

Lysine Acetylation Targets Protein Complexes and Co-Regulates Major Cellular Functions

Chunaram Choudhary,^{1,2} Chanchal Kumar,¹ Florian Gnäd,¹ Michael L. Nielsen,^{1,2} Michael Rehman,³ Tobias C. Walther,³ Jesper V. Olsen,^{1,2} Matthias Mann^{1,2*}

Lysine acetylation is a reversible posttranslational modification of proteins and plays a key role in regulating gene expression. Technological limitations have so far prevented a global analysis of lysine acetylation's cellular roles. We used high-resolution mass spectrometry to identify 3600 lysine acetylation sites on 1750 proteins and quantified acetylation changes in response to the deacetylase inhibitors suberoylanilide hydroxamic acid and MS-275. Lysine acetylation preferentially targets large macromolecular complexes involved in diverse cellular processes, such as chromatin remodeling, cell cycle, splicing, nuclear transport, and actin nucleation. Acetylation impaired phosphorylation-dependent interactions of 14-3-3 and regulated the yeast cyclin-dependent kinase Cdc28. Our data demonstrate that the regulatory scope of lysine acetylation is broad and comparable with that of other major posttranslational modifications.

Acetylation of lysine is a reversible posttranslational modification (PTM), which neutralizes the positive charge of this amino acid, changing protein function in diverse ways (1, 2). It has a key role in the regulation of gene expression through the modification of core histone tails by histone acetyltransferases (HATs) or histone deacetylases (HDACs) (3, 4). Some modified lysines specifically bind bromodomain-containing proteins, which are part of large complexes that modulate chromatin architecture. Lysine acetylation is also important for p53 functions and interactions and for microtubule stabilization (5). There are many individual reports of acetylation sites on proteins involved in diverse biological processes, suggesting that lysine acetylation has broad regulatory functions in addition to the few that are actively studied.

HDACs [or lysine deacetylases (KDACs)] are important drug targets in cancer and neurodegenerative diseases, such as Parkinson's and Alzheimer's diseases, and there are more than 80 clinical trials currently underway (6–9). Two KDAC inhibitors, suberoylanilide hydroxamic acid (SAHA) and valproic acid, are in clinical use. Furthermore, KDAC inhibitors are potent reprogramming agents for the generation of induced pluripotent stem cells (10). However, the specific mode of action of KDAC inhibitors has remained enigmatic and may involve acetylation on histones and other proteins.

Despite great biological and clinical interest in lysine acetylation, our knowledge of in vivo acetylation sites is limited. However, advances in quantitative mass spectrometry (MS)-based proteomics (11–13) allow PTMs to be studied at a proteomic scale (14, 15). In the case of phosphorylation, affinity purification of modified peptides makes it possible to quantify phosphorylation at thousands of sites (16, 17). Peptides containing acetylated lysines can also be partially enriched by an antibody directed against this modification. This strategy was used to identify 388 lysine acetylation sites, mainly in mitochondria, a cellular compartment where this modification had hardly been described before (18). We reasoned that high-resolution high-accuracy MS, when coupled to an efficient quantitation strategy, should enable a global view of the “lysine acetylome” and its changes upon KDAC inhibition.

Sequencing and quantifying the acetylome.

We enriched acetylated peptides from trypsin-digested whole-cell lysates of MV4-11 cells, a human acute myeloid leukemia cell line with an antibody against acetyl-lysine (19). We identified approximately 1000 acetylation sites by means of high-resolution high-accuracy MS (20). However, biologically important acetylation sites of low abundance remained undetected because of a large background of nonacetylated peptides. We therefore further separated peptides from immunoaffinity purifications by means of isoelectric focusing into 12 fractions (21). We used stable-isotope labeling with amino acids in cell culture (SILAC) (22, 23) so as to determine global acetylation changes in response to two KDAC inhibitors, SAHA and MS-275. Lastly, to assess reproducibility and depth of acetylome coverage, we performed similar analyses in two additional human cell lines of epithelial (A549) and lymphoid origin (Jurkat) (fig. S1A) (24). The resulting

raw data were processed with the MaxQuant suite of algorithms (25).

Overall mass accuracy was in the parts-per-billion range; high resolution was achieved for all peptides (fig. S1B), which facilitated accurate quantitation. We detected more than 3600 acetylation sites on 1750 proteins (table S1) at an overall false discovery rate (FDR) for peptides of less than 1%. The FDR for acetylated lysine-containing peptides, as opposed to all peptides, was even lower (between 0.1 and 0.3%) because SILAC determines the number of arginines and lysines directly, greatly aiding specific identification. We precisely pinpointed the site of modification for more than 95% of the peptides (table S1).

In a separate experiment without affinity enrichment, the number of acetylation sites was 60-fold lower (0.058% as compared with 3.44%). Percentage of acetylation sites C-terminal to the peptide—an in vitro artifact—was 72% as compared with 2.5% in the enriched samples. We used D₃C₁-labeled acetic acid in the sample preparation so as to independently show that C-terminal lysine acetylation was mainly caused by this chemical. We selected nine proteins for which acetylation was not previously described and one known acetylated protein and independently confirmed their acetylation by means of immunostaining (Fig. 1A). Proteins from a human cervical cancer cell line (Hela) or a human osteosarcoma cell line (U2OS) that were expressed as fusions with green fluorescent protein (GFP) under the endogenous promoter (26) were immunoprecipitated, and all showed clear lysine-acetylation signals (Fig. 1A). CBX3 was immunoprecipitated from transiently transfected human embryonic kidney 293 cells and also showed clear acetylation. Together, this demonstrates that our large-scale acetylome is of in vivo origin.

The in vivo acetylome. We first assessed the coverage of our data set for histone acetylation sites and found that we detected all sites that we were able to extract from the literature (table S1). The known in vivo acetylation sites of proteins of low abundance, such as the tumor suppressor p53, Ku70, nuclear factor κB subunit RelA, and protooncogene protein c-Myc, were also present. A large number of acetylation sites on proteins exclusively annotated as mitochondrial (381 different sites) (18), but these did not constitute the largest category. We compared the acetylomes of the three different cell types (Fig. 1B). For any one cell line, 60 to 80% of the acetylated proteins and 60 to 75% of the acetylated sites were also found in the other two. This indicates that our experimental methodology is accurate and reproducible and that acetylation patterns are similar in cells derived from different tissue types. Our experiment covers the majority of known in vivo sites and expands by at least sixfold the number of such sites reported in UniProt (27). The large overlap (fig. S2), even on proteins of low abundance, also implies sub-

¹Proteomics and Signal Transduction, Max Planck Institute for Biochemistry, Martinsried, Germany. ²The Novo Nordisk Foundation Center for Protein Research, Faculty of Health Sciences, University of Copenhagen, Blegdamsvej 3, DK-2200 Copenhagen, Denmark. ³Organelle Architecture and Dynamics, Max Planck Institute for Biochemistry, 82152 Martinsried, Germany.

*To whom correspondence should be addressed. E-mail: mmann@biochem.mpg.de

stantial coverage of the core acetylome. However, it is likely that many more sites are still to be discovered because they may only be acetylated in specific cell types or at particular developmental or cell-cycle stages.

Lysine acetylation is an ancient PTM that is conserved from prokaryotes to humans. We evaluated phylogenetic conservation of lysine-acetylated proteins and compared it with that of the entire proteome (28). Lysine-acetylated pro-

teins are significantly more conserved across the evolutionary tree (fig. S3), indicating additional selective pressure. Acetylated proteins were as conserved as phosphorylated proteins across vertebrates. For more distantly related species such as yeast (fig. S3A) and especially in prokaryotes, the acetylome was much more conserved than the phosphoproteome ($P < 10^{-20}$) (fig. S3B).

Lysine acetyltransferases (KATs) and KDACs are thought to be predominantly nuclear or mitochondrial, and indeed we did find a large number of acetylated proteins in these compartments. However, proteins with exclusive cytoplasmic annotation were also highly represented in the acetylome (Fig. 2A). HDAC6 is the only known cytoplasmic KDAC, but our results make it unlikely that it is the sole deacetylase with cytoplasmic activity.

Phosphorylation mainly occurs in unstructured regions of proteins, such as hinges and loops (28, 29). In contrast, and in concordance with an earlier report (18), acetylation sites were frequently located in regions with ordered secondary structure. Furthermore, compared with all lysines, acetylated lysines were significantly enriched in structured regions and depleted in unstructured regions (Fig. 2B).

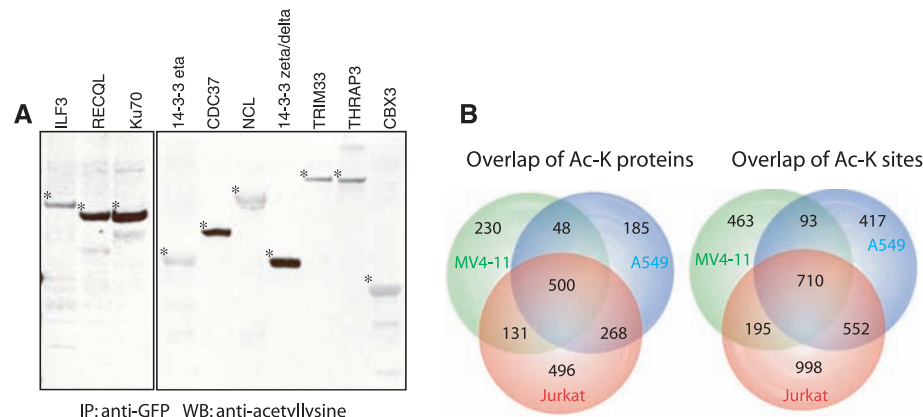
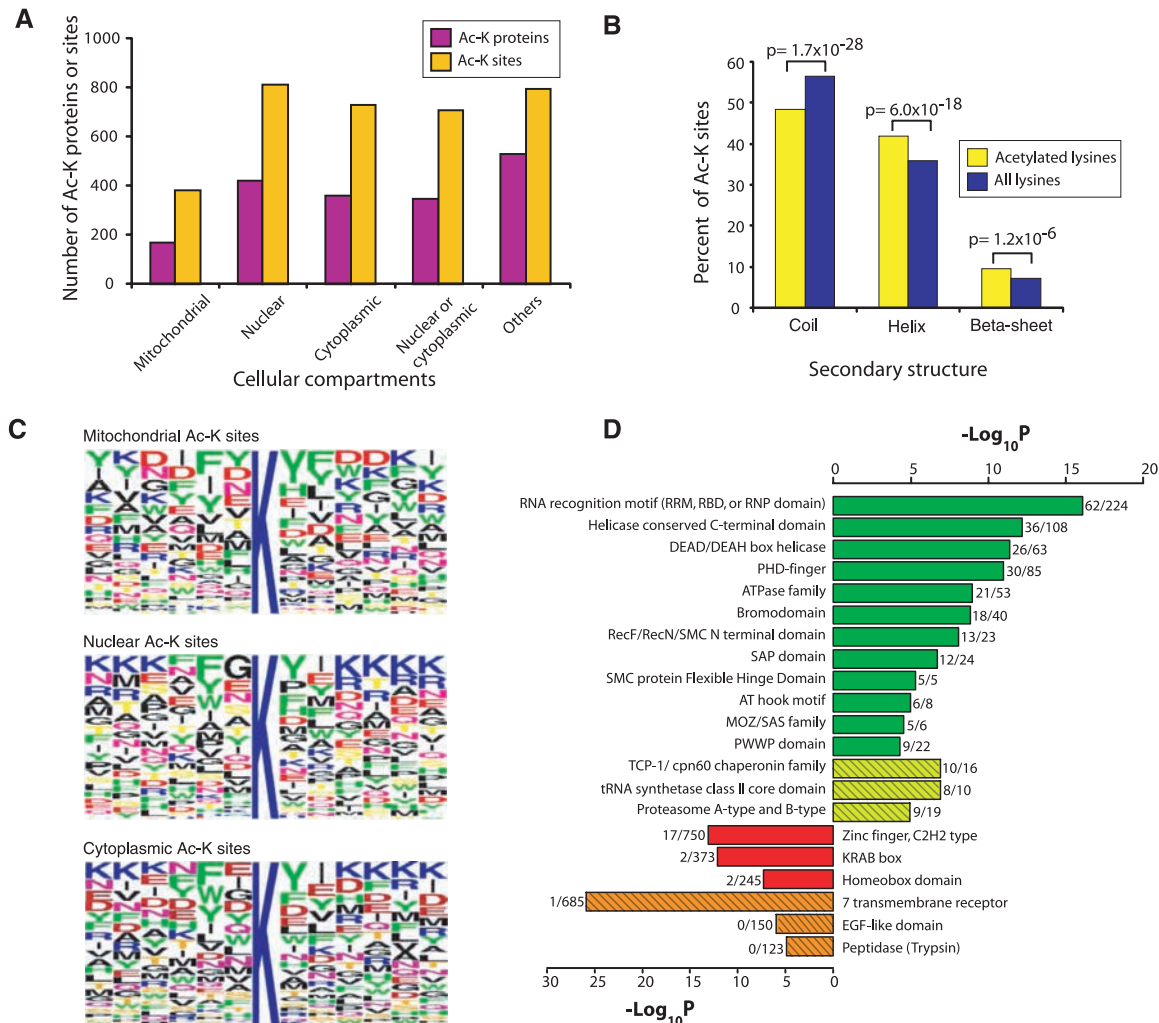


Fig. 1. Overview of in vivo acetylome analysis. **(A)** Independent validation of lysine acetylation of proteins. **(B)** Overlap of acetylated proteins and sites in three different cell lines. Ten different proteins from the acetylome data set were immunoprecipitated from GFP-tagged BAC transgenic cell lines and stained with antibody to acetyl-lysine. The bands marked with an asterisk indicate acetylated proteins.

Fig. 2. Properties of acetylated proteins and sites. **(A)** Cellular distribution of acetylated proteins and sites. Proteins were assigned based on exclusive Gene Ontology (GO) annotations. **(B)** Distribution of all lysines and acetylated lysines in structured and unstructured regions of the proteins. **(C)** Sequence logo plots represent normalized amino acid frequencies for ± 6 amino acids from the lysine acetylation site. **(D)** Domain architecture of acetylated proteins. The green bars indicate Pfam protein families and domains that are significantly overrepresented, and the red bars indicate underrepresented domains in the acetylome as compared with those in the entire proteome. The light green and orange striped bars represent cytoplasmic domains.



We analyzed local sequence context around the acetylation sites. Amino acids with a bulky side chain—mainly tyrosine and phenylalanine—were enriched in the −2 and +1 positions. This observation agrees with the frequent occurrence of acetylated lysines in ordered regions of proteins. Positively charged amino acids were almost completely excluded from the −1 position. Nuclear and cytoplasmic acetylation motifs are similar but different from mitochondrial motifs (Fig. 2C). For example, lysines were the preferred amino acids beyond the +2 and −3 positions for both nuclear and cytosolic proteins. In mitochondrial proteins, there was a preference for tyrosine (and histidine to a lesser degree) in the +1 position, and hydrophobic amino acids were enriched. Glycine at −1 was observed in vivo here and has been demonstrated on specific proteins (the GK motif) (1, 2).

Cellular localization and function of proteins is often dictated by their domains. Pfam domains (30) associated with nuclear functions were most overrepresented in our acetylome, including RNA-binding motifs, various helicases, the PWWP domain, the PHD finger, and bromodomains ($P < 10^{-5}$) (Fig. 2D). Among the most underrepresented domains were those associated with membranes or extracellular space and peptidases.

Acetylation targets macromolecular complexes. We were intrigued by the presence of a large number of acetylated proteins in large macromolecular complexes and classified acetylated proteins according to known chromatin-associated processes (fig. S4). Known nuclear lysine acetyltransferase complexes (3) were themselves highly acetylated on multiple subunits (fig. S4A). Three subunits were known to be acetylated in these nine KAT complexes; our analysis detected more than 40. Autoacetylation of p300 is a well-described example of KAT acetylation that regulates its own enzymatic activity (31). Our data demonstrate that extensive acetylation of KAT complex subunits is a general property. ATAC is a newly described KAT complex in *Drosophila melanogaster* described after our data were acquired (32). Of the human homologs of ATAC subunits, seven were acetylated. SIN3A, the main deacetylase complex, contains at least eight subunits modified on 18 distinct sites (fig. S4B).

Almost no acetylation sites are currently known on chromatin-remodeling complexes, but we found that SWI/SNF, NURD, INO80, and NURF are heavily acetylated, usually on multiple lysines (fig. S4C). Two complexes involved in regulating transcription, FACT (facilitates chromatin transcription) and TAF [TATA-box binder protein (TBP)-associated factor], were also targets of acetylation (fig. S4D).

Histones and p53 are regulated by several different PTMs, suggesting extensive crosstalk between different PTMs. Major methyltransferase complexes that contain catalytic subunits encoded by the mixed lineage leukemia (MLL) genes were extensively modified by acetylation, suggesting crosstalk between KATs and methyltransferases directly at the level of the modifying enzymes

(fig. S4E). Histone demethylases such as JARIDs (Jumonji/ARID domain-containing proteins) were also acetylated, further substantiating the regulatory link. Lastly, a large number of nuclear ubiquitin-modifying enzymes were also subject to this modification (fig. S4F). These data suggest a direct and intricate crosstalk between different enzymes involved in modulating chromatin-associated functions via PTMs.

To investigate acetylation-targeting of many macromolecular complexes in an unbiased way, we analyzed the complete acetylome data set using Comprehensive Resource of Mammalian protein complexes (CORUM), a database of manually curated and validated mammalian protein complexes (33). A total of 39 nuclear and five cytosolic complexes were enriched for acetylated proteins at $P < 0.05$ (Fisher's exact test) (table S2), including all of the complexes described above.

To quantify protein-interaction properties of the acetylome, we used the Search Tool for the Retrieval of Interacting Genes/Proteins (STRING) database of physical and functional interactions (34). Compared with similarly sized randomly selected protein data sets, the acetylome has significantly higher network connectivity ($P < 10^{-6}$) with six interactions per node as compared with less than three for random data sets (fig. S5). This further corroborates the notion of high connectivity among acetylated proteins. In the global STRING-generated protein-protein network, several complexes and cellular functions formed prominent, tightly connected clusters as assessed by means of molecular complex detection (fig. S6A) (35). In addition to the physical complexes found manually and by CORUM,

cellular processes such as RNA splicing, nuclear transport, protein folding, and cytoskeletal regulation were highlighted (fig. S6B).

