bioRxiv preprint doi: https://doi.org/10.1101/628149; this version posted May 5, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

TITLE

The structure of the yeast Ctf3 complex

AUTHORS

Stephen M. Hinshaw[1], Andrew N. Dates[2], Stephen C. Harrison[1]

AFFILIATIONS

[1] Harvard Medical School, Howard Hughes Medical Institute
[2] Harvard Chemical Biology Ph.D. Program

CONTACT

hinshaw@crystal.harvard.edu, harrison@crystal.harvard.edu

ABSTRACT

Kinetochores are the chromosomal attachment points for spindle microtubules. They are also signaling hubs that control major cell cycle transitions and coordinate chromosome folding. Most well-studied eukaryotes rely on a conserved set of factors, which are divided among two loosely-defined groups, for these functions. Outer kinetochore proteins contact microtubules or regulate this contact directly. Inner kinetochore proteins designate the kinetochore assembly site by recognizing a specialized nucleosome containing the H3 variant Cse4/CENP-A. We previously determined the structure, resolved by cryo-electron microscopy (cryo-EM), of the yeast Ctf19 complex (Ctf19c, homologous to the vertebrate CCAN), providing a high-resolution view of inner kinetochore architecture. We now extend these observations by reporting a near-atomic model of the Ctf3 complex, the outermost Ctf19c sub-assembly seen in our original cryo-EM density. The model is sufficiently well-determined by the new data to enable molecular interpretation of Ctf3 recruitment and function.

INTRODUCTION

Inner kinetochore proteins, particularly members of the Ctf19c, constitute the foundation of the kinetochore; they both recognize Cse4/CENP-A and recruit microtubule-interacting outer kinetochore proteins (Musacchio and Desai, 2017). Many of these factors were originally cloned and characterized for their contributions to faithful mitotic chromosome segregation (Hinshaw and Harrison, 2017). Specific Ctf19c activities that support chromosome segregation include specification of centromere-proximal origins as early-firing, recruitment of the Scc2/4 cohesin loading complex and its payload, the cohesin complex, and, as has been recently determined, recruitment of the chromosome passenger complex and its essential kinase subunit, Ipl1/Aurora B (Fernius et al., 2013; Garcia-Rodriguez et al., 2019; Halwachs et al., 2018; Natsume et al., 2013). Recent findings also suggest a handoff between Ctf19c-dependent microtubule contact points, whereby the link between centromeric DNA and spindle microtubules shifts between chromosomal tethers during the course of the cell cycle (Bock et al., 2012; Hara et al., 2018).

bioRxiv preprint doi: https://doi.org/10.1101/628149; this version posted May 5, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Understanding coordination of these diverse activities requires a structural picture of the inner kinetochore.

Both Scc2/4 recruitment and stable microtubule binding at later stages of the cell cycle depend on the conserved Ctf3/CENP-I complex (Ctf3c) (Hara et al., 2018; Lang et al., 2018; Natsume et al., 2013). The Ctf3c namesake, Ctf3, is a 733 amino-acid residue protein predicted to consist almost entirely of HEAT repeats (Huntingin, elongation factor 3 (EF3), protein phosphatase 2A (PP2A), and yeast kinase TOR1 repeats) (Basilico et al., 2014). Ctf3 associates with a coiled-coil dimer of Mcm16 and Mcm22 (Mcm16/22). The vertebrate Ctf3/CENP-I complex also includes the CENP-M protein, which has no known counterpart in yeast (Basilico et al., 2014). The Mcm16/22 heterodimer can be purified from recombinant sources, and the dimer is required for recovery of the full-length Ctf3 protein. Deletion of any one complex member perturbs recruitment of the other two to the kinetochore, and the intact complex binds and recruits the Cnn1-Wip1 dimer, which connects to microtubules through the four-protein Ndc80 complex (Pekgoz Altunkaya et al., 2016). In wild-type cells, kinetochore-localized Ctf3-GFP signal fluctuates throughout the cell cycle, appearing to peak in late anaphase before dropping as cells enter S phase (Pot et al., 2003).

A high-resolution Ctf3c structure, which would enable molecular interpretation of the functions listed above, has yet to be reported. Three-dimensional reconstructions of negatively-stained human CENP-I assemblies yielded an envelope for the complex consistent with the predicted HEAT repeat architecture for Ctf3/CENP-I, but the overall resolution was insufficient to support modeling beyond placement of putative domains (Basilico et al., 2014). Recent X-ray crystal structures showed the HEAT repeat architecture of an N-terminal fragment of fungal Ctf3 (Ctf3-N) bound with C-terminal fragments of Mcm16 and Mcm22 (Mcm16/22-C), providing a model for this part of the Ctf3c (Hu et al., 2019). Our previous cryo-EM reconstruction of the yeast Ctf19c enabled modeling of most of the ordered polypeptide chains, but the Ctf3c appeared flexibly attached to the core of the complex, limiting the apparent resolution for this region of the map and model (Hinshaw and Harrison, 2019). Consequently, we could not confidently assign a sequence register to the Ctf3 HEAT repeats we observed or to the Mcm16/22 coiled-coil. We now present a cryo-EM structure of the yeast Ctf3c, which resolves the overall organization of the complex, shows how Ctf3 binds Mcm16/22, and clarifies Ctf3c recruitment to the larger Ctf19c.

RESULTS

We purified the Ctf3c and determined its structure by single-particle cryo-EM (see Methods). Initial images yielded two-dimensional averages, derived from a small hand-picked particle set, that matched projections of the density assigned to the Ctf3c in the full Ctf19c reconstruction (EMDB-0523). Additional blurry density was visible at the tips of several average particle images, likely indicating the presence of a domain (probably Ctf3-N and associated Mcm16/22-C, see below) without a fixed orientation relative to the well-resolved features. We then collected a large dataset. Particles recovered from these images yielded two-dimensional averages similar to those derived from screening images (Figure 1C). Particles contributing to two-dimensional averages that showed fine features were used for calculation of an initial three-dimensional reconstruction. Except for visual comparison of cryo-EM density, we did not include structural information from the previous Ctf19c reconstruction at any stage of model generation or

bioRxiv preprint doi: https://doi.org/10.1101/628149; this version posted May 5, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

refinement. The final map shows side-chain density throughout the particle, enabling modeling at near-atomic resolution (Figure 1 – figure supplement 1).

The overall architecture of the Ctf3c agrees well with the features previously assigned to it in our Ctf19c reconstruction (Hinshaw and Harrison, 2019) (Figure 2A). Clearly-resolved side-chain density throughout the map allowed us to identify the amino-acid sequence of the Ctf3 domain visualized here and the corresponding fragments of Mcm16/22 (Figure 2 – figure supplement 1-2, Table 2). We modeled five Ctf3-C HEAT repeats (Figure 2B), which have the overall fold of canonical HEAT arrays (Perry and Kleckner, 2003). Nearly 120 amino acid residues interrupt an imperfect sixth repeat at the C-terminal end of the modeled peptide chain, making a helical knob. As we saw in the previous Ctf3c model, an Mcm16/22 coiled-coil occupies the cleft formed by the concave surface of the HEAT array. The two main components of the structure, Ctf3-C and Mcm16/22-N, have opposite polarities; the central parts of Mcm16 and Mcm22 form a parallel coiled-coil that begins near the C-terminal end of the modeled Ctf3 polypeptide. The N-terminal parts of Mcm16 and Mcm22, form a set of four interdigitated helices that pack onto the convex side of the C-terminal part of the HEAT array and contact the helical knob.

