Sgf29 binds histone H3K4me2/3 and is required for SAGA complex recruitment and histone H3 acetylation

Chuanbing Bian^{1,2,8}, Chao Xu^{2,8}, Jianbin Ruan^{1,8}, Kenneth K Lee^{3,8}, Tara L Burke^{4,8}, Wolfram Tempel², Dalia Barsyte², Jing Li¹, Minhao Wu¹, Bo O Zhou⁵, Brian E Fleharty³, Ariel Paulson³, Abdellah Allali-Hassani², Jin-Qiu Zhou⁵, Georges Mer⁶, Patrick A Grant⁴, Jerry L Workman^{3,*}, Jianye Zang^{1,*} and Jinrong Min^{2,7,*}

¹School of Life Sciences, University of Science and Technology of China, Anhui, People's Republic of China; ²Structural Genomics Consortium, University of Toronto, Toronto, Ontario, Canada; ³Stowers Institute for Medical Research, Kansas City, MO, USA; ⁴Department of Biochemistry and Molecular Genetics, University of Virginia School of Medicine, Charlottesville, VA, USA; ⁵The State Key Laboratory of Molecular Biology, Institute of Biochemistry and Cell Biology, Shanghai Institutes for Biological Sciences, The Graduate School, Chinese Academy of Sciences, Shanghai, People's Republic of China; ⁶Department of Biochemistry and Molecular Biology, Mayo Clinic College of Medicine, Rochester, MN, USA and ⁷Department of Physiology, University of Toronto, Toronto, Ontario, Canada

The SAGA (Spt-Ada-Gcn5 acetyltransferase) complex is an important chromatin modifying complex that can both acetylate and deubiquitinate histones. Sgf29 is a novel component of the SAGA complex. Here, we report the crystal structures of the tandem Tudor domains of *Saccharomyces cerevisiae* and human Sgf29 and their complexes with H3K4me2 and H3K4me3 peptides, respectively, and show that Sgf29 selectively binds H3K4me2/3 marks. Our crystal structures reveal that Sgf29 harbours unique tandem Tudor domains in its C-terminus. The tandem Tudor domains in Sgf29 tightly pack against each other face-to-face with each Tudor domain harbouring a negatively charged pocket accommodating the first residue alanine and methylated K4 residue of histone H3, respectively. The H3A1 and K4me3 binding pockets and the limited binding cleft length between these two binding pockets are the structural determinants in conferring the ability of Sgf29 to selectively recognize H3K4me2/3. Our *in vitro* and *in vivo* functional assays show that Sgf29 recognizes methylated H3K4 to recruit the SAGA complex

to its targets sites and mediates histone H3 acetylation, underscoring the importance of Sgf29 in gene regulation.

The EMBO Journal (2011) 30, 2829–2842. doi:10.1038/

emboj.2011.193; Published online 17 June 2011

Subject Categories: chromatin & transcription; structural biology

Keywords: H3K4 methylation; SAGA; Sgf29; Tudor

Introduction

The SAGA (Spt-Ada-Gcn5 acetyltransferase) complex has been extensively investigated due to its important role in the regulation of gene expression (Baker and Grant, 2007; Rodriguez-Navarro, 2009). SAGA was identified in yeast as a 1.8-MDa histone acetyltransferase complex (Grant *et al*, 1997). The complex is highly conserved from yeast to humans, further confirming its importance in transcriptional regulation. Structural integrity of the complex requires three core components, Spt7, Spt20 and Ada1 (Wu and Winston, 2002). The SAGA complex contains two catalytic proteins, Gcn5 and Ubp8. Gcn5 is the histone acetyltransferase catalytic component, while Ubp8 is the histone deubiquitinating component (Rodriguez-Navarro, 2009). The complex contains about 20 proteins. In addition to the catalytic proteins of SAGA, Gcn5 and Ubp8, the remainder of the components have been shown to be important facilitators of either Ubp8 or Gcn5 catalytic function. These other proteins modulate SAGA catalytic activities through their functional domains (Lee and Workman, 2007). These include an acetyllysine-binding bromodomain in Gcn5 and Spt7 (Zeng and Zhou, 2002); DNA-binding SWIRM domain in Ada2 (Qian *et al*, 2005; Da *et al*, 2006); and DNA- or histone-binding SANT domain in Ada2 (Boyer *et al*, 2004), and WD40 domains of unknown function in Taf5 (Durso *et al*, 2001) and Spt8 (Sermwittayawong and Tan, 2006) (Supplementary Figure S1).

In 2002, Sgf29 was identified as a new component of the yeast SAGA complex (Sanders *et al*, 2002), but very little is known about its function. Like most SAGA proteins, Sgf29 is conserved from yeast to humans, and contains an N-terminal coiled-coil domain and C-terminal putative Tudor domains. More recently, the rat orthologue of Sgf29 was shown to directly bind Ada3, a SAGA component and important modulator of Gcn5 activity, via its N-terminal coiled-coil domain and activate c-Myc targeted gene expression (Kurabe *et al*, 2007). Downregulation of Sgf29 was able to suppress genes involved in c-Myc-mediated malignant transformation, implicating the important function of Sgf29 in proper gene regulation of STAGA, the mammalian homologue of SAGA.

Most recently, a study by Vermeulen *et al* (2010) used mass spectrometry-based technologies to screen for interactors of activating and repressive tri-methyl marks on H3 and H4

*Corresponding authors. JL Workman, 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: jlw@stowers.org or J Zang, School of life sciences, University of Science and Technology of China, 96 Jinzhai Road, Hefei, Anhui 230027, People's Republic of China. Tel.: +86 551 3603433; Fax: +86 551 3600374; E-mail: zangjy@ustc.edu.cn or J Min, Structural Genomics Consortium, University of Toronto, 101 College Street, MaRs Tower, Rm 736, Toronto, Ontario, Canada M5G 1L7. Tel.: +1 416 946 3868; Fax: +1 416 946 0588; E-mail: jr.min@utoronto.ca

⁸These authors contributed equally to this work

Received: 12 January 2011; accepted: 26 May 2011; published online: 17 June 2011

in human cells. Sgf29 was shown to bind both H3K4me3 and H3K4me2 sites, with a slight preference for H3K4me3. ChIP sequencing revealed the presence of Sgf29 at gene promoters, which overlapped with the H3K4me3 mark. Conventional biochemistry methods confirmed the H3K4me2/3-specific binding by Sgf29. Sgf29, via its Tudor domains, was shown to be responsible for linking the human SAGA complex to H3K4me3 since knockdown of Sgf29 resulted in loss of H3K4me3 binding (Vermeulen *et al*, 2010). Therefore, Sgf29 acts as a molecular link between human SAGA and the post-translational modification H3K4me3.

Many chromatin-associated proteins or protein complexes have both histone writing and reading activities, which are crucial for their proper function. For example, Ctr4, the methyltransferase of the CtrC complex in fission yeast, contains both a catalytic SET domain and a chromodomain, which methylates histone H3K9 and binds H3K9me3, respectively (Min *et al*, 2002; Zhang *et al*, 2008). The ability of Ctr4 to both write and read H3K9me is essential to spread heterochromatin and facilitate heterochromatin maintenance (Zhang *et al*, 2008). Sgf29, as a reader of H3K4 methyl marks for the SAGA complex, joins the group of other proteins in both yeast and humans known to recognize methyl marks for their respective HAT complex and target its proper acetylation. For example, Yng1 is known to target NuA3 HAT activity to the 5' ORF in yeast by binding tri-methylated H3K4 via its PHD finger (Taverna *et al*, 2006). Similarly, in human cells, the PHD finger of the ING4 protein, a subunit in the HBO1 acetyltransferase complex, binds to tri-methylated H3K4, increasing acetylation by the HBO1 acetyltransferase at the promoter of target genes, leading to efficient transcription (Hung *et al*, 2009; Saksoouk *et al*, 2009). Misregulation of these targeted processes can have costly outcomes when they control genes that require tight regulation. Such is the case with ING4, a known tumour suppressor, where it drives expression of genes involved in anti-cancer activities (Hung *et al*, 2009). Therefore, these subunit reader domains of HAT complexes are important to properly target writer functions, specifically acetyltransferase activity in this case, and misregulation or loss of their function can lead to aberrations in chromatin dynamics.

In this work, our structural studies show that Sgf29 contains unique tandem Tudor domains at its C-terminus and our binding assays show that these tandem Tudor domains selectively bind H3K4me2/3. To elucidate the molecular mechanism of this specific recognition by Sgf29, we determined the crystal structures of the tandem Tudor domains of human and *Saccharomyces cerevisiae* Sgf29 in complex with different modified histone H3K4 peptides. Furthermore, our *in vivo* functional assays show that Sgf29 is required for histone H3 acetylation by the SAGA complex.

Results and discussion

Sgf29 preferentially recognizes histone H3K4me2/3 via its tandem Tudor domains

Based on secondary structure prediction, we found that Sgf29 contains a coiled-coil domain at its N-terminus and putative tandem Tudor domains at its C-terminus (Figure 1A). In contrast to the sequence diversity at its N-terminus, we found that the C-terminal region of Sgf29 has relatively higher sequence identity than the N-terminus, especially within the conserved Tudor domains (Figure 1B).

