Functions of bromodomain-containing proteins and their roles in homeostasis and cancer

Takao Fujisawa[1] *and Panagis Filippakopoulos*[1,2,#]

[1]Ludwig Institute for Cancer Research, Old Road Campus Research Building, Roosevelt Drive, Oxford, OX3 7DQ, UK

[2]Structural Genomics Consortium, Old Road Campus Research Building, Roosevelt Drive, Oxford, OX3 7DQ, UK

[#] Correspondence: Panagis Filippakopoulos

Email: panagis.filippakopoulos@sgc.ox.ac.uk

Tel: + 44 1865 617759

Fax: + 44 1865 617575

Abstract

Bromodomains (BRDs) are evolutionary conserved protein–protein interaction modules that are found in a wide range of proteins with diverse catalytic and scaffolding functions, present in most tissues. BRDs selectively recognize and bind to acetylated lysine residues, particularly acetylated histones and thereby play important roles in regulating gene expression. BRD-containing proteins are frequently deregulated in cancer, they participate in gene fusions, generating diverse, frequently pro-oncogenic protein products and many mutations have been mapped to the BRDs themselves. . Importantly, BRDs can be targeted by small molecule inhibitors. This has stimulated many translational research projects seeking to attenuate the aberrant functions of BRD-containing proteins in disease.

Introduction

Lysine acetylation (Kac) has emerged over the past decade as an important and widespread regulatory post translational modification (PTM) of proteins[1, 2], and it has been shown to be

involved in regulating transcription, metabolism and cell signalling[3, 4]. Kac results in neutralization of the positive charge on lysine residues, affecting protein–protein, protein–DNA interactions as well as protein stability and sub-cellular localization[3, 5]. Acetylation of histones in chromatin is associated with the establishment of euchromatin leading to transcriptional activation and the expression of genes, including tumour suppressor genes. In line with this,deregulation of histone acetylation patterns has been shown to drive aberrant expression of survival and growth-promoting genes resulting in proliferation and tumorigenesis[6, 7].

Acetyl groups are added to histone lysines by histone acetyltransferases (HATs) and are removed by histone deacetylases (HDACs), two enzyme classes that have also been shown to act on non-histone proteins[3]. Importantly, HDACs have been successfully targeted in the clinic[8], with 4 currently FDA-approved drugs in the market, highlighting the role of deregulated acetylation in disease.

Recognition of Kac is principally mediated by bromodomains (BRDs), evolutionary conserved protein interactions domains (also called "modules") first identified in the *Drosophila melanogaster* "brahma" gene[9]; however recently it was discovered that YEATS domains can also bind to histone Kac-motifs[10], in addition to lysines modified by crotonylation (Kcr, another acylation-based PTM)[11, 12], thus increasing the diversity of PTM readout.

The human proteome encodes 61 BRD interaction modules, found in 42 diverse proteins. These proteins primarily recognize acetylated histones and engage in regulating gene expression through a wide range of activities. First of all, they can act as scaffolds, allowing the assembly of larger protein complexes. They can also function as transcription factors and transcriptional coregulators. Finally, BRD-containing proteins can serve various catalytic functions, acting as methyltransferases, ATP-dependent chromatin remodelling complexes, HATs and helicases, and thereby engaging in various chromatin modifications (Supplementary information S1 (**Table)**). Importantly, BRD proteins are expressed in various tissues, and as indicated by recent advances in proteomics[13, 14] they show broad and variable expression profiles, suggesting that BRDs may have context-dependent significance (**Figure 1a**).

BRD-containing proteins have broad documented roles in cell homeostasis and have been implicated in human diseases, including neurological and inflammatory diseases as well as cancer (recently reviewed elsewhere[15-18]). Here we will first discuss the structural involvement of the BRD module in the recognition of acetylated lysines, and how it occurs in the context of multi-domain proteins. We will then discuss the roles of BRD-containing proteins in gene regulation and their involvement in normal cell physiology. Finally, we will discuss how

mutations, misexpression and gene fusions deregulate BRD-containing proteins, implicating these proteins in cancer development and progression. We would also like to suggest excellent previous reviews on BRD inhibitor development, particularly targeting the bromo and extra terminal (BET) sub class of BRDs[17, 19-21], which very well highlight that studying BRDs and BRD-containing proteins can have a large impact on human health, by leading to clinical applications.

Overview of BRD-containing proteins

BRD modules are often found within multi-domain proteins, where they are linked via flexible sequences to diverse interaction or catalytic domains[22]. This architectural arrangement allows for conformational flexibility which in turn facilitates interactions with diverse sequence motifs. Multi-domain arrangements combine BRDs with various protein–protein interaction domains, including methyl-reader domains such as PHD (plant homeodomain zinc finger domain) fingers, proline-tryptophan-tryptophan-proline domains (PWWP), bromo- adjacent homology domains (BAH), as well as KIX (kinase-inducible domain interacting domain) and SAND domains (named after Sp100, AIRE-1, NucP41/75, DEAF-1). In addition, BRD-containing proteins can harbour various catalytic domains, such as HAT and multiple BRDs (**Figure 1b**). These combinations suggest a complex recognition process involving the combinatorial readout of different protein-interaction motifs as well as post translational modifications beyond Kac, for example in the same or adjacent nucleosomes, or present in other proteins associating with chromatin. The diverse range of annotated modular domains, beyond BRDs themselves, found in BRD-containing proteins is listed in Supplementary information S1 (**Table)**.

The complexity introduced by such multivalent interactions makes interpretation of the histone code challenging. Recent advances in peptide array technologies have facilitated systematic studies combining multiple PTMs in short peptides immobilized on solid support, helping understand how combinatorial readout can be achieved by reader modules using recombinant proteins[22] or cellular extracts[23]. However, such techniques can only be reproducibly applied to relatively short peptide sequences (about 20-30 residues long), presenting a need for alternative technologies to capture the various Kac-containing epitopes dispersed within target proteins.

BRD module architecture

The BRD interaction module has a distinctive architecture, structurally characterized more than a decade ago[24-26], comprising four helices (αZ, αA, αB and αC), linked by two divergent loop regions (ZA- and BC-loops) (**Figure 2**). A number of conserved residues are found dispersed within the module, including a PxY motif at the end of the ZA-loop region and a tyrosine in the AB-loop which forms a salt bridge to a residue on helix B (typically an aspartate), stabilizing the fold. A conserved asparagine residue at the beginning of the BC-loop facilitates docking of Kac, while the charged surface surrounding the Kac-binding pocket presents a large charged interface which initiates interactions with the backbone of the acetylated peptide (**Figure 2a, c and d**). Some BRDs contain a tyrosine (for example the ones present in SP100, SP110 and SP140) or threonine (found in BRWD3) instead of the conserved asparagine(**Figure 2b**); however no evidence exists to date linking those unusual BRDs to Kac binding and the specific motifs that they may be able to recognize remain uncharacterized. Together, the structural topology of the BRD module presents a small hydrophobic pocket within the four helices, flanked by the ZA- and BC-loops where the neutralized side-chain of Kac can be accommodated, while binding to the entire acetylated peptide is mainly driven by the charged surface surrounding the BRD Kac-binding site.

Based on their structural topology and similarity, human BRD modules were classified into eight structural families[22] (Supplementary Information S2 (Figure)). Intriguingly, the distribution of charge on the surface of BRD modules varies dramatically, with some exhibiting highly positively charged surfaces, suggesting that acetylated histone peptides, which are typically rich in positively charged lysine and arginine residues, may not be the target of these modules. Given the wide distribution of acetylation[2] as well as the widespread expression of bromodomain containing proteins (Supplementary information S1 (**Table)**) it is tempting to speculate that several acetylated protein targets of BRDs exist, beyond histone proteins; however, despite the few but notable established examples[27-32], little structural and functional evidence exists to date, linking most BRDs to acetylated non-histone proteins. Interactions involving acetylated non-histone proteins are highlighted in **Box 1**.

Recognition of acetylated lysine

The hydrophobic nature of the cavity formed by the four helices accommodates the neutralized Kac side chain found in histone peptide sequences[24-26]. Several crystal structures of BRDs in complex with histone peptides (for a comprehensive list of structurally characterized BRD–peptide complexes see Supplementary information S3 (Table).) show a conserved binding mode,

whereby histone peptides insert one acetylated lysine within the Kac-binding cavity in the BRD, initiating contacts with water molecules within this cavity as well as the conserved asparagine (**Figure 3a**). Depending on the domains adjacent to the BRD module, the orientation of the bound peptide varies dramatically. For example, proteins harbouring a single BRD typically bind histone peptides so that the N-terminus of the peptide resides on the back of the pocket, and the peptide inserts between the ZA- and BC-loops and aligns on top of the Kac-binding cavity with an exit vector over the ZA-loop. Notably, the presence of other modular domains can largely change this arrangement. For example, the chromatin regulator TRIM24, a member of the Tripartite-motif (TRIM) proteins, (also known as Transcription intermediary factor 1-alpha, TIF1α) engages the N-terminal histone H3 tail via PHD-finger initiated recognition of lysine 4 (H3K4), while an adjacent BRD module linked via a flexible loop binds to acetylated lysine 23 (H3K23ac) on the same histone[33]. A similar PHD–BRD domain arrangement on BPTF (Nucleosome-remodelling factor subunit) facilitates binding to H3K4me$_2$ and H3K4me$_3$ via the PHD domain[34], thus allowing the BRD to gain specificity for H4K16ac found in *trans* within a single nucleosome, over other H4 acetylations[35]. Such multivalent interactions dictate the peptide orientation, aligning for example the C-terminus of the peptide between the ZA- and BC-loops on the BRD of TRIM33 in a complex of the BRD/PHD cassette with H3K9me$_3$ and H3K18ac[36] (**Figure 3b**). The presence of tandem BRDs linked by a short rigid linker found, for example on TAF1 (Transcription initiation factor TFIID subunit 1) facilitates binding to histone H4 acetylated at lysines 5 and 12 (H4K5ac and H4K12ac), allowing for higher specificity towards multiply acetylated histone H4 tails[25]. Linking of two BRD modules via long flexible linkers, as in the case of the BET subclass of bromodomains, offers conformational plasticity which in turn allows for simultaneous recognition of two different acetylated lysines within the same or even present in different proteins[30, 31]. Nevertheless, in the absence of multi-domain arrangements, single acetyl-lysine binding within histone tails seems to be relatively conserved between different BRD modules, with only rather small variations of the peptide topology spanning the groove between the BC- and ZA-loop regions (**Figure 3c**).

Some BRDs can also recognize double acetyl marks and the first described example of such a protein was mouse BRDT, which engages histone H4 peptides bearing K5ac and K8ac[37]. In this example, while K5ac directly engages the conserved asparagine, K8ac is also inserted within the Kac binding cavity and initiates interactions with K5ac as well as a number of water-assisted interactions with the protein (**Figure 3d**). Recognition of double acetyl histone marks was further suggested to be a common feature shared by members of the BET class of

bromodomains, as high resolution crystal structures of the N-terminal BRD of BRD4 with different histone H4 peptides revealed the same conserved mode of interaction involving two acetylated H4 lysines[22].

BRDs and other histone acylations

Recently it has been shown that histones can be modified by non-acetyl acylations, including propionylation, butyrylation and crotonolytation [38, 39]. Similarly to Kac, HATs were demonstrated to catalyse the deposition of these PTMs [40, 41], and it has been revealed that histone acylations have functional importance[11, 41]. This has stimulated research seeking to identify if and how these PTMs can be recognized by BRDs (recently reviewed elsewhere[3, 42]). Despite their structural conservation, it appears that several BRD modules, including ones found in BRD4, BRDT, BRD9 and TAF1, apart from binding Kac, can also accommodate the bulkier acyl-groups in the context of histone peptides (**Figure 3e; refer also to Supplementary information S3 (Table)**)[41, 43]. Notably however, at least in the case of butyrylation, it appears that distribution of this mark competes with acetylation on H4K5 and it prevents BRDT binding, thus affecting stage-specific expression programs[41]. It remains to be seen what the precise contributions of this and other acylations signals to transcriptional programmes are and how the BRD modules contribute to the relay of these signals. As other acylation PTMs are much less studied, we will focus here on the recognition of Kac by BRD modules and discuss how it contributes to cell homeostasis, highlighting the role of BRD-containing proteins in gene regulation through their interactions with chromatin.

Roles in gene regulation

BRD-containing proteins serve diverse physiological functions, alone or as part of larger protein complexes, and are most notably involved in gene regulation through modulation of transcription. First of all, these proteins are known to be involved in regulatory chromatin modifications, engaging in chromatin remodelling, and in the introduction of further histone modifications (acetylation and methylation). BRD-containing proteins can also regulate transcription through specific recognition of histones and serving as scaffolds, which control the recruitment of other transcriptional regulators to chromatin. Finally, they can modulate the transcriptional machinery itself. Here we will summarize these various activities of BRD-

containing proteins in regulating gene expression and we will discuss their functional roles and implications in normal tissue.

BRD-containing proteins as components of chromatin remodelling complexes.

BRD-containing proteins are frequently found as components of large complexes that target chromatin and control its compaction and decompaction via diverse activities, thereby leading to chromatin remodelling (**Figure 4a**). An example of such multi-subunit assemblies are SWItch/Sucrose Non-Fermentable (SWI/SNF) complexes, first identified in yeast. Mammalian SWI/SNF complexes are components of the embryonic stem cell (ESC) core pluripotency transcriptional network and play key roles controlling cell differentiation and proliferation[44]. A core subunit found in SWI/SNF complexes is responsible for ATP-dependent alteration of chromatin structure, resulting in open chromatin which in turn can be recognized by transcription factors, resulting in transcriptional stimulation. In human this core subunit consists of either SMARCA4 (named after the *Drosophila* homologue Brama Brg1 gene) or the related SMARCA2 (also known as BRM), both of which contain a C-terminal BRD module capable of recognizing acetylated histone H3 and H4[22, 45]. Despite the high degree of sequence homology, SMARCA2 and SMARCA4 have functionally distinct biological roles. For instance, homozygous deletion of SMARCA4 in mice shows embryonic lethality, while SMARCA2 null mutant mice are viable[46, 47].

