Molecular Systems Biology 7; Article number 503; doi:10.1038/msb.2011.40
Citation: ***Molecular Systems Biology* 7: 503**

www.molecularsystemsbiology.com

molecular systems biology

Combinatorial depletion analysis to assemble the network architecture of the SAGA and ADA chromatin remodeling complexes

Kenneth K Lee[1,4], Mihaela E Sardiu[1,4], Selene K Swanson[1], Joshua M Gilmore[1], Michael Torok[2], Patrick A Grant[2], Laurence Florens[1], Jerry L Workman[1] and Michael P Washburn[1,3,*]

[1] Stowers Institute for Medical Research, Kansas City, MO, USA, [2] Department of Biochemistry and Molecular Genetics, University of Virginia School of Medicine, Charlottesville, VA, USA and [3] Department of Pathology and Laboratory Medicine, The University of Kansas Medical Center, Kansas City, KS, USA
[4] These authors contributed equally to this work
* Corresponding author. Stowers Institute for Medical Research, 1000 E. 50th Street, Kansas City, MO 64110, USA. Tel.: +1 816 926 4392; Fax: +1 816 926 4692; E-mail: mpw@stowers.org

Received 24.9.10; accepted 22.5.11

Despite the availability of several large-scale proteomics studies aiming to identify protein interactions on a global scale, little is known about how proteins interact and are organized within macromolecular complexes. Here, we describe a technique that consists of a combination of biochemistry approaches, quantitative proteomics and computational methods using wild-type and deletion strains to investigate the organization of proteins within macromolecular protein complexes. We applied this technique to determine the organization of two well-studied complexes, Spt–Ada–Gcn5 histone acetyltransferase (SAGA) and ADA, for which no comprehensive high-resolution structures exist. This approach revealed that SAGA/ADA is composed of five distinct functional modules, which can persist separately. Furthermore, we identified a novel subunit of the ADA complex, termed Ahc2, and characterized Sgf29 as an ADA family protein present in all Gcn5 histone acetyltransferase complexes. Finally, we propose a model for the architecture of the SAGA and ADA complexes, which predicts novel functional associations within the SAGA complex and provides mechanistic insights into phenotypical observations in SAGA mutants.
Molecular Systems Biology **7**: 503; published online 5 July 2011; doi:10.1038/msb.2011.40
Subject Categories: proteomics; chromatin & transcription
Keywords: ADA; architecture; protein interaction network; quantitative proteomics; SAGA

Introduction

Many proteins within cells do not function as individual activities, but associate with specific partners to form multi-subunit modules with specific functions. These in turn may associate with other functional modules to form a multi-functional macromolecular complex. While the identification of subunits of such complexes can be achieved through a combination of protein purification and proteomics, it is more challenging to ascertain how individual subunits interact and are spatially arranged within these macromolecular complexes. High-resolution characterization of multi-protein assemblies using any single experimental or computational method is generally very difficult, especially since traditional methods such as X-ray crystallography or NMR have certain limitations in characterizing large dynamic protein complexes. However, even if it is not feasible to determine the structure of whole protein complexes at atomic or amino-acid levels, methods predicting lower-resolution macromolecular models that accurately position proteins and their connections will accelerate our understanding of protein complexes and their cellular functions. Here, we describe a method capable of determining the architectural organization of multi-protein complexes. It employs a combination of computational approaches and a systematic collection of quantitative proteomics data from wild-type and deletion strain purifications. We applied this approach on a data set generated in this study, which aims to gain novel insights into the *Saccharomyces cerevisiae* Spt–Ada–Gcn5 histone acetyltransferase (HAT) (SAGA) complex.

SAGA is a well-studied multi-protein complex involved in regulating histone post-translational modifications. Originally identified in yeast, the SAGA complex was subsequently shown to be evolutionarily conserved in every organism through humans (Lee and Workman, 2007). Early on, through the use of genetics and conventional biochemistry approaches, SAGA was recognized to be a multi-protein complex that is made up of smaller functional modules (Figure 1A) (Grant *et al*, 1997, 1998, 1999; Sterner *et al*, 1999). The HAT module, which carries out the HAT activity of the SAGA complex, was the first module to be described and its catalytic subunit Gcn5 was shown to harbor limited substrate recognition and specificity (Grant *et al*, 1999). Subsequently, the Ada2 and Ada3 proteins were shown to also be part of this module (Horiuchi *et al*, 1997; Saleh *et al*, 1997; Balasubramanian *et al*, 2002). Early work already recognized the existence of three

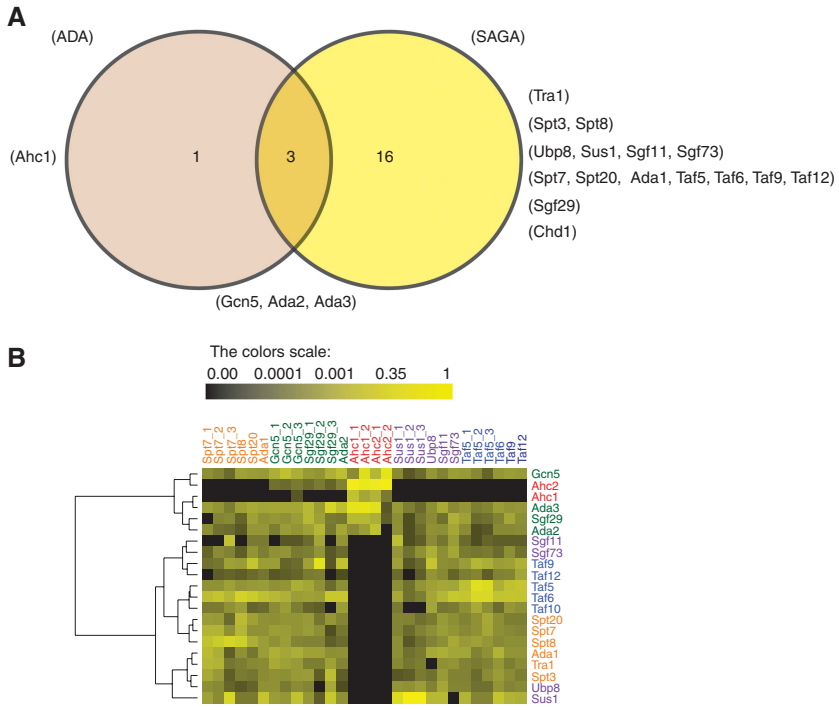

Figure 1 Proteomic analysis of wild-type purifications. (**A**) Venn diagram of previous knowledge of SAGA/ADA complexes: Using information obtained from the literature, the SAGA and ADA complexes were represented in a Venn diagram to indicate shared and specific proteins for the respective complexes. The SAGA/ADA complexes consist of distinct modules as shown by previous work, which are the recruitment module (Tra1), the acetylation module (Gcn5, Ada3 and Ada2), the TBP interaction unit (Spt3 and Spt8), the DUB module (Ubp8, Sgf11, Sgf73 and Sus1), the architecture unit (Spt7, Spt20, Ada1, Taf5, Taf6, Taf9, Taf10 and Taf12), a single subunit (Sgf29), a single subunit (Chd1) and the ADA module subunit (Ahc1) (reviewed in Koutelou *et al*, 2010). The numbers inside of the diagram represent the number of the proteins shared between the complexes. (**B**) Hierarchical clustering on the wild-type purifications. Hierarchical clustering analysis using WARD algorithm and Pearson correlation as distance metric was performed on the relative protein abundances expressed as dNSAFs normalized on the subunits of the SAGA/ADA complexes. Each column represents an isolated purification, and each row represents an individual protein (prey). The color intensity depicts the protein abundance with the brightest yellow indicating highest abundance and decreasing intensity indicating decreasing abundance. Black indicates that the protein was not detected in a particular sample. The HAT module is colored in green, the DUB module colored in violet, the SA_SPT module in orange, the SA_TAF module in blue and the two proteins unique to the ADA module were colored in red.

distinct Gcn5-containing complexes that have since been characterized as SAGA, a variant of the SAGA complex, named SLIK/SALSA, and ADA (Grant *et al*, 1997). All three complexes share the Gcn5/Ada2/Ada3 HAT module. SAGA and SLIK also share all other subunits with the exception of a C-terminal truncated form of Spt7 and Spt8 (Pray-Grant *et al*, 2002; Sterner *et al*, 2002). On the other hand, only a single unique subunit, Ahc1, was known to exist in the ADA complex (Eberharter *et al*, 1999) in addition to the HAT module. More recently, a second catalytic module, the deubiquitinylation (DUB) module, was identified within SAGA/SLIK (SALSA), which is important for the DUB of histone H2B (Henry *et al*, 2003; Daniel *et al*, 2004). Work from many laboratories has led to the identification of several subunits of this module, that is Ubp8, Sgf11, Sus1 and Sgf73 (Ingvarsdottir *et al*, 2005; Lee *et al*, 2005, 2009; Kohler *et al*, 2006, 2008). In addition, Chd1 was shown to be part of SAGA (Pray-Grant *et al*, 2005); however, it was not identified in our purifications.

Due to the complexity of the SAGA/ADA protein complex network, we reasoned that it is an ideal system to test our approach. Furthermore, partial structural information has been established for the SAGA complex, which therefore provides an objective to evaluate our method. Using electron microscopy (EM), Wu *et al* (2004) determined the first low-resolution 3D model of the SAGA complex; however, this study only localized 9 of the 19 known subunits of SAGA and the DUB module was not known to be part of SAGA at that time. On the other hand, two recent studies also determined the high-resolution structure of the four subunits of the DUB module (Kohler *et al*, 2010; Samara *et al*, 2010). Since these studies characterized only portions of the SAGA complex, there is no complete model for the architecture of SAGA. Here, we aimed to improve our understanding of the organization of proteins within the complex as well as to identify any components missing from earlier studies.

