

Structural analysis of fungal CENP-H/I/K homologs reveals a conserved assembly mechanism underlying proper chromosome alignment

Liqiao Hu^{1,†}, Hao Huang^{1,†}, Mohan Hei¹, Yang Yang², Sheng Li³, Yunshan Liu¹, Zhen Dou⁴, Mengying Wu¹, Jie Li⁵, Guang-zhong Wang⁵, Xuebiao Yao⁴, Hong Liu⁶, Xiaojing He^{1,*} and Wei Tian^{1,7,*}

¹Key Laboratory of Molecular Biophysics of the Ministry of Education, College of Life Science and Technology, Huazhong University of Science and Technology, Wuhan 430074, China, ²Department of Pharmacology, University of Texas Southwestern Medical Center, Dallas, TX 75390, USA, ³Shanghai Institute for Advanced Immunochemical Studies, ShanghaiTech University, Shanghai 201210, China, ⁴Anhui Key Laboratory for Cellular Dynamics & Chemical Biology, Hefei National Research Center for Physical Sciences at the Microscale, University of Science & Technology of China, Hefei 230027, China, ⁵CAS Key Laboratory of Computational Biology, CAS-MPG Partner Institute for Computational Biology, Shanghai Institute of Nutrition and Health, Shanghai Institutes for Biological Sciences, University of Chinese Academy of Sciences, Chinese Academy of Sciences, Shanghai 200031, China, ⁶Department of Biochemistry and Molecular Biology, Tulane University Health Science Center, New Orleans, LA 70112, USA and ⁷The Institute for Brain Research, Collaborative Innovation Center for Brain Science, Huazhong University of Science and Technology, Wuhan 430030, China

Received June 25, 2018; Revised October 18, 2018; Editorial Decision October 19, 2018; Accepted October 22, 2018

ABSTRACT

The kinetochore is a proteinaceous complex that is essential for proper chromosome segregation. As a core member of the inner kinetochore, defects of each subunit in the CENP-H/I/K complex cause dysfunction of kinetochore that leads to chromosome mis-segregation and cell death. However, how the CENP-H/I/K complex assembles and promotes kinetochore function are poorly understood. We here determined the crystal structures of CENP-I N-terminus alone from *Chaetomium thermophilum* and its complex with CENP-H/K from *Thielavia terrestris*, and verified the identified interactions. The structures and biochemical analyses show that CENP-H and CENP-K form a heterodimer through both N- and C-terminal interactions. CENP-I integrates into the CENP-H/K complex by binding to the C-terminus of CENP-H, leading to formation of the ternary complex in which CENP-H is sandwiched between CENP-K and CENP-I. Our sequence comparisons and mutational analyses showed that this architecture of the CENP-H/I/K complex is conserved in human. Mutating the binding interfaces of CENP-H for either

CENP-K or CENP-I significantly reduced their localizations at centromeres and induced massive chromosome alignment defects during mitosis, suggesting that the identified interactions are critical for CENP-H/I/K complex assembly at the centromere and kinetochore function. Altogether, our findings unveil the evolutionarily conserved assembly mechanism of the CENP-H/I/K complex that is critical for proper chromosome alignment.

INTRODUCTION

To ensure faithful segregation of chromosomes, the kinetochore, a multi-subunits protein complex, must be properly assembled at centromere during cell division (1–4). The kinetochore comprised of tens of proteins is organized into several layers. The outer layer of the kinetochore is assembled by three sub-complexes known as KMN (the KNL1, MIS12 and NDC80 complexes), which provides the microtubule-binding platform for chromosomes (5–11). At the inner layer, a number of centromere proteins form a network termed CCAN (the constitutive centromere-associated network) that binds to the centromere throughout the cell cycle (12–16). In vertebrates, the CCAN consists of 16 centromere proteins that are grouped into several

*To whom correspondence should be addressed. Tel: +86 27 87792072; Fax: +86 27 87792072; Email: tianwei@hust.edu.cn
Correspondence may also be addressed to Xiaojing He. Tel: +86 27 87792072; Fax: +86 27 87792072; Email: hexj@hust.edu.cn

†The authors wish it to be known that, in their opinion, the first two authors should be regarded as Joint First Authors.

sub-complexes: CENP-C, CENP-L/N, CENP-H/I/K/M, CENP-T/W/S/X and CENP-O/P/Q/R/U (12,16). Orthologs of some of these sub-complexes have been identified in other species, including fungi (10,11,17–19). As a substitute for canonical H3 in nucleosomes, the histone H3 variant CENP-A deposits at centromeres nucleosomes (20–23) and initiates the CCAN assembly by binding to CENP-L/N and CENP-C (24–29). Extensive studies have established the essential role of CCAN in directing the assembly of the outer kinetochore (2,4,8,14,15). CENP-C and CENP-T act as structural platform for outer kinetochore assembly by directly interacting with the MIS12 and NDC80 complexes (27,30–38).

Many components in the CCAN are held together by a complicated protein–protein interactions network (14,15,39–46). But how these interactions assemble the CCAN complex remains incompletely understood. As the core subunits of the CCAN, CENP-H, CENP-I and CENP-K (also known as Mcm16/Ctf3/Mcm22 in *Saccharomyces cerevisiae* and Fta3/Mis6/Sim4 in *Schizosaccharomyces pombe*) assemble into a ternary complex and are essential for the integrity of the kinetochore. Loss of any of these proteins significantly compromises chromosome congression (12,16). Their centromeric localizations have also been shown to be dependent on each other (12,13,16,32,43,47,48). Another CCAN subunit CENP-M has been shown by *in vitro* reconstitution to form a stable complex with CENP-H/I/K through its interaction with the C-terminus of CENP-I. This interaction is important for the CENP-I/M localization into centromere and chromosome alignment (39,43). Although the overall organization of the CENP-H/I/K/M complex has been shown by the electron microscopy analyses at low-resolution (39), the precise molecular basis for the assembly of the CENP-H/I/K complex remains largely uncharacterized. In this study, we determined the crystal structures of CENP-I N-terminus alone from *Chaetomium thermophilum* and its complex with C-termini of CENP-H/K proteins from *Thielavia terrestris*. Based on the structural and biochemical data, we identified the interacting residues that are important for the assembly of the human CENP-H/I/K complex and centromeric localization. Our findings here establish an evolutionarily conserved assembly mechanism of the CENP-H/I/K complex essential for proper chromosome alignment and segregation.

MATERIALS AND METHODS

Protein expression and purification

The full-length *T. terrestris* CENP-K (*thCENP-K*) and *C. thermophilum* CENP-I residues 1–229 (*ctCENP-I^{NT}*) were subcloned into the BamHI and XhoI sites of pGEX-6p-1 (GE Healthcare) vector with an N-terminal GST tag. The full-length *T. terrestris* CENP-H (*thCENP-H*) was subcloned into the BamHI and XhoI sites of modified pET-28a vector (Novagen) with an N-terminal His6-SUMO tag. All *thCENP-H* and *hsCENP-H* NT mutants were generated with standard two-step PCR-based methods and confirmed by DNA sequencing. *thCENP-H* and *thCENP-K* were co-expressed in *Escherichia coli* BL21 (DE3) cells, cultured in Terrific Broth medium at 37°C and induced by 0.2 mM

Isopropyl β-D-1-thiogalactopyranoside (IPTG) at 16°C overnight when OD₆₀₀ ~1.5 was reached. Cells were harvested and disrupted by high-pressure homogenizer (ATX Engineering) in the phosphate-buffered saline (PBS) buffer, and then clarified by centrifuged at 35 000 g for 45 min at 4°C. The supernatant was added GST agarose beads (GE Healthcare) and rotated at 8°C for 1 h, PBS washed the collected beads in the gravity column and tagged-free using sumo-protease overnight at 4°C. The *ctCENP-I^{NT}* was expressed and affinity purified with same method. The eluted *thCENP-H/K* complex mixed *ctCENP-I^{NT}* on ice for 1 h to form hybrid CENP-H/I/K complex in buffer containing 20 mM Tris pH 8.0, 50 mM NaCl, 1 mM Dithiothreitol (DTT) and then further purified by anion exchange chromatography (HiTrap Q FF, GE Healthcare) and gel filtration chromatography (superdex 200 10/300 GL, GE Healthcare). The purified fungal CENP-H/I/K complex was concentrated to 15–20 mg/ml in buffer containing 20 mM Tris pH 8.0, 150 mM NaCl, 1 mM DTT for crystallization.

For *ctCENP-I^{NT}*, the eluted untagged protein from gravity column in buffer containing 20 mM Hepes pH 7.0, 50 mM NaCl, 1 mM DTT further purified by cation exchange chromatography (HiTrap SP FF, GE Healthcare) and gel filtration chromatography (superdex 200 10/300 GL, GE Healthcare). The purified *ctCENP-I^{NT}* was concentrated to 5–8 mg/ml in buffer containing 20 mM Tris pH 7.6, 150 mM NaCl, 1 mM DTT for crystallization.

Crystallization, data collection and structure determination

For Crystallization trial, 2.2 mg/ml purified *ctCENP-I^{NT}* protein crystallized at 18°C using the hanging drop vapor-diffusion method by mixing 1 μl protein solution and 1 μl reservoir solution containing 0.1 M MES pH 6.5, 40% MPD, 5% PEG8000 for ~2 days. The purified CENP-H/I/K protein complex was concentrated to 10 mg/ml crystallized with same method in reservoir solution containing 0.1M HEPES pH 7.2 and 20% PEG8000 for ~10 days. The crystals were cryo-protected with reservoir solution supplemented with 20% glycerol and then flash-cooled in liquid nitrogen.