Sequence conservation analysis suggested structural similarity between the N- and C-terminal parts of human CENP-I (Basilico et al., 2014). Indeed, Ctf3/CENP-I has a tandem HEAT repeat architecture, the N-terminal part of which was resolved crystallographically (PDB 3Z07) (Hu et al., 2019), and the C-terminal part we have resolved here. Our cryo-EM densities for Ctf3 (EMDB-0523 and here) imply that the two sets of Ctf3 HEAT repeats do not have a fixed relative orientation. Nevertheless, only 7 and 12 amino acid residues for Mcm16 and Mcm22 (respectively) connect the ordered regions that coordinate the two Ctf3 globular domains (Hu et al., 2019) (Figure 2C). The final model shows that the structure determined here (Ctf3-C-Mcm16/22-N) is essentially the precise complement of the Ctf3-N-Mcm16/22-C crystal structure (Hu et al., 2019). Therefore, the Mcm16/22 dimer extends without a major interruption along the Ctf3 HEAT repeats, forming two tightly but flexibly linked structural modules: Ctf3-N-Mcm16/22-C and Ctf3-C-Mcm16/22-N.

Ctf3c recruitment to the kinetochore depends on interactions with other Ctf19c proteins (Measday et al., 2002). We previously found that Ctf3c recruitment depends on amino acid residues 1-95 of Mcm21 (Mcm21-N), a region of the peptide chain not visible in published crystal structures (Schmitzberger and Harrison, 2012; Schmitzberger et al., 2017). The high-resolution Ctf3c map we report here allowed us to identify features in the previous Ctf19c cryo-EM map that are not visible in the current model and therefore must correspond to Ctf19c proteins not included in the current reconstruction. First, a small helical fragment packs into the cleft between the convex surface of the Ctf3 HEAT array and the Mcm16/22 N-terminal helices in the Ctf19c map (Figure 3A). This fragment likely accommodates a presumed short helical motif in Mcm21-N (Mcm21-10-34) (Garcia-Rodriguez et al., 2019; Hinshaw and Harrison, 2019). Second, the Ctf3c map lacks density for an extended peptide that snakes along the solvent-exposed surface of the Mcm16/22 N-terminal helices. This extended peptide emanates from the first resolved Mcm21 residue (at its N-terminus; Mcm21-Ser153) and likely connects through a loop not visualized in the Ctf19c map to the short helix mentioned above.

The new Ctf3c map also enabled unambiguous identification the Ctf3 residues that contact Iml3 and likely contribute to Ctf3c kinetochore recruitment. Docking the new Ctf3c model determined at high resolution into the Ctf19c map shows that density for bulky Ctf3 residues can be accounted for by the new model (Figure 3B). Continuous density connecting

bioRxiv preprint doi: https://doi.org/10.1101/628149; this version posted May 5, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Ctf3 and Iml3 accommodates Ctf3-Arg366, which packs onto Iml3-Trp112, likely enabling hydrophobic contact between the tryptophan indole and the aliphatic part of the arginine side chain. The neighboring Ctf3-Lys365 is well-positioned to make a salt bridge with Iml3-Glu82 in the full Ctf19c. To test whether this interface, like the one we previously reported between Ctf3 and Mcm21-N, contributes to Ctf3c kinetochore recruitment, we mutated three protruding residues to create the *ctf3-SDD* allele (Ctf3-W362S,K365D,R366D). Yeast cells expressing Mcm22-GFP (to mark Ctf3c localization), Spc110-mCherry (to mark the spindle pole body/kinetochore), and Ctf3-WT display Mcm22-GFP localization that matches that of Ctf3-GFP (Hinshaw and Harrison, 2019). Cells lacking Iml3 (*iml3Δ*) or expressing the *ctf3-SDD* allele showed defective Mcm22-GFP localization, a phenotype that was less severe in *ctf3-SDD* cells than in *iml3Δ* cells but was reproducible in three independently-derived *ctf3-SDD* strains (Figure 3C, Figure 3 – figure supplement 1). Therefore, the Iml3-Ctf3 interface we observe in cryo-EM reconstructions of the full Ctf19c recapitulates the one used by cells to recruit the Ctf3c to the kinetochore.

DISCUSSION

There are three published Ctf3c/CENP-I complex models (Basilico et al., 2014; Hinshaw and Harrison, 2019; Hu et al., 2019). The volume derived from a negative-stain electron microscopy reconstruction of the four-protein human CENP-I complex (CENP-H/I/K/M) could likely accommodate the model we present here, along with the N-terminal part of the Ctf3/CENP-I and human CENP-M. Our recent Ctf3c structure, which we determined in the context of the full Ctf19c, provides information about the mechanism of Ctf3c recruitment but did not allow us to confidently assign amino acid sequence to the density. Finally, a recent crystal structure of Ctf3 from *Chaetomium thermophilum* in complex with Mcm16/22 from a related species (*Thielavia terrestri*) provides a high-resolution view of the Ctf3/CENP-I N-terminal domain and of the Mcm16/22 C-terminal regions but no information about the domains we report here.

Incorporation of the updated Ctf3c model into our previous model of the Ctf19c yields a nearly complete structural description of this 13-protein kinetochore anchor (Figure 4A). Newly-resolved interfaces between Ctf3 and the Ctf19c proteins Iml3 and Mcm21 enable direct visualization and, consequently, modulation of Ctf3c recruitment. A more precise understanding of Ctf3c recruitment will need to account for dynamic localization of Ctf3 and its binding partner Cnn1 during the cell cycle. Post-translational modification of Ctf3c proteins, Cnn1, Mcm21, Mif2/CENP-C, or some combination of these factors may explain this behavior.

A major open question, brought into focus by the current cryo-EM reconstruction of the Ctf3c, is whether and in what way the N-terminal domain of Ctf3 contacts the Cse4 nucleosome. Although Ctf3-N was not clearly visible in our cryo-EM reconstruction of the Ctf19c, we did see poorly-resolved density at the N-terminal tip of the Ctf3 HEAT repeat array, which we assigned to the Cnn1-Wip1 dimer. Reanalysis of this density map suggests Ctf3-N also contributes to the observed density, an arrangement that agrees with the identification of contact between Ctf3c and Cnn1-Wip1 by crosslinking-coupled mass spectrometry (Pekgoz Altunkaya et al., 2016). If so, then the entire Ctf3N-Cnn1-Wip1 module must swing outward (away from the Ctf19c central cavity) to make way for the Cse4 nucleosome, flanking DNA, and associated CBF3 complex factors (Figure 4B). This outward displacement would position Cnn1-Wip1 to contact and stabilize nucleosome-flaking DNA, a suggestion consistent with the observation that the vertebrate CENP-T/W heterodimer binds and bends DNA (Nishino et al., 2012; Takeuchi et al.,

bioRxiv preprint doi: https://doi.org/10.1101/628149; this version posted May 5, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

2014). How this reorganization works and how cell cycle-dependent enzymatic activities regulate it are matters of ongoing investigation.

METHODS

Protein Purification and Cryo-EM Sample Preparation

S. cerevisiae Ctf3c was expressed and purified as described previously (Hinshaw and Harrison, 2019; Hinshaw et al., 2017). The three protein-coding genes were expressed as a single transcript with three ribosome binding sites. After recombinant protein expression in *E. coli*, cell pellets were frozen and stored at -80 °C in lysis buffer until use. Purified protein samples (~800 µg from 24 L of bacterial growth) were supplemented with 5% glycerol by volume and stored at ~8 mg/mL at -80 °C. For cryo-EM grid preparation, an aliquot of frozen Ctf3c stock was thawed and diluted to ~1.8 mg/mL in gel filtration buffer (150 mM NaCl, 20 mM tris-HCl pH 8.5, 1 mM TCEP). The diluted sample was applied to C-flat holey carbon grids (Electron Microscopy Sciences, CF-1.2/1.3-3C for screening, CF-2/2-3C for data collection) and plunge-frozen in liquid ethane on a Cryo-Plunge 3 instrument (Gatan). Grids were prepared at ~90% relative humidity. 3.5 µL of sample was applied to negatively-discharged grids, which were subsequently blotted for 4 s from both sides with minimal delay.