Using our method, we confirmed all known components of the DUB and HAT modules, and furthermore revealed that the HAT module contains an additional protein, Sgf29, that is present in all Gcn5 complexes. Sgf29 mutants resemble those in Ada2, Ada3 and Gcn5 by displaying classic ADA phenotypes (Berger *et al*, 1992). We also identified a novel subunit of the ADA complex, which we termed Ahc2. The most intriguing observation revealed through our analysis is that the SAGA complex consists of five distinct modules. In addition to the previously described DUB, HAT/Core and ADA modules, we

identified two novel modules, which we termed SA_SPT (i.e. Saga-associated SuPpressors of Ty) and SA_TAF (i.e. Saga-associated TATA-binding protein-associated factors). Unexpectedly, these modules, which are responsible for the different functions of the SAGA complex, are capable of assembling independently from the remaining modules of the complex.

Results

Data generation for the wild-type HAT complex

A total of 15 different SAGA subunits and 2 specific ADA components were TAP tagged (hereafter referred to as 'baits'), expressed and purified by affinity purification (Supplementary Tables S1 and S2). The proteins bound to the respective subunits (i.e. 'prey' protein) were analyzed by multidimensional protein identification technology (MudPIT) (Swanson *et al*, 2009) and quantified using the distributed normalized spectral abundance factors (dNSAF) (Zhang *et al*, 2010). Since the main focus of our study is on the Gcn5 HAT complexes, we concentrated on the 21 components of the SAGA and/or ADA complexes and used these subunits for further analysis. The remaining proteins identified in the purifications are reported in Supplementary Table S2. To ensure the specificity of the prey subunits (pulled-down proteins) in each bait, we extracted non-specific proteins (contaminants) from the data by comparing the dNSAF value in each of the individual purifications with the dNSAF value from a mock control (see Supplementary information). We also ensured the reproducibility of the data set by performing multiple replicates of subunits located in different parts of the SAGA complex (Figure 1B; Supplementary Figure S1; Supplementary Tables S1–S3). Finally, a 29×21 matrix was constructed consisting of the dNSAF values for each of the 21 subunits of the complex (Figure 1B).

Since the SAGA complex consists of different functional modules (reviewed in Koutelou *et al*, 2010), we sought to determine whether a quantitative proteomics data set generated from wild-type purifications is sufficient to discern the different modules of the SAGA protein complex and to assign proteins of unknown function to the respective modules. One popular method to analyze proteomics data is to hierarchically cluster proteins based on their relative abundance level (Sardiu *et al*, 2009a). We therefore subjected the 29×21 matrix to hierarchical clustering analysis in order to identify groups of proteins that show similar abundance levels (Figure 1B). However, the dendrogram obtained from the hierarchical clustering analysis did not indicate a clear separation of the proteins into different trees, and therefore did not separate the proteins into the different modules. This is a consequence of the fact that all the dNSAF values in the wild-type network have very similar values, reflecting the stability of the intact complex.

In spite of this, novel observations were nevertheless generated from the wild-type clustering. First, a previously uncharacterized protein, YCR082W, which we termed Ahc2, was found in close proximity to Ahc1, indicating its association within the complex (Figure 1B). In addition, Ahc2 co-purified with the components of the ADA complex (Figure 1B). These results suggest that Ahc2 protein is a novel component of the ADA complex. Ahc1 and Ahc2 were only detected when components of the HAT/Core module were used as baits. Furthermore, the baits Ahc1 and Ahc2 only co-purified components of the ADA complex. Next, a protein of unknown function, Sgf29, had a similar abundance level as known subunits of the HAT/Core module and also co-purified with the proteins Ahc1 and Ahc2 (Figure 1B), also indicating its association with the ADA complex. However, additional experiments were carried out to support these novel observations.

Quantitative analysis of deletion purifications

The architecture of protein complexes can reveal important principles of cellular organization and function. The separation and the proper identification of local modules within complexes remain an outstanding problem for proteomic analysis and toward this end few methods have been developed (Sardiu *et al*, 2009b). For example, the use of a single TAP-tagged protein in different deletion strains followed by mass spectrometry (i.e. proteins dependent on the deleted protein no longer co-purify with the bait) greatly improved the insights into the modularity and interrelationship of subunits in a protein complex (Mitchell *et al*, 2008; Sardiu *et al*, 2009b). However, certain limitations exist with this method. The major constraint is that all the results obtained using a single TAP-tagged bait and different deletions can only be interpreted relative to the protein that was TAP tagged and only local information proximal to the TAP-tagged bait can be obtained.

In an effort to overcome this limitation and to comprehensively identify the protein modularity and protein interrelationships within the Gcn5 HAT complexes, we applied a more unbiased comparative approach where individual components of SAGA were deleted and combined with different TAP-tagged proteins used as baits. The rationale behind the collection of the deleted proteins and the baits was based both on known and driven (i.e. based upon observations made in this study) biology of the SAGA/ADA complexes as follows: For deletion, we selected different subunits from each of the two known functional modules (i.e. DUB and HAT/Core) as well as different subunits from outside of these modules and combined them with different baits for TAP purification (Figure 2A). In addition, since Sgf29 was a protein of unknown function, we included this deletion in the data set. Furthermore, previous studies with limited western blotting showed that the deletion of other genes such as ADA1, SPT7 and SPT20 result in the disruption of the SAGA complex (Sterner *et al*, 1999). Out of these, the deletion of SPT20 was for us of great interest, since previous work demonstrated that its deletion only yielded moderately increased levels of ubiquitylated H2B (Henry *et al*, 2003), indicating that the deletion of this single protein compromises the SAGA complex, but to a lower extent than for components of the DUB module, suggesting that it only lead to a partial loss of the complexes functionality (Henry *et al*, 2003). In order to explain these observations in more detail, a particular focus of our study sought to determine the true effect of the SPT20 deletion on the integrity of the complex, in particular on the HAT and DUB modules, and therefore we included the spt20Δ in our data set.

