Molecular Cell, Volume *82*

Supplemental information

Structure of the human inner kinetochore CCAN complex and its significance for human centromere organization

Marion E. Pesenti, Tobias Raisch, Duccio Conti, Kai Walstein, Ingrid Hoffmann, Dorothee Vogt, Daniel Prumbaum, Ingrid R. Vetter, Stefan Raunser, and Andrea Musacchio

Molecular Cell, Volume

Supplemental information

Structure of the human inner kinetochore CCAN complex and its significance for human centromere organization

Marion E. Pesenti, Tobias Raisch, Duccio Conti, Kai Walstein, Ingrid Hoffmann, Dorothee Vogt, Daniel Prumbaum, Ingrid R. Vetter, Stefan Raunser & Andrea Musacchio

Table S1 (related to Figure 1)

Cryo-EM data collection, refinement and validation statistics

	CCAN Krios2 EMDB-14098 PDB 7QOO ("dataset I")	CCAN Krios1 EMDB-14099 ("dataset II")
Data collection and processing		
Magnification	130000x	105000x
Voltage (kV)	300	300
Electron exposure (e-/Å^2)	76.8	55.8
Defocus range (μm)	-0.6 to -1.2	-0.6 to -1.2
Pixel size (native/super-resolution) (Å)	0.7/0.35	0.68/0.34
Symmetry imposed	C1	C1
Initial particle images (no.)	140910	233598
Final particle images (no.)	22853	44216
Map resolution (Å)	4.6	10
FSC threshold	0.143	0.143
Refinement		
Initial model used:	Alphafold,6EQT,4WAU,6WUC, 6QLD, 6QLE	
Model resolution (Å)	3.8	
FSC threshold	0.143	
Model composition		
Non-hydrogen atoms	24952	
Protein residues	3086	
B-factors (Å^2)		
Protein	47.8	
R.m.s. deviations		
Bond length (Å)	0.0035	
Bond angles (°)	0.84	
Validation		
MolProbity score	1.61	
Clashscore	5.14	
Poor rotamers (%)	1.53	
Ramachandran plot		
Favored (%)	96.8	
Allowed (%)	3.13	
Disallowed (%)	0.07	

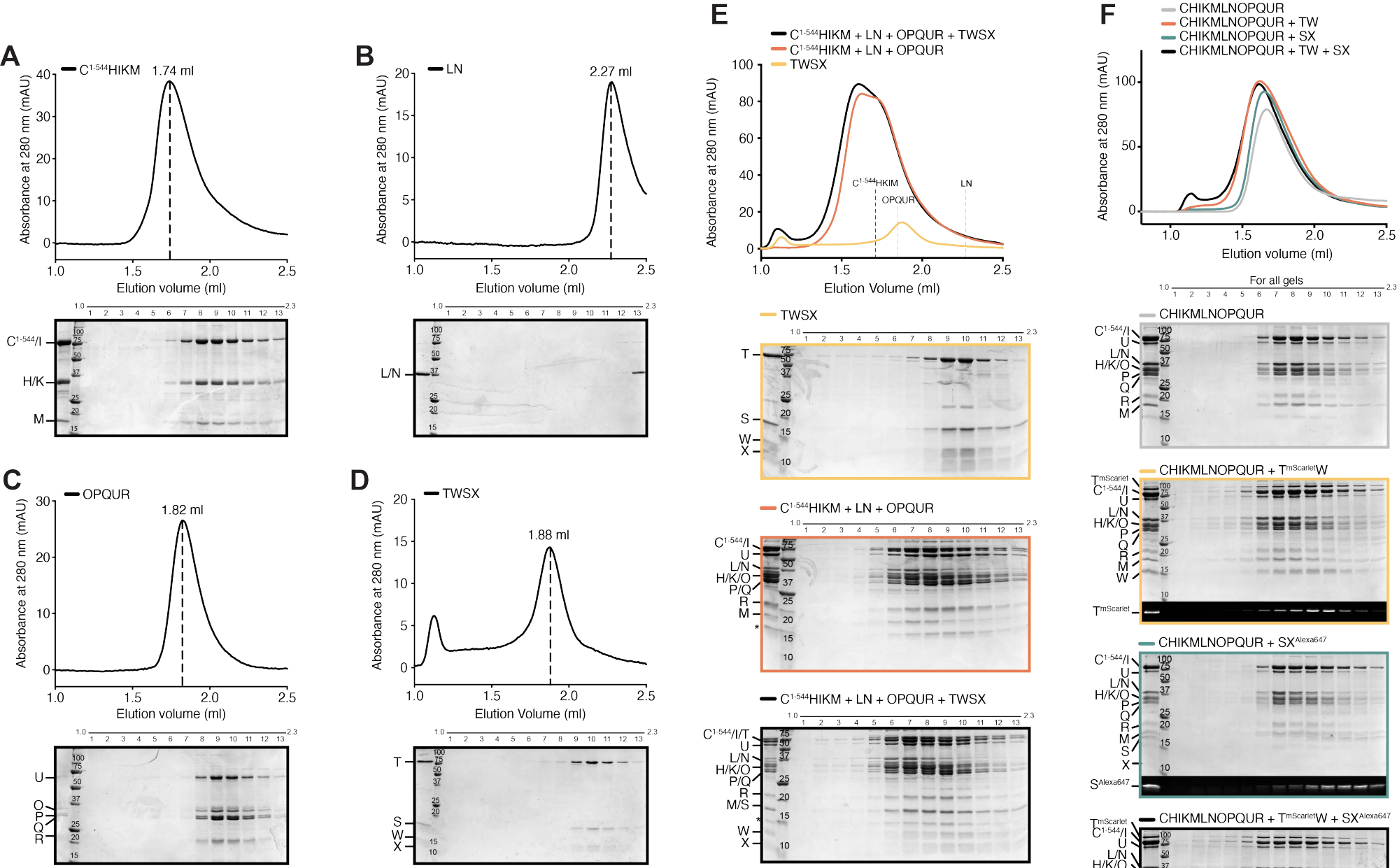


Figure S1 – Related to Figure 1

Figure S1 (related to Figure 1). Biochemical reconstitution of human CCAN

Size-exclusion chromatography on a Superose 6 5/150 column of the indicated CCAN complexes and subcomplexes. **A**) CENP-C$^{1-544}$HIKM. **B**) CENP-LN. **C**) CENP-OPQUR. **D**) CENP-TWSX. **E**) Comparison of elution profiles of CENP-TWSX, CENP-12, and CENP-16. **F**) CENP-TW binds CENP-12, but CENP-SX requires CENP-TW to interact with CENP-12. Gels and chromatograms in panels A-D also appear in Figure S2 (representative gels). Gels and chromatograms in panels A-C also appear as controls in Figure S12B-D.

