) and 1 mM DTT. Gradient dialysis was performed to allow positioning and wrapping of the DNA fragment around the histone octamer under gradual decreasing salt concentration in the buffer.

Ranges of DNA to histone octamer molar ratio were screened in small scale reconstitutions of 100 μ l or 200 μ l for every batch of new histone octamer to determine the optimal ratio for large scale reconstitution. The DNA concentration was fixed at 6 μ M for NCP and NUC while 2 μ M for diNUC reconstitution. The harvested NCP, NUCs and diNUCs were analyzed by a native PAGE to assess the quality (Figure 29, 31 & 33). The DNA: HO ratio with minimal free DNA and higher complexes was chosen to proceed with large scale reconstitution.

Figure 29. Reconstitution of NCP and NUCs. A) NCP 601L. B) NUC 169s. C) NUC 169sn. D) Biotinylated NUC 169s.

Reconstitution of diNUCs consistently displayed some higher as well as lower bands when assessed by native PAGE (Figure 30 & 32). The band migrating faster beneath the main diNUC species are apparently histone-depleted species as they disappear at higher HO ratios.

Besides being analyzed by native PAGE, the quality of diNUC reconstitution was also reflected through crystallization setups. The diNUC sample with optimal DNA: HO ratio would favor the formation of crystals, which required continuous trial and error in order to understand

the crystallization profile (morphology, size and buffer conditions) of that specific diNUC construct. The optimal DNA: HO ratios for diNUC 343 to 353 were 1:1.7 or 1:1.75 as those diNUC samples with slight lower bands appeared to increase the probabilities of crystal formation with a more robust morphology and size.

Figure 30. 6% native PAGE assessment of small scale diNUCs 343, 345, 349, 351, and 353 reconstitution to screen for the optimal DNA: HO ratio.

Figure 31. Final optimized DNA: HO ratio for large scale diNUCs 343, 345, 347, 349, 351, and 353 reconstitutions that were harvested for X-ray crystallization set ups.

Figure 32. Assessment of small scale diNUCs reconstitution for 361, 351z, 353z, and 355z constructs to screen for the optimal DNA: HO ratio by 6% native PAGE.

Figure 33. Final optimized DNA: HO ratio for large scale diNUCs reconstitution of 361, 351z, 353z, 355z, and 357z constructs that were utilized for X-ray crystallization set ups.

Chapter 4 Results and Discussion Part 2: Structural Recognition of Nucleosome by PARP1

4.1 Binding of PARP1 to nucleosome visualized by EMSA

The binding ratio of PARP1 to NCP 601L was determined using EMSA (Figure 34). As more protein was added into the nucleosome sample, the 1:1 complex band became more intense, until a certain point where another 2:1 PARP1-NCP complex band formed. In order to strike a compromise between the formation of higher complexes and completely saturated NCP, the ratio of 1.2: 1 (PARP1: NCP) was selected for further studies. Thus, PARP1 was expected to bind to NCP at 1:1 ratio but slightly higher ratio was required, which appears to be due to the simultaneous binding of two PARP1 molecules to NCP. In fact, higher order complex could be observed when there was excess PARP1 available, however this coincided with some precipitation visible in the well of the gel. Thus, the excess PARP1 might be linking NCPs together and forming aggregate.

PARP1 binding to different nucleosome substrates, NCP 601* (147 bp) and NUC CMS-1m (154 bp), was also studied using EMSA, before proceeding to electron microscopy (Figure 35). The DNA 601* and CMS-1m were a gift from colleague, Dr. Zenita Adhireksan.

Figure 34. Increasing amount of PARP1 protein was titrated to the same amount of NCP 601L sample leading to increasing intense 1:1 and 2:1 complex bands as illustrated by the 6% native PAGE.

Figure 35. The EMSA of PARP1 binding to NCP 601* and NUC CMS-1m.

4.2 Electron microscopy (EM) analysis of PARP1 protein assembled with nucleosome

4.2.1 Negative Staining Electron Microscopy

Negative staining is useful for structural studies only at low resolution due to the grain size of the stain which limits the information that can be obtained. The heavy metal atom used was uranyl acetate to generate contrast for viewing when the grid was imaged at a tilt angle. Negative staining was used to access the quality of the PARP1-NCP sample before employing cryo-electron microscopy (Ohi et al., 2004).

The concentration and the dilution of the assembly had to be optimized to achieve evenly distributed and dispersed complex particles on the grid. The binding ratio of PARP1-NCP sample was varied as well to achieve the maximum occupancy of the protein on the NCP. Furthermore, negatively stained grids were used to examine the variations of buffer salt concentration, homemade versus commercial grids and sample treatment before loading onto the grids.

NCP alone was used as a control before adding the PARP1 protein. The side view and the top view of the nucleosome could be easily distinguished from the electron microscopy image

(Figure 36). Despite optimization of the PARP1: NCP binding stoichiometry (1.2:1) based on EMSA, the ratio had to be varied in EM to optimize the occupancy of the protein on the NCP. The concentration was altered accordingly when the amount of protein added was increased for particle distribution optimization.

The binding ratio of 1.2:1 of PARP1: NCP did not show a good view of complex sample on the negative stained grid. The NCP as visualized on the control grid disappeared and only some much smaller irregular sized proteins were observed (Figure 37A). So, stoichiometric ratio of 2:1 (PARP1: NCP) was also tested and the concentration was optimized to around 85 ng/ μ l (Figure 37B). Enhanced particle distribution were visible on the grids and more particles bear resemblance of PARP1-NCP complexes.

Figure 36. Electron microscopic negative stained image of NCP alone, as a control.

Figure 37. Negative stained sample of PARP1-NCP 601L complex with T12 EM microscope. A) PARP1 and NCP 601L were added in a 1.2:1 ratio with the final PARP1-NCP 601L complex concentration of 30 ng/μl. B) PARP1 and NCP 601L were added in 2:1 ratio with the final PARP1-NCP 601L complex concentration of 85 ng/μl.

Presence of salt in the buffer had an impact on the particles density as well as the background noise level (Figure 38). The viewing of the same concentration of PARP1-NCP sample under the microscope revealed that the presence of 200 mM sodium chloride in the buffer led to denser distribution of the particles on the grid. When the sample was further diluted, the background noise level on the grid became more apparent and led to positive staining. Some speculations can be made based on these observations. The presence of salt might act as a charge-charge screening agent between the particles and the carbon layer which led to more particles sticking to the surface during blotting. Another possibility could be that the presence of salt is screening the charges on the nucleosome core particles which allowed them to come closer and resulted in a more crowded environment.

Figure 38. Effect of NaCl on the density of the PARP1-NCP 601L sample on the negative stained grid. The binding buffer contains 10 mM Tris pH7.5 and 200-250 mM NaCl. A) The buffer used for diluting (1:5) PARP1-NCP 601L complex consisted 10 mM Tris pH7.5 and 200 mM NaCl. B) PARP1-NCP 601L complex was diluted (1:5) in 10 mM Tris.

In addition, the source of the grids and the treatment of the sample before loading onto the grids were manipulated to check if those parameters could improve the image quality. From the images captured, the commercial grids and homemade grids did not make any significant difference to the sample quality (Figure 39). However, the carbon coating on the grid itself, which had to be even and complete could impact greatly the image quality, due to the efficiency of sample's adherence to the grid. Hence, duplicates were needed to control for inconsistencies in the carbon coating layer.

Some aggregates were observed on the images for PARP1-NCP 601L complex. This could have been caused by the properties of the PARP1 itself which can bind to more than one NCP at a time. The inactive PARP1 has been shown to bind to nucleosomal arrays and to promote the formation of compact chromatin structures (Kim et al., 2004). Thus, the sample containing PARP1-nucleosome complex was centrifuged at 14,000 rpm for 10 min before loading onto the grid. Uncentrifuged sample was used as a control. As shown in Figure 40, the centrifugation prior to loading onto the grid did not remove the aggregates (highly positive stained cluster).

Figure 39. The same sample of PARP1 and NCP 601L (85 ng/ μ l) were put onto two carbon-coated 200 mesh grids from different sources, A) commercial versus B) homemade.

Figure 40. Effect of centrifuging the sample before loading onto the grid. A) The PARP1-NCP 601L sample was centrifuged at 14,000 rpm for 10 mins before loading onto the grids versus B) the sample was added onto the grid without any spinning.

Consequently, longer DNA constructs were employed to examine if higher affinity interaction would result in a more homogenous and stabilized complex. Nucleosome CMS-1m (154 bp) with a 9 bp poly-A sequence linker DNA arm on only one side of the nucleosome was designed as the substrate for PARP1 binding in effort to influence its positioning (Figure 41). Since PARP1 functions as a sensor for the DNA-strand break through physical binding using the zinc finger and WGR domains (Ikejima et al., 1990), this asymmetrical design was supposed to favor PARP1 grabbing of the 9 bp linker DNA arm, instead of the other DNA end that is directly associated with core histone.

Moreover, NCP 601* (147 bp) which was widely used for protein-nucleosome structural characterization was also exploited in PARP1 EM studies (Figure 42). As compared to the PARP1-NUC CMS-1m complex, a greater ratio of nucleosome-only particles were observed on the negative-stained grid with PARP1-NCP 601* complex sample. Thus, PARP1-NUC CMS-1m complex was selected over PARP1-NCP 601* for cryo-electron microscopy studies.

Figure 41. PARP1-NUC CMS-1m complex (1.2:1). Complex concentration: 0.3 μ M.

Figure 42. PARP1-NCP 601* complex (1.2:1). Complex concentration: 0.3 μ M.

4.2.2 Cryogenic Electron Microscopy (Cryo-EM)

Single-particle cryo-EM has been widely used for macromolecular structural biology to address structure-function relationships. The stability and homogeneity of the macromolecular assemblies are important for structural characterization. On the other hand, the structural heterogeneity could be deciphered via capturing all relevant structural states of the macromolecular assemblies through 3D classifications, in which a large dataset would be required for the analysis as well as sufficient prior knowledge to make sense of the biological implications of each structural state (Lyumkis, 2019).

The concentration of the complex has to be optimized to obtain well-distributed and polydispersed particles on the grids in order to avoid oversaturation in the field of view. The concentration of the protein-nucleosome assembly had to be increased comparing to the negative staining, where particles normally have a higher affinity to adhere to the thin carbon support film, as opposed to the empty holes as in the cryo-grids. The use of filter paper for sample blotting generally resulted in the loss of the majority of the sample. A range of concentrations was tested for preparation of vitrification grids, from 0.3 μM to 1.65 μM of PARP1-NUC CMS-1m complex molarity. The cryo-grid with 0.3 μM complex concentration seemed like giving a suitable distribution of particles on grids, but the contrast was too weak to proceed (Figure 41 & 42).

Several screenings were performed but none yielded a good field of view on the vitrified grid. There were a few challenges that we faced with the cryo-EM technique which led to the inconsistency of the outcomes, for instance, the ice thickness variation, particles aggregation, weak contrast, inconsistency of the vitrified grid quality, sample sticking to the carbon surface outside of the holes, as well as nucleosome particles visualized without the protein factor bound (Figure 43). Some grid squares were crushed and not in good condition (uneven and wrinkled grid surface) which can be due to manufacturing defects or grid preparation flaws. There was ice thickness variation across different holes of the same grid. The weak contrast of particles and aggregation of the sample also leads to difficult particle boxing during data processing.

Figure 43. Challenges in structural characterization of PARP1-NUC CMS-1m complex (1.2:1) using cryo EM technique. Complex concentration: 1.65 μM . A) Crushed grid squares. B) Ice thickness variation across different holes of the same grid. C) Weak contrast of particles. D) Aggregation of the sample. E) Particle aggregation on grid holes. F) Uneven and wrinkled grid surface.

A study has showed that PARP1 binding to nucleosomal linker DNA ends resulted in a compact structure where automodification can occur, and at the meantime, causes reversible unfolding of the nucleosome DNA around the DNA entry and exit sites (Sultanov et al., 2017). Automodified PARP1 is then removed from the nucleosome, resulting in recovery of the native nucleosomal structure (Figure 44). However, the nature and extent of the conformational changes in the nucleosome remains to be investigated. Besides, poly(ADP-ribosylation) of polynucleosomes has been shown to lead to chromatin relaxation (Poirier et al., 1982a), but the underlying mechanism is unclear. Importantly, an unfolded nucleosomal state would coincide with a heterogeneous nucleosome population, which would appear to be unsuitable for high resolution structural characterization.

Figure 44. Model of PARP1-induced conformational change in nucleosomal DNA after binding to the linker DNA (detected as DNA double-stranded break) as proposed by (Sultanov et al., 2017).

Chapter 5 Results and Discussion Part 3: Structural Characterization of SIRT6 binding to nucleosomes

5.1 Binding of SIRT6 to nucleosome core particle by EMSA

The stoichiometric ratio of SIRT6 binding to NCP was estimated by using EMSA (Figure 45). From the native PAGE gel, the complete shift of the NCP band was seen to occur at a stoichiometry of 3:1 of SIRT6: NCP. Beyond that, further addition of SIRT6 did not make any significant difference, the complex band seemed to have reached equilibrium and the band intensity remained constant.

Figure 45. EMSA binding of SIRT6 to 601L NCP. Increasing concentrations of SIRT6 were added to a constant NCP amount, and an apparent 2: 1 (SIRT6: NCP) complex band was observed as indicated by red arrowhead (see BLI results below).

5.2 Bio-Layer Interferometry (BLI)

Apart from that, biochemical characterization of SIRT6-NCP assembly was carried out with the help from application scientist from Pall ForteBio by using the BLI instrument. This technique is using a dip and read method by measuring the change in light waves interference.

The detector would read the reflected light from two different interfaces, one is the reference interface and the other is the surface with substrate bound. When ligands are interacting at the surface and the phase of the light wave will change due to the change in optical density. The shift can be measured and converted into relative intensity to quantitate the binding kinetics and affinity (Figure 46). There are a few key benefits for this technique which are label-free detection, real-time results, as well as it is very simple and fast (Sultana and Lee, 2015).

Figure 46. Principles of Bio-Layer Interferometry. Picture modified and adapted from: http://www.addscro.com/en/sub/sub02_10.php.

In our experimental design, NCP 601L was used as the substrate for immobilization on the biosensor. For this purpose, the NCP 601L was biotinylated and streptavidin biosensor has been chosen for the experimental setup. Before proceeding to the kinetic measurements, qualitative tests had to be performed to confirm the interaction, while at the same time the extent of potential non-specific binding of SIRT6 protein to the blank biosensor had to be determined. As a result, the amplitude of the signal generated when Sirt6 binds to NCP 601L loaded biosensor or blank streptavidin sensor was compared.

From the analysis, SIRT6 indeed was binding to NCP 601L as concentration-dependent binding could be observed. However, the amplitude of the non-specific binding signal was around 40% of the specific binding signal (Figure 47). Thus, binding buffer had to be optimized to reduce the non-specific binding of SIRT6 for accurate quantitative data measurement.

Figure 47. Qualitative test of NCP601L-SIRT6 Interaction. A blank and an NCP 601L-loaded streptavidin biosensor were dipped into wells containing 5 nM (left) and 500 nM (right) of SIRT6 sample respectively.

The blank biosensor and the biosensor loaded with NCP 601L up to 0.4 nm signal were dipped into the wells with 5 μ M SIRT6 protein to determine the extent of non-specific binding. The initial binding buffer used was 20 mM Tris pH 7.5 with 100 mM NaCl, and the non-specific binding signal spiked to 60% when comparing the signal measured between NCP 601L loaded biosensor vs blank biosensor. After a few rounds of optimization, buffer PBS (10 mM Phosphate pH 7.4, 2.7 mM KCl, 137 mM NaCl) with the addition of 0.05% Tween-20 and 0.1% BSA was shown to reduce the non-specific binding to the biosensor from 60% to 20% (Figure 48).

Figure 48. Buffer optimization of NCP601L-SIRT6 binding assay. Binding response curves of blank and NCP 601L loaded streptavidin biosensors dipped into 5 μ M of SIRT6 in two different buffers: 20mM Tris pH 7.5, 100 mM NaCl (left) and 10 mM Phosphate pH 7.4, 2.7 mM KCl, 137 mM NaCl, 0.05% Tween-20, 0.1% BSA (right).

A quantitative binding kinetics assay was performed where five NCP601L loaded biosensors were dipped simultaneously in five wells containing two-fold serial dilution of SIRT6 protein between 250nM and 15.6nM (Figure 49). Good dose-response curves were observed with the SIRT6 protein titration. The BLI data corroborated the EMSA results, in which SIRT6 indeed displayed specific binding to NCP 601L.

Figure 49. Quantitative binding kinetics assay of Sirt6- NCP601L interaction. Five NCP 601L loaded biosensors were dipped simultaneously in five wells containing two fold serial dilution of Sirt6 protein between 250nM and 15.6nM.

The raw data was imported into the Octet Data Analysis Software and fitted with two binding models. For the 1:1 Langmuir model, the affinity of the interaction was calculated to be 6.1 nM, with average fitting for χ^2 (0.0199) and R^2 (0.9235) values (Figure 50). For the 2:1 heterogeneous ligand model, the data showed a better fit ($\chi^2= 0.0034$ and $R^2= 0.9872$) and the measured K_D values of SIRT6-NCP 601L interaction were 52.6 nM and 112.0 nM (Figure 51). From the χ^2 and R^2 values, as well as by using EMSA data as a reference, experimental results suggested that 2 molecules of SIRT6 bound to 1 molecule of NCP 601L.

The binding affinity value portrays the strength of the binding interaction between a single biomolecule to its ligand/binding partner. The binding affinity could be strong (low K_D value), but still translate to a transient interaction, in which the k_{on} and k_{off} would have to be taken into consideration. From the k_{on} and k_{off} values obtained for SIRT6-NCP interaction, the on and off rate at the binding site B is faster than the site A. The K_D values indicate that the SIRT6 binds with higher affinity at the site A, with almost half the K_D value at binding site B.

$K_D(M)$	K_D Error	$k_{on}(M^{-1}s^{-1})$	k_{on} Error	$k_{off}(s^{-1})$	k_{off} Error	χ^2	R^2
6.122E-09	5.047E-10	1.574E+06	1.265E+05	9.637E-03	1.763E-04	0.0199	0.9235

Figure 50. The raw data was fitted with 1:1 Langmuir model. The K_D of SIRT6-NCP 601L interaction was calculated to be 6.122 nM, with average fitting for χ^2 (0.0199) and R^2 (0.9235) values.

Binding Site A

$K_D(M)$	K_D Error	$k_{on}(M^{-1}s^{-1})$	k_{on} Error	$k_{off}(s^{-1})$	k_{off} Error	χ^2	R^2
5.259E-08	3.512E-09	7.737E+04	4.838E+03	4.069E-03	9.538E-05	0.0034	0.9872

Binding Site B

$K_D(M)$	K_D Error	$k_{on}(M^{-1}s^{-1})$	k_{on} Error	$k_{off}(s^{-1})$	k_{off} Error	χ^2	R^2
1.120E-07	2.294E-08	4.209E+06	8.164E+05	4.713E-01	3.118E-02	0.0034	0.9872

Figure 51. The raw data was fitted with 2:1 heterogeneous ligand model. The dissociation constant (K_D) values of SIRT6-NCP 601L interaction were calculated for binding sites A and B to be 52.6 nM and 112.0 nM respectively, with good fitting indicated: $\chi^2 = 0.0034$ and $R^2 = 0.9872$.

5.3 Electron microscopy (EM) of SIRT6-NCP 601L assembly

5.3.1 Negative stain EM

For direct visualization of a SIRT6-NCP complex under the electron microscope, a binding ratio of 4:1 of SIRT6: NCP was used in order to increase the occupancy level. No salt was present in the buffer for the sample, only 10 mM Tris pH 7.5 was included. An initial 1 μM (360 ng/ μl) complex concentration was used for grid making but the sample appeared to be too densely distributed (Figure 52A). After, the concentration of the complex was diluted to 0.65 μM (230 ng/ μl , Figure 52B), the distribution of the particles improved. However, positive staining was observed on the sample grid. Continuous optimization has led to settling on concentration of 0.5 μM for a suitable field of view on negatively stained grids, which was then followed-up with cryo-EM (Figure 53).

Figure 52. The negative stained SIRT6-NCP 601L on the carbon-coated grid. A) The complex concentration of 1 μM . B) 0.65 μM .

Figure 53. The negative stained SIRT6-NCP 601L complex on the carbon-coated grid at concentration of 0.5 μM .

5.3.2 Cryo-EM

Vitrified grids were prepared with several concentrations (0.5, 1.0, 1.5, 2.0 μM) of SIRT6-NCP 601L assembly sample. The grid with 1.5 μM concentration gave well-distributed particles and was selected for automated data collection at 53,000X magnification. The defocus was set at -1, -1.5, -2, -2.5 and the dose rate at 15 $\text{e}/\text{\AA}^2\text{s}$. Manual particle picking was done using Eman2 software and a total of 9500 particles were picked from 457 micrographs. Constant transfer function (CTF) correction, particle extraction, and subset selection were carried out via RELION-2.0. The final number of micrographs and particles left after pre-processing were 363 and 8,500, respectively, which were then subjected to 2D classification. However, most of the classes were NCP-only particles with no protein factor bound (Figure 54).

Figure 54. The SIRT6-NCP 601L assembly with a concentration of 1.5 μM . Data collection was conducted at 53,000 X magnification with defocus set at -1, -1.5, -2, -2.5. The 2D classes shown on the right.

