Universität Hamburg

Universitätsklinikum Hamburg-Eppendorf

Forschungsinstitut Kinderkrebs-Zentrum Hamburg

Using Acute Promyelocytic Leukemia (APL) to Study the PML-associated Daxx/ATRX Complex and Its Role in Pediatric Malignancies

Scientific Advisor: Dr. Thomas Martin Sternsdorf

Dissertation

To achieve a Ph.D. degree from the University of Hamburg
Faculty of Mathematics, Informatics and Natural Sciences,
Department of Biology

Submitted by
M.Med. Wencong Cui
Born in Qingdao, China
Hamburg, 2020

The following evaluators recommend the admission of the dissertation:

Dr. Thomas Sternsdorf

Prof. Dr. Thomas Dobner

Date of Disputation: 03.03.2021

This work was carried out from 2017 until 2020 in the Research Institute Children's Cancer Center Hamburg, Department of Pediatric Hematology and Oncology at Universitätsklinikum Hamburg-Eppendorf (UKE) in the research group of Dr. Thomas Sternsdorf.

Dedicated to who I love and who loves me…

Table of Contents

1. Abbreviations

Table 1-1 Abbreviation list

Abbreviation	
ADD	ATRX-Dnmt3-Dnmt3L-domain
ALT	Alternative lengthening of telomeres
AML	Acute myeloid leukemia
APL	Acute promyelocytic leukemia
ATM	Ataxia-telangiectasia mutated
ATO	Arsenic Trioxide
ATRA	All-trans retinoic acid
ATRX	ATP-dependent helicase ATRX, X-linked helicase II
BCCP	Biotin carboxyl carrier protein
Bcl-2	B-cell lymphoma-2
BET	Bromodomain extra-terminal
BETi	Bromodomain inhibitors
BioID	Proximity-dependent Biotin Identification[1]
BIR	Baculovirus IAP repeat
BPL	Biotin protein ligase
BRET	Bioluminescence resonance energy transfer
CFP	Cyan fluorescent protein
CLL	Chronic lymphocytic leukemia
CML	Chronic myelogenous leukemia
Cre	Cre recombinase
Daxx	Death-associated protein 6
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
DNMT	DNA methyltransferase
ER	Estrogen receptor
FACS	Fluorescence-activated cell sorting
FRET	Förster resonance energy transfer
FRET	Fluorescence resonance energy transfer

FWHM	Full width at half maximum
GFP	Green fluorescent protein
HDAC	Histone deacetylase
HIF-1α	Hypoxia-inducible factor-1α
HIV-1	Human immunodeficiency virus type-1
HOE	Hoechst
CBX5	Chromobox protein homolog 5
HP1α	Heterochromatin protein 1 α
HPV	Human papillomavirus
IAPs	Inhibitor of apoptosis proteins
IC50	Half-maximal inhibitory concentration
IF	Immunofluorescence
IHC	Immunohistochemistry
IN	Integrase
JNK	Jun N-terminal kinase
KD.S4	Knockdown of SMARCA4
LBR	Lamin B Receptor
LEDGF	Lens endothelial growth factor
LFA-1	Leukocyte function-associated antigen-1
mPR	mouse PML-RARα (Korf 2014)
MDM2	Murine double minute 2
NBS	Nijmegen breakage Syndrome Protein, Nibrin
NMR	Nuclear magnetic resonance
OE.S4	Overexpression of SMARCA4
PCR	Polymerase chain reaction
PVC	Pellet Cell Volume
PI	Propidium iodide
PML	Promyelocytic leukemia protein
PML-NBs	Promyelocytic leukemia nuclear bodies
PPIs	Protein-protein interactions
RA	Retinoic acid
RARα	Retinoic acid receptor alpha
SDS-PAGE	Sodium dodecyl sulfate-polyacrylamide gel electrophoresis
SILAC	Stable-isotope labeling by amino acids in cell culture

SIM	SUMO-interacting motif
Smac	Second mitochondrial activator of caspases
SMARCA4	SWI/SNF-related, matrix-associated, actin-dependent regulator of chromatin, Subfamily A, Member 4
SMCHD1	Structural maintenance of chromosomes flexible hinge domain-containing protein 1
Sp100	Nuclear autoantigen Sp-100 ("Speckled, 100 kDa")
SUMO	Small ubiquitin-like modifier
TAD	N-terminal transactivation domain
TAP-MS	Tandem affinity purification-mass spectrometry
TERT	Telomerase reverse transcriptase
TRF	Telomere repeat binding factors
TXF	Tamoxifen
Uv	Ultraviolet
Y2H	Yeast two-hybrid
YFP	Yellow fluorescent protein

2. Abstract

BACKGROUND: The Histone H3 variant H3.3 has raised significant attention due to the occurrence of specific mutations in the coding genes, namely H3F3A and H3F3B, producing so-called oncohistone. The H3.3 requires two proteins to deposit it into the heterochromatin, namely Daxx and ATRX, which are also associated with nuclear structures known as PML-nuclear bodies (ND10). The structure formed by PML, Daxx, and ATRX is called the PAX (PML-associated ARTX/Daxx) complex.

RATIONALE AND METHODS: To explain the role of the PAX complex in cancer development, we have used a system in which the complex is naturally disrupted, namely the NB4 cell line. NB4 is a line derived from acute promyelocytic leukemia (APL), which, in turn, is a form of cancer caused by the expression of an anomalous protein produced by the fusion of the genes coding PML and Retinoic acid receptor alpha (RARα). Much to the advantage of PML patients, treatment of cells with all-trans retinoic acid (ATRA), a RARα ligand, effectively degrades the PML-RARα fusion protein.

Due to the tight association of the PAX complex with the insoluble nuclear matrix, we utilized the "Proximity-mediated Biotin Identification" (BioID) method to produce a stable NB4 subclone (NB4-BD) expressing a recombinant Daxx protein fused with a promiscuous biotin ligase (BioID-Daxx). After treating NB4-BD cells with ATRA, we confirmed the physiological expression levels and the physiological localization of BioID-Daxx using immunofluorescence. We then added biotin which was covalently bound by BioID-DAXX to neighboring proteins, which were then rendered soluble by treatment under denaturing conditions, purified by the use of Streptavidin-Dynabeads, and analyzed through mass-spectrometry. Parental NB4 cells were used as negative controls to subtract nonspecific background.

RESULTS: We identified SMARCA4, which is a SWI/SNF ATPase, as well as SMCHD1, which is a protein associated with DNA methylation, as novel ATRA-dependent interactors of Daxx in NB4 cells.

CONCLUSIONS: Because SMARCA4 mutations have also been found in cancer, the ATRA-dependency of its binding to Daxx and ATRX indicates functional relevance. As these proteins interact, carcinogenesis in these tumors might follow common mechanisms.

KEYWORDS: Daxx, ATRX, H3.3 Histone Chaperone, Proximity-mediated Biotin Identification (BioID), SMARCA4, and SMCHD1

3. Introduction

3.1. Epigenetics and Cancer

3.1.1. The Hallmarks of Cancer

Back in 2000, Hanahan and Weinberg[2] suggested that all cancers share six functional capabilities that are acquired during cancer development, albeit through various mechanistic pathways, and govern the transformation of normal to cancer cells. These "hallmarks" include self-sufficiency in growth signals[3]; insensitivity to anti-growth signals[4]; evasion of programmed cell death[5]; limitless replicative potential[6,7]; sustained angiogenesis[8], and tissue invasion and metastasis[9] (Figure 3-1).

Ten years later, the same researchers added two further hallmarks, namely, abnormal metabolic pathways and immune system evasion[10].

3.1.2. Definition of Epigenetics

Epigenetics is the study of heritable changes in gene expression that do not involve changes in the DNA sequence. As a term, it was first used in the modern sense by Conrad Waddington back in 1940[11]. The epigenetics field studies a multitude of gene expression regulatory mechanisms, including DNA methylation, chromatin remodeling, and microRNA-mediated regulation. Mutations in epigenetic factors affect multiple genes, while more than one epigenetic factor may affect any given gene. For these reasons, the changes in the expression of oncosuppressive genes and proto-oncogenes that are observed in cancer are often found to be the result of a malfunction of epigenetic mechanisms rather than mutations to the coding genes themselves. Thus, it comes as no surprise that research into this field is essential for understanding cancer development and progression.

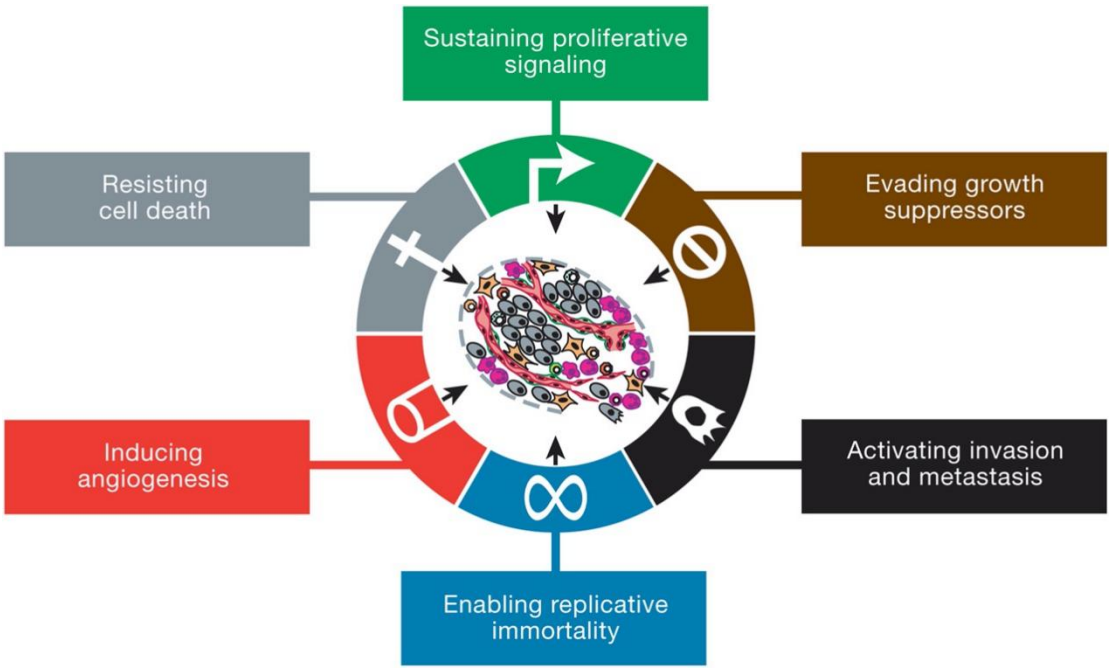


Figure 3-1 The six hallmarks of cancer. (Taken from Hanahan and Weinberg[10])

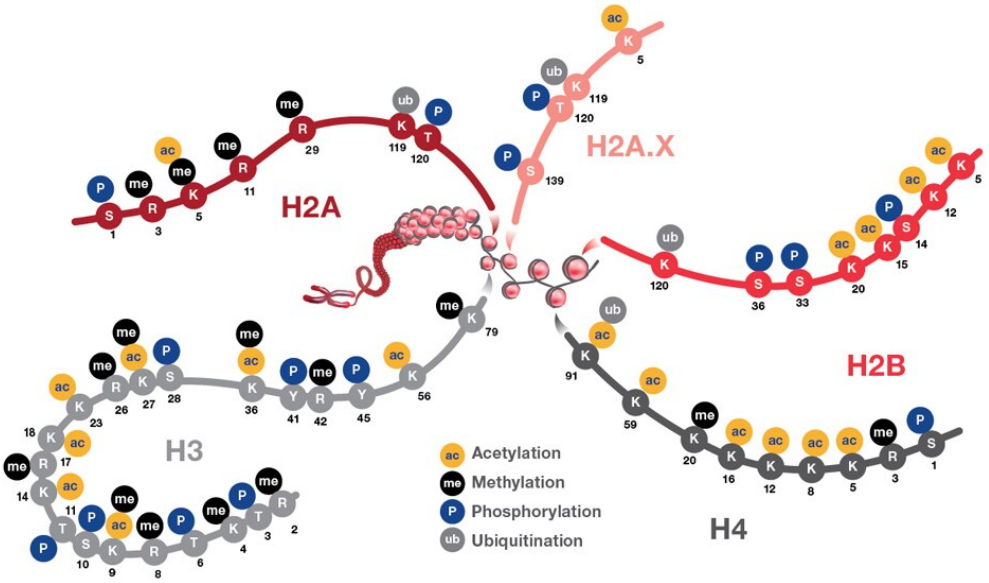


Figure 3-2 Residues of histones that can undergo one or more of the four most common modifications, namely, acetylation, methylation, phosphorylation, and ubiquitination. These histone modifications make up what is known as the histone code, which dictates the transcriptional state of the local genomic region, thus playing a crucial role in gene expression. (Taken from Thermo Fisher Scientific-Epigenetics)

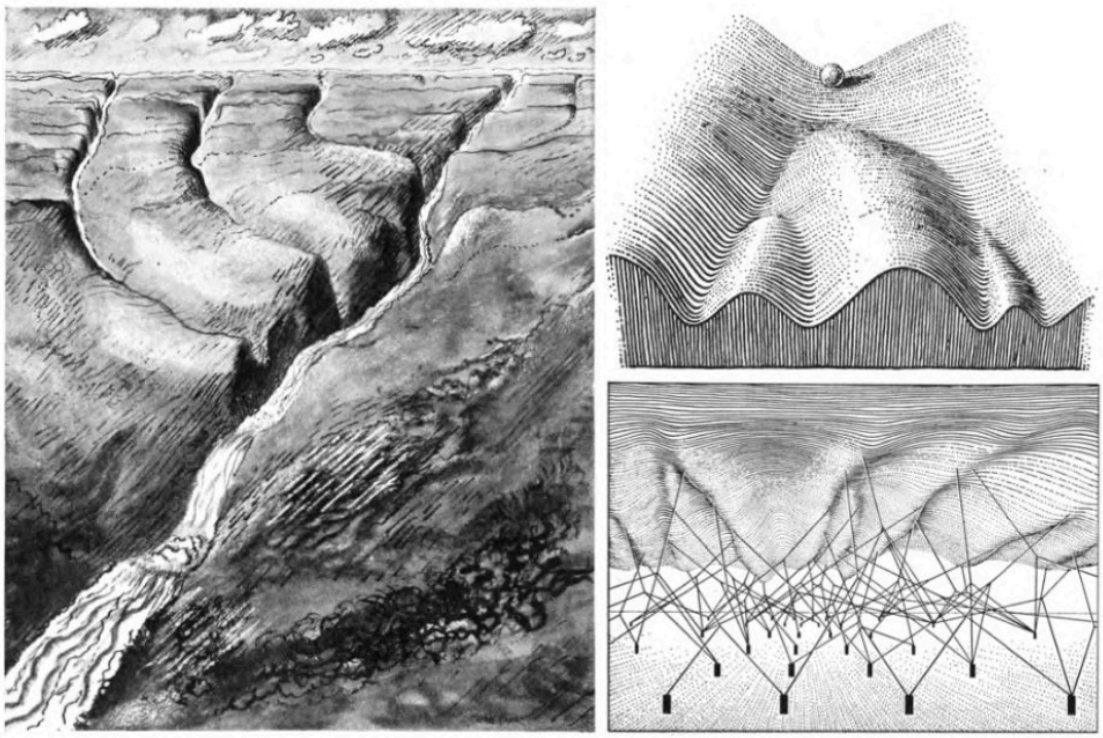


Figure 3-3 Waddington's epigenetic landscape. Left: drawing by John Piper for Waddington's "Organisers & Genes" (1940, Cambridge: Cambridge University Press). The original caption reads: *"Looking down the valley towards the sea. As the river flows away into the mountains it passes a hanging valley and then two branch valleys, on its left bank. In the distances the sides of the valleys are steeper and more canyon-like."* Right: Waddington's classical depictions of his landscape from "Strategy of the Genes" (1957, London: Allen & Unwin). Top right: top view of the landscape where the path followed by the ball represents a developmental trajectory. Bottom right: by depicting the underside of the surface, Waddington illustrated the idea that genes can change the landscape during evolution. (Taken from Gemma Anderson et al.[12])

3.1.3. DNA methylation in cancer

DNA methylation is a mechanism that typically represses transcription. In mammals, it predominantly occurs at cytosines that are located 5' to guanosine, i.e., in CpG dinucleotides[13]. The pattern of DNA methylation is heritable thanks to the existence of a "methylation maintenance" system that acts on the hemimethylated DNA double-strands that are produced due to the semi-conservative nature of DNA replication[14]. Many cancers display anomalous methylation patterns[15]. In 1983, Gama-Rosa et al.[16] were the first to show a global decrease in DNA methylation in tumor samples. In the same year, Feinberg and Vogelstein[17,18] were the first to report methylation changes in specific genes in human cancers, most notably the hypomethylation of the *RAS* oncogene. Since then, thousands of studies have been published linking DNA methylation and cancer. Changes in DNA methylation may play a role in any step of tumor progression. Hypomethylation in the promoters of oncogenes and hypermethylation in the promoters of tumor-suppressive genes are the best known, but not the only, forms of aberrant methylation associated with tumor progression[19].

3.1.4. Histone modifications

Allfrey et al.[20] were the first to demonstrate that histones are post-translationally modified. According to a hypothesis proposed by Strahl and Allis[21] back in 2000, histone modifications form a "Histone Code"[21], serving as epigenetic marks that are recognized by other proteins, leading to transcriptional activation or repression[22]. Today, it is clear that histone modifications play fundamental roles in most biological processes that are involved in the manipulation and expression of DNA[23]. There are at least twenty different types of histone modification, of which acetylation, methylation, phosphorylation, and ubiquitination are the best understood[24].

For each type of modification, there are specific sets of enzymes to add or remove the corresponding chemical group (which ranges from simple methyl groups to small proteins) to/from the targeted amino acid residues. These enzymes are often referred to as "writers" and "erasers", respectively. They play important roles in development, and their mutation or misregulation may lead to disorders and cancers. The proteins that recognize the epigenetic marks ("readers") have domains that allow their recruitment to the chromatin. The same protein may simultaneously act as a reader and a writer, allowing the reinforcement of modifications through a positive feedback loop or the inhibition of their activity by other modifications[25].

3.1.4.1. Histone Acetylation

Histone acetylation is not only one of the best-studied modifications but also the first to be linked with transcriptional activity[26,27]. A large number of human cancers display anomalous acetylation levels. In particular, the loss of H4K16 acetylation is considered a hallmark of human cancer[28]. Single-allele somatic mutations of genes encoding histone acetyltransferases such as p300 or CBP are found in various cancer types, including colorectal[29] and breast cancer[30].

As Acute promyelocytic leukemia (APL) is among the focuses of the current thesis, the work of Lin et al.[31] on APL is highly relevant to it. These researchers demonstrated that two of the RAR-containing fusion proteins characteristic of the disease, namely, PML-RAR and PLZF-RAR, use their RAR parts to bind to retinoic acid-responsive elements (RAREs) of genes blocking the differentiation of myeloid cells[31,32]. This binding results in the recruitment of histone deacetylase (HDAC) repressor complexes with a high affinity and the subsequent silencing of these genes.

3.1.4.2. Histone Methylation

Multiple methylation states on both lysine and arginine residues have been identified on histone tails[33]. Methylation represents a complex and more stable chromatin modification than acetylation. Moreover, unlike acetylation, the addition of methyl groups does not affect the electric charge, thus methylated residues retain their fundamental basic character[23]. Each type of histone methylation is regulated by different methyltransferase writers (histone methyltransferases) and eraser enzymes (histone demethylases), while specific methyl marks have been found to be critical for gene expression, cell fate determination, and genomic stability[34,35]. Busslinger and Tarakhovsky[36] suggested that aberrant changes of histone methylation levels, both increases and decreases, may play a role in tumor development.

EZH2 is the catalytic part of the Polycomb repressive complex 2 (PRC2), which is responsible for the di- and trimethylation of H3K27[37]. Khan et al.[38] suggested that loss of gene repression caused by EZH2 mutation results in reduced H3K27 trimethylation and, subsequently, increased chromatin relaxation at specific gene loci accompanied by higher transcriptional activity and may contribute to leukemogenesis. Several groups showed that H3F3A K27M mutations result in lowered/absent H3K27me3, suggesting that the EZH2-mediated trimethylation of H3K27 is inhibited in glioblastomas harboring K27M mutations[39,40].

H3K36-specific methyltransferase writers play a role in a variety of cancers[34]. Mutations were found in high-grade gliomas[41], renal cell carcinoma[42], AML, myelomas, and lung cancers[43]. Overexpression was reported in Chronic myelogenous leukemia (CML), bladder, lung, and liver cancers[44].

3.2. Acute Promyelocytic Leukemia

This study takes advantage of the unique biology of an otherwise intensively studied disease, namely Acute promyelocytic leukemia (APL), which is an aggressive type of Acute myeloid leukemia (AML) characterized by the malignant expansion of promyelocytes in the blood and bone marrow. APL represents 10-12% of AML cases[45]. It was first identified as a clinical entity in 1957[46] and was described as a hyperacute fatal illness with an average survival time of less than one week. Today, thanks to the development of targeted therapy for APL using a combination of all-trans retinoic acid (ATRA) and arsenic trioxide (ATO), 10-year survival rates are estimated to be as high as 80-90%[47]. As a result, APL nowadays is considered a "curable disease" and is mostly utilized as a model[48]. It is in this capacity that it is used in the current study. Specifically, we use the NB4 cell line, which was isolated from an APL patient back in 1989.

With respect to its pathology, APL is distinguished from other subtypes of AML by the presence of a Promyelocytic leukemia protein (PML)-Retinoic acid receptor alpha (RARα) fusion transcript[49], which results from a reciprocal t(15;17) chromosomal translocation[50]. The N- and C-terminal parts of the PML-RAR fusion protein are formed by the PML and RARα moieties, respectively. The PML protein was actually so named because it was first described as a part of this APL-characteristic fusion protein. Notably, it was the search for the origin of the fusion gene that led to the discovery of the *PML* gene on chromosome 15[51]. Even though treatment with either agent results in the degradation of the fusion protein[52], their respective targets on PML-RARα are different. ATRA targets the RARα moiety of the fusion protein[53]; unlike other forms of chemotherapy, ATRA does not directly kill the malignant cells, but induces their differentiation, after which they spontaneously undergo apoptosis[54]. Conversely, ATO targets Cysteines of the PML moiety and directly induces apoptosis[51].

3.3. Promyelocytic Leukemia Protein-Nuclear Bodies (PML-NBs)

APL did not only lead to the discovery of PML[51] but, with it, of a novel subnuclear structure[55–57] now known as ND10[57] or Promyelocytic leukemia protein nuclear bodies (PML-NBs). They are spheres 0.1-1.0 μm in diameter and are found in virtually all cell lines and tissues examined to date[58]. Their main component is PML, after which they are named. Besides PML, numerous cellular and viral proteins have been identified as their components, often by analysis of proteins associating with PML in PML-overexpressing systems. However, as PML can aggregate with many cellular components, its forced overexpression may have resulted in artifacts, notably by titration of partners or aggregation of insoluble proteins[58].

Many PML-NB proteins have been shown to be the target of post-translational modification by SUMO, namely SUMOylation, in which conjugation by the SUMO protein plays a critical role in the recruitment of partners. Overexpression of SUMO-1 prevents the stress-induced disruption of PML-NBs, showing SUMO-1 as a key regulator of their integrity[59] and the PML RING domain as critical for PML SUMOylation and PML-NB formation[60].

In APL cells, PML-NBs are disrupted, resulting in a characteristic microparticulate pattern, while retinoic acid treatment restores them[61]. Koken et al.[56], Daniel et al.[62] and Weis et al.[63] suggested that the presence of the PML-RARα fusion protein directly disrupts the localization of PML in the PML-NBs.

Regarding the experimental model, NB4 cell line. At first glance, it may seem paradoxical to study the PAX complex in a system that actually lacks PML-NBs. However, as mentioned above, treatment of these cells with ATRA leads to the reconstitution of PML-NBs. This reconstitution becomes evident when the cells are stained for PML, ATRX, and Daxx. While control cells show a dispersed localization

motif for all three proteins, ATRA-treated cells display colocalization of the three proteins in the reconstituted PML-NBs (Figure 3-4). This dependence of ATRX-Daxx association in NB4 cells on ATRA offers the unique opportunity to filter out interactions that appear even in parental cells and are thus irrelevant to PML-NBs and, by extension, to the PAX complex.

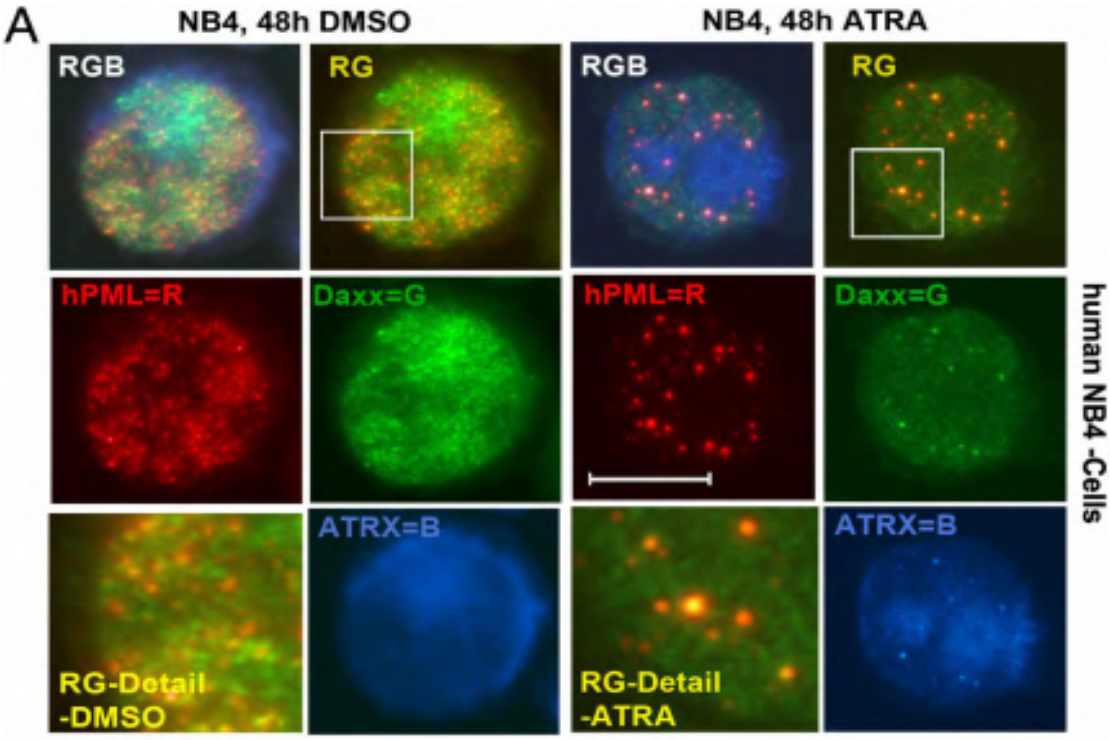


Figure 3-4 The PAX complex is disrupted by PR expression. NB4 cells were treated with 1 µM ATRA for 48 h. Cytospins were stained with the indicated antibodies. hPML is stained red, Daxx in green, and ATRX in blue. After ATRA treatment, all three antigens colocalize at the re-formed PML bodies. The detail is given of the RG merge (white boxes) in the lowest Left panel. This colocalization indicates a re-formation of the PML-NBs complex after ATRA-induced degradation of PML-RARα. (Scale bar, 5 µm.) (Taken from Korf et al.[64])

3.4. The Daxx/ATRX Chaperone

3.4.1. Histone variants

Histones are the most conserved proteins between species[65]. Besides the canonical core histones (H1, H2A, H2B, H3, and H4) numerous variant histones exist[66], which carry out different tissue-specific functions[67–69] and play specific roles in development[70,71], embryogenesis[72–74], cell fate decision[75,76], and carcinogenesis[77,78]. Even though all core histones have variant counterparts[79] (Figure 3-5), the best-studied among them are H3 variants[80].

This study will focus on H3.3. Unlike the canonical H3 variants, namely, H3.1 and H3.2, which are replication-dependent, H3.3 is DNA replication-independent, i.e., it is synthesized throughout the cell cycle[81]. H3.3 plays numerous vital roles in diverse processes, such as cell differentiation[82,83] and mammalian development[22]. As shown in Figure 3-6, H3.3 deposition into heterochromatin and transcriptionally active chromatin is catalyzed by different respective chaperones, namely, Daxx/ATRX[84,85] and HIRA[86].

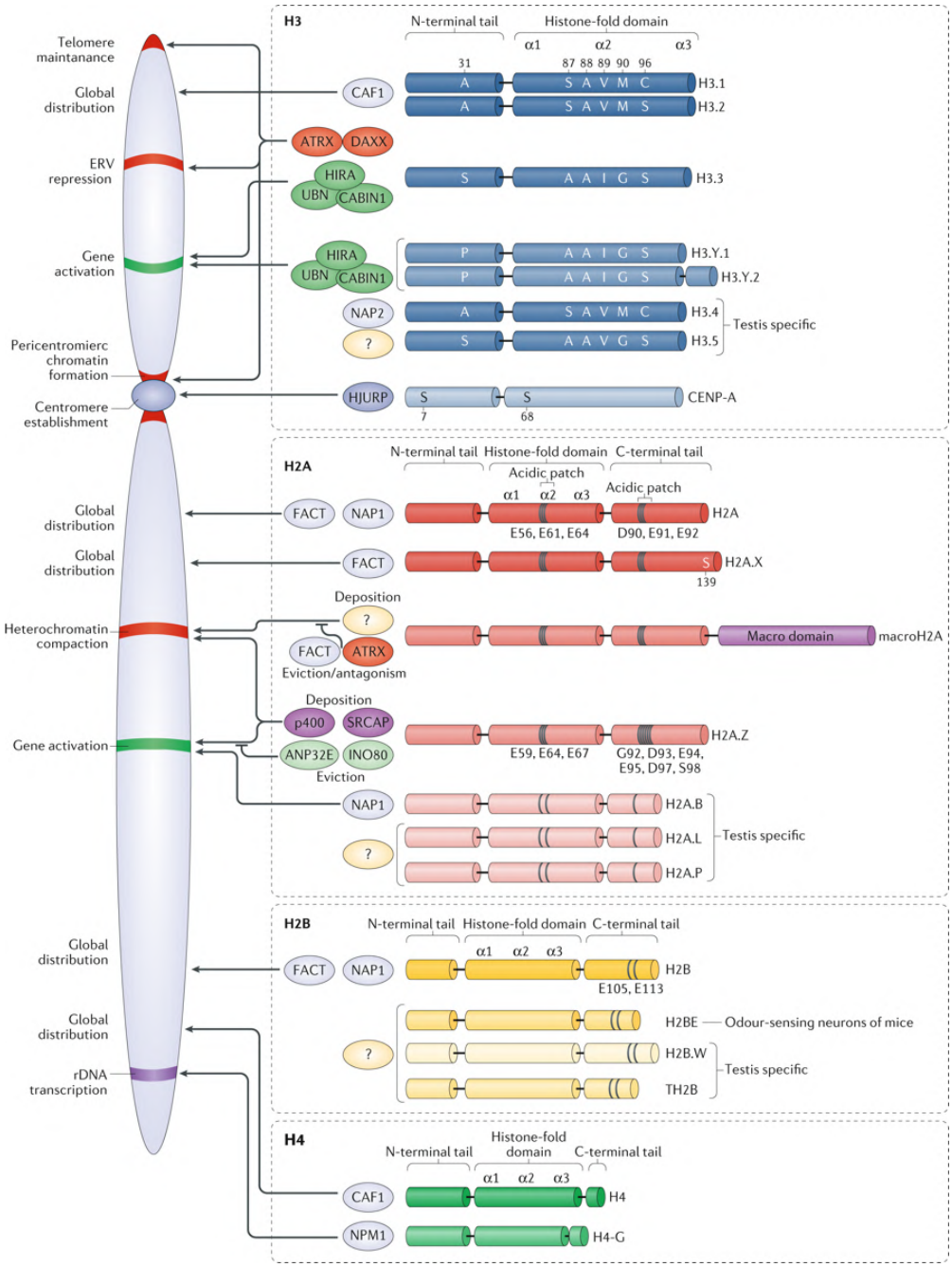


Figure 3-5 Core histones, their variants, and associated chaperone/remodeler enzymes. Replication-coupled and variant core histones with respective chaperones, genomic distribution and functional output are shown. Specific amino acid residues are illustrated at either key differences among members of a common histone protein family (for example, histone H3.3 Ser31 vs H3.1 Ala31) or at well- established histone variant-specific post-translational modifications. Different shades of color in the histone's structures are used to indicate differences within the domains compared with the replication-coupled histone. (Taken from Martire and Banaszynski[79])

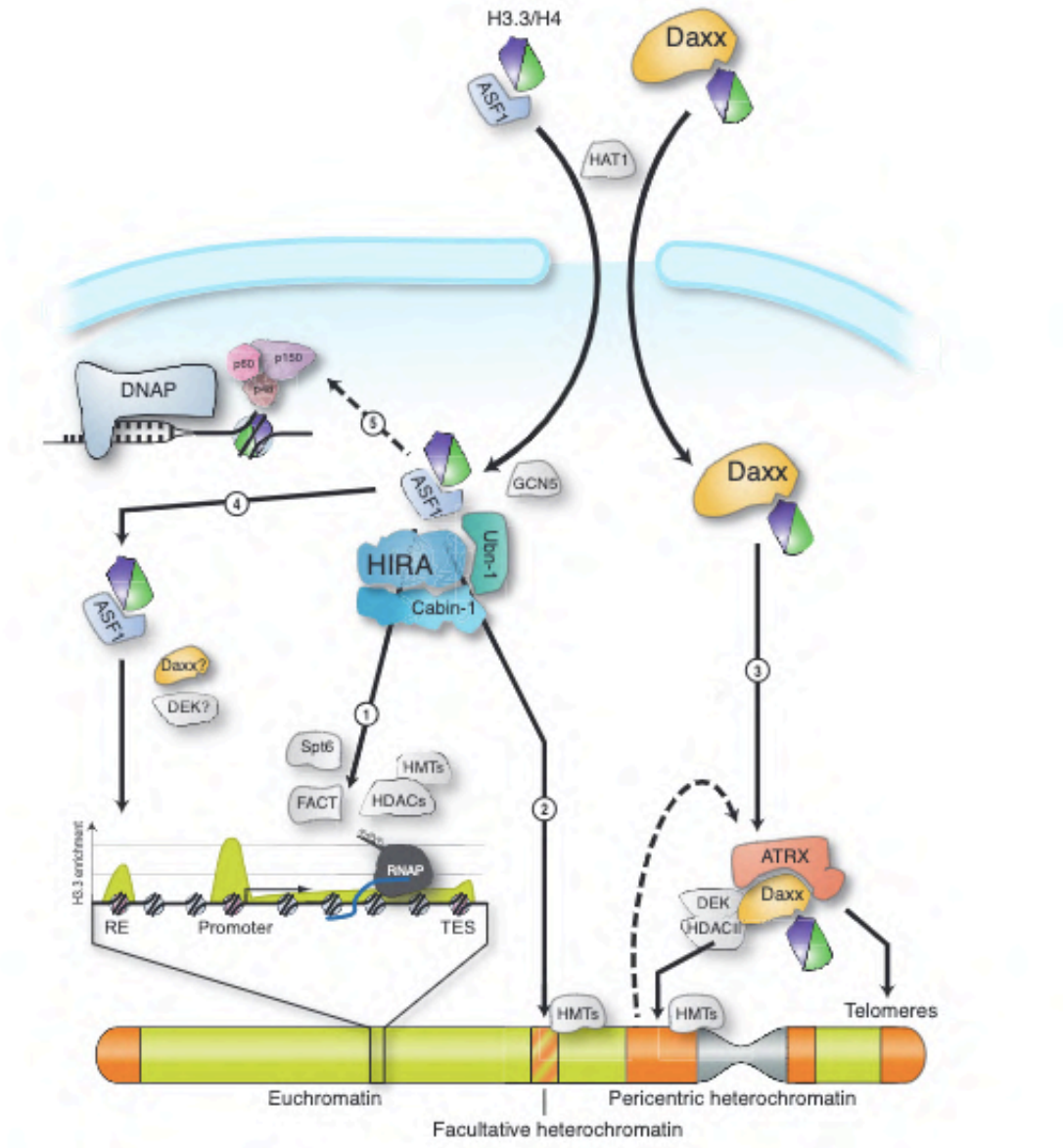


Figure 3-6 Predisposition complexes mediate H3.3 chromatin assembly. A model for the flow of histones from their synthesis in the cytosol (top) to their target genomic regions in the nucleus (bottom). Newly synthesized histones H3.3/H4 are chaperoned by ASF1 and a cytosolic pool of Daxx. HAT1 acetylates H4K5/K12 in the cytosol. Predisposition complexes HIRA.com and Daxx.com chaperone distinct pools of H3.3/H4. (1) HIRA.com mediates deposition of H3.3 at promoters and transcribed regions together with RNA polymerase (RNAP)-associated FACT (Spt16/SSRP1) and Spt6 histone chaperones. (2) HIRA.com deposits histones at facultative heterochromatin in senescent cells. Histone methyltransferases (HMTs) establish a repressive chromatin state. (3) Daxx.com cooperates with ATRX and

possibly other factors in incorporating H3.3 into pericentric heterochromatin and telomeres. The chromatin-bound deposition complex might also contain the histone chaperone Dek and the histone deacetylase HDACII (4). An unknown factor delivers H3.3 to regulatory elements (REs). (5) In the absence of Daxx.com or HIRA.com, H3.3 associates with the replication-dependent CAF-1 complex that mainly deposits histones after the DNA polymerase (DNAP) replication fork. (Taken from Elsaesser and Allis[87])

3.4.2. Daxx and its Functions

3.4.2.1. Daxx-regulated Processes

The domain architecture of Daxx is shown in Figure 3-7 B. Since its discovery, Daxx has been linked to numerous, sometimes contradictory, functions. These range from being a pro-apoptotic protein binding to the cytoplasmic domain of Fas[88], a nuclear transcriptional corepressor[89–91], a regulator of telomere maintenance[92], and a mediator of transcription activating effects[93–95].

Daxx is important for cell survival and essential for embryonic development[96,97]. Daxx overexpression protects cells from apoptosis[98], while Daxx deletion or silencing sensitizes cells to apoptosis induction[99] and promotes DNA damage[100]. Moreover, Daxx plays a crucial role in adenoviral oncogenic transformation by SUMO-1-conjugated E1B-55K-mediated degradation[101]. Finally, it may also play an antiviral role, as it associates with HIV-1-derived lentiviral DNA via interacting with HIV-1 integrase. This association results in the recruitment of HDACs to viral DNA and the subsequent repression of lentiviral gene expression[102].

SUMOylation, i.e., the covalent attachment of the SUMO protein (small ubiquitin-like modifier), is a reversible post-translational modification that has emerged as a crucial molecular regulatory mechanism[103]. Disorders in the SUMOylation pattern can lead to the development of diseases, including cancer. Interestingly, Daxx is both a SUMO binding partner and a target of SUMOylation. A SUMO-interacting motif (SIM) within Daxx had been identified[93], which has been shown to be crucial for the targeting Daxx to PML oncogenic domains, the transrepression of several SUMOylated transcription factors, and the SUMOylation of Daxx itself [93,104].

3.4.2.2. Daxx in cancer

Markedly increased expression of Daxx has been reported in a variety of cancers, with metastatic cancers displaying higher Daxx levels compared to primary tumors in breast, prostate, and colon cancer[105]. Daxx overexpression enhanced tumorigenesis capability in vivo[106], whereas Daxx depletion inhibited tumor development[100].

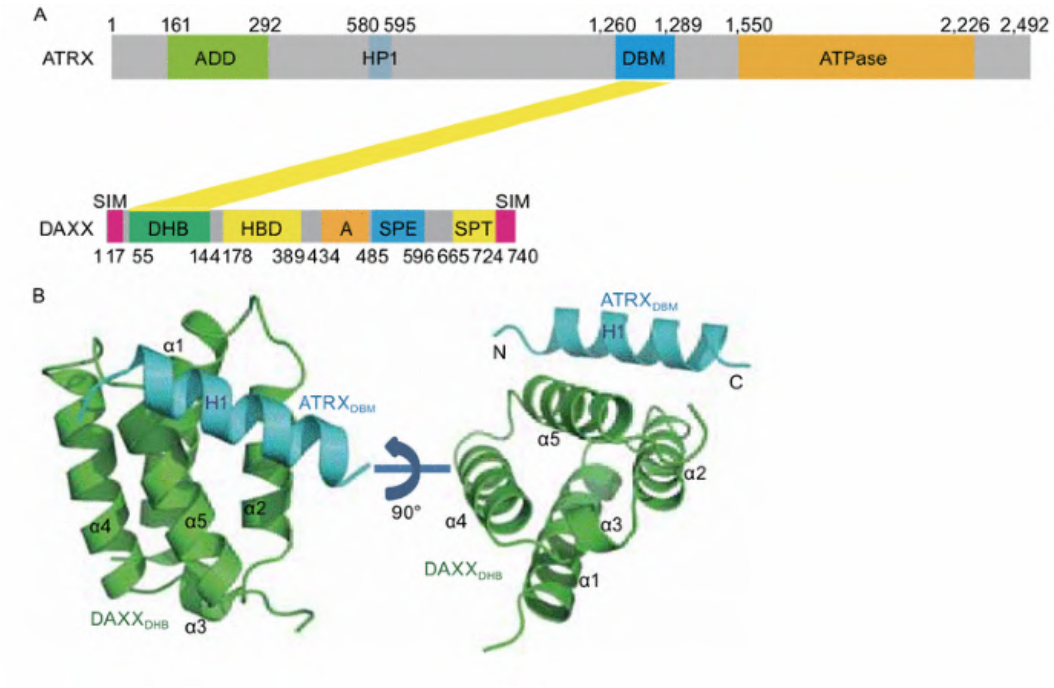


Figure 3-7 Domain architecture (A) and secondary structure (B) of Daxx and ATRX. (A) Domain organization of the ATRX and DAXX. ADD, Domain found in ATRX, Dnmt3 (DNA methyltransferase)-Dnmt3L (DNAmethyltransferase3-like protein); HP1, HP1-binding motif; DBM, DAXX binding motif; ATPase, ATPase domain; SIM, Sumo-interaction motif; DHB, DAXX helical bundle; HBD, histone binding domain; Acidic, segment rich in Glu/Asp residues; SPE, segment rich in Ser/Pro/Glu residues; SPT, segment rich in Ser/Pro/Thr residues. (B) Two orthogonal views of the DAXXDHB-ATRXDBM complex. DAXXDHB is colored in green and ATRXDBM is colored in cyan. (Taken from Wang et al.[107])

3.4.3. ATRX and its functions

3.4.3.1. ATRX-regulated Processes

The *ATRX* gene is located on the mammalian X-chromosome and its protein ATRX is a member of the SNF2 subgroup of proteins that contain ATPase activity[108,109]. ATRX may act as a regulator of gene expression[110]. The domain architecture of the ATRX protein is shown in Figure 3-7 A.

Recent discoveries have made ATRX a focus of increased interest within the fields of telomere and chromatin biology, development, and cancer research[111].

ATRX directly interacts with Heterochromatin protein 1 α (HP1α)[112], which is a reader of the H3K9me3 heterochromatin mark[113]. ATRX can also directly bind to DNA, with its binding promoted by the presence of H3K9me3 and inhibited by H3K4me3[114–116].

ATRX is required for centromere stability and the epigenetic control of heterochromatin function during mitosis and the first meiosis division[117]. It is highly expressed in both human and newborn mice; ATRX-null embryonic mice accumulate replicative damage at telomeres[118]. With respect to specific organs, various studies have shown that ATRX levels directly affect the development of various brain areas, suggesting it may be critical for brain development. Loss of ATRX leads to the apoptosis of neuronal progenitors[119,120], whereas ATRX overexpression also leads to various neurodevelopmental defects[121].

3.4.3.2. ATRX Mutations and Cancer

Loss-of-function ATRX missense mutations are known to cluster in the two ADD domains and the ATPase main functional domain (Figure 3-7)[122]. Besides compromising the function of the protein per se, ATRX mutations may hinder the recruitment of the protein to PML-NBs. This is the case in the alpha-thalassemia mental retardation syndrome, after which the protein was named[123].

Gibbons et al.[124] found that ATRX somatic mutations cause severe hematological phenotypes. Moreover, low ATRX expression was found to associate with poor-prognosis AML patients[125]. Besides leukemia, low ATRX expression was also observed in melanoma[126]. Lots of somatic mutations in *Daxx* and *ATRX* were identified in pancreatic neuroendocrine tumors; interestingly, no tumor sample was found to be mutated in both *Daxx* and *ATRX*, which suggests that they function in the same pathway[127,128], in which they play an important role in keeping genomic stability[129]. In cancers of the central nervous system, *Daxx* mutations have not been identified, unlike *ATRX*[130] and *H3.3*[131] mutations. In 20% of pediatric glioblastoma multiforme, ATRX alterations clustered in the ADD domain of ATRX, while in adult gliomas mutations were distributed evenly[131,132].

To maintain telomere length, immortal cells and cancer cells frequently express the enzyme telomerase with its Reverse Transcriptase subunit (TERT)[133]. Telomerase-negative cancer cells usually exhibit mutations in the ATRX gene[134], which are capable of maintaining telomere length through a process known as the alternative lengthening of telomeres (ALT). *TERT* mutations promoting the expression of telomerase seem to be mutually exclusive with *ATRX* mutations, which suggests that mutations in these two genes have equivalent growth advantages[135].

3.4.4. PML-Associated Daxx/ATRX (PAX) Complex

Daxx and ATRX have been shown to colocalize and form a complex[108] at PML bodies[109], which is called the PML-associated Daxx/ATRX (PAX) complex.

Concerning its localization, the field has meanwhile adopted the view of Daxx being a predominantly nuclear protein, which fits our observations. In recent years a possibly unifying concept has gained more and more attention, describing Daxx as part of a chromatin-regulating complex[136].

Daxx, together with its interaction partner ATRX[84,136], acts as an epigenetic regulator by serving as a "Histone Chaperone" that deposits the histone variant H3.3 into specific regions of chromatin in a PML-dependent manner[137] (Figure 3-8), where H3.3 has been shown to be crucial for genome integrity[22]. Also, H3.3 can stabilize Daxx protein levels and also can affect Daxx-regulated gene expression without incorporation into nucleosomes[138]. Therefore, increased H3.3 levels in cancer cells may augment Daxx oncogenic function through protein stabilization[105]. As this observation stems from multiple lines of evidence, is being adopted by numerous laboratories around the world, and has the potential to unify all previous observations about Daxx, we built our project focusing on this aspect of Daxx.

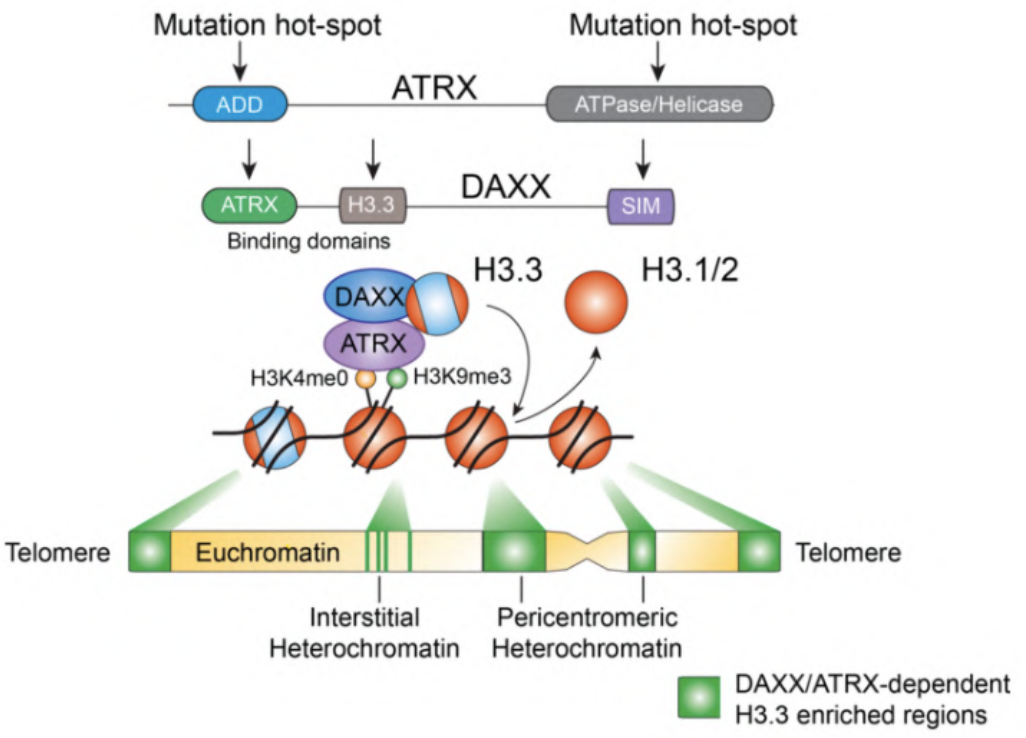


Figure 3-8 H3.3 is enriched at repetitive regions, where it is deposited by the Daxx/ATRX complex. The chromatin remodeler ATRX recognizes histone-modification states and modulates chromatin dynamics in vivo. The ATRX protein contains a N-terminal ADD and C-terminal ATPase domain, which function in the regulation of its binding and enzymatic activities, respectively. The ATRX ADD recognizes multiple histone marks, includingH3K4me0 and H3K9me3. Daxx, an H3.3 specificity factor, and ATRX complex together as a chaperone system specifically deposit the replication-independent histone variant H3.3 at telomeric and pericentromeric heterochromatic regions of the genome. Note: mutations in ATRX patients map to both the ATRX-DNMT3-DNMT3L (ADD) and catalytic ATPase/helicase domains. (Taken from https://www.banaszynskilab.org/projects and adapted from Maze et al.[139])

3.5. BioID

In our experiment, we performed the BioID method for the study of PPIs of the PML-associated Daxx/ATRX complex. The BioID method started form the Chen et al.[140] as a specific, robust and rapid method for labeling proteins with *Escherichia coli* enzyme biotin ligase (BirA), which ligates biotin to a sequence-specific 15-amino-acid acceptor peptide in an ATP-dependent manner. Slavoff et al.[141] obtained and purified the biotin ligases from nine different species and screened their ligate ability onto the human biotin acceptor domain. Biotin associates with streptavidin with extremely high affinity which allows to perform purification under very stringent conditions[142]. Roux et al.[143] have developed a new technique named BioID for proximity-dependent biotin identification, which induce lysine biotinylation on proteins neighboring the BirA-fused target protein, enabling PPI identification in live cells.

Few years later, an improved system (BioID2), using the smaller Biotin Ligase from Aquifex aeolicus, was developed[143] by the same group. As this system has proven to be superior in several aspects, we used BioID2 for our study.

4. Goal

Our main goal was the analysis and characterization of the PML-associated Daxx/ATRX (PAX) complex, which emerged as a continuation of the discovery of a PML-specific function in PML-RAR oncogenicity[64]. In particular, we decided to analyze the composition of this complex and its interaction with other proteins within the PML-NBs. This should allow a better understanding of its role in adult and pediatric cancers and possibly result in the identification of novel therapeutic targets.

The biggest challenges in this task were the difficulty in isolating the PAX-associated proteins from PML-NBs under mild lysis conditions and the effectiveness of background removal. The choice of the experimental model and the methods was made with this in mind.

With respect to the methods used to identify association partners of ATRX-Daxx, pilot proteomics-based studies indicated that conventional antibody-based affinity purification is not feasible because of the tight insoluble nuclear matrix association of Daxx and the other proteins in PML-NBs. To overcome this, we performed "Proximity-mediated Biotin Identification" (BioID)[144], where neighboring proteins are covalently biotinylated by a promiscuous Biotin ligase that is expressed as a fusion to the protein of interest. Unlike antigen-antibody interactions, biotinylation is a covalent modification, thus stable even under the harsh denaturing conditions required to solubilize PML-NBs. Another advantage of the BioID system is its utilization of Streptavidin beads to pull down the biotinylated proteins. The interaction between Streptavidin and Biotin is known to be the strongest non-covalent bond in biology, allowing the application of harsh washes to get rid of non-specific proteins. The aforementioned properties, i.e., the stability of biotinylation and the strength of the streptavidin-biotin bond, also render BioID a highly sensitive system.

5. Materials, Methods, and Techniques

5.1. Methods

5.1.1. Culture of Cell Lines

5.1.1.1. Culture Conditions

Two different cell lines, NB4 (human PML/RAR-positive neuroblastoma) and HEK293T (human embryonic kidney) were used. The two cell lines were maintained in RPMI and DMEM, respectively, supplemented with 1% L-glutamine, 1% Penicillin/Streptomycin, and 10% Fetal Bovine Serum (FBS).

Note: the extraction and culture of primary mouse bone marrow cells, which were also extensively used in this study, will be discussed in a separate chapter.

5.1.1.2. Subculture of NB4 Cells

NB4 cells grow in suspension. When the cells are ready for passaging, the flask is swirled to distribute the cells in the medium evenly. The suspension is centrifuged at 300 × *g* for 5 minutes. The supernatant, which contains the bulk of the accumulated cell debris and metabolic by-products, is disposed of, whereas the cell pellet is resuspended in a minimal volume of pre-warmed RPMI 1640. After taking a small sample to measure the cells using a hemocytometer, the rest are diluted with pre-warmed RPMI 1640 to the recommended seeding density.

5.1.1.3. Subculture of HEK293T Cells

HEK293T is an adherent cell line. The bulk of the old medium is first removed by decanting. In order to remove the last traces of medium, which is desired because the serum contained in it can inhibit the activity of trypsin in the subsequent step, cells are gently washed with PBS. After discarding the PBS, a pre-warmed trypsin solution is added to the side of the flask at a volume sufficient to cover the cell layer by gently rocking the flask container to get complete coverage of the cell layer. The culture flask is left at room temperature until at least 90% of the cells have been detached. Tapping of the flask is performed to expedite cell detachment. When ≥ 90% of the cells have been detached, the vessel is tilted for a minimal length of time to allow the cells to drain. Afterward, pre-warmed DMEM at a volume equal to 2 volumes of the trypsin solution is added to the flask and dispersed by pipetting over the cell layer surface several times. The cell suspension is transferred to a 15 mL conical tube, centrifuged at 300 × *g* for 5 minutes, and resuspended in a minimal volume of pre-warmed DMEM. A small volume (10-20 μL) is removed, mixed 1:1 with Trypan Blue 2× solution, and used to measure the cells under a microscope using a hemocytometer and a cell counter. The Trypan Blue dye allows the discrimination of dead and living cells, as it can only penetrates dead cells. After the living cell concentration is calculated, the HEK293T cell suspension is diluted to 5×10^5 cells/mL (recommended seeding density). The appropriate volume is pipetted into a new cell culture flask and the cells are returned to the incubator.

5.1.2. Cell lysis and Protein Extraction

5.1.2.1. Lysis of HEK293T Cells

The culture medium is first carefully removed, and the cell monolayer is washed twice with cold PBS. Afterward, cold RIPA Lysis Buffer, to which 1× protease inhibitor cocktail and Benzonase endonuclease are added immediately before use, is directly applied to the cells (1 mL per 5×10^6 cells). After 30 minutes on ice, the lysate is gathered to one side of the flask using a cell scraper, collected, transferred to a 1.5 mL Eppendorf tube, and centrifuged at 20,000 × *g* for 20 minutes to pellet cellular debris. The supernatant is transferred to a new tube for further analysis.

5.1.2.2. Lysis of NB4 Cells

The suspended cells are first pelleted by centrifugation at 500 × *g* for 5 minutes and then washed twice in cold PBS. The wet pellet is resuspended in cold RIPA Lysis Buffer (1 mL of RIPA buffer per 10^7 cells), to which 1× protease inhibitor cocktail and Benzonase endonuclease are added immediately before use and placed for 30 minutes in the cold room under gentle shaking. Afterward, the mixture is centrifuged at 20,000 × *g* for 20 minutes to pellet cellular debris. The supernatant is transferred to a new tube for further analysis.

5.1.3. NB4 Subcellular Protein Fractionation

5.1.3.1. Cytoplasmic Fraction

NB4 cells are collected in 50 mL Falcon tubes, centrifuged for 5 minutes at 500 × *g*, and washed twice with ice-cold PBS. The packed cells are resuspended by gentle shaking in 1× hypotonic lysis buffer (DTT and protease inhibitors are added just before use) at a volume equal to five times the estimated packed cell volume, transferred to a 1.5 mL Eppendorf tube, and incubated on ice for 30 min. Cells swell as a result of the hypotonic conditions, building up pressure to the cytoplasmic membrane. The addition of IGEPAL CA-630 solution to a final concentration of 0.6% V/V (6 µL of the 10% stock solution per 100 µL of the mixture) and subsequent vigorous vortexing breach the weakened membranes, releasing the cytoplasmic content, while the nuclear membranes should remain intact. In order to assess the degree of lysis, a sample of the lysate is viewed under the microscope after Trypan Blue staining. After confirming the absence of whole cells, the lysate is centrifuged for 1 min at 10,000 × *g* to precipitate the nuclei. The supernatant is transferred to a new tube and stored at -80 °C, while the nuclei are further processed.

5.1.3.2. Nuclear Fraction

The crude nuclear pellet is resuspended by pipetting up and down in a volume of Extraction Buffer (DTT and protease inhibitor cocktail are added just before use) equal to 2/3 × Pellet Cell Volume (PCV) (usually around 70 µL). The mixture is shaken gently for 30 minutes in the cold room and then centrifuged for 20 minutes at 20,000 × *g*. The supernatant is transferred to a clean, chilled tube and stored at -80 °C.

5.1.4. Bradford Protein Assay

A volume of 5 µL from each standard (BSA, bovine serum albumin) or unknown sample is pipetted into 96-Well Plates and mixed with 200 µL of Coomassie Reagent. After 10 minutes at room temperature, the absorbance at 595 nm is measured with a Tecan Microplate optical density reader.

The average absorbance of the replicate blank samples, i.e., those containing no protein, is subtracted from the average absorbance values of the replicates of all other standard and unknown sample replicates. The resulting values of the BSA samples are plotted against their concentration. A standard curve is calculated and used to determine the protein concentration in the unknown samples.

5.1.5. Western Blotting

5.1.5.1. Gel Preparation

The glass plates are assembled according to the manufacturer's instructions to create the gel mold. The gel is composed of two parts: a lower "resolving" gel, in which the proteins are separated, and an upper "stacking gel" with a lower acrylamide concentration, which allows the proteins to enter the former simultaneously.

The appropriate volume of resolving gel is prepared. Importantly, the polymerization catalyst, i.e., TEMED, is added last. After TEMED has been added, the mixture is swirled rapidly and poured into the gap between the glass plates, up to a height that leaves sufficient space for the stacking gel (the length of the teeth of the comb plus 1 cm). A Pasteur pipette is used to overlay the acrylamide solution with 0.1% SDS carefully. The purpose of the overlay is twofold; it inhibits polymerization by preventing oxygen from diffusing into the gel and ensures that the upper side of the

resolving gel will form a straight line. The gel is left in a vertical position at room temperature (ideally, to avoid mixing the still-liquid gel with the overlay-H_2O or isopropanol, the apparatus is moved to the final position before the overlayer is added). After the polymerization is complete (it requires around 30 minutes), the overlay is poured off, and the top of the gel is washed several times with deionized H_2O to remove any non-polymerized acrylamide. After the last wash, any remaining water is drained from the top of the gel with the edge of a paper towel.