Cellular processes regulated by acetylation. With the exception of histone modification and DNA repair, in which acetylation has well-established roles, there is little evidence for a broad role of acetylation in other nuclear processes. Our data reveal that a large number of acetylation sites are present on proteins involved in all major nuclear processes, such as splicing, cell cycle, chromatin remodeling, DNA replication, transcription, and nuclear transport, which strongly suggests that these processes may be influenced by this modification (Table 1). Previously, only a small number of well-studied proteins were known to be acetylated in these processes. For example, there are only three known acetylation sites for proteins implicated in splicing, whereas we found 130 acetylation sites on proteins involved in this process. These acetylated proteins are connected in a dense protein-protein interaction network (Fig. 3A).

Acetylation of a few proteins such as p53, Ku70, FEN1, and WRN is known to be important in the repair of damaged DNA. Acetylation of p53 regulates its stability through crosstalk with the ubiquitination machinery, modulates interactions with TAF1, and regulates its transcriptional activity (5). We confirmed acetylation of these proteins and found further sites, which may also be functional. For example, Ku70 and Ku80, which form heterodimers and regulate the activity of DNA-dependent protein kinase (DNAPK), a key kinase involved in DNA damage repair, are both acetylated. Acetylation of Ku70 is already

Table 1. Acetylation regulates diverse protein classes and biological categories. Acetylated proteins were grouped into major cellular functional categories or into protein classes on the basis of their known functions or function predicted by functional domains.

Cellular functional categories and protein classes	Number of acetylated proteins	Number of acetylation sites
DNA replication	52	98
DNA damage and repair	72	167
Chromatin remodeling	26	46
Cell cycle	132	243
RNA transcription	31	71
RNA splicing	109	206
Nuclear hormone signaling	9	22
Nuclear transport	17	41
Cytoskeleton reorganization	50	137
Nucleotide exchange factors	55	92
Endocytosis and vesicular trafficking	39	62
DNA/RNA helicases	46	105
Ubiquitin ligases and deubiquitylases	46	70
Protein kinases	47	71
Acetyltransferases and deacetylases	21	61
Methyltransferases and demethylases	12	34
Transcription factors	29	40
Histones	15	61
Adaptor proteins	14	40
Chaperones	40	127
Ribosomal proteins	75	136

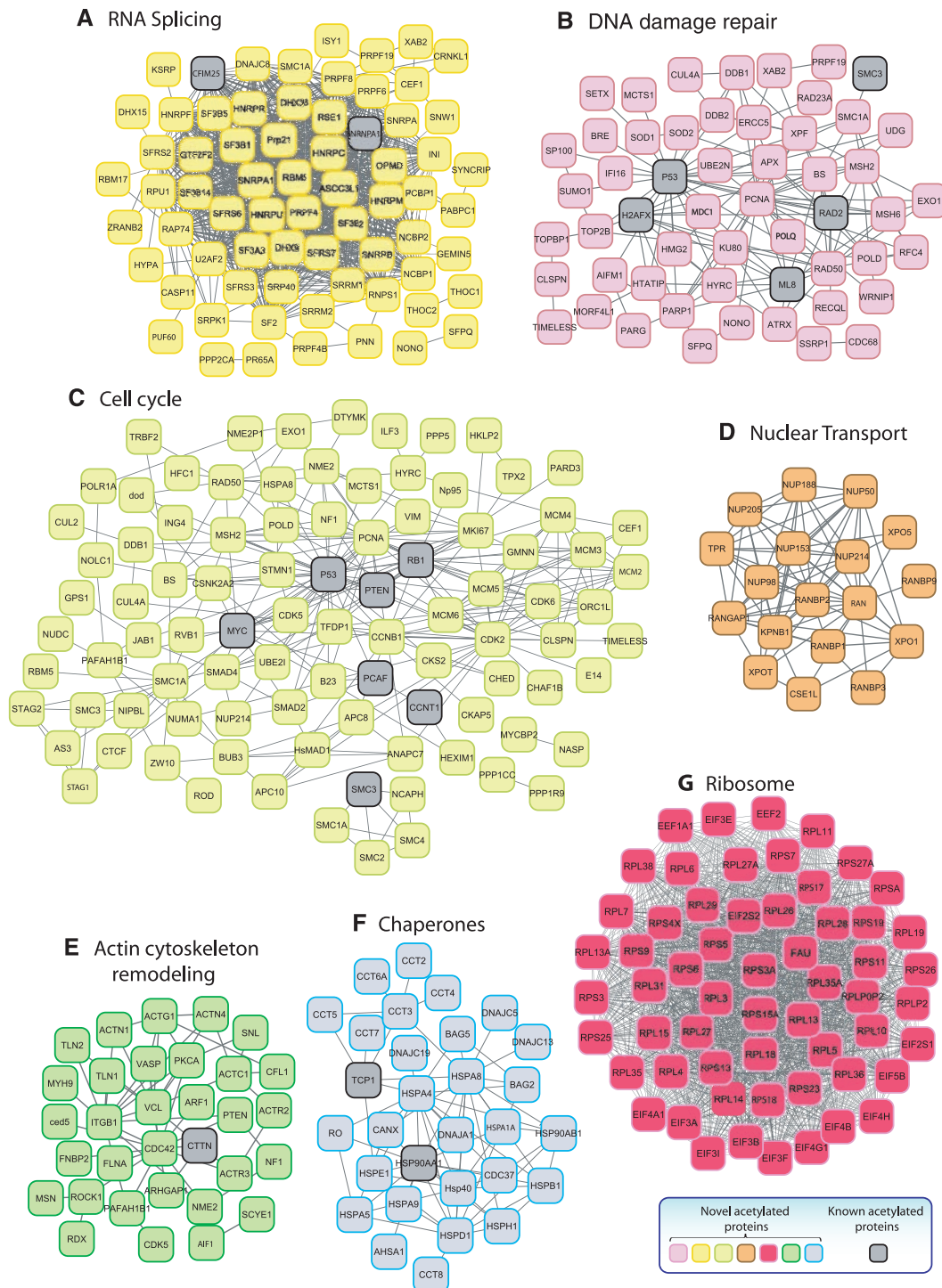
known to be important in this process. Many other proteins that are essential for DNA damage repair, such as MDC1, DDB1/2, RAD50, PCNA, WRNIP1, MSH2, BLM, and RAD54L, are acetylated (Fig. 3B). The human genome encodes six phosphoinositide-3-kinase-related protein kinases (PIKKs) that contain a conserved C-terminal FAT [FKBP12-rapamycin-associated protein (FRAP), ataxia telangiectasia mutated (ATM), transformation/transcription domain-associated protein (TRRAP)], a kinase domain, and FATC (FAT C terminus) domain structure:

ATM, ATR, SGK1, DNAPK, TRRAP, and FRAP1. PIKKs play key roles in DNA damage repair, and we found acetylation sites in SGK1, DNAPK, TRRAP, and FRAP1. Acetylation of ATM has been reported (36). The FATC domains of these kinases mediate an interaction with Tip60, an acetyltransferase that has a critical role in DNA damage repair and, in the case of ATM and DNAPK, regulates their kinase activity (36, 37). Thus, acetylation may represent a common regulatory mechanism of activation for FATC domain-containing PIKKs.

Acetylation is a widespread modification for nuclear ubiquitin ligases and deubiquitylases including UHRF1, HUWE, BRE1A, and BRE1B, which have roles in the modification of histones and chromatin-bound proteins (table S1). Similarly, HAUSP, USP22, and USP11 deubiquitylases, which have roles in chromatin biology, and Ubr37 and USP14, the two major proteasome-bound deubiquitylases, are acetylated.

The p53 interaction network controls its functions with several mechanisms. We discovered that not only p53 but several proteins that are critical for

Fig. 3. Acetylation-modulated functional networks. (A to G) Interaction networks of acetylated proteins in different cellular functions from STRING analysis of the acetylome. Individual networks were generated for each specific functional category (table S1). Gray nodes indicate proteins previously reported to be acetylated.



its functions, such as DAXX, PML, PTEN, and HAUSP (38), are acetylated (table S1). HAUSP regulates the stability of p53 and MDM2, but less is known about its own regulation. The five HAUSP acetylation sites reported here represent a possible mechanism with which to modulate its activity. Collectively, these results indicate that not only p53 but many of the core proteins of the p53 circuitry also probably are regulated by acetylation.

A large number of proteins involved in chromatin remodeling and cell cycle are acetylated (table S1). For example, CDC2, a major cyclin-dependent kinase and regulator of S-phase progression and mitosis, was acetylated at two different lysines: K6 and K33. K6 is conserved between CDC2 and CDK2 and was acetylated on both kinases. The K33 acetylation site is located within the kinase domain and is important for the coordination of adenosine 5'-triphosphate (ATP) binding. This lysine is conserved even in the budding yeast homolog. Indeed, acetylation of yeast Cdc28 was confirmed by means of protein immunoblotting (fig. S7A). MS data furthermore confirmed that Cdc28 is acetylated on K40 (fig. S7B), a residue analogous to K33 of CDC2. To address functional consequences of altering this residue, we mutated K40 in yeast Cdc28 to argi-

nine, which is charge-conserving, or to glutamine, which is noncharged and may mimic acetylated lysine. Both point mutants failed to rescue *cdc28Δ* yeast strains (Fig. 4A). This lysine therefore appears to be critical to the proper function of Cdc28 in yeast. CDK9, a non-cell-cycle kinase, is acetylated on the analogous site, and acetylation inhibits its kinase activity (39). CDK6, another cell-cycle kinase, and PRPF4B (PRP4 pre-mRNA-processing factor 4 homolog B), a kinase involved in mRNA splicing, were also acetylated on the ATP-binding lysine. Thus, lysine acetylation may have a wider role in the regulation of kinase activity. Acetylation of SMC3 (structural maintenance of chromosomes 3) in yeast and human is important for the separation of sister chromatids (40). We found 22 sites, including two recently described ones (39), on all major SMC proteins that are part of the cohesion complex (SMC1A, -2, -3, -4, and -5). These data point to a larger role of acetylation in the regulation of the cohesion complex and thus separation of sister chromatids. We also found acetylation of many other cell-cycle regulators, such as TPX2, NUMA1, cyclin T1 and B1, BUB3, CDK6, and STMN1. As shown in Fig. 3C, acetylated cell-cycle proteins are highly connected by functional and physical protein-protein interactions.

Interactions of nuclear receptor-binding proteins to nuclear receptors via LXXLL motifs (where X represents any amino acid) regulate transcriptional activation. Acetylation of nuclear receptor coactivator 3 (NcoA3) near LXXLL motifs regulates these interactions (41). We found that both NcoA2 and NcoA3 are acetylated within their LXXLL motifs, which is likely to influence binding properties and therefore the function of these important nuclear receptor coactivators.

Nuclear pores are the gatekeepers of nucleocytoplasmic transport and, together with the RAN guanine triphosphate cycle, control protein traffic. To our knowledge, acetylation has not been implicated in the regulation of nuclear transport. We discovered that several constituents of this machinery are acetylated (Fig. 3D). RANGAP1 [RAN-guanosine triphosphatase (GTPase)-activating protein 1] is acetylated on K524, the same lysine residue that has an important role in the regulation of nuclear transport when modified by sumoylation (42). Acetylation of this lysine abolishes sumoylation and thus may regulate nuclear transport.

Cytoplasmic complexes regulated by acetylation. Acetylation of tubulin regulates microtubule function (43). We found several additional acet-

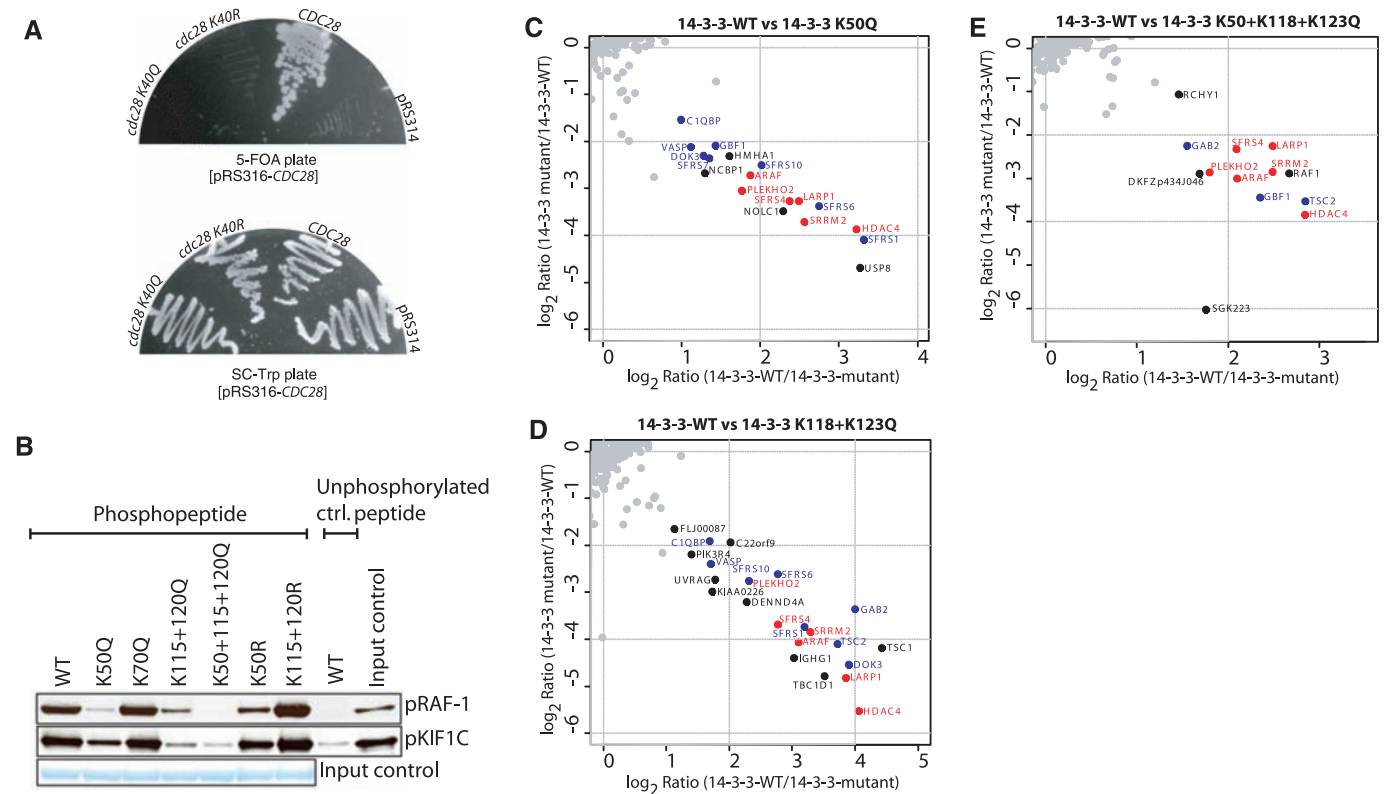


Fig. 4. Acetylation of Cdc28 and 14-3-3 impair their functions. **(A)** Mutation of acetylation site K40 on Cdc28 impairs its function. Growth of haploid yeast strains harboring empty vector or plasmids expressing *CDC28*, *cdc28-K40R*, or *cdc28-K40Q* in a *cdc28Δ* strain were tested for growth on 5-FOA plates, which select against the wild-type copy of *CDC28* on a separate URA3-marked plasmid. SC-Trp plates are shown as control. **(B)** Mutation of acetylated lysine on 14-3-3-ε abolishes binding to RAF1 and KIF1c phosphopeptide ligands. Binding of recombinant GST–14-3-3 proteins to phosphopeptides was analyzed by means of peptide pull-down assays.

(C to E) Mutation of K50Q and K118+K123Q impairs binding of 14-3-3 to full-length proteins. Proteins interacting with 14-3-3 wild type (WT) or mutants were identified and quantified by use of SILAC-based MS as described in fig. S9. Proteins in red are quantified in all six experiments, and proteins that were quantified in four or more experiments are blue. The x axis shows the relative binding of 14-3-3-WT compared with the indicated 14-3-3 mutants (log₂ SILAC ratio for WT/mutant). The y axis shows the relative binding of the indicated 14-3-3 mutants compared with 14-3-3-WT (log₂ SILAC ratio for mutant/WT).

ylation sites on various tubulin isoforms. The machinery controlling actin-based cell motility is also highly acetylated (Fig. 3E). All but one of the subunits of the ARP2/3 complex, the major actin nucleation complex, are acetylated, as are cortactin, cofilin, and coronin, which interact with the ARP2/3 complex (table S1). Of the nine acetylation sites identified on purified cortactin (44), we confirmed five and identified two others. Thus, regulation of cytoskeleton reorganization and cell motility by lysine acetylation may involve many more proteins than tubulin and cortactin.

The function of chaperone protein HSP90 is inhibited by acetylation (45). Hyperacetylation of HSP90 is thought to be a major effector through which KDAC inhibitors interfere with protein folding, and one acetylation site on HSP90- α is known (45). We could not confirm this site, but we found HSP90- α to be acetylated on at least 14 lysines and HSP90- β on a minimum of four. Likewise, all the eight subunits of TCP1 or TriC chaperone complex (CCT1 to -8) were also heavily acetylated (Fig. 3F).

14-3-3 proteins are widely expressed proteins that specifically bind to phosphoserine or phosphothreonine and regulate a diverse set of cellular processes, such as signal transduction, cell cycle, and DNA damage repair (46). 14-3-3 proteins also interact with HDAC1, HDAC4, and HDAC5. Multiple isoforms of 14-3-3 were acetylated in different cell types (table S1 and fig. S8A). Moreover, analogous sites were acetylated in multiple isoforms (fig. S8B).

