Characterization of BEND3, a novel interactor of deubiquitinase USP21

Sophia Teichmann

TESI DOCTORAL UPF / 2012

DIRECTOR DE LA TESI

Dr. Luciano Di Croce

GENE REGULATION, STEM CELLS AND CANCER, CRG

UNIVERSITAT POMPEU FABRA

Acknowledgements

In the first place, I would like to thank my supervisor Dr. Luciano Di Croce for his expert guidance. He has been a supportive and kind mentor throughout the years of my thesis always encouraging me to pursuit my personal ideas.

Also, I am grateful to my thesis committee members Dr. Salvador Aznar-Benitah, Dr. Raúl Méndez and Dr. Gabriel Gil. Their constructive advice and assistance have contributed essentially to the success of this work.

Special thanks go to all my lab mates for their assistance and sincere interest. They have been always present for discussions on scientific matter or general issues and their knowledge and helpfulness have been of great importance. Many thanks for the nice time in the lab.

Moreover I would like to thank all my colleagues from the Gene Regulation, Stem Cells and Cancer Department for their help and the inspiring atmosphere.

I am grateful for the excellent expertise of the Genomics, Proteomics and Biomolecular Screening and Protein Technologies Units and their help with the interpretation of my results.

I gratefully acknowledge the assistance of the CRG/“la Caixa” International PhD Fellowship Programme.

Finally I want to say a big thank you to my family and friends and also to the Orquestra Universitat de Barcelona (OUB).

Abstract

Monoubiquitination of histone H2A fulfills an essential role in gene repression. Several deubiquitinating enzymes (DUBs) have been identified that specifically remove ubiquitin in different cellular contexts. However, the regulation of their enzymatic activity as well as their interplay with the ubiquitination machinery are not well understood. We studied the H2A specific deubiquitinase USP21, which modulates transcriptional initiation.

We identified BEND3 as interactor of nuclear USP21. BEND3 is characterized by a quadruple repeat of BEN domains, is localized to heterochromatin, and acts as a transcriptional repressor. We validated and mapped the interaction between USP21 and BEND3 and found that BEND3 is polyubiquitinated. BEND3 protein stability and subcellular localization were independent of the catalytic activity of USP21 protein, and H2Aub levels were not influenced by BEND3. Finally microarray gene expression analysis in mouse embryonic stem cells depleted for BEND3 revealed the differential expression of genes involved in cellular growth proliferation and cell cycle amongst others.

Resumen

La monoubiquitinación de la histona H2A cumple un rol esencial en la represión génica. Diversas enzimas desubiquitinantes (DUBs) han sido identificadas, que remueven específicamente la ubiquitina de H2A en distintos contextos celulares. Sin embargo poco se sabe sobre la regulación de esta actividad enzimática y de la interrelación con la maquinaria de ubiquitinación. En el presente trabajo hemos estudiado USP21, una DUB específica de H2A, que modula la iniciación de la transcripción.

Hemos identificado BEND3 como interactor del USP21 nuclear. BEND3 se distingue por la presencia de cuatro dominios BEN, está localizado en la heterocromatina y actúa como represor de la transcripción. Hemos validado y delimitado la interacción entre USP21 y BEND3 y hemos observado que BEND3 está polyubiquitinado. La estabilidad proteica de BEND3 y su localización subcelular fueron independientes de la actividad catalítica de USP21 y los niveles de H2Aub no fueron influenciados por BEND3. Finalmente, el análisis de la expresión génica en células madre embrionarias de ratón revela la expresión diferencial de genes involucrados en crecimiento, proliferación celular y ciclo celular.

Preface

Heritable changes in gene expression do not exclusively result from changes in the nucleotide sequence but instead are also consequence of chromatin modifications. Those modifications represent the epigenome and are set in response to external and internal cues what confers a high degree of plasticity to gene regulation. In the last years it has become evident that epigenetic misregulation is a major cause for cancerogenesis and new anti-cancer drugs have been developed that target proteins executing key functions in epigenetic signaling. Stem cells are an important research field as they offer applications in regenerative medicine and understanding differentiation processes whose perturbations lead to tumorigenesis.

Monoubiquitination of H2A is a posttranslational chromatin modification associated with transcriptional repression and its global levels are decreased in several cancer types. USP21 is a deubiquitinating enzyme that acts on histone H2A by removing the ubiquitin monomer. In the study we aimed at understanding the mechanistic implication of USP21-mediated histone deubiquitination in transcription. We identified BEND3 as uncharacterized main interactor of nuclear USP21. Moreover we sought to investigate the role of both proteins in mouse embryonic stem cells and their impact on the stem cell state.

Index

List of Figures

List of Tables

1. Introduction

1.1. Mouse Embryonic Stem Cells (mESC)

1.1.1. Pluripotency and self-renewal

Mouse embryonic stem cells (mESC) are obtained from the inner cell mass of blastocysts. The first cell cultures that could be propagated *in vitro* were established in 1981 [1]. mESC are characterized by two unique features, self-renewal and pluripotency. Self-renewal is their capacity to undergo cell division giving rise to two identical daughter cells. Pluripotency means that cells can differentiate into all germ layers present in an organism [2]. Indeed, it has been shown that the mESC have the capacity to form the primitive ectoderm, the primitive endoderm and trophectoderm cells [3]. ESC are maintained in a state where they can stably keep their stem cell features but on the other hand are poised for starting rapidly the differentiation program. This is denominated as the stem cell state and its maintenance depends on the exact regulation of gene expression [2].

1.1.2. Transcriptional regulation of the stem cell state

The stem cell state is precisely controlled by the three core transcription factors OCT4, SOX2 and NANOG. OCT4, SOX2 and NANOG act together as the “core regulatory circuitry” and their interplay is highly regulated. The three factors regulate the expression of a gene network essential for the persistence of pluripotency and self-renewal of mESC. The activation and the maintenance of these transcription factors in the mESC can be influenced by extracellular signals such as the leukemia inhibitory factor (LIF). LIF signaling activates the transcription factor STAT3 that co-regulates various genes binding to the same enhancers as the core factors OCT4, SOX2 and NANOG [2]. When LIF is withdrawn from tissue culture mESC enter the differentiation program. This process starts with the formation of spherical cell agglomerates, the embryoid bodies, which correspond to early stages of embryogenesis [4]. The administration of retinoic acid (RA) drives mESC into the neuronal

differentiation program and the cells differentiate into neuronal progenitors and postmitotic neurons [5] (Figure 1).

It is important to note that the core factors mutually regulate their expression in a positive feedback loop that controls the undifferentiated state of the cells [2]. On the other hand, somatic cells can be reprogrammed into induced pluripotent stem cells (iPS) by ectopic expression of the core factors and additional factors that play a crucial role in maintaining the stem cell state. It was demonstrated that adult fibroblast are converted into iPS by the ectopic expression of OCT3/4, SOX2, c-MYC, and KLF4 [6].

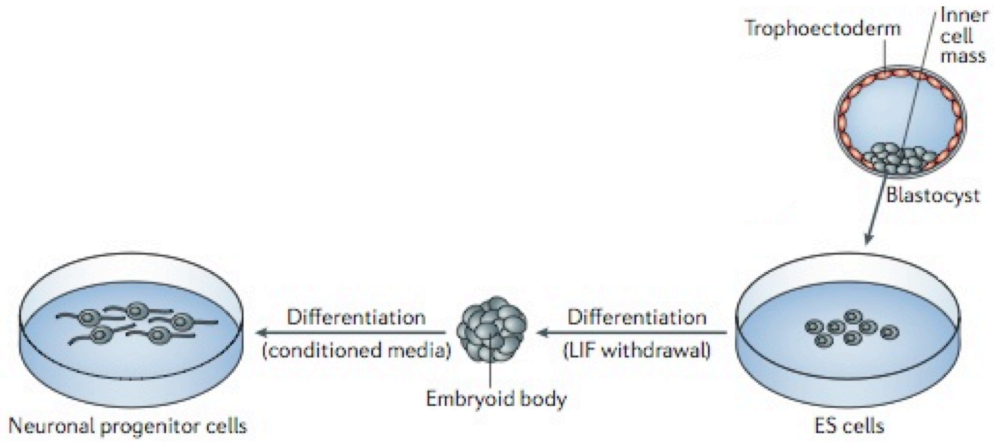


Figure 1. In vitro culture of mESC. Cell cultures of mESC are derived form the inner cell mass of blastocysts and the maintenance of pluripotency and self-renewal depends on the presence of leukemia inhibitory factor (LIF) in the cell culture medium. Without LIF mESC start differentiating. Addition of retinoic acid further induces differentiation into neuronal cells. Figure adapted from [4].

1.2. Chromatin

The mechanisms ensuring self-renewal and pluripotency of stem cells are predominantly dependent on transcriptional regulation. Understanding how the gene expression is controlled at the level of chromatin will therefore provide a deeper understanding of stem cell identity. Chromatin regions can be fundamentally distinguished into two types - heterochromatin and euchromatin. Heterochromatic regions are highly condensed and devoid of active gene transcription. A subclass is facultative heterochromatin that contains genes that are silenced after gene differentiation whereas constitutive heterochromatin in

contrast is gene-poor and genes irreversibly repressed. Genes in euchromatic regions are actively transcribed and chromatin compaction grade is low so that the transcription machinery can bind to these uncondensed gene promoters [7].

1.2.1. Composition and Structure

Chromatin is the entity of DNA and histone octamers around which the double strand DNA is wrapped for almost two turns. The histone octamers are composed of the four histone proteins H2A, H2B, H3 and H4, each of them in double copy. H2A and H2B as well as H3 and H4 form dimers (H2A/H2B and H3/H4) and two copies of each dimer type build up together the histone octamer. A nucleosome consists of one histone octamer with 147 base pairs of DNA wrapped around it. The nucleosomes themselves are connected through linker DNA and together they represent the chromatin fiber that can adopt higher condensed states. Histone H1 confers stability to the chromatin fiber by connecting nucleosomes and linker DNA [8] (Figure 2). Moreover histones can be removed and substituted by variants that fulfill specific functions. One example of such a variant is macroH2A that is present on repressed genes and increasing gene activation is accompanied by a gradual loss of macroH2A on these regions [9].

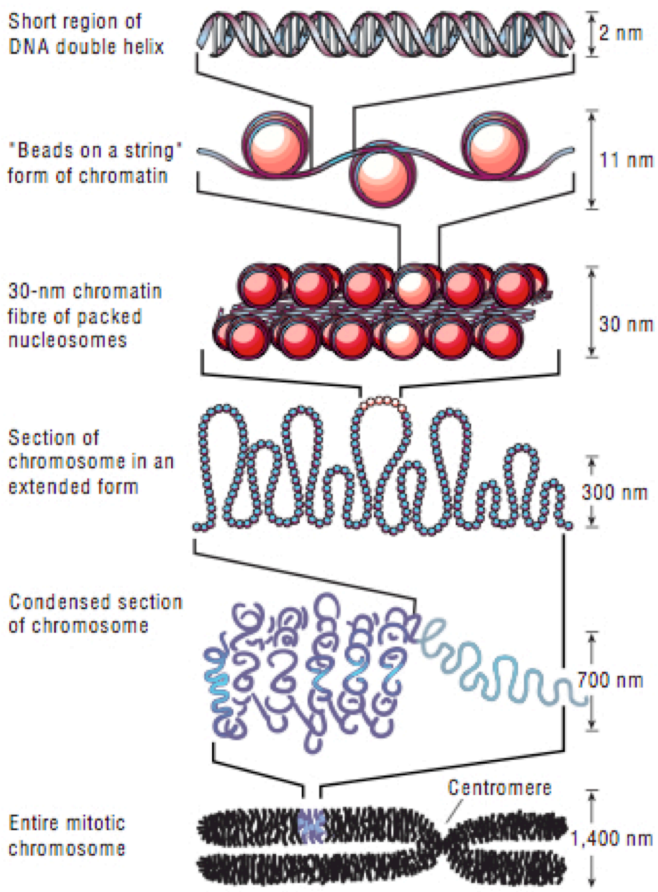


Figure 2. Chromatin organization. Nucleosomes are the structural entity of chromatin and are made up by DNA that is wound around the histone proteins. Linker DNA joins the nucleosomes and the resulting chromatin fiber further winds up and adapts higher condensation states. Figure adapted from [10].

1.2.2. Regulation of the chromatin state by posttranslational modifications (PTMs) of histones

The correct execution of many nuclear processes like DNA repair, gene expression and mitosis depends on the chromatin state [8]. The best-studied chromatin regulatory mechanism is the posttranslational modification (PTM) of the histone octamers. Histones possess a globular central structure with unstructured N-termini that protrude from the nucleosomes. Most of the modifications occur at the N-terminal tails but few are also present in the core histone. Until today more than 60 modification sites have been discovered but the actual number is estimated to be far higher. PTMs range from the attachment of small, uncharged groups up to big peptides that are covalently

bound to the histones. Different amino acids are modified by acetylation, methylation, phosphorylation, ubiquitination, sumoylation, ADP ribosylation, deimination, and proline isomerization (Table 1). PTMs have two possible outcomes on chromatin structure. On one hand they exert a direct effect on the interaction between DNA and histones or between histones of neighboring nucleosomes due to their charge and size. On the other hand they serve as docking sites for a whole plethora of chromatin associating proteins with specific binding domains [11]. PTM of one histone residue also impacts the modification states and effector binding of other residues either on the same or a proximate histone. Moreover different chromatin modifiers compete for the same residue. This phenomenon is called histone crosstalk and provides a further means of regulation [7].

Acetylation	**K**-ac
Methylation (lysines)	**K**-me1 **K**-me2 **K**-me3
Methylation (arginines)	**R**-me1 **R**-me2a **R**-me2s
Phosphorylation	**S**-ph **T**-ph
Ubiquitylation	**K**-ub
Sumoylation	**K**-su
ADP ribosylation	**E**-ar
Deimination	**R > Cit**
Proline Isomerization	**P-cis > P-trans**

Table 1. Posttranslational modifications (PTMs) of histones. Amino acids are modified by specific PTMs. Table adapted from [11].

Acetylation and Phosphorylation

Acetylation and phosphorylation are two PTMs that due to their negative charge neutralize the positive charge of histones and in consequence diminish the electrostatic attraction between negatively charged DNA and histones. Acetylation marks euchromatic regions were chromatin is in a decondensed state and easily accessible for the transcription machinery. Lysines are modified by addition of an acetyl group by histone acetyl-transferases (HATs) during gene activation. The deacetylation of histones is catalyzed by histone

deacetylases (HDACs) and linked to gene repression and the reestablishment of heterochromatic regions after transcription. Proteins bind to acetylated histones via their specific bromodomains [11]. Phosphorylation on serine, threonine and tyrosine is also reversible and dynamic and its levels are regulated by the interplay of the phosphorylating kinases and the dephosphorylating phosphatases. During gene activation histone H3 gets phosphorylated on S10. This mark is subsequently bound by domains present in members of the family of 14-3-3 proteins [7, 11].

Methylation

Lysines are mono-, di- and trimethylated and the methylation of different residues has opposing roles in transcriptional regulation as it is involved both in gene silencing as well as activation. Transcription initiation for example is characterized by trimethylation of lysine K4 on histone H3 (H3K4me3) and trimethylation of lysine K36 on histone H3 (H3K36me3) occurs during the elongation step. These methyl-marks therefore demarcate euchromatic regions in the genome. In contrast the trimethylation of histone H3 on lysine K27 (H3K27me3) and K9 (H3K9me3) play a role in gene repression [11]. H3K9me3 mainly occurs in constitutive heterochromatic regions like telomeres whereas high levels of H3K27me3 characterize regions of facultative heterochromatin that contain genes implicated in development and differentiation which get repressed in the course of differentiation [7].

Lysine residues are targeted by methyltransferases (HMTases) that are highly specific for a certain lysine. The corresponding enzymes with opposite function are the histone demethylases (HDMs). Arginine can be mono- and dimethylated and this mark is implicated like lysine methylation in gene activation and repression. The specific HMT are the protein arginine methyltransferases (PRMTs). The effect of monomethylated arginine can be abrogated by deiminating Rme1 to citrulline [11]. The only arginine specific HDM reported so far is JMJD6 that demethylates H3R2me2 and H4R3me2 [12]. In contrast to acetylation and phosphorylation methyl groups lack an overall charge but they regulate the chromatin state by serving as docking sites for proteins with

various specific methyl-binding domains like PHD domains and the chromo-like domains of the Royal family that comprises the tudor- and chromo- and MBT domains [11] (Figure 3).

Other post-translational modifications

Other modifications include ADP ribosylation or proline isomerization. In the case of sumoylation and ubiquitination a whole peptide is bound to a lysine residue [11].

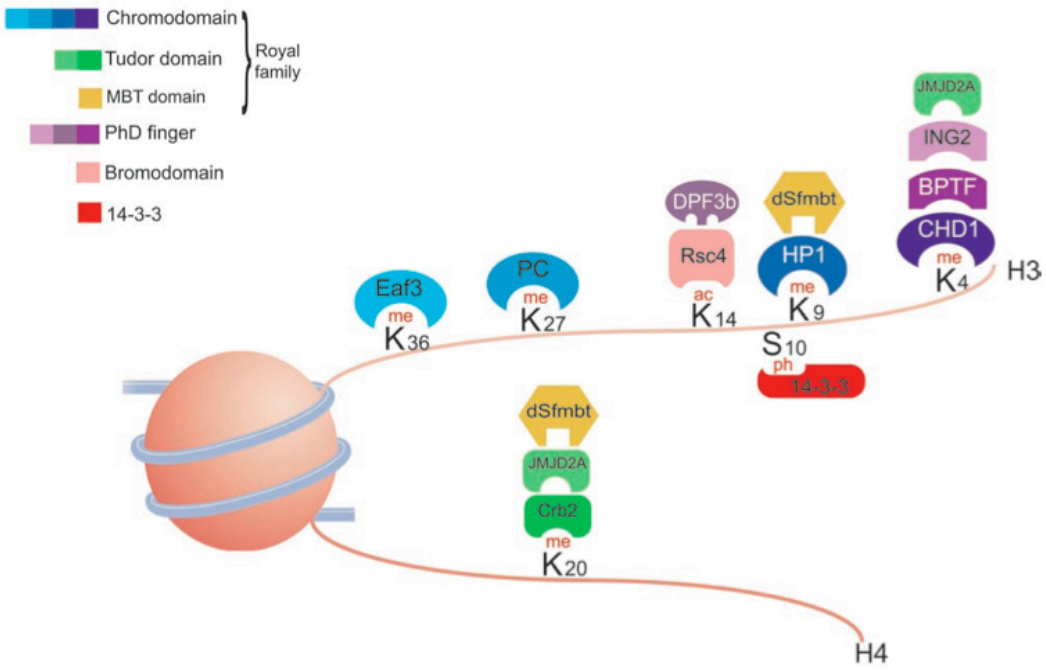


Figure 3. PTMs on histones are bound by specific protein domains. Figure from [7].

As mentioned previously PTMs can have a direct impact on chromatin conformation due to their size and charge. The second way to regulate chromatin-associated processes is by serving as docking sites for chromatin binding proteins that contain domains that specifically recognize the modification. These chromatin-binding proteins catalyze additional PTMs, recruit further regulatory proteins or remodel the chromatin structure. Chromatin remodelers are proteins that mobilize nucleosomes. The energy that is necessary for this process is obtained by hydrolysis of ATP. Multimeric protein complexes that are recruited by PTMs integrate several of these functions by

combining subunits with different domains [7]. An example of a modification-binding protein that recruits further chromatin modifiers is the yeast chromatin remodeling protein Chd1 that recognizes H3K4me3 by one of its two chromodomains. Moreover Chd1 is a subunit of the SAGA and SLIK complexes that contain the HAT Gcn5 and the histone deubiquitinase Ubp8. It was demonstrated that acetylation and activation of a gene promoter only takes place when functional Chd1 was expressed suggesting that Chd1 recruits SAGA and SLIK complexes to target genes marked by H3K4me3 [13]. Besides its role in transcriptional activation H3K4me3 is also involved in mediating gene repression. This is achieved by recruiting the repressive mSin3a-HDAC1 complex via the PHD domain of its ING2 subunit to H3K4m3 [14] and also the mammalian ING2 was shown to associate with the mammalian mSIN3A-HDAC complex [15]. Constitutive heterochromatin is characterized by H3K9me3 and HP1 proteins bind this modification by their chromodomains establishing in that way a binding platform for the HMT SUV39 that then methylates surrounding histones on H3K9. By this manner heterochromatic regions are established throughout the cell cycle or dynamically regulated [7].

1.2.3. Regulation of the chromatin state by DNA methylation

Apart from posttranscriptional modifications of histones DNA can also be modified by cytosine methylation. In general CpG poor regions with a low content of methylated cytosine can be distinguished from CpG islands that display at least 55% of methylated cytosine. CpG islands are frequent in gene promoter regions and their methylation represses gene expression. DNA methylation and transcriptional repression is directly linked to histone modification as methyl-binding proteins (MBPs) associate with the methylated cytosine and then recruit further repressors such as HDACs. Moreover HMT for H3K4 cannot bind to CpGs when cytosines are methylated [16].

1.3. Ubiquitination

1.3.1. Ubiquitin (Ub), Ubiquitin-like proteins (Ubls) and their conjugation mechanism

Ubiquitin (Ub) is composed of 76 amino acids and its covalent linkage to other proteins represents an important type of posttranslational modification. The ubiquitin mark is set by a sequential action of various enzymes (Figure 4) that display target protein specificity. The first enzyme of the cascade is the activating enzyme E1. E1 activates Ub by adenylating its C-terminus with ATP. When Ub is bound to AMP it builds a covalent thioester linkage with a cysteine residue of the E1. Ubiquitin is then transferred to the active center of the E2 conjugating enzyme that then form a complex with E3 ligases. The E3 ligases confer specificity to the reaction as they recognize the target proteins and for that reason hundreds of different E3 ligases are present in the cell. Once the target protein has been identified, the E2 enzyme catalyzes the isopeptide linkage between ubiquitin and the target. The isopeptide bond involves in the majority of cases the ε-amino group of a target lysine residue and the C-terminal glycine of ubiquitin. When a protein is monoubiquitinated, this moiety serves as starting point from which polyubiquitin chains originate by the continued action of E2/E3 or E2/E4 pairs. E4 enzymes act as E3 ligases but specifically elongate the already ubiquitinated proteins. Ubiquitination is a dynamic modification antagonized by a specific class of enzymes, the deubiquitinases (DUBs) [17]. Mass-spectrometric analyses of *S.cerevisiae* recently showed that all 7 lysine residues present in ubiquitin (K6, K11 K27, K29, K33, K48 and K63) give rise to polyubiquitin chains [18]. Polyubiquitin chains are called homotypic when the ubiquitin monomers are bound to each other by the same lysine. Linking ubiquitin moieties by a certain lysine on their surface specifically orientates the monomers to each other. In this way every homotypic chain has a characteristic three-dimensional structure [19]. A linear ubiquitin chain is generated when the C-terminus of ubiquitin forms a peptide bond with a N-terminal amino group of another ubiquitin [20]. In contrast to homotypic chains heterotypic chains that have more than one linkage type [19].

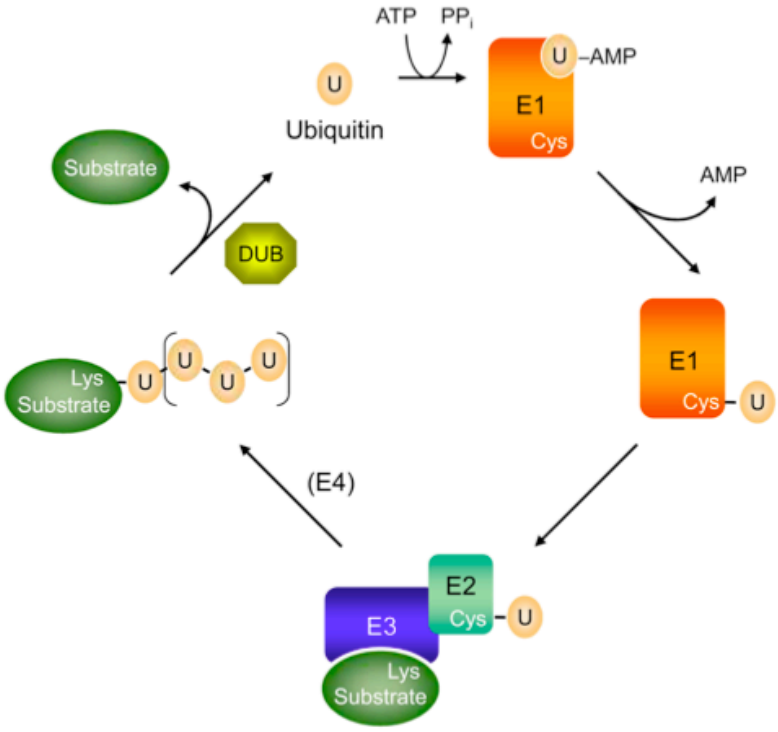


Figure 4. Ubiquitination and deubiquitination mechanisms. The modification of a lysine residue in a target protein depends on the sequential action of E1, E2 and E3 enzymes. The first step in the reaction is the activation of ubiquitin by ATP hydrolysis. E3 ligases recognize the target proteins. Monoubiquitin modifications are extended to polyubiquitin chains by E3 and E4 ligases. DUBs cleave ubiquitin from the protein and therefore ensure a dynamic equilibrium between ubiquitination and deubiquitination. Figure from [17].

The diversity in polyubiquitin chains creates several posttranslational modification subtypes that are recognized by specific binding modules, the ubiquitin-binding domains (UBD). Polyubiquitin chains are assigned to distinct cellular processes and thus have an impact on proteasomal protein degradation, cell cycle and DNA repair, amongst others. Various studies have shown that Lysine 48-linked polyubiquitin modifications plays a crucial role in marking proteins for proteasomal protein degradation. Lysine 63-linked polyubiquitin chains have two important functions that are not linked to that process [19]. One is the modification of proteins implied in signal transduction like the NF-κB pathway. This pathway depends on K63 polyubiquitination that is catalyzed by the RING domain of TRAF6 and TRAF2, signal transducer of the NF-κB pathway, and the E2 conjugating enzyme Ubc13 bound to the Ubc-like protein Uev1A [21]. Secondly, the K63 chains can modify histone H2A and its variant H2AX. The responsible E3 ligase in this case is RNF168 that binds to

chromatin upon DNA double strand breaks and leads to protein ubiquitination at their site of action [22].

Apart from ubiquitin, proteins are modified by so-called ubiquitin-like proteins (Ubls). Their name however is not due to their homology to ubiquitin in primary structure. It rather refers to their high resemblance in three-dimensional folding. The common structural unit is called the ubiquitin fold. Moreover Ubls are attached to their targets by analogue enzymatic mechanisms to that of ubiquitin conjugation. Two Ubls are SUMO (Small Ubiquitin-like Modifier) and ISG15 [23]. Sumoylation of transcription factors has been described to promote gene repression [24]. SUMO itself also constitutes a protein family in humans that consists of four members (SUMO 1-4) and the isoforms 1 to 3 have been described to modify different target proteins *in vivo*. Surprisingly the SUMO members 1-4 differ a lot in their conservation grade meaning that SUMO1 only shares about 50% identity with SUMO2 and 3, which in turn are highly identical to each other (97%) [25]. The Ubl ISG15 regulates immunity against mycobacterial diseases in humans. ISG15 is secreted by several cell types like leukocytes or granulocytes upon induction by mycobacteria and acts as an extracellular cytokine stimulating the production of interferon γ (INFγ) mainly in natural killer cells [26].

1.3.2. Ubiquitin Cycle

Ubiquitin is expressed either as a multimeric precursor or fused to ribosomal proteins. As described previously proteins can be posttranslationally modified by monoubiquitination or various types of polyubiquitination. Since ubiquitination is crucial to regulate many important processes in a cell there must be always a constant pool of free ubiquitin available. This is accomplished by maintaining the equilibrium between ubiquitination and deubiquitination. That means that the activity of E3 ligases and their counterparts, the DUBs, is strictly regulated. The illustration below (Figure 5) describes how free ubiquitin levels are maintained constant in the cell by the action of DUBs. In the part (a) of the figure is represented how ubiquitin monomers are liberated to the ubiquitin pool by DUBs that cleave the precursor molecules and fusion proteins. Part b) and

d) show the fate of a polyubiquitinated protein marked for degradation. There are two possible outcomes: first ubiquitin can be cut off and in that way the protein is saved from degradation (b). The second option is that the protein is broken down by the proteasomal proteases but DUBs previously cleave the polyubiquitin chain off from the target protein in order to impede the proteasomal degradation. Released polyubiquitin chains are subsequently recycled by DUBs to monomers (e). DUBs also act on proteins modified by ubiquitin chains that are not a signal for proteasomal degradation. In that case DUBs either remove the modification (c) or edit the chain by cleaving off ubiquitin monomers (f). Editing an ubiquitin mark is thought to modulate the function of the signal [27].