The solution for the stacking gel is usually prepared while the polymerization of the resolving gel takes place. Again, TEMED is added last, just before the solution is poured on top of the polymerized resolving gel. Without delay, a clean Teflon comb is slowly (to avoid trapping air bubbles) inserted into the stacking gel solution. The solution is allowed to polymerize in a vertical position at room temperature. Teflon combs should be cleaned with H_2O and dried with ethanol just before use. Besides preparing our own gels, we routinely use of precast NuPAGE Bis-Tris Mini Gels.

5.1.5.2. Sample Preparation, Loading, and Electrophoresis

Samples are mixed with an equal volume of 2× Laemmli sample buffer and heated at 95 °C for 5 minutes to denature proteins. The high concentration of SDS and the presence of DTT in the Laemmli buffer combined with the high temperature fully dissociate the proteins. Thanks to the glycerol contained in the Laemmli buffer, the resulting solution is denser than the running buffer of the protein gel, enabling easy loading into the gel pockets. The Laemmli buffer also contains Bromophenol blue, which allows both visualization of the sample solution during loading and observation of the progress of the electrophoresis as it runs ahead of the proteins. After the solution cools, Benzonase endonuclease is added to digest the DNA. This reduces the viscosity of the solution, making its loading easier. Any sample needs to go through the process only once. Afterward, it can be frozen and immediately used after thawing when required.

The sample solution is loaded using a pipette. The volume to be loaded depends on the desired quantity of total protein to be electrophorized. In any case, the loaded volume should not exceed the maximum capacity of the wells. A small volume (2 μL) of the PageRuler Plus Prestained 10-250 kDa Protein Ladder is loaded into one well. This contains a mix of polypeptides of known molecular weights, serving as a reference for the molecular weights of the proteins forming the bands in the sample lanes.

When using our own gels, an initial voltage of 8 V/cm was applied. After the samples (whose presence is evidenced by the presence of the dye) have passed into the resolving gel, the voltage is increased to 15 V/cm. When using precast gels, a constant voltage of 200 V is applied for the duration of the electrophoresis. In either case, the electrophoresis is terminated when the Bromophenol blue front reaches the bottom of the gel. When histones were analyzed, the front was not completely run out.

5.1.5.3. Wet Tank Transfer

Shortly (10-15 minutes) before the end of the electrophoresis, the transfer membrane (nitrocellulose), two sheets of pre-cut blotting filter paper, and two sponge pads are placed to soak in a suitable tray containing 1× transfer buffer. In the same tray, we place the gel (after trimming the wells) when the electrophoresis finishes.

The gel "sandwich" is prepared as follows: The cassette holder, gray side (i.e., the one that will face the anode during transfer) down, is placed on a clean, flat surface; one of the pre-wetted sponge pads is placed on the gray side of the cassette; one sheet of soaked filter paper is placed onto the sponge pad; the equilibrated gel is placed onto the filter paper; the membrane is carefully placed on the gel, ensuring that no gel material protrudes from the edges of the membrane; a roller, or a simple falcon tube without the lid, is used to remove air bubbles between the gel and the membrane; the second sheet of soaked filter paper is placed onto the membrane; the second soaked

sponge pad is placed on top of the stack to complete the sandwich; finally, the cassette is closed and latched carefully, in order to avoid disturbing the gel sandwich.

The cassette is placed into the electrode module, with its gray side facing the black side of the module (i.e., the one that will be connected to the anode). Thus, the negatively charged proteins inside the gel will move towards the cathode and get trapped in the membrane. The electrode module and a magnetic stir bar are placed into the buffer tank. After the tank is filled with cold 1× transfer buffer, it is transferred to the cold room and placed on a magnetic stir plate. The movement of the magnet forces the buffer to recirculate, ensuring even ion and thermal distribution throughout the transfer. After the lid is attached and the cables are plugged into the power supply, the transfer is carried out at a constant amplitude of 400 A for 1 h. At the end of the transfer run, the gel sandwich is disassembled. While the polyacrylamide gel and the filter paper sheets are discarded, the transfer buffer can be reused once or twice.

5.1.5.4. Ponceau S Staining

Ponceau S staining allows us to check the efficiency of transfer and gives us a view of how well our samples ran. The membrane is immersed into the Ponceau S Stain solution and placed on an orbital shaker for 5 minutes at room temperature. Afterward, the membrane is rinsed 3 or more times with deionized water to remove the excess stain. If the result is satisfactory, i.e., the proteins have successfully migrated from the gel to the membrane and the bands have the expected pattern, we proceed to the next stage.

5.1.5.5. Antibody Incubation and Detection

The membrane is immersed in the blocking buffer and left for1 h at room temperature under gentle shaking. Afterward, the membrane is incubated with the primary antibody, which has also been diluted in blocking buffer (depending on the antibody, we used dilutions in the range of 1:200-1:5,000), and incubated overnight at 4 °C under gentle shaking.

On the next day, the primary antibody solution is collected (most antibody solutions can be used at least 2-3 times). The membrane is washed three times (5 minutes each) with 1× TBS-T under vigorous shaking on a platform shaker at room temperature and then incubated with the secondary antibody, which is also diluted in blocking buffer (1:20,000), for 1 h at room temperature under gentle shaking. The secondary antibodies used in this study were conjugated with fluorescent dyes emitting light at 680 or 800 nm. Afterward, the secondary antibody solution is collected, and the membrane is again washed three times with TBS-T. If the 800 nm channel is going to be used, the membrane goes through additional washes with 1× TBS. The purpose of the additional washes is to remove the Tween-20 from the membrane, as this detergent also emits light at 800 nm and, thus, gives a strong background. Finally, the membrane is scanned using the Odyssey Imaging System.

5.1.6. Immunocytochemistry

5.1.6.1. Cytospin Slide Preparation

To attach NB4 cells, which grow in suspension to glass for staining, the Cytospin procedure was performed. Cytoslides to be used are labeled with a pencil and placed into Cytoclips. After the Cytoclips are gently fastened, a small volume of cell culture material is added to each clip. To achieve optimal distribution and visibility, 200 µL

containing around 10^5 cells should be used. Cytoclips are then placed inside the Cytospin Centrifuge and centrifuged at 600 × g for 10 min, using the medium acceleration setting. Once the spin is complete, the slides are removed from the Cytoclips and air-dried at room temperature.

5.1.6.2. Fixation and Permeabilization

Slides are transferred into a slide rack and then dipped into -20 °C methanol for 5 min, followed by -20 °C acetone for 15 s to rinse. Afterward, the slides are left to air-dry under the hood.

5.1.6.3. Blocking and Antibody Incubation

After drying, slides are incubated first in PBS with 2% FBS at room temperature for 30 minutes and then with the desired first antibody (also diluted in PBS with 2% FBS) for another 1 h at room temperature. After 3 washes with PBS, slides are incubated with the appropriate second antibody (also diluted in PBS with 2% FBS) for 1 h at room temperature under dark and moist conditions, washed 3 times with PBS, and stained with Hoechst for the nuclei to become visible.

5.1.6.4. Mounting and Visualization

Coverslips are mounted on the slides after the addition of a drop of mounting medium. The mounting medium seals the slide, preventing drying. The slides can now be visualized under a microscope.

5.1.7. Cloning

5.1.7.1. PCR Amplification of Desired Sequence

PCR reactions were prepared as shown in Table 5-1 in sterile thin-walled PCR tubes and run under the conditions listed in Table 5-2.

Table 5-1 Herculase Polymerase PCR reaction

COMPONENT	50 µL REACTION
DNA Template	1-30 ng vector DNA
5 × Reaction Buffer	10 µL
10 mM dNTP	2 µL
10 mM Primer Mix	2.5 µL
Herculase Polymerase	1 µL
dH2O	X µL to final 50 µL volume
Total	50 µL

Table 5-2 Herculase Polymerase PCR reaction cycling parameters

Segment	Number of cycles	Temperature	During
1	1	98 ℃	2 min
2	30	98 ℃	20 s
		Primer Tm -5℃	30 s
		72 ℃	30 seconds per kb
3	1	72 ℃	5 min
4	1	4 ℃	Pause

5.1.7.2. Digestion of Amplicon and Vector

Digestion reaction mixtures are prepared at room temperature as shown in Table 5-3, gently mixed, spun down, and incubated at 37 °C in a heat block for 15 min. Most of our cloning strategies were directional, i.e., two different enzymes were used. Depending on whether the reaction conditions required by the two enzymes are compatible or not, the two reactions are carried out simultaneously or sequentially, respectively. After inactivation by heating at 70 °C, an aliquot of the reaction is electrophorized in Agarose both to confirm the success of the digestion and to excise the desired DNA fragments. The latter may be required for one or both of the digestion reactions (amplicon and vector).

Table 5-3 FastDigest digestion reaction

	Plasmid DNA	PCR product
Water, nuclease-free	15 µL	16 µL
10 × FastDigest Green Buffer	2 µL	3 µL
DNA	2 µL (up to 1 µg)	10 µL (~0.2 µg)
FastDigest enzyme	1 µL	1 µL
Total volume	20 µL	30 µL

5.1.7.3. Gel purification

The desired bands are excised from the gel and placed in pre-weighed Eppendorf tubes. Weighting the tube again gives us the weight of each band. As we already know the density of the gel, we can use the weight to calculate the volume of the excised band. Afterward, 3 volumes of the "dissolve solution" are added to each volume of excised Agarose. The mix is incubated at 50 ℃ until the slice is completely dissolved (usually 10 minutes suffice), transferred to a Spin Column in a Collection Tube, and centrifuged for 60 seconds at 10,000 ×*g*. The column matrix is washed twice with 200 µL of "DNA Wash Buffer". Finally, the DNA is eluted using ≥ 6 µL of "DNA Elution Buffer".

5.1.7.4. Dephosphorylation (Optional)

At the ligation stage, which is described in the next chapter, the "sticky ends" of the insert and the vector are joined to form the desired molecule. When a single restrictive endonuclease is used in a cloning reaction, the sticky ends of the vector are, obviously, compatible with each other. Thus, the plasmid may recirculate. One can avoid this by dephosphorylating the digested vector prior to the ligation step. The reaction is prepared as shown in Table 5-4. After incubation at 37 °C for 60 minutes, the reaction is stopped by heating to 65 °C for 15 minutes.

Table 5-4 Dephosphorylation reaction

COMPONENT	Per 20 µL REACTION
DNA	1 µg
10 × Restriction Enzyme Buffer	2 µL
Water, Nuclease-Free	to 19 µL
Restriction Endonuclease	1 µL

5.1.7.5. Ligation

In ligation, a phosphodiester bond is formed between the 3'-OH at one end of a strand of DNA and the 5'-phosphate group of another. The reaction is prepared as shown in Table 5-5 on ice. The T4 DNA Ligase Buffer should be thawed and resuspended at room temperature. In most reactions performed in this study, a 1:3 vector-to-insert molar ratio was used. The reaction is gently mixed by pipetting up and down, spun down, and incubated at room temperature for 1 h. Following incubation, the reactions are used immediately for transformation and/or are stored at -20 °C.

Table 5-5 T4 Ligation reaction

COMPONENT	Per 20 µL REACTION
T4 DNA Ligase Buffer (10×)	2 µL
Vector DNA	50 ng
Insert DNA	Quantity containing three times the number of vector moles
Nuclease-free water	to 20 µL
T4 DNA Ligase	1 µL

5.1.7.6. Standard Mutagenesis (Overlapping Primers)

The advantage of overlapping primers is that the new product is already circular and can be directly inserted into competent bacteria. The main disadvantage is that one of the two primers must contain (in its non-overlapping part) the entirety of the mutation, which becomes problematic when one wants to add into the target sequence a number of nucleotides much larger than the typical size of the primer. Moreover, overlapping primers preclude exponential amplification.

The reaction is prepared as shown in Table 5-6 and performed as shown in Table 5-7. At the end of the PCR and after the reaction cools, 1 μL of the DpnI restriction enzyme (10 U/μL) is added directly to it. Each reaction is gently and thoroughly mixed and then incubated at 37 °C for 1 h. DpnI is a methylated-G^{m}ATC-specific endonuclease, i.e., it can only act on methylated DNA strands. Thus, it fragments (typically>1 restriction sites) the parental plasmid that was used as the template (it was isolated from a bacterial colony, so it contains methylated nucleotides), whereas it does not affect the newly synthesized circular DNA molecules, which contain the desired mutation. Following incubation, the reactions are used immediately for transformation and/or are stored at -20 °C.

Table 5-6 Mutagenesis PCR reaction

COMPONENT	50 µL REACTION
DNA Template	50 ng
10×Pfu Buffer	2 µL
10 mM dNTP	2 µL
10 mM Primer Mix	4 µL
Pfu Polymerase	1 µL
Taq Polymerase	0.5 µL
dH_2O	X
Total	50 µL

Table 5-7 Mutagenesis PCR reaction cycling parameter

Segment	Number of cycles	Temperature	During
1	1	95 ℃	2 min
2	20	95 ℃	30 s
		60 ℃	30 s
		68 ℃	1 min per kb
3	1	68 ℃	5 min
4	1	4 ℃	Pause

5.1.7.7. Q5 Site-Directed Mutagenesis (Non-Overlapping Primers)

In cases when the overlapping strategy proved non- successful (extensive GC-stretches in the template) this alternative strategy was used.

With non-overlapping primers, the desired changes can be divided into both primers, making the addition of larger numbers of nucleotides easier. Moreover, the amplification takes place exponentially. The disadvantage is that, in addition to DpnI, one must also use kinase and ligase. The latter is required to circularize the newly synthesized double strands, while the former is added because typically (i.e., when standard, non-phosphorylated primers are used) the 5'-ends of synthetic oligonucleotides are not phosphorylated, thus cannot be ligated without prior phosphorylation.

The reaction is prepared as shown in Table 5-8 and run under the conditions shown in Table 5-9. After the PCR ends and the reaction cools down, an aliquot is treated with a mix of DpnI, PNK kinase, and T4 ligase (Table 5-10) for 1 h at room temperature. Following incubation, the reactions are used immediately for transformation and/or stored at -20 °C.

Table 5-8 Q5 Site-Mutagenesis PCR reaction

COMPONENT	50 µL REACTION
DNA Template	10 ng
5×Reaction Buffer	10 µL
10 mM dNTP	1 µL
10 mM Primer Mix	5 µL
GC Enhancer	10 µL
Q5 Polymerase	1 µL
dH_2O	X
Total	50 µL

Table 5-9 Q5 Site-Mutagenesis reaction cycling parameters

Segment	Number of cycles	Temperature	During
1	1	98 °C	2 min
2	30	98 °C	10 s
		Tm*	30 s
		72 °C	30 s per kb
3	1	72 °C	2 min
4	1	4 °C	Pause

*Determined by the “NEB Tm Calculator” application (http://tmcalculator.neb.com/)

Table 5-10 Q5 Site-Mutagenesis ligation reaction

COMPONENT	20 µL REACTION
T4 DNA Ligase Buffer (10×)	2 µL
PCR product	2 µL
DpnI	1 µL
PNK	1 µL
Nuclease-free water	14 µL
T4 DNA Ligase	1 µL

5.1.7.8. HiFi Assembly

After the production of the fragments using PCR, the assembly reaction is prepared on ice as shown in Table 5-11 and then incubated at 50 °C for 60 minutes. Following incubation, the reactions are used immediately for transformation and/or stored at -20 °C.

Table 5-11 HiFi assemble reaction

	2-3 Fragment	4-6 Fragment
Recommended DNA Molar Ratio	vector:insert = 1:2	vector:insert = 1:1
Total Amount of Fragments	0.03-0.2 pmol	0.2-0.5 pmol
HiFi DNA Assembly Master Mix	10 µL	10 µL
Deionized H_2O	To 20 µL	To 20 µL

5.1.7.9. Transformation

An appropriate volume of the ligation mixture (usually 2 µL containing about 50 ng DNA) is added to a thawed 50 µL aliquot of competent cells. After mixing gently (no vortexing) by pipetting up and down or flicking the tube 4-5 times, the mixture is placed on ice for 30 minutes, heat-shocked at 42 °C for 75 s, and immediately placed back on ice. After 2 min, 200 µL of SOC medium (at room temperature) is added to the tube, which is then incubated at 37 °C for 60 minutes. After vigorous shaking, 50-100 µL of the suspension is spread on pre-warmed selection plates, which are then incubated overnight at 37 °C.

5.1.8. Transfection with FuGENE

The protocol described in this chapter was used for all transient transfection purposes except for virus production (see section 5.1.9). HEK293T cells are plated in 10 cm plates (1.5-2 × 10^6 cells/plate) one day before transfection. At this density, the confluence on the next day should be approximately 80%.

On the day of transfection, the transfection mix is prepared as follows: Opti-MEM medium pre-warmed to room temperature is added to a 96-well plate (96 µL per well). One microgram of plasmid DNA (1 µL of 1 µg/µL) is added, and the mix is vortexed. Afterward, 3 µL of FuGENE HD Transfection Reagent, which was allowed to reach room temperature and then briefly vortexed, is added to the tube and mixed immediately. After incubating the mixture for 15 minutes at room temperature, it is added to the cells, which are then placed in the incubator for 24-48 hours. The transfection efficiency can be assessed using a fluorescence microscope or by FACS.

5.1.9. Viral packaging and transduction

5.1.9.1. Virus production

HEK293T cells are plated in 10 cm dishes (2.5×10^6 in 9 mL medium per dish) in the afternoon. On the next day and at three hours before transfection, the medium is replaced with 8 mL fresh growth medium. About 5 minutes prior to transfection, Chloroquine is added to each plate to a final concentration of 25 μM. Finally, the transfection mix, which contains the DNA to be transfected, is added dropwise to the medium and evenly distributed by swirling the plate.

The transfection mix is prepared as follows (obviously, all of the following procedures take place under sterile conditions): After gentle vortexing, 500 μL of HBS 2× solution is added to a 15 mL Falcon tube. Meanwhile, the lentivirus production components displayed in Table 5-12 and ecotrophic retrovirus production components displayed in Table 5-13, which are added in a 2 mL Eppendorf tube. Typically, the total DNA quantity in the protocol we used was 27 μg (the acceptable range is 20-30 μg). The resulting solution is added dropwise to the 500 μL HBS 2× solution while vortexing to prepare 1 mL of the transfection mix, which is then added dropwise to the cell cultures. Afterword, the dishes are placed cells back in the incubator for 12 hours. Then, the medium is replaced with 10 mL of fresh normal medium. The first supernatant is collected after another 36 hours, i.e., 48 hours post-transfection, and immediately replaced with 10 mL of fresh medium (cells should remain exposed to air for as little as possible). The collected supernatant is filtered through a 0.45 μm filter and stored at -80 °C (unless immediately used for infection). The second supernatant is collected 72 hours post-transfection.

Table 5-12 Lentivirus production reaction

COMPONENT	Volume
2.5 M $CaCl_2$	62 µL
Lentiviral Plasmid	5 µg
pRSV-Rev (88)	10 µg
pMDLg-pRRE Gag/pol (87)	10 µg
phCMV-RD114env (91) Human	2 µg
MQ up to a total of	500 µL

Table 5-13 Ecotrophic retrovirus production reaction

COMPONENT	Volume
2.5 M $CaCl_2$	62 µL
Ectotrophic retroviral Plasmid	16.5 µg
pEco-Pack	14.5 µg
MQ up to a total of	500 µL

5.1.9.2. Viral Transduction via Spinoculation

One or two days before transduction, cells are seeded in 24-well plates at a density of 10^5 cells/well. On the day of transduction, plates are centrifuged at 400 × *g* for 5 minutes. Supernatants are removed, and cells are resuspended in thawed virus supernatant (500 µL/well) to which 10 µg/mL Polybrene has been added just before transduction. The cells are centrifuged for 2 h at 2,500 × *g* at room temperature. The supernatant is then discarded, while the cells are resuspended in fresh medium (500 µL/well) and then placed back in the incubator. After 12 hours, the medium is again changed. Two days post-transduction, cells can be examined under the microscope or using FACS to check GFP expression. Since the cultures may still contain contagious material, 4% Paraformaldehyde needs to be added to the samples

immediately before FACS during the first 2 weeks after transduction. Non-transduced cells are used as a negative control for FACS analyses.

5.1.9.3. Puromycin Selection

Before using Puromycin, non-transfected cells are treated with various concentrations of the drug for one week. The proliferation of the cells is checked using Trypan-Blue staining (see section 5.1.1.3). Based on the proliferation values, a "kill curve" is drawn. The concentration that kills around 90% of cells is noted.

Transfected cultures are maintained in non-selective medium for 48 hours post-transfection and then plated in medium containing Puromycin at the concentration determined by the kill curve. The reason for not choosing a higher concentration at this point is that we do not yet know how resistant the transfected cells are. The effectiveness of the selection is monitored by FACS measurement every 2-3 days. The GFP-positive percent should rise during the first 3-4 passages and then stabilize. At this point, only transfected cells and resistant clones of the non-transfected cells should remain. In order to further increase the percentage, i.e., kill even more of the non-transfected cells including very weak expressing cells.

5.1.10. Diff-Quik™ staining

We used this technique on cells immobilized on slides, fixed, and permeabilized as described in section 5.1.6.1 and 5.1.6.2. The slides are then sequentially submerged (with intermittent rinses using distilled water to minimize cross contamination between staining solutions) in the three Dispense Diff-Quik Solutions (10 seconds in each). After the incubation in the last solution, slides are rinsed with distilled water twice and left to air-dry. Afterward, the procedure described in section 5.1.6.4 is followed.

5.1.11. Flow cytometry and Fluorescence-Activated Cell Sorting

Flow cytometry is a technique used to detect and measure the physical and chemical characteristics of a population of cells or particles. One of the most typical forms uses fluorescent dyes (per se or conjugated to antibodies) to differentiate cells based on the ability of the dye or the dye-conjugated antibody to bind to specific target molecules. An example of the former is Ethidium Bromide, whose ability to intercalate between DNA base pairs allows researchers to distinguish cells based on DNA content. In the case of antibodies, the dye-conjugated antibody could be a first or a second antibody. The term "flow-cytometric immunophenotyping" is commonly used when specific (usually surface) antigens, characteristic of particular cell types, are targeted.

Fluorescence-activated cell sorting (FACS) is a specialized type of flow cytometry. It allows not just recording, as in the case of conventional flow cytometry, but actual physical separation, one cell at a time, of the subgroups comprising a heterogeneous mixture. The separation is based upon the specific light scattering and fluorescent characteristics of each cell. It is achieved by bestowing a positive or negative charge to cells based on their fluorescence and then using an electric field to convey the differentially charged cells to different containers.

In our study, aliquots of cell suspensions were diluted 1:1 in cold PBS 2× containing 2% FBS. At the resulting final concentration (1%), FBS prevents the aggregation of the cells, which would otherwise compromise the efficiency of the method. The diluted cell suspension is treated with the fluorophore or the fluorophore-treated antibody and then processed. When potentially dangerous samples need to be prepared (e.g., suspensions that may contain lentiviruses), Paraformaldehyde is also added to a final concentration of 4%.

5.1.12. BioID2 Pull-down and Streptavidin-Bead Precipitation after ATRA treatment

Cells expressing BioID2 constructs and parental NB4 control cells are treated with 2 μM ATRA or DMSO (as solvent control) for 2 days, followed by 1 day in fresh medium containing both ATRA (or DMSO) and 50 μM Biotin. On the next day, cells (usually 10^8 cells in 200 mL culture medium) are collected (400 × *g* at 4 °C for 5 min), washed 3 times with 10 mL cold PBS, and lysed in 500 μL RIPA lysis buffer (protease inhibitors and DTT are added just before use). After adding 250 U of Benzonase endonuclease, samples are incubated on a rotator at 4 °C overnight. On the next day, samples are centrifuged in a pre-cooled centrifuge at 4 °C for 20 min at 21,100 × *g*. The supernatants are loaded into Streptavidin High-Performance Spin-trap columns, which have been previously prepared by removing the storage buffer and equilibrating with 500 μL RIPA buffer (1 min at 200 × *g*). The liquid is mixed with the beads by manual inversion of the columns, followed by slow, end-over-end mixing on a rotator at 4 °C for several hours or overnight.

After the incubation, the samples are centrifuged for 1 min at 200 × *g*. The supernatant, containing the bulk of unbound proteins, is removed, while the beads are washed three times with 500 μL of Wash Buffer 1 through gentle manual inversion followed by centrifugation. This is followed by two washes with 500 μL of Wash Buffer 2 and two washes with 500 μL of Wash Buffer 3. As the wash buffers contain detergents that may impede the analysis of the bound proteins, their traces must be removed before elution. To this end, a final wash step is performed, this time with 450 μL of 50 mM Tris·HCl pH 7.4. This time, the buffer is not removed by centrifugation of the column, but the resuspended beads are transferred into 1.5 mL Eppendorf tubes. Most of the material (about 900 μL) will be used for mass-spec analysis, whereas one-ninth (about 100 μL), which is to be used for western blot

analysis, is transferred to new Eppendorf tubes. All tubes are centrifuged for 5 min at 2,000 × *g* at room temperature to precipitate the beads.

After the removal of the supernatants, the beads in the tubes that contained the 400 µL aliquots are gently resuspended in 70 µL of 50 mM NH_4HCO_3 and then stored at -80 °C, while the beads from the 100 µL aliquots are gently resuspended in 200 µL of 2× SDS-PAGE sample buffer to which 10 µL of 20× Biotin has been previously added. After heating at 95 °C for 5 min, the samples can be electrophorized (usually a volume of 10 µL is loaded) to be used for western blot and silver staining analysis, using either Streptavidin or antibodies.

5.1.13. Silver staining

The assay was performed using the Pierce Silver Stain Kit as follows: After electrophoresis, the gel is washed two times in ultrapure water (5 minutes each) and then incubated in the Fixing Solution for 15 minutes at room temperature. At the end of the first 15 minutes, the solution is replaced by fresh one, and the incubation continues for at least another 15 minutes (the gel can remain in fixing solution overnight without stain performance loss). After fixation, the gel is washed first with the Ethanol Wash solution (5 minutes plus another 5 minutes in fresh solution) and then in ultrapure water (5 minutes plus another 5 minutes in fresh water). Afterward, the gel is submerged for 1 minute in sensitizer working solution (prepared just before use by mixing 1 part Silver Stain Sensitizer with 500 parts ultrapure water), washed three times in ultrapure water for 1 minute each, and then submerged for 5 minutes in a solution prepared just before use by mixing 1 part Silver Stain Enhancer with 100 parts Silver Stain. The gel is quickly (20 s) washed in ultrapure water twice and then submerged in developer working solution. When the protein bands appear and reach the desired band intensity (usually 2 minutes suffice), the gel is placed in the Stop Solution, briefly washed, and then incubated in acetic acid for 10 minutes.

5.1.14. Nano-Glo Dual-Luciferase Reporter Assay

HEK293T cells expressing both firefly and NanoLuc luciferase are plated in 96-well plates (80 µL of suspension per well) and left to reach room temperature. Afterward, an equal volume (i.e., 80 µL) of the appropriate reagent (ONE-Glo EX Reagent for firefly luciferase, NanoDLR Stop & Glo for NanoLuc) is added to the wells. The plates are then placed on an orbital shaker at room temperature for 30 minutes, during which the cells are lysed. After incubation, the luminescence produced by each of the released luciferases is measured using the Tecan reader.

5.1.15. Forster Resonance Energy Transfer (FRET)

The following plasmid constructs were transfected into HEK293T cells using the FuGENE HD transfection Kit: donor-only control CFP-tagged protein of interest; acceptor-only control YFP-tagged protein of interest; positive control CFP- linker-YFP fusion construct producing a maximum quantity of FRET. After transfection, cells are incubated at 37 ℃ overnight. On the next day, the transfection medium is replaced by regular growth medium (i.e., DMEM plus 1% penicillin/streptomycin, 1% L-glutamine, and 10% FBS), and the cells are incubated for another 24-48 hours to obtain similar expression levels for both proteins. After incubation and before analysis, cells are washed gently three times with PBS.

The Tecan reader is set to obtain the donor background signal (light source to donor excitation channel, 431-441 nm; emission channel, 455-485 nm) and then the acceptor background signal (light source to acceptor excitation channel, 490-510 nm; emission channel, 520-550 nm. Finally, the FRET signal is obtained (light source to acceptor excitation channel, 431-441 nm; emission channel, 520-550 nm). The three sets of data are retrieved and processed.

5.1.16. Mouse Bone Marrow Cell Extraction, Transduction, and Passaging

Bone marrow was harvested from healthy, unconditioned, 3-6 month-old female Rosa26CreERT2::Brg1$^{fl/fl}$ mice by crushing isolated femurs and tibia in IMDM (Iscove's Modified Dulbecco's Medium) with a glass mortar and pestle. To measure the density of nucleated cells in the suspension, an aliquot is diluted 1:10 in 3% acetic acid with Methylene blue. This results in red blood cell lysis, allowing us to count the white blood cells using a hemocytometer. After the density is calculated, the original suspension is further diluted to reach a concentration of 2×10^7 nucleated cells per mL, seeded in triplicates to 24-well plates (100 µL, i.e., 2×10^6 cells, per plate), and mixed with 500 µL Myelocult growth medium supplemented with 1:100 Antibiotic/Antimycotic, 100 ng/mL mSCF, 6 ng/mL mIL-6, and 10 ng/mL mIL-3. After 20 hours of incubation that serves to induce cell growth (prestimulation), cells are infected by spinoculation using stored retroviral supernatant in capped round-bottom Polystyrol tubes at 2,500 × *g* for 100 minutes. For these cells, ecotrophic retroviruses were used. Transduction was performed independently for each well of the triplicate. After spinoculation, cells are resuspended in fresh MyeloCult growth medium (again containing mSCF, mIL-6, and mIL-3, as these are necessary for the proliferation of the cells) and transferred back into 24-well plates. The first sample of transfected cells (passage "zero", P0) is collected 24 hours after transduction, followed by samples every three days. At every passage, 100 µL of the culture is transferred to a new well containing 400 µL of fresh medium. The remaining 400 µL of the culture is used either partly (when an aliquot must be collected for other purposes) or wholly for FACS analysis, after first adding Propidium Iodide (PI) at 1 µg/mL final concentration (to discriminate dead cells). Both the percentage of GFP+ cells and the absolute cell numbers are obtained, the latter by performing the collection for 30 seconds at high flow rate and comparing the number with those of a sample with a known number of cells (measured in a hemocytometer).

5.1.17. Real-time PCR for Quantitative Analysis

The primers are usually kept as 100 μM stock solutions. Typically, the working solutions are prepared by diluting the stocks ten times, i.e., to 10 μM. However, as the concentrations of the forward and reverse primers in all reactions performed in this study were equal, the diluted primers were mixed at a 1:1 ratio resulting in a solution containing both primers at 5 μM. It is essential to avoid repeated freeze-thaw cycles of the "FastStart Essential DNA Master Mix 2×" reagent, as these can lead to loss of DNA polymerase activity. Moreover, the said solution should be protected from direct sunlight, as the contained fluorophore, Sybr Green, is photosensitive. Reactions are prepared in LightCycler 480 96-well plates as shown in Table 5-14. Of course, when a large number of reactions have to be prepared, it is preferable to create a master mix (not to be confused with the "FastStart Essential DNA Master Mix 2×," which is one of the solutions used) containing everything but the template. Even though mixing is necessary, care must be taken to avoid the formation of bubbles. If bubbles do appear, plates are centrifuged at 1,500 × *g* for 2 minutes.

After the reactions have been prepared, they are run in the LightCycler 96 Instrument using the program displayed on Table 5-15. However, these conditions (particularly the annealing temperature and the primer concentration) often had to be tweaked in order to achieve optimal amplification. The rule of thumb is to select the conditions at which the efficiency is maximized without compromising the specificity. The creation of a standard curve using different dilutions of the template is necessary in order to make sure that the amplification efficiency is close to 100%, both at the stage of optimization as well as afterward, since random factors can affect the efficiency of a previously optimized reaction.

With respect to the template concentration, one is considered appropriate when it gives signal strength within the range covered by the standard efficiency curve.

The abundance of the target sequence is calculated as $(1+E)^{-\Delta\Delta Ct}$: E is the efficiency of the PCR reaction. When the efficiency is 100%, 1+E is equal to 2. ΔΔCt is the difference between the normalizer and each sample of the ΔCt values, which in turn is the difference in cycle threshold (Ct) values between the target sequence and the reference sequence for a specific sample.

Table 5-14 Preparation of qPCR reactions

Regent	Volume (20 µL reaction)
template	x
Primer mix (5 µM each)	2
2× FastStart Essential DNA Master Mix	10
Water PCR grade	To 20 µL

Table 5-15 qPCR cycling parameter

Detection Format			Reaction Volume	
SYBR Green 1			20	
Programs				
	Temp. (°C)	Ramp. (°C/s)	Duration (s)	Acquisition Mode
Pre-Incubation	95	4.4	600	None
3-Step Amplification 50 cycles	95	4.4	10	None
	Primer dep	2.2	10	None
	72	4.4	10	Single
Melting	95	4.4	10	None
	65	2.2	60	None
	97	0.1	1	5 Readings/°C

5.1.18. Gene Knockout using a Tamoxifen-inducible Cre/Lox System

This was done using an inducible knockout mouse model kindly shared by U. Schüller (Forschungsinstitut Kinderkrebs-Zentrum Hamburg). These mice carry SMARCA4 alleles flanked by LoxP sites[145] and express a fusion protein (ER-Cre) coded by a sequence integrated into the ubiquitously expressed Rosa 26 locus. ER-Cre contains a Tamoxifen-responsive Cre recombinase fused with the domain of the estrogen receptor that controls its import into the nucleus in response to Estrogen[146] (Figure 5-1). Specifically, Estrogen binds to this domain and leads to the protein entering the nucleus, whereas the protein remains in the cytoplasm in the absence of Estrogen. Only in the nucleus can the Cre recombinase part of the fusion protein interact with the Loxp sites and excise the intervening sequence, i.e., the "floxed" allele. This model allows one to create knockouts both in vivo as well as ex vivo by injecting the animal or treating cell cultures, respectively, with Tamoxifen (or its active metabolite, 4-hydroxy-Tamoxifen), a selective estrogen receptor ligand.

5.1.19. Cell Cycle Analysis (Nicoletti Protocol)

A volume of culture material containing about 2×10^6 cells is collected and centrifuged at 200 × *g* for 5 min at room temperature. Cells are resuspended in 0.5 mL PBS and then fixed by the addition of 4.5 mL pre-cooled (-20 ℃) ethanol 70% (v/v). For the fixation to be completed, the 5 mL suspensions must be placed at -20 ℃ for at least 1 h. The suspensions are then centrifuged at 400 × *g* for 5 min. Fixated cells are washed with 5 mL of PBS and resuspended in 0.5 mL of PBS plus 0.5 mL of DNA extraction buffer. After incubation at room temperature for 5 min, cells are precipitated and then resuspended cells in 1 mL of DNA staining solution (containing propidium iodide at 10 mg/mL), where they are incubated for at least 30 min at room temperature in the dark. Afterward, cells are analyzed by flow cytometry using settings corresponding to propidium iodide.

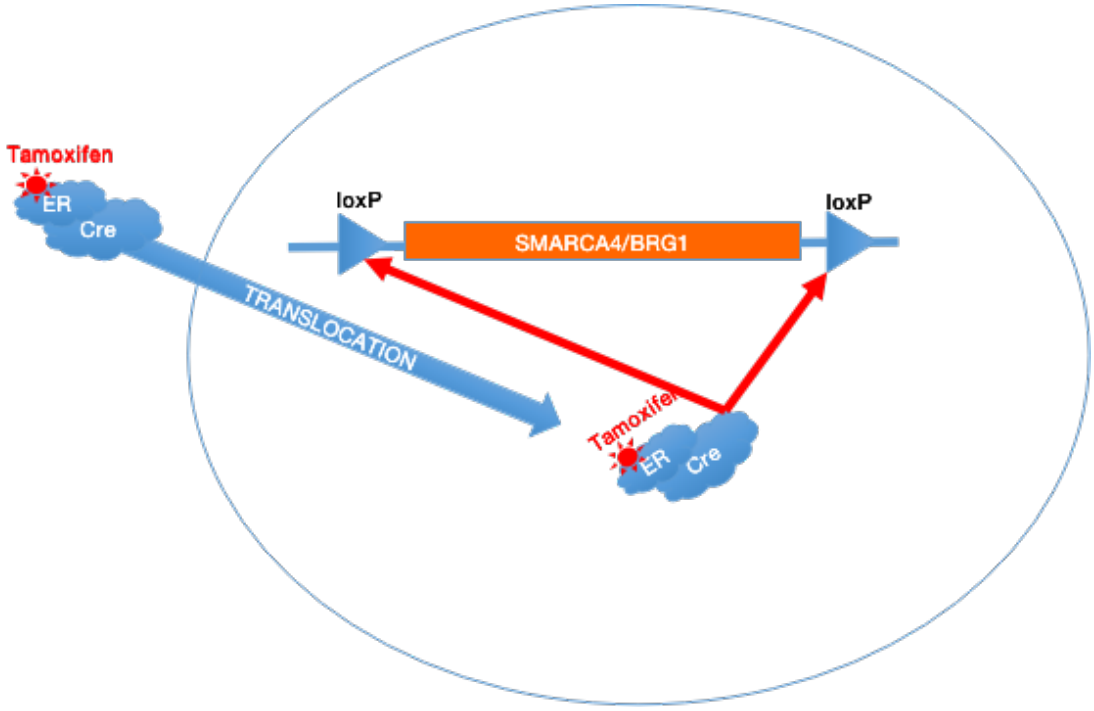


Figure 5-1 The principle of SMARCA4 knockout by 4-OH Tamoxifen in vitro. ER-Cre contains a Tamoxifen-responsive Cre recombinase fused with the domain of the estrogen receptor that controls its import into the nucleus in response to Estrogen. Specifically, Estrogen binds to this domain and leads to the protein entering the nucleus, whereas the protein remains in the cytoplasm in the absence of Estrogen.

5.1.20. NanoLC-MS/MS analysis and Raw Data Processing for the Purified Proteins

Samples were prepared for MS analysis by on-bead tryptic digestion. Solvent A was 0.1% (v/v) Formic acid and solvent B was 0.1% (v/v) formic acid in Acetonitrile. The tryptic peptides were separated using a nanoLC system. The sample was injected through the loading pump to a C18 trapping column at a flow rate of 15 µL/min with 1% solvent B and then switched to an analytical C18 column at a flow rate of 0.275 µL/min with 2% solvent B. Solvent B was increased to 20% at 115.5 min, to 32% at 135.5 min, and finally to 95% at 136.5 min. The nanoLC system was coupled with a Q Exactive mass spectrometer. For MS1 scanning, the maximum injection time was 60 ms; AGC target was 3×10^6; the m/z scan range was set from 375 to 1,600 with an orbitrap resolution of 70,000 full width at half maximum (FWHM) at m/z 200 for data acquisition. For HCD-MS/MS, the 10 highest-abundance precursor ions were selected for fragmentation with a normalized HCD collision energy of 27%. Fragment spectra were recorded using an orbitrap mass analyzer with an orbitrap resolution of 17500 FWHM at m/z 200; the maximum injection time was 50 ms; AGC target was 1×10^5. Obtained nanoLC-MS/MS raw data was visualized using the Xcalibur version 4.0.27.13 and processed with Proteome Discoverer 2.0. For peptide identification, obtained MS2 spectra were searched against theoretical fragment spectra of tryptic peptides, generated from the reviewed SWISSProt FASTA database, containing 20,239 entries, obtained in October 2018. The following parameters were used: the precursor mass tolerance was set to 10 p.p.m.; the fragment mass tolerance was set to 0.02 Da; variable modifications including oxidation of methionine (M, +15.995 Da) and acetyl (Protein N-term, +42.011 Da) were considered and the carbamidomethylation of cysteine (C, +57.021 Da) was set as a fixed modification; minimum length of considered peptides was set to 6 amino acids; 2 missed cleavages were tolerated.

5.1.21. Selection Procedure for Candidate Interactors

Besides DMSO- and ATRA-treated NB4-BD cells, our experiment utilized parental NB4 cells (both DMSO- and ATRA-treated) as additional negative controls. This allowed us a four-way comparison. Parental NB4 samples (ATRA- or DMSO-treated) allow us to filter out naturally occurring biotinylation (i.e., biotinylation not dependent on BioID2 fusion protein expression) from the corresponding NB4-BD samples.

After getting the Mass Spec result lists, the data was sorted by the total number of unique peptides (Sum) specific for a single protein in the human proteome, even though each peptide may appear more than once in the same protein, which provides information of relative abundance of that protein in the sample. As in this type of Mass spec experiments, reproducibility is typically only around 60%, we performed the purification in triplicates (Sample 1, Sample 2 and Sample 3). As an internal "positive control", we used a selection of proteins known to be associated with the complex (Sp100[147], PML, DEK, H3.3, and H4[148]). However many of them were only represented by a single peptide, such as Sp100 (in the Table 9-2 NB4-BD ATRA list), hence, we were unable to define a simple cut-off value based on unique peptides, but used cross-validation with additional bait proteins for reduction of false positives.

SMCHD1, after subtraction of background from parental NB4 cells, was the highest hit in our list. SMARCA4, on the contrary, was chosen for further study based on the fact that antibodies, cDNA, and a mouse model, were available from the group of Prof. U. Schüller in-house. After positive confirmation by western blot, we decided to take this chance for a collaboration.

5.1.22. RNA interference (SMARCA4-Knockdown)

The interfering sequence was produced based on the SMARCA4 sequence found in Plasmid pSMARCA4-H5, which was kindly procvided by the laboratory of Prof. Schneppenheim, UKE, Hamburg. We have sequenced the inserted cDNA of SMARCA4 and determined it matches the SMARCA4 isoform 3 as deposited in the NCBI database (RefSeq NM_003072). As a lentiviral vector is used to express the interfering RNA, the steps to prepare the virus and infect the cells were the same as those described in 6.1.9. “Lentivirus Packaging and Transduction”. Table 5-16 shows how the corresponding “virus production reaction” is prepared.

Table 5-16 Virus production reaction

COMPONENT	Volume
2.5 M $CaCl_2$	62 µL
Plasmid siRNA-SMARCA4	5 µg
pRSV-Rev (88)	10 µg
pMDLg-pRRE Gag/pol (87)	10 µg
phCMV-RD114env (91) Human	2 µg
MQ up to	500 µL

5.2. Material lists

Table 5-17 Lab Instruments

Name	Catalog Number	Company
Biological Safety Cabinets	51022515	Heraeus
Leica Microsystems microscope DMIL	090-135.001	Leica Microsystems
Upright Microscope	BX43	Olympus Corporation
PCR, Real-Time Instrument II	Light Cycler 480 Instrument II, 96-well	Roche
Plate reader Infinite M200	Infinite M200	Tecan Trading AG
NanoDrop 2000 spectrophotometer	ND-2000	Thermo Fisher Scientific
BD FACSCanto	657338	BD
Vortex mixer	SAB	Stuart
Overhead Shaker	Reax 2	Heidolph
Rotator	SB3	Stuart
Stuart Roller Mixer	1144170	Stuart
ThermoMixer C	5382000015	Eppendorf
Thermomixer comfort	5355000011	Eppendorf
Grant Instruments Dry Block Heating System, BTD Series	11456367	Grant Instruments BTD
Mini Centrifuge	C1301-230V-EU	Labnet International
Vented Micro Centrifuge Pico 17	75002401	Thermo Fisher Scientific
Heraeus Fresco 21 microcentrifuge	75002555	Thermo Fisher Scientific
Heraeus Multifuge 3SR Plus	75004371	Thermo Fisher Scientific
Heraeus Megafuge 40 universal centrifuge	75004503	Thermo Fisher Scientific
Thermocycler	T-Gradient	Biometra
Thermocycler	T1-Thermocycler	Biometra

Power Supply P25 Biometra	846-040-800	Biometra
Biometra Compact XS/S, M and L/XL	XS/S, M and L/XL	Biometra
BDAdigital core set	034 - 000	Biometra
Heidolph Magnetic Stirrers	MR 3001	Heidolph
Cytospin 4 cytocentrifuge	A78300003	Thermo Fisher Scientific
New Brunswick Innova 44	M1282-0002	Eppendorf
B12 Incubator, Function Line	50042307	Thermo Fisher Scientific
XCell SureLock Mini-Cell Electrophoresis System	EI0001	Thermo Fisher Scientific
TE22 Mighty Small Transfer Tank	TE 22	GE Healthcare Life Science
ProBlot 25D Economy Rocker	S2025-D-B	Labnet International
Odyssey CLx Imaging System	9120	LI-COR
BioFresh full room refrigerator		Liebherr
BioFresh full room freezer		Liebherr
BioPlus 600/660D	RF 600D	GRAM
HERAfreeze HFU T Series -86°C Ultra-Low Temperature Freezers	11650823	Thermo Scientific
Precision Balance Precisa ES 4200C 0,01g	360-9446-001	Precisa
Sartorius CP225D Analytical Balances	CP225D	Sartorius
nanoLC system	Dionex UltiMate 3000	Thermo Fisher Scientific
C_{18} trapping column	Acclaim PepMap 100 µm×2 cm, 5 µm, 100Å	Thermo Fisher Scientific
C18 column	Acclaim PepMap RSLC, 75 µm×50 cm, 2 µm, 100Å	Thermo Fisher Scientific
Q Exactive mass spectrometer		Thermo Fisher Scientific

Xcalibur	Version 4.0.27.13	Thermo Fisher Scientific
Proteome Discoverer	Version 2.0	Thermo Fisher Scientific
Snapgene	Version 4.7.3	GSL Biotech LLC
GraphPad	Prism 8.0	GraphPad Software
Image Studio Lite	Version 5.0	LI-COR Biosciences
Photoshop CC 2019	Version CC 2019	Adobe

Table 5-18 List of reagents

Name	Catalog Number	Company
RPMI 1640 (1×)	21875034	Gibco
DMEM (1×)	41965039	Gibco
FBS, Qualified	10270106	Gibco
DPBS (1×)	14190094	Gibco
Pen-Strep (100×)	15140122	Gibco
L-Glutamine 200mM (1×)	25030024	Gibco
Antibiotic-Antimycotic (100×)	15240062	Gibco
Opti-MEM	31985070	Gibco
HEPES	15630080	Gibco
Trypsin-EDTA (0.05%)	25300054	Gibco
Dimethyl sulfoxide p. a., ACS (min. 99.9%)	2347-1L	Labsolute-Chemsolute
Dimethyl sulfoxide	276855	Sigma-Aldrich
cOmplete, Mini, EDTA-free Protease Inhibitor Cocktail	4693159001	Roche
Benzonase Nuclease	E1014	Merckmillipore
DTT (dithiothreitol)	R0862	Thermo Scientific
SDS solution 20% pure A3942 BC	A3942,1000	AppliChem GmbH
Ponceau S solution BC	A2935, 0500	AppliChem GmbH
Glycerol	G7757-500ML	Honeywell
Formaldehyde solution 37%	4979.1	Carl Roth
Triton X-100	X100	Sigma-Aldrich
IGEPAL CA-630	I8896	Sigma-Aldrich
Sodium hydroxide 10 mol/L	480648	MerckMillipore
Nitrocellulose Blotting Membrane	4979.1	GE Healthcare Life Science

PageRuler Plus Prestained 10-250kDa Protein Ladder	26619	Thermo Fisher Scientific
NuPAGE MES SDS Running Buffer (20×)	NP0002	Invitrogen
Acetic acid 100%	3738.5	Carl Roth
NuPAGE™ 4-12% Bis-Tris Protein Gels, 1.0 mm, 10-well	NP0321BOX	Invitrogen
Albumin fraction V	90604-29-8	Carl Roth
TWEEN 20	P1379	Sigma-Aldrich
Tris-Solution, pH 6.8/7.4/7.8/8.0	Trizma-Base	Sigma-Aldrich
Methanol	106009	MerckMillipore
Acetone	100014	MerckMillipore
Ethanol	100983	MerckMillipore
Chloroform	1.02445.1000	MerckMillipore
Herculase II Fusion DNA Polymerase	600675	Agilent
dNTP mix (10 mM ea)	18427088	Invitrogen
UltraPure Ethidium Bromide, 10 mg / mL	15585011	Thermo Fisher Scientific
100 bp DNA Ladder	15628019	Invitrogen
1 Kb Plus DNA Ladder	10787018	Invitrogen
UltraPure Agarose	16500500	Invitrogen
Mix & Go Competent Cells - Strain JM109	T3003 / T3005	Zymo research
Shrimp Alkaline Phosphatase (rSAP)	M0371S	New England BioLabs
T4 DNA Ligase	M0202S	New England BioLabs
Pfu DNA Polymerase, recombinant	EP0501	Invitrogen
Taq DNA Polymerase, recombinant (5 U/µL)	EP0402	Invitrogen
Q5 High-Fidelity DNA Polymerase	M0491S	New England BioLabs
Puromycin Dihydrochloride	A1113803	Gibco

Trypan Blue solution	T8154	Sigma-Aldrich
Propidium iodides	P1304MP	Invitrogen
RNase A, and protease-free DNase (10 mg/mL)	EN0531	Thermo Scientific
Pierce biotin	29129	Thermo Scientific
Streptavidin High-Performance Spintrap	GE28-9031-30	GE Healthcare
MyeloCult M5300	05350	Stemcell Technologies
Recombinant Mouse Interleukin-3 (rm IL-3)	12340032	ImmunoTools
Recombinant Mouse Interleukin-6 (rm IL-6)	12340062	ImmunoTools
FastStart Essential DNA Green Master	6402712001	Roche
PFI-3	SML0939	Sigma-Aldrich
FastDigest enzymes		Thermo Fisher Scientific

Table 5-19 List of Antibodies

Antibody	Catalog Number	Company
PML antibody (H-238)	sc-5621	Santa Cruz Biotechnology
ATRX antibody (H-300)	sc-15408	Santa Cruz Biotechnology
ATRX antibody (D-5)	sc-55584	Santa Cruz Biotechnology
Anti-SMCHD1 antibody produced in rabbit	HPA039441	Sigma-Aldrich
Recombinant Anti-BRG1 antibody	ab110641	Abcam
A. aeolicus BPL/BioID2 Antibody	SS QD1	Novus Biologicals
Anti-Daxx Antibody	07-471	Sigma-Aldrich
Anti-Histone H3.3 Antibody	09-838	Sigma-Aldrich
Anti-Histone H3.3 Antibody, K27M mutant	ABE419	Sigma-Aldrich
Donkey anti-Rabbit IgG (H+L) Highly Cross-Adsorbed Secondary Antibody, Alexa Fluor 488	A-21206	Invitrogen
Donkey anti-Mouse IgG (H+L) Highly Cross-Adsorbed Secondary Antibody, Alexa Fluor 488	A-21202	Invitrogen
Donkey anti-Rabbit IgG (H+L) Highly Cross-Adsorbed Secondary Antibody, Alexa Fluor 555	A-31572	Invitrogen
Donkey anti-Mouse IgG (H+L) Highly Cross-Adsorbed Secondary Antibody, Alexa Fluor 555	A-31570	Invitrogen
Donkey anti-Rabbit IgG (H+L) Highly Cross-Adsorbed Secondary Antibody, Alexa Fluor 350	A-10039	Invitrogen
Donkey anti-Mouse IgG (H+L) Highly Cross-Adsorbed Secondary Antibody, Alexa Fluor 350	A-10035	Invitrogen
IRDye 800CW Donkey anti-Rabbit IgG Secondary Antibody	926-32213	LI-COR

IRDye 800CW Donkey anti-Mouse IgG Secondary Antibody	926-32212	LI-COR
IRDye 680LT Donkey anti-Rabbit IgG Secondary Antibody	926-68023	LI-COR
IRDye 680LT Donkey anti-Mouse IgG Secondary Antibody	926-68022	LI-COR
IRDye 800CW Streptavidin	926-32230	LI-COR
IRDye 680LT Streptavidin	926-68031	LI-COR
Antibody Flag M2, monoclonal	F1804	Sigma-Aldrich
Antibody β-Actin, polyclonal rabbit mAb	13E5	Cell Signaling

Table 5-20 List of Experiment Kits

Experiment Kit	Catalog Number	Company
Pierce Coomassie (Bradford) Protein Assay Kit	23200	Thermo Fisher Scientific
SDS Gel Preparation Kit	8091	Sigma-Aldrich
Pierce Silver Stain Kit	24612	Thermo Fisher Scientific
Zymoclean Gel DNA Recovery Kits	D4007/D4008	Zymo research
ZymoPURE Plasmid Miniprep	D4208T	Zymo research
ZymoPURE II Plasmid Maxiprep	D4202	Zymo research
ZymoPURE II Plasmid Midiprep	D4200	Zymo research
NEBuilder Hifi DNA Assembly Kit	E5520S	New England BioLabs
FuGENE® HD Transfection Reagent	E2311	Promega
ProFection® Mammalian Transfection System	E1200	Promega
Diff-Quik Stain Set	B4132-1A	Siemens
Nano-Glo® Dual-Luciferase® Reporter Assay System	N1610	Promega
Nuclear Complex Co-IP Kit	54001	Active Motif

Table 5-21 List of Plasticware, Mixed Laboratory Items

Plastic	Catalog Number	Company
Cell Culture Flask	TC, T75, Stand., Vent.	Sarstedt
Cell culture plate	6 well, 12 well, 24 well, 96 well	Nunc
PCR, Real-Time Reaction Plates	Light Cycler 480 Multiwell 96	Roche
Centrifuge 1.5 mL Tubes	EBA 12R	Eppendorf
Centrifuge 1.5 mL Tubes	Heraeus Biofuge fresco	Thermo Fisher Scientific
Centrifuge 96 Multiwell Plates	Heraeus Multifuge 3 S-R	Thermo Fisher Scientific
Pipettes	Pipetman	Gilson
Pipetus	Hirschmann® pipetus® pipette controller	Sigma-Aldrich
Pipetus Tips	2 mL, 5 mL, 10 mL, 25 mL	Sarstedt
Pipetten Biosphere Filtertips	10 µL, 100 µL, 200 µL, 1250 µL	Sarstedt
Pippetten, serologische	2 mL, 5 mL, 10 mL, 25 mL	Sarstedt
Heracell 240 CO_2 Incubator	H240-CO_2	Heraeus
Whatman	Filtration Paper	Sigma-Aldrich

Table 5-22 List of Primers

Name	Type	Sequence (5' to 3')
PCR FLAG-Daxx	Forward	GAATTCAAGGATCCCCATGGACTACAAGGATGACGATGACAAGATGGCCACCGCTAACAGCATCATCGTG
	Reverse	TTGAATTCTTAATCAGAGTCTGAGAGCACGATGATCTCTTCTGGATCG
Mutagenesis pBabe-Myc. BioID2-Xho1	Forward	CCCTCACTCCTTCTCTAGGCGCCGGCCGGATCCCACTCGAGTGGTACGTAGGAATTCGGAGACCCAAGCTGGC
	Reverse	GCCAGCTTGGGTCTCCGAATTCCTACGTACCACTCGAGTGGGATCCGGCCGGCGCCTAGAGAAGGAGTGAGGG
Mutagenesis pCMX-Sp100-Sal1	Forward	TCCCTAAGAAGAGGGTCAGGTAAAGAAGATGTCGACGATGCCAAGACTTGGCCTGCAGAATGTCAG
	Reverse	CTGACATTCTGCAGGCCAAGTCTTGGCATCGTCGACATCTTCTTTACCTGACCCTCTTCTTAGGGA
PCR Myc.BioID2-Sal	Forward	AAAGTCGACACCATGGAACAAAAACTCATCTCAGAAGAGGATCTCGAC
	Reverse	TTTGTCGACGAATTCTCACTCGAGGCTTCTTCTCAGGCTGAACTCG
Mutagenesis H3.3-FLAG Q5	Forward	GATGACGACAAGTAACCGCGGGCCCGGGAT
	Reverse	GTCTTTGTAGTCAGCACGCTCGCCACGGATG
Mutagenesis H3.3K27M	Forward	CTGGCCACCAAGGCGGCCCGTATGTCGGCGCCATCCACCGGCGGAGTG
	Reverse	CACTCCGCCGGTGGATGGCGCCGACATACGGGCCGCCTTGGTGGCCAG
HiFi assembly AB-ATRX-F1	Forward	CTCCGATTGACTGAGTCGCCCGGATCCATGACCGCTGAGCCCATG
	Reverse	GAGATGAGTTTTTGTTCCATAGCGTAATCTGGAACATCGTATGG
HiFi assembly AB-BioID-F2	Forward	ACGATGTTCCAGATTACGCTATGGAACAAAAACTCATCTCAGAAGAGGA
	Reverse	CCCTCGTCGAGGCCTGCAGGCTTGTCATCGTCATCCTTGTAGTCC
HiFi assembly BA-BioID2-F1	Forward	CTCCGATTGACTGAGTCGCCCGGATCATGGAACAAAAACTCATCTCAGAAGAGGA
	Reverse	GGCTCAGCGGTCATCTTGTCATCGTCATCCTTGTAG

HiFi assembly BA-ATRX-F2	Forward	TGACGATGACAAGATGACCGCTGAGCCCAT
	Reverse	CCCTCGTCGAGGCCTGCAGGTCAAGCGTAATCTGGAACATCGTATGG
SMARCA4.OE-F	Forward	CTCCGATTGACTGAGTCGCCCGGATCATGTCCACTCCAGACCCACCCCTGGGCGGA
	Reverse	CCCTCGTCGAGGCCTGCAGGTCAGTGATGGTGATGGTGGTGGTCTTCTTCGCTGCCAC
SMARCA4.VP16-F	Forward	GCTATTCCAGAAGTAGTGAAGAGGCTTTTTTGGAGGAGATCTAAGATGTCCACTCCAGACCCACCCCTGG
	Reverse	TTTTTAGGGCCCATGGTGGCGGCGTCAGTGATGGTGATGGTGGTGGTCTTCTTCGCTGCCAC
PCR SMARCA4	Forward	AAAAAAGCTAGCATGTCCACTCCAGACCCACCCCTGG
	Reverse	AAAAAAACCGGTTTGTGATGGTGATGGTGGTGGTCTTCTTCGCTGC

Table 5-23 List of Buffers

Buffer name	Component
Laemmli sample buffer (2×)	160 mM Tris·HCl, pH 6.8, 20% Glycerol, 4% SDS, 0.02% Bromophenol blue, 200 mM DTT.
Subcellular protein fractionation Buffers	10× Lysis Hypotonic Buffer: 100 mM HEPES, pH 7.9, 15 mM $MgCl_2$, 100 mM KCl. 5× Lysis Isotonic Buffer: 50 mM Tris·HCl, pH 7.5, 10 mM $MgCl_2$, 15 mM $CaCl_2$, 20% (v/v) Glycerol. Extraction Buffer: 20 mM HEPES, pH 7.9, 1.5 mM $MgCl_2$, 0.42 M NaCl, 0.2 mM EDTA, 25% (v/v) Glycerol. 3× Dilution and Equilibration Buffer: 60 mM HEPES, pH 7.9, 4.5 mM $MgCl_2$, 0.6 mM EDTA, 30 mM KCl. 1 mM DTT: 1 M DTT stock solution in deionized water. 10% NP-40: 10% IGEPAL CA-630 in deionized water.
BioID2 pull-down regents and solutions	Biotin: 1 mM (20 ×), dissolve 12.2 mg biotin in 50 mL of serum-free RPMI 1640. RIPA Lysis buffer: 50 mM Tris·HCl, pH 7.4, 150 mM NaCl, 2 mM $MgCl_2$, 2% (w/v) Triton X-100, 0.5% (w/v) deoxycholic acid, 0.5% (w/v) SDS, 1 mM EDTA, 1 mM DTT (add just before use), 1 × Protease Inhibitor Cocktail (EDTA-Free; add just before use). Wash buffer 1: 2% SDS (w/v). Wash buffer 2: 50 mM HEPES, pH 7.5, 500 mM NaCl, 1% (w/v) Triton X-100, 0.1% (w/v) deoxycholic acid, 1 mM EDTA. Wash buffer 3: 10 mM Tris·HCl, pH 7.4, 250 mM LiCl, 0.5% (w/v) NP-40, 0.5% (w/v) deoxycholic acid, 1 mM EDTA.
Sliver staining solution	Fixing Solution: 30% ethanol, 10% acetic acid. Stop Solution: 5% acetic acid. Ethanol Wash: 10% ethanol.
Western blotting Transfer Buffer	25 mM Tris, 192 mM glycine, 10% methanol.
20 × MES SDS Running Buffer	50 mM MES, 50 mM Tris Base, 0.1% SDS, 1 mM EDTA, pH 7.3.
DNA extraction buffer	Mix 192 mL of 0.2 M Na_2HPO_4 with 8 mL of 0.1% Triton X-100 (v/v). Adjust the pH to 7.8.
DNA staining solution	Dissolve 200 mg of PI in 10 mL of PBS. Add 2 mg of DNase free RNase.