Cryo-EM Data Collection

A screening dataset was collected manually on an F20 instrument (FEI) using the UCSF Image4 semi-automated data collection software package (Li et al., 2015). A single large dataset was collected on a Titan Krios G3i (FEI). The instrument was outfitted with a pre-camera energy filter (Gatan Image Filter) and operated at 300 kV. Images were collected in nanoprobe mode at spot size 3 with an illuminated area of ~1 µm at the sample level. Images were collected on a K3 camera (Gatan) in counting mode at a nominal magnification of 105 kx with a pixel size of 0.85 $Å^2$ (Table 1). Dose-fractionated movies were collected (3 s per movie, 50 frames, 17 electrons per $Å^2$ per second, and 1.0 electrons per $Å^2$ per frame). Five movies were collected for each hole on the grid, and nine holes were imaged at each stage position, yielding 45 movies per round of stage movement. We used SerialEM v3.7 with on-the-fly beam tilt correction of image shift-induced coma to collect the high-resolution dataset (Mastronarde, 2005). Computational beam tilt correction during data processing further improved the experimental density (Figure 1 – figure supplement 1).

Cryo-EM Data Processing

2-dimensional class average particle images were first generated from ~1,500 hand-picked particles from the screening dataset (F20). Class averages judged to represent distinct views of the Ctf3c were then used as references for automated particle picking against the full screening dataset. The resulting particles were further classified, and 11 of the best class averages were used as particle picking references for the high-resolution dataset.

For high-resolution data processing, sample displacements and defocus estimates were calculated during data collection using Motioncor2 v1.1.0 (5-by-5 patch correction) and CTFFIND4 v4.1.8 (Rohou and Grigorieff, 2015; Zheng et al., 2017). All data processing steps were carried out in RELION-3.0 unless otherwise noted (Zivanov et al., 2018) (Figure 1 – figure supplement 1). Corrected movies with maximum resolution estimations (generated by CTFFIND4) greater than 4.75 Å were discarded. Particles were picked using the class average

bioRxiv preprint doi: https://doi.org/10.1101/628149; this version posted May 5, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

images described above, and the resulting particle stack was subjected to two-dimensional classification. Following previously-described methods for reconstructing small particles, we performed a second round of two-dimensional classification to select and enrich for underrepresented projections (projections down the long aspect of the particle) (Herzik et al., 2019). Subsequent processing steps are described in Figure 1 – figure supplement 1. We noticed large particle movements with apparently poor local correlation in our movie files, leading to major improvements in map quality upon per-particle motion correction (particle polishing steps).

Model Building, Refinement, and Analysis

We used as a starting point for model building the original Ctf3c structure (PDB 6NUW), which was predominantly composed of poly-alanine chains. Manual model adjustments were done in Coot v0.8.8 (Emsley et al., 2010). We used secondary structure predictions generated by Phyre2 to guide modeling (Kelley et al., 2015). Helical fragments corresponding to HEAT repeats 1-5 fit the density well, and the higher-resolution map enabled *de novo* modeling of the Ctf3 helical knob. The Mcm16/22 coiled-coils were originally placed in the reverse orientation. Inversion of the corresponding sequence resulted in a good fit of the density throughout the chains, including the N-terminal helices external to the Ctf3 cavity. The rebuilt and manually adjusted model was subjected to real-space refinement in Phenix v1.13 with global minimization, atomic displacement parameters (ADP), and local grid searches (Afonine et al., 2018). We used the phenix.secondary_structure_restrains program to generate restraints for modeled alpha-helices, which we then edited manually. In later rounds of coordinate refinement, we also used simulated annealing.

The Ctf3c model was fit into the Ctf19c density map in Chimera using the Fit in Map command (Pettersen et al., 2004). A homology model of *S. cerevisiae* Ctf3-N was generated using Swiss Model with PDB 3Z07 as a template (Hu et al., 2019; Waterhouse et al., 2018). A crystal structure of the histone fold domain of chicken CENP-T/W (PDB 3B0C) was docked into the density adjacent to the Ctf3-N model (Nishino et al., 2012). The ResMap program, implemented in Relion-3.0, was used to determine local resolution (Figure2 – figure supplement 1A) (Kucukelbir et al., 2014).

Yeast strain construction, imaging, and Western blot

Yeast strains were created by integration of PCR products coding for 3-FLAG-tagged Ctf3 or Ctf3-SDD according to standard methods (Longtine et al., 1998). For live-cell imaging (Figure 3C), three independent clones each were examined for the *CTF3*-3FLAG and *ctf3-SDD*-3-FLAG genotypes. Three cultures were imaged for each genotype. Fluorescence imaging and image processing were carried out exactly as described previously (Hinshaw and Harrison, 2019). Cells were grown in synthetic complete medium until mid-log phase and diluted into fresh medium ~2 hrs before imaging on a Nikon Ti2 fluorescence microscope with Perfect Focus System and a Nikon Plan Apo 60 x 1.4 NA oil-immersion objective lens. The microscope was outfitted with a Tokai Hit stage-top incubator and a Hamamatsu Flash 4.0 V2 +sCMOS camera. Imaging and image segmentation protocols and settings were identical to those we described previously (Hinshaw and Harrison, 2019; Tinevez et al., 2017). At least two stage positions and data from at least 1,000 mCherry foci (fiducials used for the Mcm22-GFP focus calling routine) were averaged for each replicate. Western blotting (Figure 3 – figure supplement 1) was performed as described previously (Hinshaw and Harrison, 2019).

bioRxiv preprint doi: https://doi.org/10.1101/628149; this version posted May 5, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

ACKNOWLEDGEMENTS

We thank the staff at the Harvard Cryo-Electron Microscopy Center for Structural Biology for help with high-resolution cryo-EM data collection, particularly Richard Walsh, Sarah Sterling, and Zongli Li. We thank Shaun Rawson for real-time movie correction and for helpful comments during data processing. We thank SBGrid for computational support, particularly Mick Timony and Justin O'Connor. We thank Jennifer Waters and the staff at the Nikon Imaging Center at Harvard Medical School for light microscopy support. We thank Simon Jenni for guidance in implementing the HEALPix library and for helpful discussions. S.C.H. is an Investigator of the Howard Hughes Medical Institute. S.M.H. is an HHMI fellow of the Helen Hay Whitney Foundation.

