Histone Lysine and Genomic Targets of Histone Acetyltransferases in Mammals

Anne K. Voss and Tim Thomas

Walter and Eliza Hall Institute of Medical Research, Parkville, Melbourne, Victoria, Australia

Department of Medical Biology, The University of Melbourne, Parkville, Victoria, Australia

Key words: CBP/P300, GCN5, PCAF, MOF, TIP60, HBO1, KAT6A

Correspondence can be addressed to

Anne K. Voss
The Walter and Eliza Hall Institute of Medical Research
1G Royal Parade | Parkville, Melbourne | Victoria 3052 | Australia
Email: avoss@wehi.edu.au

Tim Thomas
The Walter and Eliza Hall Institute of Medical Research
1G Royal Parade | Parkville, Melbourne | Victoria 3052 | Australia
Email: tthomas@wehi.edu.au

This is the author manuscript accepted for publication and has undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1002/bies.201800078

Abbreviations

Acetyl-CoA, acetyl-coenzyme A
AGO, argonaute protein
APP, Amyloid-beta precursor protein
AR, androgen receptor
ATPase, adenosine triphosphatase
ATM, ataxia telangiectasia mutated
CARM1, coactivator associated arginine methyltransferase 1, PRMT4
CBP, CREB binding protein, CREBBP, KAT3A
CREB, cyclic adenosine monophosphate response element binding protein
CRD1, cell cycle regulatory domain 1
ELK1, ETS-like transcription factor 1
EP300, E1A binding protein p300, P300, KAT3B
E6.5, embryonic day 6.5
DPF, double plant homeodomain finger
DSB, double-strand DNA breaks
GCN5L2, general control of amino acid synthesis protein 5-like 2, GCN5, KAT2A
GNAT, GCN5-related N-acetyltransferases
H3K14ac, histone H3 acetylated on lysine 14
H3K14bu, histone H3 butyrylated on lysine 14
H3K14cr, histone H3 crotonylated on lysine 14
H3K14pr, histone H3 propionylated on lysine 14
HAT, histone acetyltransferases
HBO1, histone acetyltransferase binding to ORC1, MYST2, KAT7
HDACs, histone deacetylases
HP-1, heterochromatin protein 1
HOX, homeobox protein
HTATIP, TIP60, HIV-1 Tat Interactive Protein, 60kDa, KAT5
ING, inhibitor of growth protein family
KAT, lysine acetyltransferases
KD, knockdown (in RNAi experiments)
MBTD1, MBT domain-containing protein 1
MEFs, mouse embryonic fibroblasts
MLL1, myeloid/lymphoid or mixed-lineage leukaemia protein 1, lysine methyltransferase 2A
MSL, male specific lethal
MRN, Mre11/Rad50/Nbs1 complex
MCRS2, microspherule protein 2
MOF, males absent on the first
MOZ, monocytic leukaemia zinc finger protein, KAT6A, MYST3
MORF, MOZ-related factor, MYST4, QKF, KAT6B
MYOD, myogenic differentiation 1 protein
MYST, MOZ, Ybf2/Sas3, Sas2 and Tip60-related protein family
NSL, non-specific lethal
ORC1, origin of replication complex protein 1
PCAF, P300/CBP-Associated Factor, KAT2B
PHD, plant homeodomain finger domain
PR, progesterone receptor
QKF, querkopf, MYST4, MORF, KAT6B
RNAi, RNA interference
RNA Pol II, RNA polymerase II
RUNX1, runt related transcription factor 1, AML1
SAGA, Spt-Ada-Gcn5-acetyltransferase complex
siRNA, short, interfering RNA
shRNA, short hairpin RNA
TBP, TATA-box binding protein
TF, transcription factor
TBX, T-box transcription factors

TFIIB, transcription factor II B, GTF2B
TSS, transcription start site

Abstract
Histone acetylation has been recognised as an important post-translational modification of core nucleosomal histones that changes access to the chromatin to allow gene transcription, DNA replication and repair. Histone acetyltransferases were initially identified as co-activators that link DNA-binding transcription factors to the general transcriptional machinery. Over the years, more chromatin-binding modes have been discovered suggesting direct interaction of histone acetyltransferases and their protein complex partners with histone proteins. While much progress has been made in characterising histone acetyltransferase complexes biochemically, cell-free activity assay results are often at odds with in-cell histone acetyltransferase activities. In-cell studies suggest specific histone lysine targets, but broad recruitment modes, apparently not relying on specific DNA sequences, but on chromatin of a specific functional state. Here we review the evidence for general versus specific roles of individual nuclear lysine acetyltransferases in light of *in vivo* and *in vitro* data in the mammalian system.

1. Introduction

Acetylation of histone lysine residues affects genome organisation and function. Generally speaking, histone acetylation correlates with an open chromatin structure and active gene transcription.[1, 2] It is required for DNA repair[3] and can promote or inhibit the activity of DNA replication origins.[4] Although there is a substantial body of literature describing histone acetylation under a variety of conditions, the data in mammalian cells remain largely correlative and the biochemical consequences of acetylation of individual lysines remains poorly understood; even the identity of the complexes laying down specific modifications in intact cells is often unclear. In this review we will focus on the effects of histone acetylation on gene transcription and only briefly return to DNA repair and replication in the section on histone acetyltransferase (HAT) recruitment mechanisms.

Histone lysine N-acetylation is catalysed by histone acetyltransferases (HATs) or more generally termed lysine acetyltransferases (KATs).[5] It is reversible by histone deacetylases

(HDACs; for HDAC reviews see[6, 7]). Histone acetylation neutralises the positive charge of lysine residues and thereby weakens the interaction between the histones, adjacent histones and the DNA (reviewed in[8]). Furthermore, acetylated histone lysines serve as docking sites for the recruitment of chromatin-binding proteins with chromatin modifying or remodelling functions (reviewed in[9]). Cytoplasmic HATs acetylate newly synthesised histones, which is required for integration of the new histones into nucleosomes.[10] Particularly diacetylation of histone H4 on lysines 5 and 12 (H4K5, H4K12) or equivalent residues is conserved between species including humans.[10] In contrast, acetylation of newly synthesised histone H3 does not appear to be conserved. Acetylation patterns of nucleosomal histones differ from non-nucleosomal histones and are affected by nuclear HATs and HDACs.

We will concentrate on nuclear HATs, for which structurally defined histone acetyltransferase domains have been well characterised and acetyl-coenzyme A (acetyl-CoA) binding has been confirmed. Based on these criteria, three families of nuclear HATs have been characterised. They are the GNAT family (GCN5-related N-acetyltransferases),[11, 12] the closely related pair of CBP and P300[13] and the MYST family (MOZ, Ybf2/Sas3, Sas2 and Tip60).[14] In mammals, the three families comprise nine nuclear HATs (Figure 1). The lysine acetyltransferase (KAT) nomenclature,[5] although widely used, has not become generally accepted. In this review we will use the most commonly accepted names as found in the literature. However, the KAT nomenclature emphasises the close relationship between three pairs of KATs: GCN5 and PCAF are KAT2A and KAT2B; CBP and P300 are KAT3A and KAT3B; KAT6A and KAT6B are MOZ (MYST3) and MORF (MYST4 or QKF). Because the KAT nomenclature emphasises the relationship for each of these three pairs, we will briefly return to the KAT nomenclature when we consider their overlapping function of the pairs. The remaining three nuclear HATs stand alone, TIP60 (HTATIP or KAT5), HBO1 (MYST2 or KAT7), MOF (MYST1 or KAT8), although TIP60 (KAT5) and MOF (KAT8) show similarity in protein domain structure (Figure 1).

GNAT and MYST HATs are found in large, multi-subunit protein complexes, while CBP and P300 interact with a large number of proteins, reviewed in.[15] Non-histone targets have been described for some nuclear HATs.[5] In this review, we will focus on histone targets.

We will examine the question of specificity with respect to (1) histone residue acetylation targets and (2) genomic targets. Specific genomic targets require direct or indirect DNA-sequence specific recruitment mechanisms, whereas recruitment for more general functions could be mediated by interaction with chromatin of a specific state, e.g. histones with specific post-translational modifications. Generally speaking, highly specific genomic targets would also be expected to result in specific, rather than pleiotropic, loss-of-function phenotypes. We will therefore consider, if the loss-of-function phenotypes of nuclear HATs are consistent with either specific genomic targets or a more general role in regulating transcription and/or chromatin organisation. Specificity for histone acetylation targets and genomic targets is of great importance, not only to understand the biological role of these proteins, but in particular if HATs are to be considered as drug targets.

2. HAT function: discovery is limited by technical approaches and conditions tested

In this section we will briefly give examples (1) of how specific experimental approaches can reveal one function, but not another function of nuclear HATs or (2) lead to incongruent conclusions, (3) that reagents can bias the experimental outcome, (4) and that nuclear HAT function can vary between cell types.

2.1. Identification *of HAT binding to posttranslationally modified histones limited to modifications tested*

The experimental outcome is dependent on, and limited by, the experimental design dictating that conclusions be restricted to the parameters tested. For example, the double PHD finger domain (DPF) of KAT6A was found to interact with the N-terminal tail of histone H3 only if histone H3 lysine 4 (H3K4) is unmodified.[16] KAT6A induces an alpha-helical conformation of H3K4-T11, revealing a unique mode of H3 recognition. The helical structure facilitates sampling of H3K4 methylation status, and offers H3K9 and other residues for modification.[16] H3K14 crotonylation (H3K14cr) was not assessed in this, but in a subsequent study, where it was observed that the DPF of KAT6A accommodates a wide range of histone lysine acylation modifications with the strongest preference for Kcr. Crystal structures of the DPF domain of KAT6A in complex with H3K14cr, H3K14bu, and H3K14pr peptides revealed a preference for H3K14cr binding in a hydrophobic pocket of the DPF.[17] Presumably, the two binding mechanisms operate mutually exclusively and one of the two

would provide the stronger candidate for in-cell function and, since mammalian cells are phenotypically distinct, these mechanisms maybe context dependent.

2.2. Effects of depletion of a HAT by shRNA or siRNA can be at odds with genetic deletion

Different experimental approaches to loss-of-function experiments have generated disparate results that may be difficult to reconcile. For example, CBP and P300 were found to be required for the bulk, if not all of H3K18ac and H3K27ac in double *Cpb;P300* null mouse embryonic fibroblasts (MEFs), while acetylation levels at H3K56 and other histone residues were found to be unaffected.[18] In contrast, CBP and P300 were found to acetylate H3K56 using siRNA knockdown in HeLa cells.[19] shRNA and siRNA knockdown of HBO1 in HEK293T cells suggested it was required for H4K5, H4K8 and H4K12 acetylation.[20] In contrast, MEFs that lack HBO1 entirely do not show deficiencies in H4K5, H4K8 or H4K12 acetylation levels, but instead display a 10-fold reduction in H3K14ac.[21] Comparing the two experimental approaches, one might be led to believe that the RNAi experiments revealed an acute response to HBO1 depletion, whereas germline deletion of the *Hbo1* gene allowed enough time for some degree of compensation with secondary effects. However, shRNA knockdown of HBO1 in erythroblasts resulted in a 5-fold reduction in H3K14ac and less pronounced or minor changes to H4 residues,[22] suggesting that even acute depletion of HBO1 affects H3K14ac in a major way.

2.3. Low antibody specificity can thwart the discovery of the histone lysine target

A source of discrepancies between studies can stem from the choice of reagents, for example antibodies marketed with incorrect histone lysine specificity.[23, 24] Antibodies commonly used to detect H3K14ac or H3K56ac turned out to recognise H3K9ac in addition. This has led to the proposal of histone lysine targets for specific HATs that were later contested. One possibility is to test the specificity of antibodies using modified histone peptides, another to use more than one antibody to confirm histone lysine targets.

