STRUCTURAL BIOLOGY COMMUNICATIONS

ISSN 2053-230X

Structural and functional analysis of the SET3 histone deacetylase complex

Alexis A. Reyes,[a,b] Susan Fishbain[a] and Yuan He[a,b,c,d]*

[a]Molecular Biosciences, Northwestern University, Evanston, Illinois, USA, [b]Interdisciplinary Biological Sciences Program, Northwestern University, Evanston, Illinois, USA, [c]Chemistry of Life Processes Institute, Northwestern University, Evanston, Illinois, USA, and [d]Robert H. Lurie Comprehensive Cancer Center of Northwestern University, Northwestern University, Chicago, Illinois, USA. *Correspondence e-mail: yuanhe@northwestern.edu

Received 18 November 2021
Accepted 16 January 2022

Edited by K. K. Kim, Sungkyunkwan University School of Medicine, Republic of Korea

Keywords: SET3 complex; histone deacetylation; electron microscopy.

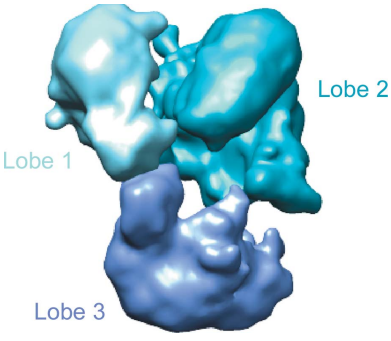
Supporting information: this article has supporting information at journals.iucr.org/f

The SET3 complex (SET3C) is a seven-subunit histone deacetylase complex that is capable of transcriptional regulation. Methylated histone 3 marks recruit SET3C to the nucleosome, and the SET3C catalytic subunits deacetylate the histone 3 and 4 tails. There is very limited structural knowledge of the SET3C subunits, with most subunits having unknown structures or functions. Here, a catalytically active SET3 complex was endogenously purified from *Saccharomyces cerevisiae* and utilized for negative-stain electron microscopy (EM) to determine an apo model for the holo complex. The negative-stain EM 3D model revealed a three-lobe architecture, with each lobe extending from a central point.

1. Introduction

Epigenetic regulation at the nucleosome level heavily relies on the post-translational modification of histone tails. Some of the most widely studied modifications are histone methylation and acetylation, which are involved in recruiting transcription regulatory factors. Histone acetylation is typically associated with open chromatin and therefore active transcription, while methylation is associated with closed chromatin, although there are exceptions (Struhl, 1998; Zhang & Reinberg, 2001; Kouzarides, 2002). The placement and removal of histone methylation and acetylation marks are tightly coordinated to ensure proper transcription regulation through the accessibility of chromatin to the transcriptional machinery. In yeast, the SET [Su(var)3–9, enhancer-of-zeste and trithorax] family of chromatin modifiers consist of complexes responsible for histone methylation, typically through their characteristic SET domains (Jenuwein, 2006). Set1 is a methyltransferase that is known to interact with RNA polymerase II (Pol II) and to deposit methylation marks after Pol II has transcribed through a nucleosome (Ng *et al.*, 2003; Terzi *et al.*, 2011; Bae *et al.*, 2020). The monomethylation, dimethylation and trimethylation of H3 lysine 4 (H3K4) by Set1 serve to recruit SET3C for histone deacetylation (Kim & Buratowski, 2009).

SET3C consists of Set3, Snt1, Sif2, Hos4, Cpr1 and two deacetylase subunits: Hst1 and Hos2. Hst1 is an NAD-dependent histone deacetylase, while Hos2-catalyzed deacetylation is NAD-independent (Pijnappel *et al.*, 2001). The Set3 subunit has two known domains: SET and PHD (plant homeodomain). The PHD finger of Set3 binds dimethylated H3K4 with high affinity (Kim & Buratowski, 2009). The function of the Set3 SET domain is currently unknown. However, SET domains typically serve as catalytic centers for histone modifications and the mediation of protein–protein



interactions (Rozenblatt-Rosen *et al.*, 1998; Dillon *et al.*, 2005; Leng *et al.*, 2021). The functions of Hos4 and Cpr1 within SET3C are not fully understood, but these subunits are likely to serve to stabilize and regulate the complex. A more complete understanding of the architecture and function of SET3C is needed to determine the specific roles of the various SET3C subunits. In this study, we provide a structural model of SET3C using negative-stain electron microscopy and use predicted secondary structures and known domains to propose a model for SET3C organization.