DATA AVAILABILITY

The final Ctf3c cryo-EM map and the soft mask used for three-dimensional refinements of fully unbinned particles have been deposited with accession code EMD-20200. The refined Ctf3c model has been deposited with accession code 6OUA.

bioRxiv preprint doi: https://doi.org/10.1101/628149; this version posted May 5, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

FIGURES

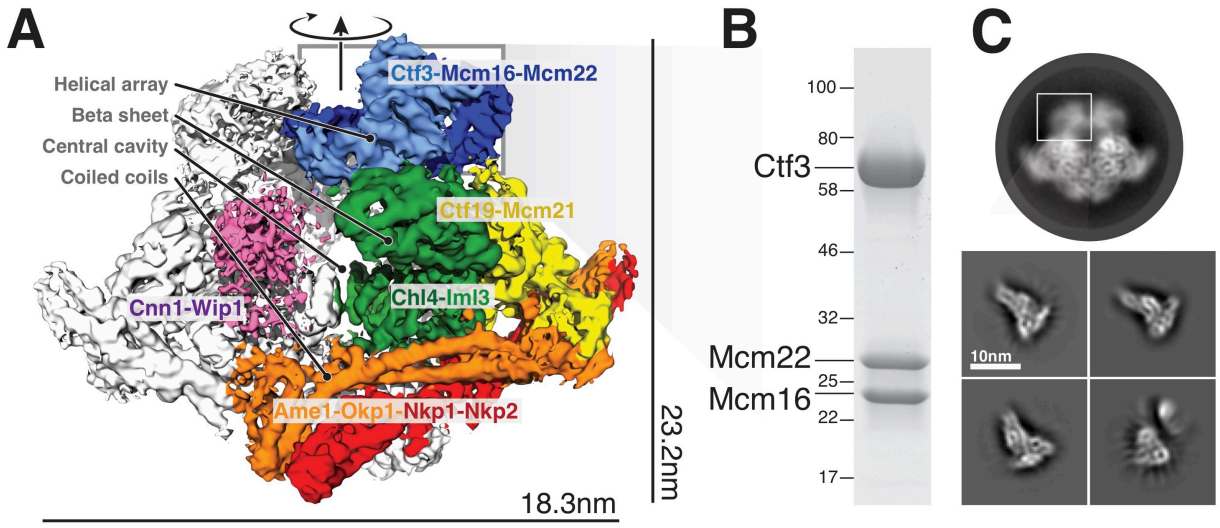

Figure 1 – Overview of the Ctf19c and position of the Ctf3c within it
(A) Ctf19c density map colored according to constituent subcomplexes (Hinshaw and Harrison, 2019). **(B)** The recombinant Ctf3c sample used for cryo-EM (size markers – kDa). **(C)** Two-dimensional class averages showing the full Ctf19c (circle) and the Ctf3c (squares). The white box indicates the fuzzy Ctf19c density assigned to the Ctf3c.

bioRxiv preprint doi: https://doi.org/10.1101/628149; this version posted May 5, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

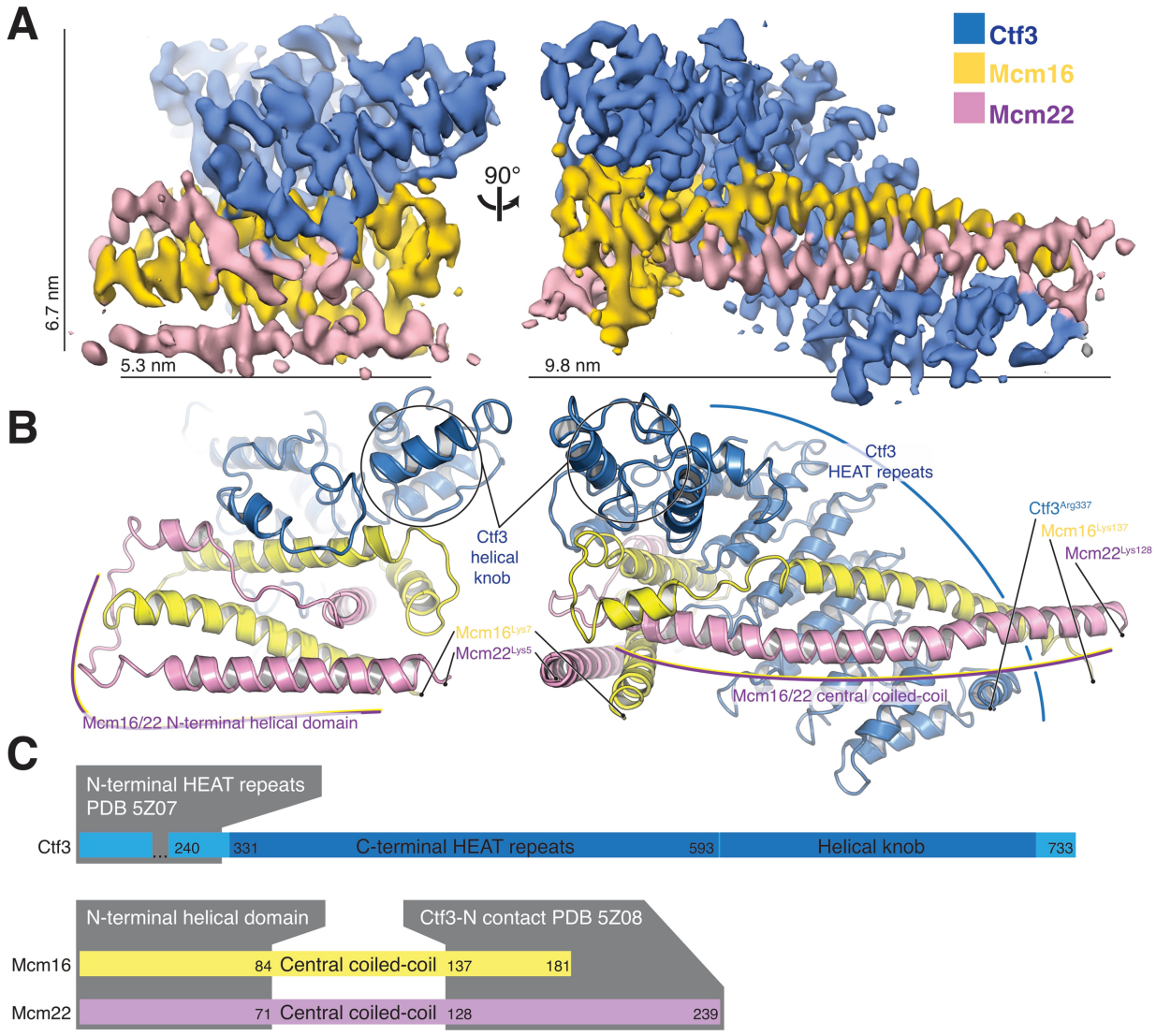

Figure 2 – Overview of the Ctf3c structure determined by cryo-EM
(A) Cryo-EM density map shown in in two orientations. The density is colored according to the atomic model of the complex. **(B)** Atomic model of the Ctf3c colored according to panel (A) and shown in the same orientations. Prominent Ctf3c features are labeled. **(C)** Domain diagram depicting Ctf3c components. PDB numbers refer to the published crystal structures of Ctf3-N and Mcm16/22-C (Hu et al., 2019).

bioRxiv preprint doi: https://doi.org/10.1101/628149; this version posted May 5, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