2.4. Requirements for individual HATs can differ between cell types

In cases where HAT null cells survive and proliferate, one may be led to conclude that the specific HAT is neither essential for cell survival or proliferation. However, requirements in different cell types can vary and so the conclusion ought to be restricted to the cell type(s) under examination. For example, *Kat6a* null MEFs undergo cellular senescence and completely cease to proliferate.[25] Deletion of the *Kat6a* gene in adult mice causes a complete

loss of haematopoietic stem cells [26]. In contrast, haematopoietic progenitor cells continue to proliferate and form mature cell types of the blood sufficient for the survival of the mice for the rest of their lives.[26] The diverse effects of loss of KAT6A in the three different cell types shows that KAT6A is required to retain haematopoietic stem cells in the stem cell state, required to repress cellular senescence in fibroblasts and not essential for proliferation or differentiation in haematopoietic progenitor cells. Overall, the diverse requirements suggest that cell-type specific mechanisms, rather than general functions common to all cells, rely on the presence of KAT6A.

In conclusion, since the perceived histone acetylation targets and cellular functions can be affected strongly by the experimental approach taken and the reagents used, it is important to consider all experimental evidence to arrive at the likely histone acetylation target(s) and cellular functions of individual HATs.

3. Broad or histone lysine residue-specific HAT activity?

The histone acetylation targets attributed to individual HATs vary considerably in identity and breadth with assay conditions (Table 1). Whether a recombinant HAT domain or protein is used or an entire HAT complex is used and whether free histones or oligonucleosomes are used as substrates affects both, the extent and perceived target of acetylation. Results derived in cell-free assays often differ substantially from results collected within cells.

In cell-free assays, HBO1 acetylates histones H2A, H3 and H4 broadly, only H3 or only H4 depending on the assay conditions.[27-29] It was suggested that HBO1 can acetylate either H4 or H3K14 depending on the composition of the HBO1 protein complex partners.[27, 28] However as mentioned above, acute reduction or lack of HBO1 in a number of cell types causes a substantial reduction specifically in H3K14ac.[21, 22, 30-32] TIP60 can acetylate H2A-K5, H2A-K15, H3K14, H4K5, H4K8, H4K12, and H4K16 in cell-free assays,[33, 34] but appears to be restricted to H2A-K5 in cells.[35] Similarly, CBP is restricted to H3K18 and H3K27 acetylation in cells,[18] but can also acetylate H4K5 in cell-free assays.[36] It is noteworthy that cell free histone acetylation assays often return histones H3 and H4 (and even H2A and H1) as potential targets, whereas in-cell assays often suggest that the target residues of an individual HAT are restricted to just a single histone protein.

The discrepancy between potential acetylation targets determined in cell-free assays and in-cell targets suggests that other components of the HAT protein complexes and the nuclear environment restrict an otherwise promiscuous acetylation function. In addition, the specific histone lysine targets may also depend on the functional state of the cells and possibly the cell type. Lastly, a HAT can be essential for cellular processes with broad effects on cellular function and a multitude of downstream effects. For example, CBP or P300 are required for RNA Pol II recruitment.[18] Therefore, it is expected that the absence of CBP and P300 not only directly affects the acetylation events specifically catalysed by CBP or P300, but also indirectly causes significant secondary effects, potentially affecting other protein acetylation events.

In summary, based on genetic loss-of-function phenotypes of nuclear HATs, which cause genome-wide reductions or losses of acetylation, strong preferences for specific histone lysine residues seem to apply in mammals, some of which are confirmed in other taxonomy classes (Table 1). MOF appears to be specific for H4K16,[37-39] CBP and P300 for H3K18 and H3K27,[18] GCN5 and PCAF for H3K9,[18] and TIP60 for H2A-K5.[35] The major acetylation target of HBO1 appears to be H3K14.[21, 22, 30] To date, no strong genome-wide effects on histone acetylation were reported in the absence of KAT6A.[40] Instead, a reduction in H3K9ac at specific target loci was observed.[40, 41] This may suggest that KAT6A acts in a locus-specific rather than global manner, a notion supported by the normal development of many organ systems in the null foetuses.[40, 41] Alternatively, the reduction in H3K9ac may be secondary to a yet to be discovered genome-wide histone lysine target of KAT6A. RNAi experiments indicate that H3K23 is a major target of KAT6A.[42, 43] Assessment of H3K23 acetylation levels has not been reported for *Kat6a* null cells to date. No histone acetylation target has been attributed to KAT6B at this point.

4. Mammalian nuclear HATs can have additional, non-HAT functions

Proteins, in particular large proteins, commonly have more than one function. This also applies to at least a subset of nuclear HATs. Protein functions other than lysine acetyltransferase activity may complicate the interpretation of loss of function experiments.

A scaffold function has been attributed either to a HAT protein itself, e.g. P300,[44] or to HAT complex protein partners, e.g. the SAGA complex.[45] An enzymatic activity other than HAT activity can reside in a HAT complex, e.g. the SAGA complex contains ubiquitin protease activity in addition to its HAT activity.[46] These non-HAT functions can be affected by the loss of the HAT protein resulting in differences between complete loss of a HAT compared to loss of its histone acetyltransferase activity.

Comparison of the effects of null loss-of-function mutations to point mutations eliminating the HAT activity are useful to delineate the HAT function from potential other functions. This has been done for KAT6A and TIP60. Null mutation of *Kat6a* gene in mice causes a complete lack of definitive haematopoietic stem cells.[47, 48] In contrast, a point mutation removing the HAT activity of KAT6A results in a reduction in definitive haematopoietic stem cells.[49] Although the reduction in haematopoietic stem cell number and function is substantial, the data suggest that KAT6A may have roles in addition to lysine acetylation. Null mutation of the *Tip60* gene in mice causes embryonic lethality at the blastocyst stage.[50] A point mutation removing the HAT activity of TIP60 causes a severe delay in the onset of expression of mesoderm and endoderm genes, as well as severe growth restriction at embryonic day 6.5 (E6.5). While TIP60 HAT dead embryos undergo gastrulation and patterning to an equivalent of E9, they remain severely growth restricted.[51] The results suggest that the HAT activity is critical for embryo growth and timely expression of developmental genes, and that apart from these roles, TIP60 has another function that is critical for general cell survival. Since apart from TIP60 HAT activity, the TIP60 complex also possesses ATPase, DNA helicase and structural DNA binding activities,[52] one can speculate that one or more of these activities are required for general cell survival and may be dependent on the presence of TIP60, with or without HAT activity.

The examples of KAT6A and TIP60 HAT-dead mutant mice indicate that functions other than the acetylation activity need to be considered for nuclear HATs.

5. Gene-specific or genome-wide functions?

A major question has been if individual HATs affect the expression of specific sets of genes or transcription in a more general manner. Approaching this question, it is useful to consider

the effects of loss of function of specific HATs (Table 2). Loss of a gene encoding a protein essential for basic cell functions in mice results in very early embryonic lethality as soon as maternally-encoded protein stores become depleted, commonly in the blastocyst stage. This includes genes encoding proteins of the basal transcriptional machinery and non-redundant histone genes. Examples of essential members of the basal transcriptional machinery are TAF8[53] and TAF10[54] and of a unique histone gene is *H2afz* encoding H2A.Z.[55] Mutants of these essential genes display a phenotype that is similar to the consequences of the loss of an essential cell survival proteins, e.g. MCL1.[56]

Loss of function of TIP60[50] and MOF[38, 39] fall into this category causing early embryonic lethality in the late blastocyst stage. It can therefore be concluded that they are either essential for basic cell functions or specifically essential in blastocyst stage cells. Conditional deletion of the *Mof* gene has provided further information. Induction of *Mof* deletion in MEFs causes a complete arrest of cells growth.[57] Similarly, while differentiated podocytes were able to tolerate the absence of MOF, podocyte proliferation required MOF.[57] These data suggest that a non-proliferating cell may be able to maintain a steady state, but cells cannot undergo growth and population expansion without MOF.

In contrast to *Tip60* and *Mof* gene deletion, mutations of genes encoding the other HATs, KAT6A,[47, 48] KAT6B,[58, 59] HBO1,[21] GCN5,[60] PCAF,[61] CBP[62-64] and P300,[64] do not cause peri-implantation lethality. KAT6A null embryos,[40, 41, 47, 48] as well as a hypomorphic KAT6B mutant that expresses 10% normal mRNA[59] can develop to birth. HBO1 null embryos develop to E8.5 and die at E10.5,[21] GCN5 null embryos develop to E7.5 and die at E10.5,[60] CBP or P300 null embryos grow to a developmental stage equivalent to between E8.5 and E10.5.[62-64] Mice lacking PCAF are viable.[61]

Null mutations that allow embryo survival to E7.5 or beyond can be interpreted as not strictly essential for basic cell functions such as cell metabolism, cell survival, DNA synthesis and cell proliferation, although any of these processes may be impaired. As a corollary, transcription of genes required for many cellular processes has to proceed at least at a sufficient level in the cells lacking KAT6A, KAT6B, HBO1, GCN5, CBP, P300 or PCAF to

allow development to E7.5 or beyond. It is noteworthy, that apart from PCAF, KAT6A and KAT6B, the loss of the other nuclear HATs appears to affect a broad range of tissues.

In summary, it appears that of the nuclear HATs, either based on their HAT activity or based on other functions, MOF and TIP60 are essential for fundamental processes in mouse preimplantation embryos. HBO1, GCN5, CBP and P300 are required in multiple cell types during embryonic patterning and organogenesis. KAT6A and KAT6B are needed for specific embryonic patterning processes and in specific tissues during foetal development. PCAF is dispensable for prenatal development.

6. Pairs of closely related HATs have shared and unique functions

Duplication of the entire genome and of individual genes is thought to be the decisive mechanism allowing functional diversification and the development of eukaryotic complexity,[65] as well as mutational robustness.[66]

The high degree of amino acid sequence identity and similarity between four pairs of HATs, GCN5 and PCAF (KAT2A and KAT2B), CBP and P300 (KAT3A and KAT3B), KAT6A and KAT6B (MYST3 and MYST4 or MOZ and MORF or MOZ and QKF) and TIP60 (KAT5) and MOF (KAT8) (Figure 1), suggests that each pair has arisen from a single ancestor in one of the genome duplication events during evolution. The fact that single null mutations of all nuclear HAT genes except KAT2B/PCAF cause embryonic or perinatal lethality, suggests that mutational robustness was not the prime driver in retaining the duplications. By contrast, analysis of pairwise double null mice suggests that partial functional overlap accompanied by some functional diversification occurred.

The *Gcn5/Pcaf (Kat2a/b*) double null mice die at E7.5, but are already abnormal at E6.5, at least one day earlier than the *Gcn5* single null embryos arrest in development.[60] Until E6.5, the embryo undergoes relatively little cellular diversification, but greatly increases in cell number. The key event of embryogenesis commencing during this time is gastrulation. This process results in the formation of the three germ layers and concurrently the basic body plan is laid down. It requires precise regulation of developmental patterning gene expression that shapes all subsequent development. The relatively early death of *Gcn5/Pcaf* double null

embryos suggests that "KAT2" function is required for fundamental developmental steps, perhaps including the initiation of novel gene expression programs. In contrast, *Gcn5/Pcaf* double null MEFs are viable and grow at a near normal rate,[18] indicating that cell survival and proliferation does not require "KAT2" function in general.

The double *heterozygous* loss of CBP and P300 (KAT3A and KAT3B) causes embryonic lethality. The combined *homozygous* loss of CBP and P300 has not been described. However, tissue-specific combined loss of CBP and P300 in the B-cell lineage[67] indicates that it is not compatible with cell survival, suggesting that "KAT3" function may be indispensable for basic cell function. Indeed, while *Gcn5/Pcaf* double null MEFs grow nearly normally and display normal morphology, *Cbp/P300* double null MEFs cease to grow.[18] These results suggest that "KAT3", but not "KAT2" function, is required for basal cell function, at least in MEFs.