Crystals of *ctCENP-I^{NT}* and CENP-H/I/K diffraction data were collected at beamline BL17U and BL18 at Shanghai Synchrotron Radiation Facility (SSRF) and processed, integrated and scaled together with HKL3000 software (49). Initial phases for seleno-methaione labeled *ctCENP-I^{NT}* were obtained by SAD in the program AutoSol of Phenix software (50). Initial phases for CENP-H/I/K protein complex were obtained by molecular replacement in the program Phaser using the structure of *ctCENP-I^{NT}* as a search model. Iterative model building and refinement were performed in the programs Coot and Phenix (50,51), respectively. Data collection and refinement statistics are summarized in Supplementary Table S1.

Analytical ultracentrifugation (AUC)

Sedimentation velocity (SV) experiments were performed using ProteomeLab XL-I analytical ultracentrifuge (Beckman Coulter, Palo Alto, CA, USA). Protein was prepared

in PBS Buffer containing 10 mM Na₂HPO₄, 1.8 mM KH₂PO₄, 2.7 mM KCl and 137 mM NaCl by size exclusion chromatography (SEC) (Superdex 200 10/300 GL, GE Healthcare) and then concentrated to 1 absorbance at 280 nm. Experiments were done at 4°C using double-sector centerpieces and sapphire windows at 42 000 rpm for about 10 h (An-60 Ti Rotor). The volume of sample loaded in the sample cell of the centerpiece was 380 µl, while the reference cell contained 400 µl PBS Buffer. Sedimentation profiles were recorded with absorbance and/or interference optics, and these data were analyzed by SEDFIT software (52) and plotted with program GUSI.

***In vitro* protein binding assays**

5 µg GST fusion protein (GST-*ct*CENP-I^{NT}, GST-*th*CENP-H, GST-*th*CENP-K full-length or truncations, and GST-*hs*CENP-K full-length or truncations) in buffer containing 20 mM Tris pH 8.0, 100 mM NaCl and 1 mM DTT was bound to equilibrated GST beads previously at 8°C for 1 h, then incubated with CENP-H or other target proteins at 8°C for 45 min, and washed three times with buffer containing 20 mM Tris pH 8.0, 500 mM NaCl, 1 mM DTT and 0.02% Triton. The proteins retained on the beads were analyzed by sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE). GST-tag bound beads were used as controls. The relative binding activity was measured by the amount of bound CENP-H or other target proteins normalized to that of GST-tagged protein on the GST beads.

Mammalian cell culture and transfection

HeLa Tet-On (Invitrogen) cells were grown in Dulbecco's modified Eagle's medium (DMEM; Invitrogen) supplemented with 10% fetal bovine serum and 10 mM L-glutamine. To arrest cells at G1/S, cells were incubated in the growth medium containing 2 mM thymidine (Sigma) for 17 h. G2 cells were collected at 7 h after the release from thymidine arrest. Mitotic cells were obtained by adding 5 mM MG132 (Sigma) at 8 h after the release from thymidine arrest and incubating for another 2 h. Plasmid transfection was performed when cells reached a confluency of about 40–60% using the Effectene reagent (Qiagen) per manufacturer's protocols. For RNAi experiments, the siRNA oligonucleotides were purchased from Thermo Scientific. HeLa cells were transfected using Lipofectamine RNAiMax (Invitrogen) and analyzed at 24–48 h after transfection. The sequences of the siRNAs used in this study are: siCENP-H: AGAUUGAUUUGGACAGUAU, siCENP-I: GAAGGUGUGUGACAUUAU from Thermo Scientific (48).

Antibodies, immunoblotting and immunoprecipitation

The commercial antibodies used in this study were: mouse anti-CENP-H (ab88593, Abcam), mouse anti-CENP-I (ab168778, Abcam), rabbit anti-CENP-K (sc-81831, Santa Cruz), rabbit anti-GFP ((in house made polyclonal antibody), mouse anti-myc (11667203001, Roche)), mouse

anti-β-tubulin (T4026, Sigma-Aldrich), human CREST autoimmune sera (HCT-0100, ImmunoVision), mouse anti-α-tubulin (T9026; Sigma-Aldrich). For immunoblotting, antibodies were used at 1 µg/ml for purified and monoclonal antibodies or at 1: 1000 dilutions for crude sera.

For immunoprecipitation, anti-MYC or anti-GFP antibodies were coupled to Affi-Prep Protein A beads (Bio-Rad) at a concentration of 1 mg/ml. HeLa cells were lysed with the Lysis Buffer (25 mM Tris-HCl pH 7.5, 75 mM NaCl, 5 mM MgCl₂, 0.1% NP-40, 1 mM DTT, 0.5 µM okadaic acid, 5 mM NaF, 0.3 mM Na₃VO₄ 10 mM β-glycerophosphate and 50 units/ml Turbo-nuclease). After 2 h incubation on ice and then 10-min incubation at 37°C, the lysate was cleared by centrifugation for 20 min at 4°C at 14 000 rpm. The supernatant was incubated with the antibody beads for 2 h at 4°C. The beads were washed four times with the lysis buffer. The proteins bound to the beads were dissolved in SDS sample buffer, separated by SDS-PAGE and blotted with the appropriate antibodies.

Immunofluorescence

For whole cell staining, cells were grown and transfected on a Lab-Tek II chamber slide. After the medium was removed, cells were pre-extracted with the PHEM buffer (60 mM PIPES, 25 mM HEPES, pH 6.9, 10 mM EGTA, 2 mM MgCl₂) containing 0.2% Triton X-100 for 2 min, and fixed by 4% paraformaldehyde in PBS (PBSP) for 4 min, and blocked for 30 min with PBS containing 3% BSA (PBSB). Cells were then incubated with indicated primary antibodies in PBS for 1 h at room temperature or overnight at 4°C. Cells were washed three times with PBS containing 0.1% Triton X-100 (PBST) and incubated with fluorescent secondary antibodies in PBSB for 1 h at room temperature. Slides were washed with PBS-T again and mounted in ProLong Gold Antifade reagent with DAPI (Invitrogen), mounted with Aqua-Poly/Mount (Polysciences, Inc.) and sealed with nail polish. Cells were visualized with a DeltaVision microscope system (Applied Precision). Alexa Fluor 488, 568 or 647, or DAPI fluorescence was observed with the appropriate filter sets. All images were acquired on the same system with a 100X NA1.4 UPLS APO objective (Olympus) in a Z-stack series of 0.2-µm intervals. All images in each experiment were taken with the same light intensity and exposure time. Images were deconvolved and projected by the Sum Intensity method in SoftWoRx (Applied Precision), and further processed and analyzed with ImageJ.

Quantification of the relative intensity of kinetochore signals was done with ImageJ. For quantification, kinetochore regions were selected based on CREST signals. The non-kinetochore region was defined as the chromatin region with the kinetochore region subtracted. A circle that enclosed CREST signals from a pair of kinetochores was drawn and set as the region of interest (ROI). The integrated density of the gray value for the selected ROI was measured from each channel. The value of object intensity was then divided by the corresponding value of CREST intensity. In total, 20 ROIs per cell chosen at random were measured. The graphs and statistics were generated with Prism

(GraphPad). In all plots, each dot represents one cell. All *P*-values were calculated using the Student's *t*-test with the Prism software (GraphPad). For presentation, images were further processed with Photoshop (Adobe).

RESULTS

Reconstitution and three-dimensional structures of the ternary complex of fungal CENP-H/I/K homologs

Our attempts to reconstitute *Homo sapiens* (*hs*) CENP-H/I/K complex for structural analyses failed due to poor stability of *hs*CENP-I. We therefore searched for orthologs of these proteins from other species. Searches in the Pfam and InterPro databases showed that the protein G2R207, G2QRQ0 and G2R3T1 from *T. terrestris* (*th*) display significant sequence homology to human CENP-H, CENP-I and CENP-K, respectively (Figure 1A) (53,54). Reciprocal Best BLAST searches strongly support the orthologous relationship between fungi and human proteins. For example, the sequence of G2QRQ0 is significantly similar to *hs*CENP-I (40% similarity, *E*-value 7.8×10^{-20}). We also constructed phylogenetic trees and sequence alignments of CENP-H, I and K from divergent species, which further demonstrated the homology relationship between fungi, yeast and human proteins (Supplementary Figures S1–4). Therefore, we annotated these three proteins in *T. terrestris* as CENP-H, CENP-I and CENP-K homologs (*th*CENP-H/I/K) (Figure 1A), respectively. The *th*CENP-H/I/K complex expressed at higher levels and showed better stability, but did not crystallize (Supplementary Figure S1D). Instead, we found that CENP-I from another fungi *C. thermophilum* (*ct*) is 62% identical in sequence to *th*CENP-I, and showed excellent biochemical behavior. However, we could not reconstitute the *ct*CENP-H/I/K complex because there is no CENP-H homolog identified in *C. thermophilum*.

Due to the high degree of sequence similarity, we tested whether *ct*CENP-I could form a hybrid complex with *th*CENP-H/K for crystallization. It has been shown that the N-terminal region of *hs*CENP-I is necessary and sufficient for interacting with *hs*CENP-H/K (32,39). We therefore purified the corresponding N-terminal region of *ct*CENP-I (*ct*CENP-I^{NT}) (Figure 1A), which indeed formed a stable hybrid complex with the *th*CENP-H/K proteins (Figure 1B). The analytical ultracentrifugation (AUC) analyses showed that the hybrid CENP-H/I/K complex had a stoichiometry of 1:1:1 (Figure 1C).

