Published in final edited form as:

Biopolymers. 2013 February ; 99(2): 98–111. doi:10.1002/bip.22128.

Histone Acetyltransferases: Rising Ancient Counterparts to Protein Kinases

Hua Yuan¹ and Ronen Marmorstein^{1,2}

¹Program in Gene Expression and Regulation, The Wistar Institute, Philadelphia, PA, 19104 USA

²Department of Chemistry, University of Pennsylvania, Philadelphia, PA, 19104 USA

Abstract

Protein kinases catalyze phosphorylation, a posttranslation modification widely utilized in cell signaling. Histone acetyltransferases (HATs) catalyze a counterpart posttranslation modification of acetylation which marks histones for epigenetic signaling but in some cases modifies non-histone proteins to mediate other cellular activities. In addition, recent proteomic studies have revealed that thousands of proteins are acetylated throughout the cell to regulate diverse biological processes, thus placing acetyltransferases on the same playing field as kinases. Emerging biochemical and structural data further supports mechanistic and biological links between the two enzyme families. In this article, we will review what is known to date about the structure, catalysis and mode of regulation of HAT enzymes and draw analogies, where relevant, to protein kinases. This comparison reveals that histone acetyltransferases may be rising ancient counterparts to protein kinases.

Introduction

Protein kinases were discovered in the 1950s in a search for an enzyme that converts the inactive phosphorylase b to its active form through phosphorylation ^{1,2}. The subsequent discovery of protein kinase A (PKA) that phosphorylates the phosphorylase kinase marked the beginning of the study of signaling cascades that are regulated by multiple phosphorylation events ³. Histone acetyltransferases (HATs) were first isolated in the 1990s, when Allis and coworkers purified a HAT from *Tetrahymena thermophila* that was orthologous to a previously identified transcriptional adaptor from yeast called Gcn5 (KAT2, nomenclature according to ref. ⁴) and conserved from yeast to man ⁵; and Sternglanz and coworkers ⁶ and Gottschling and coworkers ⁷ identified HAT1 (KAT1) that was initially proposed to be a cytoplasmic specific acetyltransferase and later shown to also harbor nuclear functions ^{8–10}. In the same year, Schreiber and colleagues isolated a mammalian histone deacetylase (HDAC) that was highly homologous to a previously characterized transcriptional repressor, Rpd3, also with conservation from yeast to man ¹¹. The findings that transcription factors can process acetyltransferase and deacetylase enzymatic activities established the connection between histone acetylation status and transcription, a concept first proposed by Allfrey and colleagues ¹². Since the initial identification of HATs and HDACs, other such enzymes were identified along with enzymes that mediate reversible histone modifications such as phosphorylation and methylation ^{13,14}. While many posttranslational histone modifications can be correlated with different DNA-

templated activities, HAT and HDAC activities are generally correlated with gene activation and repression/silencing, respectively. Aberrant HAT and HDAC activities are also correlated with diseases such as cancer and metabolic disorders^{15,16}. To date, the HATs and HDACs are the biochemically and structurally most well characterized among the histone posttranslational modification enzymes.

Although initially characterized as HATs, several of these acetyltransferase enzymes have been shown to also acetylate non-histone proteins¹⁷ such as the p53 tumor suppressor¹⁸, the importin- α adaptor¹⁹ and the E1A viral oncoprotein²⁰. In addition, recent proteomic studies on prokaryotic and eukaryotic cells have revealed that thousands of non-histone proteins are acetylated to mediate diverse biological processes in and outside of the nucleus^{21–24}. Some of the proteins that mediate these activities are just beginning to be identified^{25,26} including the MEC-17/ α TAT1 α -tubulin acetyltransferase^{27,28} and the Eco1/Esco1 cohesion acetyltransferase^{26,29,30}, with many other acetyltransferases still to be determined. The identification of so many acetylated substrates that are likely mediated by many enzymes may be reminiscent of protein kinases, as initially proposed by Kouzarides³¹, where hundreds of enzymes phosphorylate tens of thousands of substrates^{32,33}. The number of protein acetyltransferases in the genome is not yet known due to the relatively poor sequence similarity between these enzymes. As will be highlighted below, other similarities between these two enzyme families include architectural features of their structures, regulation by other protein cofactors and by automodification and the ability of their respective posttranslational modifications to be recognized by other protein modules to signal downstream activities. In addition, the aberrant activities of both enzyme families have been associated with several human diseases, thus making them both attractive drug targets. In this review, we will summarize the emerging mechanistic features of HATs and highlight their characteristics that mimic kinases in order to support the working hypothesis that protein acetyltransferases are rising ancient counterparts to protein kinases. Those who are interested in a more detailed analysis of protein kinases are directed to other excellent reviews that are exclusively focused on these enzymes^{34–36}.

Overall Structure of HATs

HATs can be grouped into at least five different families, based on the sequence divergence within the HAT domain, this includes HAT1 (KAT1) (named Histone Acetyltransferase 1 as the founding member of the superfamily), Gcn5/PCAF (named for its founding member yeast Gcn5 (KAT2) and its human ortholog, PCAF (KAT2B)), MYST (named for the founding members human MOZ (KAT6A), yeast Ybf2/Sas3 (ySAS3, KAT6), yeast Sas2 (ySas2, KAT8) and human Tip60 (hTIP60, KAT5)), p300/CBP (named for the two human paralogs hp300 (KAT3B) and hCBP (KAT3A)) and Rtt109 (yRtt109, KAT11) (named for its initial identification as a yeast regulator of Ty1 transposition gene product 109). While the Gcn5/PCAF, HAT1 and MYST families have homologs from yeast to man, p300/CBP is metazoan-specific, and Rtt109 is fungal-specific. Although other nuclear HAT families have been identified, such as the steroid receptor coactivators (human ACTR/AIB1, human SRC1)³⁷, human TAF250³⁸, human ATF-2³⁹, and human CLOCK⁴⁰, their HAT activities have not been as extensively studied as the five major HAT classes and will not be further discussed here.

Representative structures of each of the 5 families of HAT proteins have been determined by X-ray crystallography revealing the molecular characteristics of the enzymatic domains and molecular insights into catalysis and substrate acetylation (Figures 1a–1e). Despite the sequence divergence, each of the protein families contain a structurally conserved core region containing a 3-stranded β -sheet with a long α -helix parallel and spanning one side of the sheet and corresponding to the A and D regions shown to harbor sequence conservation among the Gcn5-related N-acetyltransferases (GNAT) proteins⁴¹. This core region makes conserved interactions with the AcCoA cofactor that packs along one edge of this core region; however other non-conserved regions of the HATs also make additional protein specific AcCoA interactions. This core region is flanked on opposite sides by α/β segments that are structurally distinct between the different HAT families (Figures 1a–1e). The one exception to this rule is p300/CBP and yRtt109, which show structural homology in these regions despite a lack of sequence conservation. The central core region, together with the flanking segments, form a cleft over the central core for histone substrate to bind and for catalysis to occur.

Kinases have no overall structural homology to the HAT proteins; however they also share a structurally conserved central ATP cofactor-binding site that is flanked by N- and C-terminal domains that show greater sequence divergence than the ATP cofactor-binding site. All kinases contain a β -rich N-lobe with 1 or 2 helices and α -rich C-lobe with 2 small β -strands that are connected by a series of loops that together show much greater structural conservation than do the HATs (Figure 1f). The ATP and magnesium ions that serve as substrate cofactors bind in the cleft between the N and C-lobes that contains several hallmark features including a P-loop and hinge region for ATP binding and catalytic and activation loops for regulation of substrate specific phosphorylation^{34,36} (Figure 2f). The canonical C-helix that is part of the N-lobe, together with the activation loop forms a mechanistic bridge between the two lobes playing a critical and dynamic role in kinase activity. Indeed, protein dynamics in regulating substrate binding and catalysis appears to be an important feature of kinase function³⁶. We don't yet know if protein dynamics might play a similarly important role in HAT activity.