To understand the role of 14-3-3 acetylation and its impact on the interaction of 14-3-3 with phosphorylated peptides, we mutated four lysines (K50, K69, and K118+K123) that are acetylated in vivo and are highly conserved among 14-3-3 proteins (47). Acetylated lysines in 14-3-3- ϵ were mutated either to glutamine (Q), so as to mimic acetylated lysine, or arginine (R), so as to prevent acetylation but conserve a positive charge. Mutation of K50 to glutamine reduced binding to two model phosphopeptides, and mutation of both K118 and K123 severely impaired its binding (Fig. 4B). In contrast, K118+K123R showed no such binding deficiency. A triple acetylation mimetic mutant (K50+K118+K123Q) did not bind phosphopeptides. To further confirm the importance of these sites, we performed pull-downs using glutathione *S*-transferase (GST)-13-3-3- ϵ fusion protein or its mutants and quantified proteins that specifically bound to 14-3-3 (fig. S9). The acetylation mimetic mutants of 14-3-3 (K50Q, K118+K123Q, and K50+K118+K123Q) showed impaired binding to synthetic peptides as well as to full-length proteins from whole-cell lysates (Fig. 4, C to E). Among the proteins that showed decreased binding to acetylation, mimetic mutants of 14-3-3 in these five are known to bind (RAF1, HDAC4, TSC22, GAB2, and ARAF), which independently validates the results obtained with peptide associations. The crystal structure of 14-3-3 bound to its phosphopeptide ligand confirms the importance of K50 and K123 in binding to its ligand (48). Thus, our results uncover a mechanism that modulates phosphorylation-dependent inter-

actions on the side of the phosphopeptide-binding domain and suggest crosstalk between phosphorylation and acetylation.

Our data also contain many further interesting leads for functional studies in a wide variety of biological areas. For example, an unexpected role of acetylation was recently reported for the activation of signaling from the interferon- α receptor (49). We identified acetylation sites in mediators of interferon actions such as DDX58, IRAK4, OAS2, and TRIM25. These data further strengthen the notion that acetylation may regulate innate immune responses.

Quantitative acetylation analysis upon KDAC inhibition. The clinical success of KDAC inhibitors for specific cancers as well as their promise for treating other diseases make it important to identify their targets at a site-specific level. We used SILAC to quantify downstream targets of SAHA and MS-275 in three different cell lines (fig. S1A). SAHA is reported to have broad inhibitory activity and to target many KDACs, with the exception of sirtuins (6). In contrast, MS-275 is thought to inhibit only three class I KDACs. Both inhibitors increased acetylation of specific histone sites, but neither changed acetylation of mitochondrial proteins, which are thought to be deacetylated by sirtuins. The inhibitors only up-regulated about 10% of all acetylation sites by at least a factor of two (table S4). Thus, both inhibitors were unexpectedly specific in vivo. Acetylation sites that were regulated more than twofold by both of these inhibitors in the triple-SILAC experiment in MV4-11 cells are listed in table S5. SAHA is a more potent inducer of histone acetylation than MS-275 (fig. S6). For example, a number of sites on the core histones H3 and H4 are several times more highly regulated in response to SAHA than by MS-275 (table S5). Acetylation of the variant histone H2AZ—a mark for DNA damage sites—was upregulated almost 20-fold by SAHA. The fact that acetylation of all histone sites is not equally increased suggests that these drugs target the histone deacetylases differently. Important cytoplasmic targets, such as HSP90 α and - β , were highly acetylated in the presence of SAHA but not in cells treated with MS-275 (fig. S1B). Conversely, the acetylation levels of four out of five sites on p53 were increased by MS-275 but not at all by SAHA.

Discussion. We describe a streamlined methodology for the proteome-wide identification and quantitation of acetylation sites and provide an in-depth view of the in vivo acetylome. We identified 3600 acetylation sites and implicated acetylation in the regulation of a diverse set of cellular functions. The data are available to the scientific community via Phosida (www.phosida.com) (28), a public database that is a member of ProteinExchange and provides detailed information about each acetylation site, such as evolutionary conservation and local secondary structure prediction of the protein sequence around the acetylation site.

The overall picture of lysine acetylation resulting from our experiments is that it contributes to regulation of almost all nuclear functions and to

the control of a surprisingly large array of cytoplasmic functions as well. The relatively high overlap of the measured acetylome in three cell lines and the fact that we covered most of the in vivo sites described in the literature indicate that we sampled a substantial part of all acetylation sites. The several thousand sites detected here place the size of the measured acetylome between that of the phosphoproteome, in which tens of thousands of sites are now quantifiable, and that of the spectrum of ubiquitinated proteins, for which only a few hundred sites are currently known. In common with these major PTMs, acetylation seems to have been adopted in the regulation of a wide variety of dissimilar cellular functions. A striking feature of acetylation is that it tends to occur in large macromolecular complexes.

The application of quantitative proteomics enabled us to assess the effects of perturbing the acetylation network in a global and unbiased manner. KDAC inhibitors were much more specific than generally thought, potentially shedding new light on their unexpected clinical specificity.

References and Notes

1. T. Kouzarides, *EMBO J.* **19**, 1176 (2000).
2. X. J. Yang, *Bioessays* **26**, 1076 (2004).
3. K. K. Lee, J. L. Workman, *Nat. Rev. Mol. Cell Biol.* **8**, 284 (2007).
4. M. D. Shahbazian, M. Grunstein, *Annu. Rev. Biochem.* **76**, 75 (2007).
5. X. J. Yang, E. Seto, *Mol. Cell* **31**, 449 (2008).
6. J. E. Bolden, M. J. Peart, R. W. Johnstone, *Nat. Rev. Drug Discov.* **5**, 769 (2006).
7. W. S. Xu, R. B. Parmigiani, P. A. Marks, *Oncogene* **26**, 5541 (2007).
8. K. Garber, *Nat. Biotechnol.* **25**, 17 (2007).
9. A. G. Kazantsev, L. M. Thompson, *Nat. Rev. Drug Discov.* **7**, 854 (2008).
10. D. Huangfu et al., *Nat. Biotechnol.* **26**, 795 (2008).
11. R. Aebersold, M. Mann, *Nature* **422**, 198 (2003).
12. B. Dorn, R. Aebersold, *Science* **312**, 212 (2006).
13. L. M. de Godoy et al., *Nature* **455**, 1251 (2008).
14. O. N. Jensen, *Nat. Rev. Mol. Cell Biol.* **7**, 391 (2006).
15. N. G. Ahn, A. H. Wang, *Curr. Opin. Chem. Biol.* **12**, 1 (2008).
16. J. V. Olsen et al., *Cell* **127**, 635 (2006).
17. N. C. Tedford, F. M. White, J. A. Radding, *Brief. Funct. Genomics Proteomics* **7**, 383 (2008).
18. S. C. Kim et al., *Mol. Cell* **23**, 607 (2006).
19. Materials and methods are available as supporting material on Science Online.
20. J. V. Olsen et al., *Mol. Cell. Proteomics* **4**, 2010 (2005).
21. N. C. Hubner, S. Ren, M. Mann, *Proteomics* **8**, 4862 (2008).
22. S. E. Ong et al., *Mol. Cell. Proteomics* **1**, 376 (2002).
23. M. Mann, *Nat. Rev. Mol. Cell Biol.* **7**, 952 (2006).
24. Single-letter abbreviations for the amino acid residues are as follows: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr.
25. J. Cox, M. Mann, *Nat. Biotechnol.* **26**, 1367 (2008).
26. I. Poser et al., *Nat. Methods* **5**, 409 (2008).
27. The Uniprot Consortium, *Nucleic Acids Res.* **36**, D190 (2008).
28. F. Gnäd et al., *Genome Biol.* **8**, R250 (2007).
29. R. Malik, E. A. Nigg, R. Korner, *Bioinformatics* **24**, 1426 (2008).
30. R. D. Finn et al., *Nucleic Acids Res.* **36**, D281 (2008).
31. P. R. Thompson et al., *Nat. Struct. Mol. Biol.* **11**, 308 (2004).
32. T. Suganuma et al., *Nat. Struct. Mol. Biol.* **15**, 364 (2008).
33. A. Ruepp et al., *Nucleic Acids Res.* **36**, D646 (2008).
34. L. J. Jensen et al., *Nucleic Acids Res.* **37**, D412 (2009).
35. G. D. Bader, C. W. Hogue, *BMC Bioinformatics* **4**, 2 (2003).
36. Y. Sun, X. Jiang, S. Chen, N. Fernandes, B. D. Price, *Proc. Natl. Acad. Sci. U.S.A.* **102**, 13182 (2005).
37. X. Jiang, Y. Sun, S. Chen, K. Roy, B. D. Price, *J. Biol. Chem.* **281**, 15741 (2006).

38. M. S. Song *et al.*, *Nature* **455**, 813 (2008).
39. A. Sabo, M. Lusic, A. Cereseto, M. Giacca, *Mol. Cell. Biol.* **28**, 2201 (2008).
40. J. Zhang *et al.*, *Mol. Cell* **31**, 143 (2008).
41. H. Chen, R. J. Lin, W. Xie, D. Wilpitz, R. M. Evans, *Cell* **98**, 675 (1999).
42. R. Mahajan, C. Delphin, T. Guan, L. Gerace, F. Melchior, *Cell* **88**, 97 (1997).
43. S. Westermann, K. Weber, *Nat. Rev. Mol. Cell Biol.* **4**, 938 (2003).
44. X. Zhang *et al.*, *Mol. Cell* **27**, 197 (2007).
45. B. T. Scroggins *et al.*, *Mol. Cell* **25**, 151 (2007).
46. D. K. Morrison, *Trends Cell Biol.* **19**, 16 (2009).
47. X. Yang *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **103**, 17237 (2006).
48. K. Rittinger *et al.*, *Mol. Cell* **4**, 153 (1999).
49. X. Tang *et al.*, *Cell* **131**, 93 (2007).
50. We thank the members of the Department of Proteomics and Signal Transduction at the Max Planck Institute for Biochemistry for their helpful discussions and suggestions; T. Hyman, I. Poser, and N. Hubner for providing GFP-tagged bacterial artificial chromosome (BAC) transgenic cell lines; J. Duyser and S. Raula for providing GST-14-3-3 expression plasmid; T. Misteli for providing CBX3 plasmid; Max Planck Institute for Biochemistry core facility for peptide synthesis and DNA sequencing; and H. Serve and C. Brandts for their advice with KDAC inhibitors. This work is supported by the Deutsche

Krebshilfe (108401) and the Max Planck Society. M.L.N. is supported by the European Molecular Biology Organization.

Supporting Online Material

www.sciencemag.org/cgi/content/full/1175371/DC1
Materials and Methods
Figs S1 to S9
Tables S1 to S5
References

23 April 2009; accepted 30 June 2009

Published online 16 July 2009;

10.1126/science.1175371

Include this information when citing this paper.

Detection of 16 Gamma-Ray Pulsars Through Blind Frequency Searches Using the Fermi LAT

A. A. Abdo,^{1*} M. Ackermann,² M. Ajello,² B. Anderson,³ W. B. Atwood,³ M. Axelsson,^{4,5} L. Baldini,⁶ J. Ballet,⁷ G. Barbiellini,^{8,9} M. G. Baring,¹⁰ D. Bastieri,^{11,12} B. M. Baughman,¹³ K. Bechtol,² R. Bellazzini,⁶ B. Berenji,² G. F. Bignami,¹⁴ R. D. Blandford,² E. D. Bloom,² E. Bonamente,^{15,16} A. W. Borgland,² J. Bregeon,⁶ A. Brez,⁶ M. Brigida,^{17,18} P. Bruehl,¹⁹ T. H. Burnett,²⁰ G. A. Caliendo,^{17,18} R. A. Cameron,²¹ P. A. Caraveo,²¹ J. M. Casandjian,⁷ C. Cecchi,^{15,16} Ö. Çelik,²² A. Chekhtman,^{1,23} C. C. Cheung,²² J. Chiang,² S. Ciprini,^{15,16} R. Claus,² J. Cohen-Tanugi,²⁴ J. Conrad,^{4,25,26} S. Cutini,²⁷ C. D. Dermer,¹ A. de Angelis,²⁸ A. de Luca,¹⁴ F. de Palma,^{17,18} S. W. Digel,² M. Dormody,³ E. do Couto e Silva,² P. S. Drell,² R. Dubois,² D. Dumora,^{29,30} C. Farnier,²⁴ C. Favuzzi,^{17,18} S. J. Fegan,¹⁹ Y. Fukazawa,³¹ S. Funk,² P. Fusco,^{17,18} F. Gargano,¹⁸ D. Gasparrini,²⁷ N. Gehrels,^{22,32} S. Germani,^{15,16} B. Giebels,¹⁹ N. Giglietto,^{17,18} P. Giommi,²⁷ F. Giordano,^{17,18} T. Glanzman,² G. Godfrey,² I. A. Grenier,⁷ M.-H. Grondin,^{29,30} J. E. Grove,¹ L. Guillemot,^{29,30} S. Guiriec,³³ C. Gwon,¹ Y. Hanabata,³¹ A. K. Harding,²² M. Hayashida,² E. Hays,²² R. E. Hughes,¹³ G. Jóhannesson,² R. P. Johnson,³ T. J. Johnson,^{22,32} W. N. Johnson,¹ T. Kamae,² H. Katagiri,³¹ J. Kataoka,³⁴ N. Kawai,^{35,36} M. Kerr,²⁰ J. Knödseder,³⁷ M. L. Kocian,² M. Kuss,⁶ J. Lande,² L. Latronico,⁶ M. Lemoine-Goumard,^{29,30} F. Longo,^{8,9} F. Loparco,^{17,18} B. Lott,^{29,30} M. N. Lovellette,¹ P. Lubrano,^{15,16} G. M. Madejski,² A. Makeev,^{1,23} M. Marelli,²¹ M. N. Mazziotta,¹⁸ W. McConville,^{22,32} J. E. McEnery,²² C. Meurer,^{4,26} P. F. Michelson,² W. Mitthumsiri,² T. Mizuno,³¹ C. Monte,^{17,18} M. E. Monzani,² A. Morselli,³⁸ I. V. Moskalenko,² S. Murgia,² P. L. Nolan,² J. P. Norris,³⁹ E. Nuss,²⁴ T. Ohsugi,³¹ N. Omodei,⁶ E. Orlando,⁴⁰ J. F. Ormes,³⁹ D. Paneque,² D. Parent,^{29,30} V. Pelassa,²⁴ M. Pepe,^{15,16} M. Pesce-Rollins,⁶ M. Pierbattista,⁷ F. Piron,²⁴ T. A. Porter,³ J. R. Primack,³ S. Rainò,^{17,18} R. Rando,^{11,12} P. S. Ray,¹ M. Razzano,⁶ N. Rea,^{41,42} A. Reimer,² O. Reimer,² T. Reposeur,^{29,30} S. Ritz,²² L. S. Rochester,² A. Y. Rodriguez,⁴² R. W. Romani,² F. Ryde,^{4,25} H. F.-W. Sadrozinski,³ D. Sanchez,¹⁹ A. Sander,¹³ P. M. Saz Parkinson,³ J. D. Scargle,⁴³ C. Sgrò,⁶ E. J. Siskind,⁴⁴ D. A. Smith,^{29,30} P. D. Smith,¹³ G. Spandre,⁶ P. Spinelli,^{17,18} J.-L. Starck,⁷ M. S. Strickman,¹ D. J. Suson,⁴⁵ H. Tajima,² H. Takahashi,³¹ T. Takahashi,⁴⁶ T. Tanaka,² J. G. Thayer,² D. J. Thompson,²² L. Tibaldo,^{11,12} O. Tibolla,⁴⁷ D. F. Torres,^{42,48} G. Tosti,^{15,16} A. Tramacere,^{2,49} Y. Uchiyama,² T. L. Usher,² A. Van Etten,² V. Vasileiou,^{50,51} N. Vilchez,³⁷ V. Vitale,^{38,52} A. P. Waite,² P. Wang,² K. Watters,² B. L. Winer,¹³ M. T. Wolff,¹ K. S. Wood,¹ T. Ylinen,^{4,25,53} M. Ziegler³†

Pulsars are rapidly rotating, highly magnetized neutron stars emitting radiation across the electromagnetic spectrum. Although there are more than 1800 known radio pulsars, until recently only seven were observed to pulse in gamma rays, and these were all discovered at other wavelengths. The Fermi Large Area Telescope (LAT) makes it possible to pinpoint neutron stars through their gamma-ray pulsations. We report the detection of 16 gamma-ray pulsars in blind frequency searches using the LAT. Most of these pulsars are coincident with previously unidentified gamma-ray sources, and many are associated with supernova remnants. Direct detection of gamma-ray pulsars enables studies of emission mechanisms, population statistics, and the energetics of pulsar wind nebulae and supernova remnants.

A wide variety of astrophysical phenomena, such as black holes, active galactic nuclei, gamma-ray bursts, and pulsars, are known to produce photons with energies exceeding tens of megaelectron volts. Detection and accurate lo-

calization of sources at these energies is challenging because of the low fluxes involved and the limitations of the detection techniques. The sky above 100 MeV was surveyed more than 30 years ago by the COS-B satellite (1) and more recently by the

Energetic Gamma Ray Experiment Telescope (EGRET) (2) onboard the Compton Gamma Ray Observatory. One of the main legacies of EGRET was the detection of ~300 gamma-ray sources, many of which have remained unidentified despite searches at a wide variety of wavelengths (3). Many EGRET (and COS-B) unidentified sources are thought to be of galactic origin because of their lack of variability and concentration along the galactic plane. A large fraction of these have been suspected to be pulsars [e.g., (4–6)] despite deep radio and x-ray searches often failing to uncover pulsed emission, even when the gamma-ray sources were coincident with supernova remnants (SNRs) or pulsar wind nebulae (PWNs). The lack of radio pulsations has usually been explained as the narrow radio beams missing the line of sight toward Earth (7). We refer to such pulsars as “radio-quiet”; even though they may emit radio waves, these cannot be detected at Earth. Before the launch of the Fermi Gamma-ray Space Telescope, Geminga (8) was the only known radio-quiet gamma-ray pulsar. Current models of pulsar gamma-ray emission predict that gamma-ray beams are much wider than radio beams (9), thus suggesting that there may be a large population of radio-quiet gamma-ray pulsars.