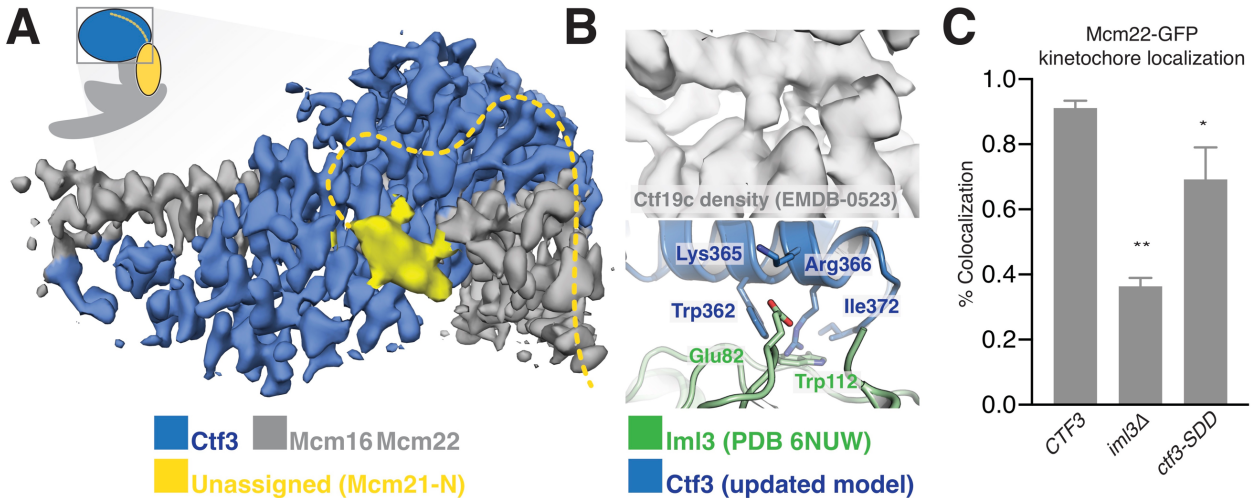

Figure 3 – Interactions between the Ctf3c and the full Ctf19c
(A) Composite cryo-EM density map showing extra density, only visible in the original Ctf19c map, which likely belongs to Mcm21-N. The image was created by combining the high-resolution Ctf3c map (blue and gray) with the displayed yellow density from the Ctf19c map (EMDB-0523). The inset shows the position of the map in the context of the full Ctf19c model. **(B)** Cryo-EM density showing the interaction between Ctf3 and Iml3. The top panel shows density determined for the interface in the context of the full Ctf19c. The bottom panel shows a composite model; Iml3 coordinates refined into the Ctf19c cryo-EM map are green, and Ctf3 coordinates refined into the current density map are blue. **(C)** The indicated *S. cerevisiae* strains (*CTF3 – CTF3-3FLAG*: SMH690, SMH691, SMH692; *iml3Δ*: SMH91; *ctf3-SDD – ctf3-SDD-3FLAG*: SMH693, SMH694, SMH695) expressing Mcm22-GFP and Spc110-mCherry were imaged during asynchronous division. Coincidence frequencies of Mcm22-GFP foci with Spc110-mCherry foci are plotted (error bars – +/- SD for frequency measurements from three independent cultures; * – $p < .05$; ** – $p < .01$, Student's t-test, two tails, unequal variance).

bioRxiv preprint doi: https://doi.org/10.1101/628149; this version posted May 5, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

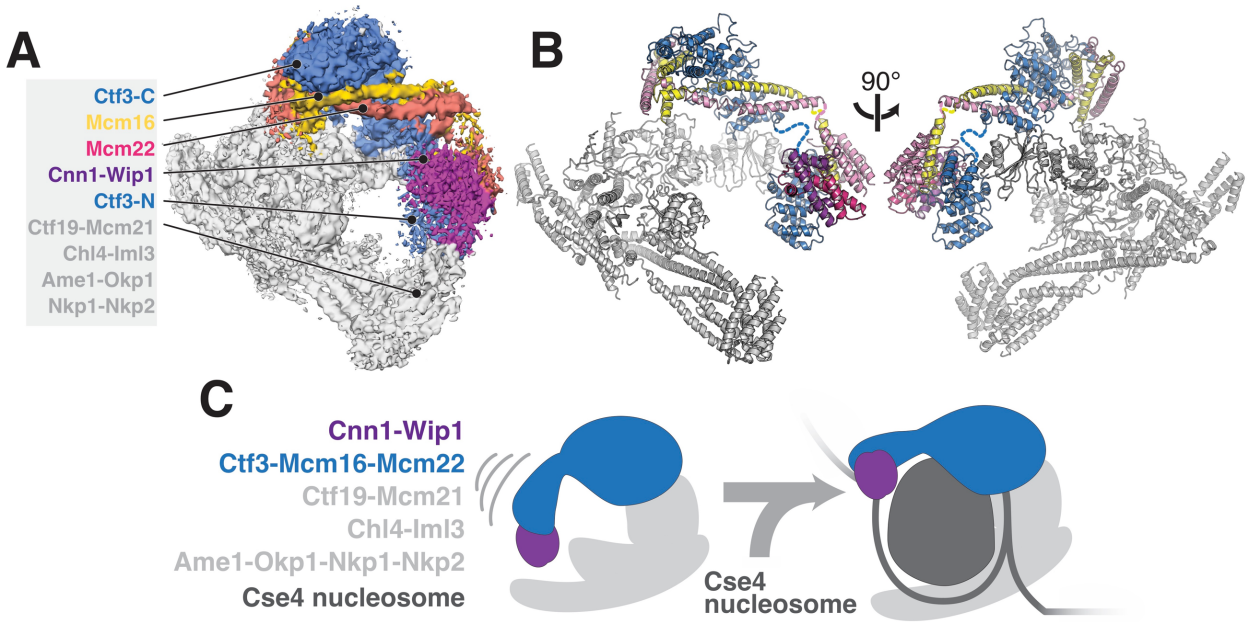

Figure 4 – Updated model of Ctf19c architecture and interactions
(A-B) Cryo-EM density map (A) and corresponding composite atomic model (B) showing the 13-protein Ctf19c in its monomeric form (EMDB-0523). The Ctf3 N-terminal HEAT repeat domain and adjacent Cnn1-Wip1 heterodimer, both of which occupy updated positions in the overall model, are labeled. The Ctf3c components visualized in the current study are also labeled. The Ctf3-N-Mcm16/22-C module (PDB 5Z08), with Ctf3-N substituted by a homology model, was positioned to minimize the gap between ordered parts of Mcm16/22 (reported here and previously) (Hu et al., 2019). The interaction between Ctf3-N and Cnn1-Wip1 has not been visualized experimentally, so the histone-fold extension of Cnn1 required for Ctf3 contact (Pekgoz Altunkaya et al., 2016) was used to position Cnn1-Wip1 with respect to Ctf3-N. **(C)** Proposed model for nucleosome accommodation by the Ctf19c in its monomeric form. Ctf3-N, along with Cnn1-Wip1, flex up and outward to make space for a Cse4/CENP-A nucleosome. Current data does not distinguish between one or two copies of the Ctf19c decorating the Cse4/CENP-A particle. A single Ctf19c was therefore drawn for clarity.

bioRxiv preprint doi: https://doi.org/10.1101/628149; this version posted May 5, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

