



A novel H2A.Z deposition pathway involving the Ku70/Ku80 heterodimer

Arslan Iftikhar

► To cite this version:

Arslan Iftikhar. A novel H2A.Z deposition pathway involving the Ku70/Ku80 heterodimer. Cellular Biology. Université de Strasbourg, 2019. English. NNT : 2019STRAJ055 . tel-03183995

HAL Id: tel-03183995

<https://theses.hal.science/tel-03183995>

Submitted on 29 Mar 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

ÉCOLE DOCTORALE DES SCIENCES DE LA VIE ET DE LA SANTÉ

Institut de Génétique et de Biologie Moléculaire et Cellulaire (IGBMC)

CNRS UMR 7104 - Inserm U 1258

THÈSE présentée par :

Arslan IFTIKHAR

Soutenue publiquement le **27 Septembre 2019**

pour obtenir le grade de : **Docteur de l'Université de Strasbourg**

Discipline : **Sciences de la Vie et de la Santé**

Spécialité : **Aspects Moléculaires et Cellulaires de la Biologie**

**Une nouvelle voie de déposition de
H2A.Z impliquant l'hétérodimère
Ku70/Ku80**

THÈSE dirigée par :

Dr. Ali HAMICHE

Directeur de Recherches, IGBMC, Université de Strasbourg.

RAPPORTEURS :

Dr. Dimitar ANGUELOV

Directeur de Recherches, École Normale Supérieure de Lyon

Dr. Regina GROISMAN

Charge de Recherches, Institut Gustave Roussy, Villejuif.

EXAMINATEUR INTERNE :

Dr. Frédéric COIN

Directeur de Recherches, IGBMC, Université de Strasbourg.

TABLE OF CONTENTS

ACKNOWLEDGEMENT.....	i
LIST OF ABBREVIATION	ii
LIST OF FIGURES	1
LIST OF TABLES	3
RESUME EN FRANÇAIS	4
SUMMARY IN ENGLISH	11

CHAPTER 1: INTRODUCTION

1.1	Chromatin	15
1.1.1	Nucleosome: The Basic Unit of Chromatin	15
1.1.2	Chromatin Organization	16
1.1.3	Histone Proteins	18
1.1.3.1	Core Histones	19
1.1.3.2	Linker Histones	21
1.2	Epigenetic Regulation	23
1.2.1	DNA Methylation	25
1.2.2	Regulatory Non-coding RNAs	27
1.2.3	Histone Post-Translational Modifications (PTMs)	28
1.2.4	Histone Variants	30
1.2.4.1	Histone H1 variants	32
1.2.4.2	Histone H2A variants	34
1.2.4.2.1	H2A.X variant	37
1.2.4.2.2	H2A.Z variant	38
1.2.4.2.3	Macro H2A variant	41
1.2.4.2.4	H2A.Bbd variant	42
1.2.4.3	Histone H2B variants	43
1.2.4.4	Histone H3 variants	44
1.2.4.5	Histone H4 variants	46
1.3	Chromatin Remodeling and H2A.Z Deposition	48
1.3.1	Histone replacement/exchange	48
1.3.1.1	Remodeling of histone variant H2A.Z at DSBs	49
1.3.1.2	Remodeling of histone variant H3.3	52
1.4	The Ku 70/80 Heterodimer	53
1.4.1	Discovery and History.....	53
1.4.2	Biochemistry and Structure	54
1.4.2.1	Structural Domains of Ku70/Ku80:	56
1.4.3	Functions of Ku Heterodimer.....	59
1.4.3.1	Ku heterodimer in DNA Repair	60

1.4.3.2	Role in V(D)J Recombination.....	63
1.4.3.3	Ku heterodimer in Apoptosis.....	64
1.4.3.4	Ku heterodimer in Telomere Maintenance.....	64
1.4.3.5	Ku heterodimer in Transcription.....	65
CHAPTER 2: OBJECTIVES		67
CHAPTER 3: MATERIALS AND METHODS		69
3.1	General Reagents	69
3.2	Cloning Strategies	71
3.2.1	Designing DNA Oligos	71
3.2.2	Polymerase Chain Reaction	71
3.2.3	Restriction Cloning	72
3.2.4	Sanger DNA Sequencing	73
3.2.5	Expression Vectors	73
3.3	Expression Procedures	75
3.4	Protein Purification	77
3.5	Protein Identification & Characterization.....	78
3.5.1	SDS-PAGE Gel Electrophoresis	78
3.5.2	Protein Gel Staining	79
3.5.3	Western Blotting	79
3.5.4	Mass Spectrometry	81
3.6	Cell Lines	81
3.6.1	HeLa	81
3.6.2	Mouse Embryonic Fibroblast (MEF)	82
3.6.3	Immunofluorescence (IF) Assay	83
3.6.4	Cryopreservation of Stable Mammalian Cell Lines	83
3.7	ChIP sequencing (ChIP-Seq) Assay	83
3.8	Ion Exchange/Gel Filtration Chromatography.....	85
3.9	Structural Modelling	86
CHAPTER 4: RESULTS AND DISCUSSION		87
4.1	Identification of Ku70/Ku80 in H2A.Z Predeposition Complex	87
4.2	Exploration of Ku70/Ku80 and H2A.Z Interaction in HeLa Cells	88
4.3	Ku Heterodimer Ku70/Ku80 Specifically Interacts with Histone Dimer H2A.Z/H2B <i>in-vitro</i>	90
4.4	von Willebrand factor type A (vWA) Domain of Ku80 is Critical for H2A.Z/H2B Interaction	93
4.5	The Specificity of Interaction of von Willebrand Factor Type A (vWA) Domain of Ku80 to H2A.Z/H2B	98

4.6	Ku80 (40-101) Motif is The Minimal Subdomain Responsible for Specific Interaction with Histone Dimer H2A.Z/H2B	99
4.6.1	Sequence Conservation of Ku80 (40-101) Motif Responsible for H2A.Z/H2B Specific Interaction	101
4.7	Identification of H2A.Z Motif Critical for Ku Heterodimer Ku70/Ku80 Interaction	102
4.7.1	Sequence Conservation of H2A.Z Motif in Various Eukaryotic Species	104
4.8	The Interaction of Ku80 (40-101) Motif with Histone Dimer H2A.Z/H2B is Mediated Favorably Through H2B	106
4.9	Evaluation of Interaction of Ku80 (40-101) Motif with Histone Dimer H2A.Z/H2B Using Size-Exclusion Chromatography (SEC)	108
4.9.1	Purification of Ku80 (40-101) and Histone Dimer H2A.Z/H2B Through Affinity and Ion-Exchange Chromatography Approaches	109
4.9.2	Gel Filtration Chromatography of Ku80 (40-101) Motif and Histone Dimer H2A.Z/H2B Complex Indicated the Favored Interaction of Ku Motif with H2B	110
4.10	Exploration of Interaction of Ku Heterodimer Ku70/Ku80 and Histone Dimer H2A.Z/H2B in Mouse Embryonic Fibroblasts (MEFs)	111
4.11	Ku70/Ku80 Heterodimer Regulates H2A.Z Deposition at Promoters	113
4.12	Structural Modeling	116
CHAPTER 5: CONCLUSION AND PERSPECTIVES		118
CHAPTER 6: REFERENCES		123
CHAPTER 6: APPENDIX		152

ACKNOWLEDGEMENT

Words are bound and expression is limited to praise the Almighty, the entire source of knowledge and wisdom, who bestowed on me His blessings and gave me the courage to complete this dissertation. Undertaking this PhD has been a truly life-changing experience for me and it would not have been possible to accomplish without the kind support and guidance of my family members, my parents, my wife, my brother. It is their moral encouragement that has always been the motivation and strength for me to set higher targets in life.

It's a perfect time to express my gratitude to my supervisor **Dr. Ali Hamiche** for giving me the opportunity to pursue my PhD in his group. It would really not have been possible for me to undergo my PhD in such a great professional environment as IGBMC without his generous consent. His contributions of time, ideas, knowledge and intellect contributed immensely to my personal and professional grooming. I am also very thankful to **Dr. Dimitar Anguelov, Dr. Regina Groisman** and **Dr. Frédéric Coin**, not only for accepting to be the member of my dissertation committee but also for valuable comments on technical aspects of the thesis.

On a very special note, I would like to thank **Christophe Papin** for providing all the technical insights to the project. With a special mention to all laboratory fellows **Christian Bronner, Philippe Romain, Catherine Romain, Isabelle Stoll, Abdulkhaleg Ibrahim, Hatem Salem, Abdul Kareem, Tajith Shaik, Maria Ignatyeva, Alexia Vidal, Naif Almalki** and **Dimitra Vlachokosta**, thanks a lot for making sense to me of various practical challenges faced and in providing encouragement at times when it looked stuck to me. It really was a great experience to share laboratory with all of you during the last four years.

I also express my special thanks to all my Pakistani friends in France **Faisal, Tanveer, Asad, Salman, Bilal, Imtiaz, Akmal, Wahid, Zahid, Waseem, Khalid, Maaz, Abubakr, Usman, Wasif, Irshad, Irfan, Hira, Sara, Fareeha** and others who made my stay in France a bouquet of lovely memory.

I gratefully acknowledge the funding received towards my PhD from **Higher Education Commission (HEC), Pakistan** and **Government College University Faisalabad, Pakistan**. I also pay my heartiest thanks to the staff of **IGBMC, University of Strasbourg, Campus France, CROUS** and other organizations for all their administrative services during my whole PhD journey.

Arslan Iftikhar

LIST OF ABBREVIATION

%: Percentage

µg: Microgram

µL: Microliter

ANP32E: Acidic Nuclear Phosphoprotein 32 family member E

APS: Ammonium Persulfate

BSA: Bovine Serum Albumin

CDK: Cyclin Dependent Kinase

cDNA: Complementary DNA

CENP-A: Centromere Protein A

ChIP-Seq: Chromatin Immunoprecipitation Sequencing

CTD: C-terminal Domain

Da: Dalton

DAPI: 4',6-Diamidino-2-Phenylindole

DLS: Dynamic Light Scattering

DNA: Deoxyribonucleic Acid

DNMTs: DNA Methyltransferases

DMEM: Dulbecco's Modified Eagle's Medium

dNTPs: Deoxynucleotides Tri-Phosphate

DSB: Double Strand Break

dsDNA: Double Stranded DNA

DTT: 1,4-Dithiothreitol

EDTA: Ethylenediaminetetraacetic Acid

FBS: Foetal Bovine Serum

GST: Glutathione S-transferase

HCl: Hydrochloric Acid

HR: Homologous Recombination

IP: Immunoprecipitation

IPTG: Isopropyl-β-D-1-Thiogalactopyranoside

MOPS: Morpholinepropanesulfonic Acid

NaCl: Sodium Chloride
NHEJ: Non-Homologous End Joining
NTD: N-terminal domain
NuRD: Nucleosome Remodeling Deacetylase
OD: Optical Density
PBS: Phosphate-Buffered Saline
PCR: Polymerase Chain Reaction
PDB: Protein Data Bank
pH: potential of Hydrogen
PMSF: Phenylmethylsulfonyl Fluoride
PTM: Post Translational Modification
RPM: Revolutions Per Minute
SDS-PAGE: Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis
ssDNA: Single-Stranded DNA
SUMO: Small Ubiquitin-like Modifier
TEMED: Tetramethylethylenediamine
TIP60: Tat Interactive Protein 60 kDa
wt: Wildtype

LIST OF FIGURES

Figure 1: Overview of Chromatin Organization.....	17
Figure 2: Schematic Assembly of Histone Octamer.	19
Figure 3: Structural Aspects of Core Histone Proteins.	20
Figure 4: DNA Damage Generate Specific Epigenetic Modifications.	24
Figure 5: Epigenetic Regulations.	25
Figure 6: Methylation of Cytosine to 5-Methylcytosine by The Action of DNMT Enzyme.	26
Figure 7: Central Dogma Indicating Role of Non-Coding RNAs.	28
Figure 8: Histone Post-Translational Modifications (PTMs).	30
Figure 9: Sequence Alignment of Major Histone H2A Variants.	35
Figure 10: Indispensable Role of Phosphorylated H2A.X (γ H2A.X) in DSB Repair Mechanism.	38
Figure 11: Multiple Alignment of Histone Variant H2A.Z from Various Eukaryotic Species.	39
Figure 12: Multiple Alignment of Various Human H2A.Z Isoforms and Sites of Various PTMs.	40
Figure 13: Sequence Alignment of Human H3 Variants.	45
Figure 14: The Mechanism of Histone Variant H2A.Z Replacement.	49
Figure 15: Replacement of H2A.Z with H2A at DSBs Contribute in H4 Acetylation.	50
Figure 16: H2A.Z Exchange Initiate a Systematic Series of Histone Modification.....	51
Figure 17: The Overview of Crystal Structure of Human Ku Heterodimer.	54
Figure 18: Secondary Structure Alignment of Subunits of Ku Heterodimer.	55
Figure 19: Schematic illustration of Ku70 and Ku80 Primary Structure.....	57
Figure 20: The Cellular Functions of Ku Heterodimer.	59
Figure 21: Types of Various DNA Damages.	60
Figure 22: Schematic Overview of Different Stages of NHEJ Repair Pathway.	62
Figure 23: Schematic Illustration of pET Expression System.	76
Figure 24: Comparison of Polypeptides Detected by Mass Spectrometry Analysis of H2A and H2A.Z.	87
Figure 25: Validation of Stable Expression of FLAG/HA Tagged wt-H2A.Z in H2A.Z Knockout HeLa Cells Through Immunofluorescence and Immunoblotting.	88
Figure 26: Ku70 & Ku80 are Associated in H2A.Z Complex.	89
Figure 27: Mass Spectrometry Analysis of H2A.Z Subcomplexes.	90
Figure 28: Ku70/Ku80 Heterodimer Specifically Interact with Histone Dimer H2A.Z/H2B <i>in- vitro</i>	92
Figure 29: Domain Characterization of Ku70/Ku80 Heterodimer.	93
Figure 30: Histone Dimer H2A.Z/H2B Exclusively Interacts with vWA Domain of Ku80 ...	95
Figure 31: Exploration of Specific Interaction of vWA Domain of Ku80 and H2A.Z/H2B Using Co-expression System.	96
Figure 32: Specificity of vWA Domain of Ku80 to H2A.Z/H2B.	99

Figure 33: Exploration of Minimal Subdomain of vWA Domain of Ku80 Responsible for Exclusive Interaction with Histone Dimer H2A.Z/H2B.	100
Figure 34: Sequence Alignment of Ku80 (40-101) Subdomain Responsible for Interaction with Histone Dimer H2A.Z/H2B.	102
Figure 35: Domain Mapping of H2A.Z Responsible for Interaction with Ku Heterodimer. .	104
Figure 36: Multiple Alignment of Histone Variant H2A.Z from Various Eukaryotic Species.	105
Figure 37: Ku80 (40-101) Domain Interacts Explicitly with H2B <i>in-vitro</i>	107
Figure 38: Interaction of Flag-tagged Ku80 (40- 101) Subdomain with His-tagged H2B <i>in-vitro</i>	108
Figure 39: Ion Exchange Chromatograms of Ku80 (40-101) and Histone Dimer H2A.Z/H2B	109
Figure 40: Purification Profile of Ku80 (40-101) and Histone Dimer H2A.Z/H2B.....	111
Figure 41: Validation of Stable Expression of Δ NKu80 and wtKu80 in Ku80 ^{-/-} MEFs Through Immunofluorescence and Immunoblotting.	112
Figure 42: Mass Spectrometry Analysis of Δ NKu80 and wtKu80 Subcomplexes.	113
Figure 43: Ku80 is Involved in Genome-wide Localization and Chromatin Enrichment of H2A.Z at Promoters.	115
Figure 44: Structural Modelling.	117

LIST OF TABLES

Table 1: Variants of linker histone H1 in mammalian cells:	33
Table 2: List of General Reagents Used.	69
Table 3: Thermocycling Conditions for Polymerase Chain Reaction.	71
Table 4: Components and Quantity of Polymerase Chain Reaction Mixture.	72
Table 5: Plasmid Constructs Used for Gene Expression.	73
Table 6: Commercial Resins Used for Immunoprecipitation.....	77
Table 7: Composition of Resolving and Stacking Gel with Variable Fraction of Acrylamide.	79
Table 8: List of Primary and Secondary Antibodies Used During Research.	80
Table 9: Composition of Washing Buffers Pertaining for Blocking Agarose.	84

RESUME EN FRANÇAIS

INTRODUCTION :

L'ADN est constamment exposé à de nombreux agents génotoxiques exogènes et endogènes pouvant entraîner différents types de dommages génomiques. Parmi ces dommages, les cassures doubles brins (DSB) forment les lésions les plus dangereuses, car elles provoquent un clivage physique du squelette de l'ADN (Downs *et al.*, 2004; Kusch *et al.*, 2004; Jha *et al.* 2008; Jackson and Bartek, 2009). La réparation de ces événements potentiellement mutagènes est très importante pour le maintien de l'intégrité du génome et d'un fonctionnement cellulaire normale. Les cellules mammifères ont développées deux types distincts de mécanismes de réparation des cassures d'ADN doubles brins : (i) La jonction d'extrémité non homologue sujette à erreur (NHEJ) qui prédomine dans la phase G1 du cycle cellulaire et (ii) la recombinaison homologue sans erreur (HR) présente pendant les phases S et G2 du cycle cellulaire.

L'hétérodimère Ku70/80 hautement conservé, principalement localisé dans le noyau, est bien connu pour son rôle dans la stabilité génomique, la signalisation cellulaire, la maintenance des télomères, la réplication de l'ADN et l'activation de la transcription (Deriano and Roth, 2013; Frit *et al.*, 2014). L'hétérodimère Ku est considéré comme le premier membre de la machinerie NHEJ à être recruté au niveau des cassures double brin, où il sert ensuite d'échafaudage moléculaire pour recruter d'autres membres du NHEJ au niveau des sites de cassures DSB (Grundy *et al.*, 2014, Williams *et al.*, 2014). Au cours des dernières années, il y a eu énormément de spéculations sur le processus de recrutement des hétérodimères Ku sur le site des DSB, mais il n'existe pratiquement aucune étude moléculaire dans les annales de la littérature sur ce processus. Par ailleurs, les lésions DSB affecte un ADN empaqueté dans une structure condensé appelée chromatine (Blier *et al.*, 1993; Mari *et al.*, 2006; Costantini *et al.*, 2007; Hammel *et al.*, 2010).

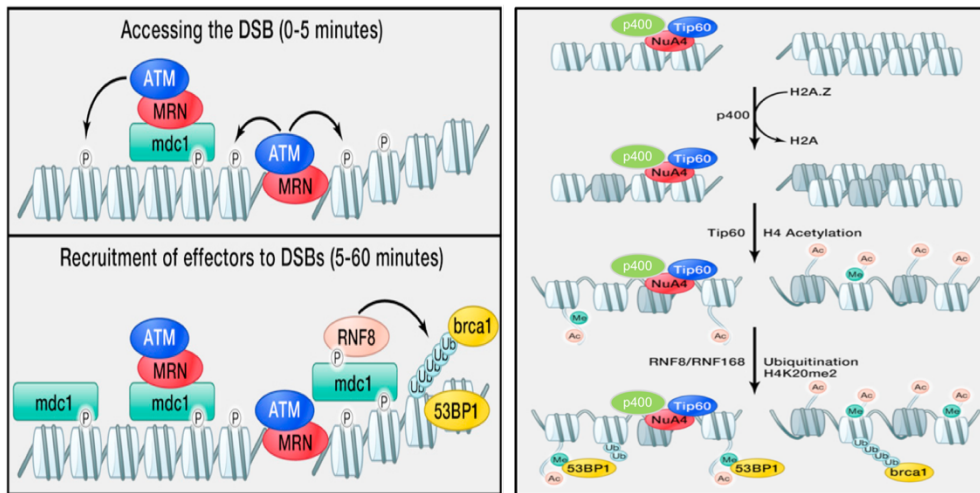


Figure A: Rôle de H2A.Z dans la réparation de l'ADN (Price and D'Andrea, 2013)

La chromatine est une structure hautement dynamique dont l'unité de base répétitive est le nucléosome. Un moyen important de réguler la dynamique de la chromatine consiste à remplacer les histones canoniques par des histones variantes. Bien que ces dernières ressemblent beaucoup à leurs homologues, les histones variantes entraînent des modifications considérables des fonctions et des processus de la chromatine (Vignali *et al.*, 2000; Clapier and Cairns, 2009). Dans le mécanisme de réparation de l'ADN, le phénomène de remodelage de la chromatine est considéré comme un processus essentiel permettant de recruter la machinerie de réparation de l'ADN au niveau des sites de cassures de l'ADN. La variante d'histone H2A.Z, fortement conservée au cours de l'évolution, possède une homologie d'environ 60% avec l'histone canonique H2A (Kennedy and D'Andrea, 2006; Jackson and Bartek, 2009; Ciccina and Elledge, 2010).

H2A.Z joue un rôle clé dans de nombreux processus cellulaires tels que l'expression et la répression génique, la ségrégation chromosomique et la réparation de l'ADN. L'incorporation de H2A.Z au niveau la chromatine est assurée par les complexes de remodelage de la chromatine p400/TIP60 et SRCAP. Ces complexes augmentent la mobilité thermique des nucléosomes au niveau de la chromatine. Cette mobilité accrue des nucléosomes contenant H2A.Z modifie les interactions histone-histone et histone/ADN à la fois dans les nucléosomes adjacents et voisins et aboutit à une conformation spécifique de la chromatine nécessaire à l'acétylation cruciale de la queue

N-terminale de l'histone H4 par l'histone acetyltransferase Tip60. Cette acétylation de l'histone H4 conduit à une ouverture de la chromatine et à son exposition à l'action de l'ubiquitine ligase RNF8 qui modifie des sites cryptiques sur le nucléosome. Cette ubiquitination permet le recrutement et la stabilisation de la protéine de réparation de l'ADN BRCA1 qui dirige le système de réparation vers la voie NHEJ en recrutant l'hétérodimère Ku (Murr *et al.*, 2006; Xu *et al.*, 2010; Courilleau *et al.*, 2012). De cette manière, les modifications initiales de la fonction des nucléosomes impliquées par le remodelage de H2A.Z jouent un rôle essentiel dans l'orientation du recrutement ordonné des protéines de réparation du DSB sur le site des lésions de l'ADN. En accord avec les études réalisées chez les vertébrés, l'éviction de H2A.Z de la chromatine chez la levure conduit à une augmentation de l'instabilité génomique, à une plus grande sensibilité aux agents endommageant l'ADN et à une réparation défectueuse des DSB. De même, chez la Drosophile, l'acétylation du seul variant de H2A, H2AV, chimère de H2A.X et H2A.Z, joue un rôle essentiel dans la réparation de l'ADN.

Notre travail sur le rôle de H2A.Z dans la réparation de l'ADN a débuté lorsque nous avons identifié par analyse protéomique l'association de l'hétérodimère Ku avec le complexe soluble H2A.Z. L'implication de H2A.Z et l'hétérodimère Ku dans les mécanismes de réparation de l'ADN nous a incité à étudier plus en détail cette interaction particulière et à explorer ses bases moléculaires en utilisant différentes approches biochimiques et génomiques.

RESULTATS et DISCUSSION :

Nous avons tout d'abord établi des lignées cellulaires HeLa stables exprimant les histones classiques H2A et H2A.Z fusionnées au niveau de leur extrémité N-terminale avec des étiquettes hémagglutinine (HA) et Flag. Nous avons ensuite isolé les complexes associés par double purification par immunoaffinité, puis nous les avons séparés sur un gel SDS-PAGE. L'analyse par spectrométrie de masse nous permis d'identifier Ku70 et Ku80 comme tant uniquement associées avec le complexe H2A.Z et pas le complexe H2A. L'analyse inverse par spectrométrie de masse des complexes

Ku70 et Ku80 nous a confirmé la présence de H2A.Z et l'absence de H2A dans ces complexes.

Pour valider cette interaction, nous avons co-exprimé individuellement les dimères d'histones H2A-H2B, H2A.X-H2B et H2A.Z-H2B avec l'hétérodimère Ku70-Ku80 en utilisant le système d'expression *Escherichia coli*. Cette approche nous a permis de confirmer que l'hétérodimère Ku70-Ku80 interagit presque exclusivement avec le dimère H2A.Z-H2B.

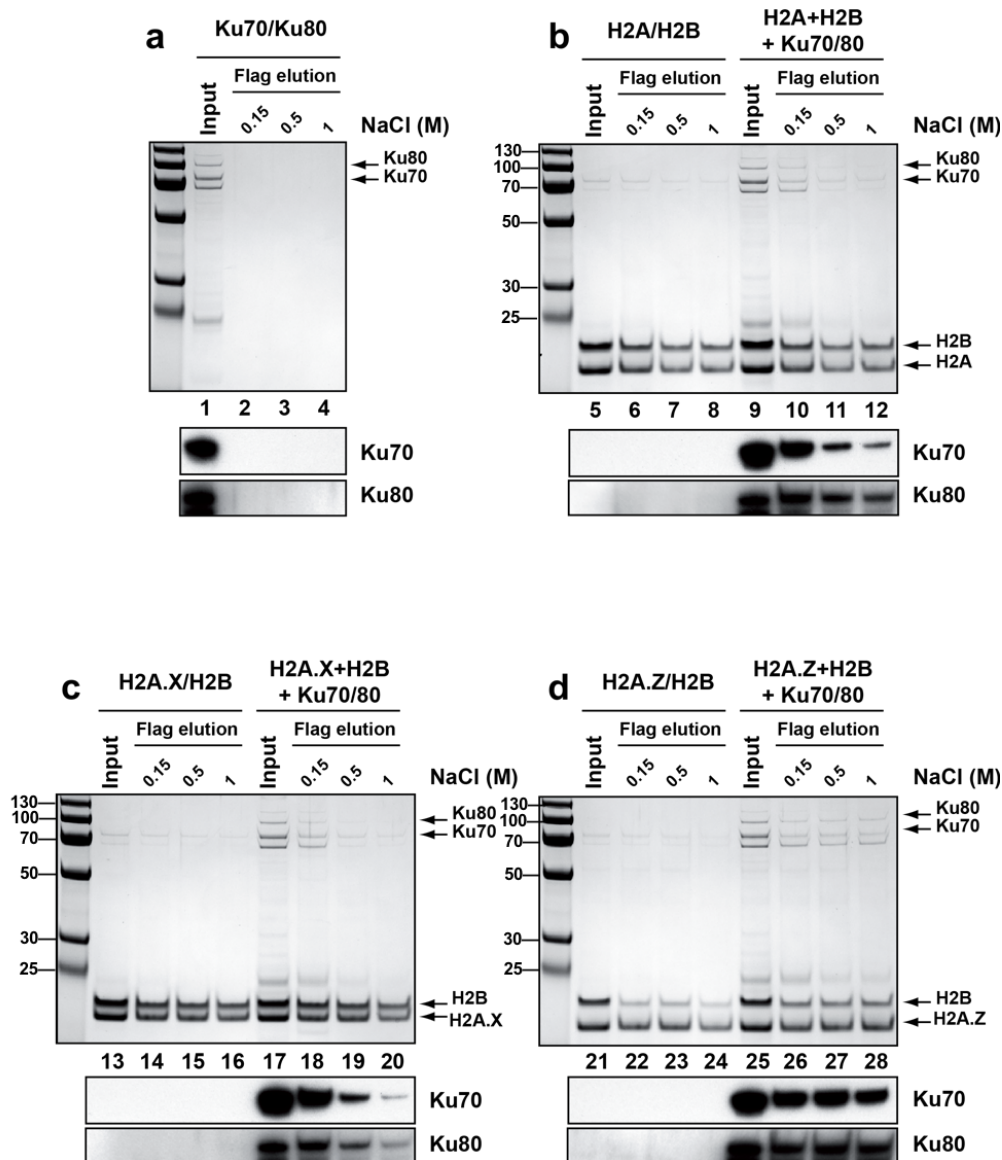


Figure D: L'hétérodimère Ku70 / Ku80 interagit spécifiquement avec le dimère d'histone H2A.Z / H2B *in vitro*

Nous avons ensuite disséqué ces interactions in vitro avec l'approché GST- Pull-down et montré que le domaine de Willebrand A (vWA) de Ku80 est spécifiquement responsable de l'interaction avec le dimère H2A.Z-H2B. Nous avons ensuite divisé ce domaine vWA de 26 kDa en trois sous-domaines et avons répété les expériences ci-dessus avec les mêmes spécifications. À travers cette expérience, nous avons identifié un sous- domaine de 7 kDa du domaine vWA exclusivement responsable de l'interaction avec les histones. Tous les résultats ont également été confirmés par une analyse Western-blot en utilisant des anticorps spécifiques. À chaque étape, nous avons évalué l'interaction spécifique des domaines exclusivement avec les dimères H2A.Z-H2B par rapport aux dimères H2A-H2B et H2A.X-H2B. Une analyse par alignement bioinformatique montre que ce domaine Ku de 7 kDa responsable de l'interaction avec les histones est hautement conservé au cours l'évolution.

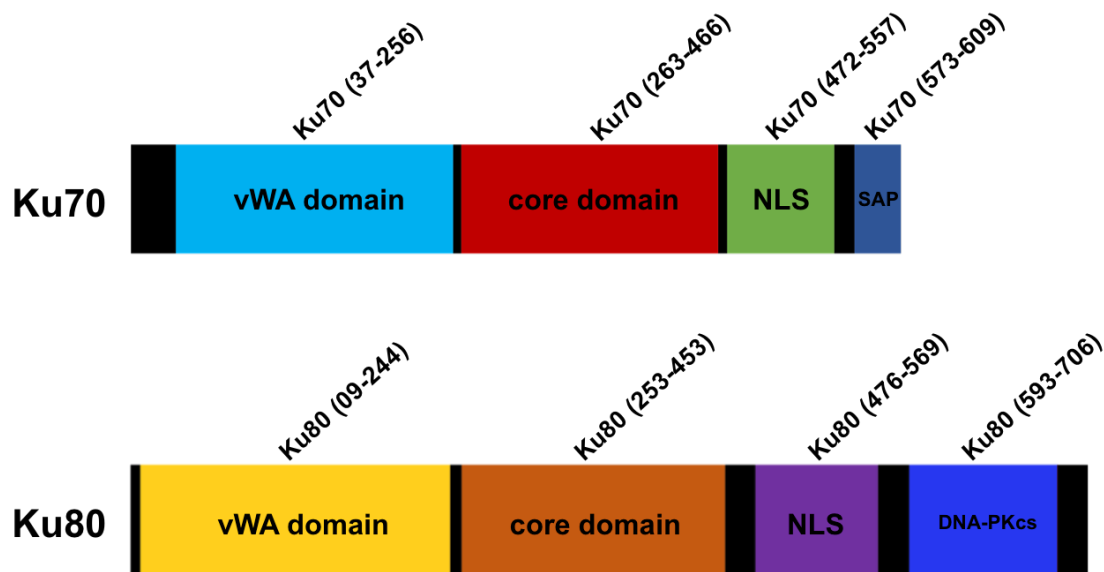


Figure C: Illustration schématique de la structure primaire Ku70 et Ku80.

Nous avons ensuite recherché le domaine histone responsable de l'interaction avec le domaine vWA de Ku80. Pour cela, nous avons créé différentes chimères de H2A.Z et H2A, conservant H2B comme second membre du dimère. Par l'approche de co-expression dans la bactérie et purification par affinité, nous avons identifié un

domaine H2A.Z de 4 kDa jouant un rôle important dans cette interaction. De manière intéressante, au cours des différentes purifications, nous avons observé la présence d'une quantité substantiellement importante de H2B par rapport à H2A.Z dans le dimère H2A.Z-H2B suggérant que la protéine Ku80 aurait une meilleure affinité pour H2B. Nous avons alors émis l'hypothèse que Ku se lie spécifiquement au dimère H2A.Z-H2B via également une interaction avec H2B. Pour valider cette hypothèse, nous avons co-exprimé le domaine 7kDa de Ku80 avec seulement l'histone H2B. Nos résultats montrent une interaction spécifique du domaine Ku 80 avec l'histone H2B.

Sur la base de nos observations biochimiques selon lesquelles une région de 7kDa de von Willebrand A (vWA) de Ku80 est responsable de la reconnaissance exclusive des dimères H2AZ-H2B, nous avons décidé de valider l'importance de cette interaction *in vivo* en utilisant la technique d'immunoprécipitation à la chromatine suivie d'un séquençage massif de l'ADN génomique (ChIP-seq). Nous avons émis l'hypothèse que l'inactivation de l'hétérodimère Ku entraînerait effectivement une réduction de H2A.Z au niveau de la chromatine. Pour étudier l'effet de Ku80 sur la déposition de H2A.Z, nous avons effectué une analyse comparative par ChIP-seq à l'échelle du génome en utilisant à la fois des fibroblastes embryonnaires de souris Ku^{-/-} (knockout) et Ku^{+/+} (sauvage). Nos données montrent que H2A.Z est spécifiquement associé aux promoteurs de gènes, aux Enhancers et aux Insulateurs. L'inactivation de Ku80 entraîne une déplétion spécifique de plus de 50% du H2A.Z localisé au niveau des promoteurs mais pas celui sur les Enhancers ou les Insulateurs. Les expériences de sauvetages du phénotype par restauration de l'expression de la protéine Ku80 avec un rétrovirus entraîne une restauration du niveau de H2A.Z au niveau des promoteurs de gènes. Le mutant Δ NKu80 (délété du domaine de liaison aux histones) n'est pas capable de restaurer le niveau de H2A.Z dans les cellules Ku80^{-/-}.

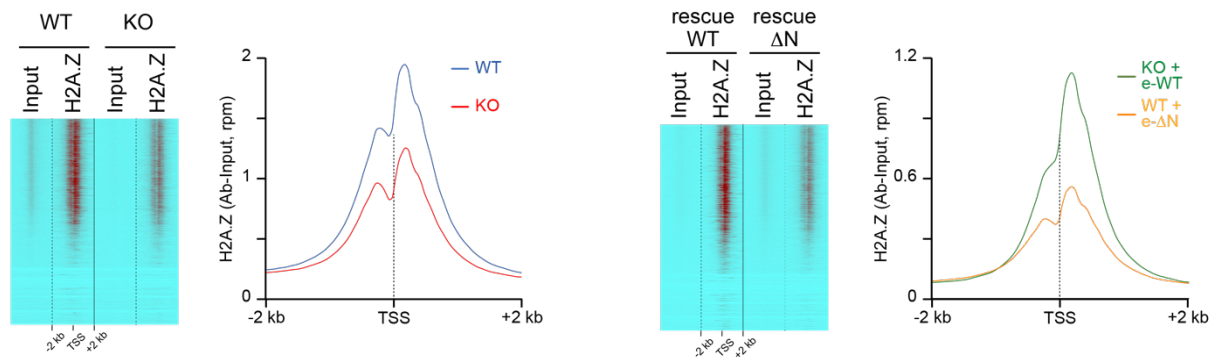


Figure D: Ku80 est impliqué dans le dépôt de H2A.Z chez les promoteurs

CONCLUSION et PERSPECTIVES :

En conclusion, nos données biochimiques, protéomiques et génomiques identifient une nouvelle voie de déposition de l'histone variante H2A.Z au niveau des promoteurs de gènes utilisant l'hétérodimère Ku70/Ku80. Nos travaux montrent que le domaine de von Willebrand A (vWA) de Ku80 est responsable de la reconnaissance du dimère d'histone H2A.Z-H2B et de sa déposition au niveau des promoteurs. Nos résultats établissent les bases moléculaires pour étudier le rôle de H2A.Z dans le remodelage, la transcription et la réparation de l'ADN se trouvant au niveau des promoteurs de gènes.

5. SUMMARY

The highly dynamic chromatin structure keeps changing its configuration in order to accomplish its multifarious roles in the cell. The process of chromatin remodeling involves the ATP-dependent modifications in the overall structure of chromatin that lead to an inconstant accessibility of various cellular proteins to the chromatin. These changes not only involve the covalent modification of histone machinery but also include the physical rearrangement of the nucleosome structure. Histone replacement is the most significant clause of chromatin remodeling as it causes the physio-chemical changes at nucleosomes level that is critical for the interaction of DNA and important cellular proteins thus affecting many important cellular mechanisms. It involves the physical exchange of conventional canonical histones with that of counterpart non- allelic, non-canonical histones (variants). Despite the fact that these variants possess a quite high resemblance to their canonical counterparts, they can cause considerable changes in chromatin dynamics and its functions.

In recent time, H2A.Z remodeling has got huge attention of the research community due to its critical role in chromatin dynamics and gene expression. Histone variant H2A.Z has been reported to be involved in a variety of important cellular processes including transcription, genomic stability, chromosome segregation, and DNA repair, etc. Bearing a 60% sequence similarity to that of other cellular H2As, H2A.Z is the most extensively studied histone variant among all its family members found primarily at promotor regions of active genes. It constitutes about 15% of all H2A variants and is distributed throughout the genome both in euchromatin and heterochromatin. With ~90% sequence conservation among diversified species, H2A.Z is considered as the most conserved histone variant during the course of evolution that justifies its important role in a variety of key cellular processes.

Likewise, the Ku heterodimer, consisting of Ku70 and Ku80 subunits (also referred as Ku86 in higher eukaryotes), is another highly abundant protein found primarily in the nuclear region of cells (Koike *et al.*, 2001). On an average, about half a million of Ku molecules are considered to be present in the nucleus of a single human cell (Mimori *et al.*, 1986; Walker *et al.*, 2001). Presence in such a huge amount certainly

advocates its important roles in several key cellular processes. Both subunits form a toroidal structure that wraps itself around the DNA to perform the dedicated functions. Moreover, in the research literature, Ku heterodimer is now being proposed not merely as a DNA binding protein rather an entity with a wide range of functions in its account. These functions include its pivotal roles in apoptosis, V(D)J recombination, telomere maintenance, DNA replication, transcriptional activation and most importantly in genomic stability through serving as a molecular scaffold to recruit the primary members of the NHEJ repair pathway at the sites of double-strand DNA breaks. Based on structural and functional evidence, both Ku70 and Ku80 have been designated to possess an N-terminal von Willebrand factor type A (vWA) domain, a central DNA binding core domain, and a nuclear localization sequence (NLS) domain. C-terminal domain of Ku70 comprises an SAP (SAF-A/B, Acinus and PIAS) domain while C-terminal domain of Ku80 contains a sequence responsible for binding to DNA-PK_{CS}. All of these structurally distinct domains have been designated to perform diversified yet particular functions, contributing to the overall role of Ku heterodimer in different cellular processes.

In the literature annals, a great deal of information has been provided regarding the mechanism of H2A.Z remodeling however with the advent of new emerging techniques to comprehend such dynamic phenomenon, new members are being unveiled playing a significant role in these mechanisms. In yeast, the ATP-dependent chromatin-remodeling complex SWR1 is responsible for the exchange of canonical H2A/H2B dimer with histone variant H2A.Z/H2B dimer mediated by histone chaperon Chz1. In higher eukaryotes, two orthologs of SWR1 complex, the p400/TIP60, and Snf2 related CBP activator protein (SRCAP) complexes are considered to be responsible for this remodeling. Here we set out to investigate the interaction of Ku heterodimer and histone variant H2A.Z and the probable role of this interaction in H2A.Z deposition in chromatin.

By virtue of double immunoaffinity technique, we first identified Ku heterodimer in the H2A.Z predeposition complex from HeLa cells. The subsequent mass spectrometric analysis also endorsed our initial findings. After that, we proceeded towards the exploration of specificity of interaction between Ku heterodimer and

H2A.Z. For that, we co-expressed Ku heterodimer with H2A/H2B, H2A.X/H2B and H2A.Z/H2B employing *E-coli* expression system. On account of this *in-vitro* analysis, we identified that Ku heterodimer interacts specifically with H2A.Z/H2B and not with counterpart H2A/H2B and H2A.X/H2B. This specificity suggested us the probable prospective of this interaction in the regulation of H2A.Z in chromatin.

Following the revelation of exclusive interaction of Ku heterodimer with histone dimer H2A.Z/H2B, we decided to further characterize the Ku70/Ku80 and H2A.Z/H2B interaction by mapping the Ku heterodimer domain involved in histone recognition and binding. By virtue of *in-vitro* experiments, we identified the von Willebrand factor type A (vWA) domain of Ku80 to be primarily responsible for this interaction with histone dimer H2A.Z/H2B. In the literature, several studies have implicated the vWA domain of Ku80 to be responsible for a wide range of protein-protein interaction carrying out significant roles in DSB repair and telomere regulation. Upon further exploration of this interaction, we identified that a 7kDa subdivision of vWA domain of Ku80, Ku80 (aa 40-101) motif, is the minimal domain sufficient for the interaction with the histone H2A.Z/H2B dimer. To re-ensure the specificity of interaction of this domain with H2A.Z/H2B, we reassessed the interaction of the identified domain with H2A.Z/H2B and counterpart H2A/H2B and H2A.X/H2B and found the consistency of specificity of interaction with H2A.Z/H2B alone.

Following the recognition of minimal domain of Ku heterodimer responsible for the interaction with H2A.Z/H2B, we aimed to identify the H2A.Z domain critical for this interaction. Towards this goal we employed chimeric proteins constituting H2A and H2A.Z portions and end up concluding that the central region of H2A.Z comprising highly conserved $\alpha 2$ helix is responsible for the specific interaction of histone dimer H2A.Z/H2B with Ku80 (aa 40-101) motif. Noteworthy, in the context of all *in vitro* experiments conducted, the amount of H2B was always found substantially higher as compared to H2A.Z suggesting that Ku80 (aa 40-101) motif might be binding to the histone H2A.Z/H2B dimer favorably through H2B partner. Employing the *in vitro* exploration through double immunoaffinity procedure and size-exclusion chromatography, we observed a strong interaction of H2B with Ku80 (aa 40-101) motif. This advocated the pivotal role of H2B in the interaction of Ku heterodimer and histone

H2A.Z/H2B dimer, though, the interaction with H2B is feasible only in the case of H2A.Z/H2B dimer and not in the case of H2A/H2B or H2A.X/H2B counterparts.

After the exploration of the interaction of Ku heterodimer and histone H2A.Z/H2B dimer through *in vivo* and *in vitro* approaches, we proceeded to evaluate the functional implications of this interaction. Employing genome-wide chromatin immunoprecipitation followed by sequencing (ChIP-seq), we identified that Ku heterodimer specifically regulates the H2A.Z deposition in chromatin, specifically at promoter regions. We found about 50% increased deposition of H2A.Z around promoter regions (TSS > 2 kilobases) with no significant variance in the deposition of H2A.Z at enhancers and insulators. This validated the critical involvement of Ku80 in genome-wide localization and chromatin enrichment of H2A.Z at promoters. Following that, we assessed the significance of Ku80 (aa 40-101) motif in H2A.Z deposition employing the same ChIP-Seq approach. Upon analysis, we observed a two-fold increase in H2A.Z deposition in case of WT-Ku80 MEFs as compared to Δ N-Ku80 MEF cells. This finding strongly advocated the critical role of the vWA domain or more precisely Ku80 (aa 40-101) motif in the deposition of H2A.Z/H2B *in vivo*.

Taken together, all these results point to a critical role of Ku heterodimer in the deposition of H2A.Z in chromatin. We suggested that von Willebrand factor type A (vWA) domain of Ku80 or more precisely Ku80 (aa 40-101) motif is substantially responsible for this interaction. We also observed that the Ku heterodimer interact exclusively with the histone H2A.Z/H2B dimer and this process is primarily mediated by H2B. This finding raised the question of the dynamic role of H2B in the specific deposition of H2A.Z. In the presumptive explanation of it, we hypothesize that Ku mediated deposition of H2A.Z/H2B may either involve the interaction of Ku with both H2A.Z and H2B members or alternatively the interaction is facilitated directly by H2B alone but only in the presence of H2A.Z as its dimer partner and not in case of H2A or H2A.X. This complexity of interaction strongly advocates the need for a crystallographic study to understand the dynamicity of the process in a more holistic manner.

5. INTRODUCTION

1.1 Chromatin:

The basic difference between the domains of life - Archaea, Bacteria, and Eukarya - is based on how they carry their genetic material. In the first two, all the genetic material is found inside the cell cytoplasm in a somewhat simpler pattern. However, in eukaryotes, the genetic information of the cell does not appear in the form of simple strands of genomic DNA rather they are found inside the cell nucleus in the form of highly complex nucleoprotein structure, the chromatin. Chromatin, a term coined by Walther Flemming in 1879, is comprised of DNA, RNA, and histone/non-histone protein macromolecules (Flemming, 1879). It's a highly dynamic structure, which changes its organization arrangement to accomplish its diversified roles in the cell.

In a diploid mammalian cell, around 6 billion base pairs are packed to construct 23 chromosome pairs. Considering the size of each base pair as 0.34 nm, around 2 meters of DNA has to be packed into a single nucleus of a diameter of just 10 μm . Similarly, on a bigger canvas, where an average 70 kg human body is estimated to have about 37 trillion cells, the total length of DNA present in a body would be around 74 trillion meters. Based on the fact that the distance between the sun and earth is 150 million kilometers, it results into an interesting fact that each of us has enough DNA to pay a return visit from earth to the sun for about 250 times (Kornberg, 1974; Bianconi *et al.*, 2013).

How nature has prepped itself to pack such an extremely huge amount of inheritable genomic material into an extremely small area, is really a nature's wonder. The understanding of the mystery of this wonder takes us to appreciate the basic structural and functional unit of chromatin, the nucleosome.

1.1.1 Nucleosome: The Basic Unit of Chromatin:

In eukaryotes, from yeast to metazoans, the structure of nucleosome is considerably similar. A nucleosome core particle is made up of around 146 base pairs

of DNA wrapped around a histone octamer (Van Hoide *et al.*, 1974; Goldman, 1988). This wrapping was first time structurally studied in the 1980s by Aaron Klug team where they hypothesized that about 1.67 left-handed turns of DNA helix are packed around a histone octamer (Kornberg and Klug, 1981; Travers and Klug, 1987). In 1997, Richmond team solved the first near atomic resolution crystal structure of nucleosome (Luger *et al.*, 1997). The building block of DNA, the base pair, constitutes of two nucleobases bound to each other through hydrogen bonding and shaping DNA into Watson-Crick double helix modal where guanine binds to cytosine and adenine to thymine. DNA has a phosphate backbone which is negatively charged and induces an electrostatic repulsion among adjacent DNA regions, resulting it impossible for DNA to fold upon itself. Subsequently, the compaction of such a large DNA is achieved through the phenomenal wrapping of DNA around positively charged small histone protein to form the repetitive unit of chromatin, the nucleosome. This electrostatic interaction is considered as the main source of energy required for wrapping of DNA around histones (Richmond *et al.*, 1984).

1.1.2 Chromatin Organization:

Nucleosomes are condensed through a sequence of successively higher order structures that finally form the chromatin. In electron micrographs, different nucleosome units can be seen binding each other through segments of linker DNA, which varies in size from 10-80 base pairs depending upon eukaryotic species, resulting into formation of ‘10 nm fiber’ structure sometimes substituted as ‘beads-on-a-string’ structure which further compacts to create ‘30 nm fiber’ structure (Olins and Olins, 1974; Woodcock et al., 1976). These 30 nm structures then form a succession of loops which fold back on themselves to manage further compaction. These fibers keep condensing into multiple loops and coils and to even higher orders giving the final structure as chromosome (Figure 1).

Structure of chromatin, however, differs depending upon the stage of the cell cycle. For instance, during mitotic division, the chromosomes are so tightly packed that their condensed form can be seen through an ordinary light microscope. During

mitotic and meiotic cell divisions, cellular machinery photocopies the chromosomes in a precise way so that each new daughter cell has an accurate number of chromosomes. The physical strength of chromatin at this stage is essential to avoid any physical damage to DNA as the daughter chromosomes get separated. On the other hand, during interphase, a stage when eukaryotic cells are not dividing, the chromatin is found in a less tightly packed pattern to allow DNA polymerases and RNA polymerases to undergo DNA replication and transcription processes respectively. During interphase, the chromatin can be easily divided into euchromatin and heterochromatin. A lesser-condensed portion, the euchromatin, is comprised of DNA regions which contain actively transcribed genes (Dame, 2005; Hansen, 2012). This looser configuration hence helps transcription machinery to access and transcribe the specific region of DNA. On the other hand, heterochromatin is more condensed region comprises transcriptionally inactive genes, hence hinders the access of enzymes involved in DNA replication, transcription, recombination, and repair etc. This composite conformation of chromatin provides cell an exemplary control on gene expression (Bednar *et al.*, 1998).

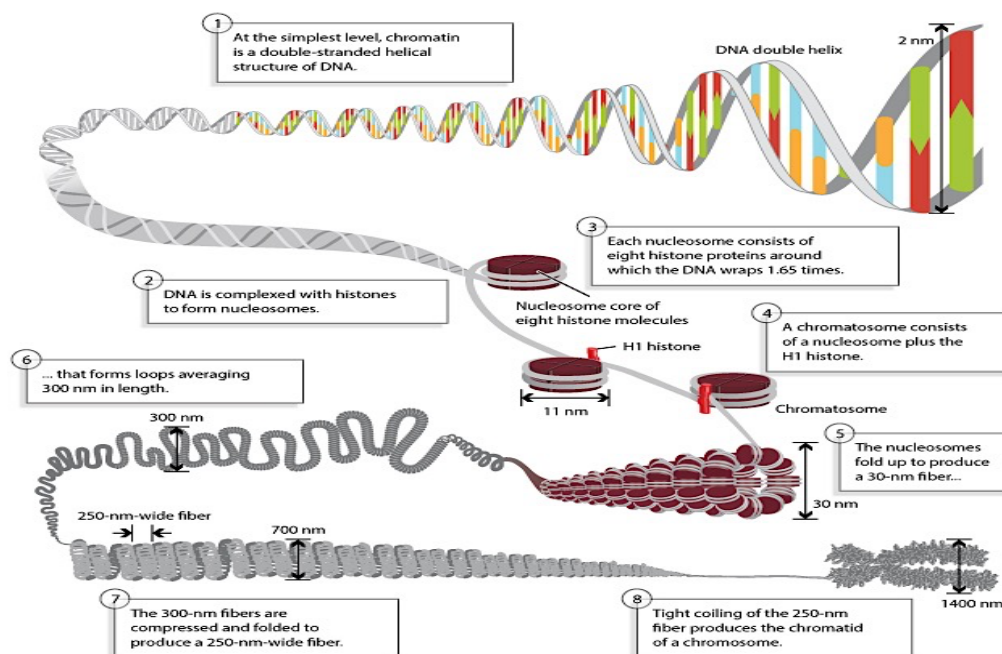


Figure 1: Overview of Chromatin Organization. The genomic DNA is wrapped around histone octamer forming fundamental building unit of chromatin, the nucleosome. The nucleosome then hierarchically condensed to finally forms the chromosome. Adapted from (Pierce, 2012) (modified).

Some particular cells of the body like spermatozoa, however, don't follow this typical organizational scheme since they have more tightly packed DNA than other eukaryotic cells (Ward and Coffey, 1991). Likewise, the genomic material of prokaryotes, preferably named as 'genophores' in literature archive, is generally presented as a single circular chromosome within the cytoplasm and has an entirely different mechanism of arrangement (Haapala and Soyer, 1975). It has a far less condensed structure as compared to its eukaryotic counterpart. Nevertheless, most of the eukaryotic cells follow the aforementioned characteristic mechanism to organize their genome.

This way, the state-of-the-art packaging of the eukaryotic genome not only permits a large amount of genomic DNA to reside in a very tiny space, but it also incorporates extraordinary functional importance. Though over the last few decades a lot of studies have tried to solve the nature's wonder, with the advent of even more advanced techniques to study chromatin structure, fact-finding will be continued to be further explored by the research community.

1.1.3 Histone Proteins:

The secret of above-mentioned chromatin organization lies in the wondrous proteins that wrap chromosomal DNA around them, the histone proteins. They are found in the nuclei of Eukarya and in certain Archaea, however, they are absent in Bacteria. Histones were first time identified in avian red blood cells by Albrecht Kossel in 1884, well before their role was discovered in chromatin packaging (Kossel, 1884). Histone genes were among the earliest eukaryotic genes to be characterized. They are the small proteins ranging from molecular weight from 10 to 14 kDa, enriched with positively charged amino acids such as arginine and lysine that constitutes about 20-24% of total structure (Lehninger *et al.*, 2005). Through facilitating chromatin assembly or disassembly, they play a central role in the transition between active and dormant states of chromatin. Based on their role, histone proteins can be divided into core and linker histones.

1.1.3.1 Core Histones:

Core histones (H2A, H2B, H3, and H4) perform the main structural functions in the chromatin assembly and are very critical for its stability. Core histones are highly conserved throughout evolutionary lineage emphasizing their importance in cell maintenance. They together form a discrete octamer around which the DNA wraps. A histone octamer comprised of two dimers of H2A-H2B and one tetramer of H3-H4 assembled together to form an approximately 63 Å diameter unit (Figure 2).

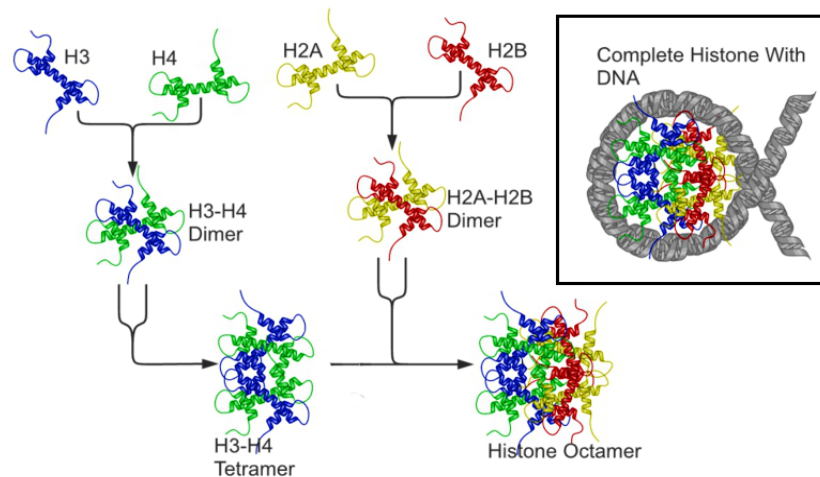


Figure 2: Schematic assembly of histone octamer: Two copies of each core histone protein (H2A, H2B, H3, and H4) forms a histone octamer around which DNA wraps to prepare the repetitive unit of chromatin, the nucleosome. Adapted from (Alberts et al., 2002) (modified)

These core histones carry moderately similar structural domains with slight differences. Each core histone is composed of an architectural motif, commonly known as the histone fold, containing two shorter α -helices ($\alpha 1$ and $\alpha 3$) along with a longer α -helix ($\alpha 2$) in the middle that is linked through two loops (L1 and L2). Two of the core histones (H2A with H2B and H3 with H4) form the basic heterodimer units (Luger et al., 1997). This heterodimer yields an organizational structure where loops (L1 and L2) comes on the opposite sides of heterodimers. The helical structure of these heterodimers allows two H3-H4 heterodimers to bind each other through

interaction of α helices of two H3 histones in a head-tail pattern forming the central H3-H4 tetramer. This head-tail pattern is more formally known in the literature as the handshake motif. Considering the same handshake motif two of H2A-H2B dimers then binds to H3-H4 tetramer, through H2B-H4 interaction, resulting in octameric histone (Mariño-Ramírez *et al.*, 2005). This way, the H3-H4 formulates the stable core while the two H2A-H2B heterodimers form the surface of the octamer, making H2A and H2B more accessible to external interactions.