6. Results

6.1. PML-NB Reconstitution after ATRA Treatment

As mentioned in 4 “Goal”, we chose to use NB4, an APL cell line, in which visible PML-association Daxx/ATRX complex is naturally disrupted but is reliably restored with ATRA treatment. To confirm that our system behaves as previously described, we performed immunofluorescence (IF) staining for ATRX (Green), Daxx (Red), and PML (Blue) to compare their localization in both ATRA-treated (Figure 6-1 Upper) and control cells (Figure 6-1 Lower). The treatment time was 7 days.

In both treated and untreated cells, ATRX is located in the nucleus, while Daxx and PML can be found in both the cytosol and the nucleus. However, distinct dots in which the three proteins colocalized only showed up in the nucleus of ATRA-treated NB4 cells. Our results suggest that our experimental model worked as expected, i.e., PML-NBs and the PAX-complex were indeed reconstituted after ATRA treatment.

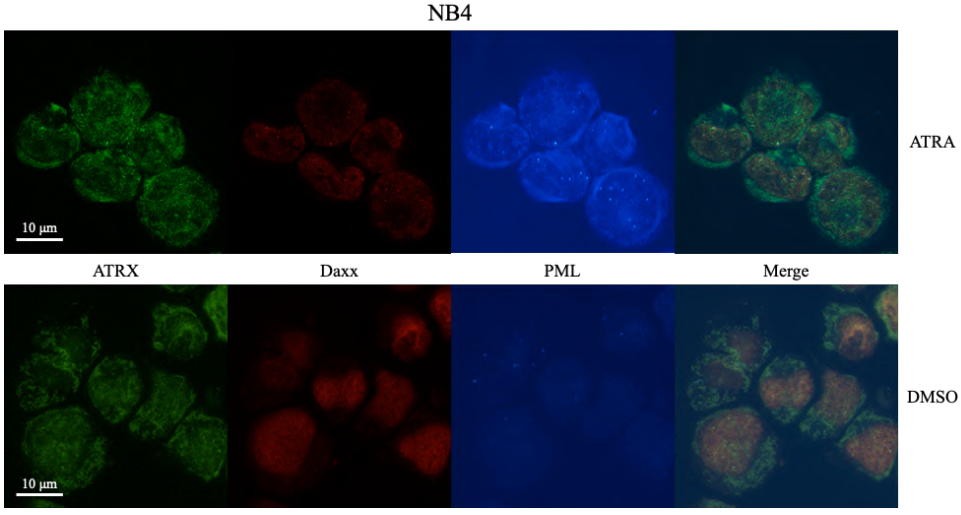


Figure 6-1 Immunofluorescence (IF) staining with ATRX (Green), Daxx (Red), and PML (Blue) antibodies in both ATRA-treated (Upper) and control (Lower) NB4 cells, showing the PAX-complex were reconstituted after ATRA treatment. Scale bar is 10 μm.

6.2. Study of PAX by Pull-Down of Biotinylated Interactors through a Fusion BioID2-Daxx Protein

6.2.1. BioID2-Daxx Generation and Validation

6.2.1.1. Creation of pLeGO-BioID2-Daxx

The generation of the pLeGO-BioID2-Daxx expression plasmid, in which Daxx is fused in-frame with the BioID2 ligase, required the use of PCR. As the Daxx is known to interact with SUMO and PML through its C-terminus, we added the ligase to the N-terminus of the protein for concerns that it would otherwise interfere with these interactions. Moreover, as seen in the corresponding plasmid map（Figure 9-1）, the fusion protein (Figure 9-7) has a myc-tag at its N-terminus and a FLAG-tag between the BioID2 and Daxx parts.

As the LeGO lentiviral vector contains a strong eukaryotic promoter taken from the spleen focus-forming virus (SFFV) U3 region, it can be directly used to transiently transfect cell lines. This allowed us to produce and perform an initial validation of the properties of the fusion protein under S1-level safety conditions, i.e., without the need of producing S2-level lentiviral particles. The purpose was to exclude unexpected side effects of BioID2 on stability of the protein and viability of transfected cells.

6.2.1.2. BioID2-Daxx validation in MCF-7 cells

For the initial characterization of the BioID2-Daxx fusion protein, we performed transient transfection experiments in MCF-7, a human breast cancer cell line with well-defined overall morphological features and PML-NBs in particular, as this would allow us to exclude unexpected effects of the BioID2 domain on the overall behavior of the exogenously expressed Daxx protein. Thus, we grew MCF-7 cells on

coverslips, transiently transfected them with pLeGO-BioID2-Daxx, and examined them by immunofluorescence microscopy on the next day. To obtain the best possible result of IF staining by using an expression level in a physiological range, a titration of the plasmid DNA amount for transfection was performed (no plasmid control, 200, 500, 1,000, and 2,000 ng).

As shown in Figure 6-2, plasmid DNA amounts lower than 1,000 ng resulted in an ambiguous IF staining pattern for FLAG and Daxx. On the other hand, when the plasmid DNA amount was 2,000 ng, the overall signal was so strong that no distinct localization could be observed. These problems were averted when 1,000 ng of pLeGO-BioID2-Daxx plasmid DNA was used, so this amount allows the best IF staining. Figure 6-3 shows a magnified view of the part of Figure 6-2 corresponding to the FLAG and Daxx staining of cultures transfected with 1,000 ng of plasmid.

Comparing the FLAG and Daxx staining patterns, we observed that all cells gave Daxx signal, but not all cells displayed FLAG staining. The latter were non-transfected cells, i.e., they only expressed the endogenous form of Daxx, and allowed us to observe its characteristic speckled pattern. In cells where both Daxx and FLAG staining was observed, i.e., successfully transfected cells expressing BioID2-Daxx, the Daxx and FLAG speckles colocalized, which is to be expected as the fusion protein is recognized by both antibodies. Importantly, the speckled pattern that the fusion protein formed was identical to the one of the endogenous Daxx in the non-transfected cells. Thus, we concluded that the fusion protein localizes physiologically in the PML-NBs and can be used to search for interactors of the PAX complex.

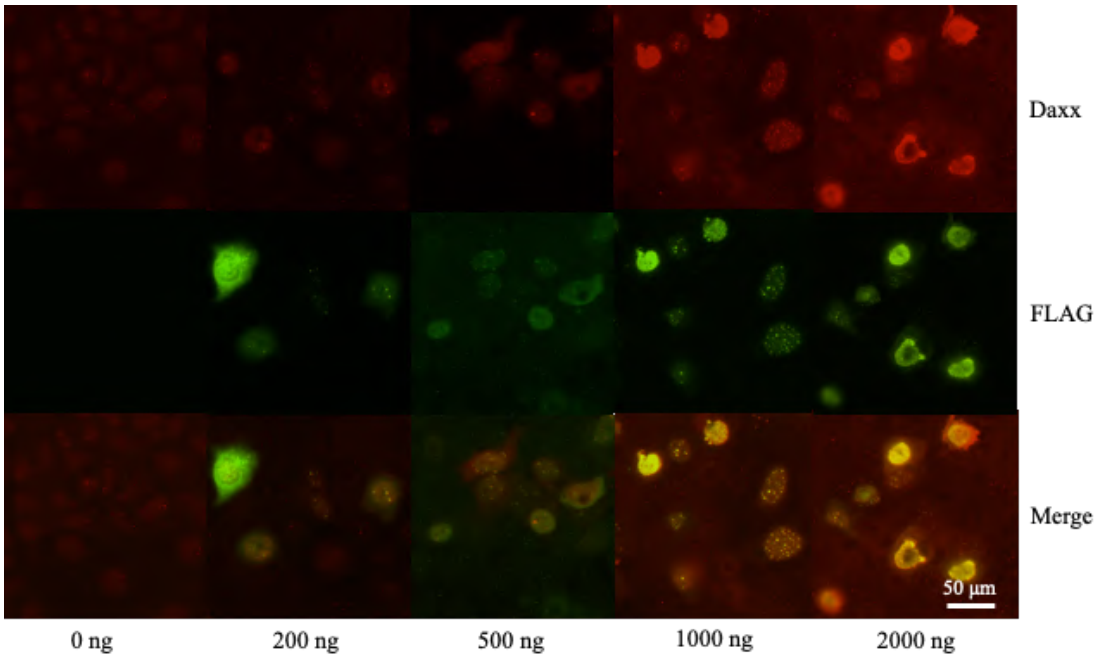


Figure 6-2 Immunofluorescence (IF) staining with anti-Daxx (Red) and anti-FLAG (Green) of MCF-7 cells transiently transfected with various amounts (0, 200, 500, 1000 and 2000 ng) of pLeGO-BioID2-Daxx, seeing when 1,000 ng of pLeGO-BioID2-Daxx plasmid DNA was used, we got the best staining. Scale bar is 50 µm.

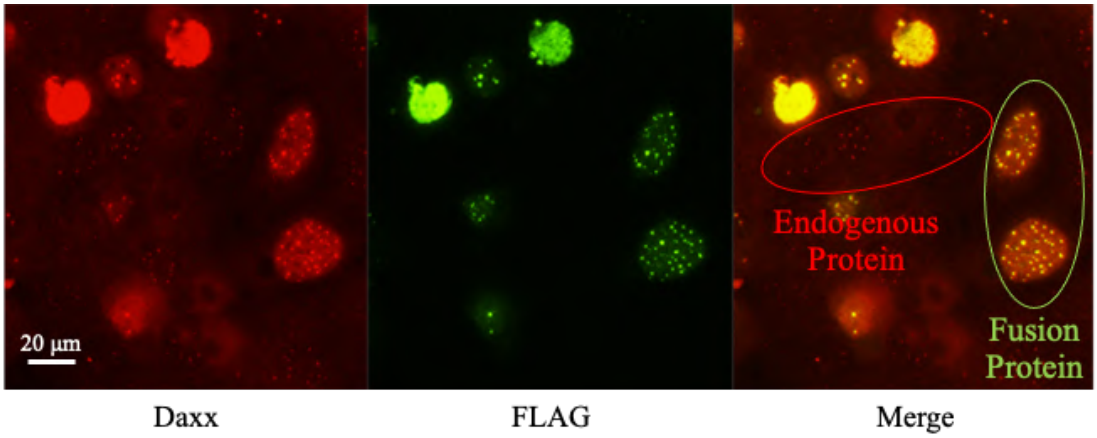


Figure 6-3 A magnified view of the part of Figure 6-2 corresponding to the FLAG (Green) and Daxx (Red) staining of cultures transfected with 1,000 ng of plasmid, seeing all cells gave Daxx signal, and non-transfected cells only expressed the endogenous Daxx but not FLAG staining; while transfected cells expressing BioID2-Daxx, showing Daxx and FLAG staining, and the Daxx and FLAG speckles colocalized. Scale bar is 20 µm.

6.2.1.3. BioID2-Daxx validation in HEK293T cells

To further validate the BioID2-Daxx fusion protein, we transiently transfected the pLeGO-BioID2-Daxx plasmid into HEK293T cells grown in 6-well plates (1,000 ng of plasmid per well). For comparison purposes, parallel cultures were transfected with pCMX-Daxx or pLeGO-FLAG-Daxx, which express non-tagged and FLAG-tagged Daxx, respectively. Four hours after transfection, 50 μM Biotin was added to allow for biotinylation. After another 20 h, cells were washed with room-temperature PBS to remove serum proteins and then lysed for immunoblot analysis using 2× Laemmli sample buffer (200 μL per 10^6 cells). After transfer, membranes were blocked and incubated with anti-cMyc, anti-Daxx, and anti-FLAG to confirm BioID2-Daxx fusion protein expression and migration by SDS-PAGE. Blots were scanned with a LI-COR Imaging System and are displayed in Figure 6-4.

As expected, all cells gave a band in the anti-Daxx blot, which corresponds to non-tagged Daxx (endogenous or expressed by the pCMX-Daxx plasmid). It must be noted that, even though the calculated molecular weight of Daxx is 81 kDa, the protein runs aberrantly at an apparent molecular weight of 110 kDa, possibly due to hyper-phosphorylation[149]. Cells transfected with the pLeGO-FLAG-Daxx construct also produce FLAG-tagged Daxx, which forms a band in both the anti-Daxx (not discernible from the one of the endogenous Daxx due to the small difference in MW) and the anti-FLAG blot. Finally, cells transfected with pLeGO-BioID2-Daxx gave a large band at around 135 kDa observed in all three antibody blots. This is expected, as BioID2-Daxx is significantly larger (by about 25 kDa, i.e., close to the MW of BioID2) than Daxx and contains Myc, FLAG, and Daxx epitopes. Even though the observed mobility (135 kDa) is significantly higher than the calculated MW of the fusion protein (111 kDa), the difference can be attributed to an anomalous electrophoretic behavior of the Daxx part of the protein, similar to the one observed in Daxx itself. These results indicated that the fusion protein remains intact, a prerequisite for being functional.

Moreover, the fact that the total signal and the number of bands in the Streptavidin blot is higher only in the fusion protein-producing cells confirms that the fusion protein exhibits the expected BioID2 enzymatic activity.

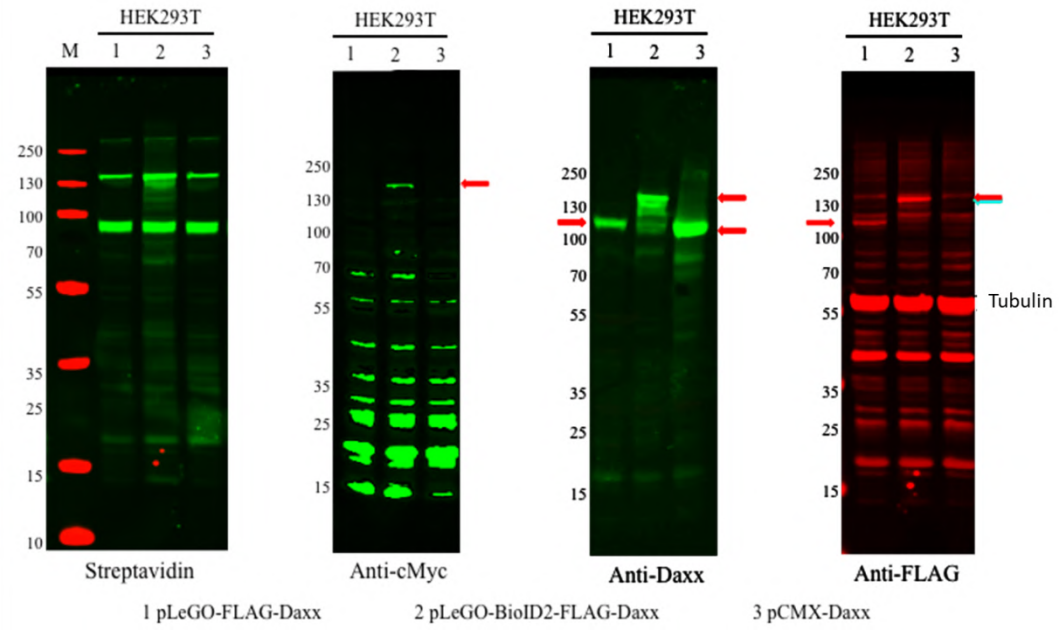


Figure 6-4 Validation of pLeGO-BioID2-Daxx in HEK293T cells by Western blot. Seeing all cells gave a band in the anti-Daxx blot, which corresponds to non-tagged Daxx. Red arrows point to the various forms of Daxx (endogenous/FLAG-tagged/BioID2-Daxx). Cells transfected with the (lane 1) pLeGO-FLAG-Daxx construct produced FLAG-tagged Daxx, which forms a band in both the anti-Daxx and the anti-FLAG blot. Cells transfected with (lane 2) pLeGO-BioID2-Daxx gave a large band at around 135 kDa observed in all three antibody blots, which contains Myc, FLAG, and Daxx epitopes. (Lane 3, overexpress untagged Daxx from Plasmid pCMX-Daxx).

6.2.1.4. Generation of NB4 Cells Stably Expressing BioID2-Daxx

After the results of the aforementioned transient transfection experiments, we were confident enough to proceed to the generation of NB4 cells stably producing BioID2-Daxx by viral transduction of the pLeGO-BioID2-Daxx vector. Lentiviral packaging and transduction were performed as described in section 5.1.9. Following transduction, we performed selection with 0.5 µg/mL Puromycin. The specific concentration was chosen because, as determined by a preparatory experiment, it is able to kill more than 90% of cells in a culture of parental (i.e., non-transfected) NB4 cells within 7 days (Figure 6-5).

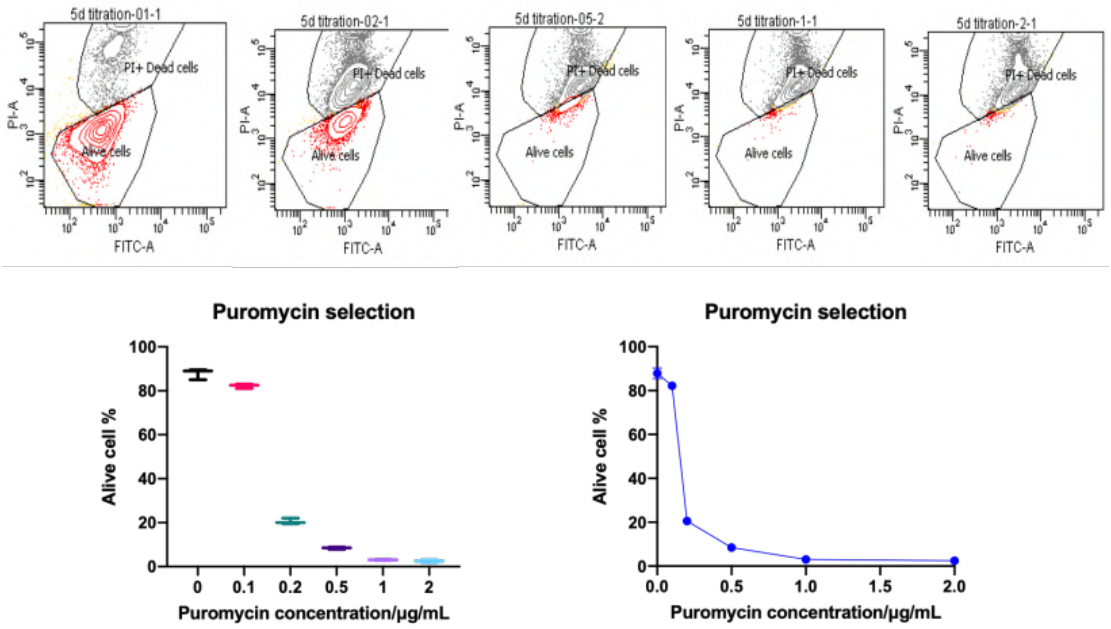


Figure 6-5 Non-transfected NB4 cells viability on Day 7 with various Puromycin concentration treatment, showing with 0.5 µg/mL Puromycin, which was able to kill more than 90% of cells in a culture of non-transfected NB4 cells within 7 days. Upper Panel: FACS analysis showed non-transfected NB4 cells viability on Day 7 using different Puromycin concentration (DMSO, 0.1, 0.2, 0.5, 1 and 2 µg/mL). Grey dots: Dead NB4 cell population; Red dots: Normal NB4 cell culture. Lower Panel: Puromycin selection Curve of FACS data. Data are represented as mean ± SD.

The Puromycin selection was kept in the medium for 60 days; during the selection, we used FACS to gate the GFP-positive cells to monitor the presence of the pLeGO-BioID2-Daxx construct. As shown in Figure 6-6, the proportion of GFP-positive cells rapidly increased during the first 10-20 days but reached a plateau of 70% after 25 days of selection.

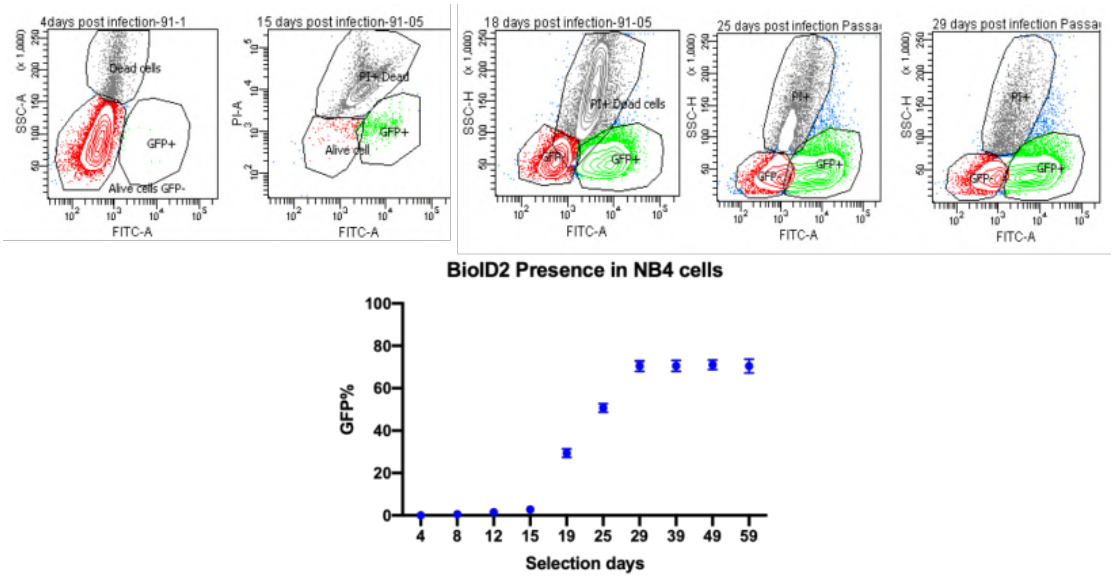


Figure 6-6 Upper: FACS analysis showing the abundance of GFP^+ cells (i.e., harboring the pLeGO-BioID2-Daxx construct) of Puromycin selection on the presence of the pLeGO-BioID2-Daxx construct in transduced NB4 cultures. Grey dots: Dead NB4 cell population; Red dots: GFP-negative in transfected NB4 cell culture; Green dots: GFP-positive in transfected NB4 cell culture; Lower: GFP+ cells population percent curve within 60 days of Puromycin selection, seeing the proportion of GFP-positive cells rapidly increased during the first 10-20 days but reached a plateau of 70% after 25 days of selection. All data are represented as mean ± SD.

Cell sorting was performed to obtain a more homogenous population (99% GFP-positive cells) (Figure 6-7). This percentage turned out to be stable in further subculturing, indicating a lack of counter-selection against BioID2-Daxx overexpression. The created stable NB4 subline was christened “NB4-BD”.

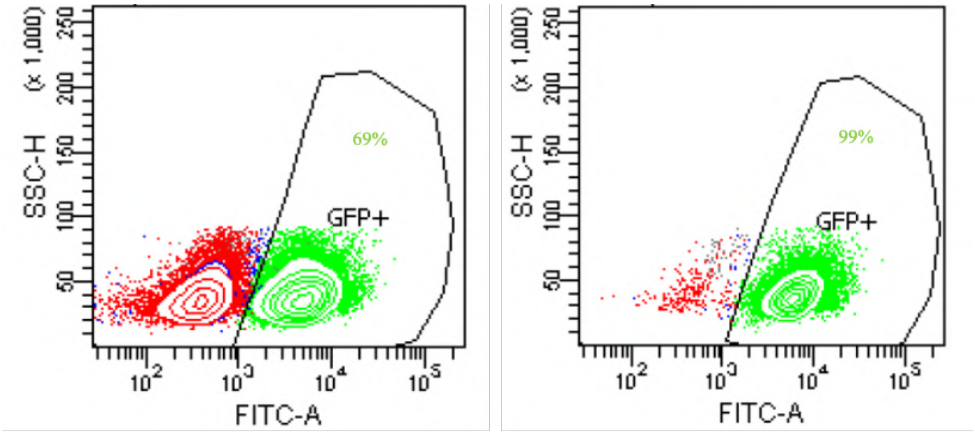


Figure 6-7 FACS analysis showing the abundance of GFP^+ cells (i.e., harboring the pLeGO-BioID2-Daxx construct) in a transfected NB4 culture before (Left) and after (Right) sorting, resulting in obtaining a homogenous population (99% GFP-positive cells) NB4-BD. Red dots: GFP-negative in transfected NB4 cell culture; Green dots: GFP-positive in transfected NB4 cell culture.

6.2.1.5. First Validation of BioID2-Daxx in NB4-BD Cells

Now that we had the NB4-BD subline, we were able to perform another validation of the BioID2-Daxx fusion protein. To this end, we compared the localization of the BioID2-Daxx fusion protein in both ATRA-treated and untreated NB4-BD cells by IF staining. The duration of treatment was 7 days.

As we can see in Figure 6-8, only ATRA-treated cells exhibited clear dot staining with all antibodies. Moreover, treated cells displayed a clear co-localization of FLAG with ATRX, Daxx, and PML, which means that reconstituted PML-NBs successfully incorporated the BioID2-Daxx fusion protein. It is worth noting that some cells with very high BioID2-Daxx expression exhibited some level of colocalization even in absence of ATRA, suggesting that the disruption of PAX can be partially overcome by an overabundance of Daxx (panels DMSO; Daxx and ATRX).

These observations strongly suggested that the NB4-BD subline is an experimental system suitable for the identification of physiological Daxx interactors. Moreover, they indicated that ATRA treatment would be a meaningful control for our further studies.

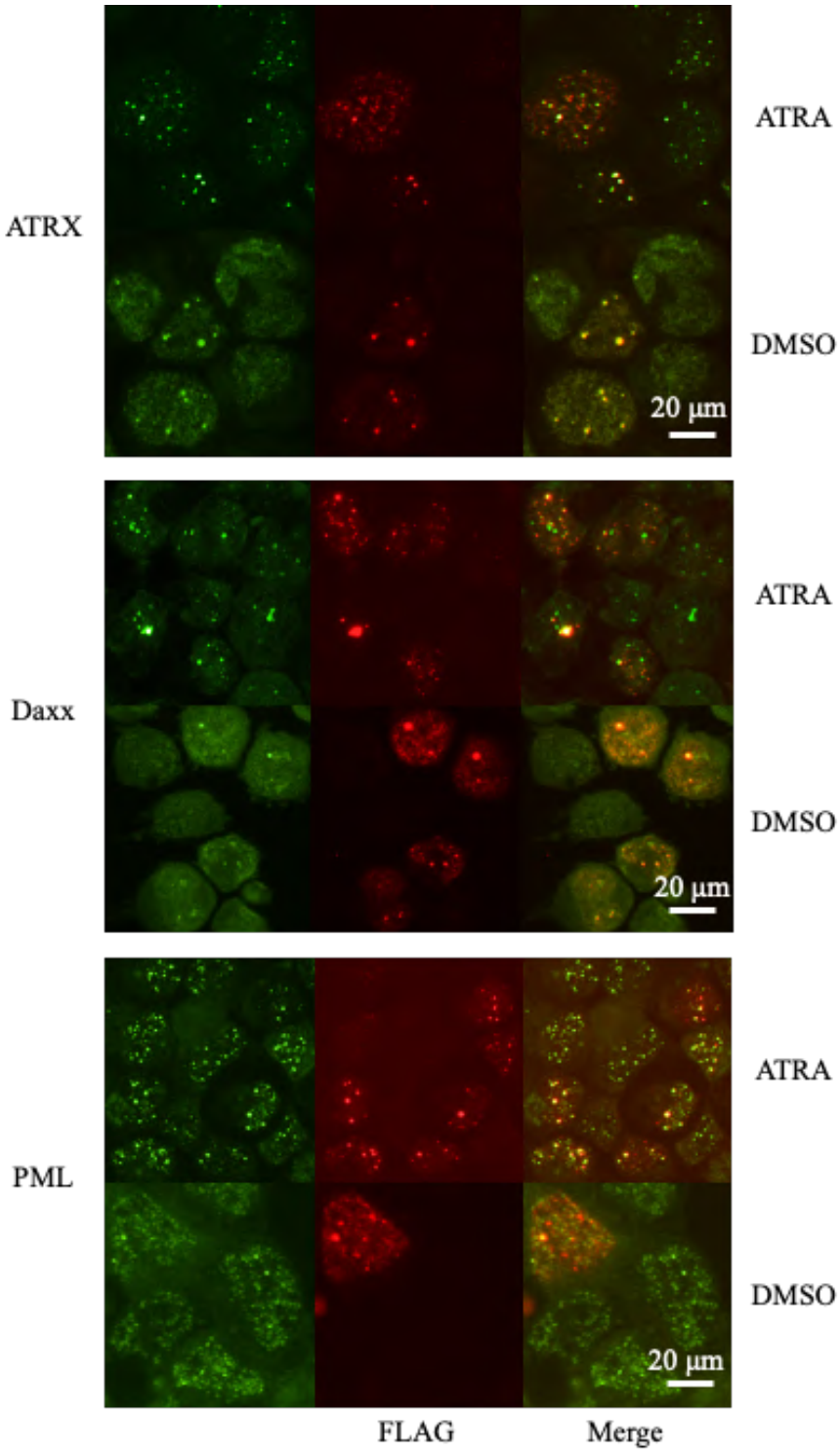


Figure 6-8 Immunofluorescence (IF) staining with anti-ATRX/Daxx/PML (Green), anti-FLAG (Red) antibody of NB4-BD cells for comparison of the localization of the BioID2-Daxx fusion protein in both ATRA treated and untreated condition, seeing only ATRA-treated cells exhibited clear dot staining, which displayed a clear co-localization of FLAG with ATRX, Daxx, and PML. Scale bar is 20 µm.

6.2.1.6. ATRA Treatment Time Course for NB4-BD Cells

In our previous experiments with NB4 and NB4-BD cells, ATRA treatment lasted 7 days. At the end of this time, NB-bodies were well reconstituted, which allowed us to make the observations mentioned in the corresponding chapters. However, this time frame was significantly longer than the one used by other researchers[64].

As mentioned above, ATRA treatment does not just reconstitute PML-NBs in NB4 cells, but also leads to their differentiation and, ultimately, cells develop signs of senecsence. As the latter two processes were bound to compromise our effort to search for interactors by comparing pro- with post-reconstitution conditions, we decided to determine the duration of ATRA treatment which resulted in our system giving the most robust differences between treated and untreated NB-BD cells.

To this end, NB4-BD cultures were treated with ATRA or DMSO for various times (1, 3, 5, 7, and 9 days), with Biotin added 20 h before the end of treatment time, and then used for protein extraction and western blot analysis with Streptavidin (Figure 6-9 Left) and Tubulin (Figure 6-9 Right).

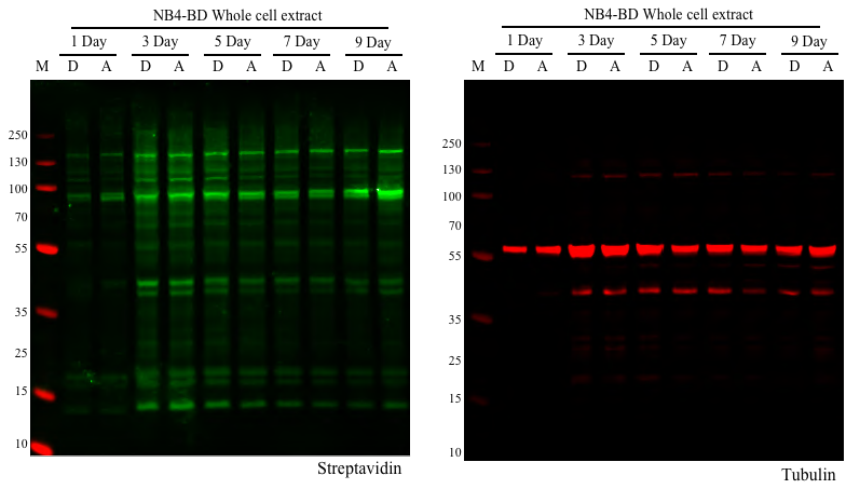


Figure 6-9 Western blots using the whole cell extract of the NB4-BD cells with ATRA- and DMSO-treatment for different time (1, 3, 5,7 and 9 days) with Streptavidin (Left, green bands) and anti-Tubulin (Right, red bands, ~60 kDa, as a loading control).

We used the Image Studio Lite software to quantify the total Streptavidin signal intensity as well as Tubulin loading control signal intensity (Figure 6-10 A), and then normalized the former using the latter (Figure 6-10 B). The resulting values were used to calculate ATRA:DMSO signal ratios for each treatment time (Figure 6-10 C). Since the highest signal ratio was the one corresponding to the 3-day treatment, this duration was adopted for all further experiments.

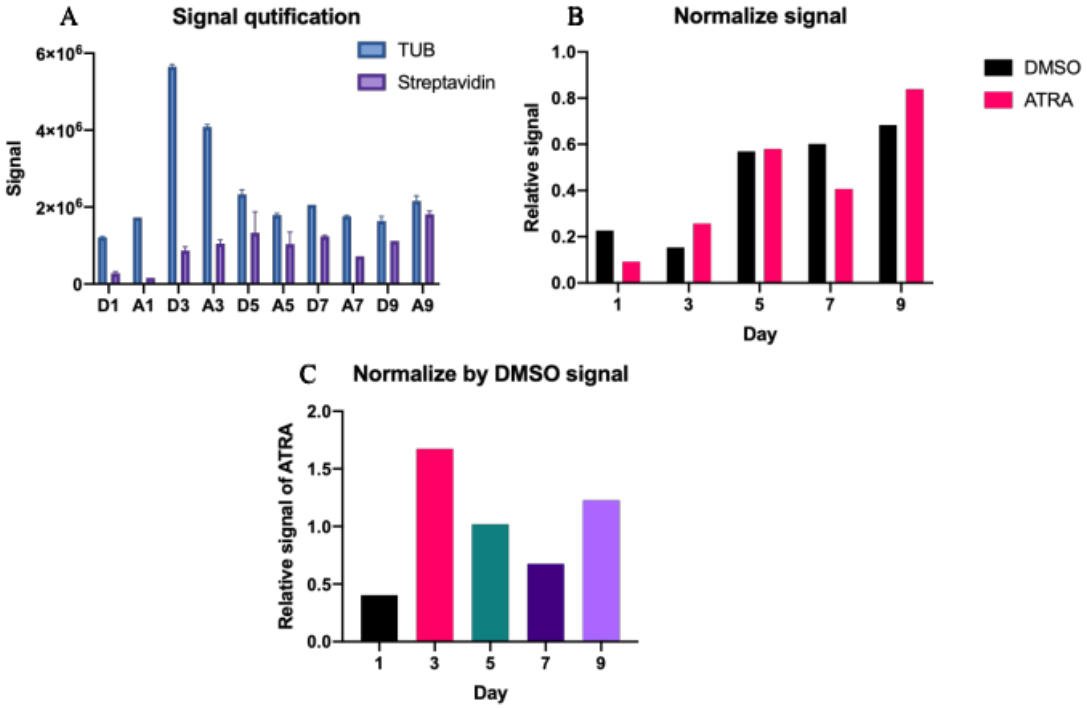


Figure 6-10 (A) Bar Graphs showing the total Streptavidin signal intensity (Purple) and the Tubulin band intensity (Blue) for each lane of the blots of Figure 6-9 (the NB4-BD cells with ATRA- and DMSO-treatment for different time (1, 3, 5,7 and 9 days)) as quantified using Image Studio Lite, all data are represented as mean ± SD; (B) Streptavidin signal for each lane (the NB4-BD cells with ATRA- and DMSO-treatment for different time (1, 3, 5,7 and 9 days)) after normalization using Tubulin; (C) the ATRA:DMSO signal intensity ratio for each treatment time (1, 3, 5,7 and 9 days), indicating the highest signal ratio was 3-day treatment.

6.2.1.7. Second Validation of BioID2-Daxx in NB4-BD cells

In section 6.2.1.5 “First Validation of BioID2-Daxx in NB4-BD Cells”, we showed that treatment of NB4-BD cells with ATRA for 7 days results in the reconstitution of PML-NBs containing the fusion protein. On the other hand, the time course in section 6.2.1.6 “ATRA Treatment Time Course for NB4-BD Cells” cells demonstrated that, with respect to the difference in biotinylation between ATRA-treated and control cells, a treatment time of 3 days is the best option. To be on the safe side, we decided to repeat the validation using the new treatment conditions before proceeding to the next stage, i.e., the pull-down experiment.

To this end, NB4-BD cells were again treated for 3 days with ATRA or DMSO, with Biotin added 20 h before collection. However, this time blotting of total extracts was not only performed with streptavidin but also with antibodies for FLAG, BioID2, and Daxx. Moreover, IF staining was performed with anti-BioID2 and anti-Daxx.

The western blot analysis (Figure 6-11) confirmed the expression and integrity of the fusion protein (135 kDa band), which is recognized by all three antibodies. As mentioned in section 6.2.1.3, the difference between the observed mobility (135 kDa) and the calculated MW of the fusion protein (111 kDa) can be explained by an anomalous electrophoretic behavior of the Daxx part of the protein, similar to the one that has been observed for Daxx itself[149]. Moreover, the Streptavidin staining gave significant differences in the biotinylation between ATRA-treated cells and controls, confirming the results of the time-course experiment.

On the other hand, the IF analysis clearly shows that tree days of ATRA treatment causes the reconstitution of NB-bodies incorporating the fusion protein (recognized by anti-Daxx and anti-FLAG staining) (Figure 6-12).

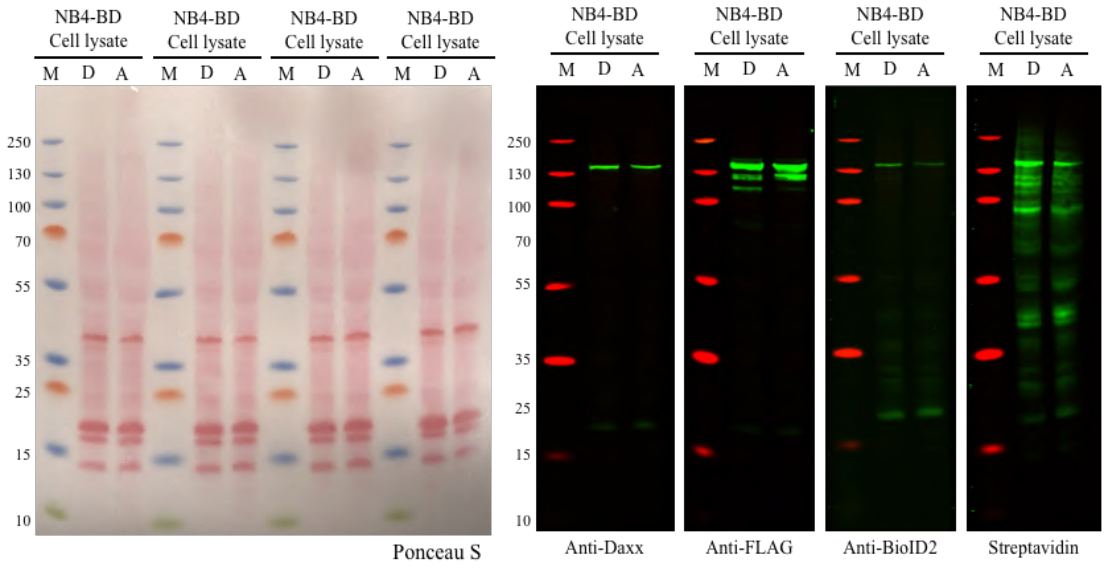


Figure 6-11 Left: Ponceau S staining of the membrane prior to blotting by the whole cell extract of the NB4-BD cells with ATRA- and DMSO-treatment, served as loading control; Right: Western blot to confirm the expression of BioID2-Daxx fusion protein by NB4-BD whole cell extract using anti-Daxx, anti-FLAG, anti-BioID2, and Streptavidin antibodies.

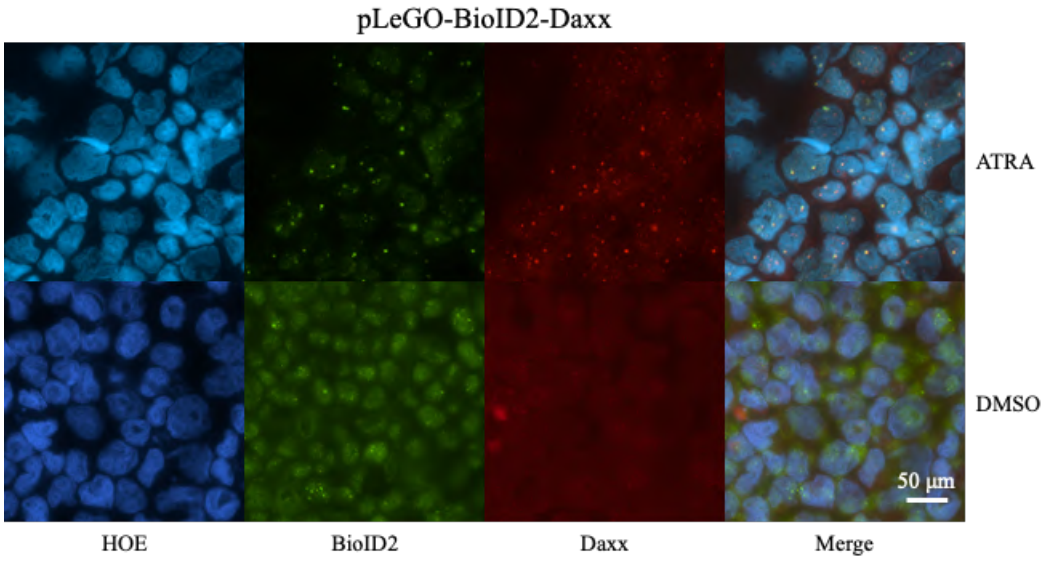


Figure 6-12 Immunofluorescence (IF) staining with Hoechst (Blue) staining for nucleus, BioID2 (Green), Daxx (Red) antibody of NB4-BD cells treated for 3 days with ATRA or DMSO, showing that 3 days of ATRA treatment causes the reconstitution of NB-bodies incorporating the fusion protein (recognized by anti-Daxx and anti-FLAG staining). Scale bar is 50 µm.

6.2.1.8. Comparison of Morphology, Cell Cycle Profile, and Proliferation between NB4 and NB4-BD cells

To verify, if NB4-BD cells still behave physiologically, we have conducted a series of experiments addressing specific properties comparing NB4-BD cells with parental NB4 cells.

Morphology

Diff-Quik staining did not reveal any differences between the parental line and the three sublines, under either control (DMSO) or ATRA-treated conditions. In both sublines, ATRA caused differentiation of cells, turning the shape of the nucleus from round to irregular (Figure 6-13). Indicating that all sublines behave physiologically, and ATRA-induced differentiation is still functioning.

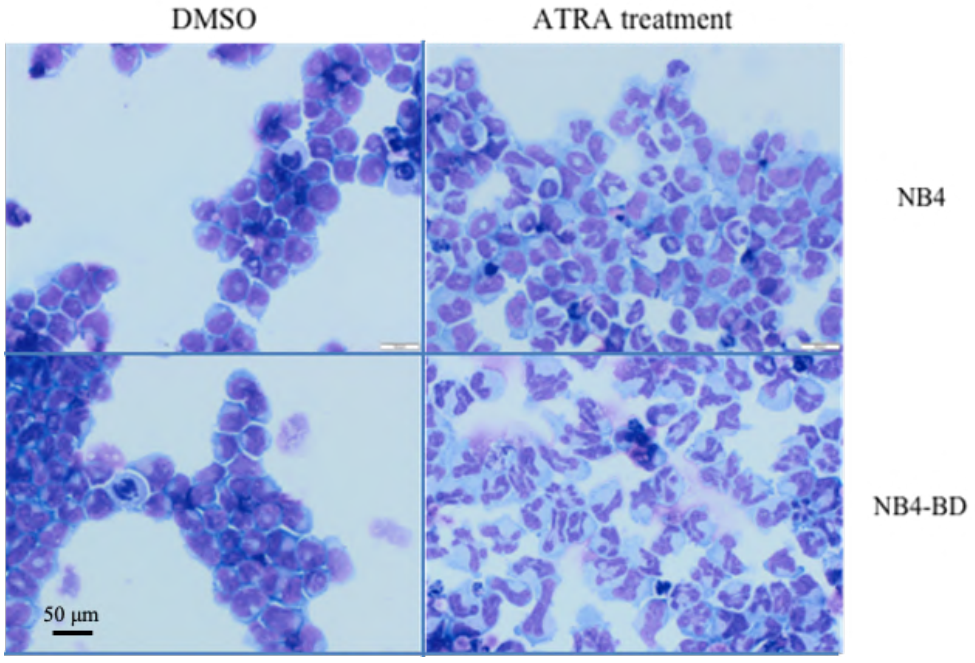


Figure 6-13 Diff-Quik staining analysis morphology of DMSO- and ATRA-treated the parental NB4 line and the NB4-BD subline, nucleus (purple) and cytoplasm (light blue), showing similar morphological changes between NB4 and NB4-BD cell after ATRA treatment. Scale bar is 50 µm.

Proliferation

Parental and subline cells were treated with DMSO or ATRA and passaged 1:2 every 2 days. At every passage, cell numbers were determined by FACS analysis.

The proliferation curves did not reveal significant changes among the various lines under control conditions (DMSO). Moreover, ATRA treatment caused the growth of all lines to stop (Figure 6-14). We conclude that the production of the fusion protein does not interfere with cell growth.

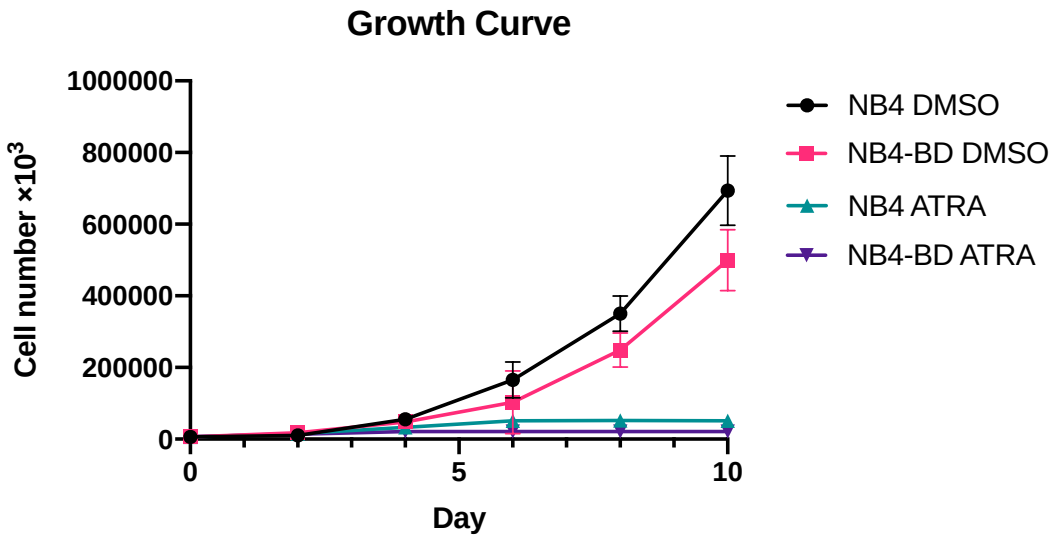


Figure 6-14 Growth curves for DMSO- and ATRA-treated parental NB4 line and the NB4-BD subline, showing identical ATRA induced growth arrest. All data are represented as mean ± SD.

Cell cycle profile

To compare the cell cycle profile of parental NB4 cells and the NB4-BD subline, we performed cell cycle analysis using FACS.

As seen in Figure 6-15, the profile of the NB4-BD subline was similar to the one of the parental cells, both under control (DMSO) and ATRA treatment conditions. In both lines, ATRA causes a G1-S growth arrest.

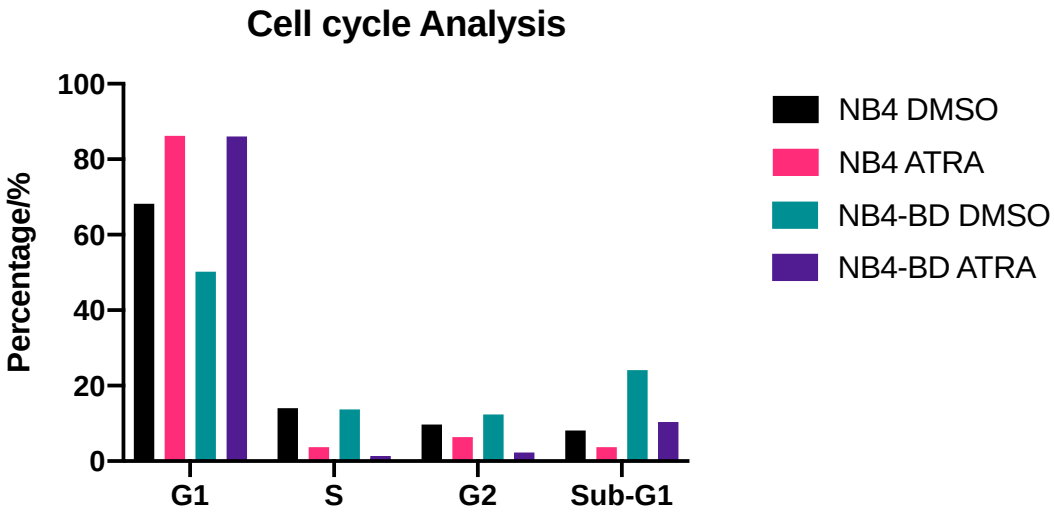


Figure 6-15 Cell cycle profiles of NB4 and NB4-BD cells treated with DMSO or ATRA for 3 days as assessed by FACS analysis, the bar graph shows DMSO- and ATRA-treated parental NB4 line and the NB4-BD subline at various stages of cell cycle, showing ATRA causes a similar G1-S growth arrest.

In summary, these control experiments have proven that in all aspects tested, NB4-BD cells exhibit a behavior similar to parental NB4 cells, meaning that expression of the BioID2 fusion protein did not substantially alter the cells in an unphysiological way.

6.2.1.9. Mass Spectrometry-Scale BioID2 Pull-Down of ATRA-Treated NB4-BD Cells and Controls

In order to isolate a sufficient amount of biotinylated proteins that can be analyzed and identified by mass-spec analysis, we performed a large scale BioID2 pull-down experiment in which 10×10^6 cells were used per condition. Besides treated and untreated NB4-BD cells, our experiment utilized parental NB4 cells (both treated and untreated) as additional negative controls.

This allowed us a four-way comparison. Parental NB4 samples (ATRA- or DMSO-treated) were used to define the background of the experimental system (negative control), allow us to filter out naturally occurring biotinylation (i.e., biotinylation not dependent on BioID2) from the corresponding NB4-BD samples and method-dependent systematic occurring signals (Background signal). Moreover, because DMSO-treated NB4-BD cells have disrupted PML-NBs, comparison of biotinylated proteins between DMSO- and ATRA-treated NB4-BD cells allows us to determine interacting proteins that are independent of PML-NB integrity and, thus, irrelevant to our PAX-focused study. As a whole, this process allows us to filter the biotinylated proteins of the ATRA-treated NB4-BD cells whose biotinylation is BioID2-dependent, thus it the result of their proximity to the fusion protein, and PAX-specific, i.e., it can only be observed in reconstituted PML-NBs.

Up to the collection of cell extracts, we followed the process described in section 5.1.12. After collecting the extracts, we incubated them with Streptavidin beads, which strongly bind biotinylated proteins. From each sample, a small aliquot of beads was eluted with 2× Laemmli sample buffer, and the eluates were electrophoresed and silver-stained (Figure 6-16, left panel). This method allowed us a first glimpse of the results of the pull-down. The differences among the samples clearly suggested that the experiment had been successful since there seemed to be several bands specific to ATRA-treated NB4-BD cells.

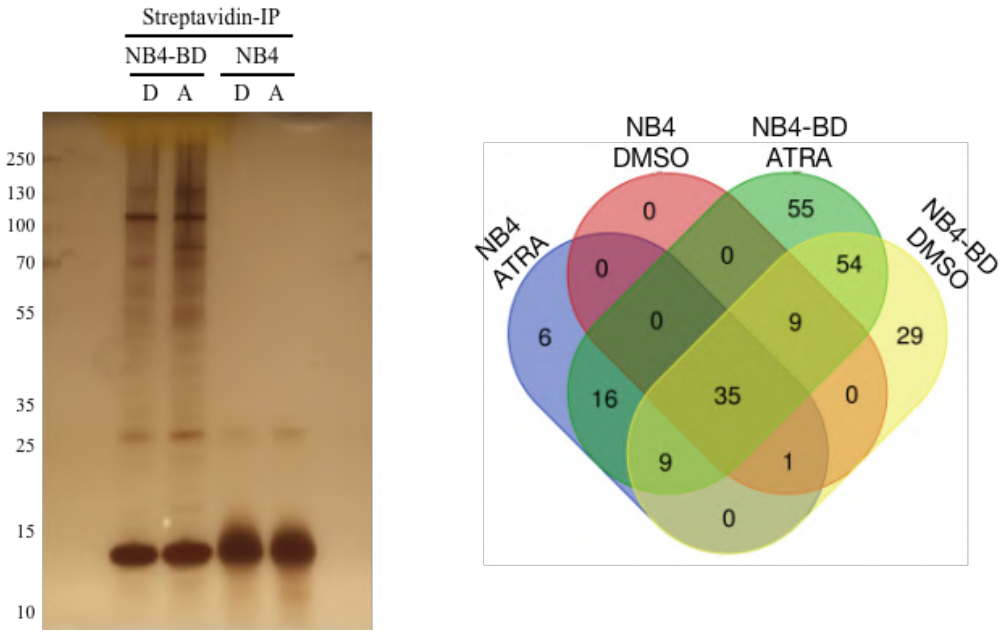


Figure 6-16 Left: Silver staining of proteins from the Streptavidin-IP experiment. Two different types of cells (parental NB4 and NB4-BD) and two different conditions (3-day treatment with ATRA or DMSO) were used, showing several bands specific to ATRA-treated NB4-BD cells. In all cases, 50 μM biotin was added to the cells 20 hours before collection. Right: Venn diagram (http://bioinformatics.psb.ugent.be/webtools/Venn/) for the proteins identified in the BioID2 pull-down samples using molecular spectrometry (Purple: ATRA-treated NB4 cells (Table 9-4 NB4 ATRA); Red: DMSO-treated NB4 cells (Table 9-3 NB4 DMSO); Green: ATRA-treated NB4-BD cells (Table 9-2 NB4-BD ATRA); Yellow: DMSO-treated NB4-BD cells (Table 9-1 NB4-BD DMSO).), showing 109 (55 and 54) and 83 (54 and 29) proteins in ATRA- and DMSO-treated NB4-BD cells, respectively, in which 54 proteins were shared.

For further analysis, the remaining beads were sent for Mass-spec analysis by on-beads digestion. Importantly, this process can remove and fragment Streptavidin-bound biotinylated proteins without releasing Streptavidin itself, which may interfere with mass spectrometry. The analysis produced a list of proteins for each sample. The list contained not only the names of the identified proteins but also the numbers of peptide sequences for each protein, i.e., the fragments from which the proteins were identified.

The lists were analyzed in parallel to produce the Venn diagram displayed in Figure 6-16 (Right Panel). After removing out the proteins in the parental NB4 samples, whose biotinylation is not a result of BioID2-Daxx, we are left with 109 and 83 proteins in ATRA- and DMSO-treated NB4-BD cells, respectively. 54 proteins were shared, which means that their interaction with BioID2-Daxx does not necessitate reconstituted PML-NBs. These included several known interactors of the Daxx protein, e.g., ATRX, HP1α, H3.3, and H4, however at a low specific unique peptide number, and not even in all three replicates which precluded the use of a specific cut-off, by peptide number, so for us the comparison between different baits proved more helpful than the use of a specific cutoff, which would have eliminated some known interactors, which proved positive when tested by western blot in the Streptavidin-pull downs.

After subtraction of parental NB4 background, one of the strongest signals in both ATRA and DMSO-samples was a protein called SMCHD1 (Structural Maintenance of Chromosomes flexible Hinge Domain Containing 1), which had not been previously reported in context with Daxx or ATRX in both cases SMCHD1 ranked even higher than the bait protein itself. The tandem mass spectrum of SMCHD1 is displayed in Figure 9-11.

To obtain additonal insight into the biological context of our identified proteins, we have used the freely available STRING-database to put the observed interactors in context with already known protein-protein interactions. For this we have applied the setting to show only experimentally verified interactions and interactions from curated databases. We found this presentation of the data more illustrative than only a list, which is still provided in 9.4 List of Mass Spec Results.

The 54 proteins were entered into the STRING-database to search for known interactions among them (Figure 6-17). It reveals two dominant groups, namely, (A) Histones and (B) ribosomal proteins.