Catalytic Mechanism of HATs

HAT proteins transfer the acetyl group from the acetyl-CoA cofactor to the N ϵ nitrogen of the target lysine side chain within histones. Structural, biochemical, mutational and enzymatic analysis has provided insights into the catalytic mechanism of these enzymes, remarkably revealing that they use different catalytic strategies for acetyl transfer. Specifically, studies on the Gcn5/PCAF family of HATs reveals that a strictly conserved glutamate in the active site (Glu173 in yGcn5 and Glu570 in hPCAF) acts as a general base for catalysis using a water molecule that is well ordered in the crystal structure (Figure 2a), through a bi-bi ternary complex mechanism where both substrates must be bound to the enzyme before catalysis can occur. An acid, if any, that protonates the CoA leaving group has not been identified.

Analogous structural and enzymatic studies of the yeast Esa1 (yEsa1) member of the MYST HAT family reveals a ping-pong catalytic mechanism involving conserved glutamate

(Glu338) and cysteine (Cys304) residues where the enzyme first forms an acetylated intermediate involving Cys304 prior to formation of the acetylated histone product through the Glu338 general base ⁴² (Figure 2b). Paradoxically, a more recent report demonstrates that yEsa1 assembled within a physiologically relevant piccolo NuA4 complex does not show a strong dependence on Cys304 for catalysis leading the authors to conclude that yEsa1 within the piccolo NuA4 complex proceeds through a ternary bi-bi catalytic mechanism similar to the Gcn5/PCAF family ⁴³, suggesting that the same HAT enzyme may employ different catalytic mechanisms within different cellular contexts.

Unlike the Gcn5/PCAF and MYST HATs, the p300/CBP HATs do not employ a glutamate residue for catalysis. Instead, mutagenesis and kinetic analysis of the potential catalytic residues in the active site uncovers only two residues, Tyr1467 and Trp1436, of hp300 that show a significant effect on catalysis when mutated (Figure 2c) ⁴⁴. Based on the position of these residues in the structure, Tyr1467 was proposed to play a role as a general acid for catalysis, while Trp1436 was proposed to help orient the target lysine into the active site. Correlating with the importance of Tyr1467 and Trp1436 for catalysis was the observation that these residues are strictly conserved within the p300/CBP HAT family. Thus, it appears that the p300/CBP family does not employ a general base for catalysis, in contrast to the Gcn5/PCAF and MYST HAT families. Taken together with the fact that hp300 is inhibited potently by the more primitive Lys-CoA inhibitor but poorly by bisubstrate inhibitors with longer peptide moieties ⁴⁵, and the observation that longer peptides are better substrates for hp300 than lysine, it was proposed that the p300/CBP family of HATs employ a “hit-and-run” or Theorell-Chance acetyl transfer mechanism that is distinct from the catalytic mechanisms employed by the Gcn5/PCAF and MYST HAT families.

The kinetic mechanism for acetylation by the Rtt109 and Hat1 families have been less well characterized than the other HAT families. Both enzyme families require association with other protein regulatory subunits for full activity. Yeast Rtt109 harbors very low acetyltransferase activity on its own ^{46,47}, but its activity is stimulated by association with either of the histone chaperone proteins yAsf1 or yVps75 ^{47–51}. Although it is not clear how these histone chaperones enhance the catalytic activity of yRtt109, the crystal structures of yRtt109/yVps75 complexes (in both 2:2 and 2:1 stoichiometry) that show no significant changes in the yRtt109 active site as a function of yVps75 binding suggests that histone chaperones may function to deliver histone substrates to yRtt109 for acetylation ^{52–54}. The crystal structure of yRtt109 revealed that despite its overall structural superposition with the hp300 HAT domain, the key catalytic residues of hp300 (Tyr1467 and Trp1436) are not conserved in yRtt109 (Figure 2d). Instead, kinetic analysis demonstrated that the yRtt109/yVps75 complex employs a sequential kinetic mechanism whereby the yRtt109–yVps75 complex, AcCoA, and histone H3 substrates form a complex prior to chemical catalysis ⁴⁸. The structure of yRtt109 and mutational and kinetic analysis points to the importance of Asp89, Trp222 and Asp288 for catalysis although mutation of these residues most significantly reduce the K_M for AcCoA binding ^{48,52,54}. Therefore, key residues involved in catalytic turnover, such as a general base or acid have not been identified for yRtt109 and dedicated residues to mediate acid/base catalysis may not be employed in this case.

HAT1 requires a general base for catalysis as suggested by a sigmoidal shaped pH profile ⁴. The recent structure determination of human HAT1 (hHAT1) bound to AcCoA and a histone H4 peptide centered around K12 reveals that three residues, Glu187, Glu276 and Asp277 are in proximity to the N ζ nitrogen of H4K12 and thus could act as potential general base residues for catalysis ⁴ (Figure 2e). A superposition of the hHAT1 active site with yGcn5 and yEsa1 reveals that residues in these enzymes that have been shown to function as a general base of catalysis, glutamates 173 and 338, respectively, superimpose with Glu276 of hHAT1 (Glu255 in yHat1) ^{4,55}, which is strictly conserved among Hat1 orthologs. In hHAT1, mutating Glu276 to Gln causes a 28.5-fold decrease in k_{cat} , while E187Q and D277N mutations cause a ~15.3 and ~8.1 fold decrease in k_{cat} , respectively. E276Q and E187Q also increase the pKa of ionization from 8.15 (± 0.06) to 8.74 (± 0.15) and 9.15 (± 0.09), respectively; while D277N only increases the pKa slightly to 8.35 (± 0.12). In addition, unlike MYST HATs, Hat1 does not contain a cysteine residue in the active site of the enzyme. Taken together, the observation described above suggests that Glu276 of hHAT1, and probably also Glu187, plays a role as general base(s) for catalysis through a ternary complex mechanism, similar to the Gcn5/PCAF HAT family.

Together, a comparison of the catalytic mechanisms of the different HAT families reveals remarkable diversity in the way each family mediates acetyl transfer. This likely reflects the relatively long evolutionary time that HATs have had to evolve and the relatively low chemical “cost” that is required to transfer an acetyl group from a thioester to an amine, as opposed to some other more demanding chemical reactions such as phosphorylation by kinases, which appears to use a more conserved catalytic mechanism involving an ordered mechanism whereby the Mg²⁺-ATP binds first before high affinity binding for protein substrate can occur. In addition, a pair of aspartic acid residues (Asp166 and Asp184 in Protein Kinase A) participate in catalysis by serving as a general base and for positioning the Mg²⁺ to facilitate phosphate transfer, respectively (Figure 2f).