In human the PBAF SWI/SNF complex also contains polybromo protein PB1 (or PBRM), a multidomain protein with six BRD modules, two bromo-adjacent homology (BAH) domains and a homeobox DNA binding domain. Mice lacking PB1 exhibit defects in heart development due to impaired epithelial-to-mesenchymal-transition (EMT) and arrested maturation of epicardium as a result of the down-regulation of several growth factors, including fibroblast, transforming and vascular endothelial growth factors (FGF, TGF, and VEGF, respectively)[48]. PB1 is also involved in mediating cellular senescence and genomic stability by regulating p53 transcriptional activity to promote the expression of a subset of genes required for replicative and oncogenic stress senescence induction[49]. Recent data also suggests that PB1 is required for the cohesion of the centromeres in mouse and human cells, thereby also promoting genome stability[50]. Other BRD-containing units of the PBAF complex include a transcriptional regulator BRD7, which was found to play a role in regulating expression of key signalling components such as Nodal, ADAMTS1, BMI-1, CRABP1, as well as thyroid releasing hormone[51]. Interestingly, BRD7 was also shown to compete with p110 (catalytic subunit of phosphatidylinositol 3-kinase (PI3K)) for

binding to the p85α (regulatory PI3K subunit), thus regulating the PI3K signalling pathway[52].Additionally, BRD7 was found in obese and diabetic mouse models to play a role in glucose homeostasis in the liver, reducing blood glucose levels[53]. The related BRD9 was also found to be part of SWI/SNF complexes in mammals, although its precise biological function remains elusive[54].

Bromodomains are also found in imitation-SWI (ISWI) chromatin remodelling complexes. ISWI complexes regulate heterochromatin replication and transcription via chromatin remodelling, and assemble into seven or more different complexes containing a central core ATPase comprising two SNF2 like mammalian homologues of yeast ISWI, SNF2L and SNF2H. The Cat eye syndrome critical region protein 2 (CECR2) is a BRD-containing proteins that is a part of the mammalianSNF2L, acting as a chromatin remodelling factor[55]. It is expressed predominantly in the nervous system where it plays a critical role in neurulation[56].

The NUcleosome Remodelling Factor (NURF) ISWI complex additionally includes Bromodomain PHD finger Transcription Factor, BPTF. BPTF was shown to be highly expressed in patients with Alzheimer's disease (in this context it was referred to as “Fetal Alz-50 reactive Clone 1” (FAC1)), as well as in fetal brain in patients with neurodegenerative diseases (referred to as Fetal Alzheimer Antigen, FALZ)[57]. Owing to the presence of a PHD-BRD tandem module , BPTF recognizes both acetylated and methylated histone tails , and notably it is able to drive multivalent engagement of single nucleosomes[35].

BRDs are also found in the bromodomain adjacent zinc finger (BAZ) sub-family of proteins encoded by four related genes in human (*BAZ1A, BAZ1B, BAZ2A* and B*AZ2B*). BAZ proteins share common domain architecture with a characteristic PHD/BRD cassette, offering the possibility of multivalent engagement of methylated and acetylated lysine sequences[58]. BAZ1A forms a chromatin remodelling complex with SNF2H, which aids DNA replication through highly condensed regions of chromatin[59]. The related protein BAZ1B (also known as Williams-Beuren syndrome transcription factor) is a subunit of the WICH complex (WSTF-ISWI ATP-dependent chromatin-remodelling complex) that was shown to have tyrosine kinase activity (despite absence of any known kinase domain), and to mediate H2AX phosphorylation, thereby playing an important role in DNA damage response[60]. BAZ2A, also known as TTF-1-interacting protein 5 (TIP5), is a key subunit of the nucleolar remodelling complex (NoRC), playing a role in non-coding RNA-dependent silencing[61]. The biological function of BAZ2B[62] remains to date elusive.

Finally, the WD-repeat protein BRWD1 (also known as WDR9), which is a member of a small family of tandem-BRD-containing proteins, has been annotated as a transcriptional regulator interacting with the SMARCA4 component of the SWI/SNF complex[63], with recent data in mice demonstrating its role in spermatogenesis and oogenesis[64, 65]. Another protein of this family, BRWD3, is involved in the regulation of the JAK-STAT signalling pathway in *Drosophila*; however its precise function in human is not yet established[66].

Participation of BRD-containing proteins in histone modifications.

Catalytic activity is found in several BRD-containing proteins, including PCAF, CREBBP, EP300, GCN5L2, BRD8, MLL, ASH1L and the BRPF family. These proteins affect chromatin structure by directly depositing PTMs (acetylation and methylation) on histones. They may be a part of larger complexes and have been implicated in various physiological processes (**Figure 4b**). For instance, HAT proteins CREBBP and EP300, which share high degree of sequence similarity and domain organization, by acting as transcriptional co-activators are involved in controlling genomic stability, development, energy homeostasis, memory formation, neuronal plasticity and cell growth (recently reviewed elsewhere[67]). Both proteins contain in addition to a catalytic HAT domain, BRD, PHD and KIX protein interaction domains. Despite the architectural similarities, the diverse phenotypes observed in knockout mice suggest that each protein has distinct biological functions. For example, homozygous deletion of EP300 results in embryonic lethality, with defects in heart development, neurulation and cell proliferation[68]. Homozygous knockout of CREBBP in mice is also lethal (*in utero*) exhibiting retardation of development and delays in both primitive and definitive haematopoiesis, with signs of defective blood vessel formation in the central nervous system[69]. Conditional *EP300* or *CREBBP* knockout in mice exhibit a degree of redundant perturbation of co-activation capacity in T cells and macrophages, and each gene was found to have unique involvement in thymocyte development[70].

The HATs PCAF and GCN5L2 also contain BRD modules, and similarly to CREBBP and EP300 play roles as transcriptional co-activators[71]. Both proteins are incorporated into the SAGA and ATAC complexes (broadly conserved large co-activator complexes that regulate the transcription of many inducible and developmentally regulated genes) in human. Despite the fact that both these complexes exhibit HAT activity, they act in a mutually exclusive manner, suggesting that each has distinct biological functions[72]. Indeed, deletion of either gene in mice has very different consequences: *PCAF* knockout affects chromatin remodelling necessary for

memory formation and response to stress[73], whereas loss of *GCN5L2* is lethal in mice, and characterized by severely retarded growth and failure to form dorsal mesoderm lineages and the failure of anterior neural tube closure[74, 75].

Some BRD-containing proteins lacking HAT activity have also been described as part of HAT complexes. BRD8, for instance, has been found to be a component of the conserved NuA4 complex, a HAT multisubunit complex responsible for acetylation of histone H4 and H2A N-terminal tails in yeast[76]. Furthermore, BRD8 has been identified as a member of the H2A.Z pre-deposition complex which stimulates H2A.Z exchange in nucleosomes[77, 78], however the precise biological function and activity of BRD8 remains unassigned. The Bromodomain and PHD finger-containing proteins:BRPF1, BRPF2 (also known as BRD1) and BRPF3 all contain BRDs as well as PWWP and assembly of Zinc-fingers (PZP) domains. BRPF1 has been characterized as a non-catalytic sub-unit of the MOZ/MORF HAT complex targeting histone H3[79]; it was recently discovered that the PZP domain of BRPF1 recruits the MOZ complex to chromatin, thus directing the HAT activity towards histones[80]. In addition, deletion of the *BRPF1* gene in mice was shown to cause embryonic lethality, indicating an essential role of this protein in embryogenesis[81]. The related protein BRPF2, associates with the HBO1 HAT complex, which regulates erythropoiesis by increasing levels of H3K14ac[82]. BRPF3 also associates with HBO1 but unlike BRPF2, BRPF3 is required for DNA replication origin activation by being involved in the deposition of H3K14ac around transcriptional start sites[83].

Apart from HAT activity, proteins that engage in protein methylation have been found to contain BRD modules. Examples of these proteins include the Absent small and homeotic disks protein 1 homolog (ASH1L) and the Myeloid/lymphoid or mixed-lineage leukemia (MLL). ASH1L is the mammalian homolog of the *Drosophila melanogaster* Trithorax group (TrxG) protein Ash1 and contains a BRD, PHD and BAH domains in addition to the catalytic SET domain, responsible for the methyltransferase activity. ASH1L occupies regulatory regions of most active genes and methylates histone H3 at a subset of genes, including *HOX* genes[84], which encode transcription factors that control angiogenesis, motility, apoptosis, differentiation and receptor signalling during haematopoiesis as well as other developmental processes. Intriguingly, ASH1L is required for myelomonocytic differentiation of haematopoietic stem cells in mice and was found to regulate both positively and negatively *HOX* gene expression[85]. MLL is also a multi-domain methyltransferase containing a BRD module in addition to a SET domain, and it plays an essential role in regulating the expression of genes implicated in self-renewal of hematopoietic stem cells[86]. MLL targets Lys4 in histone H3 and this methyltransferase activity is stimulated by

H3 acetylation, which has been associated with *HOX* gene activation and H3K4 methylation *in vivo*[87].

BRD-containing proteins as histone recognizing scaffolds.

Recognition of Kac is a shared feature of BRD-containing proteins but as discussed above the specificity and the mode of interaction of these proteins with acetylated peptides can largely vary, which makes them perfectly tailored for the specific recognition of acetylated histones. A few sub-classes of BRD-containing proteins seem to play important roles by serving as scaffolds that regulate the recruitment and segregation of transcriptional machinery components to particular loci (**Figure 4c**).

In particular the sub-family of BET proteins has been extensively studied, particularly following the recent availability of selective inhibitors targeting their dual BRD modules[88, 89] (see also **Box 2**); their roles have been recently reviewed elsewhere[16, 18]. The four members of this family (BRD2, BRD3, BRD4 and the testis specific BRDT) share a common domain architecture comprising two tandem N-terminal BRDs, and an extra-terminal (ET) protein–protein interaction domain; they play key critical roles in cell cycle progression by regulating growth promoting genes. BET proteins dock onto acetylated chromatin via their BRD modules and recruit components of the transcriptional machinery via their ET domain. Reduced levels of BRD2 resulting from a mutation at the *BRD2* promoter give rise to extreme obesity in mice, occurring without glucose intolerance[90]. However, complete knockout of *BRD2* is lethal in mice[91], as is knockdown of *BRD4*[92]. BRD3 plays a role in erythropoiesis by regulating erythroid target genes. This occurs through additional BRD-initiated interactions with the acetylated transcription factor GATA1[93] (see also Box 1). Recent data suggests that BRD2 and BRD4 may also have non-redundant functions in erythroid-gene expression[94]. BRD4 and BRDT have been shown to be key factors in transcriptional elongation, by recruiting the positive transcription elongation factor b (P-TEFb, the complex of CDK9 and cyclinT) to transcriptional start sites, where the paused polymerase is reactivated following phosphorylation by CDK9[95-97]. BRDT has also been shown to have a key role in spermatogenesis, involving the N-terminal BRD module[98]. The possibility to directly target BRDT bromodomains by drugs, established this protein as a potential target for pharmacological male contraception[99].

TAF1 is the largest component of the core sub-unit of the TFIID transcription factor and contains two tandem BRD modules that recognize and bind to multiple Kac residues on the histone H4 tail[25]. TAF1 modulates the rate of transcriptional initiation by binding to the core promoter

sequence and recruiting other transcriptional regulators, thereby regulating various biological processes such as cell cycle, apoptosis and DNA damage response[100, 101]. Interestingly, TAF1 was also found to directly associate with BRD4 exerting synergistic control over gene expression, in a mechanism that has yet to be determined[102]. The testis specific homologue of TAF1, TAF1L, and may act as a functional substitute for TAF1 during male meiosis[103].

Triple-reader assemblies involving a PHD–BRD–PWWP cassette followed by a MYND domain are found in the Zinc-finger and MYND binding domain (ZMYND) proteins. ZMYND8 was recently identified as a DNA damage response factor which recruits the Nucleosome Remodelling and Deacetylase (NuRD) complex to DNA damage sites resulting in transcriptional repression, promoting repair by homologous recombination[104]; recent data also suggests that ZMYND8 recruits the lysine demethylase KDM5C, resulting in a "molecular break" that negatively regulates enhancer activity[105]. ZMYND11 has a similar domain architecture to ZMYND8, lacking however the conserved asparagine necessary for Kac interactions within the BRD module, which is turn replaced by a tyrosine residue. ZMYND11 was shown to specifically bind to H3.3K36me3 sites, regulating transcriptional elongation (acting as a co-repressor) and pre-mRNA splicing[106, 107]

ATPase family AAA domain-containing protein 2 (ATAD2) has been previously implicated in transcriptional co-regulation of genes regulated by oestrogen receptor α (ERα)[108] and MYC[109]. It has then been shown that ATAD2 serves as a scaffold recruiting E2F transcription factors to chromatin[110], which further recruit chromatin remodelling complexes and histone modifying activities, thereby promoting gene expression. The related protein ATAD2B, is transiently expressed in developing neurons in mice, however its precise biological function remains elusive[111]. Additionally, apart from BRD ATAD2 also contains an AAA domain (ATPase Associated with diverse cellular Activities), which are broadly implicated in ATP-dependent remodelling of macromolecules. In line with this, ATAD2 has been shown to be an important modulator of chromatin dynamics in ESCs[112]. This example highlights that even a single BRD-containing protein can regulate gene expression in many ways, which are likely context-dependent.

BRD-containing proteins as modifiers of transcriptional machinery.

Various BRD-containing proteins engage in regulating gene expression by modifying transcriptional machinery, including transcription factors, the RNA polymerase as well as other chromatin modifiers (**Figure 4d**). ZMYND8 can act as a scaffold as discussed above, but it was

also shown to act as a transcriptional coactivator of ERα target genes[113], providing further evidence for the context-specificity of BRD-containing protein functions. Also some members of the TRIM sub-family act as modulators of transcriptional machinery. The TRIM sub-family of BRDs (TRIM24, TRIM28, TRIM33 and TRIM66 inhuman), contain a BRD and PHD domains together with a RING domain and one or two B-boxes (zinc binding motifs) as well as an associated coiled-coil region. TRIM24, for example, associates with nucleosomes using its BRD and PHD modules, and is involved in ligand-dependent activation of the androgen and retinoic acid receptors, thereby participating in the regulation of their target genes[33]. TRIM28 (also known as KRAB-associated protein 1 or KAP-1) plays a role as a transcriptional repressor modulating chromatin structure via association with the hetero-chromatin-associated factors HP1 (HP1α, HP1β, HP1γ), promoting silencing of euchromatic genes[114]. TRIM28 contains a PHD/BRD cassette which was shown to confer E3 SUMO ligase activity necessary for gene silencing via recruitment of the NuRD complex[115, 116]. More recently TRIM28 has also been implicated in the repair of DNA double-strand breaks in heterochromatin[117]. TRIM33 has ubiquitin ligase activity that regulates the TGFβ signalling pathway through negative modulation of the activity of transcription factor SMAD4[118]. It is recruited to acetylated histone H3 but only in the absence of R2 and K4 methylation, suggesting a tight regulation of TRIM33–DNA interactions[119]. TRIM33 has a role in development as demonstrated in knockout mice which die during early somitogenesis[120]. Finally the related TRIM66 is mainly expressed in testis where it associates with heterochromatin-associated factors (HPs) and functions as a transcriptional silencer[121].