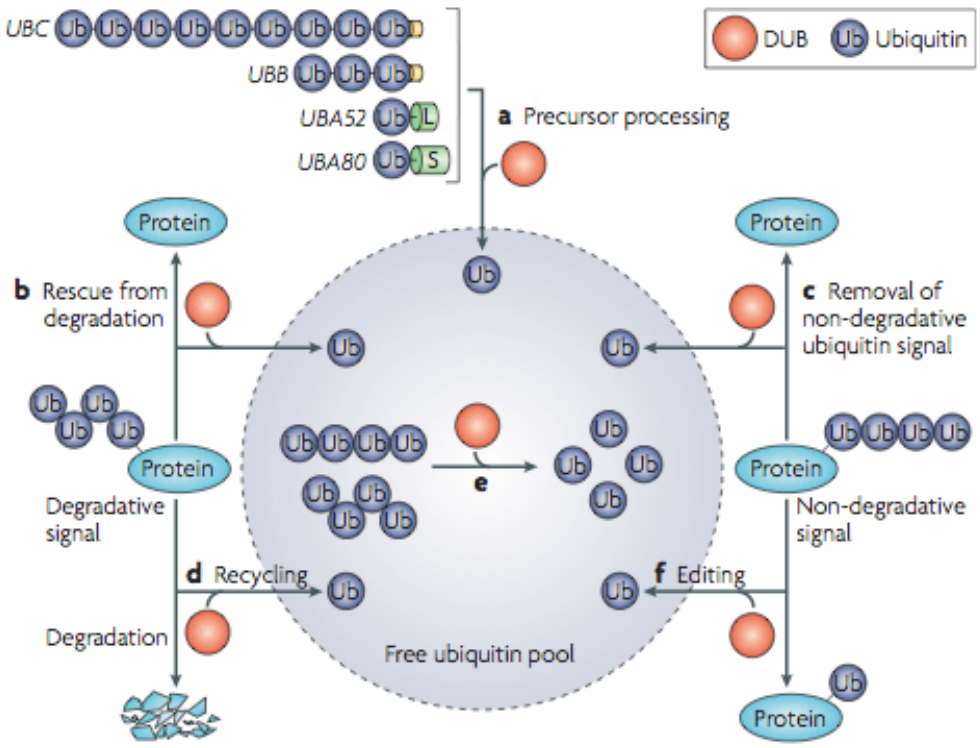


Figure 5. DUBs regulate constant cellular ubiquitin levels by reversing the ubiquitination of proteins. They act therefore as antagonists of the E3 ligases and thus maintain the pool of unconjugated ubiquitin in the cell. Deubiquitination occurs at different stages of the ubiquitin cycle as described in the text. Figure from [27].

1.3.3. Family of Deubiquitinases (DUBs)

The family of DUBs comprises proteases that cleave the isopeptide bond between the target protein and an ubiquitin moiety or within two ubiquitin moieties in a polyubiquitin chain. The huge majority of DUBs belongs to the subclass of cysteine proteases but some DUBs are members of the enzyme class of zinc metalloproteases. The cysteine proteases can be further subdivided into four more members: ubiquitin C-terminal hydrolases (UCHs), ubiquitin-specific proteases or ubiquitin-specific processing proteases (USP/UBP), ovarian tumor proteases (OTUs) and Machado-Joseph disease proteases (MJD). The group of the metalloproteases is named JAMM, the abbreviation of JAB1/MPN/Mov34 metalloenzymes. Crystallization of the catalytic domains of members of the five DUB subclasses revealed that there is a huge variety in secondary structure between the subclasses. However due to the conservation of the catalytic domains within one DUB subclass, all DUBs of the human genome could be detected by screening the ENSEMBL database for homologous proteins. In total 95 DUBs were identified, out of them 4 UCHs, 58 USPs, 14 OTUs, 5 MJDs and 14 JAMM metalloproteases USPs thus represent the largest subgroup. However there might only be 79 DUBs that are actually active as some of the retrieved candidates do either seem not to be expressed in the cell or their catalytic residues are not conserved [28].

1.3.4. Cellular processes regulated by DUBs

Various studies have shown that DUBS play an essential regulatory role in many cellular processes such as signaling pathways, cell cycle and proteasomal protein degradation. Important nuclear functions where DUBs are involved are transcription, mRNA export and DNA repair. However the mechanisms by which DUBs are regulated are still unknown in many cases. Knowing that DUB activity depends in many occasions on interaction with proteins that serve as scaffold proteins, activators of the enzymatic activity or substrates [29], Harper and collaborators identified the interactome of 75 DUBs by purifying the overexpressed enzymes from HEK 293 cells and identifying putative interacting proteins by mass-spectrometric analysis. In this way

functional contexts were assigned to DUBs whose physiological role had not been characterized till date. It turned out that DUBs are frequently part of multimeric complexes. Subsequent classification of the identified candidates according to their gene ontology (GO) revealed that DUBs could be implicated in very different cellular processes. One third of the DUBs have nuclear functions as suggested from its GO categories that include transcription and DNA repair as well as mRNA processing and rRNA transcription. In the cytoplasm DUBs participate in signal transduction, the ubiquitin cycle and mitosis, amongst others [30].

1.4. Histone Ubiquitination

Histones can be mono- and polyubiquitinated *in vivo*. All of the four core histones and also the linker histone H1 can be modified by ubiquitination. Out of those H2A and H2B are the ones that are most frequently modified and the best studied. Both histones are mainly monoubiquitinated and regulate transcription and DNA repair. However the abundance of the two marks differs. Whereas vertebrate chromatin displays between 5 and 15% of ubiquitinated H2A, H2Bub only represents 1 to 2% of total H2B. In contrast to vertebrates around 10% of H2B is modified in yeast. The modified residues are lysine 119 for H2A and lysine 120 and lysine 123 for H2B in vertebrates and yeast, respectively [31].

1.4.1. H2B monoubiquitination in transcription

H2B monoubiquitination by the E2 conjugating enzyme RAD6 in complex with the E3 ligase BRE1 occurs at the promoter regions of actively transcribed genes [32, 33]. It is involved in a trans-histone crosstalk with methylation of H3K4 and H3K79 [34, 35]. The interplay of ubiquitination and trimethylation in gene regulation is best understood in yeast.

The PAF complex regulates to a large proportion the interplay between ubiquitination and methylation [32]. This complex is bound to RNAPII. Apart from its subunit PAF1 it also contains further components like Cdc73p, Gal11p, CTR1, RTF1 and LEO1, as well as the general initiation factors TFIIB and TFIIF. It was demonstrated that the PAF complex has an impact on gene

expression [36, 37]. When PAF is bound to gene promoters it recruits the COMPASS complex that in turn can associate with the early elongating form of RNAPII and then trimethylates histone H3K4 by its SET1 subunit [38]. BRE1 recruits RAD6 to gene promoters where H2B is ubiquitinated independently from pre-bound PAF complex. However the ubiquitination activity of BRE1/RAD6 strongly depends on the PAF subunits RTF1 and PAF1. Moreover the PAF complex also mediates the binding of RAD6 to COMPASS as well as to RNAPII. PAF complex therefore integrates several mechanisms like H2Bub and H3K4me3 that are linked to transcriptional elongation [32].

In humans H2B is ubiquitinated on K120 by the RNF20/40 complex. Both proteins contain a RING domain and are homologous to each other. Their homologous yeast counterpart is the E3 ligase BRE1. The E2 conjugating enzyme UBCH6, the homologue to yeast RAD6, was identified to form a complex with the E3 dimer. In line with the mechanism described in yeast, the RNF20/40 dimer interacts with human PAF complex (hPAF) via the E2 enzyme UBCH6 and hPAF enhances the H2B ubiquitination activity in *in vitro* assays [39]. In accordance with the finding in yeast that H2Bub is linked to H3K4 and H3K79 methylation via the PAF complex [32] the experiments showed that ectopic expression of RNF20/40 leads to increased methylation of H3K4 and H3K79 in human. As mentioned previously the ubiquitination machinery localizes to H2B on transcribed genes in yeast. Chromatin immunoprecipitation (ChIP) analysis of the inducible gene MAGE-A1 in the human cells confirmed these findings by analyzing the promoter occupancy by the transcription machinery upon activation. The E2/E3 complex together with the PAF and RNAPII was recruited to the promoter and the transcribed part of the gene after induction. Interestingly this work also showed that H2Bub set by RNF20/RNF40 directly activates HOX gene expression [39].

Nevertheless H2B ubiquitination has been found as well associated to gene repression and the finding that H2Bub leads to chromatin compaction gives a possible explanation. The established model suggests that the activation of a gene rather relies on the dynamical ubiquitination and deubiquitination of H2B than on persisting H2Bub. According to the model suggesting the dynamic

changes in H2Bub during transcription the onset of gene activation would thus depend on H2B deubiquitination by the DUB Ubp8 that is integrated in the SAGA complex. That would lead to a loosening of the chromatin fiber, which would as a consequence allow the FACT complex to remove H2A/H2B dimers. New ubiquitination on the back of RNAPII by RNF20/40 could prevent erroneous cryptic transcription. The crosstalk of ubiquitination histone with H3K4 trimethylation in turn is explained by the preferable interaction of COMPASS with several histone residues when H2B is ubiquitinated [40].

1.4.2. H2A monoubiquitination in transcription

Identification of Polycomb repressive complexes 1 and 2 (PRC1 and PRC2)

Although H2A had already been found in 1975 to be monoubiquitinated [41], it took almost 30 years to identify the first E3 ligase that specifically catalyze the ubiquitination of nucleosomal H2A. Wang and collaborators [42] purified the E3 ligase RING1B of nuclear HeLa extract, which was further shown to be part of a multimeric complex. This complex contains the three subunits RING1A, BMI1 and PH2, all of which harbor a RING domain with exception of PH2 [42]. Due to its resemblance to the already described *Drosophila* PRC1 complex (Polycomb repressive complex 1) [43] the complex was named human Polycomb repressive complex 1-like (hPRC1L) [42]. The corresponding *Drosophila* homologues of RING1A and RING1B, BMI 1 and PH2 are RING, PSC and PH. The study of Francis and collaborators also identified the *Drosophila* PC protein as PRC1 subunit whose homologues are the mammalian CBX proteins.

All components of the PRC1 complex except PH2 belong to the class of polycomb-group proteins (PcGs) (Table 2). The physiological function of PcGs was first characterized in *Drosophila melanogaster* where the correct process of embryogenesis relies on the sequential expression of developmental regulators. These regulators are encoded by the homeobox genes (HOX), whose expression is tightly regulated during development by two opposing classes of proteins, the gene repressing PcGs and the trithorax-group proteins (trx-G) which ensure gene expression [44]. The conserved RING motif that is present in RING1A, RING1B and BMI1 is common to many E3 ligases [45]. Mutational

analysis of the RING domains of RING1A, RING1B B and BMI 1 revealed that only RING1B possesses deubiquitination activity towards monoubiquitinated H2A. Moreover H2Aub levels in RING1B knockdown cells dropped by 75%, pointing to RING1B importance in the global ubiquitination level of H2A in chromatin. It was further demonstrated that PRC1 mediates gene repression of the PcG target gene UBX and that the repression depends on RING1B and PC expression as well as the presence of H2Aub [42].

Drosophila[a]	Mouse[b]	Human[b]
PRC1:		
PC[c]	CBX2 (M33)	CBX2 (CDCA6, M33)
	CBX4 (MPC2, PC2)	CBX4 (hPC2, NBP16)
	CBX6	CBX6
	CBX7	CBX7
	CBX8 (PC3)	CBX8 (hPC3)
PH[d]	PHC1 (EDR1, MPH1, RAE28)	PHC1 (EDR1,HPH1,RAE28)
	PHC2 (EDR2, MPH2)	PHC2 (HPH2, HPH2)
	PHC3 (MPH3, EDR3)	PHC3 (HPH3, HPH3)
PSC[f]	BMI1 (PCGF4)	BMI1 (PCGF4, RNF51)
	MEL18 (PCGF2, RNF110, ZFP144)	MEL18 (PCGF2, RNF110)
	MBLR (PCGF6, RING6, RNF134)	MBLR (PCGF6, RING6,RNF134)
	NSPC1 (PCGF1)	NSPC1 (PCGF1, RNF68)
RING[g]	RING1 (RING1A)	RING1A (RING1, RNF1)
	RNF2 (RING1B)	RING1B (RING2, RNF2)
PRC2[h]:		
E(z)	EZH1	EZH1
	EZH2/ENX1/KMT6	EZH2/ENX1/KMT6
	EZH2 (ENX1, KMT6)	EZH2(ENX, KMT6)
Su(z)12[i]	SUZ12	SUZ12
ESC[j]	EED	EED

Table 2. PcG subunits of PRC1 and PRC2 complexes. PcGs of Drosophila PRC1 and PRC2 are listed and compared to their respective homologues in the two complexes in mouse and human. Figure adapted from [46].

Another PcG protein besides RING1B that possesses enzymatic activity is Enhancer of Zeste (EZH2) that contains a SET methyltransferase domain for which reason EZH2 was postulated to be involved in histone methylation. Purification of EZH2 showed that it represents an integral component of a multimeric complex that further contained the subunits ESC, EED, SUZ12, RBAP46 and RBAP48, all members of the class of PcG (Table 2). The complex was therefore named Polycomb Repressive Complex 2 (PRC2). As

suspected PRC2 showed methyltransferase activity towards histones based on an intact SET domain and methylated histone H3 preferentially on K27 but also on K9 [47]. This finding was supported by Cao and collaborators who showed in their assays that the PRC2 complex uniquely methylated H3K27 pointing towards K27 as the only EZH2 substrate. Apart from that H3K27me3 was necessary for the PC subunit to bind to the UBX gene [48]. PC has been shown to be involved in silencing UBX gene [49] and therefore PRC2 might act together with PRC1 in gene repression.

The interplay of PRC1 and PRC2 in gene repression

The question how PRC1 and PRC2 act together in gene repression is not completely clarified but the model suggests the sequential binding of PRC1 and PRC2 to common target genes. In Drosophila PcG target genes display specific sequences that are indispensable for Polycomb-mediated silencing, the so-called Polycomb response elements (PREs) [50]. PHO and PHO-like are two PcGs with DNA-binding property that specifically associate to those elements [51, 52]. The first step in the model is the binding of the PHO proteins to PREs. Once bound, they serve as recruiting proteins for the PRC2 subunits EZH2 and/or ESC. EZH2 subsequently trimethylates H3K27 and that PTM serves as a binding platform for the PRC1 component PC [53] that recognizes H3K27me3 by its chromodomain [47] which suggests that previous H3K27 trimethylation targets PRC1 to chromatin by its complex subunit PC. Recently, also the existence of the mammalian PRE-kr has been reported to regulate the expression of the mouse gene *MafB* and it was discovered as a highly conserved sequence between mouse and human that contains PHO binding sites. Chromatin immunoprecipitation experiments in the F9 mouse cell line revealed that PRC1 and PRC2 subunits are bound to the PRE when *MafB* is silenced [54]. Other ways to promote PRC2 binding to chromatin is via RNA molecules [55, 56] or association of proteins with DNA-binding domains [57].

RING1B containing complexes and further H2A E3 ligases

RING1B, the main H2A E3 ligase, is not exclusively present in the PRC1 complex. It has been described as well as the catalytic subunit of the RING1B-

Fbxl10 and the E2F6.com-1 complex (Figure 7). Purification of the transcription factor E2F6 identified a repressive complex that binds to E2F-, Myc-, and Brachury-binding elements and contains RING1B as well as the PcGs RING1, MBLR, h-l(3)mbt-like protein and YAF2. Further non-PcG subunits are the H3K9me3 binding protein HP1γ, two HMTs and the transcription factors MGA and MAX. Due to its complex composition and its sequence specificity it was predicted that E2F6.com-1 complex serves for stable repression of genes that respond to MYC and E2F [58]. In a proteomic study searching for interaction partners of RING1B it was shown that it interacts with the two HDMs LSD1/Aof2 (40) and FBXL10/Jhdm1B amongst other protein partners. The subsequent purification of FBXL10/Jhdm1B identified the overlap of the two interactomes and identified amongst others the corepressor BCoR as a putative subunit of a possible new RING1B/FBXL10 containing complex [59].

Apart from distinct RING1B containing complexes there are also different H2A specific E3 ligases. After RING1B, 2A-HUB is the second E3-ligase that is essential for gene repression. 2A-HUB is part of a N-CoR, HDAC1 and HDAC3 containing complex that regulates chemokine genes in which gene repression depends on the monoubiquitination activity of 2A-HUB (Figure 7) [60]. The third E3 ligase is BRCA1 that ubiquitinates H2A in silenced satellite repeats and is necessary for the establishment of heterochromatin. However when the catalytic activity of BRCA1 is disrupted the satellite repeats start getting transcribed, a phenomenon thought to be linked to genomic instability [61].

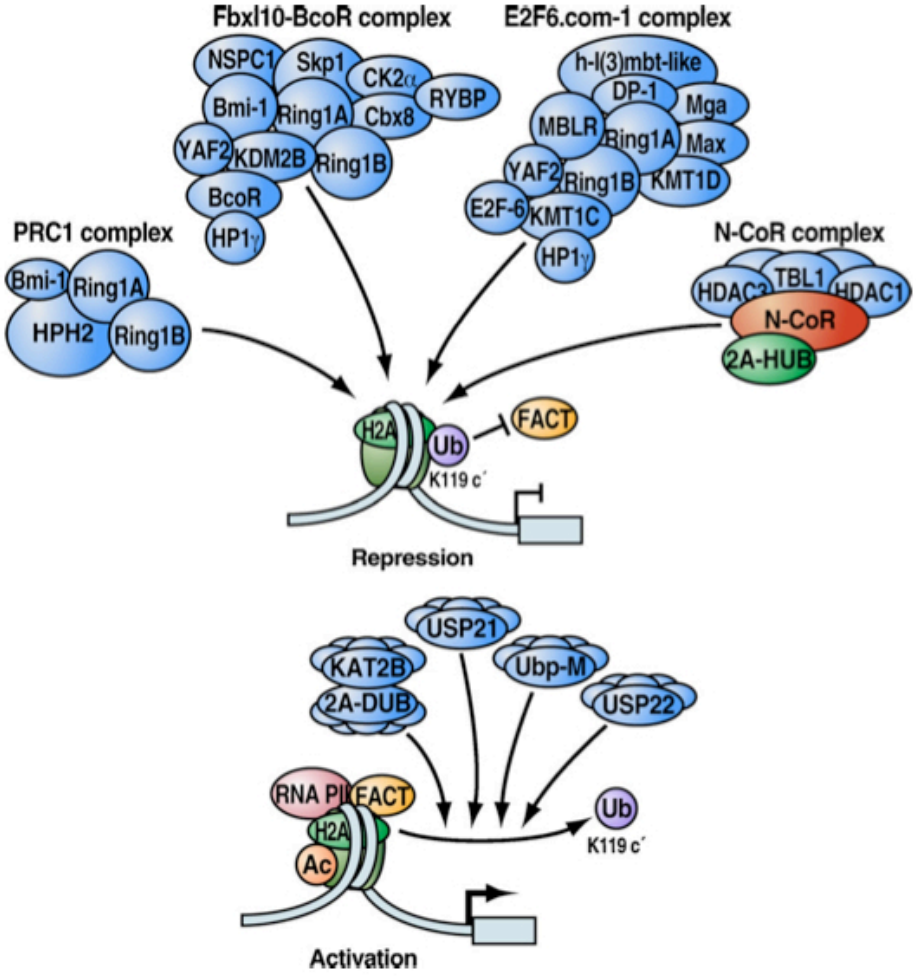


Figure 7. Role of H2Aub in gene regulation. RING1B and 2A-HUB act in multimeric complexes and mediate gene repression by ubiquitinating H2A. Removal of the ubiquitin mark by H2A specific DUBs leads to the reactivation of genes [62].

Mechanisms of PRC mediated gene repression

Until today the exact mechanisms how gene repression is achieved by PRC complexes are unclear. The *Drosophila* PRC1 complex was first reported to prevent the action of SWI/SNF chromatin remodeling complex [63]. The mechanism how ubiquitinated H2A mediates gene repression is still not fully understood and controversial. It is known that the ubiquitin mark can interfere with both the initiation and elongation step of RNAPII. Di- and trimethylation of H3K4 plays a decisive role in the initiation process by the preinitiation complex but *in vitro* transcription assays showed that no methylation could take place when H2A was ubiquitinated. As a consequence transcription was not initiated. The reason for this might be the impairment of the preinitiation complex

formation because transcription took place when the same experiment was conducted with a pre-established preinitiation complex. These findings point to a role of H2Aub in preventing the initiation step by impairing the preinitiation complex assembly (Figure 7) but not the elongation [64]. A study conducted in embryonic stem cells (ESC) however suggested the role of H2Aub in promoting elongation. ES cell promoters of genes participating in differentiation are silenced yet showing a bivalent modification state of activating and repressive histone marks in their promoters. These marks are the repressive H3K27me3, set by PRC2, and the activating H3K4me3 mark, set by HMTases. RNAPII at bivalent promoters is in a poised conformation what means that its elongation capacity is very low although it displays phosphorylation at serine 5, which is characteristic for RNAPII before it starts transcription. For this reason only very low transcript levels are observed for bivalent genes in mESC. Interestingly this poised conformation depends on the H2Aub levels set by RING1B. When changes in the ubiquitination level occur RNAPII rapidly changes its conformation and genes are derepressed [65].

RING1B moreover represses HOX gene expression by compacting chromatin in mESC and this function is independent of its E3 ligase activity for H2A. FISH analysis in mESC reported the decondensation and enhanced transcription of HOX gene loci in cells depleted for RING1B (RING1B −/−). Surprisingly when RING1B −/− cells were rescued with the catalytic mutant RING1B (I53A) that fails to monoubiquitinate H2A but integrates into PRC1 complex, the condensation and gene repression were restored to the same extent as in the rescue experiment with RING1B wildtype. Gene repression therefore rather depends on the chromatin compaction ability of RING1B than on its E3 ligase activity [66].

Further studies however contradicted the finding that gene repression is completely independent of RING1B catalytic activity. A study in mESC depleted for both RING1A and RING1B indeed confirmed that RING1B induces compaction of HOX gene clusters independently of its E3 ligase activity. ChIP-on chip analysis moreover identified genes that display H2Aub in their promoter region. Genes that displayed H3K27me3 and RING1B additionally to H2Aub

were particularly represented among the Polycomb genes. Analyzing Polycomb target gene expression levels showed the unexpected result that for their complete repression RING1B activity was indispensable. Those results suggest that the repressive function of RING1B only partially depends on its ability to locally condense chromatin but that a functional RING domain is indispensable for achieving the full repression level [67].

The induction of PcG genes upon differentiation implies the removal of PRC complexes. ZRF1 fulfills this function by binding to H2Aub. ZRF1 recruitment to chromatin goes together with a displacement of PRC1 complex from chromatin. The PRC repressed HOX genes are activated when cells are differentiated. Differentiation induction by retinoic acid was followed by ZRF1 recruitment to HOX gene promoter. This led to a concomitant loss of RING1B occupancy [68].

The next step in gene activation would then be the cleavage of ubiquitin so that RNAPII can adopt its active elongating confirmation (Figure 7). There are several deubiquitinases that have been found to be H2A specific and to activate transcription. However only for DUBs, USP16 and PR-DUB, it has been described so far that they actually act in concert with PcGs and regulate HOX gene expression [69, 70]. The compacted state of chromatin also acts as a barrier for the proceeding of RNAPII during elongation. The FACT complex removes that barrier by taking out H2A/H2B dimers and for this reason FACT is a possible co-player in gene activation. ChIP analysis in 2A-HUB depleted cells demonstrated that upon loss of 2A-HUB and H2Aub more FACT and elongating RNAPII were indeed detectable in exonic regions (Figure 7) [60]. Histone H1, the linker histone that compacts nucleosomes, is displaced upon phosphorylation [71]. Interestingly chromatin that contains a H2A mutant that cannot be ubiquitinated also displayed a low content of H1 indicating that genes could be expressed when H2A is not ubiquitinated. In contrary when cells were depleted for the H2A specific DUB 2A-DUB, the levels of H2Aub raised and at the same time the amount of phosphorylated H1 that is prone for later dissociation from chromatin decreased [72].

1.4.3. DUBs and H2A specific deubiquitination in transcription

As already mentioned a subclass of DUBs cleave ubiquitin from monoubiquitinated H2A. Those DUBs are amongst others USP16, USP21, USP22 and 2A-DUB, which are all members of the cysteine proteases with the exception of the MPN+/JAMM metalloprotease 2A-DUB [73]. Table 3 shows a summary of DUBs that deubiquitinate H2Aub in the context of transcriptional regulation.

Gene Symbol	Substrate	subclass	Nuclear Process
USP16/UbpM	H2A	USP	Activation of HoxD10 gene; G2/M transition
USP21	H2A	USP	Transcription
USP22	H2A, H2B	USP	Transcriptional coactivator with Myc, AR, ER, GR; G1/S transition
2A-DUB	H2A	JAMM/MPN+	Transcriptional coactivator with AR
BAP1	H2A	UCH	Repression of PcG gene

Table 3. H2A specific DUBs in transcription. DUBs are listed indicating their target specificity, enzymatic subclass as well as the transcriptional process they are implicated in. Table adapted from [73].

USP16 (or Ubp-M) was first reported to deubiquitinate H2Aub in chromatin extracts in 1999 in a study of Cao and coworkers. *In vitro* deubiquitination assays revealed that USP16 only deubiquitinates H2Aub when the histone is integrated into nucleosomes. Interestingly the expression of homeobox gene HOXD10 was dependent on the catalytic activity of USP16. ChIP experiments in HeLa cells demonstrated that USP16 localizes to the gene promoter and the 5' regulatory region of HOXD10 and that those regions are bound as well by the PRC1 subunit BMI1. Depletion of USP16 led to an increase of the ubiquitin mark at both regions. USP16 could be even assigned a role in *Xenopus* embryogenesis by regulating HOX gene expression because when USP16 was inactivated in *Xenopus* embryos by injection of antibodies against USP16 homeotic transformation was observed [69].

BAP1 is another DUB that is present in a multiprotein complex whose activity has been demonstrated to be necessary for high gene expression levels [74]. The corresponding homologue in *Drosophila* is Calypso, a PcG protein. Calypso associates with ASX, another PcG protein. The interaction is conserved between the human homologues BAP1 and ASXL1. The complex formed by the two interactors is denominated PR-DUB, Polycomb repressive deubiquitinase, and specifically deubiquitinates nucleosomal H2Aub. The enzymatic activity depends on the presence of ASX. Moreover PR-DUB colocalizes with PRC1 subunits on PREs in *Drosophila*. In contrast to the common model that H2A monoubiquitin represses gene transcription, the deubiquitination activity of PR-DUB though has been demonstrated to be required for HOX gene repression. This finding raises new possible mechanistic functions of H2A deubiquitination. Ubiquitination levels at promoters could be subjected for example to cyclic changes during transcription as it the case for H2B. Besides PR-DUB could fulfill the task of fine-tuning the ubiquitination density at promoters. Both options are conceivable to explain how gene regulation is achieved [70].

In addition DUBs are also implicated in activation of target genes other than HOX genes. One of those is their function as coactivator of nuclear receptors. 2A-DUB, a MPN+/JAMM metalloprotease, interacts with the androgen receptor (AR) and strongly enhances AR target gene expression. ChIP analysis demonstrated that the loss of H2Aub upon 2A-DUB recruitment after gene induction is accompanied by a loss of H1 at target gene promoter. Interestingly 2A-DUB deubiquitination activity depends to a significant proportion on the binding of the histone acetyltransferase p/CAF and its preceding acetylation of nucleosomes. In order to delimit the role of 2A-DUB in the RNAPII action upon gene induction, ChIP assays were performed on both promoter and exonic regions and demonstrated the putative involvement of the deubiquitinase in the efficient progress of both initiation and elongation [72].

USP22 which deubiquitinates both H2Aub and H2Bub, has been also observed to act as a coactivator of androgen receptor. USP22 is an integral part of the TFTC/STAGA complex and, together with the two further subunits ATXN7L3,

and ENY2, enhances androgen target gene expression *in vivo*. This activation depends on a functional catalytic center [75].

1.4.4. USP21

Gong and colleagues identified USP21 as a new deubiquitinase. Aiming at the identification of new human members of that class of proteases, they looked for candidates containing the Cys- and His-boxes that are conserved in USPs. The Cys and His boxes form part of the catalytic center [76]. In their search they identified USP21 (Ubiquitin Specific Peptidase 21) as a so far unknown USP. USP21 exists as two isoforms (Figure 8). Short isoform USP21 (AF233442_1) has 381aa and a molecular weight of 43,031kDa [77]. The full-length USP21 (NP_001014443.1) measures 565aa and has a molecular weight of 62,7kDa (ExPASy pI/Mw tool). The first 212 amino acids of full length USP21 constitute the N-terminus. This part of the protein is unstructured. The remaining 370 amino acids form the C-terminus that contains the conserved catalytic ubiquitin-hydrolase domain [78].

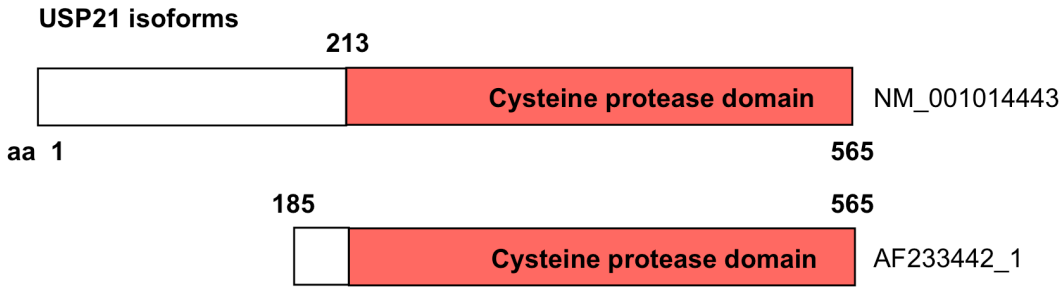


Figure 8. USP21 isoforms. Both USP21 isoforms contain the cysteine protease domain and are catalytically active.