Histone Substrate Binding by HATs

To date, the only direct molecular insights into histone binding by HATs has come from the structures of the hHAT1 HAT domain bound to AcCoA and a histone H4 peptide centered around Lys12 ⁴, and the *Tetrahymena* Gcn5 (tGcn5) HAT domain bound to CoA and several cognate substrate peptides including histone H3 (centered around Lys14) ^{56–58}, histone H4 (centered around Lys8) and p53 (centered around Lys320) ⁵⁷. These structures reveal that the peptide substrates bind across a groove formed by the central core region at the base and flanked on opposite sides by the HAT-specific N- and C-terminal regions that mediate the majority of the interactions with the substrate peptide (Figure 3a and 3b). When bound to hHAT1, the H4 peptide adopts a well-defined conformation that harbors a β -turn at its N-terminus which would otherwise be in an extended form in free H4 ⁴. Two conserved residues, Trp199 and Tyr225 of hHAT1 interact with Gly9 and Lys8 at the β -turn of the H4 peptide, respectively, to enforce the orientation of the substrate at the entrance of a groove ⁴. The C-terminus of the H4 peptide contains two positively charged residues, Arg17 and Arg19, and they make extensive hydrogen-bonds and charge-charge interactions with invariant residues of hHAT1, Glu64 and Asp62. The substrate binding groove of hHAT1 narrows where H4 Lys12 binds, allowing only Gly11 and Lys12 to be accommodated. The

strictly conserved Glu276 of hHAT1 also makes hydrogen-bond interactions with Gly11 and Lys12. Together, these specific interactions explain the preference for Lys12 of H4 as the acetylation target of hHAT1 ⁴.

A comparison of the structures of tGcn5 bound to different peptides reveals that 15 of the 19 residues of H3 are ordered in the structure rather than less than 10 ordered residues for the H4 and p53 peptide, which is consistent with the greater than about 1000-fold catalytic efficiency of Gcn5 for H3 over H4 and p53 ⁵⁹. These structures also reveal more ordered peptide and more extensive protein-peptide interaction in residues C-terminal to the reactive lysine over residues N-terminal to the reactive lysine, arguing that substrate interactions C-terminal to the target are more important for substrate binding by the Gcn5/PCAF HAT family. These interactions involve mostly hydrogen bonds to backbone residues and van der Waals interactions with side chains. Not surprisingly, the residues of tGcn5 that make contact to the peptide substrate are highly conserved within the Gcn5/PCAF family.

Both the HAT domains of hp300 ⁴⁴ and yEsa1 ⁶⁰ have been crystallized with bisubstrate inhibitors that have provided some information about histone substrate binding. hp300 has been cocrystallized with a Lys-CoA bisubstrate inhibitor that exhibits an IC₅₀ of about 400 nM, and is a more potent inhibitor than peptide-CoA inhibitors, consistent with its Theorell-Chance catalytic mechanism. The structure of the hp300/Lys-CoA complex reveals that the lysine portion of the bisubstrate inhibitor sits in a hydrophobic tunnel with the backbone of the lysine residue proximal to an electronegative groove on one side (Figure 3c). This electronegative groove connects the bisubstrate lysine binding pocket to another highly negatively charged pocket that is about 3–4 amino acid residues away. Correlating with this observation, an alignment of all known p300/CBP substrates reveals that they all contain a basic amino acid either three or four residues up- or down-stream of the target lysine. Mutagenesis of residues that form these pockets were also shown to increase the K_M for histone H3 substrate, further supporting the importance of this site for protein substrate binding by hp300 ⁴⁴. Taken together, the presented data is consistent with the more promiscuous substrate binding properties of the p300/CBP HAT family relative to the Hat1, Gcn5/PCAF, MYST and Rtt109 HATs.

The yEsa1 HAT domain was reported crystallized with an H4K16CoA bisubstrate inhibitor ⁶⁰. Although the yEsa1 HAT domain and the CoA, linker and lysine portion of the bisubstrate inhibitor is well resolved, the rest of the peptide portion of the bisubstrate inhibitor is not and is presumed to be disordered (Figure 3d). Like the hp300/Lys-CoA structure, the yEsa1/H4K16CoA complex shows a groove for peptide binding proximal to the lysine portion of the bisubstrate inhibitor for histone peptide substrate to bind, although this groove is more apolar than that of the hp300 HAT domain. This likely reflects the greater degree of substrate selectivity of the MYST proteins over the p300/CBP HAT family.

Taken together, the structures that are available of HAT domains bound to peptide substrates or CoA-peptide bisubstrate inhibitors, coupled with the fact that regions N- and C-terminal participate in protein substrate recognition by the Gcn5/PCAF HATs and are variable among the HAT families, suggest that these regions of the HAT proteins might be particularly

important for substrate-specific binding. This does draw some analogy to protein kinases, which also employ side chains from the N- and C-lobes to mediate substrate specific binding (Figure 3e). However, the C-lobe typically plays a more prominent role as a tethering surface to present the target cognate residue in the groove between the N- and C-lobes for substrate-specific phosphorylation ⁶¹.

Regulation by Protein Cofactors and Autoacetylation

Acetyltransferase activity can be regulated in at least 2 ways, through interaction with regulatory protein subunits and through autoacetylation. Many acetyltransferases function in the context of multiprotein complexes that modulate their catalytic activity and/or substrate specificity ^{62,63}. The Gcn5/PCAF HAT family are active on free histones (H3K14 and to a lesser extent H4K8/K16) or histone peptides but are much less active on histones in chromatin as recombinant proteins. Gcn5/PCAF are thought to function exclusively as multiprotein complexes in cells and their assembly in these complexes have been shown to facilitate both nucleosomal acetylation and to modulate intrinsic acetylation activity and substrate acetylation specificity ^{62,63}. For example, some yGcn5 and hPCAF complexes include SAGA/SLIK in yeast and TFTC/STAGA in human ⁶⁴. In the context of such complexes, Gcn5/PCAF broadens its specificity to several lysine residues in histones H3, H4 and H2B within chromatin. Similarly, many MYST family HATs are also assembled in multiprotein complexes in cells to modulate histone acetylation in chromatin ⁶⁵. For example, yEsa1 is part of NuA4 and piccolo/NuA4 complexes for chromatin acetylation. The molecular basis for how associated proteins within the HAT complexes modulate the HAT activity and specificity of their respective catalytic subunits is not understood.

Several HATs also require the binding of cofactors for catalytic activity. For example, the ySas2 member of the MYST HATs requires the binding of ySas4 and ySas5 for acetyltransferase activity ^{66,67}. yHat1 acetyltransferase activity is also elevated by about 10-fold when assembled into the NuB4 complex that also contains yHat2 and yHif1 ⁷. The acetyltransferase activity and substrate selectivity of yRtt109 are both modulated by its association with histone chaperones. Rtt109 exhibits very low activity on its own, but catalytic activity increases several hundred-fold when associated with either the yVps75 or yAsf1 histone chaperones ⁵², which also modulate substrate specificity; the yRtt109/yVps75 complex selectively acetylates K9 and K27 of the histone H3 tail while the yRtt109/yAsf1 complex preferentially acetylates H3K56 near the H3 core region ^{46,47,50,51,54}. The molecular basis for how the activity and substrate selectivity of these HATs are modulated by their associated protein subunits are not known, however, kinetic data suggests that associated subunits of yHat1 ⁸ and yRtt109 ^{52,54,68} contribute to catalysis, at least in part, by facilitating productive binding to the respective cognate histone substrates.