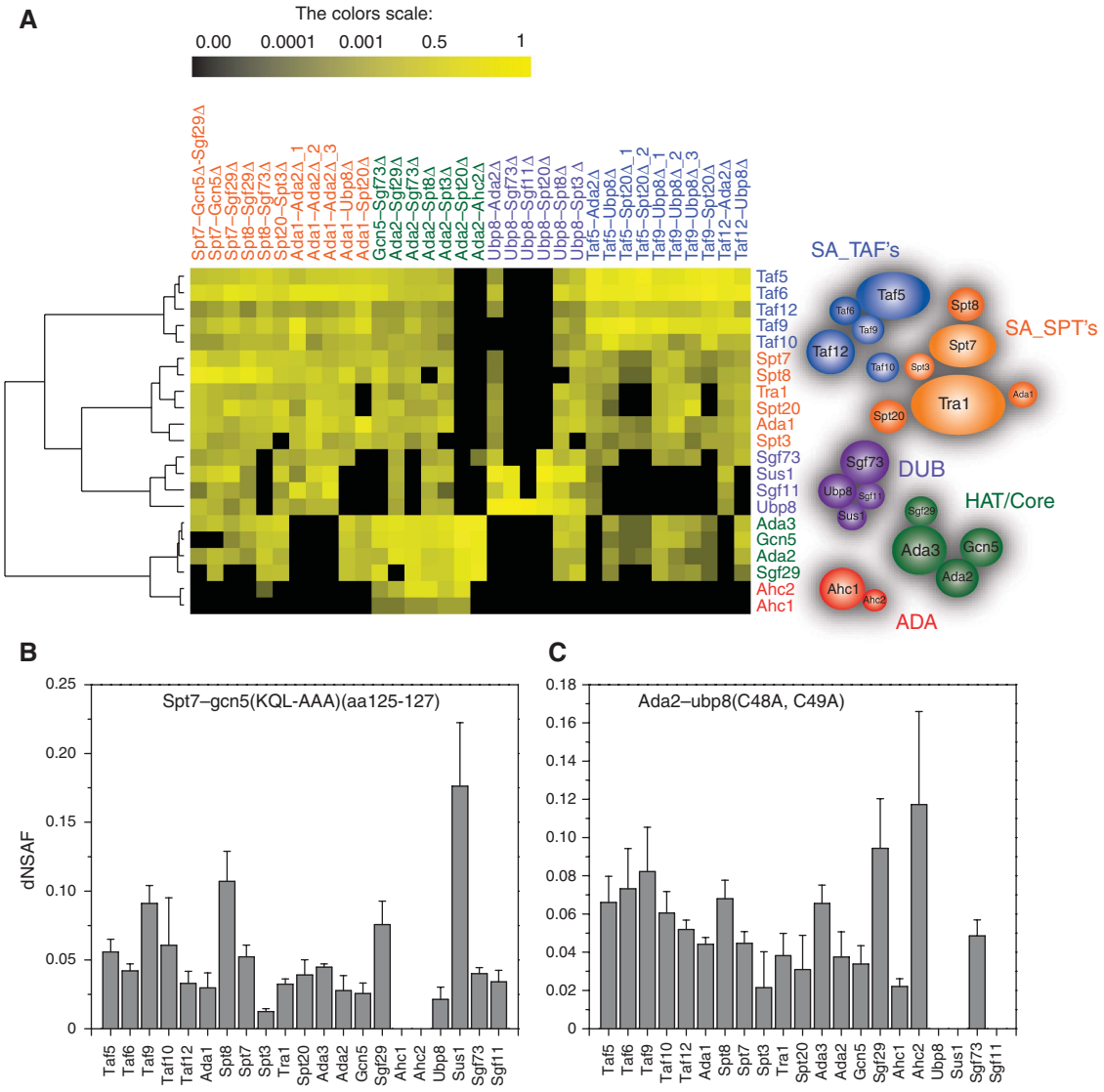

Figure 2 Hierarchical clustering on different deletion strains and analysis of catalytic mutants. (**A**) Each column represents an isolated TAP in a different deletion strain, and each row represents an individual protein (prey). The color intensity represents protein abundance (dNSAF) normalized on the subunits of the SAGA/ADA complexes with the brightest yellow indicating highest abundance and decreasing intensity indicating decreasing abundance. Black indicates that the protein was not detected in a particular purification. The proteins of the modules were colored as in Figure 1. The clustering result leads to the formation of distinct modules (represented on the right side of the cluster). Relative abundance of the 21 subunits of the SAGA/ADA complexes obtained from (**B**) purifications of the Gcn5 catalytic mutant using Spt7 as bait and (**C**) Ubp8 catalytic mutant purified by the bait Ada2. In each case, three replicate purifications were performed. The catalytic mutants of Gcn5 and Ubp8 were generated by mutating amino acids 125–127 (KQL to AAA) and by substituting the two zinc-finger amino acids C46A and C49A, respectively (Wang *et al*, 1998; Ingvarsdottir *et al*, 2005). All data is represented as average dNSAF values + s.d.

Regarding the baits, we TAP-tagged SPT proteins, proteins from the HAT module, proteins from the DUB module and TAF proteins, since strains lacking any of the TAF genes are not viable and therefore cannot be deleted. By purifying these proteins in certain deletion backgrounds, we aimed to capture architectural information from different parts of the SAGA complex. Altogether, we performed a total of 34 purifications that included 10 different TAP-tagged baits (Spt7, Spt8, Spt20, Ada1, Gcn5, Ada2, Ubp8, Taf5, Taf9 and TAf12) and 10 different deletion strains (gcn5Δ sgf29Δ double mutant, gcn5Δ, sgf29Δ, ada2Δ, sgf73Δ, sgf11Δ, ubp8Δ, spt20Δ, spt3Δ and spt8Δ) (Figure 2A). To ensure the robustness of our results, replicates were also included in our deletion analysis (Figure 2A; Supplementary Figure S1; Supplementary Tables S4–S6). After the respective purifications were conducted and processed, we first applied hierarchical clustering analysis on the entire deletion data set consisting of the 34 purifications (Figure 2A). The results of the clustering analysis indicated a clear dissociation of the SAGA complex and revealed five majors groups/modules: (1) the SA_TAF module, composed of all the SAGA's TAF proteins (Taf6, 5, 12, 9 and 10); (2) the SA_SPT module consisting of all of SAGA's SPT proteins (Spt7, 8, 3 and 20) together with Tra1 and Ada1; (3) the DUB module (Ubp8, Sgf73, Sgf11 and Sus1); (4) the HAT/Core module, which

includes all three previous described components (Gcn5, Ada3 and Ada2), together with Sgf29; and (5) the ADA module that consist of the subunits Ahc1 and Ahc2 subunits (Figure 2A). As we already observed in the wild-type purifications (Figure 1B), even after dissecting the complex by this deletion approach, the proteins Ahc2 and Sgf29 still exhibited similar abundance levels as other members of the ADA and the HAT module, respectively, further indicating that Ahc2 is part of the ADA module and Sgf29 is part of the HAT/core module (Figure 2A). Furthermore, in contrast to the wild-type purifications in which TAF proteins were separated in different branches in the wild-type cluster (Figure 1B), in the deletion purifications, all TAF subunits were now tightly grouped together in the dendrogram (Figure 2A). We also analyzed catalytic mutants of Gcn5 and Ubp8 (Figure 2B and C), which showed similar patterns to the deletion of the whole protein, which will be discussed later.

All of our results on the modularity of the SAGA/ADA complexes, together with an itemization of the similarities and discrepancies compared with previous studies, are summarized in Supplementary Table S7. The combination of different baits with several deletion strain backgrounds followed by quantitative mass spectrometric analysis and cluster analysis allowed us to determine the organization of these proteins into modules within the Gcn5 HAT complexes. To further understand the relationship between the proteins within these modules as well as between the modules, we next studied the effect of the deleted subunits on the association between prey and bait proteins within the complex.

Probabilistic deletion network and protein complex organization

The approach of purifying a protein in a deletion strain has the advantage of capturing not only information about the association between every prey protein and the bait but also between the prey protein and the deleted subunit. The bait and the deleted subunit can have similar or different locations in the complex; therefore, this relative position will affect the extent of a deletion on the preys purified by the bait. Furthermore, certain subunits will have a greater effect on the stability of the complex than others. Quantitative proteomics data is a key feature of our method, since it enables us to determine the change in associations between preys and the baits they co-precipitate with. In order to quantify these associations, we calculated the posterior probability for each prey in a deletion purification based on Bayes' rule as described previously (Sardiu *et al*, 2008). Bayes' theorem converts the observed spectral counts into discrete levels of association strength (Figure 3; Supplementary Table S8). In principle, in a single deletion purification, those preys that retain a high probability should associate stronger with the bait, while the preys that are present at a low probability or are absent from the purification associate stronger with the subunit that was deleted. The associations between each bait and the purified preys in each deletion strain are represented in Figure 3, in which the colors red, cyan and black correspond to low, medium and high probabilities, respectively (see Supplementary information for details).

With respect to the TAP-tagged proteins used in the different deletions (Figure 3), as we expected, all the proteins from the same module as the TAP-tagged protein were highly recovered and had high probabilities. For instance, in Spt7–TAP–gcn5Δ;sgf29Δ, the highest probabilities were observed for Tra1, Ada1 and all the SPTs proteins with Spt8 exhibiting the highest probability (Figure 3A). Interestingly, for Spt8–TAP–sgf29Δ, Spt7 has the highest probability (after Spt8), suggesting a strong association between these two proteins (Figure 3A). To begin, we inspected the HAT/Core module and investigated the effect of the GCN5, SGF29 and ADA2 deletions on this module as well as on the entire complex. In the specific purifications that contain these deletions, ada2Δ had a greater effect on the HAT/Core module when compared with gcn5Δ and sgf29Δ (Figures 2A and 3B). Independent of the TAP-tagged bait used, all and only the components of the HAT module were lost in ada2Δ (Figure 2A). In contrast, when GCN5 and SGF29 were deleted with any combination of TAP-tagged proteins, all components of the HAT module remained at low probabilities, except for the deleted subunit (Figure 3B). Also, as expected, for every deletion within the HAT/Core module, proteins of the module itself were most affected (Figures 2A and 3B). In addition, a catalytic mutation of gcn5 (KQL_AAA) shows a similar mild effect as TAP purifications of strains in which the whole Gcn5 protein is deleted (Spt7–TAP–gcn5Δ–sgf29Δ, Spt7–TAP–gcn5Δ; see Figure 2B; Supplementary Table S9). Taken together, these results indicate that Ada2 has a critical role in the formation of the HAT module and its association with the overall complex (Figures 2A and 3).

Next, we considered the SA_TAF module. For the TAP-tagged TAF baits, the proteins with the highest probabilities in ada2Δ also belonged to the SA_TAF module. These purifications were of particular importance, since the quantitative information obtained from the TAP-tagged TAF baits could substitute for the absence of the deletions in the TAF proteins, which are lethal, and helped group the TAF proteins into the module. Importantly, this grouping indicated that the histone-fold TAFs are associated with other TAFs and less likely dimerize with histone-fold SPT or ADA proteins (Figures 2 and 3C). Since TAF proteins are shared between SAGA and TFIID, their grouping into a discrete module suggests a similar module consisting of the same TAFs which may also exist in TFIID (Figures 2 and 3C), which is a distinct complex that contains additional proteins not observed in SAGA (Auty *et al*, 2004).

Next, we investigated the stability of the DUB module by monitoring the effect of UBP8, SGF73 and SGF11 deletions on this module as well as on the entire complex. Independent of the TAP-tagged bait used, ubp8Δ had the same effect on the DUB module, that is Sus1, Sgf11 and Ubp8 were absent from the module, while Sgf73 was still present (Figures 2 and 3D). A catalytic mutant of Ubp8 phenocopied the same effect of ubp8Δ, loss of the Sus1, Sgf11 and Ubp8, while Sgf73 was still co-purified (Figure 2C; Supplementary Table S9; Ingvarsdottir *et al*, 2005). These results suggest that a tight connectivity is present between these three proteins and additionally that Sgf73 is the anchor between the DUB module and the rest of the complex. In order to understand to which proteins Sgf73 establishes the contact, thereby attaching the DUB module to the complex, we next investigated the purifications in which SPT20 is deleted (Figures 2 and 3). For all baits from the SA_SPT and SA_TAF modules in spt20Δ, Tra1 and the whole