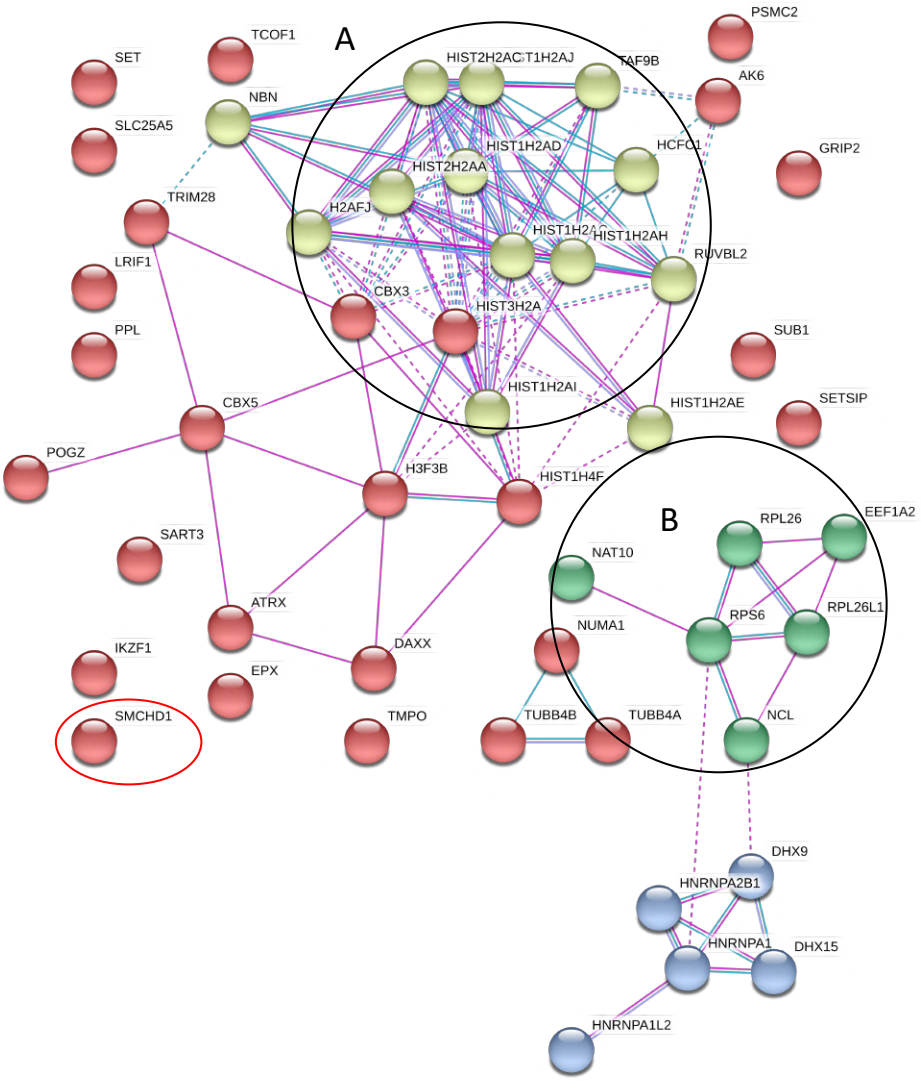


Figure 6-17 STRING analysis (http://www.string-db.org) derived protein-protein interaction networks for the 54 proteins that were identified as BioID2-Daxx interactors in both DMSO- and ATRA-treated NB4-BD cells (Figure 6-16 Right). The network nodes represent proteins. Splice isoforms or post-translational modifications are collapsed, i.e., each node represents all proteins produced by a single, protein coding gene. Lines indicate known interactions (Light blue line: from curated databases; Pink line: experimentally determined) with medium (0.400) confidence score as the analysis parameter. kmeans clustering (network is clustered to a specified number of clusters) number is 4: Two major populations can be distinguished in the figure, namely Histones (A, yellow balls) and nuclear ribonucleoproteins (B, green balls). SMCHD1 is marked in the red circle.

Our next step was to analyze the list of 55 proteins unique to ATRA-treated NB4-BD cells, which correspond to proteins that only interact with Daxx when NB-bodies have been restored, and should thus be more relevant to our study of the PAX complex. When looking at these networks, it is important to remember, that the bait protein itself interacts with all the factors shown, but itself will only show in the network, when it is indeed present in the corresponding list, i.e. Daxx is absent in the "ATRA-treated" condition, as self-biotinylation is of course independent of Daxx localization, so it will show only in the "ATRA- and DMSO-treated" group. As expected, this list included proteins known to localize in NB-bodies, such as PML, Sp100, and DEK, which we handleded as "positive controls" and these proteins only showed in a very low abundance range (Table 9-2 NB4-BD ATRA). Moreover, the list included SMARCA4 (also called BRG1), which, although only in the very low abundance range, was immediately interesting to us, as a neighboring group is actively working on this protein, in context of pediatric brain tumors, which made many reagents immediately available. So we included it in our first round of verification experiments by western blot (Figure 6-19). The spectrum for SMARCA4 is displayed in Figure 9-12. The 55 proteins were also entered into the STRING database to search for interactions among them. The resulting network is shown in Figure 6-18. Two major populations can be distinguished in the figure, namely (A) Cytoplasmic protein and (B) ribonucleoproteins.

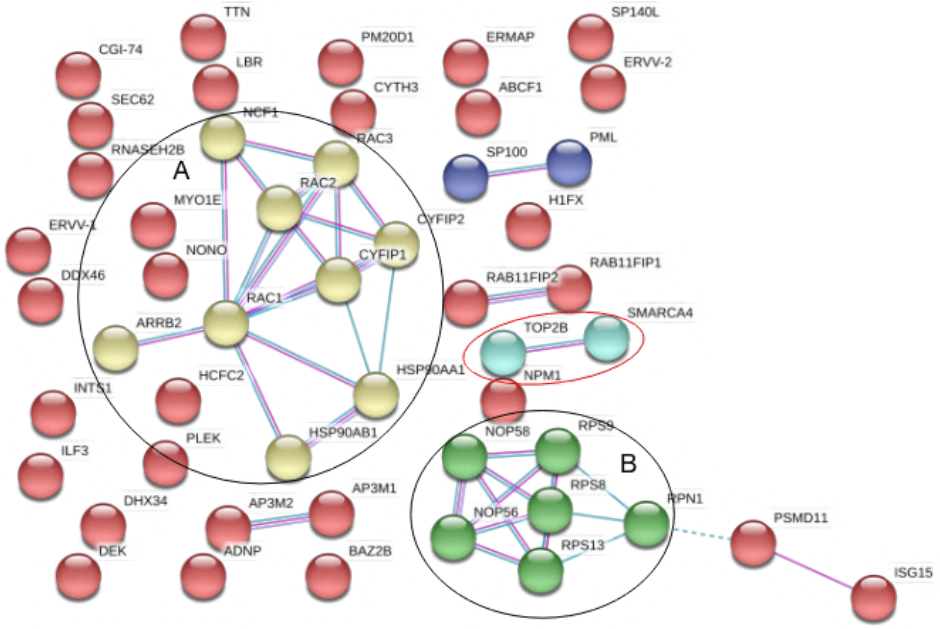


Figure 6-18 STRING analysis (http://www.string-db.org) derived protein-protein interaction networks for 55 proteins that were identified as BioID2-Daxx interactors only in ATRA-treated NB4-BD cells (Figure 6-16 Right). The network nodes represent proteins. Splice isoforms or post-translational modifications are collapsed, i.e., each node represents all proteins produced by a single, protein coding gene. Lines indicate known interactions (Light blue line: from curated databases; Pink line: experimentally determined) with medium (0.400) confidence score as the analysis parameter. kmeans clustering (network is clustered to a specified number of clusters) number is 5: Two major populations can be distinguished in the figure, namely Cytoplasmic proteins (A) and ribonucleoproteins (B). SMARCA4 is marked in the red circle. Histones were present, but do not appear here as they are also present in the DMSO-sample. The same does apply to SMCHD1.

To immunologically verify the presence of the interaction partners we identified through mass spectrometry, we performed western blotting of eluates from the Streptavidin bead-precipitated samples (DMSO- and ATRA-treated NB4 cells, DMSO and ATRA-treated NB4-BD cells) using specific antibodies for SMCHD1, Sp100, and SMARCA4 (Figure 6-19). Samples of full extracts were also included as controls.

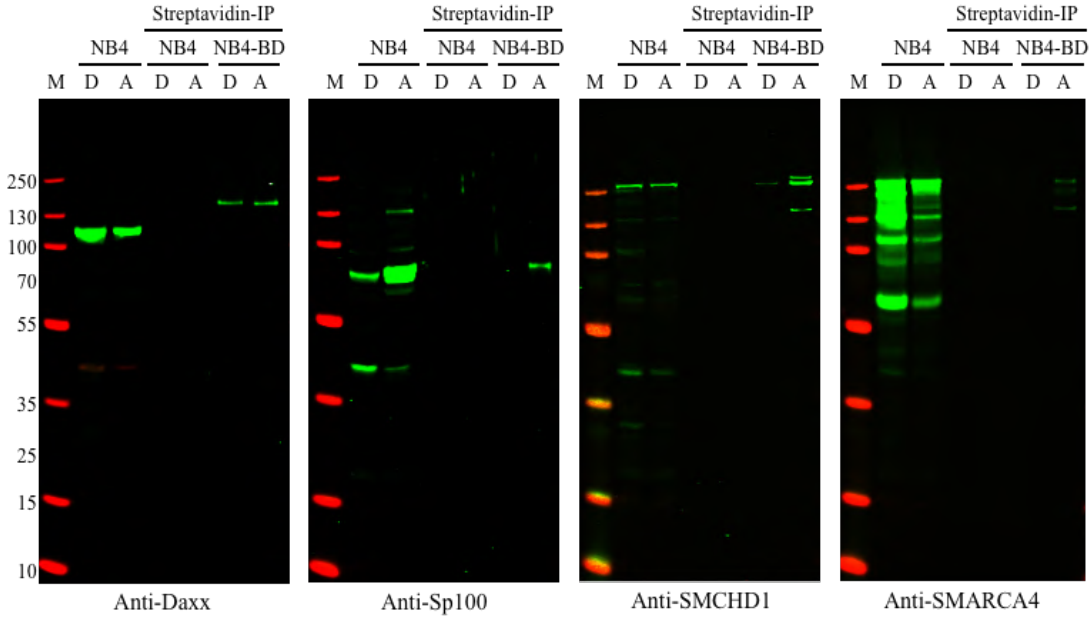


Figure 6-19 Western blotting for verification the novel interaction partners identified through mass spectrometry by the Streptavidin-IP of the NB and NB4-BD cells with ATRA- and DMSO-treatment using NB4 whole cell extract as antibody control with anti-Daxx, anti-Sp100, anti-SMCHD1 and anti-SMARCA4 antibodies, respectively, showing Sp100 as well as SMARCA4 were only observed in ATRA-treated NB4-BD cells, while SMCHD1 was observed in both DMSO- and ATRA-treated NB4-BD cells.

Sp100 in IPs was only observed in ATRA-treated NB4-BD cells, confirming that the reconstitution of PML-NBs is necessary for the protein to find itself close to BioID2-Daxx. The same was observed in the case of SMARCA4, validating it as a protein that only interacts with BioID2-Daxx in reconstituted PML-NBs. In contrast, SMCHD1 was observed in both DMSO- and ATRA-treated NB4-BD cells, confirming that it does not require reconstituted PML-NBs to find itself close to BioID2-Daxx.

Importantly, the Daxx band was similar in the lanes corresponding to the IPs of DMSO- and ATRA-treated cells, indicating that the differences in the abundance of the other three proteins were not a result of unequal loading. As expected, the full-extract samples had all three proteins.

From our novel interactors, we decided to first focus on ATRA-specific interactors for verification studies, as we expect them to have the highest chance for functional relevance. Among them SMARCA4 immediately became of interest, as through a collaboration with a neighboring group we had immediate access to key reagents, such as antibodies, the cDNA, as well as a Tamoxifen-inducible mouse model. So in our ranking SMARCA4 was placed No.1 for further confirmation and initial functional analysis.

6.3. Independent Confirmation of SMARCA4-Daxx Interaction

The BioID2 pull-down experiment described in the previous chapter identified SMARCA4 as a novel Daxx interactor within the reconstituted PML-NBs of ATRA-treated NB4-ND cells. We used a series of alternative methods to validate the results of the BioID2 pull-down assay.

6.3.1. Co-IP in HEK293T Cells Co-expressing FLAG-Daxx and SMARCA4

We transiently co-expressed FLAG-Daxx and SMARCA4 in HEK293T and performed a Co-IP using anti-FLAG. A major problem with co-IPs for proteins of the PML-NBs is their tight association with the nuclear matrix, which results in aggregation and poor solubility under mild IP conditions[58]. Since it is hard to remove insoluble proteins from the bead pellet, these proteins give false signals. To circumvent this, we created a negative control by co-expressing cells with vectors coding for FLAG-mSp100 and SMARCA4. As murine Sp100 (mSp100) has been previously shown to not directly interact with human SMARCA4, any SMARCA4 detected after performing FLAG co-IPs in these cells can be considered as "background" resulting from contamination.

After performing the anti-FLAG co-IPs in the experimental and the control cells, we eluted with two different protocols, namely “FLAG elution”, in which FLAG peptide is added in abundance in order to competitively release the FLAG-tagged proteins from the beads, and a standard elution protocol incorporating boiling of the beads. Input samples and eluates underwent western-blot analysis using antibodies for FLAG and SMARCA4 (Figure 6-20).

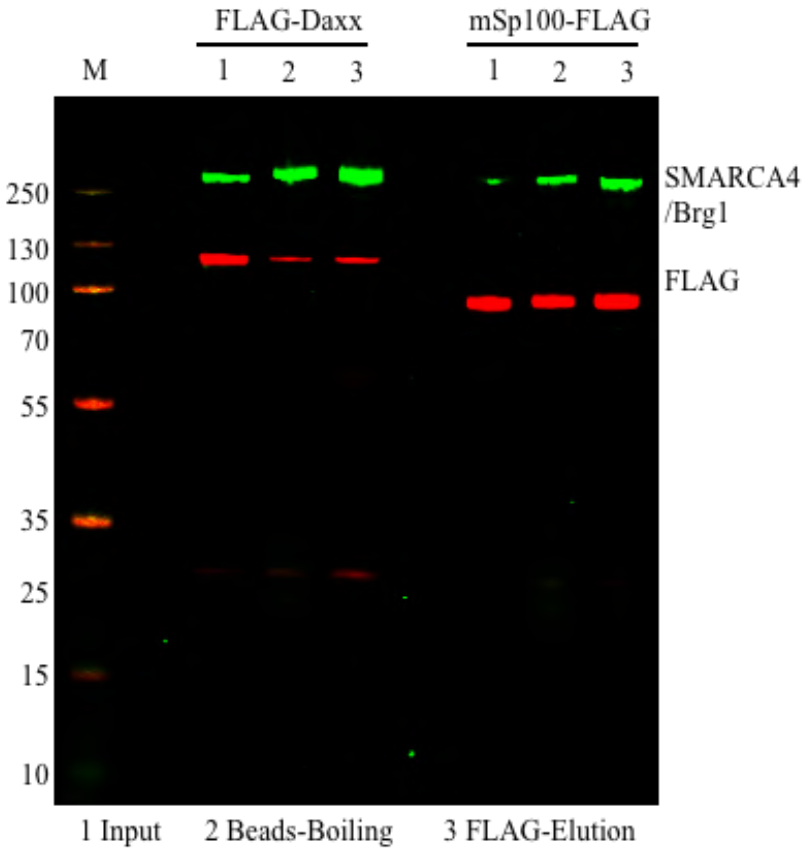


Figure 6-20 Western blots analysis of Input samples (1) and Elution samples (Two elution protocols were used, 2-“Bead-boiling elution” and 3-“FLAG elution”) after co-immunoprecipitation with anti-FLAG antibody in HEK293T cells co-expressing FLAG-Daxx and SMARCA4 or FLAG-mSp100 and SMARCA4. FLAG (red bands) was used as loading controls to normalize the SMARCA4 level (green bands).

The bands were quantified (Figure 6-21 Left) and the derived values were used to calculate "SMARCA4 signal: FLAG signal" ratios (Figure 6-21 Right). Irrespective of elution method, the SMARCA4: FLAG ratio was more than 2 times higher in the FLAG-Daxx/SMARCA4 cells compared to the FLAG-mSp100/SMARCA4, which is a strong indication that Daxx indeed interacts with SMARCA4.

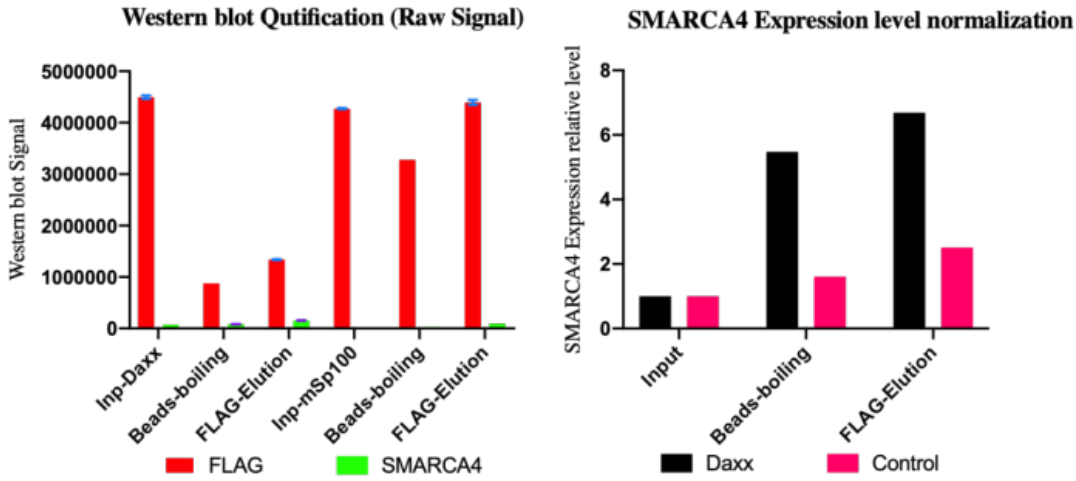


Figure 6-21 Left: Quantification of the FLAG and SMARCA4 bands of the western blots presented in Figure 6-20, all data are represented as mean ± SD; Right: SMARCA4 expression levels normalized by FLAG (control is mSp100) expression levels, both methods showing the SMARCA4: FLAG ratio was more than 2 times higher in the FLAG-Daxx/SMARCA4 cells compared to the FLAG-mSp100/SMARCA4.

6.3.2. Nano-Glo Dual-Luciferase assay

The assay was performed in HEK293T cells. The generation of the pM-Daxx and pVp16-SMARCA4 expression plasmids, which are required for testing the interaction between Daxx and SMARCA4 with this assay, were produced through PCR cloning (section 9.1.5). Furthermore, transfection with pM-Vp16 or pM+pVp16T served our positive controls, while transfection with pM, pVp16, pM-Daxx, pVp16-SMARCA4, or pM+pVp16 served negative controls; and used pM-Daxx+pVp16-SMARCA4 to confirm the interaction between Daxx and SMARCA4.

We calculated the firefly luminescence relative values for each case (Figure 6-22 Upper Panel) by normalizing the firefly luminescence with NanoLuc luminescence data. As there was no strong signal from the pM-Daxx+pVp16-SMARCA4, the assay failed to confirm the interaction between Daxx and SMARCA4. At this point, we are unable to explain the inconsistency between the results of the specific assay and the other ones.

6.3.3. FRET

The required pYFP-Daxx, pCFP-SMARCA4, and pYFP-CFP expression plasmids were constructed by PCR cloning (section 9.1.6). The pYFP-CFP linker construct served as our positive control, whereas cells transfected with the acceptor-only pYFP construct, the donor-only pCFP, pYFP-Daxx, pCFP-SMARCA4, or pYFP+pCFP were our negative controls. As seen in Figure 6-22 (Lower panel), the FRET signal of pYFP-Daxx+pCFP-SMARCA4 was higher than all of the control background signals, comparing with YFP+CFP, p-value as small as 0.015 ($0.015 < 0.05$) is strong evidence that FRET signal (statistically) significantly increased, suggesting the interaction between Daxx and SMARCA4.

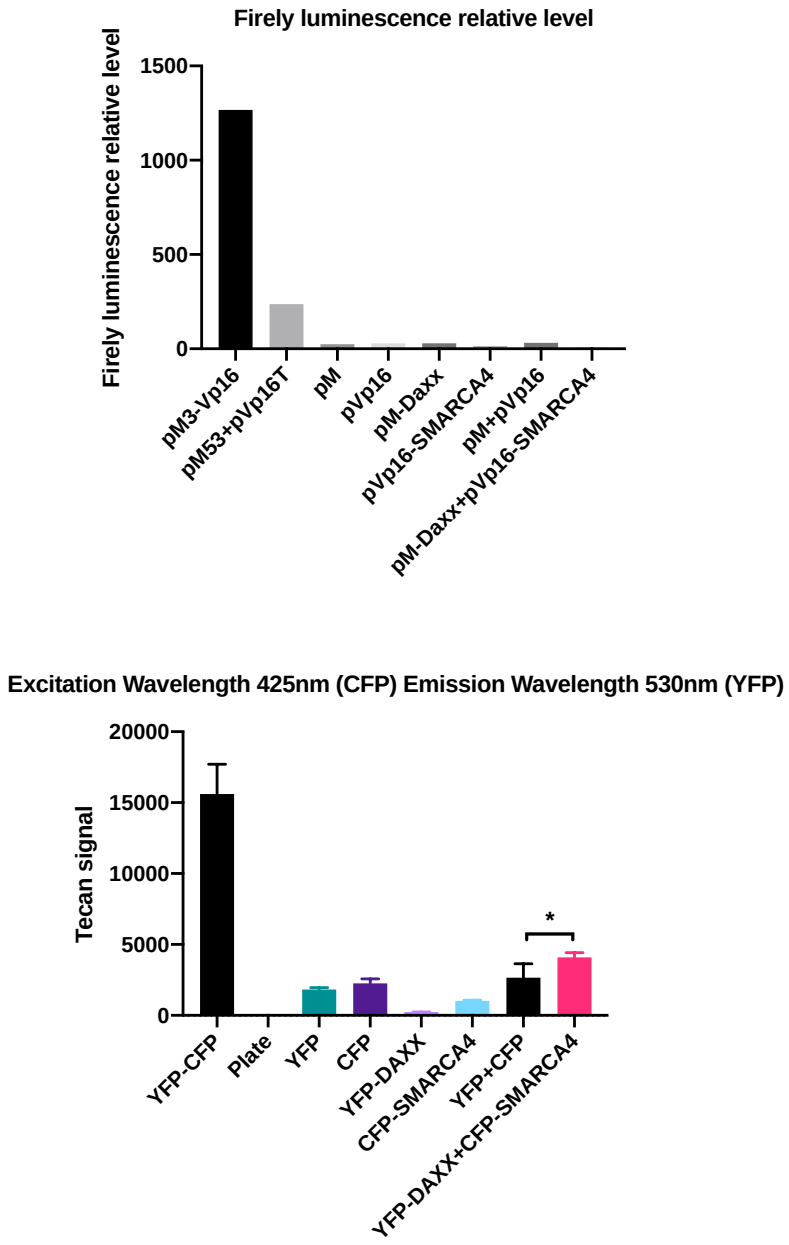


Figure 6-22 Upper panel: Bar graphs showing the results of Nano-Glo Dual-Luciferase assay, which calculated the firefly luminescence relative values for each case by normalizing the firefly luminescence with NanoLuc luminescence data, seeing no strong signal from the pM-Daxx+pVp16-SMARCA4; Lower Panel: FRET assay showing the FRET signal of pYFP-Daxx+pCFP-SMARCA4 was higher than all the control background signals (pYFP, pCFP, pYFP-Daxx, pCFP-SMARCA4, or pYFP+pCFP); comparing with YFP+CFP, p-value as small as 0.015 ($0.015 < 0.05$) is strong evidence that FRET signal (statistically) significantly increased. All data are represented as mean ± SD.

6.4. Studies on the Biological Relevance of the SMARCA4-Daxx Interaction

6.4.1. Effects of SMARCA4 Knockout on the Immortalizing Effects of mPR and H3.3K27M Expression

As was mentioned in section 3.4 “The Daxx/ATRX Chaperone”, the ATRX/Daxx dimer within the PAX complex acts as a chaperone for the integration of the H3 histone variant, H3.3. This function, combined with the interest of our lab in the effects of the oncohistone H3.3 K27M, led us to examine the effects of SMARCA4 knocking out in the presence or absence of the oncohistone. The experimental model we chose was bone marrow cells extracted from mice in which SMARCA4 alleles flanked by LoxP sites (Figure 5-1) (see section 5.1.18). These animals were chosen as they allow us to produce SMARCA4 knockout cells with exactly the same genotype (with the exemption, of course, of the floxed sequence) with their wild-type counterparts, allowing for highly reliable comparisons.

The extraction and passaging of bone marrow cells were performed as described in section 5.1.16 “Mouse Bone Marrow Cell Extraction, Transduction, and Passaging”. Importantly, 500 nM 4-hydroxy-Tamoxiphen (4-OH TXF) was added to specific wells for 6 days (P1-P3) to knock out SMARCA4 (Figure 6-23).

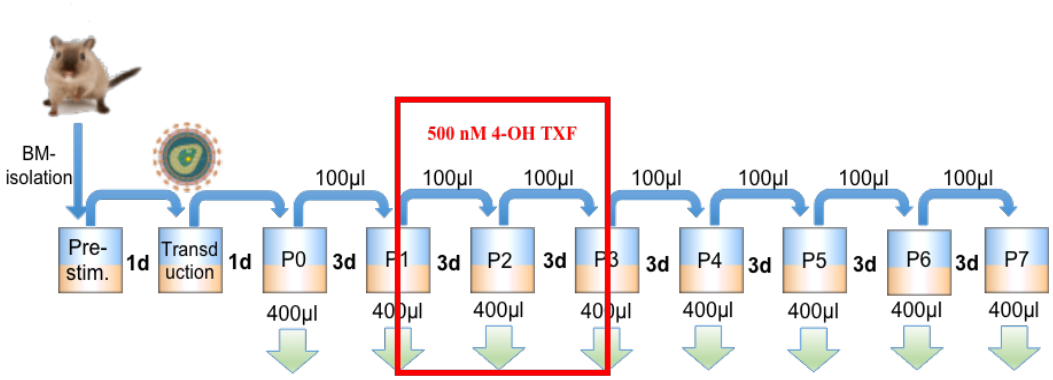


Figure 6-23 Mouse bone marrow passage assay setting of SMARCA4 knockout experiment. 500 nM 4-hydroxy-Tamoxiphen (4-OH TXF) was added to specific wells for 6 days (Passage 1-Passage 3, Red box) to knock out SMARCA4.

The results of the SMARCA4 knockout were observed in mock-transduced cells, cells expressing H3.3wt, and cells expressing oncogenes, namely the mPR oncoprotein or the K27M oncohistone. These two were chosen because of their relationship with the PAX complex. Specifically, mPR is an artificial mouse homolog of the human PR oncoprotein; it is significantly more leukemogenic in mice compared to human PR and has been shown to disrupt PML-NBs and block senescence in murine bone marrow cells, better than human PR[150]. On the other hand, H3.3 K27M is an oncogenic form of H3.3, which is known to be integrated into chromatin by ATRX/Daxx (see section 3.4 “The Daxx/ATRX Chaperone”). Previous experiments of our lab have shown that both oncoproteins can induce immortalization in murine bone marrow cells. Mock cells and cells transduced with H3.3wt, which become senescent after a small number of passages, served as a negative control. To sum up, our main goal was to see whether knocking out SMARCA4 affected the ability of the oncoproteins to immortalize the cells. Parallel to cells from floxed mice, cells from wild-type mice were subjected to the same treatment, including transduction with the same constructs, to serve as controls.

Before moving on to the results, we should clarify the terms used to describe the experimental groups: Cre+ TXF+, floxed mouse cells treated with 4-hydroxy-Tamoxifen; Cre+ TXF-, floxed mouse cells not treated with 4-hydroxy-Tamoxifen; Cre- TXF+, non-floxed mouse cells treated with 4-hydroxy-Tamoxifen; Cre- TXF-, non-floxed mouse cells not treated with 4-hydroxy-Tamoxifen. Theoretically, only Cre+ TXF+ cells should be SMARCA4 knockouts. We should also note that Cre- TXF+ cells did not differ in proliferation compared to Cre- TXF- cells, suggesting that 4-hydroxy-Tamoxifen is not toxic to mouse bone marrow cells at the used concentration.

The growth curves for all cells and conditions are shown in Figure 6-24. Mock-transduced cells and cells transduced with H3.3wt stopped dividing after some passages, irrespective of mouse origin (Cre- or Cre+) and Tamoxifen treatment (TXF- or TXF+). In contrast, that the overexpression of mPR in non- SMARCA4 knockout cells, i.e., Cre- TXF-, Cre- TXF+, and Cre+ TXF-, led to immortalization, with the successfully transducing cells becoming able to divide indefinitely. However, immortalization was not observed in the case of SMARCA4 knockouts (Cre+ TXF+).

The observations were largely similar in the case of cells transduced with the H3.3 K27M oncohistone. However, there was a remarkable exception. Cre+ TXF- cells transduced with H3.3 K27M failed to immortalize. The only plausible explanation we were able to come up with has to do with the leakiness of the system.

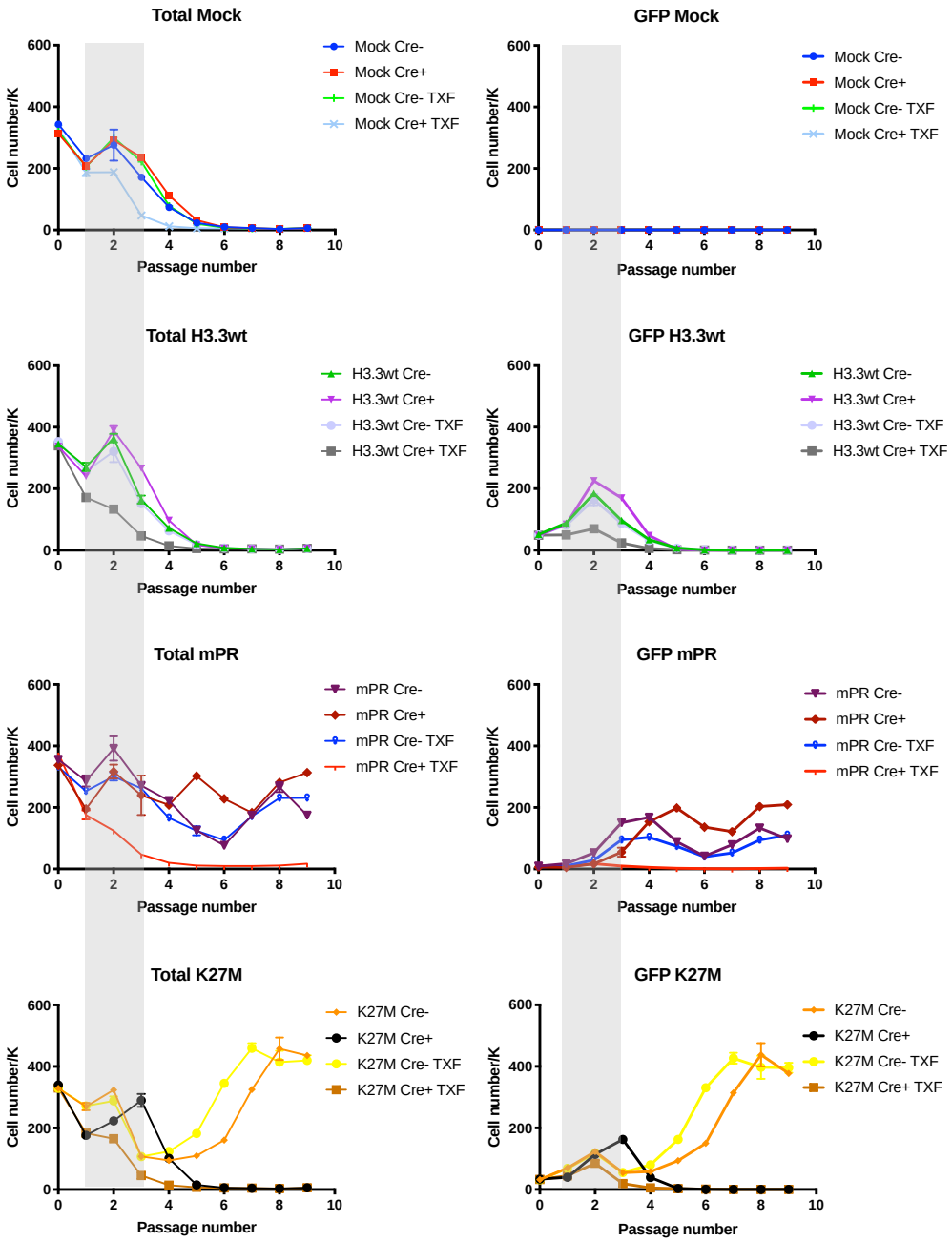


Figure 6-24 Mouse bone marrow proliferation graphs for SMARCA4 knockout experiment. Grey boxes denote the time frame of 4-OH Tamoxifen treatment. The left panel graphs show the number of all cells (i.e., both transduced and non-transduced), whereas the right panel graphs display the number of GFP-expressing, i.e., successfully transduced, cells. All data are represented as mean ± SD. Showing that the overexpression of K27M/mPR in non-SMARCA4 knockout cells led to immortalization, however, immortalization was not observed in the case of SMARCA4 knockouts (Cre+ TXF+).

To test this hypothesis, we used PCR with primers amplifying a region of the floxed *SMARCA4* gene to determine the abundance of the *SMARCA4* gene in each culture. Any gene whose presence in the genome is not affected by floxing could be used as a control. We chose *TERT*, as primers for it were already available. As seen in Figure 6-25, the results were the expected ones up to passage, i.e., rapid fall in the cultures that were treated with TXF to induce the excision of *SMARCA4*, and stability of the gene presence in non-treated cultures. However, after passage 6, a massive fall in *SMARCA4* presence was observed in the Cre+ TXF- cultures overexpressing H3.3 K27M. The change is drastic and results in these cultures practically becoming *SMARCA4* knockouts, which explains their inability to immortalize. Why this "leakage" of the flox system was observed in H3.3 K27M- but not mPR-overexpressing Cre+ TXF- is unknown.

Another strange observation is the recovery of *SMARCA4* presence in mPR-overexpressing Cre+ TXF+ cells. This could be explained by some cells evading the TMX-induced excision. These cells would ultimately outgrow their neighbors (i.e., cells in which excision had been successful), leading to the observed increase in *SMARCA4* presence. The changes in *SMARCA4* presence observed by PCR were confirmed by immunohistochemical staining of cytospin slides for the corresponding protein (Figure 6-26).

We conclude that SMARCA4 is required for the continued proliferation of cells immortalized by very different mechanisms, which suggests that SMARCA4 might be a valid therapeutic target for different malignancies.

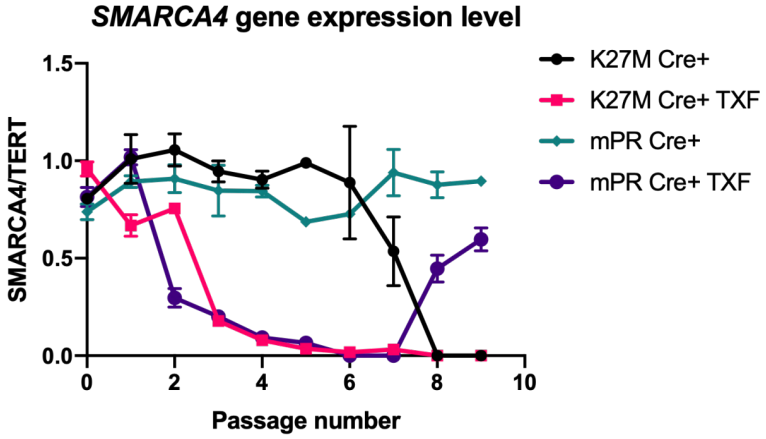


Figure 6-25 qPCR quantification of a genomic SMARCA4 sequence reveals the extent of excision under various conditions (*TERT* gene be used as a control). Showing mPR-overexpressing Cre+ TXF+ cells evading the TMX-induced excision, which ultimately outgrow their neighbors, leading to the observed increase in SMARCA4 presence. All data are represented as mean ± SD.

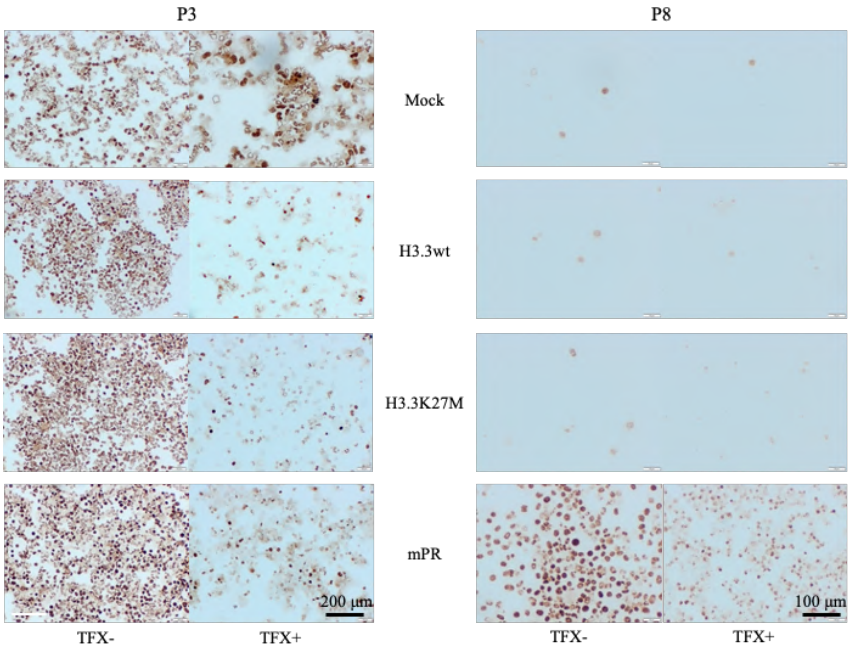


Figure 6-26 Immunohistochemistry staining of SMARCA4 (brown dots) of cytospin slides of Mock-, H3.3wt-, H3.3K27M- and mPR-transduce cells on Passage 3 and 8 (P3 and P8), confirming changes in SMARCA4 presence. Scale bar is 200 μm (Left) and 100 μm (Right), respectively.

6.4.2. Effects of SMARCA4 overexpression and knockdown in NB4 cell sublines

As shown in the previous chapter, knocking out SMARCA4 in bone marrow cells inhibited their proliferation, suggesting that this protein may be a therapeutic target in cancer. To both examine this possibility and further study the Daxx/SMARCA4 interaction, we decided to create and study SMARCA4-overexpression and SMARCA4-knockdown sublines of NB4 cells.

6.4.2.1. Generation of NB4-OE.S4 and NB4-KD.S4

NB4 cell sublines displaying SMARCA4 overexpression (NB-OE.S4) or knockdown (NB4-KD.S4) were produced by transducing NB4 cells with the pLeGO-SMARCA4 or the pLeOG-siRNA-SMARCA4 construct, respectively. Of the two constructs, the former was created by us using PCR cloning (see section 9.1.4), whereas the latter was a kind gift from Dr. med. Franziska Modemann from U. Schüller group (Forschungsinstitut Kinderkrebs-Zentrum, Hamburg). FACS sorting was performed to obtain >95% pure populations.

6.4.2.2. Validation of NB4-OE.S4 and NB4-KD.S4

After creating the two stable sublines, we proceeded to determine their levels of SMARCA4 expression through western blotting and IF staining. We were interested in those levels in both the absence and the presence of PML-NBs, thus we treated cells of the parental line and the two sublines with DMSO or ATRA for three days. The following observations can be made by studying the western blots and the derived graphs (Figure 6-27 and Figure 6-28), respectively:

a) In all cell types, ATRA treatment results in a drastic reduction of SMARCA4 levels.

b) With DMSO treatment, NB4-OE.S4 expressed 25% more SMARCA4 than its NB4 counterpart; however, under ATRA treatment, NB4-OE.S4 cells was having 3 times more SMARCA4 than ATRA-treated NB4 cells.

c) DMSO-treated NB4-KD.S4 cells have about 70% lower SMARCA4 levels than their NB4 counterparts. However, there is no significant difference in SMARCA4 abundance between these two cell types under ATRA treatment conditions.

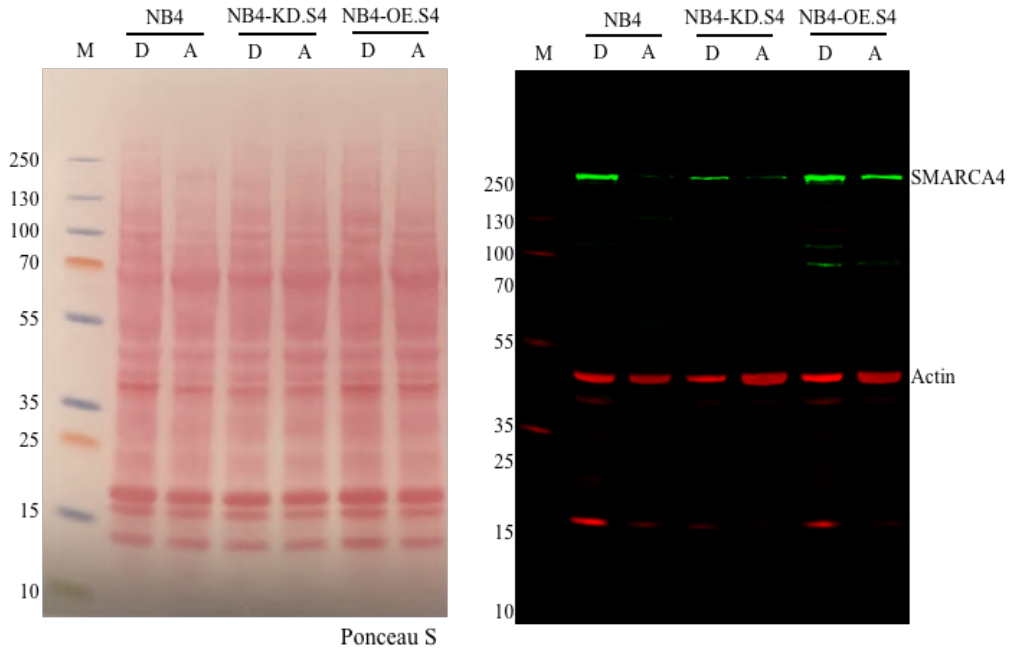


Figure 6-27 Western blot analysis of SMARCA4 levels in NB4, NB4-OE.S4, and NB4-KD.S4 cells treated with DMSO or ATRA for 3 days. Left: Ponceau S staining of the membrane prior to blotting by the whole cell extract of the NB4-AB and NB4-BA cells with ATRA- and DMSO-treatment, served as loading control; Right: Western blot to confirm the expression level of SMARCA4 (green bands at 250 kDa), and actin (red bands at 42 kDa) were used as loading controls to normalize the SMARCA4.

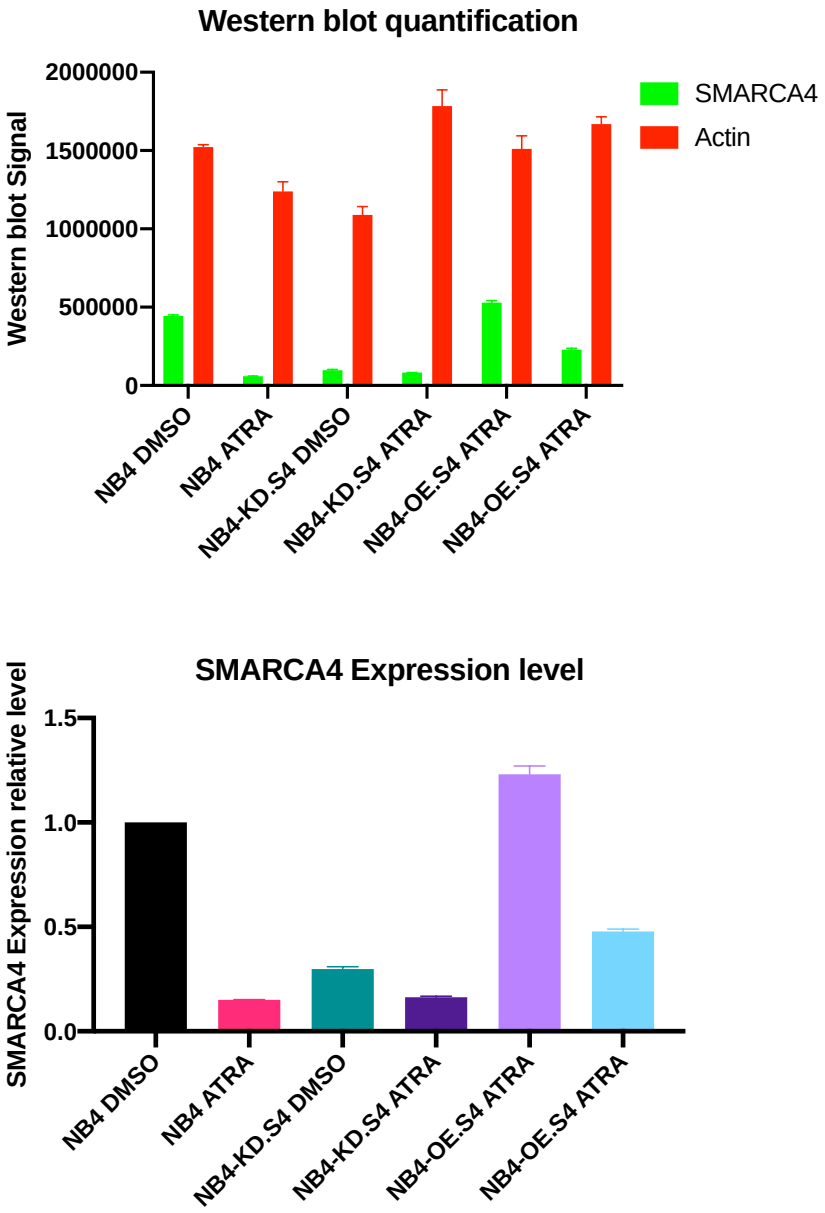


Figure 6-28 Upper: SMARCA4 and Actin western blot signal in Figure 6-27 Right, quantified by Image Studio Lite. Lower: SMARCA4 relative expression level of the NB4, NB4-OE.S4, and NB4-KD.S4 in both ATRA-treated and DMSO control condition, which normalized by the DMSO control NB4 cells. All data are represented as mean ± SD.

IF staining confirmed the results of the western blot analysis. Moreover, it revealed that the ATRA-induced reconstitution of PML-NBs takes place normally in all cell types, i.e., the parental line and the two sublines (Figure 6-29).

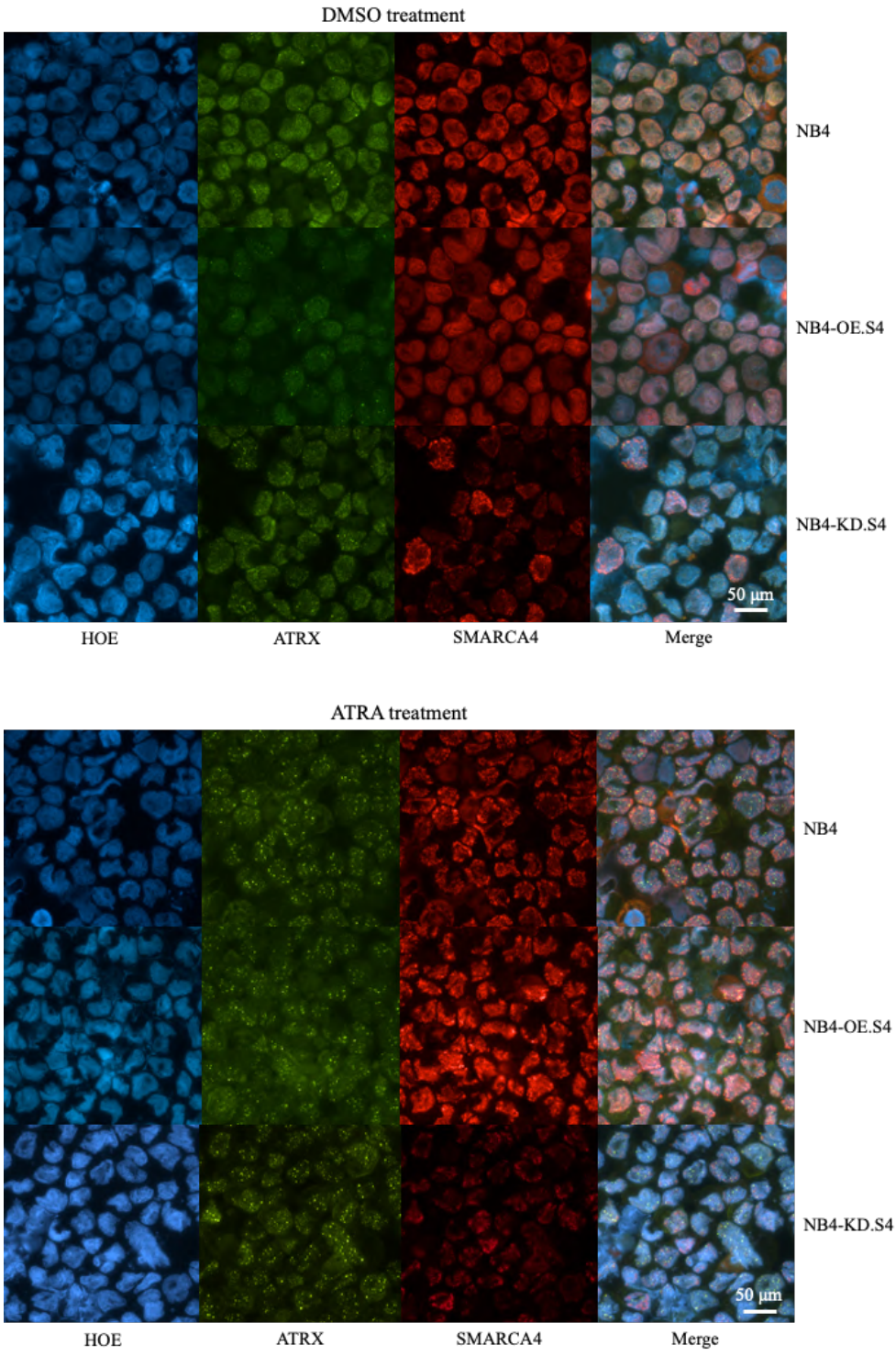


Figure 6-29 Immunofluorescence (IF) staining with Hoechst (Blue) staining for nucleus, ATRX (Green), SMARCA4 (Red) antibody confirmed the results of the western blot analysis (Figure 6-27), revealing that the reconstitution of PML-NBs takes place normally in ATRA-treated (Lower) NB4, NB4-OE.S4, and NB4-KD.S4 cells, but not in DMSO-treated (Upper) NB4, NB4-OE.S4, and NB4-KD.S4 cells. Scale bar is 50 µm.

6.4.2.3. Morphological analysis of NB4, NB4-OE.S4, and NB4-KD.S4 Cell

We used the Diff-Quik protocol to examine the morphology of NB4, NB4-OE.S4, and NB4-KD.S4 under both control (DMSO) and ATRA treatment conditions.

There were no visible differences in morphology among the three cell types (Figure 6-30 Left). Moreover, ATRA treatment in all cell types resulted in differentiation to blast cells, with the shape of the nucleus changing from round to irregular (Figure 6-30 Right). This indicates that these changes in SMARCA4 level do not visibly affect the ability for the lines to differentiate.

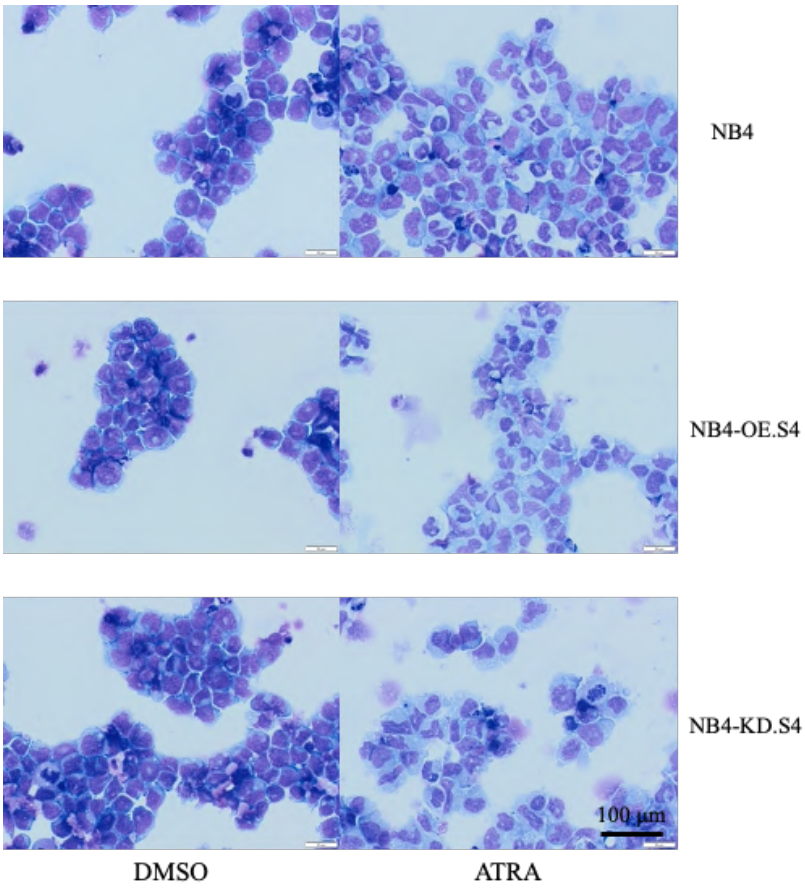


Figure 6-30 Morphology of DMSO- and ATRA-treated NB4, NB4-OE.S4, and NB4-KD.S4 cells, nucleus (purple) and cytoplasm (light blue) showing no visible differences in morphology among the three cell types; with ATRA treatment in all cell types resulted in differentiation, with the irregular shape of the nucleus. Scale bar is 100 µm.

6.4.2.4. Growth profile of NB4, NB4-OE.S4, and NB4-KD.S4 cells

Cells were grown in the presence of DMSO and ATRA. Every 2 days, cultures were passaged 1:2 and their number were calculated by FACS analysis.

Under control conditions, NB4-KD.S4 cells proliferated a bit faster than NB4 or NB4-OE.S4 cells (Figure 6-31). ATRA treatment rapidly stopped the proliferation of all three cell types, showing no significate difference between NB4, NB4-OE.S4 and NB4-KD.S4 cells lines, which indicates that ATRA-induced growth suppression is unaffected by modulation of SMARCA4 levels.

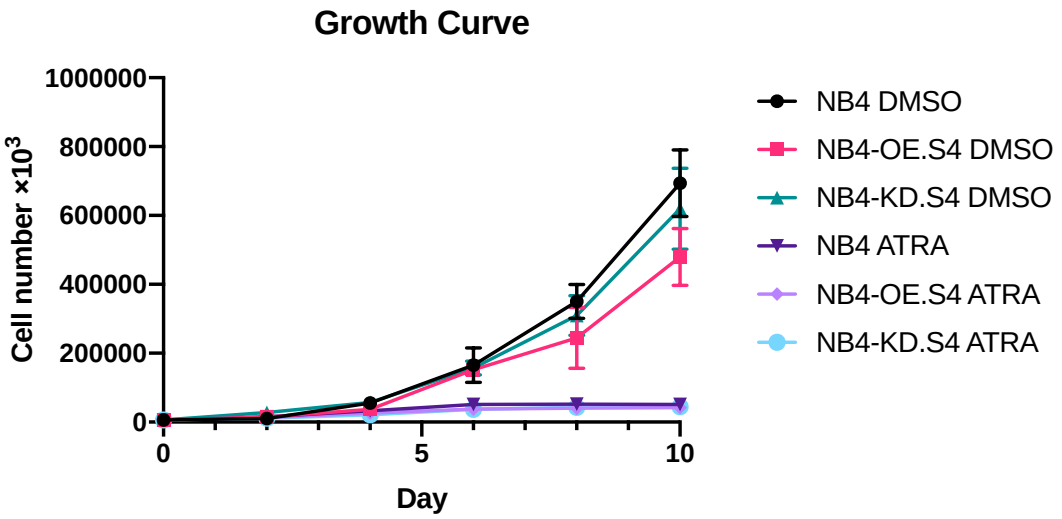


Figure 6-31 Growth curve of DMSO- and ATRA-treated NB4, NB4-OE.S4, and NB4-KD.S4 cells, indicating ATRA-induced growth suppression is unaffected by modulation of SMARCA4 levels. All data are represented as mean ± SD.

6.4.2.5. Cell cycle profile of NB4, NB4-OE.S4, and NB4-KD.S4 cells

To compare the cell cycle profiles of NB4, NB4-OE.S4, and NB4-KD.S4 cells, we performed cell cycle analysis using FACS. All three cell types had similar profiles, both under control conditions (DMSO treatment) and under ATRA treatment conditions (Figure 6-32). Both sublines entered G1-S cycle arrest after ATRA treatment similar to NB4 cells, which indicates that cycle profiles are unaffected by modulation of SMARCA4 levels showing similar G1-S cycle arrest after ATRA treatment in NB4, NB4-OE.S4, and NB4-KD.S4 cells.

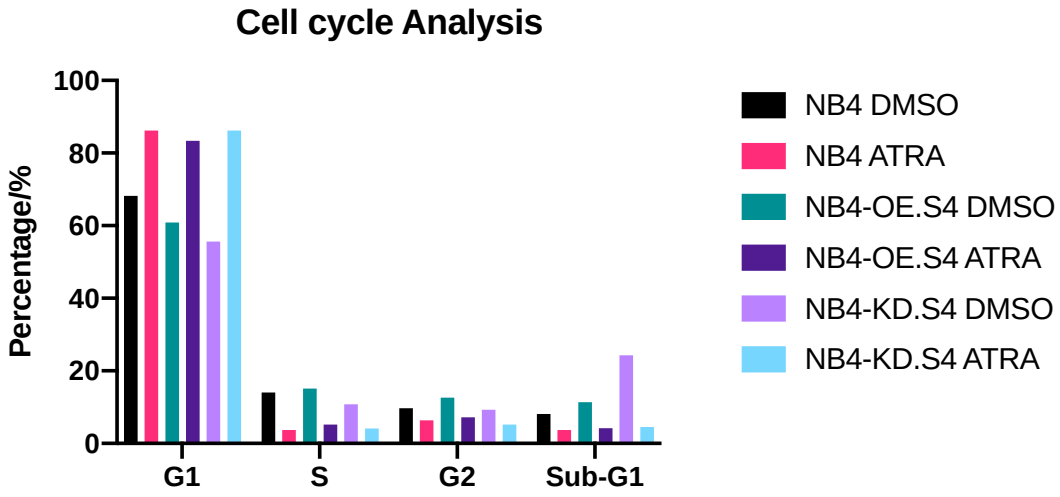


Figure 6-32 Cell cycle profiles of NB4, NB4-OE.S4, and NB4-KD.S4 cells under both control and ATRA treatment conditions, the bar graph shows cells at various stages of cell cycle, showing similar G1-S cycle arrest after ATRA treatment in NB4, NB4-OE.S4, and NB4-KD.S4 cells.

To sum up, these control experiments have proven that in all aspects tested, NB4-OE.S4 and NB4-KD.S4 cells exhibit a behavior similar to parental NB4 cells, meaning that the changes in SMARCA4 level did not substantially alter the cells.

6.4.2.6. Effects of SMARCA4 Inhibition on the Proliferation of NB4, NB4-OE.S4, and NB4-KD.S4 cells

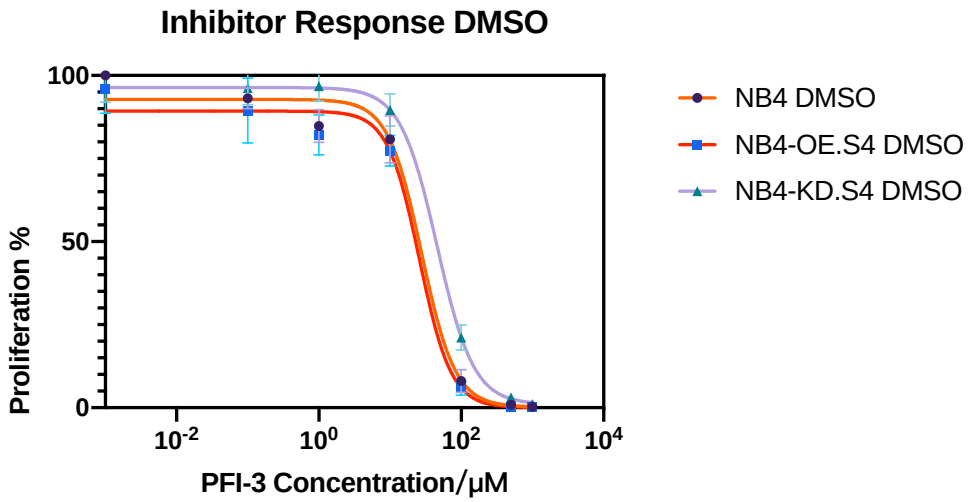
As shown in the previous two chapters, there were no major differences in proliferation between the parental NB4 cells and their knock-down subline. This could either mean that NB4 cells express SMARCA4 in a quantity so much higher than the minimum required for proliferation ("overabundance" hypothesis) that, despite the knock-down, NB4-KD.S4 cells have a sufficient quantity, or that these cells no longer require SMARCA4 for proliferation ("independence" hypothesis). In order to find an answer, we examined whether the three lines are sensitive to SMARCA4 inhibition, as well as whether there are changes in sensitivity among them. We used PFI-3, an inhibitor selective for family VIII bromodomains, which is the type found in SMARCA2 and SMARCA4[151].