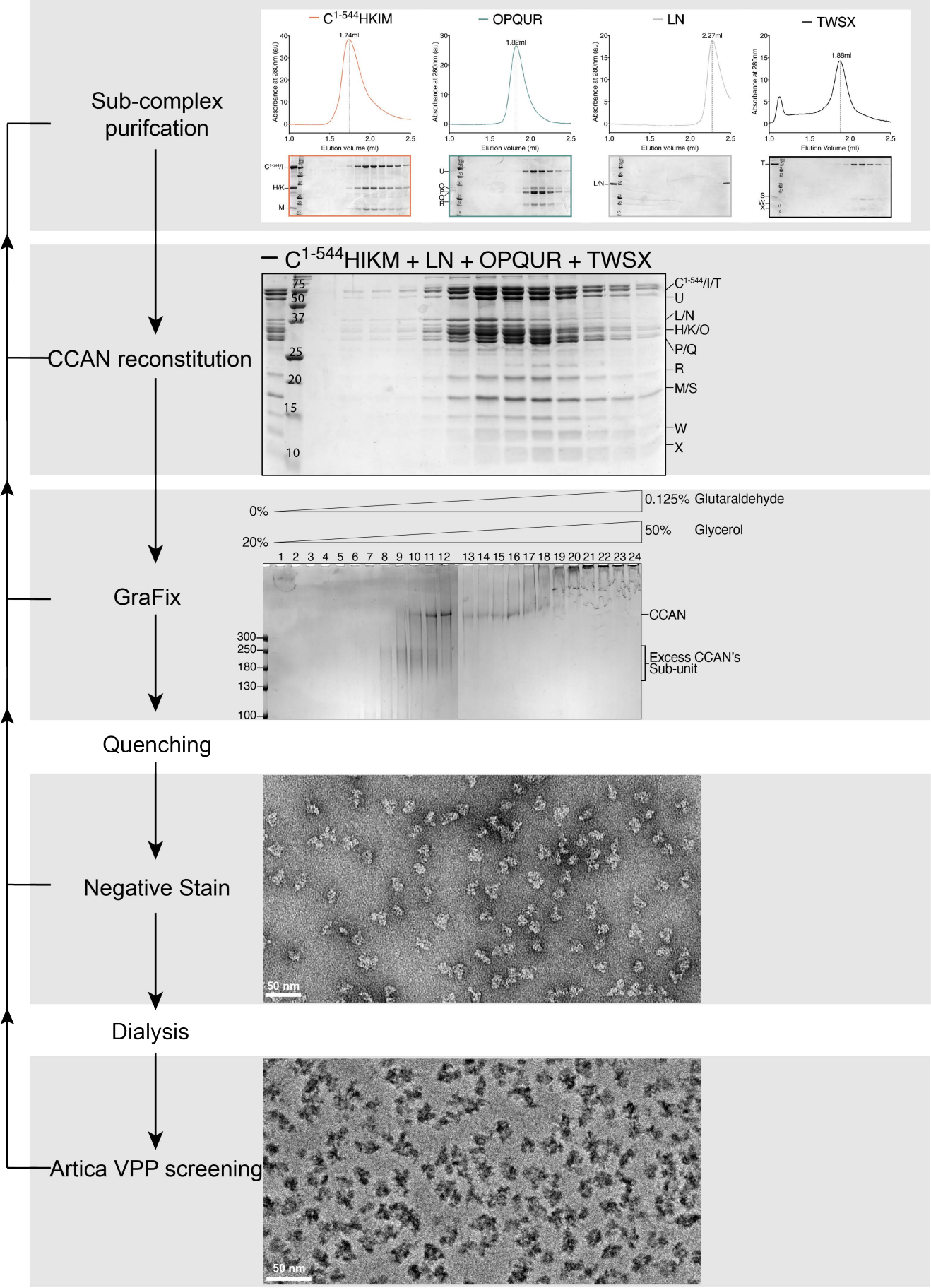


Figure S2 – Related to Figure 1

Figure S2 (related to Figure 1). Pipeline of sample preparation

The figure presents an outline of the various steps of sample preparation preceding high-resolution cryo-EM data collection. Chromatograms and gels in the upper panel are also displayed in Figure S1.