2. Materials and methods

2.1. SET3C purification

SET3C was purified from a yeast strain obtained from the High Throughput Analysis Laboratory at Northwestern University and contained a tandem affinity purification (TAP) tag on the Set3 subunit (Ghaemmaghami *et al.*, 2003). A TAP procedure was performed as described previously (Han *et al.*, 2020) with a few modifications. The yeast strain was grown to an OD_{600} of 4–5 in 16 l YPD broth (containing 2% glucose). Cells were collected by centrifugation and washed with 320 ml cold TAP extraction buffer (40 m*M* Tris pH 7.9, 250 m*M* ammonium sulfate, 1 m*M* EDTA, 10% glycerol, 0.1% Tween-20). The washed cells were resuspended in 200 ml cold TAP extraction buffer consisting of 5 m*M* DTT, 2 m*M* PMSF, 0.31 mg ml^{-1} benzamidine, 0.3 mg ml^{-1} leupeptin, 1.4 mg ml^{-1} pepstatin and 2 mg ml^{-1} chymostatin. The cell resuspension was lysed with glass beads in a BeadBeater (Biospec Products) for 15 rounds of 30 s on and 1 min off. Cell debris was removed by centrifugation at 14 000*g* at 4°C for 1.5 h. 3 ml IgG Sepharose beads (GE Healthcare) were incubated with the cleared lysate at 4°C overnight. The IgG column was washed with TAP extraction buffer and cold TEV cleavage buffer (10 m*M* Tris pH 8, 150 m*M* NaCl, 0.1% NP-40, 0.5 m*M* EDTA, 10% glycerol). For cleavage, the IgG beads were resuspended in 6 ml TEV cleavage buffer and 30 mg TEV protease at room temperature for 1 h with gentle shaking. After TEV cleavage on the IgG column, the beads were washed with three column volumes of cold calmodulin-binding buffer (cbb; 15 m*M* HEPES pH 7.8, 1 m*M* magnesium acetate, 1 m*M* imidazole, 2 m*M* $CaCl_2$, 0.1% NP-40, 10% glycerol, 300 m*M* KCl, 5 m*M* DTT, 2 m*M* PMSF, 0.31 mg ml^{-1} benzamidine, 0.3 µg ml^{-1} leupeptin, 1.4 µg ml^{-1} pepstatin, 2 µg ml^{-1} chymostatin). Calcium chloride was added to the combined IgG eluate to reach a final concentration of 2 m*M* and it was then incubated with calmodulin affinity resin (Agilent Technologies) at 4°C for 1.5 h. The calmodulin beads were then washed with cold cbb and calmodulin wash buffer (cwb; the same as cbb but with 0.05% NP-40). Bound protein complexes were eluted at room temperature with calmodulin elution buffer (ceb; 15 m*M* HEPES pH 7.8, 1 m*M* magnesium acetate, 1 m*M* imidazole, 2 m*M* EGTA, 10% glycerol, 0.05% NP-40, 300 m*M* KCl). Fractions containing SET3C were identified using SDS–PAGE, aliquoted, flash-frozen in liquid nitrogen and stored at −80°C.

2.2. Domain graphs and deacetylase activity assay

All domain graphs were generated using code available on GitHub (https://github.com/avibpatel/domainsGraph). The available code combines *ConSurf*, *PSIPRED*, *DISOPRED* and PFAM data into single images (Ashkenazy *et al.*, 2010, 2016; Buchan & Jones, 2019; Celniker *et al.*, 2013; Jones, 1999; Jones & Cozzetto, 2015; Mistry *et al.*, 2021).

Activity assays were performed as described previously (Marcum & Radhakrishnan, 2019; Schultz *et al.*, 2004) with slight modifications. An acetylated peptide containing an aminocoumarin (AMC) moiety, Ac-Gly-Ala-Lys(Ac)-AMC, was purchased from Bachem Americas and resuspended in deacetylase reaction buffer (50 m*M* HEPES pH 7.4, 100 m*M* KCl, 0.005% Tween-20, 5% DMSO). Reactions were performed in 96-well plates with 40 m*M* AMC peptide, 0.8 m*M* recombinant human HDAC8 (Novus Biologicals) or SET3C, 200 n*M* trypsin and deacetylase reaction buffer at room temperature for 1 h. There is roughly a 40-fold difference between the abundance of HDAC8 and the catalytically active subunits of SET3C. We estimate that 0.8 m*M* of our purified SET3C contains about 0.02 m*M* Hos2 and Hst1. Fluorescence was quantified using a Gemini EM fluorescence/chemiluminescence plate reader with excitation and emission wavelengths set to 360 and 460 nm, respectively. Two tailed heteroscedastic *t*-tests were performed to determine the statistical significance between the negative control and HDAC8 or SET3C averaged fluorescence.