Another mode of regulating HAT activity that has recently come to light is through the autoacetylation of HAT proteins. In particular, three HAT families have now been shown to be regulated by autoacetylation, Rtt109, p300/CBP and MYST. hp300 contains a highly basic loop of about 40 residues embedded in the middle of its HAT domain that was shown to undergo multiple autoacetylation through an intermolecular mechanism ⁶⁹ to regulate the catalytic activity of the protein, with the hyperacetylated and hypoacetylated forms

correlated with the active and inactive forms of the enzyme, respectively ⁷⁰. In addition, recombinant hp300 protein with a cleaved “autoacetylation loop” has been shown to be constitutively active ⁷⁰. The reported hp300/Lys-CoA crystal structure does not contain an intact autoacetylation loop so the molecular basis for how it regulates hp300 activity is not clear. However, Cole and colleagues proposed a model of regulation whereby the highly basic autoacetylation loop sits in the electronegative substrate binding site to directly compete with substrate binding and is released from this site upon autoacetylation allowing for cognate substrate binding ⁴⁴ (Figure 4a).

yRtt109 was also known to be autoacetylated but the molecular basis for this was not known until structures of yRtt109 were reported revealing a buried acetyl-lysine residue (Lys290) (Figure 4b) ^{68,71,72}, the acetylation of which was shown to be required for full acetylation activity ⁷³. This acetyl-lysine forms a hydrogen bond with Asp 288, a mutationally sensitive yRtt109 residue. Mass spectrometry analysis also shows that this acetylation occurs *in vitro* as well as in yeast cells ⁶⁸. Reports addressing the functional importance of Lys290 modification have yielded contradictory results with regard to genotoxic agent sensitivity *in vivo* ^{68,71,72}. A recent study by Denu and coworkers has shown that Lys290 acetylation increases the overall k_{cat} and decreases the K_M for AcCoA binding ⁷³, although the molecular basis for this is unknown.

Several recent reports have demonstrated that the hMOF MYST protein is autoacetylated at an active site lysine (Lys274) and that this autoacetylation is required for cognate acetyltransferase activity both *in vitro* and *in vivo* ^{60,74,75}. The lysine that was found to be autoacetylated in hMOF is strictly conserved in MYST proteins and Yuan et al. also showed that acetylation of this lysine residue occurs in yEsa1 and ySas2 and is required for acetylation of these proteins *in vitro* and for *in vivo* function of these enzymes in cells ⁶⁰. The structure of the yEsa1/H4K16CoA complex with this active site lysine in the acetylated form (Lys262Ac) reveals that it sits in the active site pocket with the acetyl CO group making hydrogen bonds to Tyr289 and Ser291 and the methyl group making van der Waals interactions with phenylalanines 271 and 273; and the aliphatic region of Lys262Ac making van der Waals interactions with the lysine of the H4K16CoA bisubstrate inhibitor ⁶⁰ (Figure 4c). Each of the residues that contact the acetylated lysine is strictly or highly conserved among the entire MYST protein family further arguing for the importance of this acetylated lysine for MYST function. The structure of the hMOF HAT domain has also been determined in unliganded form ⁶⁰. This structure reveals that Lys274 exists in two states. In one state it is acetylated and in the same conformation as the analogous Lys262Ac of yEsa1, making conserved intra-molecular interactions. In a second state, Lys274 is unacetylated and flipped by about 90° out of the active site such that the lysine forms a hydrogen bond with Glu350, the general base for catalysis, and in a position that would block cognate lysine binding (Figure 4c). This is consistent with biochemical studies that show that an arginine mutation of this lysine in hMOF is defective in cognate substrate binding ⁶⁰. Taken together, these structural, biochemical and functional studies demonstrate that autoacetylation of MYST proteins in the active site is required for cognate substrate acetylation.

Regulation by protein cofactors and automodification is also a property that is shared by kinases. In particular, while the final active conformation and positioning of conserved loops

and residues is highly similar in protein kinases, they show significantly more diversity in their inactive conformations and their mode of conversion to the active form⁷⁶. Kinases can be kept in inactive conformations through interaction with other associated domains within the same polypeptide chain, through interaction with other inhibitor protein cofactors or more commonly through hypophosphorylation of one or more residues (either serine, threonine or tyrosine) in the activation segment and sometimes in other regions of the kinase.

Kinase activation through activation loop phosphorylation draws the most similarity to how acetylation of HATs activates their catalytic activities. An example of this is shown in the insulin receptor kinase (IRK)^{77,78} in which an unphosphorylated tyrosine residue (Tyr1162) in the activation loop forms a hydrogen-bond with a catalytic aspartic acid residue (Asp1132) as a substrate mimic and as a result blocks the binding of the cognate substrate and cofactor ATP. Upon Tyr1162 phosphorylation, the activation loop flips out into a conformation that is compatible for cognate substrate binding and phosphorylation (Figure 4d). In the case of kinases, phosphorylation occurs by the same enzyme (autophosphorylation) or by an upstream kinase, while in the case of HATs, only autoacetylation has been noted to date.

Like many HATs, kinases can also be activated by their association with other protein cofactors forming holoenzyme complexes. Examples of these are the activation of cyclin-dependent kinases (CDKs) by cyclin protein cofactors⁷⁹ and the differential regulation of AMPK by regulatory subunits⁸⁰. Although the molecular basis for such activation is understood for kinases, it is not clear for HATs.

Beyond Protein Acetylation – Acetylation Readers and Inhibitors

In this review article, we covered what was known to date about the structure, mechanism and mode of regulation of HAT enzymes and drew analogies, where relevant, to protein kinases. Together, this comparison shows a remarkable degree of similarity between these two enzyme families. Although not covered in great detail here, these enzymes also share similarities in their mode of downstream signaling through effector domains and their importance as drug targets for therapy.

In a landmark paper in 1999, Zhou and Aggarwal showed that the bromodomain of the Gcn5/PCAF HATs are specific binding modules for acetyl-lysine⁸¹, and since that time several other bromodomains from many chromatin regulators have been shown to be employed for sequence-specific acetyl-lysine recognition⁸², with about 55 bromodomains found exclusively in chromatin regulatory proteins, so far identified in the human genome. More recently, a tandem PHD (Plant Homology Domain) finger of the human DPF3b subunit of the BAF chromatin remodeling complex was also found to bind Lys14 of histone H3 in a sequence specific manner⁸³. There are over 100 PHD fingers in the human genome and they are C₄-H-C₃ zinc-containing modules that are typically associated with chromatin regulatory modules that had previously been shown to bind trimethylated lysine residues⁸⁴. Taken together, this leaves open the possibilities that other protein modules that recognize acetyllysine residues may exist.

Many phosphorylation events that are mediated by protein kinases are also known to be recognized by phosphorylation dependent binding modules. SH2 (Src Homology 2) domains, found in over 100 human proteins have been shown to bind specific phosphotyrosine sites^{85,86} and many FHA (Forkhead Associated) domains have been shown to recognize specific phosphothreonine residues^{87,88}. In addition, other domains such as PTB⁸⁹, MH2⁹⁰, BRCT⁹¹ and others⁹² have also been shown to bind specific phosphorylated residues. In the case of kinases, these relatively high affinity phosphodependent interactions play key roles in protein recruitment and phosphorylation dependent signal transduction cascades⁹². Whether the analogous acetyllysine mediated protein interactions play an equally important role in protein recruitment and “signal transduction” remains to be elucidated.

Given the association of many activating kinase mutations in human diseases and in cancer in particular, it is not surprising that kinases are one of the most attractive drug targets, after only G-protein coupled receptors⁹³. Indeed, a poster child for kinase inhibitors is the development of Imatinib (Gleevec) for the treatment of Bcr-Abl-driven chronic myelogenous leukemias (CMLs)^{94,95}, with the more recent development of Vemurafenib for the treatment of BRAF^{V600E}-mediated metastatic melanomas^{96,97}, and at least 10 other FDA approved kinase inhibitors currently in clinical trials⁹⁸.