The Tudor domain, as an important member of the 'Royal Family' of histone-binding modules, is structurally similar to the chromo, PWWP and MBT domains (Maurer-Stroh *et al*, 2003), and has been shown to bind methylated histones (Adams-Cioaba and Min, 2009). Thus, it was compelling to speculate that Sgf29 may preserve this histone methyllysine binding ability. To better understand the binding specificity of human hsSGF29 and its yeast orthologue scSgf29, we used isothermal titration calorimetry (ITC), surface plasmon resonance (SPR) and fluorescence polarization (FP) assays to measure the binding affinity of both hsSGF29 and scSgf29 for histone H3K4, H3K9, H3K27, H3K36, H3K79 and H4K20 peptides bearing different methylation states. We found that both hsSGF29 and scSgf29 do not exhibit detectable binding to any of the H3K27, H3K36, H3K79 and H4K20 peptides, regardless of their methylation states (Table I). Instead, both Sgf29 proteins show strong binding to methylated H3K4 peptides and preferentially bind H3K4me2 and H3K4me3 marks (Table I). Yeast scSgf29 shows no detectable binding to the unmodified H3K4 peptide. Human hsSGF29 can still bind unmodified H3K4 peptide, but with nearly 50-fold weaker affinity ($K_d = 24.0 \mu\text{M}$). Interestingly, hsSGF29 can also bind to H3K9me3 peptides as strongly as to the unmodified H3K4 peptide (Table I). Because the H3K9 peptides we used share the first 11 residues with the H3K4 peptides, and scSgf29 does not bind H3K9me3 peptide similar to its inability to bind the unmodified H3K4 peptide (Table I), presumably hsSGF29 binds H3K9 peptides via the K4 containing sequence. In order to corroborate this, we synthesized a peptide H3_{6–20}K9me3 covering residues 6–20 of histone H3. The length of this peptide is the same as the wild-type (WT) H3K9me3 peptide, and the residue K9me3 is in a position corresponding to the K4me3 residue from the N-terminus of H3. Our ITC-binding assay shows that the binding to hsSGF29 is completely abolished for the H3_{6–20}K9me3 peptide (Table I). Hence, the tandem Tudor domains of Sgf29 selectively recognizes H3K4me2/3 marks.

Sgf29 contains unique tandem Tudor domains

To uncover the molecular architecture of the putative tandem Tudor domains of Sgf29, we determined the crystal structures of human hsSGF29 (residues 129–293) and *S. cerevisiae* scSgf29 (residues 113–259). The crystal structures show that both human and yeast Sgf29 indeed contain tandem Tudor domains at their C-termini. The scSgf29 and hsSGF29 structures are very conserved with an RMSD of 1.6 Å for all aligned C α atoms, although scSgf29 and hsSGF29 only have 20% amino-acid sequence identity (Figure 1B). Each Tudor domain consists of five twisted anti-parallel β strands forming a typical barrel-like fold (Figure 1C and D). scSgf29 was crystallized with a maltose-binding protein (MBP) tag fused to aid crystallization (Supplementary Figure S2A–C). scSgf29 in complex with the methylated H3K4 peptides were crystallized at pH 4.0. At such low pH, scSgf29 can still bind H3K4me2/3, although the binding affinity decreased dramatically (Supplementary Figure S2D and E). The tandem Tudor domains in Sgf29 tightly pack against each other face-to-face, which is distinct from other known tandem Tudor domain structures (Botuyan *et al*, 2006; Huang *et al*, 2006; Adams-Cioaba *et al*, 2010), which we will discuss below.

Figure 1 Crystal structures of human and yeast Sgf29 tandem Tudor domains. **(A)** Domain structures of budding yeast Sgf29 (Sc) and human Sgf29 (Hs). The coiled-coil domain is coloured in orange, and the two Tudor domains are coloured in blue and green, respectively. The starting and ending residues of each domain are numbered. **(B)** Structure-based sequence alignment of Sgf29 homologues. The two conserved Tudor domains are framed and coloured in blue and green, respectively. The secondary structure elements of scSgf29 and hsSGF29 are indicated above and below the sequence alignment, respectively. H3K4me and H3A1 binding residues are numbered and marked by stars and dots, respectively. The sequence identity of the full-length and the tandem Tudor domains of yeast Sgf29 with the other three homologues are indicated beside the first and the third row, respectively. The alignment was created with Esprict (<http://esprict.ibcp.fr/ESprict/ESprict/>). Sc, *S. cerevisiae*; Sp, *Schizosaccharomyces pombe*; Dm, *Drosophila melanogaster*; Hs, *Homo sapiens*. β and η represent β strands and 3_{10} helix, respectively. **(C, D)** Cartoon representation of the crystal structures of yeast Sgf29 and human SGF29 tandem Tudor domains, respectively. The two Tudor domains are coloured in blue and green, respectively, and the secondary structure regions in both proteins are marked.

Structural basis for the selective binding of Sgf29 to histone H3K4me2/3 peptides

To shed light on the molecular mechanism of selective binding of Sgf29 to methylated histone H3K4, we determined the crystal structures of hsSGF29 (residues 115–293) and scSgf29 (residues 113–259) in complex with di- and trimethylated H3K4 peptides, respectively. The structures of the H3K4me2–Sgf29 and H3K4me3–Sgf29 complexes are almost identical for both hsSGF29 and scSgf29 (Figure 2; Supplementary Figures S3 and S4). We used a longer hsSGF29 construct for crystallization of the complexes because crystals were of higher quality than those of the short construct (residues 129–291). The longer construct contains an extra α helix in the N-terminus, which is located between the

two Tudor domains and sits outside the histone binding cleft (Figure 2B). Hence, this extra N-terminal α helix is not directly involved in histone binding, which is also confirmed by our binding results. Our studies show that the short hsSGF29 fragment binds as tightly as the long hsSGF29 fragment to the H3K4me3 histone peptide (Supplementary Figure S5).

From the complex structures of both scSgf29 and hsSGF29, we can see that the first four residues of the H3K4me3 peptide are snugly embedded between the two Tudor domains (Figure 2A and B). The first residue in the H3K4me3 peptide (H3A1) is anchored in a small negatively charged pocket created by the first Tudor domain in both hsSGF29 and scSgf29 (Figure 2). The backbone amine group of H3A1

forms a conserved salt bridge with the carboxylate oxygen of D194 in hsSGF29 (D163 in scSgf29) (Figure 2). The importance of this H3A1-binding pocket is exemplified by the significant reduction in binding affinity (~ 70 -fold weaker) when the H3A1 residue is deleted from the H3K4me3 peptide (Table I). Because the H3A1-binding pocket

Table I Binding affinities of human and yeast Sgf29 to different histone peptides

Peptide name	K_d (μ M): hsSGF29		K_d (μ M): scSgf29	
	ITC	FP	SPR	FP
H3 ₁₋₁₁ K4me0	24 $\pm$ 2	92 $\pm$ 19	NB ^a	NB
H3 ₁₋₁₁ K4me1	4 $\pm$ 1	21 $\pm$ 4	9.2 $\pm$ 0.4	32 $\pm$ 7
H3 ₁₋₁₁ K4me2	1 $\pm$ 0.1	4 $\pm$ 1	1.1 $\pm$ 0.1	5 $\pm$ 2
H3 ₁₋₁₁ K4me3	0.5 $\pm$ 0.2	2 $\pm$ 0.3	0.8 $\pm$ 0.2	4 $\pm$ 1 (2.3 $\pm$ 0.3 ^b)
H3 ₁₋₁₅ K9me3	20 $\pm$ 1	ND ^a	ND	NB
H3 ₁₉₋₃₃ K27me0/1/2/3	ND	NB	ND	NB
H3 ₃₃₋₄₁ K36me0/1/2/3	ND	NB	ND	NB
H3 ₇₁₋₈₅ K79me0/1/2/3	ND	NB	ND	NB
H4 ₁₃₋₂₇ K20me0/1/2/3	ND	NB	ND	NB
H3 ₆₋₂₀ K9me3	NB	ND	ND	ND
H3 ₂₋₁₁ K4me3	34 $\pm$ 3	ND	ND	ND
α -N-Ac-H3 ₁₋₁₃ K4me3Rme2s	NB	ND	ND	ND
H3 ₁₋₁₃ K4me3Rme2s	0.5 $\pm$ 0.1	ND	ND	ND

^aNB, no detectable binding by ITC, SPR or FP assays; ND, not determined.

^bThe binding affinity of scSgf29 to H3K4me3 is measured by ITC in comparison with the K_d values obtained by SPR and FP.

et is negatively charged and rigidly formed, acetylation of H3A1 completely abolishes H3K4me3 binding to hsSGF29 (Table I). Hence, the short side chain H3A1 residue is tightly secured in the small negatively charged pocket formed in the first Tudor domain of Sgf29.

The tri-methyllysine K4me3 is bound in a conserved negatively charged pocket located in the second Tudor domain in both hsSGF29 and scSgf29, and is flanked by two aromatic residues (Y238 and Y245 in hsSGF29, and Y205 and Y212 in scSgf29) (Figure 2). The backbone amine group of the K4me3 also forms a hydrogen bond with the carbonyl oxygen of Y245 in hsSGF29 (Y212 in scSgf29). A third aromatic residue, F264 in hsSGF29 (F229 in scSgf29), lies underneath and buttresses Y238 (Y205 in scSgf29), although it also provides additional hydrophobic interaction with the methyllysine. A negatively charged residue D266 (E232 in scSgf29) on the other side of the K4me3-binding pocket interacts with the methyllysine via a salt bridge. Therefore, the tri-methyllysine residue K4me3 is anchored by the second Tudor domain through cation- π , van der Waals, hydrophobic, electrostatic and hydrogen bond interactions.