2.3. Negative-stain electron microscopy

Purified SET3C was diluted 100-fold in calmodulin elution buffer and cross-linked with 0.05% glutaraldehyde for 30 min at 4°C. A carbon-coated 400 mesh copper grid was plasma-cleaned with a Gatan Solarus for 10 s at 25 W before applying the cross-linked SET3C and incubating for 10 min at high humidity in a homemade humidity chamber. The grid was stained with 2% uranyl formate by massaging the grid on four 50 ml droplets for 5, 10, 15 and 20 s. Excess stain was blotted away with filter paper (Whatman No. 1) and the grids were stored until imaging. All negative-stain images were collected with a JEOL 1400 120 kV TEM and a 4K × 4K CCD camera using *Leginon* 3.2 for automated data collection (Suloway *et al.*, 2005; Cheng *et al.*, 2021).

Images were captured at 30K magnification with a pixel size of 3.71 Å and CTF-corrected with *CTFFIND3* (Mindell & Grigorieff, 2003). 17 594 particles were picked reference-free using the DoG picker and were aligned to generate 2D class averages with a box size of 96 pixels and a circular mask with a diameter of 40 pixels using iterative MSA/MRA (Voss *et al.*, 2009; Ogura *et al.*, 2003). An additional 66 836 particles were automatically picked using *RELION* 2.1 and an *ab initio* reference was generated with *cryoSPARC* 2.0 (Kimanius *et al.*, 2016; Punjani *et al.*, 2017). The final model was refined with *EMAN*2 and low-pass filtered to 15 Å (Tang *et al.*, 2007). The images in Fig. 3 were generated using the volume and coloring tools within *UCSF Chimera* (Pettersen *et al.*, 2004; Goddard *et al.*, 2007).

3. Results and discussion

3.1. SET3C is composed of highly ordered subunits

Yeast SET3C consists of seven subunits that are all highly ordered and contain significantly conserved sequences in either their known or predicted domains (Fig. 1). The one exception is Snt1, which is predicted to be highly disordered outside the two Myb DNA-binding domains (Fig. 1). Snt1 is known to interact with Sif2, which is believed to exist within the complex as a dimer or tetramer (Pijnappel *et al.*, 2001), which is consistent with the relatively stronger band that corresponds to it during biochemical characterization (Fig. 2*a*). Sif2 has an N-terminal LisH domain that mediates oligomerization and its interaction with Snt1 (Cerna & Wilson, 2005). It is unknown whether the observed Sif2 dimer or tetramer is biologically relevant on its own or how the oligomerization state of Sif2 affects the incorporation of Snt1. However, the stoichiometry of Sif2 and Snt1 has previously been shown to be 1:1, which suggests that either two or four copies of Snt1 would be incorporated to match Sif2 (Pijnappel *et al.*, 2001). The C-terminal WD40 repeats of Sif2 are capable of mediating interactions with other subunits and may act as a scaffold for Snt1 and Set3 interactions (Stirnimann *et al.*, 2010; Jain & Pandey, 2018). Set3 is required for Sif2 and Snt1 to be incorporated into SET3C (Pijnappel *et al.*, 2001). The SET domain of Set3 potentially functions to mediate protein–protein interactions, since Hos2 and Hst1 are known to be SET3C deacetylases and the Set3 PHD domain binds the methylated tails of histone 3 (Pijnappel *et al.*, 2001; Kim & Buratowski, 2009; Gatchalian *et al.*, 2017). The Set3 PHD and Snt1 Myb DNA-binding domains, stabilized by Sif2 and the Set3 SET domain, can therefore make up the nucleosome-binding region of SET3C.

Hst1 is a Sir2 class NAD-dependent histone deacetylase (Fig. 1). Hst1 is suggested to be incorporated into SET3C by Hos4, but it is not essential for complex integrity and can function independently of SET3C (Pijnappel *et al.*, 2001). Hos2 is a class I histone deacetylase, similar to Rpd3, and is potentially required for complex integrity (Wang *et al.*, 2002; Pijnappel *et al.*, 2001). The incorporation of Cpr1, the smallest subunit within SET3C, is mediated by Hos2, Sif2 and Snt1, but its function is still not fully understood (Arévalo-Rodríguez *et al.*, 2000). It has been established that Cpr1 is essential for regulating Rpd3, a histone deacetylase that is related to Hos2 (Arévalo-Rodríguez & Heitman, 2005). Cpr1 may play a similar regulatory role for Hos2 within SET3C. A recent study suggests that Cpr1 is involved in SET3C recruitment to nucleosomes through a sumoylation interacting motif (Ryu *et al.*, 2020). Lastly, Hos4 is essential for the stability of SET3C and its multiple Ank domains can mediate interactions between proteins within SET3C (Fig. 1). Considering its necessity for complex integrity, Hos4 potentially functions as a bridge between the histone deacetylases Hos2 and Hst1 and the nucleosome-binding region consisting of Set3, Cpr1, Sif2 and Snt1.