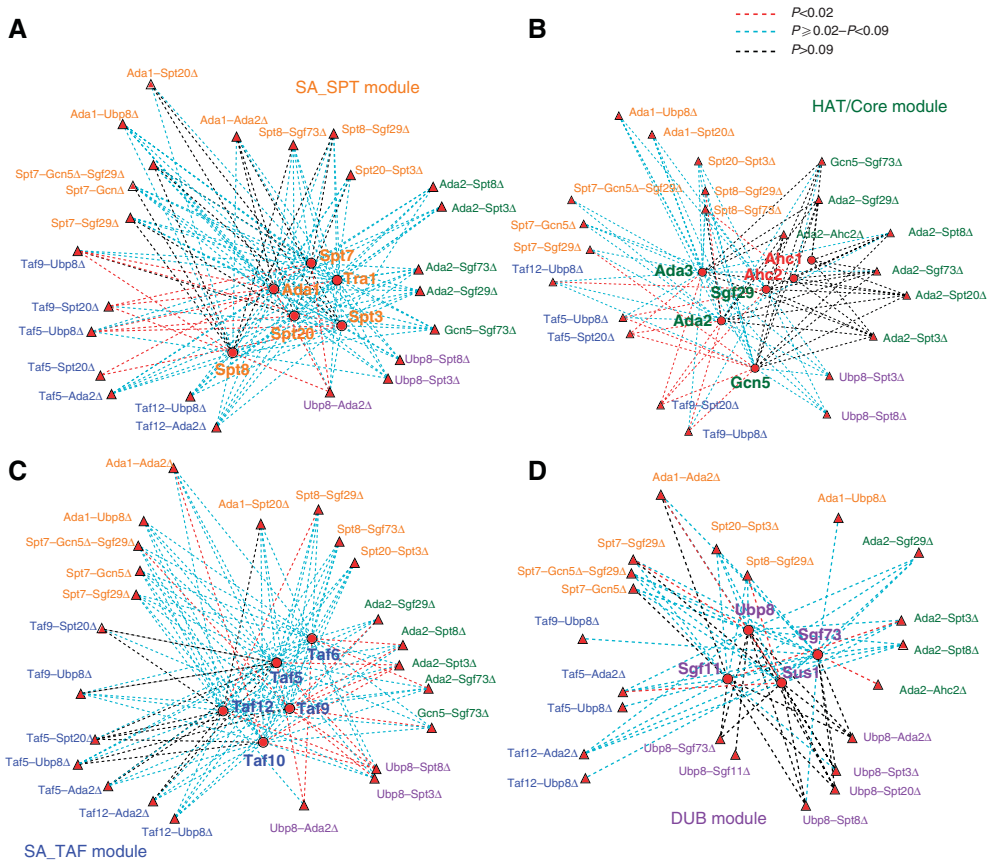

Figure 3 Deletion interaction network of the Gcn5 HAT complexes. (**A**) The probabilistic protein network of the Gcn5 HAT complexes was generated by representing proteins as nodes (baits in the respective deletion strains by triangles and preys as circles), connected by weighted edges denoting the calculated probabilities. Black dashed lines symbolize interactions with high probability, cyan dashed lines interactions with moderate probability and red dashed lines interactions with low probability. The baits are depicted as triangles and colored based on the TAP subunit: orange for the SA_SPT module, green corresponds to the ADA module, violet to the DUB module and blue to the SA_TAF module. Preys are symbolized by circles and colored as in Figure 1 (A–D). Focused probabilistic protein networks for preys of each of the four modules of the Gcn5 HAT complexes in all the baits, i.e. (A) the SA_SPT module, (**B**) the HAT/Core module, (**C**) the SA_TAF module and (**D**) the DUB module. The Cytoscape software environment was used to generate the probabilistic protein networks. In each network, only the baits that have a link (i.e. pull down the prey) with a prey are represented.

DUB module were absent (Figures 2 and 3A and C). The loss of the DUB module in spt20Δ samples indicates that Sgf73 is interacting with Spt20 in order to bring the DUB module into the complex. Furthermore, these observations also suggest a strong association between Spt20 and Tra1 (Figures 2A and 3).

The deletion of Spt20 is also a prime example to illustrate the principle of our strategy, as the choice of the TAP-tagged protein dramatically influences the modules recovered in spt20Δ (Supplementary Figure S2). When using baits from the SA_SPT (Ada1) or SA_TAF (Taf9 and 5) modules, the DUB module and Tra1 were absent (Figures 2A and 3B and D). When proteins from the DUB module were used as baits, the rest of the modules were absent except for DUB (e.g. the Ubp8–TAP–spt20Δ, which only yielded the four components of the DUB module alone; Figure 2A). In the case of baits belonging to the HAT module, all other modules were missing with the exception of the HAT/Core module (Figures 2A and 3; Supplementary Figure S3). This observation strongly suggests that even after a protein essential for the proper assembly and function of SAGA is deleted, small sub-complexes still form. This information could indicate that the assembly of the wild-type SAGA complex does not occur one protein after the other, but rather that first several modular sub-complexes form, which successively are joined together in order to form the mature complex.

Based on our results, we next assembled a macromolecular model for the SAGA and ADA complexes and combined it with previously published yeast two-hybrid and genetic complementation screens (Figure 4; Supplementary Table S10): First, the HAT/Core module contains components that are shared between SAGA and ADA. We placed Ada2 more proximal, since the effect of its deletion on the HAT/Core module was the strongest of all module-specific mutants analyzed. Conversely, Sgf29 and Gcn5, whose deletions did not reveal interdependency with the rest of the module components, were situated more peripheral. In addition, previous data from a genetic deletion screen showed a negative genetic synergism of Ada2 and Gcn5 with components of the DUB module (Costanzo *et al*, 2010) (see Supplementary Table S10), thus we positioned these two proteins closer to the DUB module. Since it was reported from yeast two-hybrid screens (Marcus *et al*, 1994; Wang *et al*, 1997; Uetz *et al*, 2000; Ito *et al*, 2001; Benecke *et al*,

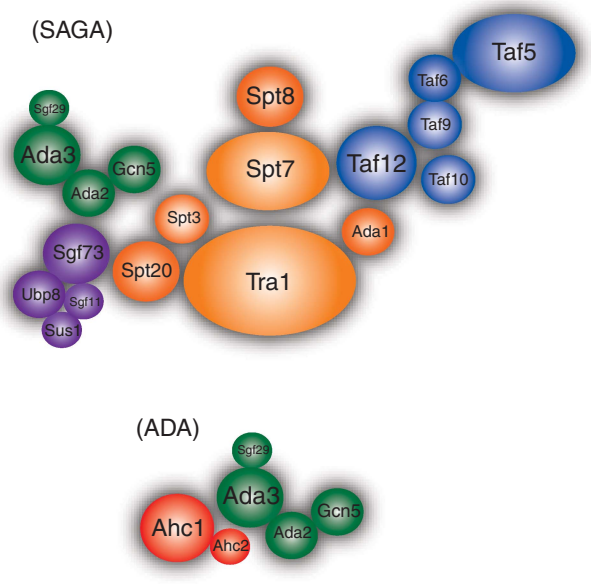

Figure 4 Deletion interaction network and the macromolecular assembly of the Gcn5 HAT complexes. Based upon all deletion purifications, all proteins of the SAGA/ADA complexes were organized into modularity and consequently a macromolecular model was assembled (for details, see main text of the manuscript). In addition to our deletion purifications, we integrated existing data from yeast two-hybrid and gene deletion experiments to further refine our model. As a result, we allowed direct contacts only between protein pairs (i.e Ada2–Gcn5; Ada2–Ada3; Ada3–Sgf29; Taf5–Taf6; and Taf6–Taf9) for which yeast two-hybrid data exist. Genetic interaction data was also used to position some of the proteins from different modules in close proximity. In particular, components of the DUB module exhibit negative genetic effects with two components of the HAT/core module, which are Ada2 and Gcn5. Therefore, these proteins were placed in close proximity. The color code is in accordance with Figures 1B and 2A. The size of the inset circle correlates with the molecular weight of each illustrated protein.

2002) that Ada2 directly interacts with Gcn5 and Ada3, we symbolized this direct interaction in the model by a direct contact between Ada2 and these two proteins. Furthermore, we positioned Ada3 in direct contact with Sgf29 based on yeast two-hybrid data (Ito *et al*, 2001). Second, we positioned the DUB module close to the SA_SPT module and located Sgf73 close to Spt20; Ubp8, Sgf11 and Sus1 were grouped together as they depend on each other. Third, Tra1 was situated close to Spt20, since the deletion of Spt20 led to the loss of Tra1. Spt3 was located closer to the ADA and DUB modules given that it led to a severe synthetic growth defect with Gcn5 (Lin *et al*, 2008) and a negative genetic effect with Sgf11 and Sgf73 (Collins *et al*, 2007; Costanzo *et al*, 2010). All remaining subunits of the SA_SPT module were added according to the order of their probabilities in the respective purifications. Fourth, for the SA_TAF module, Taf12 was placed more inside the complex, since it exhibited higher probabilities with members of the DUB module when used as a bait compared with Taf5–TAP (see Taf5–TAP and Taf12–TAP in ada2Δ). Yeast two-hybrid screens (Uetz *et al*, 2000; Ito *et al*, 2001; Yatherajam *et al*, 2003; Yu *et al*, 2008; Layer *et al*, 2010) furthermore identified direct interactions between the pairs Taf5–Taf6 and Taf6–Taf9; therefore, we permitted direct contact between these proteins in the model. Finally, for the ADA complex, we added a contact between Ahc1 and Ahc2 based on our deletion results and yeast two-hybrid screens (Uetz *et al*, 2000; Ito *et al*, 2001).