Regarding the subcellular localization of the protein, there are several reports assigning it as a predominantly cytoplasmic localization [30, 79, 80]. More recently, a nuclear role for USP21 has been described [64]. The regulation of subcellular localization of a protein is achieved by nucleocytoplasmic shuttling. Transport receptors mediate the transport by carrying the cargo proteins

through the nuclear core complex. Cargo proteins contain consensus motifs, so called nuclear export and nuclear localization signals (NES and NLS) through which they interact with the receptors. USP21 contains a NES in its N-terminus (aa134-152) by which the transport receptor CRM1 exports USP21 to the cytoplasm [80].

1.4.4.1. Enzymatic specificity

The enzymatic specificity has been studied in detail for USP21. *In vitro* assays performed with the recombinant protease domain revealed that USP21 can cleave K6-, K11-, K29-, K48-, K63- and linear Ubiquitin dimers as well as K11-, K48-, K63- and linear tetra-Ubiquitin chains. Further it acts as both an endo- and exo-DUB as it contains a secondary ubiquitin-binding site on its surface. Regarding its selectivity for ubiquitin it was observed that it also processes the ubiquitin-like protein ISG15, however with lower activity than for ubiquitin [78].

1.4.4.2. Cytoplasmic function

In the study of Sowa and colleagues the interactome of 75 DUBs was detected by mass spectrometric identification of putative interactors of the purified DUBs. Data analysis with the Proteomic Analysis Software Suite (*CompPASS*) allowed determination of high-confidence candidate interacting proteins (HCIPs) of USP21 (Figure 9). Moreover following GO analysis of the HCIPs assigned USP21 by GO analysis a putative cytoplasmic function. More precisely its possible involvement in two cell signaling pathways that depend on protein phosphorylation was predicted. The first pathway is the DNA damage response that is regulated by the ATM kinase. The ATM kinase in turn is a substrate of the protein phosphatase PPP6C whose subunits bind to USP21. Besides these, interactions were identified with ubiquitin-associated (UBA) domain containing MAP/microtubule affinity-regulating kinase family (MARK1, 2, 3 and 4), suggesting also a role in microtubule-associated processes [30]. This finding was recently confirmed by the report of the direct binding of USP21 to microtubules and centrosomes and its participation in the correct formation of microtubule networks [79]. A part from this USP21 also inhibits NF-κB signaling,

a pathway that regulates the expression of survival factors and cytokines which are activated when the TNF receptor 1 (TNFR1) is bound by its ligand tumor necrosis factor α (TNFα). The activation of downstream components of the pathway that finally activate the NF-κB transcription factor depends on the recruitment of several proteins to the activated TNFR1, such as receptor-interacting protein 1 (RIP1) and TNF receptor-associated factor 2 (TRAF2). RIP1 is modified with ubiquitin K63 chains by TRAF2 and that PTM is necessary for the signal transduction. Deubiquitination of RIP1 by USP21 therefore inhibits NF-κB signaling and target gene expression {Xu, 2009 #7}.

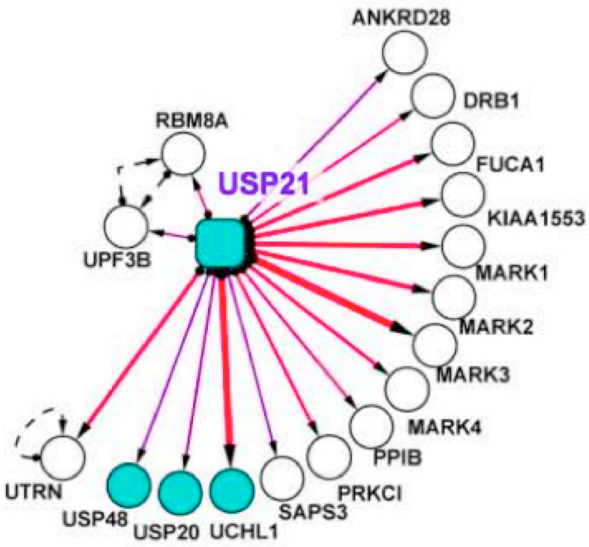


Figure 9. Topological map of USP21. Network of identified HCIPs for USP21 is shown as well as interactions among the candidate proteins [30].

1.4.4.3. Nuclear function

USP21 acts as a H2A specific DUB in the nucleus. This was discovered in a study monitoring changes in H2Aub in relation to gene expression in hepatectomy. Hepatotectomy is a convenient model system for studying the relation between gene expression and H2A monoubiquitination: during liver regeneration changes in expression of regulatory genes occur concomitantly with changes in H2A ubiquitination levels. Upon partial hepatectomy, microarray

analysis revealed up-regulation of two DUBs, USP21 and USP4, in the course of hepatocyte proliferation. Of those two, only USP21 was able to deubiquitinate chromatin. Moreover its specificity for H2A was shown. Remarkably it could be demonstrated that the long USP21 isoform deubiquitinates nucleosomes whereas the short protein is only active on free histones. This observation suggests a role of the N-terminus in directing substrate specificity. A further finding of that study was that USP21 interferes with the previously elucidated interplay between H2Aub and H3K4me3 (Figure 10). That trans-histone crosstalk regulates transcriptional initiation by impairing H3K4me3 when H2A is ubiquitinated. It was seen that methyltransferase MLL3 couldn't methylate chromatin containing the modified H2Aub in comparison to non-ubiquitinated nucleosomes. USP21 relieved the repression caused by the presence of H2Aub [64].

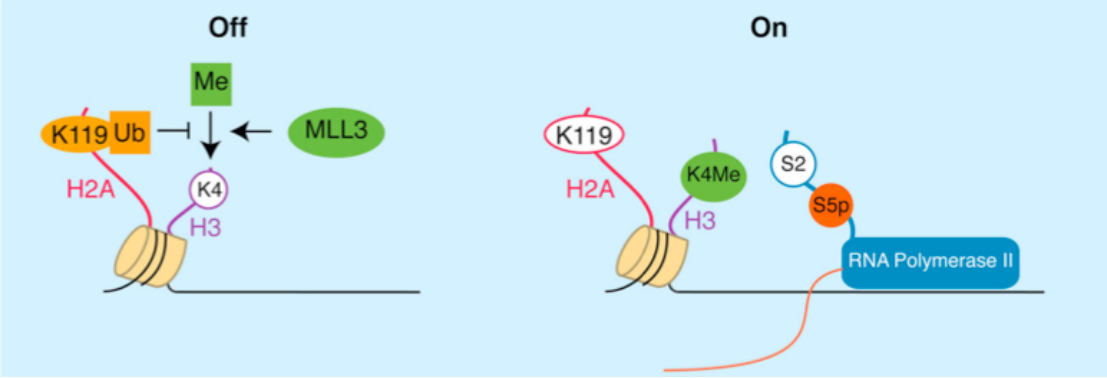


Figure 10. Crosstalk between H3K4me3 and H2Aub in transcriptional repression. RNAPII initiates transcription in vitro when H3K4 is trimethylated. However the HMT MML3 only methylates H3K4 when H2A is not ubiquitinated. Previous deubiquitination of H2Aub by USP21 enables MLL3 to set the methyl mark that is necessary for transcription [73].

1.5. Objectives

The regulation of USP21 catalytic activity is unknown however substrates are divers [30, 64, 81] suggesting that enzymatic activity is tightly regulated. Regulation of DUB specificity is accomplished by the interaction with regulatory proteins that direct protein localization and impact on the catalysis [29].

The present work therefore focuses on:

Identification of the nuclear USP21-associated interactome.

Validation and characterization of putative USP21 binding partners.

Analysis of USP21 deubiquitinating activity as a function of the identified interactors focusing on the ubiquitination state of H2Aub and binding partners.

Determination of the function of USP21 and its putative interactors in mESC.

2. Results

2.1. USP21 and its impact on global ubiquitination level

Many of the H2A-specific Deubiquitinases identified so far show an effect on global H2A monoubiquitination levels. USP16 knockdown in HeLa cells leads to higher levels of endogenous monoubiquitinated histone H2A [69]. The metalloprotease 2A-DUB has been proven to regulate the overall ubiquitination levels in HEK293T cells. While the H2A ubiquitination levels increase in 2A-DUB depleted cells, the effect is rescued by the overexpression of the DUB but not of its catalytic mutant [72]. HeLa cells depleted for the ATXN7L3 subunit of the SAGA complex resulted in an increase of H2Bub and to a lesser extent also of H2Aub [82]. Moreover the *Drosophila* PR-DUB regulates modification levels of bulk H2A in *Drosophila* embryos [70]. As a first step we sought to know whether USP21 overexpression has a repercussion on the overall H2Aub observed in chromatin. For that purpose we overexpressed Flag-USP21 in 293T cells and fractionated the cells into cytoplasm, soluble nuclear extract and chromatin fraction. Western Blot analysis of the chromatin fraction revealed that H2A monoubiquitination levels gradually decreased with increasing amounts of transfected Flag-USP21. That means that USP21 regulates the overall amount of ubiquitinated H2A in 293T cells (Figure 1).

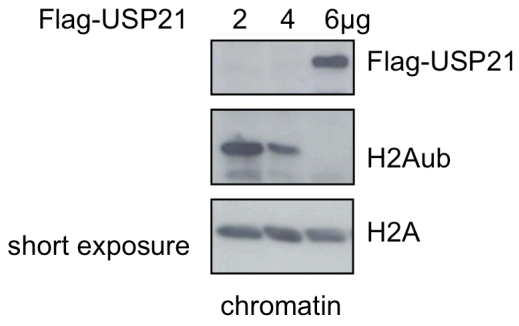


Figure 1. USP21 regulates global H2A monoubiquitination levels in 293T cells. The same number of 293T cells was seeded and transiently transfected with 2, 4 and 6µg of Flag-USP21 encoding plasmid and harvested 48h after transfection. Cells were fractionated into cytoplasm, nucleoplasm and chromatin. Chromatin was resuspended in Laemmli sample buffer and solubilized by sonication. Equal volumes of chromatin lysates were separated in SDS-PAGE and transferred to nitrocellulose membrane by Western Blot. The membrane was probed with Flag antibody to monitor USP21 overexpression and H2A antibody to detect the level of unmodified H2A and as well as monoubiquitinated H2A that shows a shift in SDS-PAGE due to the ubiquitin moiety. Unmodified H2A levels served as loading control.

In order to create tools to study the deubiquitination activity of USP21, we cloned a catalytic inactive USP21 mutant. The enzymatic activity of USP21 depends on three key amino acids in its protease domain. Those are called the catalytic triad and consist of cysteine 221, histidine 518 and aspartate 534. Mutation of the three amino acids to an alanine or asparagine leads to the loss of USP21 activity as demonstrated by *in vitro* experiments [64]. We cloned the described Flag-USP21 C221A (CA) and Flag-USP21 H518A (HA) mutants [64] in which the catalytic residues are substituted by alanine. In order to confirm the inactivity of the Flag-USP21 mutants, we overexpressed mutant and wildtype constructs in 293T cells and analyzed the level of H2A monoubiquitination in chromatin fraction (Figure 2). Chromatin of cells in which Flag-USP21 WT was expressed displayed only residual ubiquitinated H2A. In contrast, when the CA and HA mutants were expressed, H2A ubiquitination was still detectable and comparable to non-transfected control cells. This confirms that the amino acid substitution rendered the enzyme inactive.

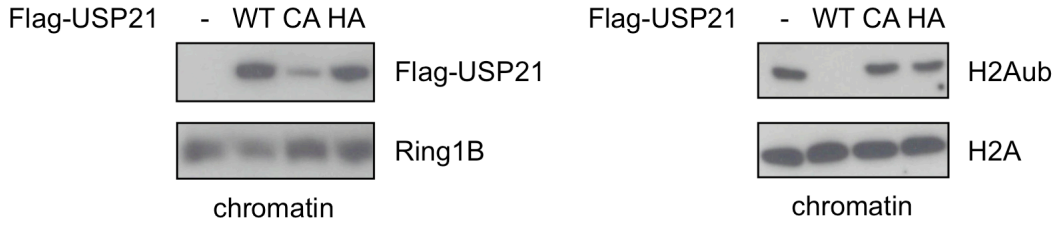


Figure 2. Flag-USP21 C221A (CA) and Flag-USP21 H518A (HA) mutants are catalytically inactive. 293T cells were transiently transfected with vector encoding Flag-USP21 or the mutants Flag-USP21 C221A or Flag-USP21 H518A and equal cell numbers were harvested 48h after transfection. Cells were fractionated into cytoplasm, nucleoplasm and chromatin. Chromatin was resuspended in equal volumes of Laemmmli loading buffer and solubilized by sonication. Chromatin lysates were separated by SDS-PAGE and transferred by Western Blot to a nitrocellulose membrane. Tubulin, RING1B and Histone H2A antibodies controlled equal loading and fractionation efficiency. Flag-USP21 constructs were detected by Flag antibody and H2Aub and H2A were detected with the H2A specific antibody.

2.2. Purification of nuclear USP21

As described in the introduction (chapter 3.4.), deubiquitinating enzymes are regulated in many cases by interaction with proteins that serve as allosteric activators, adaptors and scaffolds. We were especially interested in

understanding the function of USP21. *In vitro* assays had shown its implication in the transcription initiation process. The initiation is regulated by a trans-histone crosstalk between H3K4me3 and H2Aub. Monoubiquitinated H2A prevents the trimethylation of H3K4 by the HMT MLL3 but trimethylation of H3K4 is necessary for the onset of transcription. USP21 abolishes the repression of H3K4 methylation as it specifically deubiquitinates H2Aub [64]. In order to elucidate the role of USP21 in that crosstalk *in vivo* we decided to purify nuclear USP21. Mass spectrometry analysis of its interacting proteins would reveal what classes of proteins are bound to USP21 in the nucleus and how its enzymatic activity could be regulated. Moreover we aimed to understand how a putative USP21 containing complex could be implicated in the H2Aub and H3K4me3 trans-histone crosstalk *in vivo* and which target genes it regulates.

2.2.1. TAP-tag purification of USP21

Nuclear USP21 was purified by Tandem Affinity Purification (TAP) and putative interacting proteins were identified by mass spectrometry. The TAP-tag purification method is a two-step affinity purification in which the bait protein is fused to a double tag consisting of Protein A and the calmodulin binding peptide (CBP). Both tags are separated by the TEV protease cleavage site that is a site-specific cysteine protease. The bait protein is first purified with an affinity column of IgG coated beads. The eluted bait protein is then cleaved by TEV protease so that the CBP tag gets exposed. Purification of the eluate with the second column that contains CBP peptide retains the bait protein that is later eluted by EGTA. The complex is separated in a silver gel and the visible bands containing complex subunits are excised and subjected to mass spectrometry analysis. The advantage of this purification method is that by coupling two different affinity columns a higher degree of purification is achieved. However the amount of non-specific interacting proteins depends always to a high degree on the overexpression levels of the bait. For that reason, we chose an ecdysone-inducible mammalian expression system in which the bait protein is expressed at low levels. This system uses the EcR-293 cells that stably express

the retinoid X receptor (RXR). RXR acts as a transcriptional activator when bound to its ligand ponasterone A, a member of the ecdysteroid family of insect steroid hormones. We stably transfected a pMZI vector encoding the TAP-tagged USP21 protein into the EcR-293 cells and established stable clones expressing RXR and TAP-USP21. The pMZI vector contains the HSP promoter with a ponasterone inducible element. TAP-USP21 expression can thus be activated by RXR bound to its ligand and further the HSP promoter allows protein expression at low levels. Clonal stable cell lines were established and clone #7 was chosen for purification. Western Blot analysis of TAP-USP21 induction after ponasterone A treatment showed that protein expression was highly induced. However also the non-treated control cells showed TAP-USP21 expression at low levels which means that the system is slightly leaky (Figure 3a). Nevertheless clone #7 was utilized for purification, as the expression of TAP-USP21 upon induction was significantly higher than the non-specific background expression. The expression level could not be compared to that of endogenous USP21, as there are no suitable antibodies for endogenous protein available. Induced and non-induced EcR TAP-USP21 clone #7 cells were fractionated in cytoplasm, nucleoplasm and chromatin and nucleoplasm was subsequently used for purification (Figure 3b). In total two independent purifications of TAP-USP21 were performed.

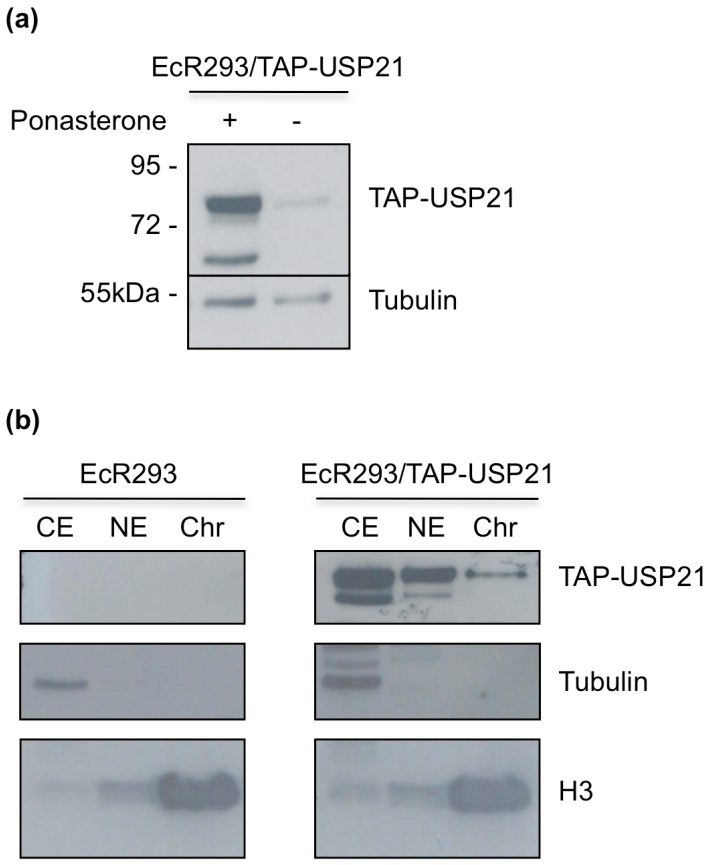


Figure 3. Purification of proteins associated to nuclear USP21. (a) TAP-USP21 expression in EcR293 cells stably transfected with TAP-USP21 plasmid was induced by ponasterone A at 3μM for 48h. Expression of TAP-USP21 was monitored by Western Blot analysis with CBP antibody. Non-induced cells served as control. (b) Fractionation of EcR293/TAP-USP21 cells used for purification. Ponasterone A induced (78h) and uninduced control cells were fractionated into cytoplasm, nucleoplasm and chromatin fraction and the same volumes of the fractions were separated by SDS-PAGE. Fractionation efficiency was checked by Western Blot using Tubulin and H3 antibody. TAP-USP21 expression was detected by CBP immunostaining.

We checked purification efficiency by analyzing the flow-through in Western Blot (Figure 4). Not all TAP-USP21 was depleted by purification, as it was still present in the flow. However it was successfully eluted from the beads as the analyzed supernatant of beads boiled after elution showed only little USP21 in Western Blot. The complexes were separated in a silver-stained gel (Figure 4). Protein bands that were present in the eluate of USP21 expressing cells were excised and subjected to mass spectrometry.

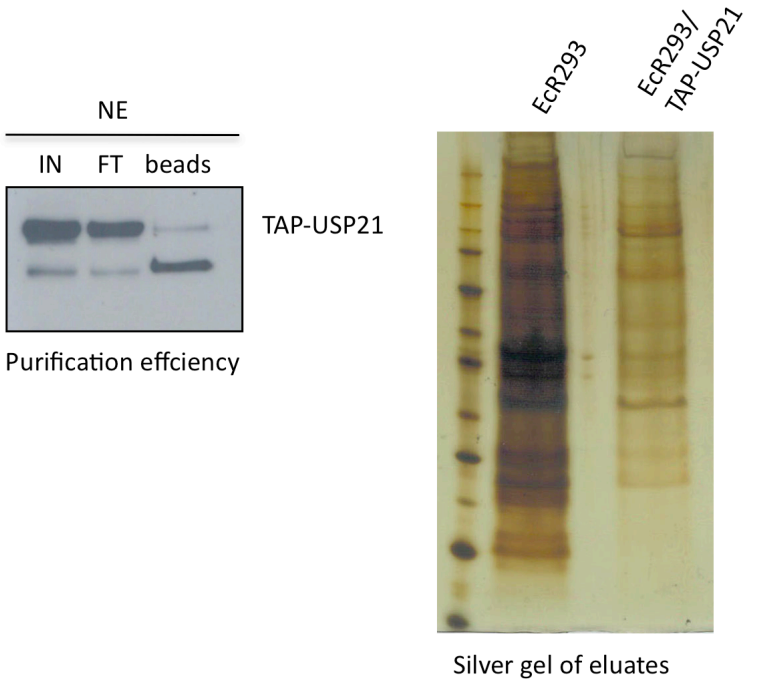


Figure 4. TAP-USP21 affinity purification of nucleoplasm. Depletion of TAP-USP21 of nucleoplasm was analyzed by Western Blot analysis of input material and the flow-through after purification. Calmodulin beads were boiled after complex elution and the obtained supernatant was analyzed as well in order to detect not eluted complexes. Eluates of the purification were separated in a silver gel. Bands were excised and proteins analyzed by mass spectrometry.

Both purifications yielded BEND3 as the main USP21 interacting protein with an outstandingly high score (Table 1). Many of the other candidate interactors that were identified only in a single purification have either an activating or repressive function in gene expression and can be subunits of multimeric protein complexes. Among those was for example BCoR, a co-repressor of the transcriptional repressor BCL-6. BCoR interacts with further histone modifiers like HDACs [83] and RING1B and has been shown to recruit a multisubunit complex to BCL-6 target genes that consists of the PcG proteins RING1B, RING1A, NSPC1 and RYBP and subunits of SCF E3 ligase complexes [84]. GATA4 is a transcription activator that was found in the USP21 purification. Studies have shown that GATA4 interacts with the HAT p300 on gene promoters and that p300 acts as a co-activator in dependence of its acetylation activity [85]. The proteins that were confirmed as putative interactors in two purifications are POLDIP3, PINX1, HOXC9 and YEATS4. Those proteins are all linked to gene transcription or chromatin regulation. POLDIP3 has been identified to interact with the p50 subunit of DNA-Polymerase [86]. PINX1 was

found to prevent telomere elongation by inhibiting telomerase activity [87]. HOXC9 induces growth arrest by repressing the transcription of cell cycle genes and can promote cell differentiation [88]. YEATS4 is a subunit of the histone acetylase complex Tip60. Moreover it represses genes that play a role in the p53 tumor suppressor pathway by binding to their promoter [89]. In one of the two TAP purifications Ring1B was amongst the candidate proteins with a low but still significant score. Ring1B was also identified in a transient Flag-USP21 purification in 293 cells conducted by C. Jeronimo in our lab (unpublished data).

	TAP purification #1		TAP purification #2	
Gene Symbol	MS score	# Peptides	MS score	# Peptides
USP21	1687	72	1153	47
POLDIP3	565	18	317	10
PINX1	74	2	47	1
HOXC9	32	1	61	1
YEATS4	27	1	37	1
BEND3	1206	46	821	28

Table 1. USP21 candidate interacting proteins. Confirmed interactors of two TAP-USP21 purifications are shown with mass spectrometry score and the number of identified peptides.

This purification approach also detected BEND3 as the main putative interacting protein of USP21 again with a very high score. The results allow the conclusion that nuclear USP21 interferes with a broad range of chromatin-associated processes like transcription, differentiation, telomeres and heterochromatin. It is involved both in gene activation and repression. As BEND3 was the protein present in distinct purifications with the highest score, we focused the following experiments on identifying the function of the interaction between USP21 and BEND3.

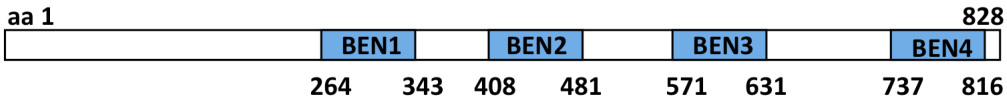


Figure 5. Schematic domain structure of BEND3.

BEND3 protein is characterized by a quadruple repeat of the BEN domain (Figure 5). That domain has been found to be present in proteins with a role in chromatin regulation and transcription. BEND3 is a member of the KIAA1553-like proteins that all have a quadruple BEN domain. Although there are several human proteins containing single BEN domains, BEND3 is the only one that harbors four copies of the domain. The KIAA1553-like family is conserved to the cnidarian Nematostella and BEND3 protein is conserved between vertebrates [90].

2.2.2. Antibody production

The human BEND3 cDNA has a total length of 2486 nucleotides and encodes the protein with a calculated mass of 95,5kDa (http://web.expasy.org/compute_pi/). No isoforms have been described. The BEND3 cDNA was cloned and was used for protein purification. Recombinant protein fragments of USP21 and BEND3 were expressed in *E.coli* Bl-21 strain and purified and the proteins were used for the immunization of rabbits. We expressed several protein fragments of different parts of USP21 and BEND3 (Figure 6a). In the case of BEND3 we chose one fragment that covers the N-terminal region and is devoid of any functional domain whereas the second fragment of BEND3 represented the C-terminus with three of its four BEN domains.

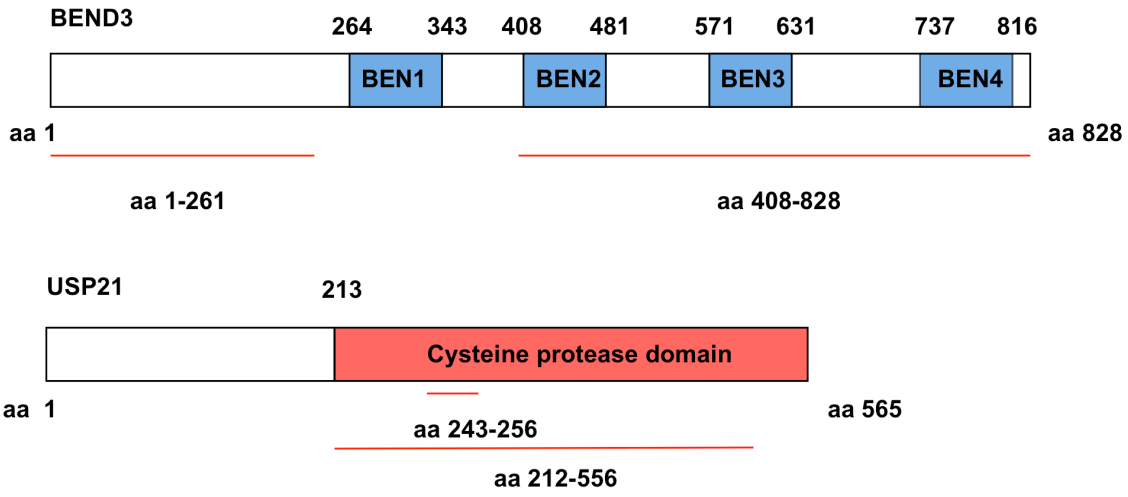


Figure 6. Scheme of the BEND3 and USP21 protein. Red lines indicate protein fragments that were used for immunization.

Two rabbits were immunized for each protein fragment of BEND3 and all four immune sera (2018, 2019, 2071 and 2072) were tested for specificity in Western Blot and immunoprecipitation. BEND3 sera were first tested for specificity on cell lysates of 293T cells expressing HA-BEND3 and control cells transfected with empty vector (Figure 7). The immune serum of rabbit 1219 was later purified with the recombinant protein. However there was no improvement of specificity observed. All sera specifically detected overexpressed BEND3 in the Western Blot. The serum 2019 detected the best the endogenous protein and only displayed low background.