Soon after launch on 11 June 2008, the Large Area Telescope (LAT) onboard Fermi began surveying the sky at energies above 20 MeV. A companion paper describes the LAT detection of a population of gamma-ray millisecond pulsars (MSPs) (10). Here we report the detection of 16 pulsars found in blind frequency searches using the LAT. Previously, gamma-ray pulsars had been detected only with the use of a radio (or, in the case of Geminga, an x-ray) ephemeris.

Observations and data analysis. The LAT is a high-energy gamma-ray telescope sensitive to photon energies from 20 MeV to >300 GeV, featuring a solid-state silicon tracker, a cesium-iodide calorimeter, and an anticoincidence detector (11). Gamma-ray events recorded in the LAT have time stamps derived from a GPS clock on the Fermi satellite with an accuracy of <1 μs (12). The LAT operates in continuous sky survey mode, covering the entire sky every 3 hours. Relative to EGRET, the LAT has a larger effective area (9500 cm² at normal incidence), larger field of view (2.4 sr), more efficient use of time on orbit for photon collection, and a finer point spread function (5° at 100 MeV, 0.8° at 1 GeV). The first three factors result in more rapid photon accumulation, and the fourth increases the signal-to-noise ratio by improving the background rejection.



www.sciencemag.org/cgi/content/full/1175371/DC1

Supporting Online Material for

Lysine Acetylation Targets Protein Complexes and Co-Regulates Major Cellular Functions

Chunaram Choudhary, Chanchal Kumar, Florian Gnad, Michael L. Nielsen,
Michael Rehman, Tobias Walther, Jesper V. Olsen, Matthias Mann*

*To whom correspondence should be addressed. E-mail: mmann@biochem.mpg.de

Published 16 July 2009 on *Science Express*
DOI: 10.1126/science.1175371

This PDF file includes:

Materials and Methods

Figs S1 to S9

Tables S3 and S5

References

Other Supporting Online Material for this manuscript includes the following:
(available at www.sciencemag.org/cgi/content/full/1175371/DC1)

Tables S1, S2, and S4 as zipped Excel files

Supporting online material (SOM) accompanying the manuscript

Lysine acetylation targets protein complexes and co-regulates major cellular functions

Chunaram Choudhary, Chanchal Kumar, Florian Gnad, Michael L. Nielsen, Michael Rehman, Tobias Walther, Jesper V. Olsen, Matthias Mann*

This SOM contains:

Experimental procedures

Seven supplementary figure (Figure S1-S9) and figure legends

Five supplementary tables (Table S1-S5) and table legends

Experimental procedures

Reagents

The lysine deacetylase inhibitors MS275 (a synthetic pyridyl carbamate compound) and suberoylanilide hydroxamic acid (SAHA) were purchased from Alexis Biochemicals (San Diego, CA). Endoproteinase Lys-C was from Wako (Japan) and modified sequencing grade trypsin was purchased from Promega (Mannheim, Germany). Agarose conjugated anti-acetyllysine antibodies were purchased from ImmuneChem Pharmaceuticals Inc. (Burnaby, British Columbia, Canada) and Sep-Pak C₁₈ cartridges were purchased from Waters Corporation. GFP-trap immunoprecipitation matrix was purchased from Chromotek (Chromotek GmbH, Martinsried, Germany), HRP conjugated anti-GST antibodies were from Abcam (Cambridge, USA), streptavidin sepharose was from GE Healthcare (Uppsala, Sweden), anti-acetyllysine antibody used for Western blotting was from Cell Signaling Technology (Cambridge, USA). Site directed mutagenesis kit was purchased from Stratagene (La Jolla, USA).

Stable isotope labeling with amino acids in cell culture (SILAC) labeling of cells

MV4-11 and Jurkat cells were cultured in RPMI1640 and A549 cells in DMEM medium supplemented with 10% FCS and 1% penicillin and streptomycin. The cells were

labeled with either L-arginine and L-lysine or L-arginine- $\text{U-}^{13}\text{C}_6\text{-}^{14}\text{N}_4$ and L-lysine- $^2\text{H}_4$ or arginine- $\text{U-}^{13}\text{C}_6\text{-}^{15}\text{N}_4$ and L-lysine- $\text{U-}^{13}\text{C}_6\text{-}^{15}\text{N}_2$ as described previously (Ong et al., 2002). All cells were cultured in a humidified incubator at 37°C with 5% CO_2 .

Preparation of cell lysates for mass spectrometric analysis

MV4-11, Jurkat, and A549 cells were cultured in the presence of MS275 (5 μM) or SAHA (2 μM) or DMSO as vehicle control for 24 hours prior to lysis. Cells were washed twice with 1X PBS and cell lysates were prepared in complete RIPA buffer, lysates were vortexed for 90 seconds and centrifuged at 17000g for 15 minutes at 4 °C. From the cleared lysates, proteins were precipitated by adding 4x excess volumes of ice-cold acetone and stored at -20 °C for overnight.

In-solution digestion of proteins and enrichment of lysine acetylated peptides

The precipitated proteins were re-dissolved in 6M urea/2M thiourea/10mM HEPES pH 8.0. For extracting proteins from the RIPA insoluble pellet, the pellets were dissolved in 8M Urea/thiourea after digesting the DNA with Benzonase (Merck). Protein disulfide bonds were reduced by 1 mM dithiothreitol, alkylated with 5.5 mM iodoacetamide, digested using endoproteinase Lys-C (1:100 w/w) and modified sequencing grade trypsin (1:100 w/w) after 4-fold dilution in deionized water, as described previously (Olsen et al., 2006). The resulting peptides were purified using reversed-phase Sep-Pak C_{18} peptide purification cartridges. The purified peptides were eluted using 50 percent acetonitrile. The eluates were lyophilized over night and lyophilized peptides were redissolved in the immunoprecipitation buffer (50 mM MOPS pH 7.2, 10 mM Sodium phosphate, 50 mM Sodium chloride) and incubated with anti-acetyllysine antibody for 12 hrs at 4 °C on a rotation wheel. The immunoprecipitates were washed 4 times with the immunoprecipitation buffer followed by two washes with distilled water. Residual water was removed and acetylated peptides bound to antibodies were eluted by acidified water (0.1% TFA in water).

Isoelectric focusing of peptides

The eluted lysine acetylated peptides were separated into 12 fractions by iso-electric focusing (IEF) using the Agilent 3100 OFFGEL Fractionator (Agilent, G3100AA) and the 3100 OFFGEL High Res kit, pH 3–10 (Agilent, 5188-6424) according to the manufacturer's instructions as described (Hubner et al., 2008). Briefly, peptides were focused for 50 kilovolt hours at a maximum current of 50 μ A and maximum power of 200 milliwatts. Peptide fractions were acidified by adding 10 μ L of a solution containing 30% acetonitrile, 10% trifluoroacetic. The samples were StageTipped as described (Rappsilber et al., 2007).

Mass spectrometric analysis

The peptides were eluted from StageTips and analyzed by online nanoflow liquid chromatography tandem mass spectrometry (LC-MS/MS) as described previously (Olsen et al., 2006) with a few modifications. All nanoLC-MS/MS-experiments were performed on a Agilent Technologies 1200 nanoflow system connected to the LTQ-Orbitrap XL (Thermo Electron, Bremen, Germany) equipped with a nanoelectrospray ion source (Proxeon Biosystems, Odense, Denmark) as described (Olsen et al., 2005). Briefly, the LTQ Orbitrap XL instrument (Olsen et al., 2007) under Xcalibur 2.0 with LTQ Orbitrap Tune Plus Developers Kit version 2.0 software was operated in the data dependent mode to automatically switch between MS and MS/MS acquisition. Survey full scan MS spectra (from m/z 300 – 2000) were acquired in the orbitrap with resolution $R=60,000$ at m/z 400 (after accumulation to a 'target value' of 1,000,000 in the linear ion trap). The ten most intense ions were sequentially isolated and fragmented in the linear ion trap by collisionally induced dissociation (CID) at a target value of 5,000 or maximum ion time of 150 ms. For all measurements with the orbitrap detector a lock-mass ion from ambient air (m/z 429.08875) was used for internal calibration as described earlier (Olsen et al., 2005). Typical mass spectrometric conditions were: spray voltage, 2.2 kV; no sheath and auxiliary gas flow; heated capillary temperature, 175°C; normalized CID collision energy 35% for MS2 in LTQ. The ion selection threshold was 500 counts for MS2. An activation $q = 0.25$ and activation time of 30 ms were used.

Assigning peptide sequences using MASCOT and MaxQuant

Raw Orbitrap full-scan MS and ion trap MS² spectra were processed by MaxQuant as previously described (Cox and Mann, 2008); all identified SILAC pairs were quantified, and MS/MS spectra were centroided and merged into peak-list files and searched against the human IPI protein database version 3.37. Peptides and proteins were identified by Mascot (Matrix Science, London, UK) via automated database searching of all tandem mass spectra against an in-house curated target/decoy database (forward and reversed version of the human International Protein Index protein sequence database (IPI, versions 3.37, 138,632 forward and reversed protein sequences, EBI (<http://www.ebi.ac.uk/IPI/>) containing all human protein entries from Swiss-Prot, TrEMBL, RefSeq and Ensembl. Spectra were initially searched with a mass tolerance of 7 ppm in MS mode and 0.5 Da in MS/MS mode and strict trypsin specificity and allowing up to 4 missed cleavage sites. Cysteine carbamidomethylation was searched as a fixed modification, whereas N-acetyl protein, N-pyroglutamine, oxidized methionine and acetylation of lysine was searched as variable modification.

Data analysis

The resulting Mascot html-output files were loaded into the MaxQuant software for further processing. In MaxQuant we fixed the estimated false discovery rate (FDR) of all peptide and protein identifications at maximum 1%, by filtering on peptide length, mass error and mascot score of all forward and reversed peptide identifications. From these only peptides with at least one internal acetylated lysine residue and minimum mascot score of 10 were accepted. Finally, to pinpoint the actual acetylated amino acid residue within the identified peptide sequence in an unbiased manner, we calculated the localization probabilities of all lysine acetylation sites using the PTM score algorithm as described (Olsen et al., 2006).

Acetylation site overlap between cell lines

For the six separate experiments (two each for MV4-11, A549 and Jurkat) we calculated the overlap of each versus all others, and calculated two metrics. (a) Overlap of the two data sets as the percent of the smaller of two data sets:

$$\frac{\cap [Data1, Data2]}{\min[Data1, Data2]} \times 100$$

and, (b) Jaccard similarity coefficient(scaled to [0,100]) as defined below:

$$\frac{\cap [Data1, Data2]}{\cup [Data1, Data2]} \times 100$$

Phylogenetic conservation analysis

To compare the degree of conservation of genes that encode acetylated proteins with the genes that do not encode acetylated proteins, we first mapped our identified peptides to gene transcripts annotated in Ensembl (Flicek et al., 2008) as described in MAPU 2.0 (Gnad et al., 2008). We used the Ensembl Compara Database, which contains phylogenetic relationships throughout eukaryotic species and cover the phylogenetic tree representatively. To avoid a possible bias for high abundant proteins in our dataset, we chose only those genes that are annotated as 'known' in the Ensembl database.

Pfam protein domain enrichment analysis

For the identified IPI protein entries the Pfam domain (Finn et al., 2008) enrichment analysis was done with respect to the whole IPI human proteome by hypergeometric test available in the GOSTATS package (Falcon and Gentleman, 2007) in the R statistical environment (Team, 2008). All domains with $p\text{-val} < 10^{-5}$ were considered significant. The analysis was done separately for over and under represented Pfam domains.

CORUM complex enrichment analysis

Protein complex annotations for human dataset were downloaded from CORUM database (Ruepp et al., 2008) <http://mips.gsf.de/genre/proj/corum/index.html>. CORUM was specifically chosen as it contains manually curated and experimentally verified protein complex annotations. The IPI identifiers were mapped to UniProt identifiers to

enable comparison with CORUM. Hypergeometric test was used to find enriched complexes with respect to the complete CORUM human dataset. All complexes with $p\text{-val} < 0.05$ were considered significant.

Network analysis for the acetylome. The IPI identifiers of our acetylation study were mapped to the Ensemble protein identifiers (2,142 in total), and then searched against the STRING database version8 (Jensen et al., 2008) for protein-protein interactions. Only interactions between the proteins belonging to the acetylation dataset were selected. STRING defines a metric called “confidence score” to define interaction confidence; we fetched all interactions for our acetylation dataset which had a confidence score ≥ 0.7 (high confidence). The resulting interactome had 1,011 nodes and 5,999 interactions which we call the acetylation interaction network (AIN) (Figure S6A).

MCODE based node cluster analysis

The acetylation interaction network (AIN) was analyzed for densely connected regions using a graph theoretic clustering algorithm - "Molecular Complex Detection" (MCODE) (Bader and Hogue, 2003). MCODE is part of the plug-in toolkit of network analysis and visualization software Cytoscape (Shannon et al., 2003). The first 7 high ranking modules were selected for further analysis and rendering (Figure S6B).

Network significance analysis

In order to ascertain the significance of the total connectivity of our AIN we employed a bootstrapping procedure to compare AIN connectivity to those of randomly sampled networks generated from STRING across 1,000,000 iterations. Firstly, we filtered the STRING version8 human interactome to include only interactions which had a confidence score ≥ 0.7 , which we call Human high confidence network (HHCN). The HHCN had 12,318 nodes with 145,221 interactions. Secondly, we sampled 2,142 (equal to the number of our acetylation dataset identifiers) unique Ensemble protein identifiers from the HHCN network node list and created a sub-network called *randomHHCN* with interactions between proteins extracted from HHCN only if both the proteins were

present in the random protein list. This was repeated 1,000,000 time and the total number of edges in $randomHHCN_i$ $i \in [1, 1000000]$ were tabulated.

The significance of AIN connectivity was calculated by the following formula:

$$p - val = \frac{No.of\ times[total\ edges(randomHHCN) \geq total\ edges(AIN)]}{Total\ no.of\ permutations}$$

We found our acetylome interaction network to have a connectivity significance of $p\text{-val} < 10^{-6}$ (Figure S5).

Yeast (*Saccharomyces cerevisiae*) strains:

Genomic deletions in diploid W303 (*ura3 trp1 leu2 his3 ade2 can1-100*) were generated by PCR-mediated genomic integration of the natNT2 marker (Janke et al., 2004). For plasmid shuffle experiments, more than ten independent transformants for each plasmid were tested.

Yeast strains list:

TWY1369	Mat a/ α	<i>ura3 trp1 leu2 his3 ade2 can1-100</i>	<i>cdc28Δ::natNT2</i>	This study
TWY1370	Mat a/ α	<i>ura3 trp1 leu2 his3 ade2 can1-100</i>	<i>cdc28Δ::natNT2</i> , [pRS316-CDC28]	This study
TWY1371	Mat α	<i>ura3 trp1 leu2 his3 ade2 can1-100</i>	<i>cdc28Δ::natNT2</i> , [pRS316-CDC28]	This study
TWY1372	Mat α	<i>ura3 trp1 leu2 his3 ade2 can1-100</i>	<i>cdc28Δ::natNT2</i> , [pRS316-CDC28; pRS314]	This study
TWY1373	Mat α	<i>ura3 trp1 leu2 his3 ade2 can1-100</i>	<i>cdc28Δ::natNT2</i> , [pRS316-CDC28; pRS314-CDC28]	This study
TWY1374	Mat α	<i>ura3 trp1 leu2 his3 ade2 can1-100</i>	<i>cdc28Δ::natNT2</i> , [pRS316-CDC28; pRS314- <i>cdc28K40R</i>]	This study
TWY1375	Mat α	<i>ura3 trp1 leu2 his3 ade2 can1-100</i>	<i>cdc28Δ::natNT2</i> , [pRS316-CDC28; pRS314- <i>cdc28K40Q</i>]	This study
TWY1376	Mat a	S288c	Cdc28-GFP::HISMx6	(Huh et al., 2003).

Acetylation mapping of yeast Cdc28

To investigate acetylation of *S. cerevisiae* Cdc28 we used a C-terminal GFP tagged Cdc28 strain (Huh et al., 2003). Total protein lysate was prepared by bead milling in a buffer containing 150mM potassium acetate, 20mM HEPES pH 7.4, 5% Glycerol. Subsequently TritonX-100 was added to a final concentration of one percent and

lysates were clarified by two consecutive spins of 5 minutes at 2000g. Cdc28-GFP was immunoprecipitated using a GFP-trap matrix and immunoprecipitates were washed 5X using the lysis buffer. Proteins were eluted by boiling the beads in SDS-PAGE gel loading buffer and separated on 1D gel, gel pieces were excised near the expected molecular weight of Cdc28-GFP. Peptides were recovered after in-gel trypsin digestion (Shevchenko et al., 1996). The resulting peptides were Stage-tipped and analyzed by LC-MS/MS as described above.

Plasmid constructs and site directed mutagenesis

The plasmid encoding human 14-3-3 epsilon fused to GST on N-terminal (pGEX-KG-GST-14-3-3 epsilon) was a kind gift from Dr. Justus Duyster (Urschel et al., 2005), and pBCHGN-CBX3 was a kind gift from Tom Misteli (Cheutin et al., 2003). To generate pRS316-*CDC28* and pRS314-*CDC28*, the promoter and open reading frame of *CDC28* was amplified using primers 5' –TATGGATCCTTCCCTTACTGGGCGCACGCAGT and 3' ATAGAATTCAATTATGATTCTTGAAGTAGGG and cloned into the EcoRI and BamHI sites of pRS316 or pRS314, respectively (Sikorski and Hieter, 1989). Mutagenesis on human 14-3-3 epsilon and yeast *CDC28* were performed using site directed mutagenesis kit (Stratagene) following manufacturer's instructions. Integrity of the mutant plasmids was verified by DNA sequencing.