We first determined the structure of *ct*CENP-I^{NT} at 2.3 Å with the Single-wavelength Anomalous Dispersion (SAD) method (Supplementary Table S1). *ct*CENP-I^{NT} adopts the HEAT repeat fold containing five pairs of helices $\alpha 1$ – $\alpha 10$. A following helix $\alpha 11$ folds back and packs against the HEAT repeat with extensive hydrophobic interactions, which likely contributes to the protein stability (Figure 2A and B). The structure of *ct*CENP-I is similar to a model of *hs*CENP-I proposed previously (39), except that helix $\alpha 11$ packs on the surface of the HEAT repeat in our structure.

We also determined the crystal structure of the hybrid complex between *ct*CENP-I^{NT} and *th*CENP-H/K at 2.2 Å with the Molecular Replacement method using the *ct*CENP-I^{NT} structure as start model (Figure 2C and D; Supplementary Table S1). The structure of *ct*CENP-I^{NT}

in the complex is very similar to the apo state (Figure 2E). In one asymmetric unit of the complex crystal, one *th*CENP-H/K heterodimer interacts with one *ct*CENP-I^{NT} homodimer (Supplementary Figure S5A). The homodimer of *ct*CENP-I^{NT} might be a crystal packing artifact, since the complex shows a 1:1:1 stoichiometry in solution (Figure 1C). Although the full-length *th*CENP-H/K proteins were used in crystallization, only the C-terminal portions of *th*CENP-H (184–228aa, *th*CENP-H^{CT}) and *th*CENP-K (161–328aa, *th*CENP-K^{CT}) are found in the structure (Figure 2C and D). The N-terminal portions of *th*CENP-H and *th*CENP-K are missing in the electron density map, which might be caused by protein degradation during crystallization, as shorter fragments were observed in electrophoresis analyses of the complex crystals (Supplementary Figure S5B).

Overall, the ternary complex adopts a sandwich-like structure. *ct*CENP-I^{NT} and *th*CENP-K^{CT} on each side pack against *th*CENP-H^{CT} in the middle (Figure 2C and D). *th*CENP-H^{CT} contains two α -helices HH1 and HH2 to interact with both *th*CENP-K^{CT} and *ct*CENP-I^{NT}. In the structure of *th*CENP-K^{CT}, three helices KH1, KH2 and KH3 wrap around the HH1 helix of *th*CENP-H^{CT} (Figure 2C and D). The loop region (192–257aa) between KH1 and KH2 is invisible in the electron density map (Figure 2F). This loop region is not conserved in *hs*CENP-K, suggesting that it may not be functional important (Supplementary Figure S3). On the other side of the complex, helices $\alpha 5$ and $\alpha 7$ of *ct*CENP-I^{NT} contact with HH2 helix of *th*CENP-H^{CT} (Figure 2C and D).

CENP-H and CENP-K form a heterodimer through both N-terminal and C-terminal interactions

To clarify the role of the N-terminal portions of *th*CENP-H and *th*CENP-K, which are missing in the structure, we purified C-terminal truncated version of *th*CENP-H and *th*CENP-K. We found that *th*CENP-H^{N164} and *th*CENP-K^{N143} form a stable complex in gel-filtration chromatography (Figure 2G, left panel). These results indicated that *th*CENP-H and *th*CENP-K form a heterodimer through both N-terminal and C-terminal interaction. Addition of *ct*CENP-I^{NT} did not alter the behavior of *th*CENP-H^{N164}/*th*CENP-K^{N143} complex in the gel-filtration analysis (Figure 2G, right panel). Consistently, no obvious interaction between *ct*CENP-I^{NT} and *th*CENP-H^{N164}/*th*CENP-K^{N143} was observed in GST pull-down experiments (Supplementary Figure S6A). These results together confirmed that *ct*CENP-I^{NT} only interact with the C-termini of *th*CENP-H/K complex as shown in the crystal structure.

To establish the generality of this dual binding mode in CENP-H/K complex, we also characterized the interactions between human CENP-H and CENP-K using GST pull-down assay. The results showed that the N-terminal fragment of *hs*CENP-K (*hs*CENP-K^{N85}) only bound to the N-terminal portion of *hs*CENP-H (*hs*CENP-H^{N120}) (Figure 2H, left panel), whereas the C-terminal fragment of *hs*CENP-K (*hs*CENP-K ^{Δ N85}) only bound to C-terminal portion of *hs*CENP-H (*hs*CENP-H ^{Δ N120}) (Figure 2H, right panel). Thus, the dual binding model represents a conserved mechanism for the formation of the CENP-H/K complex.

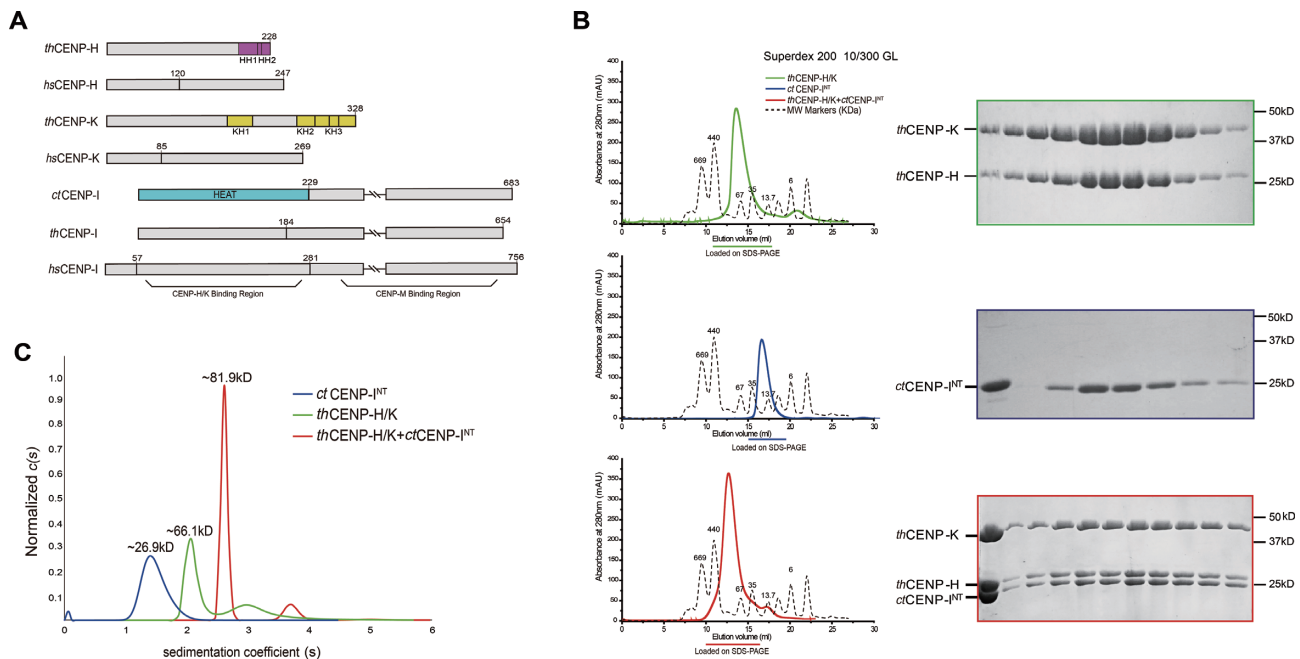


Figure 1. *In vitro* reconstitution of CENP-H/I/K complex from thermophilic fungi. (A) Schematic view of *thCENP-H/I/K*, *hsCENP-H/I/K* and *ctCENP-I* proteins. The colored regions denote the domains revealed in the structures. The numbered regions represent the constructs used in this study. It has been identified previously that N-terminus of *hsCENP-I* binds to CENP-H/K, and C-terminus interacts with CENP-M (39). (B) SEC elution profiles of *thCENP-H/K*, *ctCENP-I^{NT}* and ternary complex, respectively, from top to bottom. Peak fractions were resolved with SDS-PAGE and stained with Coomassie Blue. (C) SV AUC analysis of *ctCENP-I^{NT}* (blue), *thCENP-H/K* (green) and ternary complex (red). The molecular weights were calculated based on sedimentation coefficient, obtained with SEDFIT software using a continuous c(s) model.

Conserved residues mediate C-terminal interaction of CENP-H/K complex

The structure of fungal CENP-H/I/K complex revealed the details of the C-terminal binding interface of CENP-H/K complex. The sidechain of *thCENP-H* I205 inserts into a hydrophobic cavity surrounded by L177, F180, I270 and F300 of *thCENP-K*. On the other side, I211 and L219 of *thCENP-H* contact with the hydrophobic patch formed by W179, F180 and H184 of *thCENP-K* (Figure 3A). To validate these interactions, we mutated the interface residues in *thCENP-H* and tested the effects of these mutations in an *in vitro* binding assay. The relative binding activity was measured by the amount of bound *thCENP-H* normalized to that of GST-tagged protein on the GST beads. We found that *thCENP-H* WT robustly bound to GST-*thCENP-K^{ΔN75}*. The single mutants *thCENP-H* I205A, I211A and L219A, exhibited weaker binding to *thCENP-K^{ΔN75}*. As expected, the weakened effects of the double mutants I205A/I211A and I205A/L219A were much stronger (blue column in Figure 3B and Supplementary Figure S6B). In contrast, mutations of the neighboring residues, such as R220 or L224, did not alter the binding to *thCENP-K^{ΔN75}* (Figure 3B). Thus, the C-terminal surface revealed by our crystal structure indeed mediates the interaction of *thCENP-H/K* complex. Interestingly, the full-length *thCENP-K* (*thCENP-K^{FL}*) retained binding to the *thCENP-H* mutants (blank column in Figure 3B and Supplementary Figure S6C). For example, *thCENP-H* I205A/L219A significantly weakened the binding to

thCENP-K^{ΔN75}, but remained the binding to *thCENP-K^{FL}* (Figure 3B). These results further confirmed the N-terminal interaction between *thCENP-H* and *thCENP-K*.