We performed a series of mutagenesis experiments in hsSGF29 in order to verify the importance of the K4me3 and H3A1 binding residues. Mutating Y245 to alanine completely abolished the binding, mutating F264, Y238 or D266 to alanine markedly reduced the binding affinity (Table II), which is consistent with our structural observations that Y245 are essential in binding methylated H3K4me3, and Y238, F264 and D266 reinforce the binding.

Figure 2 Selective binding of H3K4me2/3 by the SGF29 tandem Tudor domains. Interactions between yeast (A) and human SGF29 (B) tandem Tudor domains and H3K4me3 peptide. The Sgf29 tandem Tudor domains are shown in cartoon representation and coloured in blue. The H3K4me3 peptide is shown as a stick model. The conserved and non-conserved hydrogen bonds are coloured in red and black, respectively. (C, D) Electrostatic surface representation of yeast (C) and human (D) Sgf29-H3K4me3 complex. Sgf29 is shown in surface representation, and histone peptide is shown in a stick model.

Mutating D194 to alanine or D194/D196 to alanines in hsSGF29 disrupts binding to H3K4me3 (Table II), underlying the importance of D194 in H3K4m2/3 binding through anchoring the H3A1 residue. The D196A mutant could still bind H3K4me3 although with an ~ 12 -fold weaker affinity (Table II). In human but not yeast Sgf29, D196 forms an extra hydrogen bond with the H3A1 residue. D196 is replaced by E165 in scSgf29, which does not form the hydrogen bond with H3A1 (Figure 2A). D196R mutation completely disrupts binding to H3K4me3 (Table II), which indicates that the negative charge around the H3A1-binding pocket is necessary to stabilize the complex.

The backbone of Arg2 in H3 (H3R2) forms two hydrogen bonds with the backbone amine group and the side chain carbonyl oxygen of Thr242 in hsSGF29. Mutating Thr242 to alanine severely diminishes the binding (Table II). Nevertheless, the H3R2 side chain is flexible and does not form any conserved interactions with hsSGF29 (Figure 2; Supplementary Figure S3D). This explains why methylation of H3R2 does not affect its binding to hsSGF29 ($K_d = 0.5 \mu\text{M}$; Table I).

Table II Binding affinities of hsSGF29 and its mutants to histone H3K4me3 peptide (ARTK_{me3}QTARKST)

hsSGF29 wt and mutants	K_d (μM)
Wild type	2.2 ± 0.3
D194A	NB
D194R	NB
D196A	28 ± 2
D196R	NB
D194A_D196A	NB
Y238A	123 ± 13
Q240A	5.5 ± 0.7
T242A	832 ± 207
Y245A	NB
F264A	128 ± 16
D266A	104 ± 18

NB, no detectable binding.

Sgf29 has unique tandem Tudor domains, distinct from those of JMJD2A, FMR1, 53BP1 and SND1

So far, the crystal structures of a few tandem Tudor domains have become available. Based on the architecture of these Tudor domains, tandem Tudor domains can be classified into five subfamilies, i.e., Sgf29, JMJD2A (Huang *et al*, 2006), 53BP1 (Botuyan *et al*, 2006), FMR1 (Ramos *et al*, 2006; Adams-Cioaba *et al*, 2010) and SND1 (Liu *et al*, 2010a,b) subfamilies. The two Tudor domains in Sgf29 face each other with the first two β strands from each Tudor domain packing against each other (Figure 3A and B), accordingly named as ‘face-to-face’ tandem Tudor domains. The two Tudor domains in 53BP1 line up one after the another with the first two β strands from the first Tudor packing against the last two β strands from the second Tudor (Figure 3C) (Botuyan *et al*, 2006), named as ‘lineup’ tandem Tudor domains. The two Tudor domains in the FMR1 subfamily of Tudor proteins, such as FMR1 (Ramos *et al*, 2006), FXR1/2 (Adams-Cioaba *et al*, 2010) and UHRF1, form a head-to-head architecture with the same ends of the β strands packing against each other (Figure 3D). JMJD2A has hybrid Tudor domains (Figure 3E) (Huang *et al*, 2006). The extended Tudor domain in SND1 and other TDRD members was predicted to contain a single Tudor domain, but the recent structural studies demonstrated that the N- and C-terminal extensions surrounding the canonical Tudor domain in SND1 and other TDRD members fold together and form a Tudor-like domain (Figure 3F) (Liu *et al*, 2010a,b). All these five kinds of tandem Tudor domains bind lysine/arginine methylated proteins, but only in Sgf29 do both Tudor domains extensively bind methylated histones (Supplementary Figure S6). Two negatively charged pockets are formed in each individual Tudor domain, and the length between these two pockets determines the selectivity of Sgf29. The other four kinds of tandem Tudor domains mainly use one Tudor domain to interact with their corresponding ligands. That explains why they have less strict sequence

Figure 3 Architecture of different double Tudor domains. (A) Human SGF29. (B) Yeast Sgf29. (C) 53BP1 (PDB code 2G3R). (D) FXR1 (PDB code 3O8V). (E) Hybrid Tudor domain protein JMJD2A (PDB code 2GFA). (F) The extended Tudor domain of SND1 (PDB code 3OMC). The Tudor domains are coloured in rainbow spectrum on the basis of topology, except those of JMJD2A and SND1 that are coloured based on the sequence boundaries of Tudor domains. The N-terminal and C-terminal extensions are coloured in grey.

selectivity. For example, JMJD2A cannot only bind H3K4me3, but also binds H4K20me3 in an opposite orientation (Lee *et al*, 2008).

Interaction of Sgf29 tandem Tudor domains with methylated H3K4 is important for the acetyltransferase activity of SAGA complex *in vivo*

Although recombinant Gcn5 can only acetylate histone H3 lysine 14 (H3K14) *in vitro*, the substrate specificity of Gcn5 expands to histone H3 lysine 9, 14, 18, 23 and histone H4 when it functions with other SAGA components as a complex (Zhang *et al*, 1998; Grant *et al*, 1999). To investigate whether Sgf29 is able to regulate the acetyltransferase activity of the SAGA complex in budding yeast *in vivo*, the global acetylation levels of H3 and H4 were examined. We found that similar to Gcn5 and Ada3 deletion (Figure 4A), Sgf29 deletion reduces the global acetylation levels of H3K9ac and H3K18ac and to a lesser extent H3K14ac, but has no significant effect on the

acetylation levels of H3K23ac and H4ac, which are consistent with H3K9, H3K14 and H3K18 being the primary acetylation sites for SAGA (Grant *et al*, 1999; Lee and Workman, 2007) (Figure 4A). Notably, the yeast Sas3-dependent NuA3 acetyltransferase complex also acetylates H3K14 *in vivo*, explaining the modest reduction in H3K14ac seen in our SAGA mutants (Howe *et al*, 2001; Martin *et al*, 2006; Taverna *et al*, 2006).

Since the deletion of scSgf29 affects the acetylation of H3 markedly, we introduced the scSgf29 gene into an Sgf29 knockout yeast strain to investigate if scSgf29 could rescue histone acetylation of H3 (Figure 4B, compare lanes WT, Δ Sgf29 and pSgf29_FL). The results show that the global acetylation of H3K9, H3K14 and H3K18 is restored in the rescue assay. We next made some point or deletion mutants of scSgf29, which would reduce or disrupt scSgf29 binding to H3K4me2/3, and examined the effect of these mutations on the histone H3 and H4 acetylation by transforming these scSgf29 mutants into the Sgf29 knockout yeast strains. We

Figure 4 Sgf29 is important for histone H3 acetylation. (A) Global acetylation levels of H3K9, H3K14, H3K18, H3K23 and H4, and expression of scGcn5 in various yeast mutant strains: lane 1: BY4742; lane 2: Δ Ada3; lane 3: Δ Gcn5; lane 4: Δ Sgf29. (B) Global acetylation levels of H3K9, H3K14, H3K18 and H3K23, and expression of scGcn5 in mutant strains in the yeast cells expressing Sgf29 with different deletions or point mutations. Lane 1: BY4742 (WT); lane 2: Δ Sgf29, different Sgf29 plasmids were transformed into Sgf29 deletion yeast strain, respectively; lane 3: WT Sgf29; lane 4: D163A; lane 5: E165A; lane 6: Y205A; lane 7: T209A; lane 8: T210A; lane 9: Y212A; lane 10: F229A; lane 11: Δ T1; and lane 12: Δ T2 (deletion of T1 and T2 domains). Quantified values for the acetylation levels in the above western blot panel are present in a bar chart with values as percentage of WT strain. (C) Expression levels of scSgf29 and its mutants. (D) Efficiency of human GCN5 and hsSGF29 knockdown by shRNA. (E) Global acetylation levels of histone H3K9, H3K14, H3K18, H3K23 and H4. Lane 1: knockdown of GCN5; lane 2: control; lane 3: knockdown of hsSGF29. The bands are quantified to better show the differences.

found that these mutants could not rescue H3K9, H3K14 and H3K18 global acetylation (Figure 4B). The expression levels of WT and mutant scSgf29 were monitored by quantitative RT-PCR. We found that the global reduction of H3K9, H3K14 and H3K18 acetylation is not caused by the loss of scSgf29 expression in mutant yeast strains (Figure 4C).