The tested concentrations of PFI-3 ranged from 1 nM to 500 μM and treatment lasted for 72 hours. Testing for each line was performed both under control and ATRA treatment conditions.

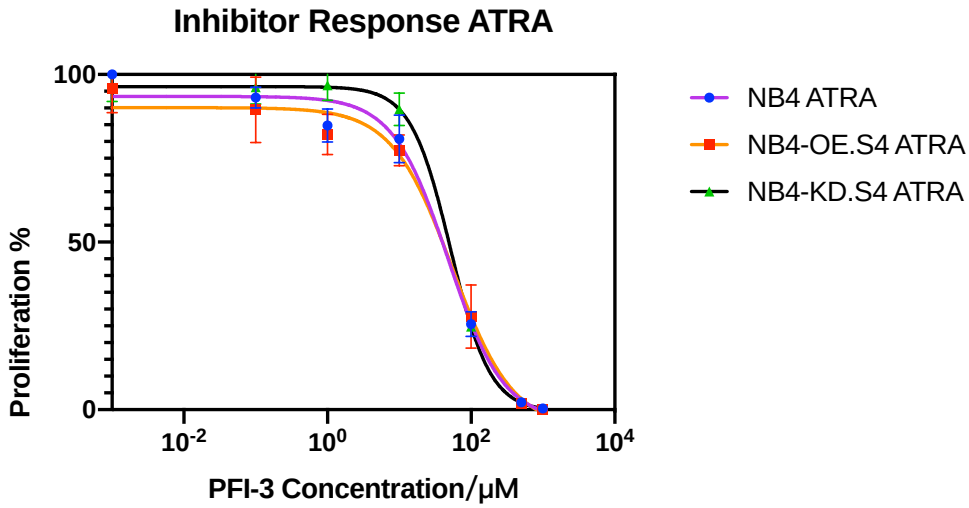
In the absence of ATRA, all sublines were sensitive to PFI-3 (Figure 6-33). However, NB4-KD.S4 cells, despite having less SMARCA4 than the other two lines, were less sensitive to the inhibitor. Moreover, there were no differences in IC50 between NB4 and NB4-OE.S4 cells, even though the latter have about 25% more SMARCA4 than the former. These data seem to disprove the "overabundance" hypothesis.

Also under ATRA treatment conditions, no significant differences among the three cell types were observed. Since ATRA-treated NB4-OE.S4 cells contain more than double SMARCA4 than ATRA-treated NB4 and NB4-KD.S4 cells, the lack of differences among them provides further disproval to the "overabundance" hypothesis.

Thus, we conclude that SMARCA4 abundance, at least at the studied level of magnitude, does not affect the growth of NB4 cells. Also, PFI-3 inhibition is not potent enough to come to a clear conclusion. Only a knockout study could indisputably demonstrate whether the growth of NB4 is truly SMARCA4-independent. Another conclusion is that, even though NB4 cells and the two sublines are sensitive to PFI-3, this most probably has more to do with the inhibition of its second specific target protein, namely SMARCA2[152].



	NB4 DMSO	NB4-OE.S4 DMSO	NB4-KD.S4 DMSO
IC50	27.75	25.88	45.79



	NB4 ATRA	NB4-OE.S4 ATRA	NB4-KD.S4 ATRA
IC50	47.73	54.85	51.78

Figure 6-33 PFI-3 dose-response curves of NB4, NB4-OE.S4, and NB4-KD.S4 under both DMSO (Upper) and ATRA-treated (Lower) conditions. The tested concentrations of PFI-3 ranged from 1 nM to 500 μM and treatment lasted for 72 hours. IC50 is showed in the table below the curve. In the absence of ATRA, NB4-KD.S4 cells, despite having less SMARCA4 than the other two lines, were less sensitive to the inhibitor. All data are represented as mean ± SD.

6.5. Study of the PAX Complex by Parallel Large-Scale BioID2-Biotin/Streptavidin Pull-Downs for Various PML-NB Components

Our next step was to perform parallel pull-downs using Bio-ID2 fusions for additional proteins that are known to localize in PML-NBs, and the study of their interactions may thus provide independent confirmation of identified factors and also more information on the PAX complex. In plain words, we wanted to repeat and expand the process (described in section 6.2) with additional bait-proteins.

6.5.1. Sp100-BioID2 generation and characterization

One of the first proteins known to locate in the structures, now known as PML-NB, besides PML itself was the Sp100 autoantigen[153]. Also we have foundand confirnmed Sp100 as ATRA-dependent interactor with our NB4-BD cells (Figure 6-19 anti-Sp100). Hence we have chosen Sp100 as a second bait.

6.5.1.1. Generation of a NB4 subline stably expressing the Sp100-BioID2 fusion protein

NB4 cells were infected with pLeGO-Sp100-BioID2 lentivirus. The generation of the pLeGO-Sp100-BioID2 expression plasmid required the use of PCR cloning (section 9.1.2). Moreover, as seen in the corresponding plasmid map（Figure 9-2）, the fusion protein (Figure 9-8) has a myc-tag between the Sp100 and BioID2 parts.

After Puromycin selection and FACS sorting, a pure (~99%) GFP-positive population was acquired. This NB4 sub clone was christened NB4-SB1.

6.5.1.2. Validation of NB4-SB1

Western blotting (Figure 6-34) and IF staining (Figure 6-35) were performed to confirm that NB4-SB1 cells indeed express the Sp100-BioID2 fusion protein, both under control (DMSO) and ATRA treatment conditions.

NB4-SB1 cells were treated with ATRA or DMSO for 3 days, with Biotin added to the cells 20 h before collection. Whole cell extracts were subjected to SDS-PAGE. After transfer, we used Ponceau S staining (Figure 6-34 Left) as a loading control, and proceeded to blot the membrane with different anti-BioID2, anti-Sp100, and Streptavidin. Our objective was to confirm both the expression of the Sp100-BioID2 fusion protein and its migration in SDS-PAGE.

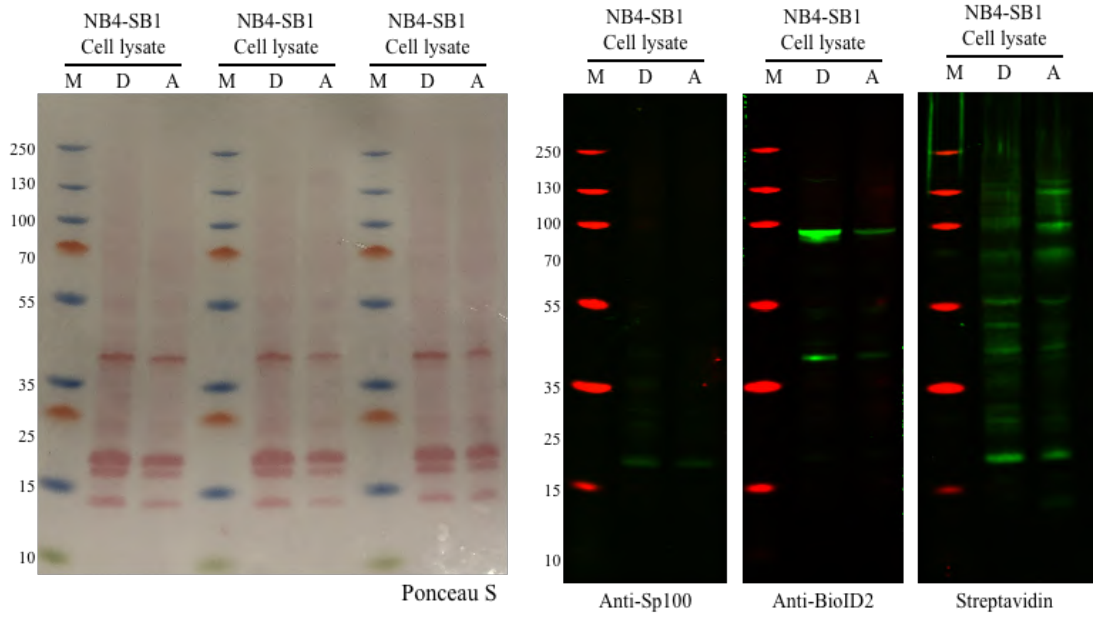


Figure 6-34 Left: Ponceau S staining of the membrane prior to blotting by the whole cell extract of the NB4-SB1 cells with ATRA- and DMSO-treatment, served as loading control; Right: Western blot to confirm the expression of Sp100-BioID2 fusion protein by NB4-SB1 whole cell extract using anti-Sp100, anti-BioID2, and Streptavidin antibodies, respectively.

Sp100 has a molecular weight of 54 kDa, but in known to exhibit an aberrant electrophoretic behavior, forming a 100 kDa band, hence its name. Thus, we expected the fusion protein to also run aberrantly, i.e., forming a band higher than its calculated molecular weight of 81 kDa (54 kDa from Sp100 and 27 kDa from BioID2).

The anti-BioID2 blot gave a strong band just below 100 kDa, however, no corresponding band was observed in the anti-Sp100 plot. This could be attributed to the Sp100 epitopes of the fusion protein being covered by the BioID2 part or disrupted due to misfolding. However, since the antibody also failed to recognize endogenous Sp100 (as mentioned above, Sp100 forms a band at 100 kDa), we concluded that this was probably a problem with the antibody. However, it should be mentioned that the same antibody gave clear Sp100 speckles in the IF step (Figure 6-35). Finally, the Streptavidin blot gave multiple bands, suggesting that the BioID2 part of the fusion protein is functional.

Our next step was to locate the Sp100-BioID2 fusion protein in cell slides using IF for anti-BioID2. Staining with Streptavidin was also performed to reveal the accumulation of biotinylated proteins. As seen in Figure 6-35, most all of the cells (99%) failed to give BioID2 signal, suggesting that most of the transduced cells did not express the fusion protein. It should be stressed that all cells of are FACS-sorted, GFP-expressing NB4-SB1, so the lack of signal cannot be attributed to unsuccessful transduction. What is most surprising is the fact that, despite fusion protein-producing cells only representing 1% of the total population, they were able to give such a strong BioID2 band in the aforementioned western blots.

Since Sp100-BioID2 is produced in only a few cells of the NB4-SB1 subline, there is no point in examining the morphology, cell cycle profile, and proliferation speed of the NB-SB1 subline. However, the subline could still be of use. The IF staining profile of the few cells that actually expressed the protein suggests that the Sp100-BioID2 fusion protein is functional, as clear dots of colocalized BioID2 and Sp100 in

these cells only showed up after ATRA treatment, i.e., the PML-NBs had been reconstituted. This means that NB-SB1 can be used in Streptavidin pull-down experiments. The small number of cells expressing the proteins compromises the sensitivity but not the specificity of this assay. In other words, even though the low percentage of fusion protein-producing cells will make it difficult to recognize interactors, it does not compromise the validity of the interactors that will be recognized.

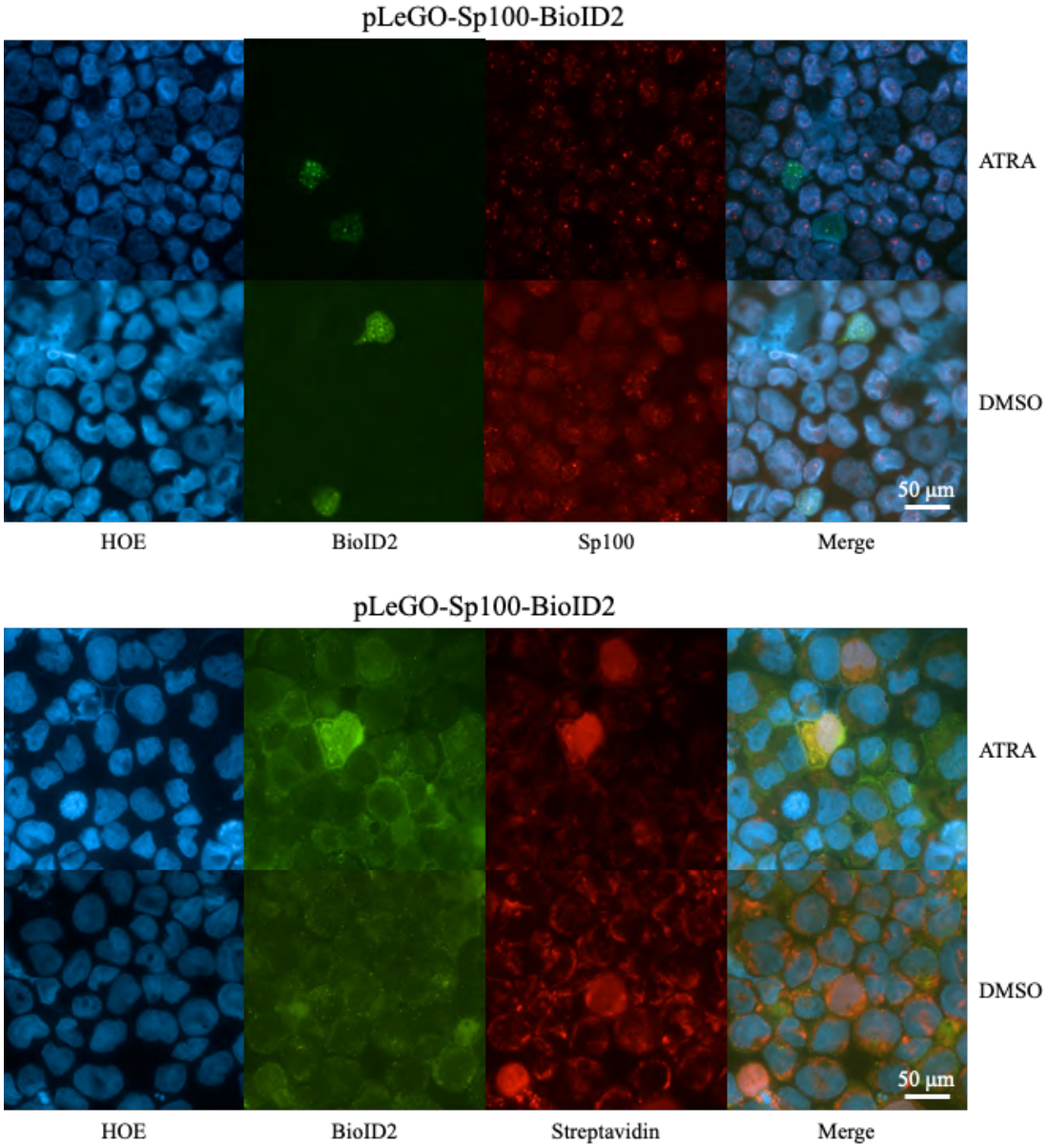


Figure 6-35 Upper panel: Immunofluorescence (IF) staining with Hoechst (Blue) staining for nucleus, BioID2 (Green), Sp100 (Red) antibody to analyze the expression of Sp100-BioID2 fusion protein in NB4-SB1 cells treated with DMSO or ATRA for 3 days, scale bar is 50 µm, showing 99% of the cells failed to give BioID2 signal, suggesting that most of the transduced cells did not express the fusion protein. Lower panel: Immunofluorescence (IF) staining with Hoechst (Blue) staining for nucleus, BioID2 (Green), Streptavidin (Red) antibody, scale bar is 50 µm. However, expressed the protein suggests that the Sp100-BioID2 fusion protein is functional, as clear dots of colocalized BioID2 and Sp100 in these cells only showed up after ATRA treatment.

6.5.2. ATRX-BioID2/BioID2-ATRX generation and characterization

As the focus of our study is the PAX complex, for a third, and last, bait protein in these continued studies, we have chosen the ATRX protein. We have generated an N-terminal and also a C-terminal BioID2 fusion with ATRX. We reasoned these complementary constructs should provide an immediate, internal control for identified factors. Also, the large size of ATRX made dual fusions a more promising approach.

6.5.2.1. Generation of NB4 sublines stably expressing ATRX-BioID2 or BioID2-ATRX

We created two constructs, pLeGO-ATRX-BioID2 and pLeGO-BioID2-ATRX, in which the ATRX CDS was fused in-frame 5' or 3' of the BioID2 CDS (see section 9.1.3, the plasmid map see Figure 9-3). Moreover, the ATRX-BioID2 fusion protein (Figure 9-9) has a HA-tag, a myc-tag between the ATRX and BioID2 parts and a FLAG-tag at its C-terminus, while the BioID2-ATRX fusion protein (Figure 9-10) has a myc-tag at its N-terminus, a FLAG-tag between the BioID2 and ATRX parts and a HA-tag at its C-terminus.

The reason for making two different fusion proteins with different orientation of the ATRX and BioID2 parts was our wish to examine whether the orientation can affect the ability of ATRX binding partners to interact with the fusion protein.

After transduction, Puromycin selection, and sorting by FACS, we obtained the pure (~99%) GFP-positive population for both constructs. The derived sublines were christened NB4-AB and NB4-BA, respectively.

6.5.2.2. Validation of the NB4-AB and NB-BA sublines

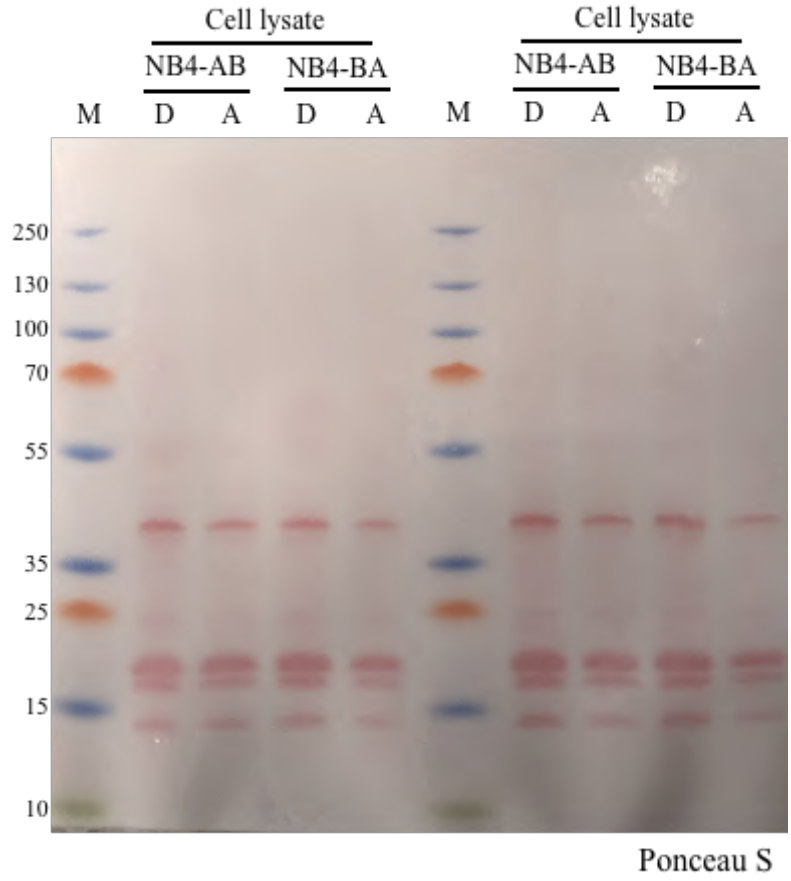
Western blotting (Figure 6-36) and IF staining (Figure 6-37 and Figure 6-38) were performed to confirm that the sublines indeed express the corresponding fusion proteins, both under control (DMSO) and ATRA treatment conditions.

NB4-AB and NB-BA cells were treated with ATRA or DMSO for 3 days. Biotin was added to the cells 20 h prior to collection. Whole cell extracts were subjected to SDS-PAGE. After the transfer, we used Ponceau S staining as a loading control and proceeded to blot the membrane with different specific antibodies, including anti-BioID2, anti-ATRX, and Streptavidin. Our objective was to confirm both the expression of the two fusion proteins and their migration in SDS-PAGE.

Blotting with anti-ATRX gave strong bands at the correct molecular weight region. However, these bands probably correspond to endogenous ATRX, as they do not appear in either the anti-BioID2 or the anti-FLAG blot. Interestingly, NB4-AB cells display a band below 55 kDa in both in the anti-BioID2 and the anti-FLAG blot. This band cannot be attributed to non-specific binding, as it does not appear in the blots of NB4-BA cells. This might suggest some orientation-derived effect in the stability of the fusion proteins, with the degradation of ATRX-BioID2 resulting in a the release of a ~50 KDa peptide that contains the entirety of the BioID2 part (thus, it reacts with both anti-FLAG and anti-BioID2) and a small piece (but not an intact epitope, thus it is not recognized by anti-ATRX) of the ATRX part.

Even though the western blots were rather disappointing, the IF images were promising. Specifically, under ATRA treatment conditions, a small number of NB4-AB (Figure 6-37) and NB4-BA (Figure 6-38) cells gave speckles in which BioID2 and ATRX staining colocalized, suggesting that they actually express the corresponding fusion proteins and, importantly, the fusion proteins integrate into the reconstituted PML-NBs. Therefore, as with NB4-SB1 cells, even though these cells only constitute small fraction in the respective sublines (10% of NB4-AB cells, 20%

of NB4-BA cells), we decided to use these cultures in Streptavidin pull-down experiments. As mentioned earlier, the small representation of these cells will affect the sensitivity but not the specificity of the specific assay.



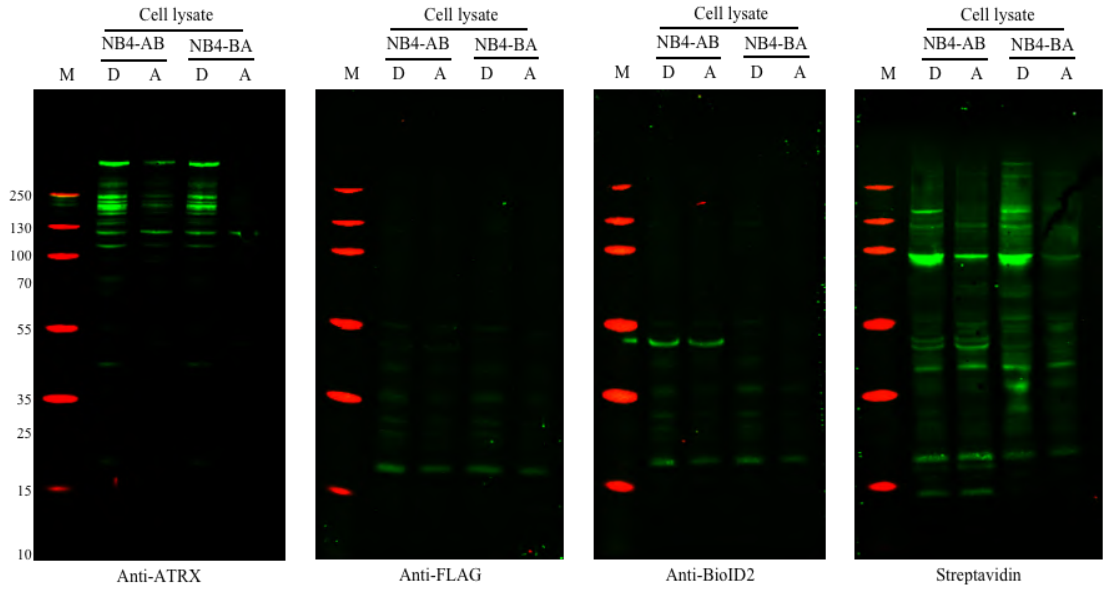


Figure 6-36 Upper: Ponceau S staining of the membrane prior to blotting by the whole cell extract of the NB4-AB and NB4-BA cells with ATRA- and DMSO-treatment, served as loading control. Lower: Western blot to confirm the expression of the ATRX-BioID2 and BioID2-ATRX fusion proteins by NB4-AB and NB4-BA cells using anti-ATRX, anti-FLAG, anti-BioID2 and Streptavidin antibodies, respectively. Blotting with anti-ATRX gave strong bands at the correct molecular weight region, however, they do not appear in either the anti-BioID2 or the anti-FLAG blot.

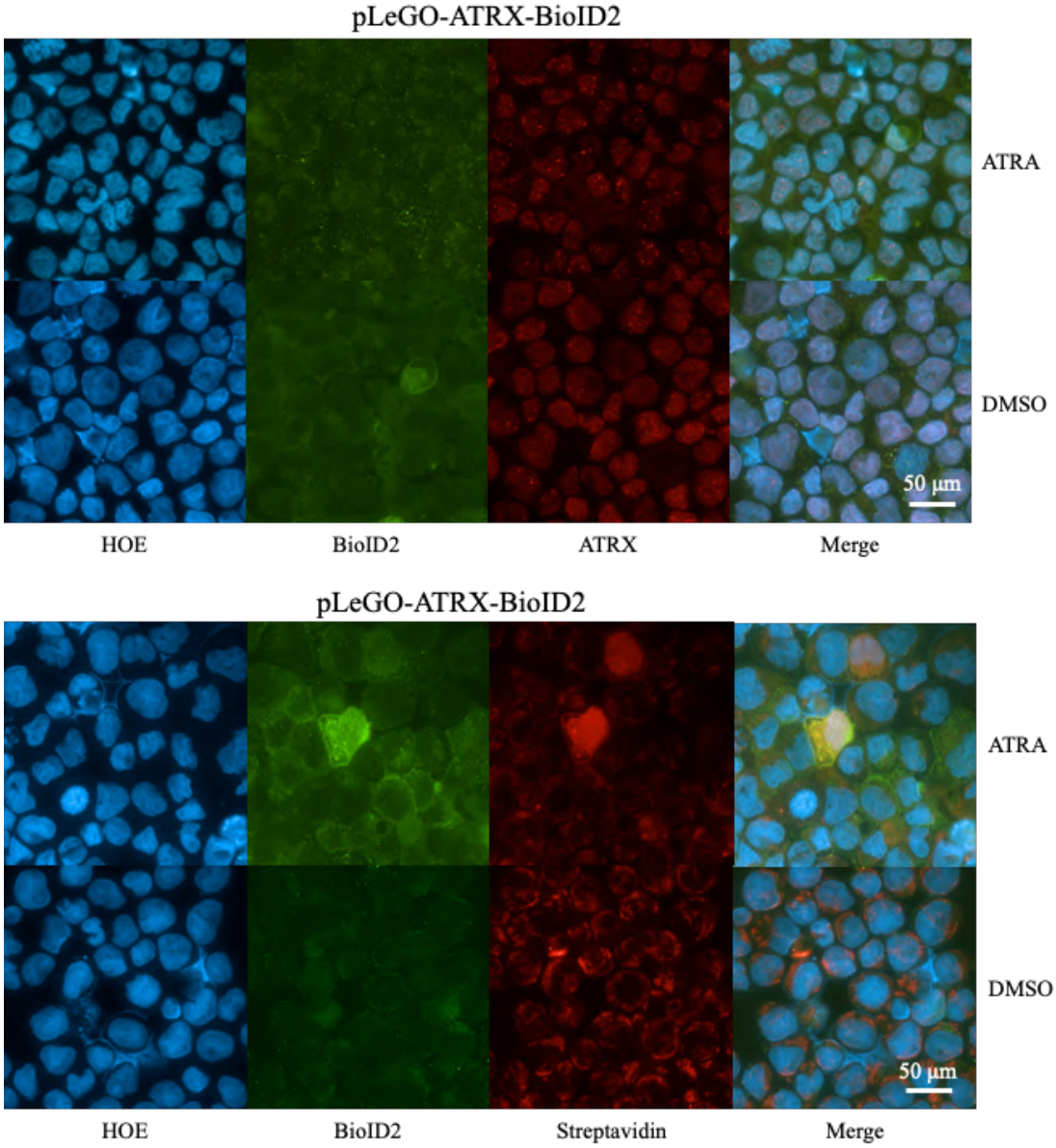


Figure 6-37 Upper panel: Immunofluorescence (IF) staining with Hoechst (Blue) staining for nucleus, BioID2 (Green), ATRX (Red) antibody to confirm the expression and the proper localization of the ATRX-BioID2 fusion protein in NB4-AB cells, giving speckles in which BioID2 and ATRX staining colocalized, suggesting that they actually express the ATRX-BioID2 fusion proteins, scale bar is 50 µm. Lower panel: Immunofluorescence (IF) staining with Hoechst (Blue) staining for nucleus, BioID2 (Green), Streptavidin (Red) antibody, scale bar is 50 µm.

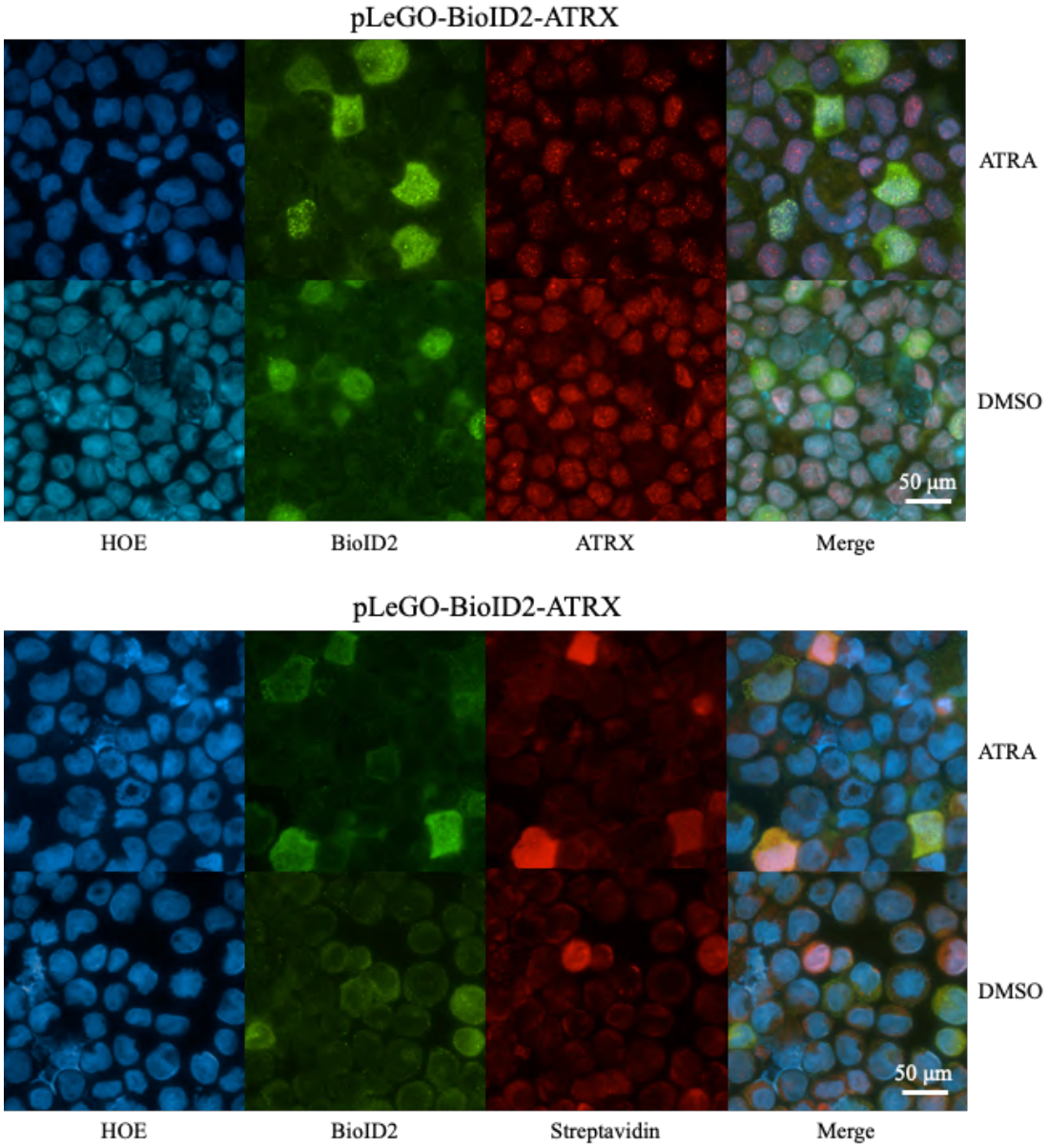


Figure 6-38 Upper panel: Immunofluorescence (IF) staining with Hoechst (Blue) staining for nucleus, BioID2 (Green), ATRX (Red) antibody to confirm the expression and the proper localization of the BioID2-ATRX fusion protein in NB4-BA cells, giving speckles in which BioID2 and ATRX staining colocalized, suggesting that they actually express the BioID2-ATRX fusion proteins, scale bar is 50 µm. Lower panel: Immunofluorescence (IF) staining with Hoechst (Blue) staining for nucleus, BioID2 (Green), Streptavidin (Red) antibody, scale bar is 50 µm.

6.5.3. Streptavidin Pull-Downs Using the NB4 Sublines expressing different BioID-fusion proteins ("Baits")

We were finally able to proceed to the parallel pull-down. Even though the NB4-BD had already been assessed using this method, we decided to include them in this experiment for reasons of comparison.

The pull-downs and their analysis were performed (as described in section 6.2.1.9 "Mass Spectrometry-Scale BioID2 Pull-Down of ATRA-Treated NB4-BD Cells and Controls"). Again, an aliquot of the Streptavidin beads of each pull-down assay underwent elution and the proteins were separated by electrophoresis and silver-stained. After the silver stains confirmed the presence of proteins (Figure 6-39), the remaining beads were sent for Mass-spec analysis by on-beads digestion.

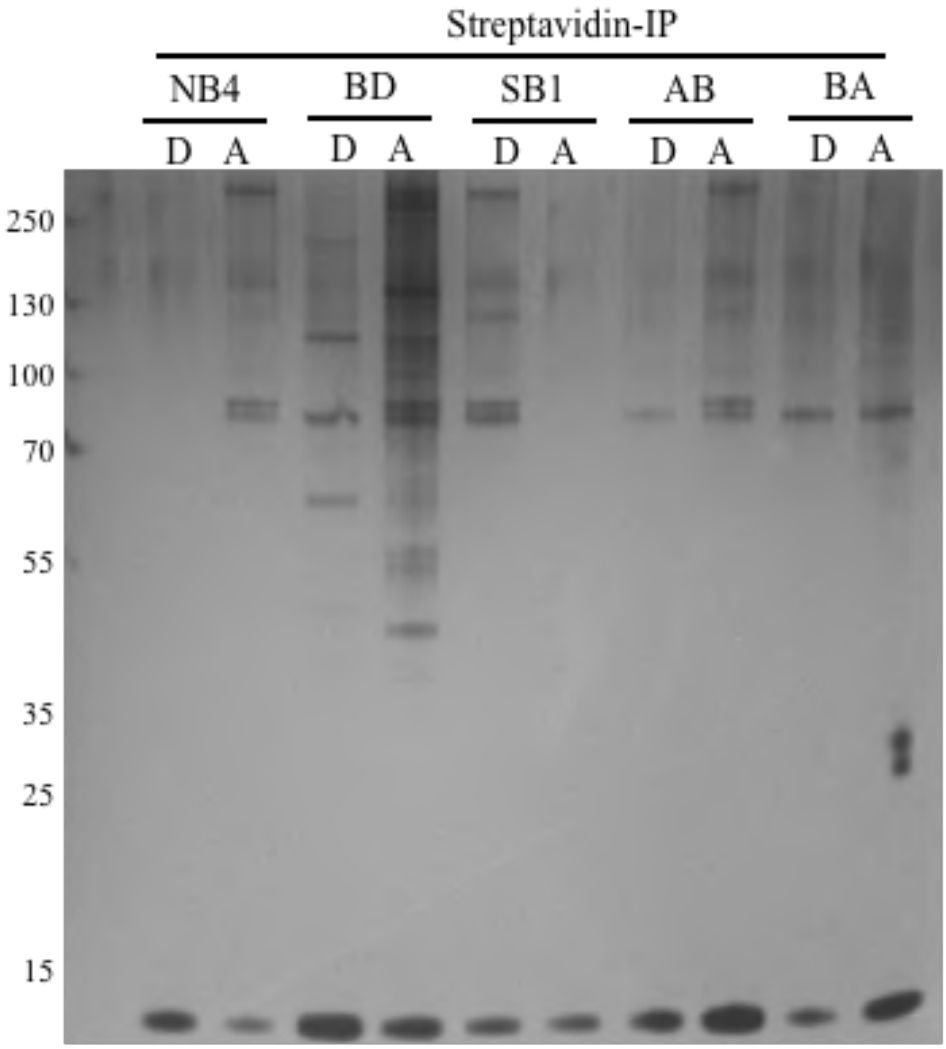


Figure 6-39 Silver staining of Streptavidin IPs from BioID2 pull-down using various NB4 sublines, including ATRA- (A) and DMSO-treated (D) parental NB4 cells, NB4-BD cells, NB4-SB1 cells, NB4-AB cells, and NB4-BA cells, confirming the presence of proteins which can be detected by mass-spec.

The Mass-Spec facility provided us with lists of identified proteins for each NB4 line and condition. Our first step to analyze these data was to create Venn diagrams, each containing the two data sets from the parental cells (NB4 DMSO and NB4 ATRA) and the two corresponding datasets from a specific subline (Figure 6-40). What interests us in each of these diagrams is the subline-specific proteins that can be found in both the DMSO- and the ATRA-treated cells of the subline ("shared") and the subline-specific protein that can only be found in the ATRA-treated cells of the subline ("ATRA-specific").

The NB4-BD subline gave 457 shared and 868 ATRA-specific proteins. These numbers are much higher than the ones of our previous BioID2 pull-down with this subline (see section 6.2.1.9). This may be attributed to fact that the number of NB4-BD cells used in this round of pull-downs was much higher than the one used in the first round, increasing the sensitivity of the method, and also due to the Streptavidin High-Performance Spintrap have a higher binding affinity.

NB4-BD cells also gave by far the largest lists of proteins among the other sublines used in the current experiment. This could represent a higher number of proteins interacting with Daxx than with Sp100 or ATRX. However, a more likely explanation is, that the amount of starting material also played an important role. In NB4-BD cultures, almost all cells expressed the desired fusion protein. In contrast, the vast majority of cells in the culture of the other sublines did not express the fusion protein, resulting in a higher amount of "background" proteins and, thus, a lower sensitivity.

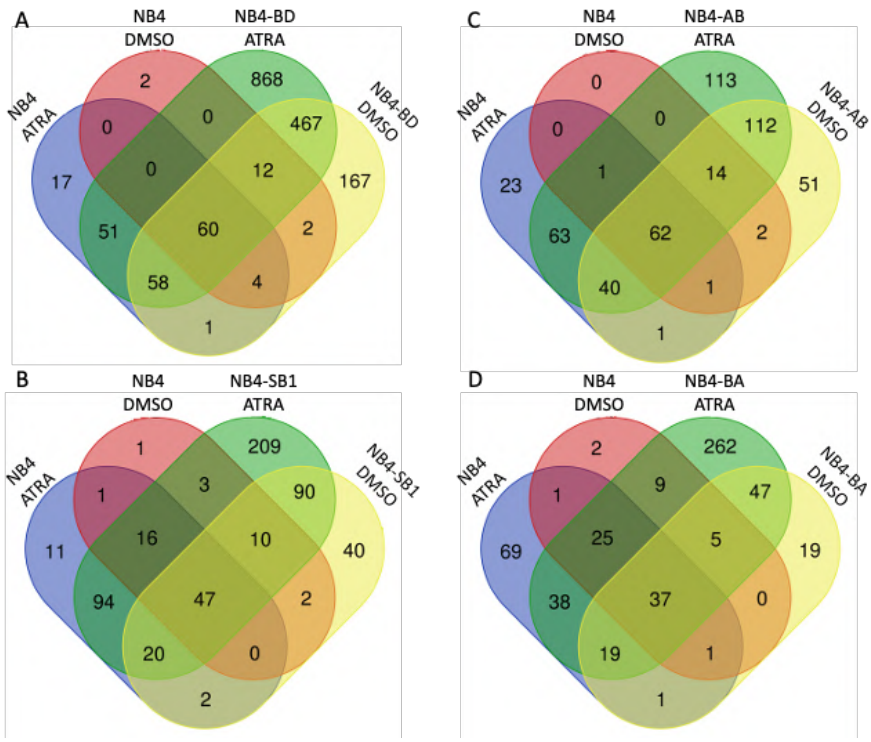


Figure 6-40 Venn diagrams (http://bioinformatics.psb.ugent.be/webtools/Venn/) for the proteins of the ATRA- and DMSO-treated NB4 cells and each of the corresponding data from the NB4-BD (A: Purple: ATRA-treated NB4 cells; Red: DMSO-treated NB4 cells; Green: ATRA-treated NB4-BD cells; Yellow: DMSO-treated NB4-BD cells), showing 1335 (868 and 467) and 634 (467 and 167) proteins in ATRA- and DMSO-treated NB4-BD cells, respectively, in which 467 proteins were shared; NB4-SB1 (B: Purple: ATRA-treated NB4 cells; Red: DMSO-treated NB4 cells; Green: ATRA-treated NB4-SB1 cells; Yellow: DMSO-treated NB4-SB1 cells), showing 299 (209 and 90) and 130 (90 and 40) proteins in ATRA- and DMSO-treated NB4-SB1 cells, respectively, in which 90 proteins were shared; NB4-AB (A: Purple: ATRA-treated NB4 cells; Red: DMSO-treated NB4 cells; Green: ATRA-treated NB4-AB cells; Yellow: DMSO-treated NB4-AB cells), showing 225 (113 and 112) and 163 (112 and 51) proteins in ATRA- and DMSO-treated NB4-BD cells, respectively, in which 112 proteins were shared; and NB4-BA cells (A: Purple: ATRA-treated NB4 cells; Red: DMSO-treated NB4 cells; Green: ATRA-treated NB4-BA cells; Yellow: DMSO-treated NB4-BA cells), showing 309 (262 and 47) and 66 (47 and 19) proteins in ATRA- and DMSO-treated NB4-BD cells, respectively, in which 47 proteins were shared.

6.5.4. Comparison of Interactors from Different Sublines

An important additional benefit of having interaction data from different, independent NB4 sublines expressing BioID2-fusion proteins of different components of the same macrocomplex is the possibility of utilizing these data, to cross-validate the interactomes obtained from each subline. This provides us, beyond ATRA-treatment with an additional way to determine high-confidence hits. We expect true interactors to be found in the positive with more than with one bait protein. Likewise, the proteome from the parental NB4 cells is the ideal negative control. So by subtracting proteins also identified in these cells (defined as Background), we can remove a large number of false positives.

6.5.4.1. Parental NB4 / NB4-BD / NB4-SB1

Figure 6-41 displays a Venn diagram including the data sets of NB4, NB4-BD, and NB4-SB. After NB4 parental cell line extraction, there is a sum of 220 shared proteins. One notable example was SMCHD1, a protein that was shown to interact with Daxx in the first round of pull-down experiments (see section 6.2.1.9 “Mass Spectrometry-Scale BioID2 Pull-Down of ATRA-Treated NB4-BD Cells and Controls”). In addition to presenting the list of identified peptides (9.4 List of Mass Spec Results) and also to check for possible interactions between the 220 proteins, we entered them into the interaction-mapping STRING-website. The resulting network is presented below (Figure 6-42). It reveals three dominant groups, namely, Histones (A), ribosomal proteins (B), and nuclear ribonucleoproteins (C).

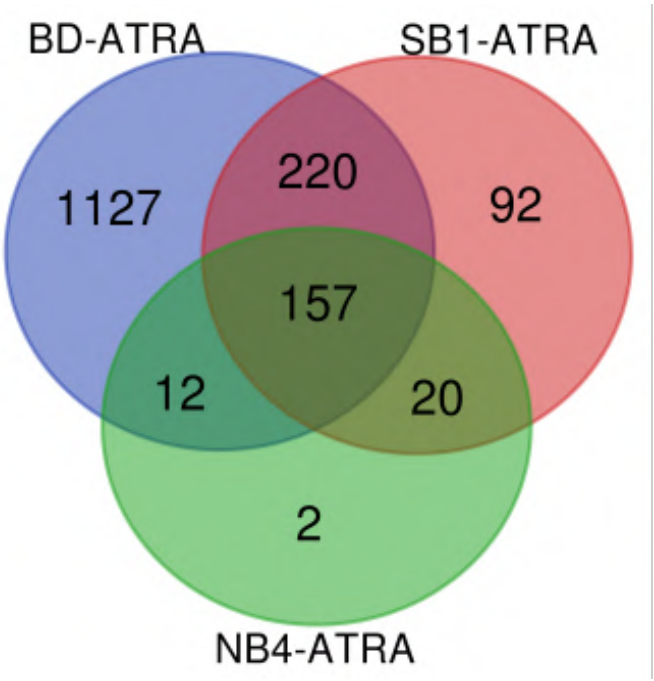


Figure 6-41 Venn diagram (http://bioinformatics.psb.ugent.be/webtools/Venn/) for the proteins of the ATRA-treated NB4 (Green), ATRA-treated NB4-BD (Purple), and ATRA-treated NB4-SB1 (Red) datasets, after NB4 parental cell line extraction, showing 1347 (1127 and 220) and 312 (220 and 92) proteins in ATRA-treated NB4-BD and NB4-SB1 cells, respectively, in which shows a sum of 220 shared proteins.

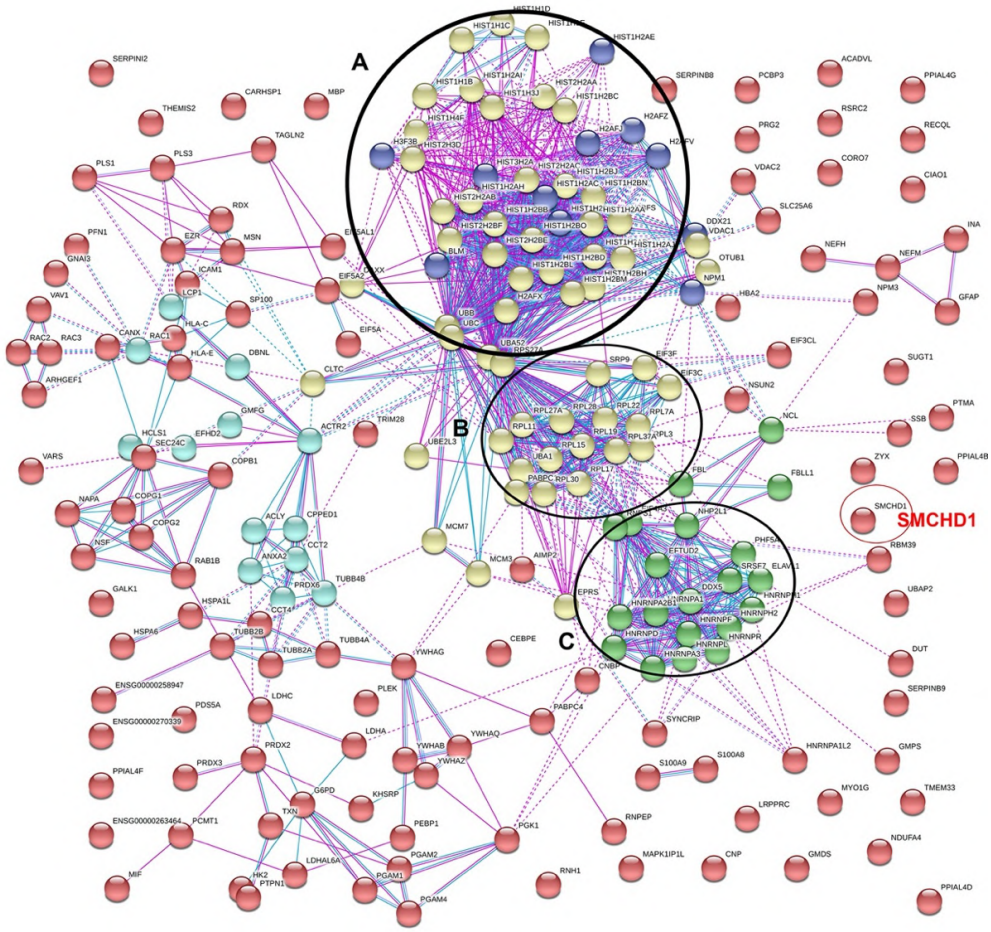


Figure 6-42 STRING analysis (http://www.string-db.org) derived protein-protein interaction networks for the 220 proteins that were identified as interactors of both Daxx and Sp100 (previous figure). The network nodes represent proteins. Splice isoforms or post-translational modifications are collapsed, i.e., each node represents all proteins produced by a single, protein coding gene. Lines indicate known interactions (Light blue line: from curated databases; Pink line: experimentally determined) with medium (0.400) confidence score as the analysis parameter. kmeans clustering (network is clustered to a specified number of clusters) number is 5: Three major groups can be distinguished in the figure, namely Histones (A, yellow and blue balls), ribosomal proteins (B, yellow balls), and nuclear ribonucleoproteins (C, green balls). Proteins from our screen with no major known interactions are shown as red, isolated circles without connecting lines. One of these is SMCHD1 (shown inside red circle), which had already been identified as a Daxx interactor by the first round of pull-downs.

6.5.4.2. NB4-AB and NB4-BA

We wanted to determine whether the orientation of the fusion protein affected the ability of the interactions. To this end, we created a Venn diagram containing the NB4-AB DMSO, NB4-AB ATRA, NB4-BA DMSO, and NB4-BA ATRA datasets (Figure 6-43). The NB4-AB DMSO and NB4-BA DMSO datasets contained 283 and 129 proteins, respectively, with 100 shared proteins. The NB4-AB ATRA and the NB4-BA ATRA datasets contained 405 and 442 proteins, respectively, with 256 common proteins.

We conclude that the orientation of the fusion protein greatly affects its ability to interact with other proteins, which is not surprising, given the large size of ATRX (2492 amino acids, calculated molecular weight 282 kDa). Thus, our choice to use both constructs was sound, as would have otherwise lost many interesting candidates.

6.5.4.3. Parental NB4 / NB4-BD / NBA4-AB / NB4-BA

Finally, we created a Venn diagram containing only the data sets for the proteins that form the ATRX/Daxx chaperone, i.e., the data from the NB4-BD, NB4-AB, and NB4-BA sublines. Moreover, as we were mostly interested in the PAX complex, only the proteins found in ATRA-treated cells (which have reconstituted PML-NBs) were included.

After parental cell line subtraction, there were 258 proteins that were shared by the NB4-BD subline and at least one of the two ATRX sublines (NB4-AB and NB4-BA) (Figure 6-44). These proteins were imported into the interaction-mapping STRING-website. The resulting network is presented below (Figure 6-45). The two dominant protein groups were found, namely Histones (A) and nuclear ribonucleoproteins (B).

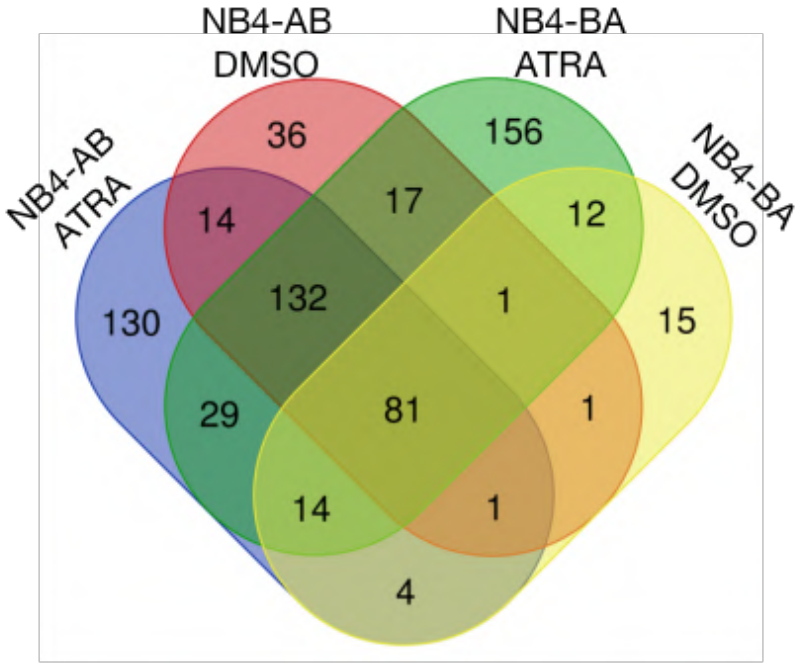


Figure 6-43 Venn diagram (http://bioinformatics.psb.ugent.be/webtools/Venn/) containing the NB4-AB DMSO (Purple), NB4-AB ATRA (Red), NB4-BA DMSO (Green), and NB4-BA ATRA (Yellow) datasets. The DMSO-treated NB4-AB (Red) and DMSO-treated NB4-BA (Yellow) datasets contained 283 and 129 proteins, respectively, with 100 shared proteins. The ATRA-treated NB4-AB (Purple) and the ATRA-treated NB4-BA (Green) datasets contained 405 and 442 proteins, respectively, with 256 common proteins.

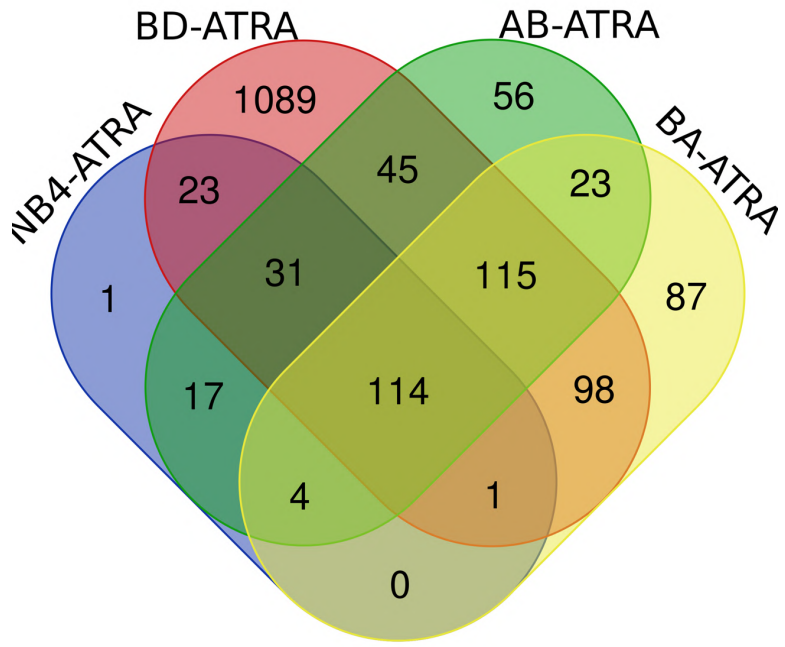


Figure 6-44 Venn diagram (http://bioinformatics.psb.ugent.be/webtools/Venn/) containing the ATRA-treated NB4 (Purple), NB4-BD (Red), NB4-AB (Green), and NB4-BA (Yellow) datasets. After parental NB4 cell line subtraction, there were 258 proteins that were shared by the NB4-BD (Red) subline and at least one of the two ATRX sublines (NB4-AB (Green) and NB4-BA (Yellow)).

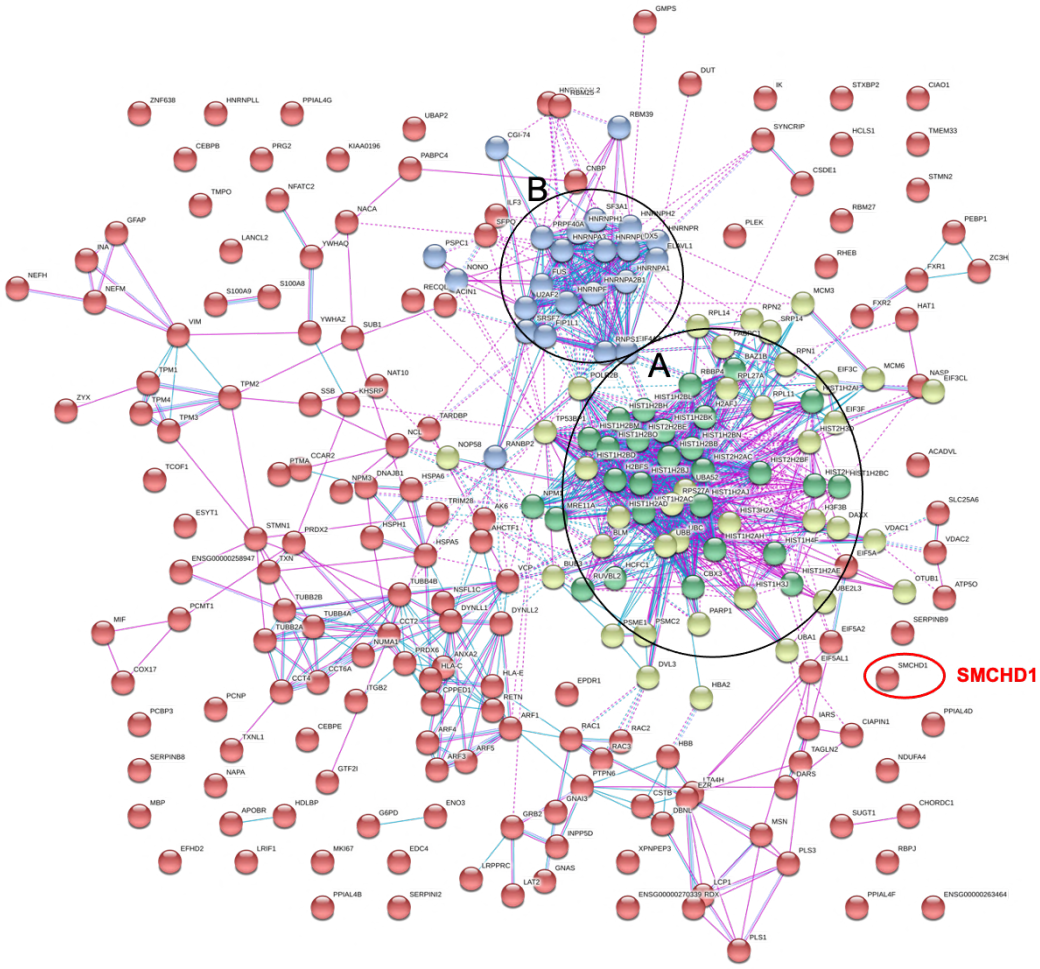


Figure 6-45 STRING analysis (http://www.string-db.org) derived protein-protein interaction networks for 258 proteins that were identified as interactors of the Daxx-BioID2 and at least one of the two ATRX fusion proteins (ATRX-BioID2 or BioID2-ATRX) under ATRA treatment conditions, after parental NB4 cell line subtraction. The network nodes represent proteins. Splice isoforms or post-translational modifications are collapsed, i.e., each node represents all proteins produced by a single, protein coding gene. Lines indicate known interactions (Light blue line: from curated databases; Pink line: experimentally determined) with medium (0.400) confidence score as the analysis parameter. kmeans clustering (network is clustered to a specified number of clusters) number is 4: Two major clustered populations can be distinguished in the figure, namely Histones (A, yellow and green balls) and Heterogeneous nuclear ribonucleoprotein (B, blue balls). SMCHD1 (shown inside red circle), which had already been identified as a Daxx interactor by the first round of pull-downs.