The nuclear body SP family of bromodomains (including SP100, SP110 and SP140) contain a SAND (Sp100, AIRE-1, NucP41/75, DEAF-1) DNA binding domain followed by a tandem PHD-BRD cassette and have been shown to regulate diverse cellular functions, including genome stability, DNA repair, transcription and chromatin organization by acting as hormone receptor transcriptional coactivators (for example for ETS1 and nuclear receptors)[122-124].

Similarly, the tandem BRD module containing protein, PH-interacting protein (PHIP), is thought to act as a transcriptional coregulator and it has been demonstrated to promote pancreatic β-cell growth and survival by regulating gene expression [125]; how exactly this protein exerts its coregulatory functions has not been determined.

BRD-containing proteins in cancer

BRD-containing proteins have been found deregulated in cancer and their aberrant expression has been shown to stimulate but also to suppress malignant phenotypes, some of which were recently reviewed elsewhere[15] (the reported roles of BRD-containing proteins in cancer are summarized in Supplementary information S4 (Table)). These dual roles of certain BRD-containing proteins (such as TRIM24) in tumour promotion as well as suppression suggest context dependent functions, making their study intriguing, albeit difficult. An additional layer of complication is introduced when many of these proteins are found participating in large complexes as well as when they engage in the formation of fusion products that result in aberrant function as detailed below. Here we will discuss the most recent advances in our knowledge of how the deregulation of BRD-containing proteins can result in cancer. We will also outline the known tumour suppressive properties of some of these proteins.

Mutations of BRD-containing proteins in cancer.

Large scale studies have contributed to identifying mutations of various proteins (including BRD-containing proteins) in cancer, although the precise mechanisms linking the observed phenotypes with protein dysfunction resulting from mutations are not completely understood most of the time. For instance, a large scale study across 21 different paediatric cancer subtypes identified recurrent mutations in 21 out of 42 BRD-containing proteins[126] (mutations of BRD-proteins registered in the COSMIC database[127] are listed in Supplementary information S5 (Table)). Similar medium-to-large scale studies have identified mutation in BRD-containing proteins in hepatocellular carcinoma (HCC)[128], squamous cell lung cancers[129], lung adenocarcinomas[130-132], small[133] and non-small cell lung cancers[134, 135], adenoid cystic carcinoma[136], spinal ependymomas[137], T-cell lymphoma[138], thymic epithelial tumours[139], cervical carcinomas[140], endometrial carcinomas[141], follicular lymphomas[142], oesophageal squamous cell carcinomas[143-145], colorectal (CRC) and gastric cancers[146]. However, no systematic studies focusing on the mutation load found on the actual BRD modules themselves exist and as such the contribution of Kac readout to the phenotypes observed remains elusive. Nonetheless, the recent advances in structural characterization of BRD modules now allow mapping mutation load in order to determine their functional outcome. In Supplementary information S6 (Figure)**,** mutations annotated in The Cancer Genome Atlas (TCGA) and found on BRD modules indicating potential loss of function were annotated on the structure of each sub-family of human BRDs. Given that the surface charge surrounding the Kac binding cavity seems to be driving binding affinity[22], we would predict that most identified mutations would alter that affinity,

negatively impacting on binding of BRD-containing proteins to their targets, but this remains to be determined for each protein affected.

Misexpression of BRD-containing proteins in cancer.

Aberrant expression of some BRD-containing proteins, for example resulting from the alterations of their gene copy numbers (for the information on copy number alterations in BRD-containing proteins registered in the COSMIC database[127] see Supplementary information S5 (Table)), has been shown to drive various malignant phenotypes. BRD9, for example, has been recently associated with HCC[128], CRC and melanoma[54], however the exact involvement and mechanism by which BRD9 contributes to disease progression remain elusive. The TRIM sub-family member TRIM24 was also recently shown to promote tumour growth and to enhance resistance to chemotherapy via stimulating PI3K-Akt signalling, and its protein levels were negatively associated with the prognosis for patients with newly diagnosed glioblastomas[147]. Similarly, TRIM66 over-expression correlates with poor prognosis in osteosarcoma patients and knockdown of *TRIM66* in osteosarcoma cells suppressed tumorigenicity in mice. This suggests that TRIM66 may act as an oncoprotein, possibly by suppressing apoptosis while promoting TGFβ[148]. Finally ZMYND8 was recently shown to promote tumour angiogenesis in zebrafish prostate cancer xenografts, as well as in prostate cancer samples from patients[149].

Oncogenic fusions of BRD-containing proteins.

Genes of several BRD-containing proteins have been found to generate fusions, resulting from chromosomal translocation between genes encoding BRD-containing proteins and other genomic regions (**Figure 5a and b,** Supplementary information S4 (Table)). Such fusion events can lead to the inactivation of the original protein(s) and/or can confer completely new functions on the fused protein product. Notably, in many cases these fusion proteins have been associated with aggressive phenotypes. For instance, the BET family members *BRD3* and *BRD4* have been found fused to the *NUT* gene, a nuclear protein expressed normally in testis. These fusions are the result of gene translocations and have been detected in midline organs, driving the so called NUT midline carcinomas[150, 151]. The BRD4-NUT product has been shown to recruit a number of factors that stimulate local chromatin hyper-acetylation, which promotes further recruitment of BRD4, thereby leading to transcriptional activation and increased expression of pro-survival genes[152] (**Figure 5c**). Interestingly, thieno-diazepine compounds can selectively target the BRDs

on the BRD4-NUT fusion product, resulting in therapeutically relevant phenotypes in mice xenograft models[88].

BETs are not the only BRD-containing proteins identified in chromosomal translocations. *CREBBP* has been found in several translocations, and was recently shown to be fused to the *SLX4* gene in lung cancers. In this case, the first two exons of *CREBBP* were fused to the 3′ portion of *SLX4* however the significance of the resulting abnormal protein is yet to be determined[153]. Another study found *CREBBP* fused to the *RHBDFL* gene in small-cell lung cancer and the product was predicted to cause loss of function of the genes involved, however this remains to be confirmed[133]. *EP300*, encoding a protein related to CREBBP, was also recently found in B-cell precursor acute lymphoblastic leukemia patients to participate in genomic fusions. Here, *EP300* was fused to *ZNF384*, whereby the N-terminal portion of *EP300* fused to the first coding exon (exon 3) of *ZNF384*. Although this fusion product is thought to play a role in the development of B cell precursor acute lymphoblastic leukemia (BCP-ALL), its precise role has not been determined yet[154].

One of the BRD-containing proteins with the largest documented number of genomic fusions is MLL, with over 80 reported fusion products occurring with a wide variety of genes in leukemias (recently reviewed elsewhere[155]) . Interestingly, *MLL* fusions found in the majority (~90%) of ALL patients seem to be caused by fusion of *MLL* with a limited number of other genes: *AF4* (also known as *AFF1)*, *AF9* (also known as *MLLT3)* or *ENL* (also known as *MLLT1)*. Genomic fusions of *MLL* are also frequent in acute myeloid leukemia (AML) patients, where *MLL* is most frequently found fused to *AF9*, *AF10* ,*ELL* or *AF6*. These frequently found fusion products seem to be involved in regulating transcriptional elongation either by directly binding RNA polymerase II (RNAPII), or by participating in super elongation complexes, controlling RNAPII.

ZMYND8 has also been recently reported to participate in a genomic translocation in acute erythroid leukemia. Here, *ZMYND8* was fused with *RELA* , whereby almost the entire *ZMYND8* fuses to the entire *RELA* gene. The product of this fusion might act as leukemia-promoting factor by modulating several cellular processes including the NF-κB pathway[156]. *ZMYND8* was previously also reported fused to *CEP250*[157] and *BCAS4*[158] in breast cancer. These fusions result in elevated expression of the fused product, however, their precise functions in cancer development remain to be determined. The related *ZMYND11* was also recently reported in fusion with *MBTD1* in paediatric acute myeloid leukemia. This fusion product does not contain the MYND interaction domain of *ZMYND11* but encodes the entire *MBTD1* and its expression seems to result in elevated expression of HOXA[159]. In glioblastomas, a recent study found

ASH1L fused to *C1ORF61* resulting in a product that lost the C-terminus of ASH1L including the SET and BRD domains[160]. In breast tumours, *CARM1* was recently found used to *SMARCA4* resulting in loss of BRD function[161] and *TRIM33* was found fused to *POSTN*, resulting in a product that did not encode any part of the TRIM33 sequence[162].

In summary, genes encoding BRD-containing proteins frequently undergo genomic translocation resulting in fusions, which when expressed generate various fusion proteins. These fusion events can influence the expression of BRD-containing proteins as well as confer new protein activities, thereby contributing to pathologic phenotypes.

BRD-containing proteins as tumour suppressors.

In previous subsections we discussed how BRD-containing proteins contribute to cancer development. Notably however, some BRD-containing proteins have been linked to suppression of tumour formation and growth as we will next see (see also Supplementary information S4 (Table)).

One example is BRD7, a component of the mammalian SWI/SNF complex, levels of which have been found reduced in CRC cell lines and cancerous tissues[163] as well as in ovarian cancer tissues[164]. Interestingly, various components of the SWI/SNF complex seem to have context-specific roles: they can potentially act protectively against cancer in specific tissues, while at the same time their mutations affecting multiple subunits seem to be prevalent in certain cancers, with a mutation frequency and pattern as complex as that seen in TP53. It has been therefore proposed that proper function of mammalian SWI/SNF may offer an important mechanism of tumour resistance[54].

Within the TRIM sub-family, TRIM24 has been found both mutated and over-expressed in different tumour types (at least in the liver tissue). On the other hand, TRIM24 also forms complexes with other TRIM family members (such as TRIM28 and TRIM33) which were shown to interact with the ligand-bound retinoic acid receptor, resulting in repression of its transcriptional activity, thereby exerting a tumour suppressive function[165]. This finding further highlights that the role of BRD-containing proteins in tumorigenesis may be highly context dependent. The related protein TRIM33 has also been shown to act as a tumour suppressor. TRIM33 levels were found to be inversely correlated with the expression of β-catenin in human glioblastoma samples, where the canonical Wnt/β-catenin pathway seems to be a significant determinant of disease initiation and progression. This led to the identification of a mechanism,

whereby β -catenin is destabilized by the E3 ubiquitin ligase activity of TRIM33, resulting in its degradation[166].

Also various other BRD-containing proteins, from diverse families have been shown to exhibit tumour-suppressive properties. ZMYND11, for example, was found to specifically repress oncogene expression *in vivo* and low levels of ZMYND11 in triple-negative breast cancer patients were found to correlate with worse disease-free survival, offering a potential biomarker in this cancer type[106]. The polybromo protein PB1 was found to regulate the expression of the cell cycle inhibitor p21; truncations and mutations in PB1 seem to be a cause of renal cell carcinoma (RCC) and other cancer types[139, 167, 168]. Finally, both SMARCA2 and SMARCA4 have been identified as tumour suppressors and their roles have been recently reviewed elsewhere[169].

Conclusions and perspectives

BRD modules are found in proteins that play key roles in cell homeostasis, linking the interpretation of acetylated lysine signals, particularly on histones, to a broad set of responses in transcription. The versatility of BRD module structures emerging from available structural data, the widespread expression of BRD-containing proteins, displaying both cell and tissue specificity as well as the large number of acetylated sites found in proteomic screens point towards a broader role of these versatile proteins, beyond the recognition of acetylated histones and transcriptional regulation (**Box 1**). It will be interesting to see how recognition of acetylation found on non-histone proteins affects signalling and how this is perturbed, if at all, when BRD-containing proteins are deregulated in disease. One of the challenges that still remains to move this research forward is the development of good and extremely selective antibodies that could recognize single Kac epitopes on BRD target proteins, beyond histones. Importantly, while histone Kac-specific antibodies show some degree of specificity[170, 171], larger acylations (for example lysine crotonylation) have already been successfully detected with specific antibodies[39]. The recent discovery of these non-acetyl acylations and their deposition by HATs also point towards novel, yet to be determined, functional roles of BRD-containing proteins, which could potentially contribute to interpreting these unusual chemical signals, linking them to novel cellular phenotypes.

The developed chemical toolset targeting BRDs offers possibilities to validate BRD-dependent interactions, to uncover new functions of BRD-containing proteins and to delineate their roles in

normal and disease tissue, demystifying their functions. It remains to be seen how the modulation of acylation-dependent circuitries linked to BRD-driven recognition of these PTMs will lead to novel, improved therapeutics. Encouragingly however, current efforts focusing on interfering with the BRD-mediated recognition of Kac (for details see Box 2 and refer to Supplementary information S7 (Table) and S8 (Table)) suggest that there is, indeed, a therapeutic window for controlling and modulating the activity of these versatile proteins. Altogether, current progress in the mechanisms of Kac recognition by BRD modules, together with elucidation of functions of BRD-containing proteins in physiology and disease hold great promise for developing novel therapeutics to complement current therapies based on the inhibition of the removal of lysine acetylation by histone de-acetylase inhibitors.

Acknowledgements

We apologize to all authors whose important contributions could not be acknowledged due to space limitations. The authors are grateful for support for their research received from the Ludwig Institute for Cancer Research and the SGC, a registered charity (number 1097737) that receives funds from AbbVie, Bayer Pharma AG, Boehringer Ingelheim, Canada Foundation for Innovation, Eshelman Institute for Innovation, Genome Canada, Innovative Medicines Initiative (EU/EFPIA) [ULTRA-DD grant no. 115766], Janssen, Merck & Co., Novartis Pharma AG, Ontario Ministry of Economic Development and Innovation, Pfizer, São Paulo Research Foundation-FAPESP, Takeda, and Wellcome Trust (092809/Z/10/Z). T.F. is supported by a Uehara Memorial Foundation Fellowship (201330102). P.F. is supported by a Wellcome Trust Career-Development Fellowship (095751/Z/11/Z).