Sustained attempts of re-producing the vitrified grids with similar sample quality were unsuccessful, no sufficient particles could be visualized on the grid holes surface. SIRT6-NCP 601L assembly sample displayed preferential sticking to the carbon matrix of the holey grids instead of to the vitreous ice in the holes. Therefore, graphene oxide coated grids were employed, as a recent report had indicated that biological macromolecules are able to better adhere to graphene oxide sheets with several advantages in single particle cryo-EM data collection (Palovcak et al., 2018).

A thin layer of graphene oxide sheet was applied on the holey carbon grids before sample was loaded, in order to increase the hydrophilicity of the grid surface for even distribution of macromolecules, as well as to keep the sample away from the air-water interface which could damage the complexes or induce specific orientations during the plunge freezing step. Nonetheless, no assembly particle was visualized on the grid holes treated with graphene oxide. In contrast, some particles were noticed outside of the holes where the carbon support film was positioned (Figure 55). This direct drop-casting method, which was utilized to apply the graphene oxide (Pantelic et al., 2010) might not be the best approach to coat the surface of holey carbon grids in mono-layer, as demonstrated by Palovcak et al., 2018. Regions with multi-sheet

aggregates or those completely lacking graphene oxide were frequently observed (Palovcak et al., 2018; Pantelic et al., 2011). Further studies would be required to optimize this methodology.

Meanwhile, different types of holey carbon grids were explored instead of the R2/2 type which was previously used. The R1.2/1.3 holey carbon grid with smaller holes and spacing was examined to assess the particle distribution on the holes (Table 4). The change in hole-size and spacing did not diminish the challenges of preferential sticking of the sample particle to the carbon support film surface which led to visualization of empty holes.

As a last resort, data collection for SIRT6-NCP 601L sample was carried out on the carbon support film that was outside of the holes. Data pre-processing was attempted, but failed at the particle picking stage. The background noise signal from the carbon film was too high and led to too poor signal-to-noise ratio for proceeding with the auto-picking process.

Quantifoil Holey Carbon Grid Type	Hole Size In μm	Spacing In μm	Period In μm (Center to center)
R2/2	1.2	1.3	2.5
R1.2/1.3	2.0	2.0	4.0

Table 4. The types of Quantifoil holey carbon grids tested in SIRT6-NCP 601L assembly cryo-EM experiments.

Figure 55. The images taken on graphene oxide coated holey carbon grids.

5.4 Co-crystallization attempts of SIRT6-NCP complex

Crystallization screening was carried out using 96 well microplates. Various screening conditions from both the commercial and homemade crystallization kits (Refer to Appendix III) were utilized to obtain optimal crystallization conditions for the complex. The conditions which yielded crystals were then reproduced with a larger drop size in 24 well plates. The reproducibility of the crystal in a larger drop size would be followed by the next stage of optimization for growing single, suitable size crystals (Figure 56). Variation of pH, buffering agent type and concentration of monovalent and divalent salts, precipitants concentration, salting in *versus* salting out approach, as well as crystallization setup (sitting *versus* hanging droplets) were the main parameters to be explored, one at a time in order to determine the effect on the crystal morphology and size.

Figure 56. Optimized SIRT6-NCP 601L assembly crystallization conditions. Crystals grown in buffer condition: A) 0.1 M NaAc pH 5, 0.2 M NaCl, 28% MPD. B) 0.1 M Na Citrate pH 5.5, 33% PEG 200. C) 0.1 M Mg Acetate, 0.05 M Na Cacodylate pH 6.5, 25% PEG 200.

Three droplets containing the crystals were harvested and analyzed by a 15% SDS-PAGE (Figure 57). Most of the SIRT6 protein was found in the drop solution, with only a faint trace of the SIRT6 protein present in the crystal. In order to rule out the possibility that there is little or no SIRT6 protein present in the crystal due to the packing arrangement in the crystal lattice, the crystals were still processed through post-crystallization soaking and mounted onto the X-ray machine. Unfortunately, the data collected showed that the space group was the same as the native NCP crystals, hence the probability that SIRT6 was incorporated into the crystals was very low.

Figure 57. 15% SDS-PAGE assessment to determine the content of the crystals.

The difficulty in crystallizing a SIRT6-NCP 601L complex might be due to the buffer condition, which favoured crystallization of NCP alone. Perhaps, longer nucleosomal DNA constructs were required for incorporation of SIRT6 protein into the crystal lattice. In order to avoid a purely NCP crystal lattice where the factor is excluded, trace fluorescence labeling of SIRT6 was utilized for the verification of complex-favoring conditions *in situ* before proceeding with further optimization work. All successful crystallization buffer conditions for protein-nucleosome complexes were compiled and modified to be used for high throughput crystallization hit screening.

In our trace fluorescence labeling setup, carboxyrhodamine fluorescence dye was used for covalent derivatization of SIRT6, and a final of 0.5% labelled SIRT6 was used for co-crystallization screening process. Crystals were subjected to fluorescence illumination for detection of SIRT6 factor bound crystals. Regrettably, the emitted red fluorescence was visible only in the drop solution but not in the crystal itself, indicating that SIRT6 protein was not being incorporated into the crystal lattice. Besides, most of the buffer conditions did not yield crystals of a desirable size despite different crystallization setups, such as microbatch and hanging vapour diffusion.

Structural characterization of SIRT6-NCP complex utilizing both EM and X-ray crystallization approaches has been challenging. This might be due to the fast on and fast off properties of SIRT6 binding and it requires a modified nucleosome substrate to bind stronger since it is an enzyme involves in modification removal. Acetylated nucleosome at the H3 tail could be attempted for co-crystallization with SIRT6. Other than that, cryo-EM method could be pursued further with optimization of buffer for complex stabilization, or cross-linking approach could also be tried to avoid the complex falling apart on the EM grids.

Chapter 6 Results and Discussion Part 4: Linker histones binding in chromatin from the structural perspective

Mammalian chromatin is highly dynamic and heterogeneous, thus the binding mode of linker histone to nucleosome has been of great interest for understanding the molecular basis of chromatin structure and dynamics. With all the latest findings and solved structures of linker histone-nucleosome assemblies by using NMR spectroscopy, X-ray crystallography, and cryo-EM (Bednar et al., 2017; Song et al., 2014; Zhou et al., 2013; Zhou et al., 2015), there are limited number of structures available and we still lack a unified understanding on the binding mode of the linker histones to the heterogeneous and crowded chromatin structures in order to impact on the cellular functionalities.

While in the process of gaining new perspectives on chromatin organization, a higher order chromatin structure without the 30 nm fiber has been proposed (Razin and Gavrilov, 2014). Since H1 has been long known to facilitate the 30 nm fiber formation (Robinson and Rhodes, 2006), the binding mode of linker histone in chromatin require to be re-investigated in facilitating higher order chromatin structure. The present studies are mainly revolving around nucleosomal substrates, but longer chromatin fiber in a crowded environment, which can better mimic the *in vivo* environment, would be a better target for studying the mode of linker histone disposition in facilitating chromatin fiber condensation.

In order to achieve the objective mentioned above, a rational design process of nucleosomal constructs is adopted by our lab to obtain well-diffracting crystals and structures that could provide us with engineered pseudo-continuous chromatin fiber as a model for investigating linker histone binding mode.

6.1 Nucleosome-Linker Histone Assembly (NLH) EMSA studies

The binding of linker histones towards nucleosome and di-nucleosomes was first assessed by utilizing EMSA before proceeding to crystallization set up for X-ray diffraction studies. When H1.0 bound to NUC 169s, the complex band migrated slower due to the higher molecular weight as shown in the native PAGE gel (Figure 58). In order to strike a compromise

between the fully shifted band and precipitation as seen in the well, a ratio of 1.2:1 H1.0: NUC 169s was adopted for crystallization set ups.

For EMSA of NLH H1.0-diNUCs (Figure 59), to achieve the binding ratio of one linker histone to one nucleosome, stoichiometry ratio of at least 2.0:1 has to be used, since there are two nucleosome particles in one diNUC. The band shift effect is little due to much smaller size of H1.0 as compared to di-nucleosome. Stoichiometry ratio of 2.4:1 (H1.0:diNUC) was used for crystallization setups in order to avoid partial occupancy due to inactive protein.

Figure 58. Linker histone H1.0 was titrated against same amount of NUC 169s.

Figure 59. EMSA of H1.0 binding to diNUC 343 and diNUC 349.

6.2 X-Ray Crystallization and Data Collection

Crystallization setups were conducted via hanging drop vapor diffusion approach and salting-in method. Salting-out was also attempted at the earlier stages but resulted in higher precipitation rates due to a lower initial divalent salt concentration in the drop. For salting-in method, the initial salt concentration in the droplet was higher and got lowered as equilibration with the well buffer occurred over time. This method was helpful in prevention of precipitation of the complex by having higher divalent salt content at the initial stage of crystallization set up.

During the crystallization drop setups, the nucleosome sample was first mixed well with the buffer that was twice the concentration of the well buffer before adding the linker histone protein and mixing the drop thoroughly. Various nucleosome and di-nucleosome constructs were utilized for co-crystallization with different linker histone variants, and the crystals were optimized to reach the desired size for X-ray data collection.

Initial crystal screening was performed with the in-house Rigaku X-ray machine. Data collection was only done if the crystal diffracted to at least 7 Å. Low resolution datasets would

still be sufficient to provide some hints on the macromolecular packing in the crystal lattice. Thereafter, more crystals would be prepared for high-resolution data collection at the synchrotron beamline. Preliminary data processing would be done with the automated processing pipeline developed at the facility and further processing would be done in-lab with iMOSFLM or XDS interface.

6.2.1 Nucleosome-Linker Histone Assembly (NLH)

Various mono-nucleosome constructs had been tested for co-crystallization with linker histone variant H1.x by Dr. Zenita Adhireksan and Dr. Bao Quiye. Out of the many that were attempted, only the NUC 169s with 4-nucleotide overhangs gave crystals that could be optimized to reach the desired size for X-ray data collection.

My project started with characterizing assembly of NUC 169s with linker histone variants H1.0, H1.3 and H1.5. According to the EMSA results, the NUC: LH ratio of 1: 1.2 was chosen for the co-crystallization setup. Both palindromic and non-palindromic constructs were explored to address the technical issues during large scale DNA purification. The shift from NUC 169s to NUC 169sn facilitated the production of DNA in a shorter time frame. Yet, the change in the palindromicity and restriction sites had resulted in a difference in the crystal unit cell dimensions (discussed later in Section 6.4).

Nonetheless, both co-crystallization of NUC 169s or NUC 169sn with linker histone yielded crystals with similar morphologies under the same crystallization buffer condition (10 mM NaAc pH 4.5, 25 mM KCl, 43-45 mM CaCl₂) (Figure 60). Crystals were harvested in half of the salt concentration of the reservoir buffer (10 mM NaAc pH 4.5, 12.5 mM KCl, 20 mM CaCl₂). Then, post-crystallization treatment was carried out in a step gradient manner by including up to 2% trehalose and 65% MPD in the harvesting buffer.

Figure 60. NLH crystals. A) H10-NUC169s. B) H10-NUC169sn.

Heavy metal soaking was tested at the post-crystallization stage in order to attempt to improve electron density maps associated with disordered linker histone tails. The single wavelength anomalous dispersion (SAD) approach was elected since good initial phases are available from molecular replacement. Several factors such as the heavy atom occupancy, characteristics of heavy metal agents and the working pH were taken into consideration to ensure sufficient anomalous signal could be obtained. Several strategies were adopted for the incorporation of heavy metals such as strontium, mercury, lead, and tantalum during the post-crystallization soaking treatment (Table 5).

Strontium cations and tantalum bromide clusters were chosen for the purpose of enhancing the phasing power. Since calcium ion was one of the endogenous metals in the stabilization buffer, strontium which is a heavier metal with similar characteristics was used to substitute it. Tantalum bromide is one of the polymetallic clusters of heavy metals that are frequently used to phase large complexes. It could provide strong anomalous signal and phasing power at low resolution via being identified as a single giant feature in the electron density map.

At the same time, these large clusters could disrupt the crystal lattice and deteriorate the resolution if binding were to induce structural changes.

For labeling purposes, lead acetate, ethyl mercury phosphate, and mercury acetate were selected to be included during the soaking stage. Due to the presence of several cysteine residues in the linker histone proteins, mercury compounds were chosen to be included during the soaking experiment. Mercury cation belongs to the class B element which forms covalent bonds with soft ligands through forming anionic complexes. While lead cation falls into the transition metal category so it typically interacts with proteins via electrostatic interactions. The solubility of lead acetate was tested at both pH 4.5 (sodium acetate buffer) and pH 6 (MES buffer). Lead acetate was observed to be well solubilized in the buffer sodium acetate pH 4.5.

One of the biggest challenges in heavy atom derivatization is the incorporation of the heavy atom into the crystal lattice often induces perturbation and leads to a change in the unit cell dimensions. Out of the five heavy atom compounds that were attempted, four of them resulted in deterioration of crystal quality and compromise of the data resolution. Only those crystals treated with lead acetate gave a similar resolution as native crystals, but the anomalous signal was only found on nucleosome sites (Table 5).

Heavy atoms	Working concentration	Strategy	Outcome
Strontium chloride	20 mM	Replace CaCl ₂ during the post-crystallization soaking step	Crystals quality deteriorated
Lead acetate	1 mM	Add into the crystals after post-crystallization soaking step for overnight	Anamalous signal found on nucleosome only.
Mercury acetate	1 mM	Add into the crystals after post-crystallization soaking step for 6 hours	Crystals quality deteriorated
Ethyl mercury phosphate	1 mM	Add into the crystals after post-crystallization soaking step for 6 hours	Crystals quality deteriorated
Tantalum cluster	1 mM	Add into the crystals after post-crystallization soaking step for 6 hours	Crystals quality deteriorated

Table 5. The strategies and outcomes of heavy metals incorporation for NLH crystals phasing and labelling.

6.2.2 Assembly of Di-Nucleosome with Linker Histone H1.0 (NLH H1.0-diNUC)

A series of di-nucleosome constructs spanning 343, 345, 347, 349, 351, 353, 357, 359, and 361 bp were tested for co-crystallization with several linker histone variants. This series of diNUC constructs consists of a constant 29 bp shared linker DNA between the two nucleosomes. The linker DNA arms located at both ends of the two nucleosomes have a varied length of increasing 1 bp for the subsequent construct of the series, with 10 bp + 4-nucleotide 3' overhangs linker DNA arms for diNUC 343 (Figure 61).

While for the z series diNUC, the DNA constructs were designed based on the X-ray crystal structure of diNUC 343 (discussed later in Section 6.3) in an attempt to recapitulate a continuous 2-start, zig-zag chromatin fiber (Song et al., 2014). The z series DNA constructs comprise a constant 31 bp shared linker DNA and varying length of linker DNA arms at both ends of the nucleosomes. There is total of four constructs for the z series diNUC: diNUC 351z, 353z, 355z, and 357z. All the work on the NLH H1.0-diNUC project was carried out collaboratively with Dr. Zenita Adhireksan and Dr. Deepti Sharma. The structures of NLH H1.0-diNUC349 and NLH H1.0-diNUC353 were modelled by Dr. Zenita Adhireksan.

Figure 61. Schematic representation for diNUC343 linker DNA.

The crystallization buffer used initially for the mono-nucleosomes did not work for all diNUC constructs. Thus, diNUC-linker histone assembly crystallization screening was carried out by substituting the calcium chloride in the crystallization buffer with different divalent salts such as magnesium chloride, manganese chloride, and spermine tetrahydrochloride. 2-8% of low molecular weight poly- γ -glutamic acid (LM-PGA) was also included as an additive in the

crystallization buffer since there is a study showing that PGA could act as a protein precipitant in facilitating macromolecule crystallization (Hu et al., 2008). Despite all the trials, diNUC 345, 347 and 351 did not give any crystal hits that were good enough to proceed with further optimization. DiNUC 357, 359, 361, and all the z series constructs are currently under the screening stage.

Only the co-crystallization of diNUC 343, 349 and 353 with linker histones yielded reasonable size crystals with decent morphology under buffer condition with calcium divalent metals and PGA (Figure 62). In spite of that, distinct crystals morphologies were obtained for different diNUC constructs assembled with linker histone H1.0. The crystals for H1.0-diNUC 343 assembly have a rod-like orthorhombic shape, while the H1.0-diNUC 349 crystals are long and monoclinic. Last but not least, the H1.0-diNUC 353 assembly shows twinned, bent and irregular crystals (Figure 62C). Modulation of length at linker DNA ends of as little as 2 bp (on either end; 349 versus 353) resulted in completely different crystal morphologies. This phenomenon suggested that there was a unique packing configuration of macromolecules in the crystal lattice.

The NLH crystals were usually treated with up to 65% 2-methyl-2,4-pentanediol (MPD) dehydration agent to improve the diffraction quality through promoting the order of macromolecule packing in the crystal lattice. Nonetheless, the NLH H1.0-diNUC crystal has a distinct unit cell dimension from the NLH H1.0-NUC crystals and the NLH H1.0-diNUC crystals soaked in 65% MPD did not diffract well. Thus, different dehydration agents such as PEG 400, glycerol and ethylene glycol were also screened to access the stability of the NLH H1.0-diNUC crystals in different post-crystallization soaking buffers. Besides functioning as the dehydration agent, these chemicals also act as cryo-protectants to prevent ice formation when the crystals are in contact with nitrogen stream (Makde and Tan, 2013; McGinty et al., 2016). The concentration of the cryo-protectant had to be optimized since ice formation could deteriorate the X-ray diffraction data, but on the contrary, the addition of too much dehydration agent might alter the ordering within the crystal lattice.

In an effort to strike a balance between the dehydration and cryo-protection effects, the X-ray diffraction spots were compared at the lowest concentration of each cryo-protectant where there was no visible ice ring pattern (Figure 63). For PEG 400 and glycerol, the concentration

had to go up to 30% while 25% of ethylene glycol was sufficient to eliminate ice ring formation. A higher concentration of ethylene glycol (up to 45%) was also tested but no improvement of diffraction could be observed. Subsequently, MPD was then attempted for crystal soaking at a lower concentration. Out of all the cryo-protectants tried, MPD at 25% gave the best diffraction spots with the highest resolution of up to 4 Å when using the in-house X-ray machine.

Figure 62. Crystals of H10-diNUCs assembly. A) H10-diNUC343 assembly. B) H10-diNUC 349 assembly. C) H10-diNUC353 assembly.

Figure 63. X-ray diffraction images of crystals H10-diNUC 349 assembly soaked in different soaking buffer conditions. A) 27% PEG 400. B) 30% PEG 400. C) 27% Glycerol. D) 30% Glycerol.

6.3 Structural Refinement and Analysis

6.3.1 NLH H1.0-NUC structures

From the crystallization setups, both the NLH H10-NUC169s and NLH H1.0-NUC169sn assemblies shared a similar crystallization buffer condition as well as crystal morphology. However, from the processed datasets, NLH H1.0-NUC169sn has double the cell dimension length at one edge as compared to NLH H1.0-NUC169s (Table 6). A unit cell is the basic repetitive geometric arrangement in the lattice that defines a crystal's characteristic symmetry. Thus, from the cell dimension parameters acquired, the unit cell of NLH H1.0-NUC169sn crystal would be predicted to have double the number of macromolecules in the unit cell of NLH H1.0-NUC169s crystal.

A)	H1.0-169s ^p	B)	H1.0-169sn
	Data collection		Data collection
	Space group		P2 ₁
	Cell dimensions		
	<i>a</i> (Å)		213.85
	<i>b</i> (Å)		102.46
	<i>c</i> (Å)		218.29
	$\alpha/\beta/\gamma$ (deg.)		90.00/100.5/90.00
	Resolution (Å)		3.51–88.28 (3.51–3.57)
	Unique reflections		117,231
	<i>R</i> _{merge} (%)		12.4 (182.1)
	<i>R</i> _{pim} (%)		5.4 (81.1)
	<i>I</i> / σI		8.0 (1.1)
	CC _{1/2} (%)		100.0 (37.0)
	Completeness (%)		100.0 (100.00)
	Redundancy		6.1 (5.9)
	Refinement		Refinement
	Resolution (Å)		3.51–88.28
	Reflections used		110,177
	<i>R</i> _{work} / <i>R</i> _{free} (%)		21.3 / 28.1
	No. atoms		52,715
	Core histone		24,396
	Linker histone		595
	DNA		27,724
	Solvent		--
	<i>B</i> -factors (Å ²)		160
	Core histone		124
	Linker histone		202
	<i>globular domain</i>		202
	<i>tail domains</i>		193
	DNA		190
	Solvent		--
	R.m.s. deviations		
	Bond lengths (Å)		0.007
	Bond angles (°)		1.30

*X-ray wavelength used for data collection was 1.0 Å. Values in parentheses are for the highest-resolution shell.

Table 6. Data collection and refinement statistics tables of A) NLH H1.0-NUC169s and B) NLH H1.0-NUC169sn.

NLH H1.0-NUC169s

The NLH of H10-NUC169s was solved and a crisscrossed closed system was observed, with two nucleosomes' cohesive ends base-paired to form a hand-shake motif from opposing direction. DNA cohesive end base pairing achieved through DNA engineering allows more controllable crystallization properties that can stabilize the crystal lattice.