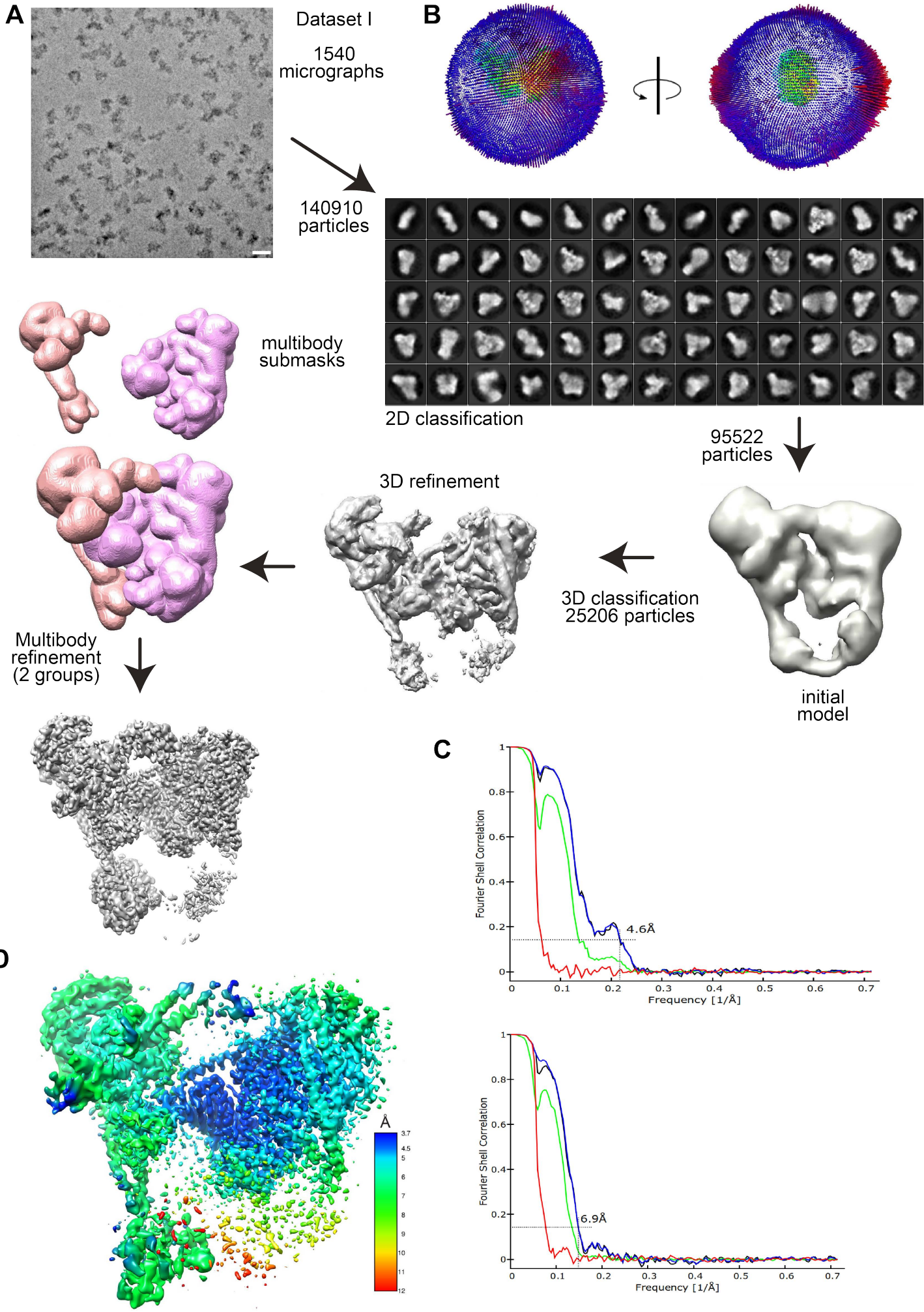


Figure S3 – Related to Figure 1

Figure S3 (related to Figure 1). Data processing flowchart for Dataset I

A) Processing flowchart for the high-resolution dataset I including an exemplary micrograph (scale bar = 20 nm) and a subset of selected 2D classes of CENP-16. The last (grey) map shows the final reconstruction and was obtained by combining the two focused maps from multibody refinement (using the 'vop maximum' command in UCSF Chimera). **B**) Angular distribution of the particles shown in two positions rotated 90° to each other. **C**) Fourier shell correlation (FSC) plots between two independent half-maps for each of the two bodies used in the multibody refinement procedure, according to the FSC=0.143 criterion. The dashed line indicates the 0.143 FSC criterion. red: phase-randomized map, green: unmasked map, blue: masked map, black: corrected map. **D**) Local resolution estimates by RELION for dataset I plotted on the two multibody reconstructions in a rainbow-colored gradient from blue (3.7 Å) to red (12 Å). The map differs from the final reconstruction depicted in panel A since the latter shows the combined sub-maps, further modified by DeepEMhancer.

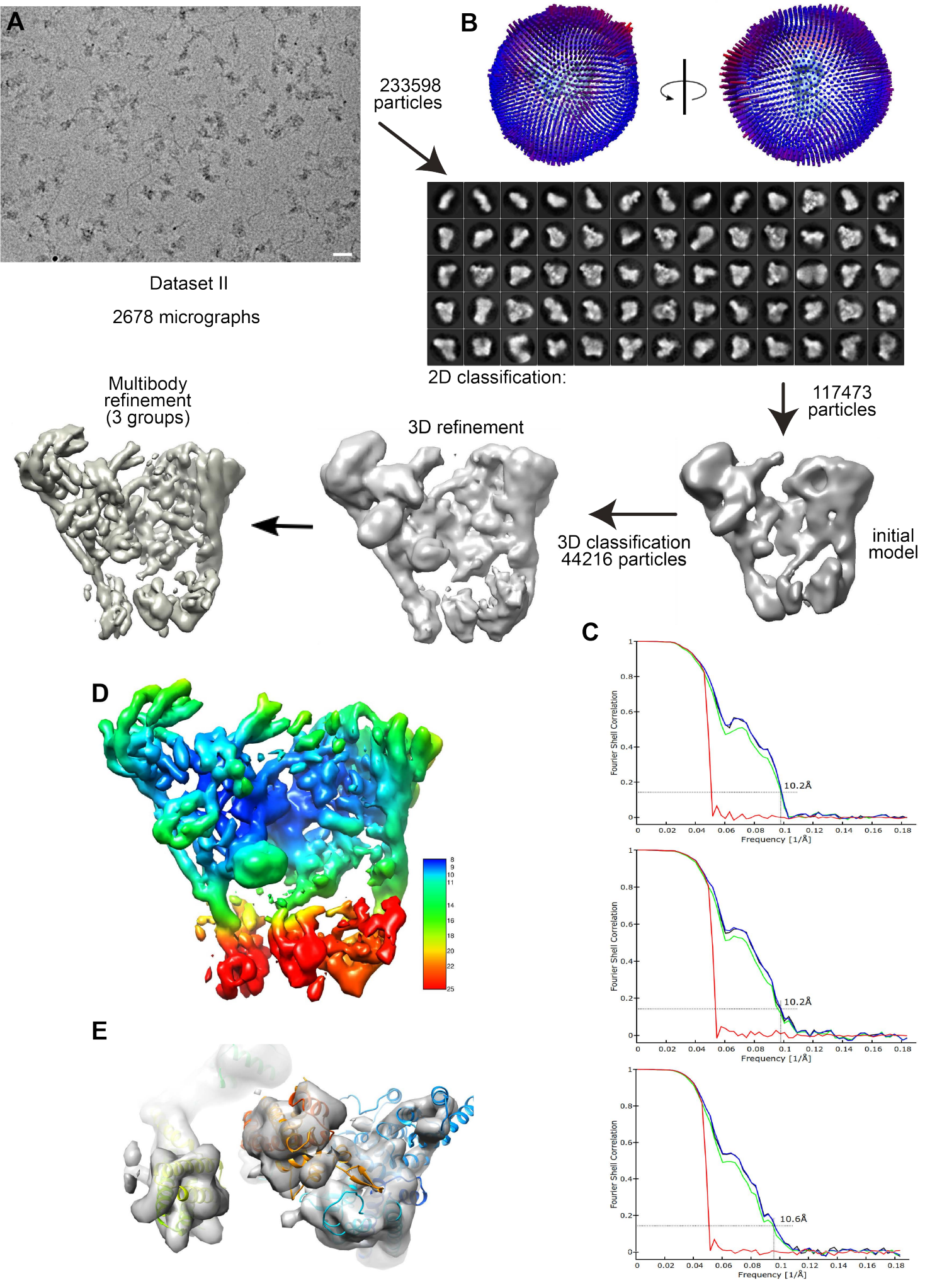


Figure S4 – Related to Figure 1

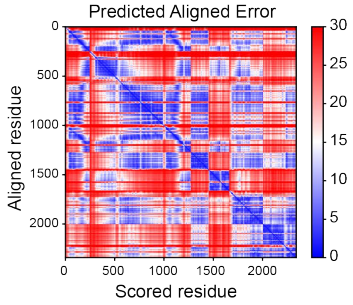
Figure S4 (related to Figure 1). Data processing flowchart for Dataset II

A) Processing flowchart for the low-resolution dataset II including an exemplary micrograph and a subset of selected 2D classes of CENP-14. **B**) Angular distribution of the particles shown in two positions rotated 90° to each other. **C**) Fourier shell correlation (FSC) plots between two independent half-maps for each of the three bodies used in the last multibody-refinement, indicating the resolution of the three groups in the multibody refinement according to the FSC=0.143 criterion. The dashed line indicates the 0.143 FSC criterion. red: phase-randomized map, green: unmasked map, blue: masked map, black: corrected map. **D**) Local resolution estimates by RELION plotted on the multibody reconstructions in a rainbow-colored gradient from blue (8 Å) to red (25 Å). **E**) Representative density of CENP-QU head (left), the CENP-TW complex (middle) and the HIK head (right), as seen from the "bottom" towards the top of the CCAN complex in D).