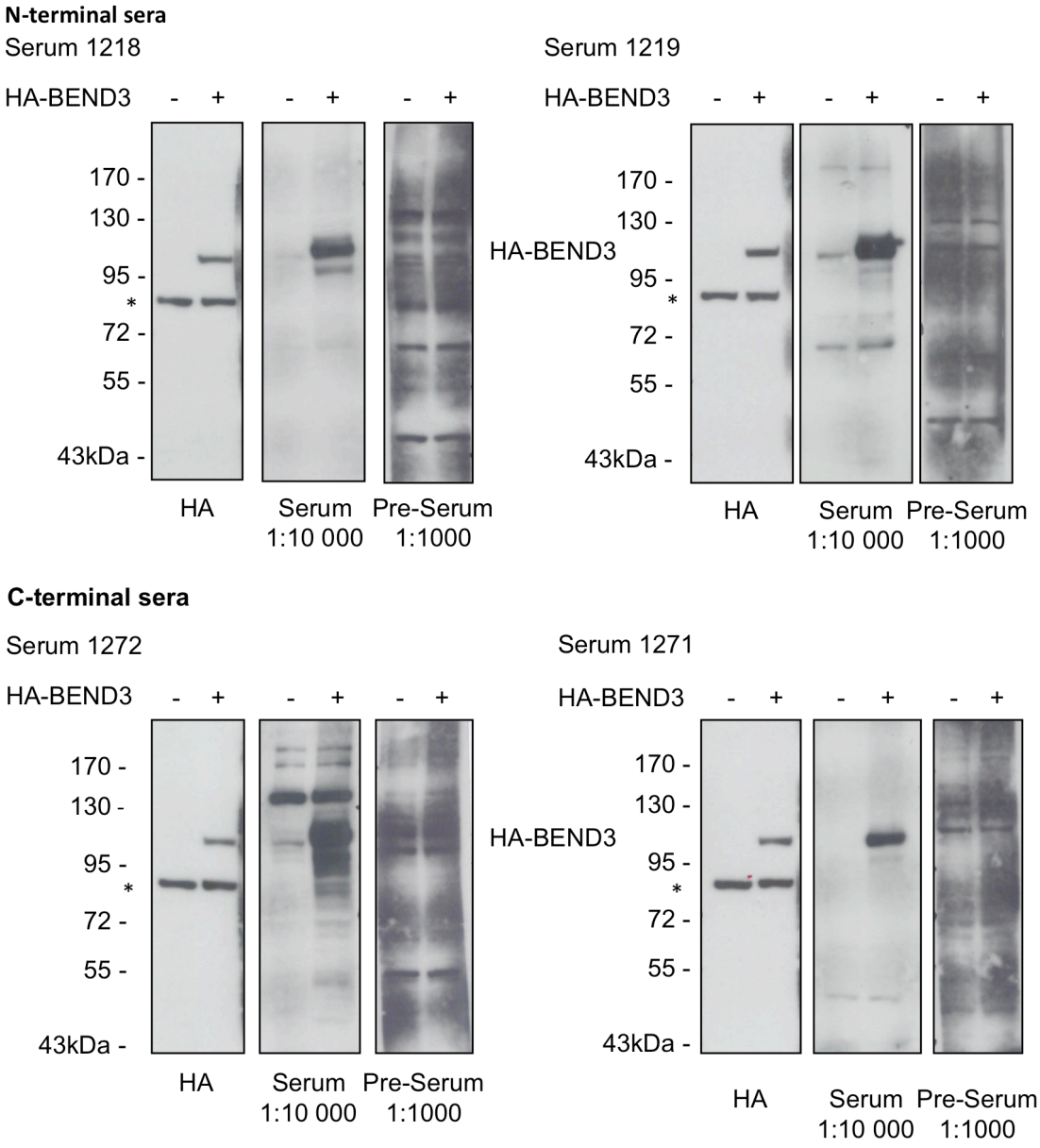


Figure 7. Specificity of BEND3 immune sera. Test of BEND3 immune sera for specificity in Western Blot. 293T cells were transiently transfected for 48h with HA-BEND3 encoding plasmid or empty vector. Equal protein amounts of whole cell extracts were separated by SDS-PAGE and transferred to a nitrocellulose membrane by Western Blot. Membranes were immunostained with immune sera, presera and HA antibody that served as a control for HA-BEND3 expression. The asterisk marks an unspecific band of the HA antibody.

Immunoprecipitation experiments were conducted of cells overexpressing HA-BEND3 and non-transfected control cells. Different amounts of control antibody and immune sera were used to precipitate overexpressed and endogenous BEND3 protein. The precipitates were separated in a SDS-PAGE. After transfer the membranes were probed with both a C-terminal as well as a N-terminal

immune serum in order to prove that the precipitated protein was indeed BEND3 (Figure 8 and 9).

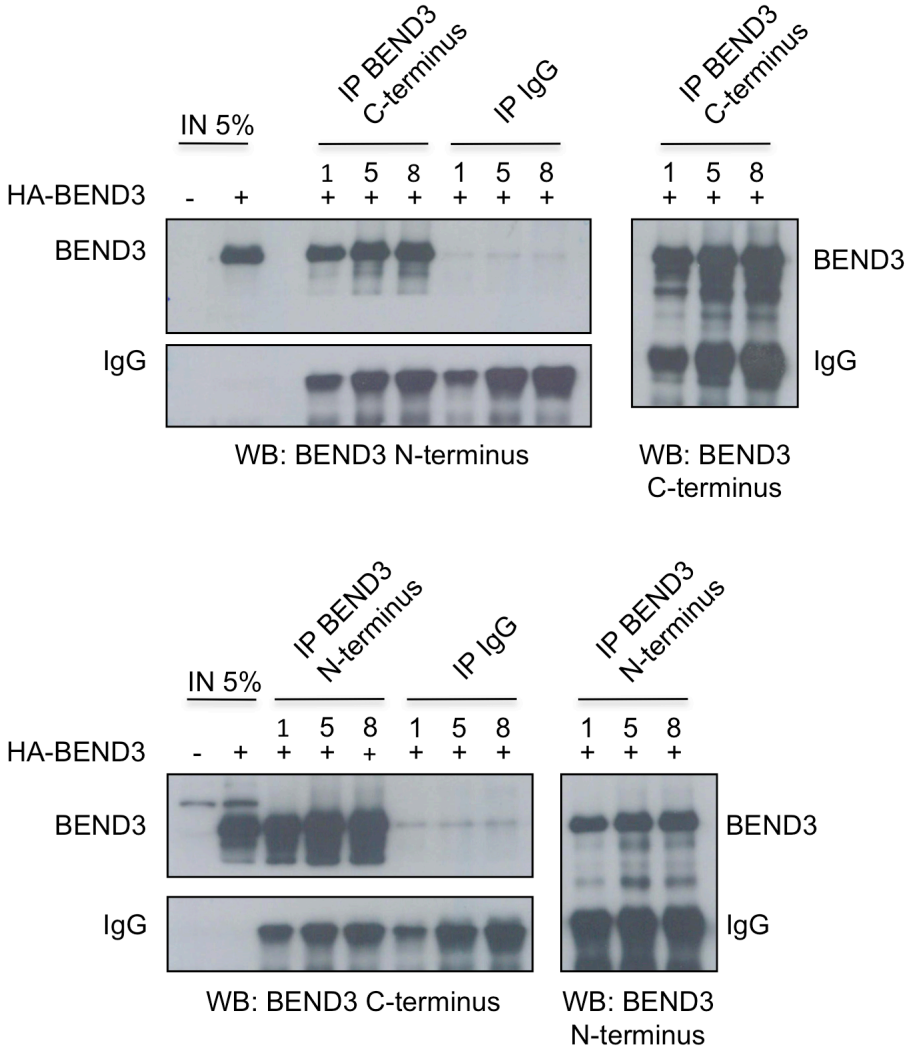


Figure 8. Test of BEND3 immune sera in immunoprecipitation of overexpressed HA-BEND3. Immunoprecipitations were conducted of 1mg whole cell protein of 293T cells transfected with HA-BEND3 plasmid. 1, 5 and 8µg of purified BEND3 N-terminal antibody and 1, 5 and 8µl of BEND3 C-terminal serum were used. 1, 5 and 8µg of unspecific IgG served as control antibody. The precipitates and lysate inputs (IN) were separated by SDS-PAGE and transferred to nitrocellulose membrane by Western Blot. Membranes were stained with N-terminal and C-terminal antibody as indicated.

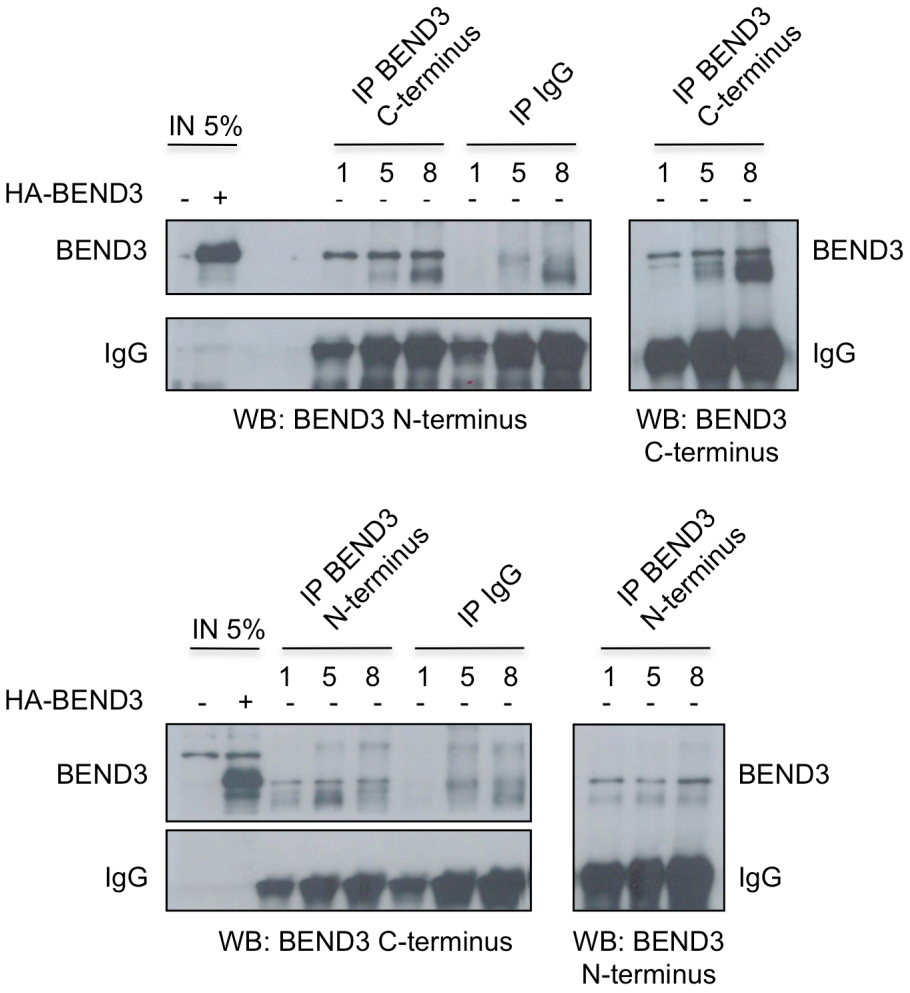


Figure 9. Test of BEND3 immune sera in immunoprecipitation of endogenous BEND3. Immunoprecipitations were conducted of 1mg whole cell protein of 293T cells. 1, 5 and 8µg of purified BEND3 N-terminal antibody and 1, 5 and 8µl of BEND3 C-terminal serum were used. 1, 5 and 8µg of unspecific IgG served as control antibody. The precipitates and lysate inputs (IN) were separated by SDS-PAGE and transferred to nitrocellulose membrane by Western Blot. Membranes were stained with N-terminal and C-terminal antibody as indicated.

In order to confirm the identity of the band detected by the BEND3 N-terminal 2019 serum we tested the antibody in Western Blot of mESC lysates depleted for BEND3. This test confirmed the identity of the band. As proven by all those experiments the purified BEND3 N-terminal serum 2019 is specific and we used this serum in the majority of the following experiments (Figure 10).

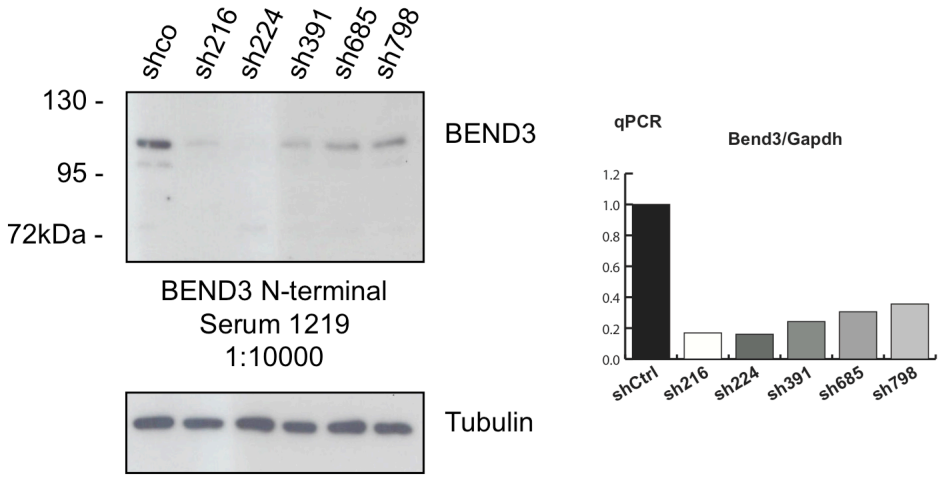
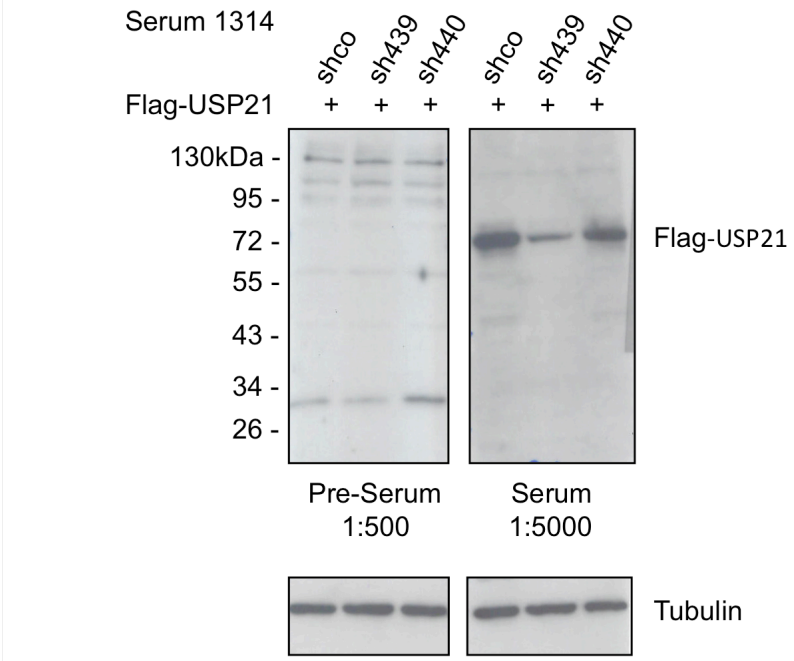


Figure 10. BEND3 N-terminal purified antibody tested on knockdown of BEND3 in mESC. Cells were stably infected with different shRNAs directed against BEND3, and an unspecific control shRNA. Knockdown efficiency was checked in qPCR. Whole cell lysates were analyzed by Western Blot. Membranes were incubated with BEND3 N-terminal purified antibody.

In the case of USP21, we chose a fragment covering the C-terminal region of USP21 that contains the protease domain for immunization. The N-terminus in contrast does not contain any functional domain and could not be expressed in bacteria. That finding was initially reported by Ye and collaborators who ascribed that fact to the high flexibility of the N-terminus [78]. Additionally we also ordered the production of a peptide antibody directed against a peptide within the catalytic domain of USP21 (Figure 6). Anti-USP21 sera raised against the C-terminal fragment were tested for specificity in Western Blot using a HEK293 cell line that stably expresses Flag-USP21. The cells were depleted for both endogenous and overexpressed USP21 by infecting them with small hairpin RNA (shRNA) directed against USP21 and the specificity of the sera could thus be controlled using Flag antibody. The only band detected by USP21 C-terminal sera indeed corresponded to Flag-USP21 as the band diminished in the knockdown cells but not in the cells infected with a control shRNA. Presera were unspecific (Figure 11).



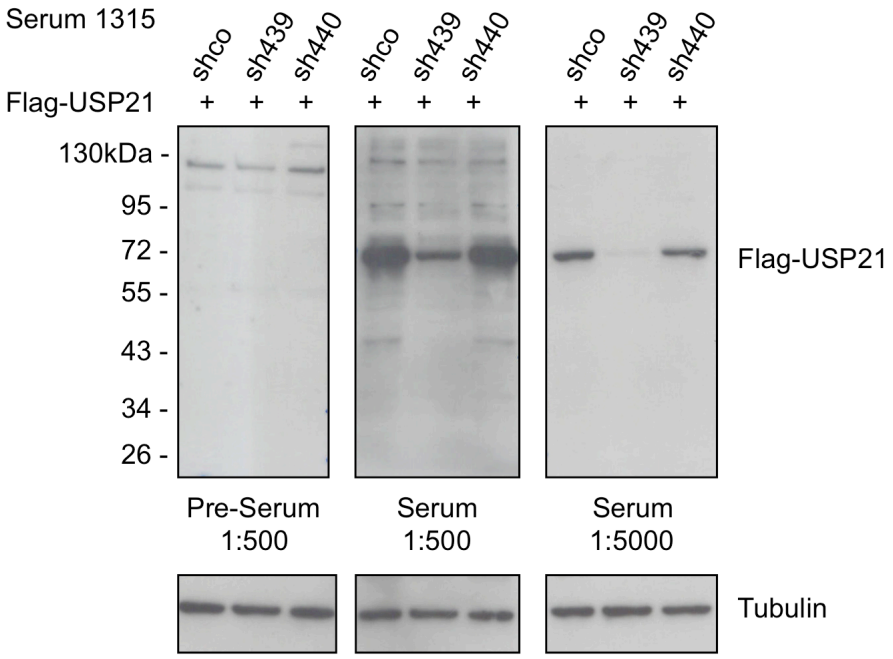


Figure 11. Specificity test of USP21 C-terminal immune sera in Western Blot. HEK293 cells stable expressing Flag-USP21 were infected with two different short hairpin RNAs directed against USP21 and an unspecific control short hairpin RNA. Equal protein amounts of whole cell extracts were separated by SDS-PAGE and transferred to a nitrocellulose membrane by Western Blot. Membranes were immunostained with two dilutions of serum and as well the unspecific preserum. Tubulin antibody is used as loading control.

When tested in immunoprecipitation both sera raised against the USP21 C-terminus only precipitated USP21 when it was overexpressed in 293T cells. Presera did not unspecifically precipitate Flag-tagged USP21 (Figure 12). This shows that the USP21 C-terminal antibody we generated is specific to the targeted overexpressed protein.

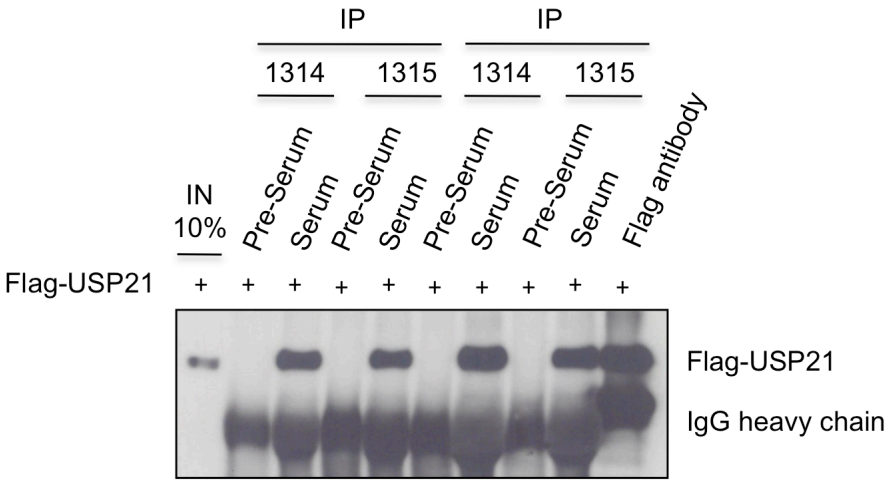


Figure 12. Specificity test of USP21 C-terminal immune sera in immunoprecipitation. 293T cells were transiently transfected with Flag-USP21. Precipitation of USP21 was performed using two different quantities of each serum. Preserum and Flag antibody served as negative and positive control respectively. Precipitates were analyzed by Western Blot and stained with Flag antibody.

In case of the peptide antibody raised against an amino acid sequence of the C-terminus its specificity was only proven in Western Blot of cell lysates with overexpressed USP21 (Figure 13). Regarding the antibodies against USP21 we have generated two antibodies that detect the overexpressed protein. However endogenous USP21 could neither be detected by immunoblot nor immunoprecipitation. Table 2 shows a table giving an overview about the specificity and the possible applications of the antibodies that we have generated.

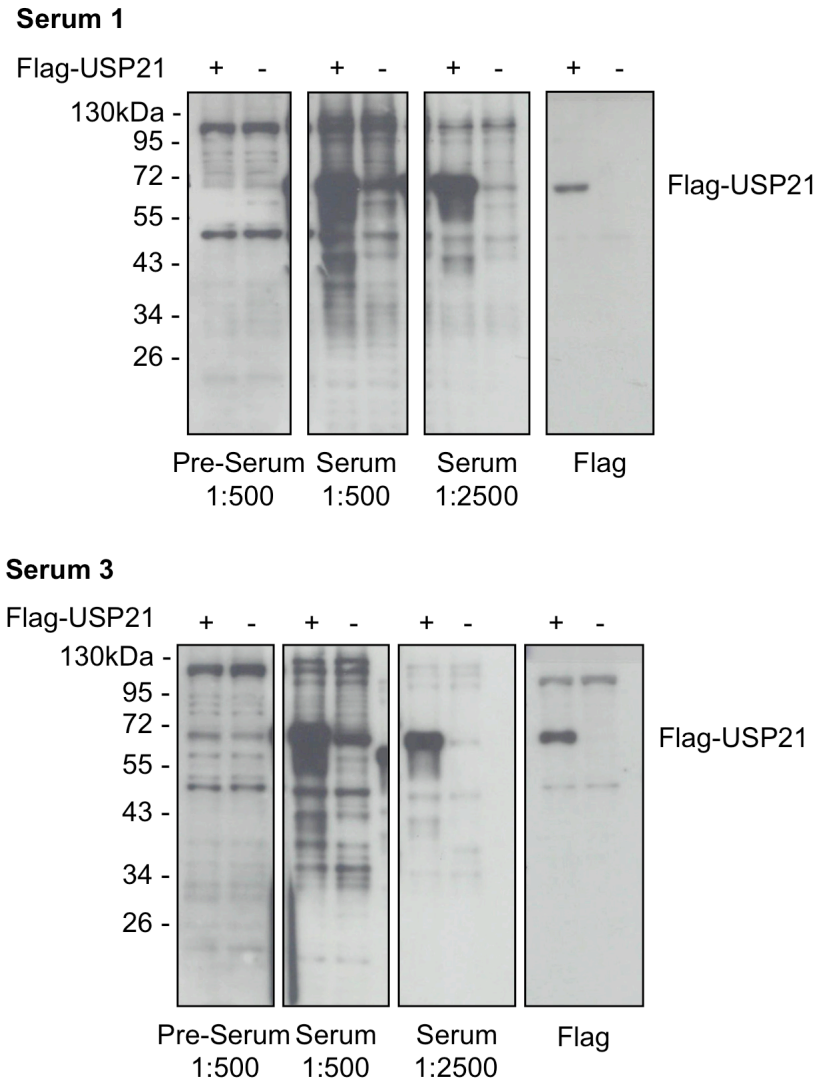


Figure 13. Specificity test of USP21 peptide immune sera in Western Blot. 293T cells were transiently transfected for 48h with Flag-USP21 encoding plasmid or empty vector. Cells were fractionated and equal protein amounts of nucleoplasm were separated by SDS-PAGE and transferred to a nitrocellulose membrane by Western Blot. Membranes were immunostained with immune sera, presera and Flag antibody that served as a control for Flag-USP21 expression.

Antibody	WB		IP	
	OE	END	OE	END
USP21 Genscript	+	-	-	-
USP21 C-terminus	+	-	+	-
BEND3 N-terminus	+	+	+	+
BEND3 C-terminus	+	+	+	+

Table 2. Antibody list. BEND3 and USP21 antibodies and their possible application.

2.2.3. Verification of BEND3 as interacting protein of USP21

2.2.3.1. Immunoprecipitations

We further verified the mass spectrometry results by immunoprecipitation. Flag-tagged USP21 was overexpressed together with HA-tagged BEND3 in 293T cells and Flag-USP21 was precipitated from whole cell extracts using Flag antibody. The whole cell extracts had been mixed with different concentrations of ethidium bromide that intercalates with DNA and prevents the formation of protein interactions due to underlying DNA. Preliminary results showed that USP21 bound BEND3 to the same extent in the presence of ethidium bromide than in non-treated control extracts what suggests that the interaction does not depend on DNA (Figure 14).

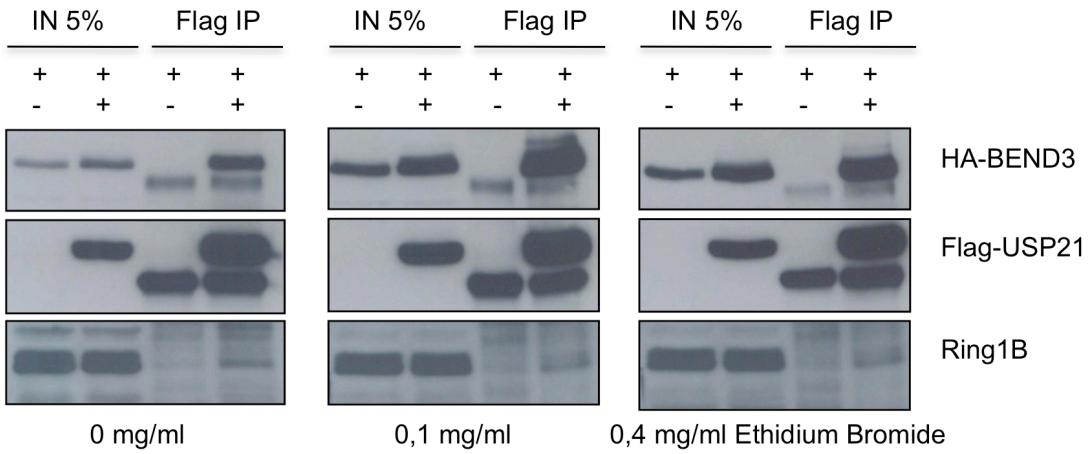


Figure 14. The interaction between USP21 and BEND3 is independent of DNA. Immunoprecipitations were conducted of 1mg whole cell protein of cells transfected with HA-BEND3 or HA-BEND3 together with Flag-USP21. Different amounts of ethidium bromide were added to the whole cell extract. The IP was performed with Flag antibodies and precipitates were separated by SDS-PAGE and transferred to nitrocellulose membrane by Western Blot. Membranes were stained with HA- and Flag-antibody as well as RING1B immune serum.

Moreover overexpressed Flag-USP21 co-precipitated as well endogenous BEND3 (Figure 5b). The same results were observed in preliminary experiments with Flag-USP21 catalytic mutants (Figure 5c) demonstrating that the interaction does not depend on USP21 catalytic activity. Together these results indicate that BEND3 binds USP21 *in vivo*.

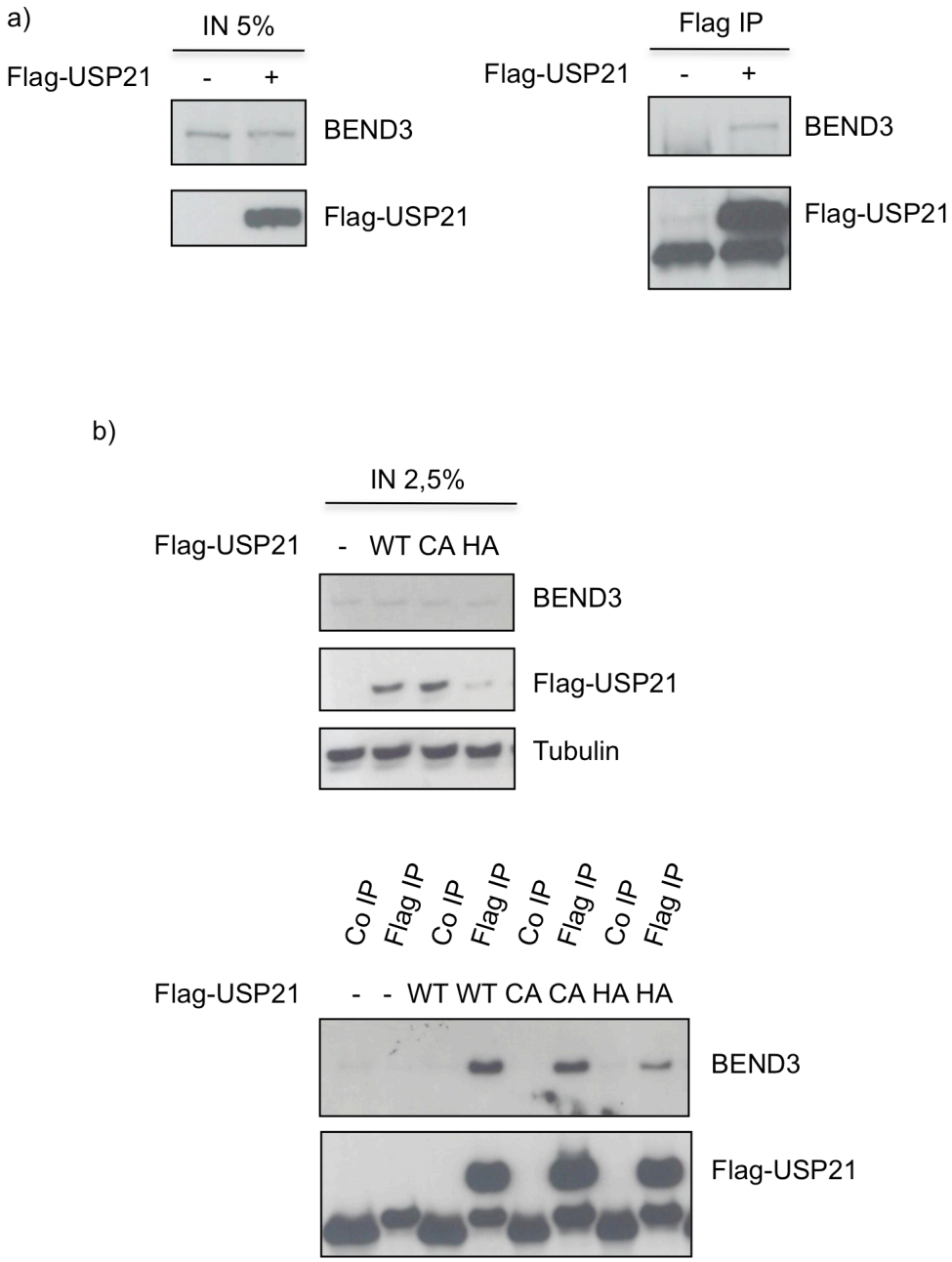


Figure 15. Characterization of the interaction between USP21 and BEND3. (a) Flag-USP21 interacts with endogenous BEND3. Flag immunoprecipitations were conducted of 1mg whole cell protein of cells transfected with Flag-USP21 or empty vector. Precipitates were separated by SDS-PAGE and transferred to nitrocellulose membrane by Western Blot. Membranes were stained with Flag and BEND3 N-terminal antibody. (b) Endogenous BEND3 interacts with Flag-USP21 catalytic mutants. 293T cells were transfected with Flag-tagged catalytic mutants and Flag-tagged wildtype USP21 and harvested 48h after transfection. Flag immunoprecipitations were performed of whole cell extract. Mouse IgG was used as unspecific control antibody. The precipitates were separated by SDS-PAGE and transferred to nitrocellulose membrane by Western Blot. Immunostaining of the membrane with BEND3 N-terminal antibody detected the interaction of endogenous BEND3 with Flag-USP21. Flag staining served as control of the precipitation of Flag-USP21. The inputs indicate equal expression of BEND3 and the overexpression level of Flag-USP21 constructs.