In one asymmetric unit, there are two nucleosome-linker histone complexes paired up through the self-complementary overhangs from each of the linker DNA ends. Layer arrangements can be seen in the crystal lattice, with almost full face-to-face overlapping of nucleosomes and a zig-zag disposition of the asymmetric unit along the stacking plane. The nucleosome packing is at a dense state where you can see phosphodiester backbone contacts via interdigitation of double helices (Figure 64).

Figure 64. NLH H1.0-NUC169s. A) The top view of nucleosome packing in crystal lattice. B) The side view of nucleosome packing in crystal lattice. C) The top view of the NLH H1.0-NUC169s asymmetric unit. D) The side view of the NLH H1.0-NUC169s asymmetric unit.

Electron density is sufficiently clear to allow H1.0 modeling (Figure 65). The H1.0 linker histones are situated asymmetrically at the periphery of the two nucleosomes, where contacts are made with these 'host' nucleosomes as well as contacts formed with the neighboring nucleosomes. Each H1.0 is in contact with 5 nucleosomes from 3 layers of nucleosome packing.

The main interaction is via the insertion of helix 3 into the minor groove around SHL 6 of the laterally situated nucleosome and helix 1 clipping the phosphate backbone while the N-terminal tail spans towards the minor groove. The loop 3 interacts with the proximal nucleosome at SHL 3 via hydrogen bonding and hydrophobic stacking. Side chains from helix 2 are contacting the host nucleosome around SHL 0.5 through hydrogen bonding with the sugar-phosphate backbone, while the C-terminal region of the helix 1 hydrogen bonds with the DNA end of the paired nucleosome. The last interaction is via the C-terminus tail that extends all the way up and interacts with the opposite layer nucleosome (Figure 66 & Table 7).

Figure 65. An F_o-F_c omit electron density map (magenta; contoured at 3σ ; H1.0 atoms omitted from the model) corresponding to the H1.0 ($\alpha 1$ helix at the front) in the H1.0-169s structure. A) Without H1.0. B) With H1.0.

Figure 66. Inter-nucleosomal contacts provide H1.0 binding sites. Each H1.0 is contacting 5 nucleosomes in the crystal packing of NLH 169s: host nucleosome, paired nucleosome, proximal nucleosome, lateral nucleosome, and distal nucleosome.

H1.0 Motifs	H1.0 Residues	Nuc169s Locations	NUC169s Positions
N-terminal tails, Loop 2 and Helix 3	K12, K17, K21, S29, E62, D63, K69	Lateral Nucleosome	SHL5 and SHL 6 bases as well as sugar phosphate backbone
Helix 1, Loop 1 and C-terminal tails	Q37, K40, K97, E100	Paired Nucleosome	Linker DNA termini bases and sugar phosphate backbone
Helix 2	K52	Host Nucleosome	SHL0.5 sugar phosphate backbone
Beta sheet loop	K85, V87	Proximal Nucleosome	SHL3 bases and sugar phosphate backbone
C-terminal tails	K116, V117, T119	Distal Nucleosome	Linker DNA termini bases and sugar phosphate backbone

Table 7. Summary of NLH H1.0-NUC169s interaction points.

NLH H1.0-NUC169sn

The structure of NLH H1.0-NUC169sn was also solved and a similar crisscrossed closed system was observed through a hand-shake motif by two base-pairing nucleosomes. In spite of that, the asymmetric unit contains four nucleosomes instead of two as seen in NLH H1.0-

NUC169s (Figure 68). The overall nucleosome packing in the crystal lattice resembles that of the NLH H1.0-NUC169s without being identical, as seen from the planes of the crisscrossed nucleosome pairs. A similar dense packing state is visible through interdigitation of double helices, almost full face-to-face overlapping of nucleosomes across layers and zig-zag disposition of the asymmetric unit along the stacking plane (Figure 67).

Figure 67. The packing of the NLH H1.0-NUC169sn asymmetric unit in the crystal lattice. A) Top view. B) Side view.

Figure 68. The asymmetric unit of NLH H1.0-NUC169sn. A) A top view of the asymmetric unit. B) The front side of a side view of the asymmetric unit. C) The back side of a side view of the asymmetric unit.

Only one copy of H1.0 (in magenta) per one asymmetric unit of NLH H1.0-NUC169sn (4 nucleosomes) could be resolved (Figure 68). H1.0 binds at a different location as compared to the H1.0 binding sites in NLH 169s. H1.0 contacts only 3 nucleosomes in the asymmetric unit of NLH 169sn, which include the host nucleosome, paired nucleosome, and proximal nucleosome (Figure 69). An inter-nucleosomal context provides a unique H1.0 binding site. The closest contact is made by loop 1, helix 2, and the beta sheet region of H1.0 with the nucleosome that is termed “host nucleosome” at the off dyad region. The second interaction point is made between the sugar phosphate backbone at the SHL-2 of the proximal nucleosome with loop 2 and helix 3 of H1.0. The third contact point is made by the helix 2 and beta sheet loop of H1.0 with the linker DNA end of the paired nucleosome (Table 8).

Figure 69. H1.0 is contacting 3 nucleosomes in the crystal packing of NLH 169sn: host nucleosome, paired nucleosome, and proximal nucleosome.

H1.0 Motifs	H1.0 Residues	Nuc169sn Locations	NUC169sn Positions
Loop 1 and beta sheet region	K40, R42, R94, K97	Host Nucleosome	Sugar phosphate backbone at the off dyad location
Helix 2 and beta sheet loop	R47, Q48, K85, S90	Paired Nucleosome	Sugar phosphate backbone at linker DNA end
Loop 2 and Helix 3	E62, N63, D65, K69	Proximal Nucleosome	Sugar phosphate backbone at the SHL-2

Table 8. Summary of NLH H1.0-NUC169sn interaction points.

The H1.0 binding in NLH H1.0-NUC169sn asymmetric unit displays a 3-contact mode, similar to the known on-dyad binding mode. Therefore, globular avian H5 in complex with a NUC 167 (PDB: 4QLC) by Zhou et al., 2015 was superimposed onto the H1.0 in the context of NLH H1.0-NUC169sn (Figure 70). From the superimposition, the inter-nucleosomal context of NLH H1.0-NUC169sn furnishes an alternative binding environment for H1.0 with respect to the on-dyad location of NUC 167, suggesting that the H1.0 has a specific but not strict predisposition towards a 3-contact DNA environment for interaction regardless of nucleosomal DNA or linker DNA.

Figure 70. gH5-NUC167 (PDB: 4QLC) was superimposed onto H1.0 in the NLH H1.0-NUC169sn.

The employment of the two different restriction sites (KpnI and SacI) was intended to bring about specific base-pairing of the linker DNA ends in the formation of a pseudo-continuous double helix. For NUC 169s, due to its palindromic sequence and identical PstI restriction sites at both DNA ends, the nucleosome halves are symmetrical and do not affect orientation in the crystal packing. On the contrary, the non-palindromic sequence and distinct restriction sites of NUC 169sn allow the nucleosome dimers to adopt two distinct pairing orientations (within the dimer), but we only observe specific pairing of the two different termini. Nonetheless, the resolution of the structure was insufficient for reliable interpretation on the specific orientation of the dimers themselves within the crystal (Figure 71).

Figure 71. The two orientations that can be adopted by one base-paired NLH H1.0-NUC169sn nucleosome pseudo-dimer due to the two distinct restriction sites: KpnI and SacI.

The changes in the sticky end sequences, palindromicity, and overall DNA sequence apparently causes H1.0 to bind at a different location in NLH H1.0-NUC169sn in comparison with NLH H1.0-NUC169s, albeit the length of linker DNA is the same between the two construct. From previous studies, linker histones have been shown to recognize some specific features of DNA sequence for interaction (An et al., 1998; Cui and Zhurkin, 2009). From the superimposition of the NLH H1.0-NUC169s and its symmetry mate onto the NLH H1.0-NUC169sn asymmetric unit (Figure 72), a deviation of the stacking planes of the paired nucleosomes can be seen, which shows that the H1.0 binding at the corresponding location would be incompatible in the opposing context. The interplay between the DNA sequence and H1.0 binding mode in dictating the crystal lattice arrangement demonstrates a versatility of H1.0 in recognizing the nucleosomal environment.

Figure 72. The superimposition of NLH H1.0-NUC169s and its symmetry mate onto NLH H1.0-NUC169sn. Superimposition was done based on one nucleosome only to illustrate stacking geometry.

High quality crystal formation and X-ray diffraction are directed by ordered macromolecular arrangement in the crystal lattice. The kinetics and thermodynamics of macromolecule packing in the crystal lattice are largely governed by geometric and conformational features, but as well stabilizing intermolecular interactions (Bukowska and Grutter, 2013). In order for nucleosomes to overcome the energy barrier in mediating nascent crystal nuclei formation, their packing in the lattice is most likely dominated by the propensity of nucleosomes to form the face to face stacking arrangements (Daban, 2011; Scheffer et al., 2012). Apart from that, the cooperative effect from the DNA sequence, the engineered DNA cohesive end pairing kinetics as well as the inter-nucleosomal DNA backbone interdigitation also have an impact on the nucleosomal arrangement in the crystal lattice as seen from the NLH structures.

Understanding the macromolecules packing and interactions in NLH 169s and 169sn crystal has propelled us to design new nucleosomal constructs that could achieve a fiber-like array throughout the crystal by self-assembly from inter-nucleosomal base-pairing, rather than crystallizing a large array itself, which might seem unfeasible. This concept of nucleosomal DNA design for crystal engineering could be used as a platform to study nuclear protein interaction in the context of different crystal lattice arrangements.

6.3.2 NLH H1.0-diNUC structures

High resolution X-ray diffraction data was collected at the synchrotron beamline for NLH H1.0-diNUC343, H1.0-diNUC349, and H1.0-diNUC353, which was later processed and proceeded with structural solution. Whereas the same space group and nearly identical unit cell dimensions are shared by NLH H1.0-diNUC343 and H1.0-diNUC353, the NLH H1.0-diNUC349 crystals have a higher symmetry and different cell dimensions (Table 9). The lattice system changes from monoclinic to orthorhombic and then back to monoclinic again as the linker DNA length at each terminus (not counting the 4-nucleotide 3' overhang) increases from 10 bp (for diNUC 343) to 13 bp (for diNUC 349) and then 15 bp (for diNUC 353).

A)	H1.0-diNUC 343	B)	H1.0-diNUC 349	C)	H1.0-diNUC 353
Data collection		Data collection		Data collection	
Space group	P2 ₁	Space group	P2 ₁ 2 ₁ 2 ₁	Space group	P2 ₁
Cell dimensions		Cell dimensions		Cell dimensions	
<i>a</i> (Å)	107.75	<i>a</i> (Å)	121.06	<i>a</i> (Å)	107.81
<i>b</i> (Å)	205.90	<i>b</i> (Å)	184.07	<i>b</i> (Å)	204.91
<i>c</i> (Å)	237.68	<i>c</i> (Å)	223.07	<i>c</i> (Å)	237.33
$\alpha/\beta/\gamma$ (deg.)	90.00/97.19/90.0	$\alpha/\beta/\gamma$ (deg.)	90.00/90.00/90.0	$\alpha/\beta/\gamma$ (deg.)	90.00/97.58/90.0
Resolution (Å)	3.89–49.30 (3.89–4.10)	Resolution (Å)	3.40–49.32 (3.40–3.59)	Resolution (Å)	4.10–94.75 (4.10–4.32)
Unique reflections	93,736	Unique reflections	69,141	Unique reflections	80,117
<i>R</i> _{merge} (%)	7.1 (174.7)	<i>R</i> _{merge} (%)	15.1 (248.4)	<i>R</i> _{merge} (%)	11.6 (163.4)
<i>R</i> _{pim} (%)	3.2 (77.4)	<i>R</i> _{pim} (%)	4.4 (72.7)	<i>R</i> _{pim} (%)	5.2 (73.9)
<i>I</i> / σI	11.4 (1.1)	<i>I</i> / σI	12.1 (1.1)	<i>I</i> / σI	5.3 (1.0)
CC½ (%)	100 (47.8)	CC½ (%)	100 (41.0)	CC½ (%)	99.6 (40.7)
Completeness (%)	99.3 (96.1)	Completeness (%)	99.8 (98.6)	Completeness (%)	99.9 (99.9)
Redundancy	6.9 (6.7)	Redundancy	13.3 (13.1)	Redundancy	5.6 (5.6)
Refinement		Refinement		Refinement	
Resolution (Å)	3.89–49.35	Resolution (Å)	3.40–49.32	Resolution (Å)	4.10–94.11
Reflections used	91,832	Reflections used	67,687	Reflections used	78,497
<i>R</i> _{work} / <i>R</i> _{free} (%)	20.1 / 26.97	<i>R</i> _{work} / <i>R</i> _{free} (%)	20.64 / 26.46	<i>R</i> _{work} / <i>R</i> _{free} (%)	24.93 / 29.93
No. atoms	53,828	No. atoms	27,058	No. atoms	54,263
Core histone	23,732	Core histone	12,149	Core histone	22,664
Linker histone	1,150	Linker histone	568	Linker histone	1150
DNA	28,130	DNA	14,311	DNA	28,950
<i>B</i> -factors (Å ²)	240.10	<i>B</i> -factors (Å ²)	161.50	<i>B</i> -factors (Å ²)	263.86
Core histone	194	Core histone	131	Core histone	188
Linker histone	335	Linker histone	199	Linker histone	447
globular domain	335	globular domain	199	globular domain	447
DNA	277	DNA	184	DNA	319
R.m.s. deviations		R.m.s. deviations		R.m.s. deviations	
Bond lengths (Å)	0.005	Bond lengths (Å)	0.006	Bond lengths (Å)	0.004
Bond angles (°)	1.22	Bond angles (°)	1.26	Bond angles (°)	1.29

*X-ray wavelength used for data collection was 1.0 Å. Values in parentheses are for the highest-resolution shell.

Table 9. Data collection and refinement statistics for the NLHs. A) H1.0-diNUC 343 assembly. B) H1.0-diNUC 349 assembly. C) H1.0-diNUC 353 assembly.

NLH H1.0-diNUC343

The X-ray structure of NLH H1.0-diNUC343 was solved to 3.89 Å resolution (Table 9). The diNUC construct self-assembles into pairs coinciding with tetraNUC asymmetric units. The two diNUCs in the tetraNUC units are connected via both linker DNA arms, yielding a closed system in which the order of interaction is NUC1-2-3-4-1-etc. (Figure 74). The tetraNUC asymmetric units are arranged in a compact, zig-zag fashion along the face-to-face stacking plane, with DNA interdigitation occurring between the nucleosome edges (Figure 73).

Figure 73. The arrangement of the NLH H1.0-diNUC343 in the crystal lattice. A) A top view. B) A side view along the stacking plane.

Figure 74. The asymmetric unit of the NLH H1.0-diNUC343. The order of the di-nucleosome connection is according to NUC1-2-3-4-1-etc.

Electron density is sufficiently clear to allow H1.0 modeling only at NUC 1 and 4 (Figure 75), which coincides with the on-dyad linker histone binding mode (Zhou et al., 2015). This can be rationalized by the fact that the linker DNA arms of NUC 2 and 3 display a more open conformation (from the distance measured), which would appear incompatible with the binding as observed for NUC 1 & 4 (Figure 77 & Table 11). Full length H1.0 was utilized for growing crystals but insufficient density is observed over the tail regions to allow modelling, meaning that the tail regions are quite disordered. From the analysis of the two H1.0 globular domains modelled at NUC 1 and NUC4 of NLH H1.0-diNUC343, contacts are mediated by electrostatic interactions between side chains of charged residues and the DNA phosphate backbone (Figure 76 & Table 10).

Figure 75. An F_0-F_C omit electron density map (magenta; contoured at 1σ ; H1.0 atoms omitted from the model) corresponding to the H1.0 ($\alpha 1$ helix at the front) in the NLH H1.0-diNUC343 structure. A) Without H1.0. B) With H1.0.

The tetranucleosomal asymmetric unit of NLH H1.0-diNUC shows a configuration that is similar to the 2-start, zig-zag model for 30 nm chromatin fiber (Dorigo et al., 2004; Schalch et al., 2005), although the structural nature of the 30 nm fiber and its relevance *in vivo* have remained uncertain (Maeshima et al., 2014). The cryo-EM structure of the nucleosome array containing human linker histone H1.4 revealed a twisted, zig-zag tetranucleosome as the repeating unit where H1.4 binds off the dyad (Song et al., 2014). However, a distinct broader

nucleosomal packing configuration is seen in our structure and H1.0 adopts the on-dyad mode instead.

Figure 76. H1.0 binding to NUC 1 and NUC 4 in the NLH H1.0-diNUC343 structure.

H1.0 Motifs	H1.0 Residues	NUC1/NUC4 Locations	NUC1/NUC4 Positions
Helix 2 and helix 3	R47, Q48, Q51, K52, K69, S90	Host Nucleosome	Sugar phosphate backbone at the on dyad location of the nucleosome
Helix 2 and beta sheet loop	K27, N63, Q67, R74		Sugar phosphate backbone at linker DNA arm
Loop 2 and Helix 3	R42, K85, R94		Sugar phosphate backbone at linker DNA arm

Table 10. Details of interaction between H1.0 and NUC1, NUC4 in the NLH H1.0-diNUC343 asymmetric unit.

Figure 77. NUC 2, 3, and 4 are superimposed onto NUC 1 to compare the linker DNA distance from H1.0.

Distance of H1.0 to Linker DNA	Key interacting residues of H1.0 in contacting linker DNA	NUC1 (Å) (H1.0 bound)	NUC2 (Å)	NUC3 (Å)	NUC4 (Å) (H1.0 bound)
To linker DNA 1	K27, N63, Q67, R74	3.8–5.5	3.5–8.6	Steric clash	5.4- 7.1
To linker DNA 2	R42, K82, K85, R94	4.0- 4.8	6.9- 9.8	6.9- 9.7	5.5- 6.2

Table 11. Distance between the H1.0 key interacting residues and the nearest linker DNA.

NLH H1.0-diNUC349

A 3.40 Å resolution structure of NLH H1.0-diNUC349 was solved by Dr. Zenita Adhireksan. From the space group and unit cell dimension, NLH H1.0-diNUC349 shares a distinct lattice configuration as compared to NLH H1.0-diNUC343 and NLH H1.0-diNUC353. The structure revealed a unique arrangement of the diNUCs into a continuous chromatin fiber. A NLH H1.0-diNUC349 is positioned on the same side, orthogonal to the base-pairing linker DNA arms which connect to two discrete NLH H1.0-diNUC349 from the opposite side to form a continuous chromatin fiber (Figure 79). The packing of NLH H1.0-diNUC349 in the crystal lattice is characterized by both face-to-face and edge-to-face stacking (Figure 78). Extensive interactions between DNA and core histones appear to be important in facilitating this stacking

pattern, which is different from the other NLH structures solved so far that all display purely planar nucleosome.

Figure 78. NLH H1.0-diNUC349 crystal lattice arrangement. A) Top view. B) Side view.

Figure 79. A) A top view and B) a side view of an asymmetric unit of NLH H1.0-diNUC349.

H1.0 binds in an inter-nucleosomal context via a 3-contact binding mode in the NLH H1.0-diNUC349 structure (Figure 80). The globular avian H5 in complex with a NUC 167 (Zhou et al., 2015) was superimposed onto the H1.0 from NLH H1.0-diNUC349 to compare their interaction environments. The beta sheet loop of H1.0 in NLH H1.0-diNUC349 sits in the minor groove of SHL 6, analogous to on-dyad binding mode depicted by the gH5-NUC167 complex. As opposed to the contacts made with both the linker DNA arms by gH5 in NUC 167, the H1.0 in NLH H1.0-diNUC349 interacts with two neighboring diNUCs at SHL 6 and SHL 3.5 & -4 respectively (Figure 81). The NLH H1.0-diNUC349 structure was another clear instance where linker histone is recognizing the nucleosome core DNA instead of linker DNA arms. However, slightly different in this situation, the nucleosomal DNA between the gyres is serving as the interaction point for H1.0 linker histone.

Alteration of the linker DNA length as little as 3 bp (NLH H1.0-diNUC349) at each terminus has resulted in a dramatically distinct configuration in the crystal lattice arrangement. The increase in DNA length at the free linker DNA would impact both the length and twist angle of the sticky ends in forming Watson-Crick base-pairing with the partner nucleosomes, thus allowing for distinct macromolecule arrangements in the crystal lattice.

Figure 81. gH5-NUC167 (PDB: 4QLC) superimposed onto H1.0 in the NLH H1.0-diNUC349 structure. A similar 3-contact binding mode is observed.

NLH H1.0-diNUC353

The structure of NLH H1.0-diNUC353 was solved to 4.10 Å resolution by Dr. Zenita Adhireksan. In the crystal lattice of NLH H1.0-diNUC353, an analogous packing system as NLH H1.0-diNUC343 is adopted. A closed tetraNUC asymmetric unit is also assumed through base-pairing of two diNUC 353 (in the order of NUC1-2-3-4-1-etc; Figure 83). However, the way in which the two diNUCs interact is altered in comparison with the NLH H1.0-diNUC343 asymmetric unit. Whereas the 343 diNUCs are interdigitated, those of the 353 are stacked one on top of the other. Nevertheless, for both systems, a nearly identical zig-zag packing along the face-to-face stacking plane of the asymmetric unit occurs (Figure 82).

Figure 82. A) A side view and B) a top view of NLH H1.0-diNUC353 crystal lattice arrangement.