A

HIKM + TW + LN

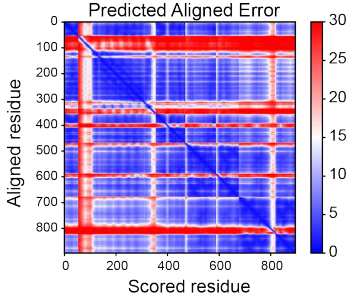
Chain	Residue		Cumulative	
H	1	247	1	247
I	1	756	248	1003
K	1	269	1004	1272
M	1	180	1273	1452
T	443	561	1453	1571
W	1	88	1572	1659
L	1	344	1660	2003
N	1	339	2004	2342



B

HIK head + TWSX

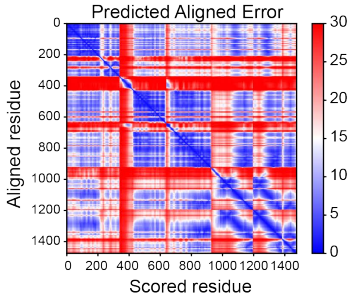
Chain	Residue		Cumulative	
H	193	247	1	55
I	1	300	56	355
K	155	269	356	470
T	443	561	471	589
W	1	88	590	677
S	1	138	678	815
X	1	81	816	896



C

N + OPQU

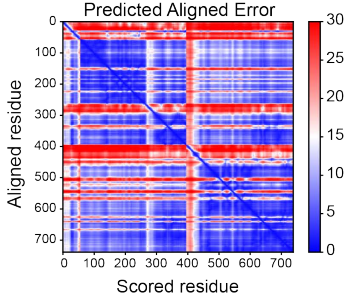
Chain	Residue		Cumulative	
N	1	339	1	339
O	1	300	340	639
P	1	288	640	927
Q	1	268	928	1195
U	232	418	1196	1382
R	66	159	1383	1476



D

C$^{275-330}$ + NL

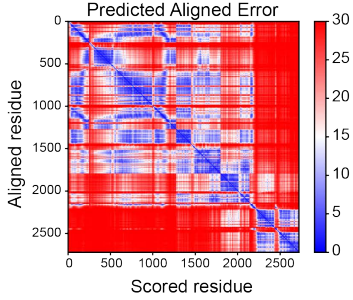
Chain	Residue		Cumulative	
C	275	330	1	55
N	1	339	56	394
L	1	344	395	739



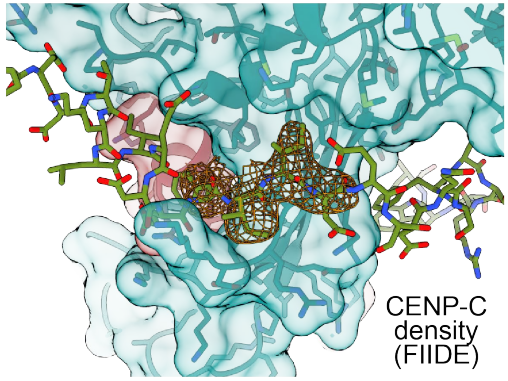
E

HIKM + LN + OP

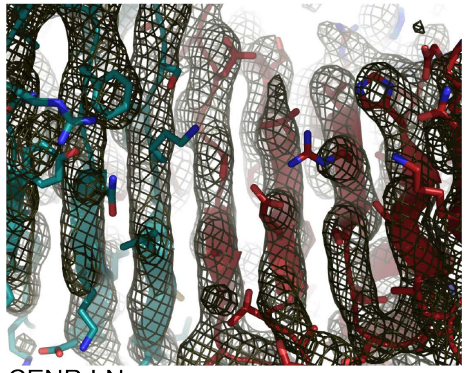
Chain	Residue		Cumulative	
H	1	247	1	247
I	1	756	248	1003
K	1	269	1004	1272
M	1	180	1273	1452
L	1	344	1453	1796
N	1	339	1797	2135
O	1	300	2136	2435
P	1	288	2436	2723