In order to explore if Sgf29 has a similar role in mammalian cells, we used shRNA to knock down hsSGF29 in MDA-MB231 cells (Figure 4D), and found that hsSGF29 knock-down reduced the acetylation levels of H3K9, H3K14, H3K18 and H3K23, but not that of H4 (Figure 4E). Hence, in both *S. cerevisiae* and humans, Sgf29 regulates H3K9, H3K14 and H3K18 acetylation. Interestingly, it was reported recently that Spt20 subunit knockdown of the SAGA complex leads to reduction in the global acetylation levels of histone H3K9 and H3K14, but not histone H4 (Nagy *et al*, 2010). Together, these results suggest that recognition of H3K4me2/3 marks by Sgf29 have an important role in H3 acetylation by the SAGA complex.

Sgf29 is important for recruiting SAGA to its target sites

Because tri-methylation of histone H3K4 is required for subsequent acetylation of histone H3 (Jiang *et al*, 2007), and Sgf29, as a component of the SAGA complex, binds H3K4me2/3 marks and is required for H3 acetylation, we speculated that the recognition of the methylated histone mark H3K4me2/3 by Sgf29 tandem Tudor domains would contribute to the targeting of the acetyltransferase activity of the SAGA complex and mediate the crosstalk between the post-translational modifications of histones. If this hypothesis holds, Sgf29 would not be directly required *per se* for the HAT activities of SAGA. To verify our hypothesis, first we asked whether the loss of Sgf29 would affect the integrity of the SAGA complex by multidimensional protein identification technology (MudPIT) mass spectrometry analysis. SAGA was purified from an Spt8TAP-tagged strain with Sgf29 deleted and compared with that purified from a WT strain. Our MudPIT mass spectrometry analysis revealed that deletion of Sgf29 resulted only in the loss of the Sgf29 protein (Supplementary Table SI), showing that Sgf29 was not required for the complex integrity. Next, in order to determine a role for Sgf29 in SAGA complex *in vitro*, a variety of substrates including core histones, nucleosome or histone H3 tails with various modifications were used to examine the HAT activity of the Δ Sgf29 SAGA complex on these substrates. We found that deleting Sgf29 had marginal effects on SAGA activity on all these substrates *in vitro* (Figure 5A). Additionally, *in vitro* HAT activity was not dependent on the H3K4 methylation and deleting the first five residues of histone H3 had no significant effect on its HAT activity (Figure 5A). Therefore, Sgf29 is required for complete and efficient HAT activity of SAGA *in vivo*, but dispensable *in vitro*.

Since deleting Sgf29 reduced H3 acetylation levels *in vivo*, we wanted to see if deletion of Sgf29 affected genome-wide H3 acetylation or localization of SAGA to genes. To do this, we employed genome-wide ChIP-Chip analysis. We found that deletion of Sgf29 resulted in a loss of H3K9 acetylation at promoters when compared with WT strain (Figure 6A and B). We also found that SAGA localization was decreased at promoters genome-wide when compared with WT strain, as determined by ChIP through the Spt8 subunit (Figure 6C and D). These results are consistent with a recent study by

Vermeulen *et al* (2010), in which ChIP-Seq data revealed that human Sgf29 is highly associated with promoter regions and this association coincides with H3K4me3 at the promoter. In addition, we looked at a specific subset of genes previously identified as requiring SAGA for their expression (Lee *et al*, 2000). This subset of 593 genes was analysed in the same manner as the genome-wide data and similar trends were found in both data sets (Figure 6A–D). In both cases, deletion of Sgf29 resulted in a decrease in both H3K9 acetylation and SAGA localization (Figure 6A–D). Furthermore, we dissected the genome into different categories based on gene length and transcription rate, and determined how deleting Sgf29 would affect acetylation and SAGA recruitment on these different categories of genes (Supplementary Figures S7–S10). In general, we found that deletion of Sgf29 displayed similar decreases in acetylation and SAGA recruitment across these different categories analyzed. Interestingly, the genome-wide reduction of SAGA was more dramatic than the SAGA-specific genes (Figure 6D). This is likely due to the fact that SAGA is present at many more genes than it has been shown to regulate (Venters *et al*, 2011). Additionally, our genome-wide analysis was conducted using the SAGA-specific subunit, Spt8, while the expression studies were done using deletions that affect not only SAGA, but also the SAGA-related complex SliK/SaLSA (Lee *et al*, 2000).

Since Sgf29 is required to maintain the global H3 acetylation, we examined what effects deleting Sgf29 or one or both of its Tudor domains would have on the well-studied SAGA target gene, GAL1. SAGA and H3 acetylation are present at the GAL1 UAS, 5' and 3' ORF under inducing conditions (Govind *et al*, 2007). SAGA targeted to the transcribed coding region of the GAL1 gene specifically increases H3 acetylation and nucleosome eviction and also promotes RNA Pol II processivity (Govind *et al*, 2007). Thus, loss of proper recruitment of SAGA to all regions of GAL1 would severely affect efficient transcription of GAL1. At the GAL1 locus, we found that scSgf29 deletion or Tudor domain deletions resulted in the H3K9 hypoacetylation across the entire GAL1 locus (Figure 5C). We next examined whether scSgf29 deletion or mutations would result in the loss of SAGA at GAL1. We found that, indeed, scSgf29 deletion or mutations resulted in the loss of SAGA recruitment to GAL1 (Figure 5D). Consistently, transcription of GAL1 was severely reduced in all the Sgf29 deletion or Tudor domain deletion strains (Figure 5B). Sgf29's recruitment of SAGA and its subsequent acetylation is dependent upon Set1-specific methylation since deletion of COMPASS subunits, Spp1 (required for tri-methylation of H3K4) and Swd1 (required for di- and tri-methyl H3K4), also results in loss of proper SAGA localization and H3 acetylation (Figure 5C and D). Previous experiments have shown that Set1-dependent H3K4me is present at the GAL1 UAS and 5' region during gene activation (Bryk *et al*, 2002; Ingvarsdottir *et al*, 2007). Supposedly, lack of Sgf29 would compromise the recruitment of the SAGA complex to the GAL1 locus, repressing the acetylation of histone H3 thereby debilitating the transcription of the GAL1 gene. Thus, Sgf29 is an important component in SAGA to bridge histone methylation and histone acetylation.

These findings contribute to the extensive reading and writing functions of chromatin-associated proteins. Specifically, in the SAGA complex, we find that Sgf29 is an important chromatin reader that helps target SAGA to promoters

Figure 5 Sgf29 is required for proper H3K9 acetylation and SAGA localization at Gal1 locus. (A) Purified SAGA from WT and Sgf29 deletion strains via Spt8-TAP was combined, respectively, with 0.5 μg of various SAGA substrates: no peptide, core histones, nucleosomes, unmodified histone H3 peptide (1–20), tetra-acetylated histone H3 peptide (1–20), mono-, di- and tri-methylated histone H3K4 peptide (1–20) and tri-methylated histone H3K9 peptide (6–20). Activity was measured as disintegrations per minute (DPM). (B) mRNA expression levels of WT, ΔSgf29, Sgf29ΔT1, Sgf29ΔT2, ΔSpp1 and ΔSwd1 were determined by RT-PCR in real time. GAL1 was normalized to ACT1 levels and values are expressed as fold induction changes with WT values in dextrose medium set at 1.0. (C) H3K9ac ChIP (anti-H3K9Ac) was used to evaluate the H3K9 acetylation levels of WT, ΔSgf29, Sgf29ΔT1, Sgf29ΔT2, Spp1Δ and Swd1Δ at the GAL1 UAS, 5' and 3' end. (D) SAGA ChIP (anti-Ada2) was used to evaluate the SAGA levels of WT, ΔSgf29, Sgf29ΔT1, Sgf29ΔT2, Spp1Δ and Swd1Δ at the GAL1 UAS, 5' and 3' end.

via its Tudor domains. Interestingly, recent studies show that human Sgf29 binding to H3K4me3 is augmented by the presence of H3K9 and H3K14 acetylation (Vermeulen *et al*, 2010). Most likely, this is the result of a synergistic effect of the methyl reader, Sgf29, and the acetyl reader, Gcn5 via its bromodomain that together may help to propagate transcription of a specific SAGA-dependent gene.

In summary, Sgf29 exhibits strong binding affinity and selectivity for H3K4me2/3 owing to the extensive intermolecular interactions between the complementary surfaces of Sgf29 and the first four residues of the H3K4me peptide. In particular, the H3A1 and K4me3 binding pockets created by each of the tandem Tudor domains and the fixed distance between these two pockets are the structural determinants in conferring Sgf29 the ability of selectively recognizing H3K4me2/3, not other histone lysine sites. Sgf29's ability to recognize and bind methylated H3K4 is required for proper histone H3 acetylation by SAGA.

Materials and methods

Protein expression and purification

Two fragments of human SGF29 (residues 115–293 and 129–291) covering the tandem Tudor domains were subcloned into a pET-28a-

MHL vector. The protocol for protein expression and purification is similar to that described in Adams-Cioaba *et al* (2010). The his-tag was cleaved from SGF29 (residues 115–293) by the addition of 0.05 mg of TEV protease per milligram of SGF29 protein, followed by incubation in ice for 12 h. The sample was then passed through a Ni-NTA column and the flow-through was collected and further purified by size exclusion chromatography (Superdex 75, GE Healthcare). In another Sgf29 construct (residues 129–291), the his-tag was not removed. The fragment of yeast Sgf29 (residues 113–259) containing the tandem Tudor domains was cloned into a pET-23b vector with N-terminal fusion to MBP (MBP-Sgf29). Recombinant MBP-Sgf29 was overexpressed in *Escherichia coli* BL21 (DE3) at 20°C for about 20 h. MBP-Sgf29 was purified by affinity chromatography (amylose resin, New England Biolabs), anion exchange chromatography (resource Q, GE Healthcare) and size exclusion chromatography (Superdex 200, GE Healthcare). Purified MBP-Sgf29 was concentrated to 10 mg/ml in Tris-HCl, pH 8.0 and 150 mM NaCl.