References

1. Kim, S.C., Sprung, R., Chen, Y., Xu, Y., Ball, H., Pei, J. et al. Substrate and functional diversity of lysine acetylation revealed by a proteomics survey. *Mol Cell* **23**, 607-18 (2006).
2. Choudhary, C., Kumar, C., Gnad, F., Nielsen, M.L., Rehman, M., Walther, T.C. et al. Lysine Acetylation Targets Protein Complexes and Co-Regulates Major Cellular Functions. *Science* **325**, 834-840 (2009).
3. Choudhary, C., Weinert, B.T., Nishida, Y., Verdin, E. & Mann, M. The growing landscape of lysine acetylation links metabolism and cell signalling. *Nat Rev Mol Cell Biol* **15**, 536-50 (2014).
4. Verdin, E. & Ott, M. 50 years of protein acetylation: from gene regulation to epigenetics, metabolism and beyond. *Nat Rev Mol Cell Biol* **16**, 258-64 (2015).
5. Bannister, A.J. & Kouzarides, T. Regulation of chromatin by histone modifications. *Cell Res* **21**, 381-95 (2011).

6. Glozak, M.A. & Seto, E. Histone deacetylases and cancer. *Oncogene* **26**, 5420-32 (2007).

7. Zhao, D., Li, F.L., Cheng, Z.L. & Lei, Q.Y. Impact of acetylation on tumor metabolism. *Mol Cell Oncol* **1**, e963452 (2014).

8. Falkenberg, K.J. & Johnstone, R.W. Histone deacetylases and their inhibitors in cancer, neurological diseases and immune disorders. *Nat Rev Drug Discov* **13**, 673-91 (2014).

9. Haynes, S.R., Dollard, C., Winston, F., Beck, S., Trowsdale, J. & Dawid, I.B. The bromodomain: a conserved sequence found in human, Drosophila and yeast proteins. *Nucleic Acids Res* **20**, 2603 (1992).

10. Li, Y., Wen, H., Xi, Y., Tanaka, K., Wang, H., Peng, D. et al. AF9 YEATS Domain Links Histone Acetylation to DOT1L-Mediated H3K79 Methylation. *Cell* **159**, 558-571 (2014).

11. Li, Y., Sabari, B.R., Panchenko, T., Wen, H., Zhao, D., Guan, H. et al. Molecular Coupling of Histone Crotonylation and Active Transcription by AF9 YEATS Domain. *Mol Cell* **62**, 181-93 (2016).

12. Andrews, F.H., Shinsky, S.A., Shanle, E.K., Bridgers, J.B., Gest, A., Tsun, I.K. et al. The Taf14 YEATS domain is a reader of histone crotonylation. *Nat Chem Biol* **12**, 396-8 (2016).

13. Kim, M.S., Pinto, S.M., Getnet, D., Nirujogi, R.S., Manda, S.S., Chaerkady, R. et al. A draft map of the human proteome. *Nature* **509**, 575-81 (2014).

14. Wilhelm, M., Schlegl, J., Hahne, H., Gholami, A.M., Lieberenz, M., Savitski, M.M. et al. Mass-spectrometry-based draft of the human proteome. *Nature* **509**, 582-584 (2014).

15. Muller, S., Filippakopoulos, P. & Knapp, S. Bromodomains as therapeutic targets. *Expert Rev Mol Med* **13**, e29 (2011).

16. Belkina, A.C. & Denis, G.V. BET domain co-regulators in obesity, inflammation and cancer. *Nat Rev Cancer* **12**, 465-77 (2012).

17. Shi, J. & Vakoc, C.R. The mechanisms behind the therapeutic activity of BET bromodomain inhibition. *Mol Cell* **54**, 728-36 (2014).

18. Wang, C.Y. & Filippakopoulos, P. Beating the odds: BETs in disease. *Trends Biochem Sci* **40**, 468-79 (2015).

19. Filippakopoulos, P. & Knapp, S. Targeting bromodomains: epigenetic readers of lysine acetylation. *Nat Rev Drug Discov* **13**, 337-56 (2014).

20. Basheer, F. & Huntly, B.J. BET bromodomain inhibitors in leukemia. *Exp Hematol* **43**, 718-31 (2015).

21. Theodoulou, N.H., Tomkinson, N.C., Prinjha, R.K. & Humphreys, P.G. Progress in the Development of non-BET Bromodomain Chemical Probes. *ChemMedChem* **11**, 477-87 (2016).

22. Filippakopoulos, P., Picaud, S., Mangos, M., Keates, T., Lambert, J.P., Barsyte-Lovejoy, D. et al. Histone recognition and large-scale structural analysis of the human bromodomain family. *Cell* **149**, 214-31 (2012).

23. Alsarraj, J., Faraji, F., Geiger, T.R., Mattaini, K.R., Williams, M., Wu, J. et al. BRD4 short isoform interacts with RRP1B, SIPA1 and components of the LINC complex at the inner face of the nuclear membrane. *PLoS One* **8**, e80746 (2013).

24. Dhalluin, C., Carlson, J.E., Zeng, L., He, C., Aggarwal, A.K. & Zhou, M.M. Structure and ligand of a histone acetyltransferase bromodomain. *Nature* **399**, 491-6 (1999).

25. Jacobson, R.H., Ladurner, A.G., King, D.S. & Tjian, R. Structure and function of a human TAFII250 double bromodomain module. *Science* **288**, 1422-5 (2000).

26. Owen, D.J., Ornaghi, P., Yang, J.C., Lowe, N., Evans, P.R., Ballario, P. et al. The structural basis for the recognition of acetylated histone H4 by the bromodomain of histone acetyltransferase gcn5p. *EMBO J* **19**, 6141-9 (2000).

27. Mujtaba, S., He, Y., Zeng, L., Farooq, A., Carlson, J.E., Ott, M. et al. Structural basis of lysine-acetylated HIV-1 Tat recognition by PCAF bromodomain. *Mol Cell* **9**, 575-86 (2002).

28. Mujtaba, S., He, Y., Zeng, L., Yan, S., Plotnikova, O., Sachchidanand et al. Structural mechanism of the bromodomain of the coactivator CBP in p53 transcriptional activation. *Mol Cell* **13**, 251-63 (2004).

29. Gamsjaeger, R., Webb, S.R., Lamonica, J.M., Billin, A., Blobel, G.A. & Mackay, J.P. Structural basis and specificity of acetylated transcription factor GATA1 recognition by BET family bromodomain protein Brd3. *Mol Cell Biol* **31**, 2632-40 (2011).

30. Schroder, S., Cho, S., Zeng, L., Zhang, Q., Kaehlcke, K., Mak, L. et al. Two-pronged binding with bromodomain-containing protein 4 liberates positive transcription elongation factor b from inactive ribonucleoprotein complexes. *J Biol Chem* **287**, 1090-9 (2012).

31. Shi, J., Wang, Y., Zeng, L., Wu, Y., Deng, J., Zhang, Q. et al. Disrupting the interaction of BRD4 with diacetylated Twist suppresses tumorigenesis in basal-like breast cancer. *Cancer Cell* **25**, 210-25 (2014).

32. Zou, Z., Huang, B., Wu, X., Zhang, H., Qi, J., Bradner, J. et al. Brd4 maintains constitutively active NF-kappaB in cancer cells by binding to acetylated RelA. *Oncogene* **33**, 2395-404 (2014).

33. Tsai, W.W., Wang, Z., Yiu, T.T., Akdemir, K.C., Xia, W., Winter, S. et al. TRIM24 links a non-canonical histone signature to breast cancer. *Nature* **468**, 927-32 (2010).

34. Li, H., Ilin, S., Wang, W., Duncan, E.M., Wysocka, J., Allis, C.D. et al. Molecular basis for site-specific read-out of histone H3K4me3 by the BPTF PHD finger of NURF. *Nature* **442**, 91-5 (2006).

35. Ruthenburg, A.J., Li, H., Milne, T.A., Dewell, S., McGinty, R.K., Yuen, M. et al. Recognition of a mononucleosomal histone modification pattern by BPTF via multivalent interactions. *Cell* **145**, 692-706 (2011).

36. Xi, Q., Wang, Z., Zaromytidou, A.I., Zhang, X.H., Chow-Tsang, L.F., Liu, J.X. et al. A poised chromatin platform for TGF-beta access to master regulators. *Cell* **147**, 1511-24 (2011).

37. Moriniere, J., Rousseaux, S., Steuerwald, U., Soler-Lopez, M., Curtet, S., Vitte, A.L. et al. Cooperative binding of two acetylation marks on a histone tail by a single bromodomain. *Nature* **461**, 664-8 (2009).

38. Chen, Y., Sprung, R., Tang, Y., Ball, H., Sangras, B., Kim, S.C. et al. Lysine propionylation and butyrylation are novel post-translational modifications in histones. *Mol Cell Proteomics* **6**, 812-9 (2007).

39. Tan, M., Luo, H., Lee, S., Jin, F., Yang, J.S., Montellier, E. et al. Identification of 67 histone marks and histone lysine crotonylation as a new type of histone modification. *Cell* **146**, 1016-28 (2011).

40. Sabari, B.R., Tang, Z., Huang, H., Yong-Gonzalez, V., Molina, H., Kong, H.E. et al. Intracellular crotonyl-CoA stimulates transcription through p300-catalyzed histone crotonylation. *Mol Cell* **58**, 203-15 (2015).

41. Goudarzi, A., Zhang, D., Huang, H., Barral, S., Kwon, O.K., Qi, S. et al. Dynamic Competing Histone H4 K5K8 Acetylation and Butyrylation Are Hallmarks of Highly Active Gene Promoters. *Mol Cell* **62**, 169-80 (2016).

42. Rousseaux, S. & Khochbin, S. Histone Acylation beyond Acetylation: Terra Incognita in Chromatin Biology. *Cell J* **17**, 1-6 (2015).

43. Flynn, E.M., Huang, O.W., Poy, F., Oppikofer, M., Bellon, S.F., Tang, Y. et al. A Subset of Human Bromodomains Recognizes Butyryllysine and Crotonyllysine Histone Peptide Modifications. *Structure* **23**, 1801-14 (2015).

44. Singhal, N., Graumann, J., Wu, G., Arauzo-Bravo, M.J., Han, D.W., Greber, B. et al. Chromatin-Remodeling Components of the BAF Complex Facilitate Reprogramming. *Cell* **141**, 943-55 (2010).

45. Singh, M., Popowicz, G.M., Krajewski, M. & Holak, T.A. Structural ramification for acetyl-lysine recognition by the bromodomain of human BRG1 protein, a central ATPase of the SWI/SNF remodeling complex. *Chembiochem* **8**, 1308-16 (2007).

46. Reyes, J.C., Barra, J., Muchardt, C., Camus, A., Babinet, C. & Yaniv, M. Altered control of cellular proliferation in the absence of mammalian brahma (SNF2alpha). *EMBO J* **17**, 6979-91 (1998).

47. Bultman, S., Gebuhr, T., Yee, D., La Mantia, C., Nicholson, J., Gilliam, A. et al. A Brg1 null mutation in the mouse reveals functional differences among mammalian SWI/SNF complexes. *Mol Cell* **6**, 1287-95 (2000).

48. Huang, X., Gao, X., Diaz-Trelles, R., Ruiz-Lozano, P. & Wang, Z. Coronary development is regulated by ATP-dependent SWI/SNF chromatin remodeling component BAF180. *Dev Biol* **319**, 258-66 (2008).

49. Burrows, A.E., Smogorzewska, A. & Elledge, S.J. Polybromo-associated BRG1-associated factor components BRD7 and BAF180 are critical regulators of p53 required for induction of replicative senescence. *Proc Natl Acad Sci U S A* **107**, 14280-5 (2010).

50. Brownlee, P.M., Chambers, A.L., Cloney, R., Bianchi, A. & Downs, J.A. BAF180 promotes cohesion and prevents genome instability and aneuploidy. *Cell Rep* **6**, 973-81 (2014).

51. Kaeser, M.D., Aslanian, A., Dong, M.Q., Yates, J.R., 3rd & Emerson, B.M. BRD7, a novel PBAF-specific SWI/SNF subunit, is required for target gene activation and repression in embryonic stem cells. *J Biol Chem* **283**, 32254-63 (2008).

52. Chiu, Y.H., Lee, J.Y. & Cantley, L.C. BRD7, a tumor suppressor, interacts with p85alpha and regulates PI3K activity. *Mol Cell* **54**, 193-202 (2014).

53. Park, S.W., Herrema, H., Salazar, M., Cakir, I., Cabi, S., Basibuyuk Sahin, F. et al. BRD7 regulates XBP1s' activity and glucose homeostasis through its interaction with the regulatory subunits of PI3K. *Cell Metab* **20**, 73-84 (2014).

54. Kadoch, C., Hargreaves, D.C., Hodges, C., Elias, L., Ho, L., Ranish, J. et al. Proteomic and bioinformatic analysis of mammalian SWI/SNF complexes identifies extensive roles in human malignancy. *Nat Genet* **45**, 592-601 (2013).

55. Fairbridge, N.A., Dawe, C.E., Niri, F.H., Kooistra, M.K., King-Jones, K. & McDermid, H.E. Cecr2 mutations causing exencephaly trigger misregulation of mesenchymal/ectodermal transcription factors. *Birth Defects Res A Clin Mol Teratol* **88**, 619-25 (2010).

56. Banting, G.S., Barak, O., Ames, T.M., Burnham, A.C., Kardel, M.D., Cooch, N.S. et al. CECR2, a protein involved in neurulation, forms a novel chromatin remodeling complex with SNF2L. *Hum Mol Genet* **14**, 513-24 (2005).

57. Bowser, R., Giambrone, A. & Davies, P. FAC1, a novel gene identified with the monoclonal antibody Alz50, is developmentally regulated in human brain. *Dev Neurosci* **17**, 20-37 (1995).

58. Tallant, C., Valentini, E., Fedorov, O., Overvoorde, L., Ferguson, F.M., Filippakopoulos, P. et al. Molecular basis of histone tail recognition by human TIP5 PHD finger and bromodomain of the chromatin remodeling complex NoRC. *Structure* **23**, 80-92 (2015).