2.2.3.2. TAP-BEND3 purification

Purification of BEND3 was performed to investigate a possible overlap of USP21 and BEND3 interactomes. BEND3 was cloned in a pcDNA plasmid with a C-terminal TAP-tag, transiently overexpressed in 293T cells and its expression levels were checked by Western Blot. The protein level of BEND3 was very high even when transfecting only small amounts of plasmid (Figure 16a). For affinity purification cells were fractionated and TAP-BEND3 was isolated from the nucleoplasmatic fraction. The purification steps were checked by Western Blot to monitor the input lysate and the flow-through after binding the protein to calmodulin beads (Figure 16b).

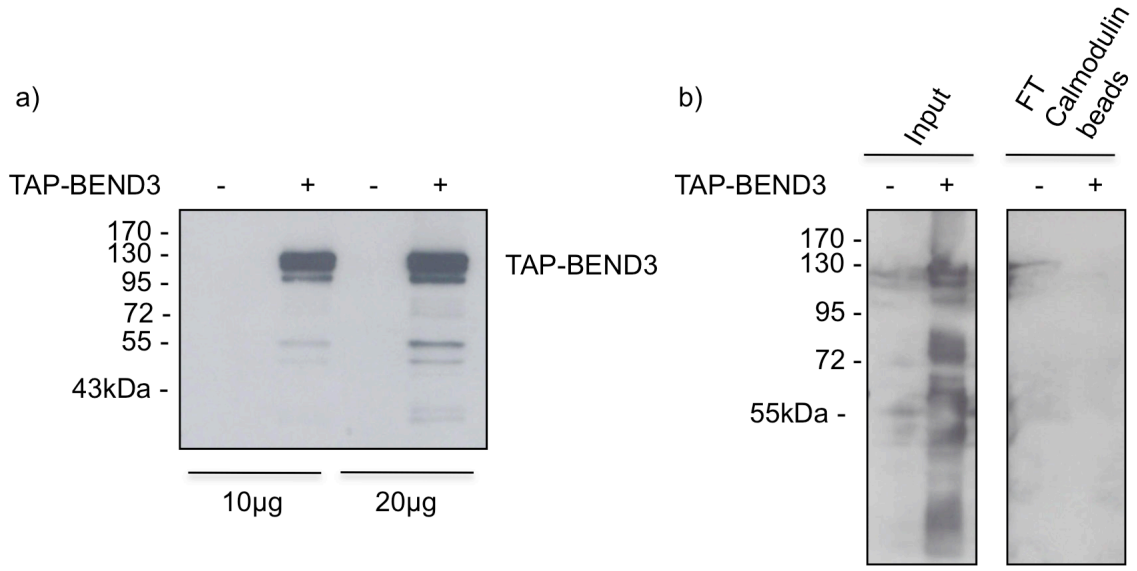


Figure 16. Purification of proteins associated to nuclear BEND3. (a) 293T cells were transiently transfected with TAP-BEND3 or empty vector. Indicated total protein amounts of whole cell extracts were analyzed by Western Blot and CBP immunostaining. (b) TAP-BEND3 affinity purification of nucleoplasm. Depletion of TAP-BEND3 of nucleoplasm was analyzed by Western Blot analysis of input material and the flow-through after purification.

The eluates containing BEND3 complexes and the supernatants of the boiled beads after elution were separated in a SDS-PAGE and silver staining revealed protein bands. Several specific bands in the TAP-BEND3 expressing cells were identified (Figure 17). Those bands were excised and subjected to mass spectrometry. However very few specific proteins were detected in the analysis. Comparison with previous purifications performed under the same conditions by C. Jeronimo (unpublished data) showed similar results and USP21 was only

detected in a single purification with a low score. Those results suggest that nuclear BEND3 is a protein that is not present in a complex and that only a subfraction interacts with USP21.

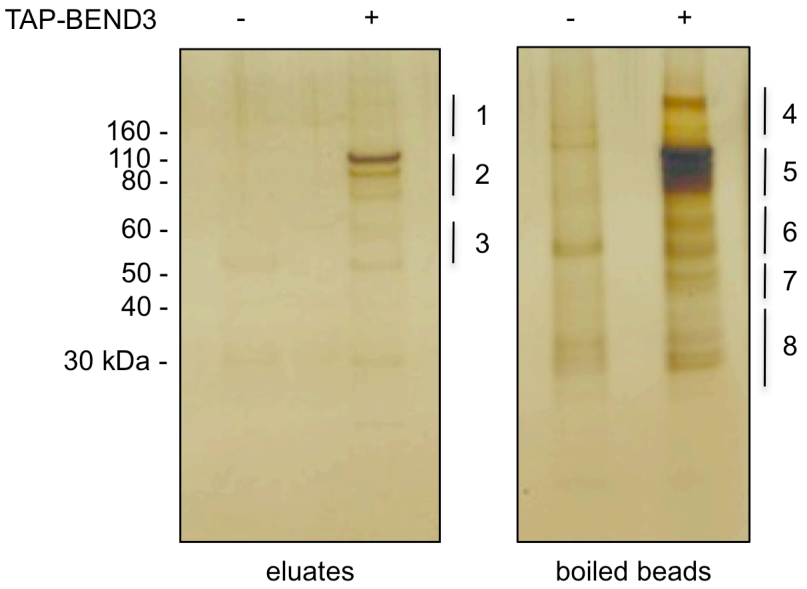


Figure 17. TAP-BEND3 purification eluates. Eluates were separated in a SDS-PAGE that was silver stained. The marked bands were excised and proteins analyzed by mass spectrometry.

2.2.4. Mapping of the interaction

Having demonstrated that BEND3 and USP21 interact *in vivo*, we next sought to delimitate the interaction in more detail. For that purpose two fragments of HA-tagged BEND3 were cloned. One encoded the N-terminus and the other comprised the C-terminus with its four BEN domains (Figure 18a). Both fragments as well as the full-length HA-BEND3 (HA-BEND3 FL) were each overexpressed together with Flag-USP21 in 293T cells. Flag-USP21 was precipitated and analyzed in Western Blot for co-immunoprecipitation of the BEND3 fragments. As expected both full-length proteins co-precipitated (Figure 18b).

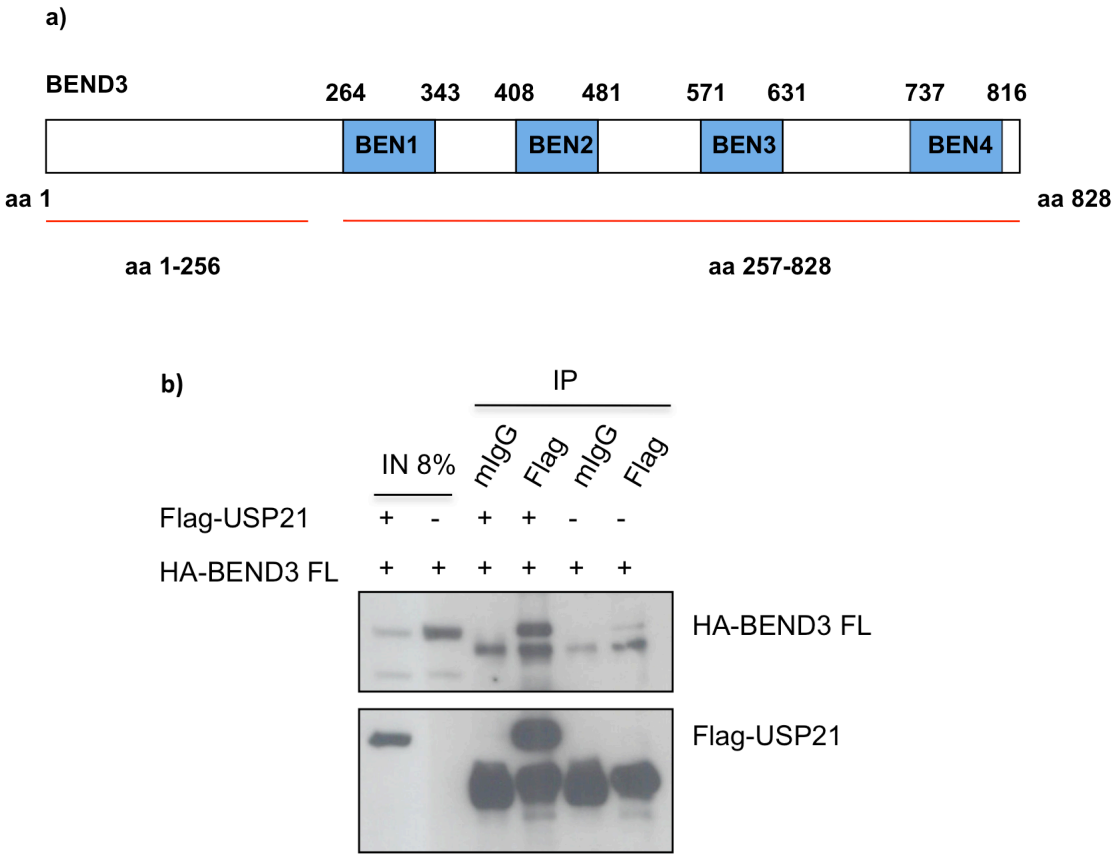


Figure 18. Mapping of the BEND3 binding domain to USP21. (a) BEND3 deletion constructs. (b) 293T cells were transfected with Flag-tagged wildtype Flag-USP21 and HA-BEND3 FL. Cells were harvested 48h after transfection. Flag immunoprecipitation was performed from whole cell extract. IgG antibody was used as unspecific control antibody. The precipitates and input (IN) lysates were separated by SDS-PAGE and transferred to nitrocellulose membrane by Western Blot. Immunostaining of the membrane with HA detected the interaction of overexpressed BEND3 with USP21. Flag staining confirmed the precipitation of Flag-USP21. The Inputs show expression of HA-BEND3 FL and Flag-USP21.

However the N-terminal region failed to bind Flag-USP21. In contrast BEND3 C-terminus that contains all four BEN domains interacted with Flag-USP21 (Figure 19). Thus the BEND3 C-terminus that contains the BEN domains directs the interaction of BEND3 with USP21.

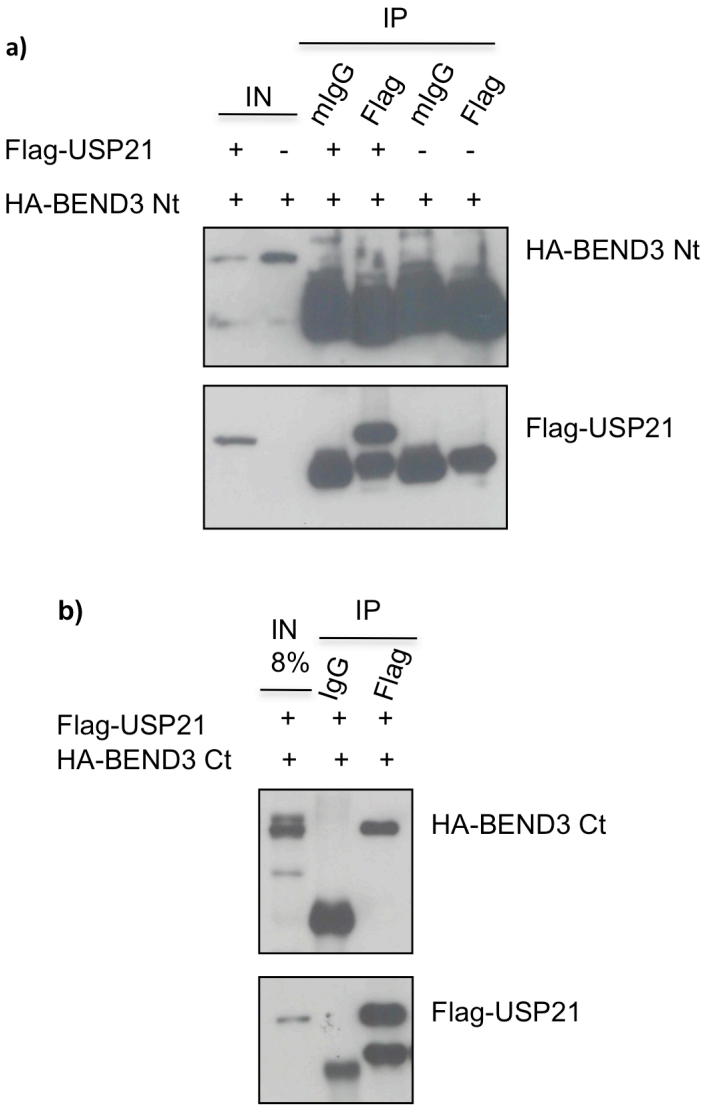


Figure 19. Mapping of the USP21 binding domain of BEND3. 293T cells were transfected with Flag-tagged wildtype Flag-USP21 and HA-BEND3 N-terminal (a) and C-terminal (b) deletion constructs and cells were harvested 48h after transfection. Flag immunoprecipitation was performed from whole cell extract. IgG antibody was used as unspecific control antibody. The precipitates and input (IN) lysates were separated by SDS-PAGE and transferred to nitrocellulose membrane by Western Blot. Immunostaining of the membrane with HA (a) and BEND3 C-terminal (b) antibody detected co-precipitated BEND3 fragments. Flag staining confirmed the precipitation of Flag-USP21. The Inputs show expression of HA-BEND3 constructs and Flag-USP21.

2.3. Functional impact of the interaction

2.3.1. Regulation of USP21 catalytic activity by BEND3

The results from the mass spectrometry analysis led us to assume that BEND3 represents the main regulator of nuclear USP21. For this reason we started studying how BEND3 impacts on USP21 protein function, more precisely its

enzymatic activity. Due to the fact that no full-length recombinant proteins could be expressed and consequently no *in vitro* deubiquitination assays could be performed, the activity of USP21 was assessed *in vivo* monitoring the global amount of monoubiquitinated H2A. For that purpose HA-BEND3 was overexpressed in 293T cells and cells were fractionated into cytoplasm, nucleoplasm and chromatin. Regarding the localization of BEND3, it was apparent that BEND3 localizes mainly in the nucleus because nucleoplasm and chromatin contained the major proportion of HA-BEND3 where as few was detected in cytoplasm. However H2A ubiquitination levels did not change compared to non-transfected cells. That indicates that BEND3 has no impact on global H2A ubiquitin levels (Figure 20).

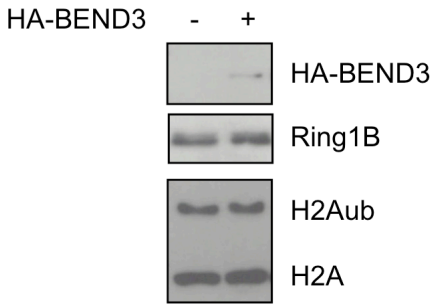


Figure 20. BEND3 does not impact on H2Aub levels. HA-BEND3 or empty vector was transiently overexpressed in the same number of 293T cells. Cells were harvested and fractionated 48h after transfection. Chromatin was sonicated and equal volumes loaded on the SDS-PAGE. The membrane was probed with HA antibody to detect BEND3 expression. Ring1B was used as loading control. H2Aub levels were detected with H2A antibody.

2.3.2. Regulation of BEND3 by USP21

2.3.2.1. Ubiquitination status of BEND3

DUBs very often associate with their target proteins. In line with that USP21 has been reported to interact with the TNFa receptor associated protein RIP1. RIP is polyubiquitinated by K63 chains and USP21 has been demonstrated to remove polyubiquitin chains from RIP1 [81]. We therefore wondered if BEND3 is posttranslationally modified by ubiquitin *in vivo*. Ubiquitination can be assessed by overexpression of HA-tagged ubiquitin together with the presumable target protein. Flag-BEND3 was overexpressed in 293T cells

together with HA-ubiquitin and was precipitated from whole cell extracts. Purified BEND3 complexes where then analyzed by Western Blot. Immunostaining with HA antibody revealed a smear starting from the size of unmodified BEND3 upwards and this smear contained the precipitated BEND3 that is modified by HA-ubiquitin (Figure 21). Overexpressed BEND3 is thus polyubiquitinated in 293T cells. Nevertheless it remains to be studied what linkage type it concerns about. That is of particular interest, as it will provide an indication of the cellular process in that ubiquitinated BEND3 is involved. Linkage types can be specified by using HA-ubiquitin mutants, that possess lysine to arginine substitutions in their key lysine residues K11, K63 or K48. Primary results of co-immunoprecipitations with K63 and K48 constructs hypothesize that BEND3 is modified neither by K48 nor by K63 ubiquitination.

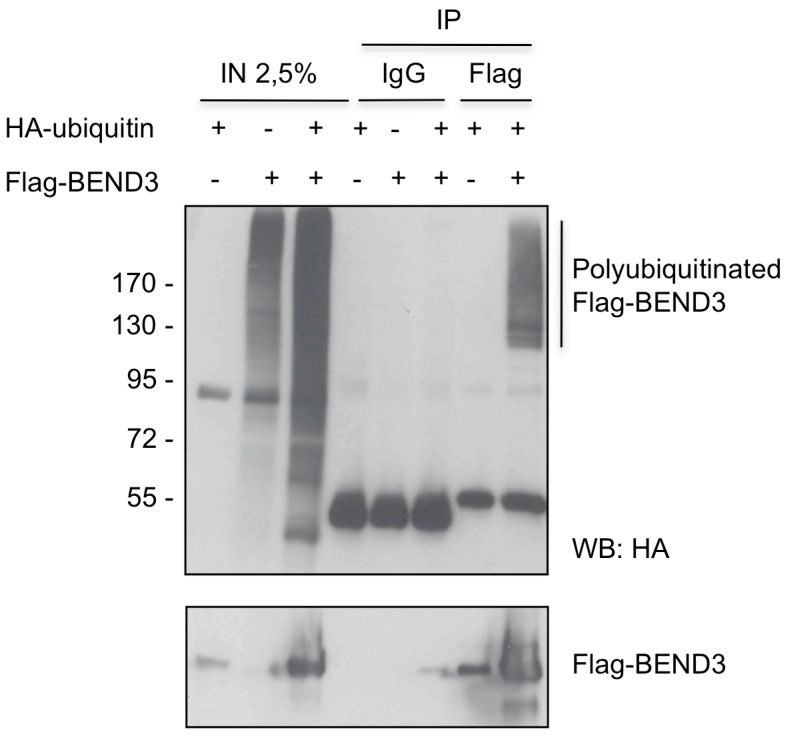


Figure 21. Flag-BEND3 is polyubiquitinated in 293T cells. Flag-BEND3 and HA-ubiquitin were overexpressed in 293T cells. Cells were harvested after 5,5h treatment with 10μM MG-132. Lysis buffer contained 10μM NEM, protease and phosphatase inhibitor. Flag immunoprecipitation was performed and proteins separated on a SDS-PAGE. Mouse IgG antibody was used as unspecific control antibody. Membranes were denatured and immunoblotted with HA antibody. The line marks the smear of polyubiquitinated Flag-BEND3. The membrane was stripped and immunoblotted with BEND3 N-terminal antibody (lower panel).

2.3.2.2. Protein stability of BEND3

Proteins are targeted for proteasomal degradation by marking them with polyubiquitin chains. Those chains consist of K48 linkages and are very abundant in cells. DUBs rescue proteins from degradation by cleaving these K48 chains and control in that way the turnover rates of their targets. Analyzing the influence of USP21 on BEND3 protein stability would therefore already give a first clue about the chain type and the function of USP21. Two different approaches were taken. First 293T cells were transfected with wildtype Flag-USP21 or its catalytic mutant CA. Comparing to non transfected cells one would expect an increase in endogenous BEND3 protein in USP21 wildtype transfected cells as the putative deubiquitination of the BEND3-linked K48 chain would lead to a diminished degradation in the proteasome. However BEND3 levels of 293T cells expressing a functional USP21 were exactly the same as in the control cells or those expressing the mutant (Figure 22a). Therefore USP21 does not seem to regulate BEND3 stability, indicating that if USP21 is the specific DUB, BEND3 polyubiquitin chains might be of another linkage type. Moreover a subset of cells was also treated with the proteasomal inhibitor MG-132. This compound can penetrate cell membranes and inhibit reversibly the proteasome. When cells are exposed to MG-132, they shut down protein degradation and as a consequence general protein levels rise and a huge portion of the proteins display K48-linked ubiquitin chains. Treating cells with MG-132 however did not influence the global amount of endogenous BEND3 what indicates that it is a very stable protein with a low proteasomal turnover. On the contrary MYC protein that has a short turnover rate was stabilized by MG-132 treatment.

The second attempt was monitoring the levels of endogenous BEND3 in stable knockdown cells for USP21 (Figure 22b). Initial experiments demonstrated that when cells were depleted for USP21, BEND3 protein levels were unaltered indicating that USP21 does not regulate its protein levels. Cells in that experiment were also treated with cycloheximide, a translational inhibitor. The use of the inhibitor causes a break off of protein production in the cell and performing a time course one can determine the half time of a determined

protein. MYC that served as positive control, decreased as expected already after 3h, unlike BEND3 whose levels stayed unchanged. This finding demonstrates that BEND3 has a very low turnover and is thus a stable protein. The time course was temporally limited because cells started dying after 9h due to impairments in cell division. It can be concluded that BEND3 is a stable protein with a low turnover rate as demonstrated by its constant protein levels after proteasome inhibition and translational block what means that only a small percentage of BEND3 is linked to K48 chains. The smear observed in the immunoprecipitations therefore might consist of polyubiquitin chains with another linkage type than K48. Moreover the overall protein levels of BEND3 are not altered by USP21 demonstrating that USP21 does not regulate BEND3 protein stability.

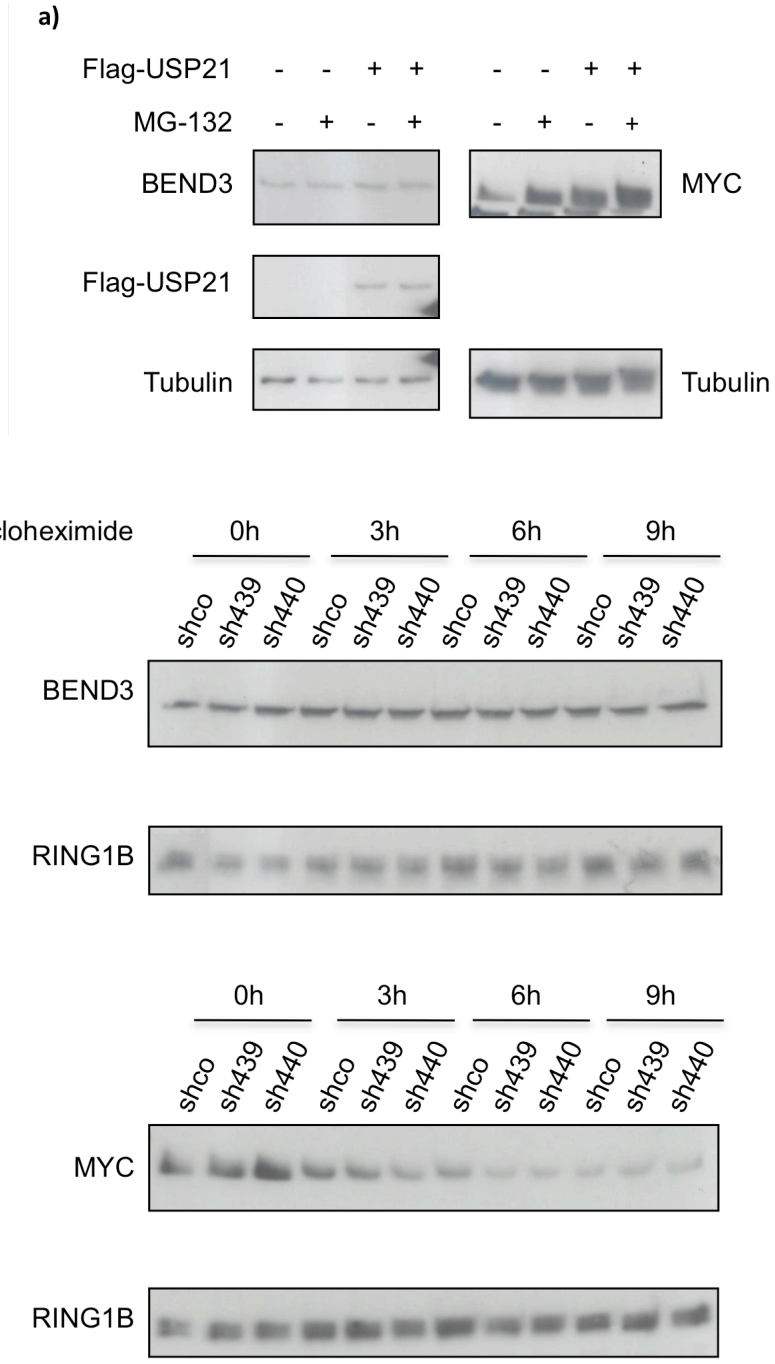


Figure 22. USP21 does not alter BEND3 protein levels in the cell. (a) BEND3 protein levels in 293T cells overexpressing USP21. 293T cell were transiently transfected with Flag-USP21 or empty vector for 48h. One set of cells was treated with 10μM MG-132 for 4h before harvesting. Whole cell extracts were separated by SDS-PAGE and immunoblotted with BEND3 N-terminal-, Flag- and MYC-antibody. Tubulin was the loading control and MYC expression served as positive control for proteasomal degradation. (b) BEND3 protein levels in 293T cells with knockdown of USP21. Cells were stably infected with two short hairpin RNAs directed against USP21 or an unspecific control short hairpin RNA. After treatment with 400μM cycloheximide for the indicated duration cells were collected and fractionated. The nucleoplasmatic fraction was loaded on a SDS-PAGE. Membranes were immunoblotted with BEND3 N-terminal antibody. MYC served as positive control. RING1B immunostaining showed equal loading.

2.3.2.3. Subcellular localization

Dynamical ubiquitination is a mean to direct proteins to their designated localization in the cell. We next investigated if subcellular localization of BEND3 depends on USP21 what would suggest an involvement of the polyubiquitin mark in directing BEND3 localization. Again we overexpressed Flag-USP21 and its CA mutant in 293T cells and that time we fractionated the cells and H2A ubiquitination levels were checked as a control for USP21 activity. Western Blot analysis demonstrated that BEND3 endogenous levels are unchanged what means that USP21 does not affect BEND3 localization (Figure 23). It cannot be ruled out that ubiquitination is still important for localization because USP21 has not been shown to be the specific DUB.

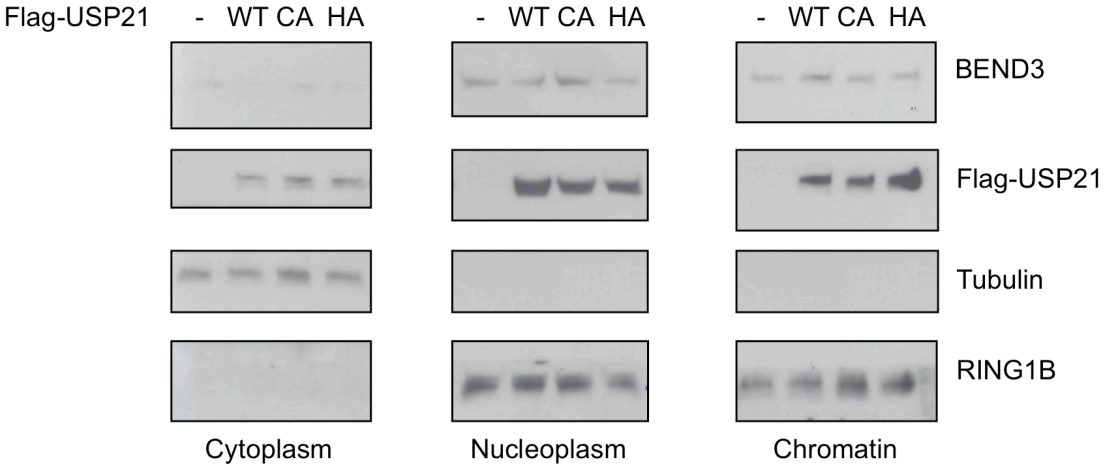


Figure 23. Subcellular localization of BEND3. 293T cells were transiently transfected with vector encoding Flag-USP21 or the mutants Flag-USP21 C221A or Flag-USP21 H518A and harvested 48h after transfection. Cells were fractionated into cytoplasm, nucleoplasm and chromatin. Chromatin was solubilized by sonication. Lysates were separated by SDS-PAGE and transferred by Western Blot to a nitrocellulose membrane. Tubulin, RING1B and Histone H3 antibodies controlled equal loading and fractionation efficiency. Flag-USP21 constructs were detected by Flag antibody, BEND3 with the N-terminal antibody and H2Aub levels with an H2Aub specific antibody (Cell signaling).