Peptide-binding assays

FP-binding assays were performed at 25°C in 20 mM Tris-HCl, pH 7.5, 50 mM NaCl, 1 mM DTT and 0.01% Tween-20 as described in Xu *et al* (2010). The ITC assays were carried out at 25°C in 25 mM Tris-HCl, pH 7.5, 50 mM NaCl and 1 mM DTT as described in Xu *et al* (2011). The SPR assays were determined at 8°C using a BIAcore 3000 instrument. The peptides were biotinylated and immobilized on a streptavidin-precoated sensor chips (SA chip) and all measurements were carried out in 50 mM Tris-HCl, pH 8.5, 350 mM NaCl and 5% glycerol (solution A). The protein at concentrations ranging

Figure 6 Both genome-wide and SAGA-specific gene array analysis indicate that Sgf29 is required for H3K9acetylation and for proper SAGA localization. (A) Whole-genome median gene analysis of H3K9 acetylation localization in WT strain (solid grey line), sgf29Δ strain (solid black line), SAGA-specific gene analysis of H3K9 acetylation localization in WT strain (dotted grey lines) and sgf29Δ strain (dotted black lines). (B) Identical analysis as in (A), but represented as mutant/WT for sgf29Δ in the H3K9Ac for the whole-genome (All) and SAGA-specific genes (SAGA). (C) Whole-genome median gene analysis of SAGA localization (via Spt8Flag) in WT strain (solid grey line) and sgf29Δ strain (solid black line) and SAGA-specific gene analysis of SAGA localization in WT strain (dotted grey lines) and sgf29Δ strain (dotted black lines). (D) Identical analysis as in (C), but represented as mutant/WT for sgf29Δ in the Spt8 ChIP for the whole-genome (All) and SAGA-specific genes (SAGA). All traces are averages of two (in the case of SAGA) and three (histone ChIPs) independent experiments, shown as IP/IN. The grey box represents the ORF, which for this analysis is divided into 40 equally sized bins. The data are represented on a log₂ scale (y axis).

from 0.625 to 3 μM was flowed for 3 min over the peptide-coated SA chip at a flow rate of 30 ml/min and the change in response unit was recorded. Protein dissociation was monitored for 3 min by flowing solution A over the sensor chip at a flow rate of 30 ml/min. By assuming a single-site equilibrium model, the K_d was determined by global non-linear regression fitting of the association and dissociation curves according to the Langmuir binding isotherm model. The peptides were purchased from Millipore.

Protein crystallization

Purified Sgf29 (residues 115–293) was mixed with methylated histone peptides by addition of a two-fold molar excess of peptide to the protein solution and crystallized at 18°C in 0.1 M Bis-Tris, pH 5.5, 20–34% PEG3350, 100 mM ammonium sulphate and 10 mM strontium chloride. Crystals of SGF29 (residues 129–291) were obtained at 18°C in 0.1 M HEPES, pH 7.5, 0.2 M NaCl, 25% PEG3350. Crystals of MBP-fused yeast Sgf29 (113–259) (MBP-Sgf29) were grown at 8°C in 0.1 M NaAc-HAC, pH 4.0, 2.0 M (NH₄)₂SO₄.

Crystals of MBP-Sgf29 in complex with methylated histone peptides were grown at 8°C in 0.1 M NaAc-HAC, pH 4.0, 20–25% PEG3350.

Structure determination

Diffraction experiments were performed at 100 K. Further experimental details are listed in Table III. Diffraction intensities were indexed and scaled with HKL2000/HKL3000 (Minor *et al*, 2006). The structure of Sgf29–H3K4me3 (115–293) was solved by the single wavelength anomalous diffraction method (Wang, 1985) using a selenomethionine derivative (Hendrickson *et al*, 1990) of the protein and data collected on the Rigaku FRE generator with a Cu target. Heavy atom substructure determination and initial phase calculation were carried out using SHELXD and SHELXE (Schneider and Sheldrick, 2002), respectively. Automated model building in ARP/warp (Perrakis *et al*, 2001) was used to generate the initial model. Manual model improvements and refinement were carried out with COOT (Emsley and Cowtan, 2004) and

Table III Data collection and refinement statistics

	Apo-scSgf29	scSgf29-H3K4me3	scSgf29-H3K4me2	
<i>Data collection</i>				
Radiation source	SSRF BL17U	SSRF BL17U	SSRF BL17U	
Radiation wavelength (Å)	0.9790	1.0000	0.9792	
<i>Data reduction</i>				
Space group	P2 ₁ 2 ₁ 2 ₁	P2 ₁ 2 ₁ 2 ₁	P2 ₁ 2 ₁ 2 ₁	
Cell dimensions				
<i>a</i> , <i>b</i> , <i>c</i> (Å)	49.76, 95.10, 114.43	50.24, 84.85, 119.15	50.17, 85.33, 119.33	
α , β , γ (deg)	90, 90, 90	90, 90, 90	90, 90, 90	
Resolution (Å) ^a	1.92 (1.92–1.95)	2.60 (2.60–2.64)	1.48 (1.48–1.51)	
<i>R</i> _{merge} (%) ^{a,b}	9.6 (23.4)	12.7 (46.9)	7.2 (44.7)	
$\langle I \rangle / \langle \sigma(I) \rangle$ ^a	16.5 (4.8)	9.5 (2.2)	25.7 (3.4)	
Completeness (%)	95.7 (79.8)	96.9 (85.1)	100.0 (100.0)	
Redundancy	5.8 (4.0)	3.2 (2.9)	7.5 (6.7)	
<i>Refinement</i>				
Resolution (Å)	1.92	2.6	1.48	
No. reflections	37 279	16 289	86 103	
<i>R</i> _{work} / <i>R</i> _{free}	19.0/21.6	22.0/25.8	18.3/20.5	
No. atoms				
Protein	4024	3961	4106	
Peptide	0	44	34	
Water	437	308	631	
Other ligands	135	27	23	
β -Factors				
Protein	21.88	43.30	13.80	
Peptide	—	35.18	22.38	
Water	29.50	33.71	24.59	
Other ligands	43.67	24.42	27.80	
RMSD ^c				
Bond lengths (Å)	0.008	0.008	0.009	
Bond angles (deg)	0.97	1.39	1.2	
PDB code	3MP8	3MP1	3MP6	

	Apo-hsSGF29	hsSGF29-H3K4me3	hsSGF29-H3K4me2	hsSGF29-H3R2me2sK4me3
<i>Data collection</i>				
Radiation source	Rigaku FR-E SuperBright	APS 19ID	Rigaku FR-E SuperBright	APS 23ID-B
Radiation wavelength (Å)	1.5418	0.9793	1.5418	0.9793
<i>Data reduction</i>				
Space group	C2	P2 ₁ 2 ₁ 2 ₁	P2 ₁ 2 ₁ 2 ₁	P2 ₁ 2 ₁ 2 ₁
Cell dimensions				
<i>a</i> , <i>b</i> , <i>c</i> (Å)	92.62, 41.42, 52.32	50.04, 65.43, 105.62	50.16, 65.41, 105.88	50.14, 65.48, 106.00
α , β , γ (deg)	90, 121.43, 90	90, 90, 90	90, 90, 90	90, 90, 90
Resolution (Å) ^a	1.92 (1.92–1.95)	1.37 (1.37–1.39)	2.0 (2.00–2.03)	1.28 (1.28–1.30)
<i>R</i> _{merge} (%) ^{a,b}	3.8 (11.9)	8.0 (97.2)	11.3 (58.2)	10.3 (55.9)
$\langle I \rangle / \langle \sigma(I) \rangle$ ^a	37.9 (15.4)	28.6 (1.7)	15.1 (7.8)	7.7 (2.4)
Completeness (%)	92.2 (88.8)	100.0 (100.0)	99.8 (97.7)	98.6 (96.3)
Redundancy	3.8 (3.8)	7.1 (6.9)	6.8 (6.2)	6.8 (5.3)
<i>Refinement</i>				
Resolution (Å)	1.92	1.37	2.00	1.28
No. reflections	11 477	71 590	22 964	84 190
<i>R</i> _{work} / <i>R</i> _{free}	19.8/23.5	19.4/21.4	19.4/24.5	21.1/23.2
No. atoms				
Protein	1245	2770	2770	2745
Peptide	0	106	97	64
Water	84	298	237	487
Other ligands	0	16	16	10
β -Factors				
Protein	31.1	14.2	17.4	11.7
Peptide	—	17.8	28.3	19.3
Water	35.1	23.0	25.6	24.1
Other ligands	—	18.6	20.8	21.8
RMSD ^c				
Bond lengths (Å)	0.024	0.016	0.009	0.012
Bond angles (deg)	2.00	1.59	1.08	1.44
PDB code	3MEW	3ME9	3MET	3MEU

^aValues in parentheses correspond to the highest resolution shells.

^b $R_{\text{merge}} = \sum_{hkl} \sum_j |I(hkl;j) - \langle I(hkl) \rangle| / (\sum_{hkl} \sum_j \langle I(hkl) \rangle)$, where $I(hkl;j)$ is the *j*th measurement of the intensity of the unique reflection (*hkl*), and $\langle I(hkl) \rangle$ is the mean overall symmetry-related measurements.