Lysate preparation for Western blotting

Total cell extract from Hela or U2OS BAC transgenic cell lines expressing GFP-tagged transgenes was prepared using high salt RIPA lysis buffer containing 420 mM NaCl. Lysates were rotated at 4°C for 60 minutes. Thereafter, lysates were sonicated and cleared by centrifugation (16,000g for 15 minutes). For immunoprecipitation, cell lysates were incubated with 25 µL of GFP-trap matrix (Chromotek), at 4 °C overnight. Immunoprecipitates were washed five times using lysis buffer. Proteins bound to beads were eluted by boiling beads in gel loading buffer. Proteins were separated on SDS-PAGE and transferred on to PVDF membrane and probed using anti-GST or anti-acetyllysine antibodies.

Transfection of HEK293 cells

For transient overexpression of CBX3, HEK293 cells were transfected using Lipofectamine and Plus reagents (Invitrogen), following manufacturer's instructions. Thirty six hours after transfection, cells were trypsinized, washed twice with PBS, and lysed as described above.

Peptide synthesis and peptide pull-down

Phosphorylated and unphosphorylated peptide sequences of RAF-1 (LSQRQRSTpSTPNVHMV) and KIF1c (RRQRpSAP) containing a desthio-biotinylated amino-group on the N-terminus, and a Serine-Glycine (SG) spacer, were synthesized using the Fmoc strategy as described (Schulze and Mann, 2004). Peptides were immobilized on to streptavidin sepharose beads (GE Healthcare) and subsequently incubated with 2 µg of recombinant GST-14-3-3 for 2 hr at 4°C on a rotational wheel. Beads were washed five times with 1 ml of RIPA lysis buffer. Bound proteins were eluted by boiling beads in SDS, and eluted proteins were separated on 4%–12% gradient gels (Invitrogen), and stained using colloidal Coomassie blue staining kit (Invitrogen).

Preparation of recombinant GST-14-3-3

Recombinant GST-tagged 14-3-3 epsilon and its different mutants were expressed in *E. coli* BL21 cells. For expression of the recombinant protein cultures were grown at 30 degrees, and expression of the fusion protein was induced for 3 hours using 1 mM Isopropyl β-D-1-thiogalactopyranoside (IPTG). Bacterial cultures were resuspended in lysis buffer 20mM Triethanolamine (TEA)-HCL, 150 mM NaCl, 10 mM EDTA, 1 mM PMSF, 0.5% Triton-X100, and 0.1% β-mercaptoethanol. Lysates were sonicated and cleared by centrifugation. Cleared lysates were incubated with GST-sepharose beads overnight. Thereafter, beads were washed four times using GST wash buffers. Finally, beads were collected as 50% slurry. GST-14-3-3 was eluted from beads using 10 mM glutathione.

GST-pull-down and in-gel digestion

To identify proteins that interact with 14-3-3 and its mutants, 500 µg of cell extracts from SILAC-labeled (‘light’ or ‘heavy’) MV4-11 cells were incubated with different GST-14-3-3 mutants immobilized on to sepharose beads and incubated 3 hours at 4°C. Thereafter, beads were washed five times with RIPA lysis buffer and proteins bound to these beads were eluted using 8M urea and separated on 4-12% gradient SDS-PAGE, and stained using colloidal blue Coomassie staining. In-gel digestion of proteins and recovery of peptides were performed as described earlier (Shevchenko et al., 2006; Shevchenko et al., 1996). Eluted peptides were stage-tipped (Rappsilber et al., 2007) and analyzed by mass spectrometry.

Supplementary Figures:

Figure S1. Experimental strategy for lysine acetylation analysis. (A) Three human cell lines MV4-11, Jurkat and A549 were SILAC labeled and treated with KDAC inhibitors SAHA or MS-275 for 24h as indicated. (B) Quantitation of acetylation sites of HSP90α and p53 showing differential regulation by the inhibitors.

Figure S2. Overlap of acetylation sites. The acetylation sites identified in six different experiments were compared by ‘one versus all’ comparisons. (A) Overlap between two data sets as percentage of the smaller data set. (B) Shows the Jaccard similarity coefficient scaled in range [0,100], calculated as described in the SOM.

Figure S3. Evolutionary conservation of lysine acetylated proteins. (A) Percentage of human homologs of acetylated (red) and non-acetylated (blue) proteins in different eukaryotic organisms. The grey, broken line indicates conservation of the phosphoproteome. (B) Percentage of homologs of acetylated proteins in prokaryotes with same color code.

Figure S4. Acetylation sites of subunits of large macromolecular complexes.

(A-F) Major functional complexes involved in regulation of chromatin functions. Complexes were extracted from the literature. Nuclear protein complexes have colored background. Each red dot indicates a novel acetylation site whereas blue dots are sites that have been found here but also reported previously. HCF is reported as a subunit specific to drosophila ATAC complex (Guelman et al., 2009; Suganuma et al., 2008), thus marked with an asterisk. The complexes shown with colored background are large protein complexes that involve direct protein-protein interactions. The figure also shows ubiquitin ligases and deubiquitylases which are not part of a specific protein complex but are involved in modifying chromatin associated proteins.

Figure S5. Network connectivity significance analysis. The connectivity significance of the acetylation interaction network was accessed by bootstrapping of node permutations across 1000,000 iterations. (A) Histogram of the average edges per network (grey shaded) and the mean [average edges] of the random permutations (blue line). The average number of edges in our acetylome is shown by the red line. (B) Histogram of the total number of edges per network (grey shaded) and the mean [total edges] of the random permutations (blue line). The total number of edges in the acetylome is shown by a red line.

Figure S6. Network topology of the acetylome. (A) The protein-protein interaction network of our acetylation dataset as extracted from STRING. The seven most highly ranked and tightly connected network clusters as obtained by MCODE analysis (Bader and Hogue, 2003) are color coded and later rendered as separate modules (B-H).

Figure S7. Yeast Cdc28 is acetylated at the ATP binding lysine K40. (A) Acetylation of cdc28 in *S. cerevisiae*. GFP-tagged Cdc28 was immunoprecipitated from *S. cerevisiae* and lysine acetylation was analyzed using anti-acetyllysine immunostaining. (B) The figure shows the fragmentation spectra of the doubly charged tryptic peptide surrounding lysine 40 of Cdc40. The peptide contains two lysines because trypsin does not cleave at acetylated lysine. The y and b series indicate fragments at amide bonds of

the peptide. The complete series is present, unambiguously verifying the acetylated peptide. The inset shows the intact peptide signal of the acetylated peptide before fragmentation.

Figure S8. Sequence alignment of human 14-3-3 isoforms. (A) Protein sequences of all seven human 14-3-3 isoforms were aligned using the ClustalW2 algorithm (Larkin et al., 2007). The lysines highlighted in red are acetylated in at least two different isoforms. Asterisks mark lysines that are acetylated in multiple 14-3-3s and are conserved across all seven isoforms. (B) Multiple 14-3-3 isoforms are acetylated in-vivo. Some of the peptides sequences are not unique to specific 14-3-3 isoform but match to multiple isoforms.

Figure S9. Schematic presentation of GST-14-3-3 pull-down experiment. GST-14-3-3 epsilon or its mutants were incubated with cell lysates prepared from MV4-11 cells labeled with 'light' or 'heavy' SILAC. In the forward experiment, 14-3-3 epsilon wild-type was incubated with 'heavy' and mutant 14-3-3 with 'light' SILAC lysates. The specific interactors that bind more strongly to 14-3-3 compared to the mutant show higher peptide intensity in their heavy forms in the SILAC pair. In a reverse experiment the SILAC label was swapped and 14-3-3 wild-type was incubated with 'light' and mutant 14-3-3 with 'heavy' labeled SILAC. This results in the reversal of SILAC ratios for proteins that interact more strongly with 14-3-3 compared to the mutants.

Supplementary Tables:

Supplementary Table S1

List of all acetylation sites identified from MV4-11, Jurkat, and A549 cells. For description of columns see Online Materials and Methods. Acetylated proteins and sites were classified into different biological categories and are given in different sheets in the Excel table.

Supplementary Table S2

Acetylated proteins are present in large protein complexes. Enrichment of acetylated proteins in known mammalian protein complexes was analyzed using CORUM. This table list the protein complexes obtained from CORUM database, whose components were found to be significantly enriched in our acetylation dataset (over-represented, p-value <0.05). Complexes that were significantly acetylated were retained (p < 0.05). These complexes were further categorized based on their cellular localization (Nuclear, Cytoplasmic, Mitochondria, Others) as provided by CORUM annotations (Ruepp et al., 2008).

Supplementary Table S3

List of proteins that interact with 14-3-3 in an acetylation dependent manner. Proteins that interact strongly with 14-3-3 compared to its acetylation mimetic mutants (K50Q, or K118+123Q, or K50+K118+K123Q) were identified and quantified by SILAC-based 14-3-3 GST-pull down experiments as described in Fig. S9.

Supplementary Table S4

List of acetylation sites quantified in response to SAHA or MS-275.

Supplementary Table S5.

Acetylation sites upregulated by SAHA and MS-275 in MV4-11 cells. Acetylation changes of peptides upon inhibitor treatments were quantified as described in the supporting online material. Peptides with >2 fold increase in acetylation in response to both inhibitors are listed.

Description of supplementary table headers (Table S1, S5)

Header used in supplementary tables	Detailed description
Position	Position of acetylated amino acid in the identified protein. Positions are numbered according to International Protein Index (IPI) IDs.
Known Site	If a lysine acetylation site is previously known and annotated in Uniprot database as acetylated
Proteins	International Protein Index (IPI) ID of the proteins to which a acetylated matches to
Protein	International Protein Index (IPI) ID of the protein to which a acetylated matches to
Protein Names	Names of proteins to which the acetylated peptide sequence is assigned to
Gene Names	Name of gene to which indentified acetylated peptide is mapped to
Uniprot	Uniprot protein ID (Swiss-prot or TrEMBL identifier)
ENSEMBL	ENSEMBL gene ID
Pfam	Protein domains and families (Pfam) ID
Pfam name	Protein domains and families names
Pfam description	Description of Pfam domains and protein families
GOCC	Gene Ontology cellular compartment ID
GOCC name	Gene Ontology cellular compartment name
GOMF	Gene Ontology molecular function ID
GOMF name	Gene Ontology molecular function names
GOBP	Gene Ontology biological process ID
GOBP names	Gene Ontology biological process names
KEGG	Kyoto Encyclopedia of Genes and Genomes ID
KEGG pathway	Kyoto Encyclopedia of Genes and Genomes pathways
KEGG pathway names	Kyoto Encyclopedia of Genes and Genomes pathway names
KEGG Pathway orthologs	Kyoto Encyclopedia of Genes and Genomes pathway orthologs
Localization Prob	Localization probability of acetylation site assignment to a specific lysine in the identified peptide based on the PTM score
Score Diff	Delta PTM score between highest scoring acetylated site assignment and second highest scoring
Number of Acetyl (K)	Number of lysine acetylated residues present on the identified lysine acetylated peptide
Amino Acid	Amino acid that is acetylated i.e Lysine
Sequence Window	A 13 amino acid sequence sequence showing 6 amino acid flanking each side of lysine acetylated site

Mascot Score	Peptide ion MS/MS Probability-based Mowse score from Matrix Science: Mowse score = $-10 \times \log_{10}(p)$, where p is the likelihood that the identification is a random event
PTM Score	Probability-based phosphosite localization score based on the binominal distribution score: PTM score = $-10 \times \log_{10}(p)$, where p is the likelihood that the identification is a random event
Modified Sequence	Acetylated peptide sequence and modifications of amino acids on this peptide sequence that were identified by mass spectrometer; M(ox) = oxidized methionine, ac = N-Protein acetylated residue, K(ac) = acetylated residue.
Acetyl (K) Probabilities	Localization probability of acetylation site assignment to specific lysines in the identified peptide based on the PTM score
Charge	Charge state of the identified lysine acetylated peptide
m/z	Mass-to-charge ratio for the identified acetylated peptide
Mass Error (ppm)	Mass accuracy of identified peptide in parts per million
Ratio MS275/Ctrl.	SILAC Ratio of acetylated peptide intensities for cells treated with MS-275 versus control cells treated with DMSO
Ratio SAHA/Ctrl.	SILAC Ratio of acetylated peptide intensities for cells treated with MS-275 versus control cells treated with DMSO
<u>Abbreviations:</u>	<u>Explanation:</u>
Ctrl.	Control

Supplementary References

Bader, G.D., and Hogue, C.W. (2003). An automated method for finding molecular complexes in large protein interaction networks. *BMC bioinformatics* 4, 2.

Cheutin, T., McNairn, A.J., Jenuwein, T., Gilbert, D.M., Singh, P.B., and Misteli, T. (2003). Maintenance of stable heterochromatin domains by dynamic HP1 binding. *Science* (New York, NY) 299, 721-725.

Cox, J., and Mann, M. (2008). MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat Biotechnol.*

Falcon, S., and Gentleman, R. (2007). Using GOSTats to test gene lists for GO term association. *Bioinformatics* 23, 257-258.

Finn, R.D., Tate, J., Mistry, J., Coghill, P.C., Sammut, S.J., Hotz, H.R., Ceric, G., Forslund, K., Eddy, S.R., Sonnhammer, E.L., *et al.* (2008). The Pfam protein families database. *Nucleic acids research* 36, D281-288.

Flicek, P., Aken, B.L., Beal, K., Ballester, B., Caccamo, M., Chen, Y., Clarke, L., Coates, G., Cunningham, F., Cutts, T., *et al.* (2008). Ensembl 2008. *Nucleic acids research* 36, D707-714.

Gnad, F., Oroshi, M., Birney, E., and Mann, M. (2008). MAPU 2.0: high-accuracy proteomes mapped to genomes. *Nucleic acids research.*

Guelman, S., Kozuka, K., Mao, Y., Pham, V., Solloway, M.J., Wang, J., Wu, J., Lill, J.R., and Zha, J. (2009). The double-histone-acetyltransferase complex ATAC is essential for mammalian development. *Molecular and cellular biology* 29, 1176-1188.

Hubner, N.C., Ren, S., and Mann, M. (2008). Peptide separation with immobilized pl strips is an attractive alternative to in-gel protein digestion for proteome analysis. *Proteomics*.

Huh, W.K., Falvo, J.V., Gerke, L.C., Carroll, A.S., Howson, R.W., Weissman, J.S., and O'Shea, E.K. (2003). Global analysis of protein localization in budding yeast. *Nature* 425, 686-691.

Janke, C., Magiera, M.M., Rathfelder, N., Taxis, C., Reber, S., Maekawa, H., Moreno-Borchart, A., Doenges, G., Schwob, E., Schiebel, E., *et al.* (2004). A versatile toolbox for PCR-based tagging of yeast genes: new fluorescent proteins, more markers and promoter substitution cassettes. *Yeast (Chichester, England)* 21, 947-962.

Jensen, L.J., Kuhn, M., Stark, M., Chaffron, S., Creevey, C., Muller, J., Doerks, T., Julien, P., Roth, A., Simonovic, M., *et al.* (2008). STRING 8--a global view on proteins and their functional interactions in 630 organisms. *Nucleic acids research*.

Larkin, M.A., Blackshields, G., Brown, N.P., Chenna, R., McGettigan, P.A., McWilliam, H., Valentin, F., Wallace, I.M., Wilm, A., Lopez, R., *et al.* (2007). Clustal W and Clustal X version 2.0. *Bioinformatics (Oxford, England)* 23, 2947-2948.

Olsen, J.V., Blagoev, B., Gnad, F., Macek, B., Kumar, C., Mortensen, P., and Mann, M. (2006). Global, in vivo, and site-specific phosphorylation dynamics in signaling networks. *Cell* 127, 635-648.

Olsen, J.V., de Godoy, L.M., Li, G., Macek, B., Mortensen, P., Pesch, R., Makarov, A., Lange, O., Horning, S., and Mann, M. (2005). Parts per million mass accuracy on an Orbitrap mass spectrometer via lock mass injection into a C-trap. *Mol Cell Proteomics* 4, 2010-2021.

Olsen, J.V., Macek, B., Lange, O., Makarov, A., Horning, S., and Mann, M. (2007). Higher-energy C-trap dissociation for peptide modification analysis. *Nature methods* 4, 709-712.

Ong, S.E., Blagoev, B., Kratchmarova, I., Kristensen, D.B., Steen, H., Pandey, A., and Mann, M. (2002). Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. *Mol Cell Proteomics* 1, 376-386.

Rappsilber, J., Mann, M., and Ishihama, Y. (2007). Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. *Nature protocols* 2, 1896-1906.

Ruepp, A., Brauner, B., Dunger-Kaltenbach, I., Frishman, G., Montrone, C., Stransky, M., Waegle, B., Schmidt, T., Doudieu, O.N., Stumpflen, V., *et al.* (2008). CORUM: the comprehensive resource of mammalian protein complexes. *Nucleic acids research* 36, D646-650.

Schulze, W.X., and Mann, M. (2004). A novel proteomic screen for peptide-protein interactions. *The Journal of biological chemistry* 279, 10756-10764.

Shannon, P., Markiel, A., Ozier, O., Baliga, N.S., Wang, J.T., Ramage, D., Amin, N., Schwikowski, B., and Ideker, T. (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res* 13, 2498-2504.

Shevchenko, A., Tomas, H., Havlis, J., Olsen, J.V., and Mann, M. (2006). In-gel digestion for mass spectrometric characterization of proteins and proteomes. *Nature protocols* 1, 2856-2860.