The development of HAT inhibitors has not progressed nearly as far as kinase inhibitors. However, HATs are attractive targets for therapy because they mediate many different biological processes including cell cycle progression, dosage compensation, repair of DNA damage and hormone signaling; and aberrant HAT function has also been correlated with several human diseases including solid tumors, leukemias, inflammatory lung disease, viral infection, diabetes, fungal infection and drug addiction^{99,100}. For example, the p300/CBP HATs have properties of both oncoproteins and tumor suppressors. The hCBP HAT forms translocation products with hMLL (mixed lineage leukemia) and hMOZ (monocytic leukemia zinc-finger protein), another HAT, in a subset of acute myeloid leukemias. In addition, the hp300 HAT is mutated in a subset of colorectal and gastric cancers making it a bona fide tumor suppressor. The yRtt109 HAT was also reported to be required for the pathogenesis of *Candida albicans*, the most prevalent cause of hospital-acquired fungal infections¹⁰¹.

Inhibitors for HATs are at the early stages and readers who would like a more comprehensive review of such inhibitors are referred to other excellent review articles on this topic^{99,102,103}. In brief, the most potent and specific inhibitors have come from the development of peptide-based bisubstrate inhibitors in which coenzyme A is directly linked to the N ϵ nitrogen of the target lysine within histone peptides⁴⁵. Using this technology, Cole and coworkers developed submicromolar inhibitors that show selectivity between the Gcn5/PCAF and p300/CBP family of HATs⁴⁵. These bisubstrate inhibitors are a proof of principle that selective HAT inhibitors can be prepared, although peptide-based inhibitors do not generally have favorable pharmacokinetic properties⁹⁹. Several natural product HAT inhibitors have been reported including the hPCAF and hp300 inhibitors anacardic acid¹⁰⁴ and garcinol¹⁰⁵; and the p300/CBP-selective inhibitors curcumin¹⁰⁶, Epigallocatechin-3-gallate (EGCG)¹⁰⁷ and plumbagin¹⁰⁸. High throughput screens have also led to the

identification of a family of isothiazolones hPCAF and hp300 inhibitors¹⁰⁹. Virtual ligand screens have also led to the development of pyrazolone-containing small molecule hp300 HAT inhibitors¹¹⁰ and MYST HAT inhibitors¹¹¹. Most recently, a series of 6-alkylsalicylates were developed as MYST inhibitors¹¹².

Together, these inhibitors exhibit IC₅₀ values in the mid to low micromolar range and some show activity in cells^{99,102}, however, none of them have been rigorously biochemically evaluated *in vitro* or *in vivo* and their structures bound to HAT proteins have not been determined, so their molecular basis for action is still unclear. These studies motivate the need to develop more potent and selective HAT inhibitors that might be developed into therapeutic agents with efficacy similar to many current kinase inhibitors.

Conclusions

Over the last decade we have learned that HATs have many fascinating and unanticipated properties. Although they all carry out the same biochemical reaction, they form subfamilies with little to no sequence homology between them. Yet they each contain a structurally conserved central core region that templates the protein and AcCoA substrates together in an analogous way. Nonetheless they employ different catalytic strategies and distinct N- and C-terminal segments that mediate distinct biological activities. Many HAT enzymes are also regulated by autoacetylation and interact with other protein subunits that also modulate their activities. Through more recent proteomics studies we have also learned that acetylation extends beyond histones to mediate diverse biological process and that HATs and possibly other protein acetyltransferases are attractive drug targets due to their association with diseases such as cancers and metabolic disorders. Given the poor sequence conservation between different acetyltransferase families, it is likely that many acetyltransferases exist that are yet to be identified. Acetylation marks are also recognized by dedicated protein modules, including bromodomains, tandem PHD finger domains and possibly other domains to mediate downstream biological signals. Interestingly, many of the properties listed above for acetyltransferases are shared by the superfamily of kinase enzymes, although kinases appear to have greater sequence conservation and their catalytic strategies are more conserved. Like kinases, acetyltransferases also appear to employ dedicated protein modules that specifically recognize the respective posttranslational modification, and like kinases, these recognition events may mediate downstream signaling. Although acetyltransferases are evolutionarily more ancient than protein kinases, kinases have been studied much longer and are therefore better understood. This may change over the coming decades as acetyltransferases become better appreciated as a second family of master regulators of signal transduction pathways in biology and validated targets for therapy.

References

1. Krebs EG, Fischer EH. J Biol Chem. 1955; 216:113–120. [PubMed: 13252011]
2. Fischer EH, Krebs EG. J Biol Chem. 1955; 216:121–132. [PubMed: 13252012]
3. Walsh DA, Perkins JP, Krebs EG. Journal of Biological Chemistry. 1968; 243:3763–3765. [PubMed: 4298072]
4. Wu H, Moshkina N, Min J, Zeng H, Joshua J, Zhou MM, Plotnikov AN. Proc Natl Acad Sci U S A. 2012; 109:8925–8930. [PubMed: 22615379]

5. Brownell JE, Zhou J, Ranalli T, Kobayashi R, Edmondson DG, Roth SY, Allis CD. *Cell*. 1996; 84:843–851. [PubMed: 8601308]
6. Kleff S, Andrulis ED, Anderson CW, Sternglanz R. *J Biol Chem*. 1995; 270:24674–24677. [PubMed: 7559580]
7. Parthun MR, Widom J, Gottschling DE. *Cell*. 1996; 87:85–94. [PubMed: 8858151]
8. Ruiz-Garcia AB, Sendra R, Galiana M, Pamblanco M, Perez-Ortin JE, Tordera V. *J Biol Chem*. 1998; 273:12599–12605. [PubMed: 9575221]
9. Ai X, Parthun MR. *Mol Cell*. 2004; 14:195–205. [PubMed: 15099519]
10. Poveda A, Pamblanco M, Tafrov S, Tordera V, Sternglanz R, Sendra R. *J Biol Chem*. 2004; 279:16033–16043. [PubMed: 14761951]
11. Taunton J, Hassig CA, Schreiber SL. *Science*. 1996; 272:408–411. [PubMed: 8602529]
12. Allfrey VG, Faulkner R, Mirsky AE. *Proc Natl Acad Sci U S A*. 1964; 51:786–794. [PubMed: 14172992]
13. Hodawadekar SC, Marmorstein R. *Oncogene*. 2007; 26:5528–5540. [PubMed: 17694092]
14. Marmorstein R, Trievel RC. *Biochim Biophys Acta*. 2009; 1789:58–68. [PubMed: 18722564]
15. Selvi RB, Kundu TK. *Biotechnol J*. 2009; 4:375–390. [PubMed: 19296442]
16. Keppler BR, Archer TK. *Expert Opin Ther Targets*. 2008; 12:1457–1467. [PubMed: 18851700]
17. Glozak MA, Sengupta N, Zhang X, Seto E. *Gene*. 2005; 363:15–23. [PubMed: 16289629]
18. Gu W, Roeder RG. *Cell*. 1997; 90:595–606. [PubMed: 9288740]
19. Bannister AJ, Miska EA, Gorlich D, Kouzarides T. *Curr Biol*. 2000; 10:467–470. [PubMed: 10801418]
20. Zhang Q, Yao H, Vo N, Goodman RH. *Proc Natl Acad Sci U S A*. 2000; 97:14323–14328. [PubMed: 11114158]
21. Kim SC, Sprung R, Chen Y, Xu Y, Ball H, Pei J, Cheng T, Kho Y, Xiao H, Xiao L. *Molecular Cell*. 2006; 23:607–618. [PubMed: 16916647]
22. Choudhary C, Kumar C, Gnad F, Nielsen ML, Rehman M, Walther TC, Olsen JV, Mann M. *Science*. 2009; 325:834–840. [PubMed: 19608861]
23. Zhao S, Xu W, Jiang W, Yu W, Lin Y, Zhang T, Yao J, Zhou L, Zeng Y, Li H, Li Y, Shi J, An W, Hancock SM, He F, Qin L, Chin J, Yang P, Chen X, Lei Q, Xiong Y, Guan KL. *Science*. 2010; 327:1000–1004. [PubMed: 20167786]
24. Kim GW, Yang XJ. *Trends in Biochemical Sciences*. 2011; 36:211–220. [PubMed: 21075636]
25. Akella JS, Wloga D, Kim J, Starostina NG, Lyons-Abbott S, Morrisette NS, Dougan ST, Kipreos ET, Gaertig J. *Nature*. 2010; 467:218–222. [PubMed: 20829795]
26. Ivanov D, Schleiffer A, Eisenhaber F, Mechtler K, Haering CH, Nasmyth K. *Current Biology*. 2002; 12:323–328. [PubMed: 11864574]
27. Akella JS, Wloga D, Kim J, Starostina NG, Lyons-Abbott S, Morrisette NS, Dougan ST, Kipreos ET, Gaertig J. *Nature*. 2010; 467:218–222. [PubMed: 20829795]
28. Shida T, Cueva JG, Xu Z, Goodman MB, Nachury MV. *Proceedings of the National Academy of Sciences*. 2010; 107:21517–21522.
29. Hou F, Zou H. *Mol Biol Cell*. 2005; 16:3908–3918. [PubMed: 15958495]
30. Williams BC, Garrett-Engele CM, Li Z, Williams EV, Rosenman ED, Goldberg ML. *Current Biology*. 2003; 13:2025–2036. [PubMed: 14653991]
31. Kouzarides T. *EMBO J*. 2000; 19:1176–1179. [PubMed: 10716917]
32. Olsen JV, Blagoev B, Gnad F, Macek B, Kumar C, Mortensen P, Mann M. *Cell*. 2006; 127:635–648. [PubMed: 17081983]
33. Villen J, Beausoleil SA, Gerber SA, Gygi SP. *Proc Natl Acad Sci U S A*. 2007; 104:1488–1493. [PubMed: 17242355]
34. McInnes C, Fischer PM. *Curr Pharm Des*. 2005; 11:1845–1863. [PubMed: 15892678]
35. Nolen B, Taylor S, Ghosh G. *Molecular cell*. 2004; 15:661–675. [PubMed: 15350212]
36. Taylor SS, Kornev AP. *Trends Biochem Sci*. 2011; 36:65–77. [PubMed: 20971646]
37. Spencer TE, Jenster G, Burcin MM, Allis CD, Zhou JX, Mizzen CA, McKenna NJ, Onate SA, Tsai SY, Tsai MJ, Omalley BW. *Nature*. 1997; 389:194–198. [PubMed: 9296499]