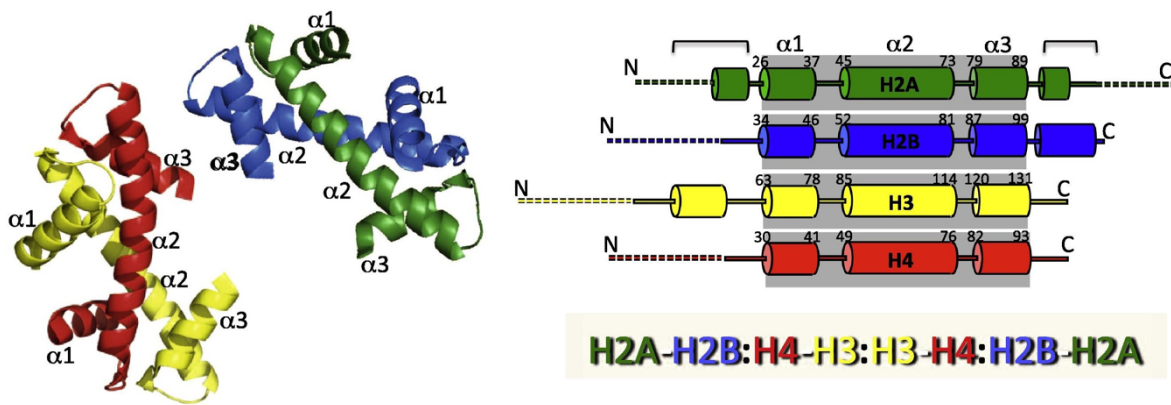


Figure 3: Structural Aspects of Core Histone Proteins: Left: Structure of H2A (green), H2B (blue), H3 (yellow) and H4 (red) showing three α -helices ($\alpha1$, $\alpha2$ and $\alpha3$) along with some of loops. Right: Secondary structure of core histone protein indicating their longer α -helix, $\alpha2$, separated from two shorter α -helices ($\alpha1$ and $\alpha3$) with the help of loops L1 and L2. Bottom illustration showing interactions among different core histones. Dashes (-) indicates histone dimerization while colon (:) indicates dimer-dimer interaction. Adapted and modified from (Cutter and Hayes, 2015).

In addition to the histone fold domain, each of the four core histone molecules contains flexible N-terminal tails of variable length that protrude out from the center of nucleosome core as a disordered structure and serve as the place for a wide variety of post-translation modifications (PTMs) such as acetylation, methylation, phosphorylation and ubiquitination, etc. Though these tail extensions represent up to 30% of the total histone mass, yet they are not visible in the crystal structure of nucleosomes as they undergo high intrinsic flexibility and are thought to be primarily unstructured. These tails are not only essential for chromatin stability but also very decisive for maintaining the condensed state of the 30 nm chromatin fiber and higher order structures hence play an important role in assembly or disassembly of chromatin.

The histone octamer through its positively charged amino acids such as arginines and lysines, physically and chemically interact with DNA. This interaction is mediated by the histone core folds and their N terminal tails. DNA binds to the positively charged histone octamer through its negatively charged phosphodiester backbone. Neutralizing the positive charge of histones decreases the intensity of their interaction with DNA, allowing the nucleosome to adopt a more open structure, a prerequisite for gene activation. Within the core histone octamer, two major DNA binding sites including the 'α1α2 site', which is formed with the help of α helices of two adjacent histones, and the 'L1L2 site', which is crafted by the loops of two adjacent histones. The lack of any site-specific interaction between DNA and histone octamer explains the phenomenon of how nucleosome is able to arrange the long DNA in a sequence-independent fashion. The N-terminal tails of core histones do not interact with specific DNA region yet play a very important role in stabilizing DNA around the octamer thus also contribute to differential genes expression. They help DNA to wrap in a left-handed pattern around the histone octamer during chromatin formation (Cutter and Hayes, 2015).

1.1.3.2 Linker Histones (H1s):

Linker histones are an important constituent of nucleosomes in higher eukaryotes. Regardless of the fact that they are considerably conserved, linker histones are the most variable histones across eukaryotic species. Structurally, histone H1 comprised of a tripartite structure which includes a protease-resistant stably folded (~80 amino acid) central domain and a short (~30 amino acid) N-terminal and a large (~100 amino acid) basic C-terminal tails which are protease-sensitive (Ramakrishnan *et al.*, 1993). Unlike core histones, linker histones are not involved in the basic 'bead on string' formation. Instead, the linker histone H1 helps to attach various primary nucleosome structures at the entry/exit sites of the DNA, consequently helps to clamp the DNA around histone protein. Furthermore, linker histone also binds to 'linker DNA' which varies in size from 10-80 base pairs depending upon eukaryotic species and hence provide assistance in the formation of higher order chromatin structure.

Through these interactions with DNA, linker histone H1 plays important role in regulating specific gene expression (Thoma and Koller, 1977; Thoma *et al.*, 1979; Routh *et al.*, 2008)

In addition to the role in the maintenance of higher-order chromatin structure, many other cellular mechanisms like mitosis and replication timing are well linked to post-translational modifications of linker H1 histones. The major post-translational modifications of histone H1 include acetylation, ubiquitination, and methylation of lysine; and phosphorylation of serine and threonine. The histone H1 family comprises of five somatic variants namely, H1.1, H1.2, H1.3, H1.4, and H1.5 which are found in almost all cell types. However, in different species of eukaryotes and even in different tissues and developmental stages of the same species, different isoforms of H1 linker histone can be found (Talbert *et al.*, 2012). An important isoform, histone H5, exist only in avian erythrocytes. Other isoforms include histone H1M, histone H10, histone H1A-E, etc.

1.2 Epigenetic Regulation:

The term ‘epigenetics’ was coined for the first time by the British geneticist Conrad Hal Waddington in 1939 (Waddington, 1939). Failure to a satisfactory explanation of the mystery that how the same embryonic stem cells having the same genetic material differentiate themselves into different specialized cells, led to the foundation of a new phenomenon named ‘epigenetics’. The Greek prefix epi- (ἐπι- "around, over") in the word ‘epigenetics’ designates to the characteristics that are "in addition to" the conventional genetic feature of inheritance. Over the time, with the enormous expansion of our understandings regarding the molecular mechanisms of gene regulation in eukaryotes, the term ‘epigenetics’ has undergone through a process of evolution. Epigenetics is usually referred to as the ‘study of changes in organism caused by modification of gene expression rather than alteration of the genetic code itself’. Since its conception, it has been redefined many times in literature archive by different researchers (Holliday, 1990; Riggs *et al.*, 1996; Riggs and Porter, 1996; Bird, 2007). However, as in Cold Spring Harbor meeting in 2008, a consensus definition of the epigenetic trait was framed as "stably heritable phenotype resulting from changes in a chromosome without alterations in the DNA sequence" (Berger *et al.*, 2009).

Traditionally, many phenomena that are not considered essentially heritable have also been designated as epigenetic. For instance, some of the modification like histone modifications of chromosomal regions have long been studied under the ‘epigenetics’ irrespective of their association with a phenotype and whether or not if these changes are heritable. According to the current consensus definition, a phenotypic trait needs to be heritable either to the progeny of organisms or of cells if it is to be considered as an epigenetic characteristic (Ledford, 2008; Berger *et al.*, 2009). Furthermore, it is also not essential to ‘epigenetics’ to represents only the modifications that affect gene expression and its activity rather it can include any heritable phenotypic change. These types of changes may occur due to any foreign or environmental factors that influence the phenotypes at the cellular and organismal level. Most of the epigenetic modifications happen within a succession of a single individual’s lifetime. Subsequently, they are preserved and transmitted to the next

generation through a process known as ‘transgenerational epigenetic inheritance’ (Chandler, 2007).

The secret of the mystery of how some genes are turned ‘off’ and ‘on’ is now well revealed with the study of epigenetic phenomena. Epigenetic factors influence the gene expression without changing the genetic code of the DNA sequence. They perform this function either through changing the microstructure of DNA or through covalently modifying the associated protein, like histones, ultimately results in activation or silencing of genes. The same phenomenon of epigenetic regulation enables the genetic material of embryonic stem cells to get differentiated into particular cells and to express only particular genes that are essential for the specific activity of those cells.

DNA damage, a very common phenomenon of the cell since on an average basis one mammalian cell can undergo more than 60,000 endogenous DNA damages per day, can cause epigenetic changes as well (Kovalchuk and Baulch, 2008; Ilnytskyy and Kovalchuk, 2011). Although these DNA damages are generally repaired yet many of epigenetic changes prevail at DNA repair sites (Cuozzo *et al.*, 2007). For instance, in double-strand breaks (DSBs), DNA break can provoke an unprogrammed and non-conventional epigenetic gene silencing through both chromatin remodeling and DNA methylation (O'Hagan *et al.*, 2008). A detailed review on the role of DNA damage and repair on epigenome would be made in the later chapters.

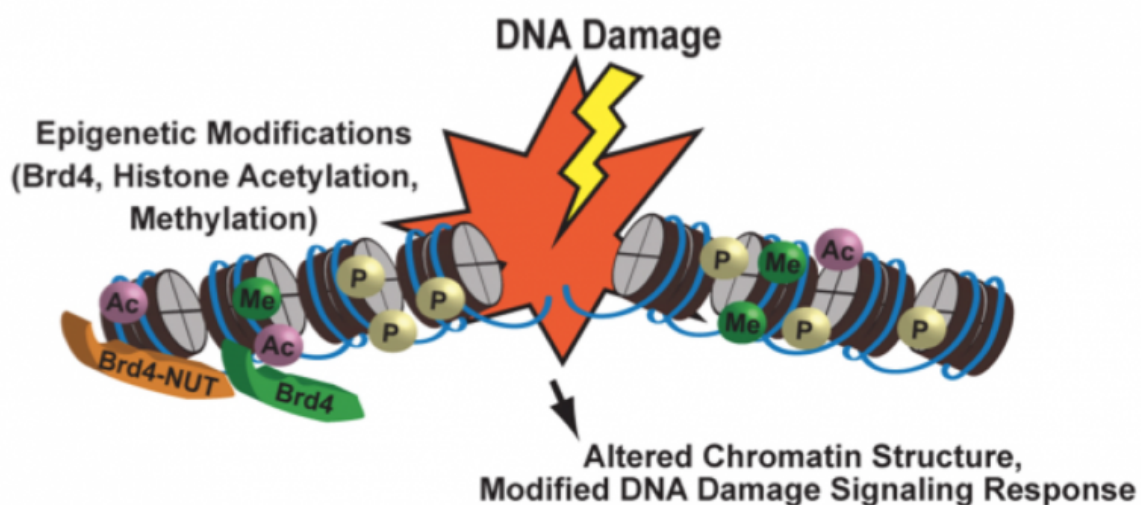


Figure 4: DNA Damage Generate Specific Epigenetic Modifications: Schematic representation of particular epigenetic modification caused as a result of DNA damage. These epigenetic modifications prevail over generations. Adapted from (Duke University Health System, 2013).

As mentioned beforehand, epigenetics undergoes covalent modifications to both the DNA and histone proteins. Some major epigenetic mechanisms include DNA methylation, non-coding RNAs associated gene silencing, histone post-translational modifications (PTMs) and nucleosome remodeling through the exchange of histone variants (Figure 4). Each of these mechanisms modifies the expression of genes without changing the primary DNA sequence of genetic material.

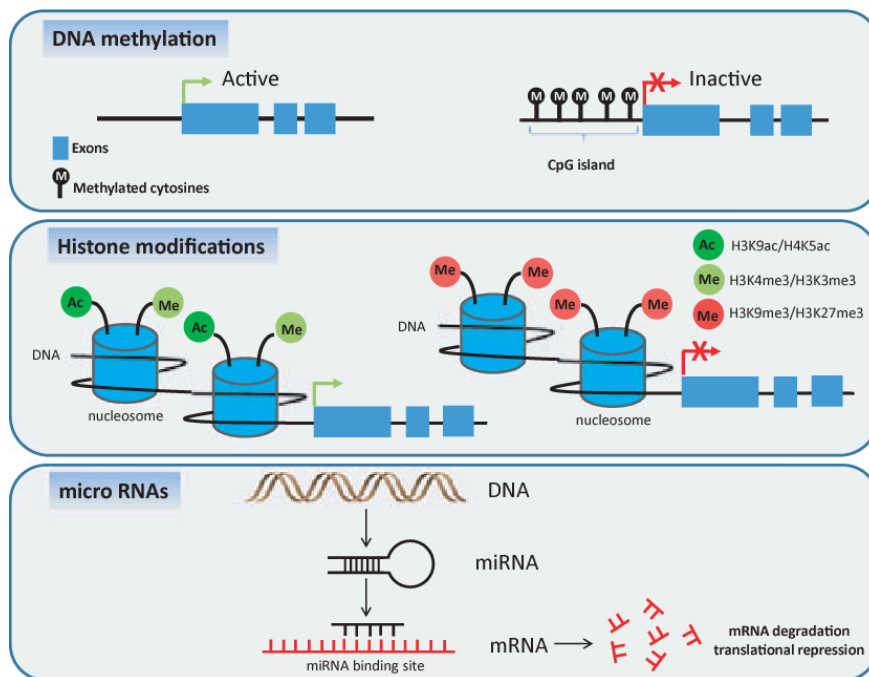


Figure 5: Epigenetic Regulations: Schematic synopsis of various epigenetic mechanisms through which the cell controls gene expression. Covalent modifications to the DNA and histone proteins result into epigenetic changes in genetic message. Adapted from (D'addario *et al.*, 2013)

1.2.1 DNA methylation:

In mammals, DNA methylation is one of extensively studied epigenetic mechanism which is critical for normal development since it is associated with a number of key processes including B cell differentiation (Kulis *et al.*, 2015), X-chromosome inactivation, genomic imprinting (Beard and Jaenisch 1995), repression of transposable elements, carcinogenesis (Daura-Oller *et al.*, 2009) and aging (Horvath, 2013) etc. The process of DNA methylation was primarily described by Rollin D. Hotchkiss in 1948 during his studies on calf thymus DNA where he first

time identified 5-methylcytosine as a distinct entity than cytosine using paper chromatography technique (Hotchkiss, 1948).

Methylation can happen on both cytosine and adenine bases of DNA. Adenine methylation is less common and has been identified in bacteria, plants and in recent times in mammalian cells (Wu *et al.*, 2016). On the other hand, the cytosine methylation is more common and is widespread throughout prokaryotic and eukaryotic species. Cytosine methylation occurs at the C5 position of cytosine pyrimidine ring converting it into 5-methylcytosine. The process is catalyzed by a group of specific enzymes known as DNA methyltransferases (DNMTs) that includes DNMT1, DNMT3A, and DNMT3B which transfer the methyl group from S-adenosyl-L-methionine to the cytosine base (Figure 6).

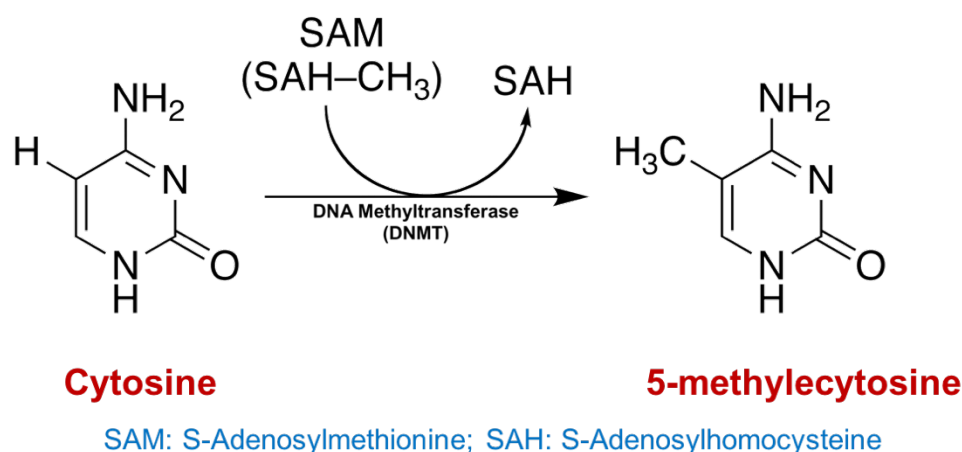


Figure 6: Methylation of cytosine to 5-methylcytosine by the action of DNMT enzyme:
The process of DNA methylation requires the transfer of methyl group from S-adenosylmethionine to cytosine nucleotide to convert it into 5 methylcytosine (5mC).

Majority of cytosine methylation occurs on the cytosines that are found adjacent to guanosine nucleotides, formally known as CpG dinucleotides. About 60-90% of the CpGs across the genome are considered methylated in mammalian cells (Gardiner-Garden and Frommer, 1987; Larsen *et al.*, 1992; Deaton and Bird, 2011). These methylated CpGs are more commonly found within the repetitive elements and centromeric regions of the genome. On the other hand, non-methylated CpGs are more abundantly found in CpG-rich genomic sequences, designated as CpG islands. Around 50% of the genes in the human genome are associated with CpG islands in their

promoter regions. Any disturbance in the methylation pattern of these promoters can lead to serious abnormalities (Guilleret *et al.*, 2002; Dawson and Kouzarides, 2012).

1.2.2 Regulatory Non-coding RNAs:

As per recent exploration by the Encyclopedia of DNA Elements (ENCODE) project, it is estimated that about 70% of genomic DNA get transcribed into RNA molecules. However, only 3% of these RNA molecules, as mRNA is annotated, ultimately get translated into a protein product and on that account known as coding RNAs. Besides these coding RNAs, the remaining huge fraction of RNA transcripts do not appear to encode for proteins and are generally referred to as non-coding RNAs. Non-coding RNAs (ncRNAs) are generally defined as the RNA molecules that are not actually translated into a protein product. Although some of the researchers suggest that these ncRNA may encode for short peptides yet considering the results from ribosome profiling (Guttman *et al.*, 2013) and evidence from different ENCODE cell lines that undergo the large-scale proteomic analysis, overall consensus lies on the fact that they are largely non-translated/non-coding transcripts.

Not getting translated doesn't suggest that they are functionally inert rather a huge fraction of these ncRNAs provide several regulatory activities in the cell and are also considered an important epigenetic component of the cell to control gene expression. Though not all of the ncRNAs are functionally important yet many of them have long been known significant for different cellular processes. These functionally crucial ncRNAs includes both small (having size less than 200 nucleotides in length) and large-sized non-coding RNAs. Small ncRNAs includes transfer RNAs (tRNAs) essential for transporting amino acids, small nuclear RNAs (snRNAs) considered important for RNA splicing and small nucleolus RNAs (snoRNAs) proposed crucial for RNA modifications. Large ncRNAs includes functionally indispensable ribosomal RNAs (rRNAs) well known for its pivotal role in protein synthesis. The ncRNAs also comprise some tiny ncRNAs like endogenous siRNAs (endo-siRNAs), microRNAs (miRNAs) and PIWI-associated small RNAs (piRNAs) which are widely studied in recent times (Malone and Hannon, 2009;

Pasquinelli, 2012; Thomson and Lin, 2009). Many of these non-coding RNAs have also been reported showing a significant role in many diseases including cancer (Shahrouki and Larsson, 2012), autism (Ziats and Rennert, 2013), cartilage-hair hypoplasia, Prader–Willi syndrome (Sahoo *et al.*, 2008) and Alzheimer's disease (Faghihi *et al.*, 2008) suggesting that ‘functional ribonomics’ is an important and integral section of functional genomics and epigenetics.

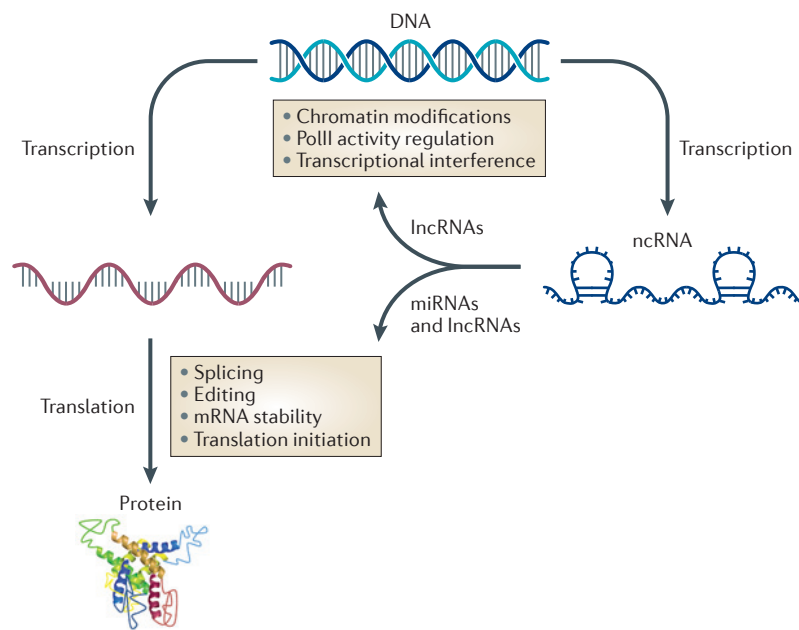


Figure 7: Central dogma indicating role of non-coding RNAs: The ncRNAs are considered vital in gene regulation at transcriptional or translational levels. Recent studies have also showed their interaction with DNA methyltransferases and histone modifying complexes hence emerged as important epigenomic regulator. Adapted from (Wahlestedt, 2013)

1.2.3 Histone Post-Translational Modifications (PTMs):

It is well explained in literature archive that how the histone proteins serve as a positively charged platform for operative binding of negatively charged DNA at the nucleosome level. Generally stated, all of the canonical histones (H2A, H2B, H3, and H4) comprise of N-terminal tail and C-terminal globular domain of variable lengths. The N-terminal tails are rich in arginine and lysine amino acids which makes them prone to different post-translational modifications such as acetylation, biotinylation, carbonylation, citrullination, glycosylation, methylation, phosphorylation, ADP-

ribosylation, SUMOylation and ubiquitination etc. These histone post-translational modifications are considered an important part of epigenome as it dynamically controls the local and global structure of chromatin.

Well before the phenomenon of DNA methylation came into debate, some of researches had already hypothesized that histone proteins could be noteworthy player in controlling gene expression (Stedman and Stedman, 1950). The scientists had an idea that within a living system, different types of cells must have different types of histones that can direct cells of so-called ‘same genetic material’ to be converted into different phenotypes. Allfrey and Mirsky (1964) first time hypothesized that ‘histone acetylation’ might be associated with gene activation. Afterward, serine phosphorylation and lysine acetylation were identified in a wide range of species sample such as calf thymus, rat liver, human lymphocytes and peas etc. (Kleinsmith *et al.*, 1966; Ord and Stocken, 1967; DeLange *et al.*, 1968; Gershey *et al.*, 1968; Vidali *et al.*, 1968). Soon after in 1975 ADP-ribosylation and ubiquitylation were revealed as important post-translational modifications that are related to histone protein (Goldknopf *et al.*, 1975; Ueda *et al.*, 1975). Later on, other modification like methylation and phosphorylation were identified, however their functional implication remained unclear for long time. In subsequent researches on yeast, scientists found N-terminal tails of histone proteins crucial for gene expression (Wallis *et al.*, 1980; Durrin *et al.*, 1991).

As of today, more than 500 distinct histone post-translational modifications have been found in literature archive present on about 250 different sites on core and linker histones (Zhao and Garcia, 2015). As a whole, these PTMs can be divided into active and repressive chromatin transcripts. For instance, methylation of K36 and H3K4, and ubiquitination of H2B are considered related to transcriptional activation (Joshi and Struhl, 2005; Shilatifard, 2008). On the other hand, methylation of H3K9, H4K20 and K27 are correlated to transcriptional repression in literature archive (Arnaudo and Garcia, 2013). Contrarily, the post-translational modifications can also be categorized the way they affect the gene expression. For instance, certain modifications on histone tails, such as phosphorylation and acetylation can modify the charge on the tails and hence change the chromatin dynamics through regulating

electrostatic mechanisms. On the other hand, other PTMs on these N-terminal tails act epigenetically through regulating the interaction chromatin with that of non-histone proteins (Jenuwein and Allis, 2001; Iizuka and Smith, 2003; Grewal and Moazed, 2003; Martin and Zhang, 2005).

All of these histone post-translational modifications constitute a “histone code” explaining the phenomenon of specific patterns of post-translational modifications. Since only specific amino acid residues of specific histone proteins undergo these PTMs, scientists have developed the hypothesis of ‘histone code’ considering that transcription of genetic information is partially regulated by these very specific chemical modifications of histone proteins (Jenuwein and Allis, 2001).

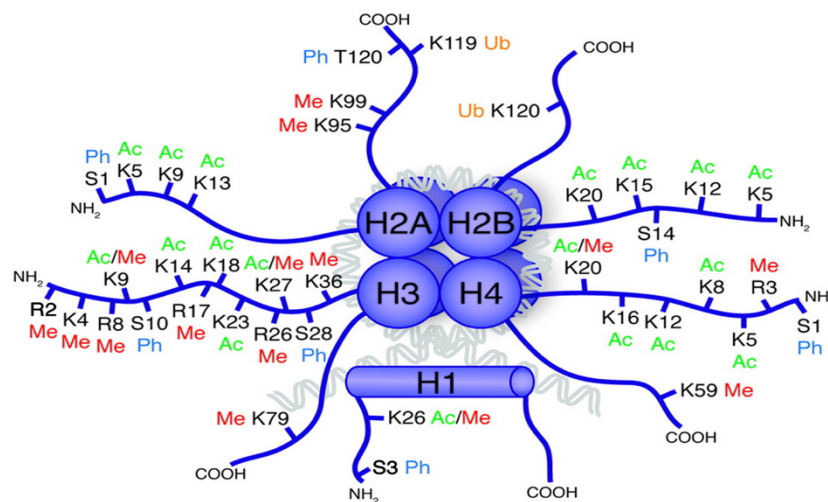


Figure 8: Histone Post-Translational Modifications: Overview of post-translational modifications of N- and C- terminals of canonical histones (H2A, H2B, H3 and H4). Most common PTMs include acetylation (Ac), methylation (Me), phosphorylation (Ph) and ubiquitination (Ub). Adapted from (Tollervey and Lunyak, 2012) (Modified)

1.2.4 Histone Variants:

Though the structure of chromatin seems pretty rigid at the nucleosome level, these structural units of chromatin are quite dynamic. Another important mean of regulating gene expression at nucleosome level is through replacing canonical histones with that of counterpart non-allelic histone variants. Histone variants are the non-canonical variants of conventional histones that encompass a difference of one or few

amino acids with that to their usual fellow histones. The term ‘histone variants’ may also include variants of histone H1 that lack a definite canonical counterpart. The debate on these non-allelic isoforms was started as early as in 1972 when during a research on calf thymus, a discrete isoform of histone H3 was first time identified (Marzluff *et al.*, 1972). Since then several histone variants have been identified throughout various eukaryotic species. Due to their role in chromatin dynamics and gene expression they have got a huge attention of research community in recent past.

Histone variants are transcribed independently from distinctive genomic loci in comparison to their respective canonical histones. At the biochemical level, some of these variants such as H2A.Z and macroH2A have highly distinctive polypeptide sequences as compared to their canonical siblings hence results in large-scale changes in chromatin structure and its function (Jackson and Gorovsky, 2000). On the other hand, other variants such as H3.3, which differ to its canonical H3 with just four amino acids, encounter only small sequence differences. Nevertheless, despite having high resemblance to their counterparts, such variants also cause considerable changes in chromatin dynamics (Szenker *et al.*, 2011).

The incorporation of canonical histones is strictly linked to cell replication. These canonical histones are synthesized only during S-Phase of the cell cycle. On the other hand, the histone variants are considered replication-independent and are kept producing throughout the whole cell cycle (Thiriet and Hayes, 2005; Dion *et al.*, 2007). Furthermore, in contrast to canonical histones that have absence of introns in their genes, the non-allelic variants may have introns in their genes. The encoding genes of canonical histones are present in multiple copies and are classically found throughout the chromosome whereas genes of histone variants are primarily found as single-copy (Weber and Henikoff, 2014). Likewise, mRNAs of canonical histones contain a short 3'- stem loop tails whereas mRNAs of corresponding variants have conventional polyadenylated tails (Marzluff *et al.*, 2008).

Acknowledging the fact that about half of the mass of eukaryotic chromosome is comprised of histone proteins, their importance in fundamental cellular processes becomes pretty obvious. This significant relevance has accompanied to evolution of huge number of distinct histone variants in different eukaryotic organisms. Literature

archive indicate the diverse roles of these non-allelic histone variants in many important cellular processes. They feature specialized cellular functions such as CENP-A mediated kinetochore assembly by (DeRop *et al.*, 2012), H2A.X attributed DNA repair (Yuan *et al.*, 2010), and H2A.Z associated anti-silencing (Hild and Paro, 2003), etc.

Histone variants exhibit a wide degree of conservation among different species. Based on conservation, these non-allelic histone variants can be divided into two main groups. One that are highly conserved like H2A.X, H2A.Z and CenH3 etc (Jackson and Gorovsky, 2000; Dubin *et al.*, 2010). Such variants are involved in various essential cellular functions such as formation of the centromere, antisilencing and DNA damage repair etc. These histone variants cannot be substituted with their respective canonical siblings as their counterparts are unable to perform these specific functions. Other histone variants like H2B.1 and H2A.Bbd are not considered highly conserved. They are tissue specific and are expressed in an on-demand manner to perform specific functions (Eirín-López *et al.*, 2008).

Among core histones, H2A and H3 have exhibits the highest number of known variants. H2B on the other hand have very limited number of corresponding non-allelic variants (Shaytan *et al.*, 2015). Only a few variants of H4 have been identified in trypanosoma and drosophila with no variant found in any other species till now (Kamakaka and Biggins, 2005). The low among of H2B and H4 variants might be associated to the fact that these histones are located in the centre of nucleosome core hence hinder their exchange for another non-allelic variant. With the increasing number of histone variant, recently a new nomenclature system has been proposed for a uniformity among the names of different variants (Talbert *et al.*, 2012). In the succeeding sections, histone variants of H1, H2A, H2B and H3 will be discussed in more details.

1.2.4.1 Histone H1 variants:

As described earlier, arginine rich linker histone H1 plays a significant role in regulating specific gene expression through maintaining higher-order chromatin

structure. Unlike core histones they lack the proper histone fold domain and bind to the linker DNA exiting from the core particle nucleosomes (Izzo *et al.*, 2008). Though highly conserved, linker histone H1 exhibits the most variable histone family with at least 11 different subtypes in mammals. These variants can be classified based on their distinct specificity to certain tissues, developmental stages, modes of expression and species in which they are expressed (Table 1).

Histone H1.1 to H1.6 are generally expressed during the S phase of cell cycle and therefore referred to as replication-dependent variants. All of them are encoded by single genes found within histone cluster 1 located on chromosome 6 in human cells. Other variants H1.0, H1.7, H1.8, H1.9 and H1.10 are expressed throughout the whole cell cycle hence referred as replication-independent variants. They have encoded genes distributed on different locations within the genome.

H1 variant	Amino Acids	Molecular Weight	Gene Symbol	Gene Location (human)	Expression Pattern
Replication Dependent					
H1.1	215	22 kDa	HIST1H1A	6p21.3	Somatic
H1.2	213	22 kDa	HIST1H1C	6p21.3	Somatic
H1.3	221	24 kDa	HIST1H1D	6p21.3	Somatic
H1.4	219	22 kDa	HIST1H1E	6p21.3	Somatic
H1.5	226	24 kDa	HIST1H1B	6p22.1	Somatic
H1.6	207	22 kDa	HIST1H1T	6p21.3	Testis
Replication Independent					
H1.0	194	21 kDa	H1F0	22q13.1	Somatic
H1.7	255	28 kDa	H1FNT	12q13.1	Testis
H1.8	346	36 kDa	H1FOO	3q22.1	Oocyte
H1.9	231	26 kDa	HILS1	17q21.33	Testis
H1.10	213	22 kDa	H1FX	3q21.3	Somatic

Table 1: Variants of linker histone H1 in mammalian cells: Representing highly diverse family of histones H1 that includes replication-dependent variants found within histone gene cluster 1 and replication-independent variants that are encoded by orphan

1.2.4.2 Histone H2A variants:

Histone H2A protein is comprised of around 130 amino acids with a molecular weight of about 14 kDa. It is an important core protein that forms the dimer with H2B through the handshake motif. Two of these dimers then interact with the H3-H4 tetramer to establish the final ‘octamer core’ around which the DNA wrap itself to form the nucleosome. Principally, the structure of histone H2A contains a ‘histone fold’ also known as ‘helix-turn-helix motif’ that is comprised of three alpha helices joined by two loops. This fold is actually responsible for handshake dimerization of H2A with that of counterpart H2B. This histone fold encompasses long extended tails on both N-terminal and C-terminal. These N-terminal and C-terminal tails are very much critical for various post translational modifications like acetylation, methylation, phosphorylation, ubiquitylation, and citrullination, etc.

In a histone octamer, histone H2A-H2B dimers are found on the surface of the octamer, the possible phenomenon that attribute H2A to exhibits the highest number of variants among all other histone protein family members. Most of histone H2A variants exhibit highest level of variability at their C-terminus. This C-terminus of histone H2A is predominantly found at DNA entry/exit positions that make the variants significantly important for chromatin functions and dynamics. In addition to C-terminus, H2A variants particularly differ with each other at Loop1 region that is considered responsible for interaction between two neighboring H2A.

Based on protein length and amino acid sequence the histone H2A family has been narrated to have about 265 distinctive members among different species. Apart from canonical H2A, it has four major class of variants namely H2A.X, H2A.Z, H2A.B (previously known as H2A.Bbd) and macroH2A. Both H2A.X and H2A.Z were discovered in 1980s (West and Bonner, 1980; Mannironi *et al.*, 1989) whereas subtype of macroH2A, mH2A1 was identified in 1990s (Fried, 1992) with the discovery of mH2A2 isotype in early 2000s (Costanzi and Pehrson, 2001). Variant H2A.Bbd was identified relatively more recently in 2000s (Chadwick and Willard, 2001). All H2A variants including canonical H2A are encoded by 26 genes in human genome. Majority of these genes are located in histone cluster 1 and 2 present on

couple the characteristics of both H2A.Z and H2A.X variants (Verni and Cenci, 2015). Similarly, the H2A variant H2A.W present specifically in plants contribute in chromatin condensation in *Arabidopsis* (Yelagandula et al., 2014).

Apart from non-canonical variants, different canonical H2A protein exist with slight sequence variation that are coded by distinctive canonical histone genes. Two different isoforms of canonical H2A includes H2A.1 and H2A.2. Both recently identified isoforms or otherwise named minor variants are undergoing different researches to study their functional importance. In literature archives, they are not marked as separate variants as their respective genes are very closely related. Furthermore, the sequence difference between these isoforms are usually lesser than their difference to other nominated H2A variants. These canonical H2A isoforms were first time identified in 1977 through electrophoretic separation proceeded by sequence analysis where the researchers found two different amino acids leucine and methionine on position 51 in H2A.1 and H2A.2 respectively (Franklin and Zweidler, 1977). Some other studies also witnessed that at position 16, about half of canonical H2A encompass a threonine while the rest half contain a serine residue. Consequent studies indicated the presence of sequence difference at few other places as well particularly at C-terminal end (Marzluff et al., 2002).

As per general rule for all canonical histones, canonical H2A is also deposited in a replication-dependent manner during S phase whereas most of its variants are expressed during the whole cell cycle. This marvel makes these variants functionally very critical for normal cellular processes. For instance, negatively charged residues of conventional H2A primarily responsible for development of acidic patch on the surface of nucleosome core which is important for the condensation of chromatin. Replacement of H2A with its variant H2A.Z that has an extended acidic patch as compared to H2A, causes a subsequent reduction in nucleosome array folding. As these acidic patches also serve as a scaffold for a variety of other cellular proteins, replacement of canonical histone with respective variants affect the binding of these crucial proteins at nucleosome surface (Fan *et al.*, 2004; Zhou *et al.*, 2007)

1.2.4.2.1 H2A.X variant:

Histone H2A.X is ~14 kDa H2A variant encoded by *H2AFX* gene. Based on cell type, it constitutes around 2-25% of total H2A variants. H2A.X contribute in a variety of important cellular functions including nucleosome construction, DNA repair and chromatin remodeling etc. At its C-terminus histone H2A.X encompasses an evolutionary conserved sequence motif 'SQ(E/D)Φ' where Φ symbolizes a variable hydrophobic residue that usually designate for tyrosine in mammalian cells. This is the sequence that get phosphorylated in response to different cellular episodes including X-chromosome inactivation, chromatin remodeling and most importantly DNA damage (Redon et al., 2002; Scully and Xie, 2013).

H2A.X converts into gamma-H2A.X (γ H2A.X) through the phosphorylation of serine present at position 139 within this SQ sequence motif in response of apoptotic digestion (Rogakou *et al.*, 2000), V(D)J splicing (Chen *et al.*, 2000), meiotic recombination (Mahadevaiah *et al.*, 2001) and most significantly in double strand DNA damage (Price and D'Andrea, 2013; Rybak *et al.*, 2016). Thousands of H2A.X molecules undergo this rapid phosphorylation to convert into γ H2A.X around the proximity of a double strand DNA damage break. The process is carried out through three critical kinases including ataxia-telangiectasia mutated (ATM), Ataxia Telangiectasia and Rad3 Related (ATR) and DNA-dependent Protein Kinase catalytic subunit (DNA-PK_{CS}) all three belongs to Phosphatidylinositol 3-Kinase-related Kinase (PIKK) family of kinases.

In literature archive, H2A.X has been rightly acknowledged as the 'genome guardian'. Gamma H2A.X is not only important for the repair of DSBs but also considered an efficient target for diagnosing double stand DNA damages in the cells. Nevertheless, the presence of gamma H2A.X in not a confirmation of DSBs in the cells. It can be deducted in the vicinity of DSBs as soon as 20 seconds post DSB formation. The mechanism how γ H2A.X help in repairing DSBs is still not completely understood yet it is known that γ H2A.X causes a decrease in DNA condensation hence provide a platform where members of DNA repair machinery can be recruited (Figure 9). Gamma H2A.X recruits the mdc1 mediator at the site of DSB. This mdc1 protein

then recruits two more proteins MRN and ATM which then initiates with the extension of γ H2A.X away from each side of DSB for hundreds of kilobase creating an extended domain of γ H2A.X. This wide array of γ H2A.X/mdc1 complex then causes the recruitment of RNF8 and RNF168 ubiquitin ligases that eventually recruits the DNA repair proteins brca1 and 53BP1.

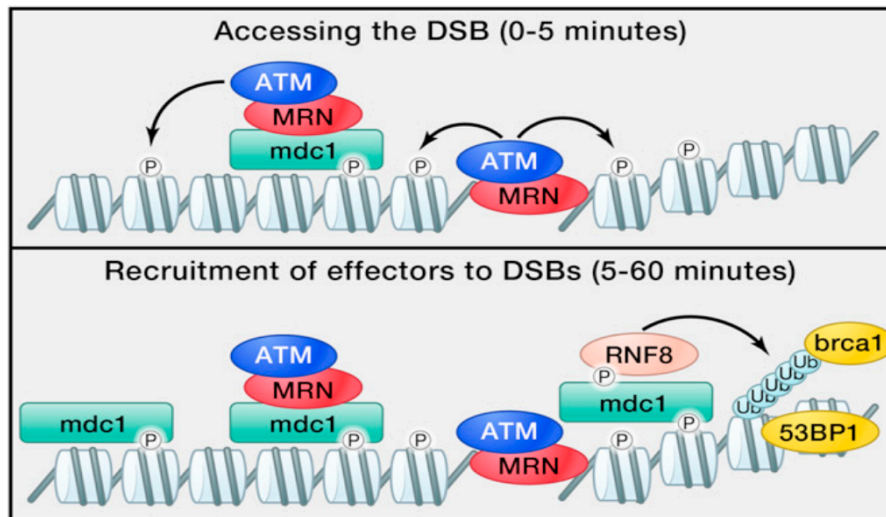


Figure 10: Indispensable Role of Phosphorylated H2A.X (γ H2A.X) in DSB Repair Mechanism: The PI3 family of enzymes phosphorylates H2A.X at DSBs that causes the recruitment of mdc1 protein. Next, ATM-MRN complex binds to mdc1 protein which then causes the spreading of γ H2A.X on both sides away from DSB. Adapted from (Price and D'Andrea, 2013).

1.2.4.2.2 H2A.Z variant:

The histone variant H2A.Z was for the first time identified in mouse in 1980s (West and Bonner, 1980). Afterward, it was identified in a protozoa *Tetrahymena thermophila* where it was found associated specifically with transcriptionally active chromatin (Allis *et al.*, 1986). Few years later, its homolog H2Av was identified in fruit fly *Drosophila melanogaster* (van Daal *et al.*, 1988). Subsequently H2A.Z variant was identified in 1990s in other species as well including yeast, mouse, etc. (Hatch and Bonner, 1990) and was evidenced to be essential for cell survival.

Until the last few years, only two isoforms of H2A.Z were described in literature archive. It includes H2A.Z.1 and H2A.Z.2 that are encoded by two separate non-allelic

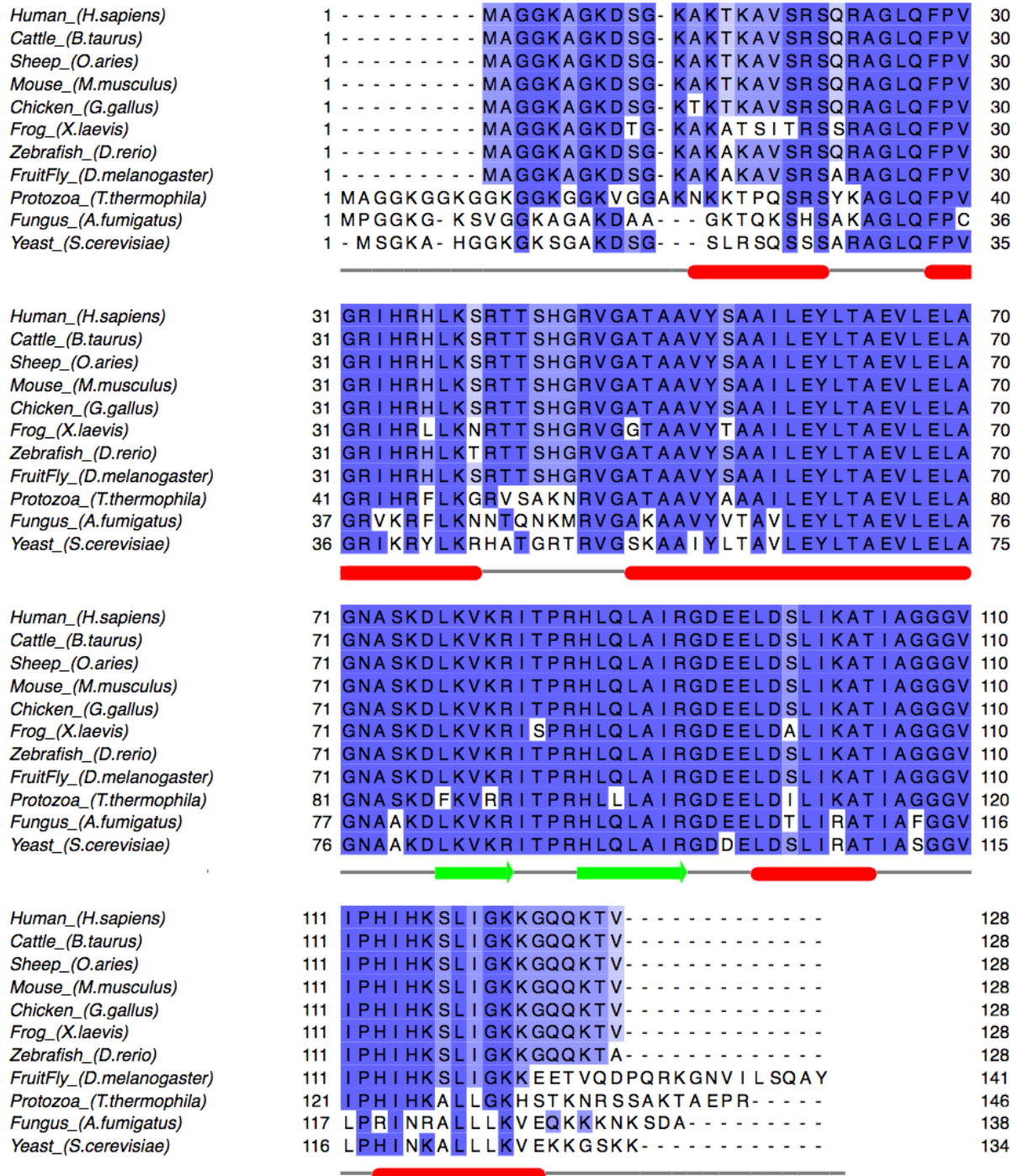


Figure 11: Multiple Alignment of Histone Variant H2A.Z from Various Eukaryotic Species: Evident sequence conservation of histone variant H2A.Z in different organisms is shown by shading, from dark blue (high) to light blue (low). The annotation of secondary structure shows the sites of α -helix (red cylinders) and β -sheet (green arrows). The sequence alignment is performed using Jalview® alignment program.

genes *H2AFZ* and *H2AFV*. Mass spectrometry analysis indicated that both of these variants differ with each other at only three amino acids (Talbert *et al.*, 2012). Even this slight difference of three amino acids bestow both entirely specialized functions.

For instance, it is H2A.Z.2 and not H2A.Z.1 which get exchanged with H2A at the site of double strand breaks (DSBs) (Nishibuchi *et al.*, 2014). More recently a primate specific new isoform named H2A.Z.2.2 has been discovered found primarily in brain cells (Talbert *et al.*, 2012). Both H2A.Z.2.1 and H2A.Z.2.2 differ with each other at C-terminus (Figure 12). H2A.Z.2.2 has relatively shorter C-terminal domain as compared to H2A.Z.2.1. Due to reduced binding of H2A.Z.2.2 to the neighboring histones, it presents a relatively less stable chromatin architecture (Bönisch *et al.*, 2012).

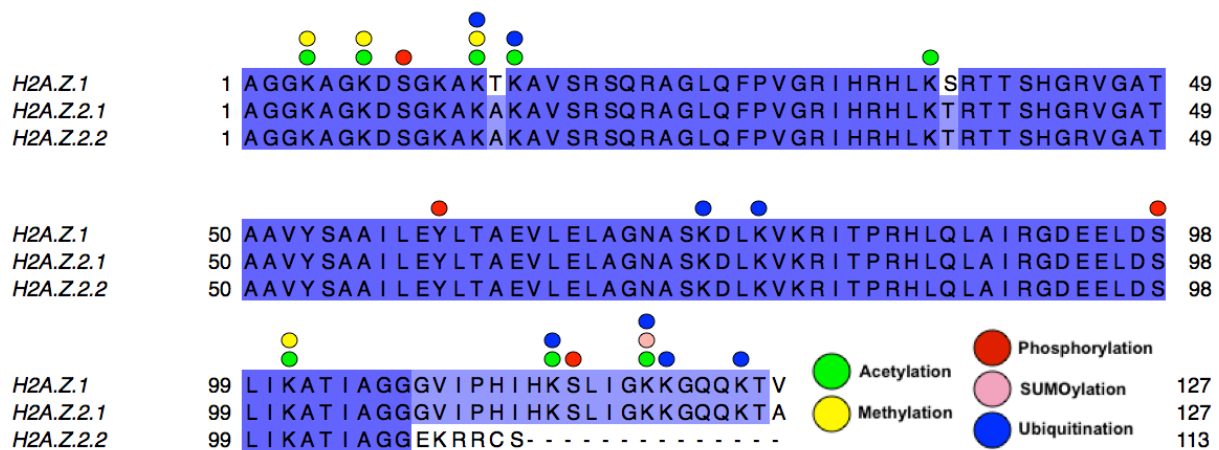


Figure 12: Multiple Alignment of various human H2A.Z isoforms and sites of various PTMs: Alignment of amino acids sequence of human H2A.Z isoforms H2A.Z.1, H2A.Z.2.1 and H2A.Z.2.2 indicate the minor differences of few amino acids. H2A.Z.2.2 specifically lacks the conventional C-terminus sequence. Sites for various types of post-translational modifications (PTMs) including Acetylation (green), Methylation (yellow), Phosphorylation (red), SUMOylation (pink) and Ubiquitination (blue) have also been indicated. The sequence alignment has been performed using Jalview® alignment program.

Histone H2A.Z is the most extensively studied histone variant among all its family members. It exhibits around 60% sequence similarity to that of other cellular H2As (Figure 9). It involves around 10% of all H2A variants. Histone variant H2A.Z remains highly or even probably most conserved histone variant during the course of evolution (Figure 11). This ~90% sequence conservation among diversified species advocate the important role of H2A.Z in critical cellular processes. Though in different species, the analogous of H2A.Z have been given different names such as H2A.F in birds, Htz1 in *Saccharomyces cerevisiae*, H2Av in fruit fly *Drosophila melanogaster*, hv1 in protozoa *Tetrahymena thermophila* and H2AF.Z in sea urchin. A sequence

alignment of H2A.Z or its analogous in different species of evolution has been framed in Figure 11. In humans, H2A.Z is encoded by H2AFZ gene. This gene got this name when cDNA of human H2A.Z (encoded by previously named H2A.Z gene) and chicken H2A.Z (encoded by H2AF) were found homolog to each other.

Histone variant H2A.Z is known to be involved in a variety of important cellular processes including transcription, genomic stability, chromosome segregation and DNA repair etc. H2A.Z has been reported to have contrary role in transcription as it has been associated to both transcription activation and suppression. Plenty of researches have indicated a negative correlation between H2A.Z and transcription (Farris *et al.*, 2005; Hardy *et al.*, 2009). However, on the other hand H2A.Z is primarily associated with increased transcription rate (Santisteban *et al.*, 2000; Gévry *et al.*, 2007). Deletion of H2A.Z from the mouse genome has also been reported to lead to an early embryonic lethality (Binda *et al.*, 2013). Several studies on H2A.Z has indicated its role in chromatin remodeling on DSBs (Papamichos-Chronakis *et al.*, 2011; Xu *et al.*, 2012). This role of histone variant H2A.Z in double strand breaks repair will be discussed in more details in the coming sections.

1.2.4.2.3 MacroH2A variant

The histone variant macroH2A was for the first time discovered in 1990s in rat hepatocytes (Fried, 1992). This is most unusual H2A variant comprised of 327 amino acids with a molecular weight of around 40 kDa, which is about three times higher to that of conventional H2A. The N-terminal domain of macroH2A resembles to that of canonical H2A, however the C-terminal portion that is also designated as non-histone region (NHR) is a macrodomain that resembles to leucine zipper in sequence structure. In mammals macroH2A is encoded by two genes *H2AFY* and *H2AFY2* that encode respectively for macroH2A.1 and macroH2A.2 subtypes. In the human genome, *H2AFY* is present on chromosome 5 while *H2AFY2* is located on chromosome 10. The macroH2A.1 isotype exhibits about 64% sequence similarity with canonical H2A while macroH2A.2 display about 68% sequence similarity to conventional counterpart H2A (Costanzi and Pehrson, 2001).

Diverse range of cellular functions has been attributed to the histone variant macroH2A. Most of these functions are credited to the highly conserved C-terminal macrodomain. macroH2A has been recognized to play significant role in X chromosome inactivation (Costanzi and Pehrson, 1998; Rasmussen *et al.*, 2000; Nusinow *et al.*, 2007). Several *in vivo* studies have also explained the role of macroH2A in transcription repression. This function is specifically attributed to the non-histone region of macroH2A that not only interfere with the binding sites of several transcriptional factors but also inhibit the ATP-dependent chromatin remodeling (Dantzer and Santoro, 2013; Chen *et al.*, 2014). Phylogenetically macroH2A is found in some elementary metazoan species and craniates, however it is absent in nematodes and insects (Talbert *et al.*, 2012).

1.2.4.2.4 H2A.B variant

Histone variant H2A.B is the smallest H2A variant that encompasses a molecular weight of around 12.7 kDa. It was first time identified by Chadwick and Willard in early 2000s (Chadwick and Willard, 2001). It was initially detected in female inactive X chromosome in humans, therefore primarily named as Barr Body Deficient H2A (H2A.Bbd). In humans, histone H2A.B represents two slightly different isotypes namely H2A.B1 and H2A.B2. encoded by three distinct intron-less genes. In sequence comparison to that of canonical H2A, the variant H2A.B exhibits many striking differences with a sequence identity of about 48% only. Being much shorter than counterpart H2A it completely lacks the characteristic C-terminus of conventional H2A and is also deprived of part of docking domain that is responsible for interface with neighboring H4 histone protein (Eirín-López *et al.*, 2008). Moreover, it also lacks the representative acidic patch present in all other H2A family members that is responsible for internucleosomal interaction with other histone members and chromatin condensation. Based on lack of this acidic patch it has also been designated as H2A.Lap1 (lack of acidic patch) (Soboleva *et al.*, 2012; Shaytan *et al.*, 2015).

H2A.B is not found uniformly throughout all the tissues of the body suggesting its tissue specific functions. Histone H2A.B is primarily found in abundance in testicular cells, which indicate its specific role in chromatin remodeling during spermatogenesis. It has also been reported to play important role in elongation of spermatids and even in embryo development subsequently as it remain present in mature cells (Ishibashi *et al.*, 2009; Nekrasov *et al.*, 2013). Several studies have revealed the co-localization of histone H2A.B to that of acetylated histone H4 suggesting its critical function in transcriptional activation. This colocalization explains the moderately less condensed chromatin that may be credited for its role in spermatogenesis. As a result of this highly unstable condensation of nucleosome resulted from H2A.B presence, several cryo-microscopy and atomic microscopy studies have concluded that H2A.B nucleosome is able to wrap only 118-130 base pairs of DNA helix (Bao *et al.*, 2004; Doyen *et al.*, 2006).

1.2.4.3 Histone H2B variants:

Histone H2B along with H2A forms the fundamental heterodimers that serves as the core components for the nucleosome structure. Most of H2B amino acids comprises positive charges which favors their interaction with negatively charged DNA. Though about sixteen variants of histone H2B have been identified in human, but, unlike H2A variants very little is known about the functional implications of histone H2B variants. Such a large number of H2B variants certainly suggests their critical significance in important cellular processes. Out of 16 identified variants in humans, 13 are expressed in somatic cells while the rest of three are expressed specifically in testis (Draizen *et al.*, 2016). These isoforms are encoded by 23 genes that are mainly located in histone cluster 1 present on chromosome 6. Other genes are located on histone cluster 2 and 3 present on chromosome 1. These genes are organized in pairs with that of counterpart H2A sequence and share a common promoter region. Unlike canonical histone H2B which is expressed solely during S-phase of cell cycle in a DNA replication dependent manner, many histone variants are expressed throughout the whole cell cycle.

As these isoforms contain almost the same number of amino acids and exhibits a very little amino acids sequence variation among each other, they are also structurally very similar to each other. Nevertheless, despite having virtually same structure and slightly divergent amino acids sequence they differ a lot in their known functions. For instance, somatic H2B variants in humans that encompass a sequence difference of just two to five amino acids have much diversified functions. In H2B1L, replacement of just single amino acid proline with that of leucine changes the structure of its N-terminal tail and hence affecting the overall nucleosome configuration (Wu *et al.*, 2016). Likewise, in H2B3B variant, serine at position 36 is critical for the regulation of cell stress response as it affects the DNA-histone interaction. The replacement of this S36 with that of alanine and threonine in other variants propose them with different functions (Bungard *et al.*, 2010; Molden *et al.*, 2015).

Testis-specific histone H2B variants have found exhibiting very specialized functions. In sperm cells, when DNA is no longer needs to be accessible for replication and transcription, it makes itself super-condensed to protect from external damaging factors. This modification is achieved through replacing canonical histones to that of their respective variants. For instance, in sperm cells of sea urchin specific H2B variants are shown to additionally bind to the DNA minor groove to make chromosomes more condensed than normal somatic cells (Talbert and Henikoff, 2010). Likewise, in rats, stem-loop of the testis-specific histone variants including that of H2B are considered essential for the protection of the genetic material (Witt *et al.*, 1996). Similar to specialized functions shown by testis-specific variants, some of somatic variants also exhibits tissue-specific functions. For instance, in mice, H2B variant H2BE is hypothesized to have a role in controlling olfactory gene expression (Santoro and Dulac, 2012).

1.2.4.4 Histone H3 variants:

In mammals, variants of histone H3 can be divided into two subcategories. First category is comprised of replicative variants that exhibits a peak expression during the S phase of cell cycle. It includes the replication dependent H3.1 and H3.2 variants and

considered as the key histone source during the DNA replication. The other category of H3 variants are replacement variants. It includes H3.3, H3.4 (H3t), H3.5 (H3.3C), H3.X (H3.Y.2), H3.Y and Centromeric Protein A (CENP-A) histone variants.