2.4. BEND3 in mESC

Maintenance of self-renewal and pluripotency in mESC depends in a large part on the tightly regulated interplay of three essential transcription factors, namely OCT4, SOX2 and NANOG that constitute together the so-called core transcriptional network [2]. OCT4 and NANOG are indispensable for maintaining the stem cell state. Loss of NANOG for example induces ESC to differentiate [91]. The same is the case for OCT4 whose expression levels have been observed to drive the differentiation state of mouse embryonic stem cells. Altering its expression level 50% up or down induced cells to differentiate to endoderm and mesoderm or trophectoderm, respectively [92]. Its binding partner SOX2 moreover synergistically enhances the transcriptional activation that is mediated by OCT4 because the complex formation between the two regulators is cooperatively enhanced by the underlying promoter sequence. The transcription factor (TF) binding sites must therefore be in close proximity to each other [93]. Identification of the target genes of OCT4, SOX2 and NANOG by chromatin immunoprecipitation (ChIP) followed by DNA microarray hybridization revealed that many promoters are co-occupied by the three regulators. Classification of the co-bound genes regarding their expression status interestingly showed that active genes comprise many components of signaling pathways that are crucial for maintenance of the stem cell state. Besides that also the three core factors were among the active genes that were co-bound by all three TFs indicating that OCT4, SOX2 and NANOG act together to activate their own expression. The repressed genes in contrast represented proteins involved in differentiation and lineage commitment, amongst them many HOX genes. The stem cell state therefore depends on an autoregulatory and feedforward circuitry regulated by the three factors that are downregulated during differentiation [94]. The transcriptional network was further studied by chromatin immunoprecipitation of several transcription factors coupled with ultra-high-throughput DNA sequencing (ChIP-sequencing). This approach identified their genome wide binding sites in mESC and moreover revealed that binding motifs of specific subsets of TFs are frequently located very close to each other forming multiple transcription factor-binding loci (MTL).

Comparing the TF binding sites as well as gene expression changes upon differentiation allowed the construction of a gene regulatory network in undifferentiated cells [95]. Identifying the interacting proteins of the core factors will allow a deeper understanding of how the transcription factors bind together to co-regulated promoters and whether that depends on the presence of further regulatory proteins. In a study analyzing the interactome of OCT4 in mESC many trancription factors were found that have an impact on stem cell pluripotency, amongst those SALL4, ESRRB and Sox2 but also subunits of chromatin remodeling complexes such as the NuRD complex. Interestingly the putative OCT4 binding proteins Sox2, NAC1, TCFCP2L1, ESRRB, DAX1 had been previously described to share target genes with OCT4 and moreover the own gene promoters of several of the purified TFs are known to be targeted and regulated by OCT4. This again demonstrates the high grade of regulation of the transcriptional network in mESC. The putative interactors TCFCP2L1, DAX1, SALL4, and ESRRB were also affinity purified in that study in order to yield more members of the regulatory OCT4 network. Mass spectrometry analysis of SALL4 showed that BEND3 was the candidate interacting protein with the second highest score. BEND3 was also detected in the purification of TCFCP2L1 and ESRRB [96]. Those findings prompted us to investigate the role of BEND3 in gene expression in mESC.

As a first approach we analyzed changes in gene expression in pluripotent and self-renewing mESC upon depletion for BEND3. For that purpose mESC were stably infected with short hairpin RNA specific for BEND3 (sh216). The efficient depletion of BEND3 in the samples used for gene expression analysis was checked in qPCR as well as Western Blot (Figure 24a and b).

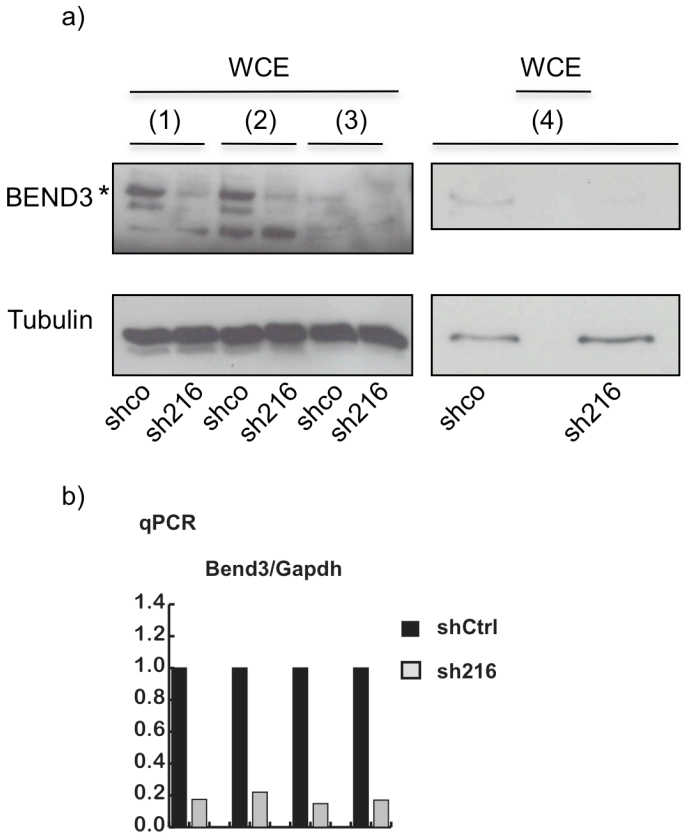


Figure 24. Gene expression analysis in mESC with BEND3 knockdown. (a) BEND3 protein levels in mESC cells used for Microarray analysis. Cells of the three samples used for analysis were lysed and equal protein amounts were loaded in a SDS-PAGE. In the case of sample 3 cells were lysed in Laemmli loading buffer and equal volumes loaded. Tubulin staining controlled equal loading; BEND3 was detected with BEND3 C-terminal immune serum. (b) qPCR analysis of BEND3 mRNA levels in mESC cells used for Microarray analysis.

The microarray showed that are total of 402 genes were downregulated whereas 354 genes were upregulated. Gene ontology analysis of the genes that significantly changed their expression after BEND3 knockdown revealed that both subsets of genes are involved in cellular growth and proliferation (Figure 25). Down-regulated genes have an additional role in cellular development. Most of the up-regulated genes participate in cell cycle but to a lesser extent also in DNA replication, recombination and repair as well as cell death.

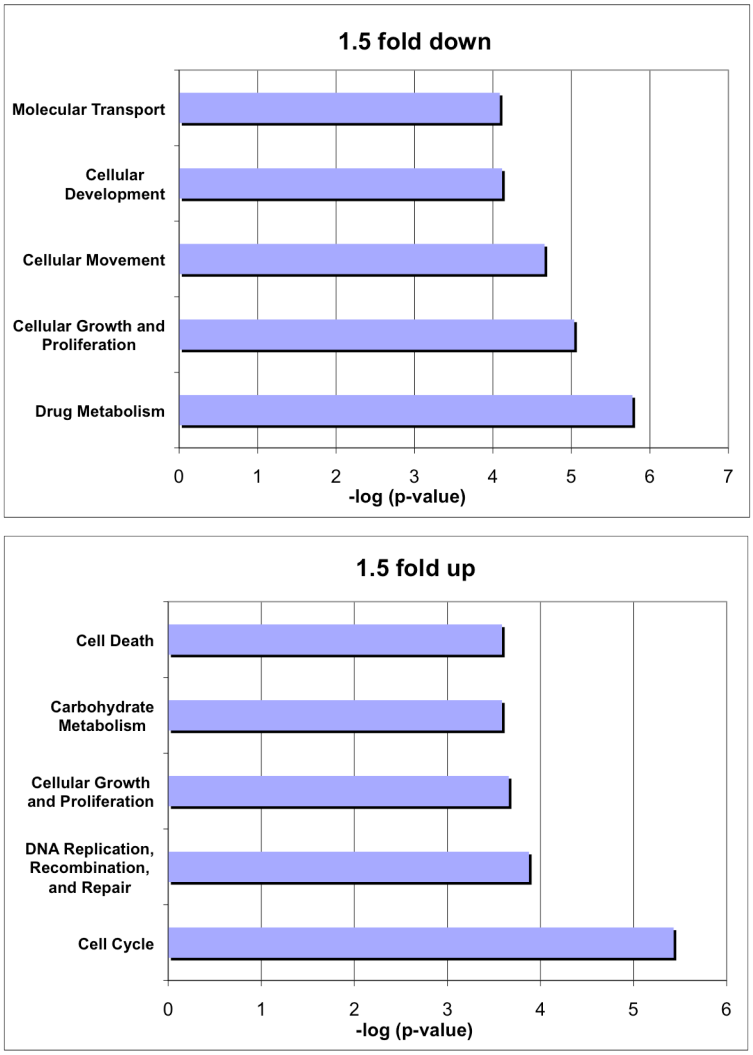


Figure 25. Gene ontology analysis. Regulated genes upon BEND3 knockdown in mESC. Genes with at least 1.5-fold change in expression and a p-value <=0.05 were considered.

We confirmed several of the most regulated genes that belong to the listed GO categories by qPCR (Figure 26). As we wanted to analyze whether USP21 and BEND3 regulate the same genes we also analyzed gene expression of the confirmed genes after knockdown of USP21 in mESC. However it turned out that after depletion of USP21 expression of the BEND3 regulated genes was not affected. Considered as whole it appears that BEND3 is involved in the expression of genes that are both up- and down regulated amongst others in cell cycle, cellular growth and proliferation, cell death, cellular development and DNA replication, recombination and repair. Those genes do not depend in their expression of USP21 in the cell. BEND3 and USP21 therefore seem not to share the same target genes in mESC.

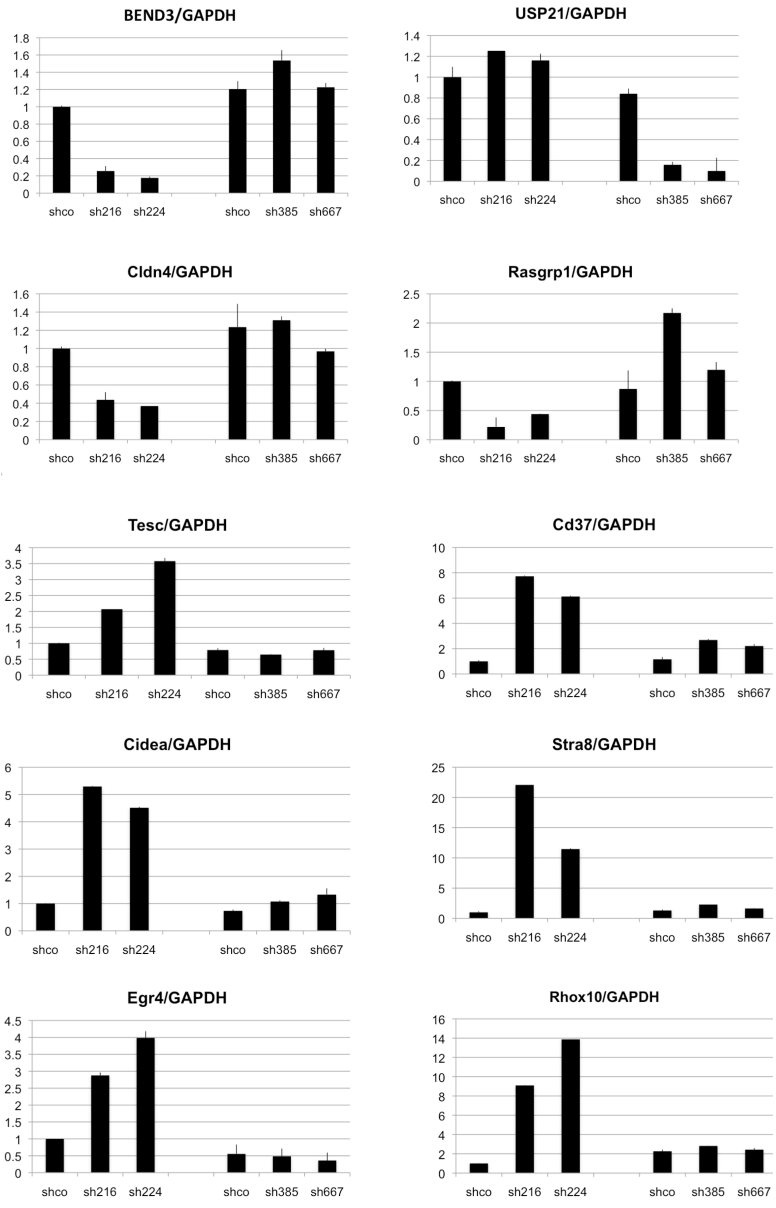


Figure 26. Validation of genes deregulated in microarray. mESC were depleted for BEND3 by shRNA and gene expression determined by qPCR. Gene expression changes were compared to expression levels in mESC with knockdown of USP21. Short hairpin constructs used in these experiments are: sh1 (sh216 and sh385 for BEND3 and USP21, respectively) and sh2 (sh224 and sh667 for BEND3 and USP21, respectively). Cldn4 and Rasgrp1 were downregulated in the microarray whereas the other represented genes were upregulated. Genes were validated in three independent experiments (2 in the case of sh667) and the results of one representative experiment are shown.

3. Discussion

3.1. USP21/BEND3 interactome

We purified stable overexpressed TAP-USP21 from nuclear extract in EcR-293 cells and these results were further confirmed by an affinity purification of Flag-tagged USP21 in 293T cells (experiment conducted by C. Jeronimo in our laboratory). In all purifications BEND3 was identified as its main interactor. In the screening for several DUBs' interactors conducted by Sowa and collaborators, USP21 was also found associated with BEND3. Moreover, they predicted by GO-analysis the putative cellular processes in which USP21 is involved. USP21 was suggested to interfere with the phosphatase PPP6C that regulates DNA damage repair, and the family of microtubule-affinity regulating kinases (MARKs) that regulate microtubules and centrosomes and moreover display the ubiquitin binding domain UBA [30]. This prediction was confirmed by a recent work that showed that USP21 binds to MARK1, MARK2 and to microtubules and centrosomes via its N-terminus [79]. Our purifications aimed at identifying a specific role for the nuclear USP21 and not its cytoplasmic function and all of our purifications identified proteins with a known role in chromatin regulation, amongst those the confirmed interactors POLDIP3, PINX1, HOXC9 and YEATS4, as mentioned in the results section (chapter 1). Interestingly also RING1B was identified in one of the TAP purifications and confirmed by the purification of Flag-USP21 what is in line with the fact that DUBs often bind E3-ligases underlining that ubiquitination and deubiquitination are interdependent processes [97]. Additionally, the chromodomain containing CBX8 was identified as the only further PcG protein in one of the TAP purifications.

On the other hand, in the purification of BEND3 by Flag-tag and by TAP-tag affinity methods USP21 was only detected in one of the Flag purifications. POLDIP3, that was identified as strong putative interactor of USP21 in both TAP-purifications was identified once to bind to TAP-BEND3 however at a low score. BEND3 has been found as interacting protein in several proteomic screens conducted by other groups. It was identified as a weak interactor with all three isoforms of the transcription factor C/EBPβ that can be both activating

and repressing [98]. It also interacts with SIRT7, a sirtuin localized in nucleoli [99]. SIRT7 is an HDAC that has recently been shown to repress gene expression by deacetylating H3K18 [100]. BEND3 has also been identified in affinity purification of SMARCAD1, a protein that plays a role in heterochromatin establishment after replication [101]. All in all it can be concluded that BEND3 interacts with transcription regulators that are in large part associated with repression. The interaction of SMARCAD1 suggests the interplay of BEND3 in the establishment of heterochromatin what is line with its capacity to induce chromatin condensation [102] and for those reasons it would be interesting to study the role of BEND3 in SMARCAD1 target gene expression.

3.2. Validation of USP21/BEND3 interaction

The interaction between BEND3 and USP21 was confirmed by co-immunoprecipitation in overexpression conditions as well as between overexpressed Flag-USP21 and the endogenous BEND3. Moreover, we demonstrated that the interaction takes place independently from the catalytic activity of USP21 and is not due to underlying DNA.

Purification of recombinant USP21 and BEND3

It remains to be ascertained if the interaction is direct or mediated by further protein partners. Several attempts were made to express recombinant USP21 and BEND3 proteins in order to assess the binding by *in vitro* pull-down assays. In the *E.coli* Bl21 strain however both full-length proteins and most of their fragments were localized in bacterial inclusion bodies, where insoluble proteins are stored and the remaining soluble portion of the proteins formed aggregates during purification. A possible explanation is that both proteins depend for their solubility on PTMs and thus when constructs of human USP21 and BEND3 are expressed in bacterial cells they are not modified in the correct manner. Another expression system for recombinant proteins is the insect cell line Sf9 that is derived from *Spodoptera frugiperda* but protein expression levels in Sf9 cells infected with USP21 and BEND3 containing plasmids were very low. It is known that BEND3 is phosphorylated as it was identified as a target in two

phosphorylation screens. One of them was conducted in HeLa cells whose signaling cascades were activated by epidermal growth factor (EGF) and temporal changes in the phosphorylation state of proteins were monitored by mass spectrometry (Olsen et al., 2006). Another study compared the phosphorylation levels of proteins in the undifferentiated human embryonic stem cell line H1 with those after differentiation by retinoic acid and withdrawal of the basic fibroblast growth factor bFGF [103]. Both studies resulted in the identification of three modified residues in BEND3, namely serine S379, S489 or S503 [102]. BEND3 is also sumoylated as found in the study of Vertegaal and colleagues [104].

However, the structure of the BEN domain could also explain the formation of aggregates. Sequence alignment of the amino acid sequences of BEN domains in various proteins revealed several conserved residues as well as the common LhxxlFs motif (l stands for aliphatic-, s for small and x for any residue). As BEND3 contains four copies of this domain it can be suspected that the BEN repeat leads to multimerization [90].

USP21 complex composition

BEND3 was purified with an outstanding high score suggesting that it is the main interactor of USP21 whereas the other confirmed candidates displayed lower scores. This might indicate that the other proteins associate with the BEND3/USP21 complex only transiently or with weaker affinity. It is also possible that the interaction of BEND3 with USP21 excludes the association of the other partners so that there is a second USP21 complex made up by the remaining candidates POLDIP3, PINX1, HOXC9 and YEATS4. A way to address this question is performing gel-filtration analysis that allows the separation of protein complexes by their size because proteins that are present in the same complex elute in the same size fractions. Moreover by increasing stringency of the binding conditions the affinity of subunit interactions can be determined. Western Blot or mass spectrometry analysis of the fractions would then reveal which proteins are present in different USP21-containing fractions. By those means it could be assessed whether there are two distinct USP21

complexes and what stoichiometry they possess, or if BEND3 is in fact the only binding partner of USP21.

3.3. Localization and recruitment of BEND3 and USP21

3.3.1. BEND3 regulates USP21 localization

BEND3 regulates USP21 subcellular localization

As DUBs have many different substrates in different subcellular compartments, their localization must be tightly regulated what is achieved by their recruitment by specific proteins. One possibility is that BEND3 fulfills the function to recruit USP21 to the nucleus. We tried to assess this question by fractionating cells that overexpressed BEND3 in cytoplasm, nucleoplasm and chromatin and analyzing the amount of overexpressed USP21 in the different compartments by Western Blot. Localization of endogenous USP21 could not be monitored due to the lack of suitable antibody and overexpressed USP21 was present in all three cellular compartments. However it was not possible to derive any quantitative conclusions from those experiments because the overexpression of BEND3 resulted in increased expression of exogenous USP21. If BEND3 were the regulator of USP21 subcellular localization it would probably act as a USP21 retaining protein in the nucleus for two reasons. First endogenous BEND3 is a mainly nuclear protein as we have seen in our Western Blot analysis and second USP21 contains a NES that regulates its nucleocytoplasmic shuttling by the CRM receptor [80] suggesting that BEND3 can either have an influence on the shuttling process or regulate USP21 localization in the nucleus independently from CRM.

BEND3 recruits USP21 to chromatin

Another possibility is that BEND3 has the function to direct USP21 that lacks a DNA binding domain to target genes where it can activate genes by deubiquitinating H2A. A DUB that is recruited to target genes by a transcription factor is USP22. This H2B specific DUB is integrated in the hSAGA coactivator complex and binds to chromatin when recruited by the transcription activator MYC. In this case the complete activation of MYC target genes depends on the

presence of the DUB [105]. Thus it is interesting to study if USP21 relies on BEND3 for its association to chromatin. Immunofluorescence studies of BEND3 deletion constructs demonstrated that BEND3 binds to heterochromatic foci and that recruitment is mediated by the fourth BEN domain that resides at the C-terminal end [102] suggesting that BEND3 can fulfill the task to direct USP21 to target genes. The function as a recruiting protein is consistent with the assumption of Abhiman and colleagues who predicted that the BEN domain probably enables protein interaction or interactions between DNA and proteins in the context of chromatin based processes. The first step to address this question is the identification of USP21 and BEND3 target genes by ChIP sequencing. So far it was not possible to conduct this experiment due to the lack of antibodies that specifically precipitate endogenous USP21 and BEND3 in chromatin immunoprecipitation. Hence a second option is to determine target genes of the ectopically expressed proteins and to identify those genes that are bound by both factors. After having validated the common targets by ChIP analysis, the mechanism of USP21 recruitment could be then addressed. For that purpose ChIP experiments will be conducted in the future at identified genes testing first if the binding of USP21 depends on the presence of BEND3 and as a second step analyzing the protein domains that are implicated in the recruitment. It will be verified if BEND3 is directed to genes by its BEN4 since it seems implicated in targeting to transcriptional inactive genomic regions. We already found out by performing co-immunoprecipitations that USP21 binds *in vivo* to the C-terminal region of BEND3 that comprises all BEN domains. Indeed several studies report other members of the BEND family to bind proteins by regions that comprise BEN domains. Apart from BEND3 it will be necessary to map also the interaction surfaces on USP21. The general structure of DUBs is modular, consisting of a catalytic domain with additional insertions and N-terminal and C-terminal extensions. Those extensions contain domains that are directly involved in the regulation of the enzymatic activity and specificity as well as the localization and integration of the DUB in multimeric protein complexes. Domains that are present in the extensions are different kinds of protein-protein interaction domains, ubiquitin-like domains (UBL) and frequently ubiquitin-

binding domains (UBDs) like UBA (ubiquitin associated domain) or Znf-UBP (zinc finger common in UBP DUBs) that confer linkage-specificity to the DUB and enhance their affinity to ubiquitin [29]. The unstructured N-terminus with the NES [80] consists of the first 212 amino acids and displays a high proportion of proline that makes it prone for interaction with other proteins [79] what is characteristic for the additional sequences of DUBs outside their catalytic domain. It was demonstrated that two distinct parts of the N-terminus interact with centrosomes and microtubules, respectively [79] and it was moreover described that it is indispensable for the recognition of H2A in the context of nucleosomes [64]. It would be interesting to investigate whether the N-terminus also mediates the interaction with BEND3. Once the interaction domains have been determined by co-immunoprecipitations the recruitment mechanism should be verified *in vivo* conducting ChIP experiments on the common target genes in cells that are depleted for endogenous wildtype proteins but that are rescued by the ectopic expression of deletion proteins.

3.3.2. USP21 regulates BEND3 localization

USP21 regulates BEND3 subcellular localization

Immunofluorescence studies of overexpressed YFP-BEND3 in NIH3T3 mouse cells showed that BEND3 is present in heterochromatic foci that are characterized by H3K9me3 and the H3K9me3-binding protein HP1 [102]. However little is known regarding the regulation of the import of BEND3 into the nucleus. Moreover, to date there have been no functional nuclear localization sequences (NLS) identified for BEND3 and it is therefore possible that USP21 has the function to retain BEND3 in the nuclear compartment.

USP21 recruits BEND3 to chromatin

The second possibility is that USP21 actually recruits BEND3 to USP21-specific target genes in order to abrogate gene expression by inducing chromatin condensation. This question can be addressed in the course of the proposed ChIP experiments in chapter 3.1. It is however unknown how USP21 itself is recruited to chromatin because it lacks chromatin binding domains.

3.4. Function of USP21 and BEND3 interaction

3.4.1. USP21 as the BEND3 specific DUB

We have demonstrated that overexpressed Flag-BEND3 is polyubiquitinated in 293T cells. In the future, it will be interesting to clarify whether USP21 is the specific DUB of BEND3. This is a likely assumption because in several cases DUBs and their substrates co-purified in complex affinity purifications as for example USP7 and USP11 that associate with the PRC1 complex and deubiquitinate the subunits MEL18 and BMI1.

As described in the introduction (chapter 3.1.) ubiquitin moieties in polyubiquitin chains display different linkage types and as a consequence give rise to different three-dimensional structures. Our Western Blot analysis of immunoprecipitated Flag-BEND3 showed the characteristic high-molecular smear of polyubiquitination. However the chain type remains to be determined and so far data point towards a linkage type different from K48 polyubiquitination. We have tested the turnover rate of endogenous BEND3 by using the translational inhibitor cycloheximide and even at 9h after treatment total BEND3 levels were unchanged indicating that BEND3 is a very stable protein. The same results were obtained when treating cells with the proteasome inhibitor MG-132 that was expected to increase BEND3 protein levels, as it cannot longer be degraded. However no changes in global BEND3 levels were observed suggesting that the fraction of BEND3 that is degraded in the proteasome, and therefore is marked by K48 chains, is too small to lead to a detectable enhancement of BEND3 after proteasome inhibition. Additionally, overexpression of a catalytic active USP21 did not have any influence on BEND3 protein levels underscoring again that in the case that USP21 was the specific DUB, other polyubiquitin chains than K48 would be attached to BEND3. Furthermore, preliminary results indicate that BEND3 is neither modified with K48 nor with K63 chains. Once the linkage type has been determined, the following experiments will focus on analyzing the physiologic function of the modification and how USP21 is involved in that process.

Interestingly overexpressed BEND3 was shown to be sumoylated on lysines K20 and K512. Regarding the impact of that modification on its repressor

function it was described that mutant protein that has both sumoylation sites inactivated is still bound to the transgene array but loses its ability to repress transcription by inducing chromatin condensation. Mechanistically it was further demonstrated that RNAPII is 48% less recruited to the transgene locus in the presence of wildtype BEND3 than in non-transfected cells whereas in cells expressing BEND3 sumoylation mutant RNAPII is recruited to the same extent than in the non-transfected control cells. The same observations were made for the locus recruitment of cyclin-dependent kinase 9 (Cdk9) [102].

Sumoylation has been attributed a regulatory role in different cellular processes but especially in transcriptional repression and in several cases repression is achieved by the recruitment of HDACs to transcriptional cofactors [106]. In line with the two other BEN domain containing proteins NAC1 and SMAR1 that repress genes by recruiting HDACs to chromatin via the BEN domains (chapter 3.1.) BEND3 interacts as well with HDAC2 and HDAC3 *in vivo* [102]. However it remains to be analyzed if the BEN domains recruit the HDACs to chromatin when BEND3 is sumoylated and what impact this has on gene expression. The other possibility how gene repression is achieved by sumoylation of transcription factors is the establishment of facultative heterochromatin. SP3 is a transcription factor that has similar characteristics to BEND3 in sumoylation dependent gene repression. Moreover SP3 acts both as activator and repressor but does not longer activate gene expression when it is sumoylated [107]. Given the possibility that BEND3 recruits USP21 to chromatin the question arises if there is a functional connection between a transcription-activating DUB that interacts with a sumoylated transcription repressor. A first explanation could be the fact that transcription factors as SALL4B [108] act as repressors or activators depending on their sumoylation state and USP21 could bind to BEND3 depending on its sumoylation and therefore activation state. To address the question if USP21 and BEND3 only form a complex when BEND3 is not sumoylated, co-immunoprecipitations should be carried out in cells expressing sumoylation deficient BEND3 mutants and target gene expression levels should be monitored under those conditions.

Moreover the sumoylation and ubiquitination pathways overlap at many points whereby both modifications can occur mutually exclusive on the same lysine residues [109]. In this context it is interesting to mention that using the ubiquitination prediction tool UbPred (http://www.ubpred.org/) BEND3 was predicted to have five high confidence ubiquitination sites amongst them the sumoylation site lysine 512. Preliminary results of co-immunoprecipitations of HA-ubiquitin and Flag-BEND3 K512R mutant in which the lysine 512 was substituted with arginine showed that BEND3 mutant remains polyubiquitinated what can however be explained by the possible ubiquitination of the other predicted lysine residues. Mutants with all lysine sites mutated will allow studying a possible interplay between sumo and ubiquitin in more detail.

Another intersection of the two modifications is the existence of sumo-target ubiquitin ligases (STUbLs) that are supposed to regulate the turnover rate of sumoylated transcription factors. This enzymatic class comprises E3-ligases that recognize sumo-modified proteins and target them via ubiquitination for proteasomal degradation [109]. It has not been studied if BEND3 protein levels are controlled by STUbLs but if so USP21 as a DUB could interfere in that process by editing the polyubiquitin chains.

3.4.2. BEND3 as a regulator of USP21 catalytic activity

Apart from a possible role of BEND3 in recruiting USP21, the interaction could have also an effect on the cysteine protease activity of USP21. Cysteine proteases cleave the isopeptide bonds that are either established between two ubiquitin molecules or a target that is bound to ubiquitin. The catalytic mechanism bases on a nucleophilic attack of the partially positive charged cysteine towards the peptide bond. The reaction step is supported by two other amino acids, namely histidine and aspartic acid or asparagine, respectively, that participate in stabilizing and augmenting the polarization of cysteine. The three amino acids located in the catalytic center are called the catalytic triad and are essential for the enzymatic activity [27]. The catalytic activity of DUBs is controlled at different steps. Firstly, the transcription of DUBs underlies a temporal control to ensure that DUBs are only present in a cell during certain

physiological conditions [29]. In the case of USP21 it was observed that USP21 expression is induced during liver regeneration after partial hepatectomy [64]. Additionally enzymatic activity of DUBs also depends on posttranslational modifications like phosphorylation, ubiquitination and sumoylation [29]. For USP21 no functional posttranslational modifications have been described so far but its phosphorylation *in vivo* on threonine, serine, tyrosine and lysine has been reported in several studies (http://www.phosphosite.org).