59. Collins, N., Poot, R.A., Kukimoto, I., Garcia-Jimenez, C., Dellaire, G. & Varga-Weisz, P.D. An ACF1-ISWI chromatin-remodeling complex is required for DNA replication through heterochromatin. *Nat Genet* **32**, 627-32 (2002).

60. Xiao, A., Li, H., Shechter, D., Ahn, S.H., Fabrizio, L.A., Erdjument-Bromage, H. et al. WSTF regulates the H2A.X DNA damage response via a novel tyrosine kinase activity. *Nature* **457**, 57-62 (2009).

61. Zhou, Y., Schmitz, K.M., Mayer, C., Yuan, X., Akhtar, A. & Grummt, I. Reversible acetylation of the chromatin remodelling complex NoRC is required for non-coding RNA-dependent silencing. *Nat Cell Biol* **11**, 1010-6 (2009).

62. Jones, M.H., Hamana, N., Nezu, J. & Shimane, M. A novel family of bromodomain genes. *Genomics* **63**, 40-5 (2000).

63. Huang, H., Rambaldi, I., Daniels, E. & Featherstone, M. Expression of the Wdr9 gene and protein products during mouse development. *Dev Dyn* **227**, 608-14 (2003).

64. Philipps, D.L., Wigglesworth, K., Hartford, S.A., Sun, F., Pattabiraman, S., Schimenti, K. et al. The dual bromodomain and WD repeat-containing mouse protein BRWD1 is required for normal spermiogenesis and the oocyte-embryo transition. *Dev Biol* **317**, 72-82 (2008).

65. Pattabiraman, S., Baumann, C., Guisado, D., Eppig, J.J., Schimenti, J.C. & De La Fuente, R. Mouse BRWD1 is critical for spermatid postmeiotic transcription and female meiotic chromosome stability. *J Cell Biol* **208**, 53-69 (2015).

66. Muller, P., Kuttenkeuler, D., Gesellchen, V., Zeidler, M.P. & Boutros, M. Identification of JAK/STAT signalling components by genome-wide RNA interference. *Nature* **436**, 871-5 (2005).

67. Dancy, B.M. & Cole, P.A. Protein lysine acetylation by p300/CBP. *Chem Rev* **115**, 2419-52 (2015).

68. Yao, T.P., Oh, S.P., Fuchs, M., Zhou, N.D., Ch'ng, L.E., Newsome, D. et al. Gene dosage-dependent embryonic development and proliferation defects in mice lacking the transcriptional integrator p300. *Cell* **93**, 361-72 (1998).

69. Tanaka, Y., Naruse, I., Hongo, T., Xu, M., Nakahata, T., Maekawa, T. et al. Extensive brain hemorrhage and embryonic lethality in a mouse null mutant of CREB-binding protein. *Mech Dev* **95**, 133-45 (2000).

70. Kasper, L.H., Fukuyama, T., Biesen, M.A., Boussouar, F., Tong, C., de Pauw, A. et al. Conditional knockout mice reveal distinct functions for the global transcriptional coactivators CBP and p300 in T-cell development. *Mol Cell Biol* **26**, 789-809 (2006).

71. Nagy, Z. & Tora, L. Distinct GCN5/PCAF-containing complexes function as co-activators and are involved in transcription factor and global histone acetylation. *Oncogene* **26**, 5341-57 (2007).

72. Krebs, A.R., Karmodiya, K., Lindahl-Allen, M., Struhl, K. & Tora, L. SAGA and ATAC histone acetyl transferase complexes regulate distinct sets of genes and ATAC defines a class of p300-independent enhancers. *Mol Cell* **44**, 410-23 (2011).

73. Maurice, T., Duclot, F., Meunier, J., Naert, G., Givalois, L., Meffre, J. et al. Altered memory capacities and response to stress in p300/CBP-associated factor (PCAF) histone acetylase knockout mice. *Neuropsychopharmacology* **33**, 1584-602 (2008).

74. Xu, W., Edmondson, D.G., Evrard, Y.A., Wakamiya, M., Behringer, R.R. & Roth, S.Y. Loss of Gcn5l2 leads to increased apoptosis and mesodermal defects during mouse development. *Nat Genet* **26**, 229-32 (2000).

75. Bu, P., Evrard, Y.A., Lozano, G. & Dent, S.Y. Loss of Gcn5 acetyltransferase activity leads to neural tube closure defects and exencephaly in mouse embryos. *Mol Cell Biol* **27**, 3405-16 (2007).

76. Doyon, Y., Selleck, W., Lane, W.S., Tan, S. & Cote, J. Structural and functional conservation of the NuA4 histone acetyltransferase complex from yeast to humans. *Mol Cell Biol* **24**, 1884-96 (2004).

77. Altaf, M., Auger, A., Monnet-Saksouk, J., Brodeur, J., Piquet, S., Cramet, M. et al. NuA4-dependent acetylation of nucleosomal histones H4 and H2A directly stimulates incorporation of H2A.Z by the SWR1 complex. *J Biol Chem* **285**, 15966-77 (2010).

78. Obri, A., Ouararhni, K., Papin, C., Diebold, M.L., Padmanabhan, K., Marek, M. et al. ANP32E is a histone chaperone that removes H2A.Z from chromatin. *Nature* **505**, 648-53 (2014).

79. Yang, X.J. MOZ and MORF acetyltransferases: Molecular interaction, animal development and human disease. *Biochim Biophys Acta* **1853**, 1818-26 (2015).

80. Klein, B.J., Muthurajan, U.M., Lalonde, M.E., Gibson, M.D., Andrews, F.H., Hepler, M. et al. Bivalent interaction of the PZP domain of BRPF1 with the nucleosome impacts chromatin dynamics and acetylation. *Nucleic Acids Res* **44**, 472-84 (2016).

81. You, L., Chen, L., Penney, J., Miao, D. & Yang, X.J. Expression atlas of the multivalent epigenetic regulator Brpf1 and its requirement for survival of mouse embryos. *Epigenetics* **9**, 860-72 (2014).

82. Mishima, Y., Miyagi, S., Saraya, A., Negishi, M., Endoh, M., Endo, T.A. et al. The Hbo1-Brd1/Brpf2 complex is responsible for global acetylation of H3K14 and required for fetal liver erythropoiesis. *Blood* **118**, 2443-53 (2011).

83. Feng, Y., Vlassis, A., Roques, C., Lalonde, M.E., Gonzalez-Aguilera, C., Lambert, J.P. et al. BRPF3-HBO1 regulates replication origin activation and histone H3K14 acetylation. *EMBO J* **35**, 176-92 (2016).

84. Gregory, G.D., Vakoc, C.R., Rozovskaia, T., Zheng, X., Patel, S., Nakamura, T. et al. Mammalian ASH1L is a histone methyltransferase that occupies the transcribed region of active genes. *Mol Cell Biol* **27**, 8466-79 (2007).

85. Tanaka, Y., Kawahashi, K., Katagiri, Z., Nakayama, Y., Mahajan, M. & Kioussis, D. Dual function of histone H3 lysine 36 methyltransferase ASH1 in regulation of Hox gene expression. *PLoS One* **6**, e28171 (2011).

86. Rao, R.C. & Dou, Y. Hijacked in cancer: the KMT2 (MLL) family of methyltransferases. *Nat Rev Cancer* **15**, 334-46 (2015).

87. Milne, T.A., Briggs, S.D., Brock, H.W., Martin, M.E., Gibbs, D., Allis, C.D. et al. MLL targets SET domain methyltransferase activity to Hox gene promoters. *Mol Cell* **10**, 1107-17 (2002).

88. Filippakopoulos, P., Qi, J., Picaud, S., Shen, Y., Smith, W.B., Fedorov, O. et al. Selective inhibition of BET bromodomains. *Nature* **468**, 1067-73 (2010).

89. Nicodeme, E., Jeffrey, K.L., Schaefer, U., Beinke, S., Dewell, S., Chung, C.W. et al. Suppression of inflammation by a synthetic histone mimic. *Nature* **468**, 1119-23 (2010).

90. Wang, F., Liu, H., Blanton, W.P., Belkina, A., Lebrasseur, N.K. & Denis, G.V. Brd2 disruption in mice causes severe obesity without Type 2 diabetes. *Biochem J* **425**, 71-83 (2010).

91. Shang, E., Wang, X., Wen, D., Greenberg, D.A. & Wolgemuth, D.J. Double bromodomain-containing gene Brd2 is essential for embryonic development in mouse. *Dev Dyn* **238**, 908-17 (2009).

92. Houzelstein, D., Bullock, S.L., Lynch, D.E., Grigorieva, E.F., Wilson, V.A. & Beddington, R.S. Growth and early postimplantation defects in mice deficient for the bromodomain-containing protein Brd4. *Mol Cell Biol* **22**, 3794-802 (2002).

93. Lamonica, J.M., Deng, W., Kadauke, S., Campbell, A.E., Gamsjaeger, R., Wang, H. et al. Bromodomain protein Brd3 associates with acetylated GATA1 to promote its chromatin occupancy at erythroid target genes. *Proc Natl Acad Sci U S A* **108**, E159-68 (2011).

94. Stonestrom, A.J., Hsu, S.C., Jahn, K.S., Huang, P., Keller, C.A., Giardine, B.M. et al. Functions of BET proteins in erythroid gene expression. *Blood* **125**, 2825-34 (2015).

95. Jang, M.K., Mochizuki, K., Zhou, M., Jeong, H.S., Brady, J.N. & Ozato, K. The bromodomain protein Brd4 is a positive regulatory component of P-TEFb and stimulates RNA polymerase II-dependent transcription. *Mol Cell* **19**, 523-34 (2005).

96. Yang, Z., Yik, J.H., Chen, R., He, N., Jang, M.K., Ozato, K. et al. Recruitment of P-TEFb for stimulation of transcriptional elongation by the bromodomain protein Brd4. *Mol Cell* **19**, 535-45 (2005).

97. Gaucher, J., Boussouar, F., Montellier, E., Curtet, S., Buchou, T., Bertrand, S. et al. Bromodomain-dependent stage-specific male genome programming by Brdt. *EMBO J* **31**, 3809-20 (2012).

98. Shang, E., Nickerson, H.D., Wen, D., Wang, X. & Wolgemuth, D.J. The first bromodomain of Brdt, a testis-specific member of the BET sub-family of double-bromodomain-containing proteins, is essential for male germ cell differentiation. *Development* **134**, 3507-15 (2007).

99. Matzuk, M.M., McKeown, M.R., Filippakopoulos, P., Li, Q., Ma, L., Agno, J.E. et al. Small-molecule inhibition of BRDT for male contraception. *Cell* **150**, 673-84 (2012).

100. Buchmann, A.M., Skaar, J.R. & DeCaprio, J.A. Activation of a DNA damage checkpoint response in a TAF1-defective cell line. *Mol Cell Biol* **24**, 5332-9 (2004).

101. Kimura, J., Nguyen, S.T., Liu, H., Taira, N., Miki, Y. & Yoshida, K. A functional genome-wide RNAi screen identifies TAF1 as a regulator for apoptosis in response to genotoxic stress. *Nucleic Acids Res* **36**, 5250-9 (2008).

102. Sdelci, S., Lardeau, C.H., Tallant, C., Klepsch, F., Klaiber, B., Bennett, J. et al. Mapping the chemical chromatin reactivation landscape identifies BRD4-TAF1 cross-talk. *Nat Chem Biol* **12**, 504-10 (2016).

103. Wang, P.J. & Page, D.C. Functional substitution for TAF(II)250 by a retroposed homolog that is expressed in human spermatogenesis. *Hum Mol Genet* **11**, 2341-6 (2002).

104. Gong, F., Chiu, L.Y., Cox, B., Aymard, F., Clouaire, T., Leung, J.W. et al. Screen identifies bromodomain protein ZMYND8 in chromatin recognition of transcription-associated DNA damage that promotes homologous recombination. *Genes Dev* **29**, 197-211 (2015).

105. Shen, H., Xu, W., Guo, R., Rong, B., Gu, L., Wang, Z. et al. Suppression of Enhancer Overactivation by a RACK7-Histone Demethylase Complex. *Cell* **165**, 331-42 (2016).

106. Wen, H., Li, Y., Xi, Y., Jiang, S., Stratton, S., Peng, D. et al. ZMYND11 links histone H3.3K36me3 to transcription elongation and tumour suppression. *Nature* **508**, 263-8 (2014).

107. Guo, R., Zheng, L., Park, J.W., Lv, R., Chen, H., Jiao, F. et al. BS69/ZMYND11 reads and connects histone H3.3 lysine 36 trimethylation-decorated chromatin to regulated pre-mRNA processing. *Mol Cell* **56**, 298-310 (2014).

108. Zou, J.X., Revenko, A.S., Li, L.B., Gemo, A.T. & Chen, H.W. ANCCA, an estrogen-regulated AAA+ ATPase coactivator for ERalpha, is required for coregulator occupancy and chromatin modification. *Proc Natl Acad Sci U S A* **104**, 18067-72 (2007).

109. Ciro, M., Prosperini, E., Quarto, M., Grazini, U., Walfridsson, J., McBlane, F. et al. ATAD2 is a novel cofactor for MYC, overexpressed and amplified in aggressive tumors. *Cancer Res* **69**, 8491-8 (2009).

110. Revenko, A.S., Kalashnikova, E.V., Gemo, A.T., Zou, J.X. & Chen, H.W. Chromatin loading of E2F-MLL complex by cancer-associated coregulator ANCCA via reading a specific histone mark. *Mol Cell Biol* **30**, 5260-72 (2010).

111. Leachman, N.T., Brellier, F., Ferralli, J., Chiquet-Ehrismann, R. & Tucker, R.P. ATAD2B is a phylogenetically conserved nuclear protein expressed during neuronal differentiation and tumorigenesis. *Dev Growth Differ* **52**, 747-55 (2010).

112. Morozumi, Y., Boussouar, F., Tan, M., Chaikuad, A., Jamshidikia, M., Colak, G. et al. Atad2 is a generalist facilitator of chromatin dynamics in embryonic stem cells. *J Mol Cell Biol* **8**, 349-62 (2016).

113. Malovannaya, A., Lanz, R.B., Jung, S.Y., Bulynko, Y., Le, N.T., Chan, D.W. et al. Analysis of the human endogenous coregulator complexome. *Cell* **145**, 787-99 (2011).