These histone H3 variants can also be grouped into somatic (H3.3, H3.X, H3.Y and CENP-A) and testis specific (H3.4 and H3.5) histone variants. With the exception of CENP-A variant, the sequence of most of H3 histone variants are considered highly conserved with a sequence difference of only a few amino acids. Replicative variants H3.1 and H3.2 have about 99% sequence identity. Whereas replacement histone H3.3 has about 96% sequence similarity to that of H3.1 differing at just five amino acids level. However, CENP-A variant shares only around 50% sequence identity to that of canonical H3.1. (Figure 9)

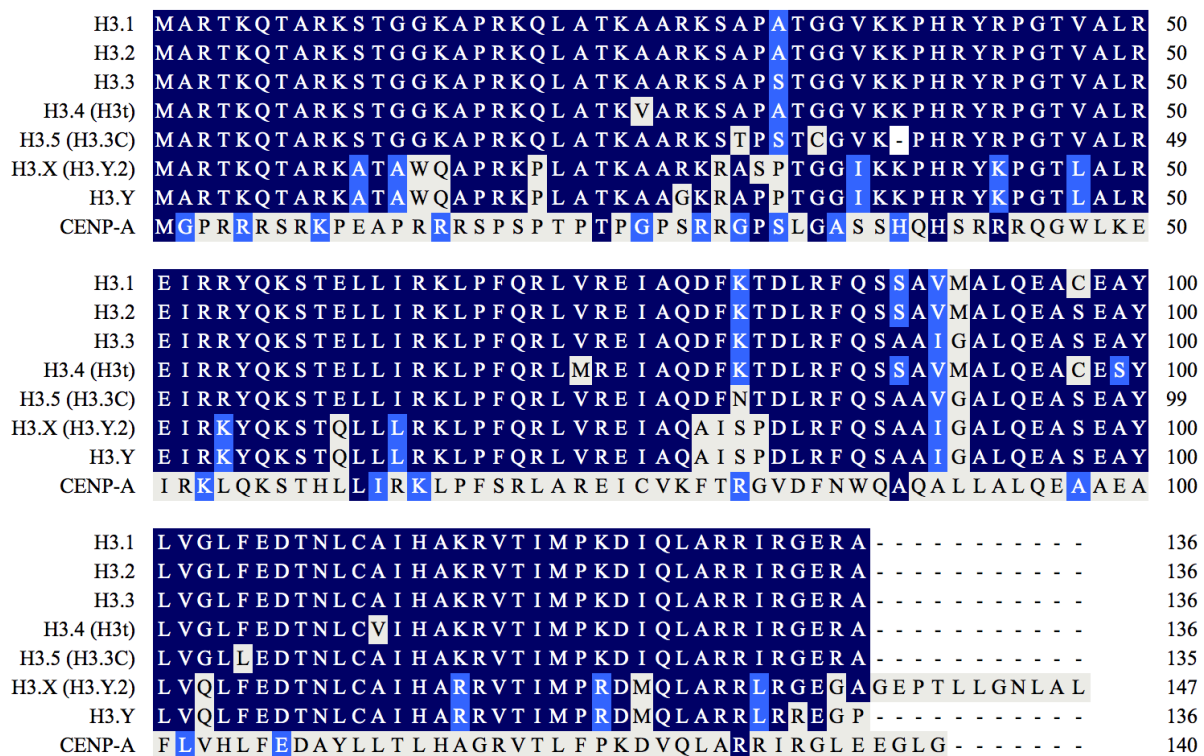


Figure 13: Sequence Alignment of Human H3 Variants: Sequence alignment of various human H3 variants (H3.1, H3.2, H3.3, H3.4, H3.5, H3.X, H3.Y and CENP-A). Identical amino acids among different variants are represented in dark blue. Notice the contrary amino acids of CENP-A from the rest of H3 variants (indicated in grey colour). The sequence alignment is performed using Sequence Manipulation Suite® alignment software.

Among mammals, H3.3 and centromeric H3 variant (cenH3, called also CENPA in humans) are considered as the most common histone variants. Histone

H3.3 encoded by the *H3f3a* and *H3f3b* genes, whose synthesis occurs throughout the cell cycle even outside the S-phase of cell cycle, differs from its respective canonical H3.1 by just five amino acids. It has Alanine, Serine, Valine, Methionine and cysteine instead of Serine, Alanine, Isoleucine, Glycine and Serine at position 31, 87, 89, 90 and 96 respectively (Wells *et al.*, 1987). Though the slight difference of H3.3 to that of H3.1 in amino acid sequences doesn't affect the overall structure of the nucleosome, yet this subtle variance bestows H3.3 nucleosome a low stability hence playing an important role in regulation of gene expression epigenetically. Mutation in H3F3A gene has recently been linked to different sort of cancers including Diffuse Intrinsic Pontine Glioma-DIPG (Wu *et al.*, 2012), chondroblastoma, astrocytomas and giant cell tumour of bone (Buczkowicz *et al.*, 2014; Behjati *et al.*, 2013).

CENP-A, encoded by the CENPA gene in humans, is present in almost all eukaryotes and found specifically in kinetochore domain of centromeric nucleosomes. In the nucleosome core of centromeric chromatin, CENP-A replaces canonical H3 where they play an important role in recruiting many kinetochore proteins (Earnshaw and Rattner, 1989). In eukaryotes, kinetochore exhibits a discrete trilaminar structure which is critical for its specific functions. CENP-A plays a significant role in achieving this specific structure therefore considered crucial for chromosome segregation, cytokinesis and eventually for mitosis. It is the highly conserved histone fold that contains the centromere specific information sequence whereas highly divergent N-terminal tails are designated for binding to other kinetochore proteins. In brief, CENP-A can be designated as an important epigenetic marker of centromeric chromatin regulating the vital cell processes (Folco *et al.*, 2008; Chueh *et al.*, 2004).

1.2.4.5 Histone H4 variants:

As described earlier, histone H4 is one of the key histone members participating in nucleosome structure. It forms with histone H3 a tetramer that establishes the core component of the nucleosome unit. In humans, there is mainly one H4 isoform coded by 14 different genes (HIST4H4, HIST2H4B, HIST1H4I, HIST1H4A, HIST1H4D, HIST1H4F, HIST1H4K, HIST1H4J, HIST1H4C, HIST1H4H, HIST1H4B,

HIST1H4E, HIST1H4L, HIST2H4A) present essentially in cluster 1. Evolutionarily, histone H4 is referred as the slowest evolving histone protein. Because of this highly conserved sequence and structure, H4 is considered as the docking site for other histone members. The evolutionary data indicates a difference of just two amino acids among the histone H4 of pea to cow. In the majority of species, histone H4 has no functional variant. Very little explanation is available in literature archive about this lack of sequence variants. Among few identified variants is H4V which is identified in *Trypanosoma brucei* and suggested to have a significant role in transcription regulation (Siegel *et al.*, 2009). More recently, another H4 variant H4G has been proposed to regulate rDNA transcription in breast cancer (Long *et al.*, 2019).

1.3 Chromatin Remodeling:

As explained earlier that chromatin is a highly dynamic structure that keep changing its organization arrangement to accomplish its diversified roles in the cell. This process of chromatin remodeling involves the ATP-dependent changes in the overall structure of chromatin that result into altered access of different cellular proteins to that of DNA. These changes not only involve the covalent modification of histone machinery but also include the physical rearrangement of nucleosome structure (Vignali *et al.*, 2000; Clapier and Cairns, 2009). This remodeling may involve a range of processes varying from the complete absence of nucleosomes to the change in position of specific nucleosomes to the interchange of histone members. The physical replacement of histone proteins alters the overall structure of nucleosome and hence the chromatin structure that eventually results into a completely different chromatin interactome. This interactome comprised of different proteins involved in critical cellular processes such as DNA replication, transcription, chromosome recombination and even more importantly DNA damage. The energy required for this physical replacement is drawn from the process of ATP hydrolysis (Krebs *et al.*, 2000; Alén *et al.*, 2002).

1.3.1 Histone replacement/exchange:

Histone replacement is the most significant section of chromatin remodeling as it causes the physio-chemical changes at nucleosomes level that is critical for interaction of DNA and important cellular proteins thus affecting many important cellular processes. It involves the physical exchange of conventional canonical histones with that of non-allelic non-canonical variants. Post translational modifications (PTMs) play a critical role in this histone exchange process. Along with other factors, these PTMs contribute in altering the conventional interaction between DNA and core histone thus facilitate chromatin remodelers to play their specific role (Gregory *et al.*, 2001; Goll and Bestor, 2002; Martinowich *et al.*, 2003). The histone remodeling is also facilitated by a variety of histone chaperones. Histone chaperones

are defined as the set of proteins that interact with histone proteins and contribute in the regulation of nucleosome architecture (Loyola and Almouzni, 2004; Burgess and Zhang, 2013). Our lab has recently identified two important histone chaperones namely the acidic nuclear phosphoprotein 32 family member E (ANP32E) and YL1 as critical histone chaperones for H2A.Z replacement (Obri *et al.*, 2014; Latrick *et al.*, 2016). Not all canonical histones are subject to the same level of substitution and exchange by variant histones. The two most replaced histones are H2A by H2A.Z and H3 by H3.3.

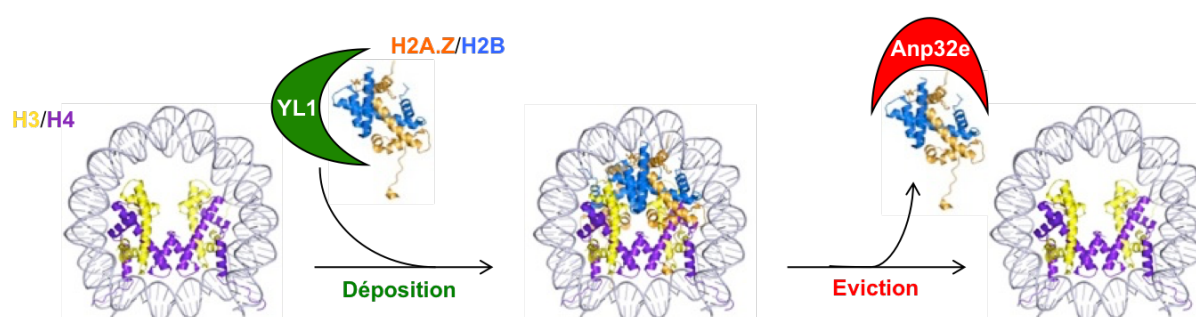


Figure 14: The mechanism of histone variant H2A.Z replacement: YL1 has been identified as specific H2A.Z deposition chaperone. On the other hand, ZID domain of ANP32e, a member of P400-TIP60 has been reported responsible for histone H2A.Z eviction through its interaction with α C helix of H2A.Z.

1.3.1.1 Remodeling of histone variant H2A.Z at DSBs:

The genetic information stored in the genome of the cell should be protected from being damaged to keep cellular integrity and to ensure faithful transmission of the genetic material to daughter cells. Over the period of evolution cells have developed several mechanisms to overturn these damages through its repair. Double strand breaks (DSBs) are considered the most cytotoxic injury as it causes the physical breakage of DNA backbone and results into damage of wide chromosomal regions (Kennedy and D'Andrea, 2006; Jackson and Bartek, 2009; Ciccia and Elledge, 2010). On an average, one dividing mammalian cell undergo about ten DSBs in one day period. Repair of DNA damage starts with rapid establishment of more open and less compact chromatin structure. Histone H4 acetylation by the Tip60 acetyltransferase

has been shown to play a key role in opening chromatin structure in response to DNA damage. This modification extends away from damage site to hundreds of kilobases (Downs *et al.*, 2004; Kusch *et al.*, 2004; Jha *et al.* 2008).

Many researches have provided a wealth of information about the creation of this open and relaxed chromatin structure. It has been now well-established fact that relatively more open and relaxed chromatin structure is achieved through the replacement of H2A with H2A.Z, a process carried out by the p400/Tip60 chromatin remodeling complex in mammalian cells (Fuchs *et al.*, 2001; Gévry *et al.*, 2007) (Figure 15). Several studies have reported that knocking out H2A.Z results into increased sensitivity of cells towards DNA damages (Morillo-Huesca *et al.*, 2010; Xu *et al.*, 2010). Loss of H2A.Z has been reported to cause a defective repair of double strand breaks of DNA (Kalocsay *et al.*, 2009). The rapid exchange of H2A by H2A.Z results in relaxed chromatin domains containing H2A.Z and prone for DNA repair (Downs *et al.*, 2004; Papamichos-Chronakis *et al.*, 2006; Murr *et al.*, 2007).

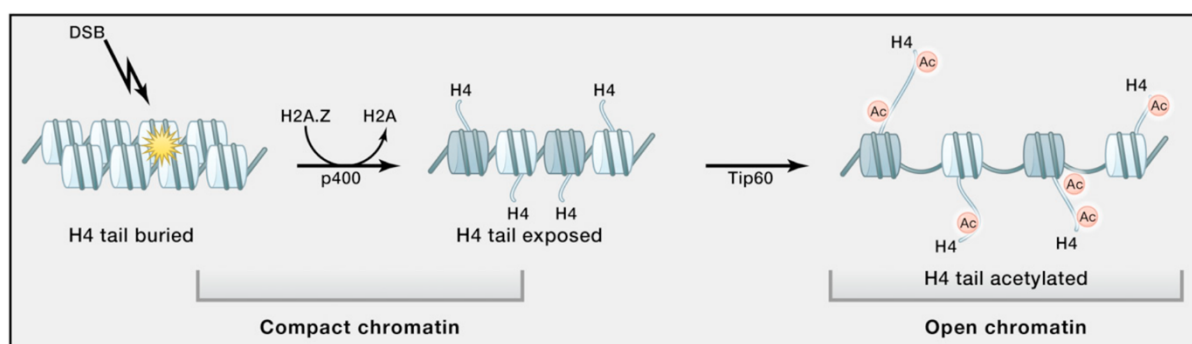


Figure 15: Replacement of H2A.Z with H2A at DSBs contribute in H4 acetylation: The exchange of H2A.Z by p400 at double strand breaks results into a relatively opened and relaxed chromatin structure that favours the H4 acetylation through Tip60. Adapted from (Price and D'Andrea, 2013) (Modified)

There are many theories explaining how the relaxed conformation of chromatin favors the H4 tails acetylation. In general, negatively charged residues of conventional H2A primarily responsible for development of acidic patch on the surface of nucleosome core that provide place for the attachment of N-terminal tails of H4 and hence contribute to the condensation of chromatin. Interchange of H2A with its variant H2A.Z that has an extended acidic patch as compared to H2A, causes a subsequent reduction in 30 nm fiber formation and hence a relatively relaxed chromatin. An

increase in salt solubility of histones after H2A.Z incorporation at DSBs has also been reported in the literature (Xu *et al.*, 2012). This advocate for the formation of a relatively more open and relaxed chromatin structure at sites of DSBs.

Both H2A.Z exchange and acetylation of N-terminal tails of H4 favors the further modification including methylation and ubiquitination of H4 through RNF8/RNF168 hence further modifying the overall nucleosome architecture. These post translational modifications (PTMs) at DSBs level induces the recruitment of DNA repair proteins i.e. brca1 and 53BP1. Both brca1 and 53BP1 proteins then destinies and recruits the machinery of NHEJ and HR repair pathways respectively, which allows DNA repair to occur (Murr *et al.*, 2006; Xu *et al.*, 2010; Courilleau *et al.*, 2012). In this way, both the phenomenon of exchange of H2A.Z and consequent acetylation of H4 tails destinies the DSBs to get repaired through appropriate repair mechanisms and therefore vital for cell survival and genome integrity. The lack of H2A.Z or H2A.Z exchange would deprive the cells to get their DSBs repaired (Figure 16).

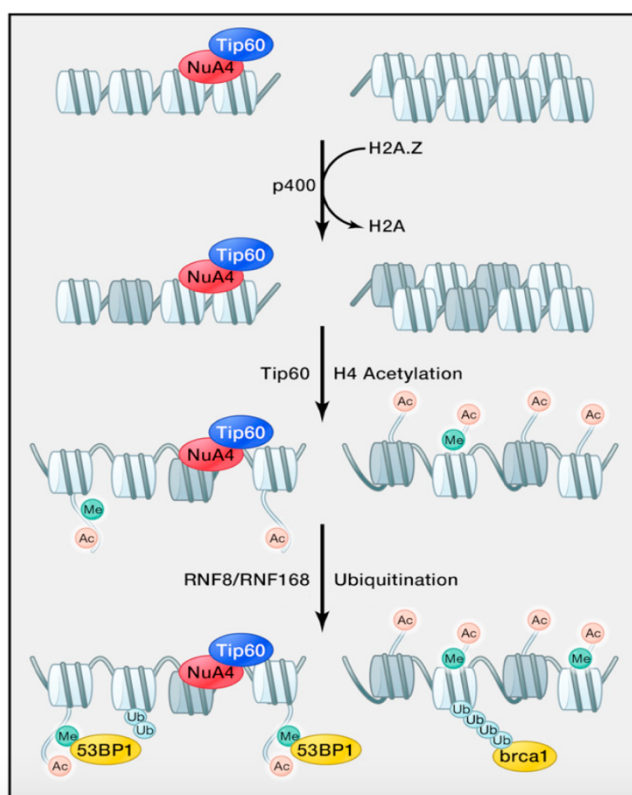


Figure 16: H2A.Z exchange initiate a systematic series of histone modification: Replacement of H2A with H2A.Z by p400 ATPase facilitates the acetylation of N-terminal tails of H4 by Tip60. This acetylation then leads to the ubiquitination of histones by RNF8/RNF168 ligases. Other post translational modifications then also occur correspondingly after this ubiquitination. All these PTMs then act as a scaffold where 53BP1 and brca1 proteins are recruited which then initiates the HR and NHEJ repair pathways. Adapted from (Price and D'Andrea, 2013) (Modified)

1.3.1.2 **Remodeling of histone variant H3.3:**

In mammals, histone variant H3.3 is encoded by two distinctive genes namely *H3f3a* and *H3f3b*. It is primarily found associated with active chromatin hence considered a consistent marker for euchromatic regions (Lin *et al.*, 2013). Remodeling of H3.3 in chromatin contribute to transcriptional activity of the targeted regions. In contrast to other H3 variants like H3.1 and H3.2 which are expressed only during S-phase of the cell cycle, H3.3 is expressed throughout the cell cycle even outside S-phase. H3.3 gradually replaces its canonical counterpart H3.1 in differentiating cells with its preferential place at active genes (Ahmad and Henikoff, 2002). The H3.3 replacement has also been recently linked to transformation of male and female genomes during the process of fertilization and subsequent embryonic development. It is found abundantly in both mature oocyte and more importantly in sperm. In sperm it has been proposed to have a role in pronucleus formation and other important cellular phenomenon (Szenker *et al.*, 2011; Wen *et al.*, 2014).

1.4 The Heterodimer Ku70/Ku80:

Ku is heterodimeric protein that is made up of two subunits Ku70 (~70 kDa) and Ku80 (~83 kDa). In the literature, Ku80 has also been referred to as Ku86 in higher eukaryotes. In humans, Ku70 is encoded by the XRCC6 (X-ray repair complementing defective repair in Chinese hamster cells 6) gene present on chromosome 22, whereas Ku80 is encoded by the XRCC5 (X-ray repair complementing defective repair in Chinese hamster cells 5) gene present on chromosome 2.

Ku heterodimer is a highly abundant protein found primarily in nuclear region of cells (Mimori *et al.*, 1981; Koike *et al.*, 2001). However, under specific condition, the presence of Ku heterodimer in other subcellular regions of cell has been well established (Fewell and Kuff, 1996; Sawada *et al.*, 2003). On an average, about half a million of Ku molecules are considered present in nucleus of a single human cell (Mimori *et al.*, 1986; Walker *et al.*, 2001). Presence in such a huge amount indisputably advocates its important roles in several key cellular processes.

1.4.1 Discovery and History:

The Ku protein was originally identified in 1981 as an autoantigen in the serum of a Japanese patient diagnosed with polymyositis-scleroderma overlap syndrome (Mimori *et al.*, 1981). The name 'Ku' was derived from the first two letters of the surname of that patient. Subsequently, autoantibodies against Ku were identified in many other autoimmune diseases including Sjorgren's syndrome, systemic lupus erythematosus, scleroderma and polymyositis (Reeves, 1985; Francoeur *et al.*, 1986; Cooley *et al.*, 1999; Mimori, 2002). From prokaryotes to eukaryotes, Ku heterodimer is considered highly conserved throughout the evolutionary process and designated to perform special roles in the maintenance of genomic integrity (Bowater and Doherty, 2006). More recently, homologues of Ku have been identified in Archaea and Bacteria domains of life.

Historically, Ku has been labelled and studied with different names such as *Transferrin Receptor Promoter Element Binding Factor* (Roberts *et al.*, 1989; Knuth *et al.*, 1990), *Nuclear Factor-IV* (de Vries *et al.*, 1989; Stuiver *et al.*, 1990), *End Binding Protein* (Falzon *et al.*, 1993), *Proximal Sequence Element 1* (Zhang and Yaneva, 1992) and *Human DNA Helicase II* (Tuteja *et al.*, 1997). Since its discovery in 1981, Ku heterodimer remained overlooked by the scientific community for its critical roles in important cellular process, however, the discovery of its essential role in DNA repair mechanism put Ku70/Ku80 heterodimer in the spotlight.

1.4.2 Biochemistry and Structure:

As mentioned before, Ku heterodimer is composed of two subunits, Ku70 and Ku80. Ku70 is composed of 609 amino acids with a molecular weight of 69843 Da and theoretical pI of 6.23. Whereas, Ku80 is comprised of 732 amino acids with a molecular weight of 82704 Da and theoretical pI of 5.55. Both members have sequential repeats of leucine or leucine with interrupted serine sequences that is characteristic to leucine zipper moiety which are considered as specific structural motif in proteins responsible for their DNA binding feature hence play important role in transcription regulation.

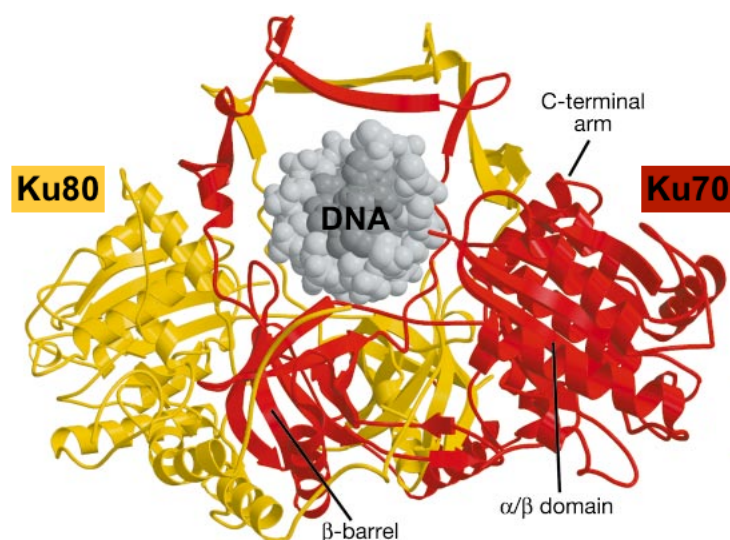


Figure 17: The overview of crystal structure of human Ku heterodimer: Both members of Ku heterodimer, Ku70 and Ku80 together form a basket shaped structure that forms a positively charged ring, a central canal that assist heterodimer to bind with negatively charged DNA. Adapted from (Walker *et al.*, 2001) (Modified)

Crystallographic studies show that both partners construct a typical basket shaped structure that wraps around negatively charged duplex DNA through its positive charge surface (Figure 10). A close relationship in crystal structure of Ku70 and Ku80 can be comprehended by looking at their secondary structure that has been highly conserved during the course of evolution (Figure 11). About 14% sequence identity is found between both subunits of Ku heterodimer (Walker *et al.*, 2001). These aspects evidently advocate that both subunits probably originated from a homodimer that subsequently diverged to form a heterodimer during the course of evolution (Aravind and Koonin, 2001).

In prokaryotes, Ku homologues is approximately 40 kDa and lacks vWA and C-terminal domains present in eukaryotic Ku heterodimer. It contains merely the central core domain responsible for DNA binding (Downs and Jackson, 2004). Though, this Ku homologues exhibits very less sequence resemblance yet it shows a good structural homology to the central core DNA binding domain of eukaryotic Ku heterodimer. The prokaryotic homologue forms a homodimer that binds to the double-stranded DNA ends for repair mechanism (Doherty *et al.*, 2001; Weller *et al.*, 2002).

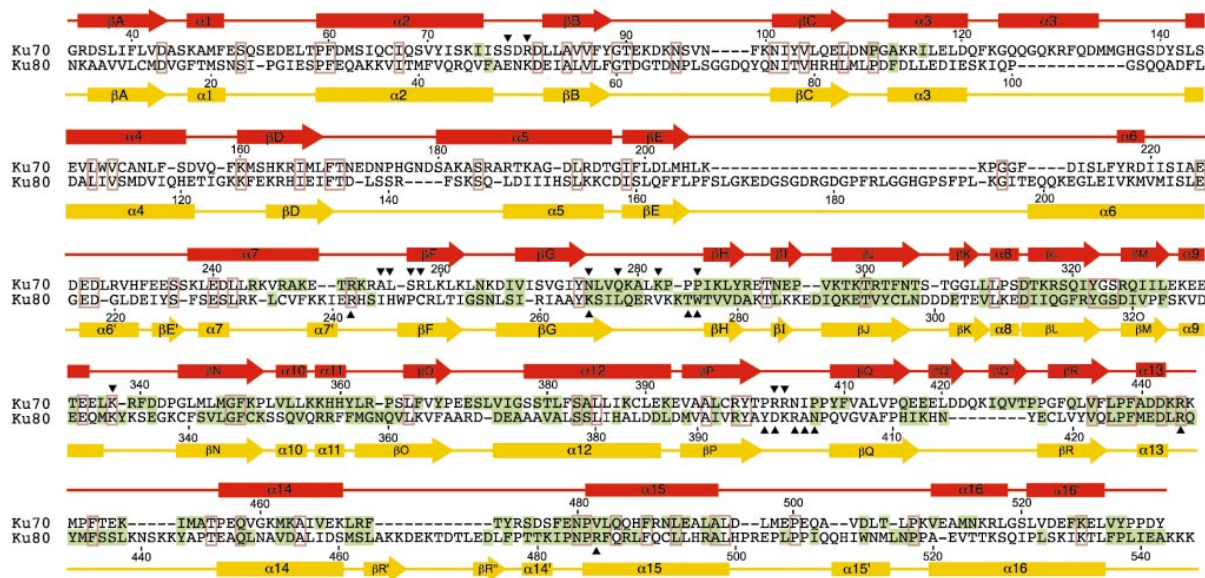


Figure 18: Secondary structure alignment of subunits of Ku heterodimer: The alpha helix and beta barrel units of Ku70 is shown in red while that of Ku80 in yellow. Adapted from (Walker *et al.*, 2001) (Modified)

Scientific community is convinced that Ku subunits are very much unstable as monomers and heterodimerization is indispensable for their correct functioning. In several researches, Ku70/Ku80 has been designated as obligatory heterodimer which determines that the subunits of this heterodimer can't sustain alone in monomeric form. Ku70 deficient embryonic stem cells have shown very low levels of Ku80 as well (Gu et al., 1997), whereas Ku80^{-/-} cells exhibits a reduced level of Ku70 (Errami et al., 1996; Singleton et al., 1997). A similar pattern has also been witnessed in yeast cells as well (Downs and Jackson, 2004).

1.4.2.1 Structural Domains of Ku70/Ku80:

Both Ku subunits Ku70 and Ku80 make a toroidal structure that wraps itself around the DNA to perform specific functions (Figure 10). Both crystal and functional studies suggest that both subunits are primarily composed of three main domains. The N-terminal domain, also known as von Willebrand A (vWA) domain, contain both alpha helix and beta barrels in their structure. The name vWA is given to this domain because of its protein prototype von Willebrand Factor type A, mainly extracellular matrix proteins that shows very vital roles in cell adhesion and interaction of cell with broad range of ligands (Whittaker and Hynes, 2002). The central β -barrel core domain is responsible for Ku70 and Ku80 dimerization and DNA binding. The C-terminal domain is comprised of three alpha helices in Ku70 and six alpha helices in Ku80 (Jones *et al.*, 2001).

The N-terminal domain of both subunits forms a Rossman fold that is composed of six stranded beta sheets. The amino-terminal end of vWA domain interacts with the DNA groove, although it contributes little to DNA binding. On the other hand, the carboxyl end of this Rossman fold protrudes out that frame and it is accessible for interaction with other cellular proteins. Over the last few years, this vWA domain has emerged as a notably important domain playing crucial roles in telomere regulation and DNA repair mechanism in different eukaryotic species. For instance, in mice, mutations in vWA domain of Ku70 at helix 5 leads to the interruption in double strand breaks (DSBs) repair mechanism as it is considered

necessary for the recognition of DNA damage induced by ionizing radiation (Ribes-Zamora *et al.*, 2013). This recognition is the primary function of Ku heterodimer which allows the DNA repair mechanism through NHEJ pathway to start (details in coming sections). Similarly, helix 4 of vWA domain of Ku70 is considered crucial for DNA damage response that leads to apoptosis in human cells. In MEF cells, mutation at position S155 and D156 of vWA domain of Ku70 represses the activation of transcriptional factor 2 (ATF2) and leads to an increased cell survival rate. Modification of S155 with A155 confer MEF cells this enhanced survival rate. This clearly indicates the importance of S155 in activation of ATF2 that consequently leads the cell to apoptosis (Fell and Schild-Poulter, 2012).

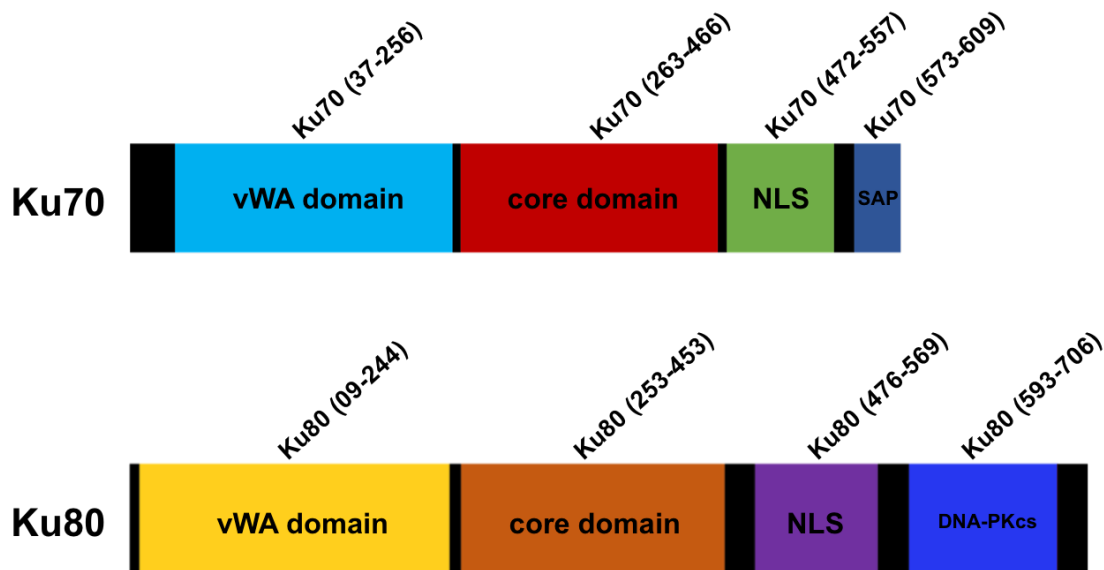


Figure 19: Schematic illustration of Ku70 and Ku80 Primary Structure: Both Ku70 and Ku80 consist of an N-terminal von Willebrand factor type A (vWA) domain, a central DNA binding core domain and a nuclear localization sequence (NLS) domain. C-terminal domain of Ku70 comprises a SAP (SAF-A/B, Acinus and PIAS) domain while C-terminal domain Ku80 contains the sequence responsible for binding to DNA-PK_{CS}.

In the same way as the vWA domain of Ku70, the N-terminal vWA domain of Ku80 is also considered important for interaction with other crucial cellular proteins like telomere repeat binding factor 2 (TRF2) and Aprataxin and PNK-like factor (APLF). Mutations in vWA domain of Ku80 in *Saccharomyces cerevisiae* at position

L240, Y49, M16I and L149 are considered responsible for the interruption of telomere silencing by affecting the interaction of Ku with TRF2, highlighting the importance of this domain in telomere maintenance (Ribes-Zamora *et al.*, 2007). Similarly, mutation in vWA domain of Ku80 at L68, Y74 and I112 leads to a reduced interaction of Ku heterodimer to APLF protein hence altering the DNA repair mechanism (Grundy *et al.*, 2013; Shirodkar *et al.*, 2013).

As mentioned earlier, Ku70 and Ku80 heterodimerizes with each other through the central core domain of both subunits and forms a toroidal ring structure that interacts with DNA helix through both of its major and minor groves. One Ku heterodimer is estimated to wraps around approximately two turns of double-stranded DNA. The inner side of this ring is lined up with positively charged amino acids of both Ku70 and Ku80 hence cause electrostatic interaction with that of negatively charged DNA in a sequence-independent manner (Mimori and Hardin, 1986; Yoo *et al.*, 1999). Quantitative binding analyses shows that the interaction of Ku heterodimer with DNA is a result of high electrostatic affinity with a dissociation constant $K_d \sim 10^{-9}$ M. that is primarily limited to double stranded DNA as it has considerably lesser affinity for single-stranded DNA, circular DNA and even for supercoiled double stranded DNA (Paillard and Strauss, 1991; Ono *et al.*, 1994). While the positively charged inner surface of Ku ring forms contact with the phosphate groups of the DNA backbone, the electrostatic affinity facilitate the easy sliding of heterodimer over a double stranded DNA end. Once electrostatically attached to the DNA ends, the Ku heterodimer move along the whole DNA molecule in an ATP-independent manner (Chan *et al.*, 2002; Douglas *et al.*, 2002). While bound upon DNA ends or breaks both the Ku70 and Ku80 subunit have a specific orientation where Ku70 is positioned proximally from the DNA end while Ku80 is placed distally to the end of DNA helix.

The C-terminal domain (CTD) domain of both Ku70 and Ku80 are quite dissimilar to each other however in general both CTDs contain alpha helices in their secondary structure. In Ku70, the CTD is composed of three alpha helix bundle that forms a highly flexible linker arm followed by a ~5kDa SAP domain. The name 'SAP' is given to this domain based on the initials of the three proteins SAF-A/B, Acinus and PIAS that contain the same motifs. In the literature, this SAP domain is well known

for its DNA binding feature that strengthen the overall DNA binding affinity of Ku (Aravind and Koonin, 2000; Hu *et al.*, 2012). This CTD of Ku70 is also considered prone to some important post-translational modifications (PTMs) like SUMOylation and acetylation that make this domain very crucial for epigenetic phenomenon (Cohen *et al.*, 2004; Kim *et al.*, 2014; Hang *et al.*, 2014). In vertebrates, the C-terminal domain of Ku80 is composed of six alpha helices in which $\alpha 2$ and $\alpha 4$ form a hydrophobic pocket that is responsible for its protein-protein interaction (Harris *et al.*, 2004; Zhang *et al.*, 2004). This CTD of Ku80 is considered responsible for the recruitment of DNA-dependent protein kinase catalytic subunit (DNA-PK_{CS}) at the site of double strand breaks (DSBs) that initiates the NHEJ pathway of DNA repair (Gell and Jackson, 1999; Singleton *et al.*, 1999). However, since DNA-PK_{CS} is not present in lower eukaryotes the CTD of Ku80 in lower eukaryotes like *Aradopsis thaliana* and *Saccharomyces cerevisiae* lacks this DNA-PK_{CS} binding region (Shrivastav *et al.*, 2008).

1.4.3 Functions of Ku heterodimer:

Both Ku70 and Ku80 subunits have been cloned and suggested to have several important cellular functions. Based on several direct and indirect observations, it is obvious that Ku is not merely a DNA binding protein it rather has a wide range of functions in diverse biological processes.

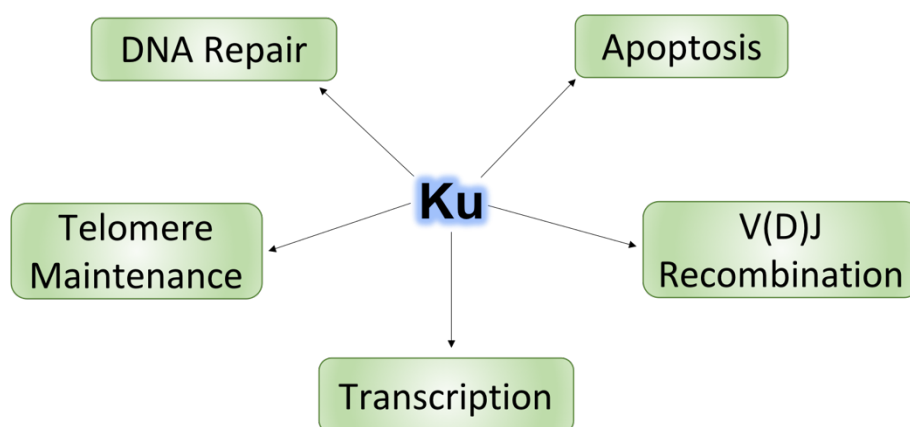


Figure 20: The cellular functions of Ku heterodimer: Ku70/Ku80 exhibits a wide range of important cellular activities. Many researches have now implicated Ku as more than a DNA binding protein since it also has very crucial roles in other different biological processes.

Though Ku is considered primarily a nuclear protein yet many studies have witnessed the non-nuclear roles of Ku heterodimer as well. Not only Ku heterodimer is critical for the genomic integrity of cell but it is also important for proper cellular and embryonic development. It has also been implicated to have specific roles in different diseases and phenomenon including cancer, aging and immune system disorders.

1.4.3.1 Ku Heterodimer in DNA Repair:

Maintaining the architecture of genetic information is of high importance for the normal cellular and organismal survival. The living cells are constantly exposed to a variety of genotoxic agents that can cause negligible effects to the lethal outcomes including the death of cells. These cytotoxic agents can be categorized into external factors like environmental toxins, ultraviolet radiation, chemotherapy, viral infections and ionizing radiation, or internal factors including replication stress and natural metabolites like reactive oxygen species etc. These factors cause a wide range of DNA damages comprising damaged bases and damaged sugar residues, development of inter-strand and intra-strand crosslinks, mismatches of bases, and more crucially single and double strand breaks (Figure 14). Among these damages, double strand breaks (DSBs) are considered most cytotoxic as it causes the physical breakage of DNA backbone and results into damage of wide chromosomal regions. On an average, one dividing mammalian cell undergo about ten DSBs in one day period (Martin et al., 1985; Lieber and Karanjawala, 2004).

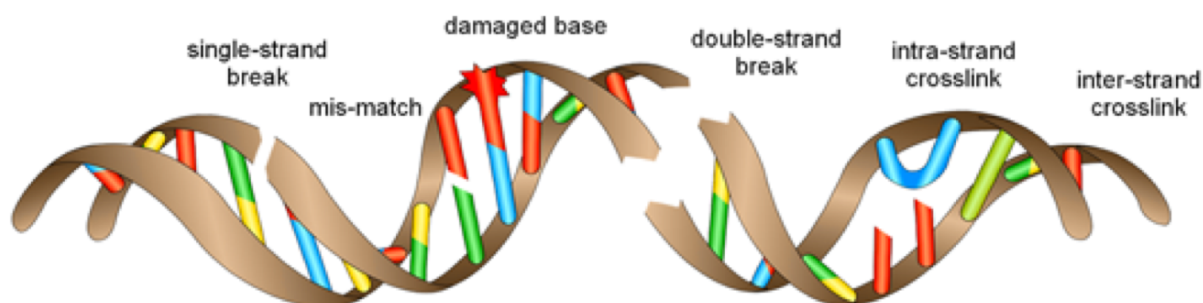


Figure 21: Types of various DNA damages: Genetic machinery is constantly exposed to variety of harmful factors. These factors cause a wide range of DNA damages that ranges from single base lesions to catastrophic breakage of DNA strand. Acquired from (Oxford University Press, 2012).

To confront these damages cells have evolved about six specialized repair mechanisms. In eukaryotes, two main repair mechanism for treating DSBs include Homologous Recombination (HR) and Non-Homologous End Joining (NHEJ) (Deriano and Roth, 2013; Frit *et al.*, 2014). In mammals, the DNA damage repair is predominantly made through the NHEJ pathway since it has the potential to resume throughout the cell cycle in comparison to HR which is only active during the S and G2 phase of cell cycle. Another preference of NHEJ over HR is because of fact that NHEJ can ligate any two DNA fragments in a sequence independent manner in comparison to HR that needs a complementary sister chromatid to execute repair process (Grundy *et al.*, 2014, Williams *et al.*, 2014).

The Non-Homologous End Joining (NHEJ) repair pathway starts with the binding of Ku molecules at the site of DSBs. Ku is present in large quantity in the cells with an average of 0.5 million molecules per human cell (Blier *et al.*, 1993; Mari *et al.*, 2006). However, many *in vivo* studies have suggested that very few Ku molecules (one to two) load at the site of a particular DSB (Roberts and Ramsden, 2007; Britton *et al.*, 2013). The Ku molecules then causes the recruitment of DNA-PK catalytic subunit (DNA-PK_{CS}), an important component of the phosphoinositide 3- kinase-related kinase (PIKK) family. This recruitment is accomplished by the C-terminal domain of Ku80 however there are evidence suggesting that Ku heterodimer makes numerous contacts with DNA-PK_{CS} (Rivera-Calzada *et al.*, 2007). Ku in association with DNA-PK_{CS} forms DNA-dependent protein kinase (DNA-PK) complex considered very critical for effective DNA repair. Loss of any member of DNA-PK complex, either Ku or DNA-PK_{CS}, causes the complete arrest of NHEJ repair pathway (Rivera-Calzada *et al.*, 2007; Lovejoy and Cortez, 2009). This strong connection between Ku and DNA-PK_{CS} ultimately results into the activation of DNA-PK_{CS} and subsequent translocation of Ku molecule from the immediate site of DSB. This translocation of Ku heterodimer permits DNA-PK_{CS} to be arranged at the terminal points of DSBs and mediate the synaptic assembly of both broken ends (Costantini *et al.*, 2007; Hammel *et al.*, 2010). However, with increasing evidences many researches have witnessed that Ku itself through its enzymatic activity also directly instigates the bridging of two broken DNA ends (Cary *et al.*, 1997; Ramsden and Gellert, 1998).

Based on the type of damage, the production of compatible ends and final repair of damaged termini may need some specific enzymes to be recruited at the site of DNA damage. It includes phosphatases, nucleases (Mre11, ExoI, WRN, Artemis), helicases (RECQ1), kinases (PNK), phosphodiesterases and polymerases (DNA polymerase λ , DNA polymerase μ), etc. Several studies have advocated that Ku act a scaffold for the recruitment of most of these enzymes at the site of DNA damage (Karmakar *et al.*, 2002; Mahajan *et al.*, 2002; Ma *et al.*, 2004; Costantini *et al.*, 2007; Shirodkar *et al.*, 2013).

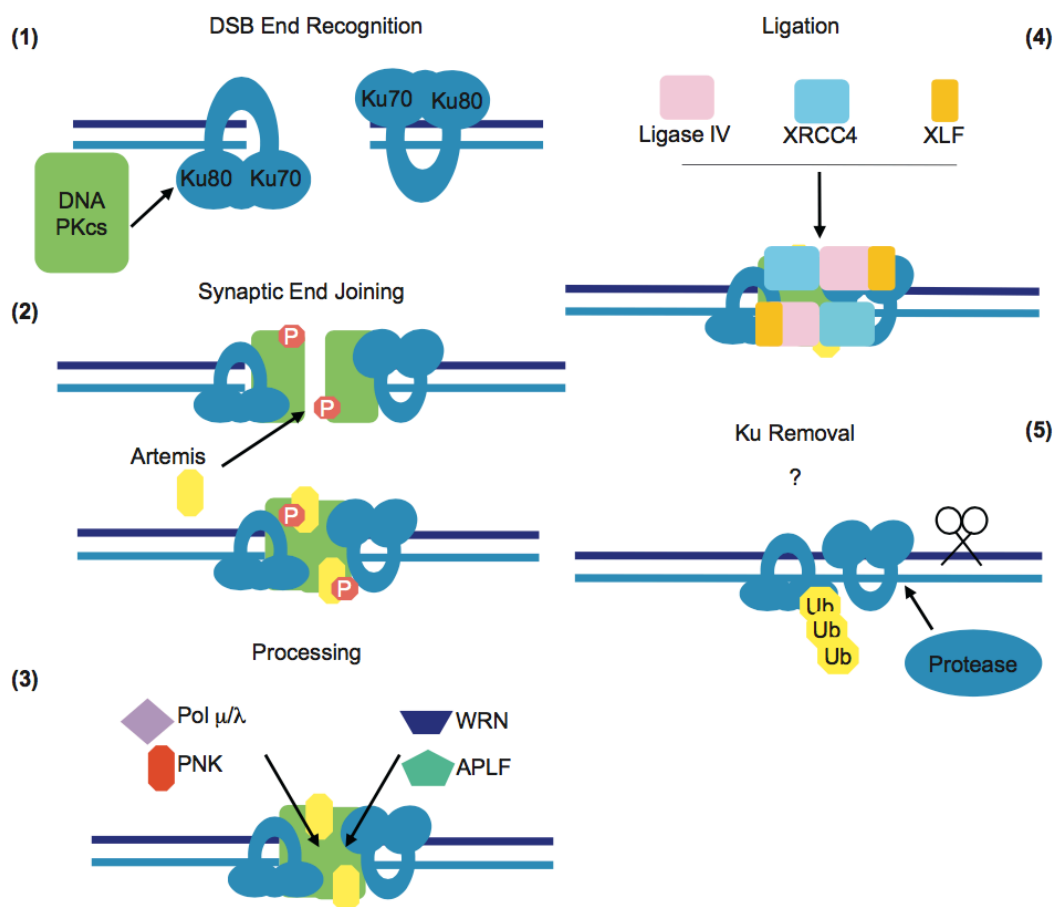


Figure 22: Schematic overview of different stages of NHEJ repair pathway: (1) Ku70/80 heterodimer recognize the DSB and recruit DNA-PKcs at damage site to form DNA-PK complex. (2) The autophosphorylation of DNA-PK causes the production of synapse between the broken DNA ends and recruits Artemis at DSB. (3) Proper bridging of ends needs the creation of compatible ends for ligation that needs the recruitment of different enzymes at damage site by the Ku heterodimer. (4) Ligation of canonical ends is performed by recruiting the ligation machinery at DSB. (5) Ku complex needs to be removed from repaired site through proteolytic degradation of Ku complex by different proteases and ubiquitinases. Diagram acquired from (Fell and Schild-Poulter, 2015)

Following the formation of compatible DNA ends, the ligation machinery including XRCC4, DNA ligase IV and XLF are recruited at the site of DSB. This recruitment is primarily believed to be performed by Ku70/Ku80 heterodimer (Hsu *et al.*, 2002; Costantini *et al.*, 2007; Yano *et al.*, 2008). The assembly of NHEJ machinery is very much dynamic. Once the damage site is repaired all the members needs to be removed from intact DNA. There have been many theories in literature archive regarding the removal of Ku heterodimer from the site of DSB. One possible mechanism suggests the ubiquitination of L48 of Ku80 that leads to ultimate degradation of Ku heterodimer (Postow *et al.*, 2008; Feng and Chen, 2012). Other studies propose the direct nicking of DNA that facilitates Ku emission from DSB site (Neale *et al.*, 2005; Wu *et al.*, 2008). Ku heterodimer is not only playing its significant role in NHEJ pathway of DNA repair rather many other studies have suggested its critical roles in other repair pathways including BER.

1.4.3.2 Ku Heterodimer in V(D)J Recombination:

V(D)J recombination is a mechanism that occur in the primary lymphoid organs through which both T cells and B cells randomly arrange variable (V), joining (J), and in some cases, diversity (D) gene segments to produce novel antigen receptors. These unique amino acid sequences in the antigen-binding regions of immunoglobulins (Igs) make them recognize a wide range of antigens from different pathogens including bacteria, viruses, and parasites etc. This V(D)J recombination occurs in thymus for T cells and in bone marrow for B cells (Malu *et al.*, 2012).

The phenomenon of V(D)J recombination starts with the development of sequence specific double strand breaks at the terminal points of two coding sections that are intended to be joined later. This breakage results into the formation of hairpin structure where desired sequence segments with blunt ends are added. It is believed that subsequent repair and processing of these structures is performed through NHEJ system where Ku heterodimer is considered to play a crucial role in binding to the hairpin structure of genomic coding regions to that of blunt ended DNA molecules (Ouyang *et al.*, 1997). This role of Ku heterodimer was for the first time identified

when cells deficient in V(D)J recombination showed low levels of Ku80 expression (Taccioli *et al.*, 1994). In subsequent studies Ku70 subunit was also considered involved in this NHEJ repair events during V(D)J recombination (Gu *et al.*, 1997). Many immune system abnormalities like arresting of T-cells, arrest of B-cell development, and failed lymphocyte differentiation have been reported in Ku70 knockout mice (Gu *et al.*, 1997; Ouyang *et al.*, 1997; Manis *et al.*, 1998). Though many researches have also indicated DNA-PKcs as the main player in these mechanisms yet Ku alone has also been witnessed many times for its significant role in V(D)J recombination. Most of these functions is primarily performed by Ku80 in particular (Li *et al.*, 2013)

1.4.3.3 Ku Heterodimer in Apoptosis:

Apart from its major roles in nuclear region, Ku heterodimer also exhibits some non-nuclear, cytosolic functions. An apoptotic protein named BAX, also known as BCL-2-like protein 4 is an important apoptosis regulator that forms a heterodimer with BCL2 and cause the opening of the mitochondrial voltage-dependent anion channel (VDAC). This causes the release of cytochrome C along with other pro-apoptotic factors from mitochondria that then leads to membrane permeabilization of outer membrane of mitochondria which then ultimately leads to apoptotic death of cell (Oltval *et al.*, 1993). Ku70 has been identified to bind this pro-apoptotic protein and to inhibit BAX-mediated apoptosis by inhibiting relocalization of BAX to the mitochondria. This function is considered mediated by C-terminal domain (CTD) of Ku70 (Sawada *et al.*, 2003). Recent studies on cancerous cells especially lung cancer cells have established that tumor cells use the interaction between Ku70 and BAX to elude cellular apoptosis (Jing-Jing *et al.*, 2013).

1.4.3.4 Ku Heterodimer in Telomere Maintenance:

Telomeres are the physical ends of chromosomes that are composed of repetitive nucleotide sequences. These telomeres protect the linear ends of

chromosomes from degradation. They have very high potential to be recognized as double strand breaks and hence can be approached by the DNA repair complexes. Telomeres not only help these linear ends of chromosomes to confront its replication difficulties but also provide help in protecting their ends from the access of DNA repair machinery (Price *et al.*, 2010; Longhese *et al.*, 2012).

The role of Ku in maintaining telomeres was the first time identified in *Saccharomyces cerevisiae*, where the depletion of either Ku members resulted into shorter length telomeres advocating the crucial role of Ku in telomere conservation (Boulton and Jackson, 1996; Porter *et al.*, 1996; Gravel *et al.*, 1998). Likewise, in a research on human somatic cells, Ku80^{-/-} cells revealed unusual telomere structure that supported the fact that Ku is involved in repression of fatal telomeric deletion (Wang *et al.*, 2009). Another way of Ku to help in telomere maintenance is through its interaction with the telomerase. Specific reverse transcriptase termed telomerase is considered responsible for the replication of telomeres. Ku has been reported to help in the recruitment of the telomerase and hence help in lengthening of telomeres (Stellwagen *et al.*, 2003; Fisher *et al.*, 2004).

In addition to important role of Ku in telomeric integrity through its physical protection, Ku is also considered involved in the mechanism of telomere position effect (TPE). TPE is a gene silencing phenomenon observed in many eukaryotic species, through which genes located near the telomere are made transcriptionally silenced. In *Saccharomyces cerevisiae*, Ku heterodimer together with Silent Information Regulator (SIR) are considered indispensable for TPE. Removal of either of Ku subunits results in total loss of telomeric silencing (Aparicio *et al.*, 1991; Boulton and Jackson, 1998; Laroche *et al.*, 1998)

1.4.3.5 Ku Heterodimer in Transcription:

Ku has been reported as an important transcription factor in several scientific studies. In a study investigating the subnuclear localization of Ku protein, Ku has been found associated with RNA polymerase II (Pol II) elongation sites. This association has been explained through the protein-protein interaction of Ku80 with that of other

transcriptional proteins (Woodard *et al.*, 2001, Mo and Dynan, 2002). Ku has also been reported to arrest the RNA polymerase-1 (Pol-I) mediated transcription (Kuhn *et al.*, 1995). A probable mechanistic explanation of this function is explained through the interaction of Ku with DNA-PK_{CS}. Both members jointly cause the phosphorylation of the transcription machinery members. Moreover, Ku has also been reported to interact with many other cellular proteins including homeodomain proteins (Schild-Poulter *et al.*, 2001), Redox Factor Protein-1 (Chung *et al.*, 1996), Epithelium-Specific ETS-1 (Wang *et al.*, 2004), Heat Shock Transcription Factor 1 (Huang *et al.*, 1997), CTC- Box Binding Factor (Genersch *et al.*, 1995), TATA-Binding Protein (Hoff *et al.*, 1994) and U5 Repressive Element (Okumura *et al.*, 1996), Yolk Protein Factor 1 (Jacoby and Wensink, 1994) etc., that affect the overall transcription process of the cell. On the other hand, some Ku-like protein like Enhancer 1 Binding Factor (Zhang and Jacob, 1990) and Constitutive HSE Binding Factor (Li *et al.*, 1995) also show their effects at different levels of cellular transcription.

2. OBJECTIVES OF STUDY

The highly dynamic chromatin structure keeps changing its configuration in order to accomplish its multifarious roles in the cell. This change in organizational design is accomplished not only through the covalent modification of histone machinery but also include the physical rearrangement of the nucleosome structure. Among these ATP-dependent changes, histone replacement is the most significant segment of chromatin remodeling as it causes the physio-chemical changes at nucleosomes level that is critical for its interaction with DNA and several important cellular proteins, therefore, affecting many critical cellular mechanisms. It involves the physical exchange of conventional canonical histones with that of counterpart non- allelic, non-canonical variants. Among histone remodeling, H2A.Z remodeling has got huge attention of the research community in the recent past by virtue of its critical role in a variety of important cellular processes including transcription, genomic stability, chromosome segregation, and very importantly in DNA repair. In the literature, a wealth of information has been provided regarding the mechanism of H2A.Z remodeling however with the advent of new emerging techniques to comprehend such dynamic phenomenon, new members are being unveiled playing a significant role in these mechanisms.

Here, we set out to investigate the interaction of Ku heterodimer and histone variant H2A.Z and the probable role of this interaction in H2A.Z deposition in chromatin. Our lab has previously witnessed the presence of Ku70 (XRCC6) and Ku80 (XRCC5) along with other DNA repair members in the H2A.Z soluble complex during one of our work on histone chaperone ANP32E. As together with various significant mechanistic roles by Ku heterodimer and H2A.Z, both have been independently designated in literature archive to undergo important functionality in DNA repair mechanisms. This correlation impelled us to explore the molecular basis of the interaction between Ku heterodimer and H2A.Z by employing different *in-vivo* and *in-vitro* biochemical and genomic approaches as well.

We will first review the specificity of interaction of Ku heterodimer with histone variant H2A.Z in comparison to its counterparts H2A and H2A.X using tandem-affinity purification methods. Subsequently, we will proceed towards the further

characterization by mapping the minimal Ku70 or Ku80 domain involved in this interaction. Later, we'll attempt towards the determination of H2A.Z domain critical for this specific interaction by using chimeric constructs of H2A.Z and counterpart H2A. Following the validation of interaction of Ku heterodimer and histone variant H2A.Z through co-immunoprecipitation and tandem affinity purification-mass spectroscopy (TAP-MS) approaches, we will attempt to validate the interaction through size-exclusion chromatographic technique.

Finally, we will proceed towards the evaluation of any possible functional implications of this interaction through genome-wide approaches using mouse embryonic fibroblasts. As anticipated, if the interaction between Ku heterodimer and histone dimer H2A.Z/H2B has some substantial engagement, knocking out the Ku heterodimer from cells should provide some piece of evidence in genome-wide alteration of H2A.Z deposition or eviction. With these means, we will be able to establish the probable noteworthy role of Ku heterodimer in H2A.Z remodeling and its functionality.

3. MATERIALS & METHODS

3.1 General Reagents:

The general chemicals and reagents used during the research are mentioned in table 2 along with their supplier details. All reagents were used and preserved as per standard recommendations of supplier and literature archive. Standard precautions and guidelines were followed while using all the reagents and disposal of surplus if needed was made as per recommendations of Prevention and Health & Safety Organization, IGBMC, University of Strasbourg.