The sequence alignment of CENP-H orthologs showed that the residues involved in binding CENP-K are highly conserved (Figure 3C). We mutated the corresponding residues in *hsCENP-H*, including L219, V225 and L233 (yellow spots in Figure 3C), and tested the binding to *hsCENP-K* full length (*hsCENP-K^{FL}*) or N-terminal truncation (*hsCENP-K^{ΔN85}*) in the GST pull-down assay (Figure 3D; Supplementary Figure S7A and B). The results are similar to those with *thCENP-H* and *thCENP-K*. For instance, *hsCENP-H* L219A/L233A lost the interaction with *hsCENP-K^{ΔN85}*, but retained partial binding to *hsCENP-K^{FL}* (Figure 3D). These results suggest the binding mode between CENP-H and CENP-K as shown by our crystal structure are conserved in human.

CENP-H directly interacts with CENP-I through the conserved C-terminal surface

thCENP-H mainly utilizes its HH2 helix to interact with the HEAT repeat surface of *ctCENP-I^{NT}*. The salt-bridge is formed between R220 in *thCENP-H* and E86 of *ctCENP-I^{NT}*. L224 in *thCENP-H* interacts with the hydrophobic concavity formed by L89, V126 and V130 in *ctCENP-I^{NT}* (Figure 4A). To validate these interactions, we mutated the interface residues in *thCENP-H* and tested the binding activity. The results showed that the R220E and L224A mutations of *thCENP-H* dramatically decreased the binding to

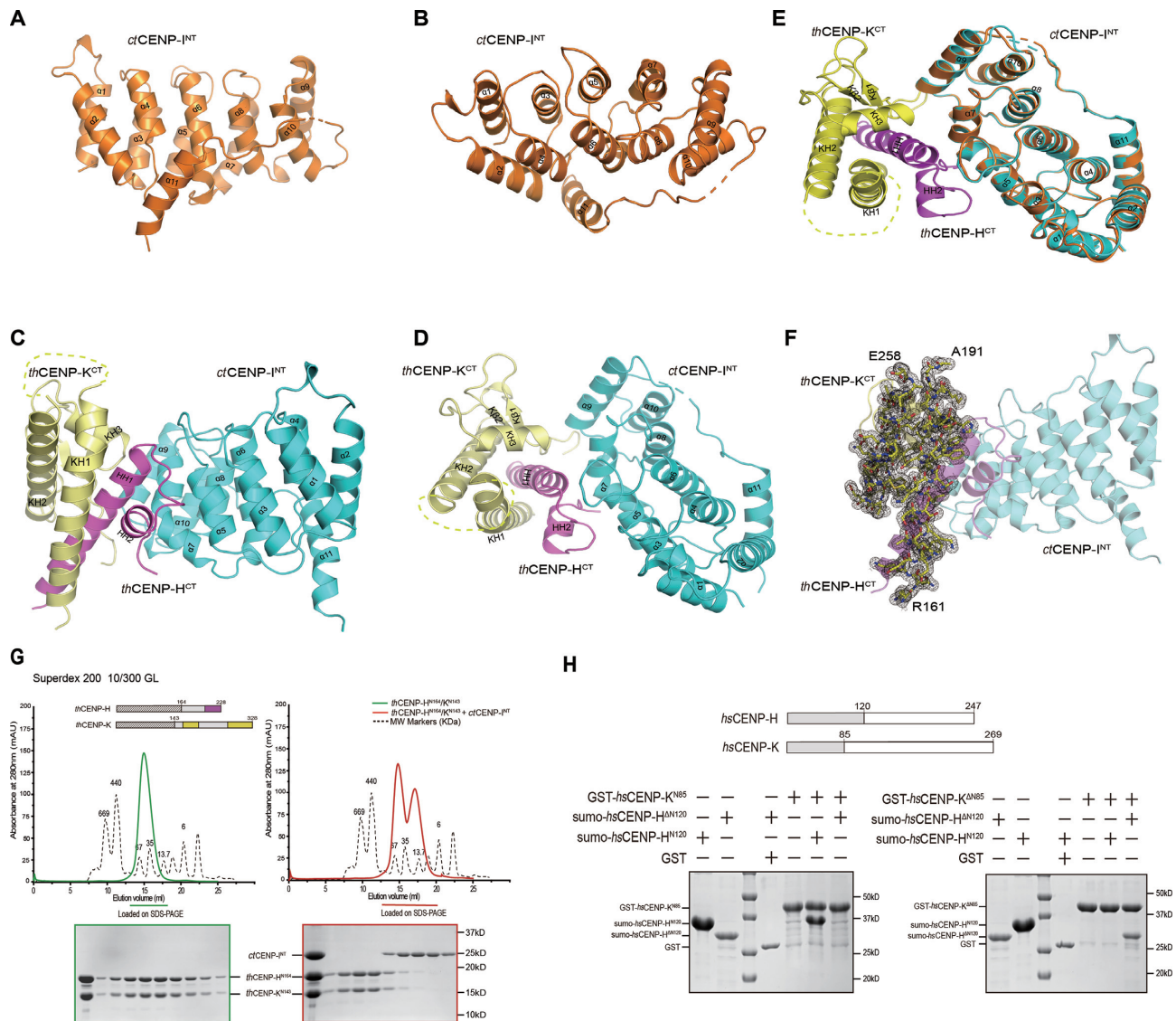


Figure 2. Crystal structures of *ctCENP-I^{NT}* alone and its complex with *thCENP-H^{CT}/K^{CT}*. (A and B) Cartoon view of the crystal structure of *ctCENP-I^{NT}* alone in side (A) and top (B) orientations. Residues 198–205 between $\alpha 10$ and $\alpha 11$ not traceable in the electron density map are shown as dotted lines. (C and D) Cartoon view of the crystal structure of fungal CENP-H/I/K complex in side (C) and top (D) orientations. The *thCENP-K^{CT}*, *thCENP-H^{CT}* and *ctCENP-I^{NT}* are colored yellow, magenta and cyan, respectively. (E) Electron density map for *thCENP-K^{CT}* in the complex structure. The composite omit map was calculated with Phenix, contoured at 1.0σ . There is no electron density for residues 192–267 between KH1 and KH2 of *thCENP-K*. (F) Overlay of the ribbon diagrams of the *ctCENP-I^{NT}* alone (orange) and its complex with *thCENP-H^{CT}/K^{CT}* (colored yellow, magenta and cyan, respectively). (G) SEC elution profiles of *thCENP-H^{NT164}/CENP-K^{NT143}* complex (left) and its mixture with *ctCENP-I^{NT}* (right). Peak fractions were resolved with SDS-PAGE and stained with Coomassie Blue. (H) The N-termini of *hsCENP-K* only bound to the N-terminal fragment of *hsCENP-H* (left panel), the C-termini of *hsCENP-K* only bound to C-terminal fragment of *hsCENP-H* (right panel). GST-*hsCENP-K^{N85}* and GST-*hsCENP-K ^{Δ N85}* were immobilized on beads and incubated with sumo-*hsCENP-H^{N120}* or sumo-*hsCENP-H ^{Δ N120}*. The bead-bound proteins were resolved with SDS-PAGE and stained with Coomassie Blue.

GST-*ctCENP-I^{NT}*, while the R220E/L224A double mutation completely abolished the binding (Figure 4B and Supplementary Figure S8A). Thus, the binding interface revealed by our structure is critical for the interaction between *thCENP-H* and *ctCENP-I^{NT}*. Furthermore, the sequence alignment of CENP-H clearly showed that K234 and L238 in *hsCENP-H* correspond to R220 and L224 in *thCENP-H* (Figure 3C), which are critical for binding to *ctCENP-I*, indicating a potential conserved interaction mode between CENP-H and CENP-I.

CENP-K enhances the interaction between CENP-H and CENP-I to form the ternary complex

We next tested how these three proteins form ternary complex using *in vitro* pull-down assays. We found that *thCENP-H* can bind to *ctCENP-I^{NT}* and *thCENP-K* individually (lanes 3 and 6, Figure 4C), consistent with the structure of the complex in which *thCENP-H* is placed at the center and make extensive interactions with both *thCENP-K* and *ctCENP-I^{NT}*. In contrast, *thCENP-K* did not directly interact with *ctCENP-I^{NT}* (lane 5, Figure 4C).