Figure 83. The asymmetric unit of NLH H1.0-diNUC353 and the order of nucleosome connectivity is shown as a guide. A) Side view. B) Top view.

Nearly identical unit cell dimensions and asymmetric unit structures are assumed by the NLH H1.0-diNUC343 and NLH H1.0-diNUC353 constructs (Figure 84), indicative of the mechanical plasticity of the DNA in being able to accommodate a difference of extra 10 bp with in essentially the same crystal packing volume. The addition of 5 bp at both of the terminal linker DNA arms of diNUC 343, to make the diNUC 353 construct, is equivalent to almost half a double helix turn advance at both termini. Within the tetraNUC asymmetric unit in either case, the longer linker DNA is positioned diagonally (internal), whereas the shorter linker DNA is situated along the same plane (external). As such, the longer linker DNA of NLH H1.0-diNUC343 is the shared 29 bp continuous tract (14.5 bp long for each of the two sharing NUCs)

that takes the longer diagonal path, and conversely, it is the resulting 34 bp conjoined linker DNA of NLH H1.0-diNUC353, which is the longer, that instead takes the diagonal pathway.

The work on different length diNUC constructs has provided some insight into favoured nucleosome packing and stacking arrangements in the crystal lattice. The close structural relationship between the 343 and 353 constructs suggests that favoured packing configurations can dominate the outcome, whereby the double helix adjusts conformationally to accommodate DNA length and twist relationships (Chua et al., 2012). These degrees of freedom would need to be taken into account in designing new DNA constructs to achieve specific structural arrangements.

Figure 84. The asymmetric unit of NLH H1.0-diNUC353 superimposed onto NLH H1.0-diNUC343.

The structural investigation of nucleosome recognition by linker histones have yielded a few proposed binding models (Bednar et al., 2017; Song et al., 2014; Zhou et al., 2013; Zhou et al., 2015). The inconsistencies of the linker histone binding mode could be due to differences in species, variant-specificity, incompleteness of linker histone sequence or unsuitable nucleosome substrate to represent the *in vivo* chromatin heterogeneity. The nucleosome models used so far are tri-nucleosomes and nucleosome arrays, but the low resolution lacks the sufficient detail of the interaction site (Syed et al., 2010). While, most of the *in vitro* studies have been focusing on avian H5 and *Drosophila* H1, these linker histone might have biophysical features distinct from the mammalian, thus their functional relevance remains to be studied. The role of linker histone in regulating the structure and function of chromatin is still poorly understood, more work would

be necessary from the structural perspective to address the linker histone versatility in coping with a wide range of cellular chromatin heterogeneity.

Through interpretation of all the structures solved by our group so far (Table 12), some insights were gained in how the linker histone H1.0 carries out its function as a chromatin architectural protein in a crowded and heterogeneous chromatin structure. We revealed that the binding of H1.0 is versatile to adapt to different nucleosomal environment contexts. The H1.0 interaction is based upon the chromatin environment via a specific yet tolerant recognition of the DNA geometry provided in an optimal spatial distance. From our structure analysis, H1.0 is recognizing at least a 3-contact DNA geometry and the key residues has to be interacting with DNA in a distance of $\leq 5\text{-}6\text{ \AA}$. In line with our findings, H1.0 accumulates in terminally-differentiated cells and implicated in malignant transformation (Zlatanova and Doenecke, 1994).

NLH structures	H1.0 binding mode
H1.0-NUC 169s	5 contact points
H1.0-NUC169sn	3 contact points
H1.0-diNUC343	On-dyad 3 contact points
H1.0-diNUC349	3 contact points
H1.0-diNUC353	On-dyad 3 contact points

Table 12. Summary of H1.0 binding modes across various NLH structures.

On top of that, the intrinsic properties of DNA sequences are likely to have indirect influence on the placement of H1.0 too. Different H1 variants and the effect of PTMs also play a role in determining H1 localization in chromatin. The physiological salt conditions and crowding effect that exist in the cellular nucleus have to be taken into consideration for studying the linker histone-chromatin interaction in our crystallization approach. This crystallographic platform appear to be helpful in deconvoluting chromatin structure and linker histone binding behavior.

Chapter 7: Current perspectives and future directions

Structural Recognition of Nucleosome by PARP1

My studies are initially aimed at obtaining the structure of two chromatin modification proteins, PARP1 and SIRT6, in association with the nucleosome to best represent how these factors interact with chromatin to achieve their nuclear activities. This kind of structural information would be critical for biological insight and to foster novel structure-based drug design. This could lead to potent and selective modulation of the specific protein as well as its enzymatic activity for the development of future therapeutics.

PARP1 has been shown to bind to NCP with nanomolar affinity and its binding stoichiometry appears to be 1:1 at limiting concentrations of PARP1. On the other hand, SIRT6 appears to associate with the NCP in a higher stoichiometry, most likely 2:1. We imagine that one molecule of SIRT6 may bind to either face of the NCP, whereas PARP1 should associate with one of the DNA termini. Binding of a second molecule of PARP1 to the other free DNA terminus may be inhibited sterically by the other associated PARP1 molecule. Structure characterizations of the macromolecular assemblies appear to be challenging and more optimization work would be required for both X-ray crystallographic and electron microscopic techniques.

For the PARP1 project, crystallization attempts carried out by Dr. Deepti Sharma have so far not been successful. This might be due to the huge size and the flexibility of the linker loops on the PARP1 protein, which hinders an ordered crystal lattice to form. Thus, crystallization of the full length PARP1-NCP complex was not favoured. Electron microscopy has come up as an alternative to obtain the structure of this huge protein complex with nucleosome. Some ideas have been gathered from the negative stain approach before proceeding with cryo-EM. Nonetheless, more work is needed to determine the suitable concentration and particle distribution on cryo grids. Optimization of buffer conditions would need to be looked into, for instance the use of surfactants for screening to evaluate the stability of the complex. Gradient fixation could also be attempted, if homogeneous complex could be obtained upon loaded onto EM grids.

Some PARP inhibitors can also be included in the complex sample to see if the conformation heterogeneity of PARP1 binding to NCP could be reduced when the cofactor binding site is occupied. Reduced heterogeneity of the complex molecules would likely facilitate complex structure resolution, in particular with the crystallographic approaches.

One of the main challenges of cryo-EM would be the low phase contrast for biological molecules. This leads to the low phase contrast of PARP1 protein and its occupancy on NCP could not yet be confirmed. Collecting such data might deteriorate the protein density during the data processing stage by averaging out the signal with noise. Recent advances in phase plate imaging could overcome this issue and enable interpreting signal over noise to facilitate the structure determination process (Chua et al., 2016). However, the phase plate imaging technology in the EM platform needs to be further developed for integration of these advances into the automated data collection process.

Structural Characterization of SIRT6 binding to nucleosomes

Both crystallization and EM approaches have been utilized for attempting to obtain a SIRT6-NCP structure. The crystallization conditions optimized so far resulted in nucleosome-only crystals. For imaging SIRT6-NCP complex using electron microscopy, more effort is needed to obtain a stable macromolecule complex for high-resolution structural characterization. For instance, optimization of the buffer conditions or use of gradient-fixation could be tested, to acquire homogeneous, high-occupancy, and single-particle SIRT6-NCP complex.

The macromolecule complex components could also be varied to increase the binding affinity and stability of the complex. Since SIRT6 is a deacetylase enzyme, acetylated substrate would be preferred for SIRT6 binding with higher affinity. Other than the acetylated nucleosome, mutation of histone H3 K9Q and K56Q can be carried out to compare the binding affinity and kinetics parameters using EMSA, SPR or BLI. The modified and mutated NCP can be used for structural determination by using both the crystallization and EM approaches in parallel. The structure of the macromolecule complex would be essential for explaining any differences in affinity values between distinct SIRT6-NCP complexes, which would also shed light on the therapeutic potential of the specific modification in various cellular functions.

Linker histones binding in chromatin from the structural perspective

Distinct nucleosome packing configurations could be achieved via modulation of linker DNA length, twist parameter, palindromicity and etc. As such, these structural investigations appear to provide a rationale and a means for understanding the heterogeneous nature of chromatin organization in the cell, since the irregular and dynamic native chromatin fiber appears to be incompatible for high-resolution structural analysis.

Based on the NLH H1.0-diNUC343 structure, a new ‘z series’ of DNA constructs was designed in an attempt to achieve a zig-zag fiber-like arrangement in the crystal lattice through the base-pairing of the sticky linker DNAs (Figure 85). A 2 bp DNA double helix was inserted at the shared linker DNA to revolve the twist angle of the two connected nucleosomes by $\sim 70^\circ$. The change in the twist parameter of the DNA was engineered to open up the closed tetraNUC system in the expectation that a continuous zig-zag configuration could be adopted. The z series diNUC constructs are currently under the crystallization screening stage.

Figure 85. Illustration of a potential fiber-like arrangement that could be achieved through sticky end base-pairing with the z series diNUC constructs.

Besides modulation of DNA length, twist angle, sequence, and palindromicity, post-translational modified H1 which could enhance the chromatin condensation process could be tested for crystallization as well to investigate if alternative binding mode would be adapted. Other than that, the H1 variant-specific binding mode also needs to look into by incorporating distinct human H1 variants, since variant-specific functional heterogeneity has been proposed.

References

- Allan, J., Hartman, P.G., Crane-Robinson, C., and Aviles, F.X. (1980a). The structure of histone H1 and its location in chromatin. *Nature* 288, 675-679.
- Allan, J., Staynov, D.Z., and Gould, H. (1980b). Reversible dissociation of linker histone from chromatin with preservation of internucleosomal repeat. *Proceedings of the National Academy of Sciences of the United States of America* 77, 885-889.
- Ame, J.C., Spenlehauer, C., and de Murcia, G. (2004). The PARP superfamily. *BioEssays : news and reviews in molecular, cellular and developmental biology* 26, 882-893.
- An, W., Leuba, S.H., van Holde, K., and Zlatanova, J. (1998). Linker histone protects linker DNA on only one side of the core particle and in a sequence-dependent manner. *Proceedings of the National Academy of Sciences of the United States of America* 95, 3396-3401.
- Anderson, J.D., and Widom, J. (2000). Sequence and position-dependence of the equilibrium accessibility of nucleosomal DNA target sites. *Journal of molecular biology* 296, 979-987.
- Armache, K.J., Garlick, J.D., Canzio, D., Narlikar, G.J., and Kingston, R.E. (2011). Structural basis of silencing: Sir3 BAH domain in complex with a nucleosome at 3.0 Å resolution. *Science* 334, 977-982.
- Arnaudo, N., Fernandez, I.S., McLaughlin, S.H., Peak-Chew, S.Y., Rhodes, D., and Martino, F. (2013). The N-terminal acetylation of Sir3 stabilizes its binding to the nucleosome core particle. *Nat Struct Mol Biol* 20, 1119-1121.
- Ausio, J. (2015). The shades of gray of the chromatin fiber: recent literature provides new insights into the structure of chromatin. *BioEssays : news and reviews in molecular, cellular and developmental biology* 37, 46-51.
- Avery, O.T., MacLeod, C.M., and McCarty, M. (1979). Studies on the chemical nature of the substance inducing transformation of pneumococcal types. Inductions of transformation by a desoxyribonucleic acid fraction isolated from pneumococcus type III. *The Journal of experimental medicine* 149, 297-326.
- Avvakumov, N., Nourani, A., and Cote, J. (2011). Histone chaperones: modulators of chromatin marks. *Molecular cell* 41, 502-514.
- Bannister, A.J., and Kouzarides, T. (2011). Regulation of chromatin by histone modifications. *Cell research* 21, 381-395.
- Barbera, A.J., Chodaparambil, J.V., Kelley-Clarke, B., Joukov, V., Walter, J.C., Luger, K., and Kaye, K.M. (2006). The nucleosomal surface as a docking station for Kaposi's sarcoma herpesvirus LANA. *Science* 311, 856-861.

- Bates, D.L., and Thomas, J.O. (1981). Histones H1 and H5: one or two molecules per nucleosome? *Nucleic acids research* 9, 5883-5894.
- Battye, T.G., Kontogiannis, L., Johnson, O., Powell, H.R., and Leslie, A.G. (2011). iMOSFLM: a new graphical interface for diffraction-image processing with MOSFLM. *Acta crystallographica Section D, Biological crystallography* 67, 271-281.
- Bednar, J., Garcia-Saez, I., Boopathi, R., Cutter, A.R., Papai, G., Reymer, A., Syed, S.H., Lone, I.N., Tonchev, O., Crucifix, C., *et al.* (2017). Structure and Dynamics of a 197 bp Nucleosome in Complex with Linker Histone H1. *Molecular cell* 66, 384-397 e388.
- Bednar, J., Horowitz, R.A., Grigoryev, S.A., Carruthers, L.M., Hansen, J.C., Koster, A.J., and Woodcock, C.L. (1998). Nucleosomes, linker DNA, and linker histone form a unique structural motif that directs the higher-order folding and compaction of chromatin. *Proceedings of the National Academy of Sciences of the United States of America* 95, 14173-14178.
- Belmont, A.S. (2014). Large-scale chromatin organization: the good, the surprising, and the still perplexing. *Current opinion in cell biology* 26, 69-78.
- Belmont, A.S., Braunfeld, M.B., Sedat, J.W., and Agard, D.A. (1989). Large-scale chromatin structural domains within mitotic and interphase chromosomes in vivo and in vitro. *Chromosoma* 98, 129-143.
- Benafif, S., and Hall, M. (2015). An update on PARP inhibitors for the treatment of cancer. *OncoTargets and therapy* 8, 519-528.
- Bheda, P., Jing, H., Wolberger, C., and Lin, H. (2016). The Substrate Specificity of Sirtuins. *Annual review of biochemistry* 85, 405-429.
- Bilokapic, S., Strauss, M., and Halic, M. (2018). Structural rearrangements of the histone octamer translocate DNA. *Nat Commun* 9, 1330.
- Blander, G., and Guarente, L. (2004). The Sir2 family of protein deacetylases. *Annual review of biochemistry* 73, 417-435.
- Bohrmann, B., and Kellenberger, E. (1994). Immunostaining of DNA in electron microscopy: an amplification and staining procedure for thin sections as alternative to gold labeling. *The journal of histochemistry and cytochemistry : official journal of the Histochemistry Society* 42, 635-643.
- Bonet-Costa, C., Vilaseca, M., Diema, C., Vujatovic, O., Vaquero, A., Omenaca, N., Castejon, L., Bernues, J., Giralt, E., and Azorin, F. (2012). Combined bottom-up and top-down mass spectrometry analyses of the pattern of post-translational modifications of *Drosophila melanogaster* linker histone H1. *Journal of proteomics* 75, 4124-4138.
- Bouchard, V.J., Rouleau, M., and Poirier, G.G. (2003). PARP-1, a determinant of cell survival in response to DNA damage. *Experimental hematology* 31, 446-454.

- Boveri, T. (2008). Concerning the origin of malignant tumours by Theodor Boveri. Translated and annotated by Henry Harris. *J Cell Sci 121 Suppl 1*, 1-84.
- Brown, D.T., Izard, T., and Misteli, T. (2006). Mapping the interaction surface of linker histone H1(0) with the nucleosome of native chromatin in vivo. *Nature structural & molecular biology 13*, 250-255.
- Brown, J.S., Kaye, S.B., and Yap, T.A. (2016). PARP inhibitors: the race is on. *British journal of cancer 114*, 713-715.
- Bryant, H.E., Schultz, N., Thomas, H.D., Parker, K.M., Flower, D., Lopez, E., Kyle, S., Meuth, M., Curtin, N.J., and Helleday, T. (2005). Specific killing of BRCA2-deficient tumours with inhibitors of poly(ADP-ribose) polymerase. *Nature 434*, 913-917.
- Bukowska, M.A., and Grutter, M.G. (2013). New concepts and aids to facilitate crystallization. *Current opinion in structural biology 23*, 409-416.
- Burkart, V., Wang, Z.Q., Radons, J., Heller, B., Herceg, Z., Stingl, L., Wagner, E.F., and Kolb, H. (1999). Mice lacking the poly(ADP-ribose) polymerase gene are resistant to pancreatic beta-cell destruction and diabetes development induced by streptozocin. *Nature medicine 5*, 314-319.
- Calabrese, C.R., Almassy, R., Barton, S., Batey, M.A., Calvert, A.H., Canan-Koch, S., Durkacz, B.W., Hostomsky, Z., Kumpf, R.A., Kyle, S., *et al.* (2004). Anticancer chemosensitization and radiosensitization by the novel poly(ADP-ribose) polymerase-1 inhibitor AG14361. *Journal of the National Cancer Institute 96*, 56-67.
- Campos, E.I., and Reinberg, D. (2009). Histones: annotating chromatin. *Annual review of genetics 43*, 559-599.
- Carafa, V., Rotili, D., Forgione, M., Cuomo, F., Serretiello, E., Hailu, G.S., Jarho, E., Lahtela-Kakkonen, M., Mai, A., and Altucci, L. (2016). Sirtuin functions and modulation: from chemistry to the clinic. *Clinical epigenetics 8*, 61.
- Cerrato, A., Morra, F., and Celetti, A. (2016). Use of poly ADP-ribose polymerase [PARP] inhibitors in cancer cells bearing DDR defects: the rationale for their inclusion in the clinic. *Journal of experimental & clinical cancer research : CR 35*, 179.
- Chen, Z.H., Zhu, M., Yang, J., Liang, H., He, J., He, S., Wang, P., Kang, X., McNutt, M.A., Yin, Y., *et al.* (2014). PTEN interacts with histone H1 and controls chromatin condensation. *Cell Rep 8*, 2003-2014.
- Chua, E.Y., Vasudevan, D., Davey, G.E., Wu, B., and Davey, C.A. (2012). The mechanics behind DNA sequence-dependent properties of the nucleosome. *Nucleic acids research 40*, 6338-6352.
- Chua, E.Y., Vogirala, V.K., Inian, O., Wong, A.S., Nordenskiold, L., Plitzko, J.M., Danev, R., and Sandin, S. (2016). 3.9 Å structure of the nucleosome core particle determined by phase-plate cryo-EM. *Nucleic acids research 44*, 8013-8019.

- Clapier, C.R., Iwasa, J., Cairns, B.R., and Peterson, C.L. (2017). Mechanisms of action and regulation of ATP-dependent chromatin-remodelling complexes. *Nat Rev Mol Cell Biol* 18, 407-422.
- Clark, D.J., and Kimura, T. (1990). Electrostatic mechanism of chromatin folding. *J Mol Biol* 211, 883-896.
- Clark, N.J., Kramer, M., Muthurajan, U.M., and Luger, K. (2012). Alternative modes of binding of poly(ADP-ribose) polymerase 1 to free DNA and nucleosomes. *J Biol Chem* 287, 32430-32439.
- Clore, G.M., Gronenborn, A.M., Nilges, M., Sukumaran, D.K., and Zarbock, J. (1987). The polypeptide fold of the globular domain of histone H5 in solution. A study using nuclear magnetic resonance, distance geometry and restrained molecular dynamics. *The EMBO journal* 6, 1833-1842.
- Collaborative Computational Project, N. (1994). The CCP4 suite: programs for protein crystallography. *Acta Crystallogr D Biol Crystallogr* 50, 760-763.
- Cui, F., and Zhurkin, V.B. (2009). Distinctive sequence patterns in metazoan and yeast nucleosomes: implications for linker histone binding to AT-rich and methylated DNA. *Nucleic acids research* 37, 2818-2829.
- Curtin, N.J. (2005). PARP inhibitors for cancer therapy. *Expert reviews in molecular medicine* 7, 1-20.
- Cutter, A.R., and Hayes, J.J. (2015). A brief review of nucleosome structure. *FEBS Lett* 589, 2914-2922.
- D'Amours, D., Desnoyers, S., D'Silva, I., and Poirier, G.G. (1999). Poly(ADP-ribosyl)ation reactions in the regulation of nuclear functions. *The Biochemical journal* 342 (Pt 2), 249-268.
- Daban, J.R. (2011). Electron microscopy and atomic force microscopy studies of chromatin and metaphase chromosome structure. *Micron* 42, 733-750.
- Dahm, R. (2005). Friedrich Miescher and the discovery of DNA. *Developmental biology* 278, 274-288.
- Dantzer, F., Ame, J.C., Schreiber, V., Nakamura, J., Menissier-de Murcia, J., and de Murcia, G. (2006). Poly(ADP-ribose) polymerase-1 activation during DNA damage and repair. *Methods in enzymology* 409, 493-510.
- Dantzer, F., de La Rubia, G., Menissier-De Murcia, J., Hostomsky, Z., de Murcia, G., and Schreiber, V. (2000). Base excision repair is impaired in mammalian cells lacking Poly(ADP-ribose) polymerase-1. *Biochemistry* 39, 7559-7569.
- Das, C., Tyler, J.K., and Churchill, M.E. (2010). The histone shuffle: histone chaperones in an energetic dance. *Trends in biochemical sciences* 35, 476-489.

Daujat, S., Zeissler, U., Waldmann, T., Happel, N., and Schneider, R. (2005). HP1 binds specifically to Lys26-methylated histone H1.4, whereas simultaneous Ser27 phosphorylation blocks HP1 binding. *The Journal of biological chemistry* 280, 38090-38095.

Davey, C.A., Sargent, D.F., Luger, K., Maeder, A.W., and Richmond, T.J. (2002). Solvent mediated interactions in the structure of the nucleosome core particle at 1.9 Å resolution. *Journal of molecular biology* 319, 1097-1113.