Reagent Name	Brand	Reference No.
1,4-Dithiothreitol (DTT)	Sigma Aldrich®	3483-12-3
2-Propanol	Sigma Aldrich®	67-63-0
β-Mercaptoethanol	Fluka®	60-24-2
Acrylamide (40%)	MP Biomedicals®	79-06-1
Acetic Acid_Glacial	Sigma Aldrich®	1099510001
Agarose LE (gel electrophoresis)	Fisher Scientific®	BP160-100
Ammonium Acetate	Sigma Aldrich®	631-61-8
Ammonium Persulfate (APS)	Merck Millipore®	7727-54-0
Ampicillin	Euromedex®	EU0400-A
Bromophenol Blue	Sigma Aldrich®	115-39-9
BSA (bovine serum albumin) Fraction V	MP Biomedicals®	9048-46-8
Calcium Chloride	Sigma Aldrich®	10035-04-8
DAPI	Sigma Aldrich®	28718-90-3
DMEM Media	Gibco™	11500416
DMSO (Dimethyl Sulfoxide)	Thermo Fisher®	67-68-5
dNTP mix	Roche™	KK1007
EDTA	Euromedex®	EU0007-A
Ethanol	VWR Chemicals®	64-17-5
Ethidium Bromide	Euromedex®	EU0170

Foetal Bovine Serum (FBS)	Sigma Aldrich [®]	-
Glycerol (87%)	GE Healthcare [®]	GE17-1325-01
Glycogen	Roche [™]	10901393001
Imidazole	Sigma Aldrich [®]	288-32-4
Isopropyl β -D-1-thiogalactopyranoside (IPTG)	Euromedex [®]	EU0008-A
Kanamycin	Euromedex [®]	EU0420-A
L-Glutathione	Sigma Aldrich [®]	70-18-8
Lysozyme (from chicken egg white)	Sigma Aldrich [®]	12650-88-3
Magnesium Chloride	Merck Millipore [®]	7786-30-3
Methanol	VWR Chemicals [®]	67-56-1
Morpholinepropanesulfonic acid (MOPS)	Sigma Aldrich [®]	1132-61-2
NP-40	Sigma Aldrich [®]	9016-45-9
Phenylmethylsulfonyl Fluoride (PMSF)	Sigma Aldrich [®]	329-98-6
Phosphate-Buffered Saline (PBS)	Hyclone [™]	SH3025601
Potassium Chloride	Sigma Aldrich [®]	7447-40-7
Protease Inhibitor Cocktail (cOmplete [™])	Roche [™]	11836153001
Proteinase K	Thermo Fisher [®]	AM2546
SDS 20%	Euromedex [®]	EU0660
Sodium Acetate	Sigma Aldrich [®]	127-09-3
Sodium Chloride	Carlo Erba [®]	7647-14-5
Sodium Hydroxide	Sigma Aldrich [®]	1310-73-2
Spermidine	Sigma Aldrich [®]	124-20-9
Spermine	Sigma Aldrich [®]	71-44-3
Tetramethylethylenediamine (TEMED)	MP Biomedicals [®]	110-18-9
Tris Base (Trizma [®])	Sigma Aldrich [®]	77-86-1
Triton X-100	Sigma Aldrich [®]	9002-93-1
Trypsin	Sigma Aldrich [®]	T4049

Table 2: List of General Reagents Used: All reagents were preserved as per standard recommendations of supplier to maintain their efficacy and reliability. Solutions where needed were also prepared and preserved as per customary guidelines. Personal protective equipment (lab coats, gloves, goggles, masks) were used wherever recommended to avoid any health concern.

3.2 Cloning Strategies:

In order to express the recombinant protein in *E-coli* expression system, the gene of interest has to be transferred into a plasmid vector which then subsequently transformed into the bacterial cells. Restriction endonuclease dependent cloning is a traditional cloning strategy of plasmid cloning that is very common in molecular laboratories due to its convenience and promptness. During this research, all the requisite plasmids were prepared exclusively in-house as per standard protocols.

3.2.1 Designing DNA oligos:

All DNA oligos were designed manually or using **Primer3Plus®** program. All oligos were ordered to **Sigma Aldrich®** for preparation. DNA oligos received were dissolved in TE buffer to get final concentration of 100 μ M. Specific restriction sites were added to oligos to make PCR product capable to be ligated in plasmid vectors.

3.2.2 Polymerase Chain Reaction:

The genes of interest were amplified using gene specific primers through PCR technique. *Phusion™ High-Fidelity DNA Polymerase* kit by Thermo Scientific® (reference: F530S) was used for all the PCR experimentation with following thermocycling conditions:

Step	Temperature	Time	Cycles
Initial Denaturation	94°C	5 mins	1
Denaturation	94°C	30 sec	30
Primer Annealing	58°C	30 sec	
Extension	72°C	1 min/kb	
Final Extension	72°C	5 mins	1
Cooling	4°C	Hold	∞

Table 3: Thermocycling conditions for Polymerase Chain Reaction.

The PCR mix was prepared as recommended by Thermo Scientific™ Phusion™ High-Fidelity DNA Polymerase with following reading:

Component	Volume (in 50 µl reaction)
PCR Grade Water	~40 µl
GC/HF Buffer	5 µl
DMSO	1.5 µl
Template DNA	~200 ng
dNTP mix	0.5 µl
Forward Primer (Conc: 100 µM)	0.5 µl
Reverse Primer (Conc: 100 µM)	0.5 µl
Phusion® High Fidelity Polymerase	0.5 µl

Table 4: Components and quantity of Polymerase Chain Reaction mixture.

3.2.3 Restriction Cloning:

The amplified PCR products were subsequently digested with **FastDigest™ Restriction Enzymes** by Thermo Scientific® to produce sticky ends. The corresponding vector template vectors were also digested with same restriction endonucleases. Following restriction enzyme digestion, these DNA fragments were separated through agarose gel electrophoresis and then subsequently purified using **NucleoSpin® Gel and PCR Clean-up** kit by Macherey-Nagel GmbH™. Hereafter, the ligation reactions were set up to ligate the sticky ends of gene inserts and plasmid template using **T4 DNA Ligase** enzyme by New England Biolabs®. This ligated product then transformed in DH5α competent cells by Invitrogen®. The *E. coli* cells were the streaked onto LB agar plates containing 50µg/mL and 100µg/mL of kanamycin and ampicillin respectively. Individual colonies were picked up and were subsequently grown in antibiotic containing LB media. Thereafter, the plasmid was purified by miniprep protocol using **NucleoSpin® Plasmid** kits by Macherey-Nagel GmbH™. To acquire larger quantity of DNA where needed, we used **NucleoBond® Xtra Midi/Maxi** kits by Macherey-Nagel GmbH™.

3.2.4 Sanger DNA Sequencing:

The quantification of purified plasmid DNA was made through using NanoDrop™ 2000 by NanoDrop Technologies, Thermo Fisher™. Measurement of optical density (OD) was performed at 260 nm. All plasmids constructed were then sent formerly to **GATC Biotech®**, Germany and more latterly to **Eurofins®** for sequencing. Where possible, conventional corresponding primers were used for sequencing, otherwise specific primers were used for Sanger sequencing. Sequences were then analyzed using **4Peaks®** and **ApE- A plasmid Editor®** software.

3.2.5 Expression Vectors:

A range of vectors with variable antibiotic resistance and diverse restriction sites were used during the course of research. It includes both low copy and high copy plasmids. List of the plasmids made and utilized during the research is as follows:

Vector Backbone	Gene Insert	Epitope Tag	Bacterial Resistance	Cloning Approach
pRSV-Rev (Addgene®)	ΔN-Ku80	FLAG/HA	Ampicillin	ΔN-Ku80 was inserted in MCS* through <i>XhoI-NotI</i> restriction sites
pRSV-Rev (Addgene®)	wt-Ku80	FLAG/HA	Ampicillin	Wild type Ku80 was inserted in MCS* through <i>XhoI-NotI</i> restriction sites
pOZ-N (Addgene®)	wt H2A.Z	FLAG/HA	Ampicillin	Wild type H2A.Z was inserted in MCS* through <i>XhoI-NotI</i> restriction sites
pOZ-N (Addgene®)	wt Ku70	FLAG/HA	Ampicillin	Wild type Ku70 was inserted in MCS* through <i>XhoI-NotI</i> restriction sites
pOZ-N (Addgene®)	wt Ku80	FLAG/HA	Ampicillin	Wild type Ku80 was inserted in MCS* through <i>XhoI-NotI</i> restriction sites
Bicistronic pET28b (Addgene®)	Ku70 & Ku80	-	Kanamycin	Ku70 was inserted using <i>XhoI-NotI</i> sites, while Ku80 was inserted using <i>BamHI-HindIII</i> sites
Bicistronic pET28b (Addgene®)	H2A & H2B	FLAG (H2A) His (H2B)	Kanamycin	H2A was inserted using <i>XhoI-NotI</i> sites, while H2B was inserted using <i>BamHI-HindIII</i> sites

Bicistronic pET28b (Addgene®)	H2A.X & H2B	FLAG (H2A.X) His (H2B)	Kanamycin	H2A.X was inserted using <i>XhoI-NotI</i> sites, while H2B was inserted using <i>BamHI-HindIII</i> sites
Bicistronic pET28b (Addgene®)	H2A.Z & H2B	FLAG (H2A.Z) His (H2B)	Kanamycin	H2A.Z was inserted using <i>XhoI-NotI</i> sites, while H2B was inserted using <i>BamHI-HindIII</i> sites
pET15b (Addgene®)	Ku70	His	Ampicillin	wt-Ku 70** was inserted through <i>XhoI-KpNI</i> restriction sites
pET15b (Addgene®)	Ku80	His	Ampicillin	wt-Ku 80** was inserted through <i>XhoI-KpNI</i> restriction sites
pGEX-5X-1 (GE Healthcare®)	Ku70	GST	Ampicillin	wt-Ku 70** was inserted through <i>XhoI-NotI</i> restriction sites
pGEX-5X-1 (GE Healthcare®)	Ku80	GST	Ampicillin	wt-Ku 80** was inserted through <i>XhoI-NotI</i> restriction sites
pGEX-5X-1 (GE Healthcare®)	Ku70 (37-256)	GST	Ampicillin	N-terminus of Ku 70 was inserted through <i>XhoI-NotI</i> restriction sites
pGEX-5X-1 (GE Healthcare®)	Ku70 (263-466)	GST	Ampicillin	Core domain of Ku 70 was inserted through <i>XhoI-NotI</i> restriction sites
pGEX-5X-1 (GE Healthcare®)	Ku70 (472-557)	GST	Ampicillin	NLS domain of Ku 70 was inserted through <i>XhoI-NotI</i> restriction sites
pGEX-5X-1 (GE Healthcare®)	Ku70 (573-609)	GST	Ampicillin	SAP domain Ku 70 was inserted through <i>XhoI-NotI</i> restriction sites
pGEX-5X-1 (GE Healthcare®)	Ku80 (09-244)	GST	Ampicillin	N-terminus of Ku 80 was inserted through <i>XhoI-NotI</i> restriction sites
pGEX-5X-1 (GE Healthcare®)	Ku80 (253-453)	GST	Ampicillin	Core domain of Ku 80 was inserted through <i>XhoI-NotI</i> restriction sites
pGEX-5X-1 (GE Healthcare®)	Ku80 (476-569)	GST	Ampicillin	NLS domain of Ku 80 was inserted through <i>XhoI-NotI</i> restriction sites
pGEX-5X-1 (GE Healthcare®)	Ku80 (593-706)	GST	Ampicillin	DNA PK _{cs} domain of Ku 80 was inserted through <i>XhoI-NotI</i> restriction sites
pGEX-5X-1 (GE Healthcare®)	Ku80 (09-125)	GST	Ampicillin	Subdomain of vWA domain of Ku 80 was inserted through <i>XhoI-NotI</i> restriction sites
pGEX-5X-1 (GE Healthcare®)	Ku80 (126-244)	GST	Ampicillin	Subdomain of vWA domain of Ku 80 was inserted through <i>XhoI-NotI</i> restriction sites

pGEX-5X-1 (GE Healthcare®)	Ku80 (40-101)	GST	Ampicillin	Subdomain of vWA domain of Ku 80 was inserted through <i>XhoI-NotI</i> restriction sites
Bicistronic pET28b (Addgene®)	H2B & Ku (40-101)	His (H2B) FLAG (Ku)	Kanamycin	H2B was inserted using <i>XhoI-NotI</i> sites, while Ku subdomain was inserted using <i>BamHI-HindIII</i> sites
Bicistronic pET28b (Addgene®)	H2A/H2A.Z Chimera & H2B	FLAG (H2A/H2A.Z) His (H2B)	Kanamycin	Chimera(s) of H2A and H2A.Z was inserted using <i>XhoI-NotI</i> sites, while H2B subdomain was inserted using <i>BamHI-HindIII</i> sites

Table 5: Plasmid constructs used for gene expression: All the plasmids used during the research were prepared altogether in laboratory. All epitope tags were attached on N-termini of inserted gene unless otherwise mentioned. Different vectors encompassed different promoters based on their backbone.

*MCS: multiple cloning site

**wt: full length wild type protein

3.3 Expression Procedures:

Among different heterologous expression systems for recombinant proteins, bacterial expression system is considered as the most rapid, simple and most importantly the cheapest the way to yield high quantity of recombinant proteins that subsequently can be used for *in-vitro* studies. Since both histones and Ku proteins have been documented to well-expressed in *E-coli* expression system, we adopted to use this technique for all our *in-vitro* studies (Figure 23).

The strains of *E. coli* used during this research include **BL21 (DE3)** by Thermo Fisher Scientific™ (C6000-03), **BL21 (DE3) pLysS** by Invitrogen™ (C602003) and **BL21 Rosetta™ (DE3)** by Sigma Aldrich® (70954). All of these commercial strains exercise T7 promotor system for gene expression and use isopropyl β-D-1-thiogalactopyranoside (IPTG) for induction. After executing transformation of these competent cells with desired clones, the cells were plated on antibiotic specific agar plates for overnight. Next day a swipe of colonies was picked up and grown in 20 mL fresh Luria-Bertani (LB) media containing specific antibiotic. Few hours later, the 20 mL preculture was inoculated in 500 mL of fresh LB media and was grown at 37°C and 180 rpm of agitation until an OD₆₀₀ of 0.5-0.7 was achieved. Expression of every clone was then compared at different induction conditions including various induction

temperatures (18°C, 25°C and 37°C) and different IPTG concentrations (0.1mM, 0.4mM and 1mM).

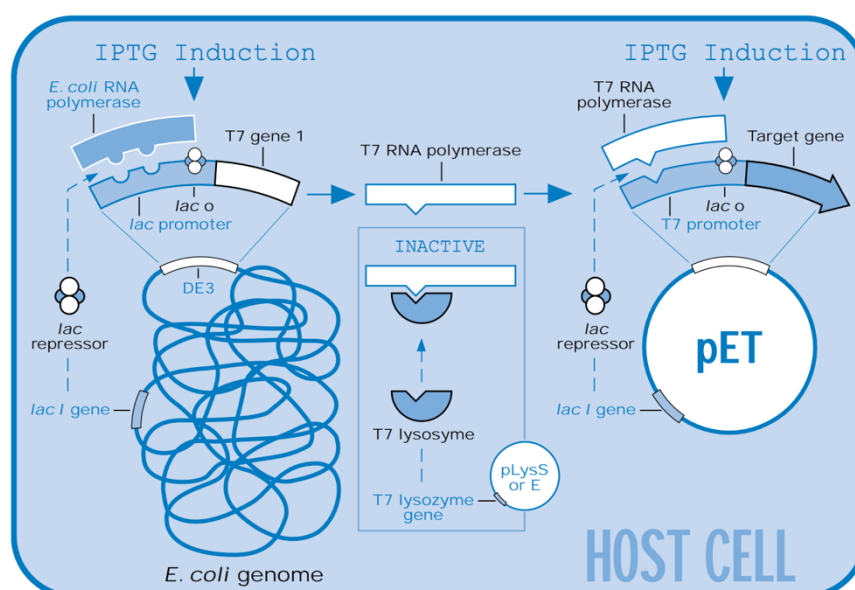


Figure 23: Schematic illustration of pET Expression System: The pET system is the most compelling technique yet established for the cloning and expression of recombinant proteins in *Escherichia coli*. Diagram acquired from Novagen's pET System Manual (11th Edition).

After induction with IPTG, the cells were harvested through high centrifugation and resuspended in appropriate buffers to generate their soluble extracts. Cells expressing histones were resuspended in 50mM Tris (pH:7.65), 10% glycerol, 1M NaCl, 0.4M ammonium acetate, 0.01% NP-40, 2mM dTT, 0.2mM PMSF and 10mM imidazole. All clones intended to be purified through GST tag were resuspended in 20mM Tris (pH:7.65), 10% glycerol, 1M NaCl, 3mM magnesium chloride, 0.1mM EDTA, 0.1% triton X-100, 0.01% NP-40 and 0.2mM PMSF (phenylmethylsulfonyl fluoride). Protease inhibitor cocktail tablets by **Roche**TM were also added in lysis buffer. Cells were lysed through freeze/thaw successions followed by sonication. Lysozyme (Sigma Aldrich[®]) was also added to assist the lysis of cells. The soluble fractions were then retrieved using high speed centrifugation at 4°C.

For large scale protein production, growth was made in 2X Luria-Bertani (LB) media supplemented with required antibiotics. Cultures were grown until optical density OD₆₀₀ of 1.0 may got achieved. The cultures were then cooled down to room temperature and induction was performed subsequently with IPTG at a concentration of

0.4mM at optimum temperature. The cells were then harvested using high centrifugation and stores at -20°C until subsequent use.

3.4 Protein Purification:

For precise biochemical studies of a particular protein, it must be purified from all other contaminants of the expression host present in the lysate. For all *in-vitro* interaction studies we used tandem affinity purification (TAP) technique of immunoprecipitation to evaluate protein-protein interaction. Commercially available resins mentioned in table 6 were used for all immunoprecipitation experiments. Appropriate protocols recommended by respective supplier were followed. All resins were first washed extensively with TGEN to get rid of supernatant. After binding with resin, the proteins were eluted using corresponding eluting methods. Appropriate binding and eluting durations were followed as per recommended protocols in literature archive. Both Flag and HA peptides were acquired from Peptide Synthesis Service, IGBMC, University of Strasbourg. Quantitative analysis of all eluted proteins was performed using Bradford measurement.

Epitope Tag	Binding Resin	Binding Time	Eluting Agent	Elution Time
FLAG	Anti-Flag [®] M2 Affinity Gel (Millipore [™] A2220)	6 hours	Flag Peptides (5mg/mL)	Overnight
HA	Pierce [™] Anti-HA Agarose (Thermo Fisher [®] 26181)	6 hours	HA Peptides (5mg/mL)	Overnight
Polyhistidine (HIS)	cOmplete [™] His-Tag Purification Resin (Roche [™] 5893682001)	2 hours	250 mM Imidazole	1 hour
GST	Glutathione Sepharose 4B Purification Resin (GE Healthcare [®] 17075605)	2 hours	L-Glutathione (20mM)	1 hour

Table 6: Commercial resins used for immunoprecipitation: All proteins are purified for subsequent *in-vitro* studies using epitope tags. Commercially available resins were used to immunoprecipitate the proteins. Resin bound proteins were then eluted using corresponding eluting agents. Binding and elution times were followed as per recommended in literature archive.

Following elution, the recombinant proteins were concentrated using **Amicon[®]** Ultra-15 filter units. After that, if required, proteins were dialyzed using **SnakeSkin[™] Dialysis Tubing** (Sigma Aldrich[®]) of appropriate sizes. To make two proteins interacted, equal stoichiometric amount of dialyzed proteins were mixed in the presence of appropriate amount of TGEN. They were incubated on standard agitation at 4°C for 2-4 hours and subsequently subjected to batch purification.

3.5 Protein Identification & Characterization:

Following protein purification, its identification and characterization is very important to understand the proteomic attributes. Different biophysical methods have been employed in biochemical research to explore these properties. Since different proteins differ from each other in their size, structure and physiochemical features, they can be identified and characterization through separation and identification.

3.5.1 SDS-PAGE Gel Electrophoresis:

Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis more commonly known as SDS-PAGE Gel Electrophoresis is the most extensively used technique to distribute different proteins from a single mixed sample based on variable molecular weights of different proteins. **Tetramethylethylenediamine** (TEMED) causes the splitting of ammonium persulfate (APS) to produce free radicals which subsequently causes the polymerization of acrylamide.

All the acrylamide gels were prepared in-house using Invitrogen[™] 1.0 mm Empty Gel Cassettes (*NC2010*). For different range of molecular weights, a specific percentage of SDS-PAGE gels were used. The composition of these SDS-PAGE gels corresponding to different percentage of acrylamide is described in Table 7. In addition to self-prepared gels, NuPAGE[™] 1.0 mm 4-12% Bis-Tris Protein Precast Gels (*NP0321PK2*) were also used where appropriate. The relative molecular weights of different proteins were estimated either through pre-stained **PageRuler[™] protein ladder** or through unstained **Mark12[™] Unstained Standard**.

Ingredient	Running					Stacking
	8%	10%	12%	15%	18%	3%
Acrylamide 40%	2 ml	2.5 ml	3 ml	3.75 ml	4.5 ml	0.375 μ l
Tris 1.5M pH:8.8	5 ml	5 ml	5 ml	5 ml	5 ml	1.25 ml
SDS 20%	50 μ l	50 μ l	50 μ l	50 μ l	50 μ l	25 μ l
APS 10%	100 μ l	100 μ l	100 μ l	100 μ l	100 μ l	50 μ l
TEMED	10 μ l	10 μ l	10 μ l	10 μ l	10 μ l	5 μ l
H ₂ O	2.84 ml	2.34 ml	1.84 ml	1.09 ml	340 μ l	3.3 ml

Table 7: Composition of resolving and stacking gel with variable fraction of acrylamide: Different percentage of SDS-PAGE gels were used as per requirement. The table indicates the component volume of each constituent to prepare a total of 10 mL solution.

3.5.2 Protein Gel Staining:

For the staining of SDS-PAGE gels both Coomassie and silver staining techniques were used as per required. For Coomassie staining, gels were initially fixed for about 20 minutes in fixation solution (10% acetic acid; 30% ethanol). Subsequently, gels were incubated in **Coomassie blue G250** staining solution (0.006% Coomassie Blue G250; 30mM HCl). Gels were destained in boiling water until the sharp protein bands were made visible. For silver staining PierceTM Silver Stain Kit (Thermo Fisher[®] 24612) was used.

3.5.3 Western Blot:

For supplementary confirmation of expression of protein, about 50 μ g of protein lysate was separated using SDS-PAGE and run in tris-glycine running buffer (25 mM Tris, 192 mM Glycine, 0.1% SDS, pH 8.3) using Bio-Rad vertical gel electrophoresis system. Hereafter, separated proteins were transferred to in-advance activated polyvinylidene difluoride (PVDF) or nitrocellulose membranes using Bio-Rad blotting system. Transfer buffer used was comprised of 25 mM Tris, 192 mM glycine and 20% methanol. Following transfer on membrane, membranes were blocked in blocking solution containing 1X PBS, 10% dried milk and 0.1% Tween-20TM (Sigma Aldrich[®]

T9039) for 1 hour in order to reduce the non-specific background. Following blocking the membranes were incubated in appropriate primary antibodies at 4°C/overnight and subsequently washed with PBS-Tween solution 3-4 times to remove unspecific binding. After that, the membranes were incubated in appropriate concentration of secondary antibodies and then were again washed in PBS-Tween solution for 3-4 times. Following last washing, the antigen-antibody complexes were identified using chemiluminescent substrates (Millipore™ WBKLS0050). The list of primary and secondary antibodies used in the research is as follows:

Name	Type	Species	Source	Reference Number
Primary Antibodies				
Anti-HA	Monoclonal (clone 3F10)	Rat	Roche® Diagnostics GmbH	11-867-423-001
Anti-FLAG	Monoclonal	Mouse	Sigma Aldrich®	A8592
Anti-His	Monoclonal	Mouse	Thermo Fisher Scientific®	MA1-21315
Anti-H2A	Polyclonal	Mouse	Kind Gift by Dr. Stefan Dimitrov	
Anti-H2A.Z	Polyclonal	Rabbit	Active Motif®	AB-2615081
Anti-H2B	Polyclonal	Rabbit	Sigma Aldrich®	07-371
Anti-Ku70	Monoclonal	Mouse	Santa Cruz Biotechnology®	sc-17789
Anti-Ku80	Monoclonal	Mouse	Santa Cruz Biotechnology®	sc-515736
Secondary Antibodies				
Anti-Mouse	Polyclonal	Mouse	Thermo Fisher Scientific®	Alexa Fluor® 568 (A-11031)
Anti-Rabbit	Polyclonal	Mouse	Thermo Fisher Scientific®	Alexa Fluor® 488 (A-11008)
Anti-Rat	Polyclonal	Mouse	Thermo Fisher Scientific®	Alexa Fluor® 488 (A-11006)

Table 8: List of primary and secondary antibodies used during research: Commercially available epitope tag specific and protein specific antibodies were used to conduct Western Blot experiments. All antibodies were assayed with positive and negative controls before using in real experiments.

3.5.4 Mass Spectrometry:

Mass spectrometric analysis of proteins was accomplished using an ion-trap mass spectrometer (ThermoFinnigan LTQ-Orbitrap Velos) or by the Taplin Biological Mass Spectrometry Facility (Harvard Medical School, Boston, Massachusetts, USA) as described previously by Orbi *et al.*, 2014.

3.6 Cell Lines:

Over the last few decades, several revolutionizing techniques including bacterial expression system, baculovirus implicating insect SF9 expression system, yeast cell lines and many eukaryotic expression techniques have developed for heterologous expression of recombinant proteins. In this study, along with *E-coli* expression system we used HeLa and Mouse Embryonic Fibroblast cell lines for to prepare protein complexes and study protein interactions.

3.6.1 HeLa:

For protein complex study and mass spectrometric analysis of Ku70, Ku80 and H2A.Z HeLa cells were procured from Cell Facility Centre, IGBMC, University of Strasbourg. The cells were cultured in Dulbecco's modified Eagle's medium (DMEM) complemented with 10% of heat inactivated fetal bovine serum and 100U/mL of penicillin/streptomycin antibiotic. Cells were grown at 37°C with 5% CO₂ in a humidified environment. Subsequently they were undergone transduction with desired plasmids (pOZ-N vector by Addgene® containing PCR amplified inserts of desired sequence) by infecting them with retroviruses produced through Phoenix helper-free retrovirus producer lines (Phoenix-E). Stably transduced HeLa cells lines were established through 3-4 times immunomagnetic selection using Dynabeads™ CD25 beads (Invitrogen™ 11157D). The proportion of positive cells was estimated using immunofluorescence (IF) assay. Subsequently the stable cell lines were grown more and

insoluble nuclear extracts (INE) were acquired as per recommended protocols in literature archive.

3.6.2 Mouse Embryonic Fibroblasts (MEFs):

For acquisition of chromatin corresponding to wt-Ku80 and Δ N-Ku80 for ChIP-Seq analyses and also for mass spectrometric analysis of same samples, mouse embryonic fibroblasts (MEFs) were used. Ku80 knockout MEF cells were a kind gift from *Professor Penny Jeggo*, Professor Emeritus, Genome Damage and Stability Centre, University of Sussex, UK. The cells were cultured in Dulbecco's modified Eagle's medium (DMEM) complemented with 10% of heat inactivated fetal bovine serum and 100U/mL of penicillin/streptomycin antibiotic. Cells were grown at 37°C with 5% CO₂ in a humidified environment. The coding sequence of full-length wild type Ku80 and Δ N Ku80 (107-732) were amplified using classic PCR and subsequently cloned in pRSV-Rev (Addgene®) retroviral plasmid containing N-terminal FLAG and HA tags.

Phoenix helper-free retrovirus producer lines (Phoenix-E) were transfected with calcium phosphate mediated transfection of pRSV-Rev plasmids to produce retroviruses. These 293T cell line based Phoenix lines are capable of generating viruses very quickly as compared to other cell lines. The viruses produced through Phoenix system then used to transduce MEF cells. Transduced stable MEF cell populations were established through 3-4 times immunomagnetic selection using Dynabeads™ CD4 beads (Invitrogen™ 11145D). The efficiency of transduction was evaluated through immunofluorescence assay as explained in coming section. Positively transfected MEF cells were then further grown to harbor suitable quantity of cells. Insoluble nuclear extracts (INE) were acquired using recommended protocols and subsequently used for ChIP-Seq and mass spectrometry samples preparation.

3.6.3 Immunofluorescence (IF) Assay:

Before proceeding to the usage of mammalian cells for ChIP-Seq and mass spectrometry all cell lines were evaluated for positive transduction through Immunofluorescence assay. For that purpose, cells were plated on cover slips and fixed with 4% paraformaldehyde. Following the termination of paraformaldehyde fixation through 100mM glycine treatment, cells were washed with PBS for 2-3 times. Subsequently cells were permeabilized with 0.5% v/v Triton X-100 incubation. After that, cells were incubated in blocking solution (1% BSA and 1% fetal bovine serum) for 20 minutes. Samples were then incubated with recommended concentration of primary antibody at room temperature for 1 hour. After that, 2-3 times washing with PBS to remove unbound antibody, cells were incubated in corresponding Alexa Flour conjugated secondary antibody. Cells were then washed again 2 times with PBS and nuclei were thereafter stained with 0.1% 4',6-diamidino-2-phenylindole DAPI followed by subsequent washing with PBS again. The coverslips were then visualized under fluorescent microscope to estimate the percentage of positive cells.

3.6.4 Cryopreservation of Stable Mammalian Cell Lines:

All mammalian cell lines were frozen in heat inactivated fetal bovine serum (FBS) supplemented with 10% dimethyl sulfoxide (DMSO). Cells were initially stored at -80°C for 72 hours and subsequently shifted to liquid nitrogen (-196°C) with proper labelling and recording in catalogs.

3.7 ChIP sequencing (ChIP-Seq) Assay:

ChIP sequencing is a combination of chromatin immunoprecipitation (ChIP) and genome wide sequencing that enable us to understand the genome wide binding sites of our protein of interest. In this research, all the ChIP-Seq samples were procured through MEFs cells. MEFs of each of four categories viz. Ku80^{-/-}, Ku80^{+/+}, rescued wtKu80 and rescued ΔNku80 were crosslinked using 0.4% paraformaldehyde for 10 mins at room

temperature followed by the treatment of 150 mM glycine at room temperature to halt paraformaldehyde action. Afterwards, cells were washed two times with PBS and then scrapped, collected and finally centrifuged at 2000rpm for 10 minutes at 4°C to get pellet. The pellet was then resuspended in cell lysis buffer containing 50mM Tris HCl (pH: 8.0), 10mM EDTA (pH: 8.0) and 1% SDS. Following that, lysate was sonicated at 4°C for 25-30 sets of 30 sec pulses and subsequently centrifuged at 18,000g for 20 mins. The sonicated chromatin samples were preserved at -80°C till further assays.

For immunoprecipitation, firstly sonicated chromatin was diluted in buffer comprising 167 mM NaCl, 16.7mM Tris HCl (pH: 8.0), 12mM EDTA (pH: 8.0), 1.1% Triton X-100 and 0.01% SDS and then 5uL of ChIP grade H2A.Z antibody by Active Motif® was added in each sample and incubated at 4°C for overnight. Simultaneously, equal amount of Protein A Sepharose and Protein G Sepharose were mixed and washed two times with TGEN. After washing, both A and G beads were incubated at 4°C for overnight in buffer containing 5μL SDS20%, 20μL tRNA and 25μL BSA (20γ). After overnight blocking of beads, 50μL of slurry was added in each chromatin sample and given 2-3hours incubation at 4°C. The bound material was then recovered and washed two times with each of following buffers in respective manner to remove non-specific binding if any.

Buffer	Composition
Low salt immune complex wash buffer	Tris-HCl 20mM pH (8.1) NaCl 150mM SDS 0.1% 1% Triton X-100 2mM EDTA
High Salt Immune Complex Wash Buffer	Tris-HCl 20mM pH (8.1) NaCl 500mM SDS 0.1% 1% Triton X-100 2mM EDTA
LiCl Immune Complex Wash Buffer	Tris-HCl 10 mM pH (8.1) 0.25 M LiCl 1% NP-40 1% Deoxycholate EDTA 1 mM
TE Buffer	10 mM Tris-HCl pH (8.1) 1 mM EDTA, pH (8.0)

Table 9: Composition of washing buffers pertaining for blocking agarose.

Following washing with above buffers, the ChIPed material was eluted in buffer containing 100mM NaHCO₃ and 1% SDS. The eluted chromatin was then reverse-crosslinked by adding 300mM final concentration of NaCl and 1μL RNase and incubated at 67°C for overnight. Following that, the DNA was presented to proteinase K digestion and then extracted through phenol-chloroform protocol. Finally, after washing the extracted product with 70% ethanol and water the quantity of DNA was estimated through NanoDrop™ and subsequently sent to the ISO 9001 certified GenomEast® Platform, IGBMC, University of Strasbourg to undergo sequencing. The sequencing by GenomEast® Platform was performed using Illumina HiSeq 4000. The computational analysis of raw data received by GenomEast® Platform was performed in laboratory using appropriate bioinformatic programs.

3.8 Ion Exchange/Gel Filtration Chromatography:

Ion exchange (IEX) chromatography involves the separation of molecules based on their total surface charge. Proteins are amphoteric molecules whose net surface charge changes with the change of environmental pH. An oppositely charged matrix is used to hold the sample while the non-specific, uncharged or same (charge with the matrix) charged species are washed out of the column without binding. The retained proteins are then eluted through changing the ionic strength or the pH.

In the current research, Flag/His tagged H2A.Z/H2B were first affinity purified and then subjected to ion exchange chromatography using 26/10 quaternary ammonium Q Sepharose (strong anion exchanger) High Performance™ column for Ku80 (aa 40-101) and 26/10 sulphopropyl SP Sepharose (strong cation exchanger) High Performance™ column for H2A.Z/H2B. Both columns were equilibrated with a buffer containing 50 mM Tris (pH 8.0), 50 mM NaCl and 2 mM DTT. Following washing with same equilibration buffer the proteins were eluted with a gradient of increasing salt concentration ranging from 50 to 1000 mM NaCl. Peak fractions were then concentrated with an Amicon™ ultra centrifugal filter units.

Gel filtration chromatography also known as molecular sieve chromatography or size-exclusion chromatography (SEC) is commonly used technique to purify protein complexes based on their molecular weights. Unlike ion exchange chromatography or affinity chromatography, gel filtration chromatography doesn't require any interaction of the sample with the charged matrix. In the current research, after the interaction of equimolar amount of the concentrated proteins, the complexes were loaded onto HiLoad™ 16/600 Superdex 200 prep grade gel filtration column (GE Healthcare). The gel filtration column was pre-equilibrated with gel filtration buffer containing 50mM Tris-Cl pH 8.0, 50mM NaCl and 2mM DTT.

3.9 Structural Modelling:

The structural modelling was performed using ClusPro® docking server which has been rated as the best online server for protein-protein docking in the last CAPRI (Critical Assessment of PRediction of Interactions) evaluation meeting in 2016 (Kozakov et al., 2017).

4. RESULTS & DISCUSSION

4.1 Identification of Ku70/Ku80 in H2A.Z predeposition complex:

The evolutionary conserved histone variant H2A.Z has been proposed to play crucial roles in many important cellular processes. Similarly, Ku heterodimer Ku70/Ku80 has also been implicated in vital functions such as DNA repair. The idea to explore the interaction between Ku heterodimer and H2A.Z was perceived when we identified Ku70 (XRCC6) and Ku80 (XRCC5) along with other DNA repair members (Figure 24) in the H2A.Z soluble complex during one of our previous work on histone chaperone ANP32E (Obri *et al.*, 2014).

e-H2A.Z.com			
Entry	Proteins	Peptides	
Q9Y4A5-1	TRRAP	229	
Q96L91-2	EP400	105	
Q9H0E9-3	BRD8	32	
Q9NV56	MRGBP	3	
Q9H2F5-1	EPC1	23	p400/TIP60.com
Q52LR7	EPC2	22	
Q92993-2	TIP60	18	
Q9NXR8-1	ING3	14	
Q15014	MRGX	3	
Q9UBU8	MRG15	3	
O95619	GAS41	5	
Q9NPF5	DMAP1	30	
Q9Y265-1	RUVBL1	23	
Q9Y230	RUVBL2	30	
Q15906	YL1	9	SRCAP.com
O96019-1	ACTL6A	16	
O43257	ZNHIT1	4	
Q6ZRS2-2	SRCAP	63	
Q9GZN1	ACTR6	6	
Q05BQ5-3	MBTD1	9	
P11142-1	HSPA8	27	
P78527-1	PRKDC	84	
Q01831	XPC	8	
P49916-2	LIG3	8	
P52701-2	MSH6	2	DNA repair
P43246	MSH2	7	
P18887	XRCC1	7	
P13010	XRCC5	18	
P12956	XRCC6	22	
P54727	RAD23B	6	
Q9BTT0	ANP32E	6	histone chaperone
P55209	NAP1L1	6	
Q99733	NAP1L4	1	
P68431	H3	4	
P62807	H2B	26	
P62805	H4	3	
P0C0S5	H2A.Z	30	

Figure 24: Mass spectrometry analysis of e-H2A.Z: List of polypeptides identified in e-H2A.Z complex purified from HeLa cells. The presence of XRCC6 (Ku70) and XRCC5 (Ku80) in the H2A.Z complex suggests a specific interaction of Ku heterodimer with the histone variant H2A.Z. Adapted from (Obri *et al.*, 2014) (modified)

Through different mechanisms, both Ku heterodimer and H2A.Z have been independently designated to undergo important functionality in DNA repair mechanisms in addition to many other significant mechanistic roles. Therefore, we decided to further investigate this specific interaction between the Ku heterodimer and H2A.Z by exploring the molecular basis of this interaction using different *in-vivo* and *in-vitro* biochemical and genomic approaches.

4.2 Exploration of Ku70/Ku80 and H2A.Z interaction in HeLa cells:

To investigate the interaction of Ku heterodimer with the histone variant H2A.Z, we established a HeLa cell line stably expressing N-terminally FLAG/HA epitope-tagged human H2A.Z by retroviral transduction, according to an established protocol (Nakatani and Ogryzko, 2003). The proportion of positive cells was then estimated by means of immunoblotting using anti-HA and anti-Flag antibodies and subsequently through immunofluorescence (IF) assay using anti-HA antibody (Figure 25).

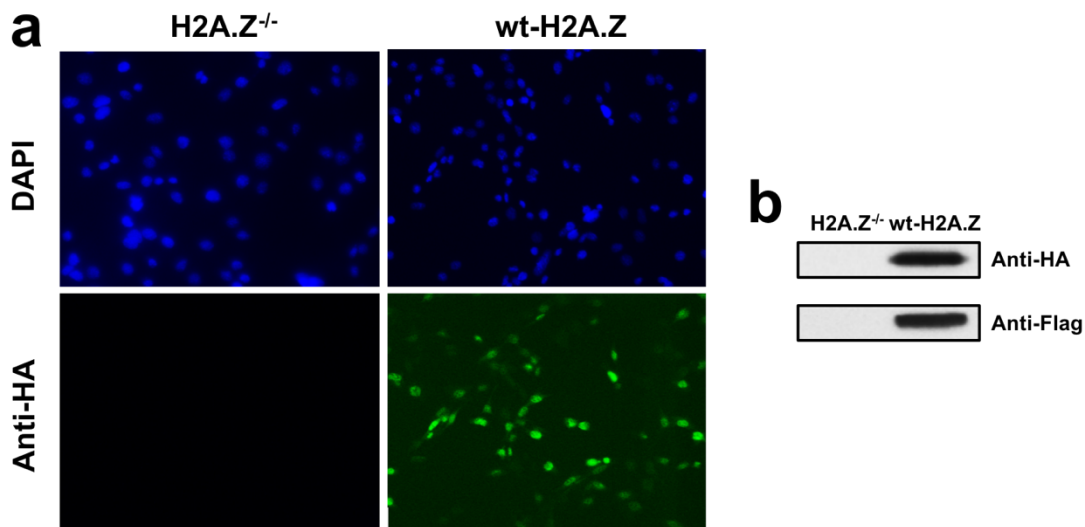


Figure 25: Validation of stable expression of FLAG/HA tagged wt-H2A.Z in H2A.Z knockout HeLa cells through immunofluorescence and immunoblotting: (a) Both control (H2A.Z^{-/-}) and wt-H2A.Z expressing HeLa cells were undergone immunofluorescence assay using DAPI (top) and anti-HA (bottom). (b) Extracts from H2A.Z^{-/-} and wt-H2A.Z expressing HeLa cell were immunoblotted with anti-HA and anti-Flag antibodies to validate the expression of H2A.Z.

Following establishment of stable HeLa cell lines, the cells were grown in large quantity (10 L). The e-H2A.Z complex was purified from the insoluble nuclear extracts (INE) using tandem immuno-affinity purification through anti-FLAG followed by anti-HA. The predeposition complex was then fractionated on a glycerol gradient of 15-35%, resolved on SDS-PAGE and subjected to silver staining (upper panel, Figure 26). The bands detected at ~70kDa and ~83kDa of medium molecular weight complex (MMWC) were presumed to be Ku70 and Ku80 respectively. Subsequently, the presence of Ku80 in the purified e-H2A.Z complex was validated using anti-Ku80 western blotting (Obri *et al.*, 2014) (bottom panel, Figure 26).

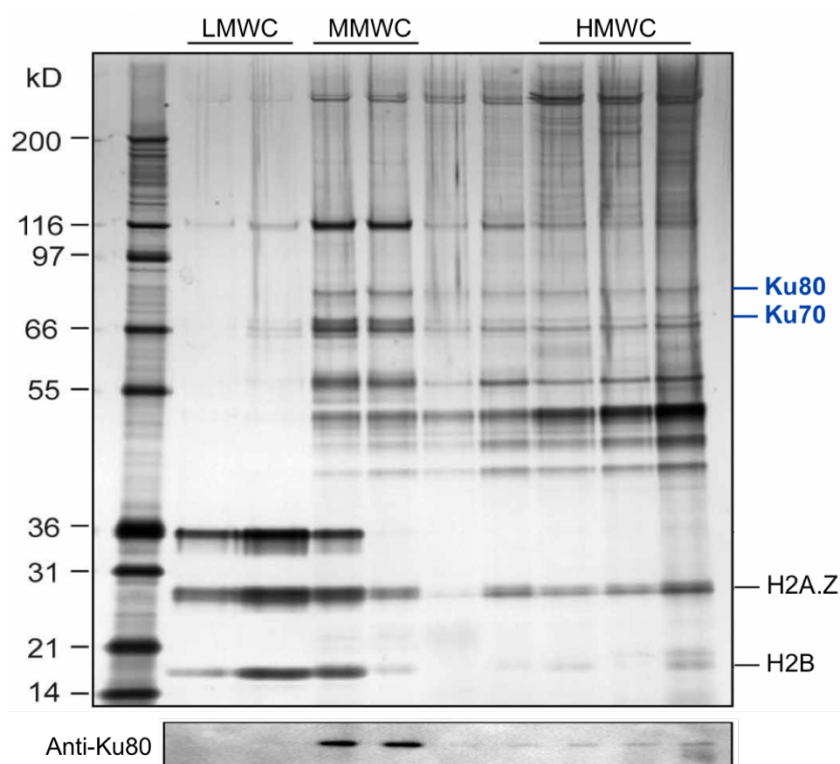


Figure 26: Ku70 & Ku80 are associated in H2A.Z complex: Silver staining of proteins associated with double immunoaffinity (HA-Flag) purified e-H2A.Z complex fractionated on 15%-35% glycerol gradient (**top**). Western blotting of glycerol fractions with anti-Ku80 antibody indicating Ku as an H2A.Z specific binding protein (**bottom**). Adapted from (Obri *et al.*, 2014) (modified)

Following the immunoblotting confirmation of Ku heterodimer in e-H2A.Z complex, the glycerol gradient samples were sent for mass spectrometry analysis. Mass spectrometric results confirmed the presence of Ku70 and Ku80 in the medium

molecular weight complex (MMWC) of H2A.Z. Histone fellows H2B, H3 and H4 were also identified in complex as expected (Obri *et al.*, 2014) (Figure 27).

gradient e-H2A.Z.com Entry	Proteins	HMWC	MMWC	LMWC
Q9Y4A5-1	TRRAP	163	56	7
Q96L91-2	p400	95	29	12
Q9H0E9-3	BRD8	21	6	2
Q92993-2	TIP60	7	0	0
Q9NPF5	DMAP1	24	16	0
Q9H2F5-1	EPC1	9	2	0
Q52LR7	EPC2	10	1	0
Q15014	MRGX	4	0	0
Q9NV56	MRGBP	3	0	0
Q9UBU8	MRG15	1	0	0
Q92993-2	TIP60	7	0	0
Q9NXR8-1	ING3	5	0	0
Q95619	GAS41	12	0	0
Q9GZN1	ACTR6	9	0	0
Q15906	YL1	8	5	0
Q9Y265-1	RUVBL1	28	8	5
Q9Y230	RUVBL2	41	9	4
Q6ZRS2-2	SRCAP	68	15	0
P12956	XRCC5	0	35	0
P13010	XRCC6	1	32	0
O43257	ZNHIT1	2	0	0
P78527-1	DNA-PK	123	47	0
P78527-1	PARP1	0	28	0
Q08945	SSRP1	4	2	0
Q9Y5B9	SUPT16H	12	7	0
P55209	NAPL1	0	7	1
Q9BTT0	ANP32E	5	2	25
Q96P70	IPO9	16	29	3
P62807	H2B	5	8	12
P68431	H3	0	2	0
P62805	H4	0	3	0
P0C0S5	H2AZ	4	6	15

Figure 27: Mass spectrometry analysis of H2A.Z subcomplexes: Mass spectrometric identification of various polypeptides associated with e-H2A.Z, purified through double immunoaffinity (Flag/HA). The associated protein complex was split into high molecular weight complex (HMWC), medium molecular weight complex (MMWC) and low molecular weight complex (LMWC). Number of specific peptides corresponding to each associated protein identified were presented respectively. Note the presence of Ku70 (XRCC6) and Ku80 (XRCC5) in the interactome of H2A.Z. Adapted from (Obri *et al.*, 2014) (modified)

4.3 Ku heterodimer Ku70/Ku80 specifically interacts with histone dimer H2A.Z/H2B *in-vitro*:

Following the *in-vivo* identification of Ku heterodimer in the H2A.Z complex, we moved to *in-vitro* characterization of this interaction. We employed *Escherichia coli* bacterial expression system for all our *in-vitro* studies of protein associations. Full length Ku70 and Ku80 were amplified and subsequently inserted in the bicistronic pET28b vector using *XhoI-NotI* and *BamHI-HindIII* restriction enzyme sites respectively. Similarly, another bicistronic vector containing His-tagged H2B at the *XhoI-NotI* sites and Flag-tagged H2A or H2A.X or H2A.Z at the *BamHI-HindIII* restriction sites was used. Following that, the bicistronic vector containing Ku70 and Ku80 was co-expressed with each of the bicistronic vectors containing H2A/H2B (lane 9-12, Figure 28), H2AX/H2B (lane 17-20, Figure 28) or H2A.Z/H2B (lane 25-28, Figure 28) using Invitrogen™ BL21 (DE3) pLysS *E.coli* strain. Protein expression

was induced with 1mM isopropyl β -D-1- thiogalactopyranoside (IPTG) for 3 hours at 37°C. Eventually, the soluble fractions of cell lysates were undergone immunoprecipitation using anti-FlagTM M2 Affinity resin and subsequently subjected to flag peptides (DYKDDDDK) elution. Cells expressing just Ku70/Ku80 were also subjected to anti-Flag elution to exclude any possibility of non-specific interaction of Ku heterodimer with anti-Flag agarose affinity resin (lane 1-4, Figure 28). Separate expression and Flag elution of all three histone dimers H2A/H2B (lane 5-8, Figure 28), H2AX/H2B (lane 13-16, Figure 28) and H2A.Z/H2B (lane 21-24, Figure 28) was performed as positive control.

Resin bound proteins were washed with TGEN buffer containing different salt concentrations (0.15M, 0.5M or 1M NaCl). The eluted proteins were fractionated on SDS-PAGE. Soluble crude extract fractions (input) of all experiments were also subjected to SDS-PAGE fractionation to ensure the stoichiometric expression of all members. Upon colloidal blue staining, Ku heterodimer exhibited specific interaction with H2A.Z/H2B histone dimers stable at high salt concentration. The presence of Ku70 and Ku80 in input and eluted fractions of all control and treated samples were also validated by western blotting using anti-Ku70 and anti-Ku80 antibodies. These *in vitro* data further validated our *in vivo* findings on the specific interaction between Ku heterodimer and H2A.Z/H2B dimers.

The exclusive interaction of Ku heterodimer with H2A.Z/H2B suggest a possible implication of Ku proteins in H2A.Z functioning. We therefore hypothesized that Ku proteins may have some special role in H2A.Z deposition at the sites of DNA damage as both Ku and H2A.Z have already been reported to play a role in DNA repair. Ku proteins are known to be involved in the initiation of NHEJ repair whereas H2A.Z has been shown to be crucial for generating an open chromatin structure which is a prerequisite for the recruitment of the repair machinery.

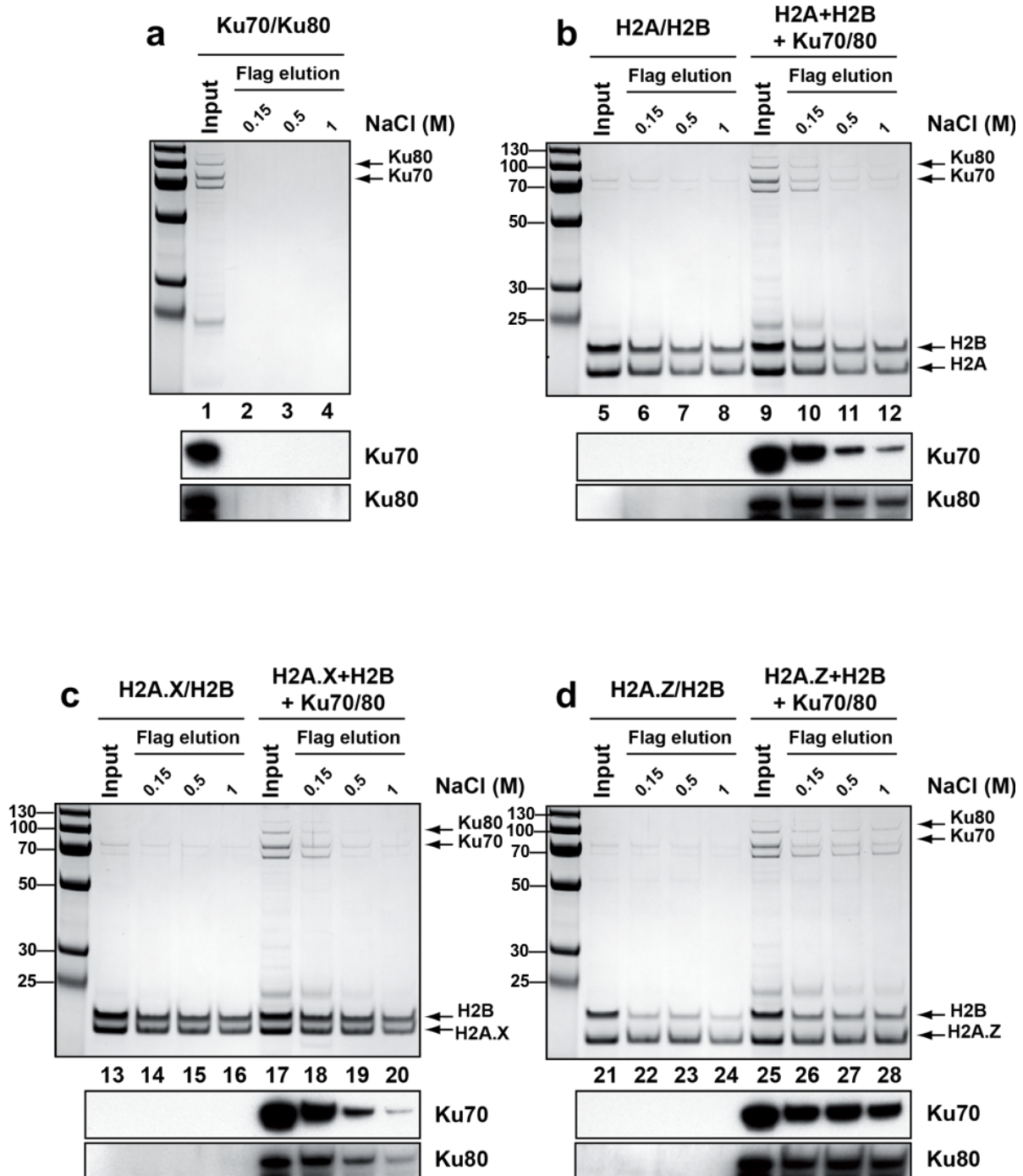


Figure 28: Ku70/Ku80 heterodimer specifically interacts with the histone H2A.Z/H2B dimer in vitro: Using *E.-coli* expression system, Ku70/Ku80 were co-expressed individually with the histone dimers: H2A/H2B (b), H2A.X/H2B (c) and H2A.Z/H2B (d), where H2B was Flag tagged while H2A, H2A.X and H2A.Z were His tagged. The co-expressed proteins were then immuno-purified using the Flag epitope. Resin bound protein(s) were washed with TGEN at different salt concentrations (0.15M, 0.5M or 1M NaCl). The eluted proteins were then fractionated on SDS-PAGE and stained with colloidal blue. The Ku heterodimer showed a specific interaction with the H2A.Z/H2B dimer stable at high salt concentration. Negative control (a) was also assessed to exclude any non-specific interaction of Ku heterodimer with anti-Flag agarose affinity gel. The presence of Ku70 and Ku80 proteins in input and eluted fractions was also assessed by Western blotting using Ku70 and Ku80 antibodies.

4.4 von Willebrand factor type A (vWA) domain of Ku80 is critical for H2A.Z/H2B interaction:

Having identified a direct interaction between the Ku proteins and the histone H2A.Z/H2B dimers, we next decided to further characterize in details this interaction by mapping the minimal Ku70 or Ku80 domain involved in this interaction. Based on their primary structures, we decided to divide both Ku70 and Ku80 proteins into four functional domains (Figure 29a) and to express and assay each domain separately for its interaction with H2A.Z/H2B dimers.

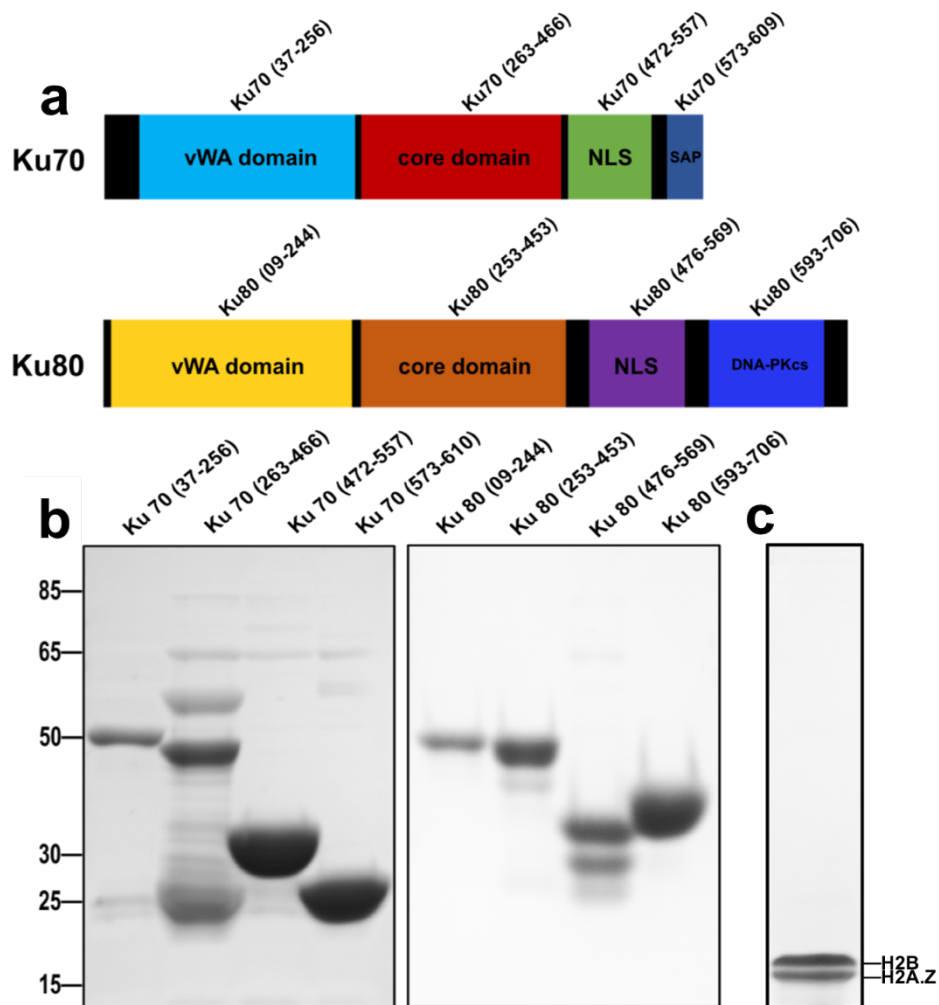


Figure 29: Domains characterization of Ku70/Ku80: (a) Primary structure of Ku70/Ku80: Both Ku70 and Ku80 consist of an N-terminal von Willebrand factor type A (vWA) domain, a central DNA binding core domain and a nuclear localization sequence (NLS) domain. C-terminal domain of Ku70 comprises an SAP (SAF-A/B, Acinus and PIAS) domain while C-terminal domain Ku80 comprises of sequence responsible for binding with DNA-PKcs. (b) Immunoprecipitation of N-terminally GST-tagged domains of Ku70/Ku80 expressed using *E. coli* expression system. (c) Histone dimer H2A.Z/H2B was expressed in *E. coli* expression system and subsequently purified from soluble fraction using double immunoaffinity (Flag/His) assays.