While *Kat6a* (*Moz)* null[40, 41, 47, 48] and *Kat6b (Qkf*) mutant pups[59] can develop to birth, *Kat6a/Kat6b* double null mice develop to E9.0 [Voss & Thomas, unpublished], indicating some functional overlap during prenatal development. Development to E9.0 of *Kat6a/Kat6b* double null suggests that "KAT6" function is not essential for cell survival and cell proliferation. It follows that, although "KAT6" function may be required for gene expression under certain circumstances, gene expression required for cell metabolism, cell survival and cell proliferation proceeds at least adequately without "KAT6".

Although the protein domain structure of TIP60 (KAT5) and MOF (KAT8) indicates that they too arose from a common ancestor gene, they are individually essential for cell survival in the mouse inner cell mass,[38, 39, 50] precluding an analysis of additional potential overlapping functions in germline null loss-of-function mice.

In conclusion, like single null mutant mice of MOF and TIP60, *Cbp/P300* double null mutant mice indicate essential roles in fundamental cell functions such as cell survival, cell proliferation and/or cell metabolism. In contrast, similar to the single *Hbo1* mutant mice, double mutant mice of *Gcn5/Pcaf* and *Kat6a/Kat6b* suggest a more restricted, albeit broad function of these pairs of closely related nuclear HATs.

7. HATs – always present and active, specifically activated or specifically recruited?

In this section we will discuss the in-cell evidence for the genomic distribution of HATs, i.e. genome-wide or at chromatin of a specific state (such as active gene loci) vs. recruited to specific gene loci, as well as activation *in situ* on the chromatin. HATs that are required for acetylation of histone lysines genome-wide would require little or no DNA sequence-specific targeting, whereas HATs that only affect the expression of a restricted subset of genes may be specifically recruited to their target gene loci (Table 3).

7.1. Broad requirement for MOF and H4K16ac

Loss of MOF in mice results in a pronounced loss of H4K16 acetylation,[38, 39] but does not affect acetylation levels at other histone lysine residues.[39] Likewise, H4K16ac in MEFs and mouse podocytes,[57] mouse ES cells,[68] mouse haematopoietic cells[69] and in human cells[37] depends on the presence of MOF. MOF was found to be essential for proliferating MEFs, but dispensable in terminally differentiated podocytes during steady state. However, to exit steady state, e.g. to mount a stress response podocytes required MOF, too.[57] Conditional deletion of the *Mof* gene in the haematopoietic system indicates that MOF is critical for the maintenance of adult haematopoietic cells.[69] In the same study, *Mof* mRNA was reduced to 50% of control at E14.5,[69] a possible reason why foetal haematopoiesis still occurred in this model. MOF occupies the regions of the transcription start site of actively transcribed genes in mouse embryonic stem cells and neural progenitor cells [68]. MOF is widely distributed in the genome at active gene loci and appears to be recruited for their activation.

7.2. Broad requirement for HBO1 and H3K14ac

HBO1 occurs in nuclear histone acetyltransferase complexes containing ING4 or ING5, BRPF1,2 or 3 or JADE1, 2 or 3 and EAF6.[6, 20, 22, 27] Within these complexes, HBO1 has been proposed to acetylate H3K14 in the region of the TSS of transcriptionally active genes and H4 throughout the gene depending on the presence or absence of ING4 or ING5.[28] Similar to the function of MOF in acetylating H4K16, HBO1 is required for more than 90% of all H3K14ac.[21] H3K14ac is enriched at transcriptionally active and poised gene loci.[70] Congruent with the distribution of H3K14ac, HBO1 occupies gene loci and its level of occupancy correlates positively with the level of gene expression.[71]

In addition to the HBO1 complex described above, a small amount of the total HBO1 protein (~5-10%) immuno-precipitates with the origin of replication complex.[29, 72] This finding and RNAi interference experiments have led to the conclusion that HBO1 is essential for DNA replication[73] and cell proliferation.[20, 71] However, genetic loss of HBO1 has no effects on DNA synthesis in mouse embryos or proliferation of MEFs, but has major effects on other vital cell functions, such as gene activation and cell survival.[21] Genetic mutation or knockdown of HBO1 has only mild effects on DNA replication and cell proliferation in Drosophila,[74] indicating the HBO1 may be beneficial, but is not essential for DNA synthesis in this organism. If cell survival is not specifically assessed, a loss of cells to cell death can be mistaken for a reduction in cell proliferation. Indeed, while the loss of the HBO1 complex protein BRD1 led to reduced proliferation of erythoblasts, it also caused an increase in cell death and a reduction in *Gata1* gene expression (among other genes).[22] Interestingly, the proliferation and survival defects were efficiently rescued by restoration of *Gata1* gene expression, suggesting that the reduction in *Gata1* gene expression mainly accounted for the impaired proliferation and survival of *Brd1*$^{-/-}$ erythroblasts.[22] Nevertheless, it has been proposed that that acetylation of H3K14 by HBO1 in the context of a BRPF3-containing complex around transcription start sites enables efficient activation of nearby replication origins.[30] Overall, there is strong *in vivo* evidence for a role of HBO1 in facilitating gene transcription and cell survival (possibly via gene transcription). Dependent on the cell type, loss of HBO1 does or does not lead to a reduction in cell growth, which may be a secondary effect of changes in gene transcription and cell survival.

7.3. Specific requirement for KAT6A for homeobox protein encoding gene transcription

In contrast to MOF and HBO1, genetic deletion of the *Kat6a* gene has no effect on genome-wide levels of histone acetylation levels tested thus far. However, KAT6A is essential for normal expression of *Hox* and *Tbx* genes and H3K9ac specifically at these gene loci,[40, 41] posing the question as to how KAT6A is specifically recruited to *Hox* and *Tbx* genes, which still remains to be determined. Whether the locus-specific reduction in H3K9ac is a direct or an indirect effect of loss of KAT6A is difficult to establish *in vivo*.

7.4. HATs regulated by posttranslational modification while already present on chromatin

Protein interaction partners have been reported to regulate the enzymatic activity of individual HATs. This regulation can take place on the chromatin and adds a layer of regulation, in

addition to active recruitment of the HAT (Table 4). In this way, sumoylation of P300 in its CRD1 transcriptional repression domain has been reported to inhibit its acetylation function[75] and pre-assembled ELK1-P300 complexes become activated following ELK1 phosphorylation by changes in ELK1-P300 interactions.[76] Similarly, CARM1-dependent CBP methylation increases HAT activity of CBP[77] and phosphorylation of HBO1 by ATM/ATR during the DNA damage response has been noted to be required for efficient acetylation of H3K14 and H4 rendering the chromatin more accessible to nucleotide excision repair factors.[78] As a corollary of regulation of HAT activity *in situ* on the chromatin, HAT proteins can remain on the chromatin after gene expression has terminated. For example, P300 and Pol II bookmark T cell immediate early gene promoters long after transcription has terminated.[79] Similarly, the H3K27ac mark appears to be maintained during mitosis and to be important for mitotic bookmarking of pluripotency genes in embryonic stem cells.[80]

8. How are HATs recruited to chromatin?

A number of mechanisms mediating recruitment to chromatin have been proposed for most of the nuclear HATs (Figure 2). Different mechanisms reported for individual HATs would be expected to result in different outcomes. Binding to unmodified histones offers a broad genome-wide recruitment mechanism. Binding to specifically modified histones or to members of the basal transcriptional machinery presents a more restricted mechanism, e.g. to transcriptionally poised or active genes. Binding to DNA-sequence-specific transcription factors would provide a mechanism potentially specific to a small subset of genes.

The recruitment of HATs by specific transcription factors to target genes has been observed (Table 4). For example, CBP[81] and P300[82] are recruited by CREB, GCN5 and PCAF by MYOD[83] and KAT6A and KAT6B by RUNX1 and RUNX2.[84] Specific domains within HAT proteins mediate interaction with DNA-binding transcription factors (Figure 1). These include the CREB interaction domain of CBP,[81] or its TAZ1 domain, which interacts with HIF1a.[85, 86] Similarly, the serine- and methionine-rich domains of KAT6A and KAT6B are able to interact with runt family transcription factors.[84] In addition, indirect binding to nuclear receptors via adaptor proteins has been observed, for example for CBP and P300[87] and for TIP60.[35]

A broader recruitment mechanism involving binding to components of the basal transcriptional machinery has been reported for CBP binding to TFIIB[88] and for GCN5, which binds indirectly to TBP via SAGA complex members, at least in yeast.[89] Other broader recruitment mechanisms involve direct or indirect binding to histone residues. HAT complexes typically contain several proteins with histone binding domains. These include bromodomains, which can bind acetylated lysines, and chromodomains and PHD domains, which can bind methylated lysines (for review see[9]). Some HAT proteins contain histone-binding domains (Figure 1), KAT6A and KAT6B both have a double PHD finger domain, through which they can bind directly to modified H3K14.[16, 17, 90, 91] The recruitment of MOF in alternative protein complexes have been reported. They involve indirect binding to histones via MLL, MSL1 or MCRS2. Binding to chromatin via MLL1 was observed in a trithorax transcriptional co-activator complex.[92] Binding to histones via MSL1 occurs in the male specific lethal (MSL) complex[93] and via MCRS2 in the non-specific lethal (NSL) complex.[94] Direct and indirect binding to modified histone residues have been observed for TIP60, binding to H3K9me3 via its chromodomain[95] or binding to H4K20me1/2 via MBTD1[34] (Table 4). In addition, recruitment of TIP60 to sites of DNA double strand breaks can be mediated by binding to both H3K9me3 and ATM.[95-97] Binding of TIP60 to AGO2 stimulated by sequence-specific non-coding RNAs could provide another mechanism.[98]

The reported mechanisms of HAT recruitment are necessarily affected by the chosen experimental approach and expectation bias. As data accumulated, the interpretations of the recruitment mechanisms appear to have shifted in some cases. While some work focussed on locus specific interaction of CBP and P300 with DNA-binding transcription factors (examples in Table 4), other work suggests more general occupancy of gene loci (examples in Table 4) and enhancers by CBP and P300.[36, 99]

While the broader binding mechanisms would imply a much less specific and more widespread association with activated or poised gene loci, they could be complemented by additional locus-specific binding.

9. The genomic distribution of individual HATs supports chromatin mediated recruitment

The physical genomic distribution of individual HATs (Table 3) may provide supporting evidence in favour of a subset of the proposed mechanisms for recruitment of HATs to the chromatin (Table 4).

The genome occupancy (Table 3) of CBP, P300, GCN5, PCAF, TIP60 and MOF individually was reported to correlate highly with RNA Pol II occupancy.[2] A high correlation between GCN5 and RNA Pol II occupancy was also observed in another study.[100] Occupancy in the region of the transcription start site of transcriptionally active gene loci was reported by others for MOF in *Drosphila*[94] and for HBO1.[27, 28, 71] In addition, CBP and P300 were found to be present at enhancers,[36, 99] where they are thought to acetylate H3K27.[101, 102] These patterns of occupancy suggest a recruitment mechanism that applies to most transcriptionally active gene loci, rather than DNA sequence and locus-specific recruitment. Some of the proposed recruitment mechanisms fit this concept (Table 4). For example, CBP has been shown to interact with basal transcription factor TFIIB.[88] GCN5 through the SAGA complex proteins Spt8 and Ada1 can bind TBP[89] or H3A1 and H3K4me2 and H3K4me3 through the tandem tudor domains of Sgf29,[103] TIP60 can bind H3K9me1 or H3K9me3 directly via its chromodomain.[35, 95, 96] MOF via MSL complex proteins MSL1 or the NSL complex protein MCRS2 can bind to histones.[93, 94] HBO1, as well as KAT6A and KAT6B, can bind H3K4me2/3 through their interaction with ING4 and/or ING5.[28, 71] Lastly, KAT6A and KAT6B can bind acetylated, crotonylated and butyrylated H3K14 directly through their double PHD fingers.[16, 17, 90, 91]

These examples suggest a more general recruitment of nuclear HATs, chiefly to sites prepared for transcription by chromatin modifying enzymes, including HATs and general transcription factors. However, other results suggest more locus-specificity, such as reported interaction with specific transcription factors, as discussed in the section above and summarised in Table 4. The two modes of recruitment may not be mutually exclusive, even within one target gene locus. There may be a general recruitment of nuclear HATs to sites of active genes or genes that are primed for activation. In addition, specific interaction with DNA-binding transcription factors may boost the recruitment of HATs to increase the level of gene transcription. In this manner, basal gene expression levels could be facilitated and also amplified, as needed.