David, K.K., Andrabi, S.A., Dawson, T.M., and Dawson, V.L. (2009). Parthanatos, a messenger of death. *Frontiers in bioscience* 14, 1116-1128.

De Blasio, A., Musmeci, M.T., Giuliano, M., Lauricella, M., Emanuele, S., D'Anneo, A., Vassallo, B., Tesoriere, G., and Vento, R. (2003). The effect of 3-aminobenzamide, inhibitor of poly(ADP-ribose) polymerase, on human osteosarcoma cells. *International journal of oncology* 23, 1521-1528.

de Murcia, J.M., Niedergang, C., Trucco, C., Ricoul, M., Dutrillaux, B., Mark, M., Oliver, F.J., Masson, M., Dierich, A., LeMeur, M., *et al.* (1997). Requirement of poly(ADP-ribose) polymerase in recovery from DNA damage in mice and in cells. *Proc Natl Acad Sci U S A* 94, 7303-7307.

Dekker, J., and Heard, E. (2015). Structural and functional diversity of Topologically Associating Domains. *FEBS letters* 589, 2877-2884.

DeLano, W.L. (2002). The PyMol molecular graphics sytem.

Deng, X., Ma, W., Ramani, V., Hill, A., Yang, F., Ay, F., Berletch, J.B., Blau, C.A., Shendure, J., Duan, Z., *et al.* (2015). Bipartite structure of the inactive mouse X chromosome. *Genome biology* 16, 152.

Dixon, J.R., Selvaraj, S., Yue, F., Kim, A., Li, Y., Shen, Y., Hu, M., Liu, J.S., and Ren, B. (2012). Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature* 485, 376-380.

Dominy, J.E., Jr., Lee, Y., Jedrychowski, M.P., Chim, H., Jurczak, M.J., Camporez, J.P., Ruan, H.B., Feldman, J., Pierce, K., Mostoslavsky, R., *et al.* (2012). The deacetylase Sirt6 activates the acetyltransferase GCN5 and suppresses hepatic gluconeogenesis. *Molecular cell* 48, 900-913.

Donawho, C.K., Luo, Y., Luo, Y., Penning, T.D., Bauch, J.L., Bouska, J.J., Bontcheva-Diaz, V.D., Cox, B.F., DeWeese, T.L., Dillehay, L.E., *et al.* (2007). ABT-888, an orally active poly(ADP-ribose) polymerase inhibitor that potentiates DNA-damaging agents in preclinical tumor models. *Clinical cancer research : an official journal of the American Association for Cancer Research* 13, 2728-2737.

Dorigo, B., Schalch, T., Kulangara, A., Duda, S., Schroeder, R.R., and Richmond, T.J. (2004). Nucleosome arrays reveal the two-start organization of the chromatin fiber. *Science* 306, 1571-1573.

- Downs, J.A., Kosmidou, E., Morgan, A., and Jackson, S.P. (2003). Suppression of homologous recombination by the *Saccharomyces cerevisiae* linker histone. *Molecular cell* *11*, 1685-1692.
- Droscher, A. (2002). Edmund B. Wilson's the cell and cell theory between 1896 and 1925. *Hist Philos Life Sci* *24*, 357-389.
- Du, J., Zhou, Y., Su, X., Yu, J.J., Khan, S., Jiang, H., Kim, J., Woo, J., Kim, J.H., Choi, B.H., *et al.* (2011). Sirt5 is a NAD-dependent protein lysine demalonylase and desuccinylase. *Science* *334*, 806-809.
- du Preez, L.L., and Patterson, H.G. (2013). Secondary structures of the core histone N-terminal tails: their role in regulating chromatin structure. *Sub-cellular biochemistry* *61*, 37-55.
- Eltsov, M., Maclellan, K.M., Maeshima, K., Frangakis, A.S., and Dubochet, J. (2008). Analysis of cryo-electron microscopy images does not support the existence of 30-nm chromatin fibers in mitotic chromosomes in situ. *Proceedings of the National Academy of Sciences of the United States of America* *105*, 19732-19737.
- Emsley, P., and Cowtan, K. (2004). Coot: model-building tools for molecular graphics. *Acta crystallographica Section D, Biological crystallography* *60*, 2126-2132.
- Eustermann, S., Wu, W.F., Langelier, M.F., Yang, J.C., Easton, L.E., Riccio, A.A., Pascal, J.M., and Neuhaus, D. (2015). Structural Basis of Detection and Signaling of DNA Single-Strand Breaks by Human PARP-1. *Molecular cell* *60*, 742-754.
- Fan, L., and Roberts, V.A. (2006). Complex of linker histone H5 with the nucleosome and its implications for chromatin packing. *Proc Natl Acad Sci U S A* *103*, 8384-8389.
- Fan, Y., Nikitina, T., Zhao, J., Fleury, T.J., Bhattacharyya, R., Bouhassira, E.E., Stein, A., Woodcock, C.L., and Skoultschi, A.I. (2005). Histone H1 depletion in mammals alters global chromatin structure but causes specific changes in gene regulation. *Cell* *123*, 1199-1212.
- Farmer, H., McCabe, N., Lord, C.J., Tutt, A.N., Johnson, D.A., Richardson, T.B., Santarosa, M., Dillon, K.J., Hickson, I., Knights, C., *et al.* (2005). Targeting the DNA repair defect in BRCA mutant cells as a therapeutic strategy. *Nature* *434*, 917-921.
- Felsenfeld, G. (2014). A brief history of epigenetics. *Cold Spring Harb Perspect Biol* *6*.
- Felsenfeld, G., and McGhee, J.D. (1986). Structure of the 30 nm chromatin fiber. *Cell* *44*, 375-377.
- Flanagan, T.W., and Brown, D.T. (2016). Molecular dynamics of histone H1. *Biochimica et biophysica acta* *1859*, 468-475.
- Fong, P.C., Boss, D.S., Yap, T.A., Tutt, A., Wu, P., Mergui-Roelvink, M., Mortimer, P., Swaisland, H., Lau, A., O'Connor, M.J., *et al.* (2009). Inhibition of poly(ADP-ribose) polymerase in tumors from BRCA mutation carriers. *The New England journal of medicine* *361*, 123-134.

- Frye, R.A. (2000). Phylogenetic classification of prokaryotic and eukaryotic Sir2-like proteins. *Biochemical and biophysical research communications* 273, 793-798.
- Fussner, E., Strauss, M., Djuric, U., Li, R., Ahmed, K., Hart, M., Ellis, J., and Bazett-Jones, D.P. (2012). Open and closed domains in the mouse genome are configured as 10-nm chromatin fibres. *EMBO reports* 13, 992-996.
- Fyodorov, D.V., Zhou, B.R., Skoultschi, A.I., and Bai, Y. (2018). Emerging roles of linker histones in regulating chromatin structure and function. *Nat Rev Mol Cell Biol* 19, 192-206.
- Gabrovsky, N., Georgieva, M., Laleva, M., Uzunov, K., and Miloshev, G. (2013). Histone H1.0- a potential molecular marker with prognostic value for patients with malignant gliomas. *Acta neurochirurgica* 155, 1437-1442.
- Gansen, A., Felekyan, S., Kuhnemuth, R., Lehmann, K., Toth, K., Seidel, C.A.M., and Langowski, J. (2018). High precision FRET studies reveal reversible transitions in nucleosomes between microseconds and minutes. *Nat Commun* 9, 4628.
- Garcia, B.A., Joshi, S., Thomas, C.E., Chitta, R.K., Diaz, R.L., Busby, S.A., Andrews, P.C., Ogorzalek Loo, R.R., Shabanowitz, J., Kelleher, N.L., *et al.* (2006). Comprehensive phosphoprotein analysis of linker histone H1 from *Tetrahymena thermophila*. *Molecular & cellular proteomics : MCP* 5, 1593-1609.
- Ghosh, S., George, S., Roy, U., Ramachandran, D., and Kolthur-Seetharam, U. (2010). NAD: a master regulator of transcription. *Biochimica et biophysica acta* 1799, 681-693.
- Gibcus, J.H., Samejima, K., Goloborodko, A., Samejima, I., Naumova, N., Nuebler, J., Kanemaki, M.T., Xie, L., Paulson, J.R., Earnshaw, W.C., *et al.* (2018). A pathway for mitotic chromosome formation. *Science* 359.
- Gibson, B.A., and Kraus, W.L. (2012). New insights into the molecular and cellular functions of poly(ADP-ribose) and PARPs. *Nature reviews Molecular cell biology* 13, 411-424.
- Gil, R., Barth, S., Kanfi, Y., and Cohen, H.Y. (2013). SIRT6 exhibits nucleosome-dependent deacetylase activity. *Nucleic acids research* 41, 8537-8545.
- Grigoryev, S.A., and Woodcock, C.L. (2012). Chromatin organization - the 30 nm fiber. *Experimental cell research* 318, 1448-1455.
- Groschel, S., Sanders, M.A., Hoogenboezem, R., de Wit, E., Bouwman, B.A.M., Erpelinck, C., van der Velden, V.H.J., Havermans, M., Avellino, R., van Lom, K., *et al.* (2014). A single oncogenic enhancer rearrangement causes concomitant EVI1 and GATA2 deregulation in leukemia. *Cell* 157, 369-381.
- Haigis, M.C., and Guarente, L.P. (2006). Mammalian sirtuins--emerging roles in physiology, aging, and calorie restriction. *Genes & development* 20, 2913-2921.

- Hansen, J.C., Connolly, M., McDonald, C.J., Pan, A., Pryamkova, A., Ray, K., Seidel, E., Tamura, S., Rogge, R., and Maeshima, K. (2018). The 10-nm chromatin fiber and its relationship to interphase chromosome organization. *Biochemical Society transactions* 46, 67-76.
- Happel, N., Warneboldt, J., Hanecke, K., Haller, F., and Doenecke, D. (2009). H1 subtype expression during cell proliferation and growth arrest. *Cell cycle* 8, 2226-2232.
- Hashimoto, H., Sonoda, E., Takami, Y., Kimura, H., Nakayama, T., Tachibana, M., Takeda, S., and Shinkai, Y. (2007). Histone H1 variant, H1R is involved in DNA damage response. *DNA repair* 6, 1584-1595.
- Hassa, P.O., Buerki, C., Lombardi, C., Imhof, R., and Hottiger, M.O. (2003). Transcriptional coactivation of nuclear factor-kappaB-dependent gene expression by p300 is regulated by poly(ADP)-ribose polymerase-1. *J Biol Chem* 278, 45145-45153.
- Hassa, P.O., Haenni, S.S., Elser, M., and Hottiger, M.O. (2006). Nuclear ADP-ribosylation reactions in mammalian cells: where are we today and where are we going? *Microbiology and molecular biology reviews* : MMBR 70, 789-829.
- Hassa, P.O., and Hottiger, M.O. (2002). The functional role of poly(ADP-ribose)polymerase 1 as novel coactivator of NF-kappaB in inflammatory disorders. *Cellular and molecular life sciences* : CMLS 59, 1534-1553.
- Heeres, J.T., and Hergenrother, P.J. (2007). Poly(ADP-ribose) makes a date with death. *Current opinion in chemical biology* 11, 644-653.
- Heinke, R., Carlino, L., Kannan, S., Jung, M., and Sippl, W. (2011). Computer- and structure-based lead design for epigenetic targets. *Bioorganic & medicinal chemistry* 19, 3605-3615.
- Hendzel, M.J., Lever, M.A., Crawford, E., and Th'ng, J.P. (2004). The C-terminal domain is the primary determinant of histone H1 binding to chromatin in vivo. *The Journal of biological chemistry* 279, 20028-20034.
- Hergeth, S.P., and Schneider, R. (2015). The H1 linker histones: multifunctional proteins beyond the nucleosomal core particle. *EMBO reports* 16, 1439-1453.
- Hershey, A.D., and Chase, M. (1952). Independent functions of viral protein and nucleic acid in growth of bacteriophage. *J Gen Physiol* 36, 39-56.
- Hottiger, M.O., Hassa, P.O., Luscher, B., Schuler, H., and Koch-Nolte, F. (2010). Toward a unified nomenclature for mammalian ADP-ribosyltransferases. *Trends in biochemical sciences* 35, 208-219.
- Houtkooper, R.H., Pirinen, E., and Auwerx, J. (2012). Sirtuins as regulators of metabolism and healthspan. *Nature reviews Molecular cell biology* 13, 225-238.

- Hu, J., Gu, L., Ye, Y., Zheng, M., Xu, Z., Lin, J., Du, Y., Tian, M., Luo, L., Wang, B., *et al.* (2018). Dynamic placement of the linker histone H1 associated with nucleosome arrangement and gene transcription in early *Drosophila* embryonic development. *Cell death & disease* 9, 765.
- Hu, T.C., Korczynska, J., Smith, D.K., and Brzozowski, A.M. (2008). High-molecular-weight polymers for protein crystallization: poly-gamma-glutamic acid-based precipitants. *Acta crystallographica Section D, Biological crystallography* 64, 957-963.
- Hubbard, B.P., and Sinclair, D.A. (2014). Small molecule SIRT1 activators for the treatment of aging and age-related diseases. *Trends in pharmacological sciences* 35, 146-154.
- Huletsky, A., de Murcia, G., Muller, S., Hengartner, M., Menard, L., Lamarre, D., and Poirier, G.G. (1989). The effect of poly(ADP-ribosylation) on native and H1-depleted chromatin. A role of poly(ADP-ribosylation) on core nucleosome structure. *J Biol Chem* 264, 8878-8886.
- Ikejima, M., Noguchi, S., Yamashita, R., Ogura, T., Sugimura, T., Gill, D.M., and Miwa, M. (1990). The zinc fingers of human poly(ADP-ribose) polymerase are differentially required for the recognition of DNA breaks and nicks and the consequent enzyme activation. Other structures recognize intact DNA. *J Biol Chem* 265, 21907-21913.
- Imai, S., Armstrong, C.M., Kaeberlein, M., and Guarente, L. (2000). Transcriptional silencing and longevity protein Sir2 is an NAD-dependent histone deacetylase. *Nature* 403, 795-800.
- Izzo, A., Kamieniarz, K., and Schneider, R. (2008). The histone H1 family: specific members, specific functions? *Biological chemistry* 389, 333-343.
- Izzo, A., and Schneider, R. (2016). The role of linker histone H1 modifications in the regulation of gene expression and chromatin dynamics. *Biochimica et biophysica acta* 1859, 486-495.
- Jackson, M.D., and Denu, J.M. (2002). Structural identification of 2'- and 3'-O-acetyl-ADP-ribose as novel metabolites derived from the Sir2 family of beta -NAD⁺-dependent histone/protein deacetylases. *J Biol Chem* 277, 18535-18544.
- Jagtap, P., and Szabo, C. (2005). Poly(ADP-ribose) polymerase and the therapeutic effects of its inhibitors. *Nature reviews Drug discovery* 4, 421-440.
- Javle, M., and Curtin, N.J. (2011). The role of PARP in DNA repair and its therapeutic exploitation. *British journal of cancer* 105, 1114-1122.
- Jeggo, P.A. (1998). DNA repair: PARP - another guardian angel? *Current biology : CB* 8, R49-51.
- Ji, Y., and Tulin, A.V. (2010). The roles of PARP1 in gene control and cell differentiation. *Current opinion in genetics & development* 20, 512-518.
- Jiang, H., Khan, S., Wang, Y., Charron, G., He, B., Sebastian, C., Du, J., Kim, R., Ge, E., Mostoslavsky, R., *et al.* (2013). SIRT6 regulates TNF-alpha secretion through hydrolysis of long-chain fatty acyl lysine. *Nature* 496, 110-113.

- Johnson, D.G., and Dent, S.Y. (2013). Chromatin: receiver and quarterback for cellular signals. *Cell* 152, 685-689.
- Joti, Y., Hikima, T., Nishino, Y., Kamada, F., Hihara, S., Takata, H., Ishikawa, T., and Maeshima, K. (2012). Chromosomes without a 30-nm chromatin fiber. *Nucleus* 3, 404-410.
- Kabsch, W. (2010). Xds. *Acta crystallographica Section D, Biological crystallography* 66, 125-132.
- Kaidi, A., Weinert, B.T., Choudhary, C., and Jackson, S.P. (2010). Human SIRT6 promotes DNA end resection through CtIP deacetylation. *Science* 329, 1348-1353.
- Kale, S., Goncarencu, A., Markov, Y., Landsman, D., and Panchenko, A.R. (2019). Molecular recognition of nucleosomes by binding partners. *Curr Opin Struct Biol* 56, 164-170.
- Kanfi, Y., Peshti, V., Gil, R., Naiman, S., Nahum, L., Levin, E., Kronfeld-Schor, N., and Cohen, H.Y. (2010). SIRT6 protects against pathological damage caused by diet-induced obesity. *Aging cell* 9, 162-173.
- Kato, H., Jiang, J., Zhou, B.R., Rozendaal, M., Feng, H., Ghirlando, R., Xiao, T.S., Straight, A.F., and Bai, Y. (2013). A conserved mechanism for centromeric nucleosome recognition by centromere protein CENP-C. *Science* 340, 1110-1113.
- Kawahara, T.L., Michishita, E., Adler, A.S., Damian, M., Berber, E., Lin, M., McCord, R.A., Ongaigui, K.C., Boxer, L.D., Chang, H.Y., *et al.* (2009). SIRT6 links histone H3 lysine 9 deacetylation to NF-kappaB-dependent gene expression and organismal life span. *Cell* 136, 62-74.
- Kawahara, T.L., Rappavoli, N.A., Wu, A.R., Qu, K., Quake, S.R., and Chang, H.Y. (2011). Dynamic chromatin localization of Sirt6 shapes stress- and aging-related transcriptional networks. *PLoS genetics* 7, e1002153.
- Khachaturov, V., Xiao, G.Q., Kinoshita, Y., Unger, P.D., and Burstein, D.E. (2014). Histone H1.5, a novel prostatic cancer marker: an immunohistochemical study. *Human pathology* 45, 2115-2119.
- Kim, K., Choi, J., Heo, K., Kim, H., Levens, D., Kohno, K., Johnson, E.M., Brock, H.W., and An, W. (2008). Isolation and characterization of a novel H1.2 complex that acts as a repressor of p53-mediated transcription. *The Journal of biological chemistry* 283, 9113-9126.
- Kim, K., Lee, B., Kim, J., Choi, J., Kim, J.M., Xiong, Y., Roeder, R.G., and An, W. (2013). Linker Histone H1.2 cooperates with Cul4A and PAF1 to drive H4K31 ubiquitylation-mediated transactivation. *Cell reports* 5, 1690-1703.
- Kim, M.Y., Mauro, S., Gevry, N., Lis, J.T., and Kraus, W.L. (2004). NAD⁺-dependent modulation of chromatin structure and transcription by nucleosome binding properties of PARP-1. *Cell* 119, 803-814.

- Ko, H.L., and Ren, E.C. (2011). Novel poly (ADP-ribose) polymerase 1 binding motif in hepatitis B virus core promoter impairs DNA damage repair. *Hepatology* 54, 1190-1198.
- Ko, H.L., and Ren, E.C. (2012). Functional Aspects of PARP1 in DNA Repair and Transcription. *Biomolecules* 2, 524-548.
- Kobayashi, W., and Kurumizaka, H. (2019). Structural transition of the nucleosome during chromatin remodeling and transcription. *Curr Opin Struct Biol* 59, 107-114.
- Koh, D.W., Dawson, T.M., and Dawson, V.L. (2005). Mediation of cell death by poly(ADP-ribose) polymerase-1. *Pharmacological research* 52, 5-14.
- Kornberg, R.D. (1974). Chromatin structure: a repeating unit of histones and DNA. *Science* 184, 868-871.
- Kornberg, R.D., and Lorch, Y. (1999). Twenty-five years of the nucleosome, fundamental particle of the eukaryote chromosome. *Cell* 98, 285-294.
- Kornberg, R.D., and Thomas, J.O. (1974). Chromatin structure; oligomers of the histones. *Science* 184, 865-868.
- Kouzarides, T. (2007). Chromatin modifications and their function. *Cell* 128, 693-705.
- Kowalski, A., and Palyga, J. (2016). Modulation of chromatin function through linker histone H1 variants. *Biol Cell* 108, 339-356.
- Koyama, M., and Kurumizaka, H. (2018). Structural diversity of the nucleosome. *J Biochem* 163, 85-95.
- Kraus, W.L., and Lis, J.T. (2003). PARP goes transcription. *Cell* 113, 677-683.
- Krishnakumar, R., and Kraus, W.L. (2010). The PARP side of the nucleus: molecular actions, physiological outcomes, and clinical targets. *Molecular cell* 39, 8-24.
- Kugel, S., and Mostoslavsky, R. (2014). Chromatin and beyond: the multitasking roles for SIRT6. *Trends in biochemical sciences* 39, 72-81.
- Kumar, A., and Chauhan, S. (2016). How much successful are the medicinal chemists in modulation of SIRT1: A critical review. *European journal of medicinal chemistry* 119, 45-69.
- Landau, D.A., and Wu, C.J. (2013). Chronic lymphocytic leukemia: molecular heterogeneity revealed by high-throughput genomics. *Genome medicine* 5, 47.
- Langelier, M.F., Planck, J.L., Roy, S., and Pascal, J.M. (2012). Structural basis for DNA damage-dependent poly(ADP-ribosylation) by human PARP-1. *Science* 336, 728-732.
- Lequieu, J., Schwartz, D.C., and de Pablo, J.J. (2017). In silico evidence for sequence-dependent nucleosome sliding. *Proc Natl Acad Sci U S A* 114, E9197-E9205.