Structural and functional studies have suggested that both the Ku70 and Ku80 consist of an N-terminal von Willebrand factor type A (vWA) domain which is comprised of 37-256 residues in Ku70 and 09-244 residues in Ku80. The central DNA binding core domain is composed of residues 263-466 in Ku70 and residues 253-453 in Ku80. Following the central domain comes the nuclear localization sequence (NLS) domain that comprised of 472-557 amino acids sequence in Ku70 and 476-569 in case of Ku80. The C-terminal domain of Ku70 (residues 573-609) comprises an SAP (SAF-A/B, Acinus and PIAS) domain while C-terminal domain Ku80 (residues 593-706) comprises of sequence responsible for binding with DNA-PK_{CS} (Figure 29a). Using *XhoI-NotI* restriction enzyme sites all eight domains were amplified and subsequently inserted in pGEX-5X-1 vector that already contained N-terminally linked GST epitope beforehand.

All pGEX vectors containing Ku domains were individually transformed into BL21 (DE3) pLysS (Invitrogen™) strain of *E.coli*. Cells were induced with 1mM isopropyl β-D-1- thiogalactopyranoside (IPTG) for 3 hours at 37°C. Eventually, the soluble fractions of cell lysate were subjected to immunoprecipitation using Glutathione Sepharose 4B purification resin and subsequently eluted using 20mM L- Glutathione (Figure 29b).

On the other hand, the histone H2A.Z (Flag)/H2B (His) dimer was expressed separately using the pET28b bicistronic vector and purified using Ni-NTA affinity purification for His-tagged H2A.Z. Both the histone dimer and Ku70/80 domains were then dialyzed against 0.15M TGEN using appropriately sized SnakeSkin™ Dialysis Tubing. Afterwards, the histone H2A.Z/H2B dimer was mixed individually with the different Ku domains in TGEN buffer containing 0.5M NaCl (0.5 TEGN). Thereafter, the bound proteins were subjected to GST pulldown. The GST bound proteins were washed with 1M TGEN to wipe out nonspecific binding and subsequently eluted proteins were fractionated on SDS-PAGE and stained with colloidal blue. Upon analysis, the von Willebrand factor type A (vWA) domain of Ku80 (amino acids 09-244) was found to interact specifically with the H2A.Z/H2B dimer. Immunoblotting of the eluted fractions with anti-Flag (H2A.Z) and anti-His (H2B) antibodies further confirmed this conclusion (Figure 30).

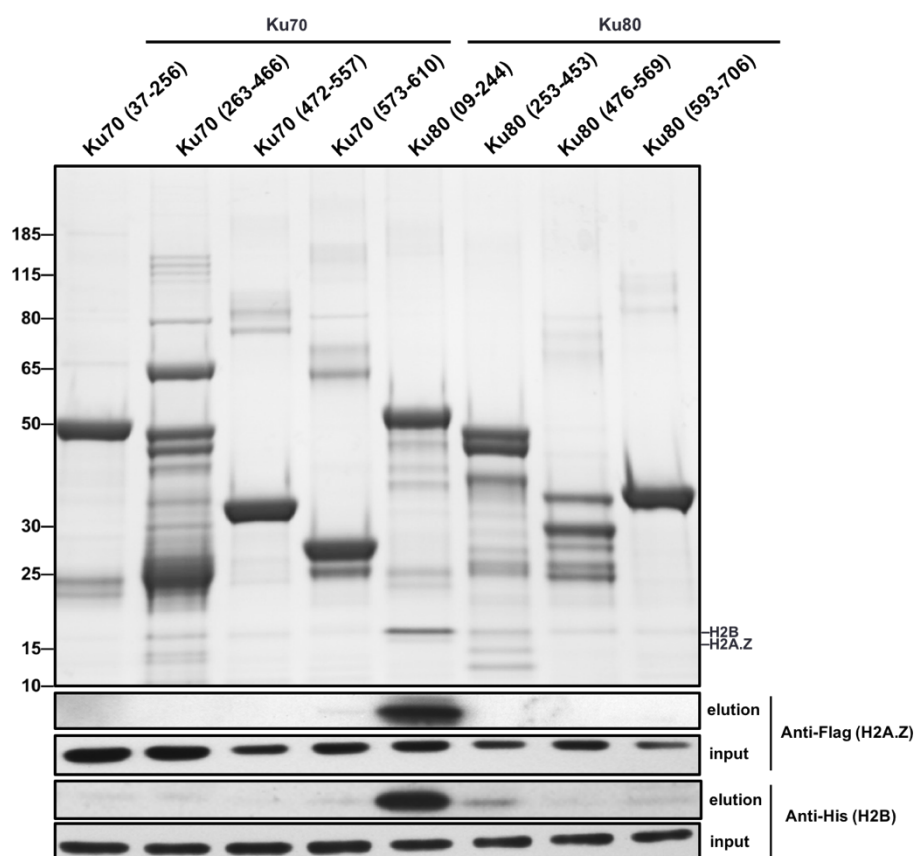


Figure 30: The histone H2A.Z/H2B dimer exclusively interacts with vWA domain of Ku80: Epitope tagged H2A.Z(Flag)/H2B(His) dimer were assayed for interaction with the different Ku70 and Ku80 domains in TGEN (0.5M NaCl) buffer. GST pulled-down proteins were fractionated on SDS-PAGE and stained with colloidal blue. The von Willebrand factor type A (vWA) domain of Ku80 was found to specifically interacts with the H2A.Z/H2B dimers (**top**). Western blotting analysis of inputs indicated that equal amount of histone dimers was used for interaction with individual Ku domains. Immunoblotting of eluted fractions validated the specific interaction of vWA domain of Ku80 with H2A.Z/H2B (**bottom**).

To further validate this interaction in a more physiological context, the histone H2A.Z/H2B dimers were co-expressed in bacteria with the different Ku70/80 domains and assayed for interaction using the GST-pull-down assay. Toward this goal, BL21 (DE3) pLysS *E. Coli* cells were co-transformed with the H2A.Z/H2B vector and each of the different Ku70/Ku80 domains vectors described above. Cells were induced with 1mM IPTG for 3 hours at 37°C. Soluble fractions of the different cell lysates were used to pull-down the different Ku domains using the GST-pull-down assay. The resin bound proteins were washed with 1M TGEN and eluted with 20mM L-Glutathione. The eluted proteins were then fractionated on SDS-PAGE and stained with colloidal blue. Results were subsequently validated by western blotting using anti-Flag and anti-His antibodies. Upon consideration of both gel results, staining and immunoblotting,

the von Willebrand factor type A (vWA) domain of Ku80 was found required for interaction with the histone H2A.Z/H2B dimer (Figure 31). This finding is consistent with the above described results where separate eluted domains were mixed and assayed for interaction *in vitro*.

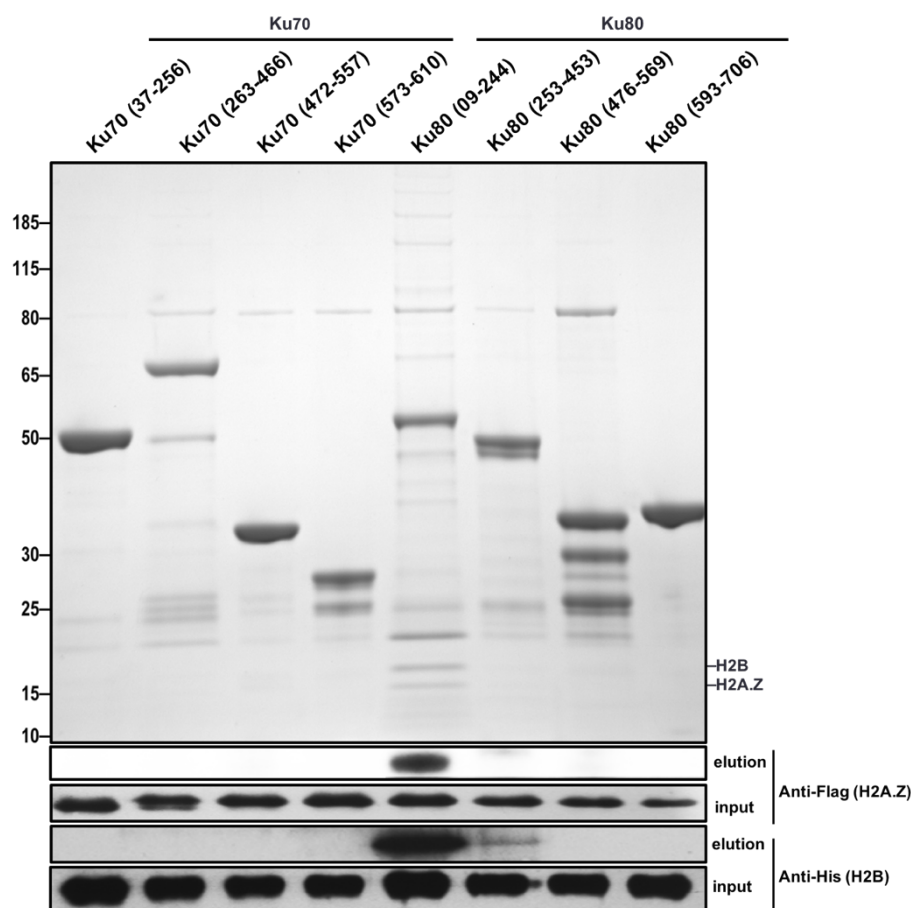


Figure 31: Exploration of specific interaction of vWA domain of Ku80 and H2A.Z/H2B using the bacterial co-expression system: *pET-28b* bicistronic vector containing H2A.Z (Flag epitope) and H2B (His epitope) was co-expressed with the different *pGEX-5X-1* vectors containing GST-tagged Ku70 or Ku80 domains. GST pulled-down proteins were extensively washed with 1M TEGN, eluted, fractionated on SDS-PAGE and stained with colloidal blue. The von Willebrand factor type A (vWA) domain of Ku80 was found to specifically interacts with the H2A.Z/H2B dimers (**top**). Immunoblotting of eluted fractions validated the specific interaction of vWA domain of Ku80 with H2A.Z/H2B (**bottom**).

The data described above established that the N-terminal von Willebrand factor type A (vWA) domain (also known as α/β domain) of Ku80 is responsible for interacting with the histone H2A.Z/H2B dimer. Structurally, the vWA domain consist of a six stranded β sheet forming a fold named Rossman fold. The domain got its name

from its prototype protein von Willebrand factor type A. vWA is present in many plasma proteins, cell surface receptors and a variety of intercellular proteins. The vWA domain of Ku is different from other classic vWA element as it doesn't need metal ion-dependent adhesion site for protein-protein interaction. Several studies have implicated the vWA domain of Ku80 to be responsible for a wide range of protein-protein interaction playing significant role in DSB repair and telomere regulation.

In the literature, a very crucial role of vWA domain of Ku80 is its interaction with aprataxin and polynucleotide kinase/phosphatase like factor (APLF) that is carried out through the hydrophobic pocket of vWA domain of Ku80 (amino acids T61-C156) (Grundy *et al.*, 2013; Grundy *et al.*, 2016; Nemoz *et al.*, 2018). Mutations in L68, Y74 and I112R of Ku80 disrupts the significant interaction of Ku heterodimer with that of APLF that is considered an important component of NHEJ complex (Shirodkar *et al.*, 2013; Grundy *et al.*, 2014; Kragelund *et al.*, 2016). The interaction of Ku heterodimer and APLF subsequently facilitates the recruitment of XRCC4 at the site of DNA damage and stabilizes the repair complex. Furthermore, vWA domain of Ku80 also interact with XLF also known as NHEJ-1 that is believed to be responsible for end-bridging or ligation steps of NHEJ repair pathway (Yano *et al.*, 2011; Menon and Povirk, 2017). The vWA domain of Ku80 has also been reported to recruit WRN helicase that play an important role in cell proliferation and DNA repair (Doherty and Jackson, 2001).

The vWA domain of Ku80 has also been reported to have significant role in telomere maintenance. In *Saccharomyces cerevisiae*, loss of the Ku80 homologue yku80 reported to cause cell death within a few cell divisions due to dramatic shortening of the telomeres (Boulton and Jackson, 1996; Wang *et al.*, 2009). Likewise, in separate studies, mutations in helix 5 of vWA domain Ku80 has been found to retract telomere silencing (Ribes-Zamora *et al.*, 2007). This helix found on the surface of Yku80 vWA domain has been reported to interact with silent information regulator (Sir), a protein responsible for the creation of telomeric heterochromatin (Nugent *et al.*, 1998; Laroche *et al.*, 1998). Furthermore, mutations in the vWA domain of yku80 including L240S, Y49H, M16I, and L149R has been suggested to play a critical role in telomere maintenance (Bertuch and Lundblad, 2003). In other studies, vWA domain of

Ku80 has been reported to interact directly with telomere repeat binding factor 2 (TRF2), a telomere complex component essential to recruit a number of factors and enzymes obligatory for telomeric defense (Ribes-Zamora et al., 2007).

4.5 The specificity of interaction of von Willebrand factor type A (vWA) domain of Ku80 with H2A.Z/H2B:

Having identified the vWA domain of Ku80 as the main interacting domain with the histone H2A.Z/H2B dimer, we next proceeded to evaluate the specificity of this interaction. Toward this goal, we co-expressed the GST-tagged vWA domain individually with H2A/H2B, H2A.X/H2B and H2A.Z/H2B dimers using the plasmids described above. Cells were induced with 1mM IPTG for 3 hours at 37°C and ultimately the soluble fractions of different cell lysates containing GST tagged vWA domain of Ku80 with probable bound histone dimers were pulled-down using Glutathione Sepharose 4B resin. The resin was washed with 1M TGEN to wipe out nonspecific binding between vWA and histones and then eluted using 20 mM L-Glutathione. The soluble fractions (input) were evaluated by western blotting to validated the equal stoichiometric amount of proteins expression in all three experiments. The eluted proteins were then fractionated through SDS-PAGE and stained with colloidal blue and results were also validated through western blotting using anti-Flag and anti-His antibodies. Both gel results, staining and immunoblotting, revealed the specificity of interaction of the vWA domain of Ku80 with the histone H2A.Z/H2B dimer. As expected no interaction could be detected between the vWA domain of Ku80 and H2A/H2B or H2A.X/H2B dimers. The vWA domain maintained the H2A.Z selectivity exhibited by the full-length Ku heterodimer (Figure 32).

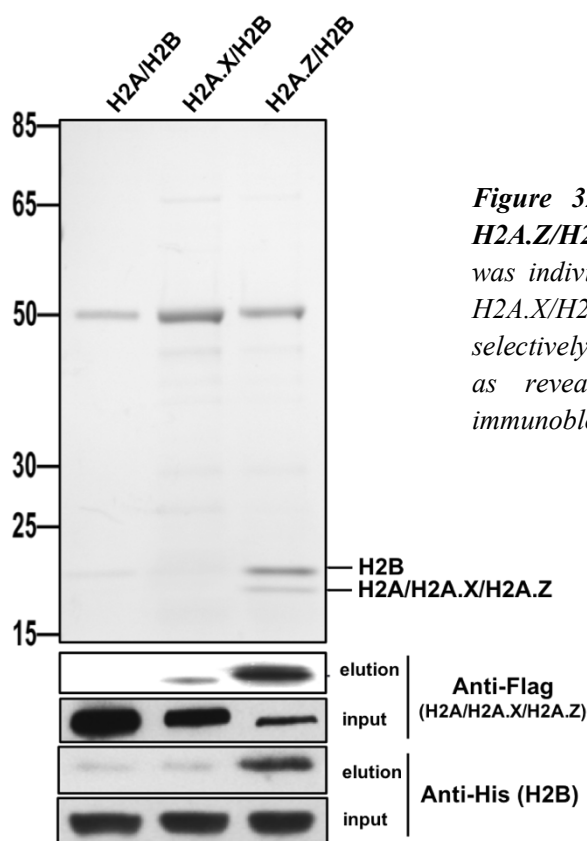


Figure 32: Specificity of vWA domain of Ku80 for H2A.Z/H2B: GST tagged vWA domain (09-244) of Ku80 was individually co-expressed in bacteria with H2A/H2B, H2A.X/H2B or H2A.Z/H2B. The GST-tagged vWA domain selectively binds H2A.Z/H2B over H2A/H2B or H2AX/H2B as revealed by colloidal blue staining (**top**) and immunoblotting (**bottom**).

4.6 Ku80 (40-101) motif is the minimal subdomain responsible for the specific interaction with the histone H2A.Z/H2B dimer:

After identifying the vWA domain of Ku80 as the H2A.Z/H2B interacting domain, we next searched for the minimal vWA domain required for H2A.Z/H2B binding. Based on the secondary and tertiary structural information available in the literature, we subdivided the vWA domain into three subdomains (Figure 33a). The first subdomain encompassed amino acid 09-125. The second subdomain included amino acids 126-244 and the third small 7kDa subdomain included amino acids 40-101.

The three subdomains were amplified and inserted into pGEX-5X-1 vector that already contained N-terminally linked GST epitope. All pGEX vectors were individually co-transformed in BL21 (DE3) pLysS (Invitrogen™) *E.coli* strain along with H2A.Z(Flag)/H2B(His) bicistronic plasmid. The co-expressed proteins were assayed for interaction using the GST-pull-down assay.

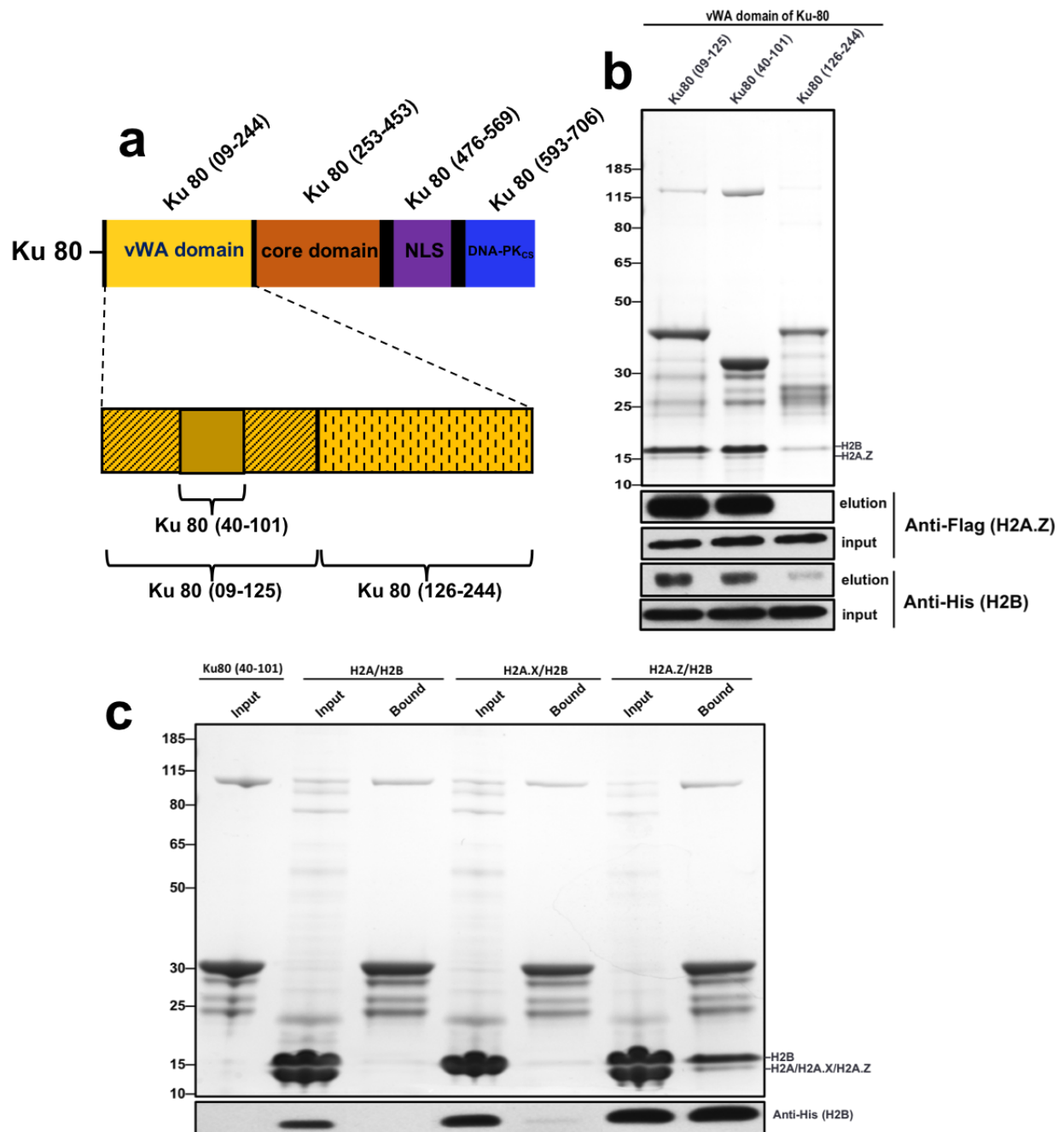


Figure 33: Exploration of minimal subdomain of vWA domain of Ku80 responsible for exclusive interaction with histone dimer H2A.Z/H2B: (a) Primary structure of vWA domain of Ku80 indicating the subdivision of domain into three sub portions including Ku80 (09-125), Ku80 (126-244) and a 7kDa Ku80 (40-101) segment. (b) Histone dimer H2A.Z(Flag)/H2B(His) was co-expressed individually with each of three GST tagged subdomains of vWA domain. Upon immunoaffinity through GST epitope, histone dimer H2A.Z/H2B indicated interaction with Ku80 (09-125) and Ku80 (40-101) subdomains of vWA of Ku80. Resin bound protein(s) were washed with 1M TGEN and the eluted proteins were then fractionated through SDS-PAGE and stained with colloidal blue (top). The stoichiometric expression of histone H2A.Z and H2B in all samples was evaluated through immunoblotting the soluble fractions (input) against respective epitope antibodies. The eluted fractions were also undergone immunoblotting assay to validate the interaction (bottom). (c) The specificity of interaction of 7kDa Ku80 (40-101) subdomain and H2A.Z/H2B was evaluated through in-vitro interaction of Ku80 (40-101) individually with H2A/H2B, H2A.X/H2B and H2A.Z/H2B in TGEN (0.5M NaCl). Subsequent GST pulldown indicated the exclusivity of H2A.Z/H2B and Ku80 (40-101)

The soluble fractions (input) were evaluated through western blotting using anti-Flag and anti-His antibodies to validate the stoichiometric expression of co-expressed proteins in all three experiments. The eluted proteins were then fractionated on SDS-PAGE and stained with colloidal blue. The obtained results were further validated by western blotting using anti-Flag and anti-His antibodies. Both gel results, staining and immunoblotting, suggested that the Ku80 domain encompassing amino acids 40-101 is the minimal domain required for interacting with H2A.Z/H2B (Figure 33b).

The specificity of interaction of H2A.Z/H2B dimer with the minimal Ku80 (aa 40-101) domain was further evaluated by assaying its interaction with the other histone dimers *i.e* H2A/H2B and H2A.X/H2B. The GST-Ku80 (40-101) pulldown assay revealed that only H2A.Z/H2B dimer is able to bind Ku80 (aa 40-101) (Figure 33c). In conclusion, our data revealed that Ku heterodimer Ku70/Ku80 specifically interacts with histone dimer H2A.Z/H2B through the Ku80 vWA domain with amino acids 40-101 forming it the critical minimal binding domain.

4.6.1 Sequence conservation of Ku80 (40-101) motif responsible for H2A.Z/H2B specific interaction:

In a given protein, all residues are not equally important for the proper structure and function of the protein rather only a small proportion of it is critical for its operation. The rest of the protein can be readily replaced without any significant change in functioning. The chunk of residues that are critical for the functionality of a particular protein is usually conserved among species over the period of evolution. Conservation analysis is one of the most widely used methods for predicting these functionally important residues in protein sequences. Sequence alignment of Ku80 amino acids 40-101 among various eukaryotic species, showed a strong conservation of this subdomain during evolution indicating its functional significance in different species. A high level of sequence consensus was found among major eukaryotic species (Figure 34). As per secondary structure, Ku80 (40-101) subdomain is comprised of highly conserved two α helices and two β sheets.

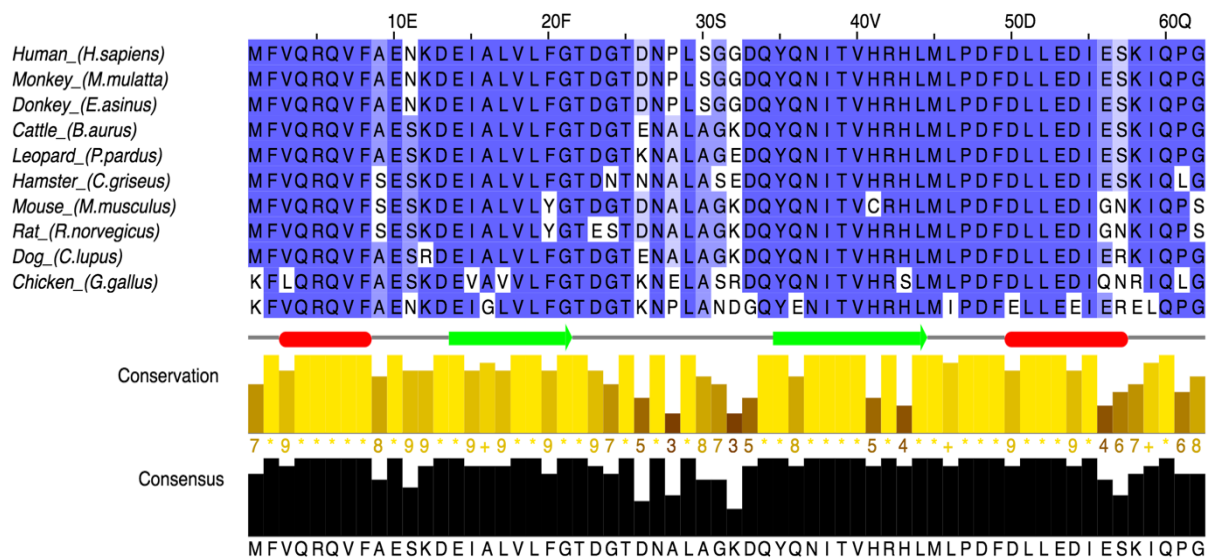


Figure 34: Sequence alignment of Ku80 (aa 40-101) subdomain responsible for the interaction with H2A.Z/H2B: Evident sequence conservation of 7kDa Ku80 (aa 40-101) in various eukaryotic species is indicated through shading of amino acids, from dark blue (high) to light blue (low) and also through annexed annotation. The secondary structure referred indicates the sites of α -helix (red cylinders) and β -sheet (green arrows). The sequence alignment is performed using Jalview® alignment program.

4.7 Identification of H2A.Z motif critical for Ku heterodimer Ku70/Ku80 interaction:

Following the identification of the minimal domain of Ku heterodimer responsible for interaction with H2A.Z/H2B, we aimed to identify the H2A.Z/H2B domains critical for this interaction. Towards this goal, we constructed a series of pET28b bicistronic vectors containing His-tagged H2B followed by different Flag-tagged chimeras of H2A/H2A.Z (Figure 35a). Chimeras were designed based on structural and functional data available in the literature. We started truncation of H2A.Z from the C-terminus (1-18, 1-44, 1-82, 1-105) and replaced it with counterpart C-terminus of H2A. An additional chimera was designed that contain the middle part of H2A.Z (45-82) that primarily comprised of α 2 helix of H2A.Z, connected with H2A brims on both N and C terminus (Figure 35a).

Following the amplification of chimeric genes into pET28b bicistronic plasmid that already contained H2B as the second partner, we transformed them in BL21 (DE3) pLysS (Invitrogen™) *E.coli* strains and induced cells with 1mM (IPTG) for 3

hours at 37°C. Subsequently, the soluble fractions of the different cell lysates were subjected to tandem-affinity purification first using anti-His and then anti-Flag purification. On the other hand, we separately purified GST-tagged 7kDa Ku80 (aa 40-101) using same expression conditions. Next, we dialyzed both the histone dimers and Ku80 (40-101) in 0.15M TGEN using appropriately sized SnakeSkin™ Dialysis Tubing. Thereafter, stoichiometric amount of histone dimers was mixed individually with dialyzed Ku80 (40-101) in TGEN buffer with final concentration of 0.5M NaCl for 3 hours at 4°C. The interacted proteins were then subjected to GST pull-down. Resin bound protein(s) were washed with 1M TGEN and eluted proteins were then fractionated on SDS-PAGE assay. Input fractions were also run to confirm the stoichiometric amounts of histones used in the assay.

Upon colloidal blue staining, the central part of H2A.Z comprising the highly conserved $\alpha 2$ helix was found to be critical for interaction with Ku80 (aa 40-101). All the three chimeric constructs (chimera 3, chimera 4 and chimera 5) containing middle part of H2A.Z (45-82) presented a significant interaction with Ku80 minimal motif. The presence of H2B in input and eluted fractions was validated through immunoblotting assay using anti-His antibody (Figure 35b). Our data highlight the importance of the $\alpha 2$ helix of H2A.Z for the interaction with Ku 80. This result is in agreement with several previous studies on the significance of $\alpha 2$ helix (Hong *et al.*, 2014; Ranjan *et al.*, 2015) and recent structural studies indicated the significance of A57 of the highly conserved $\alpha 2$ helix of H2A.Z for the recruitment of H2A.Z-specific chaperone Chz1 (Wang *et al.*, 2019).

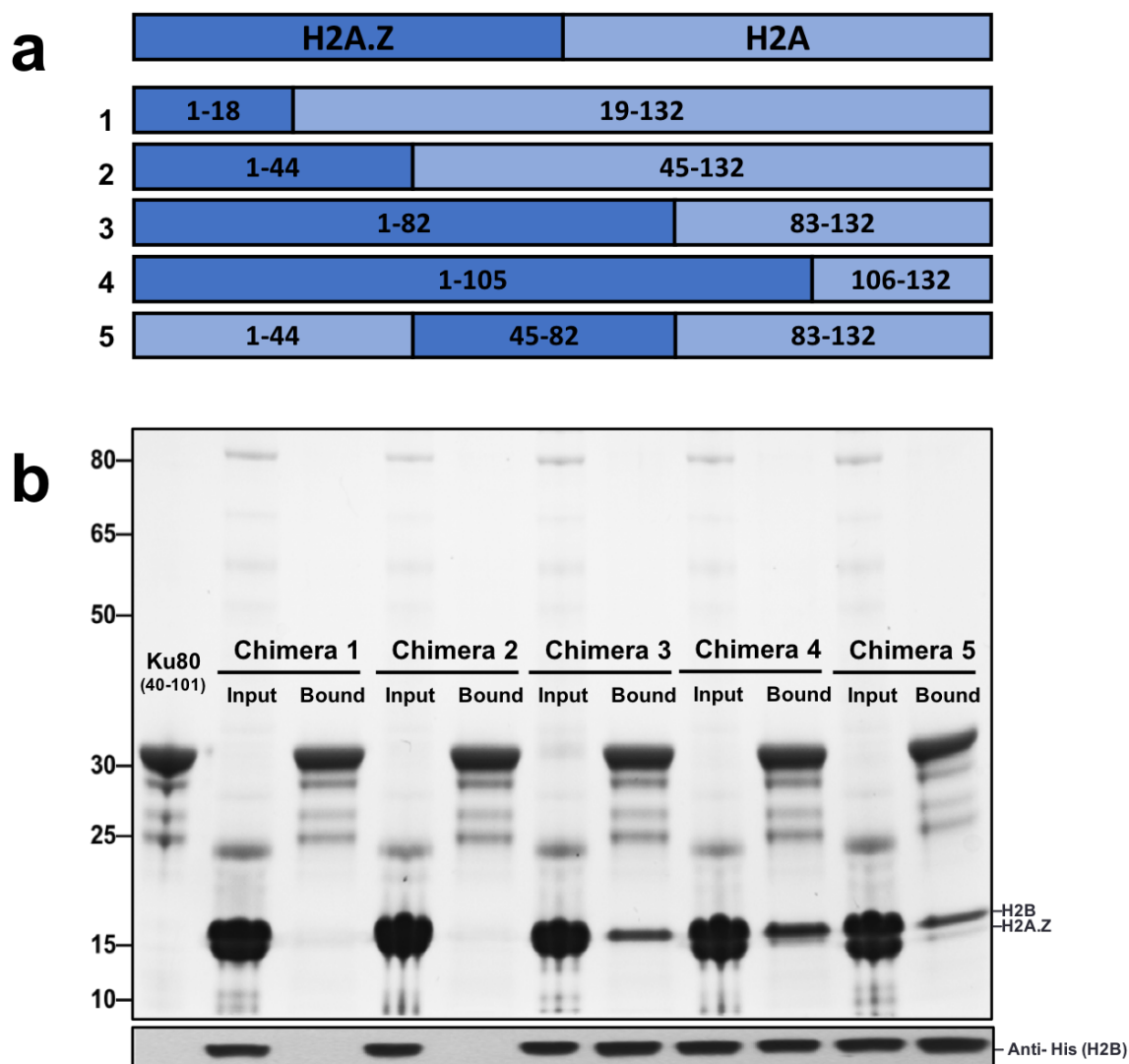


Figure 35: Domain mapping of H2A.Z responsible for interaction with Ku heterodimer: (a) Schematic representation of chimeric proteins constructed from combinations of H2A and H2A.Z. Chimeric genes were prepared by fusing the structural genes incorporating H2A.Z at N-terminal and H2A at C-terminal. The amplified chimeric genes were then inserted in suitable bicistronic expression vector that already contained H2B. (b) Purified chimeric proteins were mixed with GST tagged Ku80 (aa 40-101) in 0.5M TGEN buffer and subjected to GST-pull-down. Resin bound protein(s) were washed with 1M TGEN and eluted proteins were fractionated on SDS-PAGE and colloidal blue stained. The central part of H2A.Z comprising the $\alpha 2$ helix was found to be responsible for interaction with Ku80 (aa 40-101). The presence of H2B in the input and eluted fractions was validated by Western blotting using anti-His antibody.

4.7.1 Sequence conservation of H2A.Z motif in various eukaryotic species:

Among different computational methods established to predict functionally important residues, the analysis of multiple sequence alignment (MSA) is one of the

most widely used method for predicting the functional importance of specific amino acid residues. Despite the fact that most of H2A.Z pertaining interactions are carried out by $\alpha 3$ or αC helices of H2A.Z, the highly conserved orthologous sequences of $\alpha 2$ helix of histone variant H2A.Z among different eukaryotic species ranging from yeast to humans indicate the functional significance of this helix containing motif (Figure 36).

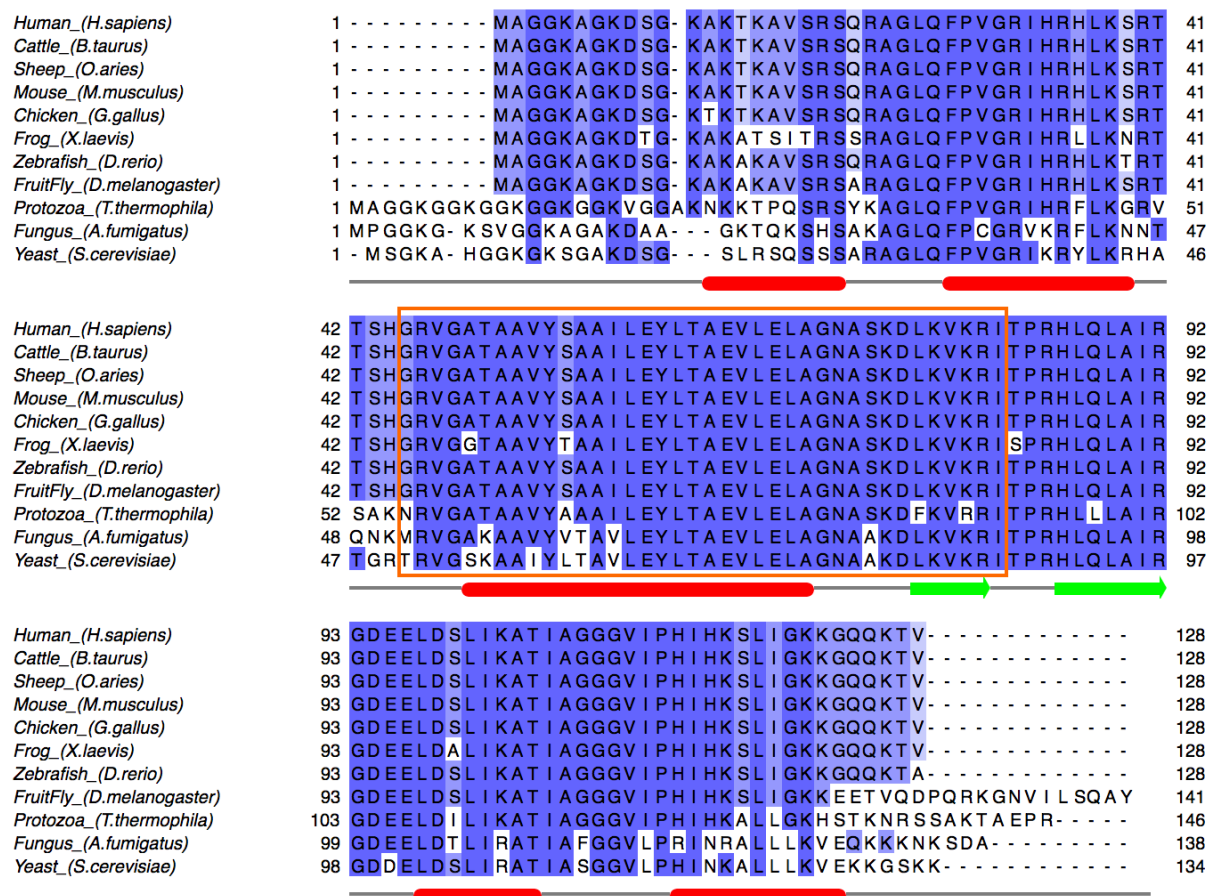


Figure 36: Multiple Alignment of Histone Variant H2A.Z from Various Eukaryotic Species: Evident sequence conservation of histone variant H2A.Z in different organisms is shown by hues of blue, from dark blue (high) to light blue (low). The annotation of secondary structure shows the sites of α -helix (red cylinders) and β -sheet (green arrows). Red box highlights the sequence of H2A.Z comprising $\alpha 2$ -helix that is responsible for interaction of histone H2A.Z/H2B dimer with Ku heterodimer. The sequence alignment is performed using Jalview® alignment program.

4.8 The interaction of Ku80 (aa 40-101) motif with histone dimer H2A.Z/H2B is mediated favorably through H2B:

During all the experiments conducted to assess the interaction of Ku80 (aa 40-101) motif with histone H2A.Z/H2B dimer, the amount of H2B was always found substantially higher as compared to H2A.Z suggesting that Ku80 (aa 40-101) motif might be binding to histone dimer H2A.Z/H2B favorably through H2B ally. This hinted us to evaluate the interaction of Ku80 (aa 40-101) motif with H2B alone. For this purpose, we separately expressed His-tagged H2B in BL21 Rosetta™ (DE3) *E.coli* cells. Considering that the histone monomer H2B alone is insoluble we decided to harvest it from insoluble inclusion bodies. Since the binding of histidine tag to immobilized divalent metal ion can occur in the presence of any chaotropic agent at high concentration, we solubilized the inclusion bodies containing H2B using 8M urea and subsequently refolded it by gradual slow exchange to non-denaturing buffer of near neutral pH and finally eluted it in buffer containing 20mM Tris-HCl, 500mM NaCl, 500mM imidazole and 2mM 2-mercaptoethanol pH 8.0. On the other hand, GST-tagged Ku80 (40-101) was purified separately using same conditions as described in previous section. The expression of both GST-tagged Ku80 (aa 40-101) and H2B was assessed by running input (control) fractions (Figure 37 lane 1 & lane 2).

Stoichiometric amount of both GST-tagged Ku80 (aa 40-101) and H2B were mixed in TGEN buffer with final concentration of 0.5M NaCl for 3 hours at 4°C. The interacted proteins were then subjected to GST pull-down. Resin bound protein(s) were washed with 1M TGEN to wipe out any non-specific binding and finally eluted proteins were fractionated on SDS-PAGE. The presence of H2B in input and eluted fractions was validated by western blotting using anti-His antibody. Taking into consideration the gel staining and immunoblotting results, Ku80 (aa 40-101) motif exhibited a very strong interaction with H2B indicating the pivotal role of H2B in Ku heterodimer and histone H2A.Z/H2B dimer interaction (Figure 37). This definite interaction also explained the high-density bands of H2B as compared to H2A.Z in all previous experiments of interaction with Ku80 (aa 40-101) motif.

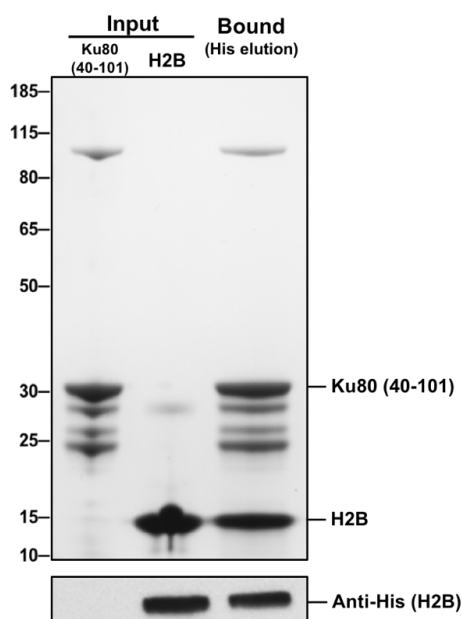


Figure 37: Ku80 (40-101) domain interacts explicitly with H2B in-vitro: GST-tagged Ku80 (40-101) and His-tagged H2B were separately expressed using *E. coli* expression system. Following their dialysis in 0.15M TGEN, stoichiometrically equal amounts of both proteins were made interacted in TGEN buffer with final concentration of 0.5M NaCl. Subsequent GST pulldown indicated the stable interaction of H2B to that of Ku80(40-101). Resin bound protein(s) were washed with 1M TGEN and the eluted proteins were then fractionated through SDS-PAGE and stained with colloidal blue. The presence of H2B in input and eluted fractions was validated through immunoblotting assay using anti-His antibody.

To validate the interaction of Ku80 (aa 40-101) motif with H2B and to rule out any non-specific GST epitope driven interaction with H2B, we aimed to re-evaluate this interaction using co-expression strategy employing bicistronic vector. We amplified His-tagged H2B and Flag-tagged Ku80 (aa 40-101) in the bicistronic pETDuetTM vector and co-expressed it in BL21 RosettaTM (DE3) *E.coli* cells. The soluble fractions of cell lysate were then first subjected to anti-Flag immunoprecipitation and thereafter anti-His purification. After both anti-Flag and anti-His purifications, the eluted proteins were fractionated on SDS-PAGE (18% precast tris-glycine gel) and stained with colloidal blue. The input (soluble) fractions were also evaluated to validate substantial expression of Ku80 (aa 40-101) and H2B. The presence of both H2B and Ku80 (aa 40-101) in input and eluted fractions were also validated through immunoblotting using anti-His and anti-Flag antibodies (Figure 38). Both gel results, staining and immunoblotting, were found in compliance with previous results suggesting the dedicated role of H2B in novel interaction of Ku heterodimer Ku70/Ku80 with histone dimer H2A.Z/H2B. However, taking into account the previous results this interaction is indicated to be possible only in case of H2A.Z/H2B and not in case of H2A/H2B and H2A.X/H2B.

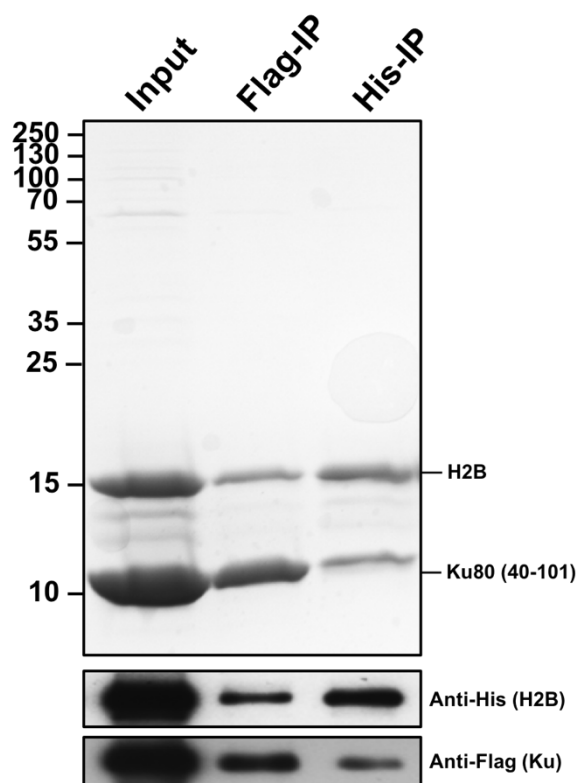


Figure 38: Interaction of Flag-tagged Ku80 (aa 40-101) subdomain with His-tagged H2B *in vitro*: Flag-tagged Ku80 (aa 40-101) and His-tagged H2B were co-expressed in *E. coli*. The co-expressed proteins were pulled-down using anti-Flag followed by anti-His purification (**top**). The presence of H2B and Ku80 (aa 40-101) in input and eluted fractions was validated through Western blotting using anti-His and anti-Flag antibodies (**bottom**).

4.9 Evaluation of interaction of Ku80 (aa 40-101) motif with histone dimer H2A.Z/H2B using size-exclusion chromatography (SEC):

A variety of *in vitro* techniques have been established to investigate protein-protein interactions (PPIs). A very customary method includes the size-exclusion chromatography in which molecules in solution are separated by their size. Gel-filtration chromatography is the most common approach for size-exclusion chromatography that take account of aqueous solution to transport the proteins through the columns. Following our previous results from co-immunoprecipitation and tandem affinity purification-mass spectroscopy (TAP-MS) approaches, we decided to validate our results through chromatographic approaches as well. For that purpose, we first separately purified GST-tagged Ku80 (aa 40-101) and histone H2A.Z(Flag)/H2B(His) dimer through affinity and ion-exchange chromatography approaches. Thereafter, we mixed both partners at high protein concentration and fractionated them on HiLoad™ 16/600 Superdex 200 prep grade gel-filtration chromatography using ÄKTA pure protein purification system™.

4.9.1 Purification of Ku80 (aa 40-101) and histone H2A.Z/H2B dimer through affinity and ion-exchange chromatography:

GST-tagged Ku80 (aa 40-101) and histone H2A.Z(Flag)/H2B(His) dimer were expressed separately in BL21 Rosetta™ (DE3) *E.coli* cells using pGEX-5X-1 and pET28b plasmids respectively. The soluble fractions of cell lysates were subjected to affinity chromatography using Glutathione S-transferase resin for Ku80 (aa 40-101) and Ni-NTA agarose for histone H2A.Z/H2B. The proteins were eluted in respective high salt concentration and dialyzed to reduce the salt concentration to the minimal level (50 mM NaCl). GST-tagged Ku80 (aa 40-101) was further purified on 26/10 quaternary ammonium Q Sepharose High Performance™ column while H2A.Z/H2B purified on 26/10 sulphopropyl SP Sepharose High Performance™ column.

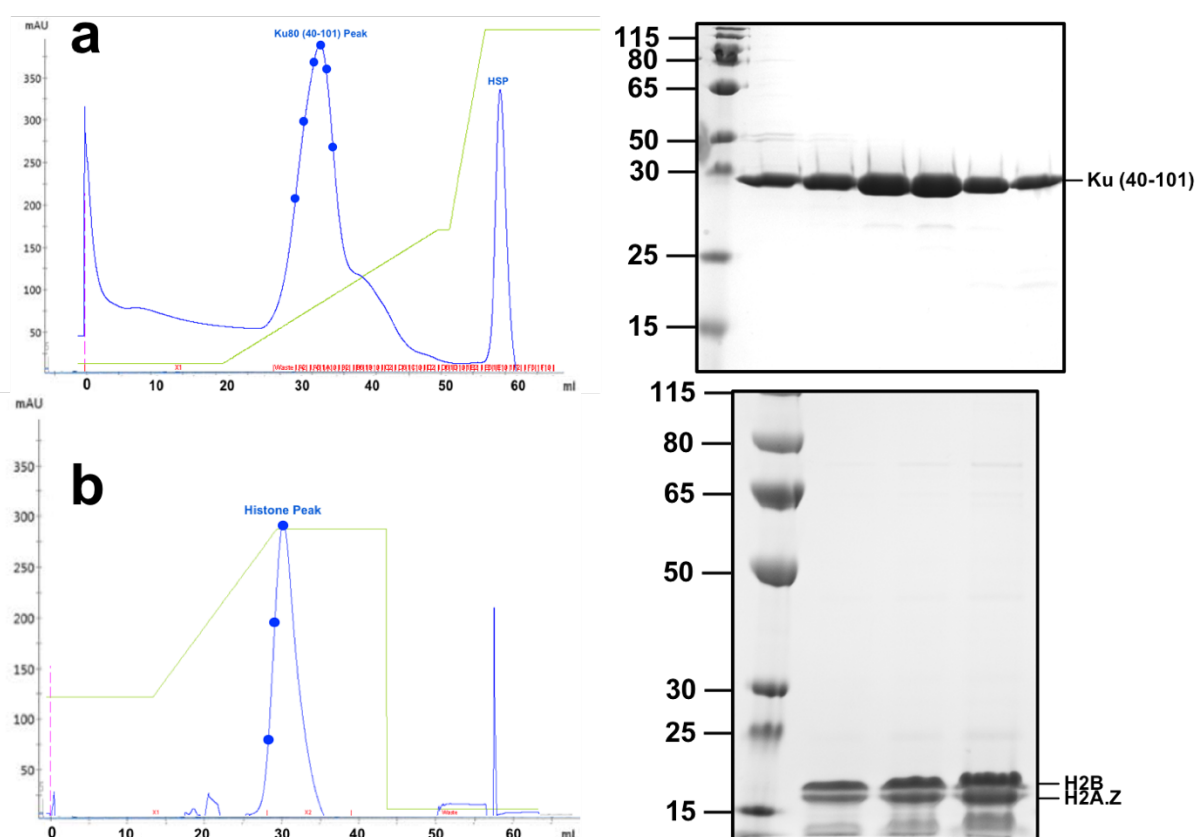


Figure 39: Ion exchange chromatograms of Ku80 (aa 40-101) and histone H2A.Z/H2B dimer: GST tagged Ku80 (aa 40-101) and Flag/His tagged H2A.Z/H2B were affinity purified and then subjected to ion exchange chromatography using 26/10 quaternary ammonium Q Sepharose High Performance™ column for Ku80 (aa 40-101) and 26/10 sulphopropyl SP Sepharose High Performance™ column for H2A.Z/H2B. Peak fractions were concentrated with an Amicon™ ultra centrifugal filter units. A single peak in both ion exchange chromatograms indicate the homogeneity of eluted proteins. The purity of the different proteins was assessed by running the samples on NuPAGE™ 4-12% Bis-Tris protein gels.

Both columns were equilibrated with a buffer containing 50 mM Tris (pH 8.0), 50 mM NaCl and 2 mM DTT. Following washing with same equilibration buffer the proteins were eluted with a gradient of increasing salt concentration ranging from 50 to 1000 mM NaCl. Peak fractions were then concentrated with an Amicon™ ultra centrifugal filter units. A single peak in both ion exchange chromatograms indicate the homogeneity of eluted proteins (Figure 39 a & b; left panels). In both cases, the purified proteins were found pure and free of any significant containment. Small fractions of eluted proteins were then analyzed on NuPAGE™ 4-12% Bis-Tris protein gels (Figure 39 a & b; right panels). The eluted proteins were further concentrated to the maximum possible levels and then mixed and incubated overnight at 4°C to allow complex formation before gel filtration analysis.

4.9.2 Gel filtration chromatography of Ku80 (aa 40-101) motif and histone dimer H2A.Z/H2B complex indicated the favored interaction of Ku motif with H2B:

Following affinity and ion exchange chromatography, equimolar amount of the concentrated Ku80 (aa 40-101) and histone H2A.Z/H2B dimer were mixed together and incubated overnight at 4°C and then subsequently loaded onto HiLoad™ 16/600 Superdex 200 prep grade gel filtration column (GE Healthcare). The gel filtration column was pre-equilibrated with gel filtration buffer containing 50mM Tris-Cl pH 8.0, 50mM NaCl and 2mM DTT. Gel filtration chromatogram clearly indicate the fractionation of the complex in two distinct peaks (Figure 40a). The peak fractions (peak 1 and peak 2) run on a 4-12% NuPAGE™ Bis-Tris protein gel validated this observation. Peak 1 represented essentially the interaction of Ku80 (aa 40-101) with H2B and Peak 2 contained the fraction of H2A.Z dissociated from the complex together with the free H2A.Z/H2B dimers (Figure 40b). The complex eluted in the second peak had a higher molecular weight than expected indicating the presence unfolded histone proteins. The dynamic light scattering (DLS) analysis of the peak fraction containing Ku indicated the presence of high molecular weight content that also suggest the presence of soluble unfolded protein aggregates.

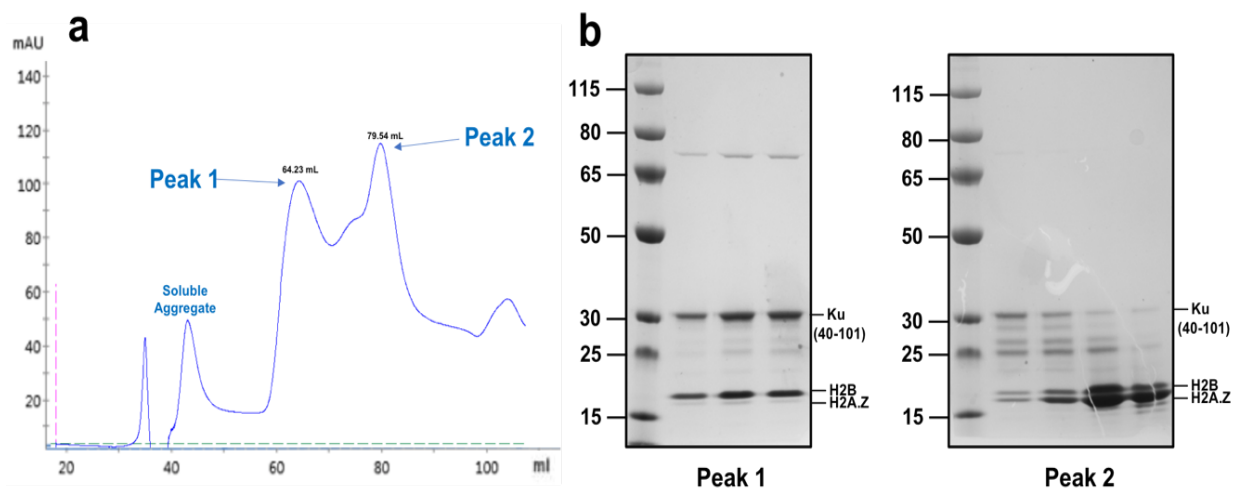


Figure 40: Purification Profile of Ku80 (aa 40-101) and Histone H2A.Z/H2B dimer: Following affinity and ion exchange chromatography, both Ku80 (aa 40-101) and histone H2A.Z/H2B dimer were incubated overnight on a roller shaker at 4°C and then subsequently loaded onto **HiLoad™ 16/600 Superdex 200 prep grade** gel filtration column (GE Healthcare). The gel filtration column was pre-equilibrated with gel filtration buffer including 50mM Tris pH 8.0, 50mM NaCl and 2mM DTT. The buffer used for gel filtration purification included 10mM Tris pH 8.0, 200mM NaCl, 2mM DTT). **(a)** Gel filtration chromatogram indicated two distinct peaks representing interaction of Ku80 (aa 40-101) with H2B in the first peak and the fraction of H2A.Z dissociated from the complex in the second peak. **(b)** The peak fractions (Peak 1 & Peak 2) were analyzed by running the eluted samples on NuPAGE™ 4-12% Bis-Tris protein gels.