7. Discussion

7.1. Daxx/ATRX Complex

7.1.1. Daxx

The Daxx protein, which is the focus of this project, was originally identified in a yeast two-hybrid (Y2H) screen as a cytoplasmic interactor of the intracellular domain of Fas[88], the latter being a membrane-bound receptor protein transmitting a pro-apoptotic signal initiated by binding of its cognate ligand FasL (Fas-Ligand)[154].

For years, a cytoplasmic function of Daxx was still postulated[155]. However, detecting cytoplasmic interactions of Daxx using Y2H was problematic because Y2H proteins are targeted into the nucleus by means of fusion with NLS-containing DNA-binding or transactivation domains[156]. As, for instance, SUMOylation of Sp100 depends on an active NLS[157], it could be speculated that cytoplasmic protein domains may have undergone unphysiological SUMOylation, which may have led to binding to the SUMO interaction motif to Daxx, producing a false positive signal.

7.1.2. Co-Immunoprecipitation (Co-IP)

In the original outline for this project, it was planned to perform a classical complex purification using the FLAG epitope tag to pull-down the components. However, as noticed in the very first experiments, the Daxx/ATRX complex always ends up in the insoluble nuclear matrix fraction, as its matrix-attached components keep the whole complex insoluble. Releasing Daxx and ATRX into the soluble fraction demanded conditions too harsh for performing IP using FLAG antibodies. This also explains, why, although an intensely studied topic for years, the composition of PML-NBs has not yet been conclusively been determined by Mass spectrometry, as the same problems exist there.

To overcome this problem, we adopted the BioID system, as biotinylation persists even in the harsh conditions required to solubilize the proteins of the complex. This strategy not only allowed us to study the interactions of the Daxx/ATRX project but also produced, through the use of Sp100, provides many novel proteins that may also be constituents of PML-NBs.

7.2. Confirmation of Interaction

7.2.1. FRET

FRET is a physical phenomenon that is being widely used in biomedical fields. FRET relies on the distance-dependent transfer of energy from a donor molecule to an acceptor molecule. Due to its sensitivity to distance, FRET has been used to investigate molecular interactions in which distance needs to be less than 10 nm. In our BioID method, the Biotin labelling radius is around 10 nm[158], which is comparable with the distance in the FRET assay.

In our FRET assay, the pYFP-CFP linker construct served as our positive control, whereas cells transfected with the acceptor-only pYFP construct, the donor-only pCFP, pYFP-Daxx, pCFP-SMARCA4, or pYFP+pCFP were our negative controls. The FRET signal of pYFP-Daxx+pCFP-SMARCA4 was statistically significantly higher than all of the control background signals, suggesting an interaction indeed exists between Daxx and SMARCA4. Unfortunately, though the difference was statistically significant, the FRET signal was not large enough to give a definite answer. We suspect that the inability of FRET to provide a strong distance is due to the fact that the two proteins are big (Daxx 740 amino acids-82.2 kDa; SMARCA4 1647 amino acids-184.6 kDa).

7.2.2. Nano-Glo Dual-Luciferase assay

In Nano-Glo Dual-Luciferase assay, cells transfected with pM-Daxx+pVp16-SMARCA4 were used to test the interaction between Daxx and SMARCA4, whereas cells transfected with pM-Vp16 or pM+pVp16 served as our positive controls and cells transfected with pM, pVp16, pM-Daxx, pVp16-SMARCA4, or pM+pVp16 were

negative controls. We calculated the firefly luminescence relative values by normalizing the firefly luminescence with NanoLuc luminescence data. As there was no strong signal from the pM-Daxx+pVp16-SMARCA4-transduced cells, the assay failed to confirm the interaction between the two proteins.

Again, the big size of the two proteins may be responsible for our lack of success. Because of their size, both proteins have huge surface areas. This means that pM53 or pVp16 may be too far from the interacting areas, which would preclude the production of signal from the assay. In the current work, we only tested pM-Daxx (Daxx with C-terminal pM) and pVp16-SMARCA4 (SMARCA4 with N-terminal pVp16), which did not give us confirmation of the interaction between Daxx and SMARCA4. In the future, alternative fusion sites should be tested.

7.2.3. Proximity ligation assay

Among the methods used to determine and confirm protein-protein interactions, Proximity ligation assay (PLA) has been the most popular and promising in recent years thanks to its ability to directly detect proteins, protein interactions, and post-translational modifications with high sensitivity and specificity[159]. PLA can be used to quickly (within hours) confirm results from conventional co-immunoprecipitation and co-localization techniques, even those having to do with transient or weak interactions, using only a few cells or sub-cellular events[160]. PLA has been used to study cancer[161,162] and development[163,164].

Due to these features, PLA will in all probability be used in our future studies to verify the interaction between Daxx and SMARCA4 and other interesting candidates.

7.3. Protein-protein Interactions

7.3.1. Protein-protein Interactions and Diseases

The human interactome is composed of a huge number of different types of PPIs[165], which can be classified into families based on their structural and functional diversity[166]. PPIs participate in a large number of biological reactions, and act through the protein networks or protein-protein complexes[167], and are regulated by the interactions of a great number of noncovalent side chains of amino acid residues[168], which allows them to bind to each other to prevent or enhance its biological activities[169].

Extracellular PPIs and cell-cell adhesion are necessary for the interaction between bacteria such as the Salmonella and their hosts during infection[170], which renders the inhibition of these interactions a useful strategy for the prevention of such diseases[171]. PPIs are also central in the life cycle of viruses. For example, the remarkable genetic diversity of the HIV-1 allows it to produce a similarly diverse set of proteins whose PPI with host proteins modulate pathogenesis and are thus critical for diagnosis, drug susceptibility, and vaccine development[172].

7.3.2. Protein-protein Interaction Inhibitors

Unlike traditional drug targets, whose inhibitors block binding on the cofactors or substrates to assert their functions, PPIs are inhibited by chemical agents that disrupt the conformation of protein complexes or cut off the PPI network[167]. Being essential for almost every biological process, PPIs provide a vast of therapeutic targets[173]. Therapeutic antibodies showed that PPIs would be of great value for these kinds of targets[174].

High-resolution imaging has shown that PPI interfaces are large and flat[175]. In contrast, small molecules use “deep pockets” to bind their targets[176]. For these reasons, it is important to find small molecules with the ability to bind to or even be recruited to the protein complex. “Hot spots” for this are important for the binding of the protein complex[177–180]. The small molecules bind deeper on the surface hot spots of the target protein and with higher efficiencies than the natural protein partner[181]. In recent years, several databases with PPI targets emerged[182–184]. The hot spots of most PPI-targeting inhibitors used in clinical trials are concentrated near the small binding area[183,185].

For the identification of more inhibitors, we need not only basically screening approaches such as NMR[186,187], tethering[188], and virtual screening[189], but also synthesis and modified the already exit compound with their structural insight.

7.4. Identification of Daxx Interactors

Our study applied the second generation of the BioID method (BioID2)[143] to assess the protein-protein interactions (PPIs) of the PML-associated Daxx/ATRX complex. This method was chosen because the unusual harsh lysis conditions required to break the PAX complex's tight association with the nuclear matrix compromised antibody-binding, making standard co-immunoprecipitation impossible. Mass Spec analysis of the streptavidin pull-downs from controls and ATRA-treated NB4-BD cells gave us a large number of proteins that interacted with Daxx independently (detected in samples of DMSO- and ATRA-treated cells) or dependently (detected only in samples of ATRA-treated cells) of PML-NBs. Both lists contained several previously described Daxx interactors, confirming the validity of the method. Examples of such proteins in the first list (ATRA-independent) were ATRX[138], HP1α[190], Histone H3.3[84], and Histone H4[5]. Examples from the second list (ATRA-dependent) are PML[191], Sp100[192], and DEK[193]. Even though the latter three proteins have been previously reported as Daxx interactors, our data prove for the first time that these interactions are disrupted by PML-RAR expression alongside with the integrity of PML-NBs. Besides detecting already described interactors, we have identified a considerable number of proteins that have not been previously found in this complex. We decided to initially focus on two proteins, namely, SMCHD1 and SMARCA4. These two candidates were chosen because the high number of identified peptides made SMCHD1 a high-confidence interactor. SMARCA4 was chosen because it opened an unexpected collaborative opportunity with the neighboring group of Prof. U. Schüller which gave us immediate access to materials, such as antibodies, the SMARCA4 cDNA, and even a mouse model for functional studies. Western blotting had immediately confirmed both proteins to have increased association after ATRA treatment, serving as first specificity control and suggesting functional relevance.

7.4.1. SMCHD1

SMCHD1 (Structural Maintenance of Chromosomes Hinge Domain containing protein 1) was one of our highest peptide count candidate Daxx-interactors in both DMSO- and ATRA-treated NB4-BD cells (see 9.4 List of Mass Spec Results), but absent in parental NB4 cells, suggesting that this interaction takes place irrespective of the presence of PML-NBs. Confirmation of these interactions with methods other than BioID2 largely proved problematic, as the high-molecular-weight of the protein (predicted 226 kDa) and its tight association with the nuclear matrix made their application extremely difficult. Thus, we had to independently confirm this interaction with BioID2 experiments using other "bait" fusion proteins. Indeed, streptavidin pull-down experiments in NB4-SB1 (i.e., cells transduced to express Sp100-BioID) cells treated with ATRA also identified SMCHD1, thus independently confirming that SMCHD1 is indeed a novel interactor of the PML-associated Daxx/ATRX complex.

This protein is an epigenetic regulator[194] that creates a transcriptionally repressive environment[195,196], which is critical for X-chromosome inactivation and hence essential for female embryo viability[197]. SMCHD1 variants are associated with facioscapulohumeral muscular dystrophy (FSHD) and the craniofacial disorder Bosma arhinia microphthalmia syndrome (BAMS)[198–200]. The fact that we find this protein as a PAX-complex-associated factor may offer novel leads to understand how the PAX complex may exert a repressive function in the chromatin context. SMCHD1's function in the context of X chromosome inactivation provides a highly interesting novel aspect in connecting the heterochromatin-association of Daxx/ATRX-mediated H3.3 deposition with X inactivation, an RNA-guided epigenetic silencing event[201]. This could provide new mechanistic clues for how the repressive state is achieved by Daxx/ATRX-mediated depositioning.

7.4.2. SMARCA4/BRG1

SMARCA4 (also called BRG1) was found to interact with Daxx only under ATRA treatment conditions, suggesting that this interaction only takes place in the presence of intact PML-NBs. This specificity serves as an indication that SMARCA4 has functional relevance with the PAX complex. Our efforts to confirm the pull-down findings using alternative methods had mixed success. On one side, despite the difficulties deriving from the large size of SMARCA4, we successfully confirmed the Daxx-SMARCA4 interaction in vitro and in vivo[202] using co-IP and FRET, respectively. On the other side, IF staining did not reveal a co-localization of SMARCA4 and Daxx in reconstituted PML-NBs. This could be due to multiple factors such as suboptimal staining conditions or issues with the sensitivity or epitope accessibility of the anti-SMARCA4 monoclonal antibody. Another explanation is the possibility of the interaction requiring the existence of intact PML-NBs but taking place outside them. After all, PML-NBs have been hypothesized to serve as scaffolds for assembly of multiprotein complexes[203]. This protein belongs to the large SWI/SNF ATPase family, whose members associate with chromatin directly or indirectly[204]. They consume ATP to move nucleosomes along the DNA, an ability that was discovered due to the resulting changes in the accessibility of the affected regions to nuclease digestion[205,206].

Global cancer genomics projects revealed that SWI/SNF ATPases are mutated in approximately 20% of all human cancers[207,208], with SMARCA4 being the most frequently mutated among them[209]. Cancer types in which loss of SMARCA4 expression has been observed include non-small cell lung cancer[210–213], rhabdoid tumor[214,215], medulloblastoma[216], Burkitt's lymphoma[217], and acute lymphoblastic leukemia (ALL)[218]. These data support the hypothesis that SMARCA4 is a tumor suppressor, which originated from studies in knock-out mice[145,219]. These studies showed that, even though SMARCA4-null mice were embryonic lethal, heterozygotes survived but had a much-increased probability of developing mammary carcinomas.

However, Watanabe et al.[220] observed that SMARCA4 was frequently overexpressed in colorectal cancer specimens, whereas the knockdown of SMARCA4 impaired cellular growth. Despite the aforementioned fact SMARCA4 heterozygous mice develop mammary tumors, Qiong et al.[152] reported that SMARCA4 is actually overexpressed in human primary breast cancers, while knocking down of SMARCA4 in a triple-negative breast cancer cell line reduced tumor formation both in vivo and in vitro. Other cancers also showed an oncogenic role of SMARCA4[221,222]. As a whole, these studies suggest that the SMARCA4 role varies depending on the type or even the stage of cancer.

In our study, knocking-out SMARCA4 in mouse bone marrow cells using the Cre-Lox recombination system caused the termination of cell proliferation. Moreover, the overexpression of the oncoproteins cMyc, mPR or H3.3 K27M, which under non-knockout conditions leads to the immortalization of these cells, was unable to overcome the loss of SMARCA4, i.e., these cells also stopped growing. This is consistent with the previously reported embryonic lethality in null mice[220].

As the Daxx-SMARCA4 interaction is affected by PML-RAR and ATRA treatment in NB4 cells, we established NB4-SMARCA4 knockdown cells to further analyze the role of SMARCA4 in the context of disrupted or ATRA-restored PML-NBs.

From the Cre/Lox studies, we expected that SMARCA4 reduction would not be tolerated by NB4 cells. However, we readily obtained a line stably expressing the knockdown construct. The line displayed SMARCA4 levels equal to about 30% of the ones observed in the parental cells. Surprisingly, the proliferation rate of the knockdown cells did not differ significantly from the one of the parental cells, indicating that this level of SMARCA4 reduction can be tolerated by NB4 cells. If this is related to the disruption of the Daxx-SMARCA4 interaction needs to be addressed in future studies. As in parental NB4 cells, ATRA-treatment led to a halt in proliferation. This is not affected by the lower levels of SMARCA4.

7.4.3. Other Potentially interesting Interactors

7.4.3.1. Lamin B Receptor (LBR)

Although not yet verified by independent means, LBR is in many ways a promising candidate through its functional properties. Like SMCHD1, LBR is involved in X inactivation. It directly interacts with the XIST long non-coding RNA, which mediates spreading of heterochromatin across the inactive X chromosome in the process known as X inactivation in female individuals. This again emphasizes a possible functional relationship between the PAX complex, which, through the Daxx/ATRX H3.3 chaperone, serves depositioning H3.3 into Heterochromatin[87].

7.4.3.2. Nijmegen breakage Syndrome Protein (Nibrin, NBS)

NBS1 is a protein involved in DNA double strand break repair. Remarkably, it also plays an essential role in host restriction of Herpes simplex virus 1 via direct interaction with the viral transactivator protein ICP0[223]. Remarkably, ICP0 is also the protein responsible for disassembly of PML-NBs during infection[224], hence NBS might represent the “docking site” for ICP0 to target PML-NBs.

7.4.3.3. Heterochromatin protein 1 α (HP1α)

The HP1α is a key player in the establishment and maintenance of heterochromatin[225], again emphasizing the role of the PAX complex in the regulation of heterochromatin.

7.5. Drug Target

One of the important goals of our research is always to try and identify potential novel targets for cancer therapy.

7.5.1. SMARCA4 as a Potential Drug Target

Our experiments eliminating SMARCA4 in bone marrow cells immortalized by the expression of various oncogenes clearly demonstrate that SMARCA4 would be an excellent candidate for therapeutic intervention, as inactivation of SMARCA4 stopped growth even of cells transformed with cMyc, one of the most powerful oncoproteins.

Besides the ATPase domain, SMARCA4 also contains a bromodomain. These modules can be found in many transcription factors. Most of the bromodomain-harboring proteins, including SMARCA4, are classified as members of the Bromodomain and Extra-Terminal Domain (BET) family. Bromodomains are epigenetic "readers" that recognize acetylated Lysines. This recognition leads to the binding of the transcription factor and, ultimately, the activation of gene transcription[226]. BET-containing proteins are often overexpressed in cancers[227]. As a result, the inhibition of BET-mediated binding to the chromatin can reduce cell proliferation and induce apoptosis, rendering such inhibition a potential therapeutic approach[228].

Since SMARCA4 contains a BET, we decided to examine whether its inhibition can stop the proliferation of NB4 cells. For this purpose, we used PFI-3, an inhibitor specific for the bromodomains of SMARCA4 and SMARCA2[229]. Treating NB4 cells and our two NB4-derived sublines, namely, the SMARCA4-overexpressing NB4-KD.S4 and the SMARCA4-knockdown NB4-OE.S4, with PFI-3 did stop the proliferation of all three lines, a result that differs from what Vangamudi et al. observed in four lung cancer cell lines[230]. However, the IC50 values for the three NB4-based lines we used (parental, NB4-KD.S4, and NB4-OE.S4) did not correlate

with their relative SMARCA4 content, strongly suggesting that the cessation of growth had little to do with the inhibition of the bromodomain of SMARCA4.

Despite the results of the inhibitor experiments, we still consider the development of more effective inhibitors or inhibitors targeting other domains of SMARCA4 as a promising approach for novel cancer therapeutics. This belief is based on the drastic anti-proliferative effect of knocking out SMARCA4 in bone marrow cell lines, both non-transduced controls as well as transduced to express oncoproteins that normally (i.e., under non-knockout conditions) are able to immortalize the cells, among them cMyc, one of the most powerful oncoproteins.

7.5.2. Relevance of SMARCA4 for Proliferation

As shown by our Inducible Cre/Lox SMARCA4 knockout experiments, SMCARCA4 is essential for proliferation in murine bone marrow.

SMARCA4/Brg1, belongs to the "Bromo- and Extra-Terminal domain" (BET)-family of proteins, which can be inhibited by the BET inhibitor PFI-3.

Bromodomains, present in many proteins, are "readers" of acetylated lysine epigenetic marks on histone tails, typically activating gene transcription[226]. BET proteins are key players in many types of cancer. For instance, they are required and overexpressed in glioblastoma[227] and also leads to aberrant expression of the *MYC* oncogene in various hematologic malignancies such as mixed lineage leukemia[231]-fusion AML (MLL-AML)[232], AML[233], Burkitt's lymphoma[234], and B-ALL[235]. Use of BET inhibitors in these malignancies can reduce cell proliferation and induce apoptosis. Therefore, targeting the acetyl-lysine binding property of BET proteins is a potential therapeutic approach of cancer[228].

Filippaopoulos and Knapp[236] gave a review on recent progress in the development of several potent bromodomain inhibitors and their potential applications in cancer inflammation drug discovery. In 2009, Triazolothienodiazepines were identified as the first BETi by an anti-inflammatory phenotypic assay[237]. JQ1 binds to bromodomains with high potency and specificity to BET family member BRD4[238],reducing cancer cell viability in vitro and in vivo by G1 cell cycle arrest and S phase percentage decrease[239]. Ceribelli M et al.[240] reported that pharmacologic interference of BET proteins reduces NF-κB activity and reduces the viability of the B-cell-like subtype of diffuse large B-cell lymphoma.

For NB4 cells, however, there were no major differences in proliferation between the parental NB4 cells and their SMARCA4 knock-down subline. This could either mean that NB4 cells express SMARCA4 in a quantity much higher than the minimum required for proliferation so that, despite the knock-down, NB4-KD.S4 cells have a sufficient quantity, or, alternatively, that these cells no longer require SMARCA4 for proliferation, possibly because they lack intact PML-NBs.

Under both DMSO- and ATRA treatment conditions, no significant differences among the three cell types were observed (Figure 6-36), clearly showing that restoration of PML-NBs does not affect the effects of SMARCA4 reduction by knockdown.

Thus, we conclude that SMARCA4 abundance, at least at the level of magnitude achieved in our experiments, does not affect the growth of NB4 cells. Also, the BET-inhibitor PFI-3 did not produce an effect relating to SMARCA4 abundance. We believe, only a knockout study could indisputably demonstrate whether the growth of NB4 is truly SMARCA4-independent. Although NB4 cells and the two sublines (knockdown/overexpression) exhibit some sensitivity to PFI-3, the response of the two sublines does not differ. This might be explained, if the growth-reducing effect is actually caused by inhibition of the second PFI-3-specific target protein,

SMARCA2[152]. If so, naturally, modulation of the amount of SMARCA4 would of course not produce an effect.

Concerning the biological role of the Daxx/SMARCA4 interaction, SMARCA4/Brg1 and SMARCA2/Brm are alternative ATPase subunits of the SWI/SNF Chromatin remodeling complex[241]. It is tempting to speculate the PAX-interaction could orchestrate the exchange of the SMARCA4/SMARCA2 subunits in this complex, as cells exit the proliferative phase and enter differentiation, a hypothesis that could be promising to follow up.

7.6. Technical Considerations

NB4 sublines expressing different bait proteins had surprising differences in the percentage of cells that actually expressed the desired protein. BioID2 fusion protein expression was observed in more than 90% of NB4-BD cells, while in only around 1% in NB4-SB1 cells and 10–20% of NB4-AB and NB4-BA cells. This is bound to lead to differences in sensitivity among the corresponding pull-down experiments. Lower percentages of BioID2 fusion protein-expressing cells lead to lower relative abundance BioID2 fusion protein percentage in the cell extracts, which in turn increases the relative amount of background proteins in the results, compromising the sensitivity (but not the specificity) of the Mass-spec analysis. That is likely the reason we got five times more specific proteins in the BioID2-Daxx cell subline pull-down sample, i.e., the subline with the highest percentage of BioID2 fusion protein-expressing cells. Creating new sublines with higher percentages of fusion protein-expressing cells is the obvious way to increase the sensitivity of the pull-down experiments. As the GFP protein is expressed as a fusion protein with the Puromycin -resistance marker, through an internal Ribosomal entry site (IRES), hence from the same mRNA as the corresponding bait protein, the reasons for this discrepancy are difficult to explain.

For the ATRX protein, we applied both N-terminal and C-terminal tagging, i.e., both BioID2-ATRX and ATRX-BioID2 fusion protein-expressing cells were used. We found that the corresponding list of identified proteins shared about 1/3 of total proteins, confirming specificity. However, the large number of non-common proteins proved that using both orientations was necessary in order to avoid missing some interesting targets of our interaction candidates, which could be due to steric effects. As this is probably true for all BioID2 fusion proteins, a Daxx-BioID2-expressing subline will also have to be produced and used in pull-down experiments to complement the work we have already done with BioID2-Daxx and get a more complete list of Daxx interaction candidates.

But for now, if one interaction candidate showed up in the pull-down experiments of at least two different BioID2 fusion proteins, we preliminarily categorized the candidate as a PAX-complex interactor. Nevertheless, further independent verification methods will need to be attempted to verify the interaction for a journal publication.

There were also some limitations to our Mass-spec analysis. On one hand, for SMARCA4 and Sp100, which are regulatory proteins often be found in low abundance, both showed a very low unique peptide number in our Mass-spec results; however, we were readily able to verify them by western blot analysis, which means we must not ignore low-abundance interaction candidates, even in cases where a single unique peptide was observed, as we may still be able to verify it by using other methods. Besides utilizing other methods, starting with higher cell numbers and using more stringent conditions may be a way to overcome this problem in future studies.

On the other hand, Mass-spec gave a similarly high number of SMCHD1 unique peptides in the pull-down samples of both DMSO- and ATRA-treated NB4-BD cells. However, the western blot verification revealed that, even though the protein did show up in both DMSO- and ATRA-treated NB4-BD samples, the signal was much stronger in the ATRA-treated cells. This result inferred that our Mass-spec results can only give an indication of the occurrence of interaction candidates but cannot precisely reveal their absolute or relative abundance among samples. However, their relative amount can be determined by performing western blot, just as we did in the case of SMCHD1. The low cost of western blotting makes it much more practical than the “absolute-quantification” obtained by Mass-spec.

Last but not least, the Mass-spec list of identified proteins also included other exciting interactor candidates, such as Non-POU domain-containing octamer-binding protein (NONO), Nibrin, and DEK, which, for unknown reasons, could not be verified by western blot. Possible explanations could be the insufficient quality of the antibodies used or that the amount in our sample was too low for antibodies to detect.

Here would be a long way until we figure out what was happened in our study, which unfortunately was not possible within the timeframe available.

7.7. Perspective

For the future, we would consider performing functional studies with SMARCA4 and SMCHD1 in the context of the NB4 model as most promising. NB4, as a cell line, would enable us to perform more detailed, genome-wide studies on effects on PAX-related phenomena, such as H3.3 distribution and chromatin-state changes after ATRA, which induces a myeloid differentiation program.

A potential unifying motif in the Daxx-interactome is the fact that Daxx contains a SUMO interaction motif (SIM)[93]. Many of the prospective interacting proteins contain consensus motives for SUMOylation. Many DNA binding transcription factors undergo SUMOylation[242], which could trigger the association of Daxx. Lin et al.[203] suggested that the SUMOylation of PML-NB-associated factors might play a role in assembling multi-protein transcription regulatory complexes.

With respect to SMARCA4, our results from the knock-out blood marrow-derived cells immortalized by exogenous oncogene expression were inconsistent with the ones we received from knock-down NB4 cells. Specifically, the former stopped proliferating, whereas the latter did not display apparent differences compared to the parental in any tested properties. As this is probably a result of the fact that the knock-down cells still had 30% of normal SMARCA4 content, a rational next step would be to generate NB4 SMARCA4 knockouts, possibly using the CRISPR-Cas system.

In the current thesis, most of our attention was focused on SMARCA4. As a result, no functional analysis of the SMCHD1-Daxx interaction was performed. Future studies might include CRISPR-mediated knockout of SMCHD1 in NB4 cells or analysis of

the consequences of the chromatin changes after treatment with ATRA and DMSO control conditions. Using “Hi-C technology”[243], we could investigate the role of SMCHD1 in changes in chromosome distribution and structure in our NB4 system when we induce myeloid differentiation of the cells with ATRA.

Our work provides an in-depth analysis of proteins associated with a highly cancer-associated epigenetic regulator complex; it exposes unexpected connections of H3.3 depositioning with processes such as X-inactivation and SWI/SNF remodelers. It is a resource for further studies aiming to target the complex for cancer therapy. Likewise, many more promising drug targets may be identified in the future among the multiple yet uncharacterized interactors. In addition, this study has produced numerous new candidates for novel PML-NB resident proteins.

8. Summary

Using NB4 cells as a tool we have studied the PML-associated Daxx/ATRX (PAX) complex. As its constituents play a role in cancer, many of the novel components identified here may represent promising novel drug targets for cancer therapy.

Using the BioID method, we have defined dynamics of the interactome of the Daxx/ATRX complex, which will be an invaluable resource for further studies.

The Mass-spec results of the pull-down experiments using ATRA-treated NB4-BD and NB4-SB1 sublines, revealed SMCHD1 strongly suggest that SMCHD1 is a novel interactors of the PML-associated Daxx/ATRX complex.

Our pull-down experiments also identified SMARCA4 (also called BRG1) as a Daxx interactor. We confirmed this interaction both in vitro and in cells by performing Daxx-SMARCA4 Co-IP and FRET experiments. Moreover, the results of our 4-OH TXF-induced SMARCA4 knockout experiments indicated the dependency of cells immortalized by various oncogenes on SMARCA4 for their proliferation. We believe that finding novel inhibitors and/or seeking for alternative methods to inhibit SMARCA4 are still a promising approach for novel cancer therapeutics.

9. Appendix

9.1. List of Plasmid Maps

9.1.1. pLeGO-BioID2-Daxx

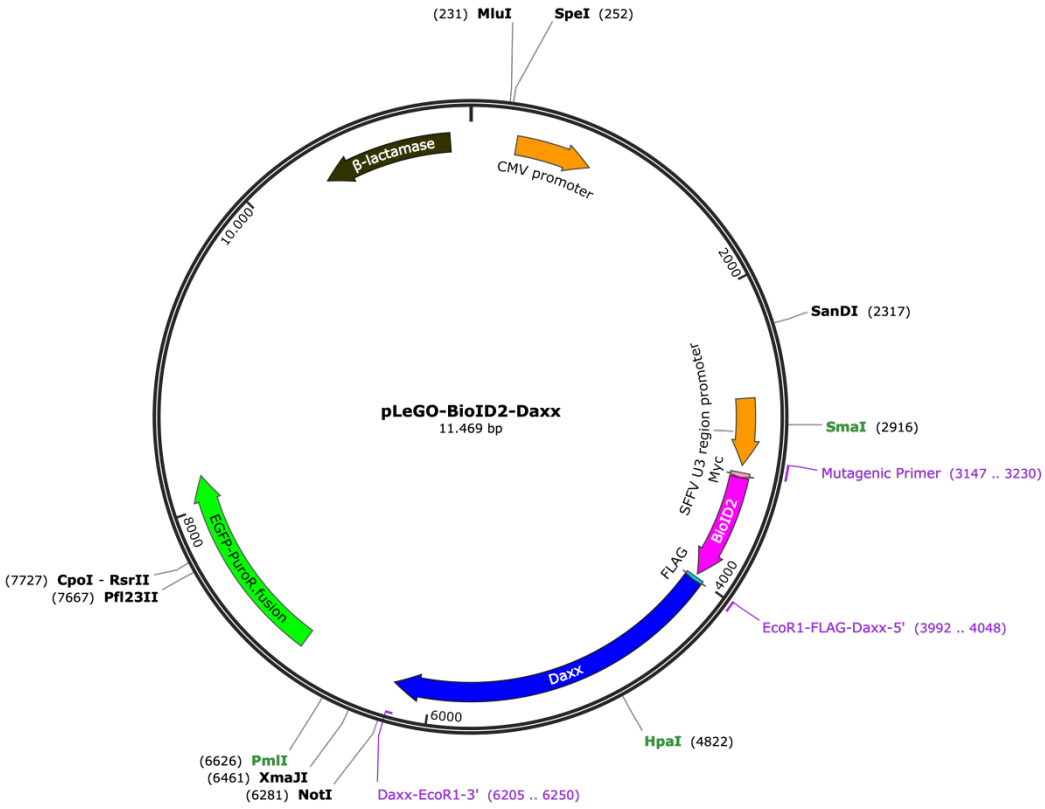


Figure 9-1 pLeGO[244]-BioID2-Daxx expression plasmid vector; the encoded BioID2-Daxx fusion protein carries a FLAG tag and a Myc tag, so it can be recognized by multiple antibodies to ensure specificity in whole-cell lysates. Expression of the fusion protein is driven by the Spleen Focus Forming Virus U3 promoter, the CMV promoter drives expression of the lentiviral pregenomic RNA.

9.1.2. pLeGO-Sp100-BioID2

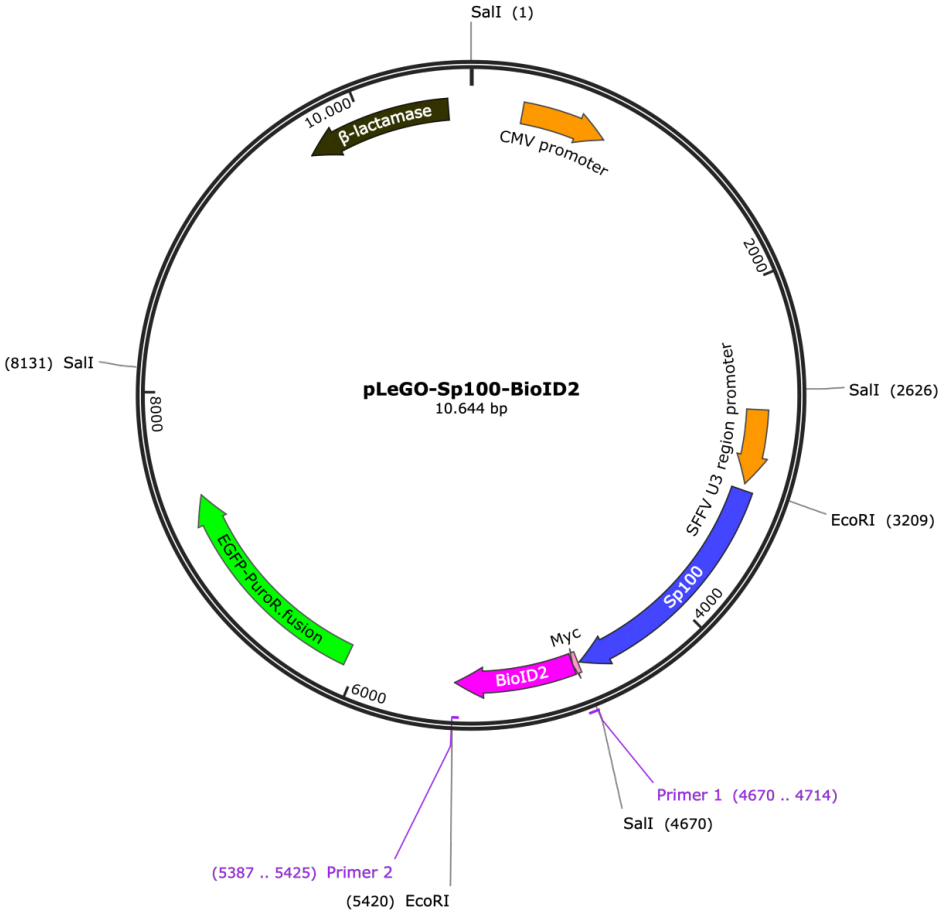


Figure 9-2 Map of the pLeGO-Sp100-BioID2 expression plasmid vector; the encoded BioID2-Sp100 fusion protein carries a Myc tag, so it can be recognized by multiple antibodies to ensure specificity in whole-cell lysates. Expression of the fusion protein is driven by the Spleen Focus Forming Virus U3 promoter, the CMV promoter drives expression of the lentiviral pregenomic RNA.

9.1.3. pLeGO-ATRX-BioID2 and pLeGO-BioID2-ATRX

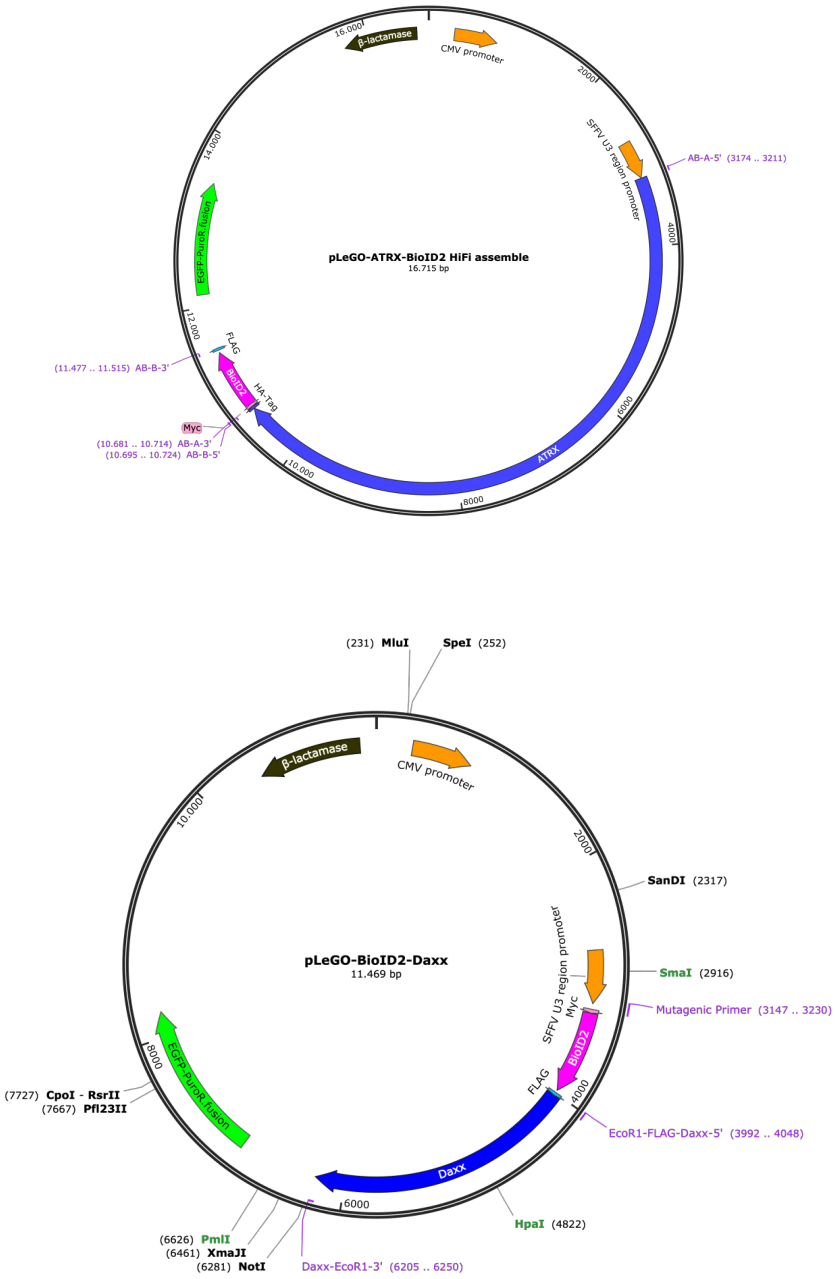


Figure 9-3 pLeGO-ATRX-BioID2 (Left) and pLeGO-BioID2-ATRX (Right) expression plasmid vector; the encoded ATRX-BioID2/BioID2-ATRX fusion proteins carry a FLAG tag and a Myc tag, so they can be recognized by multiple antibodies to ensure specificity in whole-cell lysates. Expression of the fusion protein is driven by the Spleen Focus Forming Virus U3 promoter, the CMV promoter drives expression of the lentiviral pregenomic RNA.

9.1.4. SMARCA4 Overexpression

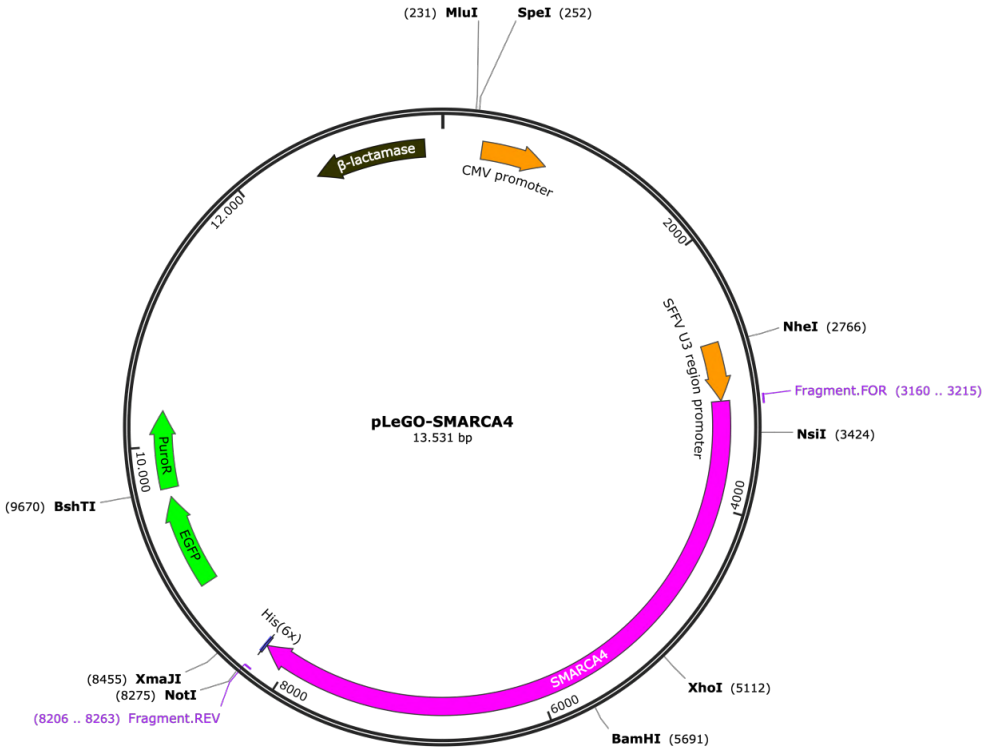


Figure 9-4 pLeGO-SMARCA4 expression plasmid vector; the encoded SMARCA4/BRG1 carries a His(6) tag. Expression of the fusion protein is driven by the Spleen Focus Forming Virus U3 promoter, while the CMV promoter drives expression of the lentiviral pregenomic RNA.

9.1.5. Nano-Glo Dual-Luciferase Reporter Assay

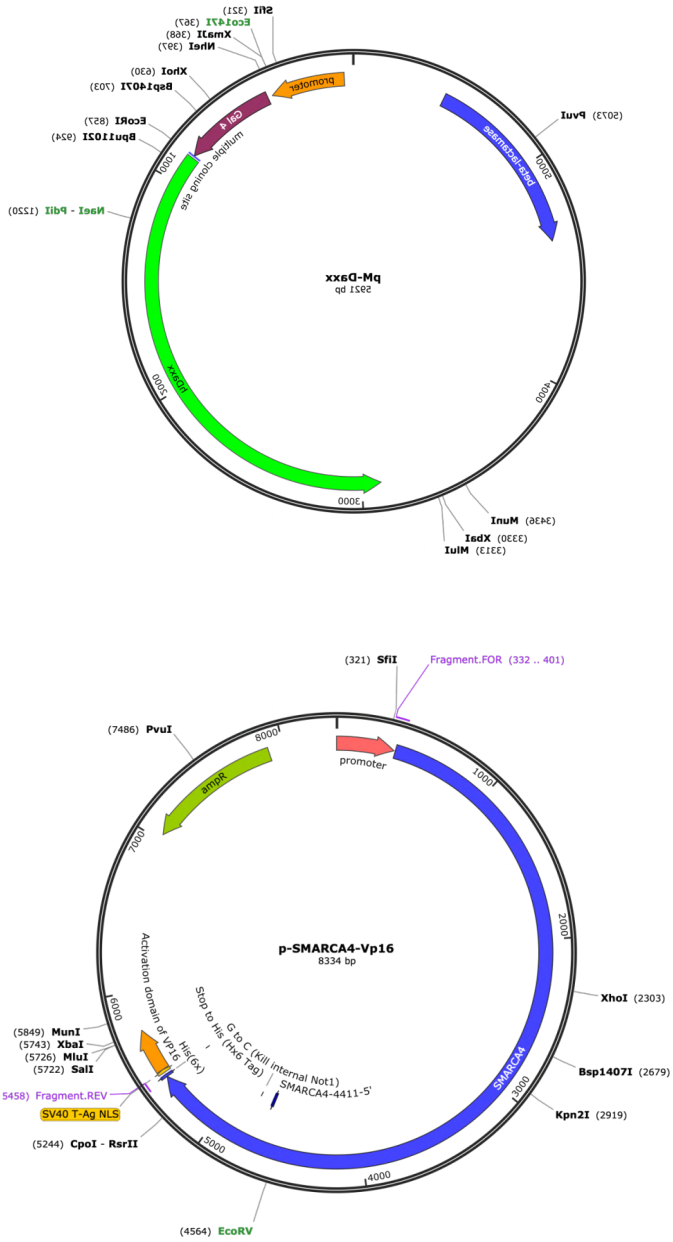


Figure 9-5 Nano-Glo Dual-Luciferase assay pM-Daxx and pVp16-SMARCA4 plasmid maps.

9.1.6. FRET

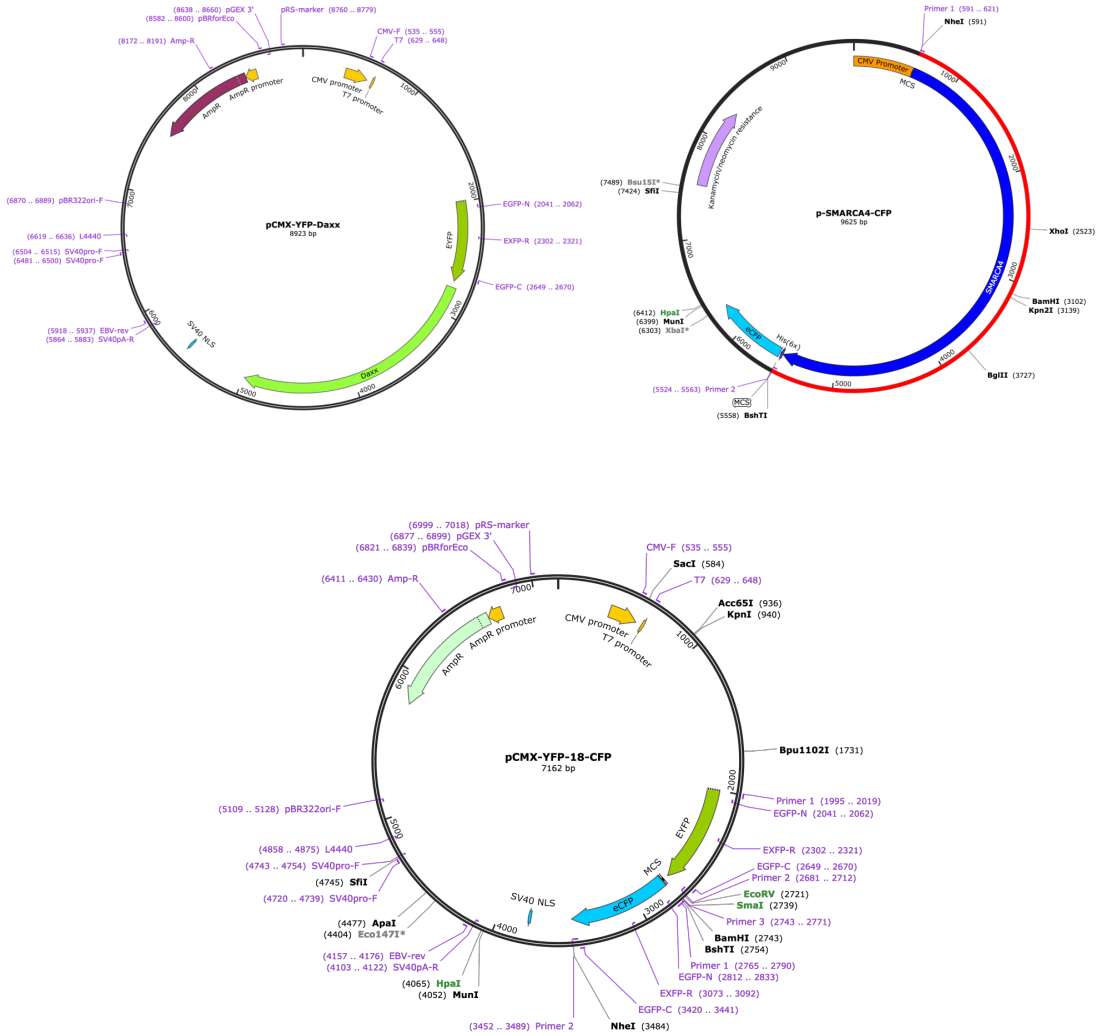


Figure 9-6 FRET assay pYFP-Daxx (Upper Left), pCFP-SMARCA4 (Upper Right), and pCFP-YFP linker (Lower) plasmid maps.

9.2. List of BioID Fusion Proteins

9.2.1. BioID2-Daxx

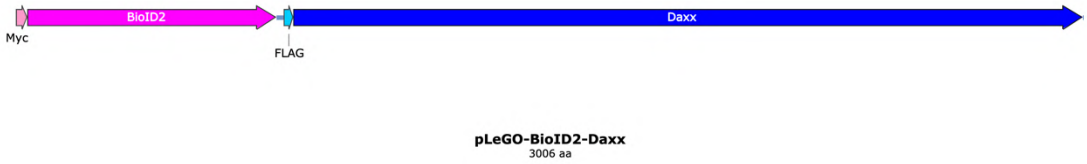


Figure 9-7 BioID2-Daxx fusion protein, made by Snapgene, which contains a myc-tag at its N-terminus and a FLAG-tag between the BioID2 and Daxx parts.

9.2.2. Sp100-BioID2

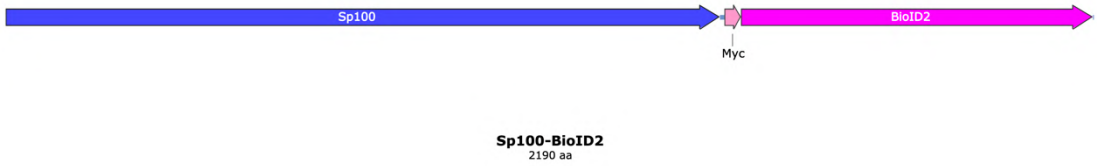


Figure 9-8 Sp100-BioID2 fusion protein, made by Snapgene, which contains a myc-tag between the Sp100 and BioID2 parts.

9.2.3. ATRX-BioID2

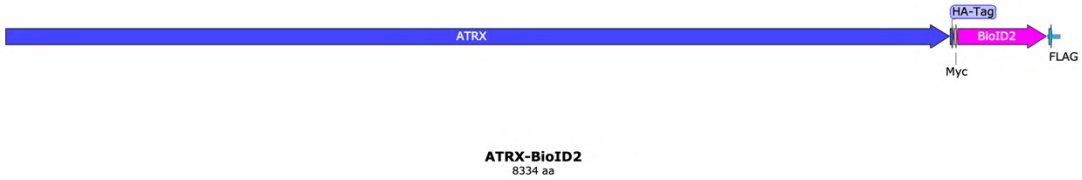


Figure 9-9 ATRX-BioID2 fusion protein, made by Snapgene, which contains a HA-tag, a myc-tag between the ATRX and BioID2 parts and a FLAG-tag at its C-terminus.

9.2.4. BioID2-ATRX

Figure 9-10 BioID2-ATRX fusion protein, made by Snapgene, which contains a myc-tag at its N-terminus, a FLAG-tag between the BioID2 and ATRX parts and a HA-tag at its C-terminus.

9.3. List of MS2 Spectra

9.3.1. MS2 Spectrum of SMCHD1

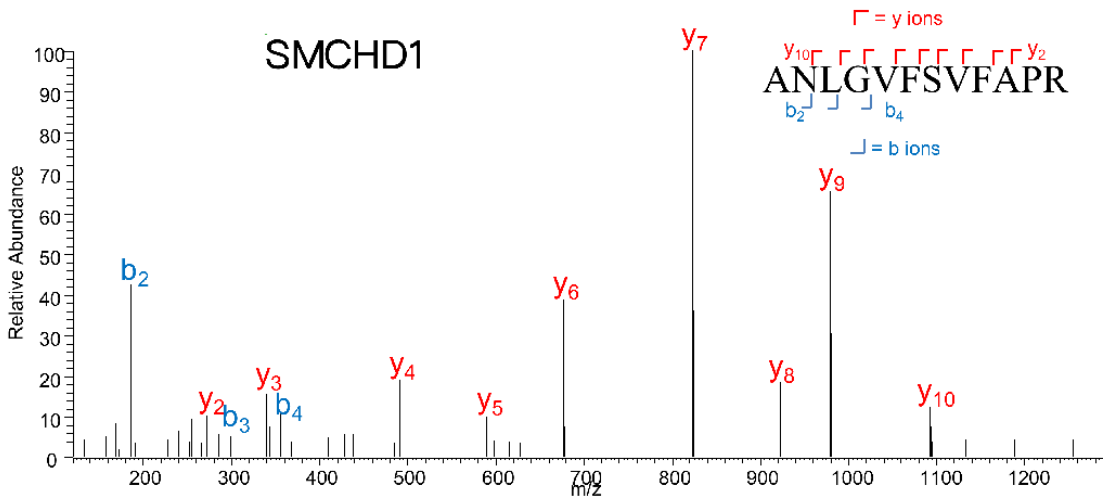


Figure 9-11 MS2 spectrum of SMCHD1. The X-axis gives m/z (mass per charge) to identify amino acids in the analyzed peptide, read from the N-t ("b-ions") or the C-terminus ("y-ions"), resulting in high-confidence identification of the 12-meric peptide shown.

9.3.2. MS2 Spectrum of SMARCA4

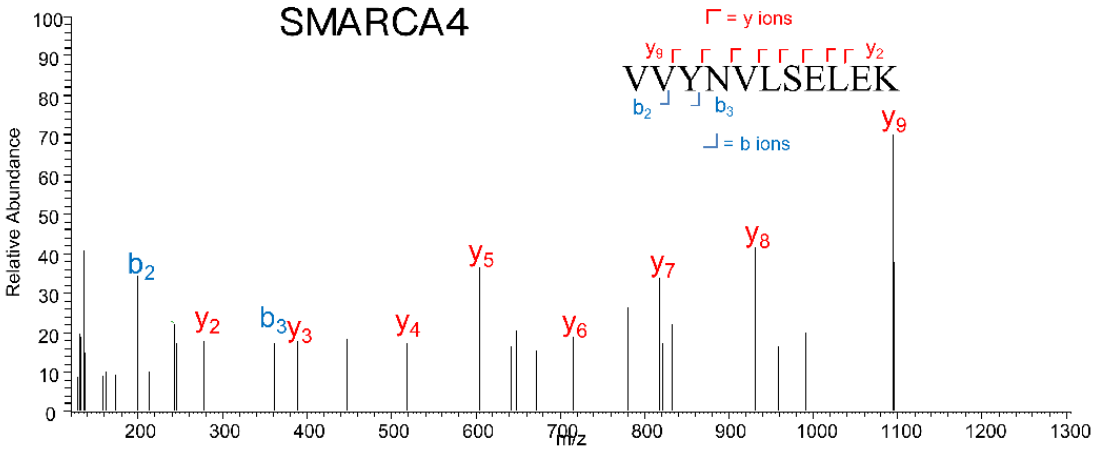


Figure 9-12 SMARCA4 MS2 spectrometry, X-axis gives m/z (mass per charge) to identify amino acids in the analyzed peptide, read from N-terminus (“b-ions”) or C-terminus (“y-ions”), resulting in high-confidence identification of the 11-meric peptide shown.