- Lever, M.A., Th'ng, J.P., Sun, X., and Hendzel, M.J. (2000). Rapid exchange of histone H1.1 on chromatin in living human cells. *Nature* 408, 873-876.
- Li, B., Carey, M., and Workman, J.L. (2007). The role of chromatin during transcription. *Cell* 128, 707-719.
- Li, G., Margueron, R., Hu, G., Stokes, D., Wang, Y.H., and Reinberg, D. (2010). Highly compacted chromatin formed in vitro reflects the dynamics of transcription activation in vivo. *Molecular cell* 38, 41-53.
- Li, G., and Reinberg, D. (2011). Chromatin higher-order structures and gene regulation. *Current opinion in genetics & development* 21, 175-186.
- Li, H., Kaminski, M.S., Li, Y., Yildiz, M., Ouillette, P., Jones, S., Fox, H., Jacobi, K., Saiya-Cork, K., Bixby, D., *et al.* (2014). Mutations in linker histone genes HIST1H1 B, C, D, and E; OCT2 (POU2F2); IRF8; and ARID1A underlying the pathogenesis of follicular lymphoma. *Blood* 123, 1487-1498.
- Li, Y., You, L., Huang, W., Liu, J., Zhu, H., and He, B. (2015). A FRET-based assay for screening SIRT6 modulators. *European journal of medicinal chemistry* 96, 245-249.
- Lilyestrom, W., van der Woerd, M.J., Clark, N., and Luger, K. (2010). Structural and biophysical studies of human PARP-1 in complex with damaged DNA. *Journal of molecular biology* 395, 983-994.
- Liszt, G., Ford, E., Kurtev, M., and Guarente, L. (2005). Mouse Sir2 homolog SIRT6 is a nuclear ADP-ribosyltransferase. *J Biol Chem* 280, 21313-21320.
- Lohr, J.G., Stojanov, P., Lawrence, M.S., Auclair, D., Chapuy, B., Sougnez, C., Cruz-Gordillo, P., Knoechel, B., Asmann, Y.W., Slager, S.L., *et al.* (2012). Discovery and prioritization of somatic mutations in diffuse large B-cell lymphoma (DLBCL) by whole-exome sequencing. *Proceedings of the National Academy of Sciences of the United States of America* 109, 3879-3884.
- Lowary, P.T., and Widom, J. (1998a). New DNA sequence rules for high affinity binding to histone octamer and sequence-directed nucleosome positioning. *Journal of molecular biology* 276, 19-42.
- Lowary, P.T., and Widom, J. (1998b). New DNA sequence rules for high affinity binding to histone octamer and sequence-directed nucleosome positioning. *Journal of molecular biology* 276, 19-42.
- Loyola, A., and Almouzni, G. (2004). Histone chaperones, a supporting role in the limelight. *Biochimica et biophysica acta* 1677, 3-11.
- Loyola, A., and Almouzni, G. (2007). Marking histone H3 variants: how, when and why? *Trends in biochemical sciences* 32, 425-433.

- Lu, A., Zougman, A., Pudelko, M., Bebenek, M., Ziolkowski, P., Mann, M., and Wisniewski, J.R. (2009). Mapping of lysine monomethylation of linker histones in human breast and its cancer. *Journal of proteome research* 8, 4207-4215.
- Luger, K. (2006). Dynamic nucleosomes. *Chromosome research : an international journal on the molecular, supramolecular and evolutionary aspects of chromosome biology* 14, 5-16.
- Luger, K., Mader, A.W., Richmond, R.K., Sargent, D.F., and Richmond, T.J. (1997). Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature* 389, 251-260.
- Luger, K., Rechsteiner, T.J., and Richmond, T.J. (1999). Preparation of nucleosome core particle from recombinant histones. *Methods in enzymology* 304, 3-19.
- Lupianez, D.G., Kraft, K., Heinrich, V., Krawitz, P., Brancati, F., Klopocki, E., Horn, D., Kayserili, H., Opitz, J.M., Laxova, R., *et al.* (2015). Disruptions of topological chromatin domains cause pathogenic rewiring of gene-enhancer interactions. *Cell* 161, 1012-1025.
- Lyumkis, D. (2019). Challenges and opportunities in cryo-EM single-particle analysis. *The Journal of biological chemistry* 294, 5181-5197.
- Mabley, J.G., Jagtap, P., Perretti, M., Getting, S.J., Salzman, A.L., Virag, L., Szabo, E., Soriano, F.G., Liaudet, L., Abdelkarim, G.E., *et al.* (2001). Anti-inflammatory effects of a novel, potent inhibitor of poly (ADP-ribose) polymerase. *Inflammation research : official journal of the European Histamine Research Society [et al]* 50, 561-569.
- Maeshima, K., Imai, R., Tamura, S., and Nozaki, T. (2014). Chromatin as dynamic 10-nm fibers. *Chromosoma* 123, 225-237.
- Makde, R.D., England, J.R., Yennawar, H.P., and Tan, S. (2010). Structure of RCC1 chromatin factor bound to the nucleosome core particle. *Nature* 467, 562-566.
- Makde, R.D., and Tan, S. (2013). Strategies for crystallizing a chromatin protein in complex with the nucleosome core particle. *Analytical biochemistry* 442, 138-145.
- Makowski, M.M., Grawe, C., Foster, B.M., Nguyen, N.V., Bartke, T., and Vermeulen, M. (2018). Global profiling of protein-DNA and protein-nucleosome binding affinities using quantitative mass spectrometry. *Nat Commun* 9, 1653.
- Mao, Z., Hine, C., Tian, X., Van Meter, M., Au, M., Vaidya, A., Seluanov, A., and Gorbunova, V. (2011). SIRT6 promotes DNA repair under stress by activating PARP1. *Science* 332, 1443-1446.
- Masutani, M., Nakagama, H., and Sugimura, T. (2005). Poly(ADP-ribosyl)ation in relation to cancer and autoimmune disease. *Cellular and molecular life sciences : CMLS* 62, 769-783.
- McCord, R.A., Michishita, E., Hong, T., Berber, E., Boxer, L.D., Kusumoto, R., Guan, S., Shi, X., Gozani, O., Burlingame, A.L., *et al.* (2009). SIRT6 stabilizes DNA-dependent protein kinase at chromatin for DNA double-strand break repair. *Aging* 1, 109-121.

- McGinty, R.K., Henrici, R.C., and Tan, S. (2014). Crystal structure of the PRC1 ubiquitylation module bound to the nucleosome. *Nature* 514, 591-596.
- McGinty, R.K., Makde, R.D., and Tan, S. (2016). Preparation, Crystallization, and Structure Determination of Chromatin Enzyme/Nucleosome Complexes. *Methods Enzymol* 573, 43-65.
- McGinty, R.K., and Tan, S. (2015). Nucleosome Structure and Function. *Chem Rev* 115, 2255-2273.
- McNicholas, S., Potterton, E., Wilson, K.S., and Noble, M.E. (2011). Presenting your structures: the CCP4mg molecular-graphics software. *Acta crystallographica Section D, Biological crystallography* 67, 386-394.
- Medrzycki, M., Zhang, Y., McDonald, J.F., and Fan, Y. (2012). Profiling of linker histone variants in ovarian cancer. *Frontiers in bioscience* 17, 396-406.
- Michishita, E., McCord, R.A., Berber, E., Kioi, M., Padilla-Nash, H., Damian, M., Cheung, P., Kusumoto, R., Kawahara, T.L., Barrett, J.C., *et al.* (2008). SIRT6 is a histone H3 lysine 9 deacetylase that modulates telomeric chromatin. *Nature* 452, 492-496.
- Michishita, E., McCord, R.A., Boxer, L.D., Barber, M.F., Hong, T., Gozani, O., and Chua, K.F. (2009). Cell cycle-dependent deacetylation of telomeric histone H3 lysine K56 by human SIRT6. *Cell cycle* 8, 2664-2666.
- Misteli, T., Gunjan, A., Hock, R., Bustin, M., and Brown, D.T. (2000). Dynamic binding of histone H1 to chromatin in living cells. *Nature* 408, 877-881.
- Mostoslavsky, R., Chua, K.F., Lombard, D.B., Pang, W.W., Fischer, M.R., Gellon, L., Liu, P., Mostoslavsky, G., Franco, S., Murphy, M.M., *et al.* (2006). Genomic instability and aging-like phenotype in the absence of mammalian SIRT6. *Cell* 124, 315-329.
- Multani, A.S., and Chang, S. (2007). WRN at telomeres: implications for aging and cancer. *Journal of cell science* 120, 713-721.
- Murga, M., Jaco, I., Fan, Y., Soria, R., Martinez-Pastor, B., Cuadrado, M., Yang, S.M., Blasco, M.A., Skoultschi, A.I., and Fernandez-Capetillo, O. (2007). Global chromatin compaction limits the strength of the DNA damage response. *The Journal of cell biology* 178, 1101-1108.
- Naumova, N., Imakaev, M., Fudenberg, G., Zhan, Y., Lajoie, B.R., Mirny, L.A., and Dekker, J. (2013). Organization of the mitotic chromosome. *Science* 342, 948-953.
- Nayernia, K., Drabent, B., Meinhardt, A., Adham, I.M., Schwandt, I., Muller, C., Sancken, U., Kleene, K.C., and Engel, W. (2005). Triple knockouts reveal gene interactions affecting fertility of male mice. *Molecular reproduction and development* 70, 406-416.
- Ng, M.K., and Cheung, P. (2016). A brief histone in time: understanding the combinatorial functions of histone PTMs in the nucleosome context. *Biochemistry and cell biology = Biochimie et biologie cellulaire* 94, 33-42.

- Nishino, Y., Eltsov, M., Joti, Y., Ito, K., Takata, H., Takahashi, Y., Hihara, S., Frangakis, A.S., Imamoto, N., Ishikawa, T., *et al.* (2012). Human mitotic chromosomes consist predominantly of irregularly folded nucleosome fibres without a 30-nm chromatin structure. *The EMBO journal* *31*, 1644-1653.
- Nishiyama, M., Oshikawa, K., Tsukada, Y., Nakagawa, T., Iemura, S., Natsume, T., Fan, Y., Kikuchi, A., Skoultchi, A.I., and Nakayama, K.I. (2009). CHD8 suppresses p53-mediated apoptosis through histone H1 recruitment during early embryogenesis. *Nature cell biology* *11*, 172-182.
- Nishizuka, Y., Ueda, K., Nakazawa, K., and Hayaishi, O. (1967). Studies on the polymer of adenosine diphosphate ribose. I. Enzymic formation from nicotinamide adenine dinucleotide in mammalian nuclei. *J Biol Chem* *242*, 3164-3171.
- Nora, E.P., Lajoie, B.R., Schulz, E.G., Giorgetti, L., Okamoto, I., Servant, N., Piolot, T., van Berkum, N.L., Meisig, J., Sedat, J., *et al.* (2012). Spatial partitioning of the regulatory landscape of the X-inactivation centre. *Nature* *485*, 381-385.
- North, B.J., and Verdin, E. (2004). Sirtuins: Sir2-related NAD-dependent protein deacetylases. *Genome biology* *5*, 224.
- Nowak, S.J., and Corces, V.G. (2004). Phosphorylation of histone H3: a balancing act between chromosome condensation and transcriptional activation. *Trends in genetics : TIG* *20*, 214-220.
- Nozaki, T., Imai, R., Tanbo, M., Nagashima, R., Tamura, S., Tani, T., Joti, Y., Tomita, M., Hibino, K., Kanemaki, M.T., *et al.* (2017). Dynamic Organization of Chromatin Domains Revealed by Super-Resolution Live-Cell Imaging. *Molecular cell* *67*, 282-293 e287.
- Oberg, C., and Belikov, S. (2012). The N-terminal domain determines the affinity and specificity of H1 binding to chromatin. *Biochemical and biophysical research communications* *420*, 321-324.
- Oberg, C., Izzo, A., Schneider, R., Wrangé, O., and Belikov, S. (2012). Linker histone subtypes differ in their effect on nucleosomal spacing in vivo. *Journal of molecular biology* *419*, 183-197.
- Ohi, M., Li, Y., Cheng, Y., and Walz, T. (2004). Negative Staining and Image Classification - Powerful Tools in Modern Electron Microscopy. *Biological procedures online* *6*, 23-34.
- Okosun, J., Bodor, C., Wang, J., Araf, S., Yang, C.Y., Pan, C., Boller, S., Cittaro, D., Bozek, M., Iqbal, S., *et al.* (2014). Integrated genomic analysis identifies recurrent mutations and evolution patterns driving the initiation and progression of follicular lymphoma. *Nature genetics* *46*, 176-181.
- Olins, A.L., Moyer, B.A., Kim, S.H., and Allison, D.P. (1989). Synthesis of a more stable osmium ammine electron-dense DNA stain. *The journal of histochemistry and cytochemistry : official journal of the Histochemistry Society* *37*, 395-398.
- Olins, A.L., and Olins, D.E. (1974). Spheroid chromatin units (v bodies). *Science* *183*, 330-332.

- Olins, D.E., and Olins, A.L. (2003). Chromatin history: our view from the bridge. *Nature reviews Molecular cell biology* 4, 809-814.
- Ou, H.D., Phan, S., Deerinck, T.J., Thor, A., Ellisman, M.H., and O'Shea, C.C. (2017). ChromEMT: Visualizing 3D chromatin structure and compaction in interphase and mitotic cells. *Science* 357.
- Oudet, P., Gross-Bellard, M., and Chambon, P. (1975). Electron microscopic and biochemical evidence that chromatin structure is a repeating unit. *Cell* 4, 281-300.
- Palovcak, E., Wang, F., Zheng, S.Q., Yu, Z., Li, S., Betegon, M., Bulkley, D., Agard, D.A., and Cheng, Y. (2018). A simple and robust procedure for preparing graphene-oxide cryo-EM grids. *Journal of structural biology* 204, 80-84.
- Pan, P.W., Feldman, J.L., Devries, M.K., Dong, A., Edwards, A.M., and Denu, J.M. (2011). Structure and biochemical functions of SIRT6. *J Biol Chem* 286, 14575-14587.
- Pantelic, R.S., Suk, J.W., Magnuson, C.W., Meyer, J.C., Wachsmuth, P., Kaiser, U., Ruoff, R.S., and Stahlberg, H. (2011). Graphene: Substrate preparation and introduction. *Journal of structural biology* 174, 234-238.
- Parenti, M.D., Grozio, A., Bauer, I., Galeno, L., Damonte, P., Millo, E., Sociali, G., Franceschi, C., Ballestrero, A., Bruzzone, S., *et al.* (2014). Discovery of novel and selective SIRT6 inhibitors. *Journal of medicinal chemistry* 57, 4796-4804.
- Peng, C., Lu, Z., Xie, Z., Cheng, Z., Chen, Y., Tan, M., Luo, H., Zhang, Y., He, W., Yang, K., *et al.* (2011). The first identification of lysine malonylation substrates and its regulatory enzyme. *Molecular & cellular proteomics : MCP* 10, M111 012658.
- Peralta-Leal, A., Rodriguez-Vargas, J.M., Aguilar-Quesada, R., Rodriguez, M.I., Linares, J.L., de Almodovar, M.R., and Oliver, F.J. (2009). PARP inhibitors: new partners in the therapy of cancer and inflammatory diseases. *Free radical biology & medicine* 47, 13-26.
- Perez-Montero, S., Carbonell, A., and Azorin, F. (2016). Germline-specific H1 variants: the "sexy" linker histones. *Chromosoma* 125, 1-13.
- Perisic, O., Collepardo-Guevara, R., and Schlick, T. (2010). Modeling studies of chromatin fiber structure as a function of DNA linker length. *Journal of molecular biology* 403, 777-802.
- Petrucchio, S., and Percudani, R. (2008). Structural recognition of DNA by poly(ADP-ribose)polymerase-like zinc finger families. *The FEBS journal* 275, 883-893.
- Petty, E., and Pillus, L. (2013). Balancing chromatin remodeling and histone modifications in transcription. *Trends in genetics : TIG* 29, 621-629.
- Poirier, G.G., de Murcia, G., Jongstra-Bilen, J., Niedergang, C., and Mandel, P. (1982a). Poly(ADP-ribosyl)ation of polynucleosomes causes relaxation of chromatin structure.

Proceedings of the National Academy of Sciences of the United States of America 79, 3423-3427.

Poirier, G.G., de Murcia, G., Jongstra-Bilen, J., Niedergang, C., and Mandel, P. (1982b). Poly(ADP-ribosyl)ation of polynucleosomes causes relaxation of chromatin structure. *Proc Natl Acad Sci U S A* 79, 3423-3427.

Pommier, Y., O'Connor, M.J., and de Bono, J. (2016). Laying a trap to kill cancer cells: PARP inhibitors and their mechanisms of action. *Science translational medicine* 8, 362ps317.

Pruss, D., Bartholomew, B., Persinger, J., Hayes, J., Arents, G., Moudrianakis, E.N., and Wolffe, A.P. (1996). An asymmetric model for the nucleosome: a binding site for linker histones inside the DNA gyres. *Science* 274, 614-617.

Pusey, M., Barcena, J., Morris, M., Singhal, A., Yuan, Q., and Ng, J. (2015). Trace fluorescent labeling for protein crystallization. *Acta crystallographica Section F, Structural biology communications* 71, 806-814.

Ramakrishnan, V., Finch, J.T., Graziano, V., Lee, P.L., and Sweet, R.M. (1993). Crystal structure of globular domain of histone H5 and its implications for nucleosome binding. *Nature* 362, 219-223.

Rao, S.S., Huntley, M.H., Durand, N.C., Stamenova, E.K., Bochkov, I.D., Robinson, J.T., Sanborn, A.L., Machol, I., Omer, A.D., Lander, E.S., *et al.* (2014). A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell* 159, 1665-1680.

Rao, S.S.P., Huang, S.C., Glenn St Hilaire, B., Engreitz, J.M., Perez, E.M., Kieffer-Kwon, K.R., Sanborn, A.L., Johnstone, S.E., Bascom, G.D., Bochkov, I.D., *et al.* (2017). Cohesin Loss Eliminates All Loop Domains. *Cell* 171, 305-320 e324.

Razin, S.V., and Gavrilov, A.A. (2014). Chromatin without the 30-nm fiber: constrained disorder instead of hierarchical folding. *Epigenetics* 9, 653-657.

Ricci, M.A., Manzo, C., Garcia-Parajo, M.F., Lakadamyali, M., and Cosma, M.P. (2015). Chromatin fibers are formed by heterogeneous groups of nucleosomes in vivo. *Cell* 160, 1145-1158.

Richmond, T.J., and Davey, C.A. (2003). The structure of DNA in the nucleosome core. *Nature* 423, 145-150.

Richmond, T.J., Finch, J.T., Rushton, B., Rhodes, D., and Klug, A. (1984). Structure of the nucleosome core particle at 7 Å resolution. *Nature* 311, 532-537.

Robinson, P.J., Fairall, L., Huynh, V.A., and Rhodes, D. (2006). EM measurements define the dimensions of the "30-nm" chromatin fiber: evidence for a compact, interdigitated structure. *Proceedings of the National Academy of Sciences of the United States of America* 103, 6506-6511.

- Robinson, P.J., and Rhodes, D. (2006). Structure of the '30 nm' chromatin fibre: a key role for the linker histone. *Current opinion in structural biology* *16*, 336-343.
- Roque, A., Sortino, R., Ventura, S., Ponte, I., and Suau, P. (2015). Histone H1 Favors Folding and Parallel Fibrillar Aggregation of the 1-42 Amyloid-beta Peptide. *Langmuir : the ACS journal of surfaces and colloids* *31*, 6782-6790.
- Rottenberg, S., Jaspers, J.E., Kersbergen, A., van der Burg, E., Nygren, A.O., Zander, S.A., Derksen, P.W., de Bruin, M., Zevenhoven, J., Lau, A., *et al.* (2008). High sensitivity of BRCA1-deficient mammary tumors to the PARP inhibitor AZD2281 alone and in combination with platinum drugs. *Proc Natl Acad Sci U S A* *105*, 17079-17084.
- Rouleau, M., Aubin, R.A., and Poirier, G.G. (2004). Poly(ADP-ribosyl)ated chromatin domains: access granted. *Journal of cell science* *117*, 815-825.
- Sanchez-Fidalgo, S., Villegas, I., Sanchez-Hidalgo, M., and de la Lastra, C.A. (2012). Sirtuin modulators: mechanisms and potential clinical implications. *Current medicinal chemistry* *19*, 2414-2441.
- Sarg, B., Lopez, R., Lindner, H., Ponte, I., Suau, P., and Roque, A. (2015). Identification of novel post-translational modifications in linker histones from chicken erythrocytes. *Journal of proteomics* *113*, 162-177.
- Sbodio, J.I., Lodish, H.F., and Chi, N.W. (2002). Tankyrase-2 oligomerizes with tankyrase-1 and binds to both TRF1 (telomere-repeat-binding factor 1) and IRAP (insulin-responsive aminopeptidase). *The Biochemical journal* *361*, 451-459.
- Schalch, T., Duda, S., Sargent, D.F., and Richmond, T.J. (2005). X-ray structure of a tetranucleosome and its implications for the chromatin fibre. *Nature* *436*, 138-141.
- Scheffer, M.P., Eltsov, M., Bednar, J., and Frangakis, A.S. (2012). Nucleosomes stacked with aligned dyad axes are found in native compact chromatin in vitro. *Journal of structural biology* *178*, 207-214.
- Schett, G., Smole, J., Zimmermann, C., Hiesberger, H., Hoefler, E., Fournel, S., Muller, S., Rubin, R.L., and Steiner, G. (2002). The autoimmune response to chromatin antigens in systemic lupus erythematosus: autoantibodies against histone H1 are a highly specific marker for SLE associated with increased disease activity. *Lupus* *11*, 704-715.
- Schreiber, V., Dantzer, F., Ame, J.C., and de Murcia, G. (2006). Poly(ADP-ribose): novel functions for an old molecule. *Nature reviews Molecular cell biology* *7*, 517-528.
- Shaytan, A.K., Armeev, G.A., Goncarencu, A., Zhurkin, V.B., Landsman, D., and Panchenko, A.R. (2016). Coupling between Histone Conformations and DNA Geometry in Nucleosomes on a Microsecond Timescale: Atomistic Insights into Nucleosome Functions. *J Mol Biol* *428*, 221-237.