114. Cammas, F., Oulad-Abdelghani, M., Vonesch, J.L., Huss-Garcia, Y., Chambon, P. & Losson, R. Cell differentiation induces TIF1beta association with centromeric heterochromatin via an HP1 interaction. *J Cell Sci* **115**, 3439-48 (2002).

115. Ivanov, A.V., Peng, H., Yurchenko, V., Yap, K.L., Negorev, D.G., Schultz, D.C. et al. PHD domain-mediated E3 ligase activity directs intramolecular sumoylation of an adjacent bromodomain required for gene silencing. *Mol Cell* **28**, 823-37 (2007).

116. Zeng, L., Yap, K.L., Ivanov, A.V., Wang, X., Mujtaba, S., Plotnikova, O. et al. Structural insights into human KAP1 PHD finger-bromodomain and its role in gene silencing. *Nat Struct Mol Biol* **15**, 626-33 (2008).

117. Tubbs, A.T., Dorsett, Y., Chan, E., Helmink, B., Lee, B.S., Hung, P. et al. KAP-1 promotes resection of broken DNA ends not protected by gamma-H2AX and 53BP1 in G(1)-phase lymphocytes. *Mol Cell Biol* **34**, 2811-21 (2014).

118. Dupont, S., Mamidi, A., Cordenonsi, M., Montagner, M., Zacchigna, L., Adorno, M. et al. FAM/USP9x, a deubiquitinating enzyme essential for TGFbeta signaling, controls Smad4 monoubiquitination. *Cell* **136**, 123-35 (2009).

119. Agricola, E., Randall, R.A., Gaarenstroom, T., Dupont, S. & Hill, C.S. Recruitment of TIF1gamma to chromatin via its PHD finger-bromodomain activates its ubiquitin ligase and transcriptional repressor activities. *Mol Cell* **43**, 85-96 (2011).

120. Kim, J. & Kaartinen, V. Generation of mice with a conditional allele for Trim33. *Genesis* **46**, 329-33 (2008).

121. Khetchoumian, K., Teletin, M., Mark, M., Lerouge, T., Cervino, M., Oulad-Abdelghani, M. et al. TIF1delta, a novel HP1-interacting member of the transcriptional intermediary factor 1 (TIF1) family expressed by elongating spermatids. *J Biol Chem* **279**, 48329-41 (2004).

122. Bloch, D.B., Nakajima, A., Gulick, T., Chiche, J.D., Orth, D., de La Monte, S.M. et al. Sp110 localizes to the PML-Sp100 nuclear body and may function as a nuclear hormone receptor transcriptional coactivator. *Mol Cell Biol* **20**, 6138-46 (2000).

123. Wasylyk, C., Schlumberger, S.E., Criqui-Filipe, P. & Wasylyk, B. Sp100 interacts with ETS-1 and stimulates its transcriptional activity. *Mol Cell Biol* **22**, 2687-702 (2002).

124. Yordy, J.S., Li, R., Sementchenko, V.I., Pei, H., Muise-Helmericks, R.C. & Watson, D.K. SP100 expression modulates ETS1 transcriptional activity and inhibits cell invasion. *Oncogene* **23**, 6654-65 (2004).

125. Podcheko, A., Northcott, P., Bikopoulos, G., Lee, A., Bommareddi, S.R., Kushner, J.A. et al. Identification of a WD40 repeat-containing isoform of PHIP as a novel regulator of beta-cell growth and survival. *Mol Cell Biol* **27**, 6484-96 (2007).

126. Huether, R., Dong, L., Chen, X., Wu, G., Parker, M., Wei, L. et al. The landscape of somatic mutations in epigenetic regulators across 1,000 paediatric cancer genomes. *Nat Commun* **5**, 3630 (2014).

127. Forbes, S.A., Beare, D., Gunasekaran, P., Leung, K., Bindal, N., Boutselakis, H. et al. COSMIC: exploring the world's knowledge of somatic mutations in human cancer. *Nucleic Acids Res* **43**, D805-11 (2015).

128. Cleary, S.P., Jeck, W.R., Zhao, X., Chen, K., Selitsky, S.R., Savich, G.L. et al. Identification of driver genes in hepatocellular carcinoma by exome sequencing. *Hepatology* **58**, 1693-702 (2013).

129. Cancer Genome Atlas Research, N. Comprehensive genomic characterization of squamous cell lung cancers. *Nature* **489**, 519-25 (2012).

130. Liu, J., Lee, W., Jiang, Z., Chen, Z., Jhunjhunwala, S., Haverty, P.M. et al. Genome and transcriptome sequencing of lung cancers reveal diverse mutational and splicing events. *Genome Res* (2012).

131. Comprehensive molecular profiling of lung adenocarcinoma. *Nature* **511**, 543-50 (2014).

132. Liu, L., Huang, J., Wang, K., Li, L., Li, Y., Yuan, J. et al. Identification of hallmarks of lung adenocarcinoma prognosis using whole genome sequencing. *Oncotarget* **6**, 38016-28 (2015).

133. Peifer, M., Fernandez-Cuesta, L., Sos, M.L., George, J., Seidel, D., Kasper, L.H. et al. Integrative genome analyses identify key somatic driver mutations of small-cell lung cancer. *Nat Genet* **44**, 1104-10 (2012).

134. Grunwald, C., Koslowski, M., Arsiray, T., Dhaene, K., Praet, M., Victor, A. et al. Expression of multiple epigenetically regulated cancer/germline genes in nonsmall cell lung cancer. *Int J Cancer* **118**, 2522-8 (2006).

135. Zheng, C.X., Gu, Z.H., Han, B., Zhang, R.X., Pan, C.M., Xiang, Y. et al. Whole-exome sequencing to identify novel somatic mutations in squamous cell lung cancers. *Int J Oncol* **43**, 755-64 (2013).

136. Ho, A.S., Kannan, K., Roy, D.M., Morris, L.G., Ganly, I., Katabi, N. et al. The mutational landscape of adenoid cystic carcinoma. *Nat Genet* **45**, 791-8 (2013).

137. Lourdusamy, A., Rahman, R. & Grundy, R.G. Expression alterations define unique molecular characteristics of spinal ependymomas. *Oncotarget* **6**, 19780-91 (2015).

138. Odejide, O., Weigert, O., Lane, A.A., Toscano, D., Lunning, M.A., Kopp, N. et al. A targeted mutational landscape of angioimmunoblastic T-cell lymphoma. *Blood* **123**, 1293-6 (2014).

139. Petrini, I., Meltzer, P.S., Kim, I.K., Lucchi, M., Park, K.S., Fontanini, G. et al. A specific missense mutation in GTF2I occurs at high frequency in thymic epithelial tumors. *Nat Genet* **46**, 844-9 (2014).

140. Ojesina, A.I., Lichtenstein, L., Freeman, S.S., Pedamallu, C.S., Imaz-Rosshandler, I., Pugh, T.J. et al. Landscape of genomic alterations in cervical carcinomas. *Nature* **506**, 371-5 (2014).

141. Shang, P., Meng, F., Liu, Y. & Chen, X. Overexpression of ANCCA/ATAD2 in endometrial carcinoma and its correlation with tumor progression and poor prognosis. *Tumour Biol* **36**, 4479-85 (2015).

142. Okosun, J., Bodor, C., Wang, J., Araf, S., Yang, C.Y., Pan, C. et al. Integrated genomic analysis identifies recurrent mutations and evolution patterns driving the initiation and progression of follicular lymphoma. *Nat Genet* **46**, 176-81 (2014).

143. Song, Y., Li, L., Ou, Y., Gao, Z., Li, E., Li, X. et al. Identification of genomic alterations in oesophageal squamous cell cancer. *Nature* **509**, 91-5 (2014).

144. Gao, Y.B., Chen, Z.L., Li, J.G., Hu, X.D., Shi, X.J., Sun, Z.M. et al. Genetic landscape of esophageal squamous cell carcinoma. *Nat Genet* **46**, 1097-102 (2014).

145. Zhang, L., Zhou, Y., Cheng, C., Cui, H., Cheng, L., Kong, P. et al. Genomic analyses reveal mutational signatures and frequently altered genes in esophageal squamous cell carcinoma. *Am J Hum Genet* **96**, 597-611 (2015).

146. Oh, H.R., An, C.H., Yoo, N.J. & Lee, S.H. Somatic mutations of amino acid metabolism-related genes in gastric and colorectal cancers and their regional heterogeneity--a short report. *Cell Oncol (Dordr)* **37**, 455-61 (2014).

147. Zhang, L.H., Yin, A.A., Cheng, J.X., Huang, H.Y., Li, X.M., Zhang, Y.Q. et al. TRIM24 promotes glioma progression and enhances chemoresistance through activation of the PI3K/Akt signaling pathway. *Oncogene* **34**, 600-10 (2015).

148. Chen, Y., Guo, Y., Yang, H., Shi, G., Xu, G., Shi, J. et al. TRIM66 overexpresssion contributes to osteosarcoma carcinogenesis and indicates poor survival outcome. *Oncotarget* **6**, 23708-19 (2015).

149. Kuroyanagi, J., Shimada, Y., Zhang, B., Ariyoshi, M., Umemoto, N., Nishimura, Y. et al. Zinc finger MYND-type containing 8 promotes tumour angiogenesis via induction of vascular endothelial growth factor-A expression. *FEBS Lett* **588**, 3409-16 (2014).

150. Alekseyenko, A.A., Walsh, E.M., Wang, X., Grayson, A.R., Hsi, P.T., Kharchenko, P.V. et al. The oncogenic BRD4-NUT chromatin regulator drives aberrant transcription within large topological domains. *Genes Dev* **29**, 1507-23 (2015).

151. Andreasen, S., French, C.A., Josiassen, M., Hahn, C.H. & Kiss, K. NUT Carcinoma of the Sublingual Gland. *Head Neck Pathol* (2015).

152. Reynoird, N., Schwartz, B.E., Delvecchio, M., Sadoul, K., Meyers, D., Mukherjee, C. et al. Oncogenesis by sequestration of CBP/p300 in transcriptionally inactive hyperacetylated chromatin domains. *EMBO J* **29**, 2943-52 (2010).

153. Pleasance, E.D., Stephens, P.J., O'Meara, S., McBride, D.J., Meynert, A., Jones, D. et al. A small-cell lung cancer genome with complex signatures of tobacco exposure. *Nature* **463**, 184-90 (2010).

154. Gocho, Y., Kiyokawa, N., Ichikawa, H., Nakabayashi, K., Osumi, T., Ishibashi, T. et al. A novel recurrent EP300-ZNF384 gene fusion in B-cell precursor acute lymphoblastic leukemia. *Leukemia* **29**, 2445-8 (2015).

155. Marschalek, R. Systematic Classification of Mixed-Lineage Leukemia Fusion Partners Predicts Additional Cancer Pathways. *Ann Lab Med* **36**, 85-100 (2016).

156. Panagopoulos, I., Micci, F., Thorsen, J., Haugom, L., Buechner, J., Kerndrup, G. et al. Fusion of ZMYND8 and RELA genes in acute erythroid leukemia. *PLoS One* **8**, e63663 (2013).

157. Edgren, H., Murumagi, A., Kangaspeska, S., Nicorici, D., Hongisto, V., Kleivi, K. et al. Identification of fusion genes in breast cancer by paired-end RNA-sequencing. *Genome Biol* **12**, R6 (2011).

158. Kalyana-Sundaram, S., Shankar, S., Deroo, S., Iyer, M.K., Palanisamy, N., Chinnaiyan, A.M. et al. Gene fusions associated with recurrent amplicons represent a class of passenger aberrations in breast cancer. *Neoplasia* **14**, 702-8 (2012).

159. de Rooij, J.D., van den Heuvel-Eibrink, M.M., Kollen, W.J., Sonneveld, E., Kaspers, G.J., Beverloo, H.B. et al. Recurrent translocation t(10;17)(p15;q21) in minimally differentiated acute myeloid leukemia results in ZMYND11/MBTD1 fusion. *Genes Chromosomes Cancer* **55**, 237-41 (2016).

160. Frattini, V., Trifonov, V., Chan, J.M., Castano, A., Lia, M., Abate, F. et al. The integrated landscape of driver genomic alterations in glioblastoma. *Nat Genet* **45**, 1141-9 (2013).

161. Robinson, D.R., Kalyana-Sundaram, S., Wu, Y.M., Shankar, S., Cao, X., Ateeq, B. et al. Functionally recurrent rearrangements of the MAST kinase and Notch gene families in breast cancer. *Nat Med* **17**, 1646-51 (2011).

162. Asmann, Y.W., Necela, B.M., Kalari, K.R., Hossain, A., Baker, T.R., Carr, J.M. et al. Detection of redundant fusion transcripts as biomarkers or disease-specific therapeutic targets in breast cancer. *Cancer Res* **72**, 1921-8 (2012).

163. Wu, W.J., Hu, K.S., Chen, D.L., Zeng, Z.L., Luo, H.Y., Wang, F. et al. Prognostic relevance of BRD7 expression in colorectal carcinoma. *Eur J Clin Invest* **43**, 131-40 (2013).

164. Park, Y.A., Lee, J.W., Kim, H.S., Lee, Y.Y., Kim, T.J., Choi, C.H. et al. Tumor suppressive effects of bromodomain-containing protein 7 (BRD7) in epithelial ovarian carcinoma. *Clin Cancer Res* **20**, 565-75 (2014).

165. Herquel, B., Ouararhni, K., Khetchoumian, K., Ignat, M., Teletin, M., Mark, M. et al. Transcription cofactors TRIM24, TRIM28, and TRIM33 associate to form regulatory complexes that suppress murine hepatocellular carcinoma. *Proc Natl Acad Sci U S A* **108**, 8212-7 (2011).

166. Xue, J., Chen, Y., Wu, Y., Wang, Z., Zhou, A., Zhang, S. et al. Tumour suppressor TRIM33 targets nuclear beta-catenin degradation. *Nat Commun* **6**, 6156 (2015).

167. Varela, I., Tarpey, P., Raine, K., Huang, D., Ong, C.K., Stephens, P. et al. Exome sequencing identifies frequent mutation of the SWI/SNF complex gene PBRM1 in renal carcinoma. *Nature* **469**, 539-42 (2011).

168. Jiao, Y., Pawlik, T.M., Anders, R.A., Selaru, F.M., Streppel, M.M., Lucas, D.J. et al. Exome sequencing identifies frequent inactivating mutations in BAP1, ARID1A and PBRM1 in intrahepatic cholangiocarcinomas. *Nat Genet* **45**, 1470-3 (2013).