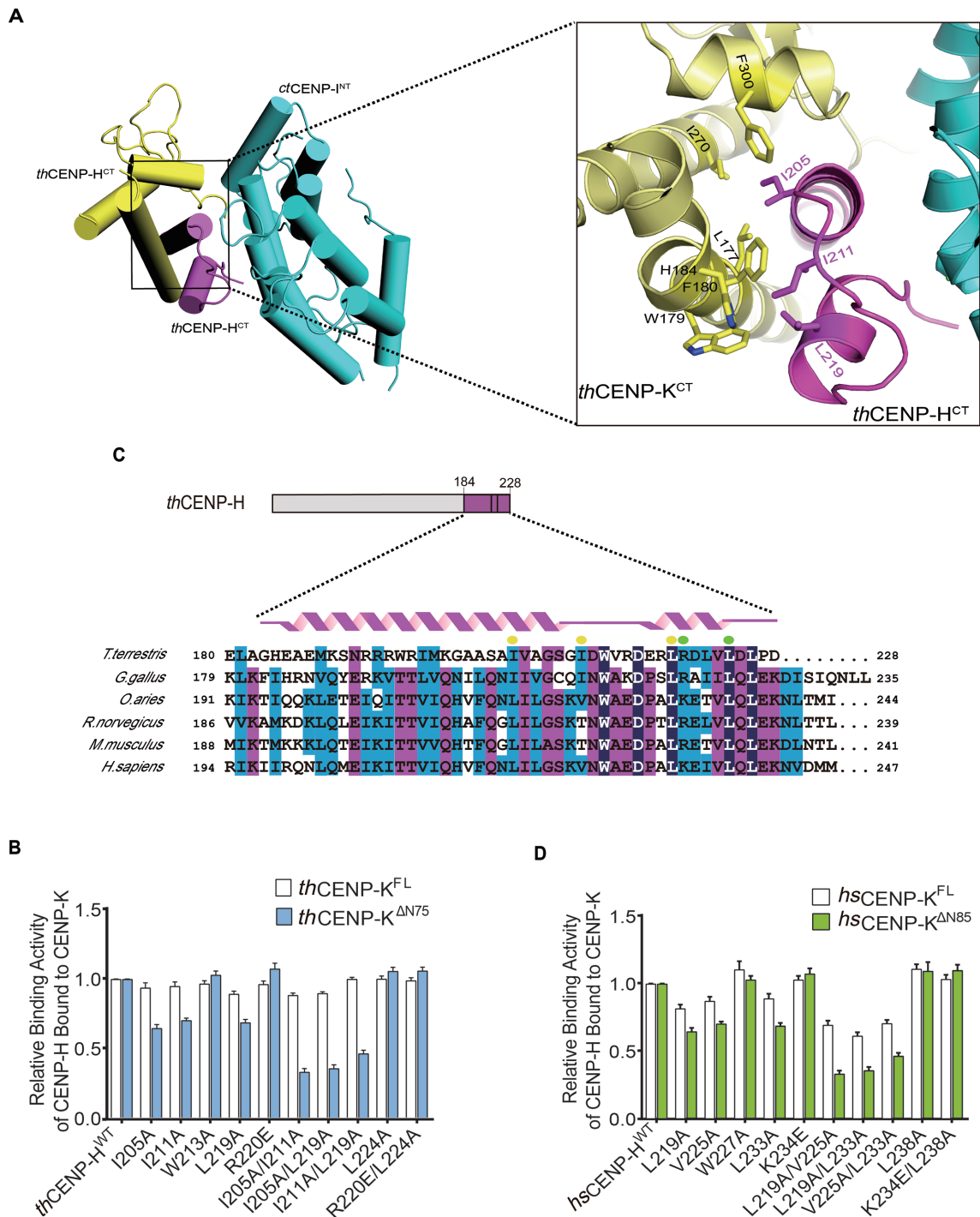


Figure 3. Validate the C-terminal binding interface between CENP-H and CENP-K. (A) Zoomed-in view of the binding interface between *thCENP-H*^{CT} and *thCENP-K*^{CT} in the crystal structure of ternary complex. The interacting residues were drawn in stick and marked in number. (B) The relative binding activity of *thCENP-H* (WT and mutants) bound to GST-*thCENP-K*^{FL} and GST-*thCENP-K*^{ΔN75} was assessed using GST pull-down assays. Error bars represent standard deviations, which were obtained from three independent experiments. The representative results of the pull-down assays were also available in Supplementary Figure S6B and C. (C) Alignment of the C-terminus sequences of CENP-H orthologs across species using Clustal Omega Program. The conserved residues were colored. The interacting residues identified in our structures were marked with yellow dots (binding to *thCENP-K*^{CT}) and green dots (binding to *ctCENP-INT*). (D) The relative binding activity of *hsCENP-H* (WT and mutants) bound to GST-*hsCENP-K*^{FL} and GST-*hsCENP-K*^{ΔN85} was assessed using GST pull-down assays. Error bars represent standard deviations, which were obtained from three independent experiments. The representative results of the pull-down assays were also available in Supplementary Figure S7A and B.

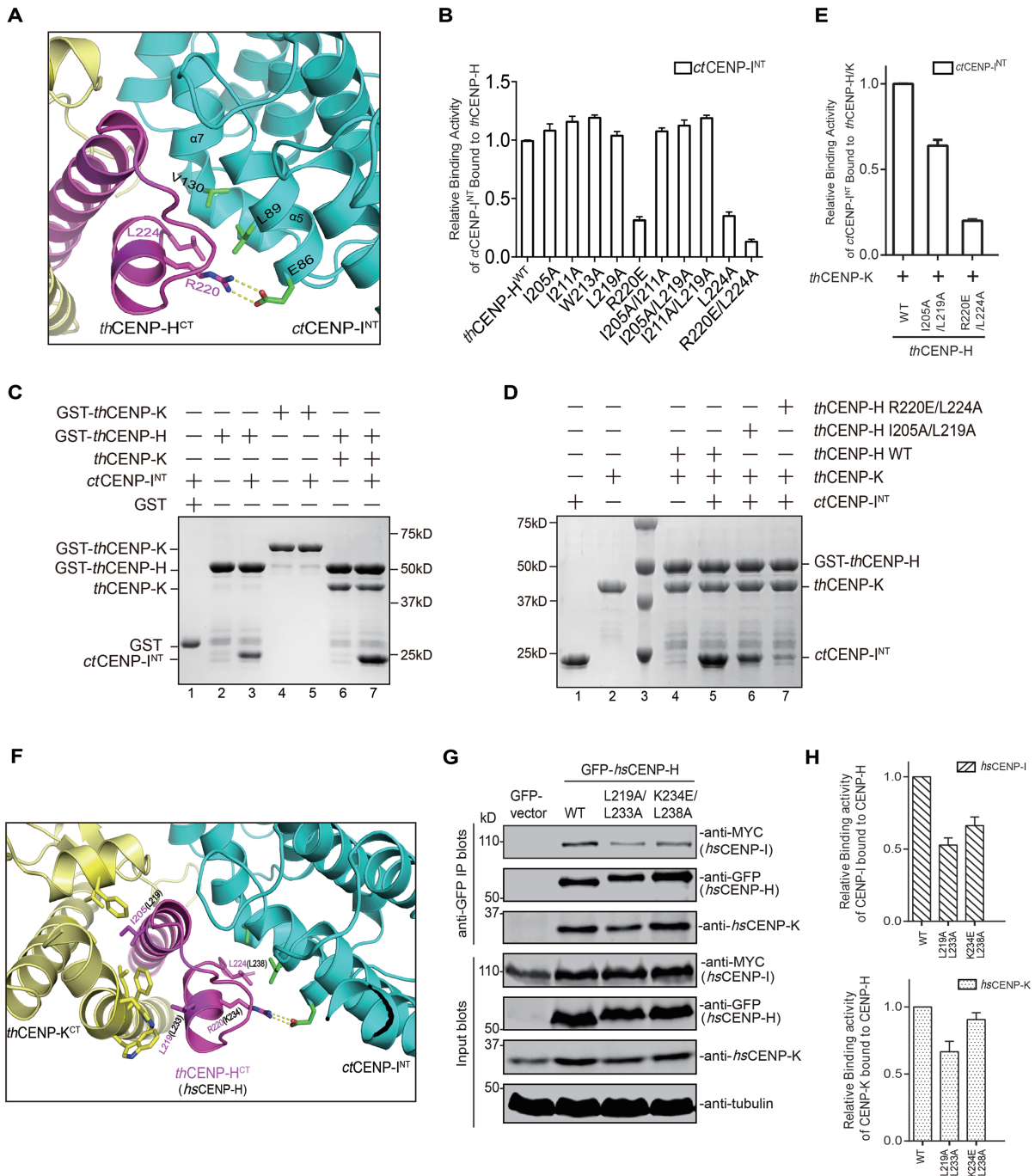


Figure 4. Validate the binding interface between CENP-H and CENP-I, and the assembly mode of CENP-H/I/K ternary complex. **(A)** Zoomed-in view of the binding interface between *thCENP-H^{CT}* and *ctCENP-I^{NT}*. The interacting residues were drawn in stick and marked with number. **(B)** The relative binding activity of *thCENP-H* (WT and mutants) bound to GST-*ctCENP-I^{NT}* was assessed using GST pull-down assays. Error bars represent standard deviations, which were obtained from three independent experiments. The representative results of the pull-down assays were also available in Supplementary Figure S8A. **(C)** *In vitro* binding assays were performed to examine the binding activity among the ternary complex. Proteins with distinct combinations as indicated were mixed and GST pull-down assays were performed. The bead-bound proteins were resolved with SDS-PAGE and stained with Coomassie Blue. **(D)** *In vitro* binding assays were performed to assess the effects of *thCENP-H* mutants on the formation of the fungal CENP-H/I/K ternary complex. Proteins with distinct combinations as indicated were mixed and GST pull-down assays were performed. The bead-bound proteins were resolved with SDS-PAGE and stained with Coomassie Blue. **(E)** Quantification of the relative binding activity between *ctCENP-I^{NT}* and *thCENP-H* (WT and mutants) in the presence of *thCENP-K* based on the results from lanes 5, 6 and 7 in **(D)**. **(F)** Zoomed-in view of the binding interfaces of *thCENP-H^{CT}* in the complex structure. The interacting residues in *thCENP-H* were drawn in stick and marked with numbers in red. And the counterpart residues in *hsCENP-H* were shown with numbers in black. **(G)** HeLa Tet-On cells transfected with GFP-*hsCENP-H* (WT or mutants), non-tagged *hsCENP-K* and MYC-*hsCENP-I*. Cell lysates were treated with anti-GFP antibody. Immunoprecipitated samples were resolved with SDS-PAGE and blotted with the indicated antibodies. **(H)** Quantification of the relative binding activity between *hsCENP-I* and *hsCENP-H* WT or mutants (upper panel), and between *hsCENP-K* and *hsCENP-H* WT or mutants (lower panel) based on the results from Figure 4G.