SGF29 is a bona fide ADA family member and a core subunit of Gcn5/HAT complexes

During our proteomic analysis of 12 different wild-type baits, Sgf29 was found to segregate together with components of the HAT/Core module of the Gcn5 complexes (Figure 1B). Our analysis on various subunit deletions of these complexes strengthened our conclusion that Sgf29 is indeed a member of the HAT/core module that is part of all Gcn5 HAT complexes and not just SAGA (Figure 2A). In contrast to other well-characterized components of the HAT complexes, Sgf29 is a poorly characterized protein, whose deregulated expression is implicated in malignant transformation (Kurabe *et al*, 2007). Therefore, we set out to test whether the deletion of SGF29 resulted in similar phenotypes as deletion of GCN5, ADA2 or ADA3. We first analyzed the transcriptional coactivation capacity of the SGF29 deletion strain in order to assay for similarities with ADA gene function (Berger *et al*, 1992; McMahon *et al*, 2005). All ADA gene products isolated to date are known to incorporate into the SAGA and SLIK complexes. We assayed for the cells' ability to survive overexpression of Gal4-VP16, which is toxic to wild-type cells, but not lethal for deletions in ADA components. Overexpression of VP16 has been suggested to cause misdirection of SAGA to inappropriately activate a number of cellular genes, and to sequester general transcription factors away from productive transcription complexes (Horiuchi *et al*, 1997). Mutations in SAGA that alter functional interaction with VP16 allow the cells to overcome the toxic growth defect and constitute an ADA phenotype. WT and sgf29Δ yeast strains, along with an ada3Δ strain as a control, were transformed with a high-copy plasmid containing Gal4-VP16 (McMahon *et al*, 2005). Figure 5A shows that the sgf29Δ strain behaved in the same manner as the ada3Δ strain in this assay, suppressing VP16 toxicity (Figure 5A). This finding indicates that Sgf29 is a functional ADA family member, consistent with our observation that it is part of SAGA and SLIK. The suppression of VP16 toxicity in ADA mutants is accompanied by the inability to activate an artificial LacZ reporter gene that is driven by Gal4-VP16 (McMahon *et al*, 2005). In agreement with a suppression of VP16 toxicity, the sgf29Δ yeast strain was also deficient in low-copy Gal4-VP16-dependent expression of the LacZ reporter gene (Figure 5B), similar to other ADA family members (McMahon *et al*, 2005). Overall, our biochemical analysis of Sgf29 revealed that it behaves like a classic ADA gene, as its deletion rescued GAL4-VP16-mediated toxicity, while also being required for SAGA-mediated transcriptional activity (Figure 5A and B).

Deletion of a number of SAGA subunits results in a decreased fitness when yeast are grown on carbon sources other than dextrose. Therefore, we decided to assay whether deletion of SGF29 also compromised growth on various carbon sources. We indeed found that the deletion of SGF29 phenocopied a deletion of SPT7, a SAGA subunit, resulting in a severe growth defect when grown on plates containing only galactose, acetate, ethanol or glycerol as the sole carbon sources (Figure 5C). These phenotypes indicate that the

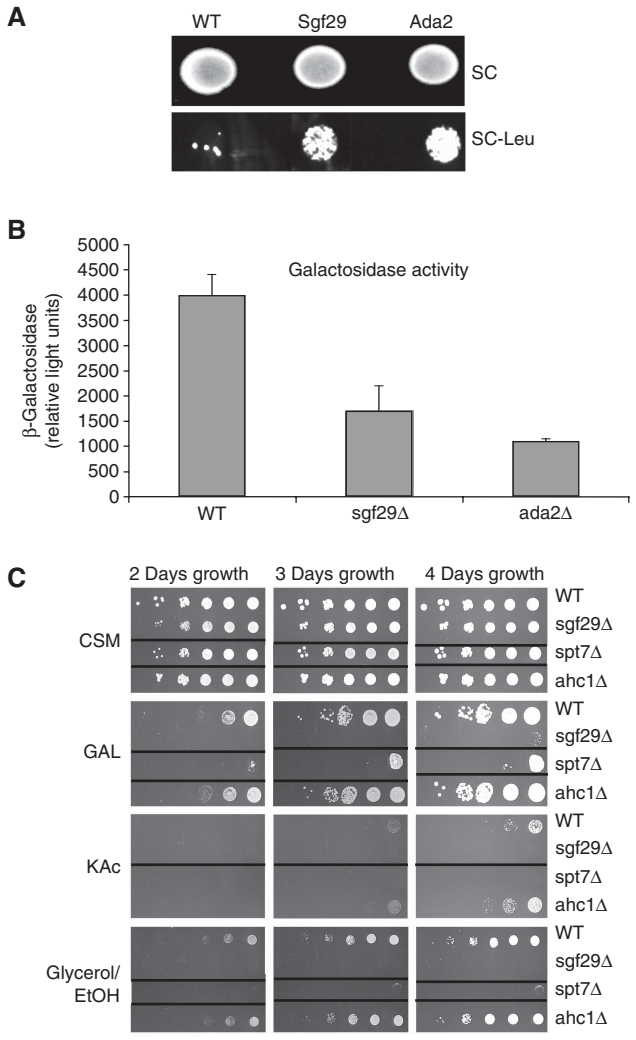

Figure 5 Sgf29 exhibits the characteristics of other known ADA proteins, including Ada2 and Gcn5. (**A**) Deletion of SGF29 rescue Gal4-VP16-mediated toxicity in yeast, similar to the deletion of ADA2. (**B**) β-Galactosidase activation by VP16 in yeast is compromised by the deletion of SGF29. This phenotype is similar to what is seen for the deletion of ADA2 as seen in the graph. (**C**) Yeast lacking SGF29 is compromised for growth on alternative carbon sources, similar to what is observed for other SAGA subunits, including SPT7 (separated by black lines). Yeast were serially diluted on the indicated plates and imaged at the indicated times (see Materials and methods for details).

deletion of SGF29 results in an inability to activate the pathways required to use galactose (GAL1), acetate (CIT2) or ethanol/glycerol (ADH1) as the sole carbon source. Taken together, these observations indicate a functional similarity of Sgf29 with other members of the SAGA complex and the ADA gene family.

AHC2 is a novel component of the ADA HAT complex required for the presence of the ADA module

Previous studies have shown that the ADA complex contains Ada2, Ada3, Gcn5 and a unique subunit Ahc1 (Eberharter *et al*, 1999). However, our analysis revealed that ADA is actually composed of the additional two subunits, Sgf29 (as a member of the HAT/Core) and a previously unidentified polypeptide, YCR082W, which we termed Ahc2 (Figures 1B and 2). Unlike Sgf29, purification of Ahc2 only purified the ADA complex and none of the other components of SAGA or SLIK/SALSA (Figures 1B and 6A and B). In order to confirm our findings that Ahc2 and Sgf29 associate with other ADA complex members, we immunoprecipitated yeast containing a TAP tag on Ahc2 or Sgf29 and probed with an antibody to Ada3, a known component of the HAT/core module (Figure 6A). We found that similar to Ada2–TAP, both Sgf29 and Ahc2 associated with Ada3 (Figure 6A, compare lane 2 with lanes 3 and 4). We next aimed to identify all proteins associated with Ahc2. Purification of the ADA complex using an Ahc2–TAP tag strain followed by MudPIT analysis revealed that Ahc2 only associated with components of the ADA complex (Figures 1B and 6B). Since the Ahc1 deletion was previously shown to not affect the integrity of the rest of the ADA complex, we tested whether the same was true for Ahc2. We performed an Ada2–TAP purification in an AHC2 deletion strain and found that the two specific proteins to the ADA complex, Ahc1 and Ahc2, were lost, while the shared proteins of the ADA complex remained intact (i.e. Gcn5, Ada2, Sgf29 and Ada3) (Figure 2A). This implies that Ahc2 is responsible for tethering Ahc1 into the ADA complex. Since the hallmark of these complexes is their ability to acetylate substrates such as histones, we next tested the ADA complex purified through Ahc2–TAP for HAT activity. To our surprise, we found that the Ahc2-purified ADA complex strongly preferred to acetylate nucleosomes as opposed to core histones (Figure 6C; Supplementary Figure S6A–C). Although this is in contrast to a previous report (Eberharter *et al*, 1999), this discrepancy could be explained by the fact that our experiment for the first time purified the ADA complex through a specified ADA subunit prior to the assay ensuring no cross-contamination of other Gcn5 HAT complexes, such as SAGA and SLIK/SALSA.

Discussion

In order to comprehend how a multi-protein complex functions, it is crucial to first understand how the subunits of the complex are organized and assembled. To this end, we employed a combination of biochemistry approaches, quantitative proteomics and computational methods to better understand the architectural organization of the Gcn5 HAT complexes in *S. cerevisiae*. In a limited previous approach, insights about tight protein complexes were achieved with yeast deletion strains using only a TAP-tagged bait (Sardiu *et al*, 2009b). This approach only provided insights into the local architecture of the complex around the TAP-tagged protein and not the whole complex (Sardiu *et al*, 2009b). For example, if a certain deletion results in the loss of many proteins from the complex, it cannot be determined if the deletion simply prevented the bait protein from binding to an otherwise intact complex or if the whole complex dissociated. Here, the key for the new methodology was to utilize several TAP-tagged baits and deletions to clearly define modules and their interconnectivities. As exemplified by Spt20, a protein essential for the function of the SAGA complex, its central role