^cRMSD, root mean squared deviation.

REFMAC (Vagin *et al*, 2004), respectively. The MOLPROBITY server (Davis *et al*, 2004) was used for the validation of protein model geometry. The other human SGF29 structures were solved by molecular replacement using the program MOLREP (Vagin and Teplyakov, 2000). X-ray diffraction data of apo MBP-Sgf29 and MBP-Sgf29 in complex with methylated histone peptides were collected at 100 K at beamline BL17U of SSRF and 3W1A of BSRF and processed with HKL2000 (Minor, 1997; Otwinowski and Minor, 1997). The initial phases were obtained by molecular replacement with the program MOLREP using the MBP structure (PDB ID: 1HSJ) as template (Vagin and Teplyakov, 2000), and refinement were carried out with program COOT (Emsley and Cowtan, 2004) and REFMAC (Vagin *et al*, 2004). Crystal diffraction data and refinement statistics are displayed in Table III.

TAP purification

Purification of TAP-tagged complexes was carried out as previously described (Lee and Workman, 2007) with the following modifications: Elutions were carried out in a volume of 500 µl and repeated five times for a total volume of 3 ml of purified complexes. TAP-purified complexes were resolved on a 10% SDS-PAGE gel and visualized by silver staining.

Yeast strains and media

Yeast strain BY4742_ΔSgf29 (Sgf29 knockout) (*MATα his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 YCL010c::kanMX4*) and its isogenic parent strain BY4742 (*MATα his3Δ1 leu2Δ0 met15Δ0 ura3Δ0*) were obtained from the EUROSCARF. Rich medium (YP, 2% peptone, 1% yeast extract) plus 2% glucose (YPD) and YC complete medium plus 2% glucose or 3% glycerol, respectively, were used for growth as required. Standard protocols for transformation were used in strains constructions (Gietz and Woods, 2002).

Plasmid constructs

The pRS316-scSgf29 centromere plasmid was constructed by cloning a 1607-bp fragment, which contains the scSgf29 coding region and its upstream 400 bp and downstream 400 bp regions into the *Bam*HI-*Xho*I sites of pRS316. The plasmid was confirmed by DNA sequencing. Mutants of scSgf29 were created by site-directed mutagenesis with primer-specific two-step PCR using WT pRS316-scSgf29 plasmid as template to change the desired amino acids into alanine (A). All the mutants were cloned in yeast centromere plasmid pRS316 and confirmed by DNA sequencing.

Semi-quantitative growth assay

Phenotypic analysis was performed using serial dilutions of yeast cells grown at 30°C to early mid-log phase (OD₆₀₀ of 1.0–1.5). Serial dilutions were carried out by spotting cells and diluting four-fold for each spot onto the indicated plates with starting cell density of 10⁷ cells/ml. Plates were grown at 30°C for 2–3 days and imaged.

MudPIT analysis

TCA-precipitated proteins were urea-denatured, reduced, alkylated and digested with endoproteinase Lys-C (Roche) followed by modified trypsin (Promega) as described (Washburn *et al*, 2001). Peptide mixtures were loaded onto 100 µm fused silicamicrocapillary columns packed with 5-µm C18 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 a 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 (Florens and Washburn, 2006). Tandem mass (MS/MS) spectra were interpreted using SEQUEST (Eng *et al*, 1994) against a database of 11 986 amino-acid sequences, consisting of 5880 *S. cerevisiae* proteins (non-redundant entries from NCBI 2008-02-11 release), 177 usual contaminants (such as human keratins, IgGs and proteolytic enzymes), and, to estimate false discovery rates, 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 ΔCn 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.

Antibodies and western blotting for scSgf29

Yeast strains were grown to stable log phase (0.6–1.0) in YC complete medium plus 2% glucose. Then, cells were harvested and washed in 1 ml PBS twice. The pellets were resuspended with 500 µl Lyticase Buffer (50 mM Tris-HCl at pH 7.5, 10 mM MgCl₂, 1 M sorbitol) containing 50 U lyticase and incubated at 30°C for 1 h. Then spheroblasts were collected and washed in LB twice and resuspended with 500 µl EBX (20 mM Tris-HCl at pH 7.4, 100 mM NaCl, 0.25% Triton X-100, 5 mM DTT, 50 mM Na-butyrate⁺, protease inhibitors cocktail). Triton X-100 was added up to 0.5% to lyse the outer cell membrane and the lysate was layered over 1 ml NIB Buffer (20 mM Tris-HCl at pH 7.4, 100 mM NaCl, 1.2 M sucrose, 5 mM DTT, 50 mM Na-butyrate⁺, protease inhibitors cocktail) and centrifuged. Then the nuclear pellets were resuspended in 500 µl EBX and Triton X-100 was added up to 1% to lyse the nuclear membrane and centrifuged. Then the chromatin and nuclear debris were obtained. Washed the chromatins with EBX three times and suspended in 50 µl 1 M Tris (pH 8.0). SDS-PAGE loading buffer was added and samples were incubated at 95°C. For western blotting assays, histone samples were electrophoresed on SDS-15% polyacrylamide gels, electroblots to PVDF membrane and visualized immunochemically by standard methods. Immunodetection was performed with an enhanced chemiluminescence kit (34079, Thermo). The antibodies used to probe were Primary antibody anti-acetyl-H3K9 (07-352, Upstate), anti-acetyl-H3K14 (07-353, Upstate), anti-acetyl-H3K18 (ab1191, Abcam), anti-acetyl-H3K23 (ab46982, Abcam), anti-acetyl-H4 (06-598, Upstate), anti-monomethy-H3K4 (ab8895, Abcam), anti-di-methy-H3K4 (ab32356, Abcam), anti-tri-methy-H3K4 (9727s, Cell Signaling) and anti-Histone H3 (ab1791, Abcam), anti-Histone H4 (ab10158, Abcam), anti-GCN5p (ab63810, Abcam) and secondary antibody, goat anti-rabbit IgG, HRP conjugate (12-348, Millipore).

Antibodies and western blotting for hsSGF29

MDA-MB231 cells were cultured in RPMI with 10% FBS. shRNAs were obtained from the RNAi Consortium (TRC) courtesy Dr J Moffat, processed and used as outlined in the TRC protocols <http://www.broadinstitute.org/rnai/trc>. Sgf29 and Gcn5l2 shRNA chosen on the basis of best knockdown were clones NM_138414.1-269s1c1 and NM_021078.2-1848s21c1, respectively. For western blot experiments, histones were extracted by standard acid extraction method using 0.1 N HCl overnight, acetone precipitated, protein amount quantified and 0.1 µg ran on SDS Tris-Bis gels (Invitrogen). The antibodies used to probe were from Abcam: total histone H3 (#1791), H3K18ac (#1191) and Upstate: H3K9ac (#06-942).

Chromatin immunoprecipitation

Chromatin immunoprecipitation assays were performed as described by Kuo and Allis (1999). The wild-type, ΔSgf29, sgf29ΔT1, sgf29ΔT2 strains were derivatives of BY4742 (Research Genetics, Huntsville, AL). Yeast cultures (50 ml) were grown in SC medium containing 2% dextrose to an absorbance at 600 nm of 0.5, cells were then reseeded in SC medium containing 2% galactose until they reached an absorbance at 600 nm of 1.0. Equal amounts of yeast extracts (1.5 mg total protein) were immunoprecipitated with 1 µl of antibody against acetyl histone H3, lysine 9 (Millipore, 07-352) or antibody against Ada2 in a total reaction volume of 400 µl. DNA was amplified using oligonucleotide primers (Invitrogen) specific for *GAL1* UAS and *GAL1* 5' and 3' intergenic regions. The recovered DNA was quantified in triplicate by real-time PCR using Power SYBR Green PCR Master Mix (Applied Biosystems) and the Bio-Rad MyiQ Single Color real-time PCR detection system. ChIP values obtained are represented as percent input; where signals obtained from the ChIP are divided by the signal obtained from the respective input sample.

HAT assays

HAT assays were carried out as previously described (Eberharther *et al*, 1996), with the following changes: in the peptide HAT assay, 0.5 µg of each individual histone peptide synthesized by Tufts University Core Facility were used in a 30-µl reaction. Both the core histones and the nucleosomes were isolated from HeLa nuclear extract and 0.5 µg were used in the assays described. Additionally, the reactions were carried out for 30 min at 30°C and activity was

monitored as DPM (disintegrations per minute). Equal amounts of Spt8TAP and Spt8TAP (Δ Sgf29) were used as the HAT enzyme in the reaction.

mRNA isolation and qRT-PCR

The WT, Δ Sgf29, sgf29 Δ T1, sgf29 Δ T2, Δ Spp1 strains were derivatives of BY4742 (Research Genetics). Yeast cultures (50 ml) were grown in SC medium containing 2% dextrose to an absorbance at 600 nm of 0.5, cells were then reseeded in SC medium containing 2% galactose until they reached an absorbance at 600 nm of 1.0. Total RNA was extracted from cell pellets, RNA was then DNase treated using TURBO DNA-free kit from Ambion. RNA (1 μ g) was reverse transcribed using SuperScript VILO cDNA synthesis kit (Invitrogen, # 11750). Generated cDNA was amplified in real time using iQ SYBR Green Supermix (Bio-Rad, # 2012-03) and the Bio-Rad MyiQ Single Color real-time PCR detection system. cDNA was amplified using primers specific for *GAL1* and *ACT1* genes.