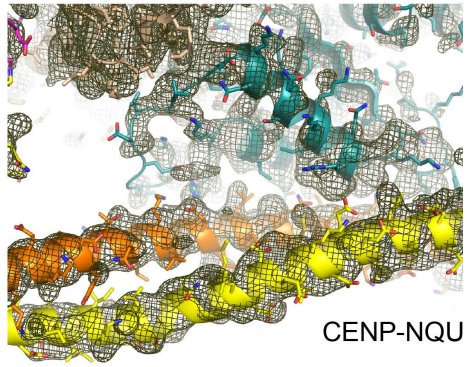


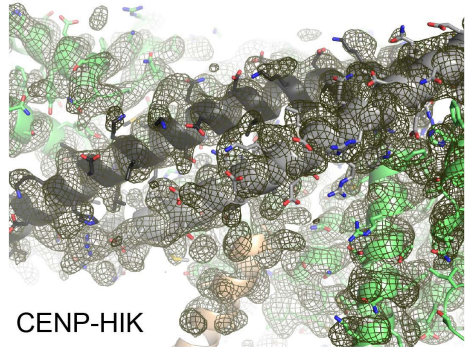
F

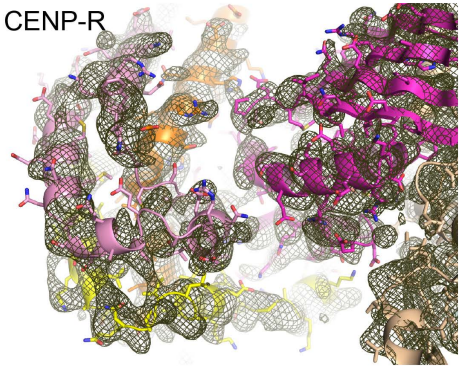


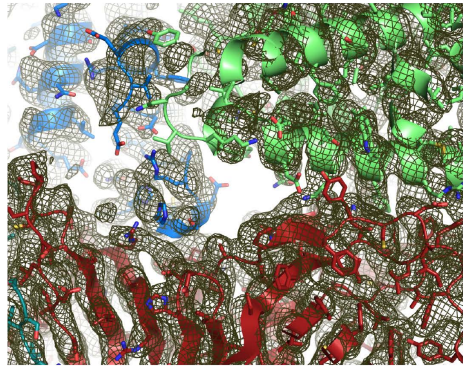
G











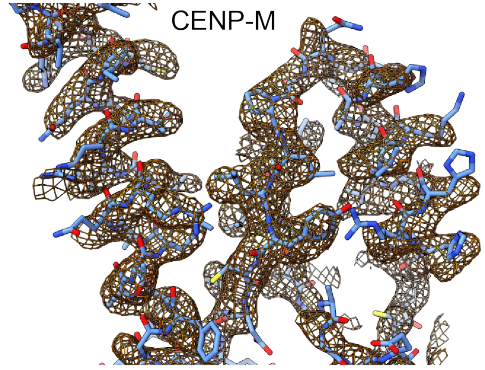
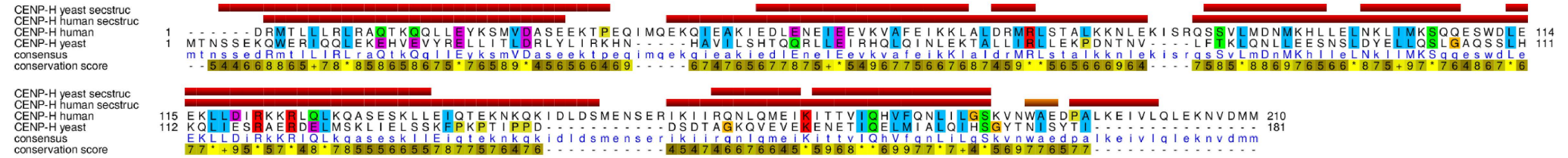


Figure S5 – Related to Figure 1

Figure S5 (related to Figure 1). AlphaFold2 model quality and close-ups of density maps

A-**E**) Predicted aligned error (PAE) plots for AF2 models of the indicated CCAN sub-complexes. Blue color indicates low error and high confidence of the relative positioning of the subdomains, red low confidence. **F**) Density for CENP-C 303-FIID-306 bound to CENP-N. The sequence assignment of CENP-C into this short stretch of density was based on a highly significant AF2 prediction for this interaction. **G**) Representative ap densities at the indicated subunits.

CENP-H



CENP-I

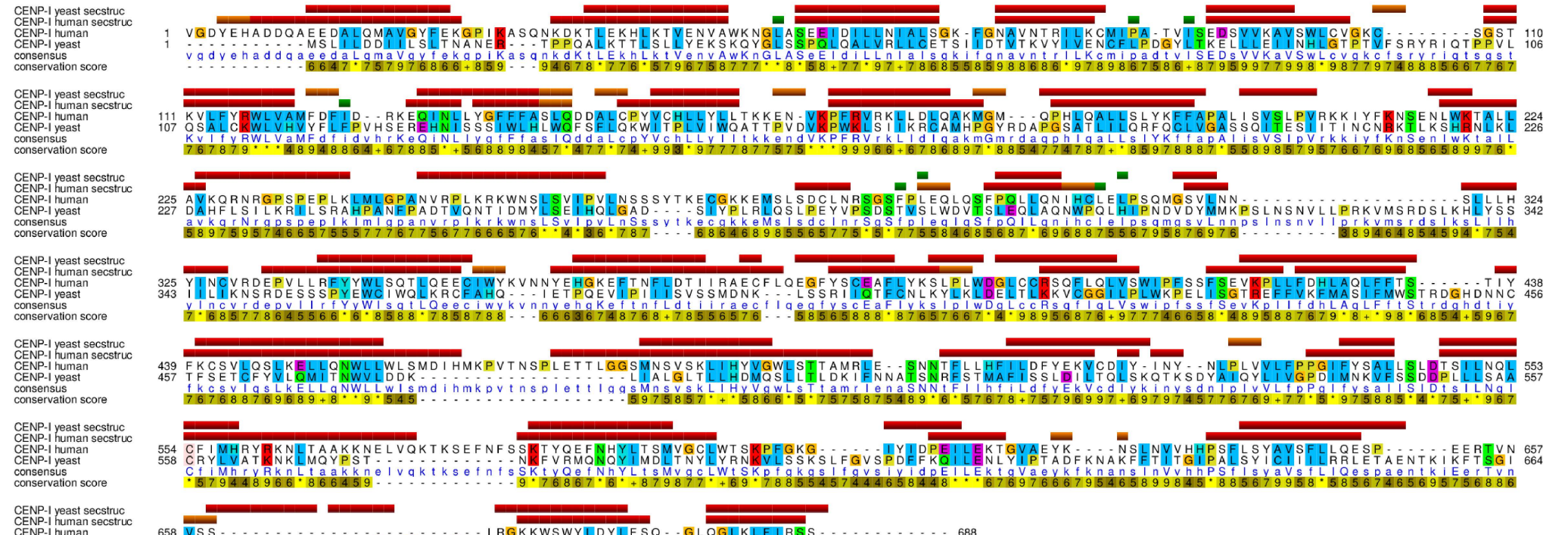


Figure S6 – Related to Figure 1

Figure S6 (related to Figure 1). Sequence alignments of human and yeast CENP-H and CENP-I

Secondary structure elements of yeast CCAN (PDB ID 6QLE) were aligned to the human CCAN structure by keeping the molecular shapes and relative orientations as similar as possible. Structure-based alignments were computed by Chimera. The secondary structures for the yeast proteins were assigned by DSSP (Kabsch and Sander, 1983) from the original yeast structures with PDB ID 6QLE and 6WUC. α-helices are shown in red, 3_{10} helices in orange, and β-sheets in green. Sequences are colored according to the ClustalX coloring scheme.