4.10 Exploration of interaction of Ku heterodimer Ku70/Ku80 and histone H2A.Z/H2B dimer in mouse embryonic fibroblasts (MEFs):

Primary mouse embryonic fibroblasts (MEFs) are a popular tool for molecular and cell biology studies. After the determination of the *in vivo* interaction of Ku heterodimer Ku70/Ku80 with the histone H2A.Z/H2B dimer in HeLa cells followed by the exploration of this interaction *in vitro*, we next proceeded to validate the functionality of this interaction *in vivo* by using Ku80 knockout mouse embryonic fibroblasts (MEFs). We established two distinct stable MEF cell lines expressing N-terminally FLAG/HA epitope-tagged wild type human Ku80 (1-732) or ΔN Ku80 (107-732) by retroviral transduction and anti-CD4 immunomagnetic selection of positive cells.

The proportion of positive cells was then estimated by immunofluorescence using anti-HA antibody (Figure 41a) and by immunoblotting using anti-Flag antibody (Figure 41b). DAPI (4',6-diamidino-2-phenylindole) staining was used as a positive

control in immunofluorescence assays (Figure 41a). The images shown in Figure 41a were acquired using ZEISS Axioplan-2™ Upright Fluorescent Microscope with Photometrics CoolSNAP HQ2 digital camera using Openlab software.

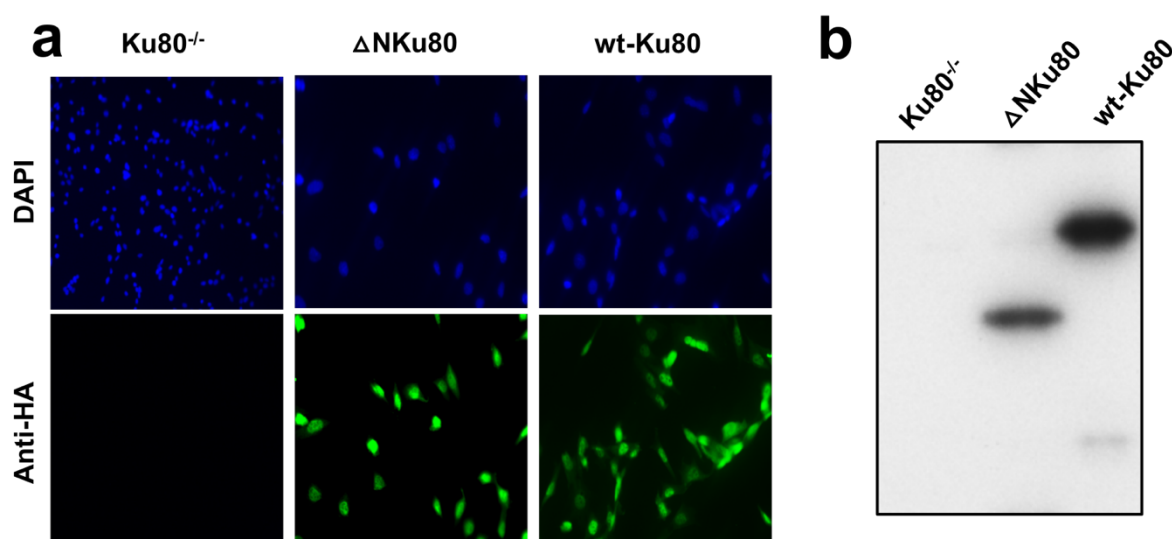


Figure 41: Validation of stable expression of Δ NKu80 and wtKu80 in $Ku80^{-/-}$ MEFs by immunofluorescence and immunoblotting: (a) MEFs expressing Flag/HA tagged Δ NKu80 (107-732) and full-length WT-Ku80 along with control $Ku80^{-/-}$ cells were analyzed by immunofluorescence and Western blotting for positive expression. (a) DAPI staining (top) and anti-HA (bottom). (b) Control $Ku80^{-/-}$ (lane1), Δ NKu80 (lane2) and WT-Ku80 (lane 3) MEF cell line extracts were immunoblotted with anti-Flag antibody to confirm positive expression of the transduced cDNAs.

Following establishment of stable MEF cell lines for both wt-Ku80 (1-732) and Δ NKu80 (107-732), the cells were grown in large quantity and subsequently both insoluble nuclear extracts (INE) and soluble nuclear extracts (SNE) were prepared according to the procedure described previously. The *e*-wt-Ku80 and *e*- Δ NKu80 complexes were purified from the soluble nuclear extracts (SNE) and the insoluble nuclear extracts (INE) using tandem immuno-affinity purification anti-FLAG followed by anti-HA. The purified complex was then fractionated on SDS-PAGE and subjected to silver staining (Figure 42). Mass spectrometric analysis identified several NHEJ related proteins as part of the *e*-WT-Ku80 complex (such as DNA-PKcs, Ku70, PARP1, PARP2, HLTF, CHD1L, H2A.Z, H2B, H3, H4, etc) (Table 42). Interestingly, *e*- Δ NKu80 complex has lost all the NHEJ proteins and the histones particularly H2A.Z and H2B (Table 42). These data further validated the functional importance of the

vWA domain of Ku80 in recruiting the H2A.Z/H2B dimer and stabilizing the NHEJ complex.

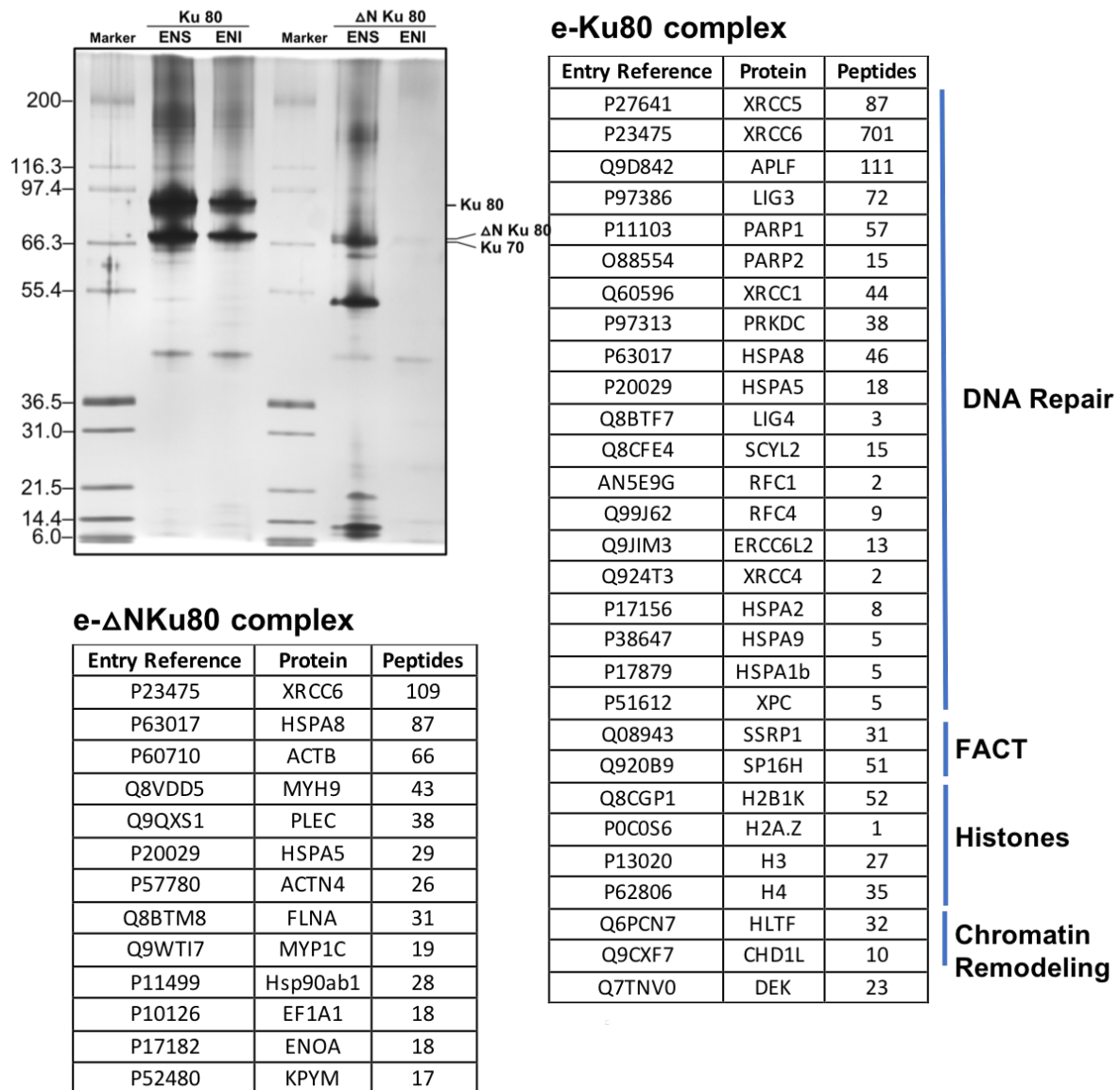


Figure 42: Mass spectrometry analysis of e-ΔNku80 and e-wtKu80 subcomplexes: Many proteins related to DNA repair mechanism are found in subcomplex of e-wtKu80. Histone members were also found in the complex. However, both the DNA repair proteins and histones were found absent in e-ΔNku80 subcomplex indicating the important role of vWA domain of Ku80 in important cellular processes.

4.11 Ku70/Ku80 heterodimer regulates H2A.Z deposition at promoters but not at enhancers:

If Ku heterodimer Ku70/Ku80, as strongly suggested by our data, is involved in H2A.Z functioning *in vivo*, its depletion should affect its dynamic or impair its

recruitment to chromatin. Our data point to an important role of Ku80 (aa 40-101) subdomain in H2A.Z/H2B functioning. Therefore, to study the effect of Ku80 depletion on H2A.Z localization *in vivo*, we performed a genome-wide comparative ChIP-seq analysis using both wild type and knockout Ku80 MEF cells.

Our data clearly indicated that Ku80 is involved in genome-wide localization and chromatin enrichment of H2A.Z at promoters. The absence of Ku80 resulted in strong reduction of H2A.Z deposition particularly at promoters (distance to nearest transcription start site TSS > 2 kilobases) (Figure 43a). However, no significant variance was found in the deposition of H2A.Z at enhancers and insulators. Interestingly, although no new H2A.Z peaks were observed at promoters in knockout cells (Ku80^{-/-} MEFs) the net amount of promoter-associated H2A.Z increased significantly by 50% and presented similar distribution around TSS compared with wild-type cells (Ku80^{+/+} MEFs). As described previously (Guillemette *et al.*, 2005; Mavrich *et al.* 2008 and Nekrasov *et al.* 2012), two histone H2A.Z-containing nucleosomes were found positioned on either side of the TSS (Figure 43b). These data strongly suggested that Ku80 is a critical H2A.Z interacting protein playing an important role in the genome-wide deposition of H2A.Z at its normal sites of deposition particularly in promoter regions.

After determining the role of Ku80 in H2A.Z deposition we proceeded towards analyzing the role of Ku80 (aa 40-101) motif in this deposition. Toward this goal, we rescued the Ku80^{-/-} MEF cells with Δ NKu80 (aa 107-732) using retroviral transduction, as described above. Afterwards we performed genome-wide comparative ChIP-seq analysis using both WT-Ku80 and Δ NKu80 (aa 107-732) rescued Ku80^{-/-} cell lines. Quite interestingly, Ku80 knockout MEF cell line rescued with full-length wild type Ku80 showed a higher level of H2A.Z deposition at promoters when compared to Ku80 knockout MEF cell line rescued with Δ NKu80 (107-732). The density of H2A.Z deposition at promoters (distance to nearest transcription start site TSS > 2 kilobases) was estimated to be almost double in Ku80^{-/-} MEFs rescued with full-length wild type Ku80 compared to Δ NKu80 (107-732) (Figure 43c). This finding undoubtedly advocated the critical role of vWA domain and more precisely Ku80 (40-101) motif in the deposition of H2A.Z at promotor regions.

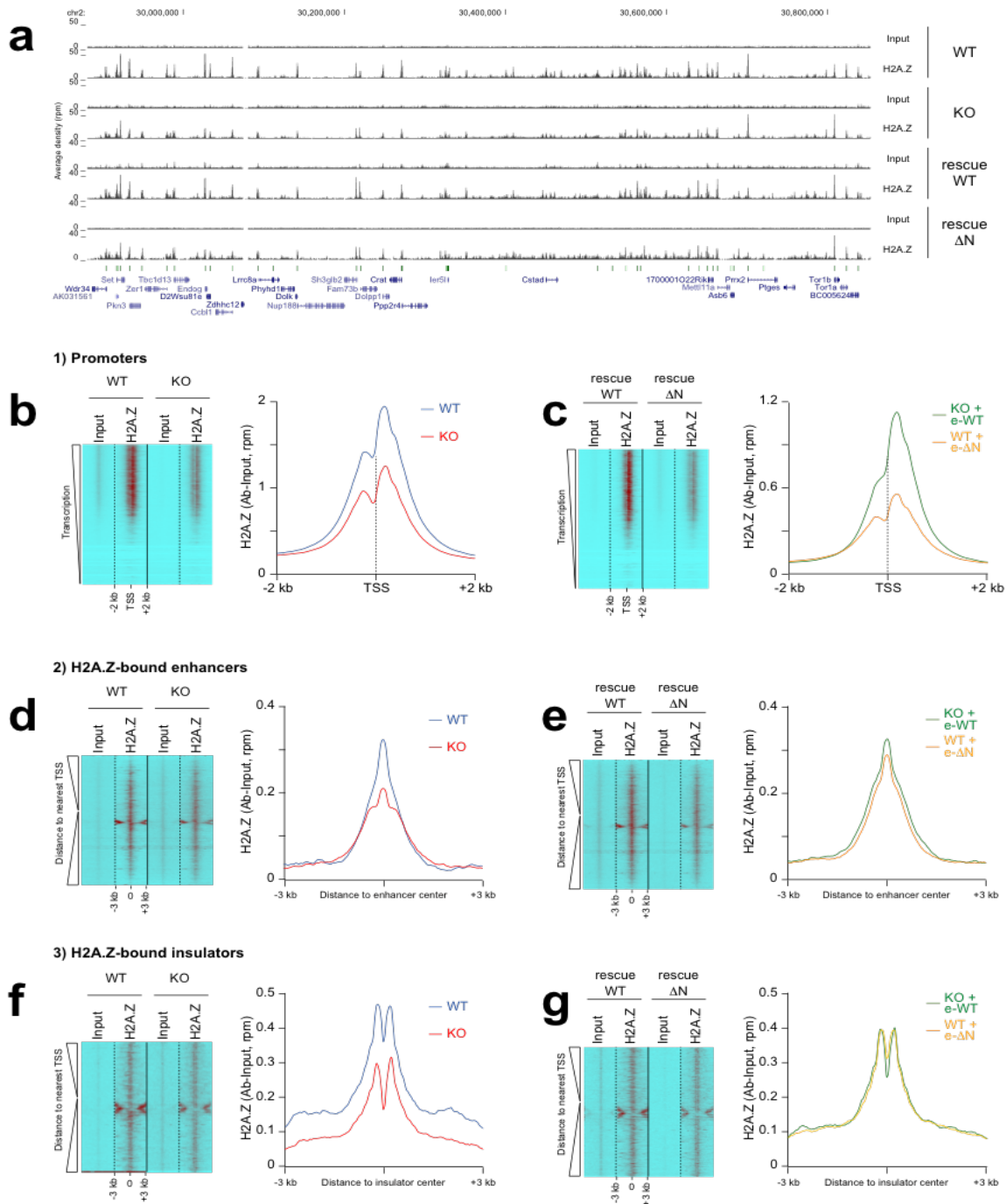


Figure 43: Ku80 is Involved in Genome-wide Localization and Chromatin Enrichment of H2A.Z at Promoters: (a) The genomic browser views indicating the depleted H2A.Z deposition at promoters (distance to nearest transcription start site TSS > 2 kilobases) in Ku80 knockout ($Ku80^{-/-}$) mouse embryonic fibroblasts in comparison to wild type $Ku80^{+/+}$ MEFs. The rescuing of full-length wild type Ku80 in $Ku80^{-/-}$ MEFs indicated higher level of H2A.Z deposition at promoters as compared to $Ku80^{-/-}$ MEFs rescued with ΔN Ku80 (107-732). (b) Density of H2A.Z deposition across transcription start site in Ku80 knockout MEFs (red) compared with wild type $Ku80^{+/+}$ MEFs (blue). The inactivation of Ku80 in $Ku80^{-/-}$ MEFs indicated more than 50% decrease in the level of H2A.Z deposition at promoters as compared to wild type MEFs. (c) The level of H2A.Z deposition at promoters is about double in $Ku80^{-/-}$ MEFs rescued with full-length wild type Ku80 (green) as compared to $Ku80^{-/-}$ MEFs rescued with ΔN Ku80 (107-732) (orange). This suggested the significance of Ku80 (40-101) subdomain in deposition of H2A.Z at promoters. (d) (e) (f) (g) The level of H2A.Z deposition at enhancers and insulators.

4.12 Structural Modeling:

In a nutshell, our both *in-vivo* and *in-vitro* results indicated that Ku80 (40-101) motif specifically interact with histone dimer H2A.Z/H2B and not to counterpart H2A/H2B or H2A.X/H2B dimers. In order to strengthen our hypothesis, with the collaboration of our partner research laboratories, we are already progressing positively towards solving the crystal structure of this complex. However, for now, we performed the *in-silico* docking of Ku80 (40-101) motif independently with H2A/H2B and H2A.Z/H2B histone dimers using ClusPro® docking server, a most commonly used online server for protein-protein docking. For this purpose, we retrieved the crystal structure of Ku80 (40-101) motif from an already presented structure of Ku heterodimer (PDB ID: 6ERH). The crystal structures of histone dimer H2A/H2B (PDB ID: 3MGR) and H2A.Z/H2B (PDB ID: 1AOI) were also acquired from RCSB Protein Data Bank. All structures were then docked using ClusPro® docking server. Majority of top predicted binding models revealed that Ku80 (40-101) motif favorably binds with H2A.Z/H2B (Figure 44). Based on all the major interactions (electrostatic, hydrophobic, and vander waal force + electrostatic), a stronger interaction for H2A.Z/H2B/Ku80(40-101) complex was predicted as compared to H2A/H2B/Ku80(40-101) complex.

After that, the weighed scores of predicted models were compared using the same ClusPro® server. The ‘weighed scores’ are dependent on the size of cluster (number of docked structures) and PIPER energies of the models. Higher number of docked structures in conjunction with lower free energies predict a favorable binding between the two proteins. When compared, models including the interaction of Ku80 (40-101) motif with H2A.Z/H2B dimer had better scores as compared to models comprising the interaction of Ku80 (40-101) motif with canonical H2A/H2B dimer. This predicts a relatively higher chances of interaction of Ku80 (40-101) motif with H2A.Z/H2B as compared to H2A/H2B. In this way the *in-silico* modeling of interacting members supports our hypothesis that Ku heterodimer favorably binds to H2A.Z/H2B dimer (Figure 44 a & b).

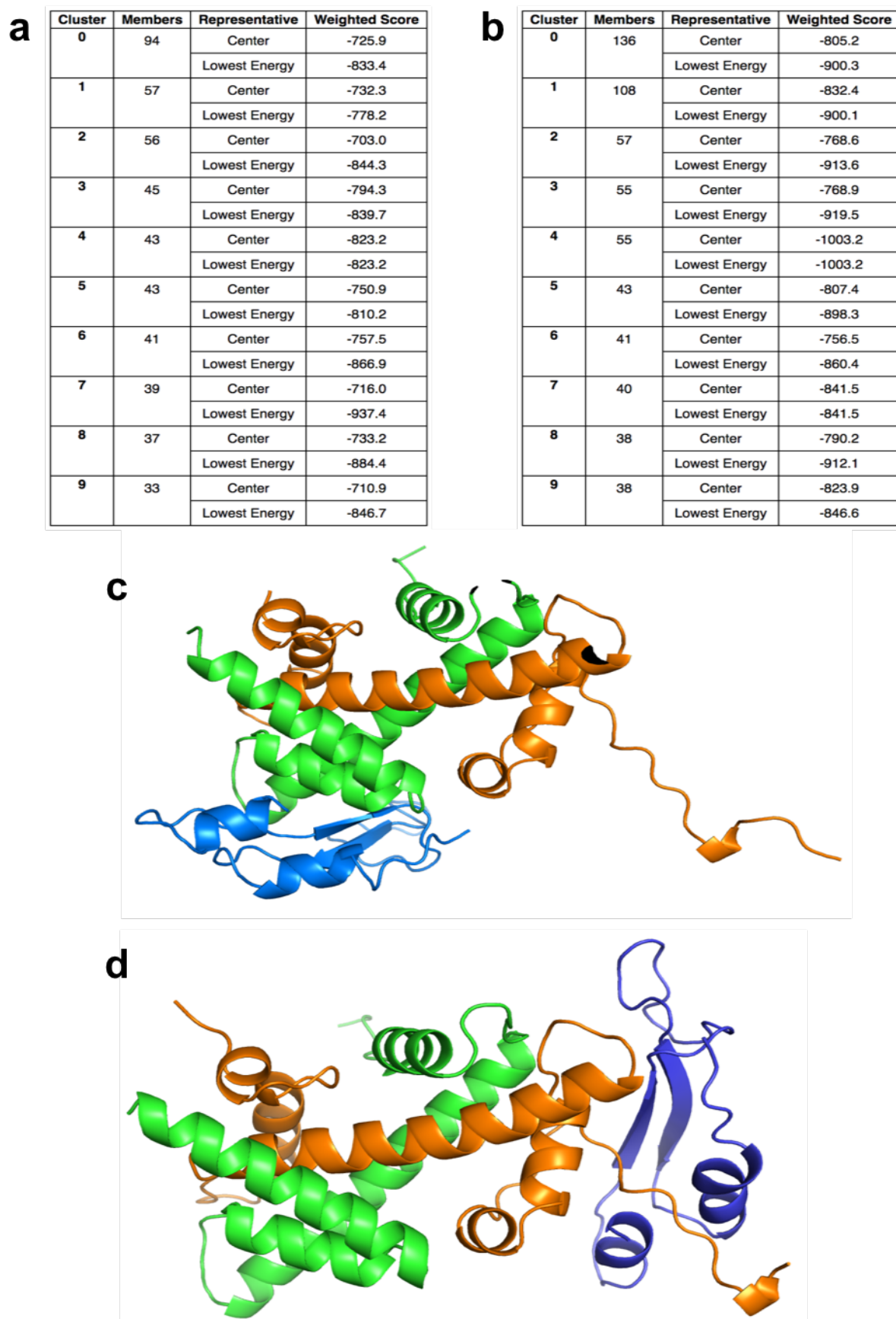


Figure 44: Structural Modelling: Comparison of **(a)** H2A/H2B/Ku80 (40-101) and **(b)** H2A.Z/H2B/Ku80 (40-101) weighted scores for the best fit modals as predicted by ClusPro[®] server. **(c)** Sample of predicted docking of H2A (orange) / H2B (green) / Ku80_40-101 (blue). **(d)** Sample of predicted docking of H2A.Z (orange) / H2B (green) / Ku80_40-101 (blue).

5. CONCLUSION AND PERSPECTIVES

The Ku heterodimer, consisting of Ku70 and Ku80 subunits (also referred as Ku86 in higher eukaryotes), is a highly abundant protein found primarily in the nuclear region of cells (Koike *et al.*, 2001). On an average, about half a million of Ku molecules are considered to be present in the nucleus of a single human cell (Mimori *et al.*, 1986; Walker *et al.*, 2001). Presence in such a huge amount certainly advocates its important roles in several key cellular processes. Both subunits form a toroidal structure that wraps itself around the DNA to perform the dedicated functions. Moreover, in the research literature, Ku heterodimer is now being proposed not merely as a DNA binding protein rather an entity with a wide range of functions in its account. These functions include its pivotal roles in apoptosis (Jing-Jing *et al.*, 2013), V(D)J recombination (Manis *et al.*, 1998), telomere maintenance (Wang *et al.*, 2009), DNA replication (Mukherjee *et al.*, 2016), transcriptional activation (Woodard *et al.*, 2001) and most importantly in genomic stability through serving as a molecular scaffold to recruit the primary members of the NHEJ repair pathway at the sites of double-strand DNA breaks (Grundy *et al.*, 2014). Based on structural and functional evidence, both Ku70 and Ku80 have been designated to possess an N-terminal von Willebrand factor type A (vWA) domain, a central DNA binding core domain, and a nuclear localization sequence (NLS) domain. C-terminal domain of Ku70 comprises an SAP (SAF-A/B, Acinus and PIAS) domain while C-terminal domain of Ku80 contains a sequence responsible for binding to DNA-PK_{CS}. All of these structurally distinct domains have been designated to perform diversified yet particular functions, contributing to the overall role of Ku heterodimer in different cellular processes.

Chromatin is considered a highly dynamic structure that keeps changing its organizational layout to accomplish its diversified roles in the cell. The process of chromatin remodeling involves the ATP-dependent changes in the overall structure of chromatin that result in altered access of different cellular proteins to the chromatin. These changes not only involve the covalent modification of histone machinery but also include the physical rearrangement of the nucleosome structure. Histone replacement is the most significant clause of chromatin remodeling as it causes the physio-chemical

changes at nucleosomes level that is critical for the interaction of DNA and important cellular proteins thus affecting many important cellular mechanisms. It involves the physical exchange of conventional canonical histones with that of counterpart non-allelic, non-canonical histones (variants). Despite the fact that these variants possess a quite high resemblance to their canonical counterparts, they can cause considerable changes in chromatin dynamics and its functions.

Among histone remodeling, H2A.Z remodeling has got huge attention of the research community in the recent past by virtue of its critical role in chromatin dynamics and gene expression. Histone variant H2A.Z has been reported to be involved in a variety of important cellular processes including transcription (Hardy *et al.*, 2009), genomic stability (Papamichos-Chronakis *et al.*, 2011), chromosome segregation (Hou *et al.*, 2010), and DNA repair (Xu *et al.*, 2012), etc. Bearing 60% sequence similarity to that of other cellular H2As, H2A.Z is the most extensively studied histone variant among all its family members found primarily at promotor regions of active genes. It constitutes about 15% of all H2A variants and is distributed throughout the genome both in euchromatin and heterochromatin. With ~90% sequence conservation among diversified species, H2A.Z is considered as the most conserved histone variant during the course of evolution that justifies its important role in a variety of key cellular processes.

In the literature, a wealth of information has been provided regarding the mechanism of H2A.Z remodeling however with the advent of new emerging techniques to comprehend such dynamic phenomenon, new members are being unveiled playing a significant role in these mechanisms. In yeast, the ATP-dependent chromatin-remodeling complex SWR1 is responsible for the exchange of canonical H2A/H2B dimer with histone variant H2A.Z/H2B dimer mediated by histone chaperon Chz1. In higher eukaryotes, two orthologs of SWR1 complex, the p400/TIP60, and Snf2 related CBP activator protein (SRCAP) complexes are considered to be responsible for this remodeling. Here we set out to investigate the interaction of Ku heterodimer and histone variant H2A.Z and the probable role of this interaction in H2A.Z deposition in chromatin.

Using double immunoaffinity procedure, we first identified Ku heterodimer in the H2A.Z predeposition complex from HeLa cells. The subsequent mass spectrometric analysis also endorsed our initial findings. After that, we proceeded towards the exploration of specificity of interaction between Ku heterodimer and H2A.Z. For that, we co-expressed Ku heterodimer with H2A/H2B, H2A.X/H2B and H2A.Z/H2B employing *E-coli* expression system. On account of this *in-vitro* analysis, we identified that Ku heterodimer interacts specifically with H2A.Z/H2B and not with counterpart H2A/H2B and H2A.X/H2B. This specificity suggested us the probable prospective of this interaction in the regulation of H2A.Z in chromatin.

Following the revelation of exclusive interaction of Ku heterodimer with histone dimer H2A.Z/H2B, we decided to further characterize the Ku70/Ku80 and H2A.Z/H2B interaction by mapping the Ku heterodimer domain involved in histone recognition and binding. By virtue of *in-vitro* experiments, we identified the von Willebrand factor type A (vWA) domain (also known as α/β domain) of Ku80 to be primarily responsible for this interaction with histone dimer H2A.Z/H2B. In the literature, several studies have implicated the vWA domain of Ku80 to be responsible for a wide range of protein-protein interaction carrying out significant roles in DSB repair and telomere regulation. Upon further exploration of this interaction, we identified that a 7kDa subdivision of vWA domain of Ku80, Ku80 (aa 40-101) motif, is the minimal domain sufficient for the interaction with the histone H2A.Z/H2B dimer. To re-ensure the specificity of interaction of this domain with H2A.Z/H2B, we reassessed the interaction of the identified domain with H2A.Z/H2B and counterpart H2A/H2B and H2A.X/H2B and found the consistency of specificity of interaction with H2A.Z/H2B alone.

Following the recognition of minimal domain of Ku heterodimer responsible for the interaction with H2A.Z/H2B, we aimed to identify the H2A.Z domain critical for this interaction. Towards this goal we employed chimeric proteins constituting H2A and H2A.Z portions and end up concluding that the central region of H2A.Z comprising highly conserved $\alpha 2$ helix is responsible for the specific interaction of histone dimer H2A.Z/H2B with Ku80 (aa 40-101) motif. Noteworthy, in the context of all *in vitro* experiments conducted, the amount of H2B was always found substantially higher as compared to H2A.Z suggesting that Ku80 (aa 40-101) motif might be binding to the

histone H2A.Z/H2B dimer favorably through H2B partner. Employing the *in vitro* exploration through double immunoaffinity procedure and size-exclusion chromatography, we observed a strong interaction of H2B with Ku80 (aa 40-101) motif. This advocated the pivotal role of H2B in the interaction of Ku heterodimer and histone H2A.Z/H2B dimer, though, the interaction with H2B is feasible only in the case of H2A.Z/H2B dimer and not in the case of H2A/H2B or H2A.X/H2B counterparts.

After the exploration of the interaction of Ku heterodimer and histone H2A.Z/H2B dimer through *in vivo* and *in vitro* approaches, we proceeded to evaluate the functional implications of this interaction. Employing genome-wide chromatin immunoprecipitation followed by sequencing (ChIP-seq), we identified that Ku heterodimer specifically regulates the H2A.Z deposition in chromatin, specifically at promoter regions. We found about 50% increased deposition of H2A.Z around promoter regions (TSS > 2 kilobases) with no significant variance in the deposition of H2A.Z at enhancers and insulators. This validated the critical involvement of Ku80 in genome-wide localization and chromatin enrichment of H2A.Z at promoters. Following that, we assessed the significance of Ku80 (aa 40-101) motif in H2A.Z deposition employing the same ChIP-Seq approach. Upon analysis, we observed a two-fold increase in H2A.Z deposition in case of WT-Ku80 MEFs as compared to Δ N-Ku80 MEF cells. This finding strongly advocated the critical role of the vWA domain or more precisely Ku80 (aa 40-101) motif in the deposition of H2A.Z/H2B *in vivo*.

Taken together, all these results point to a critical role of Ku heterodimer in the deposition of H2A.Z in chromatin. We suggested that von Willebrand factor type A (vWA) domain of Ku80 or more precisely Ku80 (aa 40-101) motif is substantially responsible for this interaction. We also observed that the Ku heterodimer interact exclusively with the histone H2A.Z/H2B dimer and this process is primarily mediated by H2B. This finding raised the question of the dynamic role of H2B in the specific deposition of H2A.Z. In the presumptive explanation of it, we hypothesize that Ku mediated deposition of H2A.Z/H2B may either involve the interaction of Ku with both H2A.Z and H2B members or alternatively the interaction is facilitated directly by H2B alone but only in the presence of H2A.Z as its dimer partner and not in case of H2A or H2A.X. This complexity of interaction strongly advocates the need for a

crystallographic study to understand the dynamicity of the process in a more holistic manner.

Since in literature annals, both the Ku heterodimer and histone variant H2A.Z have been implicated for a dedicated role in DNA repair process at double-strand breaks, our findings strongly propose to review the interaction of Ku heterodimer and histone H2A.Z/H2B in the context of DNA repair mechanisms particularly in non-homologous end joining (NHEJ) pathway. In that direction, our outcomes would turn out to be a worthy contribution in understanding the critical role of H2A.Z remodeling in DNA repair mechanism and will open many new horizons for future research. By virtue of X-ray crystallography, the structural characterization of the interaction would greatly enhance our understanding of Ku/H2A.Z/H2B complex formation. In this direction, with the collaboration of our partner research laboratories, we are already progressing positively towards solving the crystal structure of this complex.

The revelation of genome-wide localization of Ku heterodimer through ChIP-sequencing approach would also help us understand the association of Ku proteins with the histone variant H2A.Z in more details. The ChIP-Sequencing of Ku in H2A.Z knock-out and wild type cells would also direct us to understand the importance of this interaction more comprehensively. Likewise, employing H2A.Z knockout cells, the assessment of recruitment of Ku heterodimer at the site of DSB would suggest us the probable critical role of this unique interaction in DNA repair system. In short, this thesis work has set a substantial basis for a variety of future researches that would help the scientific community to better understand many dynamic cellular processes.

6. REFERENCES

- Ahmad, K., & Henikoff, S. (2002). The histone variant H3. 3 marks active chromatin by replication-independent nucleosome assembly. *Molecular Cell*, 9(6), 1191-1200.
- Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K., & Walter, P. (2002). DNA-binding motifs in gene regulatory proteins. In *Molecular Biology of the Cell*. 4th edition. Garland Science.
- Alén, C., Kent, N. A., Jones, H. S., O'Sullivan, J., Aranda, A., & Proudfoot, N. J. (2002). A role for chromatin remodeling in transcriptional termination by RNA polymerase II. *Molecular Cell*, 10(6), 1441-1452.
- Allfrey, V. G., Faulkner, R., & Mirsky, A. E. (1964). Acetylation and methylation of histones and their possible role in the regulation of RNA synthesis. *Proceedings of the National Academy of Sciences*, 51(5), 786-794.
- Allis, C. D., Richman, R., Gorovsky, M. A., Ziegler, Y. S., Touchstone, B., Bradley, W. A., & Cook, R. G. (1986). hv1 is an evolutionarily conserved H2A variant that is preferentially associated with active genes. *Journal of Biological Chemistry*, 261(4), 1941-1948.
- Aparicio, O. M., Billington, B. L., & Gottschling, D. E. (1991). Modifiers of position effect are shared between telomeric and silent mating-type loci in *S. cerevisiae*. *Cell*, 66(6), 1279-1287.
- Aravind, L., & Koonin, E. V. (2000). SAP—a putative DNA-binding motif involved in chromosomal organization. *Trends in Biochemical Sciences*, 25(3), 112-114.
- Aravind, L., & Koonin, E. V. (2001). Prokaryotic homologs of the eukaryotic DNA-end-binding protein Ku, novel domains in the Ku protein and prediction of a prokaryotic double-strand break repair system. *Genome Research*, 11(8), 1365-1374.
- Arnaudo, A. M., & Garcia, B. A. (2013). Proteomic characterization of novel histone post-translational modifications. *Epigenetics & Chromatin*, 6(1), 24.
- Bao, Y., Konesky, K., Park, Y.J., Rosu, S., Dyer, P.N., Rangasamy, D., Tremethick, D.J., Laybourn, P.J. and Luger, K. (2004). Nucleosomes containing the histone variant H2A. Bbd organize only 118 base pairs of DNA. *The EMBO Journal*, 23(16), 3314-3324.

- Beard, C., Li, E., & Jaenisch, R. (1995). Loss of methylation activates Xist in somatic but not in embryonic cells. *Genes & Development*, 9(19), 2325-2334.
- Bednar, J., Horowitz, R. A., Grigoryev, S. A., Carruthers, L. M., Hansen, J. C., Koster, A. J., & Woodcock, C. L. (1998). Nucleosomes, linker DNA, and linker histone form a unique structural motif that directs the higher-order folding and compaction of chromatin. *Proceedings of the National Academy of Sciences*, 95(24), 14173-14178.
- Behjati, S., Tarpey, P.S., Presneau, N., Scheipl, S., Pillay, N., Van Loo, P., Wedge, D.C., Cooke, S.L., Gundem, G., Davies, H. and Nik-Zainal, S. (2013). Distinct H3F3A and H3F3B driver mutations define chondroblastoma and giant cell tumor of bone. *Nature genetics*, 45(12), 1479.
- Berger, S. L., Kouzarides, T., Shiekhata, R., & Shilatifard, A. (2009). An operational definition of epigenetics. *Genes & development*, 23(7), 781-783.
- Bertinato, J., Schild-Poulter, C., & Hache, R. J. (2001). Nuclear localization of Ku antigen is promoted independently by basic motifs in the Ku70 and Ku80 subunits. *Journal of cell science*, 114(1), 89-99.
- Bertuch, A. A., & Lundblad, V. (2003). The Ku heterodimer performs separable activities at double-strand breaks and chromosome termini. *Molecular and cellular biology*, 23(22), 8202-8215.
- Bianconi, E., Piovesan, A., Facchin, F., Beraudi, A., Casadei, R., Frabetti, F., Vitale, L., Pelleri, M.C., Tassani, S., Piva, F. and Perez-Amodio, S. (2013). An estimation of the number of cells in the human body. *Annals of human biology*, 40(6), 463-471.
- Binda, O., Sevilla, A., LeRoy, G., Lemischka, I. R., Garcia, B. A., & Richard, S. (2013). SETD6 monomethylates H2AZ on lysine 7 and is required for the maintenance of embryonic stem cell self-renewal. *Epigenetics*, 8(2), 177-183.
- Bird, A. (2007). Perceptions of epigenetics. *Nature*, 447(7143), 396-398.
- Blier, P. R., Griffith, A. J., Craft, J., & Hardin, J. A. (1993). Binding of Ku protein to DNA. Measurement of affinity for ends and demonstration of binding to nicks. *Journal of Biological Chemistry*, 268(10), 7594-7601.
- Bönisch, C., Schneider, K., Pünzeler, S., Wiedemann, S.M., Bielmeier, C., Bocola, M., Eberl, H.C., Kuegel, W., Neumann, J., Kremmer, E. and Leonhardt, H. (2012). H2A. Z.

- 2.2 is an alternatively spliced histone H2A. Z variant that causes severe nucleosome destabilization. *Nucleic acids research*, 40(13), 5951-5964.
- Boulton, S. J., & Jackson, S. P. (1996). Identification of a *Saccharomyces cerevisiae* Ku80 homologue: roles in DNA double strand break rejoining and in telomeric maintenance. *Nucleic acids research*, 24(23), 4639-4648.
- Boulton, S. J., & Jackson, S. P. (1998). Components of the Ku-dependent non-homologous end-joining pathway are involved in telomeric length maintenance and telomeric silencing. *The EMBO journal*, 17(6), 1819-1828.
- Bowater, R., & Doherty, A. J. (2006). Making ends meet: repairing breaks in bacterial DNA by non-homologous end-joining. *PLoS genetics*, 2(2), e8.
- Britton, S., Coates, J., & Jackson, S. P. (2013). A new method for high-resolution imaging of Ku foci to decipher mechanisms of DNA double-strand break repair. *J Cell Biol*, 202(3), 579-595.
- Buczkowicz, P., Hoeman, C., Rakopoulos, P., Pajovic, S., Letourneau, L., Dzamba, M., Morrison, A., Lewis, P., Bouffet, E., Bartels, U. and Zuccaro, J. (2014). Genomic analysis of diffuse intrinsic pontine gliomas identifies three molecular subgroups and recurrent activating ACVR1 mutations. *Nature genetics*, 46(5), 451-456.
- Bungard, D., Fuerth, B.J., Zeng, P.Y., Faubert, B., Maas, N.L., Viollet, B., Carling, D., Thompson, C.B., Jones, R.G. and Berger, S.L. (2010). Signaling kinase AMPK activates stress-promoted transcription via histone H2B phosphorylation. *Science*, 329(5996), 1201-1205.
- Burgess, R. J., & Zhang, Z. (2013). Histone chaperones in nucleosome assembly and human disease. *Nature structural & molecular biology*, 20(1), 14.
- Cary, R. B., Peterson, S. R., Wang, J., Bear, D. G., Bradbury, E. M., & Chen, D. J. (1997). DNA looping by Ku and the DNA-dependent protein kinase. *Proceedings of the National Academy of Sciences*, 94(9), 4267-4272.
- Chadwick, B. P., & Willard, H. F. (2001). A novel chromatin protein, distantly related to histone H2A, is largely excluded from the inactive X chromosome. *The Journal of cell biology*, 152(2), 375-384.
- Chan, D. W., Chen, B. P. C., Prithivirajasingh, S., Kurimasa, A., Story, M. D., Qin, J., & Chen, D. J. (2002). Autophosphorylation of the DNA-dependent protein kinase catalytic

- subunit is required for rejoining of DNA double-strand breaks. *Genes & development*, 16(18), 2333-2338.
- Chandler, V. L. (2007). Paramutation: from maize to mice. *Cell*, 128(4), 641-645.
- Chen, H. T., Bhandoola, A., Difilippantonio, M. J., Zhu, J., Brown, M. J., Tai, X., Rogakou, E. P., Brotz, T. M., Bonner, W. M., Ried, T. and Nussenzweig, A. (2000). Response to RAG-mediated V (D) J cleavage by NBS1 and γ -H2AX. *Science*, 290(5498), 1962-1964.
- Chen, H., Ruiz, P. D., Novikov, L., Casill, A. D., Park, J. W., & Gamble, M. J. (2014). MacroH2A1. 1 and PARP-1 cooperate to regulate transcription by promoting CBP-mediated H2B acetylation. *Nature structural & molecular biology*, 21(11), 981.
- Chueh, A. C., Wong, L. H., Wong, N., & Choo, K. A. (2004). Variable and hierarchical size distribution of L1-retroelement-enriched CENP-A clusters within a functional human neocentromere. *Human molecular genetics*, 14(1), 85-93.
- Chung, U. I., Igarashi, T., Nishishita, T., Iwanari, H., Iwamatsu, A., Suwa, A., Mimori, T., Hata, K., Ebisu, S., Ogata, E. and Fujita, T. (1996). The interaction between Ku antigen and REF1 protein mediates negative gene regulation by extracellular calcium. *Journal of Biological Chemistry*, 271(15), 8593-8598.
- Ciccia, A., & Elledge, S. J. (2010). The DNA damage response: making it safe to play with knives. *Molecular cell*, 40(2), 179-204.
- Clapier, C. R., & Cairns, B. R. (2009). The biology of chromatin remodeling complexes. *Annual review of biochemistry*, 78, 273-304.
- Cohen, H.Y., Lavu, S., Bitterman, K.J., Hekking, B., Imahiyerobo, T.A., Miller, C., Frye, R., Ploegh, H., Kessler, B.M. and Sinclair, D.A. (2004). Acetylation of the C terminus of Ku70 by CBP and PCAF controls Bax-mediated apoptosis. *Molecular cell*, 13(5), 627-638.
- Cooley, H. M., Melny, B. J., Gleeson, R., Greco, T., & Kay, T. W. (1999). Clinical and serological associations of anti-Ku antibody. *The Journal of rheumatology*, 26(3), 563-567.
- Costantini, S., Woodbine, L., Andreoli, L., Jeggo, P. A., & Vindigni, A. (2007). Interaction of the Ku heterodimer with the DNA ligase IV/Xrcc4 complex and its regulation by DNA-PK. *DNA repair*, 6(6), 712-722.

- Costanzi, C., & Pehrson, J. R. (1998). Histone macroH2A1 is concentrated in the inactive X chromosome of female mammals. *Nature*, 393(6685), 599.
- Costanzi, C., & Pehrson, J. R. (2001). MACROH2A2, a new member of the MACROH2A core histone family. *Journal of Biological Chemistry*, 276(24), 21776-21784.
- Courilleau, C., Chailleux, C., Jauneau, A., Grimal, F., Briois, S., Boutet-Robinet, E., Boudsocq, F., Trouche, D. and Canitrot, Y. (2012). The chromatin remodeler p400 ATPase facilitates Rad51-mediated repair of DNA double-strand breaks. *Journal of Cell Biology*, 199(7), 1067-1081.
- Cuozzo, C., Porcellini, A., Angrisano, T., Morano, A., Lee, B., Di Pardo, A., Messina, S., Iuliano, R., Fusco, A., Santillo, M.R. and Muller, M.T. (2007). DNA damage, homology-directed repair, and DNA methylation. *PLoS genetics*, 3(7), 1144-1162
- Cutter, A. R., & Hayes, J. J. (2015). A brief review of nucleosome structure. *FEBS letters*, 589(20), 2914-2922.
- D'addario, C., Di Francesco, A., Pucci, M., Finazzi Agrò, A., & Maccarrone, M. (2013). Epigenetic mechanisms and endocannabinoid signalling. *The FEBS journal*, 280(9), 1905-1917.
- Dame, R. T. (2005). The role of nucleoid-associated proteins in the organization and compaction of bacterial chromatin. *Molecular microbiology*, 56(4), 858-870.
- Dantzer, F., & Santoro, R. (2013). The expanding role of PARPs in the establishment and maintenance of heterochromatin. *The FEBS journal*, 280(15), 3508-3518.
- Daura-Oller, E., Cabre, M., Montero, M. A., Paternain, J. L., & Romeu, A. (2009). Specific gene hypomethylation and cancer: new insights into coding region feature trends. *Bioinformatics*, 3(8), 340.
- Dawson, M. A., & Kouzarides, T. (2012). Cancer epigenetics: from mechanism to therapy. *Cell*, 150(1), 12-27.
- De Rop, V., Padeganeh, A., & Maddox, P. S. (2012). CENP-A: the key player behind centromere identity, propagation, and kinetochore assembly. *Chromosoma*, 121(6), 527-538.
- de Vries, E., van Driel, W., Bergsma, W. G., Arnberg, A. C., & van der Vliet, P. C. (1989). HeLa nuclear protein recognizing DNA termini and translocating on DNA

- forming a regular DNA-multimeric protein complex. *Journal of molecular biology*, 208(1), 65-78.
- Deaton, A. M., & Bird, A. (2011). CpG islands and the regulation of transcription. *Genes & development*, 25(10), 1010-1022.
- DeLange, R. J., Fambrough, D. M., Smith, E. L., & Bonner, J. (1968). Calf and Pea Histone IV I. Amino acid compositions and the identical COOH-terminal 19-residue sequence. *Journal of Biological Chemistry*, 243(22), 5906-5913.
- Deriano, L., & Roth, D. B. (2013). Modernizing the nonhomologous end-joining repertoire: alternative and classical NHEJ share the stage. *Annual review of genetics*, 47, 433-455.
- Dion, M. F., Kaplan, T., Kim, M., Buratowski, S., Friedman, N., & Rando, O. J. (2007). Dynamics of replication-independent histone turnover in budding yeast. *Science*, 315(5817), 1405-1408.
- Doherty, A. J., & Jackson, S. P. (2001). DNA repair: how Ku makes ends meet. *Current Biology*, 11(22), R920-R924.
- Doherty, A. J., Jackson, S. P., & Weller, G. R. (2001). Identification of bacterial homologues of the Ku DNA repair proteins. *FEBS letters*, 500(3), 186-188.
- Douglas, P., Sapkota, G.P., Morrice, N., Yaping, Y.U., Goodarzi, A.A., Merkle, D., Katheryn, M.E.E.K., Alessi, D.R. and Lees-Miller, S.P. (2002). Identification of in vitro and in vivo phosphorylation sites in the catalytic subunit of the DNA-dependent protein kinase. *Biochemical Journal*, 368(1), 243-251.
- Downs, J. A., & Jackson, S. P. (2004). A means to a DNA end: the many roles of Ku. *Nature Reviews Molecular Cell Biology*, 5(5), 367-378.
- Downs, J.A., Allard, S., Jobin-Robitaille, O., Javaheri, A., Auger, A., Bouchard, N., Kron, S.J., Jackson, S.P. and Côté, J. (2004). Binding of chromatin-modifying activities to phosphorylated histone H2A at DNA damage sites. *Molecular cell*, 16(6), 979-990.
- Doyen, C. M., Montel, F., Gautier, T., Menoni, H., Claudet, C., Delacour-Larose, M., Angelov, D., Hamiche, A., Bednar, J., Faivre-Moskalenko, C. and Bouvet, P (2006). Dissection of the unusual structural and functional properties of the variant H2A. Bbd nucleosome. *The EMBO journal*, 25(18), 4234-4244.

- Draizen, E. J., Shaytan, A. K., Mariño-Ramírez, L., Talbert, P. B., Landsman, D., & Panchenko, A. R. (2016). HistoneDB 2.0: a histone database with variants-an integrated resource to explore histones and their variants. Database, 2016.
- Dubin, M., Fuchs, J., Gräf, R., Schubert, I., & Nellen, W. (2010). Dynamics of a novel centromeric histone variant CenH3 reveals the evolutionary ancestral timing of centromere biogenesis. *Nucleic acids research*, 38(21), 7526-7537.
- Durrin, L. K., Mann, R. K., Kayne, P. S., & Grunstein, M. (1991). Yeast histone H4 N-terminal sequence is required for promoter activation in vivo. *Cell*, 65(6), 1023-1031.
- Earnshaw, W. C., & Rattner, J. B. (1989). A map of the centromere (primary constriction) in vertebrate chromosomes at metaphase. *Progress in clinical and biological research*, 318, 33-42.
- Eirín-López, J. M., Ishibashi, T., & Ausió, J. (2008). H2A. Bbd: a quickly evolving hypervariable mammalian histone that destabilizes nucleosomes in an acetylation-independent way. *The FASEB Journal*, 22(1), 316-326.
- Errami, A., Smider, V., Rathmell, W. K., He, D. M., Hendrickson, E. A., Zdzienicka, M. Z., & Chu, G. (1996). Ku86 defines the genetic defect and restores X-ray resistance and V (D) J recombination to complementation group 5 hamster cell mutants. *Molecular and cellular biology*, 16(4), 1519-1526.
- Faghihi, M.A., Modarresi, F., Khalil, A.M., Wood, D.E., Sahagan, B.G., Morgan, T.E., Finch, C.E., Laurent III, G.S., Kenny, P.J., & Wahlestedt, C. (2008). Expression of a noncoding RNA is elevated in Alzheimer's disease and drives rapid feed-forward regulation of β -secretase. *Nature medicine*, 14(7), 723-730.
- Falzon, M., Fewell, J. W., & Kuff, E. L. (1993). EBP-80, a transcription factor closely resembling the human autoantigen Ku, recognizes single-to double-strand transitions in DNA. *Journal of Biological Chemistry*, 268(14), 10546-10552.
- Fan, J. Y., Rangasamy, D., Luger, K., & Tremethick, D. J. (2004). H2A. Z alters the nucleosome surface to promote HP1 α -mediated chromatin fiber folding. *Molecular cell*, 16(4), 655-661.
- Farris, S. D., Rubio, E. D., Moon, J. J., Gombert, W. M., Nelson, B. H., & Krumm, A. (2005). Transcription-induced chromatin remodeling at the c-myc gene involves the

- local exchange of histone H2A. *Z. Journal of Biological Chemistry*, 280(26), 25298-25303.
- Fell, V. L., & Schild-Poulter, C. (2012). Ku regulates signaling to DNA damage response pathways through the Ku70 von Willebrand A domain. *Molecular and cellular biology*, 32(1), 76-87.
- Fell, V. L., & Schild-Poulter, C. (2015). The Ku heterodimer: function in DNA repair and beyond. *Mutation Research/Reviews in Mutation Research*, 763, 15-29.
- Feng, L., & Chen, J. (2012). The E3 ligase RNF8 regulates KU80 removal and NHEJ repair. *Nature structural & molecular biology*, 19(2), 201.
- Fewell, J. W., & Kuff, E. L. (1996). Intracellular redistribution of Ku immunoreactivity in response to cell-cell contact and growth modulating components in the medium. *Journal of Cell Science*, 109(7), 1937-1946.
- Fisher, T. S., Taggart, A. K., & Zakian, V. A. (2004). Cell cycle-dependent regulation of yeast telomerase by Ku. *Nature structural & molecular biology*, 11(12), 1198.
- Flemming, W. (1879). *Beitrage zur Kenntniss der Zelle und ihrer Lebenserscheinungen. Archiv für mikroskopische Anatomie*, 16(1), 302-436.
- Folco, H. D., Pidoux, A. L., Urano, T., & Allshire, R. C. (2008). Heterochromatin and RNAi are required to establish CENP-A chromatin at centromeres. *Science*, 319(5859), 94-97.
- Francoeur, A.M., Peebles, C. L., Gompfer, P. T., and Tan, E. M. (1986). Identification of Ki (Ku, p70/p80) autoantigens and analysis of anti-Ki autoantibody reactivity. *Journal of Immunology*, 136, 1648–1653.
- Franklin, S. G., & Zweidler, A. (1977). Non-allelic variants of histones 2a, 2b and 3 in mammals. *Nature*, 266(5599), 273.
- Fried, V. A. (1992). MacroH2A, a core histone containing a large nonhistone region. *Science*, 257(5075), 1398-1400.
- Frit, P., Barboule, N., Yuan, Y., Gomez, D., & Calsou, P. (2014). Alternative end-joining pathway (s): bricolage at DNA breaks. *DNA repair*, 17, 81-97.
- Fuchs, M., Gerber, J., Drapkin, R., Sif, S., Ikura, T., Ogryzko, V., Lane, W.S., Nakatani, Y. and Livingston, D.M. (2001). The p400 complex is an essential E1A transformation target. *Cell*, 106(3), 297-307.

- Gardiner-Garden, M., & Frommer, M. (1987). CpG islands in vertebrate genomes. *Journal of molecular biology*, 196(2), 261-282.
- Gell, D., & Jackson, S. P. (1999). Mapping of protein-protein interactions within the DNA-dependent protein kinase complex. *Nucleic acids research*, 27(17), 3494-3502.
- Genersch, E., Eckerskorn, C., Lottspeich, F., Herzog, C., Kühn, K., & Pöschl, E. (1995). Purification of the sequence-specific transcription factor CTCBF, involved in the control of human collagen IV genes: subunits with homology to Ku antigen. *The EMBO Journal*, 14(4), 791-800.
- Gershey, E. L., Vidali, G. and Allfrey, V. G. (1968) Chemical studies of histone acetylation. The occurrence of epsilon-N-acetyllysine in the f2a1 histone. *Journal of Biological Chemistry*, 243, 5018–5022.
- Gévry, N., Chan, H. M., Laflamme, L., Livingston, D. M., & Gaudreau, L. (2007). p21 transcription is regulated by differential localization of histone H2A.Z. *Genes & development*, 21(15), 1869-1881.
- Goldknopf, I. L., Taylor, C. W., Baum, R. M., Yeoman, L. C., Olson, M. O., Prestayko, A. W., & Busch, H. (1975). Isolation and characterization of protein A24, a "histone-like" non-histone chromosomal protein. *Journal of Biological Chemistry*, 250(18), 7182-7187.
- Goldman, M. A. (1988). The chromatin domain as a unit of gene regulation. *Bioessays*, 9(2-3), 50-55.
- Goll, M. G., & Bestor, T. H. (2002). Histone modification and replacement in chromatin activation. *Genes & development*, 16(14), 1739-1742.
- Gravel, S., Larrivée, M., Labrecque, P., & Wellinger, R. J. (1998). Yeast Ku as a regulator of chromosomal DNA end structure. *Science*, 280(5364), 741-744.
- Gregory, P. D., Wagner, K., & Hörz, W. (2001). Histone acetylation and chromatin remodeling. *Experimental cell research*, 265(2), 195-202.
- Grewal, S. I., & Moazed, D. (2003). Heterochromatin and epigenetic control of gene expression. *Science*, 301(5634), 798-802.
- Grundy, G. J., Moulding, H. A., Caldecott, K. W., & Rulten, S. L. (2014). One ring to bring them all - the role of Ku in mammalian non-homologous end joining. *DNA repair*, 17, 30-38.