3.2. SET3C activity and architecture

To determine the molecular architecture of SET3C, we purified the full-length endogenous complex from yeast using the well established tandem affinity purification (TAP) method. The TAP tag on the Set3 subunit allowed the co-purification of all complex components (Fig. 2*a*). Consistent with previous reports, Hos2 and Hst1 run together on the denaturing gel, with Hos2 likely to be in excess relative to Hst1. Cpr1 is the smallest and least intense band, which is

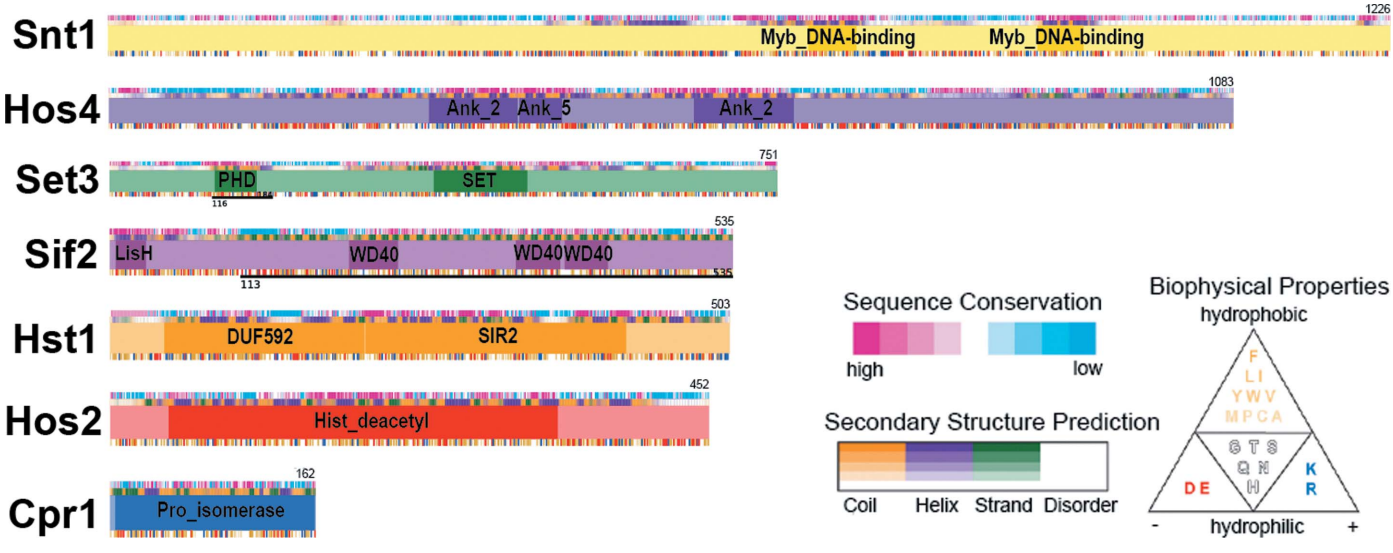


Figure 1
SET3C domain maps. Cartoon representation of the seven subunits that constitute SET3C. Each map contains the protein length in amino acids, the identification of known or predicted domains, sequence conservation, secondary-structure prediction and biophysical properties for each residue within the respective protein. The keys show the color spectrum for sequence conservation, secondary-structure prediction and biophysical properties. Domains with known 3D structures are underlined with black lines. Domain maps were generated using sequences, conservation scores, structure prediction and domain identification from UniProt, *ConSurf*, *PSIPRED* 4.0, *DISOPRED*3 and PFAM.

consistent with a previous publication revealing the low stoichiometry of Cpr1 within SET3C (Pijnappel *et al.*, 2001). A few contaminants are present in the final purified product, which is consistent with previous reports of nonspecific contamination by various heat-shock and glycolysis proteins under the purification conditions described above (Fig. 2*a*; Pijnappel *et al.*, 2001; Kim & Buratowski, 2009).

To determine the deacetylase ability of purified SET3C, we used an acetylated peptide that mimics histone tails with an AMC moiety on the C-terminal end. Upon lysine deacetylation, trypsin can cleave AMC, allowing fluorescence from the freed moiety to be detected. For each reaction condition, the substrate, trypsin and enzymes were incubated for 60 min before fluorescence readings were taken (Fig. 2*b*). We used reconstituted human HDAC8 as a positive control since it was previously determined to work well with the AMC peptide (Schultz *et al.*, 2004). In the presence of HDAC8, the relative fluorescence reading significantly increases compared with a negative control, signifying that HDAC8 efficiently deacetylates the peptide. While SET3C is not as efficient at deacetylating the AMC peptide as HDAC8, SET3C also shows a significant increase in fluorescence over the negative control (Fig. 2*b*).