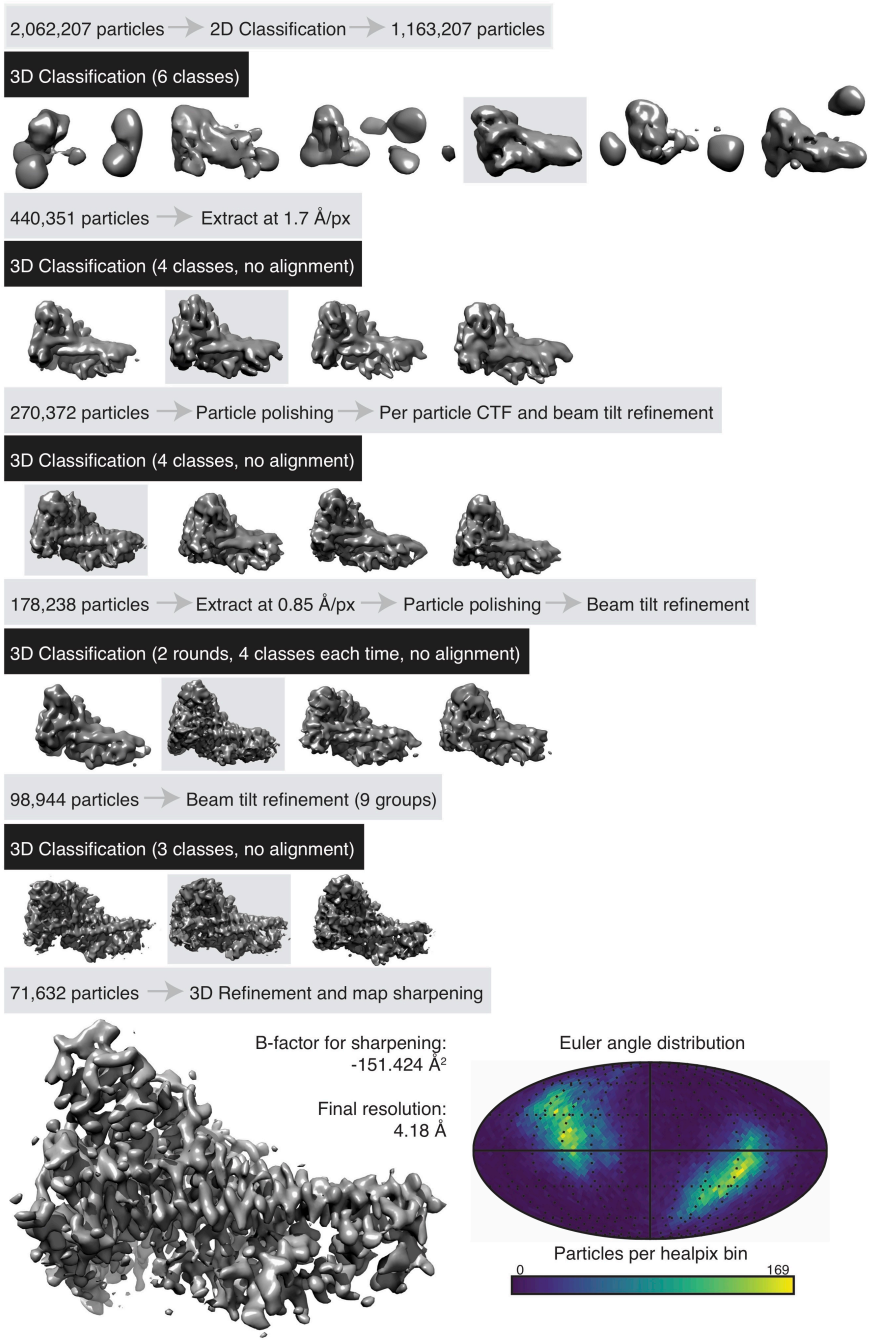

Figure 1 – figure supplement 1 – Cryo-EM data processing
The data processing procedure for Ctf3c structure determination is shown. Three-dimensional refinement using a soft mask around the particle volume was carried out before each three-dimensional classification step, particle polishing step, and round of particle parameter refinement (CTF or beam tilt parameters). For the final round of beam tilt refinement, data was segmented into groups according to the relative hole position from where the particle data originated. The Euler angle plot (bottom right) shows the Euler angular distribution according to real-space HEALPix binning (Gorski et al., 2005).

bioRxiv preprint doi: https://doi.org/10.1101/628149; this version posted May 5, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

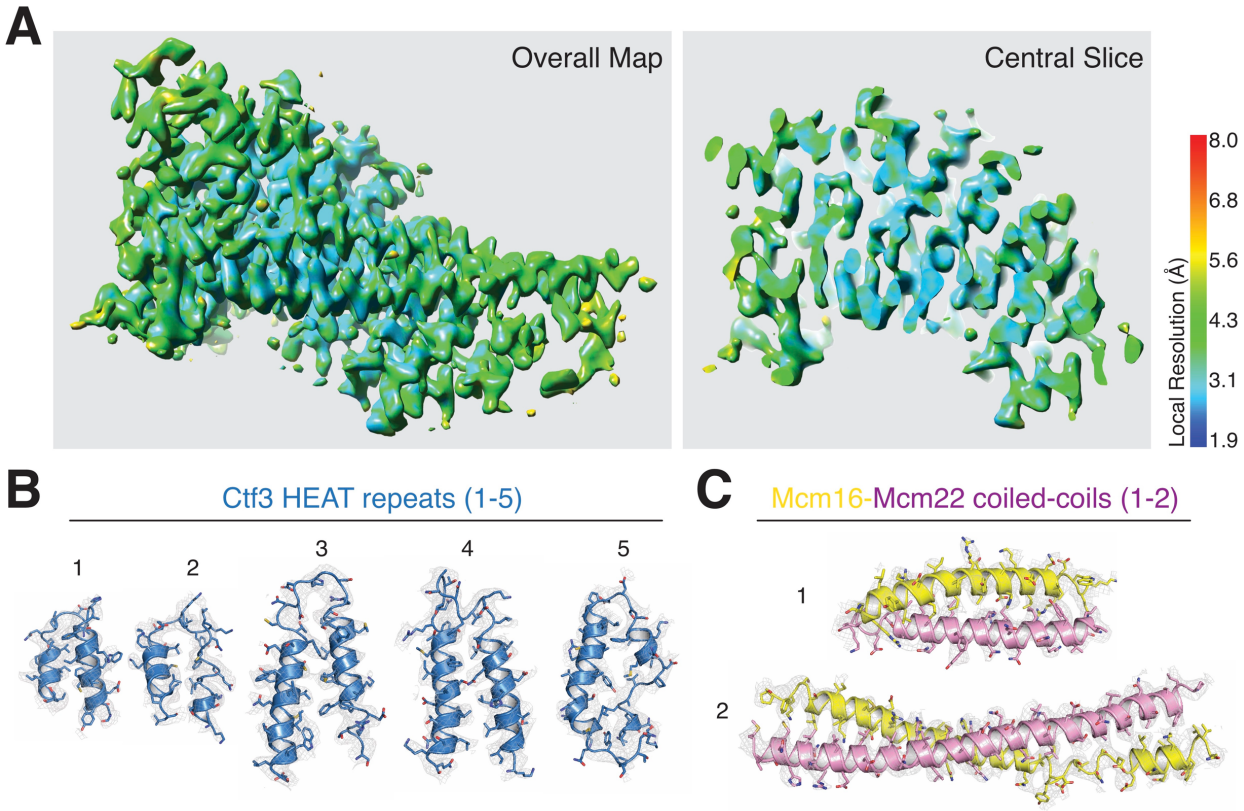

Figure 2 – figure supplement 1 – Analysis of cryo-EM density map quality and model fitting **(A)** Local resolution estimation for the Ctf3c density map (pValue 0.05, minRes 1.90 Å, maxRes 8.00 Å, step 1.0 Å). **(B)** Cryo-EM density and model fitting for Ctf3-C HEAT repeats 1-5 (designated from N- to C-terminus of the modeled polypeptide chain; 1 – 354-386, 2 – 391-423, 3 – 430-475, 4 – 477-525, 5 – 531-576). **(C)** Cryo-EM density and model fitting for the Mcm16/22-N parallel coiled-coils (1 – Mcm16-7-38 and Mcm22-7-38, 2 – Mcm16-80-136 and Mcm22-69-130).

bioRxiv preprint doi: https://doi.org/10.1101/628149; this version posted May 5, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