Several DUBs (*e.g.* HAUSP) are activated by substrate interaction. When crystallizing the catalytic domain of HAUSP (USP7) it was observed that the catalytic domain is a made up by a finger, thumb and a palm domain what is the common feature of USPs. Thumb and palm domain form the catalytic cleft with the catalytic triad in their interface. Amino acid residues that have key functions for the establishment of the three-dimensional subunit arrangement are highly conserved in different USPs pointing towards a common structure and catalytic mechanism. However, the catalytic residues are at too big distance to each other so that their necessary polarization for the nucleophilic attack would not be possible. Crystallization of HAUSP in complex with ubiquitin aldehyde, a modified ubiquitin dimer, revealed that ubiquitin binds with its C-terminus inside the cleft and that the distal finger domain is involved in binding the substrate. HAUSP therefore adopts a structure like a hand grabbing around the ubiquitin. In this binding state localization of the catalytic residues is rearranged in a way that they were now in the right position to permit the hydrolysis [110]. Besides the strength of the interaction with ubiquitin and the linkage type, the number and kind of UBDs that are present in a DUB [76] further enhance specificity. Regarding USP21 however there is no information about UBDs available.

DUBs also interact with proteins other than the substrates and these interactors often act as allosteric activators as for example UAF1 that forms a complex with USP1, USP12 and USP46 via its 8 repeats of the WD40 domain. The complex formation with UAF1 enhanced the catalytic activity for all three USPs [111, 112]. For USP21 there no interactors are known that impact on its catalytic activity and therefore we started to study the influence of BEND3 on USP21 mediated deubiquitination. As we could not purify recombinant BEND3 and

USP21 proteins to perform *in vitro* deubiquitination assays we focused on USP21 deubiquitination activity on H2Aub *in vivo* as a reporter system to monitor the catalytic activity. Regarding H2Aub changes on global a level we have observed that overexpression of BEND3 does not alter the total amount of H2Aub associated to chromatin. Nevertheless it is conceivable that BEND3 modulates H2Aub only on selected target genes and that would not be apparent by looking at the bulk chromatin. The identification of direct USP21 target genes however will provide a new tool to analyze by ChIP if USP21 activity on H2Aub is modulated by BEND3 and which domains are implicated.

3.4.3. USP21 and BEND3 in H2Aub-mediated gene expression

In chapter 3 we have discussed the possible mutual recruitment of USP21 and BEND3 to gene promoters. Nevertheless BEND3 and USP21 have opposing functions in gene regulation: USP21 was described as a putative activator by regulating the trans-histone crosstalk between H2Aub and H3K4me3 [64] whereas BEND3 was reported to suppress gene transcription by inducing chromatin condensation and preventing the recruitment of the RNAPII. More precisely it was shown that BEND3 modulates the transcription of an array of reporter loci stably introduced into U2OS cells. Experiments demonstrated that ectopically expressed YFP-BEND3 represses promoter activity by inducing chromatin condensation and moreover impaired the transcription onset by preventing the recruitment of the RNAPII pre-initiation complex [102].

This apparent paradox of the interaction of an H2Aub-specific DUB that activates transcription with a transcription repressor (BEND3) might be explained by a new role of H2Aub in gene expression. Scheuermann and collaborators reported the discovery of CALYPSO, a *Drosophila* PcG protein with specific deubiquitination activity for H2Aub that colocalizes with PcG proteins at PREs and surprisingly represses HOX genes by its DUB activity. In contrast to the common model of DUBs counteracting E3-ligase function, SCE, the *Drosophila* PRC1 E3 ligase, and PR-DUB both contribute to the correct execution of embryonic development. Assuming that SCE and PR-DUB

regulate the same subset of genes in that context one can think that H2Aub levels must be accurately regulated determining in that way the adequate gene expression level [70].

Another H2A specific DUB that regulates Polycomb-repressed HOX genes is USP16. It was shown to localize together with the PRC2 subunit BMI1 to HOXD10 promoter and 5' regulatory region and to deubiquitinate H2Aub in that area. In this case however loss of catalytic active USP16 was followed by an increase in HOXD10 mRNA expression levels what indicates that the transcriptional outcome of H2Aub levels might be different. Interestingly the presence of BMI1 protein on HOXD10 promoter and 5' regulatory region was independent of USP16 [69].

A DUB that might have a dual regulatory function in gene expression is USP22 that deubiquitinates H2A and H2B [105, 113] and is integrated in the human SAGA complex, a multimeric protein complex that possesses the histone acetyltransferase subunit GCN5 and is recruited to active genes [114]. USP22 has been demonstrated to activate transcription when being recruited to gene promoters by MYC [105] and it is also part of the death of cancer signature that consists of 11 proteins with role in cancer biology, among them BMI1 and RNF2 and several BMI1-regulated genes [115]. Those findings give rise to the suspicion that USP22 interferes in the expression of PcG genes during metastasis and it remains to be clarified if it regulates the same genes as BMI1 and which role its DUB activity plays in that context.

Taking in mind the previously mentioned examples, USP21 could have the task to fine-tune BEND3 repressed genes. Interestingly BEND3 has the ability to induce premature chromatin condensation when it is ectopically expressed and immunofluorescence showed that EZH2 and H3K27me3 colocalize to those heterochromatin foci [102]. This is in line with the finding that also EZH1 has been shown to compact chromatin [116]. From that observation one can suspect that BEND3 target genes might overlap with PRC2 targets and further one can speculate that BEND3 recruits USP21 to those sites where the DUB together with PRC1 would then determine gene expression levels by fine-tuning the ubiquitination state of H2A. In order to address that question it would be first

necessary to correlate H2Aub at the common target genes with their expression levels in ChIP analysis by using catalytic mutants of USP21 or depleting cells for BEND3 or USP21. From that one could conclude if USP21 and BEND3 act together in an activating or repressive context. Secondly overlapping those targets with known PRC1 target genes would show if USP21 and BEND3 also bind to PcG target genes. In favor of that is the fact that RING1B was identified as a putative interactor of USP21 in the TAP- and Flag purifications and that interaction was also observed in a co-immunoprecipitation of endogenous RING1B with overexpressed USP21 in 293T cells (Figure 14). Once a subset of co-regulated genes would be identified it could be further assessed how gene expression is regulated by USP21 and BEND3 and if the two proteins are involved in PRC1-mediated silencing mechanism.

3.5. BEND3 and USP21 in mESC

3.5.1. BEND3 as a part of the OCT4 interactome

The proteomic analysis of the OCT4 interactome in mECS led to the identification of many proteins with a role in the regulation of the stem cell state. Amongst them were most of the subunits of the NuRD complex and also many transcription factors like SALL4, ESRRB, TCFCP2L1, KLF5 or SOX2 whose essential function in mESC has already been characterized [96]. Analysis of the genome-wide binding sites of SALL4 in mESC revealed that this transcription factor targets more than 3000 genes involved in stem cell biology and many of those form part of signaling pathways that govern the differentiation and lineage commitment process. It was further demonstrated that SALL4 forms a complex with the core transcription factors OCT4 and NANOG and all three proteins co-occupy a subset of SALL4 target genes [117]. The involvement of SALL4 in the core transcriptional network and differentiation process was further confirmed by the finding that SALL4 and OCT4 positively regulate each others expression by direct binding to the respective promoters and that loss of SALL4 leads to enhanced expression of endodermal, ectodermal and mesodermal lineage markers. Besides that SALL4 turned out to terminate the positive feedback loop with OCT4 by repressing its own transcription. SALL4 therefore acts as an

activator and a repressor on different genes [118]. With respect to NANOG target genes it has been recently shown that ESRRB (estrogen-related receptor b) plays a pivotal role for LIF-independent self-renewal of mESC. Cells that overexpressed ESRRB were unable to differentiate when cultured in conditions when the LIF pathway was not activated [119]. Another transcription factor identified in the OCT4 interactome identified by Berg and colleagues is TCFCP2L1 and interestingly many PcG target genes contain its DNA binding motif [120].

The reciprocal complex purification of the putative OCT4 interactors DAX1, SALL4, ESRRB and TCFCP2L1 revealed that they interact directly with each other or via shared subunits reinforcing the idea of an intertwined transcriptional network. Interestingly co-purification of BEND3 was observed in SALL4, ESRRB and TCFCP2L1 purifications pointing towards a role of BEND3 in the regulated interplay of those transcription factors. BEND3 was the candidate with the second highest score in the SALL4 purification after SALL1, another member of the family of spalt-like transcription factors, and for that reason BEND3 could have an important role in SALL4 mediated transcription. As BEND3 is a transcription repressor it can be suspected that it might act together with SALL4 on repressed genes. Surprisingly in the purifications of all four factors subunits of the repressive NuRD complex that contains two HDACs were detected [96]. The binding of SALL4 to NuRD has also been documented in another study showing that SALL4 represses gene expression at promoters that are co-occupied by the NuRD complex [121]. NuRD has a crucial function in maintaining the stem cell state as it was described that cells lacking the complex subunit MBD3 kept their capacity to self-renew upon LIF withdrawal and were unable to run through the differentiation program [122]. NuRD shares with PRC2 a subset of target genes that are essential for embryonic development and signaling in mESC. Moreover, it was demonstrated that PRC2 only binds to target genes when the NuRD complex with intact HDAC activity is present [123]. The results of Berg and colleagues showed an interaction of OCT4 and TCFCP2L1 with subunits of the PRC1 complex besides the identified NuRD subunits and in the context with the work of Reynolds thus imply that

target genes of the OCT4 network can be repressed by the transcription factors of the OCT4 interactome together with PRC complexes and NuRD. The presumable role of PRC complexes in the gene regulation mediated by OCT4-associated factors suggests that H2Aub levels are also implicated in that regulation.

3.5.2. DUBs in stem cell differentiation

Until now no H2A specific DUBs have been described that are important for pluripotency or differentiation of mESC. However two different studies reported recently that H2A specific 2A-DUB is a regulator of hematopoietic stem cells, particularly of B-cells [124, 125]. 2A-DUB regulates the expression of EBF1, the key transcription factor for B cell differentiation and analysis of the changes on EBF1 promoter upon 2A-DUB loss revealed an increase in H2Aub and as well of H3K27me3 but a loss of H3K4me3. The changes in the histone modification states were accompanied by a loss of transcription factor binding and at the same time the recruitment of the PRC1 components BMI1 and RING1B [125]. This example reveals that DUBs are implicated in stem cell differentiation in concert with PcGs. Elucidating the role of USP21 in mESC as an interactor of BEND3 - which is part of the OCT4 interactome - is very promising as in 293T cells global effects on H2Aub levels were observed upon overexpression of USP21 what suggests its important regulatory role.

BEND3 is a transcription factor whose repressive function depends on sumoylation of K20 and K512 [102] and interestingly the transcriptional activities of several core transcription factors in mESC rely on this modification, as well. As specified previously SALL4 interacts with OCT4 on protein level and additionally acts as a transcription activator of the OCT4 gene whereas it represses its own gene promoter. Besides that SALL4B, one of the two isoforms of SALL4, is subject of several PTMs among those sumoylation on four lysine residues and this modification enhances the interaction with OCT4. On transcriptional level sumoylation decreases the activation of OCT4 promoter as it also decreases the repressive function of SALL4B on its own gene promoter [108]. Another example for modulation of the stem cell state by

sumoylation is the control of NANOG expression that fulfills a key function in the maintenance of self-renewal. A heterodimeric complex consisting of OCT4 and SOX2 activates NANOG promoter and their interaction is disrupted upon sumoylation. Opposing effects of the sumoylation were observed at gene expression levels as covalently modified OCT4 entailed an increase in NANOG transcription whereas NANOG expression was repressed when SOX2 was bound to SUMO [126]. Considering the importance of sumoylation in the regulatory network of mESC it will be interesting to study how BEND3 sumoylation impacts on its transcriptional activity on target genes in mESC and how its integration into the OCT4 network is influenced.

3.5.3. Physiological role of BEND3 in mESC

Global gene expression analysis in mESC depleted for endogenous BEND3 was performed to understand the physiological context of BEND3. As BEND3 has been shown to be a transcription repressor [102] we expected more genes to be upregulated than downregulated upon BEND3 knockdown. However 402 genes were downregulated compared to 354 upregulated genes what is explainable by the fact that the observed deregulation of genes is also due to secondary effects on transcription upon knockdown. Classification of the deregulated genes in BEND3 depleted mESC corresponding to their GO categories suggested that both up-regulated and down-regulated genes assume roles in the category “cell growth and proliferation” and down-regulated genes are additionally assigned to the group of “cellular development”. Preliminary experiments revealed that BEND3 does not significantly change its expression upon differentiation by retinoic acid or LIF withdrawal but that does not exclude that BEND3 might still have a regulatory function during the differentiation process. mESC are characterized by a high proliferation rate and interestingly many of the upregulated genes belong to the category “cell cycle” and among those genes were for example cyclins. Moreover Sathyan and coworkers reported that U2OS cells upon overexpression of YFP-BEND3 are arrested in early S-phase and thus our future experiments will focus on analyzing a possible role of BEND3 in cell cycle regulation and also in

differentiation. Finally it will be important to know which of the genes identified in the microarray that fulfill essential functions in the above mentioned categories actually are direct BEND3 targets what will be determined by comparing them to the genes identified in ChIP sequencing.

4. Conclusions

USP21 regulates global H2Aub levels in 293T cells.

BEND3 is the main interactor of nuclear USP21.

USP21 interacts with BEND3 via BEND3 C-terminus.

Endogenous BEND3 localizes predominantly at chromatin and in nucleoplasm.

Ectopically expressed BEND3 does not alter global H2Aub levels.

Ectopically expressed BEND3 is polyubiquitinated in cells.

Endogenous BEND3 protein displays a low turnover rate.

BEND3 subcellular localization and protein levels are independent from USP21 protein and catalytic activity.

BEND3 regulates the expression of hundreds genes in mESC that are mainly involved in cellular growth and proliferation, cell death and cell cycle.

5. Material and Methods

Cell lines and culture conditions

Human cell lines were grown in DMEM medium supplemented with 10% fetal bovine serum (FBS), penicillin and streptomycin at 37°C and 5% CO_2. mESC cells were cultured in Glasgow medium with 2µl/ml LIF and supplemented with 20% fetal bovine serum (FBS) at 37°C and 5% CO_2. Tissue culture plates for mESC were previously covered with a 5% of gelatin solution in PBS by incubating them for 20min at 37°C.

Cells were trypsinized for freezing and the reaction was stopped by medium addition. Cells were then washed with PBS, resuspended in cold freezing medium (90% FBS and 10% dimethyl sulfoxide (DMSO)) and frozen at -80%. For long-term storage cells were transferred to the liquid nitrogen tank. Frozen cells were thawed quickly in a 37°C warm water bath, washed in 10ml preheated medium and seeded on culture plates.

Transfections

For the calcium phosphate method cells were seeded the day before obtaining about 75% confluency at transfection. The adjusted total DNA amounts were resuspended in water and calcium phosphate was added to a 0,15M final concentration. The same volume of 2x HBS buffer was added dropwise while vortexing the tube and tubes were subsequently incubated for 10min at room temperature before adding the transfection mix to the cells. Medium was replaced after 14h and cells were harvested after 36h for expression analysis.

Lipofectamine transfections were made following the instructions of the Invitrogen manual.

Stable cell lines

HEK 293 cells were transfected by calcium phosphate method and the medium was changed 14h after transfection. Selection with 1,5µg/µl puromycin or 250µg/µl hygromycin was started 24h after removing the medium. Selection with 500-600µg/µl geneticin was started 36h after transfection.

mESC were transfected by lipofectamine method and selection with 500µg/µl geneticin was started 24h after transfection.

EcR-293 cells that stably expressed the pVRXR construct with bleocin resistance were cultured in DMEM with 10% FBS and 30 µg/ml bleocin and transfected in 100mm dishes by calcium phosphate method with the ponasterone-inducible expression vector pMZI that contains geneticin resistance and the gene of interest. 14h later medium was replaced with medium containing additionally 300µg/µl geneticin and half of the cells were passed to a 150mm dish. Clonal cell lines were established and gene expression was induced with 3µM ponasterone.

Lentiviral infections

293T cells were seeded for virus production at a density of 10^5 cells per 100mm dish and transfected by calcium-phosphate method with a DNA mix consisting of the corresponding short hairpin plasmid (7µg) and the packaging vectors encoding VSV-G (5µg) and dR8,91 (6µg). Medium was changed 14h after transfection, in the case of mESC infections with Glasgow medium. The viral supernatant was taken after 72h for over night infection at 37°C. mESC infections were done using polybrene in a 1:1000 dilution.

Short hairpin constructs used in this work (Sigma, MISSION shRNA)

shRNA	**sh USP21 human,** pLKO.1-puro
349	CCGGGTGCTCCATCTGAATCGATTTCTCGAGAAATCGATTCAGAT GGAGCACTTTTTG
439	CCGGGTGCTCCATCTGAATCGATTTCTCGAGAAATCGATTCAGAT GGAGCACTTTTTG
440	CCGGGACCCTCTGCAATATCACTTTCTCGAGAAAGTGATATTGCA GAGGGTCTTTTTG
897	CCGGCCACTTTGAGACGTAGCACTTCTCGAGAAGTGCTACGTCT CAAAGTGGTTTTTG
949	CCGGTCCCATCTCGGACCAACTTAGCTCGAGCTAAGTTGGTCCG AGATGGGATTTTTG
	sh USP21 mouse, pLKO hygro
667	shUSP21_mouse_667_F CCGGCCACTTTGAGACGTAGTACTTCTCGAGAAGTACTACGTCTC AAAGTGGTTTTTG shUSP21_mouse_ 667_R AATTCAAAAACCACTTTGAGACGTAGTACTTCTCGAGAAGTACTA CGTCTCAAAGTGG
385	shUSP21_mouse_385_F CCGGCCCAGATGAAAGGCTCAAGAACTCGAGTTCTTGAGCCTTT CATCTGGGTTTTTG shUSP21_mouse_ 385_R AATTCAAAAACCCAGATGAAAGGCTCAAGAACTCGAGTTCTTGAG CCTTTCATCTGGG
	sh BEND3, mouse pLKO.1-puro
216	CCGGCCTAGCAAGTTACGCAATCAACTCGAGTTGATTGCGTAACT TGCTAGGTTTTTTG
224	CCGGGATCCAGAAGATGTTCTACATCTCGAGATGTAGAACATCTT CTGGATCTTTTTTG
391	CCGGCAATGGACTCACGACTAATATCTCGAGATATTAGTCGTGAG TCCATTGTTTTTTG
685	CCGGCGTGATAACTTTAATGGCTTACTCGAGTAAGCCATTAAAGT TATCACGTTTTTTG
798	CCGGCTATGAATGTATACCTAGCATCTCGAGATGCTAGGTATACA TTCATAGTTTTTTG

Reagents used in cell culture

Polybrene	Sigma
MG-132	Calbiochem
Cycloheximide	Sigma
LIF	Produced in the lab

Cell extracts and immunoprecipitations

Whole cell extract

Cells were resuspended in cold lysis buffer with protease inhibitors (cOmplete Protease Inhibitor Cocktail Tablets, Roche). Two different lysis buffers were used in the experiments:

50mM Tris pH7.4, 150mM NaCl, 1mM EDTA, 1mM EGTA and 0,5% Triton

or

50mM Hepes pH7.5, 150mM NaCl, 1mM EDTA, 2,5mM EGTA and 0,1% Tween.

Resuspended cells were sonicated twice in Branson sonicator for 15sec at 15% and samples were incubated afterwards for 30min on ice. Cellular debris was removed by centrifugation at 13000rpm for 20min. Protein concentrations were determined by Bradford Assay.

Immunoprecipitations were performed over night with the respective specific antibodies or control IgG of 500µg -1mg total protein. Samples were incubated with blocked Protein A or Protein G beads for 2h and precipitates washed abundantly with lysis buffer. Beads were resuspended in 2x Laemmli buffer and boiled. Eluted proteins were loaded on a SDS-Page and analyzed by Western Blot.

Immunoprecipitation of ubiquitinated proteins of whole extract [127]

293T were transfected in 100mm dishes with Flag-BEND3 and HA-ubiquitin (5-fold excess) by calcium-phosphate method in the morning and medium was changed in the evening. The next day morning cells were treated with 5-10µM MG-132 for 5-6h before harvesting one 100mm dish in 150µl cold lysis buffer (50mM Tris-HCl, pH7.4, 0.25M NaCl, 0.1% Triton X-100, 1mM EDTA, 50mM NaF, 1mM DTT, 0.1mM Na3VO4) with 1mM o-vanadate, 20mM glycerophosphate, 10µm *N*-ethylmaleimide (NEM, Gibco), and protease inhibitors. SDS was added to a 1% final concentration and samples were boiled for 10min. Samples were diluted by adding 75µl 20% Triton and 1260µl lysis buffer, sonicated at 15% and incubated 30min on ice before pelleting the cellular debris at 13000rpm for 20min. Immunoprecipitations were performed

over night using Flag-beads or IgG of 500µg -1mg total protein. Samples were incubated with blocked Protein A beads for 2h and precipitates were washed abundantly with lysis buffer. Beads were resuspended in 2x Laemmli buffer and boiled. Eluted proteins were loaded on a SDS-Page.

Cell fractionation

Cells expressing the bait protein and control cells expressing the empty vector were harvested and washed in PBS. Cells were resuspended in a volume of low salt buffer A (10mM Hepes pH7.9, 1,5mM $MgCl_2$, 10mM KCl, 0,5mM DTT and protease inhibitor) corresponding to 10x of the packed cell volume (PCV) and incubated 5min on ice. NP-40 was added to a final concentration of 0,3%, cells were carefully resuspended by pipetting and incubated for further 3min After that cells were resuspended again by pipetting in order break the cell membranes. Nuclei were pelleted either for 10min at 16000 g at 4°C (TAP purifications) or at 3000 rpm at 4°C (remaining fractionation experiments). The cytoplasmic supernatant was stored on ice and nuclei were lysed in one volume of high salt buffer C (20mM HEPES pH7.9, 25% glycerol, 420mM NaCl, 1,5mM $MgCl_2$, 0,2mM EDTA, 0,5mM DTT and protease inhibitor) while vortexing every ten minutes thoroughly during 40min in order to break the nuclear membrane. Nuclear extract was separated from chromatin by centrifugation 16000 rpm at 4°C. Chromatin in the cell fractionation experiments was either resuspended in Laemmli loading buffer or in lysis buffer (25mM Tris-HCl pH 7.6; 1% SDS; 1mM EDTA; 1mM EGTA and protease inhibitors) and sonicated and boiled before loading it on a SDS-PAGE.

Protein detection

Protein extracts were separated on a SDS-PAGE (sodium dodecyl sulfate polyacrylamide gel electrophoresis) and coomassie or silver staining directly visualized proteins.

Western Blot analysis and immunostaining

Proteins were transferred by Western Blot from the SDS-PAGE to a nitrocellulose membrane. The membrane was blocked with TBS with 2%Tween and 5mg/ml bovine serum albumin (BSA) and subsequently incubated with primary antibody against the specific proteins. The membrane was incubated with secondary antibodies directed against the specie of the first antibody after washing with TBS with 2% Tween. Secondary antibodies were coupled to horseradish peroxidase (HRP) and proteins were detected by ECL chemiluminescence reaction.

In case of detecting ubiquitinated proteins the membrane was denatured by Western Blot denaturing Buffer (6M guanidine-HCl, 20mM Tris-HCl, pH7.5, 5mM ß-mercaptoethanol) for 30min at 4°C previous to immunostaining in order to expose the ubiquitin antigen.

Antibodies used in this work

Primary antibody	**Specie**	**Company**
HA	mouse	Lab Di Croce
HA	rabbit	HA.11, PRB101P, Covance
RING1B serum	rabbit	Lab Di Croce
FLAG	mouse	Flag M2, Sigma
Tubulin	mouse	ab7291, Abcam
c-MYC	rabbit	N-262, Santa Cruz
H3	rabbit	Abcam
H2A	rabbit	Abcam
H2Aub	rabbit	Cell signaling
BEND3 N-terminal (Serum 1219, purified)	rabbit	Lab Di Croce
BEND3 N-terminal (Serum 1218)	rabbit	Lab Di Croce
BEND3 C-terminal (Sera 1271 and 1272)	rabbit	Lab Di Croce
USP21 C-terminal (Sera 1314 and 1315)	rabbit	Lab Di Croce
USP21 Genscript (Sera 1 and 3)	rabbit	Lab Di Croce
Secondary antibody	**Specie**	**Company**
anti mouse	polyclonal, rabbit	Dako
anti rabbit	polyclonal, goat	Dako

Stripping of nitrocellulose membranes

Primary and secondary antibodies were removed from nitrocellulose membranes by incubation for 30min on a shaker with stripping buffer (100mM 2-mercaptoethanol; 2% sodium dodecyl sulphate; 62,5mM Tris-HCl pH6.7). The membrane was blocked again with TBS with 2%Tween and 5mg/ml bovine serum albumin (BSA) after washing with TBS with 2% Tween, and reused for immunostaining.

Tandem affinity purification (TAP)

Cells were fractionated as described in chapter 5.3.1. USP21 and BEND3 were purified from 75mg and 78mg nuclear extract, respectively, and the high salt concentrations of the nuclear extracts were diluted with 1 volume 2x NEEB buffer (1,5mM $MgCl_2$, 20µM $ZnCl_2$, 0,6% NP-40 and 0,2mM EDTA). Proteins were bound for 4h to IgG beads that had been previously washed in IPP buffer (10mM Tris pH8, 100mM NaCl, 0,1% Triton X100, 0,5mM EDTA and 10% glycerol) and unbound proteins were removed by washing with IPP buffer. TEV cleavage of proteins from the IgG beads was realized overnight in TEV buffer (10mM Hepes-KOH pH8, 150mM NaCl, 0,1% NP-40, 0,5mM EDTA and 1mM DTT). Calmodulin beads equilibrated in CCB buffer (10mM Tris pH8, 100mM NaCl, 1mM imidazole, 1mM magnesium acetate, 2mM $CaCl_2$, 0,1% Triton X, 10% glycerol and ß-mercaptoethanol (7µl/10ml buffer)) were incubated with the TEV eluate and 150µl TEV buffer obtained from washing the IgG beads. 5µl of 1M $CaCl_2$ was added to enable binding during 2h. Calmodulin beads were washed in CBB buffer and proteins eluted in CEB buffer (10mM Tris pH8, 100mM NaCl, 1mM imidazole, 1mM magnesium acetate, 2mM EGTA, 10% glycerol and ß-mercaptoethanol (7µl/10ml buffer)). The eluates were separated on a SDS-PAGE and protein bands were visualized by silver staining.

Recombinant protein expression

GST and His fusion expression constructs were transformed in *E.coli* Bl21 strain and protein expression was induced at OD 0,6 with 0,1 to 2mM Isopropyl

β-D-1-thiogalactopyranoside (IPTG) at temperatures between 17°C and 37°C up to 20h. For His constructs cells were lysed by sonication in 50mM NaH_2PO_4, 300mM NaCl, 10mM Imidazole, 0,1% glycerol, protease inhibitors and either 0,05% Tween or 0,5% Triton at pH8. For GST constructs cells were lysed by sonication in 50mM Tris-HCl pH8, 10mM EDTA, 150mM NaCl and protease inhibitors. His fusion proteins were fused to Ni-NTA beads and GST fusion proteins to glutathione-sepharose beads. His fusion proteins were eluted with 250mM imidazole and GST fusion proteins were eluted with 50mM Tris-HCl pH8, 10mM DTT and 15mM to 200mM reduced glutathione at pH8. Eluates were changed to 1x PBS with 5% glycerol using Viva Spin columns. Recombinant proteins were checked by SDS-PAGE with subsequent Coomassie staining.

Antibody purification

Immune sera where purified by affinity purification using recombinant protein coupled to glutathione-sepharose beads. Recombinant protein was bound to beads in 1x PBS with 1mg/ml BSA and after that cross-linked in 0,1M triethanolamine in PBS with 6,5 mg/ml dimethyl pimelimidate (DMP) pH8. Quenching buffer contained 50mM ethanolamine in PBS and uncrosslinked protein was later eluted twice from beads with 0,2M glycine in PBS at pH2.5 for 5min. Beads were first washed in 1M Tris-HCl pH8 and then in PBS with 0,2% Triton between washing and stored in the same buffer with sodium acetate. Immune serum was purified with the protein-coupled beads. Beads were washed (50mM Tris-HCl pH8; 500mM NaCl) after passing the serum through the column and bound antibody was eluted with 100mM glycine, pH2.5. The buffer of the eluted antibodies was changed to PBS with 1% Triton and 10% glycerol in Viva Spin columns. Purified antibodies were checked by SDS-PAGE followed by coomassie staining.

Isolation of total RNA and reverse transcription

Total RNA was isolated from cells previously washed with PBS following the instructions of the RNeasy Mini Kit (Qiagen). RNA used in gene expression microarray was purified using the miRNeasy Mini Kit (Qiagen).

cDNA was obtained by reverse transcription from 1µg RNA per reaction using the first Strand cDNA Synthesis Kit and oligo-dT primer according to the manufacturer's instructions.