Open questions and tasks that remain:

- Discrepancies of results between different experimental approaches (cell-free assays, RNAi knock down, null deletions and point mutations) suggest it would be useful to generate mutants that specifically address the different functions of the HAT proteins (HAT activity, scaffold function and possibly more).
- HAT mutants that specifically address different functions of the HAT proteins need to be assessed in different cell types and also in different functional cell states.
- Consequences to consider include, but are not restricted to, the histone acetylation levels, the acetylation of other proteins, acylation modification other than acetylation, as well as effects on recruitment of other chromatin-associated proteins.

10. Conclusions and outlook

The following conclusions based on the in-cell phenotypes of specific HAT gene mutations of HAT may be considered:

- Cell-free biochemical assays display a potential histone lysine acetylation/acylation spectrum, but do not predict the in-cell HAT target(s).
- The reported distribution of the nuclear HATs in the mammalian genome supports recruitment processes that do not solely rely on specific DNA sequence elements, perhaps in addition to DNA-binding transcription factor specific recruitment.
- The combined null deletions of the genes encoding the pairs of CBP/p300, GCN5/PCAF and KAT6A/KAT6B suggest evolutionary functional diversification, but also overlapping roles.
- Different cell types and functional states may require different complements of nuclear HATs.
- Nuclear HATs can have functions in addition to lysine acetylation so that a HAT null phenotype and a HAT-dead mutant do not necessarily display the same phenotype.

Precision of the conclusions ceases to be an academic problem and gains significantly in importance, as the research community becomes increasingly interested in generating small molecule inhibitors specific to individual HATs (or pairs of HATs) as lead compounds for

drug development.[104, 105] Here it becomes relevant if inhibition of the HAT activity confers the desired therapeutic outcome or if instead a protein-protein interaction surface ought to be targeted to interfere with a scaffold function. Conversely, the adverse side effects observed in the complete absence of a HAT protein may not materialise if only the HAT activity is blocked. Therefore the risk that effective drug targets are excluded from consideration can be avoided by discriminating between the HAT function and the totality of all functions of the intact complex. That inhibition of specific HATs can result in tumour cell growth arrest via cellular senescence without DNA damage[104] suggests that a new class of cancer therapeutics may be within reach.

Acknowledgements

We would like to thank Dr Tanya de Jong-Curtain for outstanding work proofreading. The authors and their work were supported by the National Health and Medical Research Council through project grants (1030704, 1084248, 1084352 to AKV and 1084504, 1084509, 1080146 to TT) and fellowships (1081421, 575512 to AKV and 1003435 to TT). This work was made possible by operational infrastructure grants through the Australian Government IRISS and the Victorian State Government OIS.

References

[1] T. J. Stasevich, Y. Hayashi-Takanaka, Y. Sato, K. Maehara, Y. Ohkawa, K. Sakata-Sogawa, M. Tokunaga, T. Nagase, N. Nozaki, J. G. McNally, H. Kimura, *Nature* **2014**, 516, 272.

[2] Z. Wang, C. Zang, K. Cui, D. E. Schones, A. Barski, W. Peng, K. Zhao, *Cell* **2009**, 138, 1019.

[3] R. Murr, J. I. Loizou, Y. G. Yang, C. Cuenin, H. Li, Z. Q. Wang, Z. Herceg, *Nat Cell Biol* **2006**, 8, 91.

[4] P. Norio, *EMBO Rep* **2006**, 7, 779.

[5] C. D. Allis, S. L. Berger, J. Cote, S. Dent, T. Jenuwien, T. Kouzarides, L. Pillus, D. Reinberg, Y. Shi, R. Shiekhattar, A. Shilatifard, J. Workman, Y. Zhang, *Cell* **2007**, 131, 633.

[6] C. Choudhary, B. T. Weinert, Y. Nishida, E. Verdin, M. Mann, *Nat Rev Mol Cell Biol* **2014**, 15, 536.

[7] M. Haberland, R. L. Montgomery, E. N. Olson, *Nat Rev Genet* **2009**, 10, 32.

[8] G. E. Zentner, S. Henikoff, *Nat Struct Mol Biol* **2013**, 20, 259.

[9] M. Yun, J. Wu, J. L. Workman, B. Li, *Cell Res* **2011**, 21, 564.

[10] R. E. Sobel, R. G. Cook, C. A. Perry, A. T. Annunziato, C. D. Allis, *Proc Natl Acad Sci U S A* **1995**, 92, 1237.

[11] A. Clements, J. R. Rojas, R. C. Trievel, L. Wang, S. L. Berger, R. Marmorstein, *Embo j* **1999**, 18, 3521.

[12] Y. Lin, C. M. Fletcher, J. Zhou, C. D. Allis, G. Wagner, *Nature* **1999**, 400, 86.

[13] M. Delvecchio, J. Gaucher, C. Aguilar-Gurrieri, E. Ortega, D. Panne, *Nat Struct Mol Biol* **2013**, 20, 1040.

[14] Y. Yan, N. A. Barlev, R. H. Haley, S. L. Berger, R. Marmorstein, *Mol Cell* **2000**, 6, 1195.

[15] K. K. Lee, J. L. Workman, *Nat Rev Mol Cell Biol* **2007**, 8, 284.

[16] I. Dreveny, S. E. Deeves, J. Fulton, B. Yue, M. Messmer, A. Bhattacharya, H. M. Collins, D. M. Heery, *Nucleic Acids Res* **2014**, 42, 822.

[17] X. Xiong, T. Panchenko, S. Yang, S. Zhao, P. Yan, W. Zhang, W. Xie, Y. Li, Y. Zhao, C. D. Allis, H. Li, *Nat Chem Biol* **2016**, 12, 1111.

[18] Q. Jin, L. R. Yu, L. Wang, Z. Zhang, L. H. Kasper, J. E. Lee, C. Wang, P. K. Brindle, S. Y. Dent, K. Ge, *Embo j* **2011**, 30, 249.

[19] C. Das, M. S. Lucia, K. C. Hansen, J. K. Tyler, *Nature* **2009**, 459, 113.

[20] Y. Doyon, C. Cayrou, M. Ullah, A. J. Landry, V. Cote, W. Selleck, W. S. Lane, S. Tan, X. J. Yang, J. Cote, *Mol Cell* **2006**, 21, 51.

[21] A. J. Kueh, M. P. Dixon, A. K. Voss, T. Thomas, *Mol Cell Biol* **2011**, 31, 845.

[22] Y. Mishima, S. Miyagi, A. Saraya, M. Negishi, M. Endoh, T. A. Endo, T. Toyoda, J. Shinga, T. Katsumoto, T. Chiba, N. Yamaguchi, I. Kitabayashi, H. Koseki, A. Iwama, *Blood* **2011**, 118, 2443.

[23] T. A. Egelhofer, A. Minoda, S. Klugman, K. Lee, P. Kolasinska-Zwierz, A. A. Alekseyenko, M. S. Cheung, D. S. Day, S. Gadel, A. A. Gorchakov, T. Gu, P. V. Kharchenko, S. Kuan, I. Latorre, D. Linder-Basso, Y. Luu, Q. Ngo, M. Perry, A. Rechtsteiner, N. C. Riddle, Y. B. Schwartz, G. A. Shanower, A. Vielle, J. Ahringer, S. C. Elgin, M. I. Kuroda, V. Pirrotta, B. Ren, S. Strome, P. J. Park, G. H. Karpen, R. D. Hawkins, J. D. Lieb, *Nat Struct Mol Biol* **2011**, 18, 91.

[24] S. B. Rothbart, B. M. Dickson, J. R. Raab, A. T. Grzybowski, K. Krajewski, A. H. Guo, E. K. Shanle, S. Z. Josefowicz, S. M. Fuchs, C. D. Allis, T. R. Magnuson, A. J. Ruthenburg, B. D. Strahl, *Mol Cell* **2015**, 59, 502.

[25] B. N. Sheikh, B. Phipson, F. El-Saafin, H. K. Vanyai, N. L. Downer, M. J. Bird, A. J. Kueh, R. E. May, G. K. Smyth, A. K. Voss, T. Thomas, *Oncogene* **2015**, 34, 5807.

[26] B. N. Sheikh, Y. Yang, J. Schreuder, S. K. Nilsson, R. Bilardi, S. Carotta, H. M. McRae, D. Metcalf, A. K. Voss, T. Thomas, *Blood* **2016**, 128, 2307.

[27] M. E. Lalonde, N. Avvakumov, K. C. Glass, F. H. Joncas, N. Saksouk, M. Holliday, E. Paquet, K. Yan, Q. Tong, B. J. Klein, S. Tan, X. J. Yang, T. G. Kutateladze, J. Cote, *Genes Dev* **2013**, 27, 2009.

[28] N. Saksouk, N. Avvakumov, K. S. Champagne, T. Hung, Y. Doyon, C. Cayrou, E. Paquet, M. Ullah, A. J. Landry, V. Cote, X. J. Yang, O. Gozani, T. G. Kutateladze, J. Cote, *Mol Cell* **2009**, 33, 257.
[29] M. Iizuka, B. Stillman, *J Biol Chem* **1999**, 274, 23027.
[30] Y. Feng, A. Vlassis, C. Roques, M. E. Lalonde, C. Gonzalez-Aguilera, J. P. Lambert, S. B. Lee, X. Zhao, C. Alabert, J. V. Johansen, E. Paquet, X. J. Yang, A. C. Gingras, J. Cote, A. Groth, *Embo j* **2016**, 35, 176.
[31] R. Matsunuma, H. Niida, T. Ohhata, K. Kitagawa, S. Sakai, C. Uchida, B. Shiotani, M. Matsumoto, K. I. Nakayama, H. Ogura, N. Shiiya, M. Kitagawa, *Mol Cell Biol* **2016**, 36, 394.
[32] C. Zou, Y. Chen, R. M. Smith, C. Snavely, J. Li, T. A. Coon, B. B. Chen, Y. Zhao, R. K. Mallampalli, *J Biol Chem* **2013**, 288, 6306.
[33] A. Kimura, M. Horikoshi, *Genes Cells* **1998**, 3, 789.
[34] K. Jacquet, A. Fradet-Turcotte, N. Avvakumov, J. P. Lambert, C. Roques, R. K. Pandita, E. Paquet, P. Herst, A. C. Gingras, T. K. Pandita, G. Legube, Y. Doyon, D. Durocher, J. Cote, *Mol Cell* **2016**, 62, 409.
[35] K. W. Jeong, K. Kim, A. J. Situ, T. S. Ulmer, W. An, M. R. Stallcup, *Nat Struct Mol Biol* **2011**, 18, 1358.
[36] D. A. Bose, G. Donahue, D. Reinberg, R. Shiekhattar, R. Bonasio, S. L. Berger, *Cell* **2017**, 168, 135.
[37] M. Taipale, S. Rea, K. Richter, A. Vilar, P. Lichter, A. Imhof, A. Akhtar, *Mol Cell Biol* **2005**, 25, 6798.
[38] A. Gupta, T. G. Guerin-Peyrou, G. G. Sharma, C. Park, M. Agarwal, R. K. Ganju, S. Pandita, K. Choi, S. Sukumar, R. K. Pandita, T. Ludwig, T. K. Pandita, *Mol Cell Biol* **2008**, 28, 397.
[39] T. Thomas, M. P. Dixon, A. J. Kueh, A. K. Voss, *Mol Cell Biol* **2008**, 28, 5093.
[40] A. K. Voss, C. Collin, M. P. Dixon, T. Thomas, *Dev Cell* **2009**, 17, 674.
[41] A. K. Voss, H. K. Vanyai, C. Collin, M. P. Dixon, T. J. McLennan, B. N. Sheikh, P. Scambler, T. Thomas, *Dev Cell* **2012**, 23, 652.
[42] L. Simo-Riudalbas, M. Perez-Salvia, F. Setien, A. Villanueva, C. Moutinho, A. Martinez-Cardus, S. Moran, M. Berdasco, A. Gomez, E. Vidal, M. Soler, H. Heyn, A. Vaquero, C. de la Torre, S. Barcelo-Batllori, A. Vidal, L. Roz, U. Pastorino, K. Szakszon, G. Borck, C. S. Moura, F. Carneiro, I. Zondervan, S. Savola, R. Iwakawa, T. Kohno, J. Yokota, M. Esteller, *Cancer Res* **2015**, 75, 3936.
[43] D. Lv, F. Jia, Y. Hou, Y. Sang, A. A. Alvarez, W. Zhang, W. Q. Gao, B. Hu, S. Y. Cheng, J. Ge, Y. Li, H. Feng, *Cancer Res* **2017**, 77, 6190.
[44] C. Mertens, R. G. Roeder, *Mol Cell Biol* **2008**, 28, 5764.
[45] A. M. Dudley, C. Rougeulle, F. Winston, *Genes Dev* **1999**, 13, 2940.
[46] V. M. Weake, J. L. Workman, *Trends Cell Biol* **2012**, 22, 177.
[47] T. Thomas, L. M. Corcoran, R. Gugasyan, M. P. Dixon, T. Brodnicki, S. L. Nutt, D. Metcalf, A. K. Voss, *Genes Dev* **2006**, 20, 1175.
[48] T. Katsumoto, Y. Aikawa, A. Iwama, S. Ueda, H. Ichikawa, T. Ochiya, I. Kitabayashi, *Genes Dev* **2006**, 20, 1321.
[49] F. M. Perez-Campo, J. Borrow, V. Kouskoff, G. Lacaud, *Blood* **2009**, 113, 4866.
[50] Y. Hu, J. B. Fisher, S. Koprowski, D. McAllister, M. S. Kim, J. Lough, *Dev Dyn* **2009**, 238, 2912.
[51] D. Acharya, S. J. Hainer, Y. Yoon, F. Wang, I. Bach, J. A. Rivera-Perez, T. G. Fazzio, *Cell Rep* **2017**, 19, 671.
[52] T. Ikura, V. V. Ogryzko, M. Grigoriev, R. Groisman, J. Wang, M. Horikoshi, R. Scully, J. Qin, Y. Nakatani, *Cell* **2000**, 102, 463.
[53] A. K. Voss, T. Thomas, P. Petrou, K. Anastassiadis, H. Scholer, P. Gruss, *Development* **2000**, 127, 5449.
[54] W. S. Mohan, Jr., E. Scheer, O. Wendling, D. Metzger, L. Tora, *Mol Cell Biol* **2003**, 23, 4307.