Interestingly, addition of *thCENP-K* significantly increased the binding of *ctCENP-I^{NT}* to *thCENP-H* (lanes 3 and 7 in Figure 4C). The enhancement effect of *thCENP-K* is lost for the *thCENP-H* I205A/L219A (Figure 4D and E), which disrupts the C-terminal interaction of *thCENP-H/K* (Figure 3B). In contrast, *ctCENP-I^{NT}* did not affect the interactions between *thCENP-H* and *thCENP-K* (Figure 4C), and the *thCENP-H* R220E/L224A also disrupted the binding of *ctCENP-I^{NT}* without affecting the *thCENP-H/K* interactions (Figure 4D and E). These results together suggest that formation of CENP-H/K complex enhances the recruitment of CENP-I. It is possible that binding of *thCENP-K* to *thCENP-H* might induce conformational changes in *thCENP-H*, thus presenting a better position of R220 and L224 for *ctCENP-I^{NT}* binding (Figure 4F).

We then asked whether *hsCENP-K* can also promote the assembly of the human CENP-H/I/K complex as *thCENP-K* does. The sequence alignment of CENP-H showed that K234/L238 in *hsCENP-H* correspond to R220/L224 in *thCENP-H* for *ctCENP-I^{NT}* binding, while L219/L233 in *hsCENP-H* correspond to I205/L219 in *thCENP-H* for *thCENP-K* binding (Figures 3C and 4F). We mutated these two potential interfaces in *hsCENP-H* to test the assembly of CENP-H/I/K ternary complex in HeLa cells. HeLa cells were co-transfected with MYC-*hsCENP-I*, non-tagged *hsCENP-K* and GFP-*hsCENP-H* WT or mutants. The formation of CENP-H/I/K complex was assessed by immunoprecipitation experiments using an anti-GFP antibody. The results showed that GFP-*hsCENP-H* WT robustly bound to both *hsCENP-I* and *hsCENP-K* (Figure 4G). As expected, GFP-*hsCENP-H* K234E/L238A showed a diminished binding to *hsCENP-I*, but retained the ability to bind *hsCENP-K* (Figure 4G and H), which is consistent with the binding results of fungal protein (Figure 4B and D). These results strongly support the notion that the binding interface between CENP-H and CENP-I are conserved from fungi to human. Interestingly, GFP-*hsCENP-H* L219A/L233A not only exhibited a decreased binding to *hsCENP-K*, also exhibited a significantly decreased binding to *hsCENP-I* (Figure 4G and H), as *thCENP-H* I205A/L219A did to *thCENP-I^{NT}* for the ternary complex formation (Figure 4D and E). These results highlighted that the enhancement of CENP-K is also critical for human CENP-H/I/K complex formation. Thus, all these results indicated that human CENP-H/I/K complex likely adopts a similar interaction mode as its thermophile fungus orthologs, suggesting that this assembly mode may be conserved across species.

Maintaining the proper assembly of CENP-H/I/K complex is essential for centromeric localization and faithful chromosome segregation

We also examined the centromeric localization of *hsCENP-H* and *hsCENP-I* in mitotic cells. Both *hsCENP-H* and *hsCENP-I* co-localized with the centromere marker CREST in the control cells (Figure 5A and B). Depletion of *hsCENP-H* by treatment with *hsCENP-H* siRNA almost completely abolished the centromeric localizations of both *hsCENP-H* and *hsCENP-I* (Figure 5A–C). This was

not caused by cross effects of the siRNA since *hsCENP-H* siRNA significantly decreased the protein level of *hsCENP-H*, but not *hsCENP-I*, and vice versa (Supplementary Figure S8B).

We next sought to investigate whether the interactions among the CENP-H/I/K complex affect their centromeric localization during mitosis. siRNA-resistant GFP-*hsCENP-H* WT expressed in *hsCENP-H* depleted cells, showed the normal centromeric localization and restored the localizations of *hsCENP-I* to centromeres (Figure 5D–F). In contrast, both *hsCENP-H* L219A/L233A and K234E/L238A failed to localize into centromeres, and therefore could not restore *hsCENP-I* localization (Figure 5D–F). The localization defects of *hsCENP-H* mutants was unlikely caused by decreased expression levels or protein stability, since they were expressed at levels comparable to the wild-type protein and interacted with *hsCENP-K* (Supplementary Figure S8C and Figure 4G). Thus, these results suggested that the interactions among CENP-H/I/K complex are essential for maintaining their centromeric localizations during mitosis, and confirmed previous reports showing that centromeric localization of these proteins are interdependent (12,13,16,43,47,48). Therefore, efficient incorporation of these proteins into the kinetochore requires the proper formation of the CENP-H/I/K ternary complex.

Moreover, we examined the chromosome alignment phenotype during mitosis. Approximately 80% of *hsCENP-H* depleted cells ($n > 50$) exhibited significant chromosome alignment defects, while chromosomes aligned properly in most of the control cells (Figure 5A, B and G). The expression of GFP-*hsCENP-H* WT significantly rescued the chromosome alignment defects in *hsCENP-H* depleted cells (Figure 5D and G). In contrast, expressions of the *hsCENP-H* L219A/L233A and K234E/L238A mutants failed to do so (Figure 5D and G). These results show that the interactions among the CENP-H/I/K complex are critical for proper chromosome alignment and segregation in mitosis.

DISCUSSION

Determining how the CCAN subunits are organized is important for understanding the kinetochore assembly and function. Among the CCAN subunits, CENP-H, CENP-I, CENP-K and CENP-M have been previously shown to form a stable sub-complex based on proteomic analyses and reconstitution experiments (12,13,16,24,30,35,36, 39). In addition, a previous study has presented the low-resolution structure and model of the CENP-H/I/K/M quaternary complex (39). Our study for the first time provides the high-resolution structure of the fungal CENP-H/I/K complex. Further analyses suggest that the interactions revealed in this complex represent the evolutionarily conserved mechanism for the assembly of the CENP-H/I/K complex.

CENP-A nucleosome directly recruits CENP-C and CENP-L/N to the centromere, and initiates assembly of the CCAN (25,26,29,32,43,44). CENP-C has been shown to directly interact with CENP-H/K, not CENP-I, which contributes to centromeric recruitment of the CENP-H/I/K/M complex (32). CENP-L/N and CENP-H/I/K/M are mutually required for centromeric localization (43), although the detailed interactions remain unclear.