169. Wilson, B.G. & Roberts, C.W. SWI/SNF nucleosome remodellers and cancer. *Nat Rev Cancer* **11**, 481-92 (2011).

170. Egelhofer, T.A., Minoda, A., Klugman, S., Lee, K., Kolasinska-Zwierz, P., Alekseyenko, A.A. et al. An assessment of histone-modification antibody quality. *Nat Struct Mol Biol* **18**, 91-3 (2011).

171. Rothbart, S.B., Dickson, B.M., Raab, J.R., Grzybowski, A.T., Krajewski, K., Guo, A.H. et al. An Interactive Database for the Assessment of Histone Antibody Specificity. *Mol Cell* **59**, 502-11 (2015).

172. Zhang, G., Liu, R., Zhong, Y., Plotnikov, A.N., Zhang, W., Zeng, L. et al. Down-regulation of NF-kappaB transcriptional activity in HIV-associated kidney disease by BRD4 inhibition. *J Biol Chem* **287**, 28840-51 (2012).

173. Li, A.G., Piluso, L.G., Cai, X., Gadd, B.J., Ladurner, A.G. & Liu, X. An acetylation switch in p53 mediates holo-TFIID recruitment. *Mol Cell* **28**, 408-21 (2007).

174. Ciceri, P., Muller, S., O'Mahony, A., Fedorov, O., Filippakopoulos, P., Hunt, J.P. et al. Dual kinase-bromodomain inhibitors for rationally designed polypharmacology. *Nat Chem Biol* **10**, 305-12 (2014).

175. Martin, M.P., Olesen, S.H., Georg, G.I. & Schonbrunn, E. Cyclin-dependent kinase inhibitor dinaciclib interacts with the acetyl-lysine recognition site of bromodomains. *ACS Chem Biol* **8**, 2360-5 (2013).

176. Dittmann, A., Werner, T., Chung, C.W., Savitski, M.M., Falth Savitski, M., Grandi, P. et al. The Commonly Used PI3-Kinase Probe LY294002 Is an Inhibitor of BET Bromodomains. *ACS Chem Biol* **9**, 495-502 (2014).

177. Ember, S.W., Zhu, J.Y., Olesen, S.H., Martin, M.P., Becker, A., Berndt, N. et al. Acetyl-lysine Binding Site of Bromodomain-Containing Protein 4 (BRD4) Interacts with Diverse Kinase Inhibitors. *ACS Chem Biol* (2014).

178. Allen, B.K., Mehta, S., Ember, S.W., Schonbrunn, E., Ayad, N. & Schurer, S.C. Large-Scale Computational Screening Identifies First in Class Multitarget Inhibitor of EGFR Kinase and BRD4. *Sci Rep* **5**, 16924 (2015).

179. Kurimchak, A.M., Shelton, C., Duncan, K.E., Johnson, K.J., Brown, J., O'Brien, S. et al. Resistance to BET Bromodomain Inhibitors Is Mediated by Kinome Reprogramming in Ovarian Cancer. *Cell Rep* **16**, 1273-86 (2016).

180. Dawson, M.A., Prinjha, R.K., Dittmann, A., Giotopoulos, G., Bantscheff, M., Chan, W.I. et al. Inhibition of BET recruitment to chromatin as an effective treatment for MLL-fusion leukaemia. *Nature* **478**, 529-33 (2011).

181. Stathis, A., Zucca, E., Bekradda, M., Gomez-Roca, C., Delord, J.P., de La Motte Rouge, T. et al. Clinical Response of Carcinomas Harboring the BRD4-NUT Oncoprotein to the Targeted Bromodomain Inhibitor OTX015/MK-8628. *Cancer Discov* (2016).

182. Bolden, J.E., Tasdemir, N., Dow, L.E., van Es, J.H., Wilkinson, J.E., Zhao, Z. et al. Inducible in vivo silencing of Brd4 identifies potential toxicities of sustained BET protein inhibition. *Cell Rep* **8**, 1919-29 (2014).

183. Di Micco, R., Fontanals-Cirera, B., Low, V., Ntziachristos, P., Yuen, S.K., Lovell, C.D. et al. Control of embryonic stem cell identity by BRD4-dependent transcriptional elongation of super-enhancer-associated pluripotency genes. *Cell Rep* **9**, 234-47 (2014).

184. Korb, E., Herre, M., Zucker-Scharff, I., Darnell, R.B. & Allis, C.D. BET protein Brd4 activates transcription in neurons and BET inhibitor Jq1 blocks memory in mice. *Nat Neurosci* **18**, 1464-73 (2015).

185. Fong, C.Y., Gilan, O., Lam, E.Y., Rubin, A.F., Ftouni, S., Tyler, D. et al. BET inhibitor resistance emerges from leukaemia stem cells. *Nature* **525**, 538-42 (2015).

186. Rathert, P., Roth, M., Neumann, T., Muerdter, F., Roe, J.S., Muhar, M. et al. Transcriptional plasticity promotes primary and acquired resistance to BET inhibition. *Nature* **525**, 543-7 (2015).

187. Shu, S., Lin, C.Y., He, H.H., Witwicki, R.M., Tabassum, D.P., Roberts, J.M. et al. Response and resistance to BET bromodomain inhibitors in triple-negative breast cancer. *Nature* **529**, 413-7 (2016).

188. Togel, L., Nightingale, R., Chueh, A.C., Jayachandran, A., Tran, H., Phesse, T. et al. Dual targeting of bromodomain and extra-terminal domain proteins, and WNT or MAPK signaling, inhibits c-MYC expression and proliferation of colorectal cancer cells. *Mol Cancer Ther* (2016).

Display items

Box 1 | Recognition of non-histone acetylated proteins

Recognition and binding of BRDs to non-histone acetylated motifs has also been described in a number of studies involving transcription factors such as GATA1[29], the RelA subunit of NFκB[32, 172] or TWIST[31]. In these cases, BET BRDs stabilize the complex of the transcription factor with acetylated chromatin, affecting the transcription of its target genes. A similar mode of interaction was described for the complex of cyclin dependent kinase 9 (CDK9) and cyclinT, whereby the recruitment of this complex to transcriptional start sites is facilitated by BRD4 and involves simultaneous interactions via the C-terminal motif of BRD4 as well as its BRDs which recognize

and bind to acetylated cyclinT[30]. The detailed interactions of BET BRDs with non-histone acetylated factors have been recently reviewed elsewhere[18]. Beyond BRD-containing proteins of the BET sub-family , TAF1 has also been shown to bind to di-acetylated p53, regulating p53-mediated transcription[173]. CREBBP has also been shown to bind to C-terminally acetylated p53 following DNA damage, controlling G1 cell cycle arrest[28].

The various interactions of BRD-containing proteins with non-histone targets are shown in the accompanying Figure, depicting structures of: the complex of CREBBP with p53 (acetylated at K382, PDB: 1JSP) determined by NMR; the first BRD4 module BRD4(1) in complex with RelAK310ac (PDB: 4KV1); the NMR-resolved complex of the first BRD module of BRD3, BRD3(1), with GATA1 (acetylated at K312 and K315, PDB: 2L5E) and the complex of the second BRD4 module BRD4(2) with TWIST (acetylated at K73 and K76, PDB: 2MJV). The location of the ZA-loop and the visible helices are annotated in the figure. Acteyl-lysines inserted in the BRD cavities are highlighted. Peptides are shown as ball and stick representations with a transparent surface indicating the radii of van der Waals interactions.

Box 2 | BRD inhibitor development & Clinical translation

Inhibitor development

The hydrophobic nature of the acetyl-lysine binding cavity of BRDs was recently found to be tractable for the development of small molecule inhibitors, yielding many potent and selective classes of compounds. The main target so far was the BET family of BRD-containing proteins. BET proteins are linked to transcriptional stimulation, by providing a tissue-specific scaffolding platform of recruitment, tethering transcriptionally active complexes to acetylated histones and chromatin, thus controlling lineage specific genes linked to cell cycle progression[18]. BET inhibition therefore results in very strong phenotypes, justifying the discovery and development of tools to attenuate their function. The potent and selective structural classes of BET inhibitors, thieno- and benzo- diazepines, emerged from phenotypic screening following long years of development of these compounds. This was followed by the rapid translation of these highly optimized molecules to the clinic (see below). This rapidly growing area has been recently reviewed elsewhere[17, 19-21].

Many non-BET directed chemical tool compounds exhibiting high potency and selectivity while retaining cellular activity have also been developed, targeting distinct subsets of BRD modules within different sub-families. A non-exhaustive list of these tools is summarized in

Supplementary information S7 (Table). Making these compounds available for research will be of extreme importance in the immediate future, and their use will allow uncovering novel roles for the respective BRD-containing proteins, as well as validating their contribution to already established phenotypes.

We envision that this toolkit will be instrumental in validating targets and activities of BRD-containing proteins, informing on potential drug targets in diverse biological settings. Of particular interest were recent reports suggesting that BET family BRDs can be potently inhibited by chemical tools developed specifically against kinases[174-178], offering novel opportunities to explore poly-pharmacology, simultaneously targeting both protein classes. This is an appealing possibility, since recent data suggests that resistance to BET inhibition (at least in ovarian cancer models) seems to be mediated by the activation of pro-survival receptor tyrosine kinase networks which oppose the BET-inhibition-induced apoptotic programmes[179]. Poly-pharmacology may be a useful tool for overcoming this type of resistance.

Clinical translation of BRD inhibitors

The initial success of BET inhibitors, particularly the thieno-diazepine JQ1[88], the benzo-diazepine iBET[89] and the isoxazole iBET151[180], resulted in rapid clinical translation; currently 16 active clinical trials in various phases of development (summarized in Supplementary information S8 (Table)) seek to target BET proteins in haematological and solid tumour settings[19]. Initial data published from the NUT midline carcinoma trials reported improved patient response to inhibition with minor toxicity, establishing the proof of concept of clinical activity towards the BRD4-NUT onco-fusion[181]. It remains to be seen whether and how BET inhibition will apply to other tumour types.

Importantly, recent reports in animal models revealed that prolonged BET inhibition resulted in developmental defects, decreased cellular diversity and stem cell depletion in the small intestine[182]. Furthermore, depletion of stem cell pools affecting self-renewal ability and pluripotency of embryonic stem cells[183] as well as indications of memory-deficits in mice[184], have raised concerns. This collectively indicates that BET inhibitors can be toxic and that better molecules as well as better understanding of the precise biological functions of these proteins are urgently needed. In addition, resistance to BET inhibition (resulting in re-activation of the *MYC* oncogene) was recently reported in leukemia cell lines; suppression of MYC expression has been recognized as a hallmark of BET inhibition, however in the inhibitor-resistant cells, MYC expression was re-established in a mechanism involving Wnt signalling machinery, which acts independently of BRD4 transcriptional stimulation[185, 186]. Resistance was also observed in triple

negative breast tumours where cells remained dependent on BRD4, however cell proliferation and transcription were found to be decoupled from BRD4-bromodomain functions[187]. Interestingly a recent report suggests that combinatorial targeting of BRD4 and factors that occupy the *MYC* super-enhancer, belonging to the WNT-β-catenin-TCF or MEK-ERK pathways, enhances the effects of therapeutic suppression of MYC expression achieved by BET inhibitors[188]. In the future it will be interesting to see if resistance to BET inhibition can be alleviated or prevented also by other combination treatments.

Figures

Figure 1 | **Tissue expression and domain organization of BRD-containing proteins. a** | Recent advances in proteomics have allowed determining tissue-specific expression of proteins. BRD-containing proteins exhibit wide distribution between tissues, potentially informing on their context and cell-dependent roles in cellular homeostasis. Data compiled from the draft proteome (http://www.humanproteomemap.org)[13] and clustered according to BRD structural families (for details on families of BRD-containing proteins see Supplementary information S1 (Table) and S2 (Figure)). **b** | BRD modules are components of multi-domain proteins with diverse functions (see also Supplementary information S1 (Table)). The modular architecture combines protein-protein interaction domains, such as the Plant Homeodomain (PHD), Kinase-inducible interacting domain (KIX), Bromo-Adjacent Homology domain (BAH), as well as catalytic domains such as histone acetyltransferase (HAT) and Helicase domains. The combinatorial existence of such modular functions within each protein facilitates interactions within complexes, recruitment to specific sites as well as directed catalytic activity. Characteristic domain organizations are shown for the P300/CBP-associated factor (PCAF), the bromo and extra-terminal 4 (BRD4), CREB-binding protein (CREBBP), bromodomain and PHD finger-containing protein 1 (BRPF1), tripartite motif-containing protein 33 (TRIM33), Speckled 140 kDa (SP140), poly-bromo containing protein 1 (PB1) and SWI/SNF-related matrix-associated actin-dependent regulator of chromatin subfamily A member 4 (SMARCA4) proteins. Domain abbreviations are explained in the inset.

Figure 2 | **Structure and Plasticity of BRD modules. a** | The structure of the first BRD2 module (BRD2(1))(PDB: 2DQV) is shown and conserved structural elements are coloured and annotated. The four BRD helices (αZ, αA, αB and αC) are linked by flexible loop regions (AB-,

BC- and ZA-loops). A conserved tyrosine (Y135 in BRD2) found on the AB-loop. forms a salt bridge to an aspartate residue (D144), keeping the two helices together. A conserved PxY motif is found on the ZA-loop region, offering another structural template common to most BRD modules. A conserved asparagine on the BC-loop region is responsible for docking o acetylated lysine (Kac) peptides onto BRD modules. **b** | A subset of BRD modules lack the conserved asparagine residue responsible for peptide docking. The structure of SP100 is shown (PDB: 4PTB) where the asparagine is replaced by a tyrosine residue (Y848). **c** | Histone H4 peptide binding onto the N-terminal domain of BRD2. K12ac inserts in the cavity between the body of helixes αB and αC and the ZA-loop while initiating a hydrogen bond to the conserved asparagine on the BC-loop. **d** | Electrostatic charge of the BRD2(1) surface lining up the Kac binding groove where K12ac inserts. Interactions of the peptide backbone with the charged surface lining the Kac binding site drive the binding of histone peptides to the BRD module. Charge on the surface is plotted as indicated in the inset (k: boltzman constant; T: temperature; e: electron charge).