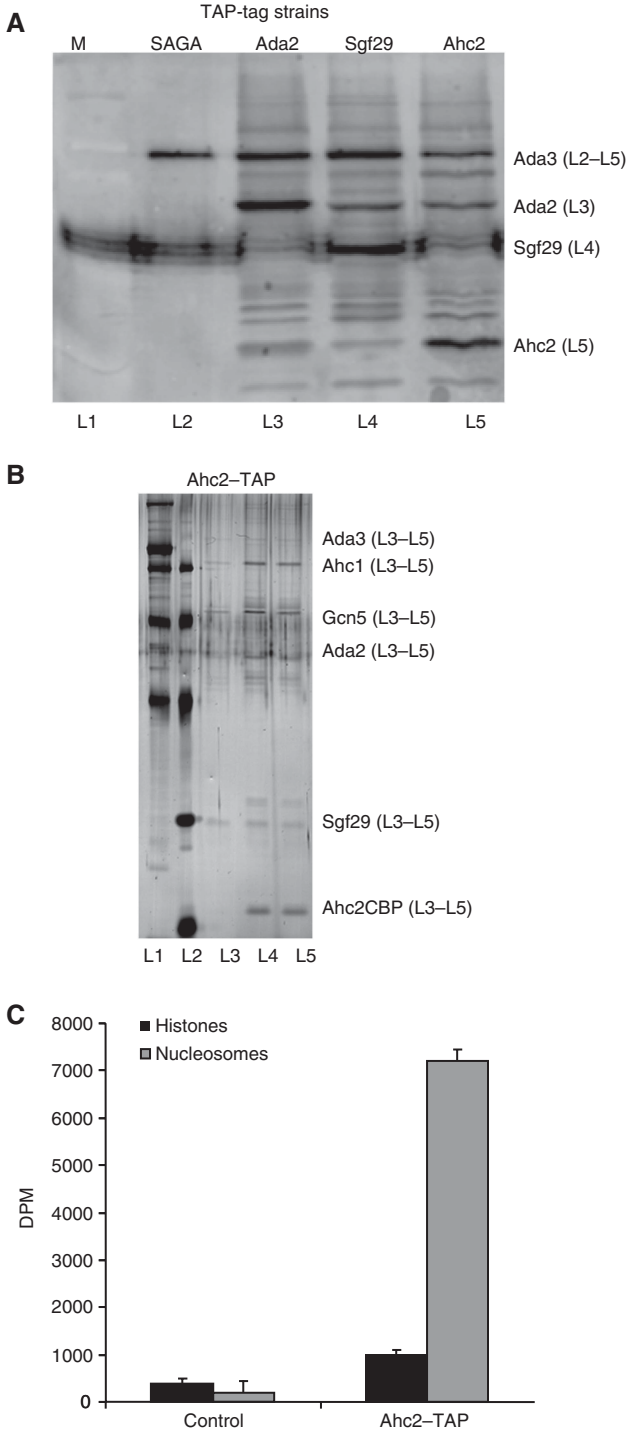

Figure 6 Ahc2 is a bona fide member of the ADA HAT complex. (**A**) Western blot analysis of calmodulin pull-down experiments indicates that both Ahc2 and Sgf29 precipitate Ada3, a known component of Gcn5 HAT complexes (see lanes 3 and 4). (**B**) Silver stain of the TAP tag purification of Ahc2 identified only the six components of the ADA complex. Each of the components are indicated on the gel. (**C**) *In vitro* HAT assay using Ahc2–TAP-purified ADA complex demonstrates that the ADA complex preferentially acetylates nucleosomes compared with histones.

in the assembly of the complex can only be captured by analyzing its deletion in different TAP strains, as distinct modules were purified depending on the component that was chosen for TAP purification (Figure 2A; Supplementary Figure S2). Moreover, our method also permits evaluating lethal components of protein complexes like proteins belonging to the TAF family, for which no deletion analysis can be performed. Through the use of several TAP-tagged TAF proteins in combination with different deletions from outside the module, we still acquired sufficient information to separate and discriminate the SA_TAF module from the remaining proteins of the complex.

A macromolecular model for the SAGA and ADA protein complexes

The macromolecular model proposed upon the results of our analysis extends earlier studies like a single particle EM reconstruction (Wu *et al*, 2004), which only localized 9 of the now 19 known subunits, and two recent studies resolving the structure of the DUB module, which contains 4 subunits, using X-ray crystallography (Kohler *et al*, 2010; Samara *et al*, 2010). There is a need for methods that can provide alternative architectural information to bridge this gap in the knowledge of SAGA. Our study, for example, places Ada2, which was not mapped in the EM study, into the center of the HAT/Core module. Similarly, it brings the SA_SPT module in close proximity to the DUB module, and our model predicts that this link is established through Sgf73, which is in striking agreement with the above-mentioned crystallographic study of Kohler *et al* (2010). Our model also incorporates the two novel ADA subunits identified in this study, Ahc2 and Sgf29, and its placement is supported both by functional experiments performed in this study and by previous large-scale yeast studies, which reported interaction for protein pairs Ahc1–Ahc2, Ahc2–Gcn5 and Sgf29–Ada3 (Uetz *et al*, 2000; Ito *et al*, 2001; Krogan *et al*, 2006). Through additional experimentation, we demonstrated that Sgf29 is an ADA family member and a core subunit of the HAT/Core module. In addition, we demonstrated that Ahc2 is a *bona fide* novel component of the ADA and HAT/Core modules, which can preferentially acetylate nucleosomes over core histones. Since the ADA complex does not contain Tra1 to target it to gene activators, it is intriguing to speculate that the ADA complex may function in a similar fashion with the piccolo NuA4 complex to help maintain overall H3 acetylation in the genome (Selleck *et al*, 2005; Berndsen *et al*, 2007).

Contrary to the previous EM-based view, our model also proposes a modularity of the SAGA and ADA complexes. This modular view, which assigns the different functions of the complexes to distinct modules, is strongly supported by deletions of non-catalytic units, which affect only some, but not all of the complexes' functions, like the deletion of Spt20, which leaves the DUB function almost intact. This observation suggests that the distinct functional modules of the SAGA complex can persist separately. It is intriguing to speculate that such a modular buildup of different functional units could also be observed in other multi-protein complexes beyond SAGA and ADA, and could be a common mechanism to utilize the same functional modules in distinct protein complexes. Despite the modularity of SAGA and ADA, the SA_SPT module, which according to our analysis is centrally located in the complex, seems to be necessary for multiple if not all

functions of the complexes, as deletions of SPT20, SPT7 and ADA1 were previously shown to disrupt the complexes to an extent, which compromises its multiple functions (Grant *et al*, 1997; Horiuchi *et al*, 1997; Roberts and Winston, 1997).

Identification of stable SAGA sub-complexes

One of the most interesting findings in our analysis revolved around the purification of SAGA from strains lacking SPT20. The deletion of SPT20 is well known to compromise the integrity of the SAGA and SLIK/SALSA complexes (Sterner *et al*, 1999). However, the exact nature of this disruption had not been addressed until now. We were intrigued by the finding that the deletion of SPT20 leads to only a slight increase in H2B ubiquitination (Henry *et al*, 2003). If SAGA were disrupted, one would assume that the DUB module would also be compromised. However, the analysis of our proteomic data obtained from purifications through both Ada2 and Ubp8 in the absence of SPT20 revealed that the individual HAT/Core module and the DUB module were intact in the SPT20 deletion (Figure 2A). This finding is consistent with only a partial loss of H2B DUB seen in this deletion (Henry *et al*, 2003), as the DUB module can probably still carry out a subset of its activity when it is not part of SAGA. Since our deletion analysis of the components of SAGA demonstrated the stability of the modules even after perturbing the complex, it is important to take this into consideration when discussing protein complex integrity. Although SAGA as a whole may be disrupted, there could still be residual activities associated with isolated intact HAT and DUB modules that could lead to spurious acetylation and DUB, which could be detrimental to the cell.

The application of our method to the SAGA and ADA complexes highlights the ability of this approach to generate architectural insights into multi-protein complexes. It not only provides architectural information, but also facilitates the identification of subunits, which are essential for the integrity of specific modules as well as of the whole complex. Compared with other structural studies, which mapped 9 of the 19 known SAGA subunits using single EM reconstruction (Wu *et al*, 2004) or resolved the structure of the 4 subunits of the DUB module using X-ray crystallography (Kohler *et al*, 2010; Samara *et al*, 2010), our approach is not limited to a maximum number of complex subunits. Consequently, we were able to construct a macromolecular model consisting of all 21 SAGA/ADA subunits, which bridges the gap between the previous limited EM analysis and focused on X-ray crystallography analysis. Our analysis also emphasizes the benefit of architectural information for the functional characterization of multi-protein complexes. Especially in the case of protein complexes composed of multiple functional modules, this information eases the prediction of phenotypic outcomes due to targeted deletions or mutations observed in clinical diseases. Given the enormous challenges in generating high-resolution structures of multi-protein complexes with traditional structural biology tools, our method, which can be carried out in any system where gene depletions are possible, provides an alternative approach to generating novel insight into the organization and architecture of multi-protein complexes.

Materials and methods

S. cerevisiae strains

TAP tag and Mat, a knockout strains, were obtained from Open Biosystems. Gene deletions in the TAP tag strains were carried out by homologous recombination using a kanamycin gene cassette flanked by 200 base pairs of gene-specific sequence. Strains containing mutants in either UBP8 or GCN5 were constructed as follows: Both wild-type plasmids were obtained for the MoBY-ORF collection (Open Biosystems). The plasmids were subsequently subjected to site-directed mutagenesis using the Quick-Change mutagenesis kit (Stratagene). The mutated plasmids were sequence verified and then transformed into strains either lacking UBP8 or GCN5. A total of 3 l of the transformed strains were grown in media lacking uracil to maintain the plasmid and subsequent TAP purification was carried out as described earlier.