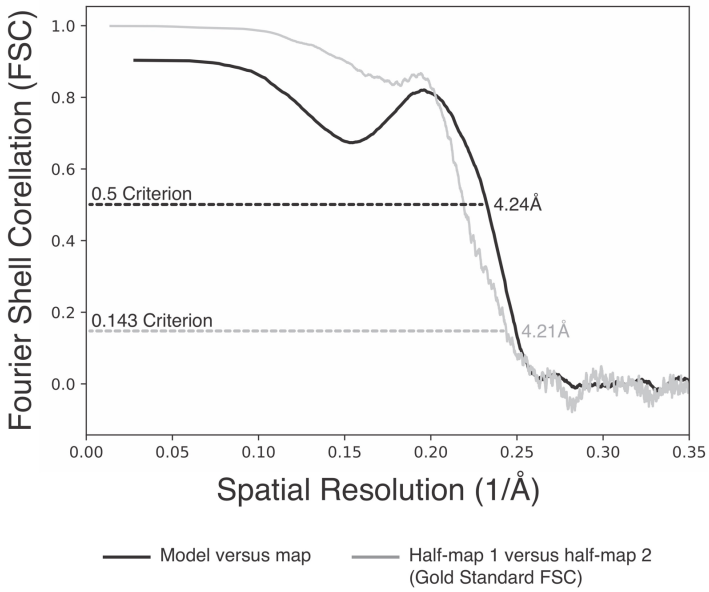

Figure 2 – figure supplement 2 – Fourier shell correlation curves describing the Ctf3c density map (half-map to half-map) and map-to-model correlations.

bioRxiv preprint doi: https://doi.org/10.1101/628149; this version posted May 5, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

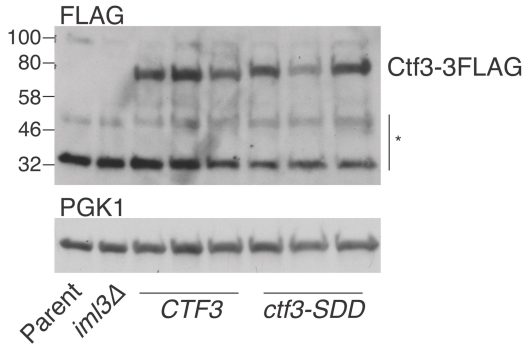

Figure 3 – figure supplement 1 – Expression levels of Ctf3 WT and mutant proteins
Total cellular protein extract was analyzed by Western blot for the presence of Ctf3-3FLAG in the indicated strains (* – non-specific bands; Parent: SMH81; *iml3Δ*: SMH91; *CTF3 – CTF3-3FLAG*: SMH690, SMH691, SMH692; *ctf3-SDD – ctf3-SDD-3FLAG*: SMH693, SMH694, SMH695).

bioRxiv preprint doi: https://doi.org/10.1101/628149; this version posted May 5, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Table 1 – Cryo-EM data collection, refinement, and validation

	CTF3C (EMDB-20200) (PDB 6OUA)
Data collection and processing	
Magnification	105,000
Voltage (kV)	300
Electron exposure (e–/Å^2)	51
Defocus range (μm)	-1.5 to -3.5
Pixel size (Å)	.85
Symmetry imposed	None
Initial particle images (no.)	2,062,207
Final particle images (no.)	71,632
Map resolution (Å) FSC threshold	4.18 (0.143)
Refinement	
Initial model used (PDB code)	6NUW
Model resolution (Å) FSC threshold	4.24 (0.5)
Model resolution range (Å)	40-4.18
Map sharpening *B* factor (Å^2)	-151.424
Model composition Non-hydrogen atoms Protein residues Ligands	 5208 637 0
B factors (Å^2)	Min: 36.02 Max: 141.08 Mean: 72.17
R.m.s. deviations Bond lengths (Å) Bond angles (°)	 0.008 1.351
Validation MolProbity score Clashscore Poor rotamers (%)	 2.37 7.09 2.35
Ramachandran plot Favored (%) Allowed (%) Disallowed (%)	 82.25 99.68 0.32

bioRxiv preprint doi: https://doi.org/10.1101/628149; this version posted May 5, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Table 2 – Model information

PROTEIN	CHAIN ID	HOMOLOGS (HUMAN; POMBE)	CHAIN LENGTH (TOTAL; MODELED)	TEMPLATE	PROCEDURE	MODELLED RESIDUES
Mcm16	H	CENP-H; Fta3	181; 130	--	Build de novo	7-137
Ctf3	I	CENP-I; Mis6	733; 381	6NUW	Modify, build de novo	337-718
Mcm22	K	CENP-K; Sim4	239; 124	--	Build de novo	5-128

bioRxiv preprint doi: https://doi.org/10.1101/628149; this version posted May 5, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Table 3 – Yeast strains used in this study

STRAIN NUMBER	GENOTYPE	REFERENCE
SMH81	*MATa MCM22-GFP::HisMX SPC110-mCherry::hphMX*	(Huh et al., 2003)
SMH91	*MATa MCM22-GFP::HisMX SPC110-mCherry::hphMX iml3Δ::KanMX*	This study
SMH690	*MATa MCM22-GFP::HisMX SPC110-mCherry::hphMX CTF3-3FLAG::KanMX*	This study
SMH691	*MATa MCM22-GFP::HisMX SPC110-mCherry::hphMX CTF3-3FLAG::KanMX*	This study
SMH692	*MATa MCM22-GFP::HisMX SPC110-mCherry::hphMX CTF3-3FLAG::KanMX*	This study
SMH693	*MATa MCM22-GFP::HisMX SPC110-mCherry::hphMX ctf3-SDD-3FLAG::KanMX*	This study
SMH694	*MATa MCM22-GFP::HisMX SPC110-mCherry::hphMX ctf3-SDD-3FLAG::KanMX*	This study
SMH695	*MATa MCM22-GFP::HisMX SPC110-mCherry::hphMX ctf3-SDD-3FLAG::KanMX*	This study

bioRxiv preprint doi: https://doi.org/10.1101/628149; this version posted May 5, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Table 4 – Plasmids used in this study

PLASMID NUMBER	GENOTYPE	REFERENCE
pSMH145	pLIC-Tra His6-TEV-Ctf3; His6-TEV-Mcm16; His6-TEV-Mcm22	(Hinshaw et al., 2017)
pSMH1269	pFA6a-*CTF3*-6Gly-3FLAG-KanMX6 (BsiWI/SalI)	This study
pSMH1577	pFA6a-*ctf3-SDD*-6Gly-3FLAG-KanMX6 (BsiWI/SalI)	This study

bioRxiv preprint doi: https://doi.org/10.1101/628149; this version posted May 5, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

REFERENCES

Afonine, P.V., Poon, B.K., Read, R.J., Sobolev, O.V., Terwilliger, T.C., Urzhumtsev, A., and Adams, P.D. (2018). Real-space refinement in PHENIX for cryo-EM and crystallography. Acta Crystallogr D Struct Biol *74*, 531-544.

Basilico, F., Maffini, S., Weir, J.R., Prumbaum, D., Rojas, A.M., Zimniak, T., De Antoni, A., Jeganathan, S., Voss, B., van Gerwen, S.*, et al.* (2014). The pseudo GTPase CENP-M drives human kinetochore assembly. Elife *3*, e02978.

Bock, L.J., Pagliuca, C., Kobayashi, N., Grove, R.A., Oku, Y., Shrestha, K., Alfieri, C., Golfieri, C., Oldani, A., Dal Maschio, M.*, et al.* (2012). Cnn1 inhibits the interactions between the KMN complexes of the yeast kinetochore. Nat Cell Biol *14*, 614-624.

Emsley, P., Lohkamp, B., Scott, W.G., and Cowtan, K. (2010). Features and development of Coot. Acta crystallographica Section D, Biological crystallography *66*, 486-501.

Fernius, J., Nerusheva, O.O., Galander, S., Alves Fde, L., Rappsilber, J., and Marston, A.L. (2013). Cohesin-dependent association of scc2/4 with the centromere initiates pericentromeric cohesion establishment. Curr Biol *23*, 599-606.