9.4. List of Mass Spec Results

9.4.1. NB4-BD DMSO List

Table 9-1 NB4-BD DMSO

Accession	Protein Names	# Unique Peptides			
		SUM	Sa.1	Sa.2	Sa.3
Q13085	Acetyl-CoA carboxylase 1 (ACC1) (EC 6.4.1.2) (ACC-alpha) [Includes: Biotin carboxylase (EC 6.3.4.14)]	60	22	16	22
A6NHR9	Structural maintenance of chromosomes flexible hinge domain-containing protein 1 (SMC hinge domain-containing protein 1)	48	17	15	16
Q9UER7	Death domain-associated protein 6 (Daxx) (hDaxx) (ETS1-associated protein 1) (EAP1) (Fas death domain-associated protein)	42	14	12	16
P19338	Nucleolin (Protein C23)	31	9	10	12
P11498	Pyruvate carboxylase, mitochondrial (EC 6.4.1.1) (Pyruvic carboxylase) (PCB)	29	11	8	10
P46100	Transcriptional regulator ATRX (EC 3.6.4.12) (ATP-dependent helicase ATRX) (X-linked helicase II) (X-linked nuclear protein) (XNP) (Znf-HX)	25	10	7	8
Q96RQ3	Methylcrotonoyl-CoA carboxylase subunit alpha, mitochondrial (MCCase subunit alpha) (EC 6.4.1.4) (3-methylcrotonyl-CoA carboxylase 1) (3-methylcrotonyl-CoA carboxylase biotin-containing subunit) (3-methylcrotonyl-CoA:carbon dioxide ligase subunit alpha)	24	8	9	7
O00763	Acetyl-CoA carboxylase 2 (EC 6.4.1.2) (ACC-beta) [Includes: Biotin carboxylase (EC 6.3.4.14)]	15	5	4	6
P05164	Myeloperoxidase (MPO) (EC 1.11.2.2) [Cleaved into: Myeloperoxidase; 89 kDa myeloperoxidase; 84 kDa myeloperoxidase; Myeloperoxidase light chain; Myeloperoxidase heavy chain]	14	6	3	5
Q13428	Treacle protein (Treacher Collins syndrome protein)	12	5	4	3
P05165	Propionyl-CoA carboxylase alpha chain, mitochondrial (PCCase subunit alpha) (EC 6.4.1.3) (Propanoyl-CoA:carbon dioxide ligase subunit alpha)	12	6	2	4
Q9BVA6	Adenosine monophosphate-protein transferase FICD (EC 2.7.7.n1) (AMPylator FICD) (De-AMPylase FICD) (EC 3.1.4.-) (FIC domain-containing protein) (Huntingtin yeast partner E) (Huntingtin-interacting protein 13) (HIP-13) (Huntingtin-interacting protein E)	10	3	3	4
Q9NX58	Cell growth-regulating nucleolar protein	9	4	2	3
P46776	60S ribosomal protein L27a (Large ribosomal subunit protein uL15)	9	3	2	4
P16401	Histone H1.5 (Histone H1a) (Histone H1b) (Histone H1s-3)	9	3	3	3
Q9H0A0	RNA cytidine acetyltransferase (EC 2.3.1.-) (18S rRNA cytosine acetyltransferase) (N-acetyltransferase 10)	8	3	1	4
P83731	60S ribosomal protein L24 (60S ribosomal protein L30) (Large ribosomal subunit protein eL24)	8	3	3	2
P16403	Histone H1.2 (Histone H1c) (Histone H1d) (Histone H1s-1)	7	2	3	2
P16402	Histone H1.3 (Histone H1c) (Histone H1s-2)	7	2	3	2
P10412	Histone H1.4 (Histone H1b) (Histone H1s-4)	7	2	3	2
Q99880	Histone H2B type 1-L (Histone H2B.c) (H2B/c)	6	1	2	3
Q99879	Histone H2B type 1-M (Histone H2B.e) (H2B/e)	6	1	2	3
Q99877	Histone H2B type 1-N (Histone H2B.d) (H2B/d)	6	1	2	3
Q93079	Histone H2B type 1-H (Histone H2B.j) (H2B/j)	6	1	2	3

Q5QNW6	Histone H2B type 2-F	6	1	2	3
Q13185	Chromobox protein homolog 3 (HECH) (Heterochromatin protein 1 homolog gamma) (HP1 gamma) (Modifier 2 protein)	6	3	0	3
P62829	60S ribosomal protein L23 (60S ribosomal protein L17) (Large ribosomal subunit protein uL14)	6	2	2	2
P62807	Histone H2B type 1-C/E/F/G/I (Histone H2B.1 A) (Histone H2B.a) (H2B/a) (Histone H2B.g) (H2B/g) (Histone H2B.h) (H2B/h) (Histone H2B.k) (H2B/k) (Histone H2B.l) (H2B/l)	6	1	2	3
P62805	Histone H4	6	2	2	2
P58876	Histone H2B type 1-D (HIRA-interacting protein 2) (Histone H2B.1 B) (Histone H2B.b) (H2B/b)	6	1	2	3
P57053	Histone H2B type F-S (Histone H2B.s) (H2B/s)	6	1	2	3
P51610	Host cell factor 1 (HCF) (HCF-1) (C1 factor) (CFF) (VCAF) (VP16 accessory protein) [Cleaved into: HCF N-terminal chain 1; HCF N-terminal chain 2; HCF N-terminal chain 3; HCF N-terminal chain 4; HCF N-terminal chain 5; HCF N-terminal chain 6; HCF C-terminal chain 1; HCF C-terminal chain 2; HCF C-terminal chain 3; HCF C-terminal chain 4; HCF C-terminal chain 5; HCF C-terminal chain 6]	6	2	2	2
P42766	60S ribosomal protein L35 (Large ribosomal subunit protein uL29)	6	2	2	2
P42166	Lamina-associated polypeptide 2, isoform alpha (Thymopoietin isoform alpha) (TP alpha) (Thymopoietin-related peptide isoform alpha) (TPRP isoform alpha) [Cleaved into: Thymopoietin (TP) (Splenin); Thymopentin (TP5)]	6	2	1	3
O60814	Histone H2B type 1-K (H2B K) (HIRA-interacting protein 1)	6	1	2	3
Q9Y230	RuvB-like 2 (EC 3.6.4.12) (48 kDa TATA box-binding protein-interacting protein) (48 kDa TBP-interacting protein) (51 kDa erythrocyte cytosolic protein) (ECP-51) (INO80 complex subunit J) (Repressing pontin 52) (Reptin 52) (TIP49b) (TIP60-associated protein 54-beta) (TAP54-beta)	5	2	1	2
Q8N257	Histone H2B type 3-B (H2B type 12)	5	1	2	2
Q16778	Histone H2B type 2-E (Histone H2B-GL105) (Histone H2B.q) (H2B/q)	5	1	2	2
P84098	60S ribosomal protein L19 (Large ribosomal subunit protein eL19)	5	2	1	2
P33778	Histone H2B type 1-B (Histone H2B.1) (Histone H2B.f) (H2B/f)	5	1	2	2
P23527	Histone H2B type 1-O (Histone H2B.2) (Histone H2B.n) (H2B/n)	5	1	2	2
P06899	Histone H2B type 1-J (Histone H2B.1) (Histone H2B.r) (H2B/r)	5	1	2	2
P06733	Alpha-enolase (EC 4.2.1.11) (2-phospho-D-glycerate hydro-lyase) (C-myc promoter-binding protein) (Enolase 1) (MBP-1) (MPB-1) (Non-neural enolase) (NNE) (Phosphopyruvate hydratase) (Plasminogen-binding protein)	5	1	2	2
Q9UPN9	E3 ubiquitin-protein ligase TRIM33 (EC 2.3.2.27) (Ectodermin homolog) (RET-fused gene 7 protein) (Protein Rfg7) (RING-type E3 ubiquitin transferase TRIM33) (Transcription intermediary factor 1-gamma) (TIF1-gamma) (Tripartite motif-containing protein 33)	4	1	1	2
Q9P2N5	RNA-binding protein 27 (RNA-binding motif protein 27)	4	1	1	2
Q9BQE3	Tubulin alpha-1C chain (Alpha-tubulin 6) (Tubulin alpha-6 chain) [Cleaved into: Detyrosinated tubulin alpha-1C chain]	4	2	0	2
Q5VTE0	Putative elongation factor 1-alpha-like 3 (EF-1-alpha-like 3) (Eukaryotic elongation factor 1 A-like 3) (eEF1A-like 3) (Eukaryotic translation elongation factor 1 alpha-1 pseudogene 5)	4	2	1	1
P68363	Tubulin alpha-1B chain (Alpha-tubulin ubiquitous) (Tubulin K-alpha-1) (Tubulin alpha-ubiquitous chain) [Cleaved into: Detyrosinated tubulin alpha-1B chain]	4	2	0	2
P68104	Elongation factor 1-alpha 1 (EF-1-alpha-1) (Elongation factor Tu) (EF-Tu) (Eukaryotic elongation factor 1 A-1) (eEF1A-1) (Leukocyte receptor cluster member 7)	4	2	1	1
P42167	Lamina-associated polypeptide 2, isoforms beta/gamma (Thymopoietin, isoforms beta/gamma) (TP beta/gamma) (Thymopoietin-related peptide isoforms beta/gamma) (TPRP isoforms beta/gamma) [Cleaved into: Thymopoietin (TP) (Splenin); Thymopentin (TP5)]	4	0	1	3

P26373	60S ribosomal protein L13 (Breast basic conserved protein 1) (Large ribosomal subunit protein eL13)	4	1	1	2
P16949	Stathmin (Leukemia-associated phosphoprotein p18) (Metablastin) (Oncoprotein 18) (Op18) (Phosphoprotein p19) (pp19) (Prosolin) (Protein Pr22) (pp17)	4	2	1	1
Q96A08	Histone H2B type 1-A (Histone H2B, testis) (TSH2B.1) (hTSH2B) (Testis-specific histone H2B)	3	1	1	1
Q7Z3K3	Pogo transposable element with ZNF domain (Suppressor of hairy wing homolog 5) (Zinc finger protein 280E) (Zinc finger protein 635)	3	1	1	1
Q71U36	Tubulin alpha-1A chain (Alpha-tubulin 3) (Tubulin B-alpha-1) (Tubulin alpha-3 chain) [Cleaved into: Detyrosinated tubulin alpha-1A chain]	3	2	0	1
Q71DI3	Histone H3.2 (Histone H3/m) (Histone H3/o)	3	1	1	1
Q6NXT2	Histone H3.3C (Histone H3.5)	3	1	1	1
Q16695	Histone H3.1t (H3/t) (H3t) (H3/g)	3	1	1	1
Q14980	Nuclear mitotic apparatus protein 1 (Nuclear matrix protein-22) (NMP-22) (Nuclear mitotic apparatus protein) (NuMA protein) (SP-H antigen)	3	1	0	2
Q13422	DNA-binding protein Ikaros (Ikaros family zinc finger protein 1) (Lymphoid transcription factor LyF-1)	3	1	1	1
Q13263	Transcription intermediary factor 1-beta (TIF1-beta) (E3 SUMO-protein ligase TRIM28) (EC 2.3.2.27) (KRAB-associated protein 1) (KAP-1) (KRAB-interacting protein 1) (KRIP-1) (Nuclear corepressor KAP-1) (RING finger protein 96) (RING-type E3 ubiquitin transferase TIF1-beta) (Tripartite motif-containing protein 28)	3	1	1	1
Q01105	Protein SET (HLA-DR-associated protein II) (Inhibitor of granzyme A-activated DNase) (IGAAD) (PHAPII) (Phosphatase 2A inhibitor I2PP2A) (I-2PP2A) (Template-activating factor I) (TAF-I)	3	1	1	1
P84243	Histone H3.3	3	1	1	1
P68431	Histone H3.1 (Histone H3/a) (Histone H3/b) (Histone H3/c) (Histone H3/d) (Histone H3/f) (Histone H3/h) (Histone H3/i) (Histone H3/j) (Histone H3/k) (Histone H3/l)	3	1	1	1
P62750	60S ribosomal protein L23a (Large ribosomal subunit protein uL23)	3	1	1	1
P61513	60S ribosomal protein L37a (Large ribosomal subunit protein eL43)	3	1	1	1
P45973	Chromobox protein homolog 5 (Antigen p25) (Heterochromatin protein 1 homolog alpha) (HP1 alpha)	3	2	0	1
P25705	ATP synthase subunit alpha, mitochondrial (ATP synthase F1 subunit alpha)	3	1	2	0
P22626	Heterogeneous nuclear ribonucleoproteins A2/B1 (hnRNP A2/B1)	3	2	0	1
P18621	60S ribosomal protein L17 (60S ribosomal protein L23) (Large ribosomal subunit protein uL22) (PD-1)	3	1	1	1
P0DME0	Protein SETSIP (SET pseudogene protein 18) (SET similar protein) (Similar to SET translocation protein)	3	1	1	1
P0C7P4	Putative cytochrome b-c1 complex subunit Rieske-like protein 1 (Ubiquinol-cytochrome c reductase Rieske iron-sulfur subunit pseudogene 1)	3	1	1	1
Q9P258	Protein RCC2 (RCC1-like protein TD-60) (Telophase disk protein of 60 kDa)	2	0	1	1
Q9NY65	Tubulin alpha-8 chain (Alpha-tubulin 8) (Tubulin alpha chain-like 2)	2	1	0	1
Q8ND82	Zinc finger protein 280C (Suppressor of hairy wing homolog 3) (Zinc finger protein 633)	2	1	1	0
Q8IX01	SURP and G-patch domain-containing protein 2 (Arginine/serine-rich-splicing factor 14) (Splicing factor, arginine/serine-rich 14)	2	1	0	1
Q5T3J3	Ligand-dependent nuclear receptor-interacting factor 1 (Receptor-interacting factor 1)	2	1	0	1
Q15020	Squamous cell carcinoma antigen recognized by T-cells 3 (SART-3) (Tat-interacting protein of 110 kDa) (Tip110) (p110 nuclear RNA-binding protein)	2	1	0	1
Q08211	ATP-dependent RNA helicase A (EC 3.6.4.13) (DEAH box protein 9) (DExH-box helicase 9) (Leukophysin) (LKP) (Nuclear DNA helicase II) (NDH II) (RNA helicase A)	2	1	1	0

Q05639	Elongation factor 1-alpha 2 (EF-1-alpha-2) (Eukaryotic elongation factor 1 A-2) (eEF1A-2) (Statin-S1)	2	1	1	0
P68371	Tubulin beta-4B chain (Tubulin beta-2 chain) (Tubulin beta-2C chain)	2	1	0	1
P68366	Tubulin alpha-4A chain (Alpha-tubulin 1) (Testis-specific alpha-tubulin) (Tubulin H2-alpha) (Tubulin alpha-1 chain)	2	1	0	1
P62753	40S ribosomal protein S6 (Phosphoprotein NP33) (Small ribosomal subunit protein eS6)	2	1	0	1
P53999	Activated RNA polymerase II transcriptional coactivator p15 (Positive cofactor 4) (PC4) (SUB1 homolog) (p14)	2	1	0	1
P46013	Proliferation marker protein Ki-67 (Antigen identified by monoclonal antibody Ki-67) (Antigen KI-67) (Antigen Ki67)	2	1	0	1
P35998	26S proteasome regulatory subunit 7 (26S proteasome AAA-ATPase subunit RPT1) (Proteasome 26S subunit ATPase 2) (Protein MSS1)	2	0	1	1
P17480	Nucleolar transcription factor 1 (Autoantigen NOR-90) (Upstream-binding factor 1) (UBF-1)	2	0	1	1
P07437	Tubulin beta chain (Tubulin beta-5 chain)	2	1	0	1
P05141	ADP/ATP translocase 2 (ADP,ATP carrier protein 2) (ADP,ATP carrier protein, fibroblast isoform) (Adenine nucleotide translocator 2) (ANT 2) (Solute carrier family 25 member 5) [Cleaved into: ADP/ATP translocase 2, N-terminally processed]	2	1	1	0
P04350	Tubulin beta-4A chain (Tubulin 5 beta) (Tubulin beta-4 chain)	2	1	0	1
A6NJT0	Homeobox protein unc-4 homolog (Homeobox protein Uncx4.1)	2	1	0	1
Q9UNX3	60S ribosomal protein L26-like 1 (Large ribosomal subunit protein uL24-like 1)	1	0	1	0
Q9NW82	WD repeat-containing protein 70	1	0	0	1
Q9HBM6	Transcription initiation factor TFIID subunit 9B (Neuronal cell death-related protein 7) (DN-7) (Transcription initiation factor TFIID subunit 9-like) (Transcription-associated factor TAFII31L)	1	0	1	0
Q9H853	Putative tubulin-like protein alpha-4B (Alpha-tubulin 4B)	1	0	0	1
Q9C0E4	Glutamate receptor-interacting protein 2 (GRIP-2)	1	0	1	0
Q9BVP2	Guanine nucleotide-binding protein-like 3 (E2-induced gene 3 protein) (Novel nucleolar protein 47) (NNP47) (Nucleolar GTP-binding protein 3) (Nucleostemin)	1	1	0	0
Q9BTM1	Histone H2A.J (H2a/j)	1	1	0	0
Q99878	Histone H2A type 1-J (Histone H2A/e)	1	1	0	0
Q96KK5	Histone H2A type 1-H (Histone H2A/s)	1	1	0	0
Q93077	Histone H2A type 1-C (Histone H2A/l)	1	1	0	0
Q93045	Stathmin-2 (Superior cervical ganglion-10 protein) (Protein SCG10)	1	1	0	0
Q92541	RNA polymerase-associated protein RTF1 homolog	1	0	0	1
Q8IY92	Structure-specific endonuclease subunit SLX4 (BTB/POZ domain-containing protein 12)	1	0	0	1
Q7L7L0	Histone H2A type 3	1	1	0	0
Q6PEY2	Tubulin alpha-3E chain (Alpha-tubulin 3E) [Cleaved into: Detyrosinated tubulin alpha-3E chain]	1	1	0	0
Q6FI13	Histone H2A type 2-A (Histone H2A.2) (Histone H2A/o)	1	1	0	0
Q5T8P6	RNA-binding protein 26 (CTCL tumor antigen se70-2) (RNA-binding motif protein 26)	1	0	0	1
Q5HYC2	Uncharacterized protein KIAA2026	1	0	0	1
Q32P51	Heterogeneous nuclear ribonucleoprotein A1-like 2 (hnRNP A1-like 2) (hnRNP core protein A1-like 2)	1	1	0	0
Q17RG1	BTB/POZ domain-containing protein KCTD19	1	0	1	0
Q16777	Histone H2A type 2-C (Histone H2A-GL101) (Histone H2A/q)	1	1	0	0

Q16594	Transcription initiation factor TFIID subunit 9 (RNA polymerase II TBP-associated factor subunit G) (STAF31/32) (Transcription initiation factor TFIID 31 kDa subunit) (TAFII-31) (TAFII31) (Transcription initiation factor TFIID 32 kDa subunit) (TAFII-32) (TAFII32)	1	0	1	0
Q13748	Tubulin alpha-3C/D chain (Alpha-tubulin 2) (Alpha-tubulin 3C/D) (Tubulin alpha-2 chain) [Cleaved into: Detyrosinated tubulin alpha-3C/D chain]	1	1	0	0
Q02878	60S ribosomal protein L6 (Large ribosomal subunit protein eL6) (Neoplasm-related protein C140) (Tax-responsive enhancer element-binding protein 107) (TaxREB107)	1	0	0	1
P62861	40S ribosomal protein S30 (Small ribosomal subunit protein eS30)	1	0	1	0
P61254	60S ribosomal protein L26 (Large ribosomal subunit protein uL24)	1	0	1	0
P49321	Nuclear autoantigenic sperm protein (NASP)	1	0	0	1
P47881	Olfactory receptor 3A1 (Olfactory receptor 17-40) (OR17-40) (Olfactory receptor OR17-15)	1	1	0	0
P36578	60S ribosomal protein L4 (60S ribosomal protein L1) (Large ribosomal subunit protein uL4)	1	0	1	0
P35268	60S ribosomal protein L22 (EBER-associated protein) (EAP) (Epstein-Barr virus small RNA-associated protein) (Heparin-binding protein HBp15) (Large ribosomal subunit protein eL22)	1	0	1	0
P28370	Probable global transcription activator SNF2L1 (EC 3.6.4.-) (ATP-dependent helicase SMARCA1) (Nucleosome-remodeling factor subunit SNF2L) (SWI/SNF-related matrix-associated actin-dependent regulator of chromatin subfamily A member 1)	1	0	1	0
P20671	Histone H2A type 1-D (Histone H2A.3) (Histone H2A/g)	1	1	0	0
P11678	Eosinophil peroxidase (EPO) (EC 1.11.1.7) [Cleaved into: Eosinophil peroxidase light chain; Eosinophil peroxidase heavy chain]	1	1	0	0
P0C0S8	Histone H2A type 1 (H2A.1) (Histone H2A/ptl)	1	1	0	0
P09651	Heterogeneous nuclear ribonucleoprotein A1 (hnRNP A1) (Helix-destabilizing protein) (Single-strand RNA-binding protein) (hnRNP core protein A1) [Cleaved into: Heterogeneous nuclear ribonucleoprotein A1, N-terminally processed]	1	1	0	0
P09104	Gamma-enolase (EC 4.2.1.11) (2-phospho-D-glycerate hydro-lyase) (Enolase 2) (Neural enolase) (Neuron-specific enolase) (NSE)	1	0	0	1
P04908	Histone H2A type 1-B/E (Histone H2A.2) (Histone H2A/a) (Histone H2A/m)	1	1	0	0
O94776	Metastasis-associated protein MTA2 (Metastasis-associated 1-like 1) (MTA1-L1 protein) (p53 target protein in deacetylase complex)	1	1	0	0
O75928	E3 SUMO-protein ligase PIAS2 (EC 2.3.2.-) (Androgen receptor-interacting protein 3) (ARIP3) (DAB2-interacting protein) (DIP) (E3 SUMO-protein transferase PIAS2) (Msx-interacting zinc finger protein) (Miz1) (PIAS-NY protein) (Protein inhibitor of activated STAT x) (Protein inhibitor of activated STAT2)	1	1	0	0
O75925	E3 SUMO-protein ligase PIAS1 (EC 2.3.2.-) (DEAD/H box-binding protein 1) (E3 SUMO-protein transferase PIAS1) (Gu-binding protein) (GBP) (Protein inhibitor of activated STAT protein 1) (RNA helicase II-binding protein)	1	1	0	0
O75923	Dysferlin (Dystrophy-associated fer-1-like protein) (Fer-1-like protein 1)	1	0	0	1
O75096	Low-density lipoprotein receptor-related protein 4 (LRP-4) (Multiple epidermal growth factor-like domains 7)	1	0	0	1
O60934	Nibrin (Cell cycle regulatory protein p95) (Nijmegen breakage syndrome protein 1)	1	0	0	1
O60437	Periplakin (190 kDa paraneoplastic pemphigus antigen) (195 kDa cornified envelope precursor protein)	1	0	0	1
O60264	SWI/SNF-related matrix-associated actin-dependent regulator of chromatin subfamily A member 5 (SWI/SNF-related matrix-associated actin-dependent regulator of chromatin A5) (EC 3.6.4.-) (Sucrose nonfermenting protein 2 homolog) (hSNF2H)	1	0	1	0
O43143	Pre-mRNA-splicing factor ATP-dependent RNA helicase DHX15 (EC 3.6.4.13) (ATP-dependent RNA helicase #46) (DEAH box protein 15)	1	1	0	0

9.4.2. NB4-BD ATRA List

Table 9-2 NB4-BD ATRA

Accession	Protein Names	# Unique Peptides			
		SUM	Sa.1	Sa.2	Sa.3
Q13085	Acetyl-CoA carboxylase 1 (ACC1) (EC 6.4.1.2) (ACC-alpha) [Includes: Biotin carboxylase (EC 6.3.4.14)]	77	25	31	21
A6NHR9	Structural maintenance of chromosomes flexible hinge domain-containing protein 1 (SMC hinge domain-containing protein 1)	41	15	15	11
Q9UER7	Death domain-associated protein 6 (Daxx) (hDaxx) (ETS1-associated protein 1) (EAP1) (Fas death domain-associated protein)	32	12	11	9
Q96RQ3	Methylcrotonoyl-CoA carboxylase subunit alpha, mitochondrial (MCCase subunit alpha) (EC 6.4.1.4) (3-methylcrotonyl-CoA carboxylase 1) (3-methylcrotonyl-CoA carboxylase biotin-containing subunit) (3-methylcrotonyl-CoA:carbon dioxide ligase subunit alpha)	31	11	11	9
P19338	Nucleolin (Protein C23)	26	9	11	6
P46100	Transcriptional regulator ATRX (EC 3.6.4.12) (ATP-dependent helicase ATRX) (X-linked helicase II) (X-linked nuclear protein) (XNP) (Znf-HX)	25	8	9	8
O00160	Unconventional myosin-If (Myosin-Ie)	21	8	9	4
P05165	Propionyl-CoA carboxylase alpha chain, mitochondrial (PCCase subunit alpha) (EC 6.4.1.3) (Propanoyl-CoA:carbon dioxide ligase subunit alpha)	20	6	7	7
P41218	Myeloid cell nuclear differentiation antigen	17	6	7	4
O00763	Acetyl-CoA carboxylase 2 (EC 6.4.1.2) (ACC-beta) [Includes: Biotin carboxylase (EC 6.3.4.14)]	16	5	5	6
P09874	Poly [ADP-ribose] polymerase 1 (PARP-1) (EC 2.4.2.30) (ADP-ribosyltransferase diphtheria toxin-like 1) (ARTD1) (NAD(+) ADP-ribosyltransferase 1) (ADPRT 1) (Poly[ADP-ribose] synthase 1)	13	5	5	3
P16401	Histone H1.5 (Histone H1a) (Histone H1b) (Histone H1s-3)	12	4	4	4
Q13428	Treacle protein (Treacher Collins syndrome protein)	12	5	4	3
P10412	Histone H1.4 (Histone H1b) (Histone H1s-4)	11	3	4	4
P16402	Histone H1.3 (Histone H1c) (Histone H1s-2)	11	3	4	4
P16403	Histone H1.2 (Histone H1c) (Histone H1d) (Histone H1s-1)	11	3	4	4
Q9NX58	Cell growth-regulating nucleolar protein	11	4	4	3
P05164	Myeloperoxidase (MPO) (EC 1.11.2.2) [Cleaved into: Myeloperoxidase; 89 kDa myeloperoxidase; 84 kDa myeloperoxidase; Myeloperoxidase light chain; Myeloperoxidase heavy chain]	10	2	5	3
Q9BVA6	Adenosine monophosphate-protein transferase FICD (EC 2.7.7.n1) (AMPylator FICD) (De-AMPylase FICD) (EC 3.1.4.-) (FIC domain-containing protein) (Huntingtin yeast partner E) (Huntingtin-interacting protein 13) (HIP-13) (Huntingtin-interacting protein E)	10	4	3	3
P62701	40S ribosomal protein S4, X isoform (SCR10) (Single copy abundant mRNA protein) (Small ribosomal subunit protein eS4)	9	4	3	2
P08567	Pleckstrin (Platelet 47 kDa protein) (p47)	8	2	4	2
Q13185	Chromobox protein homolog 3 (HECH) (Heterochromatin protein 1 homolog gamma) (HP1 gamma) (Modifier 2 protein)	8	2	3	3
P62750	60S ribosomal protein L23a (Large ribosomal subunit protein uL23)	7	3	3	1
P68104	Elongation factor 1-alpha 1 (EF-1-alpha-1) (Elongation factor Tu) (EF-Tu) (Eukaryotic elongation factor 1 A-1) (eEF1A-1) (Leukocyte receptor cluster member 7)	7	3	2	2

Q5VTE0	Putative elongation factor 1-alpha-like 3 (EF-1-alpha-like 3) (Eukaryotic elongation factor 1 A-like 3) (eEF1A-like 3) (Eukaryotic translation elongation factor 1 alpha-1 pseudogene 5)	7	3	2	2
Q6WKZ4	Rab11 family-interacting protein 1 (Rab11-FIP1) (Rab-coupling protein)	7	3	2	2
Q9H0A0	RNA cytidine acetyltransferase (EC 2.3.1.-) (18S rRNA cytosine acetyltransferase) (N-acetyltransferase 10)	7	2	3	2
P32121	Beta-arrestin-2 (Arrestin beta-2) (Non-visual arrestin-3)	6	3	2	1
P42766	60S ribosomal protein L35 (Large ribosomal subunit protein uL29)	6	2	2	2
P46776	60S ribosomal protein L27a (Large ribosomal subunit protein uL15)	6	2	1	3
P61247	40S ribosomal protein S3a (Small ribosomal subunit protein eS1) (v-fos transformation effector protein) (Fte-1)	6	2	3	1
P62805	Histone H4	6	2	2	2
P62829	60S ribosomal protein L23 (60S ribosomal protein L17) (Large ribosomal subunit protein uL14)	6	2	2	2
P63173	60S ribosomal protein L38 (Large ribosomal subunit protein eL38)	6	2	2	2
P68363	Tubulin alpha-1B chain (Alpha-tubulin ubiquitous) (Tubulin K-alpha-1) (Tubulin alpha-ubiquitous chain) [Cleaved into: Detyrosinated tubulin alpha-1B chain]	6	1	3	2
Q9BQE3	Tubulin alpha-1C chain (Alpha-tubulin 6) (Tubulin alpha-6 chain) [Cleaved into: Detyrosinated tubulin alpha-1C chain]	6	1	3	2
P18621	60S ribosomal protein L17 (60S ribosomal protein L23) (Large ribosomal subunit protein uL22) (PD-1)	5	2	2	1
P25705	ATP synthase subunit alpha, mitochondrial (ATP synthase F1 subunit alpha)	5	1	2	2
P26373	60S ribosomal protein L13 (Breast basic conserved protein 1) (Large ribosomal subunit protein eL13)	5	2	1	2
P42166	Lamina-associated polypeptide 2, isoform alpha (Thymopoietin isoform alpha) (TP alpha) (Thymopoietin-related peptide isoform alpha) (TPRP isoform alpha) [Cleaved into: Thymopoietin (TP) (Splenin); Thymopentin (TP5)]	5	2	1	2
P42167	Lamina-associated polypeptide 2, isoforms beta/gamma (Thymopoietin, isoforms beta/gamma) (TP beta/gamma) (Thymopoietin-related peptide isoforms beta/gamma) (TPRP isoforms beta/gamma) [Cleaved into: Thymopoietin (TP) (Splenin); Thymopentin (TP5)]	5	2	1	2
P62861	40S ribosomal protein S30 (Small ribosomal subunit protein eS30)	5	2	2	1
Q5SSJ5	Heterochromatin protein 1-binding protein 3 (Protein HP1-BP74)	5	1	1	3
P68366	Tubulin alpha-4A chain (Alpha-tubulin 1) (Testis-specific alpha-tubulin) (Tubulin H2-alpha) (Tubulin alpha-1 chain)	4	1	2	1
Q9NY65	Tubulin alpha-8 chain (Alpha-tubulin 8) (Tubulin alpha chain-like 2)	4	1	2	1
O60814	Histone H2B type 1-K (H2B K) (HIRA-interacting protein 1)	3	2	1	0
P04350	Tubulin beta-4A chain (Tubulin 5 beta) (Tubulin beta-4 chain)	3	1	1	1
P06733	Alpha-enolase (EC 4.2.1.11) (2-phospho-D-glycerate hydro-lyase) (C-myc promoter-binding protein) (Enolase 1) (MBP-1) (MPB-1) (Non-neural enolase) (NNE) (Phosphopyruvate hydratase) (Plasminogen-binding protein)	3	1	1	1
P06899	Histone H2B type 1-J (Histone H2B.1) (Histone H2B.r) (H2B/r)	3	2	1	0
P07437	Tubulin beta chain (Tubulin beta-5 chain)	3	1	1	1
P0DMV9	Heat shock 70 kDa protein 1B (Heat shock 70 kDa protein 2) (HSP70-2) (HSP70.2)	3	1	1	1
P11142	Heat shock cognate 71 kDa protein (Heat shock 70 kDa protein 8) (Lipopolysaccharide-associated protein 1) (LAP-1) (LPS-associated protein 1)	3	1	1	1
P17066	Heat shock 70 kDa protein 6 (Heat shock 70 kDa protein B')	3	1	1	1
P22626	Heterogeneous nuclear ribonucleoproteins A2/B1 (hnRNP A2/B1)	3	1	2	0
P23527	Histone H2B type 1-O (Histone H2B.2) (Histone H2B.n) (H2B/n)	3	2	1	0

P33778	Histone H2B type 1-B (Histone H2B.1) (Histone H2B.f) (H2B/f)	3	2	1	0
P34931	Heat shock 70 kDa protein 1-like (Heat shock 70 kDa protein 1L) (Heat shock 70 kDa protein 1-Hom) (HSP70-Hom)	3	1	1	1
P48741	Putative heat shock 70 kDa protein 7 (Heat shock 70 kDa protein B)	3	1	1	1
P51610	Host cell factor 1 (HCF) (HCF-1) (C1 factor) (CFF) (VCAF) (VP16 accessory protein) [Cleaved into: HCF N-terminal chain 1; HCF N-terminal chain 2; HCF N-terminal chain 3; HCF N-terminal chain 4; HCF N-terminal chain 5; HCF N-terminal chain 6; HCF C-terminal chain 1; HCF C-terminal chain 2; HCF C-terminal chain 3; HCF C-terminal chain 4; HCF C-terminal chain 5; HCF C-terminal chain 6]	3	1	1	1
P54652	Heat shock-related 70 kDa protein 2 (Heat shock 70 kDa protein 2)	3	1	1	1
P57053	Histone H2B type F-S (Histone H2B.s) (H2B/s)	3	2	1	0
P58876	Histone H2B type 1-D (HIRA-interacting protein 2) (Histone H2B.1 B) (Histone H2B.b) (H2B/b)	3	2	1	0
P61254	60S ribosomal protein L26 (Large ribosomal subunit protein uL24)	3	1	1	1
P61513	60S ribosomal protein L37a (Large ribosomal subunit protein eL43)	3	1	1	1
P62244	40S ribosomal protein S15a (Small ribosomal subunit protein uS8)	3	1	1	1
P62266	40S ribosomal protein S23 (Small ribosomal subunit protein uS12)	3	1	1	1
P62753	40S ribosomal protein S6 (Phosphoprotein NP33) (Small ribosomal subunit protein eS6)	3	1	1	1
P62807	Histone H2B type 1-C/E/F/G/I (Histone H2B.1 A) (Histone H2B.a) (H2B/a) (Histone H2B.g) (H2B/g) (Histone H2B.h) (H2B/h) (Histone H2B.k) (H2B/k) (Histone H2B.l) (H2B/l)	3	2	1	0
P68371	Tubulin beta-4B chain (Tubulin beta-2 chain) (Tubulin beta-2C chain)	3	1	1	1
Q05639	Elongation factor 1-alpha 2 (EF-1-alpha-2) (Eukaryotic elongation factor 1 A-2) (eEF1A-2) (Statin-S1)	3	1	1	1
Q16778	Histone H2B type 2-E (Histone H2B-GL105) (Histone H2B.q) (H2B/q)	3	2	1	0
Q5QNW6	Histone H2B type 2-F	3	2	1	0
Q5T3J3	Ligand-dependent nuclear receptor-interacting factor 1 (Receptor-interacting factor 1)	3	1	2	0
Q71U36	Tubulin alpha-1A chain (Alpha-tubulin 3) (Tubulin B-alpha-1) (Tubulin alpha-3 chain) [Cleaved into: Detyrosinated tubulin alpha-1A chain]	3	0	2	1
Q8N257	Histone H2B type 3-B (H2B type 12)	3	2	1	0
Q93079	Histone H2B type 1-H (Histone H2B.j) (H2B/j)	3	2	1	0
Q99877	Histone H2B type 1-N (Histone H2B.d) (H2B/d)	3	2	1	0
Q99879	Histone H2B type 1-M (Histone H2B.e) (H2B/e)	3	2	1	0
Q99880	Histone H2B type 1-L (Histone H2B.c) (H2B/c)	3	2	1	0
Q9H853	Putative tubulin-like protein alpha-4B (Alpha-tubulin 4B)	3	1	1	1
Q9UNX3	60S ribosomal protein L26-like 1 (Large ribosomal subunit protein uL24-like 1)	3	1	1	1
Q9Y2X3	Nucleolar protein 58 (Nucleolar protein 5)	3	0	2	1
O00567	Nucleolar protein 56 (Nucleolar protein 5A)	2	0	1	1
O43143	Pre-mRNA-splicing factor ATP-dependent RNA helicase DHX15 (EC 3.6.4.13) (ATP-dependent RNA helicase #46) (DEAH box protein 15)	2	1	1	0
O60934	Nibrin (Cell cycle regulatory protein p95) (Nijmegen breakage syndrome protein 1)	2	1	1	0
P04843	Dolichyl-diphosphooligosaccharide--protein glycosyltransferase subunit 1 (Dolichyl-diphosphooligosaccharide--protein glycosyltransferase 67 kDa subunit) (Ribophorin I) (RPN-I) (Ribophorin-1)	2	1	1	0
P04908	Histone H2A type 1-B/E (Histone H2A.2) (Histone H2A/a) (Histone H2A/m)	2	1	0	1
P06748	Nucleophosmin (NPM) (Nucleolar phosphoprotein B23) (Nucleolar protein NO38) (Numatrin)	2	0	1	1

P07900	Heat shock protein HSP 90-alpha (Heat shock 86 kDa) (HSP 86) (HSP86) (Lipopolysaccharide-associated protein 2) (LAP-2) (LPS-associated protein 2) (Renal carcinoma antigen NY-REN-38)	2	1	1	0
P08238	Heat shock protein HSP 90-beta (HSP 90) (Heat shock 84 kDa) (HSP 84) (HSP84)	2	1	1	0
P09651	Heterogeneous nuclear ribonucleoprotein A1 (hnRNP A1) (Helix-destabilizing protein) (Single-strand RNA-binding protein) (hnRNP core protein A1) [Cleaved into: Heterogeneous nuclear ribonucleoprotein A1, N-terminally processed]	2	0	0	2
P0C0S8	Histone H2A type 1 (H2A.1) (Histone H2A/ptl)	2	1	0	1
P0C7P4	Putative cytochrome b-c1 complex subunit Rieske-like protein 1 (Ubiquinol-cytochrome c reductase Rieske iron-sulfur subunit pseudogene 1)	2	1	1	0
P0DME0	Protein SETSIP (SET pseudogene protein 18) (SET similar protein) (Similar to SET translocation protein)	2	0	1	1
P20671	Histone H2A type 1-D (Histone H2A.3) (Histone H2A/g)	2	1	0	1
P29590	Protein PML (Promyelocytic leukemia protein) (RING finger protein 71) (Tripartite motif-containing protein 19)	2	1	0	1
P35659	Protein DEK	2	1	1	0
P35998	26S proteasome regulatory subunit 7 (26S proteasome AAA-ATPase subunit RPT1) (Proteasome 26S subunit ATPase 2) (Protein MSS1)	2	1	0	1
P45973	Chromobox protein homolog 5 (Antigen p25) (Heterochromatin protein 1 homolog alpha) (HP1 alpha)	2	0	2	0
P53999	Activated RNA polymerase II transcriptional coactivator p15 (Positive cofactor 4) (PC4) (SUB1 homolog) (p14)	2	1	0	1
P62277	40S ribosomal protein S13 (Small ribosomal subunit protein uS15)	2	1	0	1
P62280	40S ribosomal protein S11 (Small ribosomal subunit protein uS17)	2	0	1	1
P68431	Histone H3.1 (Histone H3/a) (Histone H3/b) (Histone H3/c) (Histone H3/d) (Histone H3/f) (Histone H3/h) (Histone H3/i) (Histone H3/j) (Histone H3/k) (Histone H3/l)	2	0	1	1
P83731	60S ribosomal protein L24 (60S ribosomal protein L30) (Large ribosomal subunit protein eL24)	2	0	1	1
P84243	Histone H3.3	2	0	1	1
Q01105	Protein SET (HLA-DR-associated protein II) (Inhibitor of granzyme A-activated DNase) (IGAAD) (PHAPII) (Phosphatase 2A inhibitor I2PP2A) (I-2PP2A) (Template-activating factor I) (TAF-I)	2	0	1	1
Q02880	DNA topoisomerase 2-beta (EC 5.99.1.3) (DNA topoisomerase II, beta isozyme)	2	1	1	0
Q12965	Unconventional myosin-Ie (Myosin-Ic) (Unconventional myosin 1E)	2	0	1	1
Q13422	DNA-binding protein Ikaros (Ikaros family zinc finger protein 1) (Lymphoid transcription factor LyF-1)	2	1	1	0
Q14568	Heat shock protein HSP 90-alpha A2 (Heat shock 90 kDa protein 1 alpha-like 3) (Heat shock protein HSP 90-alpha A2 pseudogene)	2	1	1	0
Q15233	Non-POU domain-containing octamer-binding protein (NonO protein) (54 kDa nuclear RNA- and DNA-binding protein) (55 kDa nuclear protein) (DNA-binding p52/p100 complex, 52 kDa subunit) (NMT55) (p54(nrb)) (p54nrb)	2	0	1	1
Q16594	Transcription initiation factor TFIID subunit 9 (RNA polymerase II TBP-associated factor subunit G) (STAF31/32) (Transcription initiation factor TFIID 31 kDa subunit) (TAFII-31) (TAFII31) (Transcription initiation factor TFIID 32 kDa subunit) (TAFII-32) (TAFII32)	2	1	0	1
Q16695	Histone H3.1t (H3/t) (H3t) (H3/g)	2	0	1	1
Q16777	Histone H2A type 2-C (Histone H2A-GL101) (Histone H2A/q)	2	1	0	1
Q32P51	Heterogeneous nuclear ribonucleoprotein A1-like 2 (hnRNP A1-like 2) (hnRNP core protein A1-like 2)	2	0	0	2

Q58FF8	Putative heat shock protein HSP 90-beta 2 (Heat shock protein 90-beta b) (Heat shock protein 90Bb)	2	1	1	0
Q6FI13	Histone H2A type 2-A (Histone H2A.2) (Histone H2A/o)	2	1	0	1
Q6NXT2	Histone H3.3C (Histone H3.5)	2	0	1	1
Q71DI3	Histone H3.2 (Histone H3/m) (Histone H3/o)	2	0	1	1
Q7L7L0	Histone H2A type 3	2	1	0	1
Q7L804	Rab11 family-interacting protein 2 (Rab11-FIP2) (NRip11)	2	1	0	1
Q93077	Histone H2A type 1-C (Histone H2A/l)	2	1	0	1
Q96KK5	Histone H2A type 1-H (Histone H2A/s)	2	1	0	1
Q99878	Histone H2A type 1-J (Histone H2A/e)	2	1	0	1
Q9BTM1	Histone H2A.J (H2a/j)	2	1	0	1
Q9HBM6	Transcription initiation factor TFIID subunit 9B (Neuronal cell death-related protein 7) (DN-7) (Transcription initiation factor TFIID subunit 9-like) (Transcription-associated factor TAFII31L)	2	1	0	1
Q9Y230	RuvB-like 2 (EC 3.6.4.12) (48 kDa TATA box-binding protein-interacting protein) (48 kDa TBP-interacting protein) (51 kDa erythrocyte cytosolic protein) (ECP-51) (INO80 complex subunit J) (Repressing pontin 52) (Reptin 52) (TIP49b) (TIP60-associated protein 54-beta) (TAP54-beta)	2	1	1	0
A6NI72	Putative neutrophil cytosol factor 1B (NCF-1B) (Putative SH3 and PX domain-containing protein 1B)	1	0	1	0
A8MVU1	Putative neutrophil cytosol factor 1C (NCF-1C) (Putative SH3 and PX domain-containing protein 1C)	1	0	1	0
B6SEH8	Endogenous retrovirus group V member 1 Env polyprotein (HERV-V_19q13.41 provirus ancestral Env polyprotein 1)	1	0	0	1
B6SEH9	Endogenous retrovirus group V member 2 Env polyprotein (HERV-V_19q13.41 provirus ancestral Env polyprotein 2)	1	0	0	1
O00231	26S proteasome non-ATPase regulatory subunit 11 (26S proteasome regulatory subunit RPN6) (26S proteasome regulatory subunit S9) (26S proteasome regulatory subunit p44.5)	1	0	1	0
O43739	Cytohesin-3 (ARF nucleotide-binding site opener 3) (Protein ARNO3) (General receptor of phosphoinositides 1) (Grp1) (PH, SEC7 and coiled-coil domain-containing protein 3)	1	0	0	1
O60437	Periplakin (190 kDa paraneoplastic pemphigus antigen) (195 kDa cornified envelope precursor protein)	1	1	0	0
P05141	ADP/ATP translocase 2 (ADP,ATP carrier protein 2) (ADP,ATP carrier protein, fibroblast isoform) (Adenine nucleotide translocator 2) (ANT 2) (Solute carrier family 25 member 5) [Cleaved into: ADP/ATP translocase 2, N-terminally processed]	1	1	0	0
P05161	Ubiquitin-like protein ISG15 (Interferon-induced 15 kDa protein) (Interferon-induced 17 kDa protein) (IP17) (Ubiquitin cross-reactive protein) (hUCRP)	1	0	1	0
P11678	Eosinophil peroxidase (EPO) (EC 1.11.1.7) [Cleaved into: Eosinophil peroxidase light chain; Eosinophil peroxidase heavy chain]	1	0	1	0
P14598	Neutrophil cytosol factor 1 (NCF-1) (47 kDa autosomal chronic granulomatous disease protein) (47 kDa neutrophil oxidase factor) (NCF-47K) (Neutrophil NADPH oxidase factor 1) (Nox organizer 2) (Nox-organizing protein 2) (SH3 and PX domain-containing protein 1A) (p47-phox)	1	0	1	0
P15153	Ras-related C3 botulinum toxin substrate 2 (GX) (Small G protein) (p21-Rac2)	1	0	1	0
P23497	Nuclear autoantigen Sp-100 (Nuclear dot-associated Sp100 protein) (Speckled 100 kDa)	1	1	0	0
P36578	60S ribosomal protein L4 (60S ribosomal protein L1) (Large ribosomal subunit protein uL4)	1	0	1	0
P46781	40S ribosomal protein S9 (Small ribosomal subunit protein uS4)	1	0	1	0

P51532	Transcription activator BRG1 (EC 3.6.4.-) (ATP-dependent helicase SMARCA4) (BRG1-associated factor 190A) (BAF190A) (Mitotic growth and transcription activator) (Protein BRG-1) (Protein brahma homolog 1) (SNF2-beta) (SWI/SNF-related matrix-associated actin-dependent regulator of chromatin subfamily A member 4)	1	0	1	0
P53677	AP-3 complex subunit mu-2 (Adaptor-related protein complex 3 subunit mu-2) (Clathrin assembly protein assembly protein complex 3 mu-2 medium chain) (Clathrin coat assembly protein AP47 homolog 2) (Clathrin coat-associated protein AP47 homolog 2) (Golgi adaptor AP-1 47 kDa protein homolog 2) (HA1 47 kDa subunit homolog 2) (Mu3B-adaptin) (P47B)	1	0	0	1
P60763	Ras-related C3 botulinum toxin substrate 3 (p21-Rac3)	1	0	1	0
P62241	40S ribosomal protein S8 (Small ribosomal subunit protein eS8)	1	1	0	0
P63000	Ras-related C3 botulinum toxin substrate 1 (Cell migration-inducing gene 5 protein) (Ras-like protein TC25) (p21-Rac1)	1	0	1	0
P84098	60S ribosomal protein L19 (Large ribosomal subunit protein eL19)	1	1	0	0
Q08211	ATP-dependent RNA helicase A (EC 3.6.4.13) (DEAH box protein 9) (DExH-box helicase 9) (Leukophysin) (LKP) (Nuclear DNA helicase II) (NDH II) (RNA helicase A)	1	1	0	0
Q12906	Interleukin enhancer-binding factor 3 (Double-stranded RNA-binding protein 76) (DRBP76) (M-phase phosphoprotein 4) (MPP4) (Nuclear factor associated with dsRNA) (NFAR) (Nuclear factor of activated T-cells 90 kDa) (NF-AT-90) (Translational control protein 80) (TCP80)	1	1	0	0
Q13263	Transcription intermediary factor 1-beta (TIF1-beta) (E3 SUMO-protein ligase TRIM28) (EC 2.3.2.27) (KRAB-associated protein 1) (KAP-1) (KRAB-interacting protein 1) (KRIP-1) (Nuclear corepressor KAP-1) (RING finger protein 96) (RING-type E3 ubiquitin transferase TIF1-beta) (Tripartite motif-containing protein 28)	1	0	1	0
Q13748	Tubulin alpha-3C/D chain (Alpha-tubulin 2) (Alpha-tubulin 3C/D) (Tubulin alpha-2 chain) [Cleaved into: Detyrosinated tubulin alpha-3C/D chain]	1	0	1	0
Q14147	Probable ATP-dependent RNA helicase DHX34 (EC 3.6.4.13) (DEAH box protein 34)	1	0	0	1
Q14739	Lamin-B receptor (Integral nuclear envelope inner membrane protein) (LMN2R)	1	0	1	0
Q14980	Nuclear mitotic apparatus protein 1 (Nuclear matrix protein-22) (NMP-22) (Nuclear mitotic apparatus protein) (NuMA protein) (SP-H antigen)	1	1	0	0
Q15020	Squamous cell carcinoma antigen recognized by T-cells 3 (SART-3) (Tat-interacting protein of 110 kDa) (Tip110) (p110 nuclear RNA-binding protein)	1	0	1	0
Q5TBB1	Ribonuclease H2 subunit B (RNase H2 subunit B) (Aicardi-Goutieres syndrome 2 protein) (AGS2) (Deleted in lymphocytic leukemia 8) (Ribonuclease HI subunit B)	1	0	1	0
Q6GTS8	N-fatty-acyl-amino acid synthase/hydrolase PM20D1 (EC 3.5.1.-) (EC 4.3.-.-) (Peptidase M20 domain-containing protein 1)	1	0	0	1
Q6PEY2	Tubulin alpha-3E chain (Alpha-tubulin 3E) [Cleaved into: Detyrosinated tubulin alpha-3E chain]	1	0	1	0
Q7L014	Probable ATP-dependent RNA helicase DDX46 (EC 3.6.4.13) (DEAD box protein 46) (PRP5 homolog)	1	0	1	0
Q7L576	Cytoplasmic FMR1-interacting protein 1 (Specifically Rac1-associated protein 1) (Sra-1) (p140sra-1)	1	0	1	0
Q7Z3K3	Pogo transposable element with ZNF domain (Suppressor of hairy wing homolog 5) (Zinc finger protein 280E) (Zinc finger protein 635)	1	1	0	0
Q8N201	Integrator complex subunit 1 (Int1)	1	0	1	0
Q8NE71	ATP-binding cassette sub-family F member 1 (ATP-binding cassette 50) (TNF-alpha-stimulated ABC protein)	1	0	0	1
Q8NFZ3	Neuroligin-4, Y-linked (Neuroligin Y)	1	0	0	1

Q8WZ42	Titin (EC 2.7.11.1) (Connectin) (Rhabdomyosarcoma antigen MU-RMS-40.14)	1	0	0	1
Q92522	Histone H1x	1	0	0	1
Q96A08	Histone H2B type 1-A (Histone H2B, testis) (TSH2B.1) (hTSH2B) (Testis-specific histone H2B)	1	1	0	0
Q96F07	Cytoplasmic FMR1-interacting protein 2 (p53-inducible protein 121)	1	0	1	0
Q96PL5	Erythroid membrane-associated protein (hERMAP) (Radin blood group antigen) (Scianna blood group antigen)	1	0	0	1
Q99442	Translocation protein SEC62 (Translocation protein 1) (TP-1) (hTP-1)	1	0	1	0
Q9C0E4	Glutamate receptor-interacting protein 2 (GRIP-2)	1	0	0	1
Q9H2P0	Activity-dependent neuroprotector homeobox protein (Activity-dependent neuroprotective protein)	1	1	0	0
Q9H930	Nuclear body protein SP140-like protein	1	1	0	0
Q9UIF8	Bromodomain adjacent to zinc finger domain protein 2B (hWALp4)	1	0	0	1
Q9Y2T2	AP-3 complex subunit mu-1 (AP-3 adaptor complex mu3A subunit) (Adaptor-related protein complex 3 subunit mu-1) (Mu-adaptin 3A) (Mu3A-adaptin)	1	0	0	1
Q9Y383	Putative RNA-binding protein Luc7-like 2	1	0	1	0
Q9Y5Z7	Host cell factor 2 (HCF-2) (C2 factor)	1	0	0	1

9.4.3. NB4 DMSO List

Table 9-3 NB4 DMSO

Accession	Protein Names	# Unique Peptides			
		SUM	Sa.1	Sa.2	Sa.3
Q13085	Heat shock protein HSP 90-alpha A2 (Heat shock 90 kDa protein 1 alpha-like 3) (Heat shock protein HSP 90-alpha A2 pseudogene)	132	42	45	45
P05165	Histone H2B type 1-J (Histone H2B.1) (Histone H2B.r) (H2B/r)	33	11	11	11
P11498	Histone H1.2 (Histone H1c) (Histone H1d) (Histone H1s-1)	30	11	9	10
Q96RQ3	Tubulin alpha-1C chain (Alpha-tubulin 6) (Tubulin alpha-6 chain) [Cleaved into: Detyrosinated tubulin alpha-1C chain]	29	10	9	10
Q9BVA6	Cell growth-regulating nucleolar protein	10	3	3	4
P83731	Heat shock protein 75 kDa, mitochondrial (HSP 75) (TNFR-associated protein 1) (Tumor necrosis factor type 1 receptor-associated protein) (TRAP-1)	9	2	4	3
P68363	60S ribosomal protein L24 (60S ribosomal protein L30) (Large ribosomal subunit protein eL24)	8	3	2	3
Q9BQE3	Putative uncharacterized protein encoded by AGPAT4-IT1 (AGPAT4 intronic transcript 1)	8	3	2	3
P46776	Heat shock-related 70 kDa protein 2 (Heat shock 70 kDa protein 2)	7	2	3	2
P16401	Heat shock 70 kDa protein 6 (Heat shock 70 kDa protein B')	6	2	2	2
Q71U36	Histone H2B type 1-H (Histone H2B.j) (H2B/j)	6	2	2	2
O60814	Histone H2B type 1-K (H2B K) (HIRA-interacting protein 1)	5	2	2	1
P16402	60S ribosomal protein L17 (60S ribosomal protein L23) (Large ribosomal subunit protein uL22) (PD-1)	5	1	2	2
P16403	Histone H2B type 1-O (Histone H2B.2) (Histone H2B.n) (H2B/n)	5	1	2	2
P23527	Histone H2B type 1-B (Histone H2B.1) (Histone H2B.f) (H2B/f)	5	2	2	1
P26373	60S ribosomal protein L35 (Large ribosomal subunit protein uL29)	5	2	1	2
P33778	60S ribosomal protein L27a (Large ribosomal subunit protein uL15)	5	2	2	1
P42766	Voltage-dependent calcium channel subunit alpha-2/delta-1 (Voltage-gated calcium channel subunit alpha-2/delta-1) [Cleaved into: Voltage-dependent calcium channel subunit alpha-2-1; Voltage-dependent calcium channel subunit delta-1]	5	2	2	1
P57053	60S ribosomal protein L37a (Large ribosomal subunit protein eL43)	5	2	2	1
P58876	40S ribosomal protein S23 (Small ribosomal subunit protein uS12)	5	2	2	1
P62807	Elongation factor 1-alpha 1 (EF-1-alpha-1) (Elongation factor Tu) (EF-Tu) (Eukaryotic elongation factor 1 A-1) (eEF1A-1) (Leukocyte receptor cluster member 7)	5	2	2	1
P68104	Histone H3.1 (Histone H3/a) (Histone H3/b) (Histone H3/c) (Histone H3/d) (Histone H3/f) (Histone H3/h) (Histone H3/i) (Histone H3/j) (Histone H3/k) (Histone H3/l)	5	1	2	2
P68366	60S ribosomal protein L19 (Large ribosomal subunit protein eL19)	5	2	1	2
P84098	Acetyl-CoA carboxylase 1 (ACC1) (EC 6.4.1.2) (ACC-alpha) [Includes: Biotin carboxylase (EC 6.3.4.14)]	5	1	2	2
Q16778	Histone H2B type 2-F	5	2	2	1
Q5QNW6	Ovostatin homolog 1	5	2	2	1
Q5VTE0	Tubulin alpha-3E chain (Alpha-tubulin 3E) [Cleaved into: Detyrosinated tubulin alpha-3E chain]	5	1	2	2

Q8N257	Methylcrotonoyl-CoA carboxylase subunit alpha, mitochondrial (MCCase subunit alpha) (EC 6.4.1.4) (3-methylcrotonyl-CoA carboxylase 1) (3-methylcrotonyl-CoA carboxylase biotin-containing subunit) (3-methylcrotonyl-CoA:carbon dioxide ligase subunit alpha)	5	2	2	1
Q93079	Histone H2B type 1-M (Histone H2B.e) (H2B/e)	5	2	2	1
Q99877	Tubulin beta-2B chain	5	2	2	1
Q99879	Adenosine monophosphate-protein transferase FICD (EC 2.7.7.n1) (AMPylator FICD) (De-AMPylase FICD) (EC 3.1.4.-) (FIC domain-containing protein) (Huntingtin yeast partner E) (Huntingtin-interacting protein 13) (HIP-13) (Huntingtin-interacting protein E)	5	2	2	1
Q99880	Nuclear pore complex protein Nup85 (85 kDa nucleoporin) (FROUNT) (Nucleoporin Nup75) (Nucleoporin Nup85) (Pericentrin-1)	5	2	2	1
P10412	Histone H1.5 (Histone H1a) (Histone H1b) (Histone H1s-3)	4	0	2	2
P06899	Heat shock protein HSP 90-alpha (Heat shock 86 kDa) (HSP 86) (HSP86) (Lipopolysaccharide-associated protein 2) (LAP-2) (LPS-associated protein 2) (Renal carcinoma antigen NY-REN-38)	3	0	2	1
P07437	Heat shock protein HSP 90-beta (HSP 90) (Heat shock 84 kDa) (HSP 84) (HSP84)	3	1	2	0
P25705	Heat shock 70 kDa protein 1-like (Heat shock 70 kDa protein 1L) (Heat shock 70 kDa protein 1-Hom) (HSP70-Hom)	3	0	1	2
P68431	Elongation factor 1-alpha 2 (EF-1-alpha-2) (Eukaryotic elongation factor 1 A-2) (eEF1A-2) (Statin-S1)	3	1	1	1
P84243	Tubulin beta-3 chain (Tubulin beta-4 chain) (Tubulin beta-III)	3	1	1	1
Q13748	Histone H2B type 2-E (Histone H2B-GL105) (Histone H2B.q) (H2B/q)	3	1	1	1
Q16695	Uncharacterized protein KIAA2026	3	1	1	1
Q6NXT2	Mucin-19 (MUC-19)	3	1	1	1
Q6PEY2	Histone H2B type 3-B (H2B type 12)	3	1	1	1
Q71DI3	Coiled-coil domain-containing protein 168	3	1	1	1

9.4.4. NB4 ATRA List

Table 9-4 NB4 ATRA

Accession	Protein Names	# Unique Peptides			
		SUM	Sa.1	Sa.2	Sa.3
Q13085	Acetyl-CoA carboxylase 1 (ACC1) (EC 6.4.1.2) (ACC-alpha) [Includes: Biotin carboxylase (EC 6.3.4.14)]	91	32	26	33
P05165	Propionyl-CoA carboxylase alpha chain, mitochondrial (PCCase subunit alpha) (EC 6.4.1.3) (Propanoyl-CoA:carbon dioxide ligase subunit alpha)	32	9	11	12
Q96RQ3	Methylcrotonoyl-CoA carboxylase subunit alpha, mitochondrial (MCCase subunit alpha) (EC 6.4.1.4) (3-methylcrotonyl-CoA carboxylase 1) (3-methylcrotonyl-CoA carboxylase biotin-containing subunit) (3-methylcrotonyl-CoA:carbon dioxide ligase subunit alpha)	32	11	9	12
O00160	Unconventional myosin-If (Myosin-Ie)	19	7	5	7
O00763	Acetyl-CoA carboxylase 2 (EC 6.4.1.2) (ACC-beta) [Includes: Biotin carboxylase (EC 6.3.4.14)]	18	6	6	6
Q9BVA6	Adenosine monophosphate-protein transferase FICD (EC 2.7.7.n1) (AMPylator FICD) (De-AMPylase FICD) (EC 3.1.4.-) (FIC domain-containing protein) (Huntingtin yeast partner E) (Huntingtin-interacting protein 13) (HIP-13) (Huntingtin-interacting protein E)	13	4	3	6
P16401	Histone H1.5 (Histone H1a) (Histone H1b) (Histone H1s-3)	12	5	3	4
Q9NX58	Cell growth-regulating nucleolar protein	12	4	4	4
P10412	Histone H1.4 (Histone H1b) (Histone H1s-4)	11	4	3	4
P16402	Histone H1.3 (Histone H1c) (Histone H1s-2)	11	4	3	4
P16403	Histone H1.2 (Histone H1c) (Histone H1d) (Histone H1s-1)	11	4	3	4
P41218	Myeloid cell nuclear differentiation antigen	9	2	3	4
P46776	60S ribosomal protein L27a (Large ribosomal subunit protein uL15)	9	3	3	3
P62701	40S ribosomal protein S4, X isoform (SCR10) (Single copy abundant mRNA protein) (Small ribosomal subunit protein eS4)	8	3	2	3
O60814	Histone H2B type 1-K (H2B K) (HIRA-interacting protein 1)	7	3	2	2
P18621	60S ribosomal protein L17 (60S ribosomal protein L23) (Large ribosomal subunit protein uL22) (PD-1)	7	3	2	2
P57053	Histone H2B type F-S (Histone H2B.s) (H2B/s)	7	3	2	2
P58876	Histone H2B type 1-D (HIRA-interacting protein 2) (Histone H2B.1 B) (Histone H2B.b) (H2B/b)	7	3	2	2
P62807	Histone H2B type 1-C/E/F/G/I (Histone H2B.1 A) (Histone H2B.a) (H2B/a) (Histone H2B.g) (H2B/g) (Histone H2B.h) (H2B/h) (Histone H2B.k) (H2B/k) (Histone H2B.l) (H2B/l)	7	3	2	2
Q5QNW6	Histone H2B type 2-F	7	3	2	2
Q93079	Histone H2B type 1-H (Histone H2B.j) (H2B/j)	7	3	2	2
Q99877	Histone H2B type 1-N (Histone H2B.d) (H2B/d)	7	3	2	2
Q99879	Histone H2B type 1-M (Histone H2B.e) (H2B/e)	7	3	2	2
Q99880	Histone H2B type 1-L (Histone H2B.c) (H2B/c)	7	3	2	2
P06899	Histone H2B type 1-J (Histone H2B.1) (Histone H2B.r) (H2B/r)	6	2	2	2
P23527	Histone H2B type 1-O (Histone H2B.2) (Histone H2B.n) (H2B/n)	6	2	2	2
P33778	Histone H2B type 1-B (Histone H2B.1) (Histone H2B.f) (H2B/f)	6	2	2	2
P68104	Elongation factor 1-alpha 1 (EF-1-alpha-1) (Elongation factor Tu) (EF-Tu) (Eukaryotic elongation factor 1 A-1) (eEF1A-1) (Leukocyte receptor cluster member 7)	6	2	2	2
P83731	60S ribosomal protein L24 (60S ribosomal protein L30) (Large ribosomal subunit protein eL24)	6	1	2	3
Q16778	Histone H2B type 2-E (Histone H2B-GL105) (Histone H2B.q) (H2B/q)	6	2	2	2

Q5SSJ5	Heterochromatin protein 1-binding protein 3 (Protein HP1-BP74)	6	3	1	2
Q5VTE0	Putative elongation factor 1-alpha-like 3 (EF-1-alpha-like 3) (Eukaryotic elongation factor 1 A-like 3) (eEF1A-like 3) (Eukaryotic translation elongation factor 1 alpha-1 pseudogene 5)	6	2	2	2
Q8N257	Histone H2B type 3-B (H2B type 12)	6	2	2	2
P05164	Myeloperoxidase (MPO) (EC 1.11.2.2) [Cleaved into: Myeloperoxidase; 89 kDa myeloperoxidase; 84 kDa myeloperoxidase; Myeloperoxidase light chain; Myeloperoxidase heavy chain]	5	2	2	1
P11142	Heat shock cognate 71 kDa protein (Heat shock 70 kDa protein 8) (Lipopolysaccharide-associated protein 1) (LAP-1) (LPS-associated protein 1)	5	2	1	2
P26373	60S ribosomal protein L13 (Breast basic conserved protein 1) (Large ribosomal subunit protein eL13)	5	2	0	3
P42766	60S ribosomal protein L35 (Large ribosomal subunit protein uL29)	5	1	2	2
P54652	Heat shock-related 70 kDa protein 2 (Heat shock 70 kDa protein 2)	5	2	1	2
P61247	40S ribosomal protein S3a (Small ribosomal subunit protein eS1) (v-fos transformation effector protein) (Fte-1)	5	1	1	3
P62829	60S ribosomal protein L23 (60S ribosomal protein L17) (Large ribosomal subunit protein uL14)	5	2	1	2
P63173	60S ribosomal protein L38 (Large ribosomal subunit protein eL38)	5	2	1	2
P06576	ATP synthase subunit beta, mitochondrial (EC 3.6.3.14) (ATP synthase F1 subunit beta)	4	2	1	1
P11498	Pyruvate carboxylase, mitochondrial (EC 6.4.1.1) (Pyruvic carboxylase) (PCB)	4	2	0	2
P25705	ATP synthase subunit alpha, mitochondrial (ATP synthase F1 subunit alpha)	4	0	2	2
P62750	60S ribosomal protein L23a (Large ribosomal subunit protein uL23)	4	1	1	2
P68363	Tubulin alpha-1B chain (Alpha-tubulin ubiquitous) (Tubulin K-alpha-1) (Tubulin alpha-ubiquitous chain) [Cleaved into: Detyrosinated tubulin alpha-1B chain]	4	2	0	2
Q9BQE3	Tubulin alpha-1C chain (Alpha-tubulin 6) (Tubulin alpha-6 chain) [Cleaved into: Detyrosinated tubulin alpha-1C chain]	4	2	0	2
O15042	U2 snRNP-associated SURP motif-containing protein (140 kDa Ser/Arg-rich domain protein) (U2-associated protein SR140)	3	1	1	1
P06733	Alpha-enolase (EC 4.2.1.11) (2-phospho-D-glycerate hydro-lyase) (C-myc promoter-binding protein) (Enolase 1) (MBP-1) (MPB-1) (Non-neural enolase) (NNE) (Phosphopyruvate hydratase) (Plasminogen-binding protein)	3	2	0	1
P09874	Poly [ADP-ribose] polymerase 1 (PARP-1) (EC 2.4.2.30) (ADP-ribosyltransferase diphtheria toxin-like 1) (ARTD1) (NAD(+) ADP-ribosyltransferase 1) (ADPRT 1) (Poly[ADP-ribose] synthase 1)	3	1	2	0
P0DMV9	Heat shock 70 kDa protein 1B (Heat shock 70 kDa protein 2) (HSP70-2) (HSP70.2)	3	1	1	1
P17066	Heat shock 70 kDa protein 6 (Heat shock 70 kDa protein B')	3	1	1	1
P18124	60S ribosomal protein L7 (Large ribosomal subunit protein uL30)	3	1	1	1
P34931	Heat shock 70 kDa protein 1-like (Heat shock 70 kDa protein 1L) (Heat shock 70 kDa protein 1-Hom) (HSP70-Hom)	3	1	1	1
P36578	60S ribosomal protein L4 (60S ribosomal protein L1) (Large ribosomal subunit protein uL4)	3	1	1	1
P48741	Putative heat shock 70 kDa protein 7 (Heat shock 70 kDa protein B)	3	1	1	1
P61513	60S ribosomal protein L37a (Large ribosomal subunit protein eL43)	3	1	1	1
P62244	40S ribosomal protein S15a (Small ribosomal subunit protein uS8)	3	1	1	1
P62266	40S ribosomal protein S23 (Small ribosomal subunit protein uS12)	3	1	1	1
P62280	40S ribosomal protein S11 (Small ribosomal subunit protein uS17)	3	1	1	1
P62854	40S ribosomal protein S26 (Small ribosomal subunit protein eS26)	3	1	1	1
P62861	40S ribosomal protein S30 (Small ribosomal subunit protein eS30)	3	1	0	2
P62910	60S ribosomal protein L32 (Large ribosomal subunit protein eL32)	3	1	1	1
P68366	Tubulin alpha-4A chain (Alpha-tubulin 1) (Testis-specific alpha-tubulin) (Tubulin H2-alpha) (Tubulin alpha-1 chain)	3	2	0	1

Q5JNZ5	Putative 40S ribosomal protein S26-like 1	3	1	1	1
Q71U36	Tubulin alpha-1A chain (Alpha-tubulin 3) (Tubulin B-alpha-1) (Tubulin alpha-3 chain) [Cleaved into: Detyrosinated tubulin alpha-1A chain]	3	1	0	2
Q96A08	Histone H2B type 1-A (Histone H2B, testis) (TSH2B.1) (hTSH2B) (Testis-specific histone H2B)	3	1	1	1

10. Acknowledgment

First of all, I would like to sincerely thank Dr. Thomas Sternsdorf for supervising my work in his group at the Research Institute Children's Cancer Center Hamburg at Universitätsklinikum Hamburg-Eppendorf (UKE). He cared about me and taught me everything about how to be a good scientist.