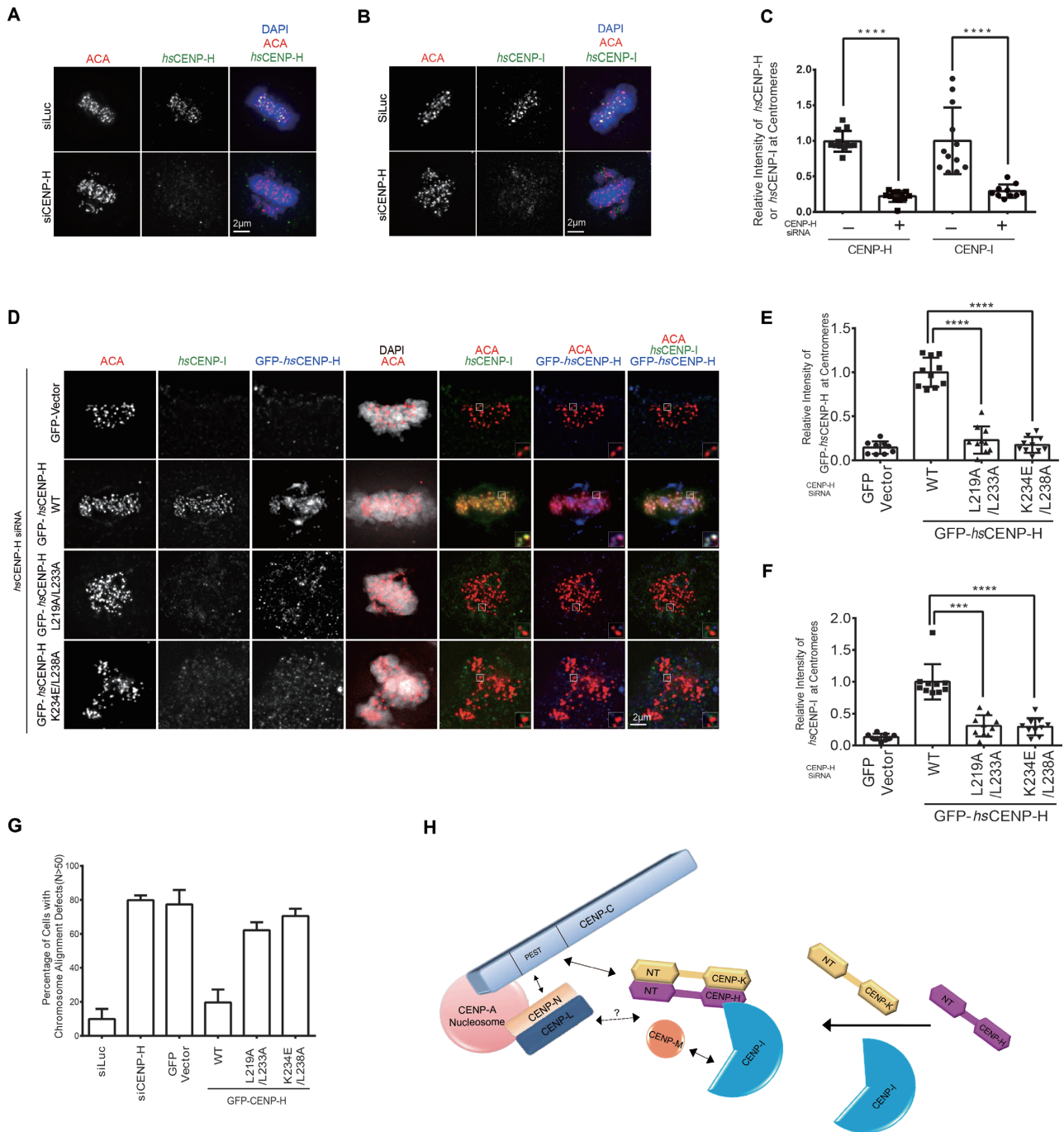


Figure 5. Mutating the interface residues of CENP-H abolished its centromeric localization and induced massive chromosome mis-alignment. (A and B) Representative images of interphase HeLa Tet-On cells treated with Luciferase and *hsCENP-H* siRNA. The centromeric localization of endogenous *hsCENP-H* (A) or *hsCENP-I* (B) was detected by indicated antibody. (C) Quantification of the centromeric intensities of *hsCENP-H* in (A) and *hsCENP-I* in (B) normalized to the ones of ACA (CREST). At least 10 cells (20 kinetochores per cell) were quantified for each condition. Mean ± SD (standard deviation) was shown here. (D) HeLa Tet-On cells transfected with RNAi-resistant GFP-*hsCENP-I* WT, L219A/L233A and K234E/L238A were treated with *hsCENP-H* siRNA. Cells were briefly treated with MG132 before subjected to staining. Representative images of mitotic cells were shown here. (E and F) Quantification of the centromeric intensities of GFP-*hsCENP-H* (E) and *hsCENP-I* (F) normalized to the ones of ACA (CREST) kinetochore signal. At least 10 cells (20 kinetochores per cell) were quantified for each condition. Mean ± SD was shown here. (G) Quantification of mitotic cells with chromosome alignment defects for the experiments in (A, B and D). At least 50 cells were counted for each condition. Mean ± SD was shown here. (H) The assembly mode of CENP-H/I/K ternary complex among the CCAN. CENP-A directly recruits CENP-C and CENP-L/N to the centromere (25,26,29,32,43,44). Both CENP-C and CENP-L/N are required for kinetochore recruitment of CENP-H/I/K/M (32,43). CENP-C directly interacts with CENP-H/K, not CENP-I (32). How CENP-L/N interacts with CENP-H/I/K/M remains unclear. CENP-H/K form a heterodimer through both N-termini and C-termini, and the C-termini of CENP-H/K heterodimer binds to N-terminus of CENP-I. CENP-M directly interacts with the C-terminus of CENP-I and integrate into the CENP-H/I/K/M complex (39).

CENP-M directly interact with the C-terminus of CENP-I and integrate into the CENP-H/I/K/M complex (39). Our findings demonstrate that CENP-H/K form a heterodimer through both N-termini and C-termini, and the C-termini of CENP-H/K heterodimer binds to the N-terminus of CENP-I (Figure 5H). The roles of the central regions of CENP-H/K complex are currently unclear. They might be involved in the ternary complex formation or interacting with other CCAN components. In comparison with the previously co-linear interaction model based on the low-resolution structure (39), our results suggest a slightly distinct model explaining how the CENP-H/I/K complex is assembled. The difference might be due to divergent experimental conditions or species difference. Alternatively, it might reflect a possibility that the CENP-H/I/K complex exists in distinct conformations in cells.

Although loss of each component of the CENP-H/I/K complex has been shown to cause chromosome mis-segregation (12,13,16,43,47,48), the exact roles of the individual protein–protein interfaces in chromosome segregation had never been investigated. Our results indicate that any flaw in these interactions can result in massive chromosome mis-alignment, suggestive of importance of the integrity of the CENP-H/I/K complex in proper kinetochore function. The CENP-H/I/K/M complex has also been shown to interact with CENP-T/W and CENP-L/N (32,43). In future, it will be important to identify the interacting surfaces and evaluate their functional importance in kinetochore function.

DATA AVAILABILITY

Structures and crystallographic data have been deposited at the wwPDB: 5Z07 (*ct*CENP-I^{NT}) and 5Z08 (*ct*CENP-I^{NT} complex with *th*CENP-H^{CT}/K^{CT}).

SUPPLEMENTARY DATA

[Supplementary Data](#) are available at NAR Online.

ACKNOWLEDGEMENTS

We appreciate Dr Hongtao Yu, Dr Xuewu Zhang and Dr Xuelian Luo for providing important reagents and insightful suggestions, and thank the staff at beamline BL17U1 and BL18U1 of the Shanghai Synchrotron Radiation Facility (SSRF) for assistance with the data collection, Dr Mingzhu Wang and Dr Shilong Fan for on-site assistance and technical support. We also thank Xiaoxia Yu for analysis of the AUC data.

FUNDING

National Natural Science Foundation of China [31500629, 31470763, 31500605]; Natural Science Foundation of Hubei Province of China [2015CFB484, 2015CFB689]. Funding for open access charge: National Natural Science Foundation of China [31500629, 31470763, 31500605]; Natural Science Foundation of Hubei Province of China [2015CFB484, 2015CFB689].

Conflict of interest statement. None declared.

REFERENCES

- Cleveland,D.W., Mao,Y. and Sullivan,K.F. (2003) Centromeres and kinetochores: from epigenetics to mitotic checkpoint signaling. *Cell*, **112**, 407–421.
- Cheeseman,I.M. (2014) The kinetochore. *Cold Spring Harb. Perspect. Biol.*, **6**, a015826.
- Fukagawa,T. and Earnshaw,W.C. (2014) The centromere: chromatin foundation for the kinetochore machinery. *Dev. Cell*, **30**, 496–508.
- Musacchio,A. and Desai,A. (2017) A molecular view of kinetochore assembly and function. *Biology (Basel)*, **6**, 5.
- Akiyoshi,B., Sarangapani,K.K., Powers,A.F., Nelson,C.R., Reichow,S.L., Arellano-Santoyo,H., Gonen,T., Ranish,J.A., Asbury,C.L. and Biggins,S. (2010) Tension directly stabilizes reconstituted kinetochore-microtubule attachments. *Nature*, **468**, 576–579.
- Bharadwaj,R., Qi,W. and Yu,H. (2004) Identification of two novel components of the human NDC80 kinetochore complex. *J. Biol. Chem.*, **279**, 13076–13085.
- Cheeseman,I.M., Chappie,J.S., Wilson-Kubalek,E.M. and Desai,A. (2006) The conserved KMN network constitutes the core microtubule-binding site of the kinetochore. *Cell*, **127**, 983–997.
- Cheeseman,I.M. and Desai,A. (2008) Molecular architecture of the kinetochore-microtubule interface. *Nat. Rev. Mol. Cell Biol.*, **9**, 33–46.
- Cheeseman,I.M., Niessen,S., Anderson,S., Hyndman,F., Yates,J.R. 3rd, Oegema,K. and Desai,A. (2004) A conserved protein network controls assembly of the outer kinetochore and its ability to sustain tension. *Genes Dev.*, **18**, 2255–2268.
- De Wulf,P., McAinsh,A.D. and Sorger,P.K. (2003) Hierarchical assembly of the budding yeast kinetochore from multiple subcomplexes. *Genes Dev.*, **17**, 2902–2921.
- Nekrasov,V.S., Smith,M.A., Peak-Chew,S. and Kilmartin,J.V. (2003) Interactions between centromere complexes in *Saccharomyces cerevisiae*. *Mol. Biol. Cell*, **14**, 4931–4946.
- Foltz,D.R., Jansen,L.E.T., Black,B.E., Bailey,A.O., Yates,J.R. and Cleveland,D.W. (2006) The human CENP-A centromeric nucleosome-associated complex. *Nat. Cell Biol.*, **8**, 458–469.
- Hori,T., Amano,M., Suzuki,A., Backer,C.B., Welburn,J.P., Dong,Y., McEwen,B.F., Shang,W.H., Suzuki,E., Okawa,K. *et al.* (2008) CCAN makes multiple contacts with centromeric DNA to provide distinct pathways to the outer kinetochore. *Cell*, **135**, 1039–1052.
- McKinley,K.L. and Cheeseman,I.M. (2016) The molecular basis for centromere identity and function. *Nat. Rev. Mol. Cell Biol.*, **17**, 16–29.
- Nagpal,H. and Fukagawa,T. (2016) Kinetochore assembly and function through the cell cycle. *Chromosoma*, **125**, 645–659.
- Okada,M., Cheeseman,I.M., Hori,T., Okawa,K., McLeod,I.X., Yates,J.R., Desai,A. and Fukagawa,T. (2006) The CENP-H-I complex is required for the efficient incorporation of newly synthesized CENP-A into centromeres. *Nat. Cell Biol.*, **8**, 446–457.
- Amano,M., Suzuki,A., Hori,T., Backer,C., Okawa,K., Cheeseman,I.M. and Fukagawa,T. (2009) The CENP-S complex is essential for the stable assembly of outer kinetochore structure. *J. Cell Biol.*, **186**, 173–182.
- Izuta,H., Ikeno,M., Suzuki,N., Tomonaga,T., Nozaki,N., Obuse,C., Kisu,Y., Goshima,N., Nomura,F., Nomura,N. *et al.* (2006) Comprehensive analysis of the ICEN (Interphase Centromere Complex) components enriched in the CENP-A chromatin of human cells. *Genes Cells*, **11**, 673–684.
- Westermann,S. and Schleiffer,A. (2013) Family matters: structural and functional conservation of centromere-associated proteins from yeast to humans. *Trends Cell Biol.*, **23**, 260–269.
- Black,B.E., Brock,M.A., Bedard,S., Woods,V.L. and Cleveland,D.W. (2007) An epigenetic mark generated by the incorporation of CENP-A into centromeric nucleosomes. *Proc. Natl. Acad. Sci. U.S.A.*, **104**, 5008–5013.
- Conde e Silva,N., Black,B.E., Sivobol,A., Filipinski,J., Cleveland,D.W. and Prunell,A. (2007) CENP-A-containing nucleosomes: easier disassembly versus exclusive centromeric localization. *J. Mol. Biol.*, **370**, 555–573.
- Palmer,D., O’Day,K., Wener,M., Andrews,B. and Margolis,R. (1987) A 17-kD centromere protein (CENP-A) copurifies with nucleosome core particles and with histones. *J. Cell Biol.*, **104**, 805–815.