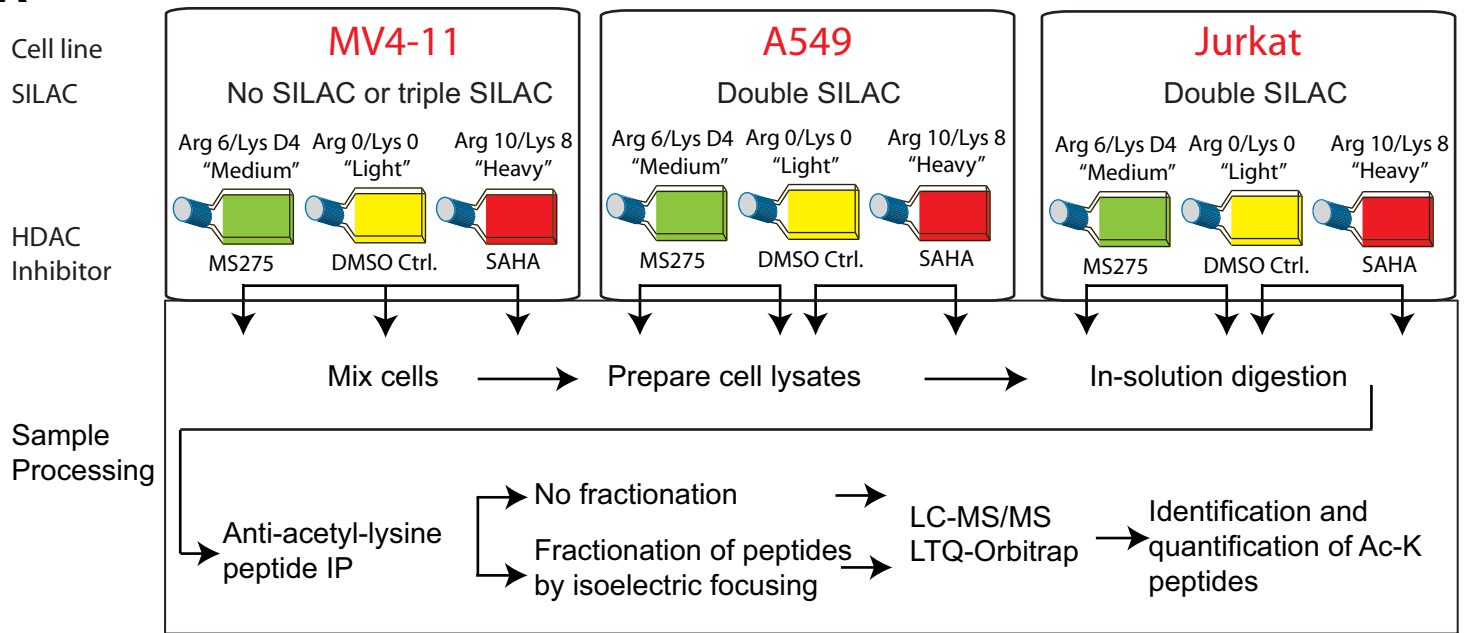
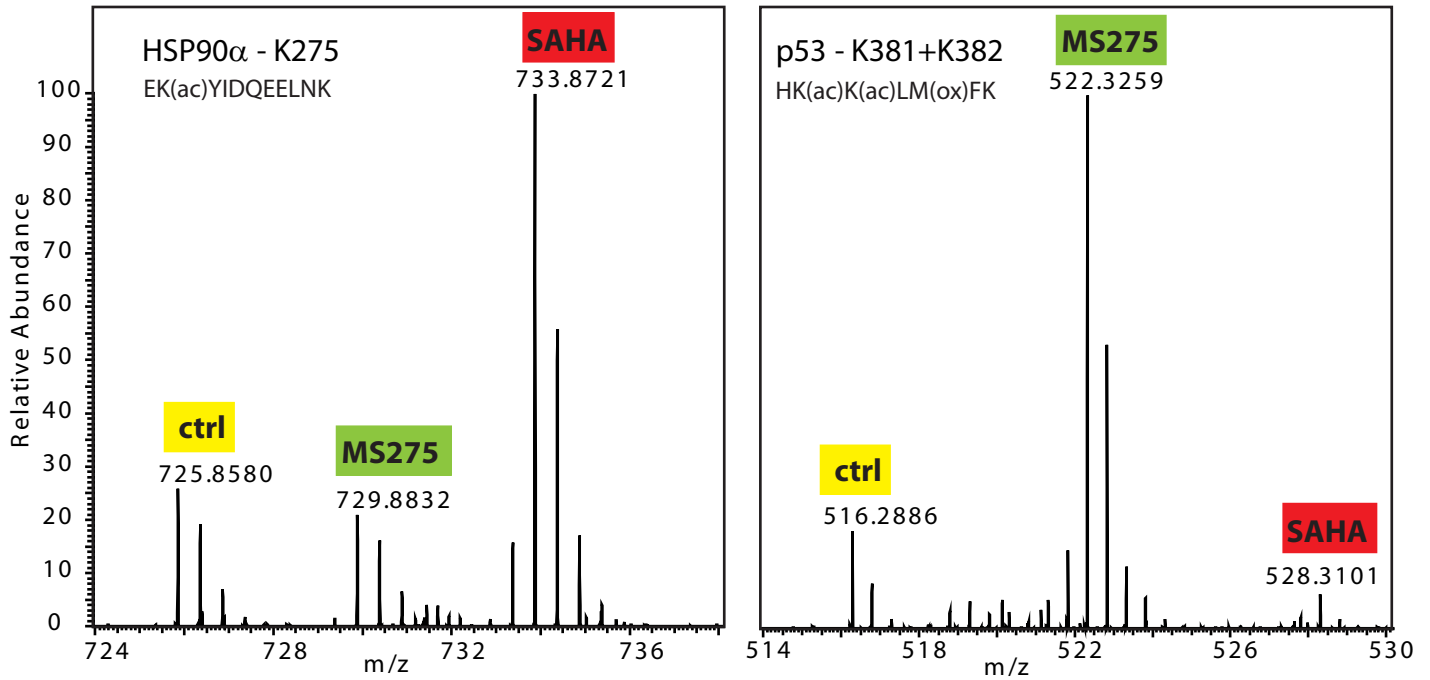
Shevchenko, A., Wilm, M., Vorm, O., and Mann, M. (1996). Mass spectrometric sequencing of proteins silver-stained polyacrylamide gels. *Analytical chemistry* 68, 850-858.

Sikorski, R.S., and Hieter, P. (1989). A system of shuttle vectors and yeast host strains designed for efficient manipulation of DNA in *Saccharomyces cerevisiae*. *Genetics* 122, 19-27.

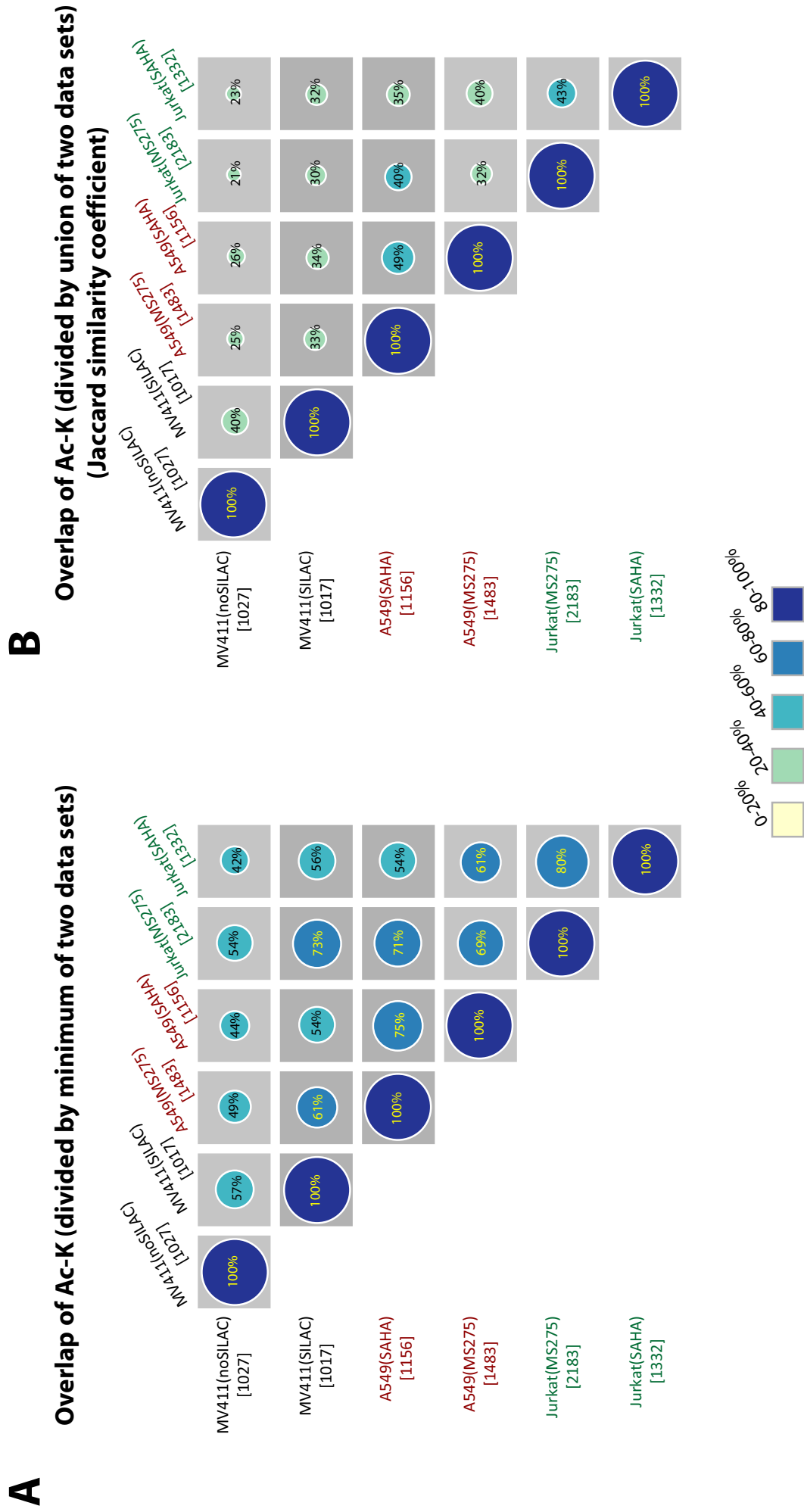
Suganuma, T., Gutierrez, J.L., Li, B., Florens, L., Swanson, S.K., Washburn, M.P., Abmayr, S.M., and Workman, J.L. (2008). ATAC is a double histone acetyltransferase complex that stimulates nucleosome sliding. *Nature structural & molecular biology* 15, 364-372.

Team, R.D.C. (2008). R: A Language and Environment for Statistical Computing (Vienna, Austria, R Foundation for Statistical Computing).

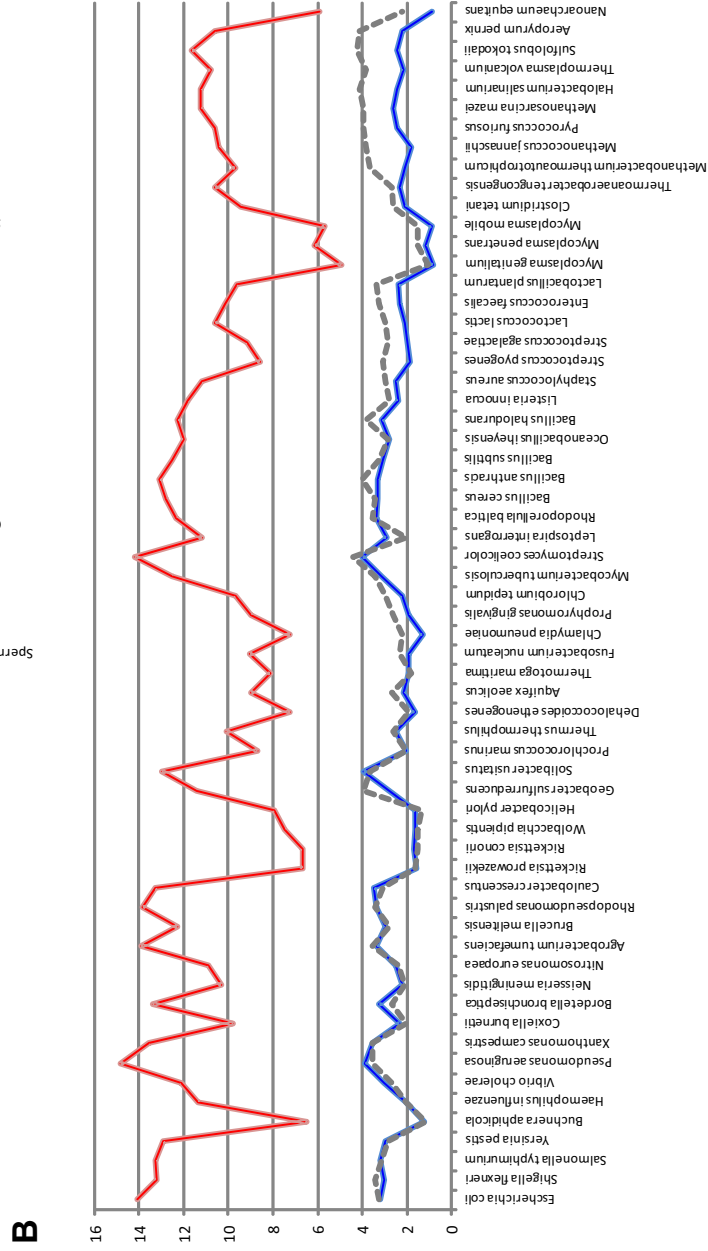
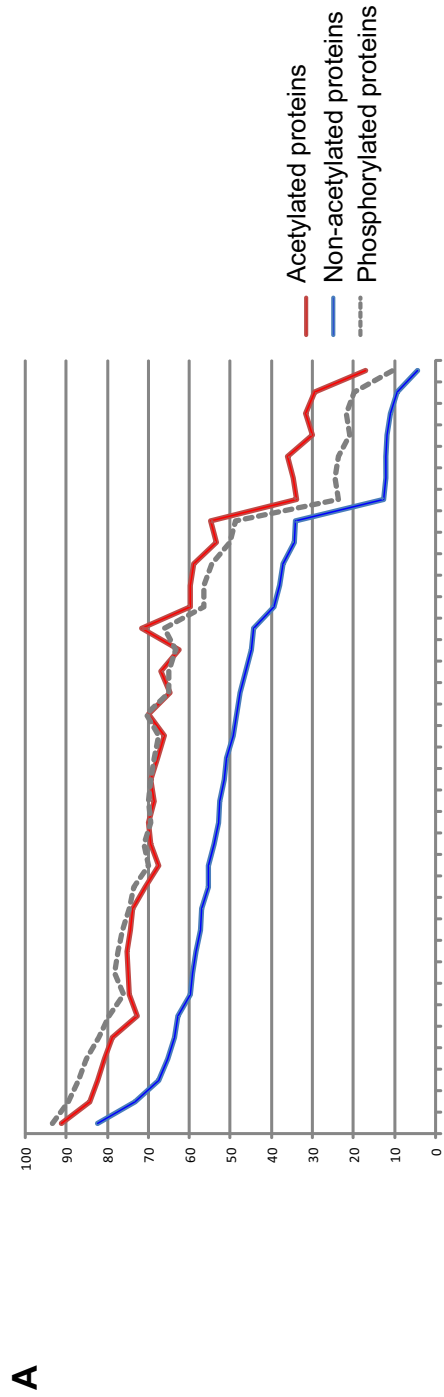
Urschel, S., Bassermann, F., Bai, R.Y., Munch, S., Peschel, C., and Duyster, J. (2005). Phosphorylation of grb10 regulates its interaction with 14-3-3. *The Journal of biological chemistry* 280, 16987-16993.