Next, we used negative-stain electron microscopy to determine the overall architecture of functional SET3C (Fig. 3). A reference-free 2D classification of SET3C particle images reveals multiple orientations, with SET3C having three lobes (Fig. 3*a*). The three distinct lobes of SET3C were confirmed upon 3D classification and model refinement (Figs. 3*b* and 3*c*). Our model shows that SET3C adopts a Y-shaped architecture, with lobes 2 and 3 having extra density extending from the ends that does not interact with lobe 1 (Fig. 3*c*). We used the measure volume tool within *UCSF ChimeraX* to estimate the volume of the full density depicted in Figs. 3(*b*) and 3(*c*) (Pettersen *et al.*, 2021; Goddard *et al.*, 2018). The total volume was calculated to be 7.04×10^5 Å^3 at a threshold of 2.59. To determine the theoretical volume of SET3C, we utilized the *AlphaFold* fetch command within *UCSF ChimeraX* and depicted the predicted structure of each subunit as a surface using the molmap command, before measuring the volumes (Jumper *et al.*, 2021). Since the molmap command generates density maps for atomic structures that are proportional to the resolution of the structure, we calculated the volumes of maps generated at 4.5, 10 and 15 Å resolution. Assuming that SET3C contains two Sif2–Snt1 heterodimers, the theoretical volumes of SET3C based on *AlphaFold* structures were calculated to be 1.04×10^6, 1.61×10^6 and 1.78×10^6 Å^3 at 4.5, 10 and 15 Å resolution, respectively (Supplementary Table S1). Given these calculations, about 40–68% of SET3C is predicted to fit into our 3D reconstruction. However, the predicted *AlphaFold* structures do not consider how the inter-subunit interactions within

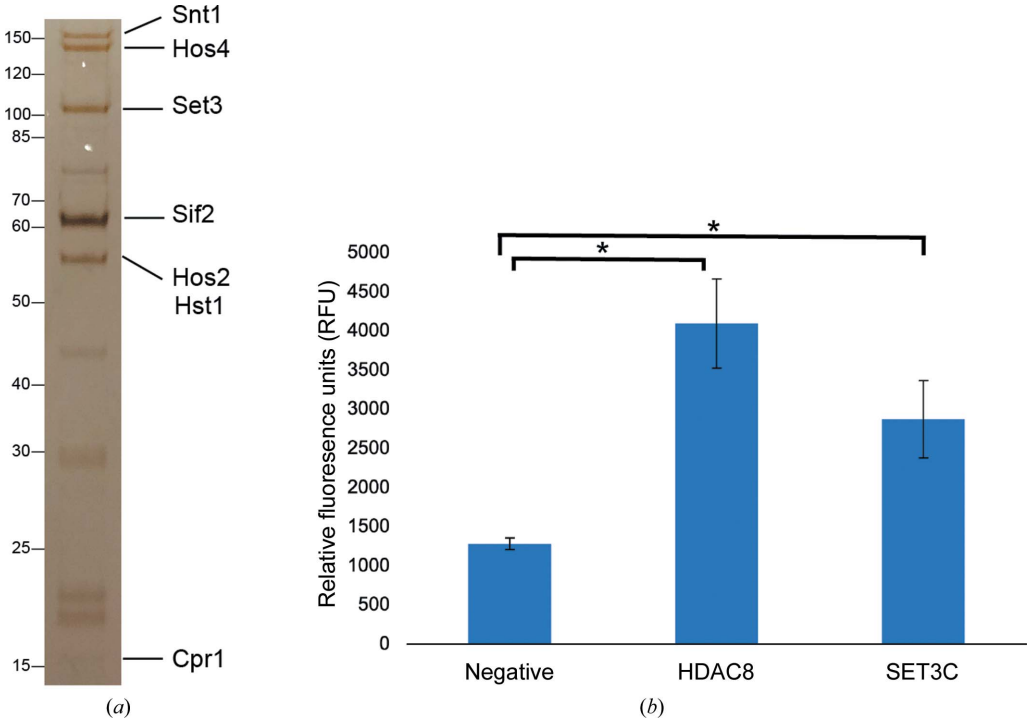


Figure 2
SET3C protein purification and activity. (*a*) Tris–glycine denaturing gel (4–12%) obtained by silver staining of the SET3C TAP purification final product. Since Set3 still has the calmodulin tag, it runs on the gel closer to 100 kDa rather than the expected 85 kDa. Unlabeled bands correspond to common nonspecific contaminants, including Ssa2, Ssb1, HSP60 and enolase 2. (*b*) Fluorescence-based deacetylase assay. Negative-control, HDAC8 and SET3C reaction conditions all contain 40 m*M* AMC peptide and 200 m*M* trypsin and were incubated for 60 min at room temperature. Standard deviations are shown as error bars. Asterisks indicate statistically significant differences between conditions with *p*-values less than 0.05. Previous work has tested the ability of NAD to enhance histone deacetylation by SET3C (Pijnappel *et al.*, 2001). However, our assay did not show any significant difference in deacetylation in the presence of NAD (data not shown). This discrepancy is likely to be due to Hos2 being in excess relative to the NAD-dependent Hst1. The similar SET3C activity seen in the presence or absence of NAD can therefore be caused by the prominence of Hos2-mediated deacetylation.