23. Sullivan, K.F., Hechenberger, M. and Masri, K. (1994) Human CENP-A contains a histone H3 related histone fold domain that is required for targeting to the centromere. *J. Cell Biol.*, **127**, 581–592.
24. Carroll, C.W., Milks, K.J. and Straight, A.F. (2010) Dual recognition of CENP-A nucleosomes is required for centromere assembly. *J. Cell Biol.*, **189**, 1143–1155.
25. Carroll, C.W., Silva, M.C.C., Godek, K.M., Jansen, L.E.T. and Straight, A.F. (2009) Centromere assembly requires the direct recognition of CENP-A nucleosomes by CENP-N. *Nat. Cell Biol.*, **11**, 896–902.
26. Chittori, S., Hong, J., Saunders, H., Feng, H., Ghirlando, R., Kelly, A.E., Bai, Y. and Subramaniam, S. (2018) Structural mechanisms of centromeric nucleosome recognition by the kinetochore protein CENP-N. *Science*, **359**, 339–343.
27. Falk, S.J., Guo, L.Y., Sekulic, N., Smoak, E.M., Mani, T., Logsdon, G.A., Gupta, K., Jansen, L.E.T., Van Duyne, G.D., Vinogradov, S.A. *et al.* (2015) Chromosomes. CENP-C reshapes and stabilizes CENP-A nucleosomes at the centromere. *Science*, **348**, 699–703.
28. Guse, A., Carroll, C.W., Moree, B., Fuller, C.J. and Straight, A.F. (2011) In vitro centromere and kinetochore assembly on defined chromatin templates. *Nature*, **477**, 354–358.
29. Kato, H., Jiang, J.S., Zhou, B.R., Rozendaal, M., Feng, H.Q., Ghirlando, R., Xiao, T.S., Straight, A.F. and Bai, Y.W. (2013) A conserved mechanism for Centromeric Nucleosome Recognition by Centromere Protein CENP-C. *Science*, **340**, 1110–1113.
30. Gascoigne, K.E., Takeuchi, K., Suzuki, A., Hori, T., Fukagawa, T. and Cheeseman, I.M. (2011) Induced ectopic kinetochore assembly bypasses the requirement for CENP-A nucleosomes. *Cell*, **145**, 410–422.
31. Huis In 't Veld, P.J., Jeganathan, S., Petrovic, A., Singh, P., John, J., Krenn, V., Weissmann, F., Bange, T. and Musacchio, A. (2016) Molecular basis of outer kinetochore assembly on CENP-T. *Elife*, **5**, e21007.
32. Klare, K., Weir, J.R., Basilico, F., Zimniak, T., Massimiliano, L., Ludwigs, N., Herzog, F. and Musacchio, A. (2015) CENP-C is a blueprint for constitutive centromere-associated network assembly within human kinetochores. *J. Cell Biol.*, **210**, 11–22.
33. Malvezzi, F., Litos, G., Schleiffer, A., Heuck, A., Mechtler, K., Clausen, T. and Westermann, S. (2013) A structural basis for kinetochore recruitment of the Ndc80 complex via two distinct centromere receptors. *EMBO J.*, **32**, 409–423.
34. Nishino, T., Rago, F., Hori, T., Tomii, K., Cheeseman, I.M. and Fukagawa, T. (2013) CENP-T provides a structural platform for outer kinetochore assembly. *EMBO J.*, **32**, 424–436.
35. 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.
36. Petrovic, A., Keller, J., Liu, Y.H., Overlack, K., John, J., Dimitrova, Y.N., Jenni, S., van Gerwen, S., Stege, P., Wohlgemuth, S. *et al.* (2016) Structure of the MIS12 complex and molecular basis of its interaction with CENP-C at human kinetochores. *Cell*, **167**, 1028–1040.
37. Przewloka, M.R., Venkei, Z., Bolanos-Garcia, V.M., Debski, J., Dadlez, M. and Glover, D.M. (2011) CENP-C is a structural platform for kinetochore assembly. *Curr. Biol.*, **21**, 399–405.
38. Screpanti, E., De Antoni, A., Alushin, G.M., Petrovic, A., Melis, T., Nogales, E. and Musacchio, A. (2011) Direct binding of Cenp-C to the Mis12 complex joins the inner and outer kinetochore. *Curr. Biol.*, **21**, 391–398.
39. 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.
40. Fang, J.N., Liu, Y.T., Wei, Y., Deng, W.Q., Yu, Z.L., Huang, L., Teng, Y., Yao, T., You, Q.L., Ruan, H.H. *et al.* (2015) Structural transitions of centromeric chromatin regulate the cell cycle-dependent recruitment of CENP-N. *Gene Dev.*, **29**, 1058–1073.
41. Hara, M. and Fukagawa, T. (2018) Kinetochore assembly and disassembly during mitotic entry and exit. *Curr. Opin. Cell Biol.*, **52**, 73–81.
42. Hinshaw, S.M. and Harrison, S.C. (2018) Kinetochore function from the bottom up. *Trends Cell Biol.*, **28**, 22–33.
43. McKinley, K.L., Sekulic, N., Guo, L.Y., Tsinman, T., Black, B.E. and Cheeseman, I.M. (2015) The CENP-L-N complex forms a critical node in an integrated meshwork of interactions at the Centromere-Kinetochore interface. *Mol. Cell*, **60**, 886–898.
44. Pentakota, S., Zhou, K., Smith, C., Maffini, S., Petrovic, A., Morgan, G.P., Weir, J.R., Vetter, I.R., Musacchio, A. and Luger, K. (2017) Decoding the centromeric nucleosome through CENP-N. *Elife*, **6**, e33442.
45. Pesenti, M.E., Weir, J.R. and Musacchio, A. (2016) Progress in the structural and functional characterization of kinetochores. *Curr. Opin. Struct. Biol.*, **37**, 152–163.
46. Tian, T., Li, X.R., Liu, Y.Y., Wang, C.L., Liu, X., Bi, G.Q., Zhang, X., Yao, X.B., Zhou, Z.H. and Zang, J.Y. (2018) Molecular basis for CENP-N recognition of CENP-A nucleosome on the human kinetochore. *Cell Res.*, **28**, 374–378.
47. Cheeseman, I.M., Hori, T., Fukagawa, T. and Desai, A. (2008) KNL1 and the CENP-H/I/K complex coordinately direct kinetochore assembly in vertebrates. *Mol. Biol. Cell*, **19**, 587–594.
48. Kim, S. and Yu, H.T. (2015) Multiple assembly mechanisms anchor the KMN spindle checkpoint platform at human mitotic kinetochores. *J. Cell Biol.*, **208**, 181–196.
49. Minor, W., Cymborowski, M., Otwinowski, Z. and 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*, **62**, 859–866.
50. Adams, P.D., Afonine, P.V., Bunkoczi, G., Chen, V.B., Davis, I.W., Echols, N., Headd, J.J., Hung, L.W., Kapral, G.J., Grosse-Kunstleve, R.W. *et al.* (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.*, **66**, 213–221.
51. Emsley, P., Lohkamp, B., Scott, W.G. and Cowtan, K. (2010) Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.*, **66**, 486–501.
52. Schuck, P. (2000) Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and lamm equation modeling. *Biophys. J.*, **78**, 1606–1619.
53. Finn, R.D., Coggill, P., Eberhardt, R.Y., Eddy, S.R., Mistry, J., Mitchell, A.L., Potter, S.C., Punta, M., Qureshi, M., Sangrador-Vegas, A. *et al.* (2016) The Pfam protein families database: towards a more sustainable future. *Nucleic Acids Res.*, **44**, D279–D285.
54. Jones, P., Binns, D., Chang, H.Y., Fraser, M., Li, W., McAnulla, C., McWilliam, H., Maslen, J., Mitchell, A., Nuka, G. *et al.* (2014) InterProScan 5: genome-scale protein function classification. *Bioinformatics*, **30**, 1236–1240.