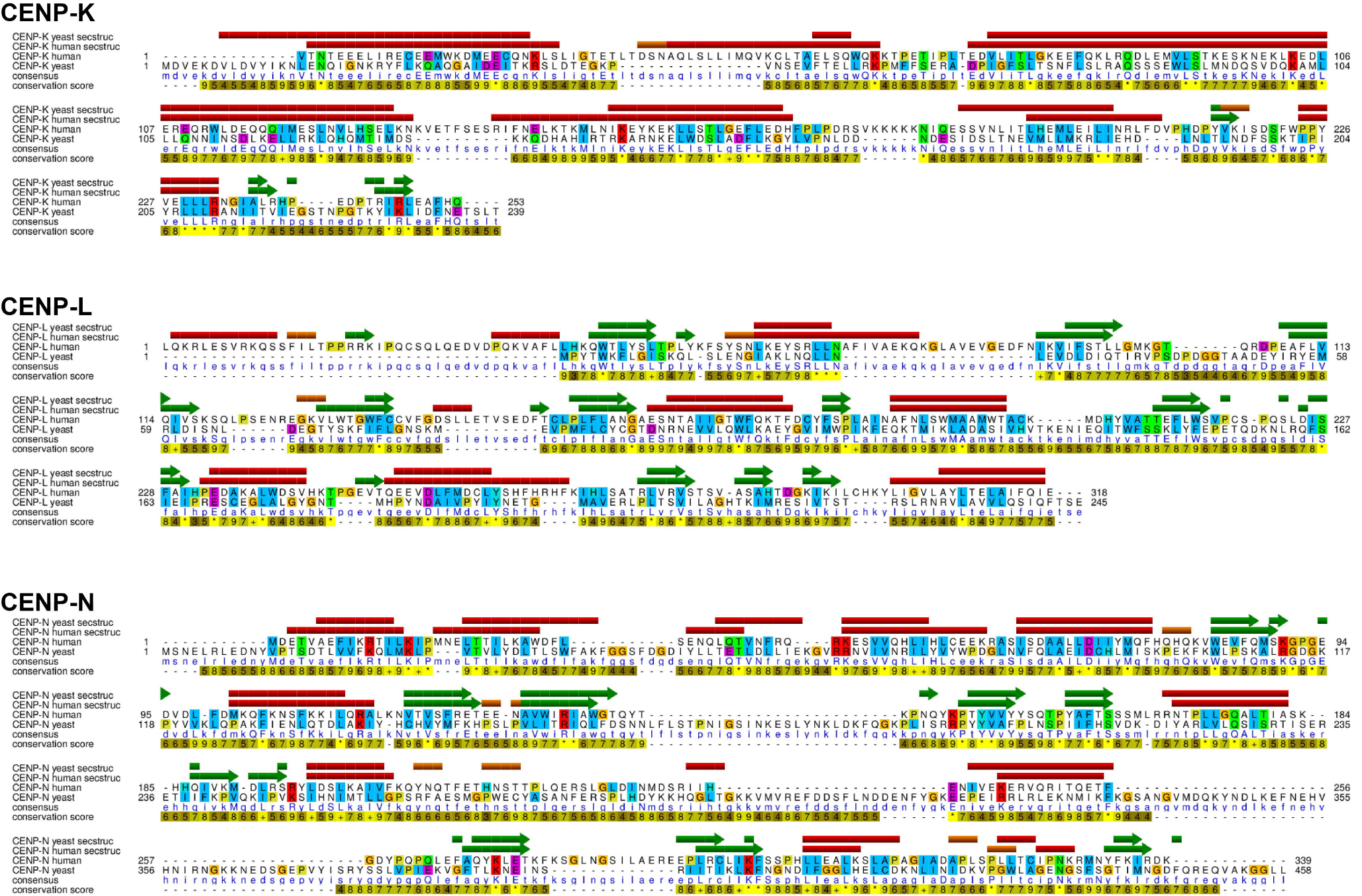


Figure S7 – Related to Figure 1

Figure S7 (related to Figure 1). Sequence alignments of human and yeast CENP-K, CENP-L, and CENP-N

For details, please refer to legend of Figure S6.

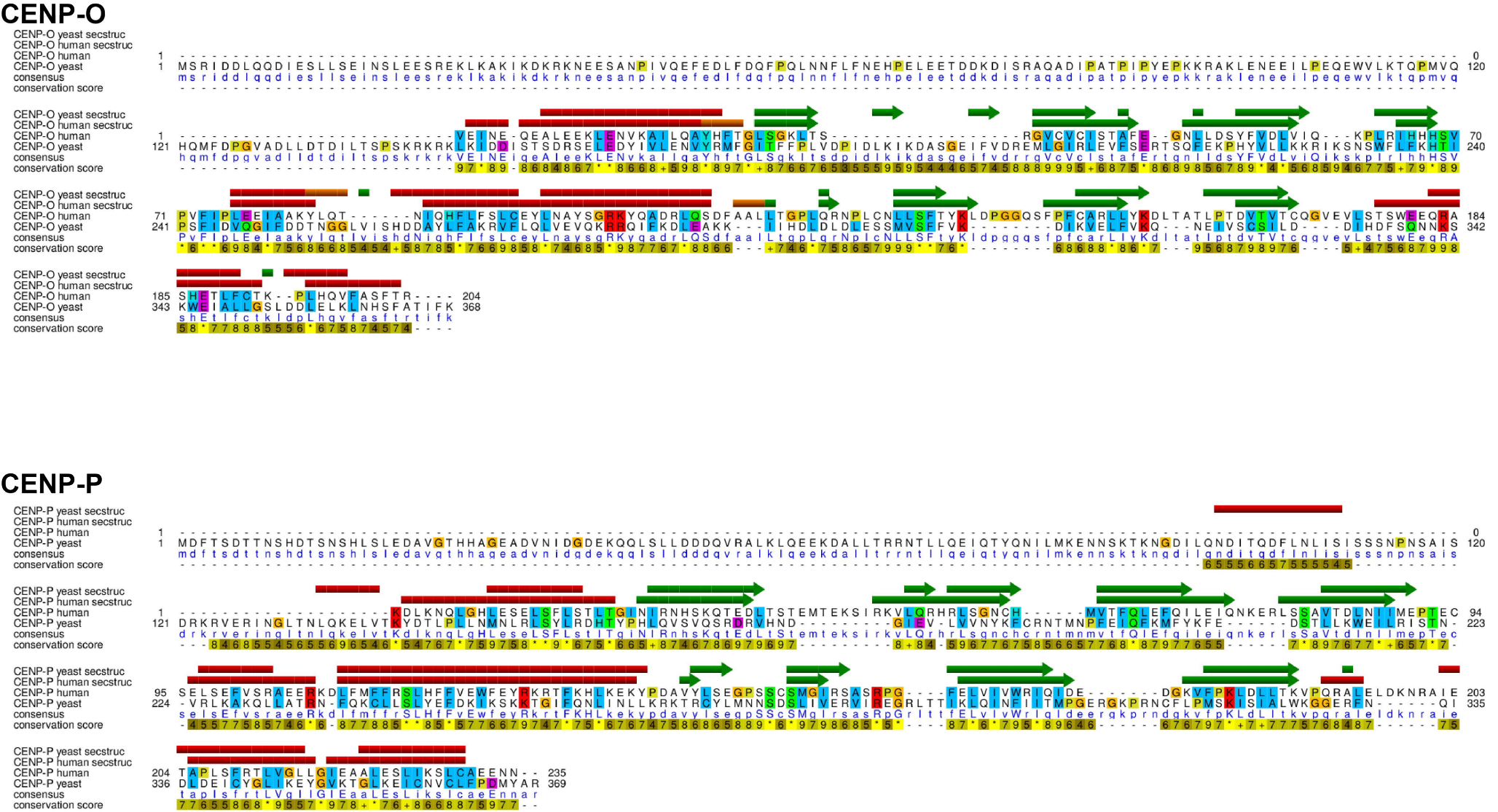


Figure S8 – Related to Figure 1

Figure S8 (related to Figure 1). Sequence alignments of human and yeast CENP-O and CENP-P

For details, please refer to legend of Figure S6.