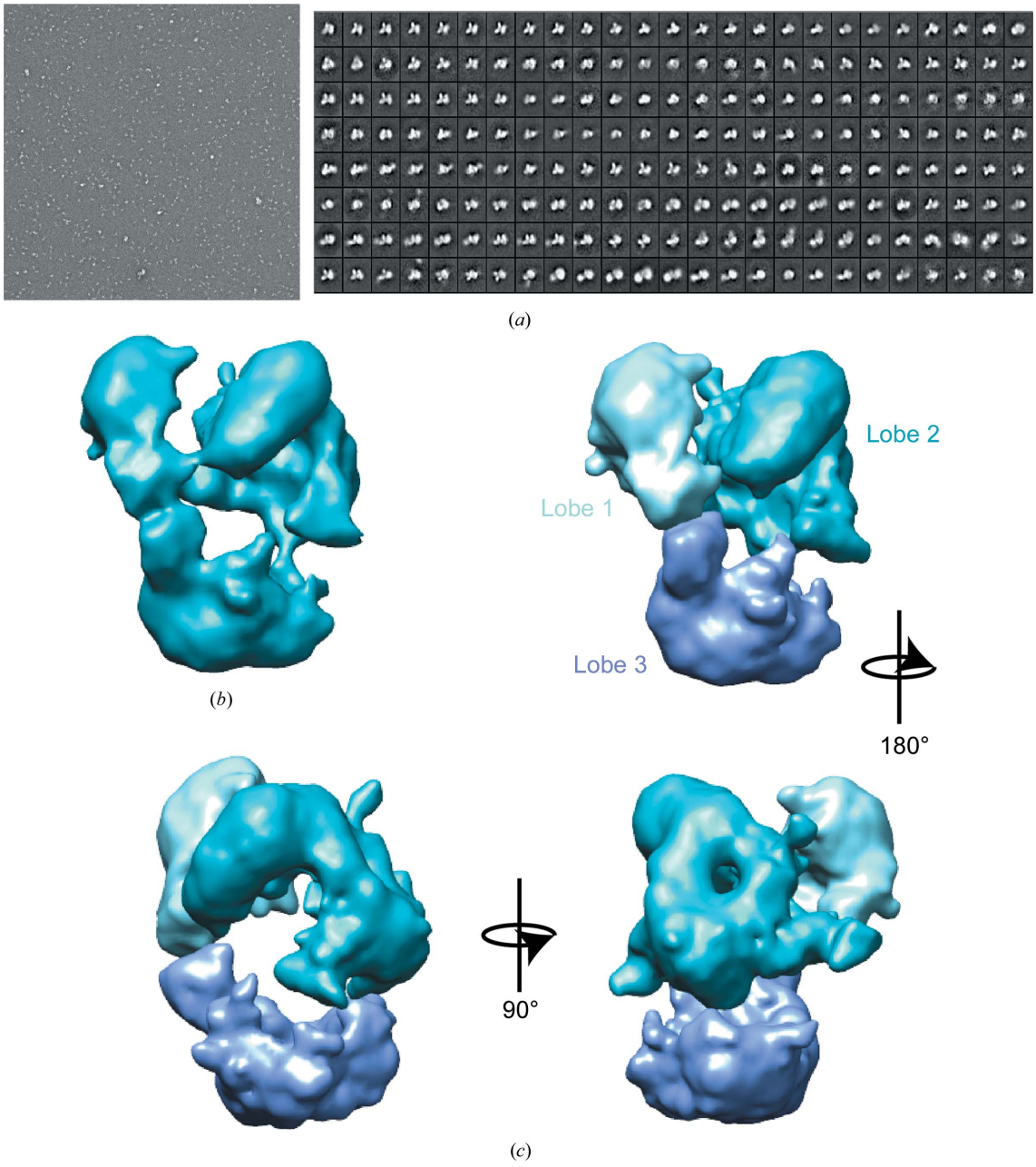


Figure 3
3D reconstruction of SET3C. (*a*) Representative micrograph and 2D class averages generated from reference-free alignment by iterative MSA/MRA. (*b*) Auto-refinement of the most populated 3D model identified through 3D classification using an *ab initio* reference model. (*c*) Multiple views of the 3D reconstruction of SET3C, with each lobe differentiated by color.

SET3C may change the conformations of the large, disordered regions during complex assembly. Nor do the corresponding maps account for the likelihood that disordered regions are not visible in our negative-stain reconstruction. Assuming that disordered regions are not visible in our negative-stain density, we removed large, disordered regions within the *AlphaFold* structures of Hos4, Snt1, Sif2 and Set3 and used these clipped models to estimate the volumes of SET3C at different resolutions (Supplementary Table S1). We found that 60–99% of SET3C without large, disordered regions fitted into our 3D reconstruction. Further structural investigation with cryo-EM will improve the resolution, allowing atomic model building and identification of the various subunits within the model.