Identification of proteins by MudPIT

MudPIT analysis of purified complexes was carried out as previously described (Lee *et al*, 2009). TCA-precipitated proteins were urea-denatured, reduced, alkylated and digested with endoproteinase Lys-C (Roche) followed by modified trypsin (Promega) as described in Florens and Washburn (2006). Peptide mixtures were loaded onto 100 μm fused silica microcapillary columns packed with 5 μm C_{18} reverse phase (Aqua, Phenomenex), strong cation exchange particles (Partisphere SCX, Whatman) and reverse phase (McDonald *et al*, 2002). Loaded microcapillary columns were placed in-line with a Quaternary 1100 series HPLC pump (± Agilent) and an LTQ or XP linear ion trap mass spectrometer equipped with a nano-LC electrospray ionization source (ThermoFinnigan). Fully automated 10-step MudPIT runs were carried out on the electrosprayed peptides, as described in Florens and Washburn (2006). Tandem mass (MS/MS) spectra were interpreted using SEQUEST (Eng *et al*, 1994) against a database of 11 982 amino-acid sequences, consisting of 5877 *S. cerevisiae* proteins (non-redundant entries from NCBI 2007-03-04 release), 177 usual contaminants (such as human keratins, IgGs and proteolytic enzymes) and, to estimate false discovery rates (FDR), 5993 randomized sequences for each non-redundant protein entry. Peptide/spectrum matches were selected and compared using DTASelect/CONTRAST (Tabb *et al*, 2002) with the following criteria set: spectra/peptide matches were only retained if they had a DeltCn of at least 0.08, and minimum XCorr of 1.8 for singly, 2.5 for doubly and 3.5 for triply charged spectra. In addition, peptides had to be fully tryptic and at least seven amino acids long. Combining all runs, proteins had to be detected by at least two such peptides, or one peptide with two independent spectra. Under these criteria, the FDR is $<1\%$ (Supplementary Tables S1 and S4). To estimate relative protein levels, normalized spectral abundance factors (NSAFs) were calculated for each non-redundant protein, as described in Zybailov *et al* (2006). Spectral counts for peptides shared between proteins are counted only once, and distributed according to the spectral count contribution of peptides unique to each isoform. NSAF are then calculated based on distributed spectral counts (*dSpC*) with shared spectral counts distributed among protein isoforms (Zhang *et al*, 2010). The protein interactions from this publication have been submitted to the IMEx (http://imex.sf.net) consortium through IntAct (pmid: 19850723) and assigned the identifier IM-15346. The data associated with this manuscript may be downloaded from ProteomeCommons.org Tranche using the following hash: ERr + h3ogpfy2X6FxP4mDtSCfxk8LcZ7HTe7l87ecEnv + cgtpOIxluBlXE/OOFlm/JLXi8k3oAwTSUcb1R1GhzvpIHfYAAAAAAAACTA==. In addition, all RAW files are available from ftp://ftp.stowers-institute.org/pub/washburn/Lee_SAGA_MSB/.

ADA phenotype

In order to assay for the classic ADA phenotype, wild-type and yeast strains deleted for ADA3 and SGF29 were transformed with a high-copy GAL4-VP16 plasmid and grown on LEU plates for 3 days at 30°C (McMahon *et al*, 2005).

β-Galactosidase assays

β-Galactosidase assays were performed in WT, sgf29Δ and ada2Δ yeast strains as described in McMahon *et al* (2005). Yeast strains were transformed with a vector containing a GAL1 promoter element fused to LacZ and a second low-copy expression vector containing Gal4-VP16. If SAGA is present, Gal4-VP16 bound to the GAL1 promoter drives LacZ expression.

Protein techniques

TAP purifications were carried out as previously described (Lee *et al*, 2009), with the exception of the ADA complex used in Figure 6A, which was purified as described in Berger *et al* (1992). For calmodulin pull-down experiments, 50 ml of YPD were grown with the TAP-tagged strains, 1 mg of whole cell extract was added to 25 μl of calmodulin beads and incubated at 4°C overnight. The next day, the beads were washed three times with 300 mM calmodulin-binding buffer, then 2 × SDS sample was added and the samples were boiled and analyzed by western blotting for Ada3, which also detects the IgG tag in the TAP tag, allowing for simultaneous visualization of both the tag and the interacting protein.

In vitro HAT assay

HeLA core histones and nucleosomes were used to perform the *in vitro* HAT assay as described previously (Eberharter *et al*, 1998).

Supplementary information

Supplementary information is available at the *Molecular Systems Biology* website (www.nature.com/msb).

Acknowledgements

This work was supported by the Stowers Institute for Medical Research and NIH Grant R37GM047867 to JLW.
Author contributions: KKL, MES, MPW, PAG and JLW conceived and designed the research. KKL, MES and MPW wrote the paper. KKL, SKS, JMG and MT developed and carried out the experiments. MES, SKS and LF analyzed the data. MES carried out computational proteomic analyses.

Conflict of interest

The authors declare that they have no conflict of interest.

References

Auty R, Steen H, Myers LC, Persinger J, Bartholomew B, Gygi SP, Buratowski S (2004) Purification of active TFIID from Saccharomyces cerevisiae. Extensive promoter contacts and co-activator function. *J Biol Chem* **279:** 49973–49981

Balasubramanian R, Pray-Grant MG, Selleck W, Grant PA, Tan S (2002) Role of the Ada2 and Ada3 transcriptional coactivators in histone acetylation. *J Biol Chem* **277:** 7989–7995

Benecke A, Gaudon C, Garnier JM, vom Baur E, Chambon P, Losson R (2002) ADA3-containing complexes associate with estrogen receptor alpha. *Nucleic Acids Res* **30:** 2508–2514

Berger SL, Pina B, Silverman N, Marcus GA, Agapite J, Reigier JL, Triezenberg SJ, Guarente L (1992) Genetic isolation of ADA2: a potential transcriptional adaptor required for function of certain acidic activation domains. *Cell* **70:** 251–265

Berndsen CE, Selleck W, McBryant SJ, Hansen JC, Tan S, Denu JM (2007) Nucleosome recognition by the piccolo NuA4 histone acetyltransferase complex. *Biochemistry* **46:** 2091–2099

Collins SR, Miller KM, Maas NL, Roguev A, Fillingham J, Chu CS, Schuldiner M, Gebbia M, Recht J, Shales M, Ding H, Xu H, Han J, Ingvarsdottir K, Cheng B, Andrews B, Boone C, Berger SL, Hieter P, Zhang Z *et al* (2007) Functional dissection of protein complexes involved in yeast chromosome biology using a genetic interaction map. *Nature* **446:** 806–810

Costanzo M, Baryshnikova A, Bellay J, Kim Y, Spear ED, Sevier CS, Ding H, Koh JL, Toufighi K, Mostafavi S, Prinz J, St Onge RP, VanderSluis B, Makhnevych T, Vizeacoumar FJ, Alizadeh S, Bahr S, Brost RL, Chen Y, Cokol M *et al* (2010) The genetic landscape of a cell. *Science* **327:** 425–431

Daniel JA, Torok MS, Sun ZW, Schieltz D, Allis CD, Yates III JR, Grant PA (2004) Deubiquitination of histone H2B by a yeast acetyltransferase complex regulates transcription. *J Biol Chem* **279:** 1867–1871

Eberharter A, John S, Grant PA, Utley RT, Workman JL (1998) Identification and analysis of yeast nucleosomal histone acetyltransferase complexes. *Methods (San Diego, Calif)* **15:** 315–321

Eberharter A, Sterner DE, Schieltz D, Hassan A, Yates III JR, Berger SL, Workman JL (1999) The ADA complex is a distinct histone acetyltransferase complex in saccharomyces cerevisiae. *Mol Cell Biol* **19:** 6621–6631

Eng JK, McCormack AL, Yates III JR (1994) An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database. *J Am Soc Mass Spectrom* **5:** 976–989

Florens L, Washburn MP (2006) Proteomic analysis by multidimensional protein identification technology. *Methods Mol Biol* **328:** 159–175

Grant PA, Duggan L, Cote J, Roberts SM, Brownell JE, Candau R, Ohba R, Owen-Hughes T, Allis CD, Winston F, Berger SL, Workman JL (1997) Yeast Gcn5 functions in two multisubunit complexes to acetylate nucleosomal histones: characterization of an Ada complex and the SAGA (Spt/Ada) complex. *Genes Dev* **11:** 1640–1650

Grant PA, Eberharter A, John S, Cook RG, Turner BM, Workman JL (1999) Expanded lysine acetylation specificity of Gcn5 in native complexes. *J Biol Chem* **274:** 5895–5900

Grant PA, Schieltz D, Pray-Grant MG, Steger DJ, Reese JC, Yates III JR, Workman JL (1998) A subset of TAF(II)s are integral components of the SAGA complex required for nucleosome acetylation and transcriptional stimulation. *Cell* **94:** 45–53

Henry KW, Wyce A, Lo WS, Duggan LJ, Emre NC, Kao CF, Pillus L, Shilatifard A, Osley MA, Berger SL (2003) Transcriptional activation via sequential histone H2B ubiquitylation and deubiquitylation, mediated by SAGA-associated Ubp8. *Genes Dev* **17:** 2648–2663

Horiuchi J, Silverman N, Pina B, Marcus GA, Guarente L (1997) ADA1, a novel component of the ADA/GCN5 complex, has broader effects than GCN5, ADA2, or ADA3. *Mol Cell Biol* **17:** 3220–3228

Ingvarsdottir K, Krogan NJ, Emre NC, Wyce A, Thompson NJ, Emili A, Hughes TR, Greenblatt JF, Berger SL (2005) H2B ubiquitin protease Ubp8 and Sgf11 constitute a discrete functional module within the Saccharomyces cerevisiae SAGA complex. *Mol Cell Biol* **25:** 1162–1172