Real-Time quantitative PCR (qPCR)

qPCR was performed using LightCycler 480 SYBR Green I Master (Roche) and cDNA was diluted 1:10 in water.

Real-Time PCR program

Temperature	Time	Cycles
95°C	8min	x1
95°C	10sec	x55
60°C	10sec	
72°C	8sec	
95°C	5sec	melting curve
65°C	1min	
98°C	continuous	
4°C	∞	cooling

Table of primer sequences used in this work

	human
USP21	RT2_USP21_ F_874: CCCCTACCCTGTTCAGCAT
	RT2_USP21_ R_1001: GCCCAGAAGGAGTGTGTGAT
BEND3	BEND3_qPCR_3_F: GAGCCTCTCCGTGGGCAACTTTG
	BEND3_qPCR_3_R: TTCTCCACCGGGTAGACGGCTTC
BEND3	BEND3_qPCR_4_F: AGATCCCCTTGGACGAGCTGGTG
	BEND3_qPCR_4_R: CAAAGTTGCCCACGGAGAGGCTC
GAPDH	F TCTTCTTTTGCGTCGCCAG
	R AGCCCCAGCCTTCTCCA

	mouse
USP21	Ms_RT_USP21_815_F CAAGATGGCTCACCACACAC
	Ms_RT_USP21_892_R CAAACACTGTAGCACGGCAT
USP21	Ms_RT_USP21_1736_F GTATGCCCTTTGCAACCACT
	Ms_RT_USP21_1817_R AAACGCGGGAGTCATTGTAG
BEND3	Ms_RT_BEND3_2671_F CAACAGGAAAAAGTGCGACA
	Ms_RT_BEND3_2888_R CTTGTGGCTGGGTAGGTGTT
BEND3	Ms_RT_BEND3_2869_F AACACCTACCCAGCCACAAG
	Ms_RT_BEND3_3021_R GAAGGGACCCTACCTACCA
CLDN4	mCldn4_RT_440_F TCTACAACCCTATGGTGGCTTC
	mCldn4_584_R GAGTAGGGCTTGTCGTTGCTA
RASGRP1	mRasgrp1_RT_2200_F AAACCAAAACCAGAACTTCACCT
	mRasgrp1_RT_2304_R CTTCAGTTGGATCTTCAGAGCAT
TESC	mTesc_RT_460_F GCTCGGTCCATTGCAGAC
	mTesc_RT_609_R ACGAATGTGCATCTTGGTCTC
CIDEA	mCidea_RT_444_F AGCCACGATGTACGAGATGTACT
	mCidea_RT_544_R GTGCAGCATAGGACATAAACCTC
RHOX10	mRhox10_RT_287_F ATGCCCAAATGTGTGAACTAGAG
	mRhox10_RT_387_R GCATTCGTCCACATCAATAAGTT
EGR4	mEgr4_RT_31_F GACGCGCTTCTCTCCAAG
	mEgr4_RT_158_R CTCAAAGCCCAGCTCAAGAAG
CD37	mCd37_RT_582_F AAGCTCTTTTTCTCCCAGCTAAG
	mCd37_RT_726_R GATATTGTTGTGCAGCCACTTCT
STRA8	mStra8_RT_1121_F CAACTCAGAAAATCCAGAGGAGA
	mStra8_RT_1251_R AAGGTCTCCAGGCACTTCAG
GAPDH	F GTATGACTCCACTCACGGCAAA
	R TTCCCATTCTCGGCCTTG

Gene expression microarray

Microarray analysis was performed of mESC infected with short hairpin RNA against BEND3 and a non-specific control hairpin. Of each condition four biological replicas were taken at different days after infection and RNA was isolated and hybridized to Agilent Mouse 8x60K expression arrays. Expression data were analyzed with the ingenuity software and genes with a p-value lower than 0,05 and an expression change greater than 1,5 fold were subjected to gene ontology analysis.

6. References

1. Evans, M.J. and M.H. Kaufman, *Establishment in culture of pluripotential cells from mouse embryos.* Nature, 1981. **292**(5819): p. 154-6.
2. Young, R.A., *Control of the embryonic stem cell state.* Cell, 2011. **144**(6): p. 940-54.
3. Niwa, H., *How is pluripotency determined and maintained?* Development, 2007. **134**(4): p. 635-46.
4. Meshorer, E. and T. Misteli, *Chromatin in pluripotent embryonic stem cells and differentiation.* Nat Rev Mol Cell Biol, 2006. **7**(7): p. 540-6.
5. Okada, Y., et al., *Retinoic-acid-concentration-dependent acquisition of neural cell identity during in vitro differentiation of mouse embryonic stem cells.* Dev Biol, 2004. **275**(1): p. 124-42.
6. Takahashi, K. and S. Yamanaka, *Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors.* Cell, 2006. **126**(4): p. 663-76.
7. Bannister, A.J. and T. Kouzarides, *Regulation of chromatin by histone modifications.* Cell Res, 2011. **21**(3): p. 381-95.
8. Quina, A.S., M. Buschbeck, and L. Di Croce, *Chromatin structure and epigenetics.* Biochem Pharmacol, 2006. **72**(11): p. 1563-9.
9. Buschbeck, M., et al., *The histone variant macroH2A is an epigenetic regulator of key developmental genes.* Nat Struct Mol Biol, 2009. **16**(10): p. 1074-9.
10. Felsenfeld, G. and M. Groudine, *Controlling the double helix.* Nature, 2003. **421**(6921): p. 448-53.
11. Kouzarides, T., *Chromatin modifications and their function.* Cell, 2007. **128**(4): p. 693-705.
12. Chang, B., et al., *JMJD6 is a histone arginine demethylase.* Science, 2007. **318**(5849): p. 444-7.
13. Pray-Grant, M.G., et al., *Chd1 chromodomain links histone H3 methylation with SAGA- and SLIK-dependent acetylation.* Nature, 2005. **433**(7024): p. 434-8.
14. Shi, X., et al., *ING2 PHD domain links histone H3 lysine 4 methylation to active gene repression.* Nature, 2006. **442**(7098): p. 96-9.
15. Doyon, Y., et al., *ING tumor suppressor proteins are critical regulators of chromatin acetylation required for genome expression and perpetuation.* Mol Cell, 2006. **21**(1): p. 51-64.
16. Miranda, T.B. and P.A. Jones, *DNA methylation: the nuts and bolts of repression.* J Cell Physiol, 2007. **213**(2): p. 384-90.
17. Hochstrasser, M., *Origin and function of ubiquitin-like proteins.* Nature, 2009. **458**(7237): p. 422-9.
18. Xu, P., et al., *Quantitative proteomics reveals the function of unconventional ubiquitin chains in proteasomal degradation.* Cell, 2009. **137**(1): p. 133-45.
19. Trempe, J.F., *Reading the ubiquitin postal code.* Curr Opin Struct Biol, 2011. **21**(6): p. 792-801.
20. Li, W. and Y. Ye, *Polyubiquitin chains: functions, structures, and mechanisms.* Cell Mol Life Sci, 2008. **65**(15): p. 2397-406.

21. Deng, L., et al., *Activation of the IkappaB kinase complex by TRAF6 requires a dimeric ubiquitin-conjugating enzyme complex and a unique polyubiquitin chain.* Cell, 2000. **103**(2): p. 351-61.
22. Pinato, S., et al., *RNF168, a new RING finger, MIU-containing protein that modifies chromatin by ubiquitination of histones H2A and H2AX.* BMC Mol Biol, 2009. **10**: p. 55.
23. Pickart, C.M. and M.J. Eddins, *Ubiquitin: structures, functions, mechanisms.* Biochim Biophys Acta, 2004. **1695**(1-3): p. 55-72.
24. Stielow, B., et al., *SUMO-modified Sp3 represses transcription by provoking local heterochromatic gene silencing.* EMBO Rep, 2008. **9**(9): p. 899-906.
25. Gareau, J.R. and C.D. Lima, *The SUMO pathway: emerging mechanisms that shape specificity, conjugation and recognition.* Nat Rev Mol Cell Biol, 2010. **11**(12): p. 861-71.
26. Bogunovic, D., et al., *Mycobacterial disease and impaired IFN-gamma immunity in humans with inherited ISG15 deficiency.* Science, 2012. **337**(6102): p. 1684-8.
27. Komander, D., M.J. Clague, and S. Urbe, *Breaking the chains: structure and function of the deubiquitinases.* Nat Rev Mol Cell Biol, 2009. **10**(8): p. 550-63.
28. Nijman, S.M., et al., *A genomic and functional inventory of deubiquitinating enzymes.* Cell, 2005. **123**(5): p. 773-86.
29. Reyes-Turcu, F.E., K.H. Ventii, and K.D. Wilkinson, *Regulation and cellular roles of ubiquitin-specific deubiquitinating enzymes.* Annu Rev Biochem, 2009. **78**: p. 363-97.
30. Sowa, M.E., et al., *Defining the human deubiquitinating enzyme interaction landscape.* Cell, 2009. **138**(2): p. 389-403.
31. Cao, J. and Q. Yan, *Histone ubiquitination and deubiquitination in transcription, DNA damage response, and cancer.* Front Oncol, 2012. **2**: p. 26.
32. Wood, A., et al., *The Paf1 complex is essential for histone monoubiquitination by the Rad6-Bre1 complex, which signals for histone methylation by COMPASS and Dot1p.* J Biol Chem, 2003. **278**(37): p. 34739-42.
33. Wood, A., et al., *Bre1, an E3 ubiquitin ligase required for recruitment and substrate selection of Rad6 at a promoter.* Mol Cell, 2003. **11**(1): p. 267-74.
34. Dover, J., et al., *Methylation of histone H3 by COMPASS requires ubiquitination of histone H2B by Rad6.* J Biol Chem, 2002. **277**(32): p. 28368-71.
35. Ng, H.H., et al., *Ubiquitination of histone H2B by Rad6 is required for efficient Dot1-mediated methylation of histone H3 lysine 79.* J Biol Chem, 2002. **277**(38): p. 34655-7.
36. Shi, X., et al., *Cdc73p and Paf1p are found in a novel RNA polymerase II-containing complex distinct from the Srbp-containing holoenzyme.* Mol Cell Biol, 1997. **17**(3): p. 1160-9.

37. Mueller, C.L. and J.A. Jaehning, *Ctr9, Rtf1, and Leo1 are components of the Paf1/RNA polymerase II complex.* Mol Cell Biol, 2002. **22**(7): p. 1971-80.
38. Krogan, N.J., et al., *The Paf1 complex is required for histone H3 methylation by COMPASS and Dot1p: linking transcriptional elongation to histone methylation.* Mol Cell, 2003. **11**(3): p. 721-9.
39. Zhu, B., et al., *Monoubiquitination of human histone H2B: the factors involved and their roles in HOX gene regulation.* Mol Cell, 2005. **20**(4): p. 601-11.
40. Chandrasekharan, M.B., F. Huang, and Z.W. Sun, *Histone H2B ubiquitination and beyond: Regulation of nucleosome stability, chromatin dynamics and the trans-histone H3 methylation.* Epigenetics, 2010. **5**(6): p. 460-8.
41. Goldknopf, I.L., et al., *Isolation and characterization of protein A24, a "histone-like" non-histone chromosomal protein.* J Biol Chem, 1975. **250**(18): p. 7182-7.
42. Wang, H., et al., *Role of histone H2A ubiquitination in Polycomb silencing.* Nature, 2004. **431**(7010): p. 873-8.
43. Francis, N.J., et al., *Reconstitution of a functional core polycomb repressive complex.* Mol Cell, 2001. **8**(3): p. 545-56.
44. Jacobs, J.J. and M. van Lohuizen, *Cellular memory of transcriptional states by Polycomb-group proteins.* Semin Cell Dev Biol, 1999. **10**(2): p. 227-35.
45. Joazeiro, C.A. and A.M. Weissman, *RING finger proteins: mediators of ubiquitin ligase activity.* Cell, 2000. **102**(5): p. 549-52.
46. Morey, L. and K. Helin, *Polycomb group protein-mediated repression of transcription.* Trends Biochem Sci, 2010. **35**(6): p. 323-32.
47. Kuzmichev, A., et al., *Histone methyltransferase activity associated with a human multiprotein complex containing the Enhancer of Zeste protein.* Genes Dev, 2002. **16**(22): p. 2893-905.
48. Cao, R., et al., *Role of histone H3 lysine 27 methylation in Polycomb-group silencing.* Science, 2002. **298**(5595): p. 1039-43.
49. Breiling, A., et al., *General transcription factors bind promoters repressed by Polycomb group proteins.* Nature, 2001. **412**(6847): p. 651-5.
50. Simon, J., et al., *Elements of the Drosophila bithorax complex that mediate repression by Polycomb group products.* Dev Biol, 1993. **158**(1): p. 131-44.
51. Brown, J.L., et al., *The Drosophila pho-like gene encodes a YY1-related DNA binding protein that is redundant with pleiohomeotic in homeotic gene silencing.* Development, 2003. **130**(2): p. 285-94.
52. Brown, J.L., et al., *The Drosophila Polycomb group gene pleiohomeotic encodes a DNA binding protein with homology to the transcription factor YY1.* Mol Cell, 1998. **1**(7): p. 1057-64.
53. Wang, L., et al., *Hierarchical recruitment of polycomb group silencing complexes.* Mol Cell, 2004. **14**(5): p. 637-46.
54. Sing, A., et al., *A vertebrate Polycomb response element governs segmentation of the posterior hindbrain.* Cell, 2009. **138**(5): p. 885-97.

55. Guil, S., et al., *Intronic RNAs mediate EZH2 regulation of epigenetic targets.* Nat Struct Mol Biol, 2012. **19**(7): p. 664-70.
56. Rinn, J.L., et al., *Functional demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs.* Cell, 2007. **129**(7): p. 1311-23.
57. Peng, J.C., et al., *Jarid2/Jumonji coordinates control of PRC2 enzymatic activity and target gene occupancy in pluripotent cells.* Cell, 2009. **139**(7): p. 1290-302.
58. Ogawa, H., et al., *A complex with chromatin modifiers that occupies E2F- and Myc-responsive genes in G0 cells.* Science, 2002. **296**(5570): p. 1132-6.
59. Sanchez, C., et al., *Proteomics analysis of Ring1B/Rnf2 interactors identifies a novel complex with the Fbxl10/Jhdm1B histone demethylase and the Bcl6 interacting corepressor.* Mol Cell Proteomics, 2007. **6**(5): p. 820-34.
60. Zhou, W., et al., *Histone H2A monoubiquitination represses transcription by inhibiting RNA polymerase II transcriptional elongation.* Mol Cell, 2008. **29**(1): p. 69-80.
61. Zhu, Q., et al., *BRCA1 tumour suppression occurs via heterochromatin-mediated silencing.* Nature, 2011. **477**(7363): p. 179-84.
62. Zhou, W., X. Wang, and M.G. Rosenfeld, *Histone H2A ubiquitination in transcriptional regulation and DNA damage repair.* Int J Biochem Cell Biol, 2009. **41**(1): p. 12-5.
63. Shao, Z., et al., *Stabilization of chromatin structure by PRC1, a Polycomb complex.* Cell, 1999. **98**(1): p. 37-46.
64. Nakagawa, T., et al., *Deubiquitylation of histone H2A activates transcriptional initiation via trans-histone cross-talk with H3K4 di- and trimethylation.* Genes Dev, 2008. **22**(1): p. 37-49.
65. Stock, J.K., et al., *Ring1-mediated ubiquitination of H2A restrains poised RNA polymerase II at bivalent genes in mouse ES cells.* Nat Cell Biol, 2007. **9**(12): p. 1428-35.
66. Eskeland, R., et al., *Ring1B compacts chromatin structure and represses gene expression independent of histone ubiquitination.* Mol Cell, 2010. **38**(3): p. 452-64.
67. Endoh, M., et al., *Histone H2A Mono-Ubiquitination Is a Crucial Step to Mediate PRC1-Dependent Repression of Developmental Genes to Maintain ES Cell Identity.* PLoS Genet, 2012. **8**(7): p. e1002774.
68. Richly, H., et al., *Transcriptional activation of polycomb-repressed genes by ZRF1.* Nature, 2010. **468**(7327): p. 1124-8.
69. Joo, H.Y., et al., *Regulation of cell cycle progression and gene expression by H2A deubiquitination.* Nature, 2007. **449**(7165): p. 1068-72.
70. Scheuermann, J.C., et al., *Histone H2A deubiquitinase activity of the Polycomb repressive complex PR-DUB.* Nature, 2010. **465**(7295): p. 243-7.
71. Contreras, A., et al., *The dynamic mobility of histone H1 is regulated by cyclin/CDK phosphorylation.* Mol Cell Biol, 2003. **23**(23): p. 8626-36.

72. Zhu, P., et al., *A histone H2A deubiquitinase complex coordinating histone acetylation and H1 dissociation in transcriptional regulation.* Mol Cell, 2007. **27**(4): p. 609-21.
73. Vissers, J.H., et al., *The many faces of ubiquitinated histone H2A: insights from the DUBs.* Cell Div, 2008. **3**: p. 8.
74. Yu, H., et al., *The ubiquitin carboxyl hydrolase BAP1 forms a ternary complex with YY1 and HCF-1 and is a critical regulator of gene expression.* Mol Cell Biol, 2010. **30**(21): p. 5071-85.
75. Zhao, Y., et al., *A TFTC/STAGA module mediates histone H2A and H2B deubiquitination, coactivates nuclear receptors, and counteracts heterochromatin silencing.* Mol Cell, 2008. **29**(1): p. 92-101.
76. Reyes-Turcu, F.E. and K.D. Wilkinson, *Polyubiquitin binding and disassembly by deubiquitinating enzymes.* Chem Rev, 2009. **109**(4): p. 1495-508.
77. Gong, L., et al., *Identification of a novel isopeptidase with dual specificity for ubiquitin- and NEDD8-conjugated proteins.* J Biol Chem, 2000. **275**(19): p. 14212-6.
78. Ye, Y., et al., *Polyubiquitin binding and cross-reactivity in the USP domain deubiquitinase USP21.* EMBO Rep, 2011. **12**(4): p. 350-7.
79. Urbe, S., et al., *Systematic survey of deubiquitinase localization identifies USP21 as a regulator of centrosome- and microtubule-associated functions.* Mol Biol Cell, 2012. **23**(6): p. 1095-103.
80. Garcia-Santisteban, I., S. Banuelos, and J.A. Rodriguez, *A global survey of CRM1-dependent nuclear export sequences in the human deubiquitinase family.* Biochem J, 2012. **441**(1): p. 209-17.
81. Xu, G., et al., *Ubiquitin-specific peptidase 21 inhibits tumor necrosis factor alpha-induced nuclear factor kappaB activation via binding to and deubiquitinating receptor-interacting protein 1.* J Biol Chem, 2009 . **285**(2): p. 969-78.
82. Lang, G., et al., *The tightly controlled deubiquitination activity of the human SAGA complex differentially modifies distinct gene regulatory elements.* Mol Cell Biol, 2011. **31**(18): p. 3734-44.
83. Huynh, K.D., et al., *BCoR, a novel corepressor involved in BCL-6 repression.* Genes Dev, 2000. **14**(14): p. 1810-23.
84. Gearhart, M.D., et al., *Polycomb group and SCF ubiquitin ligases are found in a novel BCOR complex that is recruited to BCL6 targets.* Mol Cell Biol, 2006. **26**(18): p. 6880-9.
85. Dai, Y.S. and B.E. Markham, *p300 Functions as a coactivator of transcription factor GATA-4.* J Biol Chem, 2001. **276**(40): p. 37178-85.
86. Liu, L., et al., *Identification of a novel protein, PDIP38, that interacts with the p50 subunit of DNA polymerase delta and proliferating cell nuclear antigen.* J Biol Chem, 2003. **278**(12): p. 10041-7.
87. Chen, G., et al., *C-terminal amino acids 290-328 of LPTS/PinX1 confer telomerase inhibition.* Biochem Biophys Res Commun, 2010. **398**(4): p. 683-9.
88. Mao, L., et al., *HOXC9 links cell-cycle exit and neuronal differentiation and is a prognostic marker in neuroblastoma.* Cancer Res, 2011. **71**(12): p. 4314-24.

89. Park, J.H. and R.G. Roeder, *GAS41 is required for repression of the p53 tumor suppressor pathway during normal cellular proliferation.* Mol Cell Biol, 2006. **26**(11): p. 4006-16.
90. Abhiman, S., L.M. Iyer, and L. Aravind, *BEN: a novel domain in chromatin factors and DNA viral proteins.* Bioinformatics, 2008. **24**(4): p. 458-61.
91. Mitsui, K., et al., *The homeoprotein Nanog is required for maintenance of pluripotency in mouse epiblast and ES cells.* Cell, 2003. **113**(5): p. 631-42.
92. Niwa, H., J. Miyazaki, and A.G. Smith, *Quantitative expression of Oct-3/4 defines differentiation, dedifferentiation or self-renewal of ES cells.* Nat Genet, 2000. **24**(4): p. 372-6.
93. Ambrosetti, D.C., C. Basilico, and L. Dailey, *Synergistic activation of the fibroblast growth factor 4 enhancer by Sox2 and Oct-3 depends on protein-protein interactions facilitated by a specific spatial arrangement of factor binding sites.* Mol Cell Biol, 1997. **17**(11): p. 6321-9.
94. Boyer, L.A., et al., *Core transcriptional regulatory circuitry in human embryonic stem cells.* Cell, 2005. **122**(6): p. 947-56.
95. Chen, X., et al., *Integration of external signaling pathways with the core transcriptional network in embryonic stem cells.* Cell, 2008. **133**(6): p. 1106-17.
96. van den Berg, D.L., et al., *An Oct4-centered protein interaction network in embryonic stem cells.* Cell Stem Cell, 2010. **6**(4): p. 369-81.
97. Ventii, K.H. and K.D. Wilkinson, *Protein partners of deubiquitinating enzymes.* Biochem J, 2008. **414**(2): p. 161-75.
98. Steinberg, X.P., et al., *Human CCAAT/enhancer-binding protein beta interacts with chromatin remodeling complexes of the imitation switch subfamily.* Biochemistry, 2012. **51**(5): p. 952-62.
99. Tsai, Y.C., et al., *Functional proteomics establishes the interaction of SIRT7 with chromatin remodeling complexes and expands its role in regulation of RNA polymerase I transcription.* Mol Cell Proteomics, 2012. **11**(5): p. 60-76.
100. Barber, M.F., et al., *SIRT7 links H3K18 deacetylation to maintenance of oncogenic transformation.* Nature, 2012. **487**(7405): p. 114-8.
101. Rowbotham, S.P., et al., *Maintenance of silent chromatin through replication requires SWI/SNF-like chromatin remodeler SMARCAD1.* Mol Cell, 2011. **42**(3): p. 285-96.
102. Sathyan, K.M., et al., *A BEN-domain-containing protein associates with heterochromatin and represses transcription.* J Cell Sci, 2011. **124**(Pt 18): p. 3149-63.
103. Brill, L.M., et al., *Phosphoproteomic analysis of human embryonic stem cells.* Cell Stem Cell, 2009. **5**(2): p. 204-13.
104. Vertegaal, A.C., et al., *Distinct and overlapping sets of SUMO-1 and SUMO-2 target proteins revealed by quantitative proteomics.* Mol Cell Proteomics, 2006. **5**(12): p. 2298-310.
105. Zhang, X.Y., et al., *The putative cancer stem cell marker USP22 is a subunit of the human SAGA complex required for activated transcription and cell-cycle progression.* Mol Cell, 2008. **29**(1): p. 102-11.

106. Garcia-Dominguez, M. and J.C. Reyes, *SUMO association with repressor complexes, emerging routes for transcriptional control.* Biochim Biophys Acta, 2009. **1789**(6-8): p. 451-9.
107. Sapetschnig, A., et al., *Transcription factor Sp3 is silenced through SUMO modification by PIAS1.* EMBO J, 2002. **21**(19): p. 5206-15.
108. Yang, F., et al., *Sumoylation is important for stability, subcellular localization and transcriptional activity of Sall4, an essential stem cell transcription factor.* J Biol Chem, 2012.
109. Denuc, A. and G. Marfany, *SUMO and ubiquitin paths converge.* Biochem Soc Trans, 2010. **38**(Pt 1): p. 34-9.
110. Hu, M., et al., *Crystal structure of a UBP-family deubiquitinating enzyme in isolation and in complex with ubiquitin aldehyde.* Cell, 2002. **111**(7): p. 1041-54.
111. Cohn, M.A., et al., *UAF1 is a subunit of multiple deubiquitinating enzyme complexes.* J Biol Chem, 2009. **284**(8): p. 5343-51.
112. Cohn, M.A., et al., *A UAF1-containing multisubunit protein complex regulates the Fanconi anemia pathway.* Mol Cell, 2007. **28**(5): p. 786-97.
113. Zhang, X.Y., et al., *USP22, an hSAGA subunit and potential cancer stem cell marker, reverses the polycomb-catalyzed ubiquitylation of histone H2A.* Cell Cycle, 2008. **7**(11): p. 1522-4.
114. Krebs, A.R., et al., *SAGA and ATAC histone acetyl transferase complexes regulate distinct sets of genes and ATAC defines a class of p300-independent enhancers.* Mol Cell, 2011. **44**(3): p. 410-23.
115. Glinsky, G.V., O. Berezovska, and A.B. Glinskii, *Microarray analysis identifies a death-from-cancer signature predicting therapy failure in patients with multiple types of cancer.* J Clin Invest, 2005. **115**(6): p. 1503-21.
116. Margueron, R., et al., *Ezh1 and Ezh2 maintain repressive chromatin through different mechanisms.* Mol Cell, 2008. **32**(4): p. 503-18.
117. Yang, J., et al., *Genome-wide analysis reveals Sall4 to be a major regulator of pluripotency in murine-embryonic stem cells.* Proc Natl Acad Sci U S A, 2008. **105**(50): p. 19756-61.
118. Yang, J., et al., *A novel SALL4/OCT4 transcriptional feedback network for pluripotency of embryonic stem cells.* PLoS One, 2010. **5**(5): p. e10766.
119. Festuccia, N., et al., *Esrrb Is a Direct Nanog Target Gene that Can Substitute for Nanog Function in Pluripotent Cells.* Cell Stem Cell, 2012. **11**(4): p. 477-90.
120. Liu, Y., Z. Shao, and G.C. Yuan, *Prediction of Polycomb target genes in mouse embryonic stem cells.* Genomics, 2010. **96**(1): p. 17-26.
121. Lu, J., et al., *Stem cell factor SALL4 represses the transcriptions of PTEN and SALL1 through an epigenetic repressor complex.* PLoS One, 2009. **4**(5): p. e5577.
122. Kaji, K., et al., *The NuRD component Mbd3 is required for pluripotency of embryonic stem cells.* Nat Cell Biol, 2006. **8**(3): p. 285-92.
123. Reynolds, N., et al., *NuRD suppresses pluripotency gene expression to promote transcriptional heterogeneity and lineage commitment.* Cell Stem Cell, 2012. **10**(5): p. 583-94.

124. Nijnik, A., et al., *The critical role of histone H2A-deubiquitinase Mysm1 in hematopoiesis and lymphocyte differentiation.* Blood, 2011. **119**(6): p. 1370-9.
125. Jiang, X.X., et al., *Control of B cell development by the histone H2A deubiquitinase MYSM1.* Immunity, 2011. **35**(6): p. 883-96.
126. Wu, Y., et al., *SUMOylation represses Nanog expression via modulating transcription factors Oct4 and Sox2.* PLoS One, 2012. **7**(6): p. e39606.
127. Bloom, J. and M. Pagano, *Experimental tests to definitively determine ubiquitylation of a substrate.* Methods Enzymol, 2005. **399**: p. 249-66.

Abbreviations

mESC	mouse embryonic stem cells
LIF	leukemia inhibitory factor
RA	retinoic acid
iPS	induced pluripotent stem cells
PTM	posttranslational modification
ATP	adenosine triphosphate
ADP	adenosine diphosphate
AMP	adenosine monophosphate
HAT	histone acetyl-transferases
HDAC	histone deacetylase
HMTase	methyltransferases
PRMTs	protein arginine methyltransferases
MBP	methyl-binding protein
Ub	ubiquitin
Ubl	ubiquitin-like protein
UBD	ubiquitin-binding domain
UBA	ubiquitin-associated domain
UBL	ubiquitin-like domains
H2Aub	monoubiquitinated histone H2A
H2Bub	monoubiquitinated histone H2B
SUMO	Small Ubiquitin-like Modifier
DUB	deubiquitinating enzyme
UCH	ubiquitin C-terminal hydrolases
USP	ubiquitin-specific processing proteases
USP21	Ubiquitin Specific Peptidase 21
OTU	ovarian tumor proteases
JAMM	JAB1/MPN/Mov34 metalloenzymes
MJD	Machado-Joseph disease proteases
BEND3	BEN domain containing 3
RNAPII	RNA polymerase II
PcG	polycomb-group protein
PRC1	Polycomb repressive complex 1
PRC2	Polycomb repressive complex 2
RING1	ring finger protein 1
HOX genes	homeobox genes
Trx-G	trithorax-group proteins
EZH2	Enhancer of Zeste
PRE	Polycomb response element
ChIP	chromatin immunoprecipitation
ChIP seq	chromatin immunoprecipitation coupled with DNA sequencing

NES	nuclear export signal
NLS	nuclear localization signal
RIP1	receptor-interacting protein 1
OCT4	Octamer binding transcription factor 4
SOX2	SRY (sex determining region Y)-box 2
NANOG	Homeobox protein NANOG
SALL4	Sal-like protein 4
ESRRB	estrogen-related receptor beta
TCFCP2L1	transcription factor CP2-like 1