[55] R. Faast, V. Thonglairoam, T. C. Schulz, J. Beall, J. R. Wells, H. Taylor, K. Matthaei, P. D. Rathjen, D. J. Tremethick, I. Lyons, *Curr Biol* **2001**, 11, 1183.
[56] J. L. Rinkenberger, S. Horning, B. Klocke, K. Roth, S. J. Korsmeyer, *Genes Dev* **2000**, 14, 23.
[57] B. N. Sheikh, W. Bechtel-Walz, J. Lucci, O. Karpiuk, I. Hild, B. Hartleben, J. Vornweg, M. Helmstadter, A. H. Sahyoun, V. Bhardwaj, T. Stehle, S. Diehl, O. Kretz, A. K. Voss, T. Thomas, T. Manke, T. B. Huber, A. Akhtar, *Oncogene* **2016**, 35, 2698.
[58] T. D. Merson, M. P. Dixon, C. Collin, R. L. Rietze, P. F. Bartlett, T. Thomas, A. K. Voss, *J Neurosci* **2006**, 26, 11359.
[59] T. Thomas, A. K. Voss, K. Chowdhury, P. Gruss, *Development* **2000**, 127, 2537.
[60] W. Xu, D. G. Edmondson, Y. A. Evrard, M. Wakamiya, R. R. Behringer, S. Y. Roth, *Nat Genet* **2000**, 26, 229.
[61] T. Yamauchi, J. Yamauchi, T. Kuwata, T. Tamura, T. Yamashita, N. Bae, H. Westphal, K. Ozato, Y. Nakatani, *Proc Natl Acad Sci U S A* **2000**, 97, 11303.
[62] Y. Tanaka, I. Naruse, T. Maekawa, H. Masuya, T. Shiroishi, S. Ishii, *Proc Natl Acad Sci U S A* **1997**, 94, 10215.
[63] Y. Oike, N. Takakura, A. Hata, T. Kaname, M. Akizuki, Y. Yamaguchi, H. Yasue, K. Araki, K. Yamamura, T. Suda, *Blood* **1999**, 93, 2771.
[64] T. P. Yao, S. P. Oh, M. Fuchs, N. D. Zhou, L. E. Ch'ng, D. Newsome, R. T. Bronson, E. Li, D. M. Livingston, R. Eckner, *Cell* **1998**, 93, 361.
[65] L. Aravind, L. M. Iyer, E. V. Koonin, *Curr Opin Struct Biol* **2006**, 16, 409.
[66] Y. Van de Peer, E. Mizrachi, K. Marchal, *Nat Rev Genet* **2017**, 18, 411.
[67] W. Xu, T. Fukuyama, P. A. Ney, D. Wang, J. Rehg, K. Boyd, J. M. van Deursen, P. K. Brindle, *Blood* **2006**, 107, 4407.
[68] T. Chelmicki, F. Dundar, M. J. Turley, T. Khanam, T. Aktas, F. Ramirez, A. V. Gendrel, P. R. Wright, P. Videm, R. Backofen, E. Heard, T. Manke, A. Akhtar, *Elife* **2014**, 3, e02024.
[69] D. G. Valerio, H. Xu, M. E. Eisold, C. M. Woolthuis, T. K. Pandita, S. A. Armstrong, *Blood* **2017**, 129, 48.
[70] K. Karmodiya, A. R. Krebs, M. Oulad-Abdelghani, H. Kimura, L. Tora, *BMC Genomics* **2012**, 13, 424.
[71] N. Avvakumov, M. E. Lalonde, N. Saksouk, E. Paquet, K. C. Glass, A. J. Landry, Y. Doyon, C. Cayrou, G. A. Robitaille, D. E. Richard, X. J. Yang, T. G. Kutateladze, J. Cote, *Mol Cell Biol* **2012**, 32, 689.
[72] T. W. Burke, J. G. Cook, M. Asano, J. R. Nevins, *J Biol Chem* **2001**, 276, 15397.
[73] B. Miotto, K. Struhl, *Mol Cell* **2010**, 37, 57.
[74] K. H. McConnell, M. Dixon, B. R. Calvi, *Development* **2012**, 139, 3880.
[75] D. Girdwood, D. Bumpass, O. A. Vaughan, A. Thain, L. A. Anderson, A. W. Snowden, E. Garcia-Wilson, N. D. Perkins, R. T. Hay, *Mol Cell* **2003**, 11, 1043.
[76] Q. J. Li, S. H. Yang, Y. Maeda, F. M. Sladek, A. D. Sharrocks, M. Martins-Green, *Embo j* **2003**, 22, 281.
[77] D. G. Ceschin, M. Walia, S. S. Wenk, C. Duboe, C. Gaudon, Y. Xiao, L. Fauquier, M. Sankar, L. Vandel, H. Gronemeyer, *Genes Dev* **2011**, 25, 1132.
[78] H. Niida, R. Matsunuma, R. Horiguchi, C. Uchida, Y. Nakazawa, A. Motegi, K. Nishimoto, S. Sakai, T. Ohhata, K. Kitagawa, S. Moriwaki, H. Nishitani, A. Ui, T. Ogi, M. Kitagawa, *Nat Commun* **2017**, 8, 16102.
[79] J. S. Byun, M. M. Wong, W. Cui, G. Idelman, Q. Li, A. De Siervi, S. Bilke, C. M. Haggerty, A. Player, Y. H. Wang, M. J. Thirman, J. J. Kaberlein, C. Petrovas, R. A. Koup, D. Longo, K. Ozato, K. Gardner, *Proc Natl Acad Sci U S A* **2009**, 106, 19286.
[80] Y. Liu, B. Pelham-Webb, D. C. Di Giammartino, J. Li, D. Kim, K. Kita, N. Saiz, V. Garg, A. Doane, P. Giannakakou, A. K. Hadjantonakis, O. Elemento, E. Apostolou, *Cell Rep* **2017**, 19, 1283.
[81] J. C. Chrivia, R. P. Kwok, N. Lamb, M. Hagiwara, M. R. Montminy, R. H. Goodman, *Nature* **1993**, 365, 855.

[82] J. R. Lundblad, R. P. Kwok, M. E. Laurance, M. L. Harter, R. H. Goodman, *Nature* **1995**, 374, 85.
[83] P. L. Puri, V. Sartorelli, X. J. Yang, Y. Hamamori, V. V. Ogryzko, B. H. Howard, L. Kedes, J. Y. Wang, A. Graessmann, Y. Nakatani, M. Levrero, *Mol Cell* **1997**, 1, 35.
[84] N. Pelletier, N. Champagne, S. Stifani, X. J. Yang, *Oncogene* **2002**, 21, 2729.
[85] S. A. Dames, M. Martinez-Yamout, R. N. De Guzman, H. J. Dyson, P. E. Wright, *Proc Natl Acad Sci U S A* **2002**, 99, 5271.
[86] S. J. Freedman, Z. Y. Sun, F. Poy, A. L. Kung, D. M. Livingston, G. Wagner, M. J. Eck, *Proc Natl Acad Sci U S A* **2002**, 99, 5367.
[87] Y. Kamei, L. Xu, T. Heinzel, J. Torchia, R. Kurokawa, B. Gloss, S. C. Lin, R. A. Heyman, D. W. Rose, C. K. Glass, M. G. Rosenfeld, *Cell* **1996**, 85, 403.
[88] R. P. Kwok, J. R. Lundblad, J. C. Chrivia, J. P. Richards, H. P. Bachinger, R. G. Brennan, S. G. Roberts, M. R. Green, R. H. Goodman, *Nature* **1994**, 370, 223.
[89] D. Sermwittayawong, S. Tan, *Embo j* **2006**, 25, 3791.
[90] Y. Qiu, L. Liu, C. Zhao, C. Han, F. Li, J. Zhang, Y. Wang, G. Li, Y. Mei, M. Wu, J. Wu, Y. Shi, *Genes Dev* **2012**, 26, 1376.
[91] B. J. Klein, J. Simithy, X. Wang, J. Ahn, F. H. Andrews, Y. Zhang, J. Cote, X. Shi, B. A. Garcia, T. G. Kutateladze, *Structure* **2017**, 25, 650.
[92] Y. Dou, T. A. Milne, A. J. Tackett, E. R. Smith, A. Fukuda, J. Wysocka, C. D. Allis, B. T. Chait, J. L. Hess, R. G. Roeder, *Cell* **2005**, 121, 873.
[93] V. Morales, T. Straub, M. F. Neumann, G. Mengus, A. Akhtar, P. B. Becker, *Embo j* **2004**, 23, 2258.
[94] S. J. Raja, I. Charapitsa, T. Conrad, J. M. Vaquerizas, P. Gebhardt, H. Holz, J. Kadlec, S. Fraterman, N. M. Luscombe, A. Akhtar, *Mol Cell* **2010**, 38, 827.
[95] Y. Sun, X. Jiang, Y. Xu, M. K. Ayrapetov, L. A. Moreau, J. R. Whetstine, B. D. Price, *Nat Cell Biol* **2009**, 11, 1376.
[96] A. Kaidi, S. P. Jackson, *Nature* **2013**, 498, 70.
[97] Y. Sun, X. Jiang, S. Chen, N. Fernandes, B. D. Price, *Proc Natl Acad Sci U S A* **2005**, 102, 13182.
[98] Q. Wang, M. Goldstein, *Cancer Res* **2016**, 76, 1904.
[99] A. Boija, D. B. Mahat, A. Zare, P. H. Holmqvist, P. Philip, D. J. Meyers, P. A. Cole, J. T. Lis, P. Stenberg, M. Mannervik, *Mol Cell* **2017**, 68, 491.
[100] C. L. Hirsch, Z. Coban Akdemir, L. Wang, G. Jayakumaran, D. Trcka, A. Weiss, J. J. Hernandez, Q. Pan, H. Han, X. Xu, Z. Xia, A. P. Salinger, M. Wilson, F. Vizeacoumar, A. Datti, W. Li, A. J. Cooney, M. C. Barton, B. J. Blencowe, J. L. Wrana, S. Y. Dent, *Genes Dev* **2015**, 29, 803.
[101] I. B. Hilton, A. M. D'Ippolito, C. M. Vockley, P. I. Thakore, G. E. Crawford, T. E. Reddy, C. A. Gersbach, *Nat Biotechnol* **2015**, 33, 510.
[102] A. Visel, M. J. Blow, Z. Li, T. Zhang, J. A. Akiyama, A. Holt, I. Plajzer-Frick, M. Shoukry, C. Wright, F. Chen, V. Afzal, B. Ren, E. M. Rubin, L. A. Pennacchio, *Nature* **2009**, 457, 854.
[103] C. Bian, C. Xu, J. Ruan, K. K. Lee, T. L. Burke, W. Tempel, D. Barsyte, J. Li, M. Wu, B. O. Zhou, B. E. Fleharty, A. Paulson, A. Allali-Hassani, J. Q. Zhou, G. Mer, P. A. Grant, J. L. Workman, J. Zang, J. Min, *Embo j* **2011**, 30, 2829.
[104] J. B. Baell, D. J. Leaver, S. J. Hermans, G. L. Kelly, M. S. Brennan, N. L. Downer, N. Nguyen, J. Wichmann, H. M. McRae, Y. Yang, B. Cleary, H. R. Lagiakos, S. Mieruszynski, G. Pacini, H. K. Vanyai, M. I. Bergamasco, R. E. May, B. K. Davey, K. J. Morgan, A. J. Sealey, B. Wang, N. Zamudio, S. Wilcox, A. L. Garnham, B. N. Sheikh, B. J. Aubrey, K. Doggett, M. C. Chung, M. de Silva, J. Bentley, P. Pilling, M. Hattarki, O. Dolezal, M. Dennis, H. Falk, B. Ren, S. A. Charman, K. L. White, J. Rautela, A. Newbold, E. D. Hawkins, R. W. Johnstone, N. D. Huntington, T. S. Peat, J. K. Heath, A. Strasser, M. W. Parker, G. K. Smyth, I. P. Street, B. J. Monahan, A. K. Voss, T. Thomas, *Nature* **2018**, (in press).