38. Mizzen CA, Yang XJ, Kokubo T, Brownell JE, Bannister AJ, Owen-Hughes T, Workman J, Wang L, Berger SL, Kouzarides T, Nakatani Y, Allis CD. *Cell*. 1996; 87:1261–1270. [PubMed: 8980232]
39. Kawasaki H, Schiltz L, Chiu R, Itakura K, Taira K, Nakatani Y, Yokoyama KK. *Nature*. 2000; 405:195–200. [PubMed: 10821277]
40. Doi M, Hirayama J, Sassone-Corsi P. *Cell*. 2006; 125:497–508. [PubMed: 16678094]
41. Neuwald AF, Landsman D. *Trends Biochem Sci*. 1997; 22:154–155. [PubMed: 9175471]
42. Yan Y, Harper S, Speicher DW, Marmorstein R. *Nature Struct Biol*. 2002; 9:862–869. [PubMed: 12368900]
43. Berndsen CE, Albaugh BN, Tan S, Denu JM. *Biochemistry*. 2007; 46:623–629. [PubMed: 17223684]
44. Liu X, Wang L, Zhao K, Thompson PR, Hwang Y, Marmorstein R, Cole PA. *Nature*. 2008; 451:846–850. [PubMed: 18273021]
45. Lau OD, Kundu TK, Soccio RE, Ait-Si-Ali S, Khalil EM, Vassilev A, Wolffe AP, Nakatani Y, Roeder RG, Cole PA. *Mol Cell*. 2000; 5:589–595. [PubMed: 10882143]
46. Driscoll R, Hudson A, Jackson SP. *Science*. 2007; 315:649–652. [PubMed: 17272722]
47. Tsubota T, Berndsen CE, Erkmann JA, Smith CL, Yang L, Freitas MA, Denu JM, Kaufman PD. *Mol Cell*. 2007; 25:703–712. [PubMed: 17320445]
48. Albaugh BN, Kolonko EM, Denu JM. *Biochemistry*. 2010; 49:6375–6385. [PubMed: 20560668]
49. Berndsen CE, Tsubota T, Lindner SE, Lee S, Holton JM, Kaufman PD, Keck JL, Denu JM. *Nat Struct Mol Biol*. 2008; 15:948–956. [PubMed: 19172748]
50. Han J, Zhou H, Li Z, Xu RM, Zhang Z. *Journal of Biological Chemistry*. 2007; 282:28587–28596. [PubMed: 17690098]
51. Han J, Zhou H, Li Z, Xu RM, Zhang Z. *Journal of Biological Chemistry*. 2007; 282:14158–14164. [PubMed: 17369253]
52. Kolonko EM, Albaugh BN, Lindner SE, Chen Y, Satyshur KA, Arnold KM, Kaufman PD, Keck JL, Denu JM. *Proc Natl Acad Sci U S A*. 2010; 107:20275–20280. [PubMed: 21057107]
53. Su D, Hu Q, Zhou H, Thompson JR, Xu RM, Zhang Z, Mer G. *J Biol Chem*. 2011; 286:15625–15629. [PubMed: 21454705]
54. Tang Y, Holbert MA, Delgosaie N, Wurtele H, Guillemette B, Meeth K, Yuan H, Drogaris P, Lee EH, Durette C, Thibault P, Verreault A, Cole PA, Marmorstein R. *Structure*. 2011; 19:221–231. [PubMed: 21256037]
55. Yan Y, Barlev NA, Haley RH, Berger SL, Marmorstein R. *Mol Cell*. 2000; 6:1195–1205. [PubMed: 11106757]
56. Clements A, Marmorstein R. *Methods Enzymol*. 2003; 371:545–564. [PubMed: 14712728]
57. Poux AN, Marmorstein R. *Biochemistry*. 2003; 42:14366–14374. [PubMed: 14661947]
58. Rojas JR, Trievel RC, Zhou J, Mo Y, Li X, Berger SL, Allis CD, Marmorstein R. *Nature*. 1999; 401:93–98. [PubMed: 10485713]
59. Trievel RC, Li FY, Marmorstein R. *Anal Biochem*. 2000; 287:319–328. [PubMed: 11112280]
60. Yuan H, Rossetto D, Mellert H, Dang W, Srinivasan M, Johnson J, Hodawadekar S, Ding EC, Speicher K, Abshiru N, Perry R, Wu J, Yang C, Zheng YG, Speicher DW, Thibault P, Verreault A, Johnson FB, Berger SL, Sternglanz R, McMahon SB, Cote J, Marmorstein R. *EMBO J*. 2012; 31:58–70. [PubMed: 22020126]
61. Kobe B, Kampmann T, Forwood JK, Listwan P, Brinkworth RI. *Biochim Biophys Acta*. 2005; 1754:200–209. [PubMed: 16172032]
62. Lee KK, Workman JL. *Nat Rev Mol Cell Biol*. 2007; 8:284–295. [PubMed: 17380162]
63. Carrozza MJ, Utley RT, Workman JL, Cote J. *Trends in Genetics*. 2003; 19:321–329. [PubMed: 12801725]
64. Tanner KG, Trievel RC, Kuo MH, Howard RM, Berger SL, Allis CD, Marmorstein R, Denu JM. *J Biol Chem*. 1999; 274:18157–18160. [PubMed: 10373413]
65. Sapountzi V, Cote J. *Cellular and Molecular Life Sciences*. 2011; 68:1147–1156. [PubMed: 21132344]