I also would like to thank Prof. Thomas Dobner who served as my co-supervisor and provided suggestions and opportunities for my project.

For the Mass Spec analysis, many thanks to my collaborators Dr. Yudong Guan, Min Zhang, and Prof. Hartmut Schlüter who worked in the department of clinical chemistry, for their generosity and hard work.

Also, I am grateful to Dr. Marios Xydous who not only taught me many scientific techniques, especially q-PCR, but also helped me a lot with my thesis.

For SMARCA4 cDNA, antibodies, and technical tips, I would like to thank Prof. Ulrich Schüller, Dr. Dörthe Holdhof and Dr. Franziska Modemann, who helped me a lot in a completely new field.

I also would like to thank Dr. TuLan-VuHan, Dr. Zoya Eskandarian, Dr. Marcos Seoane Souto, Dr. Pablo Iglesias Vázquez, and Diego González López who always supported me, cared about me, and felt like families.

Many thanks to Prof. Martin Horstmann, Dr. Peter Nollau, who improved the research a lot. Thanks to Julia Strauss and Marianne Klokow, I have learned so many techniques; and for help with the mouse work, I would like to thank Dr. Antonina Wrzeszcz.

I would like to thank all of our members in the institute for the nice and friendly atmosphere in the lab.

Financial support by the China Scholarship Council is gratefully acknowledged (CSC No. 201708080210).

Finally, I would like to thank my beloved family, my lovely mom, my great dad, my husband, my aunt, my uncle-in-law, and my landlords here in Hamburg for their support and endless love in my life.

11. Publications

11.1. Poster Presentations

1. **Cui W,** Guan Y, Holdhof D, Modemann F, Schlüter H, Sternsdorf T.

The SWI/SNF ATPase SMARCA4/BRG1 interacts with the Histone H3.3 Chaperone Daxx, physically interconnecting two seemingly distinct epigenetic Mechanisms.

Keystone Symposia on Molecular and Cellular Biology-Epigenetics and Human Disease, March 2019, Banff, Alberta, Canada.

2.**Cui W,** Guan Y, Holdhof D, Modemann F, Schlüter H, Sternsdorf T.

Identification of novel interactors of the H3.3 histone chaperone subunit Daxx using proximity-mediated biotin identification (BioID), reveals unexpected crosstalk between seemingly different epigenetic processes.

1st EMBO Workshop-Histone chaperones: Structure, function and role in development and disease, October 2019, Crete, Greece.

11.2. Conference Papers

1.**Cui W,** Guan Y, Xydous M, Schlüter H, Sternsdorf T. (June 2019)

Using Acute Promyelocytic Leukemia to study the H3.3 Histone chaperone system and its role in pediatric malignancies.

Klinische Pädiatrie 231(03): 163-163.

11.3. Journal Articles in Preparation

1. **Cui W**, Xydous M, Haschke A, Korf K, Pohlmann D, Spohn M, Sweet S, Kruchen A, Ip H-W, Wrzeszcz A, Müller J, Sternsdorf T.

The Polycomb Protein BMI-1 is a Key Effector of the H3.3 K27M Oncohistone in a Murine Bone Marrow Model.

2.**Cui W**, Guan Y, Holdhof D, Modemann F, Schlüter H, Sternsdorf T.

Definition of the PML-associated Daxx/ATRX using Proximity-mediated Biotin Identification, Reveals unexpected Crosstalk between seemingly Different Epigenetic Processes.

12. Curriculum Vitae

PERSONAL INFORMATION

Name: Wencong Cui

Gender: Female

Date of Birth: Aug 7, 1990

E-Mail: wencongcui@gmai.com

Phone： 017656977669

Address： Farnhornweg 35, 22547, Hamburg, Germany

EDUCATION

Bachelor of Science

Duration: Sep, 2009 - Jul, 2013

Specialty: Biological Science

University: Ocean University of China

Master of Medicine

Duration: Sep, 2013 - Jun, 2016

Specialty: Pharmaceutical Chemistry

University: Ocean University of China

Supervisor: Prof. Xia Zhao

Master Thesis: Extraction, isolation, structural characterization and bioactivity evaluation of polysaccharides from Asterias amurensis viscera

PhD Study

Duration: Mar, 2017 - Aug, 2020

Research Topics: Cancer Research; Epigenetics of Carcinogenesis; Biology of Oncohistone; Composition and Function of the Histone Chaperone complex

University: Universität Hamburg-Universitätsklinikum Hamburg-Eppendorf (UKE)

Institution: Children's Cancer Research Institute Hamburg

Supervisor: Dr. Thomas Sternsdorf

PhD Thesis: Using Acute Promyelocytic Leukemia (APL) to study the PML-associated Daxx/ATRX Complex and its role in pediatric malignancies.

EXPERIMENTAL SKILLS

PhD: Cell culture; Cell fractionation; Agarose Gel Electrophoresis; Polymerase Chain Reaction (PCR); Western blotting (WB); Immunofluorescence (IF) microscope; Quantitative Polymerase Chain Reaction (qPCR); Isolate and analyze DNA, RNA and protein; FACS; BioID; Tissue sectioning; Mass-spec (LC/MS) data analysis.

Master: High-performance liquid chromatography (HPLC); Infrared spectroscopy (IR); Gas chromatography-mass spectrometry (GC/MS); Nuclear magnetic resonance spectral analysis (NMR); Multi-stage mass spectra (MSn) samples pre-treatment and spectrum analysis; N/O-glycans of biochemical degradation and purification; Liquid chromatography-mass spectrometry (LC/MS); and Data processing.

AWARDS

2009-2010, Second-class scholarship, science and technology scholarships, outstanding members of Communist Youth League, outstanding student leader.

2010-2011, Third-class scholarship.

2013-2014, Exceptional scholarship.

2017-2020, China Scholarship Council Scholarship.

13. Statutory Declaration

Eidesstattliche Versicherung

Hiermit erkläre ich an Eides statt, dass ich die vorliegende Dissertationsschrift selbst verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe.

Unterschrift: .....

14. Language Proof

Yolanda M. Collier, San Diego, USA

Letter of Confirmation

To whom it may concern,

My name is Yolanda Collier, I am US citizen and a native English speaker

Hereby I certify that the Doctoral thesis of Ms. Wencong Cui is written in correct English.

Sincerely,

(Signature, Date and place)

15. References

1. Roux, K. J., Kim, D. I., Burke, B. & May, D. G. BioID: A Screen for Protein-Protein Interactions. *Curr Protoc Protein Sci* **91**, 19.23.1-19.23.15 (2018).

2. Hanahan, D. & Weinberg, R. A. The Hallmarks of Cancer. *Cell* **100**, 57–70 (2000).

3. Han, Z.-J., Feng, Y.-H., Gu, B.-H., Li, Y.-M. & Chen, H. The post-translational modification, SUMOylation, and cancer (Review). *Int J Oncol* **52**, 1081–1094 (2018).

4. McClatchey, A. I. & Yap, A. S. Contact inhibition (of proliferation) redux. *Curr Opin Cell Biol* **24**, 685–94 (2012).

5. Elmore, S. Apoptosis: A Review of Programmed Cell Death. *Toxicol Pathol* **35**, 495–516 (2007).

6. Greenberg, R. Telomeres, Crisis and Cancer. *Curr Mol Med* **5**, 213–218 (2005).

7. Cesare, A. J. & Reddel, R. R. Alternative lengthening of telomeres: models, mechanisms and implications. *Nat Rev Genet* **11**, 319–330 (2010).

8. Bergers, G. & Benjamin, L. E. Tumorigenesis and the angiogenic switch. *Nat Rev Cancer* **3**, 401–410 (2003).

9. Zijl, F. van, Krupitza, G. & Mikulits, W. Initial steps of metastasis: cell invasion and endothelial transmigration. *Mutat Res* **728**, 23–34 (2011).

10. Hanahan, D. & Weinberg, R. A. Hallmarks of Cancer: The Next Generation. *Cell* **144**, 646–674 (2011).

11. Waddingdon, C. H. Reviews: The Strategy of the Genes: A Discussion of Some Aspects of Theoretical Biology. *Med J Australia* **1**, 493–493 (1958).

12. Anderson, G., Verd, B. & Jaeger, J. Drawing to Extend Waddington's Epigenetic Landscape. *Leonardo* 1–14 (2019) doi:10.1162/leon_a_01738.

13. Holliday, R. & Grigg, G. W. DNA methylation and mutation. *Mutat Res Fundam Mol Mech Mutagen* **285**, 61–67 (1993).

14. HOLLIDAY, R. MECHANISMS FOR THE CONTROL OF GENE ACTIVITY DURING DEVELOPMENT. *Biol Rev* **65**, 431–471 (1990).

15. Herman, J. G. & Baylin, S. B. Gene Silencing in Cancer in Association with Promoter Hypermethylation. *New Engl J Med* **349**, 2042–2054 (2003).

16. Gama-Sosa, M. A. *et al.* The 5-methylcytosine content of DNA from human tumors. *Nucleic Acids Res* **11**, 6883–6894 (1983).

17. Feinberg, A. P. & Vogelstein, B. Hypomethylation distinguishes genes of some human cancers from their normal counterparts. *Nature* **301**, 89–92 (1983).

18. Feinberg, A. P. & Vogelstein, B. Hypomethylation of ras oncogenes in primary human cancers. *Biochem Bioph Res Co* **111**, 47–54 (1983).

19. Liang, G. & Weisenberger, D. J. DNA methylation aberrancies as a guide for surveillance and treatment of human cancers. *Epigenetics* **12**, 0 (2017).

20. Allfrey, V. G., Faulkner, R. & Mirsky, A. E. ACETYLATION AND METHYLATION OF HISTONES AND THEIR POSSIBLE ROLE IN THE REGULATION OF RNA SYNTHESIS. *Proc National Acad Sci* **51**, 786–794 (1964).

21. Strahl, B. D. & Allis, C. D. The language of covalent histone modifications. *Nature* **403**, 41–45 (2000).

22. Jang, C.-W., Shibata, Y., Starmer, J., Yee, D. & Magnuson, T. Histone H3.3 maintains genome integrity during mammalian development. *Gene Dev* **29**, 1377–92 (2015).

23. Bannister, A. J. & Kouzarides, T. Regulation of chromatin by histone modifications. *Cell research* **21**, 381–395 (2011).

24. Zhao, Y. & Garcia, B. A. Comprehensive Catalog of Currently Documented Histone Modifications. *Csh Perspect Biol* **7**, a025064 (2015).

25. Zhang, T., Cooper, S. & Brockdorff, N. The interplay of histone modifications - writers that read. *EMBO reports* **16**, 1467–1481 (2015).

26. Kuo, M.-H. *et al.* Transcription-linked acetylation by Gcn5p of histones H3 and H4 at specific lysines. *Nature* **383**, 269–272 (1996).

27. Brownell, J. E. *et al.* Tetrahymena Histone Acetyltransferase A: A Homolog to Yeast Gcn5p Linking Histone Acetylation to Gene Activation. *Cell* **84**, 843–851 (1996).

28. Fraga, M. F. *et al.* Loss of acetylation at Lys16 and trimethylation at Lys20 of histone H4 is a common hallmark of human cancer. *Nat Genet* **37**, 391–400 (2005).

29. Gayther, S. A. *et al.* Mutations truncating the EP300 acetylase in human cancers. *Nat Genet* **24**, 300–303 (2000).

30. Muraoka, M. *et al.* p300 gene alterations in colorectal and gastric carcinomas. *Oncogene* **12**, 1565–9 (1996).

31. Lin, R. J. & Evans, R. M. Acquisition of Oncogenic Potential by RAR Chimeras in Acute Promyelocytic Leukemia through Formation of Homodimers. *Mol Cell* **5**, 821–830 (2000).

32. Lin, R. J. *et al.* Role of the histone deacetylase complex in acute promyelocytic leukaemia. *Nature* **391**, 811–814 (1998).

33. Zhang, Y. & Reinberg, D. Transcription regulation by histone methylation: interplay between different covalent modifications of the core histone tails. *Gene Dev* **15**, 2343–2360 (2001).

34. Audia, J. E. & Campbell, R. M. Histone Modifications and Cancer. *Cold Spring Harbor perspectives in biology* **8**, a019521 (2016).

35. Black, J. C., Rechem, C. V. & Whetstine, J. R. Histone lysine methylation dynamics: establishment, regulation, and biological impact. *Mol Cell* **48**, 491–507 (2012).

36. Busslinger, M. & Tarakhovsky, A. Epigenetic Control of Immunity. *Csh Perspect Biol* **6**, a019307 (2014).

37. Ruta, V. *et al.* Inhibition of Polycomb Repressive Complex 2 activity reduces trimethylation of H3K27 and affects development in Arabidopsis seedlings. *Bmc Plant Biol* **19**, 429 (2019).

38. Khan, S. N. *et al.* Multiple mechanisms deregulate EZH2 and histone H3 lysine 27 epigenetic changes in myeloid malignancies. *Leukemia* **27**, 1301–1309 (2013).

39. Venneti, S. *et al.* Evaluation of histone 3 lysine 27 trimethylation (H3K27me3) and enhancer of Zest 2 (EZH2) in pediatric glial and glioneuronal tumors shows decreased H3K27me3 in H3F3A K27M mutant glioblastomas. *Brain pathology (Zurich, Switzerland)* **23**, 558–564 (2013).

40. Lewis, P. W. *et al.* Inhibition of PRC2 activity by a gain-of-function H3 mutation found in pediatric glioblastoma. *Sci New York N Y* **340**, 857–61 (2013).

41. Fontebasso, A. M. *et al.* Mutations in SETD2 and genes affecting histone H3K36 methylation target hemispheric high-grade gliomas. *Acta Neuropathol* **125**, 659–669 (2013).

42. Hakimi, A. A. *et al.* Clinical and Pathologic Impact of Select Chromatin-modulating Tumor Suppressors in Clear Cell Renal Cell Carcinoma. *Eur Urol* **63**, 848–854 (2013).

43. Wang, G. G., Cai, L., Pasillas, M. P. & Kamps, M. P. NUP98–NSD1 links H3K36 methylation to Hox-A gene activation and leukaemogenesis. *Nat Cell Biol* **9**, 804–812 (2007).

44. Kang, D. *et al.* The histone methyltransferase Wolf–Hirschhorn syndrome candidate 1-like 1 (WHSC1L1) is involved in human carcinogenesis. *Genes Chromosomes Cancer* **52**, 126–139 (2013).

45. Ravandi, F. *et al.* Gemtuzumab Ozogamicin: Time to Resurrect? *J Clin Oncol* **30**, 3921–3923 (2012).

46. HILLESTAD, L. K. Acute Promyelocytc Leukemia. *Acta Med Scand* **159**, 189–194 (2009).

47. Coombs, C. C., Tavakkoli, M. & Tallman, M. S. Acute promyelocytic leukemia: where did we start, where are we now, and the future. *Blood Cancer J* **5**, e304–e304 (2015).

48. Thé, H. de, Pandolfi, P. P. & Chen, Z. Acute Promyelocytic Leukemia: A Paradigm for Oncoprotein-Targeted Cure. *Cancer Cell* **32**, 552–560 (2017).

49. Grignani, F. *et al.* The acute promyelocytic leukemia-specific PML-RARα fusion protein inhibits differentiation and promotes survival of myeloid precursor cells. *Cell* **74**, 423–431 (1993).

50. Rowley, J., Golomb, H. & Dougherty, C. 15/17 TRANSLOCATION, A CONSISTENT CHROMOSOMAL CHANGE IN ACUTE PROMYELOCYTIC LEUKAEMIA. *Lancet* **309**, 549–550 (1977).

51. Thé, H. de, Chomienne, C., Lanotte, M., Degos, L. & Dejean, A. The t(15;17) translocation of acute promyelocytic leukaemia fuses the retinoic acid receptor α gene to a novel transcribed locus. *Nature* **347**, 558–561 (1990).

52. Ablain, J. *et al.* Uncoupling RARA transcriptional activation and degradation clarifies the bases for APL response to therapies. *J Exp Medicine* **210**, 647–53 (2013).

53. Breitman, T., Collins, S. & Keene, B. Terminal differentiation of human promyelocytic leukemic cells in primary culture in response to retinoic acid. *Blood* **57**, 1000–1004 (1981).

54. Wang, R., Xia, L., Gabrilove, J., Waxman, S. & Jing, Y. Sorafenib Inhibition of Mcl-1 Accelerates ATRA-Induced Apoptosis in Differentiation-Responsive AML Cells. *Clin Cancer Res Official J Am Assoc Cancer Res* **22**, 1211–21 (2015).

55. Dyke, D. L. V. *et al.* Recurrent cytogenetic abnormalities in squamous cell carcinomas of the head and neck region. *Genes Chromosomes Cancer* **9**, 192–206 (1994).

56. Koken, M. H. *et al.* The t(15;17) translocation alters a nuclear body in a retinoic acid-reversible fashion. *Embo J* **13**, 1073–1083 (1994).

57. Ascoli, C. A. & Maul, G. G. Identification of a novel nuclear domain. *J Cell Biology* **112**, 785–795 (1991).

58. Lallemand-Breitenbach, V. & Thé, H. de. PML Nuclear Bodies. *Csh Perspect Biol* **2**, a000661 (2010).

59. Eskiw, C. H., Dellaire, G., Mymryk, J. S. & Bazett-Jones, D. P. Size, position and dynamic behavior of PML nuclear bodies following cell stress as a paradigm for supramolecular trafficking and assembly. *J Cell Sci* **116**, 4455–4466 (2003).

60. Shen, T. H., Lin, H.-K., Scaglioni, P. P., Yung, T. M. & Pandolfi, P. P. The Mechanisms of PML-Nuclear Body Formation. *Mol Cell* **24**, 331–339 (2006).

61. Zhu, J. *et al.* Arsenic-induced PML targeting onto nuclear bodies: Implications for the treatment of acute promyelocytic leukemia. *Proc National Acad Sci* **94**, 3978–3983 (1997).

62. Daniel, M. *et al.* PML protein expression in hematopoietic and acute promyelocytic leukemia cells. *Blood* **82**, 1858–1867 (1993).

63. Weis, K. *et al.* Retinoic acid regulates aberrant nuclear localization of PML-RAR alpha in acute promyelocytic leukemia cells. *Cell* **76**, 345–56 (1994).

64. Korf, K. *et al.* The PML domain of PML-RARα blocks senescence to promote leukemia. *Proceedings of the National Academy of Sciences of the United States of America* **111**, 12133–12138 (2014).

65. Molecular Biology of the Cell (4th Ed). *J Biol Educ* **37**, 45–47 (2002).

66. Henikoff, S. & Smith, M. M. Histone variants and epigenetics. *Csh Perspect Biol* **7**, a019364 (2015).

67. Maehara, K. *et al.* Tissue-specific expression of histone H3 variants diversified after species separation. *Epigenet Chromatin* **8**, 35 (2015).

68. Hake, S. B. & Allis, C. D. Histone H3 variants and their potential role in indexing mammalian genomes: The "H3 barcode hypothesis." *Proc National Acad Sci* **103**, 6428–6435 (2006).

69. Harada, A. *et al.* Histone H3.3 sub-variant H3mm7 is required for normal skeletal muscle regeneration. *Nat Commun* **9**, 1400 (2018).

70. Banaszynski, L. A., Allis, C. D. & Lewis, P. W. Histone variants in metazoan development. *Dev Cell* **19**, 662–74 (2010).

71. Chen, P., Zhao, J. & Li, G. Histone Variants in Development and Diseases. *J Genet Genomics* **40**, 355–365 (2013).

72. Turinetto, V. & Giachino, C. Histone variants as emerging regulators of embryonic stem cell identity. *Epigenetics* **10**, 563–73 (2015).

73. Pasque, V. *et al.* Histone variant macroH2A marks embryonic differentiation in vivo and acts as an epigenetic barrier to induced pluripotency. *J Cell Sci* **125**, 6094–104 (2012).

74. Kong, Q. *et al.* Histone variant H3.3-mediated chromatin remodeling is essential for paternal genome activation in mouse preimplantation embryos. *J Biological Chem* **293**, 3829–3838 (2018).

75. Buschbeck, M. & Hake, S. B. Variants of core histones and their roles in cell fate decisions, development and cancer. *Nature reviews. Molecular cell biology* **18**, 299–314 (2017).

76. Orsi, G. A., Naughtin, M. & Almouzni, G. The Epigenome and Cancer Stem Cell Fate: Connected by a Linker Histone Variant. *Cell Stem Cell* **19**, 567–568 (2016).

77. Vardabasso, C. *et al.* Histone variants: emerging players in cancer biology. *Cell Mol Life Sci Cmls* **71**, 379–404 (2013).

78. Pritchard, A. L. Clinical Epigenetics. 133–151 (2019) doi:10.1007/978-981-13-8958-0_6.

79. Martire, S. & Banaszynski, L. A. The roles of histone variants in fine-tuning chromatin organization and function. *Nat Rev Mol Cell Biology* 1–20 (2020) doi:10.1038/s41580-020-0262-8.

80. Szenker, E., Ray-Gallet, D. & Almouzni, G. The double face of the histone variant H3.3. *Cell research* **21**, 421–434 (2011).

81. Piazzesi, A. *et al.* Replication-Independent Histone Variant H3.3 Controls Animal Lifespan through the Regulation of Pro-longevity Transcriptional Programs. *Cell Reports* **17**, 987–996 (2016).

82. Yuen, B. T. K., Bush, K. M., Barrilleaux, B. L., Cotterman, R. & Knoepfler, P. S. Histone H3.3 regulates dynamic chromatin states during spermatogenesis. *Development* **141**, 3483–3494 (2014).

83. Banaszynski, L. A. *et al.* Hira-Dependent Histone H3.3 Deposition Facilitates PRC2 Recruitment at Developmental Loci in ES Cells. *Cell* **155**, 107–120 (2013).

84. Lewis, P. W., Elsaesser, S. J., Noh, K.-M., Stadler, S. C. & Allis, C. D. Daxx is an H3.3-specific histone chaperone and cooperates with ATRX in replication-independent chromatin assembly at telomeres. *Proceedings of the National Academy of Sciences of the United States of America* **107**, 14075–14080 (2010).

85. Wong, L. H. *et al.* ATRX interacts with H3.3 in maintaining telomere structural integrity in pluripotent embryonic stem cells. *Genome Research* **20**, 351–360 (2010).

86. Goldberg, A. D. *et al.* Distinct Factors Control Histone Variant H3.3 Localization at Specific Genomic Regions. *Cell* **140**, 678–691 (2010).

87. Elsaesser, S. J. & Allis, C. D. HIRA and Daxx Constitute Two Independent Histone H3.3-Containing Predeposition Complexes. *Cold Spring Harb Sym* **75**, 27–34 (2010).

88. Yang, X., Khosravi-Far, R., Chang, H. Y. & Baltimore, D. Daxx, a Novel Fas-Binding Protein That Activates JNK and Apoptosis. *Cell* **89**, 1067–1076 (1997).

89. Croxton, R. *et al.* Daxx Represses Expression of a Subset of Antiapoptotic Genes Regulated by Nuclear Factor-κB. *Cancer Res* **66**, 9026–9035 (2006).

90. Puto, L. A. & Reed, J. C. Daxx represses RelB target promoters via DNA methyltransferase recruitment and DNA hypermethylation. *Gene Dev* **22**, 998–1010 (2008).

91. Puto, L. A., Benner, C. & Hunter, T. The DAXX co-repressor is directly recruited to active regulatory elements genome-wide to regulate autophagy programs in a model of human prostate cancer. *Oncoscience* **2**, 362–72 (2015).

92. Tang, M. *et al.* Disease mutant analysis identifies a new function of DAXX in telomerase regulation and telomere maintenance. *J Cell Sci* **128**, 331–41 (2014).

93. Lin, D.-Y. *et al.* Role of SUMO-Interacting Motif in Daxx SUMO Modification, Subnuclear Localization, and Repression of Sumoylated Transcription Factors. *Mol Cell* **24**, 341–354 (2006).

94. Boellmann, F. *et al.* DAXX interacts with heat shock factor 1 during stress activation and enhances its transcriptional activity. *Proc National Acad Sci* **101**, 4100–4105 (2004).

95. Michod, D. *et al.* Calcium-Dependent Dephosphorylation of the Histone Chaperone DAXX Regulates H3.3 Loading and Transcription upon Neuronal Activation. *Neuron* **74**, 122–135 (2012).

96. Michaelson, J. S., Bader, D., Kuo, F., Kozak, C. & Leder, P. Loss of Daxx, a promiscuously interacting protein, results in extensive apoptosis in early mouse development. *Gene Dev* **13**, 1918–1923 (1999).

97. Michaelson, J. S. & Leder, P. RNAi reveals anti-apoptotic and transcriptionally repressive activities of DAXX. *J Cell Sci* **116**, 345–352 (2002).

98. Zobalova, R., Swettenham, E., Chladova, J., Dong, L.-F. & Neuzil, J. Daxx inhibits stress-induced apoptosis in cardiac myocytes. *Redox Rep* **13**, 263–270 (2008).

99. Chen, L.-Y. & Chen, J. D. Daxx Silencing Sensitizes Cells to Multiple Apoptotic Pathways. *Mol Cell Biol* **23**, 7108–7121 (2003).

100. Pan, W.-W. *et al.* DAXX silencing suppresses mouse ovarian surface epithelial cell growth by inducing senescence and DNA damage. *Gene* **526**, 287–294 (2013).

101. Schreiner, S. *et al.* Adenovirus type 5 early region 1B 55K oncoprotein-dependent degradation of cellular factor Daxx is required for efficient transformation of primary rodent cells. *J Virol* **85**, 8752–65 (2011).

102. Huang, L. *et al.* Daxx interacts with HIV-1 integrase and inhibits lentiviral gene expression. *Biochem Bioph Res Co* **373**, 241–5 (2008).

103. Seeler, J.-S. & Dejean, A. SUMO and the robustness of cancer. *Nature Reviews Cancer* **17**, 184–197 (2017).

104. Santiago, A., Godsey, A. C., Hossain, J., Zhao, L. Y. & Liao, D. Identification of two independent SUMO-interacting motifs in Daxx: Evolutionary conservation from Drosophila to humans and their biochemical functions. *Cell Cycle* **8**, 76–87 (2009).

105. Mahmud, I. & Liao, D. DAXX in cancer: phenomena, processes, mechanisms and regulation. *Nucleic Acids Res* **15**, 593 (2019).

106. Pan, W.-W. *et al.* Death Domain-associated Protein DAXX Promotes Ovarian Cancer Development and Chemoresistance. *J Biol Chem* **288**, 13620–13630 (2013).

107. Wang, X., Zhao, Y., Zhang, J. & Chen, Y. Structural basis for DAXX interaction with ATRX. *Protein Cell* **8**, 767–771 (2017).

108. Tang, J. *et al.* A Novel Transcription Regulatory Complex Containing Death Domain-associated Protein and the ATR-X Syndrome Protein. *J Biol Chem* **279**, 20369–20377 (2004).

109. Xue, Y. *et al.* The ATRX syndrome protein forms a chromatin-remodeling complex with Daxx and localizes in promyelocytic leukemia nuclear bodies. *Proc National Acad Sci* **100**, 10635–10640 (2003).

110. Picketts, D. J. *et al.* ATRX encodes a novel member of the SNF2 family of proteins: mutations point to a common mechanism underlying the ATR-X syndrome. *Hum Mol Genet* **5**, 1899–1907 (1996).

111. Watson, L. A., Goldberg, H. & Bérubé, N. G. Emerging roles of ATRX in cancer. *Epigenomics-uk* **7**, 1365–1378 (2015).

112. Bérubé, N. G., Smeenk, C. A. & Picketts, D. J. Cell cycle-dependent phosphorylation of the ATRX protein correlates with changes in nuclear matrix and chromatin association. *Hum Mol Genet* **9**, 539–547 (2000).

113. Kourmouli, N., Sun, Y., Sar, S. van der, Singh, P. B. & Brown, J. P. Epigenetic regulation of mammalian pericentric heterochromatin in vivo by HP1. *Biochem Bioph Res Co* **337**, 901–907 (2005).

114. Dhayalan, A. *et al.* The ATRX-ADD domain binds to H3 tail peptides and reads the combined methylation state of K4 and K9. *Human Molecular Genetics* **20**, 2195–2203 (2011).

115. Iwase, S. *et al.* ATRX ADD domain links an atypical histone methylation recognition mechanism to human mental-retardation syndrome. *Nat Struct Mol Biol* **18**, 769–76 (2011).

116. Eustermann, S. *et al.* Combinatorial readout of histone H3 modifications specifies localization of ATRX to heterochromatin. *Nat Struct Mol Biol* **18**, 777–782 (2011).

117. Baumann, C., Viveiros, M. M. & Fuente, R. D. L. Loss of maternal ATRX results in centromere instability and aneuploidy in the mammalian oocyte and pre-implantation embryo. *PLOS Genetics* **6**, e1001137 (2010).

118. Watson, L. A. *et al.* Atrx deficiency induces telomere dysfunction, endocrine defects, and reduced life span. *J Clin Invest* **123**, 2049–2063 (2013).

119. Bérubé, N. G. *et al.* The chromatin-remodeling protein ATRX is critical for neuronal survival during corticogenesis. *J Clin Invest* **115**, 258–267 (2005).

120. Ritchie, K., Watson, L. A., Davidson, B., Jiang, Y. & Bérubé, N. G. ATRX is required for maintenance of the neuroprogenitor cell pool in the embryonic mouse brain. *Biology open* **3**, 1158–1163 (2014).

121. Bérubé, N. G. *et al.* Neurodevelopmental defects resulting from ATRX overexpression in transgenic mice. *Hum Mol Genet* **11**, 253–261 (2002).

122. Gibbons, R. J. *et al.* Mutations in the chromatin-associated protein ATRX. *Hum Mutat* **29**, 796–802 (2008).

123. Bérubé, N. G. *et al.* Patient mutations alter ATRX targeting to PML nuclear bodies. *Eur J Hum Genet* **16**, 192–201 (2007).

124. Gibbons, R. J. *et al.* Identification of acquired somatic mutations in the gene encoding chromatin-remodeling factor ATRX in the α-thalassemia myelodysplasia syndrome (ATMDS). *Nat Genet* **34**, 446–449 (2003).

125. Lacayo, N. J. *et al.* Gene expression profiles at diagnosis in de novo childhood AML patients identify FLT3 mutations with good clinical outcomes. *Blood* **104**, 2646–2654 (2004).

126. Qadeer, Z. A. *et al.* Decreased expression of the chromatin remodeler ATRX associates with melanoma progression. *J Investigative Dermatology* **134**, 1768–72 (2014).

127. Jiao, Y. *et al.* DAXX/ATRX, MEN1, and mTOR pathway genes are frequently altered in pancreatic neuroendocrine tumors. *Sci New York N Y* **331**, 1199–203 (2011).

128. Yachida, S. *et al.* Small cell and large cell neuroendocrine carcinomas of the pancreas are genetically similar and distinct from well-differentiated pancreatic neuroendocrine tumors. *Am J Surg Pathology* **36**, 173–84 (2012).

129. Marinoni, I. *et al.* Loss of DAXX and ATRX Are Associated With Chromosome Instability and Reduced Survival of Patients With Pancreatic Neuroendocrine Tumors. *Gastroenterology* **146**, 453-460.e5 (2014).

130. Heaphy, C. M. *et al.* Altered telomeres in tumors with ATRX and DAXX mutations. *Sci New York N Y* **333**, 425 (2011).

131. Schwartzentruber, J. *et al.* Driver mutations in histone H3.3 and chromatin remodelling genes in paediatric glioblastoma. *Nature* **482**, 226–231 (2012).

132. Jiao, Y. *et al.* Frequent ATRX, CIC, FUBP1 and IDH1 mutations refine the classification of malignant gliomas. *Oncotarget* **3**, 709–722 (2012).

133. Jafri, M. A., Ansari, S. A., Alqahtani, M. H. & Shay, J. W. Roles of telomeres and telomerase in cancer, and advances in telomerase-targeted therapies. *Genome Med* **8**, 69 (2016).

134. Amorim, J. P., Santos, G., Vinagre, J. & Soares, P. The Role of ATRX in the Alternative Lengthening of Telomeres (ALT) Phenotype. *Genes-basel* **7**, 66 (2016).

135. Killela, P. J. *et al.* TERT promoter mutations occur frequently in gliomas and a subset of tumors derived from cells with low rates of self-renewal. *Proc National Acad Sci* **110**, 6021–6026 (2013).

136. Drané, P., Ouararhni, K., Depaux, A., Shuaib, M. & Hamiche, A. The death-associated protein DAXX is a novel histone chaperone involved in the replication-independent deposition of H3.3. *Genes & Development* **24**, 1253–1265 (2010).

137. Delbarre, E. *et al.* PML protein organizes heterochromatin domains where it regulates histone H3.3 deposition by ATRX/DAXX. *Genome Research* **27**, 913–921 (2017).

138. Hoelper, D., Huang, H., Jain, A. Y., Patel, D. J. & Lewis, P. W. Structural and mechanistic insights into ATRX-dependent and -independent functions of the histone chaperone DAXX. *Nature Communications* **8**, 1193 (2017).

139. Maze, I., Noh, K.-M. & Allis, C. D. Histone regulation in the CNS: basic principles of epigenetic plasticity. *Neuropsychopharmacol Official Publ Am Coll Neuropsychopharmacol* **38**, 3–22 (2012).

140. Chen, I., Howarth, M., Lin, W. & Ting, A. Y. Site-specific labeling of cell surface proteins with biophysical probes using biotin ligase. *Nat Methods* **2**, 99–104 (2005).

141. Slavoff, S. A., Chen, I., Choi, Y.-A. & Ting, A. Y. Expanding the Substrate Tolerance of Biotin Ligase through Exploration of Enzymes from Diverse Species. *J Am Chem Soc* **130**, 1160–1162 (2008).

142. Liu, W., Samanta, S. K., Smith, B. D. & Isaacs, L. Synthetic mimics of biotin/(strept)avidin. *Chem Soc Rev* **46**, 2391–2403 (2017).

143. Kim, D. I. *et al.* An improved smaller biotin ligase for BioID proximity labeling. *Molecular biology of the cell* **27**, 1188–1196 (2016).

144. Li, P., Li, J., Wang, L. & Di, L.-J. Proximity Labeling of Interacting Proteins: Application of BioID as a Discovery Tool. *Proteomics* **17**, 1700002 (2017).

145. Bultman, S. *et al.* A Brg1 Null Mutation in the Mouse Reveals Functional Differences among Mammalian SWI/SNF Complexes. *Mol Cell* **6**, 1287–1295 (2000).

146. Metzger, D. & Chambon, P. Site- and Time-Specific Gene Targeting in the Mouse. *Methods* **24**, 71–80 (2001).

147. Koken, M. H. *et al.* The t(15;17) translocation alters a nuclear body in a retinoic acid-reversible fashion. *Embo J* **13**, 1073–1083 (1994).

148. Ivanauskiene, K. *et al.* The PML-associated protein DEK regulates the balance of H3.3 loading on chromatin and is important for telomere integrity. *Genome Research* **24**, 1584–1594 (2014).

149. Hollenbach, A. D., Sublett, J. E., McPherson, C. J. & Grosveld, G. The Pax3-FKHR oncoprotein is unresponsive to the Pax3-associated repressor hDaxx. *Embo J* **18**, 3702–3711 (1999).

150. Korf, K., Wodrich, H., Haschke, A., Evans, R. M. & Sternsdorf, T. M. The Acute Promyelocytic Leukemia-Oncoprotein PML-Raralpha Blocks Senescence and Disrupts The Atrx/Daxx Chromatin Remodeling Complex To Promote Leukemia. *Blood* **122**, 1267–1267 (2013).

151. Gerstenberger, B. S. *et al.* Identification of a Chemical Probe for Family VIII Bromodomains through Optimization of a Fragment Hit. *J Med Chem* **59**, 4800–4811 (2016).

152. Wu, Q. *et al.* The SWI/SNF ATPases Are Required for Triple Negative Breast Cancer Cell Proliferation. *J Cell Physiol* **230**, 2683–94 (2015).

153. Szostecki, C., Guldner, H. H., Netter, H. J. & Will, H. Isolation and characterization of cDNA encoding a human nuclear antigen predominantly recognized by autoantibodies from patients with primary biliary cirrhosis. *J Immunol Baltim Md 1950* **145**, 4338–47 (1990).

154. Yamada, A., Arakaki, R., Saito, M., Kudo, Y. & Ishimaru, N. Dual Role of Fas/FasL-Mediated Signal in Peripheral Immune Tolerance. *Front Immunol* **8**, 403 (2017).

155. Ko, Y.-G. *et al.* Apoptosis Signal-regulating Kinase 1 Controls the Proapoptotic Function of Death-associated Protein (Daxx) in the Cytoplasm. *J Biol Chem* **276**, 39103–39106 (2001).

156. Lindsay, C. R., Giovinazzi, S. & Ishov, A. M. Daxx is a predominately nuclear protein that does not translocate to the cytoplasm in response to cell stress. *Cell Cycle* **8**, 1544–1551 (2009).

157. Sternsdorf, T., Jensen, K., Reich, B. & Will, H. The Nuclear Dot Protein Sp100, Characterization of Domains Necessary for Dimerization, Subcellular Localization, and Modification by Small Ubiquitin-like Modifiers. *J Biol Chem* **274**, 12555–12566 (1999).

158. Varnaitė, R. & MacNeill, S. A. Meet the neighbors: Mapping local protein interactomes by proximity-dependent labeling with BioID. *Proteomics* **16**, 2503–2518 (2016).

159. Gullberg, M. *et al.* Cytokine detection by antibody-based proximity ligation. *P Natl Acad Sci Usa* **101**, 8420–8424 (2004).

160. Söderberg, O. *et al.* Direct observation of individual endogenous protein complexes in situ by proximity ligation. *Nat Methods* **3**, 995–1000 (2006).

161. Karamouzis, M. V. *et al.* HER-3 targeting alters the dimerization pattern of ErbB protein family members in breast carcinomas. *Oncotarget* **7**, 5576–5597 (2015).

162. Vincent, A., Berthel, E., Dacheux, E., Magnard, C. & Venezia, N. L. D. BRCA1 affects protein phosphatase 6 signalling through its interaction with ANKRD28. *Biochem J* **473**, 949–960 (2016).

163. Wang, S. *et al.* Detection of In Situ Protein-protein Complexes at the Drosophila Larval Neuromuscular Junction Using Proximity Ligation Assay. *J Vis Exp* 52139 (2015) doi:10.3791/52139.

164. Kwon, J., Jeong, S., Choi, I. & Kim, N.-H. ADAM10 Is Involved in Cell Junction Assembly in Early Porcine Embryo Development. *Plos One* **11**, e0152921 (2016).

165. Cierpicki, T. & Grembecka, J. Targeting protein-protein interactions in hematologic malignancies: still a challenge or a great opportunity for future therapies? *Immunol Rev* **263**, 279–301 (2014).

166. Nooren, I. M. A. & Thornton, J. M. NEW EMBO MEMBER'S REVIEW: Diversity of protein-protein interactions. *Embo J* **22**, 3486–3492 (2003).

167. Li, B., Rong, D. & Wang, Y. Targeting Protein-Protein Interaction with Covalent Small-Molecule Inhibitors. *Curr Top Med Chem* **19**, 1872–1876 (2019).

168. Ofran, Y. & Rost, B. Analysing Six Types of Protein–Protein Interfaces. *J Mol Biol* **325**, 377–387 (2003).

169. Rabbani, G., Baig, M. H., Ahmad, K. & Choi, I. Protein-protein Interactions and their Role in Various Diseases and their Prediction Techniques. *Curr Protein Pept Sc* **19**, 948–957 (2018).

170. Schleker, S., Kshirsagar, M. & Klein-Seetharaman, J. Comparing human-Salmonella with plant-Salmonella protein-protein interaction predictions. *Front Microbiol* **6**, 45 (2015).

171. Schleker, S. *et al.* The current Salmonella-host interactome. *Proteom - Clin Appl* **6**, 117–133 (2011).

172. Kanki, P. J. *et al.* Human Immunodeficiency Virus Type 1 Subtypes Differ in Disease Progression. *J Infect Dis* **179**, 68–73 (1999).

173. Arkin, M. R., Tang, Y. & Wells, J. A. Small-molecule inhibitors of protein-protein interactions: progressing toward the reality. *Chem Biol* **21**, 1102–14 (2014).

174. Stockwin, L. & Holmes, S. Antibodies as therapeutic agents: vive la renaissance! *Expert Opin Biol Th* **3**, 1133–1152 (2003).

175. Hwang, H., Vreven, T., Janin, J. & Weng, Z. Protein-protein docking benchmark version 4.0. *Proteins Struct Funct Bioinform* **78**, 3111–3114 (2010).

176. Fuller, J. C., Burgoyne, N. J. & Jackson, R. M. Predicting druggable binding sites at the protein–protein interface. *Drug Discov Today* **14**, 155–161 (2009).

177. Clackson, T. & Wells, J. A hot spot of binding energy in a hormone-receptor interface. *Science* **267**, 383–386 (1995).

178. Bogan, A. A. & Thorn, K. S. Anatomy of hot spots in protein interfaces. *J Mol Biol* **280**, 1–9 (1998).

179. DeLano, W. L. Unraveling hot spots in binding interfaces: progress and challenges. *Curr Opin Struc Biol* **12**, 14–20 (2002).

180. Ma, B., Elkayam, T., Wolfson, H. & Nussinov, R. Protein-protein interactions: Structurally conserved residues distinguish between binding sites and exposed protein surfaces. *Proc National Acad Sci* **100**, 5772–5777 (2003).

181. Wells, J. A. & McClendon, C. L. Reaching for high-hanging fruit in drug discovery at protein–protein interfaces. *Nature* **450**, 1001–1009 (2007).

182. Higueruelo, A. P. *et al.* Atomic Interactions and Profile of Small Molecules Disrupting Protein-Protein Interfaces: the TIMBAL Database. *Chem Biol Drug Des* **74**, 457–467 (2009).

183. Basse, M. J. *et al.* 2P2Idb: a structural database dedicated to orthosteric modulation of protein–protein interactions. *Nucleic Acids Res* **41**, D824–D827 (2012).

184. Labbé, C. M., Laconde, G., Kuenemann, M. A., Villoutreix, B. O. & Sperandio, O. iPPI-DB: a manually curated and interactive database of small non-peptide inhibitors of protein–protein interactions. *Drug Discov Today* **18**, 958–968 (2013).

185. Smith, M. C. & Gestwicki, J. E. Features of protein–protein interactions that translate into potent inhibitors: topology, surface area and affinity. *Expert Rev Mol Med* **14**, e16 (2012).

186. Maurer, T. *et al.* Small-molecule ligands bind to a distinct pocket in Ras and inhibit SOS-mediated nucleotide exchange activity. *Proc National Acad Sci* **109**, 5299–5304 (2012).

187. Sun, Q. *et al.* Discovery of Small Molecules that Bind to K-Ras and Inhibit Sos-Mediated Activation. *Angewandte Chemie Int Ed* **51**, 6140–6143 (2012).

188. Ostrem, J. M., Peters, U., Sos, M. L., Wells, J. A. & Shokat, K. M. K-Ras(G12C) inhibitors allosterically control GTP affinity and effector interactions. *Nature* **503**, 548–551 (2013).

189. Shima, F. *et al.* In silico discovery of small-molecule Ras inhibitors that display antitumor activity by blocking the Ras-effector interaction. *Proc National Acad Sci* **110**, 8182–8187 (2013).

190. Luciani, J. J. *et al.* PML nuclear bodies are highly organised DNA-protein structures with a function in heterochromatin remodelling at the G2 phase. *J Cell Sci* **119**, 2518–2531 (2006).

191. Ishov, A. M. *et al.* Pml Is Critical for Nd10 Formation and Recruits the Pml-Interacting Protein Daxx to This Nuclear Structure When Modified by Sumo-1. *J Cell Biology* **147**, 221–234 (1999).

192. Everett, R. D. *et al.* A dynamic connection between centromeres and ND10 proteins. *J Cell Sci* **112 (Pt 20)**, 3443–54 (1999).

193. Hollenbach, A. D., McPherson, C. J., Mientjes, E. J., Iyengar, R. & Grosveld, G. Daxx and histone deacetylase II associate with chromatin through an interaction with core histones and the chromatin-associated protein Dek. *J Cell Sci* **115**, 3319–30 (2002).

194. Jansz, N., Chen, K., Murphy, J. M. & Blewitt, M. E. The Epigenetic Regulator SMCHD1 in Development and Disease. *Trends Genet* **33**, 233–243 (2017).

195. Jansz, N. *et al.* Smchd1 regulates long-range chromatin interactions on the inactive X chromosome and at Hox clusters. *Nat Struct Mol Biol* **25**, 766–777 (2018).

196. Gdula, M. R. *et al.* The non-canonical SMC protein SmcHD1 antagonises TAD formation and compartmentalisation on the inactive X chromosome. *Nat Commun* **10**, 30 (2019).

197. Blewitt, M. E. *et al.* SmcHD1, containing a structural-maintenance-of-chromosomes hinge domain, has a critical role in X inactivation. *Nat Genet* **40**, 663–669 (2008).

198. Gordon, C. T. *et al.* De novo mutations in SMCHD1 cause Bosma arhinia microphthalmia syndrome and abrogate nasal development. *Nature Genetics* **49**, 249–255 (2017).

199. Lemmers, R. J. L. F. *et al.* Digenic inheritance of an SMCHD1 mutation and an FSHD-permissive D4Z4 allele causes facioscapulohumeral muscular dystrophy type 2. *Nat Genet* **44**, 1370–1374 (2012).

200. Shaw, N. D. *et al.* SMCHD1 mutations associated with a rare muscular dystrophy can also cause isolated arhinia and Bosma arhinia microphthalmia syndrome. *Nat Genet* **49**, 238–248 (2017).

201. Galupa, R. & Heard, E. X-Chromosome Inactivation: A Crossroads Between Chromosome Architecture and Gene Regulation. *Annu Rev Genet* **52**, 1–32 (2016).

202. Svadlenka, J., Brazina, J., Hanzlikova, H., Cermak, L. & Andera, L. Multifunctional adaptor protein Daxx interacts with chromatin-remodelling ATPase Brg1. *Biochem Biophysics Reports* **5**, 246–252 (2016).

203. Lin, R. J., Sternsdorf, T., Tini, M. & Evans, R. M. Transcriptional regulation in acute promyelocytic leukemia. *Oncogene* **20**, 7204–7215 (2001).

204. Quinn, J., Fyrberg, A. M., Ganster, R. W., Schmidt, M. C. & Peterson, C. L. DNA-binding properties of the yeast SWI/SNF complex. *Nature* **379**, 844–847 (1996).

205. Fryer, C. J. & Archer, T. K. Chromatin remodelling by the glucocorticoid receptor requires the BRG1 complex. *Nature* **393**, 88–91 (1998).

206. Serna, I. L. de la, Carlson, K. A. & Imbalzano, A. N. Mammalian SWI/SNF complexes promote MyoD-mediated muscle differentiation. *Nat Genet* **27**, 187–190 (2001).

207. Kadoch, C. *et al.* Proteomic and bioinformatic analysis of mammalian SWI/SNF complexes identifies extensive roles in human malignancy. *Nature Genetics* **45**, 592–601 (2013).

208. Shain, A. H. & Pollack, J. R. The spectrum of SWI/SNF mutations, ubiquitous in human cancers. *PLOS ONE* **8**, e55119 (2013).

209. Hodges, C., Kirkland, J. G. & Crabtree, G. R. The Many Roles of BAF (mSWI/SNF) and PBAF Complexes in Cancer. *Csh Perspect Med* **6**, a026930 (2016).

210. Orvis, T. *et al.* BRG1/SMARCA4 Inactivation Promotes Non-Small Cell Lung Cancer Aggressiveness by Altering Chromatin Organization. *Cancer Res* **74**, 6486–6498 (2014).

211. Yoshimoto, T. *et al.* Frequent loss of the expression of multiple subunits of the SWI/SNF complex in large cell carcinoma and pleomorphic carcinoma of the lung. *Pathol Int* **65**, 595–602 (2015).

212. Romero, O. A. *et al.* The tumour suppressor and chromatin-remodelling factor BRG1 antagonizes Myc activity and promotes cell differentiation in human cancer. *EMBO molecular medicine* **4**, 603–616 (2012).

213. Reisman, D. N., Sciarrotta, J., Wang, W., Funkhouser, W. K. & Weissman, B. E. Loss of BRG1/BRM in human lung cancer cell lines and primary lung cancers: correlation with poor prognosis. *Cancer Res* **63**, 560–6 (2003).

214. Schneppenheim, R. *et al.* Germline Nonsense Mutation and Somatic Inactivation of SMARCA4/BRG1 in a Family with Rhabdoid Tumor Predisposition Syndrome. *The American Journal of Human Genetics* **86**, 279–284 (2010).

215. Hasselblatt, M. *et al.* Nonsense mutation and inactivation of SMARCA4 (BRG1) in an atypical teratoid/rhabdoid tumor showing retained SMARCB1 (INI1) expression. *Am J Surg Pathology* **35**, 933–5 (2011).

216. Pugh, T. J. *et al.* Medulloblastoma exome sequencing uncovers subtype-specific somatic mutations. *Nature* **488**, 106–10 (2012).

217. Love, C. *et al.* The genetic landscape of mutations in Burkitt lymphoma. *Nat Genet* **44**, 1321–5 (2012).

218. Pottier, N. *et al.* The SWI/SNF chromatin-remodeling complex and glucocorticoid resistance in acute lymphoblastic leukemia. *J Natl Cancer I* **100**, 1792–803 (2008).

219. Bultman, S. J. *et al.* Characterization of mammary tumors from Brg1 heterozygous mice. *Oncogene* **27**, 460–468 (2007).

220. Watanabe, T., Semba, S. & Yokozaki, H. Regulation of PTEN expression by the SWI/SNF chromatin-remodelling protein BRG1 in human colorectal carcinoma cells. *Brit J Cancer* **104**, 146–154 (2010).

221. Shi, J. *et al.* Role of SWI/SNF in acute leukemia maintenance and enhancer-mediated Myc regulation. *Gene Dev* **27**, 2648–62 (2013).

222. Buscarlet, M. *et al.* Essential role of BRG, the ATPase subunit of BAF chromatin remodeling complexes, in leukemia maintenance. *Blood* **123**, 1720–8 (2014).

223. Lou, D. I. *et al.* An Intrinsically Disordered Region of the DNA Repair Protein Nbs1 Is a Species-Specific Barrier to Herpes Simplex Virus 1 in Primates. *Cell Host Microbe* **20**, 178–188 (2016).

224. Maul, G. G. & Everett, R. D. The nuclear location of PML, a cellular member of the C3HC4 zinc-binding domain protein family, is rearranged during herpes simplex virus infection by the C3HC4 viral protein ICP0. *J Gen Virol* **75**, 1223–1233 (1994).

225. Maison, C. & Almouzni, G. HP1 and the dynamics of heterochromatin maintenance. *Nat Rev Mol Cell Bio* **5**, 296–305 (2004).

226. Vollmuth, F., Blankenfeldt, W. & Geyer, M. Structures of the dual bromodomains of the P-TEFb-activating protein Brd4 at atomic resolution. *J Biological Chem* **284**, 36547–56 (2009).

227. Pastori, C. *et al.* BET bromodomain proteins are required for glioblastoma cell proliferation. *Epigenetics* **9**, 611–620 (2014).

228. Alqahtani, A. *et al.* Bromodomain and extra-terminal motif inhibitors: a review of preclinical and clinical advances in cancer therapy. *Future Science OA* **5**, 1–20 (2019).

229. Fedorov, O. *et al.* Selective targeting of the BRG/PB1 bromodomains impairs embryonic and trophoblast stem cell maintenance. *Science advances* **1**, e1500723 (2015).

230. Vangamudi, B. *et al.* The SMARCA2/4 ATPase Domain Surpasses the Bromodomain as a Drug Target in SWI/SNF-Mutant Cancers: Insights from cDNA Rescue and PFI-3 Inhibitor Studies. *Cancer research* **75**, 3865–3878 (2015).

231. Dawson, M. A. *et al.* Inhibition of BET recruitment to chromatin as an effective treatment for MLL-fusion leukaemia. *Nature* **478**, 529–533 (2011).

232. Stewart, H. J. S., Horne, G. A., Bastow, S. & Chevassut, T. J. T. BRD4 associates with p53 in DNMT3A-mutated leukemia cells and is implicated in apoptosis by the bromodomain inhibitor JQ1. *Cancer Med-us* **2**, 826–835 (2013).

233. Zuber, J. *et al.* RNAi screen identifies Brd4 as a therapeutic target in acute myeloid leukaemia. *Nature* **478**, 524–528 (2011).

234. Mertz, J. A. *et al.* Targeting MYC dependence in cancer by inhibiting BET bromodomains. *Proc National Acad Sci* **108**, 16669–16674 (2011).

235. Ott, C. J. *et al.* BET bromodomain inhibition targets both c-Myc and IL7R in high-risk acute lymphoblastic leukemia. *Blood* **120**, 2843–52 (2012).

236. Filippakopoulos, P. & Knapp, S. Targeting bromodomains: epigenetic readers of lysine acetylation. *Nature reviews. Drug discovery* **13**, 337–356 (2014).

237. Tough, D. F., Rioja, I. & Prinjha, R. K. Histone Recognition. 199–223 (2015) doi:10.1007/978-3-319-18102-8_10.

238. Filippakopoulos, P. *et al.* Selective inhibition of BET bromodomains. *Nature* **468**, 1067–1073 (2010).

239. Shi, X. *et al.* JQ1: a novel potential therapeutic target. *Die Pharmazie* **73**, 491–493 (2018).

240. Ceribelli, M. *et al.* Blockade of oncogenic I B kinase activity in diffuse large B-cell lymphoma by bromodomain and extraterminal domain protein inhibitors. *Proc National Acad Sci* **111**, 11365–11370 (2014).

241. Raab, J. R., Runge, J. S., Spear, C. C. & Magnuson, T. Co-regulation of transcription by BRG1 and BRM, two mutually exclusive SWI/SNF ATPase subunits. *Epigenetics & chromatin* **10**, 1–15 (2017).

242. Melchior, F. SUMO-NONCLASSICAL UBIQUITIN. *Annu Rev Cell Dev Bi* **16**, 591–626 (2000).

243. Oluwadare, O., Highsmith, M. & Cheng, J. An Overview of Methods for Reconstructing 3-D Chromosome and Genome Structures from Hi-C Data. *Biol Proced Online* **21**, 7 (2019).

244. Weber, K., Mock, U., Petrowitz, B., Bartsch, U. & Fehse, B. Lentiviral gene ontology (LeGO) vectors equipped with novel drug-selectable fluorescent proteins: new building blocks for cell marking and multi-gene analysis. *Gene Ther* **17**, 511–520 (2010).