Ito T, Chiba T, Ozawa R, Yoshida M, Hattori M, Sakaki Y (2001) A comprehensive two-hybrid analysis to explore the yeast protein interactome. *Proc Natl Acad Sci USA* **98:** 4569–4574

Kohler A, Pascual-Garcia P, Llopis A, Zapater M, Posas F, Hurt E, Rodriguez-Navarro S (2006) The mRNA export factor Sus1 is involved in Spt/Ada/Gcn5 acetyltransferase-mediated H2B deubiquitinylation through its interaction with Ubp8 and Sgf11. *Mol Biol Cell* **17:** 4228–4236

Kohler A, Schneider M, Cabal GG, Nehrbass U, Hurt E (2008) Yeast Ataxin-7 links histone deubiquitination with gene gating and mRNA export. *Nat Cell Biol* **10:** 707–715

Kohler A, Zimmerman E, Schneider M, Hurt E, Zheng N (2010) Structural basis for assembly and activation of the heterotetrameric SAGA histone H2B deubiquitinase module. *Cell* **141:** 606–617

Koutelou E, Hirsch CL, Dent SY (2010) Multiple faces of the SAGA complex. *Curr Opin Cell Biol* **22:** 374–382

Krogan NJ, Cagney G, Yu H, Zhong G, Guo X, Ignatchenko A, Li J, Pu S, Datta N, Tikuisis AP, Punna T, Peregrin-Alvarez JM, Shales M, Zhang X, Davey M, Robinson MD, Paccanaro A, Bray JE, Sheung A, Beattie B *et al* (2006) Global landscape of protein complexes in the yeast Saccharomyces cerevisiae. *Nature* **440:** 637–643

Kurabe N, Katagiri K, Komiya Y, Ito R, Sugiyama A, Kawasaki Y, Tashiro F (2007) Deregulated expression of a novel component of TFTC/STAGA histone acetyltransferase complexes, rat SGF29, in hepatocellular carcinoma: possible implication for the oncogenic potential of c-Myc. *Oncogene* **26:** 5626–5634

Layer JH, Miller SG, Weil PA (2010) Direct transactivator-transcription factor IID (TFIID) contacts drive yeast ribosomal protein gene transcription. *J Biol Chem* **285:** 15489–15499

Lee KK, Florens L, Swanson SK, Washburn MP, Workman JL (2005) The deubiquitylation activity of Ubp8 is dependent upon Sgf11 and its association with the SAGA complex. *Mol Cell Biol* **25:** 1173–1182

Lee KK, Swanson SK, Florens L, Washburn MP, Workman JL (2009) Yeast Sgf73/Ataxin-7 serves to anchor the deubiquitination module into both SAGA and Slik(SALSA) HAT complexes. *Epigenetics Chromatin* **2:** 2

Lee KK, Workman JL (2007) Histone acetyltransferase complexes: one size doesn't fit all. *Nat Rev* **8:** 284–295

Lin YY, Qi Y, Lu JY, Pan X, Yuan DS, Zhao Y, Bader JS, Boeke JD (2008) A comprehensive synthetic genetic interaction network governing yeast histone acetylation and deacetylation. *Genes Dev* **22:** 2062–2074

Marcus GA, Silverman N, Berger SL, Horiuchi J, Guarente L (1994) Functional similarity and physical association between GCN5 and ADA2: putative transcriptional adaptors. *EMBO J* **13:** 4807–4815

McDonald WH, Ohi R, Miyamoto DT, Mitchison TJ, Yates III JR (2002) Comparison of three directly coupled HPLC MS/MS strategies for identification of proeins from complex mixtures: single-dimension LC-MS/MS, 2-phase MudPIT, and 3-phase MudPIT. *Int J Mass Spectrom* **219:** 245–251

McMahon SJ, Pray-Grant MG, Schieltz D, Yates III JR, Grant PA (2005) Polyglutamine-expanded spinocerebellar ataxia-7 protein disrupts normal SAGA and SLIK histone acetyltransferase activity. *Proc Natl Acad Sci USA* **102:** 8478–8482

Mitchell L, Lambert JP, Gerdes M, Al-Madhoun AS, Skerjanc IS, Figeys D, Baetz K (2008) Functional dissection of the NuA4 histone acetyltransferase reveals its role as a genetic hub and that Eaf1 is essential for complex integrity. *Mol Cell Biol* **28:** 2244–2256

Pray-Grant MG, Daniel JA, Schieltz D, Yates III JR, Grant PA (2005) Chd1 chromodomain links histone H3 methylation with SAGA- and SLIK-dependent acetylation. *Nature* **433:** 434–438

Pray-Grant MG, Schieltz D, McMahon SJ, Wood JM, Kennedy EL, Cook RG, Workman JL, Yates III JR, Grant PA (2002) The novel SLIK histone acetyltransferase complex functions in the yeast retrograde response pathway. *Mol Cell Biol* **22:** 8774–8786

Roberts SM, Winston F (1997) Essential functional interactions of SAGA, a Saccharomyces cerevisiae complex of Spt, Ada, and Gcn5 proteins, with the Snf/Swi and Srb/mediator complexes. *Genetics* **147:** 451–465

Saleh A, Lang V, Cook R, Brandl CJ (1997) Identification of native complexes containing the yeast coactivator/repressor proteins NGG1/ADA3 and ADA2. *J Biol Chem* **272:** 5571–5578

Samara NL, Datta AB, Berndsen CE, Zhang X, Yao T, Cohen RE, Wolberger C (2010) Structural insights into the assembly and function of the SAGA deubiquitinating module. *Science* **328:** 1025–1029

Sardiu ME, Cai Y, Jin J, Swanson SK, Conaway RC, Conaway JW, Florens L, Washburn MP (2008) Probabilistic assembly of human protein interaction networks from label-free quantitative proteomics. *Proc Natl Acad Sci USA* **105:** 1454–1459

Sardiu ME, Florens L, Washburn MP (2009a) Evaluation of clustering algorithms for protein complex and protein interaction network assembly. *J Proteome Res* **8:** 2944–2952

Sardiu ME, Gilmore JM, Carrozza MJ, Li B, Workman JL, Florens L, Washburn MP (2009b) Determining protein complex connectivity using a probabilistic deletion network derived from quantitative proteomics. *PloS One* **4:** e7310

Selleck W, Fortin I, Sermwittayawong D, Cote J, Tan S (2005) The Saccharomyces cerevisiae Piccolo NuA4 histone acetyltransferase complex requires the enhancer of polycomb A domain and chromodomain to acetylate nucleosomes. *Mol Cell Biol* **25:** 5535–5542

Sterner DE, Belotserkovskaya R, Berger SL (2002) SALSA, a variant of yeast SAGA, contains truncated Spt7, which correlates with activated transcription. *Proc Natl Acad Sci USA* **99:** 11622–11627

Sterner DE, Grant PA, Roberts SM, Duggan LJ, Belotserkovskaya R, Pacella LA, Winston F, Workman JL, Berger SL (1999) Functional organization of the yeast SAGA complex: distinct components involved in structural integrity, nucleosome acetylation, and TATA-binding protein interaction. *Mol Cell Biol* **19:** 86–98

Swanson SK, Florens L, Washburn MP (2009) Generation and analysis of multidimensional protein identification technology datasets. *Methods Mol Biol* **492:** 1–20

Tabb DL, McDonald WH, Yates III JR (2002) DTASelect and contrast: tools for assembling and comparing protein identifications from shotgun proteomics. *J Proteome Res* **1:** 21–26

Uetz P, Giot L, Cagney G, Mansfield TA, Judson RS, Knight JR, Lockshon D, Narayan V, Srinivasan M, Pochart P, Qureshi-Emili A, Li Y, Godwin B, Conover D, Kalbfleisch T, Vijayadamodar G, Yang M, Johnston M, Fields S, Rothberg JM (2000) A comprehensive analysis of protein-protein interactions in Saccharomyces cerevisiae. *Nature* **403:** 623–627

Wang L, Liu L, Berger SL (1998) Critical residues for histone acetylation by Gcn5, functioning in Ada and SAGA complexes, are also required for transcriptional function *in vivo*. *Genes Dev* **12:** 640–653

Wang L, Mizzen C, Ying C, Candau R, Barlev N, Brownell J, Allis CD, Berger SL (1997) Histone acetyltransferase activity is conserved between yeast and human GCN5 and is required for complementation of growth and transcriptional activation. *Mol Cell Biol* **17:** 519–527

Wu PY, Ruhlmann C, Winston F, Schultz P (2004) Molecular architecture of the S. cerevisiae SAGA complex. *Mol Cell* **15:** 199–208

Yatherajam G, Zhang L, Kraemer SM, Stargell LA (2003) Protein-protein interaction map for yeast TFIID. *Nucleic Acids Res* **31:** 1252–1260

Yu H, Braun P, Yildirim MA, Lemmens I, Venkatesan K, Sahalie J, Hirozane-Kishikawa T, Gebreab F, Li N, Simonis N, Hao T, Rual JF, Dricot A, Vazquez A, Murray RR, Simon C, Tardivo L, Tam S, Svrzikapa N, Fan C *et al* (2008) High-quality binary protein interaction map of the yeast interactome network. *Science* **322:** 104–110

Zhang Y, Wen Z, Washburn MP, Florens L (2010) Refinements to label free proteome quantitation: how to deal with peptides shared by multiple proteins. *Anal Chem* **82:** 2272–2281

Zybailov B, Mosley AL, Sardiu ME, Coleman MK, Florens L, Washburn MP (2006) Statistical analysis of membrane proteome expression changes in Saccharomyces cerevisiae. *J Proteome Res* **5:** 2339–2347