66. Shia WJ, Osada S, Florens L, Swanson SK, Washburn MP, Workman JL. *Journal of Biological Chemistry*. 2005; 280:11987–11994. [PubMed: 15659401]
67. Sutton A, Shia WJ, Band D, Kaufman PD, Osada S, Workman JL, Sternglanz R. *The Journal of biological chemistry*. 2003; 278:16887–16892. [PubMed: 12626510]
68. Tang Y, Holbert MA, Wurtele H, Meeth K, Rocha W, Gharib M, Jiang E, Thibault P, Verreault A, Cole PA, Marmorstein R. *Nat Struct Mol Biol*. 2008; 15:738–745. [PubMed: 18568037]
69. Karanam B, Wang L, Wang D, Liu X, Marmorstein R, Cotter R, Cole PA. *Biochemistry*. 2007; 46:8207–8216. [PubMed: 17590016]
70. Thompson PR, Wang D, Wang L, Fulco M, Pediconi N, Zhang D, An W, Ge Q, Roeder RG, Wong J, Levrero M, Sartorelli V, Cotter RJ, Cole PA. *Nat Struct Mol Biol*. 2004; 11:308–315. [PubMed: 15004546]
71. Lin C, Yuan YA. *Structure*. 2008; 16:1503–1510. [PubMed: 18707894]
72. Stavropoulos P, Nagy V, Blobel G, Hoelz A. *Proc Natl Acad Sci U S A*. 2008; 105:12236–12241. [PubMed: 18719104]
73. Albaugh BN, Arnold KM, Lee S, Denu JM. *The Journal of biological chemistry*. 2011
74. Lu L, Li L, Lv X, Wu XS, Liu DP, Liang CC. *Cell Res*. 2011; 21:1182–1185. [PubMed: 21502975]
75. Sun B, Guo S, Tang Q, Li C, Zeng R, Xiong Z, Zhong C, Ding J. *Cell Res*. 2011; 21:1262–1266. [PubMed: 21691301]
76. Huse M, Kuriyan J. *Cell*. 2002; 109:275–282. [PubMed: 12015977]
77. Hubbard SR, Wei L, Ellis L, Hendrickson WA. *Nature*. 1994; 372:746–754. [PubMed: 7997262]
78. Hubbard SR. *EMBO J*. 1997; 16:5572–5581. [PubMed: 9312016]
79. Pavletich NP. *Journal of Molecular Biology*. 1999; 287:821–828. [PubMed: 10222191]
80. Carling D. *Trends in Biochemical Sciences*. 2004; 29:18–24. [PubMed: 14729328]
81. Dhalluin C, Carlson JE, Zeng L, He C, Aggarwal AK, Zhou MM. *Nature*. 1999; 399:491–496. [PubMed: 10365964]
82. Sanchez R, Zhou MM. *Curr Opin Drug Discov Devel*. 2009; 12:659–665.
83. Zeng L, Zhang Q, Li S, Plotnikov AN, Walsh MJ, Zhou MM. *Nature*. 2010; 466:258–262. [PubMed: 20613843]
84. Aasland R, Gibson TJ, Stewart AF. *Trends Biochem Sci*. 1995; 20:56–59. [PubMed: 7701562]
85. Filippakopoulos P, Muller S, Knapp S. *Curr Opin Struct Biol*. 2009; 19:643–649. [PubMed: 19926274]
86. Kasembeli MM, Xu X, Tweardy DJ. *Front Biosci*. 2009; 14:1010–1022.
87. Mahajan A, Yuan C, Lee H, Chen ES, Wu PY, Tsai MD. *Sci Signal*. 2008; 1:re12. [PubMed: 19109241]
88. Hammet A, Pike BL, McNees CJ, Conlan LA, Tennis N, Heierhorst J. *IUBMB Life*. 2003; 55:23–27. [PubMed: 12716058]
89. Kavanaugh WM, Turck CW, Williams LT. *Science*. 1995; 268:1177–1179. [PubMed: 7539155]
90. Wu JW, Hu M, Chai J, Seoane J, Huse M, Li C, Rigotti DJ, Kyin S, Muir TW, Fairman R, Massague J, Shi Y. *Mol Cell*. 2001; 8:1277–1289. [PubMed: 11779503]
91. Yu X, Chini CC, He M, Mer G, Chen J. *Science*. 2003; 302:639–642. [PubMed: 14576433]
92. Seet BT, Dikic I, Zhou MM, Pawson T. *Nature Reviews Molecular Cell Biology*. 2006; 7:473–483.
93. Cohen P. *Nat Rev Drug Discov*. 2002; 1:309–315. [PubMed: 12120282]
94. O'Brien SG, Guilhot F, Larson RA, Gathmann I, Baccarani M, Cervantes F, Cornelissen JJ, Fischer T, Hochhaus A, Hughes T, Lechner K, Nielsen JL, Rousselot P, Reiffers J, Saglio G, Shepherd J, Simonsson B, Gratwohl A, Goldman JM, Kantarjian H, Taylor K, Verhoef G, Bolton AE, Capdeville R, Druker BJ. *N Engl J Med*. 2003; 348:994–1004. [PubMed: 12637609]
95. Druker BJ, Guilhot F, O'Brien SG, Gathmann I, Kantarjian H, Gattermann N, Deininger MW, Silver RT, Goldman JM, Stone RM, Cervantes F, Hochhaus A, Powell BL, Gabrilove JL, Rousselot P, Reiffers J, Cornelissen JJ, Hughes T, Agis H, Fischer T, Verhoef G, Shepherd J,