To monitor the expression of WT scSgf29 and mutants, yeast strain BY4742 with WT scSgf29 or mutants was grown in YPD overnight. Overnight cultures were reseeded into fresh medium to an optical density at OD₆₀₀ of 0.4–0.5 and grown at the same temperature to early mid-log phase (OD₆₀₀ of 1.0–1.5). Then cells were collected and RNA was obtained with RNeasy® Mini Kit (Qiagen, 74104).

To analyse *GAL1* expression, yeast strains BY4742_SGF29 Δ and BY4742, were grown in YPD overnight. Then, cells were collected and washed with ddH₂O (auto-claved) twice. After washing, yeast cells were diluted in YP-raffinose (2%) or YC-raffinose (2%) to mid-log phase and grown for 2–3 h. In all, 2% galactose was added to induce the expression of *GAL1*. Cells were collected after induction for 2 h and RNA was extracted with TRIzol (Invitrogen, 15596-026).

Reverse transcription was performed using TAKARA Prime-Script® Reverse Transcription System with 200–500 ng total RNA to yield 20 μ l cDNA. cDNA was amplified in real time using TAKARA SYBR® Premix EX Taq™ with an ABI StepOnePlus™ Real-Time PCR Systems (Applied Biosystems). The primers used for RT-PCR are listed below.

Gene	Primer name	Sequence	Location
<i>GAL1</i>	GAL1_F	CAGCGAGCTTTACTGCCGAC	1238–1257
	GAL1_R	TCAAGGCACCAAATTGCTTG	1269–1288
<i>scSgf29</i>	scSgf29_F	TGTTGTAGTGTCTCCTTG	18–36
	scSgf29_R	TTCTCTCTGCTGACTGATTC	225–243

Median gene plots

GPR files were loaded into R and normalized within arrays with the Limma package, using default methods. For each set of strain/IP replicates, normalized expression ratios were averaged per probe. These were converted into wiggle-format track files, yielding one track per strain/IP combination. These average tracks were used to create two sets of derivative tracks, one set being mutant/WT ratios for each IP, the other being mark/H3 for each non-H3 IP. These three sets of tracks (Averages, Mutant/WT, Mark/H3) were summarized using the average gene analysis pipeline described (Li *et al*, 2007), with one modification: for each bin, instead of averaging the expression values, the median of non-zero values was taken, as we wanted to reduce sensitivity to outliers.

To observe the variation in H3K9Ac and Spt8 profiles among expressed versus non-expressed genes, or short versus long genes, the median gene profiles were divided into subsets based on gene

properties (Supplementary Figures S7–S10). Genes were divided into five quantiles based on microarray expression values for sgf29 Δ , and median gene profiles were generated for these quantiles in three conditions (mutant, WT and mutant/WT), for both K9Ac and Spt8 data. Genes were then divided into five quantiles by length, and the process was repeated. To observe changes in the above effects on SAGA-dependent genes, we selected a list of genes whose expression levels fell more than two-fold on deletion of SAGA, from the Young lab website (<http://web.wi.mit.edu/young/pub/holstege.html>). The above analyses were then repeated using only this subset of genes. Final gene counts for whole-genome and SAGA-dependent genes are slightly shorter in these analyses (6229/6693) because we only used genes that were measured on both ChIP and expression array. This also shortened the SAGA-dependent list slightly (593/611).

Accession codes

Protein Data Bank: Atomic coordinates and structure factors have been deposited under the following accession codes: 3MP8, 3MP6, 3MP1 for scSgf29 and its complexes with histone H3K4me peptides; 3LX7, 3ME9, 3MEA, 3MET, 3MEU, 3MEV, 3MEW for hsSGF29 and its complexes with histone H3K4me peptides.

Supplementary data

Supplementary data are available at *The EMBO Journal* Online (<http://www.embojournal.org>).

Acknowledgements

We thank SK Swanson, S Nakanishi, MP Washburn, A Shilatfard, G Wasney, M Vedadi, Z Zhu, J Kania, F MacKenzie and A Bochkarev for advice and technical assistance. We also thank Aled Edwards and Cheryl Arrowsmith for critical reading of the manuscript. This research was supported by the Structural Genomics Consortium, a registered charity (number 1097737) that receives funds from the Canadian Institutes for Health Research, the Canadian Foundation for Innovation, Genome Canada through the Ontario Genomics Institute, GlaxoSmithKline, Karolinska Institute, the Knut and Alice Wallenberg Foundation, the Ontario Innovation Trust, the Ontario Ministry for Research and Innovation, Merck & Co., Inc., the Novartis Research Foundation, the Swedish Agency for Innovation Systems, the Swedish Foundation for Strategic Research and the Wellcome Trust to JM. This work was also supported by the grants from the Chinese Ministry of Science and Technology (Nos. 2009CB825502, 2006CB806501), ‘100 talents program’ of the Chinese Academy of Sciences, and the grant from National Natural Science Foundation of China (No. 30970576) to JZ, by a post-doctoral fellowship to KKL from the Damon Runyon Cancer Research Foundation (1751-03) and NIH Grants GM46787 to JLW and CA132878 to GM. Results shown in this report are derived from work performed at Argonne National Laboratory, Structural Biology Center at the Advanced Photon Source. Argonne is operated by UChicago Argonne, LLC, for the US Department of Energy, Office of Biological and Environmental Research under contract DE-AC02-06CH11357.

Author contributions: The overall study was conceived and designed by CX, JR, JZ and JM with important contributions from CB, KKL, TLB, PAG and JLW. CB, CX, JR, KKL, TLB, WT, DB, JL, MW, BOZ, BEF, AP and AAH performed the experiments. All authors analyzed the data. JZ and JM wrote the paper with substantial contributions from CX, JR, KKL and TLB.

Conflict of interest

The authors declare that they have no conflict of interest.

References

- Adams-Cioaba MA, Guo Y, Bian C, Amaya MF, Lam R, Wasney GA, Vedadi M, Xu C, Min J (2010) Structural studies of the tandem Tudor domains of fragile X mental retardation related proteins FXR1 and FXR2. *PLoS One* **5**: e13559
- Adams-Cioaba MA, Min J (2009) Structure and function of histone methylation binding proteins. *Biochem Cell Biol* **87**: 93–105

- Baker SP, Grant PA (2007) The SAGA continues: expanding the cellular role of a transcriptional co-activator complex. *Oncogene* **26**: 5329–5340
- Botuyan MV, Lee J, Ward IM, Kim JE, Thompson JR, Chen J, Mer G (2006) Structural basis for the methylation state-specific recognition of histone H4-K20 by 53BP1 and Crb2 in DNA repair. *Cell* **127**: 1361–1373