Shilatifard, A. (2006). Chromatin modifications by methylation and ubiquitination: implications in the regulation of gene expression. *Annual review of biochemistry* 75, 243-269.

Shimizu, S., Nomura, F., Tomonaga, T., Sunaga, M., Noda, M., Ebara, M., and Saisho, H. (2004). Expression of poly(ADP-ribose) polymerase in human hepatocellular carcinoma and analysis of biopsy specimens obtained under sonographic guidance. *Oncology reports* 12, 821-825.

Simon, M., North, J.A., Shimko, J.C., Forties, R.A., Ferdinand, M.B., Manohar, M., Zhang, M., Fishel, R., Ottesen, J.J., and Poirier, M.G. (2011). Histone fold modifications control nucleosome unwrapping and disassembly. *Proceedings of the National Academy of Sciences of the United States of America* 108, 12711-12716.

Simpson, R.T. (1978). Structure of the chromatosome, a chromatin particle containing 160 base pairs of DNA and all the histones. *Biochemistry* 17, 5524-5531.

Sjoblom, T., Jones, S., Wood, L.D., Parsons, D.W., Lin, J., Barber, T.D., Mandelker, D., Leary, R.J., Ptak, J., Silliman, N., *et al.* (2006). The consensus coding sequences of human breast and colorectal cancers. *Science* 314, 268-274.

Skaper, S.D. (2003). Poly(ADP-Ribose) polymerase-1 in acute neuronal death and inflammation: a strategy for neuroprotection. *Annals of the New York Academy of Sciences* 993, 217-228; discussion 287-218.

Soldani, C., Lazze, M.C., Bottone, M.G., Tognon, G., Biggiogera, M., Pellicciari, C.E., and Scovassi, A.I. (2001). Poly(ADP-ribose) polymerase cleavage during apoptosis: when and where? *Experimental cell research* 269, 193-201.

Song, F., Chen, P., Sun, D., Wang, M., Dong, L., Liang, D., Xu, R.M., Zhu, P., and Li, G. (2014). Cryo-EM study of the chromatin fiber reveals a double helix twisted by tetranucleosomal units. *Science* 344, 376-380.

Sterner, D.E., and Berger, S.L. (2000). Acetylation of histones and transcription-related factors. *Microbiology and molecular biology reviews* : MMBR 64, 435-459.

Struhl, K., and Segal, E. (2013). Determinants of nucleosome positioning. *Nature structural & molecular biology* 20, 267-273.

Studier, F.W. (2005). Protein production by auto-induction in high density shaking cultures. *Protein expression and purification* 41, 207-234.

Sultana, A., and Lee, J.E. (2015). Measuring protein-protein and protein-nucleic Acid interactions by biolayer interferometry. *Current protocols in protein science* 79, 19 25 11-26.

Sultanov, D.C., Gerasimova, N.S., Kudryashova, K.S., Maluchenko, N.V., Kotova, E.Y., Langelier, M.F., Pascal, J.M., Kirpichnikov, M.P., Feofanov, A.V., and Studitsky, V.M. (2017). Unfolding of core nucleosomes by PARP-1 revealed by spFRET microscopy. *AIMS genetics* 4, 21-31.

Swindall, A.F., Stanley, J.A., and Yang, E.S. (2013). PARP-1: Friend or Foe of DNA Damage and Repair in Tumorigenesis? *Cancers* 5, 943-958.

Syed, S.H., Goutte-Gattat, D., Becker, N., Meyer, S., Shukla, M.S., Hayes, J.J., Everaers, R., Angelov, D., Bednar, J., and Dimitrov, S. (2010). Single-base resolution mapping of H1-nucleosome interactions and 3D organization of the nucleosome. *Proceedings of the National Academy of Sciences of the United States of America* 107, 9620-9625.

Szabo, G., Liaudet, L., Hagl, S., and Szabo, C. (2004). Poly(ADP-ribose) polymerase activation in the reperfused myocardium. *Cardiovascular research* 61, 471-480.

Talasz, H., Sarg, B., and Lindner, H.H. (2009). Site-specifically phosphorylated forms of H1.5 and H1.2 localized at distinct regions of the nucleus are related to different processes during the cell cycle. *Chromosoma* 118, 693-709.

Talbert, P.B., and Henikoff, S. (2010). Histone variants--ancient wrap artists of the epigenome. *Nature reviews Molecular cell biology* 11, 264-275.

Tallis, M., Morra, R., Barkauskaite, E., and Ahel, I. (2014). Poly(ADP-ribosyl)ation in regulation of chromatin structure and the DNA damage response. *Chromosoma* 123, 79-90.

Tan, S., and Davey, C.A. (2011). Nucleosome structural studies. *Current opinion in structural biology* 21, 128-136.

Tanaka, M., Kihara, M., Mieczekalski, B., King, G.J., and Adashi, E.Y. (2003). H1oo: a pre-embryonic H1 linker histone in search of a function. *Molecular and cellular endocrinology* 202, 5-9.

Tennen, R.I., Berber, E., and Chua, K.F. (2010). Functional dissection of SIRT6: identification of domains that regulate histone deacetylase activity and chromatin localization. *Mechanisms of ageing and development* 131, 185-192.

Th'ng, J.P., Sung, R., Ye, M., and Hendzel, M.J. (2005). H1 family histones in the nucleus. Control of binding and localization by the C-terminal domain. *The Journal of biological chemistry* 280, 27809-27814.

Thraves, P., Mossman, K.L., Brennan, T., and Dritschilo, A. (1985). Radiosensitization of human fibroblasts by 3-aminobenzamide: an inhibitor of poly(ADP-ribosylation). *Radiation research* 104, 119-127.

Tremethick, D.J. (2007). Higher-order structures of chromatin: the elusive 30 nm fiber. *Cell* 128, 651-654.

Trojer, P., Li, G., Sims, R.J., 3rd, Vaquero, A., Kalakonda, N., Boccuni, P., Lee, D., Erdjument-Bromage, H., Tempst, P., Nimer, S.D., *et al.* (2007). L3MBTL1, a histone-methylation-dependent chromatin lock. *Cell* 129, 915-928.

- Tsompana, M., and Buck, M.J. (2014). Chromatin accessibility: a window into the genome. *Epigenetics Chromatin* 7, 33.
- van Emmerik, C.L., and van Ingen, H. (2019). Unspinning chromatin: Revealing the dynamic nucleosome landscape by NMR. *Progress in nuclear magnetic resonance spectroscopy* 110, 1-19.
- van Holde, K., and Zlatanova, J. (2007). Chromatin fiber structure: Where is the problem now? *Seminars in cell & developmental biology* 18, 651-658.
- Vignali, M., and Workman, J.L. (1998). Location and function of linker histones. *Nat Struct Biol* 5, 1025-1028.
- Villalba, J.M., and Alcain, F.J. (2012). Sirtuin activators and inhibitors. *BioFactors* 38, 349-359.
- Villar-Garea, A., and Imhof, A. (2008). Fine mapping of posttranslational modifications of the linker histone H1 from *Drosophila melanogaster*. *PloS one* 3, e1553.
- Virag, L., and Szabo, C. (2002). The therapeutic potential of poly(ADP-ribose) polymerase inhibitors. *Pharmacological reviews* 54, 375-429.
- Voss, T.C., and Hager, G.L. (2014). Dynamic regulation of transcriptional states by chromatin and transcription factors. *Nat Rev Genet* 15, 69-81.
- Vyas, S., Matic, I., Uchima, L., Rood, J., Zaja, R., Hay, R.T., Ahel, I., and Chang, P. (2014). Family-wide analysis of poly(ADP-ribose) polymerase activity. *Nature communications* 5, 4426.
- Wacker, D.A., Ruhl, D.D., Balagamwala, E.H., Hope, K.M., Zhang, T., and Kraus, W.L. (2007). The DNA binding and catalytic domains of poly(ADP-ribose) polymerase 1 cooperate in the regulation of chromatin structure and transcription. *Molecular and cellular biology* 27, 7475-7485.
- Wang, Y., An, R., Umanah, G.K., Park, H., Nambiar, K., Eacker, S.M., Kim, B., Bao, L., Harraz, M.M., Chang, C., *et al.* (2016). A nuclease that mediates cell death induced by DNA damage and poly(ADP-ribose) polymerase-1. *Science* 354.
- Wang, Y., Kim, N.S., Haince, J.F., Kang, H.C., David, K.K., Andrabi, S.A., Poirier, G.G., Dawson, V.L., and Dawson, T.M. (2011). Poly(ADP-ribose) (PAR) binding to apoptosis-inducing factor is critical for PAR polymerase-1-dependent cell death (parthanatos). *Science signaling* 4, ra20.
- Wang, Z.Q., Stingl, L., Morrison, C., Jantsch, M., Los, M., Schulze-Osthoff, K., and Wagner, E.F. (1997). PARP is important for genomic stability but dispensable in apoptosis. *Genes & development* 11, 2347-2358.
- Watson, J.D., and Crick, F.H. (2003). A structure for deoxyribose nucleic acid. 1953. *Nature* 421, 397-398; discussion 396.

Weber, C.M., and Henikoff, S. (2014). Histone variants: dynamic punctuation in transcription. *Genes & development* 28, 672-682.

Widom, J. (1992). A relationship between the helical twist of DNA and the ordered positioning of nucleosomes in all eukaryotic cells. *Proceedings of the National Academy of Sciences of the United States of America* 89, 1095-1099.

Widom, J. (1998). Structure, dynamics, and function of chromatin in vitro. *Annual review of biophysics and biomolecular structure* 27, 285-327.

Widom, J. (2001). Role of DNA sequence in nucleosome stability and dynamics. *Quarterly reviews of biophysics* 34, 269-324.

Wilkins, M.H., Stokes, A.R., and Wilson, H.R. (1953). Molecular structure of deoxypentose nucleic acids. *Nature* 171, 738-740.

Winn, M.D., Ballard, C.C., Cowtan, K.D., Dodson, E.J., Emsley, P., Evans, P.R., Keegan, R.M., Krissinel, E.B., Leslie, A.G., McCoy, A., *et al.* (2011). Overview of the CCP4 suite and current developments. *Acta crystallographica Section D, Biological crystallography* 67, 235-242.

Winogradoff, D., and Aksimentiev, A. (2019). Molecular Mechanism of Spontaneous Nucleosome Unraveling. *J Mol Biol* 431, 323-335.

Wisniewski, J.R., Zougman, A., Kruger, S., and Mann, M. (2007). Mass spectrometric mapping of linker histone H1 variants reveals multiple acetylations, methylations, and phosphorylation as well as differences between cell culture and tissue. *Molecular & cellular proteomics : MCP* 6, 72-87.

Wisniewski, J.R., Zougman, A., and Mann, M. (2008). Nepsilon-formylation of lysine is a widespread post-translational modification of nuclear proteins occurring at residues involved in regulation of chromatin function. *Nucleic acids research* 36, 570-577.

Woodcock, C.L. (1994). Chromatin fibers observed in situ in frozen hydrated sections. Native fiber diameter is not correlated with nucleosome repeat length. *The Journal of cell biology* 125, 11-19.

Woodcock, C.L., and Dimitrov, S. (2001). Higher-order structure of chromatin and chromosomes. *Current opinion in genetics & development* 11, 130-135.

Woodcock, C.L., and Ghosh, R.P. (2010). Chromatin higher-order structure and dynamics. *Cold Spring Harbor perspectives in biology* 2, a000596.

Woodcock, C.L., Safer, J.P., and Stanchfield, J.E. (1976). Structural repeating units in chromatin. I. Evidence for their general occurrence. *Experimental cell research* 97, 101-110.

Woodcock, C.L., Skoultschi, A.I., and Fan, Y. (2006). Role of linker histone in chromatin structure and function: H1 stoichiometry and nucleosome repeat length. *Chromosome research* :

an international journal on the molecular, supramolecular and evolutionary aspects of chromosome biology *14*, 17-25.

Woodhouse, B.C., and Dianov, G.L. (2008). Poly ADP-ribose polymerase-1: an international molecule of mystery. *DNA repair* *7*, 1077-1086.

Yang, D., Fang, Q., Wang, M., Ren, R., Wang, H., He, M., Sun, Y., Yang, N., and Xu, R.M. (2013a). Nalpa-acetylated Sir3 stabilizes the conformation of a nucleosome-binding loop in the BAH domain. *Nat Struct Mol Biol* *20*, 1116-1118.

Yang, S.M., Kim, B.J., Norwood Toro, L., and Skoultchi, A.I. (2013b). H1 linker histone promotes epigenetic silencing by regulating both DNA methylation and histone H3 methylation. *Proceedings of the National Academy of Sciences of the United States of America* *110*, 1708-1713.

Yang, W. (2006). Poor base stacking at DNA lesions may initiate recognition by many repair proteins. *DNA repair* *5*, 654-666.

Ye, X., Feng, C., Gao, T., Mu, G., Zhu, W., and Yang, Y. (2017). Linker Histone in Diseases. *International journal of biological sciences* *13*, 1008-1018.

You, W., Rotili, D., Li, T.M., Kambach, C., Meleshin, M., Schutkowski, M., Chua, K.F., Mai, A., and Steegborn, C. (2016). Structural Basis of Sirtuin 6 Activation by Synthetic Small Molecules. *Angewandte Chemie*.

Yu, S.W., Wang, H., Poitras, M.F., Coombs, C., Bowers, W.J., Federoff, H.J., Poirier, G.G., Dawson, T.M., and Dawson, V.L. (2002). Mediation of poly(ADP-ribose) polymerase-1-dependent cell death by apoptosis-inducing factor. *Science* *297*, 259-263.

Zaremba, T., and Curtin, N.J. (2007). PARP inhibitor development for systemic cancer targeting. *Anti-cancer agents in medicinal chemistry* *7*, 515-523.

Zhang, Y., and Reinberg, D. (2001). Transcription regulation by histone methylation: interplay between different covalent modifications of the core histone tails. *Genes & development* *15*, 2343-2360.

Zhang, Z., Hildebrandt, E.F., Simbulan-Rosenthal, C.M., and Anderson, M.G. (2002). Sequence-specific binding of poly(ADP-ribose) polymerase-1 to the human T cell leukemia virus type-I tax responsive element. *Virology* *296*, 107-116.

Zhong, L., D'Urso, A., Toiber, D., Sebastian, C., Henry, R.E., Vadysirisack, D.D., Guimaraes, A., Marinelli, B., Wikstrom, J.D., Nir, T., *et al.* (2010). The histone deacetylase Sirt6 regulates glucose homeostasis via Hif1alpha. *Cell* *140*, 280-293.

Zhou, B.R., Feng, H., Kato, H., Dai, L., Yang, Y., Zhou, Y., and Bai, Y. (2013). Structural insights into the histone H1-nucleosome complex. *Proceedings of the National Academy of Sciences of the United States of America* *110*, 19390-19395.

- Zhou, B.R., Jiang, J., Feng, H., Ghirlando, R., Xiao, T.S., and Bai, Y. (2015). Structural Mechanisms of Nucleosome Recognition by Linker Histones. *Molecular cell* 59, 628-638.
- Zhou, Y.B., Gerchman, S.E., Ramakrishnan, V., Travers, A., and Muyldermans, S. (1998). Position and orientation of the globular domain of linker histone H5 on the nucleosome. *Nature* 395, 402-405.
- Zhu, F., Farnung, L., Kaasinen, E., Sahu, B., Yin, Y., Wei, B., Dodonova, S.O., Nitta, K.R., Morgunova, E., Taipale, M., *et al.* (2018). The interaction landscape between transcription factors and the nucleosome. *Nature* 562, 76-81.
- Zhu, P., and Li, G. (2016). Higher-order structure of the 30-nm chromatin fiber revealed by cryo-EM. *IUBMB life* 68, 873-878.
- Zlatanova, J., Caiafa, P., and Van Holde, K. (2000). Linker histone binding and displacement: versatile mechanism for transcriptional regulation. *FASEB journal : official publication of the Federation of American Societies for Experimental Biology* 14, 1697-1704.
- Zlatanova, J., and Doenecke, D. (1994). Histone H1 zero: a major player in cell differentiation? *FASEB journal : official publication of the Federation of American Societies for Experimental Biology* 8, 1260-1268.
- Zuin, J., Dixon, J.R., van der Reijden, M.I., Ye, Z., Kolovos, P., Brouwer, R.W., van de Corput, M.P., van de Werken, H.J., Knoch, T.A., van, I.W.F., *et al.* (2014). Cohesin and CTCF differentially affect chromatin architecture and gene expression in human cells. *Proceedings of the National Academy of Sciences of the United States of America* 111, 996-1001.

Appendix I – DNA sequences

Nucleosome constructs

1. 601L (145 bp)

70 60 50 40 30 20 10 0 10 20 30 40 50 60 70

ATCA**CAAT**C**CCGGT****GCCGA**GGCCG**CTCAA**TTGGT**CGTAG**ACAGC**TCTAG**CACCG**TTAAA**CGCAC**GTACG**GAATC**CGTAC**GTGCG**TTTAA**GCGGT**GCTAG**AGTGT**CTACG**ACCAAT**TTGAG**CGGCC**TCGGC**ACCGG**GATTG**TGAT

2. 601* (147 bp)

70 60 50 40 30 20 10 0 10 20 30 40 50 60 70
+ + + + + + + + + + + + + + + + +
atcGAGAATCCCGGTGCCGAGGCCGCTCAATTGGTCGTAGACAGCTCTAGACCGCTTAAACGCACGTACGCGCTGTCCCCGCGTTTTAACCGCCAAGGGGATTACTCCCTAGTCTCCAGGCACGTGTCTCAGATATATACATCCgat

3. CMS-1m (154 bp)

70 60 50 40 30 20 10 0 10 20 30 40 50 60 70
 ---+---+---+---+---+---+---+---O---+---+---+---+---+---+---+---
 ATCA**CAATC**CCGGT**GCCGA**GGCCGCT CAAATTGGT**CGTAG**ACAGCT**TCTAG**CACCGCT**TAAAA**CGCAC**GTAGG**GAATG**TTTGT**TCTTA**TTTAA**GCGCAC**CTAGAG**CTCG**CTACT**CGCAT**TCT AC**GATCC**GCAAG**GGATA**TTTGG**AGAAA
 AAAAAAGAT

4. CMS-169s

GCTTTTTTTTTTTTCACAATCCCGGTGCCGAGGCCGCTCAATTGGTCGTAGACAGCTCTAGCACCGCTTAAACGCACGTACGGAATCCGTACGTGCGTTTAAAGCGGTGCTAGAGCTGTCTACGACCAATTGAGCGGCCCTCGGCACCGGGATTGT
 GAAAAAAAAAAGCTGCA

5. CMS-169sn

GGTAC | CCAAAAAAAAAACAGCATCCCGGTGCCGAGGCCGCTCAATTGGTCGTAGACAGCTCTAGCACCGCTTAAACGCACGTACGCGTGCTTACCGCGTTTAAACCGCCACTAGAAGCGCTTACTAGTCTCCAGGCACGTGTGAGACCGG
CACATGCAAAAAAAAAACGAGCT | C

di-Nucleosome constructs

1. diNUC- 343 bp,
KpnI site on both ends (3' overhang)
AvaI site on shared linker

```

      3      13      23      33      43      53      63      73      83      93      103      113      123      133      143      153      163
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
GgctgaaaaaaaaacGCATCCCGGTGCCGAGGCCGCTCAATTGGTCGTAGACAGCTCTAGCACCGCTTAAACGCACGTACGCGCTGTCTACCGCGTTTAAACCGCCACTAGAAAGCGCTTACTAGTCTCCAGGCACGTGTGAGACCGGCACATgaaaaaaaaaagcatgCTCGA
      177      187      197      207      217      227      237      247      257      267      277      287      297      307      317      327      337
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
GtatgaaaaaaaaaacGCATCCCGGTGCCGAGGCCGCTCAATTGGTCGTAGACAGCTCTAGCACCGCTTAAACGCACGTACGCGCTGTCTACCGCGTTTAAACCGCCACTAGAAAGCGCTTACTAGTCTCCAGGCACGTGTGAGACCGGCACATgaaaaaaacagcGGTAC