- Saglio G, Gratwohl A, Nielsen JL, Radich JP, Simonsson B, Taylor K, Baccarani M, So C, Letvak L, Larson RA. *N Engl J Med*. 2006; 355:2408–2417. [PubMed: 17151364]
96. Tsai J, Lee JT, Wang W, Zhang J, Cho H, Mamo S, Bremer R, Gillette S, Kong J, Haass NK, Sproesser K, Li L, Smalley KS, Fong D, Zhu YL, Marimuthu A, Nguyen H, Lam B, Liu J, Cheung I, Rice J, Suzuki Y, Luu C, Settachatgul C, Shellooe R, Cantwell J, Kim SH, Schlessinger J, Zhang KY, West BL, Powell B, Habets G, Zhang C, Ibrahim PN, Hirth P, Artis DR, Herlyn M, Bollag G. *Proceedings of the National Academy of Sciences of the United States of America*. 2008; 105:3041–3046. [PubMed: 18287029]
97. Chapman PB, Hauschild A, Robert C, Haanen JB, Ascierto P, Larkin J, Dummer R, Garbe C, Testori A, Maio M, Hogg D, Lorigan P, Lebbe C, Jouary T, Schadendorf D, Ribas A, O'Day SJ, Sosman JA, Kirkwood JM, Eggermont AM, Dreno B, Nolop K, Li J, Nelson B, Hou J, Lee RJ, Flaherty KT, McArthur GA. *N Engl J Med*. 2011; 364:2507–2516. [PubMed: 21639808]
98. Dar AC, Shokat KM. *Annu Rev Biochem*. 2011; 80:769–795. [PubMed: 21548788]
99. Heery DM, Fischer PM. *Drug Discov Today*. 2007; 12:88–99. [PubMed: 17198977]
100. Renthall W, Nestler EJ. *Seminars in cell & developmental biology*. 2009; 20:387–394. [PubMed: 19560043]
101. Lopes da Rosa J, Boyartchuk VL, Zhu LJ, Kaufman PD. *Proc Natl Acad Sci U S A*. 2010; 107:1594–1599. [PubMed: 20080646]
102. Furdas SD, Kannan S, Sippl W, Jung M. *Arch Pharm (Weinheim)*. 2012; 345:7–21. [PubMed: 22234972]
103. Yang XJ, Seto E. *Oncogene*. 2007; 26:5310–5318. [PubMed: 17694074]
104. Sung B, Pandey MK, Ahn KS, Yi T, Chaturvedi MM, Liu M, Aggarwal BB. *Blood*. 2008; 111:4880–4891. [PubMed: 18349320]
105. Balasubramanyam K, Altaf M, Varier RA, Swaminathan V, Ravindran A, Sadhale PP, Kundu TK. *J Biol Chem*. 2004; 279:33716–33726. [PubMed: 15155757]
106. Balasubramanyam K, Varier RA, Altaf M, Swaminathan V, Siddappa NB, Ranga U, Kundu TK. *J Biol Chem*. 2004; 279:51163–51171. [PubMed: 15383533]
107. Choi KC, Jung MG, Lee YH, Yoon JC, Kwon SH, Kang HB, Kim MJ, Cha JH, Kim YJ, Jun WJ, Lee JM, Yoon HG. *Cancer Res*. 2009; 69:583–592. [PubMed: 19147572]
108. Ravindra KC, Selvi BR, Arif M, Reddy BA, Thanuja GR, Agrawal S, Pradhan SK, Nagashayana N, Dasgupta D, Kundu TK. *The Journal of biological chemistry*. 2009; 284:24453–24464. [PubMed: 19570987]
109. Gorsuch S, Bavetsias V, Rowlands MG, Aherne GW, Workman P, Jarman M, McDonald E. *Bioorg Med Chem*. 2009; 17:467–474. [PubMed: 19101154]
110. Bowers EM, Yan G, Mukherjee C, Orry A, Wang L, Holbert MA, Crump NT, Hazzalin CA, Liszczak G, Yuan H, Larocca C, Saldanha SA, Abagyan R, Sun Y, Meyers DJ, Marmorstein R, Mahadevan LC, Alani RM, Cole PA. *Chemistry & Biology*. 2010; 17:471–482. [PubMed: 20534345]
111. Wu J, Wang J, Li M, Yang Y, Wang B, Zheng YG. *Bioorg Chem*. 2011; 39:53–58. [PubMed: 21186043]
112. Ghizzoni M, Wu J, Gao T, Haisma HJ, Dekker FJ, Zheng YG. *European Journal of Medicinal Chemistry*. 2012; 47:337–344. [PubMed: 22100137]
113. Schrodinger, LLC. 2010.

Figure 1. Overall Structure of HAT Proteins and Kinases

Representative members of the 5 HAT families and kinase are illustrated as cartoons. The structurally conserved core region of HATs are colored in blue and flanking N- and C-terminal regions are colored in aqua. The cofactor is shown in stick figure in yellow (**a**) Yeast HAT1 (yHAT1, KAT1)/Ac-CoA (PDB 1BOB), (**b**) *Tetrahymena* Gcn5 (tGcn5, KAT2)/CoA/histone H3 peptide (PDB 1PUA), (**c**) Yeast Esa1 (yEsa1, KAT5)/H4K16-CoA (PDB 3TO6) (residues flanking K16 in the peptide are disordered in the structure), (**d**) Human p300 (hp300, KAT3B)/Lys-CoA (PDB 3BIY), (**e**) Yeast Rtt109 (yRtt109, KAT11)/CoA (PDB 3D35) and (**f**) Human insulin receptor tyrosine kinase (with the N-lobe

in green and C-lobe in aqua) in complex with the non-hydrolyzable ATP analogue, AMP-PNP (stick) and substrate-peptide (purple) (PDB 1IR3)

Figure 2. Catalytic Mechanism of HAT Proteins and Kinases

Active sites of representative members of the HAT families (same as Figure 1) and cAMP kinase are illustrated highlighting the relevant side chains on a backbone cartoon of the active site. The cofactors are shown in yellow sticks and catalytic residues in green sticks. **(a)** *tGcn5/CoA/histone H3*. Key catalytic residues are labeled and hydrophobic residues of the active site that likely raise the pKa of Glu173 are shown in white sticks. The numbering is for *yGcn5*. **(b)** *yEsa1/H4K16CoA*. Key catalytic residues are labeled and hydrophobic residues of the active site that likely raise the pKa of Glu338 are shown in white sticks. Residues flanking Lys16 in the histone H4 peptide are disordered in the structure. **(c)** *hp300/Lys-CoA*. Residues demonstrated to play catalytic roles are labeled with other potential catalytic residues shown in dark green. **(d)** *yRtt109/CoA* with potential catalytic residues in

the corresponding position of hp300 shown in sticks. **(e)** hHAT1/AcCoA/histone H4. The three general base candidate residues are represented as green sticks. **(f)** cAMP-dependent protein kinase in complex with ATP highlighting the P-loop (blue), A-loop (red), catalytic loop (green) and hinge region (orange) (PDB 1ATP).

Figure 3. Substrate Binding by HAT Proteins and kinases

Close-up electrostatic surface view of HAT domains (same as in Figure 1) and cAMP kinase structures with peptide substrates or CoA-peptide bisubstrate inhibitors shown in ball and stick model in cpk color. Electrostatic potential is color coded with red representing negative charge (-4 kT/e) and blue representing positive charge (4 kT/e). (a) Structure of tGcn5 bound to CoA (yellow) and a 19-residue histone H3 peptide (white) centered around K14 (purple). (b) Structure of hHAT1 bound to AcCoA (yellow) and a histone H4 peptide (1–20) (white) centered around K12 (purple). (c) Structure of the hp300/LysCoA complex. The LysCoA inhibitor is shown in yellow for the CoA moiety and purple for the lysine moiety. (d) Structure of the yEsa1/H4K16CoA complex (only the lysine residue of the H4K16 peptide component of the bisubstrate inhibitor is ordered in the crystal structure and shown).

in purple). **(e)** Structure of cAMP-dependent protein kinase in complex with a substrate peptide showing the serine residue in purple (PDB 1JBP).

Figure 4. Auto Regulation of HAT Proteins and Kinases

Close-up views of the autoacetylation site of HAT proteins and kinases. **(a)** Model for hp300 activation by autoacetylation (electrostatic potential maps are contoured at ± 4.0 kT/e, the same as in Figure 3), **(b)** The Lys290 autoacetylation site of yRtt109 highlighting the environment around acetylated Lys290. **(c)** Structure of hMOF highlighting the environment around acetylated Lys274 (green stick). The loop harboring Lys274 is highlighted in red for the acetylated conformation and blue for the unacetylated conformation. Superposition of the yEsa1/H4K16CoA structure with hMOF indicates that the unacetylated conformation would clash with binding of the cognate substrate lysine (as represented by the lysine of the H4K16CoA bisubstrate inhibitor shown in purple). **(d)** Structures of insulin receptor tyrosine kinase showing the autoinhibited conformation (left, PDB 1IRK) and autoactivated conformation (right, PDB 1IR3). The activation loop harboring unphosphorylated Tyr1162 adopts an autoinhibited conformation (blue) that would clash with the binding of both the cofactor (yellow stick) and cognate peptide substrate (purple).