- Boyer LA, Latek RR, Peterson CL (2004) The SANT domain: a unique histone-tail-binding module? *Nat Rev Mol Cell Biol* **5**: 158–163
- Bryk M, Briggs SD, Strahl BD, Curcio MJ, Allis CD, Winston F (2002) Evidence that Set1, a factor required for methylation of histone H3, regulates rDNA silencing in *S. cerevisiae* by a Sir2-independent mechanism. *Curr Biol* **12**: 165–170
- Da G, Lenkart J, Zhao K, Shiekhatter R, Cairns BR, Marmorstein R (2006) Structure and function of the SWIRM domain, a conserved protein module found in chromatin regulatory complexes. *Proc Natl Acad Sci USA* **103**: 2057–2062
- Davis IW, Murray LW, Richardson JS, Richardson DC (2004) MOLPROBITY: structure validation and all-atom contact analysis for nucleic acids and their complexes. *Nucleic Acids Res* **32**: W615–W619
- Durso RJ, Fisher AK, Albright-Frey TJ, Reese JC (2001) Analysis of TAF90 mutants displaying allele-specific and broad defects in transcription. *Mol Cell Biol* **21**: 7331–7344
- Eberharter A, Lechner T, Goralik-Schramel M, Loidl P (1996) Purification and characterization of the cytoplasmic histone acetyltransferase B of maize embryos. *FEBS Lett* **386**: 75–81
- Emsley P, Cowtan K (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr* **60**: 2126–2132
- Eng JK, McCormack AL, Yates 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 (Clifton, NJ)* **328**: 159–175
- Gietz RD, Woods RA (2002) Transformation of yeast by lithium acetate/single-stranded carrier DNA/polyethylene glycol method. *Methods Enzymol* **350**: 87–96
- Govind CK, Zhang F, Qiu H, Hofmeyer K, Hinnebusch AG (2007) Gcn5 promotes acetylation, eviction, and methylation of nucleosomes in transcribed coding regions. *Mol Cell* **25**: 31–42
- 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
- Hendrickson WA, Horton JR, LeMaster DM (1990) Selenomethionyl proteins produced for analysis by multiwavelength anomalous diffraction (MAD): a vehicle for direct determination of three-dimensional structure. *EMBO J* **9**: 1665–1672
- Howe L, Auston D, Grant P, John S, Cook RG, Workman JL, Pillus L (2001) Histone H3 specific acetyltransferases are essential for cell cycle progression. *Genes Dev* **15**: 3144–3154
- Huang Y, Fang J, Bedford MT, Zhang Y, Xu RM (2006) Recognition of histone H3 lysine-4 methylation by the double Tudor domain of JMJD2A. *Science (New York, NY)* **312**: 748–751
- Hung T, Binda O, Champagne KS, Kuo AJ, Johnson K, Chang HY, Simon MD, Kutateladze TG, Gozani O (2009) ING4 mediates crosstalk between histone H3 K4 trimethylation and H3 acetylation to attenuate cellular transformation. *Mol Cell* **33**: 248–256
- Ingvarsdottir K, Edwards C, Lee MG, Lee JS, Schultz DC, Shilatifard A, Shiekhatter R, Berger SL (2007) Histone H3 K4 demethylation during activation and attenuation of GAL1 transcription in *Saccharomyces cerevisiae*. *Mol Cell Biol* **27**: 7856–7864
- Jiang L, Smith JN, Anderson SL, Ma P, Mizzen CA, Kelleher NL (2007) Global assessment of combinatorial post-translational modification of core histones in yeast using contemporary mass spectrometry. LYS4 trimethylation correlates with degree of acetylation on the same H3 tail. *J Biol Chem* **282**: 27923–27934
- Kuo MH, Allis CD (1999) *In vivo* cross-linking and immunoprecipitation for studying dynamic protein:DNA associations in a chromatin environment. *Methods* **19**: 425–433
- Kurabe N, Katagiri K, Komiya Y, Ito R, Sugiyama A, Kawasaki Y, Tashiro F (2007) Deregulated expression of a novel component of TFC/STAGA histone acetyltransferase complexes, rat SGF29, in hepatocellular carcinoma: possible implication for the oncogenic potential of c-Myc. *Oncogene* **26**: 5626–5634
- Lee J, Thompson JR, Botuyan MV, Mer G (2008) Distinct binding modes specify the recognition of methylated histones H3K4 and H4K20 by JMJD2A-Tudor. *Nat Struct Mol Biol* **15**: 109–111
- Lee KK, Workman JL (2007) Histone acetyltransferase complexes: one size doesn't fit all. *Nat Rev* **8**: 284–295
- Lee TI, Causton HC, Holstege FC, Shen WC, Hannett N, Jennings EG, Winston F, Green MR, Young RA (2000) Redundant roles for the TFIID and SAGA complexes in global transcription. *Nature* **405**: 701–704
- Li B, Gogol M, Carey M, Pattenden SG, Seidel C, Workman JL (2007) Infrequently transcribed long genes depend on the Set2/Rpd3S pathway for accurate transcription. *Genes Dev* **21**: 1422–1430
- Liu H, Wang JY, Huang Y, Li Z, Gong W, Lehmann R, Xu RM (2010a) Structural basis for methylarginine-dependent recognition of Aubergine by Tudor. *Genes Dev* **24**: 1876–1881
- Liu K, Chen C, Guo Y, Lam R, Bian C, Xu C, Zhao DY, Jin J, Mackenzie F, Pawson T, Min J (2010b) Structural basis for recognition of arginine methylated Piwi proteins by the extended Tudor domain. *Proc Natl Acad Sci USA* **107**: 18398–18403
- Martin DG, Grimes DE, Baetz K, Howe L (2006) Methylation of histone H3 mediates the association of the NuA3 histone acetyltransferase with chromatin. *Mol Cell Biol* **26**: 3018–3028
- Maurer-Stroh S, Dickens NJ, Hughes-Davies L, Kouzarides T, Eisenhaber F, Ponting CP (2003) The Tudor domain 'Royal Family': Tudor, plant Ageten, Chromo, PWWP and MBT domains. *Trends Biochem Sci* **28**: 69–74
- McDonald WH, Ohi R, Miyamoto DT, Mitchison TJ, Yates JR (2002) Comparison of three directly coupled HPLC MS/MS strategies for identification of proteins from complex mixtures: single-dimension LC-MS/MS, 2-phase MudPIT, and 3-phase MudPIT. *Int J Mass Spectrom* **219**: 245–251
- Min J, Zhang X, Cheng X, Grewal SI, Xu RM (2002) Structure of the SET domain histone lysine methyltransferase Clr4. *Nat Struct Biol* **9**: 828–832
- Minor W, Cymborowski M, Otwinowski Z, Chruszcz M (2006) HKL-3000: the integration of data reduction and structure solution from diffraction images to an initial model in minutes. *Acta Crystallogr D Biol Crystallogr* **62**: 859–866
- Nagy Z, Riss A, Fujiyama S, Krebs A, Orpinell M, Jansen P, Cohen A, Stunnenberg HG, Kato S, Tora L (2010) The metazoan ATAC and SAGA coactivator HAT complexes regulate different sets of inducible target genes. *Cell Mol Life Sci* **67**: 611–628
- Otwinowski Z, Minor W (1997) Processing of X-ray diffraction data collected in oscillation mode. In *Methods in Enzymology*, Vol. 276: *Macromolecular Crystallography*. Carter Jr CW, Sweet RM (eds) Part A, pp 307–326. New York: Academic Press
- Perrakis A, Harkiolaki M, Wilson KS, Lamzin VS (2001) ARP/wARP and molecular replacement. *Acta Crystallogr* **57**: 1445–1450
- Qian C, Zhang Q, Li S, Zeng L, Walsh MJ, Zhou MM (2005) Structure and chromosomal DNA binding of the SWIRM domain. *Nat Struct Mol Biol* **12**: 1078–1085
- Ramos A, Hollingworth D, Adinolfi S, Castets M, Kelly G, Frenkiel TA, Bardoni B, Pastore A (2006) The structure of the N-terminal domain of the fragile X mental retardation protein: a platform for protein-protein interaction. *Structure* **14**: 21–31
- Rodriguez-Navarro S (2009) Insights into SAGA function during gene expression. *EMBO Rep* **10**: 843–850
- Saksouk N, Avvakumov N, Champagne KS, Hung T, Doyon Y, Cayrou C, Paquet E, Ullah M, Landry AJ, Cote V, Yang XJ, Gozani O, Kutateladze TG, Cote J (2009) HBO1 HAT complexes target chromatin throughout gene coding regions via multiple PHD finger interactions with histone H3 tail. *Mol Cell* **33**: 257–265
- Sanders SL, Jennings J, Canutescu A, Link AJ, Weil PA (2002) Proteomics of the eukaryotic transcription machinery: identification of proteins associated with components of yeast TFIID by multidimensional mass spectrometry. *Mol Cell Biol* **22**: 4723–4738
- Schneider TR, Sheldrick GM (2002) Substructure solution with SHELXD. *Acta Crystallogr D Biol Crystallogr* **58**: 1772–1779
- Sermwittayawong D, Tan S (2006) SAGA binds TBP via its Spt8 subunit in competition with DNA: implications for TBP recruitment. *EMBO J* **25**: 3791–3800
- 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
- Taverna SD, Ilin S, Rogers RS, Tanny JC, Lavender H, Li H, Baker L, Boyle J, Blair LP, Chait BT, Patel DJ, Aitchison JD, Tackett AJ, Allis CD (2006) Yng1 PHD finger binding to H3 trimethylated at K4 promotes NuA3 HAT activity at K14 of H3 and transcription at a subset of targeted ORFs. *Mol Cell* **24**: 785–796

- Vagin A, Teplyakov A (2000) An approach to multi-copy search in molecular replacement. *Acta Crystallogr* **56**: 1622–1624
- Vagin AA, Steiner RA, Lebedev AA, Potterton L, McNicholas S, Long F, Murshudov GN (2004) REFMAC5 dictionary: organization of prior chemical knowledge and guidelines for its use. *Acta Crystallogr D Biol Crystallogr* **60**: 2184–2195
- Venters BJ, Wachi S, Mavrich TN, Andersen BE, Jena P, Sinnamoni AJ, Jain P, Roller NS, Jiang C, Hemeryck-Walsh C, Pugh BF (2011) A comprehensive genomic binding map of gene and chromatin regulatory proteins in *Saccharomyces*. *Mol Cell* **41**: 480–492
- Vermeulen M, Eberl HC, Matarese F, Marks H, Denissov S, Butter F, Lee KK, Olsen JV, Hyman AA, Stunnenberg HG, Mann M (2010) Quantitative interaction proteomics and genome-wide profiling of epigenetic histone marks and their readers. *Cell* **142**: 967–980
- Wang BC (1985) Resolution of phase ambiguity in macromolecular crystallography. *Methods Enzymol* **115**: 90–112
- Washburn MP, Wolters D, Yates III JR (2001) Large-scale analysis of the yeast proteome by multidimensional protein identification technology. *Nat Biotechnol* **19**: 242–247
- Wu PY, Winston F (2002) Analysis of Spt7 function in the *Saccharomyces cerevisiae* SAGA coactivator complex. *Mol Cell Biol* **22**: 5367–5379
- Xu C, Bian C, Lam R, Dong A, Min J (2011) The structural basis for selective binding of non-methylated CpG islands by the CFP1 CXXC domain. *Nat Commun* **2**: 227
- Xu C, Bian C, Yang W, Galka M, Ouyang H, Chen C, Qiu W, Liu H, Jones AE, MacKenzie F, Pan P, Li SS, Wang H, Min J (2010) Binding of different histone marks differentially regulates the activity and specificity of polycomb repressive complex 2 (PRC2). *Proc Natl Acad Sci USA* **107**: 19266–19271
- Zeng L, Zhou MM (2002) Bromodomain: an acetyl-lysine binding domain. *FEBS Lett* **513**: 124–128
- Zhang K, Mosch K, Fischle W, Grewal SI (2008) Roles of the Ctr4 methyltransferase complex in nucleation, spreading and maintenance of heterochromatin. *Nat Struct Mol Biol* **15**: 381–388
- Zhang W, Bone JR, Edmondson DG, Turner BM, Roth SY (1998) Essential and redundant functions of histone acetylation revealed by mutation of target lysines and loss of the Gcn5p acetyltransferase. *EMBO J* **17**: 3155–3167