```

2. diNUC- 345 bp,

```

      4      14      24      34      44      54      64      74      84      94      104      114      124      134      144      154      164
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
CgctgaaaaaaaaaacGCATCCCGGTGCCGAGGCCGCTCAATTGGTCGTAGACAGCTCTAGCACCGCTTAAACGCACGTACGCGCTGTCTACCGCGTTTAAACCGCCACTAGAAAGCGCTTACTAGTCTCCAGGCACGTGTGAGACCGGCACATgaaaaaaaaaagcatgCTCGA
      178      188      198      208      218      228      238      248      258      268      278      288      298      308      318      328      338
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
GtatgaaaaaaaaaacGCATCCCGGTGCCGAGGCCGCTCAATTGGTCGTAGACAGCTCTAGCACCGCTTAAACGCACGTACGCGCTGTCTACCGCGTTTAAACCGCCACTAGAAAGCGCTTACTAGTCTCCAGGCACGTGTGAGACCGGCACATgaaaaaaacagcGGTAC

```

3. diNUC- 347 bp,

```

      5      15      25      35      45      55      65      75      85      95      105      115      125      135      145      155      165
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
-
CgctgaaaaaaaaaacGCATCCCGGTGCCGAGGCCGCTCAATTGGTCGTAGACAGCTCTAGCACCGCTTAAACGCACGTACGCGCTGTCTACCGCGTTTAAACCGCCACTAGAAAGCGCTTACTAGTCTCCAGGCACGTGTGAGACCGGCACATgaaaaaaaaaagcatgCTCG
A
      179      189      199      209      219      229      2239      249      259      269      279      289      299      309      319      329      339
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
GtatgaaaaaaaaaacGCATCCCGGTGCCGAGGCCGCTCAATTGGTCGTAGACAGCTCTAGCACCGCTTAAACGCACGTACGCGCTGTCTACCGCGTTTAAACCGCCACTAGAAAGCGCTTACTAGTCTCCAGGCACGTGTGAGACCGGCACATgaaaaaaacagcGGTAC

```

4. diNUC- 349 bp,

6	16	26	36	46	56	66	76	86	96	106	116	126	136	146	156	166
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----																
Cgctg <u>g</u> aaaaaaaaaacGCATCCCGGTGCCGAGGCCGCTCAATTGGTCGTAGACAGCTCTAGCACCGCTTAAACGCACGTACGCGTGCTACCGCGTTTAAACCGCCACTAGAAAGCGCTTACTAGTCTCCAGGCACGTGTGAGACCGGCACATgaaaaaaaaaagcatgCTCGA																
180	190	200	210	220	230	240	250	260	270	280	290	300	310	320	330	340
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----																
GtatgaaaaaaaaaacGCATCCCGGTGCCGAGGCCGCTCAATTGGTCGTAGACAGCTCTAGCACCGCTTAAACGCACGTACGCGTGCTACCGCGTTTAAACCGCCACTAGAAAGCGCTTACTAGTCTCCAGGCACGTGTGAGACCGGCACATgaaaaaaaaaacgagcGGTAC																

5. diNUC- 351 bp,

7	17	27	37	47	57	67	77	87	97	107	117	127	137	147	157	167
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----																
Cgctg <u>g</u> aaaaaaaaaacGCATCCCGGTGCCGAGGCCGCTCAATTGGTCGTAGACAGCTCTAGCACCGCTTAAACGCACGTACGCGTGCTACCGCGTTTAAACCGCCACTAGAAAGCGCTTACTAGTCTCCAGGCACGTGTGAGACCGGCACATgaaaaaaaaaagcatgCTCGA																
181	191	201	211	221	231	241	251	261	271	281	291	301	311	321	331	341
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----																
GtatgaaaaaaaaaacGCATCCCGGTGCCGAGGCCGCTCAATTGGTCGTAGACAGCTCTAGCACCGCTTAAACGCACGTACGCGTGCTACCGCGTTTAAACCGCCACTAGAAAGCGCTTACTAGTCTCCAGGCACGTGTGAGACCGGCACATgaaaaaaaaaacgagcGGTAC																

6. diNUC- 353 bp,

8	18	28	38	48	58	68	78	88	98	108	118	128	138	148	158	168
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----																
Cgctg <u>g</u> aaaaaaaaaacGCATCCCGGTGCCGAGGCCGCTCAATTGGTCGTAGACAGCTCTAGCACCGCTTAAACGCACGTACGCGTGCTACCGCGTTTAAACCGCCACTAGAAAGCGCTTACTAGTCTCCAGGCACGTGTGAGACCGGCACATgaaaaaaaaaagcatgCTCGA																
182	192	202	212	222	232	242	252	262	272	282	292	302	312	322	332	342
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----																
GtatgaaaaaaaaaacGCATCCCGGTGCCGAGGCCGCTCAATTGGTCGTAGACAGCTCTAGCACCGCTTAAACGCACGTACGCGTGCTACCGCGTTTAAACCGCCACTAGAAAGCGCTTACTAGTCTCCAGGCACGTGTGAGACCGGCACATgaaaaaaaaaacgtcagcGGTAC																

7. diNUC- 351z (349 + 2)

6	16	26	36	46	56	66	76	86	96	106	116	126	136	146	156	166
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----																
CgctggaaaaaaaaaacGCATCCCGGTGCCGAGGCCGCTCAATTGGTCGTAGACAGCTCTAGCACC GCCTTAAACGCACGTACGCGCTGCTACCGCGTTTAAACGCCACTAGAAAGCGCTTACTAGTCTCCAGGCACGTGTGAGACCGGCACATgaaaaaaaaaatgcaTgCTCGA																
182	192	202	212	222	232	242	252	262	272	282	292	302	312	322	332	342
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----																
GtatgaaaaaaaaaTcGCATCCCGGTGCCGAGGCCGCTCAATTGGTCGTAGACAGCTCTAGCACC GCCTTAAACGCACGTACGCGCTGCTACCGCGTTTAAACGCCACTAGAAAGCGCTTACTAGTCTCCAGGCACGTGTGAGACCGGCACATgaaaaaaaaaacgcgGTAC																

8. diNUC- 353z (351 + 2)

7	17	27	37	47	57	67	77	87	97	107	117	127	137	147	157	167
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----																
CgctgcgaaaaaaaaaacGCATCCCGGTGCCGAGGCCGCTCAATTGGTCGTAGACAGCTCTAGCACC GCCTTAAACGCACGTACGCGCTGCTACCGCGTTTAAACGCCACTAGAAAGCGCTTACTAGTCTCCAGGCACGTGTGAGACCGGCACATgaaaaaaaaaTgcatgCTCGA																
183	193	203	213	223	233	243	253	263	273	283	293	303	313	323	333	343
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----																
GtatgaaaaaaaaaTcGCATCCCGGTGCCGAGGCCGCTCAATTGGTCGTAGACAGCTCTAGCACC GCCTTAAACGCACGTACGCGCTGCTACCGCGTTTAAACGCCACTAGAAAGCGCTTACTAGTCTCCAGGCACGTGTGAGACCGGCACATgaaaaaaaaaacgcagcGGTAC																

9. diNUC- 355z (353 + 2)

8	18	28	38	48	58	68	78	88	98	108	118	128	138	148	158	168
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----																
CgctgacgaaaaaaaaaacGCATCCCGGTGCCGAGGCCGCTCAATTGGTCGTAGACAGCTCTAGCACC GCCTTAAACGCACGTACGCGCTGCTACCGCGTTTAAACGCCACTAGAAAGCGCTTACTAGTCTCCAGGCACGTGTGAGACCGGCACATgaaaaaaaaaTgcatgCTCGA																
184	194	204	214	224	234	244	254	264	274	284	294	304	314	324	334	344
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----																
GtatgaaaaaaaaaTcGCATCCCGGTGCCGAGGCCGCTCAATTGGTCGTAGACAGCTCTAGCACC GCCTTAAACGCACGTACGCGCTGCTACCGCGTTTAAACGCCACTAGAAAGCGCTTACTAGTCTCCAGGCACGTGTGAGACCGGCACATgaaaaaaaaaacgtcagcGGTAC																

10. diNUC- 357z (355 + 2)

9	19	29	39	49	59	69	79	89	99	109	119	129	139	149	159	169	
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----																	
CgctgcacgaaaaaaaaacGCATCCCGGTGCCGAGGCCGCTCAATTGGTCGTAGACAGCTCTAGCACCAGCTTAAACGCACGTACGCGCTGTCTACCGCGTTTAAACCGCCACTAGAAAGCGCTTACTAGTCTCCAGGCACGTGTGAGACCGGCACATgaaaaaaaaaatgcatgCTCGA																	
185	195	205	215	225	235	245	255	265	275	285	295	305	315	325	335	345	355
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----																	
GtatgaaaaaaaaaatcGCATCCCGGTGCCGAGGCCGCTCAATTGGTCGTAGACAGCTCTAGCACCAGCTTAAACGCACGTACGCGCTGTCTACCGCGTTTAAACCGCCACTAGAAAGCGCTTACTAGTCTCCAGGCACGTGTGAGACCGGCACATgaaaaaaaaaacgtccagcGGTAC																	

Appendix II – Protein sequences

H2A

GSSHHHHHHSSGLVPRGSH
MSGRGKQGKARAKAKTRSSRAGLQFPVGRVHLLRKGNYSERVGAGAPVYLAHVLEYLTAEILELAGNAARDNKKTRIIPRHLQLAIRNDEELNKLGRVTIAQGGVLPNIQAV
LLPKKTESHKAKGK

H2B

GSSHHHHHHSSGLVPRGSH
MPEPAKSAPAPKKGSKKAVTKAQKKDGKKRKRSRKESYSIYVYKVLKQVHPDTGISSKAMGIMNSFVNDIFERIAGEASRLAHYNKRSTITSREIQTAVRLLLPGELAKHAVSEG
KAVTKYTSK

H3

GSSHHHHHHSSGLVPRGSH
MARTKQTARKSTGGKAPRKQLATKAARKSAPATGGVKKPHRYRPGTVALREIRRYQKSTELLIRKLPFQRLVREIAQDFKTDLRFQSSAVMALQEACEAYLVGLFEDTNLCIHA
KRVTIMPKDIQLARRIRGERA

H4

GSSHHHHHHSSGLVPRGSH
MSGRGKGGKGLGKGGAKRHRKVLRDNIQGITKPAIRRLARRGGVKRISGLIYEETRGVLKVFLENVIRDAVITYTEHAKRKTVTAMDVVYALKRQGRTLYGFGG

PARP1

MGSSHHHHHHSSGLVPRGSH
MAESSDKLYRVEYAKSGRASCKKCESEIPKDSLRLMAIMVQSPMFDGKVPWHYHFSCFWKVGHSIRHPDVEVDGFSELRWDDQQKVKKTAEEAGGVTGKGQDGIGSKAE
KTLGDFAAEYAKSNRSTCKGCMKIEKGQVRLSKMVDPEKPQLGMIDRWYHPGCFVKNREELGFRPEYSASQLKGFSLLATEDKEALKKQLPGVKSEGKRKGDEVD
GVDEVAKKKSKKEKDKSKLEKALKAQNDLIWNIKDELKKVCSTNDLKELLIFNKQQVPSGESAILDRVADGMVFGALLPCEECGQLVFKSDAYYCTGDVTAWTKC
MVKTQTPNRKEWVTPKEFREISYLKKLVKKQDRIFPPETSASVAATPPPSTASAPAAVNSSASADKPLSNMKILTGLKLSRNKDEVKAMIEKLGGKLTGTANKASL
CISTKKEVEKMNMKMEVKEANIRVVSDFLQDVSASTKSLQELFLAHILSPWGAEVKAEPVEVVAPRGKSGAALSKKSKGQVKEEGINKSEKRMKLTGGAAVDP
DSGLEHSAHVLEKGGKVSATLGLVDIVKGTNSYYKLQLLEDDKENRYWIFRSWGRVGTVIGSNKLEQMPSKEDAEHFMKLYEEKTGNAWHSKNFTKYPKKFYPLE
IDYGQDEEAVKKLTVNPGTKSKLPKPVQDLIKMIFDVESMKKAMVEYEIDLQKMPLGKLSKRQIQAAYSILSEVQQAVSQGSSDSQILDLSNRFYTLIPHDFGMKKP
PLLNNADSVQAKVEMLDNLLDIEVAYSLLRGGSDSSKDPIDVNYEKLKTDIKVVDRDSEAEIIRKYVKNTHATTHNAYDLEVIDIFKIEREGECQRYKPFKQLHN
RRLWLHGSRTTNFAGILSQGLRIAPPEAPVTGYMFGKGIYFADMVSKSANYCHTSQGDPIGLILLGEVALGNMYELKHASHISKLPKGKHSVKGLGKTTDPDSANIS
LDGVDVPLGTGISSGVNDTSLLYNEYIVYDIAQVNLKYLKLFNFKTSLW

SIRT6

MGSSHHHHHHSSGLVPRGSH

MSVNYAAGLSPYADKGKCGLPEIFDPPEELERKVWELARLVWQSSSVVFHTGAGISTASGIPDFRGPBGVWTMEERGLAPKFDTTTFESARPTQTHMALVQLERVGLL
RFLVSQNVDGLHVRSGFPRDKLAELHGNMFVEECAKCKTQYVRDTPVGTMLKATGRLCTVAKARGLRACRGELRDTILDWEDSLPDRDLALADEASRNADLSITLG
TSLQIRPSGNLPLATKRRGGRLVIVNLQPTKHDRHADLRIHGYVDEVMTRLMKHLGLEIPAWDGPRVLERALPPLPRPPTPKLEPKESPTRINGSIPAGPKQEP
CAQHNGSEPASPKRERPTSPAPHRPPKRVKAKAVPS

H1.0

MHHHHHHSSGLEVLFGP

TENSTSAPAAKPKRAKASKKSTDHPKYSDMIVAAIQAEKNRAGSSRQSIQYIKSHYKVGGENADSQIKLSIKRLVTTGVLKQTKGVGASGSFRLAKSDEPKKSVAFK
KTKKEIKKVATPKKASKPKKAASKAPTCKPKATPVKKAKKKLAATPKKAKKPKTVKAKPVKASKPKKAKPVKPAKSSAKRAGKKK

H1.2

MGMHHHHHHSSGVDLGATENLYFQ

SSETAPAAPAAAPPAEKAPVKKKAACKAGGTTPRKASGPPVSELITKAVAASKERSGVSLAALKKALAAAGYDVEKNNSRIKLGLKSLVSKGTLVQTKGTGASGSFKL
NKKAASGEAKPKVKKAGGTTPKKPVGAACKPKKAAGGATPKKSAKKTPKKAKKPAAATVTKKVAKSPKKAKVAKPKKAASAAKAVKPKAAKPKVVKPKKAAPKKK

H1.3

MHHHHHHSSGVDLGATENLYFQS

MSETAPLAPTIPAPAEKTPVKKKAKKAGATAGKRKASGPPVSELITKAVAASKERSGVSLAALKKALAAAGYDVEKNNSRIKLGLKSLVSKGTLVQTKGTGASGSFK
LNKKAASGEGKPKAKKAGAAKPRKPAGAAKPKKVAGAATPKKSICKTPKKVKKPATAAGTKKVAKSAAKVKTPQPKKAASPAKAKAPKPKAAKPKSGKPKVTKAK
KAAPKKK

H1.5

MHHHHHHSSGVDLGATENLYFQ

SSETAPAEATATPAPVEKSPAKKKATKKAAGAGAAKPKATGPPVSELITKAVAASKERNGLSLAALKKALAAGGYDVEKNNSRIKLGLKSLVSKGTLVQTKGTGASGS
FKLNKKAASGEAKPKAKKAGAAKAPAGATPKKAKKAAGAKKAVKKTPKKAKKPAAAGVKKVAKSPKKAKAAAKPKKATKSPAKPKAVKPKAAKPKAAKPKAAKPK
AAKAKKAAAKK

Appendix III– Protein-Nucleosome crystallization screen conditions

	1	2	3	4	5	6	7	8	9	10	11	12
A	100mM NaAc pH5.0, 10% PEG400	100mM NaAc pH5.0, 200mM NaCl, 10% PEG400	100mM NaAc pH5.45, 10% PEG400	100mM NaAc pH5.45, 200mM NaCl, 10% PEG400	100mM KCaco pH5.57, 10% PEG400	100mM KCaco pH5.57, 200mM NaCl, 10% PEG400	100mM KCaco pH6.68, 10% PEG400	100mM KCaco pH6.68, 200mM NaCl, 10% PEG400	100mM MES pH6.79, 10% PEG400	100mM MES pH6.79, 200mM NaCl, 10% PEG400	100mM MES pH7.49, 10% PEG400	100mM MES pH7.49, 200mM NaCl, 10% PEG400
B	100mM NaAc pH5.0, 20% PEG400	100mM NaAc pH5.0, 200mM NaCl, 20% PEG400	100mM NaAc pH5.45, 20% PEG400	100mM NaAc pH5.45, 200mM NaCl, 20% PEG400	100mM KCaco pH5.57, 20% PEG400	100mM KCaco pH5.57, 200mM NaCl, 20% PEG400	100mM KCaco pH6.68, 20% PEG400	100mM KCaco pH6.68, 200mM NaCl, 20% PEG400	100mM MES pH6.79, 20% PEG400	100mM MES pH6.79, 200mM NaCl, 20% PEG400	100mM MES pH7.49, 20% PEG400	100mM MES pH7.49, 200mM NaCl, 20% PEG400
C	100mM NaAc pH5.0, 30% PEG400	100mM NaAc pH5.0, 200mM NaCl, 30% PEG400	100mM NaAc pH5.45, 30% PEG400	100mM NaAc pH5.45, 200mM NaCl, 30% PEG400	100mM KCaco pH5.57, 30% PEG400	100mM KCaco pH5.57, 200mM NaCl, 30% PEG400	100mM KCaco pH6.68, 30% PEG400	100mM KCaco pH6.68, 200mM NaCl, 30% PEG400	100mM MES pH6.79, 30% PEG400	100mM MES pH6.79, 200mM NaCl, 30% PEG400	100mM MES pH7.49, 30% PEG400	100mM MES pH7.49, 200mM NaCl, 30% PEG400
D	100mM NaAc pH5.0, 40% PEG400	100mM NaAc pH5.0, 200mM NaCl, 40% PEG400	100mM NaAc pH5.45, 40% PEG400	100mM NaAc pH5.45, 200mM NaCl, 40% PEG400	100mM KCaco pH5.57, 40% PEG400	100mM KCaco pH5.57, 200mM NaCl, 40% PEG400	100mM KCaco pH6.68, 40% PEG400	100mM KCaco pH6.68, 200mM NaCl, 40% PEG400	100mM MES pH6.79, 40% PEG400	100mM MES pH6.79, 200mM NaCl, 40% PEG400	100mM MES pH7.49, 40% PEG400	100mM MES pH7.49, 200mM NaCl, 40% PEG400
E	100mM NaAc pH5.0, 10% MPD	100mM NaAc pH5.0, 200mM NaCl, 10% MPD	100mM NaAc pH5.45, 10% MPD	100mM NaAc pH5.45, 200mM NaCl, 10% MPD	100mM KCaco pH5.57, 10% MPD	100mM KCaco pH5.57, 200mM NaCl, 10% MPD	100mM KCaco pH6.68, 10% MPD	100mM KCaco pH6.68, 200mM NaCl, 10% MPD	100mM MES pH6.79, 10% MPD	100mM MES pH6.79, 200mM NaCl, 10% MPD	100mM MES pH7.49, 10% MPD	100mM MES pH7.49, 200mM NaCl, 10% MPD
F	100mM NaAc pH5.0, 20% MPD	100mM NaAc pH5.0, 200mM NaCl, 20% MPD	100mM NaAc pH5.45, 20% MPD	100mM NaAc pH5.45, 200mM NaCl, 20% MPD	100mM KCaco pH5.57, 20% MPD	100mM KCaco pH5.57, 200mM NaCl, 20% MPD	100mM KCaco pH6.68, 20% MPD	100mM KCaco pH6.68, 200mM NaCl, 20% MPD	100mM MES pH6.79, 20% MPD	100mM MES pH6.79, 200mM NaCl, 20% MPD	100mM MES pH7.49, 20% MPD	100mM MES pH7.49, 200mM NaCl, 20% MPD
G	100mM NaAc pH5.0, 30% MPD	100mM NaAc pH5.0, 200mM NaCl, 30% MPD	100mM NaAc pH5.45, 30% MPD	100mM NaAc pH5.45, 200mM NaCl, 30% MPD	100mM KCaco pH5.57, 30% MPD	100mM KCaco pH5.57, 200mM NaCl, 30% MPD	100mM KCaco pH6.68, 30% MPD	100mM KCaco pH6.68, 200mM NaCl, 30% MPD	100mM MES pH6.79, 30% MPD	100mM MES pH6.79, 200mM NaCl, 30% MPD	100mM MES pH7.49, 30% MPD	100mM MES pH7.49, 200mM NaCl, 30% MPD
H	100mM NaAc pH5.0, 40% MPD	100mM NaAc pH5.0, 200mM NaCl, 40% MPD	100mM NaAc pH5.45, 40% MPD	100mM NaAc pH5.45, 200mM NaCl, 40% MPD	100mM KCaco pH5.57, 40% MPD	100mM KCaco pH5.57, 200mM NaCl, 40% MPD	100mM KCaco pH6.68, 40% MPD	100mM KCaco pH6.68, 200mM NaCl, 40% MPD	100mM MES pH6.79, 40% MPD	100mM MES pH6.79, 200mM NaCl, 40% MPD	100mM MES pH7.49, 40% MPD	100mM MES pH7.49, 200mM NaCl, 40% MPD