[105] L. M. Lasko, C. G. Jakob, R. P. Edalji, W. Qiu, D. Montgomery, E. L. Digiammarino, T. M. Hansen, R. M. Risi, R. Frey, V. Manaves, B. Shaw, M. Algire, P. Hessler, L. T. Lam, T. Uziel, E. Faivre, D. Ferguson, F. G. Buchanan, R. L. Martin, M. Torrent, G. G. Chiang, K. Karukurichi, J. W. Langston, B. T. Weinert, C. Choudhary, P. de Vries, J. H. Van Drie, D. McElligott, E. Kesicki, R. Marmorstein, C. Sun, P. A. Cole, S. H. Rosenberg, M. R. Michaelides, A. Lai, K. D. Bromberg, *Nature* **2017**, 550, 128.
[106] Y. Wang, Y. R. Guo, K. Liu, Z. Yin, R. Liu, Y. Xia, L. Tan, P. Yang, J. H. Lee, X. J. Li, D. Hawke, Y. Zheng, X. Qian, J. Lyu, J. He, D. Xing, Y. J. Tao, Z. Lu, *Nature* **2017**, 552, 273.
[107] T. Kusch, L. Florens, W. H. Macdonald, S. K. Swanson, R. L. Glaser, J. R. Yates, 3rd, S. M. Abmayr, M. P. Washburn, J. L. Workman, *Science* **2004**, 306, 2084.
[108] A. L. Kung, V. I. Rebel, R. T. Bronson, L. E. Ch'ng, C. A. Sieff, D. M. Livingston, T. P. Yao, *Genes Dev* **2000**, 14, 272.
[109] C. Gorrini, M. Squatrito, C. Luise, N. Syed, D. Perna, L. Wark, F. Martinato, D. Sardella, A. Verrecchia, S. Bennett, S. Confalonieri, M. Cesaroni, F. Marchesi, M. Gasco, E. Scanziani, M. Capra, S. Mai, P. Nuciforo, T. Crook, J. Lough, B. Amati, *Nature* **2007**, 448, 1063.
[110] K. M. Gehrking, J. M. Andresen, L. Duvick, J. Lough, H. Y. Zoghbi, H. T. Orr, *Hum Mol Genet* **2011**, 20, 2204.
[111] F. M. Perez-Campo, G. Costa, A. L. M. Lie, S. Stifani, V. Kouskoff, G. Lacaud, *Stem Cells* **2014**, 32, 1591.
[112] B. N. Sheikh, S. C. Lee, F. El-Saafin, H. K. Vanyai, Y. Hu, S. H. Pang, S. Grabow, A. Strasser, S. L. Nutt, W. S. Alexander, G. K. Smyth, A. K. Voss, T. Thomas, *Blood* **2015**, 125, 1910.
[113] M. Adamowicz, J. Vermezovic, F. d'Adda di Fagagna, *Cell Rep* **2016**, 16, 2068.
[114] B. Hanstein, R. Eckner, J. DiRenzo, S. Halachmi, H. Liu, B. Searcy, R. Kurokawa, M. Brown, *Proc Natl Acad Sci U S A* **1996**, 93, 11540.
[115] P. Dai, H. Akimaru, Y. Tanaka, D. X. Hou, T. Yasukawa, C. Kanei-Ishii, T. Takahashi, S. Ishii, *Genes Dev* **1996**, 10, 528.
[116] M. L. Avantaggiati, V. Ogryzko, K. Gardner, A. Giordano, A. S. Levine, K. Kelly, *Cell* **1997**, 89, 1175.
[117] P. J. Kallio, K. Okamoto, S. O'Brien, P. Carrero, Y. Makino, H. Tanaka, L. Poellinger, *Embo j* **1998**, 17, 6573.
[118] T. K. Kim, T. H. Kim, T. Maniatis, *Proc Natl Acad Sci U S A* **1998**, 95, 12191.
[119] M. Merika, A. J. Williams, G. Chen, T. Collins, D. Thanos, *Mol Cell* **1998**, 1, 277.
[120] L. Morris, K. E. Allen, N. B. La Thangue, *Nat Cell Biol* **2000**, 2, 232.
[121] Z. Liu, J. Wong, S. Y. Tsai, M. J. Tsai, B. W. O'Malley, *Proc Natl Acad Sci U S A* **2001**, 98, 12426.
[122] Z. Q. Huang, J. Li, L. M. Sachs, P. A. Cole, J. Wong, *Embo j* **2003**, 22, 2146.
[123] J. Vervoorts, J. M. Luscher-Firzlaff, S. Rottmann, R. Lilischkis, G. Walsemann, K. Dohmann, M. Austen, B. Luscher, *EMBO Rep* **2003**, 4, 484.
[124] S. Mujtaba, Y. He, L. Zeng, S. Yan, O. Plotnikova, Sachchidanand, R. Sanchez, N. J. Zeleznik-Le, Z. Ronai, M. M. Zhou, *Mol Cell* **2004**, 13, 251.
[125] F. Wang, C. B. Marshall, K. Yamamoto, G. Y. Li, G. M. Gasmi-Seabrook, H. Okada, T. W. Mak, M. Ikura, *Proc Natl Acad Sci U S A* **2012**, 109, 6078.
[126] Z. Tang, W. Y. Chen, M. Shimada, U. T. Nguyen, J. Kim, X. J. Sun, T. Sengoku, R. K. McGinty, J. P. Fernandez, T. W. Muir, R. G. Roeder, *Cell* **2013**, 154, 297.
[127] Y. Cherasse, A. C. Maurin, C. Chaveroux, C. Jousse, V. Carraro, L. Parry, C. Deval, C. Chambon, P. Fafournoux, A. Bruhat, *Nucleic Acids Res* **2007**, 35, 5954.
[128] X. Cao, T. C. Sudhof, *Science* **2001**, 293, 115.
[129] S. R. Frank, T. Parisi, S. Taubert, P. Fernandez, M. Fuchs, H. M. Chan, D. M. Livingston, B. Amati, *EMBO Rep* **2003**, 4, 575.
[130] T. Hattori, F. Coustry, S. Stephens, H. Eberspaecher, M. Takigawa, H. Yasuda, B. de Crombrugghe, *Nucleic Acids Res* **2008**, 36, 3011.

[131] T. Kusch, A. Mei, C. Nguyen, *Proc Natl Acad Sci U S A* **2014**, 111, 4850.
[132] Y. Xiao, Y. Nagai, G. Deng, T. Ohtani, Z. Zhu, Z. Zhou, H. Zhang, M. Q. Ji, J. W. Lough, A. Samanta, W. W. Hancock, M. I. Greene, *Cell Rep* **2014**, 7, 1471.
[133] M. Sharma, M. Zarnegar, X. Li, B. Lim, Z. Sun, *J Biol Chem* **2000**, 275, 35200.
[134] M. Georgiakaki, N. Chabbert-Buffet, B. Dasen, G. Meduri, S. Wenk, L. Rajhi, L. Amazit, A. Chauchereau, C. W. Burger, L. J. Blok, E. Milgrom, M. Lombes, A. Guiochon-Mantel, H. Loosfelt, *Mol Endocrinol* **2006**, 20, 2122.

Figures legends and Tables

Figure 1: Sequence relationship tree and schematic drawing of the protein domain structures of the mammalian nuclear HATs

The nine mammalian nuclear HATs with structurally defined acetyltransferase domains are displayed. The different protein domains are labelled "1" to "17". The common theme is the histone acetyltransferase domain (labelled with "1"). Other domains commonly confer binding either to histones directly, e.g bromodomains ("4"), PHD domains ("9") and chromodomains ("16"), or to the DNA via DNA-binding transcription factors, e.g. CBP, via its KIX domain ("7") binds to CREB (also see Table 4 for details on interactions). The full-length amino acid sequences of the human proteins were aligned and the tree drawn using Clustal Omega. Comparison of the effects of single and double loss of function mutations of the pairs of KAT2A and KAT2B, as well as KAT3A and KAT3B indicate overlapping roles. Requirement for KAT3A and KAT3B combined is similar to the essential need for TIP60 and MOF individually, while the requirement for KAT2A and KAT2B combined is more restricted.

Figure 2: Proposed mechanisms for HAT recruitment to the chromatin

References to specific studies supporting the mechanisms shown here are cited in Table 4. Recruitment to specific gene loci would imply DNA sequence-specific mechanisms such as interaction with DNA-binding transcription factors (TF) either directly (A) or indirectly (B). Recruitment through binding to components of the basal transcriptional machinery, such as TFIIB (C) could result in a general recruitment to all TFIIB bound promoters or more specific recruitment, if combined with binding to TF (C). Binding to specifically modified histone residues, such as H3K4me3 in the case of some PHD domains, either directly (D) or via an adaptor protein (E), could result in recruitment to transcriptionally active and poised genes loci. Recruitment to sites of DNA double strand breaks (DSBs) has been reported for TIP60 to involve binding to a specific modified histone residue (e.g. H3K9me3) and the DNA damage repair protein ATM (either directly or indirectly via FOXO3).