The negative-stain model and deacetylase assay presented here reveal that the apo structure of SET3C is stable and active. However, incorporating a methylated and acetylated mononucleosome into the EM sample preparation will provide more information regarding the binding and catalytic abilities of SET3C.

Acknowledgements

We thank R. Marcum for guidance with developing the deacetylase activity assay, A. Talyzina for assistance in determining initial EM processing parameters, the staff at the Structural Biology Facility and the Keck Biophysics Facility,

which are shared resources of the Robert H. Lurie Comprehensive Cancer Center of Northwestern University supported in part by NCI Cancer Support Grant No. P30 CA060553, for technical and computer support, R. Abdella, K. Berman and S. Chen for helpful comments and edits to the manuscript, and C. Koo for being the best. Molecular graphics and analyses were obtained using *UCSF ChimeraX*, developed by the Resource for Biocomputing, Visualization and Informatics at the University of California, San Francisco with support from National Institutes of Health R01-GM129325 and the Office of Cyber Infrastructure and Computational Biology, National Institute of Allergy and Infectious Diseases.

Funding information

Funding for this research was provided by a Cornew Innovation Award from the Chemistry of Life Processes Institute Northwestern University (Yuan He); a Catalyst Award from Chicago Biomedical Consortium (Yuan He); American Cancer Society (grant No. IRG-15-173-21 to Yuan He); an H Foundation Core Facility Pilot Project Award (award No. U54-CA193419 to Yuan He); National Institutes of Health (grant Nos. R01-GM135651 and P01-CA092584 to Yuan He; grant No. 1S10OD026963-01); National Institutes of Health, National Institute of General Medical Sciences – Molecular Biophysics Training Program (grant No. T32-GM008382 to Alexis A. Reyes); Robert H. Lurie Comprehensive Cancer Center (grant No. NCI CCSG P30 CA060553).