Figure 3 | **Engagement of target sequences by BRD modules. a** | Structure of a histone H4 peptide acetylated at lysine 12 (H4K12ac) bound to the first BRD module of BRD2 (BRD2(1); from PDB: 2DVQ). The acetyl-lysine (Kac) inserts into the hydrophobic-cavity of the BRD directly engaging the conserved asparagine (N156 in BRD2(1)) while initiating a number of water-mediated bridges to the protein backbone. The ZA- and BC-loops (coloured in magenta and purple respectively) direct binding of the N-terminal peptide portion towards helix B, while the C-terminus of the peptide is directed towards the Kac cavity , and inserts between a WPF motif (W97/P98/F99 in BRD2(1)), which is conserved between BET proteins and caps the front of the Kac cavity . Water molecules within the Kac cavity are shown as small spheres with their van der Waals radii highlighted as red transparent spheres. Residues defining the groove within which the peptide binds are annotated. **b** | Structure of the BRD–PHD (plant homeodomain zinc finger domain) module of TRIM33 in complex with a histone H3 peptide (PDB: 3U5O). Binding of the H3 N-terminus (K4 and K9me3 residues) to the PHD finger directs the C-terminal H3 tail towards the Kac binding groove of the BRD. The binding occurs in a reverse orientation as compared to the single BRD–peptide binding mode shown in (a), but engages similar interactions, with H3K18ac inserting in the BRD Kac-cavity. **c** | Mode of binding of a single Kac is conserved between BRD modules and histone tails. The main difference is the orientation of the exit vector (defined by the N- and C- termini) of the peptide with respect to the BRD binding

cavity exemplified here by: the BRPF3–H2AK5ac complex (PDB: 4QYL); the BAZ2B–H3K14ac complex (PDB: 4QC1); the second BRD module of mouse BRDT (BRDT(2))–H3K18ac complex (PDB: 2WP1); and the BRD2(1)–H4K12ac complex (PDB: 2DVQ). Peptides are shown in cartoon (loop) representation with their N- and C-termini annotated, highlighting the entry- and exit-vectors to and from the Kac cavity. Although the central Kac binds in the same position, the topology of the peptide backbone differs in every case, suggesting that contributions from the surface of the cavity direct the relative peptide orientation. **d** | Structures of di-acetylated histone H4 peptides in complex with members of the BET family (mouse BRDT with H4K5acK8ac (PDB: 2WP2); first BRD module of BRD4 (BRD4(1)) with H4K5acK8ac (PDB: 3UVW); BRD4(1) with H4K12acK16ac (PDB: 3UVX); and BRD4(1) with H4K16acK20ac (PDB: 3UVY). While the first Kac initiates direct contact with the conserved asparagine, the second Kac fills the empty space on the front of the BRD cavity and by forming hydrogen bonds with the first Kac, sterically fills the entire space of the cavity. Structural elements are annotated as in (c) and peptides are shown as ball and stick representations with a transparent surface indicating van der Waals radii. **e** | Recognition of non-acetyl lysine modifications by BRDs has also been reported, including crotonylation (complex of BRD9 with an H4K5crot peptide, PDB: 4YYH); propionylation (complex of BRD4(1) with an H3K23prop peptide, PDB: 3MUK); and butyrylation (complex of the second BRD module of TAF1 (TAF1(2)) with an H4K5butK8but peptide, PDB: 4YYM). Structural elements are annotated as in (c).

Figure 4 | **Roles of BRD-containing proteins in gene regulation** BRD-containing proteins regulate gene expression through various activities and mechanisms. **a** | They are frequently found as components of chromatin remodelling complexes. For example, bromodomain adjacent to zinc finger domain protein 1A (BAZ1A) and BAZ2A are components of the assembly factor/Williams syndrome transcription factor (ACF/WSTF) complex and nucleolar remodelling complex (NoRC) respectively, which promote chromatin compaction leading to gene silencing. On the other hand, other BRD-containing proteins including SWI/SNF-related matrix-associated actin-dependent regulator of chromatin subfamily A member 2 (SMARCA2) and SMARCA4 together with BRWD1, BAZ1B, Cat eye syndrome critical region protein 2 (CECR2) and Bromodomain PHD finger Transcription Factor (BPTF) are components of chromatin remodelling complexes that promote chromatin decompaction and acetylation, enhancing transcription (respectively: Switch/Sucrose non-fermenting (SWI/SNF), WICH, CECR2-

containing Remodeling Factor (CERF), NUcleosome Remodelling Factor (NURF) complexes). ATPase family AAA domain-containing protein 2 (ATAD2) has also been implicated in chromatin dynamics, positively modulating transcription (owing to the presence of an AAA-ATPase domain it may act as a histone chaperone promoting nucleosome disassembly). **b** | BRD-containing proteins engage in acetylation and methylation of histones. They can either modify histones directly through their acetyltransferase (CREBBP and EP300) or methyltransferase (mixed lineage leukemia (MLL) and absent small and homeotic disks protein 1 homolog (ASH1L)) activities, or they can be a part of larger histone-modifying complexes (this includes: general control of amino acid synthesis protein 5-like 2 (GCN5L2), P300/CBP-associated factor (PCAF), BRD8 and bromodomain and PHD finger-containing proteins (BRPF)). **c** | BRD-containing proteins can serve important roles in transcription by functioning as histone-recognizing scaffolds that promote the assembly of transcriptional complexes For example bromo and extra-terminal (BET) proteins by recruiting components of the transcriptional machinery positively regulate growth promoting genes, whereas zinc-finger and MYND binding domain 8 (ZMYND8) recruits the Nucleosome Remodelling and Deacetylase (NuRD) complex or lysine demethylase KDM5C, leading to transcriptional repression. Another BRD-containing protein, transcription initiation factor TFIID subunit 1 (TAF1), is the largest component of the core sub-unit of the TFIID transcription factor, which promotes transcriptional initiation by binding to the core promoter sequence and recruiting other transcriptional regulators. **d** | BRD-containing proteins can modulate the activity of the components of the transcriptional machinery. They can directly control transcription factor activity by acting as transcriptional coactivators (for example proteins of the SP family, ZMYND8, TRIM24 and TRIM33). They can also impact transcription by modulating other components. ZMYND11 for example, negatively regulates the activity of RNA Polymerase II.

Figure 5 | **Deregulation of BRD-containing proteins in cancer. a** | Chromosomal translocations reported to contain BRD-containing genes. The resulting fusion products maintain or lose the BRD function(s) resulting in aberrant behaviour and function of the resulting proteins. Notably *MLL* (mixed lineage leukemia gene) is known to participate in more than 80 fusions, however only the most commonly found translocations are annotated here, with *AF4*, *AF6*, *AF9*, *AF10*, *ENL* and *ELL*. Known cases of loss of BRD-function are annotated with a "*". **b** | Examples of translocations resulting in shuffling of modular domains and catalytic functions. The fusion of *BRD4* and *NUT* results in a product that retains the BRD and ET domains of *BRD4*

while gaining almost the entire *NUT* gene. Translocations involving the N-terminal portion of *CREBBP* and the C-terminal portion of *MLL* (CREBBP-MLL fusion products) retain the methyltransferase (SET) activity of *MLL* while losing the N-terminal DNA-binding portion of *MLL*. In contrast, an *MLL-CREBBP* fusion retains the DNA-binding portion of *MLL* while gaining the histone acetyltransferase (HAT) activity of *CREBBP*. **c** | Mechanism of the BRD4-NUT fusion product in the pathology of NUT midline Carcinomas. The fusion protein is recruited to acetylated chromatin via the BRD modules of BRD4 and further recruits via the NUT portion the acetyltransferase P300; this results in local hyper-acetylation of chromatin, further recruiting endogenous BRD4 which acts as a platform for the recruitment of transcriptional regulatory complexes such as Mediator or the positive transcription elongation factor b (P-TEF-b, the complex of the kinase CDK9 and CyclinT), resulting in aberrant transcriptional stimulation.

Figure S1 | **Structural Classification of the human BRD family.** The family of human BRD modules was classified in eight distinct classes, based on structural alignments and secondary structure prediction. Despite high sequence homology, the determined crystal structures (shown with electrostatic potential plotted on the surface of each module) exhibit diverse distribution of surface charge, suggesting great variability of potential recognition-sequences carrying acetylated lysine residues. For example, members of sub-family VIII (eg PB1(3)) contain positively charged regions flanking the central Kac recognition site, suggesting that histone tails may not be the target sequences of these BRD modules. All modules are oriented in the same way, with the Kac-recognition site facing up, as highlighted in the case of PCAF with a dotted circle.

Figure S2 | **Mutation load of BRD modules** | Mutations found on the BRD module of BRD-containing proteins, annotated by residue as a function of frequency, mapped on the structure of representative BRDs from each structural family. Each panel contains a front view (left part) and one rotated by 180˚. Data were compiled using TCGA data harvested via the cBioPortal (http://www.cbioportal.org/). The four BRD helices (αZ, αA, αB and αC) are annotated together with the conserved asparagine, loop regions (ZA- and BC-loops) and the conserved PxY motif. Mutation frequency is indicated by sphere size and colour. The location of mutations as well as their frequency varies between BRD families, suggesting diversity in the perturbations introduced.

Further information

GENT - Gene expression across normal and tumour tissue

http://medicalgenome.kribb.re.kr/

cBioPortal - access to large-scale cancer genomics data sets

http://www.cbioportal.org/

COSMIC - Catalogue of Somatic Mutations in Cancer

http://cancer.sanger.ac.uk

Human Proteome Map - access to peptide sequencing result from the draft map of the human proteome project

http://www.humanproteomemap.org/

AbMiner – antibody specificity database from the National Cancer Institute (NCI)

https://discover.nci.nih.gov/abminer/

Histone Antibodies – interactive histone antibody specificity database

http://www.histoneantibodies.com/

SGC Chemical Probes Portal – available chemical probes selectively targeting BRD modules

http://www.thesgc.org/chemical-probes/epigenetics

Supplementary information

Supplementary information S1 (Table): BRD-containing proteins and their function(s) in cell homeostasis

Supplementary information S2 (Figure): Structural phylogeny of BRD modules

Supplementary information S3 (Table): Peptide complexes of BRD modules

Supplementary information S4 (Table): BRD-containing proteins and their role(s) in cancer

Supplementary information S5 (Table): BRD-proteins and their mutations (COSMIC) in different Tissues

Supplementary information S6 (Figure): TCGA mutations mapped on the BRD structure of each subfamily

Supplementary information S7 (Table): Chemical Probes Targeting BRD modules

Supplementary information S8 (Table): Clinical Trials focusing on BRD-containing proteins

Copyright permissions

None

Glossary terms [Au: Please see my suggestions for additional Glossary terms]

YEATS domain: protein interaction domain with an immunoglobin-like fold that can bind to acetylated and crotonylated lysines

Transcriptional coregulators: Proteins that acts to repress or stimulate transcription via modulation of transcription factors and their activity through various mechanisms.

Histone code: The combination of histone modifications that affect transcription of genes in a hereditary fashion.

Epitopes: the part of a protein that can be recognized by a protein interaction domain. In the context of antibodies, the part of an antigen molecule to which an antibody attaches itself.

Epithelial-to-mesenchymal-transition (EMT): The process by which epithelial cells lose polarity and adhesion while gaining invasive and migratory properties acquire mesenchymal properties.

Epicardium: Connective tissue forming protective layer outside the heart muscle.

Centromeres: portions of the chromosome that link sister chromatids. Cohesion of the centromeres is achieved by associating with the protein cohesin which is subsequently modified in order to keep the two sister chromatids together.

γ-HOX genes: A group of (homeotic) related genes controlling the body plan of embryos along the head-axis tail.

Enhancer: short DNA region that provides a docking site for activating proteins, resulting in increased likelihood of transcription of particular genes.

RING (Really Interesting New Gene) domain: zinc-finger domain containing a Cys3HisCys4 motif which binds two zinc cations.

NUT- midline Carcinomas: aggressive, epithelial cancers that typically arise in organs in the midline of the body, and result from a fusion of the NUT gene together with BRD3, BRD4 or NSD3

Van der Waals radius: the radius of the sphere around an atom which represents the closest distance that another atom can approach.

Mediator complex: large multiprotein complex which acts as a transcriptional co-activator in eukaryotes, binding to the C-terminal domain of RNA polymerase II, bridging the polymerase together with transcription factors.

Author biographies

Takao Fujisawa is a Post-doctoral Associate at the Ludwig Institute for Cancer Research at Oxford University. He received his Ph.D. in Pharmaceutical Sciences under the supervision of Prof. Hidenori Ichijo at the University of Tokyo. He is interested in employing structural and molecular biology approaches to interrogate and interpret the molecular basis of epigenetic signalling.

Panagis Filippakopoulos is an Associate Professor at the Nuffield Department of Medicine at Oxford University and a principal investigator at the Structural Genomics Consortium and the Ludwig Institute for Cancer Research at Oxford. He received his Ph.D. in Inorganic Chemistry from the University of Michigan in Ann Arbor, USA, and carried out postdoctoral research in structural and chemical biology at the Structural Genomics Consortium at Oxford University. His research interests focus on structural comparisons of entire protein families and the discovery of shared and distinct mechanisms that determine substrate recognition and protein regulation, as well as the structure-guided design of specific inhibitors that modulate the function of proteins involved in epigenetic signalling, aiming to target them in cancer.
http://www.ndm.ox.ac.uk/principal-investigators/researcher/panagis-filippakopoulos

Key points

- Bromodomains (BRDs) are evolutionally conserved protein interaction modules. Structure-based alignments clustered human BRDs into eight distinct sub-families.
- BRD modules share a conserved helical bundle (helices αZ, αA, αB, αC) which is linked by diverse loop segments of variable length (ZA- and BC-loops).

- BRDs primarily recognize acetylated lysine residues on histones. Several BRDs are found to recognize acetylated non-histone proteins.
- BRD-containing proteins regulate gene expression, alone or as part of larger protein complexes, through chromatin remodeling, histone modification, histone recognition and transcriptional machinery regulation.
- BRD-containing proteins are frequently deregulated in cancer, and mutations in BRDs themselves are frequently identified. BRD-containing proteins are also identified in oncogenic rearrangement, leading to highly oncogenic fusion proteins.
- Development of BRD inhibitors as anticancer agents are being intensively investigated. Especially, small molecule inhibitors targeting BET bromodomain are now in clinical trials for treatment of various types of cancers.