Table 1: Reported histone acetylation targets of individual nuclear HATs

HAT	Category of recruitment	Assay	Details	Refs
CBP P300	H3K56	siRNA in cells	siRNA KD of CBP and P300 in HeLa cells abolishes H3K56ac	[19]
CBP P300	H3K18ac H3K27ac	Genetic deletion, mouse cells	Deletion of CBP/P300 in MEFs specifically and dramatically reduces H3K18ac and H3K27ac	[18]
CBP	H3K27ac H4K5ac	Cell-free HAT assay	Enhancer RNAs interact with CBP, stimulate CBP's acetylation of H3K27, reconstituted nucleosomes	[36]
GCN5 PCAF	H3K9ac	Genetic deletion, mouse cells	Deletion of GCN5/PCAF in MEFs specifically and dramatically reduces H3K9ac	[18]
GCN5	H3K9ac H3K14ac H3K79succ	Cell-free assay, shRNA in cells	GCN5 succinylates H3K79 in cell-free assays and shRNA KD of GCN5 or a GCN5 HAT dead mutant reduce H3K79succ in U87 cells, and H3K79succ, H3K9ac and H3K14ac in U251 cells	[106]
TIP60	H2A-K5, H3K14 H4K5, H4K8 H4K12, H4K16	Cell-free assay	H2A-K5, H3K14, H4K5, 8, 12, 16 in cell-free assay, N-terminal histone peptides assessed by mass spectrometry	[33]
TIP60	H2Av	Cell-free assay	*Drosophila* Tip60 acetylates H2Av in cell-free nucleosomal histones	[107]
TIP60	H2A-K5	In cells	TIP60 occupancy of ER-regulated genes correlates with H2A-K5ac in MCF-7 cells	[35]
TIP60	H2A-K5 H2A-K15	Cell-free assay	TIP60 acetylates H2A-K5, 15 and H4 in cell-free nucleosomes, H2A-K5ac and 15ac is reduced after siRNA KD of TIP60 in U2OS and HEK293T cells	[34]
MOF	H4K16	In cells, siRNA	siRNA KD of MOF in HeLa or HepG2 cells, reduction in H4K16ac, no differences in acetylation of H3K14, H3K23, H4K12	[37]
MOF	H4K16	Genetic deletion, mouse embryo	*Mof*$^{-/-}$ mouse embryos lack H4K16ac	[38]
MOF	H4K16	Genetic deletion, mouse embryo	*Mof*$^{-/-}$ mouse embryos lack H4K16ac, no difference in H3K9, H3K14, H4K5, H4K8, H4K12 acetylation	[39]
HBO1	H3, H4	Cell-free assay	HBO1 complex acetylates H3 and H4 as histone tetrameres, but H3 only in reconstituted chromatin	[29]
HBO1	H4K5, H4K8 H4K12	In cells, shRNA, siRNA	shRNA and siRNA KD in HEK293T cells, reduction in acetylation of H4K5,8,12	[20]
HBO1	H3 H4	Cell-free HAT assay	Isolated HBO1-BRPF1 complex acetylates only H3K14 and 23 on chromatin, but both H3 and H4 on free histones, HBO1-JADE complex acetylates H4	[27]
HBO1	H2A, H3, H4	Cell-free assay	HBO1-JADE complex	[28]
HBO1	H3K14 H3K9 H4K5, H4K8	In cells, shRNA KD	shRNA depletion of HBO1 profound reduction in H3K14ac and H3K9ac, acetylation of H4K5 and 8 reduced to lesser degree	[22]
HBO1	H3K14	Genetic deletion, mouse cells	Cells from *Hbo1* null embryos, >90% reduction of H3K14ac, no reduction in acetylation of H4K5, 8, and 12, and upregulation of H4K16ac and H3K9ac	[21]
HBO1	H3K14	In cells, shRNA	Depletion of HBO1 ablated H3K14ac in MLE cells	[32]
HBO1	H3K14	In cells, siRNA	Depletion of HBO1 caused reduction in H3K14ac	[30]
HBO1	H3K14	Constitutive HBO1	HBO1$^{Ser50/53Ala}$ mutant cells, deregulated H3K14ac	[31]
HBO1	H3K14 H4	In cells, shRNA	UV irradiation increases H3K14ac and H4, but not when is HBO1 depleted in HeLa cells	[78]
KAT6A KAT6B	H3K14	Indirect evidence via ING5 in cells	Isolated KAT6B complex acetylates H3, KAT6A/ KAT6B complex member ING5 affinity to H3K14ac	[20]

KAT6A	H3K9	Genetic deletion, mouse embryos	*KAT6A* null embryos, reduced H3K9ac and expression at *Hox* genes, rescued by retinoic acid treatment in utero	[40]
KAT6A	H3K9	Genetic deletion, mouse embryos	*KAT6A* null embryos, reduced H3K9ac and expression at *Tbx* genes, rescued by transgenic overexpression of *Tbx1*	[41]
KAT6B	H3K23	In cells, shRNA KD, overexpression	shRNA KD of KAT6B in NCI-N417 cells, reduction in H3K23ac, no effects on H3K14ac or H4K16ac	[42]
KAT6A	H3K23 H3K9 H3K14	In cells, shRNA KD	shRNA KD of KAT6A in U87 and LN229 cells, reduced H3K23ac and to a lesser degree H3K9ac and H3K14ac	[43]

Table 2: Loss of function phenotypes of HATs in mice

HAT	Category of mutation	Details	Refs
GCN5	Null	Developmental arrest between E7.5 and E8.5, high incidence of apoptosis in the Gcn5l2 mutants before the onset of morphological abnormality	[60]
PCAF	Null	Mice lacking PCAF are viable	[61]
GCN5 PCAF	*Gcn5*$^{+/-}$;*Pcaf*$^{-/-}$	Viable	[60]
GCN5 PCAF	Double null	Double null embryos are abnormal at E6.5, arrested at the egg cylinder stage, dead at E7.5	[60]
GCN5 PCAF	Double null	GCN5/PCAF DKO MEFs grow with normal morphology	[18]
CBP	Null	Lethality between E8 and E10, neural tube closure defects, haemorrhage	[62-64]
CBP	Heterozygotes	Multilineage defects in hematopoietic differentiation, increased incidence of hematologic malignancies	[108]
P300	Heterozygotes	Partial penetrance exencephaly and lethality beginning E10.5	[64]
P300	Null	Developmental arrest between E8.5 and E10.5, neural tube closure defects, cardiac defects	[64]
CBP P300	Double heterozygous	Developmental arrest between E8.5 and E9.5, neural tube closure defects	[64]
CBP P300	Double null	CBP/P300 DKO MEFs cease to grow and appear flat	[18]
TIP60	Null	Embryonic death at E3.5 with inner cell mass apoptosis	[50]
TIP60	Heterozygous for null allele	*Tip60* heterozygous loss accelerates *Eμ-Myc*-induced lymphomogenesis	[109]
TIP60	Heterozygous for null allele	Slows mid-stage neurodegeneration in a spinocerebellar ataxia type 1 (SCA1)	[110]
TIP60	HAT activity-dead point mutation	Growth restriction at day 6.5, ESCs can proliferate, minimal alterations in gene expression, chromatin accessibility, impaired differentiation into mesoderm and endoderm	[51]
MOF	Null	Embryonic death at E3.5.	[38]
MOF	Null	Embryonic death at E3.5. Without MOF, mouse embryos undergo first global chromatin condensation and then apoptosis	[39]
HBO1	Null	*Hbo1* null embryos arrest in development at E8.5 and die two days later	[21]
KAT6A	HAT activity-dead point mutation	KAT6A promotes proliferation of haematopoietic and neural stem cells	[111]
KAT6A	Null	Complete absence of definitive haematopoietic stem cells	[47, 48]
KAT6A	HAT activity-dead point mutation	Number of haematopoietic stem and progenitor cells reduced and B-cell development defect	[49]
KAT6A	Null	Death at birth, homeotic transformation of 19 body segments typical of mis-patterned HOX code, rescued by upregulation of *Hox* gene expression with retinoic acid in utero	[40]
KAT6A	Null	Death at birth, cleft palate, heart defect typical of haploinsufficient loss of TBX1, rescued by re-expression of *Tbx1*	[41]
KAT6A	Heterozygous for null allele	Retards *Eμ-Myc*-induced lymphomogenesis 4-fold	[112]
KAT6A	Null	Null MEFs undergo cellular senescence	[25]
KAT6B	Hypomorph (10%) residual activity	80% death at birth, remaining 20% failure to thrive, dwarfism, brain developmental defect, reduced numbers of neurons of various types	[59]
KAT6B	Hypomorph (10%) residual activity	Reduction in development and self-renewal of adult neural stem cells and deficient neuronal differentiation	[58]

Table 3: Actual or predicted distribution of individual nuclear HATs in the genome

HAT	Category of genome distribution	Actual/ predicted	Details	Refs
CBP	At specific gene loci	Predicted	Based on interactions with specific DNA-binding TF	See Table 4
P300	At specific gene loci	Predicted	Based on interactions with specific DNA-binding TF	See Table 4
CBP	At all active gene loci	Predicted	CBP interacts with TFIIB through a domain that is conserved in the yeast co-activator ADA-1	[88]
CBP P300	At all active gene loci	Actual	ChIP-seq experiments detect high positive correlation with RNA Pol II binding throughout the genome	[2]
CBP P300	At most active gene loci	Predicted	Based on pronounced genome-wide loss of H3K18ac and H3K27ac when both are mutated	[18]
PCAF	At specific gene loci	Predicted	Based on interactions with specific DNA-binding TF	See Table 4
GCN5 PCAF	At all active gene loci	Actual	ChIP-seq experiments detect high positive correlation with RNA Pol II binding throughout the genome	[2, 100]
GCN5	At all active gene loci	Predicted	GCN5 complex proteins Sgf29 binds H3K4me2/3	[103]
GCN5 PCAF	At most active gene loci	Predicted	Based on pronounced genome-wide loss of H3K9ac when both are mutated	[18]
TIP60	At specific gene loci	Predicted	Based on interactions with specific DNA-binding TF	See Table 4
TIP60	At all active gene loci	Actual	ChIP-seq experiments detect high positive correlation with RNA Pol II binding throughout the genome	[2]
TIP60	At sites of DNA damage	Predicted	Via TIP60 chromodomain binding to H3K9me3	[95]
TIP60	At specific promoters	Predicted	Based on TIP60-Flag ChIP-seq and binding via MBTD1 to H4K20me1/2	[34]
TIP60	At sites of DNA damage	Predicted	Based on a binding via FOXO3 to ATM	[113]
MOF	X chromosome	Predicted	Simultaneous contact of MOF with MSL1 and MSL3 leads to recruitment to chromatin, X chromosome in *Drosophila*	[93]
MOF	At MLL1 bound promoters	Predicted	Based on interactions between MLL1 C-terminal domain and MOF zinc finger domain	[92]
MOF	At all active gene loci	Predicted	Based on pronounced genome-wide loss of H4K16ac when mutated	[38, 39]
MOF	At all active gene loci	Actual	ChIP-seq experiments detect high positive correlation with RNA Pol II binding throughout the genome	[2]
MOF	All promoters	Actual	ChIP-Seq shows NSL1 and MCRS2 bind to promoters genome-wide	[94]
HBO1	At origins of replication	Predicted	A fraction of the relatively abundant HBO1 protein associates with ORC1 in human cell extracts	[29]
HBO1	At most active gene loci	Actual	ChIP-seq experiments and based on interactions with chromatin binding proteins ING4, ING5, JADE1, JADE2, JADE3, BRPFs	[27, 28]
HBO1	At most gene loci	Actual	Anti-HBO1 ChIP-seq shows binding to gene body and promoter, increased levels of binding correlate with higher gene expression	[71]
HBO1	At origins of replication	Actual	Enrichment of HBO1 at ORC1-binding sites and origins of replication	[30]
HBO1	At most active gene loci	Predicted	Based on pronounced genome-wide loss of H3K14ac when mutated	[21]
KAT6B	At specific active gene loci	Predicted	Based on normal development of many organs and cell types in the KAT6B deficient state	[58, 59]
KAT6A	At specific active gene loci	Predicted	KAT6A null cells and embryos, absence of changes in histone acetylation genome-wide, moderate locus-specific loss of H3K9ac, normal development of many organs and cell types	[26, 40, 41]