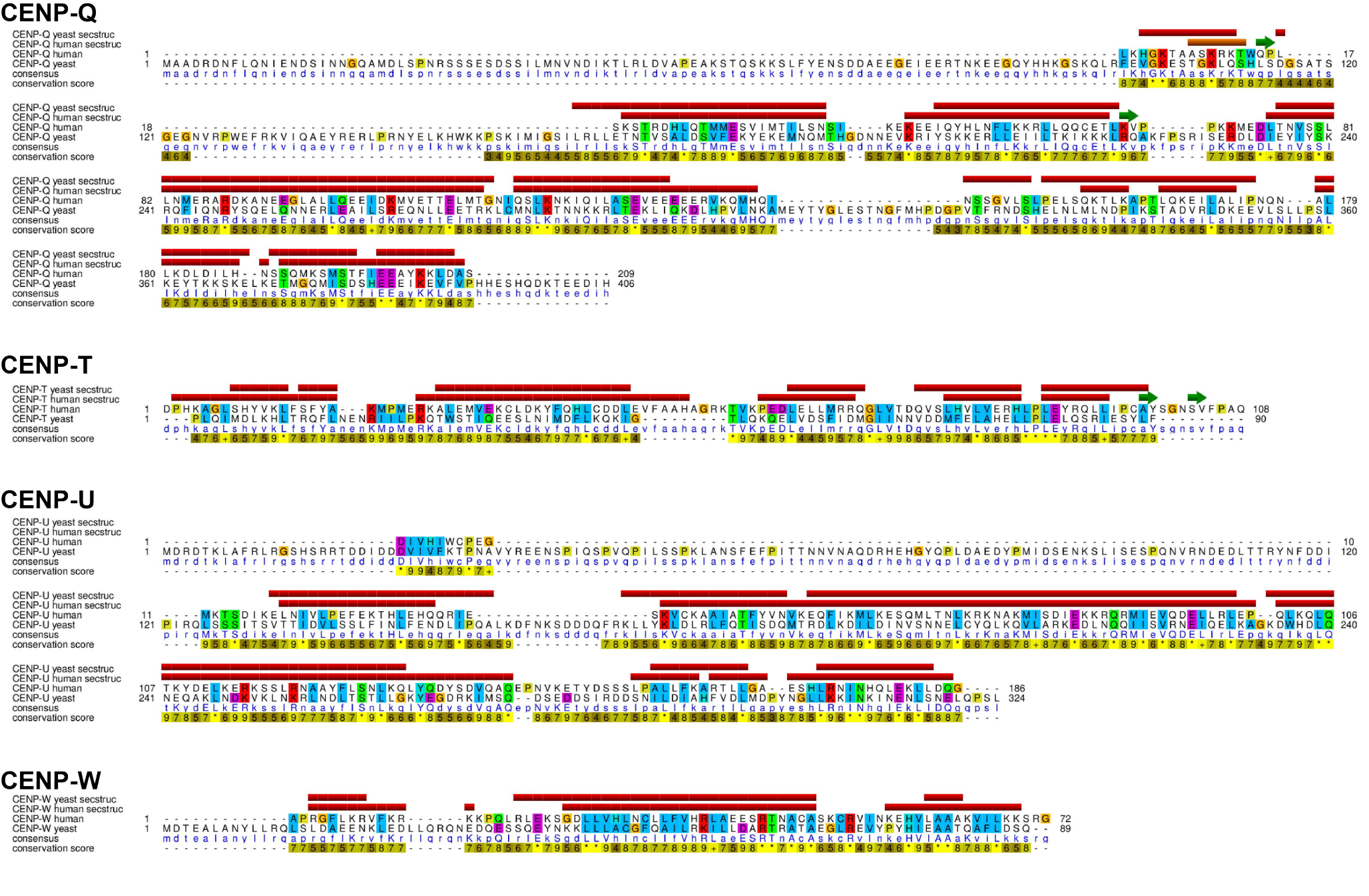


Figure S9 – Related to Figure 1

Figure S9 (related to Figure 1). Sequence alignments of human and yeast CENP-Q, CENP-U, CENP-T, and CENP-W

For details, please refer to legend of Figure S6.

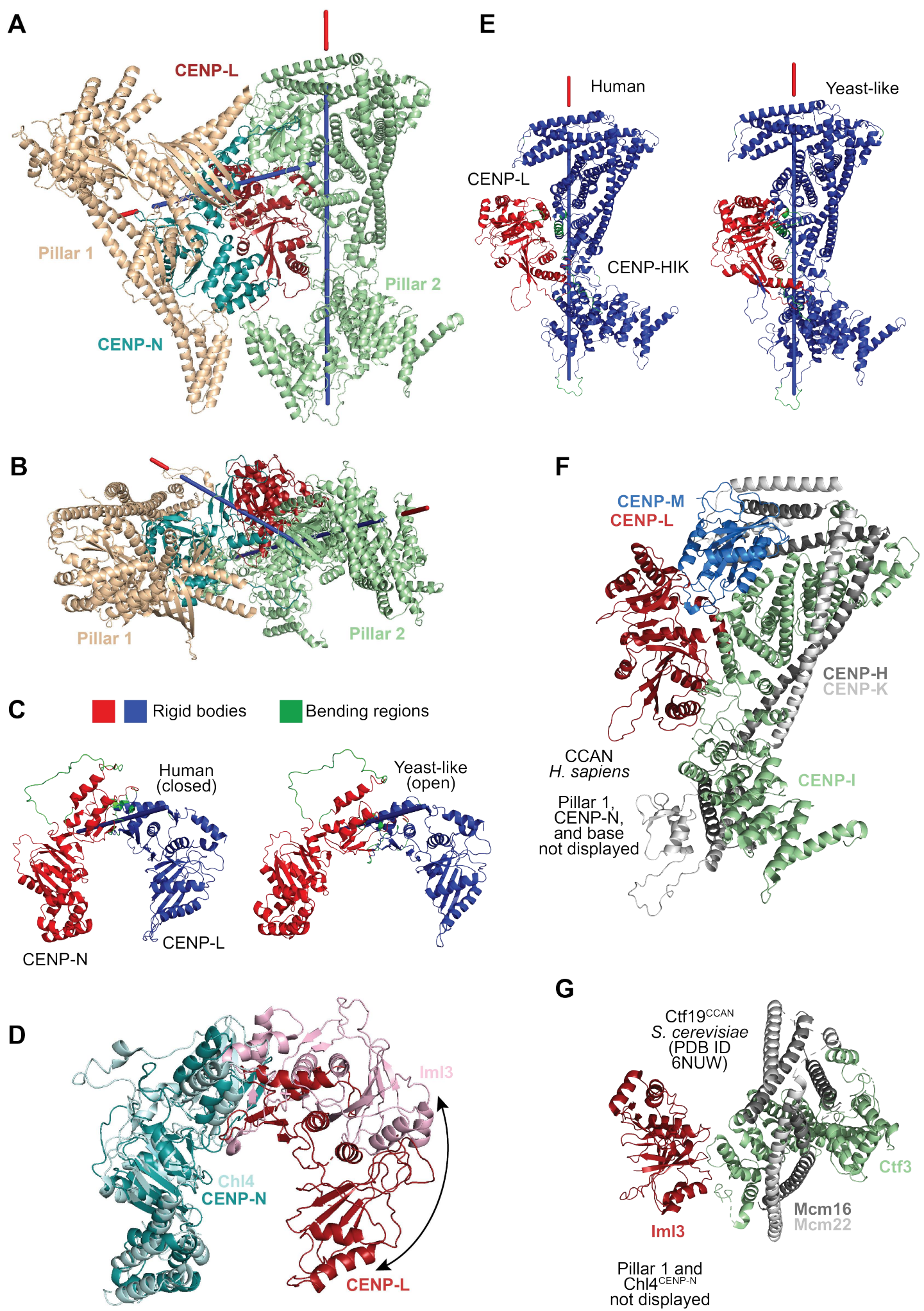


Figure S10 – Related to Figure 1

Figure S10 (related to Figure 1). Structural differences between yeast and human CCAN^{Ctf19}