Garcia-Rodriguez, L.J., Kasciukovic, T., Denninger, V., and Tanaka, T.U. (2019). Aurora B-INCENP Localization at Centromeres/Inner Kinetochores Is Required for Chromosome Bi-orientation in Budding Yeast. Curr Biol.

Gorski, K.M., Hivon, E., Banday, A.J., Wandelt, B.D., and Hansen, F.K. (2005). HEALPix – a framework for high resolution discretization, and fast analysis of data distributed on the sphere. The Astrophysical Journal *622*, 759-771.

Halwachs, J.F., Singh, S., Potocnjak, M., Hagemann, G., Solis, V., Woike, S., Steger, M.G., Feuillet, J.A., and Herzog, F. (2018). The COMA complex is required for positioning Ipl1 activity proximal to
Cse4 nucleosomes in budding yeast. biorxiv.

Hara, M., Ariyoshi, M., Okumura, E.I., Hori, T., and Fukagawa, T. (2018). Multiple phosphorylations control recruitment of the KMN network onto kinetochores. Nat Cell Biol *20*, 1378-1388.

Herzik, M.A., Jr., Wu, M., and Lander, G.C. (2019). High-resolution structure determination of sub-100 kDa complexes using conventional cryo-EM. Nat Commun *10*, 1032.

Hinshaw, S.M., and Harrison, S.C. (2017). Kinetochore Function from the Bottom Up. Trends Cell Biol *28*, 22-33.

Hinshaw, S.M., and Harrison, S.C. (2019). The structure of the Ctf19c/CCAN from budding yeast. Elife *8*.

Hinshaw, S.M., Makrantoni, V., Harrison, S.C., and Marston, A.L. (2017). The kinetochore receptor for the cohesin loading complex. Cell *171*, 72-84 e13.

Hu, L., Huang, H., Hei, M., Yang, Y., Li, S., Liu, Y., Dou, Z., Wu, M., Li, J., Wang, G.Z.*, et al.* (2019). Structural analysis of fungal CENP-H/I/K homologs reveals a conserved assembly mechanism underlying proper chromosome alignment. Nucleic Acids Res *47*, 468-479.

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

Kelley, L.A., Mezulis, S., Yates, C.M., Wass, M.N., and Sternberg, M.J. (2015). The Phyre2 web portal for protein modeling, prediction and analysis. Nat Protoc *10*, 845-858.

Kucukelbir, A., Sigworth, F.J., and Tagare, H.D. (2014). Quantifying the local resolution of cryo-EM density maps. Nature methods *11*, 63-65.

bioRxiv preprint doi: https://doi.org/10.1101/628149; this version posted May 5, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Lang, J., Barber, A., and Biggins, S. (2018). An assay for de novo kinetochore assembly reveals a key role for the CENP-T pathway in budding yeast. Elife *7*.
Li, X., Zheng, S., Agard, D.A., and Cheng, Y. (2015). Asynchronous data acquisition and on-the-fly analysis of dose fractionated cryoEM images by UCSFImage. J Struct Biol *192*, 174-178.
Longtine, M.S., McKenzie, A., 3rd, Demarini, D.J., Shah, N.G., Wach, A., Brachat, A., Philippsen, P., and Pringle, J.R. (1998). Additional modules for versatile and economical PCR-based gene deletion and modification in Saccharomyces cerevisiae. Yeast *14*, 953-961.
Mastronarde, D.N. (2005). Automated electron microscope tomography using robust prediction of specimen movements. J Struct Biol *152*, 36-51.
Measday, V., Hailey, D.W., Pot, I., Givan, S.A., Hyland, K.M., Cagney, G., Fields, S., Davis, T.N., and Hieter, P. (2002). Ctf3p, the Mis6 budding yeast homolog, interacts with Mcm22p and Mcm16p at the yeast outer kinetochore. Genes Dev *16*, 101-113.
Musacchio, A., and Desai, A. (2017). A Molecular View of Kinetochore Assembly and Function. Biology (Basel) *6*.
Natsume, T., Muller, C.A., Katou, Y., Retkute, R., Gierlinski, M., Araki, H., Blow, J.J., Shirahige, K., Nieduszynski, C.A., and Tanaka, T.U. (2013). Kinetochores coordinate pericentromeric cohesion and early DNA replication by Cdc7-Dbf4 kinase recruitment. Mol Cell *50*, 661-674.
Nishino, T., Takeuchi, K., Gascoigne, K.E., Suzuki, A., Hori, T., Oyama, T., Morikawa, K., Cheeseman, I.M., and Fukagawa, T. (2012). CENP-T-W-S-X forms a unique centromeric chromatin structure with a histone-like fold. Cell *148*, 487-501.
Pekgoz Altunkaya, G., Malvezzi, F., Demianova, Z., Zimniak, T., Litos, G., Weissmann, F., Mechtler, K., Herzog, F., and Westermann, S. (2016). CCAN Assembly Configures Composite Binding Interfaces to Promote Cross-Linking of Ndc80 Complexes at the Kinetochore. Curr Biol *26*, 2370-2378.
Perry, J., and Kleckner, N. (2003). The ATRs, ATMs, and TORs are giant HEAT repeat proteins. Cell *112*, 151-155.
Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., and Ferrin, T.E. (2004). UCSF Chimera--a visualization system for exploratory research and analysis. J Comput Chem *25*, 1605-1612.
Pot, I., Measday, V., Snydsman, B., Cagney, G., Fields, S., Davis, T.N., Muller, E.G., and Hieter, P. (2003). Chl4p and iml3p are two new members of the budding yeast outer kinetochore. Mol Biol Cell *14*, 460-476.
Rohou, A., and Grigorieff, N. (2015). CTFFIND4: Fast and accurate defocus estimation from electron micrographs. J Struct Biol *192*, 216-221.
Schmitzberger, F., and Harrison, S.C. (2012). RWD domain: a recurring module in kinetochore architecture shown by a Ctf19-Mcm21 complex structure. EMBO Rep *13*, 216-222.
Schmitzberger, F., Richter, M.M., Gordiyenko, Y., Robinson, C.V., Dadlez, M., and Westermann, S. (2017). Molecular basis for inner kinetochore configuration through RWD domain-peptide interactions. EMBO J *36*, 3458-3482.
Takeuchi, K., Nishino, T., Mayanagi, K., Horikoshi, N., Osakabe, A., Tachiwana, H., Hori, T., Kurumizaka, H., and Fukagawa, T. (2014). The centromeric nucleosome-like CENP-T-W-S-X complex induces positive supercoils into DNA. Nucleic Acids Res *42*, 1644-1655.
Tinevez, J.Y., Perry, N., Schindelin, J., Hoopes, G.M., Reynolds, G.D., Laplantine, E., Bednarek, S.Y., Shorte, S.L., and Eliceiri, K.W. (2017). TrackMate: An open and extensible platform for single-particle tracking. Methods *115*, 80-90.

bioRxiv preprint doi: https://doi.org/10.1101/628149; this version posted May 5, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Waterhouse, A., Bertoni, M., Bienert, S., Studer, G., Tauriello, G., Gumienny, R., Heer, F.T., de Beer, T.A.P., Rempfer, C., Bordoli, L.*, et al.* (2018). SWISS-MODEL: homology modelling of protein structures and complexes. Nucleic Acids Res *46*, W296-W303.

Zheng, S.Q., Palovcak, E., Armache, J.P., Verba, K.A., Cheng, Y., and Agard, D.A. (2017). MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. Nature methods *14*, 331-332.

Zivanov, J., Nakane, T., Forsberg, B.O., Kimanius, D., Hagen, W.J., Lindahl, E., and Scheres, S.H. (2018). New tools for automated high-resolution cryo-EM structure determination in RELION-3. Elife *7*.