- Grundy, G.J., Rulten, S.L., Arribas-Bosacoma, R., Davidson, K., Kozik, Z., Oliver, A.W., Pearl, L.H. and Caldecott, K.W. (2016). The Ku-binding motif is a conserved module for recruitment and stimulation of non-homologous end-joining proteins. *Nature communications*, 7, 11242.
- Grundy, G.J., Rulten, S.L., Zeng, Z., Arribas-Bosacoma, R., Iles, N., Manley, K., Oliver, A. and Caldecott, K.W. (2013). APLF promotes the assembly and activity of non-homologous end joining protein complexes. *The EMBO journal*, 32(1), 112-125.
- Gu, Y., Jin, S., Gao, Y., Weaver, D. T., & Alt, F. W. (1997). Ku70-deficient embryonic stem cells have increased ionizing radiosensitivity, defective DNA end-binding activity, and inability to support V (D) J recombination. *Proceedings of the National Academy of Sciences*, 94(15), 8076-8081.
- Gu, Y., Seidl, K.J., Rathbun, G.A., Zhu, C., Manis, J.P., van der Stoep, N., Davidson, L., Cheng, H.L., Sekiguchi, J.M., Frank, K. and Stanhope-Baker, P. (1997). Growth retardation and leaky SCID phenotype of Ku70-deficient mice. *Immunity*, 7(5), 653-665.
- Guillemette, B., Bataille, A. R., Gévry, N., Adam, M., Blanchette, M., Robert, F., & Gaudreau, L. (2005). Variant histone H2A. Z is globally localized to the promoters of inactive yeast genes and regulates nucleosome positioning. *PLoS biology*, 3(12), e384.
- Guilleret, I., Yan, P., Grange, F., Braunschweig, R., Bosman, F. T., & Benhattar, J. (2002). Hypermethylation of the human telomerase catalytic subunit (hTERT) gene correlates with telomerase activity. *International journal of cancer*, 101(4), 335-341.
- Guttman, M., Russell, P., Ingolia, N. T., Weissman, J. S., & Lander, E. S. (2013). Ribosome profiling provides evidence that large noncoding RNAs do not encode proteins. *Cell*, 154(1), 240-251.
- Haapala, O. K., & Soyer, M. O. (1975). Organization of chromosome fibrils in *Euglena gracilis*. *Hereditas*, 80(2), 185-193.
- Hammel, M., Yu, Y., Mahaney, B.L., Cai, B., Ye, R., Phipps, B.M., Rambo, R.P., Hura, G.L., Pelikan, M., So, S. and Abolfath, R.M. (2010). Ku and DNA-dependent protein kinase dynamic conformations and assembly regulate DNA binding and the initial non-homologous end joining complex. *Journal of Biological Chemistry*, 285(2), 1414-1423.

- Hang, L.E., Lopez, C.R., Liu, X., Williams, J.M., Chung, I., Wei, L., Bertuch, A.A. and Zhao, X. (2014). Regulation of Ku-DNA association by Yku70 C-terminal tail and SUMO modification. *Journal of Biological Chemistry*, 289(15), 10308-10317.
- Hansen, J. C. (2012). Human mitotic chromosome structure: what happened to the 30-nm fibre?. *The EMBO journal*, 31(7), 1621-1623.
- Hardy, S., Jacques, P.É., Gévry, N., Forest, A., Fortin, M.È., Laflamme, L., Gaudreau, L. and Robert, F. (2009). The euchromatic and heterochromatic landscapes are shaped by antagonizing effects of transcription on H2A. Z deposition. *PLoS genetics*, 5(10), e1000687.
- Harris, R., Esposito, D., Sankar, A., Maman, J. D., Hinks, J. A., Pearl, L. H., & Driscoll, P. C. (2004). The 3D solution structure of the C-terminal region of Ku86 (Ku86CTR). *Journal of molecular biology*, 335(2), 573-582.
- Hatch, C. L., & Bonner, W. M. (1990). The human histone H2A. Z gene. Sequence and regulation. *Journal of Biological Chemistry*, 265(25), 15211-15218.
- Hild, M., & Paro, R. (2003). Anti-silencing from the core: a histone H2A variant protects euchromatin. *Nature cell biology*, 5(4), 278.
- Hoff, C. M., Ghosh, A. K., Prabhakar, B. S., & Jacob, S. T. (1994). Enhancer 1 binding factor, a Ku-related protein, is a positive regulator of RNA polymerase I transcription initiation. *Proceedings of the National Academy of Sciences*, 91(2), 762-766.
- Holliday, R. (1990). DNA methylation and epigenetic inheritance. Philosophical Transactions of the Royal Society of London, *Biological Sciences*, 326(1235), 329-338.
- Hong, J., Feng, H., Wang, F., Ranjan, A., Chen, J., Jiang, J., Ghirlando, R., Xiao, T.S., Wu, C. and Bai, Y. (2014). The catalytic subunit of the SWR1 remodeler is a histone chaperone for the H2A. Z-H2B dimer. *Molecular cell*, 53(3), 498-505.
- Horvath, S. (2013). DNA methylation age of human tissues and cell types. *Genome biology*, 14(10), 3156.
- Hotchkiss, R. D. (1948). The quantitative separation of purines, pyrimidines, and nucleosides by paper chromatography. *Journal of Biological Chemistry*, 175(1), 315-332.

- Hou, H., Wang, Y., Kallgren, S. P., Thompson, J., Yates, J. R., & Jia, S. (2010). Histone variant H2A. Z regulates centromere silencing and chromosome segregation in fission yeast. *Journal of Biological Chemistry*, 285(3), 1909-1918.
- Hsu, H. L., Yannone, S. M., & Chen, D. J. (2002). Defining interactions between DNA-PK and ligase IV/XRCC4. *DNA repair*, 1(3), 225-235.
- Hu, S., Pluth, J. M., & Cucinotta, F. A. (2012). Putative binding modes of Ku70-SAP domain with double strand DNA: a molecular modeling study. *Journal of molecular modeling*, 18(5), 2163-2174.
- Huang, J., Nueda, A., Yoo, S., & Dynan, W. S. (1997). Heat shock transcription factor 1 binds selectively in vitro to Ku protein and the catalytic subunit of the DNA-dependent protein kinase. *Journal of Biological Chemistry*, 272(41), 26009-26016.
- Iizuka, M., & Smith, M. M. (2003). Functional consequences of histone modifications. *Current opinion in genetics & development*, 13(2), 154-160.
- Ilnytsky, Y., & Kovalchuk, O. (2011). Non-targeted radiation effects—an epigenetic connection. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis*, 714(1-2), 113-125.
- Ishibashi, T., Li, A., Eirín-López, J. M., Zhao, M., Missiaen, K., Abbott, D. W., Meistrich, M., Hendzel, M. J. and Ausio, J. (2009). H2A. Bbd: An X-chromosome-encoded histone involved in mammalian spermiogenesis. *Nucleic acids research*, 38(6), 1780-1789.
- Izzo, A., Kamieniarz, K., & Schneider, R. (2008). The histone H1 family: specific members, specific functions. *Biological chemistry*, 389(4), 333-343.
- Jackson, J. D., & Gorovsky, M. A. (2000). Histone H2A. Z has a conserved function that is distinct from that of the major H2A sequence variants. *Nucleic acids research*, 28(19), 3811-3816.
- Jackson, S. P., & Bartek, J. (2009). The DNA-damage response in human biology and disease. *Nature*, 461(7267), 1071.
- Jacoby, D. B., & Wensink, P. C. (1994). Yolk protein factor 1 is a Drosophila homolog of Ku, the DNA-binding subunit of a DNA-dependent protein kinase from humans. *Journal of Biological Chemistry*, 269(15), 11484-11491.

- Jenuwein, T., & Allis, C. D. (2001). Translating the histone code. *Science*, 293(5532), 1074-1080.
- Jha, S., Shibata, E., & Dutta, A. (2008). Human Rvb1/Tip49 is required for the histone acetyltransferase activity of Tip60/NuA4 and for the downregulation of phosphorylation on H2AX after DNA damage. *Molecular and cellular biology*, 28(8), 2690-2700.
- Jing-Jing, L., Qi-Hua, G., Min, L., Hua-Ping, Y., Li-Ming, C., & Cheng-Ping, H. (2013). Role of Ku70 and Bax in epigallocatechin-3-gallate-induced apoptosis of A549 cells in vivo. *Oncology letters*, 5(1), 101-106.
- Jones, J. M., Gellert, M., & Yang, W. (2001). A Ku bridge over broken DNA. *Structure*, 9(10), 881-884.
- Joshi, A. A., & Struhl, K. (2005). Eaf3 chromodomain interaction with methylated H3-K36 links histone deacetylation to Pol II elongation. *Molecular cell*, 20(6), 971-978.
- Kalocsay, M., Hiller, N. J., & Jentsch, S. (2009). Chromosome-wide Rad51 spreading and SUMO-H2A. Z-dependent chromosome fixation in response to a persistent DNA double-strand break. *Molecular cell*, 33(3), 335-343.
- Kamakaka, R. T., & Biggins, S. (2005). Histone variants: deviants. *Genes & development*, 19(3), 295-316.
- Karmakar, P., Snowden, C. M., Ramsden, D. A., & Bohr, V. A. (2002). Ku heterodimer binds to both ends of the Werner protein and functional interaction occurs at the Werner N-terminus. *Nucleic acids research*, 30(16), 3583-3591.
- Kennedy, R. D., & D'Andrea, A. D. (2006). DNA repair pathways in clinical practice: lessons from pediatric cancer susceptibility syndromes. *Journal of Clinical Oncology*, 24(23), 3799-3808.
- Kim, K.B., Kim, D.W., Park, J.W., Jeon, Y.J., Kim, D., Rhee, S., Chae, J.I. and Seo, S.B. (2014). Inhibition of Ku70 acetylation by INHAT subunit SET/TAF-I β regulates Ku70-mediated DNA damage response. *Cellular and molecular life sciences*, 71(14), 2731-2745.
- Kleinsmith, L. J., Allfrey, V. G., & Mirsky, A. E. (1966). Phosphoprotein metabolism in isolated lymphocyte nuclei. *Proceedings of the National Academy of Sciences of the United States of America*, 55(5), 1182.

- Knuth, M. W., Gunderson, S. I., Thompson, N. E., Strasheim, L. A., & Burgess, R. R. (1990). Purification and characterization of proximal sequence element-binding protein 1, a transcription activating protein related to Ku and TREF that binds the proximal sequence element of the human U1 promoter. *Journal of Biological Chemistry*, 265(29), 17911-17920.
- Koike, M., Shiomi, T., & Koike, A. (2001). Dimerization and nuclear localization of Ku proteins. *Journal of Biological Chemistry*, 276(14), 11167-11173.
- Kornberg, R. D. (1974). Chromatin structure: a repeating unit of histones and DNA. *Science*, 184(4139), 868-871.
- Kornberg, R. D., & Klug, A. (1981). The nucleosome. *Scientific American*, 244(2), 52-65.
- Kossel, A. (1884). Ueber einen peptonartigen Bestandtheil des Zellkerns. *Zeitschrift für physiologische Chemie*, 8(6), 511-515.
- Kovalchuk, O., & Baulch, J. E. (2008). Epigenetic changes and nontargeted radiation effects-is there a link?. *Environmental and molecular mutagenesis*, 49(1), 16-25.
- Kragelund, B. B., Weterings, E., Hartmann-Petersen, R., & Keijzers, G. (2016). The Ku70/80 ring in Non-Homologous End-Joining: Easy to slip on, hard to remove. *Frontiers in Bioscience*, 21, 514-527.
- Krebs, J. E., Fry, C. J., Samuels, M. L., & Peterson, C. L. (2000). Global role for chromatin remodeling enzymes in mitotic gene expression. *Cell*, 102(5), 587-598.
- Kuhn, A., Gottlieb, T. M., Jackson, S. P., & Grummt, I. (1995). DNA-dependent protein kinase: a potent inhibitor of transcription by RNA polymerase I. *Genes & development*, 9(2), 193-203.
- Kulis, M., Merkel, A., Heath, S., Queirós, A.C., Schuyler, R.P., Castellano, G., Beekman, R., Raineri, E., Esteve, A., Clot, G. and Verdaguer-Dot, N. (2015). Whole-genome fingerprint of the DNA methylome during human B cell differentiation. *Nature genetics*, 47(7), 746-756.
- Kusch, T., Florens, L., MacDonald, W.H., Swanson, S.K., Glaser, R.L., Yates, J.R., Abmayr, S.M., Washburn, M.P. and Workman, J.L. (2004). Acetylation by Tip60 is required for selective histone variant exchange at DNA lesions. *Science*, 306(5704), 2084-2087.

- Laroche, T., Martin, S. G., Gotta, M., Gorham, H. C., Pryde, F. E., Louis, E. J., & Gasser, S. M. (1998). Mutation of yeast Ku genes disrupts the subnuclear organization of telomeres. *Current biology*, 8(11), 653-657.
- Larsen, F., Gundersen, G., Lopez, R., & Prydz, H. (1992). CpG islands as gene markers in the human genome. *Genomics*, 13(4), 1095-1107.
- Latrick, C.M., Marek, M., Ouararhni, K., Papin, C., Stoll, I., Ignatyeva, M., Obri, A., Ennifar, E., Dimitrov, S., Romier, C. and Hamiche, A. (2016). Molecular basis and specificity of H2A. Z-H2B recognition and deposition by the histone chaperone YL1. *Nature structural & molecular biology*, 23(4), 309.
- Ledford, H. (2008). Disputed definitions. *Nature*, 455 (7216), 1023–1028.
- Lehninger, A. L., Nelson, D. L., Cox, M. M., & Cox, M. M. (2005). *Lehninger principles of biochemistry*. Macmillan.
- Li, G.C., Yang, S.H., Kim, D., Nussenzweig, A., Ouyang, H., Wei, J., Burgman, P. and Li, L. (1995). Suppression of heat-induced hsp70 expression by the 70-kDa subunit of the human Ku autoantigen. *Proceedings of the National Academy of Sciences*, 92(10), 4512-4516.
- Li, H., Marple, T., & Hasty, P. (2013). Ku80-deleted cells are defective at base excision repair. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis*, 745, 16-25.
- Lieber, M. R., & Karanjawala, Z. E. (2004). Ageing, repetitive genomes and DNA damage. *Nature Reviews Molecular Cell Biology*, 5(1), 69.
- Lin, C. J., Conti, M., & Ramalho-Santos, M. (2013). Histone variant H3. 3 maintains a decondensed chromatin state essential for mouse preimplantation development. *Development*, 140(17), 3624-3634.
- Long, M., Sun, X., Shi, W., Yanru, A., Leung, S.T., Ding, D., Cheema, M.S., MacPherson, N., Nelson, C.J., Ausio, J. and Yan, Y. (2019). A novel histone H4 variant H4G regulates rDNA transcription in breast cancer. *Nucleic Acids Research*, 1, 11.
- Longhese, M. P., Anbalagan, S., Martina, M., & Bonetti, D. (2012). The role of shelterin in maintaining telomere integrity. *Frontiers in Bioscience*, 17, 1715-28.
- Lovejoy, C. A., & Cortez, D. (2009). Common mechanisms of PIKK regulation. *DNA repair*, 8(9), 1004-1008.

- Loyola, A., & Almouzni, G. (2004). Histone chaperones, a supporting role in the limelight. *Biochimica et Biophysica Acta (BBA)-Gene Structure and Expression*, 1677(1-3), 3-11.
- Luger, K., Mäder, A. W., Richmond, R. K., Sargent, D. F., & Richmond, T. J. (1997). Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature*, 389(6648), 251.
- Ma, Y., Lu, H., Tippin, B., Goodman, M.F., Shimazaki, N., Koiwai, O., Hsieh, C.L., Schwarz, K. and Lieber, M.R. (2004). A biochemically defined system for mammalian nonhomologous DNA end joining. *Molecular cell*, 16(5), 701-713.
- Mahadevaiah, S. K., Turner, J. M., Baudat, F., Rogakou, E. P., de Boer, P., Blanco-Rodríguez, J., Jasin, M., Keeney, S., Bonner, W. M. and Burgoyne, P. S. (2001). Recombinational DNA double-strand breaks in mice precede synapsis. *Nature genetics*, 27(3), 271.
- Mahajan, K. N., McElhinny, S. A. N., Mitchell, B. S., & Ramsden, D. A. (2002). Association of DNA polymerase μ (pol μ) with Ku and ligase IV: role for pol μ in end-joining double-strand break repair. *Molecular and cellular biology*, 22(14), 5194-5202.
- Malone, C. D., & Hannon, G. J. (2009). Small RNAs as guardians of the genome. *Cell*, 136(4), 656-668.
- Malu, S., Malshetty, V., Francis, D., & Cortes, P. (2012). Role of non-homologous end joining in V (D) J recombination. *Immunologic research*, 54(1-3), 233-246.
- Manis, J.P., Gu, Y., Lansford, R., Sonoda, E., Ferrini, R., Davidson, L., Rajewsky, K. and Alt, F.W. (1998). Ku70 is required for late B cell development and immunoglobulin heavy chain class switching. *Journal of experimental Medicine*, 187(12), 2081-2089.
- Mannironi, C., Bonner, W. M., & Hatch, C. L. (1989). H2A. X. a histone isoprotein with a conserved C-terminal sequence, is encoded by a novel mRNA with both DNA replication type and polyA 3' processing signals. *Nucleic acids research*, 17(22), 9113-9126.
- Mari, P. O., Florea, B. I., Persengiev, S. P., Verkaik, N. S., Brüggewirth, H. T., Modesti, M., Giglia-Mari, G., Bezstarosti, K., Demmers, J. A., Luiders, T. M. and Houtsmuller, A. B. (2006). Dynamic assembly of end-joining complexes requires

- interaction between Ku70/80 and XRCC4. *Proceedings of the National Academy of Sciences*, 103(49), 18597-18602.
- Mariño-Ramírez, L., Kann, M. G., Shoemaker, B. A., & Landsman, D. (2005). Histone structure and nucleosome stability. *Expert review of proteomics*, 2(5), 719-729.
- Martin, C., & Zhang, Y. (2005). The diverse functions of histone lysine methylation. *Nature reviews Molecular cell biology*, 6(11), 838.
- Martin, G. M., Smith, A. C., Ketterer, D. J., Ogburn, C. E., & Disteché, C. M. (1985). Increased chromosomal aberrations in first metaphases of cells isolated from the kidneys of aged mice. *Israel journal of medical sciences*, 21(3), 296-301.
- Martinowich, K., Hattori, D., Wu, H., Fouse, S., He, F., Hu, Y., Fan, G. and Sun, Y.E. (2003). DNA methylation-related chromatin remodeling in activity-dependent BDNF gene regulation. *Science*, 302(5646), 890-893.
- Marzluff, W. F., Gongidi, P., Woods, K. R., Jin, J., & Maltais, L. J. (2002). The human and mouse replication-dependent histone genes. *Genomics*, 80(5), 487-498.
- Marzluff, W. F., Sanders, L. A., Miller, D. M., & McCarty, K. S. (1972). Two chemically and metabolically distinct forms of calf thymus histone F3. *Journal of Biological Chemistry*, 247(7), 2026-2033.
- Marzluff, W. F., Wagner, E. J., & Duronio, R. J. (2008). Metabolism and regulation of canonical histone mRNAs: life without a poly (A) tail. *Nature Reviews Genetics*, 9(11), 843.
- Mavrich, T.N., Ioshikhes, I.P., Venters, B.J., Jiang, C., Tomsho, L.P., Qi, J., Schuster, S.C., Albert, I. and Pugh, B.F. (2008). A barrier nucleosome model for statistical positioning of nucleosomes throughout the yeast genome. *Genome research*, 18(7), 1073-1083.
- Menon, V., & Povirk, L. F. (2017). XLF/Cernunnos: an important but puzzling participant in the nonhomologous end joining DNA repair pathway. *DNA repair*, 58, 29-37.
- Mimori, T. (2002). Clinical significance of anti-Ku autoantibodies-A serologic marker of overlap syndrome?. *Internal medicine*, 41(12), 1096-1098.
- Mimori, T., & Hardin, J. A. (1986). Mechanism of interaction between Ku protein and DNA. *Journal of Biological Chemistry*, 261(22), 10375-10379.

- Mimori, T., Hardin, J. A., & Steitz, J. A. (1986). Characterization of the DNA-binding protein antigen Ku recognized by autoantibodies from patients with rheumatic disorders. *Journal of Biological Chemistry*, 261(5), 2274-2278.
- Mimori, T.S., Akizuki, M.A., Yamagata, H.A., Inada, S.H., Yoshida, S., Homma, M. (1981). Characterization of a high molecular weight acidic nuclear protein recognized by autoantibodies in sera from patients with polymyositis-scleroderma overlap. *The Journal of clinical investigation*, 68(3), 611-20.
- Mo, X., & Dynan, W. S. (2002). Subnuclear localization of Ku protein: functional association with RNA polymerase II elongation sites. *Molecular and cellular biology*, 22(22), 8088-8099.
- Molden, R. C., Bhanu, N. V., LeRoy, G., Arnaudo, A. M., & Garcia, B. A. (2015). Multi-faceted quantitative proteomics analysis of histone H2B isoforms and their modifications. *Epigenetics & chromatin*, 8(1), 15.
- Morillo-Huesca, M., Clemente-Ruiz, M., Andújar, E., & Prado, F. (2010). The SWR1 histone replacement complex causes genetic instability and genome-wide transcription misregulation in the absence of H2A. *Z. PloS one*, 5(8), e12143.
- Murr, R., Loizou, J. I., Yang, Y. G., Cuenin, C., Li, H., Wang, Z. Q., & Herceg, Z. (2006). Histone acetylation by Trapp-Tip60 modulates loading of repair proteins and repair of DNA double-strand breaks. *Nature cell biology*, 8(1), 91.
- Murr, R., Vaissiere, T., Sawan, C., Shukla, V., & Herceg, Z. (2007). Orchestration of chromatin-based processes: mind the TRRAP. *Oncogene*, 26(37), 5358.
- Nakatani, Y., & Ogryzko, V. (2003). Immunoaffinity purification of mammalian protein complexes. *Methods in enzymology*, 370, 430-444.
- Neale, M. J., Pan, J., & Keeney, S. (2005). Endonucleolytic processing of covalent protein-linked DNA double-strand breaks. *Nature*, 436(7053), 1053.
- Nekrasov, M., Amrichova, J., Parker, B.J., Soboleva, T.A., Jack, C., Williams, R., Huttley, G.A. and Tremethick, D.J. (2012). Histone H2A. Z inheritance during the cell cycle and its impact on promoter organization and dynamics. *Nature structural & molecular biology*, 19(11), 1076.
- Nekrasov, M., Soboleva, T. A., Jack, C., & Tremethick, D. J. (2013). Histone variant selectivity at the transcription start site: H2A. Z or H2A. Lap1. *Nucleus*, 4(6), 431-437.

- Nemoz, C., Ropars, V., Frit, P., Gontier, A., Drevet, P., Yu, J., Guerois, R., Pitois, A., Comte, A., Delteil, C. and Barboule, N. (2018). XLF and APLF bind Ku80 at two remote sites to ensure DNA repair by non-homologous end joining. *Nature structural & molecular biology*, 25(10), 971.
- Nishibuchi, I., Suzuki, H., Kinomura, A., Sun, J., Liu, N.A., Horikoshi, Y., Shima, H., Kusakabe, M., Harata, M., Fukagawa, T. and Ikura, T. (2014). Reorganization of damaged chromatin by the exchange of histone variant H2A. Z-2. *International Journal of Radiation Oncology, Biology, Physics*, 89(4), 736-744.
- Nugent, C.I., Bosco, G., Ross, L.O., Evans, S.K., Salinger, A.P., Moore, J.K., Haber, J.E. and Lundblad, V. (1998). Telomere maintenance is dependent on activities required for end repair of double-strand breaks. *Current Biology*, 8(11), 657-662.
- Nusinow, D. A., Sharp, J. A., Morris, A., Salas, S., Plath, K., & Panning, B. (2007). The histone domain of macroH2A1 contains several dispersed elements that are each sufficient to direct enrichment on the inactive X chromosome. *Journal of molecular biology*, 371(1), 11-18.
- O'Hagan, H. M., Mohammad, H. P., & Baylin, S. B. (2008). Double strand breaks can initiate gene silencing and SIRT1-dependent onset of DNA methylation in an exogenous promoter CpG island. *PLoS genetics*, 4(8), 1-14
- Obri, A., Ouararhni, K., Papin, C., Diebold, M.L., Padmanabhan, K., Marek, M., Stoll, I., Roy, L., Reilly, P.T., Mak, T.W. and Dimitrov, S. (2014). ANP32E is a histone chaperone that removes H2A. Z from chromatin. *Nature*, 505(7485), 648.
- Ochi, T., Wu, Q., & Blundell, T. L. (2014). The spatial organization of non-homologous end joining: from bridging to end joining. *DNA repair*, 17, 98-109.
- Okumura, K., Sakaguchi, G., Takagi, S., Naito, K., Mimori, T., & Igarashi, H. (1996). Sp1 family proteins recognize the U5 repressive element of the long terminal repeat of human T cell leukemia virus type I through binding to the CACCC core motif. *Journal of Biological Chemistry*, 271(22), 12944-12950.
- Olins, A. L., & Olins, D. E. (1974). Spheroid chromatin units (v bodies). *Science*, 183(4122), 330-332.

- Oltval, Z. N., Milliman, C. L., & Korsmeyer, S. J. (1993). Bcl-2 heterodimerizes in vivo with a conserved homolog, Bax, that accelerates programmed cell death. *Cell*, 74(4), 609-619.
- Ono, M., Tucker, P. W., & Capra, J. D. (1994). Production and characterization of recombinant human Ku antigen. *Nucleic acids research*, 22(19), 3918-3924.
- Ord, M. G., & Stocken, L. A. (1967). Phosphate and thiol groups in histone f3 from rat liver and thymus nuclei. *Biochemical Journal*, 102(2), 631.
- Ouyang, H., Nussenzweig, A., Kurimasa, A., da Costa Soares, V., Li, X., Cordon-Cardo, C., Li, W.H., Cheong, N., Nussenzweig, M., Iliakis, G. and Chen, D.J. (1997). Ku70 is required for DNA repair but not for T cell antigen receptor gene recombination in vivo. *Journal of Experimental Medicine*, 186(6), 921-929.
- Paillard, S., & Strauss, F. (1991). Analysis of the mechanism of interaction of simian Ku protein with DNA. *Nucleic acids research*, 19(20), 5619-5624.
- Papamichos-Chronakis, M., Krebs, J. E., & Peterson, C. L. (2006). Interplay between Ino80 and Swr1 chromatin remodeling enzymes regulates cell cycle checkpoint adaptation in response to DNA damage. *Genes & development*, 20(17), 2437-2449.
- Papamichos-Chronakis, M., Watanabe, S., Rando, O. J., & Peterson, C. L. (2011). Global regulation of H2A. Z localization by the INO80 chromatin-remodeling enzyme is essential for genome integrity. *Cell*, 144(2), 200-213.
- Pasquinelli, A. E. (2012). MicroRNAs and their targets: recognition, regulation and an emerging reciprocal relationship. *Nature Reviews Genetics*, 13(4), 271.
- Pierce, B. A. (2012). Genetics: A conceptual approach. 2nd Edition. *Macmillan*.
- Porter, S. E., Greenwell, P. W., Ritchie, K. B., & Petes, T. D. (1996). The DNA-binding protein Hdf1p (a putative Ku homologue) is required for maintaining normal telomere length in *Saccharomyces cerevisiae*. *Nucleic acids research*, 24(4), 582-585.
- Postow, L., Ghenoiu, C., Woo, E. M., Krutchinsky, A. N., Chait, B. T., & Funabiki, H. (2008). Ku80 removal from DNA through double strand break-induced ubiquitylation. *The Journal of cell biology*, 182(3), 467-479.
- Price, B. D., & D'Andrea, A. D. (2013). Chromatin remodeling at DNA double-strand breaks. *Cell*, 152(6), 1344-1354.

- Price, C., Boltz, K. A., Chaiken, M. F., Stewart, J. A., Beilstein, M. A., & Shippen, D. E. (2010). Evolution of CST function in telomere maintenance. *Cell Cycle*, 9(16), 3177-3185.
- Ramakrishnan, V., Finch, J. T., Graziano, V., Lee, P. L., & Sweet, R. M. (1993). Crystal structure of globular domain of histone H5 and its implications for nucleosome binding. *Nature*, 362(6417), 219-223.
- Ramsden, D. A., & Gellert, M. (1998). Ku protein stimulates DNA end joining by mammalian DNA ligases: a direct role for Ku in repair of DNA double-strand breaks. *The EMBO Journal*, 17(2), 609-614.
- Ranjan, A., Wang, F., Mizuguchi, G., Wei, D., Huang, Y., & Wu, C. (2015). H2A histone-fold and DNA elements in nucleosome activate SWR1-mediated H2A. Z replacement in budding yeast. *Elife*, 4, e06845.
- Rasmussen, T. P., Mastrangelo, M. A., Eden, A., Pehrson, J. R., & Jaenisch, R. (2000). Dynamic relocation of histone MacroH2A1 from centrosomes to inactive X chromosomes during X inactivation. *The Journal of cell biology*, 150(5), 1189-1198.
- Redon, C., Pilch, D., Rogakou, E., Sedelnikova, O., Newrock, K., & Bonner, W. (2002). Histone H2A variants H2AX and H2AZ. *Current opinion in genetics & development*, 12(2), 162-169.
- Reeves, W. H. (1985). Use of monoclonal antibodies for the characterization of novel DNA-binding proteins recognized by human autoimmune sera. *Journal of Experimental Medicine*, 161(1), 18-39.
- Ribes-Zamora, A., Mihalek, I., Lichtarge, O., & Bertuch, A. A. (2007). Distinct faces of the Ku heterodimer mediate DNA repair and telomeric functions. *Nature structural & molecular biology*, 14(4), 301.
- Ribes-Zamora, A., Indiviglio, S. M., Mihalek, I., Williams, C. L., & Bertuch, A. A. (2013). TRF2 interaction with Ku heterotetramerization interface gives insight into c-NHEJ prevention at human telomeres. *Cell reports*, 5(1), 194-206.
- Richmond, T. J., Finch, J. T., Rushton, B., Rhodes, D., & Klug, A. (1984). Structure of the nucleosome core particle at 7 Å resolution. *Nature*, 311(5986), 532-537.
- Riggs AD, Russo VE, Martienssen RA (1996). Epigenetic mechanisms of gene regulation. Plainview, NY: Cold Spring Harbor Laboratory Press.

- Riggs, A. D., & Porter, T. N. (1996). Overview of epigenetic mechanisms. *Cold Spring Harbor monograph series*, 32, 29-46.
- Rivera-Calzada, A., Spagnolo, L., Pearl, L. H., & Llorca, O. (2007). Structural model of full-length human Ku70–Ku80 heterodimer and its recognition of DNA and DNA-PKcs. *EMBO reports*, 8(1), 56-62.
- Roberts, M. R., Miskimins, W. K., & Ruddle, F. H. (1989). Nuclear proteins TREF1 and TREF2 bind to the transcriptional control element of the transferrin receptor gene and appear to be associated as a heterodimer. *Cell regulation*, 1(1), 151-164.
- Roberts, S. A., & Ramsden, D. A. (2007). Loading of the nonhomologous end joining factor, Ku, on protein-occluded DNA ends. *Journal of Biological Chemistry*, 282(14), 10605-10613.
- Rogakou, E. P., Nieves-Neira, W., Boon, C., Pommier, Y., & Bonner, W. M. (2000). Initiation of DNA fragmentation during apoptosis induces phosphorylation of H2AX histone at serine 139. *Journal of Biological Chemistry*, 275(13), 9390-9395.
- Routh, A., Sandin, S., & Rhodes, D. (2008). Nucleosome repeat length and linker histone stoichiometry determine chromatin fiber structure. *Proceedings of the National Academy of Sciences*, 105(26), 8872-8877.
- Rybak, P., Hoang, A., Bujnowicz, L., Bernas, T., Berniak, K., Zarębski, M., Darzynkiewicz, Z. and Dobrucki, J. (2016). Low level phosphorylation of histone H2AX on serine 139 (γ H2AX) is not associated with DNA double-strand breaks. *Oncotarget*, 7(31), 49574.
- Sahoo, T., del Gaudio, D., German, J.R., Shinawi, M., Peters, S.U., Person, R.E., Garnica, A., Cheung, S.W., & Beaudet, A.L. (2008). Prader-Willi phenotype caused by paternal deficiency for the HBII-85 C/D box small nucleolar RNA cluster. *Nature genetics*, 40(6), 719-721.
- Santisteban, M. S., Kalashnikova, T., & Smith, M. M. (2000). Histone H2A. Z regulates transcription and is partially redundant with nucleosome remodeling complexes. *Cell*, 103(3), 411-422.
- Santoro, S. W., & Dulac, C. (2012). The activity-dependent histone variant H2BE modulates the life span of olfactory neurons. *Elife*, 1, e00070.

- Sawada, M., Sun, W., Hayes, P., Leskov, K., Boothman, D. A., & Matsuyama, S. (2003). Ku70 suppresses the apoptotic translocation of Bax to mitochondria. *Nature cell biology*, 5(4), 320.
- Schild-Poulter, C., Pope, L., Giffin, W., Kochan, J. C., Ngsee, J. K., Traykova-Andonova, M., & Haché, R. J. (2001). The binding of Ku antigen to homeodomain proteins promotes their phosphorylation by DNA-dependent protein kinase. *Journal of Biological Chemistry*, 276(20), 16848-16856.
- Scully, R., & Xie, A. (2013). Double strand break repair functions of histone H2AX. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis*, 750(1-2), 5-14.
- Shahrouki, P., & Larsson, E. (2012). The non-coding oncogene: a case of missing DNA evidence?. *Frontiers in genetics*, 3, 170.
- Shaytan, A. K., Landsman, D., & Panchenko, A. R. (2015). Nucleosome adaptability conferred by sequence and structural variations in histone H2A–H2B dimers. *Current opinion in structural biology*, 32, 48-57.
- Shilatifard, A. (2008). Molecular implementation and physiological roles for histone H3 lysine 4 (H3K4) methylation. *Current opinion in cell biology*, 20(3), 341-348.
- Shirodkar, P., Fenton, A. L., Meng, L., & Koch, C. A. (2013). Identification and functional characterization of a Ku-binding motif in aprataxin polynucleotide kinase/phosphatase-like factor (APLF). *Journal of Biological Chemistry*, 288(27), 19604-19613.
- Shrivastav, M., De Haro, L. P., & Nickoloff, J. A. (2008). Regulation of DNA double-strand break repair pathway choice. *Cell research*, 18(1), 134.
- Siegel, T.N., Hekstra, D.R., Kemp, L.E., Figueiredo, L.M., Lowell, J.E., Fenyo, D., Wang, X., Dewell, S. and Cross, G.A. (2009). Four histone variants mark the boundaries of polycistronic transcription units in *Trypanosoma brucei*. *Genes & development*, 23(9), 1063-1076.
- Singleton, B. K., Torres-Arzayus, M. I., Rottinghaus, S. T., Taccioli, G. E., & Jeggo, P. A. (1999). The C terminus of Ku80 activates the DNA-dependent protein kinase catalytic subunit. *Molecular and cellular biology*, 19(5), 3267-3277.

- Singleton, B.K., Priestley, A., Steingrimsdottir, H., Gell, D., Blunt, T., Jackson, S.P., Lehmann, A.R. and Jeggo, P.A. (1997). Molecular and biochemical characterization of xrs mutants defective in Ku80. *Molecular and cellular biology*, 17(3), 1264-1273.
- Soboleva, T. A., Nekrasov, M., Pahwa, A., Williams, R., Huttley, G. A., & Tremethick, D. J. (2012). A unique H2A histone variant occupies the transcriptional start site of active genes. *Nature structural & molecular biology*, 19(1), 25.
- Song, X., Bowen, J., Miao, W., Liu, Y., & Gorovsky, M. A. (2012). The nonhistone, N-terminal tail of an essential, chimeric H2A variant regulates mitotic H3-S10 dephosphorylation. *Genes & development*, 26(6), 615-629.
- Stedman, E., & Stedman, E. (1950). Cell specificity of histones. *Nature*, 166(4227), 780-781.
- Stellwagen, A. E., Haimberger, Z. W., Veatch, J. R., & Gottschling, D. E. (2003). Ku interacts with telomerase RNA to promote telomere addition at native and broken chromosome ends. *Genes & development*, 17(19), 2384-2395.
- Stuiver, M. H., Coenjaerts, F. E., & van der Vliet, P. C. (1990). The autoantigen Ku is indistinguishable from NF IV, a protein forming multimeric protein-DNA complexes. *Journal of Experimental Medicine*, 172(4), 1049-1054.
- Szenker, E., Ray-Gallet, D., & Almouzni, G. (2011). The double face of the histone variant H3. 3. *Cell research*, 21(3), 421.
- Taccioli, G.E., Gottlieb, T.M., Blunt, T., Priestley, A., Demengeot, J., Mizuta, R., Lehmann, A.R., Alt, F.W., Jackson, S.P. and Jeggo, P.A. (1994). Ku80: product of the XRCC5 gene and its role in DNA repair and V (D) J recombination. *Science*, 265(5177), 1442-1445.
- Talbert, P. B., & Henikoff, S. (2010). Histone variants-ancient wrap artists of the epigenome. *Nature reviews Molecular cell biology*, 11(4), 264-275.
- Talbert, P.B., Ahmad, K., Almouzni, G., Ausió, J., Berger, F., Bhalla, P.L., Bonner, W.M., Cande, W.Z., Chadwick, B.P., Chan, S.W. and Cross, G.A. (2012). A unified phylogeny-based nomenclature for histone variants. *Epigenetics & chromatin*, 5(1), 7.
- Thiriet, C., & Hayes, J. J. (2005). Replication-independent core histone dynamics at transcriptionally active loci in vivo. *Genes & Development*, 19(6), 677-682.

- Thoma, F., & Koller, T. (1977). Influence of histone H1 on chromatin structure. *Cell*, 12(1), 101-107.
- Thoma, F., Koller, T., & Klug, A. (1979). Involvement of histone H1 in the organization of the nucleosome and of the salt-dependent superstructures of chromatin. *The Journal of cell biology*, 83(2), 403-427.
- Thomson, T., & Lin, H. (2009). The biogenesis and function of PIWI proteins and piRNAs: progress and prospect. *Annual Review of Cell and Developmental*, 25, 355-376.
- Tollervey, J. R., & Lunyak, V. V. (2012). Epigenetics: judge, jury and executioner of stem cell fate. *Epigenetics*, 7(8), 823-840.
- Travers, A. A., & Klug, A. (1987). The bending of DNA in nucleosomes and its wider implications. *Philosophical Transactions of the Royal Society of London. B, Biological Sciences*, 317(1187), 537-561.
- Tuteja, N., Phan, T. N., Tuteja, R., Ochem, A., & Falaschi, A. (1997). Inhibition of DNA unwinding and ATPase activities of human DNA helicase II by chemotherapeutic agents. *Biochemical and biophysical research communications*, 236(3), 636-640.
- Ueda, K., Omachi, A., Kawaichi, M., & Hayaishi, O. (1975). Natural occurrence of poly (ADP-ribosyl) histones in rat liver. *Proceedings of the National Academy of Sciences*, 72(1), 205-209.
- Van Daal, A., White, E. M., Gorovsky, M. A., & Elgin, C. R. (1988). Drosophila has a single copy of the gene encoding a highly conserved histone H2A variant of the H2A. F/Z type. *Nucleic acids research*, 16(15), 7487-7497.
- Van Hoide, K. E., Sahasrabudhe, C. G., & Shaw, B. R. (1974). A model for particulate structure in chromatin. *Nucleic Acids Research*, 1(11), 1579-1586.
- Vernì, F., & Cenci, G. (2015). The Drosophila histone variant H2A. V works in concert with HP1 to promote kinetochore-driven microtubule formation. *Cell Cycle*, 14(4), 577-588.
- Vidali, G., Gershey, E. L. and Allfrey, V. G. (1968) Chemical studies of histone acetylation. The distribution of epsilon-N-acetyllysine in calf thymus histones. *Journal of Biological Chemistry*, 243, 6361–6366.

- Vignali, M., Hassan, A. H., Neely, K. E., & Workman, J. L. (2000). ATP-dependent chromatin-remodeling complexes. *Molecular and cellular biology*, 20(6), 1899-1910.
- Waddington, C. H. (1939). An introduction to modern genetics. George Allen And Unwin Ltd Museum Street; London, 154-156.
- Wahlestedt, C. (2013). Targeting long non-coding RNA to therapeutically upregulate gene expression. *Nature reviews Drug discovery*, 12(6), 433.
- Walker, J. R., Corpina, R. A., & Goldberg, J. (2001). Structure of the Ku heterodimer bound to DNA and its implications for double-strand break repair. *Nature*, 412(6847), 607-614.
- Wallis, J. W., Hereford, L., & Grunstein, M. (1980). Histone H2B genes of yeast encode two different proteins. *Cell*, 22(3), 799-805.
- Wang, H., Fang, R., Cho, J. Y., Libermann, T. A., & Oettgen, P. (2004). Positive and negative modulation of the transcriptional activity of the ETS factor ESE-1 through interaction with p300, CREB-binding protein, and Ku 70/86. *Journal of Biological Chemistry*, 279(24), 25241-25250.
- Wang, Y., Ghosh, G., & Hendrickson, E. A. (2009). Ku86 represses lethal telomere deletion events in human somatic cells. *Proceedings of the National Academy of Sciences*, 106(30), 12430-12435.
- Wang, Y., Liu, S., Sun, L., Xu, N., Shan, S., Wu, F., Liang, X., Huang, Y., Luk, E., Wu, C. and Zhou, Z. (2019). Structural insights into histone chaperone Chz1-mediated H2A. Z recognition and histone replacement. *PLoS biology*, 17(5), 3000277.
- Ward, W. S., & Coffey, D. S. (1991). DNA packaging and organization in mammalian spermatozoa: comparison with somatic cells. *Biology of Reproduction*, 44(4), 569-74.
- Weber, C. M., & Henikoff, S. (2014). Histone variants: dynamic punctuation in transcription. *Genes & development*, 28(7), 672-682.
- Weller, G.R., Kysela, B., Roy, R., Tonkin, L.M., Scanlan, E., Della, M., Devine, S.K., Day, J.P., Wilkinson, A., di Fagagna, F.D.A. and Devine, K.M. (2002). Identification of a DNA nonhomologous end-joining complex in bacteria. *Science*, 297(5587), 1686-1689.

- Wells, D., Hoffman, D., & Kedes, L. (1987). Unusual structure, evolutionary conservation of non-coding sequences and numerous pseudogenes characterize the human H3. 3 histone multigene family. *Nucleic acids research*, 15(7), 2871-2889.
- Wen, D., Noh, K. M., Goldberg, A. D., Allis, C. D., Rosenwaks, Z., Rafii, S., & Banaszynski, L. A. (2014). Genome editing a mouse locus encoding a variant histone, H3. 3B, to report on its expression in live animals. *Genesis*, 52(12), 959-966.
- West, M. H., & Bonner, W. M. (1980). Histone 2A, a heteromorphous family of eight protein species. *Biochemistry*, 19(14), 3238-3245.
- Whittaker, C. A., & Hynes, R. O. (2002). Distribution and evolution of von Willebrand/integrin A domains: widely dispersed domains with roles in cell adhesion and elsewhere. *Molecular biology of the cell*, 13(10), 3369-3387.
- Williams, G. J., Hammel, M., Radhakrishnan, S. K., Ramsden, D., Lees-Miller, S. P., & Tainer, J. A. (2014). Structural insights into NHEJ: building up an integrated picture of the dynamic DSB repair super complex, one component and interaction at a time. *DNA repair*, 17, 110-120.
- Witt, O., Albig, W., & Doenecke, D. (1996). Testis-specific expression of a novel human H3 histone gene. *Experimental cell research*, 229(2), 301-306.
- Woodard, R. L., Lee, K. J., Huang, J., & Dynan, W. S. (2001). Distinct roles for Ku protein in transcriptional reinitiation and DNA repair. *Journal of Biological Chemistry*, 276(18), 15423-15433.
- Woodcock, C. L. F., Safer, J. P., & Stanchfield, J. E. (1976). Structural repeating units in chromatin: I. Evidence for their general occurrence. *Experimental cell research*, 97(1), 101-110.
- Wu, D., Topper, L. M., & Wilson, T. E. (2008). Recruitment and dissociation of nonhomologous end joining proteins at a DNA double-strand break in *Saccharomyces cerevisiae*. *Genetics*, 178(3), 1237-1249.
- Wu, G., Broniscer, A., McEachron, T.A., Lu, C., Paugh, B.S., Becksfort, J., Qu, C., Ding, L., Huether, R., Parker, M. and Zhang, J. (2012). Somatic histone H3 alterations in pediatric diffuse intrinsic pontine gliomas and non-brainstem glioblastomas. *Nature genetics*, 44(3), 251-253.

- Wu, J., Mu, S., Guo, M., Chen, T., Zhang, Z., Li, Z., Li, Y. and Kang, X. (2016). Histone H2B gene cloning, with implication for its function during nuclear shaping in the Chinese mitten crab, *Eriocheir sinensis*. *Gene*, 575(2), 276-284.
- Wu, T.P., Wang, T., Seetin, M.G., Lai, Y., Zhu, S., Lin, K., Liu, Y., Byrum, S.D., Mackintosh, S.G., Zhong, M. and Tackett, A. (2016). DNA methylation on N 6-adenine in mammalian embryonic stem cells. *Nature*, 532(7599), 329-333.
- Xu, Y., Sun, Y., Jiang, X., Ayrapetov, M.K., Moskwa, P., Yang, S., Weinstock, D.M. and Price, B.D. (2010). The p400 ATPase regulates nucleosome stability and chromatin ubiquitination during DNA repair. *The Journal of cell biology*, 191(1), 31-43.
- Xu, Y., Ayrapetov, M. K., Xu, C., Gursoy-Yuzugullu, O., Hu, Y., & Price, B. D. (2012). Histone H2A. Z controls a critical chromatin remodeling step required for DNA double-strand break repair. *Molecular cell*, 48(5), 723-733.
- Yano, K. I., Morotomi-Yano, K., Lee, K. J., & Chen, D. J. (2011). Functional significance of the interaction with Ku in DNA double-strand break recognition of XLF. *FEBS letters*, 585(6), 841-846.
- Yano, K.I., Morotomi-Yano, K., Wang, S.Y., Uematsu, N., Lee, K.J., Asaithamby, A., Weterings, E. and Chen, D.J. (2008). Ku recruits XLF to DNA double-strand breaks. *EMBO reports*, 9(1), 91-96.
- Yelagandula, R., Stroud, H., Holec, S., Zhou, K., Feng, S., Zhong, X., Muthurajan, U.M., Nie, X., Kawashima, T., Groth, M. and Luger, K. (2014). The histone variant H2A. W defines heterochromatin and promotes chromatin condensation in Arabidopsis. *Cell*, 158(1), 98-109.
- Yoo, S., Kimzey, A., & Dynan, W. S. (1999). Photocross-linking of an Oriented DNA Repair Complex Ku bound at a single DNA end. *Journal of Biological Chemistry*, 274(28), 20034-20039.
- Yuan, J., Adamski, R., & Chen, J. (2010). Focus on histone variant H2AX: to be or not to be. *FEBS letters*, 584(17), 3717-3724.
- Zhang, J., & Jacob, S. T. (1990). Purification and characterization of a novel factor which stimulates rat ribosomal gene transcription in vitro by interacting with enhancer and core promoter elements. *Molecular and cellular biology*, 10(10), 5177-5186.

- Zhang, W. W., & Yaneva, M. (1992). On the mechanisms of Ku protein binding to DNA. *Biochemical and biophysical research communications*, 186(1), 574-579.
- Zhang, Z., Hu, W., Cano, L., Lee, T. D., Chen, D. J., & Chen, Y. (2004). Solution structure of the C-terminal domain of Ku80 suggests important sites for protein-protein interactions. *Structure*, 12(3), 495-502.
- Zhao, Y., & Garcia, B. A. (2015). Comprehensive catalog of currently documented histone modifications. *Cold Spring Harbor perspectives in biology*, 7(9), a025064.
- Zhou, J., Fan, J. Y., Rangasamy, D., & Tremethick, D. J. (2007). The nucleosome surface regulates chromatin compaction and couples it with transcriptional repression. *Nature structural & molecular biology*, 14(11), 1070.
- Ziats, M. N., & Rennert, O. M. (2013). Aberrant expression of long noncoding RNAs in autistic brain. *Journal of molecular neuroscience*, 49(3), 589-593.

Ku70/Ku80 heterodimer is involved in H2A.Z deposition in chromatin

Arslan Iftikhar¹, Christophe Papin¹, Ali Hamiche^{1*}

¹Institut de Génétique et de Biologie Moléculaire et Cellulaire, the Centre National de la Recherche Scientifique UMR 7104, the Institut National de la Santé et de la Recherche Médicale U964, Université de Strasbourg, Parc d'innovation, 1 rue Laurent Fries, 67404 Illkirch Cedex, France.

*To whom correspondence should be addressed:

Dr. Ali Hamiche

IGBMC

1 rue Laurent Fries,

67404 Illkirch Cedex, France.

E-mail: hamiche@igbmc.fr

Phone: +33 388 65 32 50

Keywords: Ku Heterodimer, Chromatin Remodeling, Histone Deposition, H2A.Z/H2B.

Abstract:

The Ku heterodimer, comprising of Ku70 and Ku80 subunits, is an evolutionary conserved highly abundant chromatin associated complex with a prominent role in NHEJ repair. Using tandem-affinity purification procedure, we identified the Ku heterodimer as one of the major component of the H2A.Z predeposition complex. Using deletion mapping, we identified the von Willebrand factor type A (vWA) domain of Ku80 as the primary motif responsible for its interaction with H2A.Z/H2B dimers. This specific interaction is mediated essentially by H2B in the presence of H2A.Z but not in the presence of H2A or H2AX. Analysis of H2A.Z localization in Ku80 knockout cells by chromatin immunoprecipitation followed by sequencing indicate specific depletion of H2A.Z from promoters but not from enhancers or insulators. Collectively, our data uncover a novel function of the Ku complex in the mechanism of H2A.Z/H2B deposition at promoters.

Introduction:

The Ku heterodimer, consisting of Ku70 and Ku80 subunits (also referred as Ku86 in higher eukaryotes), is a highly abundant protein found primarily in the nuclear region of cells (Koike *et al.*, 2001). Both subunits form a toroidal structure that wraps itself around the DNA to perform the dedicated functions. However, recent researches are now suggesting Ku heterodimer not merely as a DNA binding protein rather an entity with a wide range of functions in its account. These functions include its pivotal roles in apoptosis (Jing-Jing *et al.*, 2013), V(D)J recombination (Manis *et al.*, 1998), telomere maintenance (Wang *et al.*, 2009), DNA replication (Mukherjee *et al.*, 2016), transcriptional activation (Woodard *et al.*, 2001) and most importantly in genomic stability through serving as a molecular scaffold to recruit the primary members of the NHEJ repair pathway at the sites of double-strand DNA breaks (Grundy *et al.*, 2014). On an average, about half a million of Ku molecules are considered to be present in the nucleus of a single human cell (Mimori *et al.*, 1986; Walker *et al.*, 2001). Presence in such a huge amount certainly advocates its important roles in several key cellular processes.

Research annals have indicated chromatin as a highly dynamic structure that keeps changing its organizational design to accomplish its diversified roles in the cell. The process of chromatin remodeling involves the ATP-dependent changes in the overall structure of chromatin that result in altered access of different cellular proteins to the chromatin. These changes not only involve the covalent modifications of nucleosome structure but also include the physical rearrangement of the histone machinery that leads to alteration in nucleosome stability, dynamics and eventually the accessibility of genetic material, thus affecting many important cellular mechanisms. Histone replacement involves the physical exchange of conventional canonical histones with that of corresponding non- allelic, non-canonical variants. Despite the fact that these variants possess a quite high resemblance to their canonical counterparts, they can cause considerable changes in chromatin configuration and its functions.

Among histone remodeling, H2A.Z remodeling has got huge attention of the research community in the recent past by virtue of its critical role in overall chromatin dynamics and gene expression. Histone variant H2A.Z has been reported to be involved in a variety of important cellular processes including transcription (Hardy *et al.*, 2009), genomic stability (Papamichos-Chronakis *et al.*, 2011), chromosome segregation (Hou *et al.*, 2010), and DNA repair (Xu *et al.*, 2012), etc. Bearing 60% sequence similarity to that of other cellular H2As, H2A.Z is the most extensively studied histone variant among all its family members. It constitutes about 15% of all H2A variants and is primarily found at promotor regions of active genes. With ~90% sequence conservation among diversified species, H2A.Z is considered as the most conserved histone variant during the course of evolution that justifies its important role in a variety of key cellular processes.

In literature archive, a wealth of information has been provided regarding the mechanism of H2A.Z remodeling, however, with the advent of new emerging techniques to comprehend such dynamic phenomenon, new members are being unveiled playing a significant role in these mechanisms. In yeast, the ATP-dependent chromatin-remodeling complex SWR1 is responsible for the exchange of canonical H2A/H2B dimer with histone variant H2A.Z/H2B dimer mediated by histone chaperon Chz1. In higher eukaryotes, two orthologs of SWR1 complex, the p400/TIP60, and Snf2 related CBP activator protein (SRCAP) complexes are considered responsible for replacement of H2A with H2A.Z at specific loci. Here, we set out to investigate the interaction of Ku heterodimer and histone variant H2A.Z and the probable functional implications of this interaction in H2A.Z deposition in chromatin. Through both *in vivo* and *in vitro* experiments, we identified the von Willebrand factor type A (vWA) domain of Ku80 as the primary motif responsible for its interaction with H2A.Z/H2B dimers. This specific interaction is mediated essentially by H2B in the presence of H2A.Z but not in the presence of H2A or H2AX. Finally, the genome-wide chromatin immunoprecipitation followed by sequencing (ChIP-seq) indicated that Ku80 regulates H2A.Z deposition at promoters that indicated the functional significance of this specific interaction.

Results:

Ku heterodimer specifically interacts with H2A.Z/H2B

We have previously identified Ku70 (XRCC6) and Ku80 (XRCC5) along with other DNA repair members in the H2A.Z soluble complex (**Supplementary Fig. 1**) during one of our previous work on histone chaperone ANP32E (Obri *et al.*, 2014). To investigate this interaction further, we proceeded towards the exploration of specificity of interaction. Towards this goal, we co-expressed Ku heterodimer with H2A/H2B, H2A.X/H2B and H2A.Z/H2B using *E-coli* expression system. Upon subsequent immunoprecipitation using Flag epitope, the eluted proteins were fractionated using SDS-PAGE. Upon colloidal blue staining, Ku heterodimer exhibited specific interaction with H2A.Z/H2B histone dimers stable at high salt concentration whereas no interaction was shown with counterpart H2A/H2B and H2A.X/H2B dimers (**Fig. 1**). This specificity suggested us the probable prospective of this interaction in the regulation of H2A.Z in chromatin.

von Willebrand factor type A (vWA) domain of Ku80 specifies H2A.Z/H2B binding

Following the revelation of exclusive interaction of Ku heterodimer with histone dimer H2A.Z/H2B, we next decided to further characterize in details this interaction by mapping the minimal Ku70 or Ku80 domain involved in this interaction. Based on their primary structures, we decided to divide both Ku70 and Ku80 proteins into four functional domains and to express and assay each domain separately for its interaction with H2A.Z/H2B dimers. Structural and functional studies have suggested that both the Ku70 and Ku80 consist of an N-terminal von Willebrand factor type A (vWA) domain which is comprised of 37-256 residues in Ku70 and 09-244 residues in Ku80. The central DNA binding core domain is composed of residues 263-466 in Ku70 and residues 253-453 in Ku80. Following the central domain comes the nuclear localization sequence (NLS) domain that comprised of 472-557 amino acids sequence in Ku70 and 476-569 in case of Ku80. The C-terminal domain of Ku70 (residues 573-609) comprises an SAP (SAF- A/B, Acinus and PIAS) domain while C-terminal domain Ku80 (residues 593-706) comprises of sequence

responsible for binding with DNA-PK_{CS} (**Fig. 2a**). Upon separate expression of both the histone dimer and GST tagged Ku domains, they were mixed individually in 0.5M TGEN buffer. Thereafter, the bound proteins were subjected to GST pulldown and eluted proteins were fractionated on SDS-PAGE and stained with colloidal blue. The eluted fractions were also immunoblotted with anti-Flag (H2A.Z) and anti-His (H2B) antibodies to validate the results. Both gel results, staining and immunoblotting, indicated von Willebrand factor type A (vWA) domain (also known as α/β domain) of Ku80 to be primarily responsible for this interaction with histone dimer H2A.Z/H2B (**Fig. 2b**). To further validate this interaction in a more physiological context, the histone H2A.Z/H2B dimers were co-expressed in bacteria with the different Ku70/80 domains and assayed for interaction using the GST-pull-down assay. Likewise, the eluted fractions were fractionated on SDS-PAGE and stained with colloidal blue and blotted for anti-Flag and anti-His antibodies. Upon consideration of both gel results, staining and immunoblotting, the vWA domain of Ku80 was found required for interaction with the histone H2A.Z/H2B dimer (**Fig. 2c**). This finding is consistent with the above described results where separate eluted domains were mixed and assayed for interaction *in vitro*.