Figure_S1**A****B**

Figure_S2



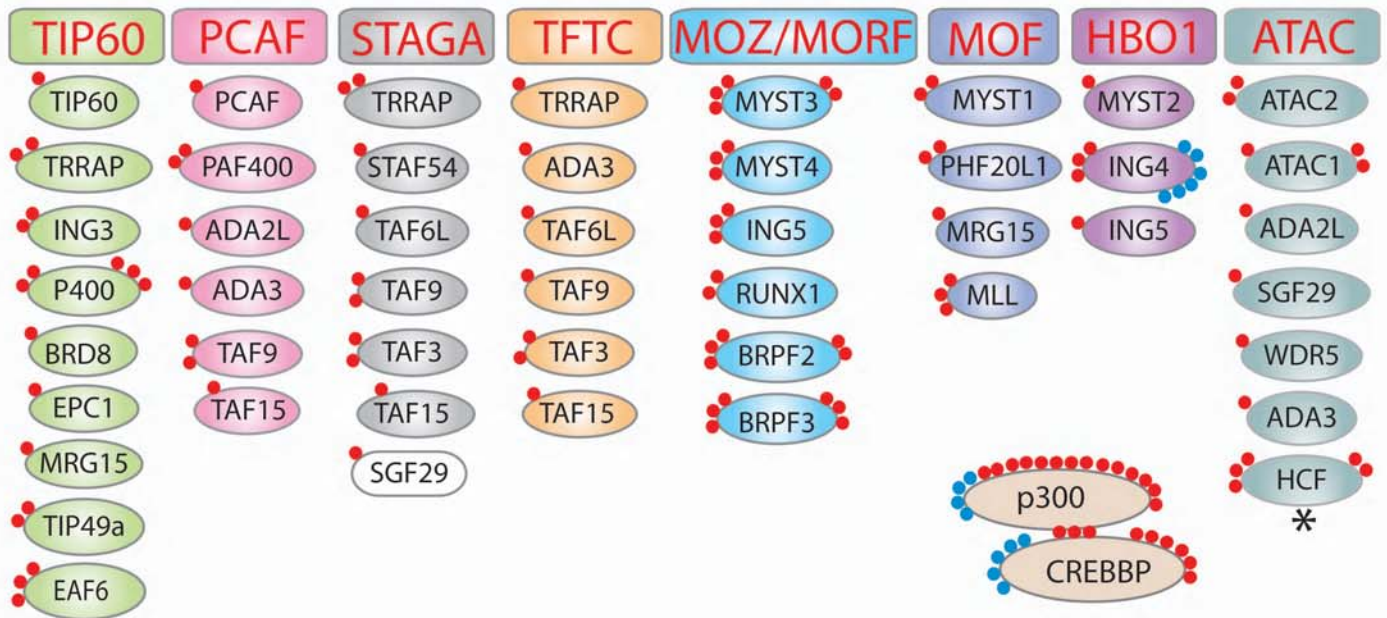
Figure_S3



Figure_S4

A

Acetylation of KAT complexes

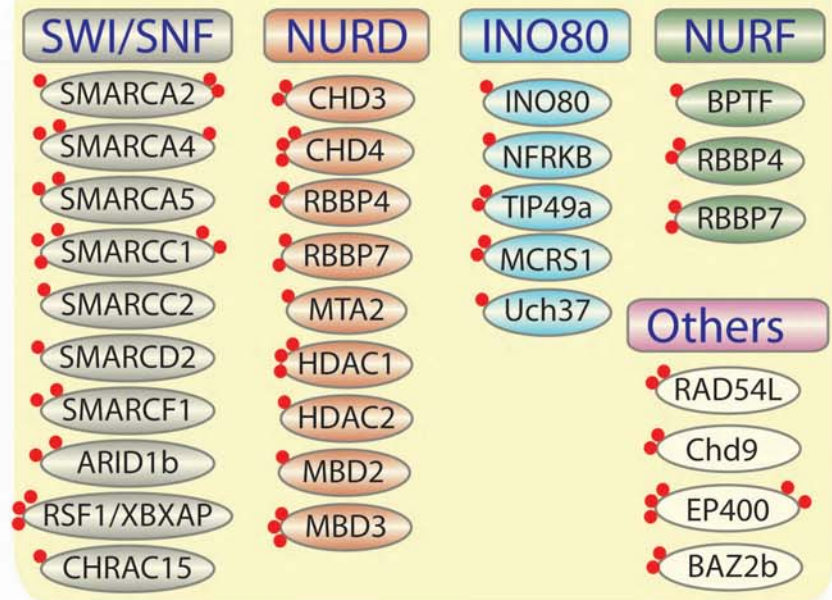


B



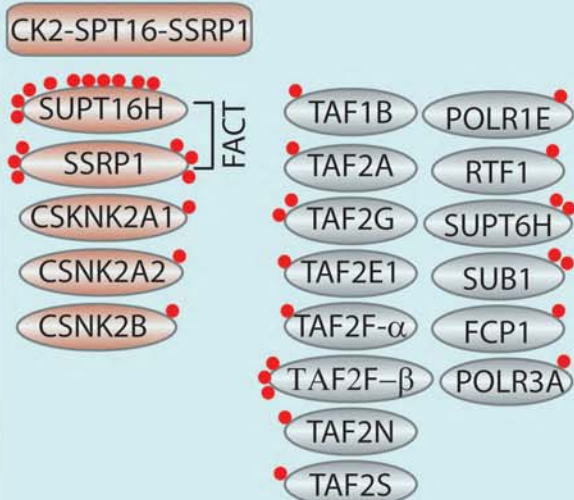
C

CHROMATIN REMODELING COMPLEXES



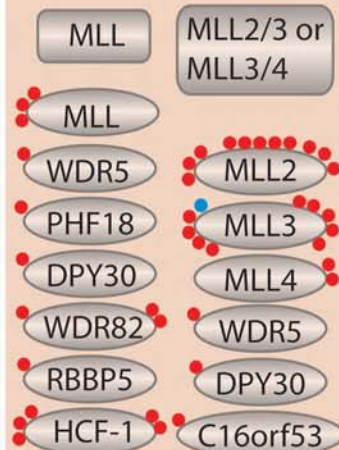
D

TRANSCRIPTIONAL REGULATION



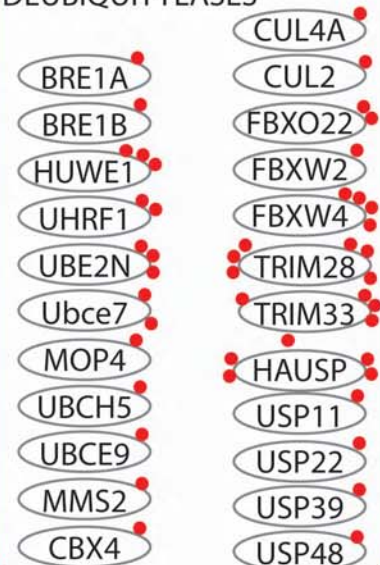
E

METHYLTRANSFERASE COMPLEXES

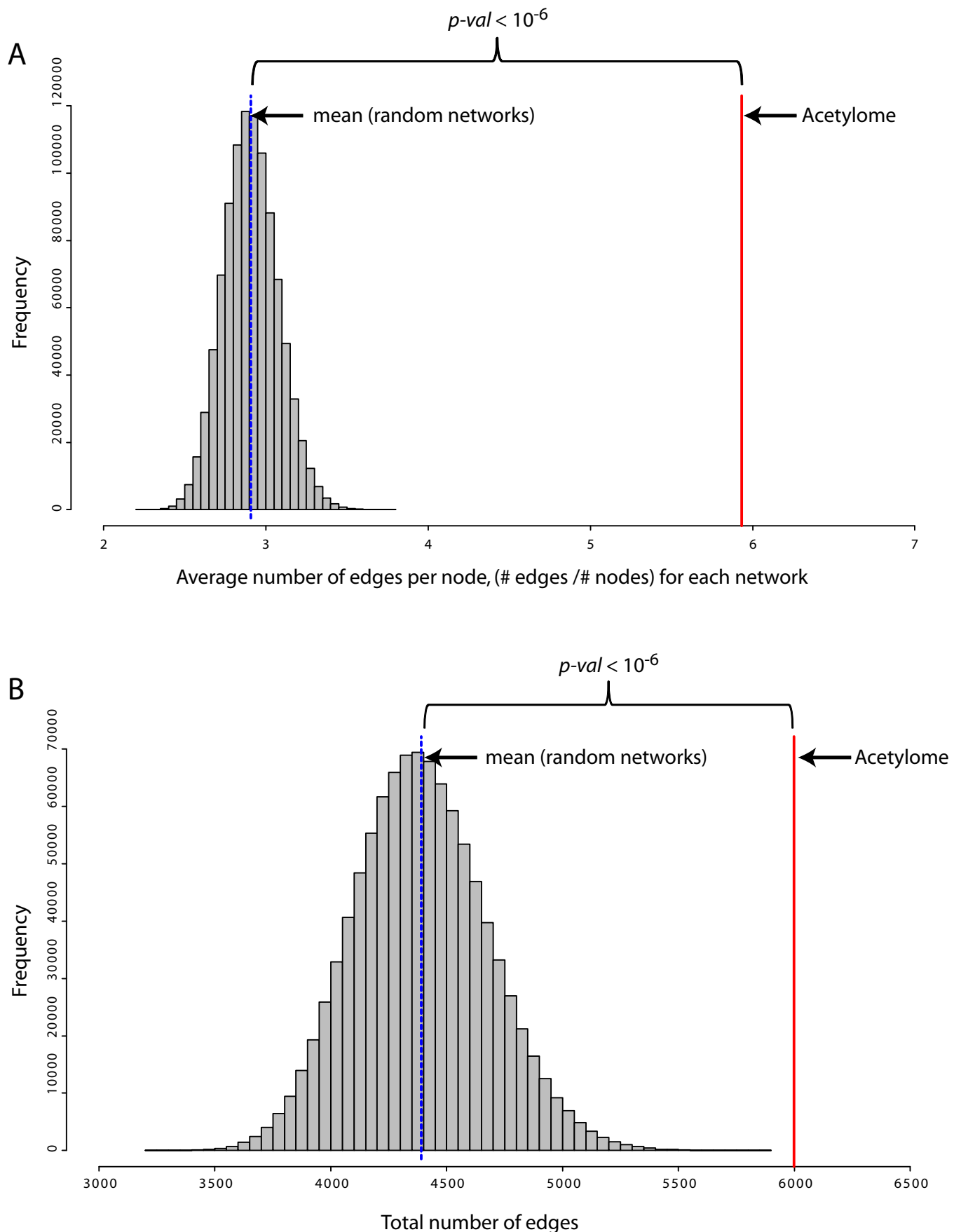


F

UBIQUITIN LIGASES AND DEUBIQUITYLASES



Figure_S5



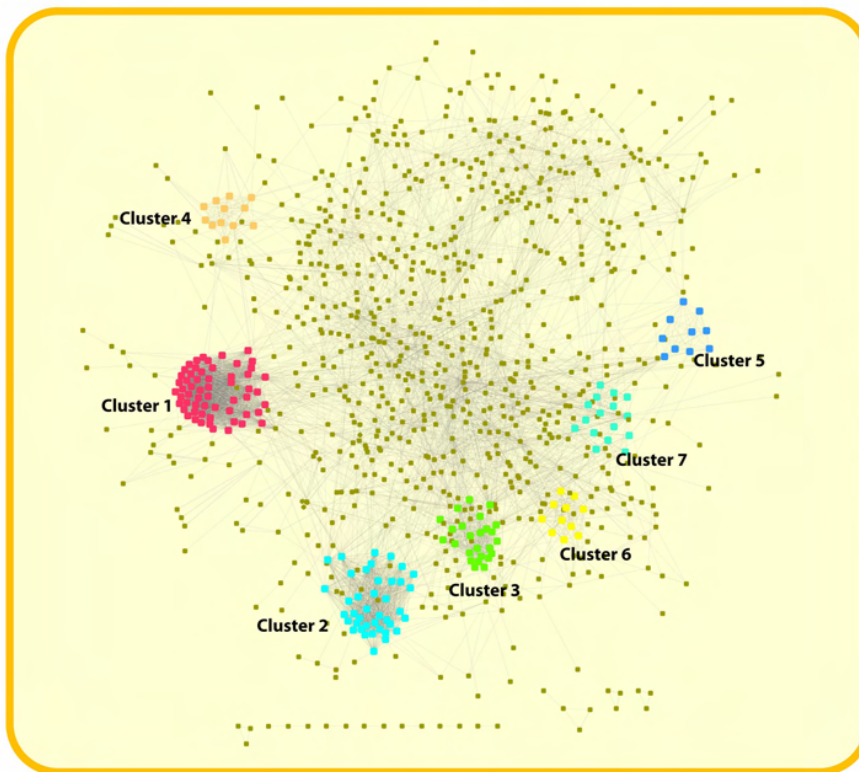
Summary statistics for acetylation interaction network connectivity

* node permutation based network randomizations

* calculated across 1,000,000 iterations

Figure_S6

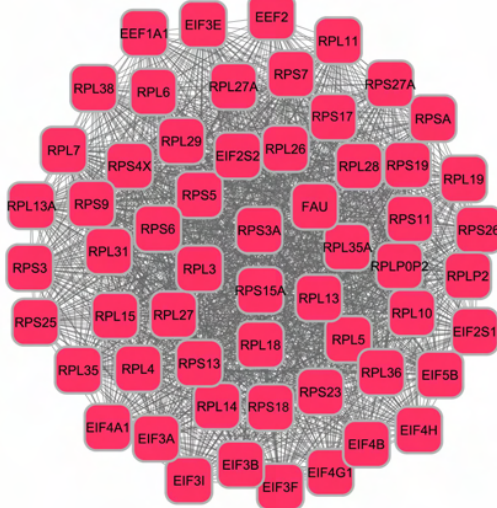
A



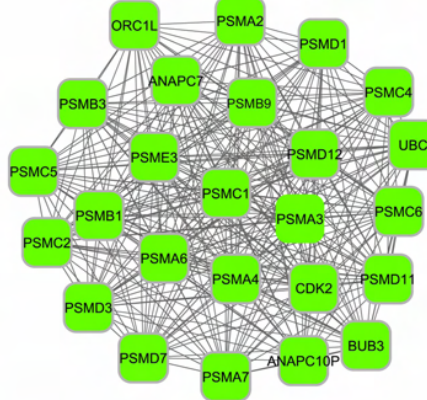
The complete acetylation interaction network(AIN) obtained from STRING with confidence score ≥ 0.7
No. of nodes = 1,011
No. of edges = 5,999

B

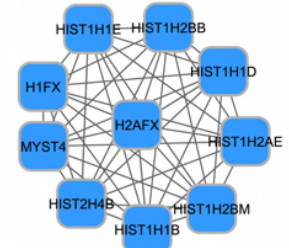
Cluster 1 (MCODE score = 28.121)
nodes = 58, edges = 1631



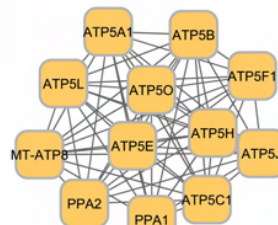
Cluster 3 (MCODE score = 11.84)
nodes = 25, edges = 296



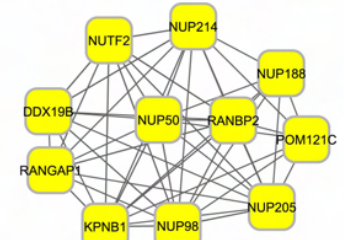
Cluster 5 (MCODE score = 4.5)
nodes = 10, edges = 45



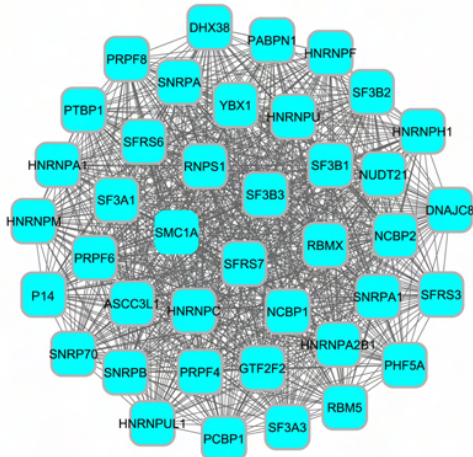
Cluster 4 (MCODE score = 5.417)
nodes = 12, edges = 65



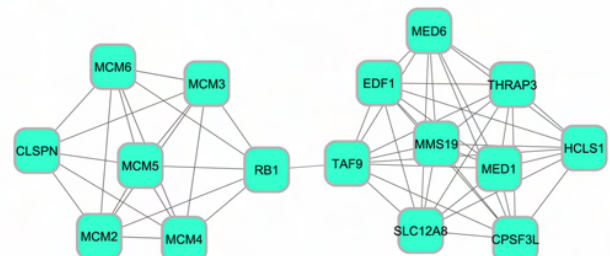
Cluster 6 (MCODE score = 4.45)
nodes = 11, edges = 49