Table 4: Examples of proposed mechanisms for nuclear HAT recruitment to the chromatin

HAT	Category of recruitment	Binding partner	Details	Refs
KAT3A (CBP, CREBBP) and KAT3B (P300, EP300)				
CBP	Binding to transcription factor (TF)	CREB	CREB phosphorylated by protein kinase A binds to cAMP response element (RE) and recruits CBP	[81]
CBP	Binding to basal transcription factor	TFIIB	CBP interacts with TFIIB through a domain that is conserved in the yeast coactivator ADA-1	[88]
P300	Binding to TF	CREB	CBP and P300 have similar binding affinity for the PKA-phosphorylated form of CREB, E1A interferes with binding	[82]
P300	Indirect binding to nuclear receptor (NR)	ER via ERAP160	Hormone-dependent interaction between oestrogen receptor, its associated protein, ERAP160, resulting in the recruitment of P300	[114]
CBP P300	Indirect binding to NR	NR via p160/SRC1	Activation by NRs requires CBP, limiting amounts of CBP/P300, CBP interacts with p160 variants of the SRC1	[87]
CBP	Binding to TF	MYB	CBP binds transcriptional activation domain of c-Myb	[115]
P300	Binding to TF	P53	p53 protein complex with P300	[116]
CBP P300 PCAF	Binding to TF	MyoD	MyoD interacts directly with both P300/CBP and PCAF, forming a multimeric protein complex	[83]
CBP P300	Binding to TF	HIF1a	HIF1a requires hypoxic stimuli for function as a CBP-dependent transcription factor	[117]
CBP	Binding to TF	NF-kappa B, IRF1, ATF2/c-Jun	ATF2/c-JUN, IRF1, p50/p65 of NF-kappaB, HMGI(Y) required CBP to activate interferon b gene	[118]
CBP P300	Binding to TF	NF-kappa B, IRF1, ATF2/c-Jun	CBP/P300 coactivator to the enhanceosome, via a new activating surface assembled from the novel p65 domain and the activation domains of all of the activators.	[119]
CBP P300	Binding to TF	E2F5	phosphorylated E2F5 recruits P300/CBP	[120]
P300	Indirect binding to NR	PR via SRC1	Sequential recruitment of SRC-1 and P300 to liganded PR, efficient recruitment of P300 required SRC-1	[121]
CBP	Binding to TF	HIF1a	Interaction of carboxy-terminal activation domain of HIF1a and the zinc-binding motif (TAZ1)/ cysteine/histidine-rich 1 (CH1) domain of CBP	[85] [86]
P300	Regulation of HAT *in situ*	NA	Sumoylation of P300 in its CRD1 transcriptional repression domain inhibits its acetylation function	[75]
P300	Indirect binding to NR	AR via SRC	SWI/SNF and Mediator complexes can be targeted to chromatin by P300 after it is recruited via SRC to sites of AR binding	[122]
CBP P300	Binding to TF	MYC	MYC is stabilised by acetylation	[123]
CBP	Binding to TF	P53	Bromodomain of CBP binds p53 at acetylated lysine 382	[124]
CBP	Indirect binding to NR, regulated by methylation	ER via p160	CARM1-dependent CBP methylation and p160 coactivator required for oestrogen-induced recruitment to chromatin targets, methylation increased HAT activity of CBP	[77]
CBP P300	Binding to TF, regulated by phosphorylation	FOXO3a	Conserved region 3 (CR3) and CR2 of FOXO3a bind KIX domain of CBP/P300, phosphorylation of S626 in CR3, increased affinity	[125]
P300	Binding to TF	P53	SET1 complex (SET1C)-mediated H3K4me3 dependent on p53- and P300-mediated H3ac, SET1C-mediated enhancement of p53- and P300-dependent transcription	[126]
KAT2A (GCN5) and KAT2B (PCAF)				
CBP/P300 PCAF	Binding to TF	MyoD	MyoD interacts directly with P300/CBP and PCAF	[83]
GCN5	Indirect binding to basal transcription factor	Spt8 and Ada1 bind TBP	SAGA (Spt-Ada-Gcn5-acetyltransferase) complex acts as a coactivator to recruit the TATA-binding protein (TBP) to the TATA box	[89]
PCAF	Binding to TF	ATF4	N-terminal region of ATF4 is required for a direct interaction with PCAF	[127]
GCN5	Histone binding via chromatin binding	Sgf29 binds H3A1 and	Tandem Tudor domains of SAGA complex component Sgf29 bind H3A1 and H3K4me2/3 peptides	[103]

	protein	H3K4me3							

KAT5 (TIP60, HTATIP)				
TIP60	Indirect binding to TF via adaptor	Fe65	Cytoplasmic tail of APP forms a multimeric complex with the nuclear adaptor protein Fe65	[128]
TIP60	Binding to TF	MYC	MYC associates with TIP60 and recruits it to chromatin in vivo with four other components of the TIP60 complex: TRRAP, p400, TIP48 and TIP49	[129]
TIP60	Binding to TF	SOX9	Endogenous Sox9 interacts with Tip60, Tip60 enhanced the transcriptional activity of Sox9	[130]
TIP60	Binding to NR and to histone	ERa H3K4me1	Oestrogen-induced recruitment, direct binding of TIP60 to ERalpha and chromatin-remodelling ATPase BRG1, increased recruitment of MLL1, increased H3K4me1, binding of TIP60 chromodomain to H3K4me1	[35]
TIP60	Binding to histone	H3K9me3	Direct interaction between the chromodomain of TIP60 and histone H3 trimethylated on lysine 9 (H3K9me3) at DSBs	[95]
TIP60	Binding to histone	H3K9me3	Tyrosine phosphorylation of TIP60 increases after DNA damage, promotes TIP60 binding to H3K9me3	[96]
TIP60	Binding to TF	HSF	Heat Shock Factor recruits the dTip60 complex to the hsp70 loci	[131]
TIP60	Binding to TF	FOXP3	TIP60 interacts with and acetylates FOXP3	[132]
TIP60	Binding via FOXO3/ATM	ATM	FOXO3a binding to ATM leads to TIP60 association with ATM	[113]
TIP60	Binding to RNA processing protein	AGO2	diRNA-associated AGO2 interacts with TIP60 to the DSB	[98]
TIP60	Histone binding via reader protein	MBTD1	MBTD1 histone reader domain for H4K20me1/2 recruits TIP60 to gene promoters (and promotes DNA double-strand break repair)	[34]
KAT8 (MOF, MYST1)				
MOF	Histone binding via chromatin associated protein	MSL1	MSL1 interacts with zinc finger within MOF's HAT domain to recruit MOF	[93]
MOF	Histone binding via chromatin associated protein	MLL1	MLL1 C-terminal interacts with MOF zinc finger domains to recruit MOF to chromatin	[92]
MOF	Histone binding via chromatin associated protein	MCRS2	MCRS2 recruits MOF to DNA	[94]
KAT7 (HBO1, MYST2)				
HBO1	Binding to ORC1	ORC1	A fraction of the relatively abundant HBO1 protein associates with ORC1 in human cell extracts	[29]
HBO1	Binding to NR	AR	Ligand-enhanced interaction between AR and HBO1	[133]
HBO1	Binding to NR	PR	N-terminal domain of PR binds C-terminal MYST domain of HBO1; HBO1 interacts through its NTD with SRC1a in the absence of steroid receptor	[134]
HBO1	Histone binding via chromatin binding protein	ING4/5 JADE1/2/3 H3K4me	Three PHD finger domains in two different subunits of the HBO1 complex: tumour suppressor proteins ING4/5 and JADE1/2/3 with different specificities to methylated H3K4	[28]
HBO1	Histone binding via chromatin binding protein	ING4 ING5 methylated H3K4	Multiple PHD finger domains present in different subunits bind methylated H3K4 methylation; ING4/5 PHD domain with HBO1-JADE	[71]
KAT6A (MOZ, MYST3) and KAT6B (MORF, QKF, MYST4)				
KAT6A KAT6B	Binding to TF	RUNX1 RUNX2	KAT6A and KAT6B interact with Runx2 through a C-terminal SM (serine- and methionine-rich) domain, the SM domain of KAT6B also binds to Runx1 (AML1)	[84]
KAT6A	Binding to histone	H3K14ac	Double PHD finger (DPF) of KAT6A in crystal structure with H3K14ac peptide, recognition of unmodified R2 and acetylated K14 on histone H3	[90]
KAT6A	Binding to histone	H3K14ac	DPF of KAT6A induces alpha-helical conformation of H3K4-T11, helical structure facilitates sampling of H3K4 methylation status, proffers H3K9 and other residues for modification	[16]
KAT6A	Binding to histone	H3K14cr	DPF domain of KAT6A binds wide range of histone lysine	[17]

			acylations, preference for Kcr, crystal structures of the DPF domain of KAT6A in complex with H3K14cr, H3K14bu, and H3K14pr peptides, non-acetyl acylations anchored in a hydrophobic pocket	
KAT6B	Histone binding	H3K14bu	KAT6B DPF binds H3K14bu, -ac, -succ, and -hib	[91]

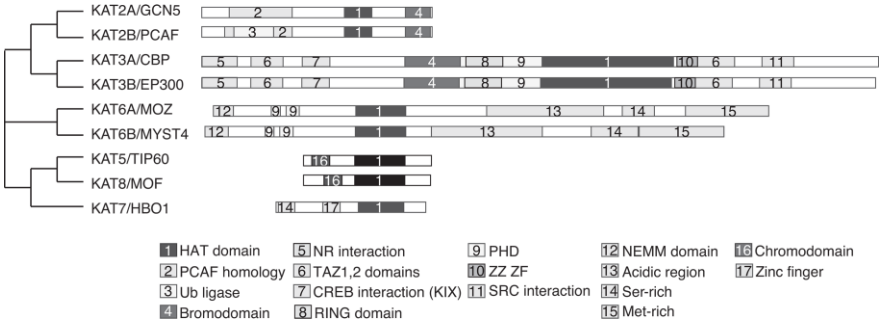

Figure 1

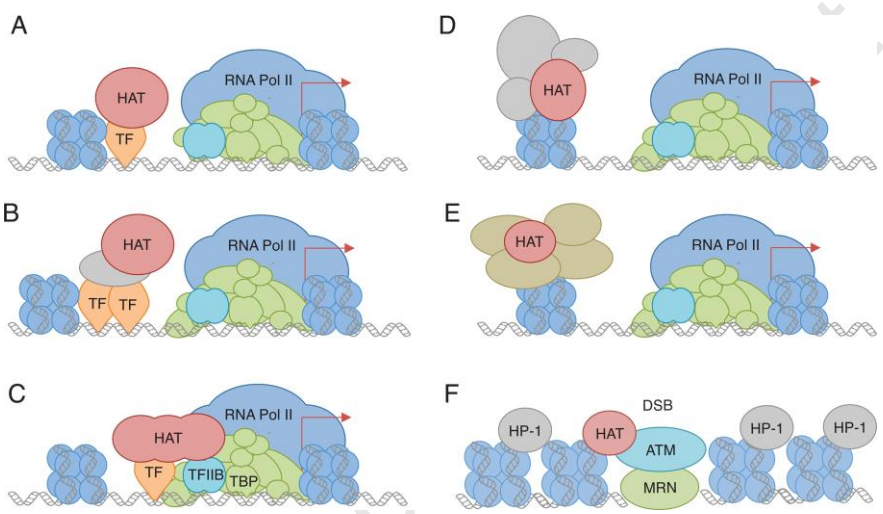

Figure 2