Following that, in the search of minimal vWA domain required for H2A.Z/H2B binding, we subdivided the vWA domain of Ku80 into three subdomains. The first subdomain encompassed amino acid 09-125. The second subdomain included amino acids 126-244 and the third small 7kDa subdomain included amino acids 40-101 (**Fig. 3a**). All subdomains were co-expressed with histone H2A.Z/H2B the same *E. coli* expression system. The co-expressed proteins were assayed for interaction using subsequent GST-pull-down assay. The eluted proteins were fractionated on SDS-PAGE and stained with colloidal blue. The obtained results were further validated by western blotting using anti-Flag and anti-His antibodies. Both gel results, staining and immunoblotting, suggested that highly conserved Ku80 domain encompassing amino acids 40-101 is the minimal domain required for interacting with H2A.Z/H2B (**Fig. 3b and Supplementary Fig. 3**). Upon

every domain identification, the specificity of its interaction with H2A.Z/H2B was assessed by its comparative interaction with counterpart H2A/H2B and H2A.X/H2B (**Supplementary Fig. 2**).

H2A.Z α 2 helix is essential for Ku heterodimer interaction

Following the recognition of minimal domain of Ku heterodimer responsible for the interaction with H2A.Z/H2B, we aimed to identify the H2A.Z domain critical for this interaction. Towards this goal, chimeras constituting H2A and H2A.Z portions were designed based on structural and functional data available in the literature. We started truncation of H2A.Z from the C-terminus (1-18, 1-44, 1-82, 1-105) and replaced it with counterpart C-terminus of H2A. An additional chimera was designed that contain the middle part of H2A.Z (45-82) that primarily comprised of α 2 helix of H2A.Z, connected with H2A brims on both N and C terminus. Upon colloidal blue staining, the central part of H2A.Z comprising the highly conserved α 2 helix was found to be critical for interaction with Ku80 (aa 40-101). All the three chimeric constructs (chimera 3, chimera 4 and chimera 5) containing middle part of H2A.Z (45-82) presented a significant interaction with Ku80 minimal motif (**Fig. 4 and Supplementary Fig. 4**).

Ku heterodimer and H2A.Z/H2B interaction is mediated essentially by H2B

Noteworthy, in the context of all *in vitro* experiments conducted, the amount of H2B was always found substantially higher as compared to H2A.Z suggesting that Ku80 (aa 40-101) motif might be binding to the histone H2A.Z/H2B dimer favorably through H2B partner. Employing the *in vitro* exploration through double immunoaffinity procedure and size-exclusion chromatography, we observed a strong interaction of H2B with Ku80 (aa 40-101) motif (**Fig. 5b and Supplementary Fig. 5**). To further validate the interaction, we mixed separately expressed GST tagged Ku80 (aa 40-101) and H2B and upon subsequent GST pulldown, a strong interaction was found between both proteins (**Fig. 5a**). This advocated the pivotal role of H2B in the interaction of Ku heterodimer and histone H2A.Z/H2B dimer, though, the interaction with H2B is feasible only

in the case of H2A.Z/H2B dimer and not in the case of H2A/H2B or H2A.X/H2B counterparts. Both gel results, staining and immunoblotting, were found in compliance with previous results suggesting the dedicated role of H2B in novel interaction of Ku heterodimer Ku70/Ku80 with histone dimer H2A.Z/H2B. This strong interaction also explained the high-density bands of H2B as compared to H2A.Z in all previous experiments of interaction with Ku80 (aa 40-101) motif.

Identification of histones in Ku80 predeposition complex

After the *in vitro* exploration of the interaction of Ku80 (aa 40-101) motif and histone H2A.Z/H2B dimer, we proceeded to evaluate this specific interaction through *in vivo* examination. For that purpose, we established two distinct stable MEF cell lines expressing N-terminally FLAG/HA epitope-tagged wild type human Ku80 (1-732) or Δ NKu80 (107-732) by retroviral transduction and anti-CD4 immunomagnetic selection of positive cells (**Fig. 6**). The *e-wtKu80* and *e- Δ NKu80* complexes were purified from the soluble nuclear extracts (SNE) and the insoluble nuclear extracts (INE) using tandem immuno-affinity purification anti-FLAG followed by anti-HA. The purified complex was then fractionated on SDS-PAGE and subjected to silver staining (Figure 42). Mass spectrometric analysis identified several NHEJ related proteins as part of the *e-wtKu80* complex (such as DNA-PK_{CS}, Ku70, PARP1, PARP2, HLTF, CHD1L, H2A.Z, H2B, H3, H4, etc) (**Fig. 7**). Interestingly, *e- Δ NKu80* complex has lost all the NHEJ proteins and the histones particularly H2A.Z and H2B. These data highlighted the indispensable importance of the vWA domain of Ku80 in recruiting the H2A.Z/H2B dimer and stabilizing the NHEJ complex.

Ku heterodimer is involved in genomic localization and chromatin enrichment of H2A.Z at promoters

After the exploration of the interaction of Ku heterodimer and histone H2A.Z/H2B dimer through *in vivo* and *in vitro* approaches, we proceeded to evaluate the functional implications of this interaction. If Ku heterodimer Ku70/Ku80 is involved in H2A.Z functioning *in vivo*, its

depletion should affect impair the H2A.Z recruitment to chromatin. For that purpose, we performed a genome-wide comparative ChIP-seq analysis using both wild type and knockout Ku80 MEF cells. Upon bioinformatic analysis, we identified that Ku heterodimer specifically regulates the H2A.Z deposition in chromatin, specifically at promoter regions. We found about 50% increased deposition of H2A.Z around promoter regions (TSS > 2 kilobases) with no significant variance in the deposition of H2A.Z at enhancers and insulators (**Fig. 8a,b**). This validated the critical involvement of Ku80 in genome- wide localization and chromatin enrichment of H2A.Z at promoters. Following that, we assessed the significance of Ku80 (aa 40-101) motif in H2A.Z deposition employing the same ChIP-Seq approach. Upon analysis, we observed a two-fold increase in H2A.Z deposition in case of wt-Ku80 MEFs as compared to DN-Ku80 MEF cells (**Fig. 8a,c**). This finding strongly advocated the critical role of the vWA domain or more precisely Ku80 (aa 40- 101) motif in the deposition of H2A.Z/H2B *in vivo*.

Discussion:

In literature archive, chromatin remodeling has been indicated as an essential phenomenon required for the repair of double-strand breaks (DSB). H2A.Z exchange causes the formation of an open relaxed chromatin conformation that is considered a prerequisite for the recruitment of repair machinery at the site of DNA damage. This ATP-dependent remodeling that specifically loads H2A.Z onto chromatin and exchanges it with H2A has been designated as a highly complex process. With the advent of contemporary researches, new members are being emerged that are crucial for the deposition of H2A.Z at chromatin. Here, we identified that Ku heterodimer plays a fundamental role in the deposition of H2A.Z at promoters. Through both *in vivo* and *in vitro* experiments, we demonstrated the specific interaction of Ku heterodimer and histone variant H2A.Z and identified the von Willebrand factor type A (vWA) domain or more precisely Ku80 (aa 40- 101) motif as the primary motif responsible for the interaction of Ku heterodimer with H2A.Z/H2B dimers.

Towards determining the biochemical characterization of interaction, we found this specific interaction mediated primarily by H2B in the presence of H2A.Z but not in the presence of H2A or H2A.X. Finally, the analysis of H2A.Z localization in Ku80 knockout cells by chromatin immunoprecipitation followed by sequencing indicated specific depletion of H2A.Z from promoters but not from enhancers or insulators suggesting us the specific role of this association in H2A.Z deposition at promoter region that leads this chromatin region becoming more permissive to transcription initiation and elongation.

In conclusion, our study establishes the molecular basis and specificity of H2A.Z recognition by the Ku70/Ku80 heterodimer and provides a framework for a better understanding of H2A.Z deposition process. Our findings also provide insights into the additional functions of Ku heterodimer in mediating histone transactions that regulate gene expression and genome integrity.

Acknowledgements:

This work was supported by institutional funds from the Centre Nationale de la Recherche Scientifique (CNRS, France), the Institut National de la Santé et de la Recherche Médicale (INSERM, France) and the Université de Strasbourg (UdS).

References:

Grundy, G. J., Moulding, H. A., Caldecott, K. W., & Rulten, S. L. (2014). One ring to bring them all - the role of Ku in mammalian non-homologous end joining. *DNA repair*, 17, 30-38.

Hardy, S., Jacques, P.É., Gévry, N., Forest, A., Fortin, M.È., Laflamme, L., Gaudreau, L. and Robert, F. (2009). The euchromatic and heterochromatic landscapes are shaped by antagonizing effects of transcription on H2A. Z deposition. *PLoS genetics*, 5(10), e1000687.

Hou, H., Wang, Y., Kallgren, S. P., Thompson, J., Yates, J. R., & Jia, S. (2010). Histone variant H2A.Z regulates centromere silencing and chromosome segregation in fission yeast. *Journal of Biological Chemistry*, 285(3), 1909-1918.

- Jing-Jing, L., Qi-Hua, G., Min, L., Hua-Ping, Y., Li-Ming, C., & Cheng-Ping, H. (2013). Role of Ku70 and Bax in epigallocatechin-3-gallate-induced apoptosis of A549 cells in vivo. *Oncology letters*, 5(1), 101-106.
- Manis, J.P., Gu, Y., Lansford, R., Sonoda, E., Ferrini, R., Davidson, L., Rajewsky, K. and Alt, F.W. (1998). Ku70 is required for late B cell development and immunoglobulin heavy chain class switching. *Journal of experimental Medicine*, 187(12), 2081-2089.
- Mimori, T., Hardin, J. A., & Steitz, J. A. (1986). Characterization of the DNA-binding protein antigen Ku recognized by autoantibodies from patients with rheumatic disorders. *Journal of Biological Chemistry*, 261(5), 2274-2278.
- Mukherjee, S., Chakraborty, P., & Saha, P. (2016). Phosphorylation of Ku70 subunit by cell cycle kinases modulates the replication related function of Ku heterodimer. *Nucleic acids research*, 44(16), 7755-7765.
- Obri, A., Ouararhni, K., Papin, C., Diebold, M.L., Padmanabhan, K., Marek, M., Stoll, I., Roy, L., Reilly, P.T., Mak, T.W. and Dimitrov, S. (2014). ANP32E is a histone chaperone that removes H2A. Z from chromatin. *Nature*, 505(7485), 648.
- Papamichos-Chronakis, M., Watanabe, S., Rando, O. J., & Peterson, C. L. (2011). Global regulation of H2A. Z localization by the INO80 chromatin-remodeling enzyme is essential for genome integrity. *Cell*, 144(2), 200-213.
- Walker, J. R., Corpina, R. A., & Goldberg, J. (2001). Structure of the Ku heterodimer bound to DNA and its implications for double-strand break repair. *Nature*, 412(6847), 607-614.
- Wang, X., Dewell, S. and Cross, G.A. (2009). Four histone variants mark the boundaries of polycistronic transcription units in *Trypanosoma brucei*. *Genes & development*, 23(9), 1063-1076.
- Woodard, R. L., Lee, K. J., Huang, J., & Dynan, W. S. (2001). Distinct roles for Ku protein in transcriptional reinitiation and DNA repair. *Journal of Biological Chemistry*, 276(18), 15423-15433.
- Xu, Y., Ayrapetov, M. K., Xu, C., Gursoy-Yuzugullu, O., Hu, Y., & Price, B. D. (2012). Histone H2A. Z controls a critical chromatin remodeling step required for DNA double- strand break repair. *Molecular cell*, 48(5), 723-733.

Figures:

Figure 1:

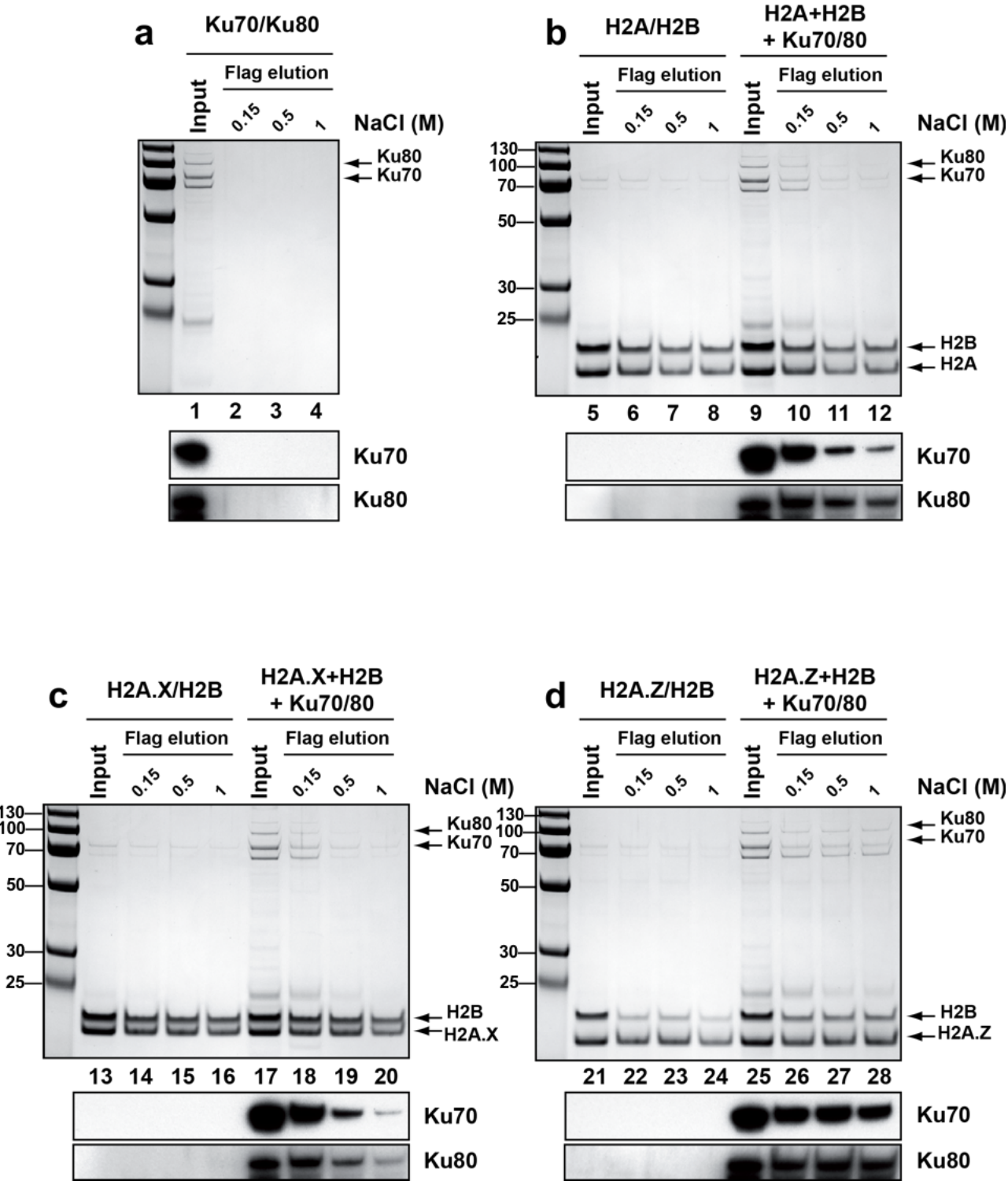


Figure 2:

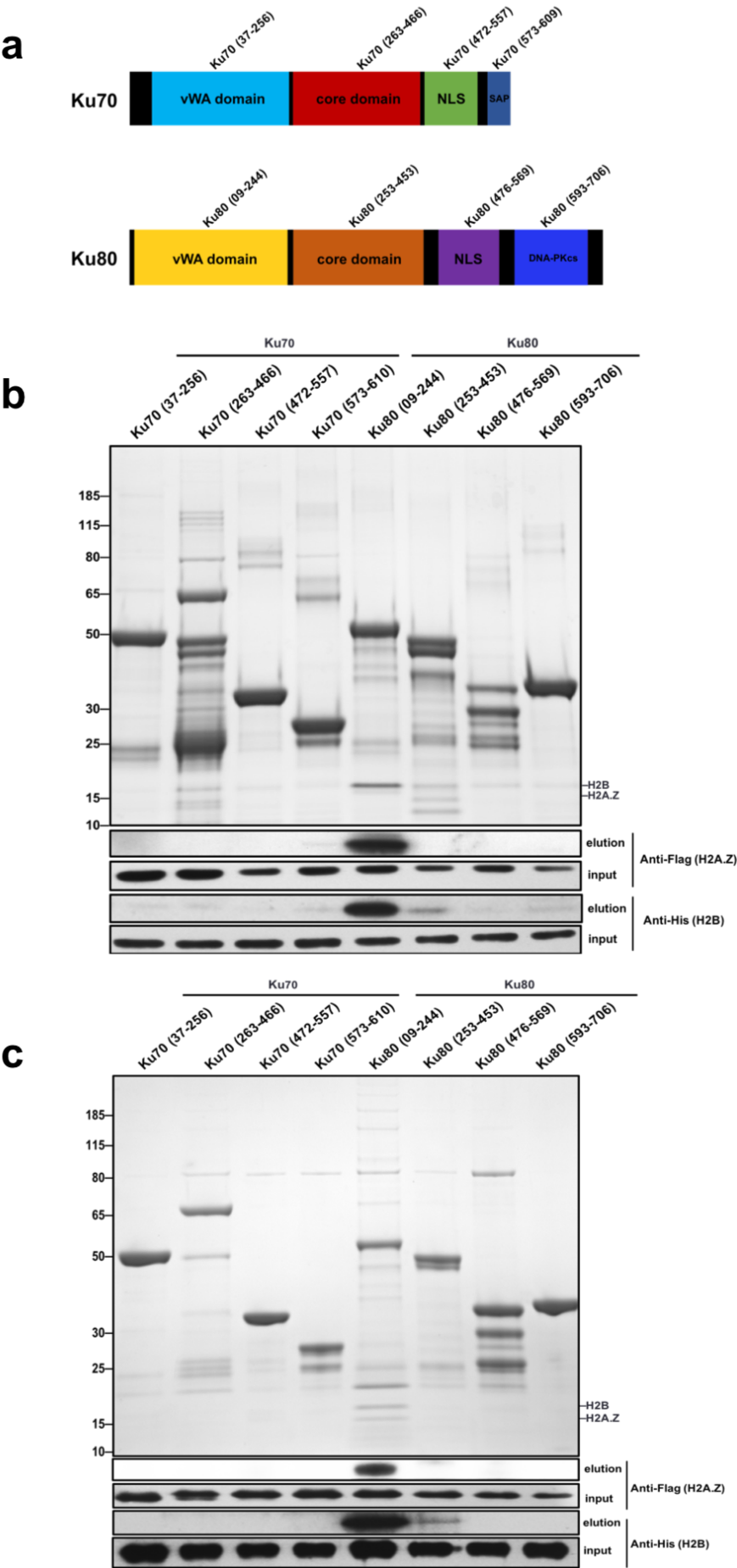


Figure 3:

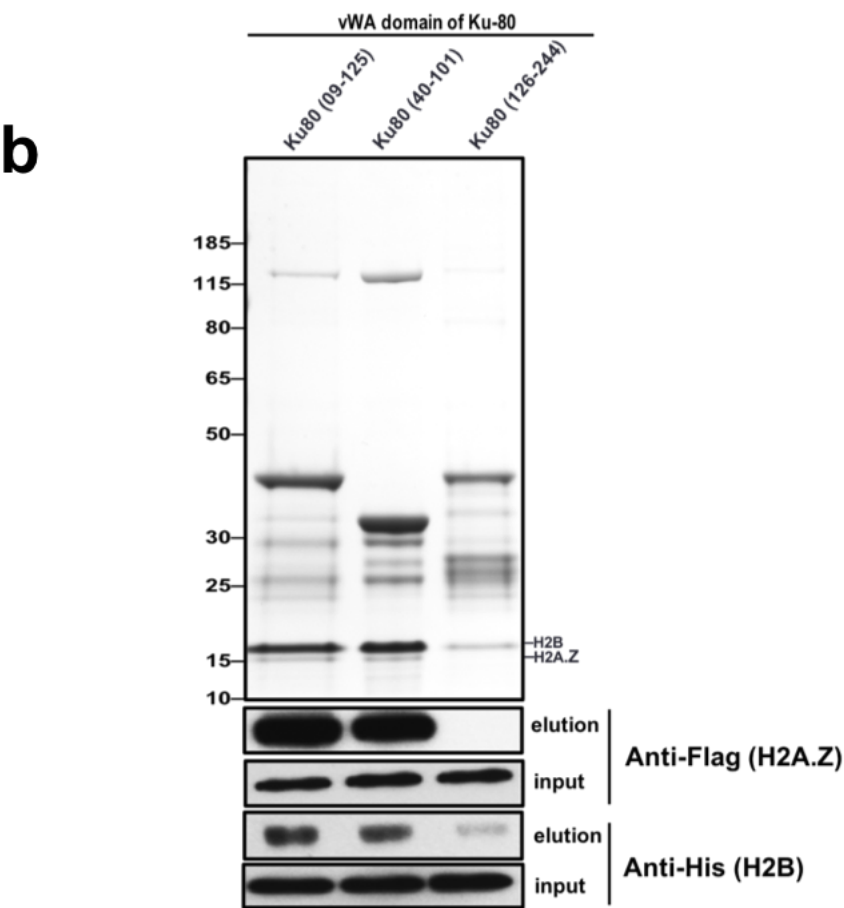
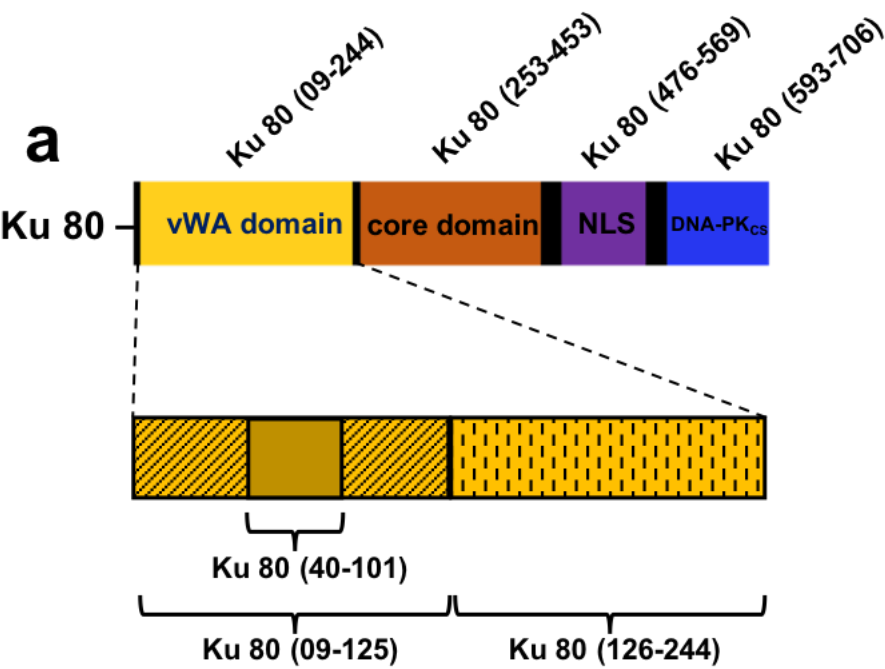


Figure 4:

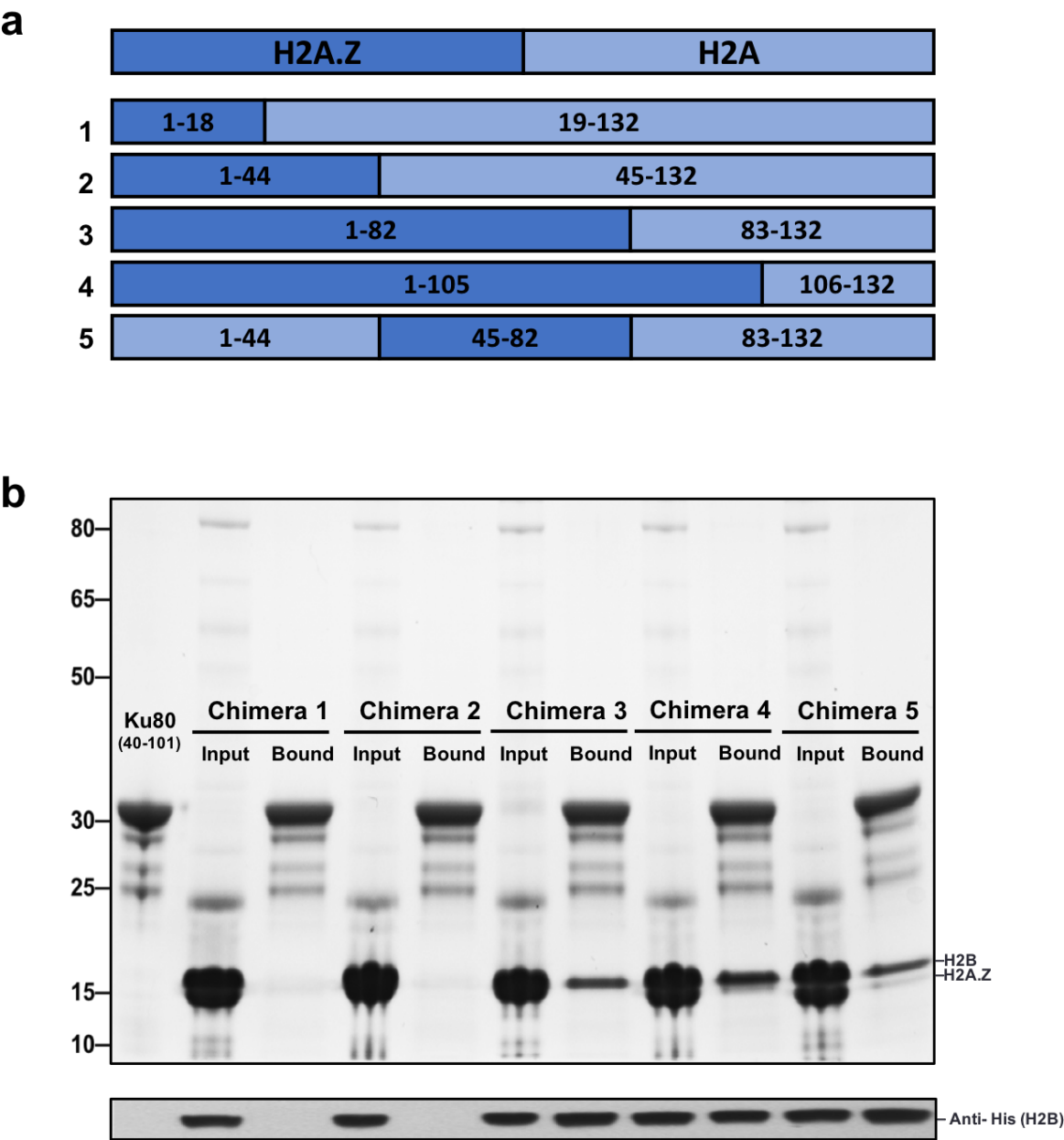


Figure 5:

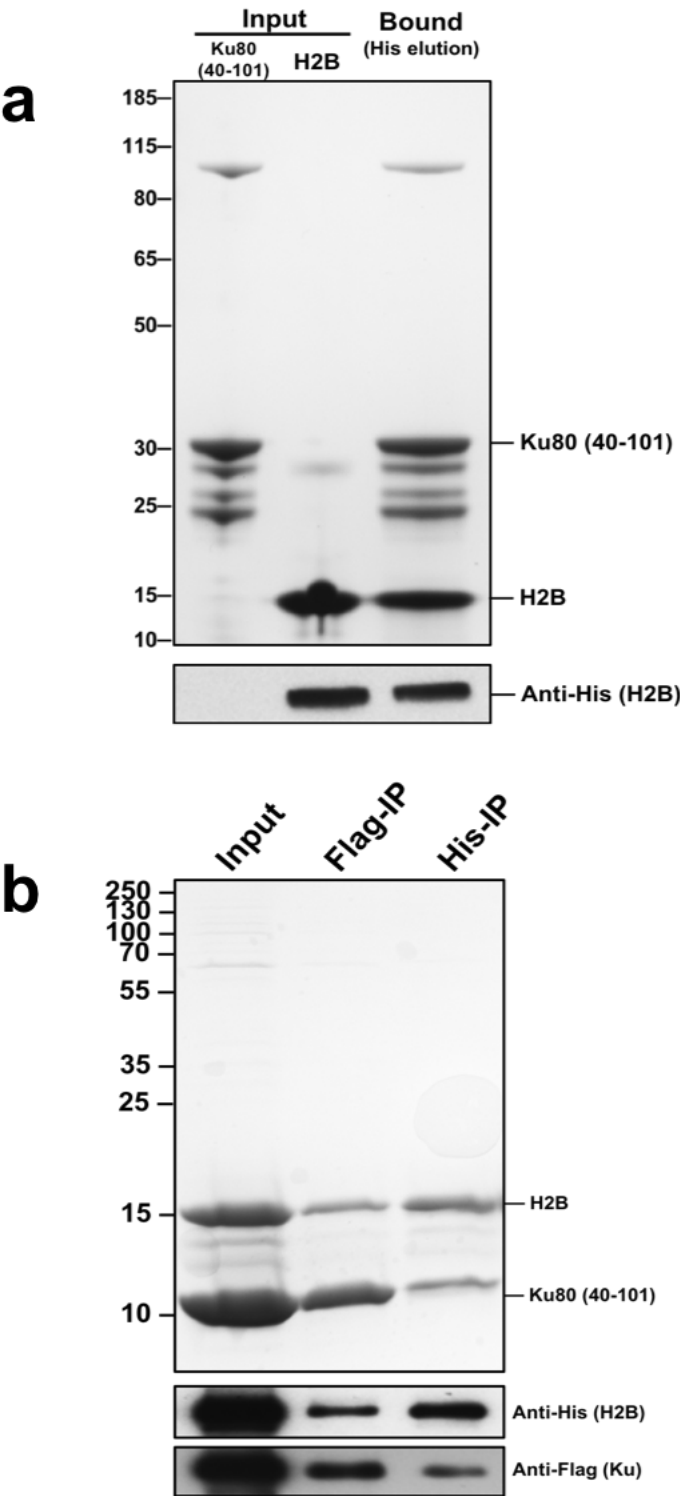


Figure 6:

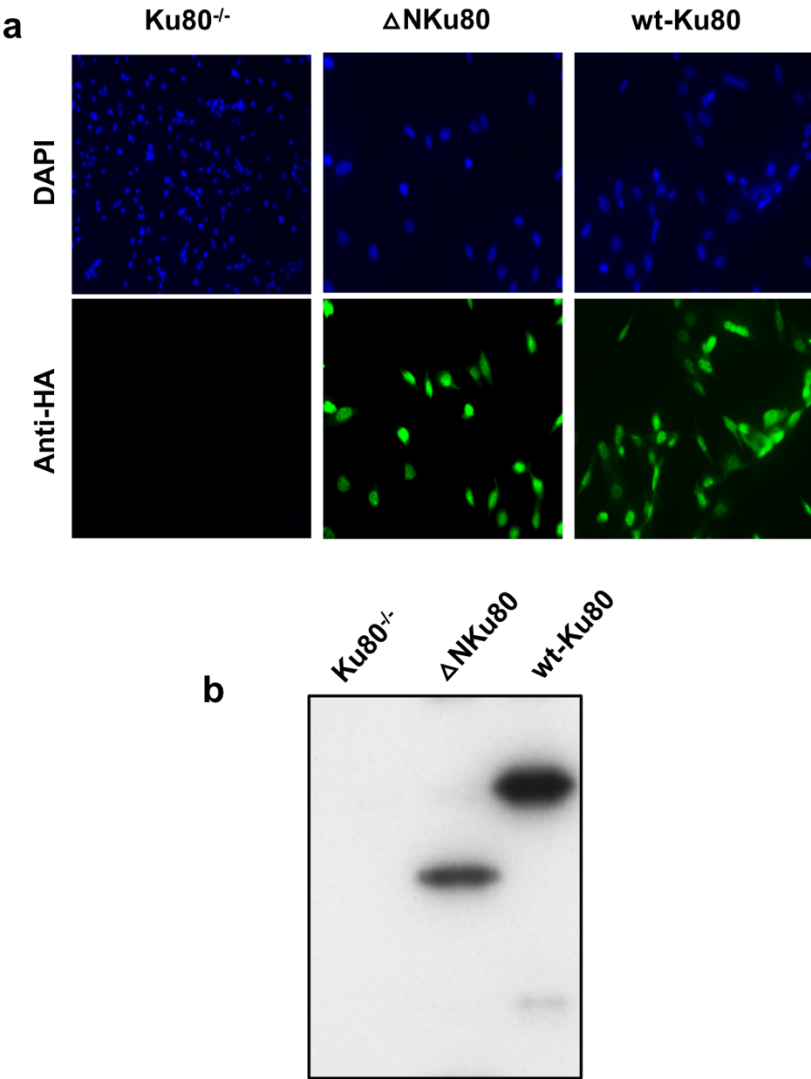


Figure 7:

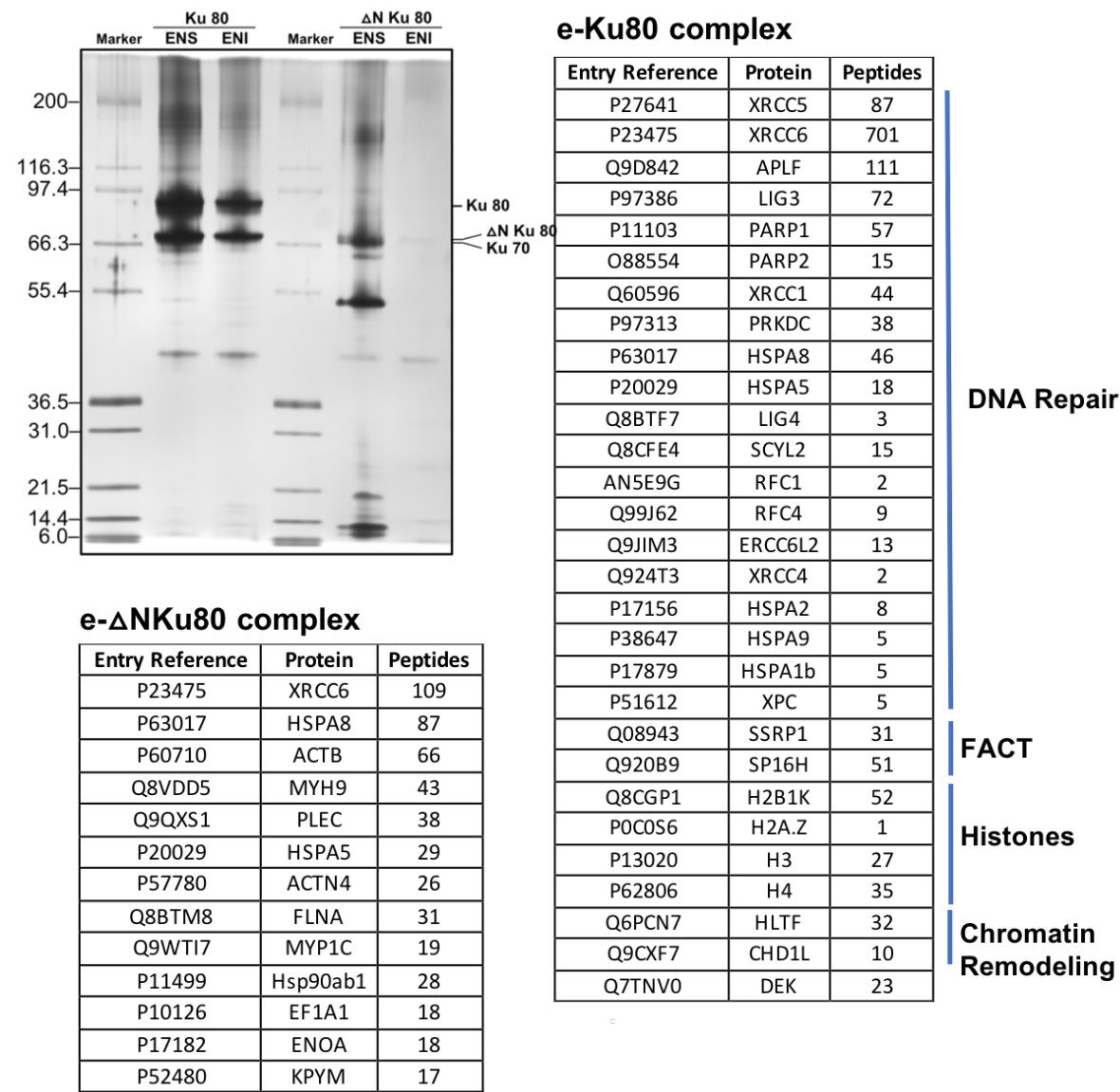
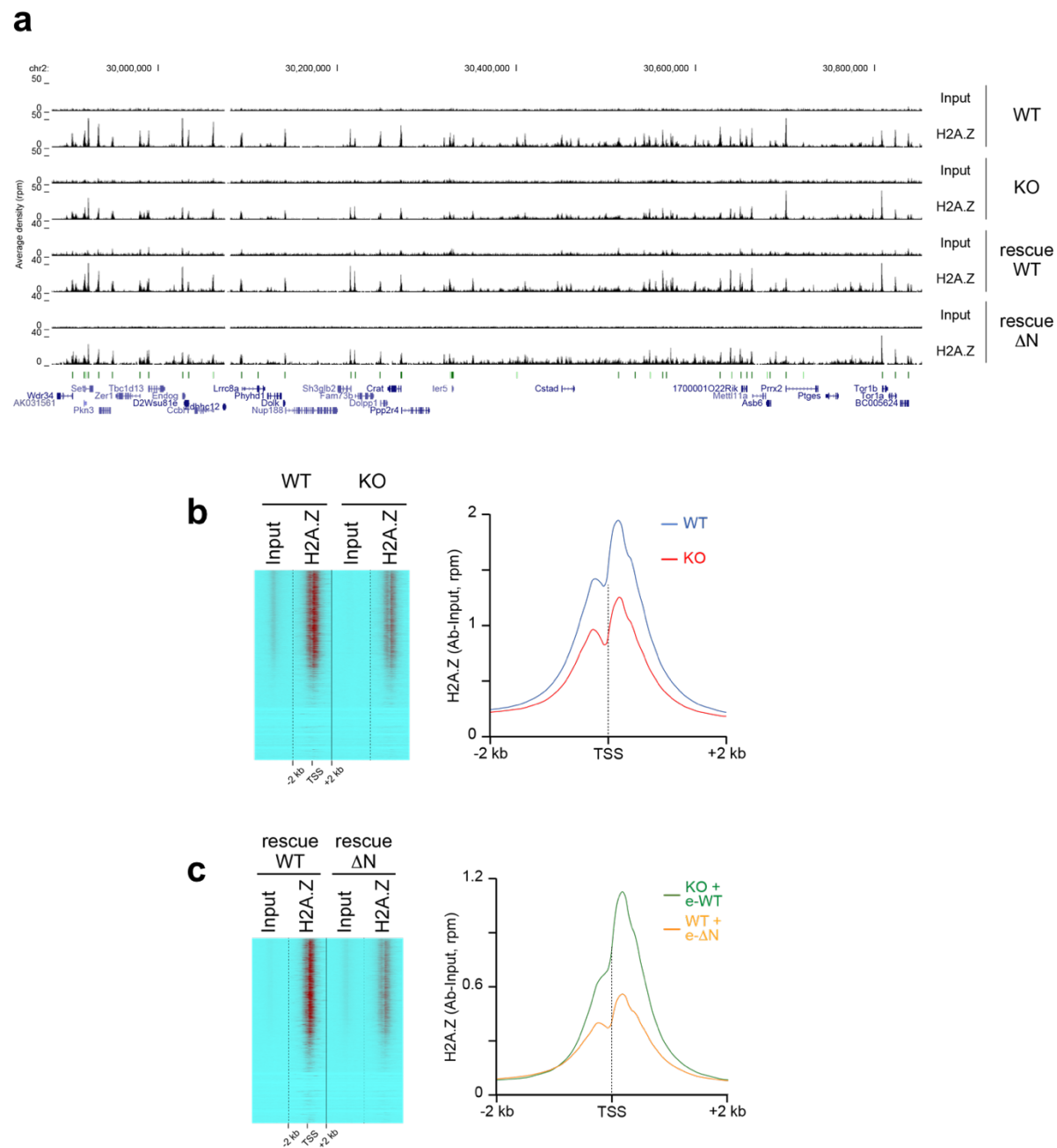
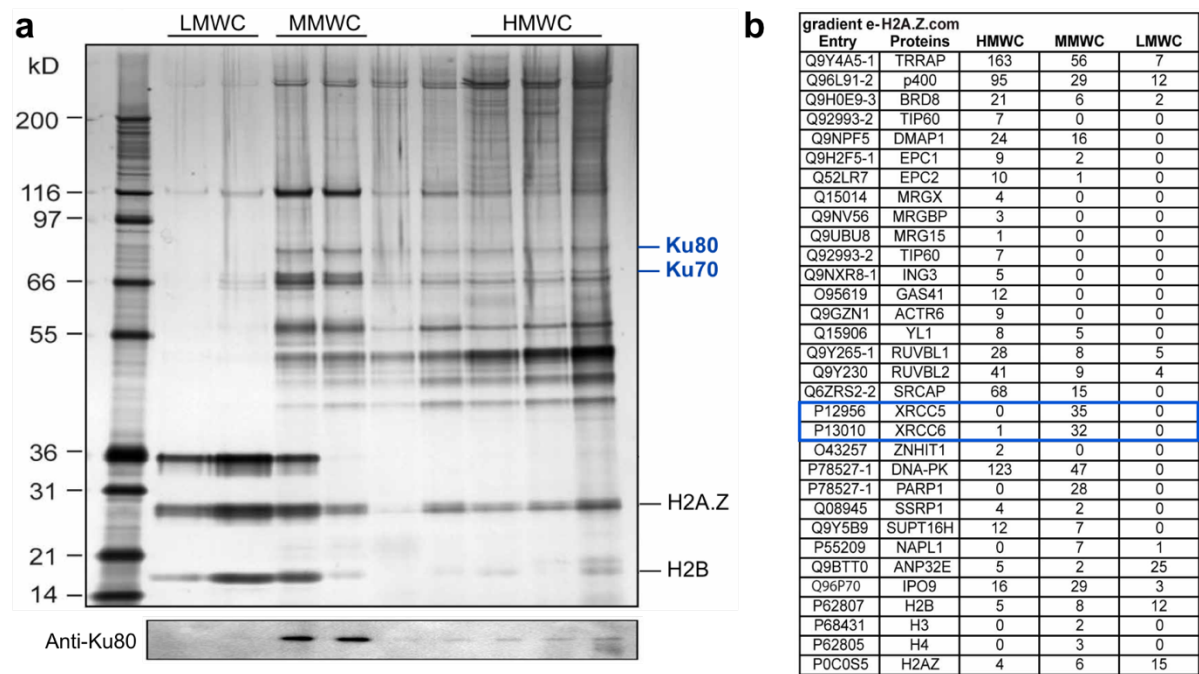


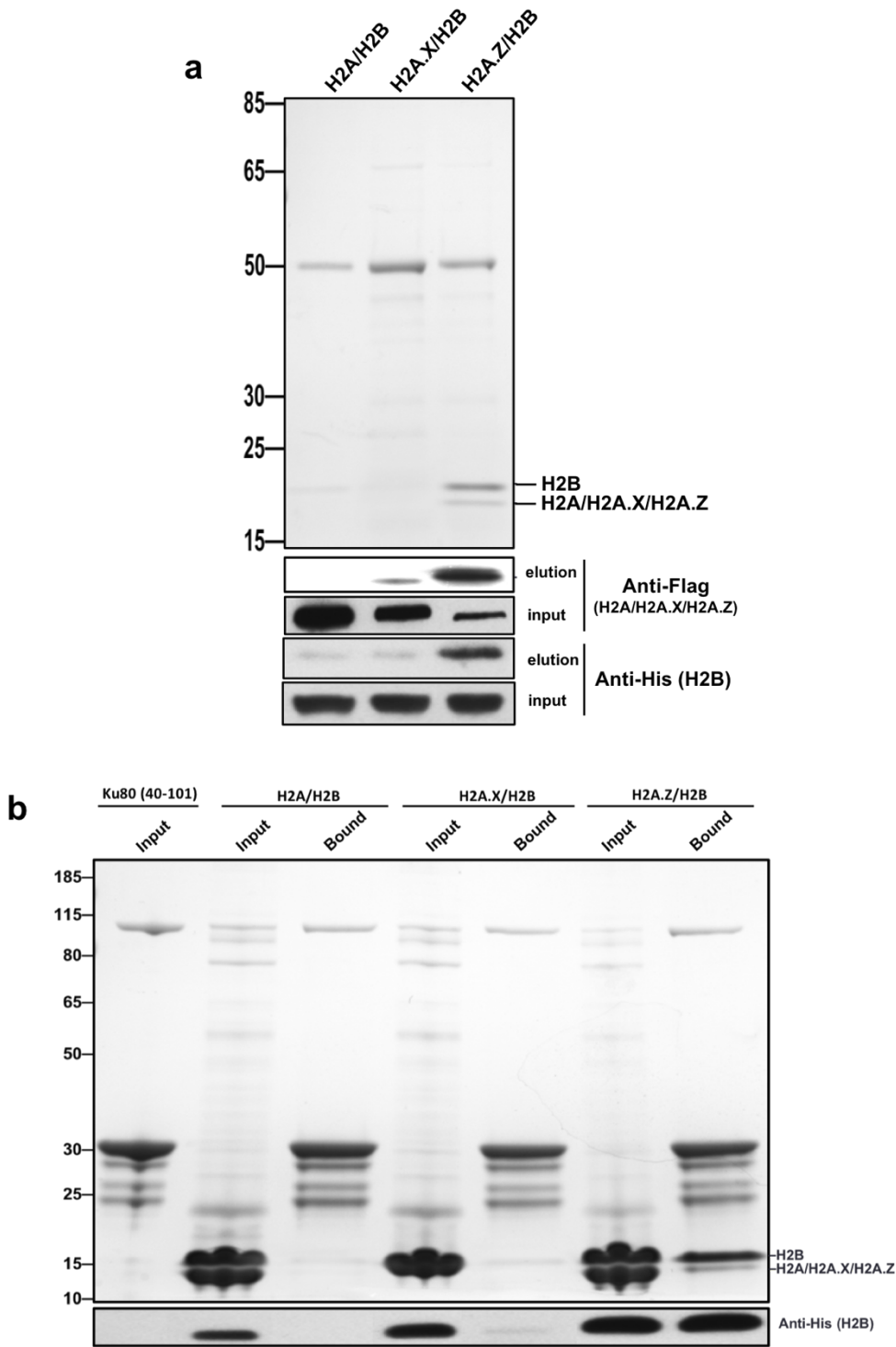
Figure 8:



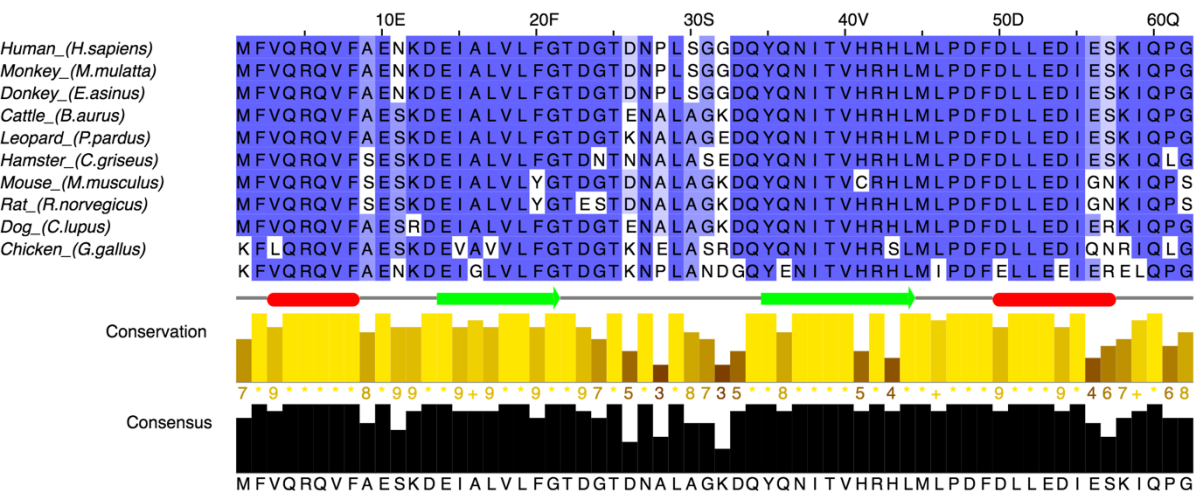
Supplementary Figure 1:



Supplementary Figure 2:



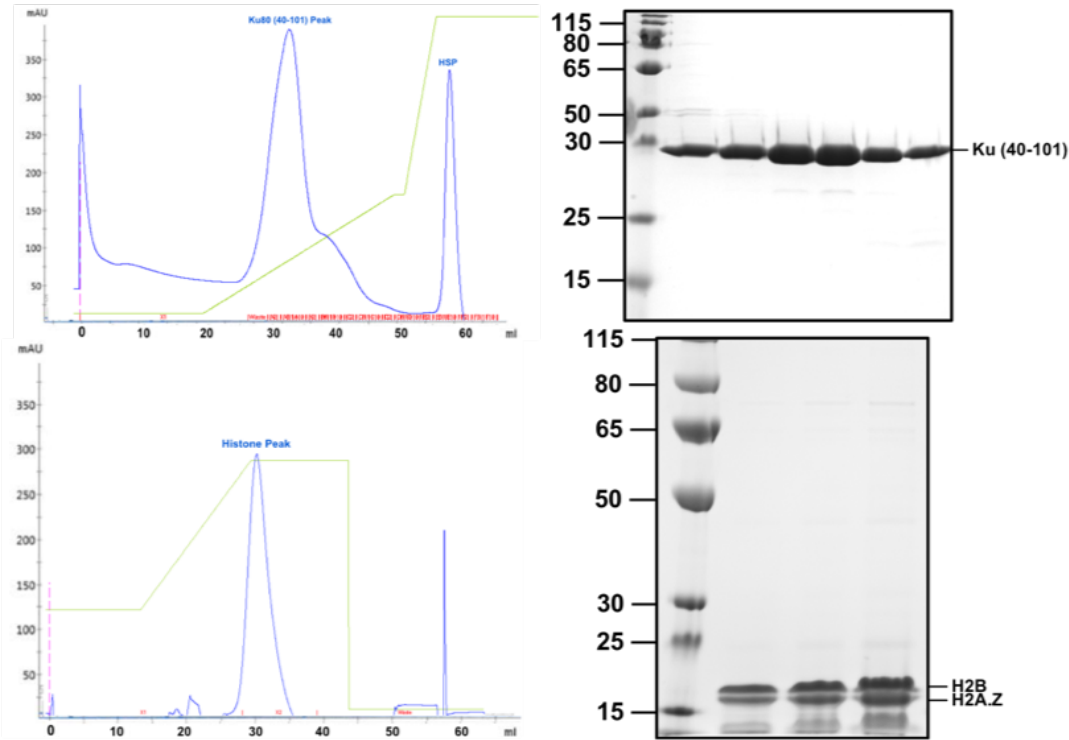
Supplementary Figure 3:



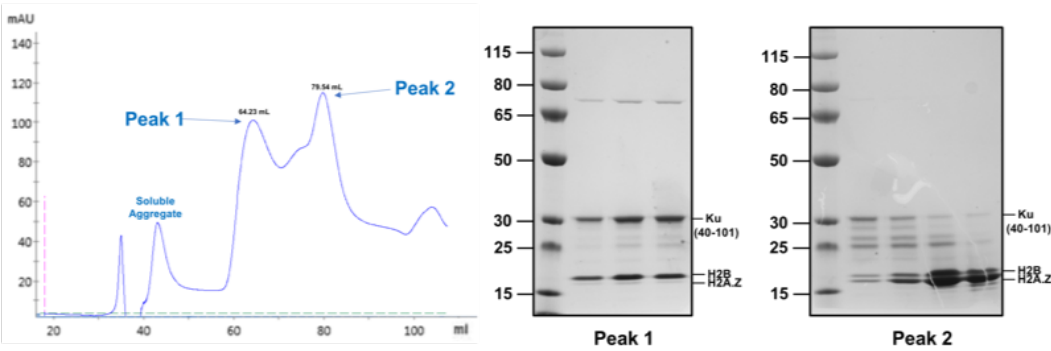
The timeline illustrates the study's progression. It begins with a baseline assessment (T0) where participants are recruited and assigned to either the intervention or control group. The intervention group receives the intervention (T1) and is then followed up (T2). The control group receives the control (T1) and is also followed up (T2). The timeline is marked with a horizontal axis and a vertical axis indicating time points.

Supplementary Figure 5:

a



b



Une nouvelle voie de déposition de H2A.Z impliquant l'hétérodimère Ku70/Ku80

Résumé

L'hétérodimère Ku, composé des sous-unités Ku70 et Ku80, est un complexe évolutivement conservé jouant un rôle important dans la réparation par NHEJ. L'hétérodimère Ku est un complexe très abondant associé à la chromatine. La façon dont les protéines Ku sont recrutées à la chromatine en l'absence de dommages à l'ADN est encore inconnue. En utilisant une procédure de purification par affinité en tandem, nous avons identifié l'hétérodimère Ku comme l'un des composants majeurs du complexe de prédéposition de H2A.Z. En utilisant la cartographie par délétion, nous avons identifié le domaine vWA de Ku80 comme étant principalement responsable de son interaction avec les dimères H2A.Z/H2B. Cette interaction est essentiellement médiée par H2B en présence de H2A.Z mais pas en présence de H2A ou H2AX. L'analyse de la localisation de H2A.Z dans des cellules Ku80^{-/-} par la technique de ChIP-seq montre une déplétion spécifique de H2A.Z au niveau des promoteurs mais pas au niveau des enhancers ou insulateurs. Collectivement, nos données révèlent une nouvelle fonction du complexe Ku dans le mécanisme de déposition de H2A.Z/H2B au niveau des promoteurs.

Mots clef : Heterodimer Ku, déposition d'histone, Histone variante, H2A.Z/H2B, chromatine.

Résumé en anglais

The Ku heterodimer, consisting of Ku70 and Ku80 subunits, is an evolutionary conserved complex with a prominent role in NHEJ repair. The Ku heterodimer is a highly abundant chromatin associated complex. How the Ku proteins are recruited to chromatin in the absence of DNA damage is still unknown. Using tandem-affinity purification procedure, we identified the Ku heterodimer as one of the major component of the H2A.Z predeposition complex. Using deletion mapping, we identified the vWA domain of Ku80 as the primary motif responsible for its interaction with H2A.Z/H2B dimers. This interaction is mediated essentially by H2B in the presence of H2A.Z but not in the presence of H2A or H2AX. Analysis of H2A.Z localization in Ku80^{-/-} cells by chromatin immunoprecipitation followed by sequencing shows specific depletion of H2A.Z from promoters but not from enhancers or insulators. Collectively, our data uncover a novel function of the Ku complex in the mechanism of H2A.Z/H2B deposition at promoters.

Key words: Ku heterodimer, Histone deposition, Histone variant, H2A.Z/H2B, chromatin.