Cluster 2 (MCODE score = 19.5)
nodes = 40, edges = 780

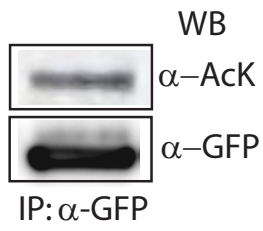


Cluster 7 (MCODE score = 3.567)
nodes = 16, edges = 57

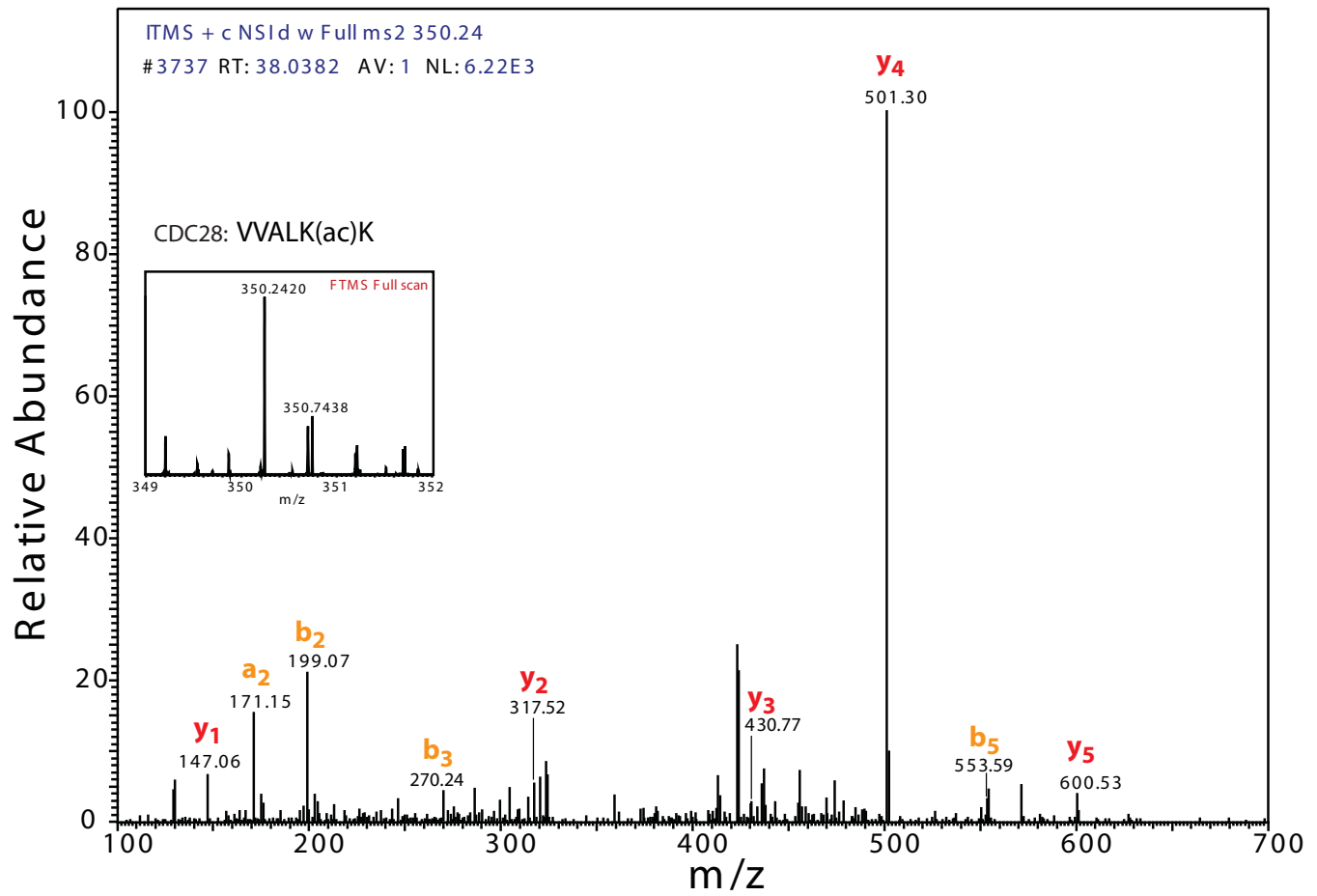


Figure_S7

A



B



Figure_S8

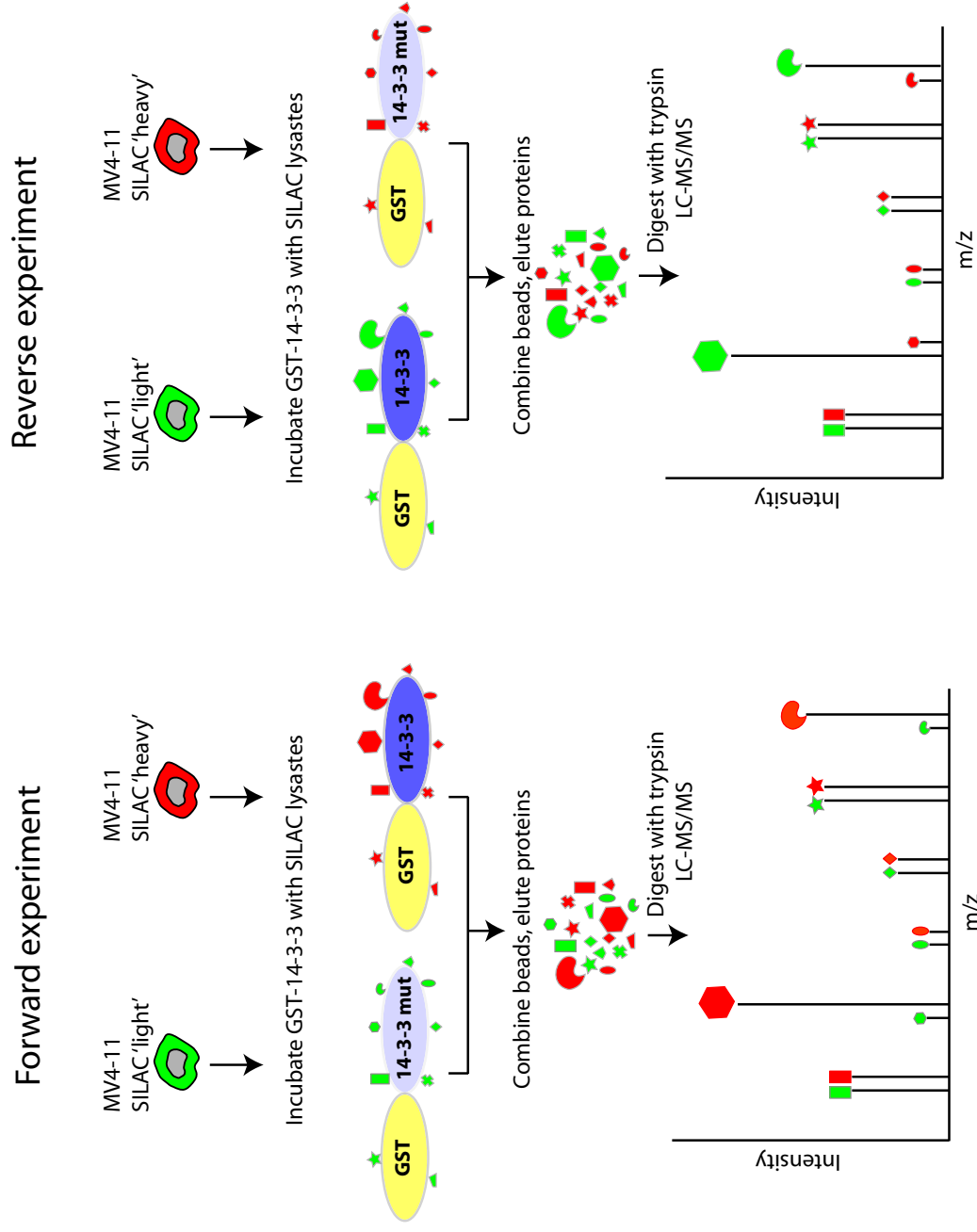
A

Modified peptide sequence	14-3-3 isoform (Ac-K)
(ac)MDK(ac)NELVQK	ζ/δ (K3)
(ac)MEK(ac)TELIQK	θ/τ (K3)
TELIQK(ac)AK	θ/τ (K9)
NLLSVAYK(ac)NVIGAR	ε (K50)
NLLSVAYK(ac)NVVGAR	ζ/δ (K49), η (K50), γ (K50), α/β (K51)
VISSIEQK(ac)TDTSDK	θ/τ (K68)
VVSSIEQK(ac)TEGAEK	ζ/δ (K68)
IISIEQK(ac)EENK	ε (K69)
VISSIEQK(ac)TER	α/β (K70)
YLIANATNPESK(ac)VFYLK	θ/τ (K115)
FLIPNASQAESK(ac)VFYLK	ζ/δ (K115)
YLIPNATQPESK(ac)VFYLK	α/β (K117)
HLIPAANTGESK(ac)VFYYK	ε (K118)
VFYLK(ac)MK	ζ/δ (K120), θ/τ (K120), σ (K122), α/β (K122), η (K125), γ (K125)
VFYYK(ac)MK	ε (K123)

B

14-3-3_zeta/delta	--MDKNELVQKAKLAEQAERYDDMAACMSVTEQGAELSNEERNLLSVAYKNVVGARRSS	58
14-3-3_beta/alpha	MTMDKSELVQKAKLAEQAERYDDMAAAMKAVTEQGHELSNEERNLLSVAYKNVVGARRSS	60
14-3-3_theta	--MEKTELIQKAKLAEQAERYDDMATCMKAVTEQGAELSNEERNLLSVAYKNVVGARRSA	58
14-3-3_sigma	--MERASLIQKAKLAEQAERYEDMAAFMKGAVEKGEELSCEERNLLSVAYKNVVGGQRAA	58
14-3-3_eta	-MGDREQLLQARLAEQAERYDDMASAMKAVTELNEPLSNEDRNLLSVAYKNVVGARRSS	59
14-3-3_gamma	-MVDREQLVQKARLAEQAERYDDMAAMKNVTELNEPLSNEERNLLSVAYKNVVGARRSS	59
14-3-3_epsilon	-MDDREDLVYQAKLAEQAERYDEMVESMKKVAGMDVELTVEERNLLSVAYKNVIGARRAS	59
14-3-3_zeta/delta	WRVSSIEQKT--EGAEEKQOMAREYREKIETELRDICNDVLSLLEKFLIPNASQA--ES	114
14-3-3_beta/alpha	WRVISSIEQKT--ERNEKKQMGKEYREKIEAELQDICNDVLELLDKYLIPNATQP--ES	116
14-3-3_theta	WRVISSIEQKT--DTSDDKLQLIKDYREKVESELRSICTTVLELLDKYLIANATNP--ES	114
14-3-3_sigma	WRVLSSIEQKSNEEGSEEKGPEVREYREKVELTELQGVCDTVLGLLDShlikeAGDA--ES	116
14-3-3_eta	WRVISSIEQKTMDAGNEKKLEKVKAYREKIEKELETVCNDVLSLLDKFLIKCNDFFQYES	119
14-3-3_gamma	WRVISSIEQKTSADGNEKKIEMVRAYREKIEKELEAVCQDVLSSLNDNYLIKNCSETQYES	119
14-3-3_epsilon	WRISSIEQKEENKGGEDLKLMIREYRQMVETELKLICCDILDVLDKHLIPAANTG--ES	117
14-3-3_zeta/delta	KVFYLMKMGDYRYRLAEVAGDDKKGIVDQSQAYQEAFEISKKEMQPTHPIRLGLALNF	174
14-3-3_beta/alpha	KVFYLMKMGDYFRYLSEVASGDNKQTTVSNSQQAYQEAFEISKKEMQPTHPIRLGLALNF	176
14-3-3_theta	KVFYLMKMGDYFRYLAEVACGDDRQKQIDNSQAYQEAFDISKKEMQPTHPIRLGLALNF	174
14-3-3_sigma	RVFYLKMGDYRYRLAEVATGDDKKRIIDSARSAYQEAMDISKKEMPPTNPPIRLGLALNF	176
14-3-3_eta	KVFYLMKMGDYRYRLAEVASGEKKNSVVEASEAAYKEAFEISKEQMPTHPIRLGLALNF	179
14-3-3_gamma	KVFYLMKMGDYRYRLAEVATGEKRATVVESEKAYSEAHEISKEHMQPTHPIRLGLALNF	179
14-3-3_epsilon	KVFYYKMGGDYHRYLAEFATGNDRKEAAENSLVAYKAASDIAMTELPPTHPIRLGLALNF	177
14-3-3_zeta/delta	SVFYIEILNSPEKACSLAKTAFDEAIAELDTLSEESYKDSTLIMQLLRDNLTLWTSDTQG	234
14-3-3_beta/alpha	SVFYIEILNSPEKACSLAKTAFDEAIAELDTLNEESYKDSTLIMQLLRDNLTLWTSENQG	236
14-3-3_theta	SVFYIEILNPELACTLAKTAFDEAIAELDTLNEDSYKDSTLIMQLLRDNLTLWTSDSAG	234
14-3-3_sigma	SVFYIEIANSPEEAISLAKTTFFDEAMADLHTLSEDSYKDSTLIMQLLRDNLTLWTADNAG	236
14-3-3_eta	SVFYIEIQNAPEQACLLAKQAFDDAIAELDTLNEDSYKDSTLIMQLLRDNLTLWTSDQQD	239
14-3-3_gamma	SVFYIEIQNAPEQACHLAKTAFDDAIAELDTLNEDSYKDSTLIMQLLRDNLTLWTSDQQD	239
14-3-3_epsilon	SVFYIEILNSPDRACRLAKAFAFDDAIAELDTLSEESYKDSTLIMQLLRDNLTLWTSDMQG	237
14-3-3_zeta/delta	---DEA-EAGEGGEN---	245
14-3-3_beta/alpha	---DEG-DAGEG-----	244
14-3-3_theta	---EEC-DAAEGAEN---	245
14-3-3_sigma	---EEGGEAPQEPQSEN-	250
14-3-3_eta	---EEA---GEGN-----	246
14-3-3_gamma	---DDG---GEGNN----	247
14-3-3_epsilon	DGEEQNKEALQDVEDENQ	255

Figure_S9



Supplementary Table S3

Gene Names	Ratio 14-3-3-WT/14-3-3-K50Q	Ratio 14-3-3-K50Q/14-3-3-WT	Ratio 14-3-3-WT/14-3-3-K118+K123Q	Ratio 14-3-3-K5118+K123Q/14-3-3-WT	Ratio 14-3-3-WT/14-3-3-K50+K118+K123Q	Ratio 14-3-3-K50+K118+K123Q/14-3-3-WT
SFRS1	9.81	0.06	4.19	3.23	9.18	0.07
USP8	9.51	0.04	NA	NA	NA	NA
HDAC4	9.32	0.07	7.17	0.07	16.69	0.02
SFRS6	6.71	0.10	4.15	3.30	6.81	0.16
SRRM2	5.89	0.08	5.61	0.14	9.76	0.07
LARP1	5.62	0.10	5.59	0.21	14.50	0.04
SFRS4	5.19	0.10	4.25	0.20	6.81	0.08
NOLC1	4.90	0.09	NA	NA	7.65	NA
SFRS10	4.03	0.17	3.34	7.89	5.27	0.16
ARAF	3.67	0.15	4.28	0.12	8.56	0.06
PLEKHO2	3.39	0.12	3.48	0.14	5.02	0.15
HMHA1	3.07	0.19	NA	NA	NA	NA
GBF1	2.66	0.23	5.09	0.09	NA	0.14
SFRS7	2.49	0.20	2.61	2.83	1.74	0.50
NCBP1	2.41	0.16	NA	NA	NA	NA
DOK3	2.41	0.21	1.65	3.23	15.03	0.04
VASP	2.19	0.23	0.47	1.47	3.23	0.19
C1QBP	2.01	0.33	1.61	0.98	3.20	0.26
FLJ00087	1.24	0.46	NA	0.39	2.19	0.32
C22orf9	1.18	NA	1.11	2.75	4.05	0.25
GAB2	1.18	0.28	2.93	0.21	16.29	0.10
KIAA0226	0.80	0.44	NA	0.12	3.34	0.12
TBC1D1	0.67	0.57	NA	0.27	11.11	0.04
PIK3R4	0.63	0.45	NA	0.12	2.61	0.21
SGK223	0.62	0.37	3.38	0.02	6.44	NA
RCHY1	0.60	NA	2.75	0.47	NA	NA
TSC1	0.44	0.96	NA	0.05	21.71	0.05
TSC2	0.37	1.10	7.19	0.09	12.97	0.06
DENND4A	0.27	1.17	1.06	0.74	4.89	0.11
IGHG1	0.04	NA	0.01	NA	8.30	0.05
RAF1	NA	0.14	6.35	0.13	NA	0.08
UVRAG	NA	0.41	NA	0.10	3.40	0.15
SFRS2	NA	0.18	2.25	3.65	NA	0.41

Table S5

Position	Protein	Gene Names	Uniprot	Modified Sequence	Ratio MS275/Ctrl.	Ratio SAHA/Ctrl.
K647	IPI00444646	AP1GBP1	A6NEV4	ALPSMDK(ac)YAVFK	2.7	20.9
K27	IPI00852708	BAT2D1	Q9HA05	YATLSLFNTYK(ac)GK	3.8	9.1
K108	IPI00328798	BCL9L	Q67FY1	NPQAGVPPFSSLK(ac)GK	14.1	5.5
K149	IPI00657642	CCDC75	Q8N954-1	AEEK(ac)LESYR	5.9	2.6
K556	IPI00456970	CDC2L5	Q14004-1	VENNLIVDK(ac)ATK	3.3	2.7
K6	IPI00025974	CHMP4B	Q9H444	(ac)SVFGK(ac)LFGAGGGK	2.8	2.5
K108	IPI00442073	CSRP1	P21291	HEEAPGHRPTTNPNASK(ac)FAQK	2.5	2.5
K328	IPI00298058	DHTKD1	Q96HY7	DWFAK(ac)R	3.8	2.0
K9	IPI00059279	EXOC4	Q96A65	(ac)AAEAAGGK(ac)YR	3.4	2.3
K5*	IPI00218448	H2AFZ	P0C0S5	AGGK(ac)AGK(ac)DSGK(ac)AK	2.7	19.6
K8*	IPI00218448	H2AFZ	P0C0S5	(ac)AGGKAGK(ac)DSGK(ac)AK	2.7	19.6
K12*	IPI00218448	H2AFZ	P0C0S5	(ac)AGGKAGK(ac)DSGK(ac)AK	2.7	19.6
K14*	IPI00218448	H2AFZ	P0C0S5	AGK(ac)DSGK(ac)AK(ac)TK	2.7	19.6
K13*	IPI00220403	HIST1H2BB	P33778	K(ac)GSK(ac)K(ac)AITK	9.2	14.2
K6*	IPI00453473	HIST1H4A	P62805	GK(ac)GGK(ac)GLGK	6.8	27.2
K9*	IPI00453473	HIST1H4A	P62805	GGK(ac)GLGK	2.3	59.4
K13*	IPI00329665	HIST2H2BF	Q5QNW6	K(ac)GSK(ac)K(ac)AVTK	9.2	14.2
K21*	IPI00329665	HIST2H2BF	Q5QNW6	AVTK(ac)VQK	2.8	2.4
K24	IPI00455457	HIST2H3PS2	Q5TEC6	QLATK(ac)AAR	2.2	3.3
K15*	IPI00216402	HIST3H3	Q16695	STGGK(ac)APR	3.1	7.2
K10*	IPI00216402	HIST3H3	Q16695	K(ac)STGGK(ac)APR	3.1	7.2
K15	IPI00179700	HMGA1	P17096-2	SSQPLASK(ac)QEK	3.0	2.5
K82	IPI00419258	HMGB1	P09429	TYIPPK(ac)GETK	20.1	4.4
K460	IPI00418313	ILF3	Q12906-4	LHVAVK(ac)VLQDMGLPTGAEGR	2.1	4.6
K1148	IPI00785015	KIAA2030	Q6ZU65-1	NLQAPSK(ac)LTNSSSTGTGK	7.6	2.2
K729	IPI00847241	LARP5	A6NEL6	RPAGGRPPSPSAMGK(ac)R	2.3	8.9
K88	IPI00010471	LCP1	P13796	STDVAK(ac)TFR	2.5	2.6
K812	IPI00552897	MDC1	Q14676-1	QTVDK(ac)VMGIPIK	2.4	2.7
K241	IPI00102495	MED6	O75586	NVQQTVSAK(ac)GPPEK(ac)R	3.5	2.2
K236	IPI00102495	MED6	O75586	NVQQTVSAK(ac)GPPEK(ac)R	3.5	2.2
K102	IPI00604620	NCL	P19338-1	TVTPAK(ac)AVTTPGK	10.4	3.7
K577	IPI00604620	NCL	P19338-1	TLFVK(ac)GLSEDTTEETLK	2.8	2.6
K572	IPI00604620	NCL	P19338-1	SQPSK(ac)TLFVK	2.2	2.4
K5	IPI00304596	NONO	Q15233	(ac)MQSNK(ac)TFNLEK	2.1	2.3
K23	IPI00646917	NUDT21	O43809	GVTQFGNK(ac)YIQQTK	2.7	2.2
K150	IPI00291608	PCNP	Q8WW12-1	DTPTSAGPNSFNK(ac)GK	10.0	2.3
K165	IPI00011268	RALY	Q53GL6	TNVPVK(ac)LFAR	3.0	6.9
K895	IPI00871780	RBM12B	A8MYB5	SHRPFVNFGRPEGGK(ac)FDFGK	4.5	2.1
K542	IPI00005036	RBM5	P52756	TAQQIAK(ac)DMER	3.5	2.5
K295	IPI00245135	RNASEH2B	Q8N451	SGM(ox)K(ac)SIDTFFGVK	3.2	2.4
K167	IPI00020127	RPA1	P27694	TFGK(ac)AAGPSLSHTSGGTQSK	4.9	5.0
K185	IPI00003377	SFRS7	Q16629-1	SRSGSIK(ac)GSR	3.0	114.2
K354	IPI00797830	SMARCC1	Q58EY4	GQASLYGK(ac)R	17.8	2.4

*Known acetylation sites