References

Arévalo-Rodríguez, M., Cardenas, M. E., Wu, X., Hanes, S. D. & Heitman, J. (2000). *EMBO J.* **19**, 3739–3749.
Arévalo-Rodríguez, M. & Heitman, J. (2005). *Eukaryot. Cell*, **4**, 17–29.
Ashkenazy, H., Abadi, S., Martz, E., Chay, O., Mayrose, I., Pupko, T. & Ben-Tal, N. (2016). *Nucleic Acids Res.* **44**, W344–W350.
Ashkenazy, H., Erez, E., Martz, E., Pupko, T. & Ben-Tal, N. (2010). *Nucleic Acids Res.* **38**, W529–W533.
Bae, H. J., Dubarry, M., Jeon, J., Soares, L. M., Dargemont, C., Kim, J., Geli, V. & Buratowski, S. (2020). *Nat. Commun.* **11**, 2181.
Buchan, D. W. A. & Jones, D. T. (2019). *Nucleic Acids Res.* **47**, W402–W407.
Celniker, G., Nimrod, G., Ashkenazy, H., Glaser, F., Martz, E., Mayrose, I., Pupko, T. & Ben-Tal, N. (2013). *Isr. J. Chem.* **53**, 199–206.
Cerna, D. & Wilson, D. K. (2005). *J. Mol. Biol.* **351**, 923–935.
Cheng, A., Negro, C., Bruhn, J. F., Rice, W. J., Dallakyan, S., Eng, E. T., Waterman, D. G., Potter, C. S. & Carragher, B. (2021). *Protein Sci.* **30**, 136–150.
Dillon, S. C., Zhang, X., Trievel, R. C. & Cheng, X. (2005). *Genome Biol.* **6**, 227.
Gatchalian, J., Ali, M., Andrews, F. H., Zhang, Y., Barrett, A. S. & Kutateladze, T. G. (2017). *J. Mol. Biol.* **429**, 2066–2074.
Ghaemmaghami, S., Huh, W. K., Bower, K., Howson, R. W., Belle, A., Dephoure, N., O'Shea, E. K. & Weissman, J. S. (2003). *Nature*, **425**, 737–741.
Goddard, T. D., Huang, C. C. & Ferrin, T. E. (2007). *J. Struct. Biol.* **157**, 281–287.
Goddard, T. D., Huang, C. C., Meng, E. C., Pettersen, E. F., Couch, G. S., Morris, J. H. & Ferrin, T. E. (2018). *Protein Sci.* **27**, 14–25.
Han, Y., Reyes, A. A., Malik, S. & He, Y. (2020). *Nature*, **579**, 452–455.
Jain, B. P. & Pandey, S. (2018). *Protein J.* **37**, 391–406.
Jenuwein, T. (2006). *FEBS J.* **273**, 3121–3135.
Jones, D. T. (1999). *J. Mol. Biol.* **292**, 195–202.
Jones, D. T. & Cozzetto, D. (2015). *Bioinformatics*, **31**, 857–863.
Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A., Bridgland, A., Meyer, C., Kohl, S. A. A., Ballard, A. J., Cowie, A., Romera-Paredes, B., Nikolov, S., Jain, R., Adler, J., Back, T., Petersen, S., Reiman, D., Clancy, E., Zielinski, M., Steinegger, M., Pacholska, M., Berghammer, T., Bodenstein, S., Silver, D., Vinyals, O., Senior, A. W., Kavukcuoglu, K., Kohli, P. & Hassabis, D. (2021). *Nature*, **596**, 583–589.
Kim, T. & Buratowski, S. (2009). *Cell*, **137**, 259–272.
Kimanius, D., Forsberg, B. O., Scheres, S. H. W. & Lindahl, E. (2016). *eLife*, **5**, e18722.
Kouzarides, T. (2002). *Curr. Opin. Genet. Dev.* **12**, 198–209.
Leng, H., Liu, S., Lei, Y., Tang, Y., Gu, S., Hu, J., Chen, S., Feng, J. & Li, Q. (2021). *Nucleic Acids Res.* **49**, 5502–5519.
Marcum, R. D. & Radhakrishnan, I. (2019). *J. Biol. Chem.* **294**, 13928–13938.
Mindell, J. A. & Grigorieff, N. (2003). *J. Struct. Biol.* **142**, 334–347.
Mistry, J., Chuguransky, S., Williams, L., Qureshi, M., Salazar, G., Sonnhammer, E. L. L., Tosatto, S. C. E., Paladin, L., Raj, S., Richardson, L. J., Finn, R. D. & Bateman, A. (2021). *Nucleic Acids Res.* **49**, D412–D419.
Ng, H. H., Robert, F., Young, R. A. & Struhl, K. (2003). *Mol. Cell*, **11**, 709–719.
Ogura, T., Iwasaki, K. & Sato, C. (2003). *J. Struct. Biol.* **143**, 185–200.
Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C. & Ferrin, T. E. (2004). *J. Comput. Chem.* **25**, 1605–1612.
Pettersen, E. F., Goddard, T. D., Huang, C. C., Meng, E. C., Couch, G. S., Croll, T. I., Morris, J. H. & Ferrin, T. E. (2021). *Protein Sci.* **30**, 70–82.
Pijnappel, W. W. M. P., Schaft, D., Roguev, A., Shevchenko, A., Tekotte, H., Wilm, M., Rigaut, G., Séraphin, B., Aasland, R. & Stewart, A. F. (2001). *Genes Dev.* **15**, 2991–3004.
Punjani, A., Rubinstein, J. L., Fleet, D. J. & Brubaker, M. A. (2017). *Nat. Methods*, **14**, 290–296.
Rozenblatt-Rosen, O., Rozovskaia, T., Burakov, D., Sedkov, Y., Tillib, S., Blechman, J., Nakamura, T., Croce, C. M., Mazo, A. & Canaani, E. (1998). *Proc. Natl Acad. Sci. USA*, **95**, 4152–4157.
Ryu, H.-Y., Zhao, D., Li, J., Su, D. & Hochstrasser, M. (2020). *Nucleic Acids Res.* **48**, 12151–12168.
Schultz, B. E., Misialek, S., Wu, J., Tang, J., Conn, M. T., Tahilramani, R. & Wong, L. (2004). *Biochemistry*, **43**, 11083–11091.
Stirnimann, C. U., Petsalaki, E., Russell, R. B. & Müller, C. W. (2010). *Trends Biochem. Sci.* **35**, 565–574.
Struhl, K. (1998). *Genes Dev.* **12**, 599–606.
Suloway, C., Pulokas, J., Fellmann, D., Cheng, A., Guerra, F., Quispe, J., Stagg, S., Potter, C. S. & Carragher, B. (2005). *J. Struct. Biol.* **151**, 41–60.
Tang, G., Peng, L., Baldwin, P. R., Mann, D. S., Jiang, W., Rees, I. & Ludtke, S. J. (2007). *J. Struct. Biol.* **157**, 38–46.
Terzi, N., Churchman, L. S., Vasiljeva, L., Weissman, J. & Buratowski, S. (2011). *Mol. Cell. Biol.* **31**, 3569–3583.
Voss, N. R., Yoshioka, C. K., Radermacher, M., Potter, C. S. & Carragher, B. (2009). *J. Struct. Biol.* **166**, 205–213.
Wang, A., Kurdistani, S. K. & Grunstein, M. (2002). *Science*, **298**, 1412–1414.
Zhang, Y. & Reinberg, D. (2001). *Genes Dev.* **15**, 2343–2360.