A-**B**) The main structural differences between the yeast and human complexes are described by two screw axes, shown from two rotated views ≈90º apart, as identified by program DynDom6D (Veevers and Hayward, 2019). **C**) Output of DynDom6D for the first main rotation axis, which crosses the CENP-LN dimerization interface. The program's input for comparison were human CCAN subcomplexes and models of the same subcomplexes obtained by optimal subunit superposition on the yeast complex (and therefore identified as "yeast-like"). As shown, the program identifies two rigid bodies (shown in red and blue) in the comparison of open and closed CENP-LN (roughly corresponding to CENP-L and CENP-N). Bending regions are shown in green and cluster at the dimerization interface near the displayed screw axis. The operation is a rotation close to 45º with minimal translation. **D**) Superposition of CENP-N and Chl4 in human CENP-LN (respectively in deepteal and firebrick) and Iml3^{CENP-L}:Chl4^{CENP-N} (respectively in cyan and pink). The curved arrow emphasizes the relative rotation of CENP-L about an axis running through the dimerization domain. **E**) Output of DynDom6D for the second main screw axis, which runs through the interface between CENP-L and CENP-HIK. The two rigid bodies in the comparison correspond to CENP-L and CENP-HIK. The rotation angle is again close to 45º with minimal translation. **F**) Cartoon model of pillar 1 and CENP-L in the vault for human CCAN demonstrates extensive interactions. **G**) The equivalent region in the *S. cerevisiae* complex shows a much more modest interface that is limited to the Iml3^{CENP-L}:Ctf3^{CENP-I} pair, further emphasizing the role of CENP-M as stabilizer of the human complex.

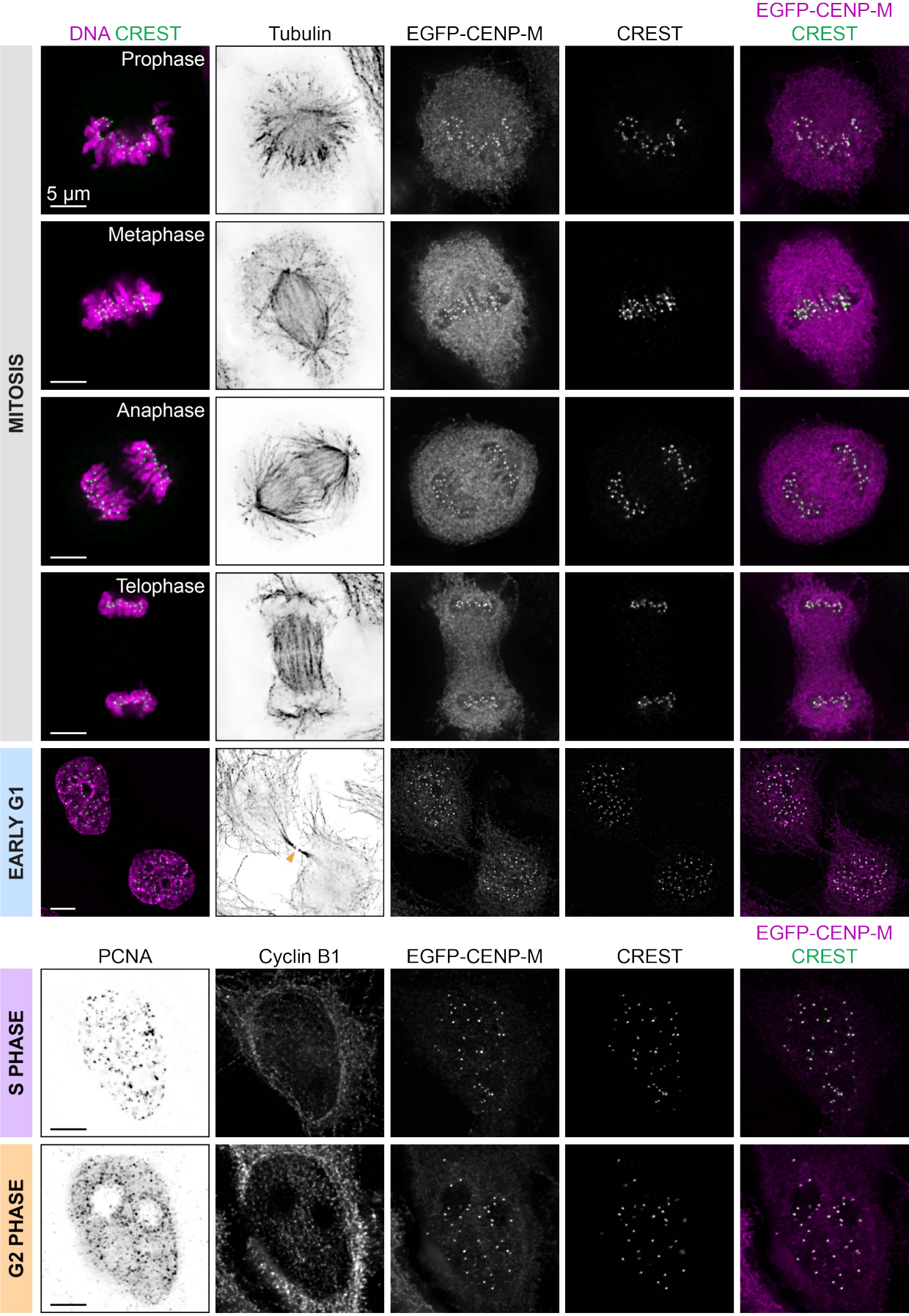


Figure S11 – Related to Figure 2

Figure S11 (related to Figure 2). Gallery of EGFP-CENP-M cell cycle localization

Maximum intensity projections of 5 x 0.2 μm Z-stacks. For Early G1 cells, the orange arrowhead indicates the spindle midbody remnant. Tubulin and DNA were the chosen markers to illustrate mitotic figures. PCNA and cytosolic Cyclin B1 were the chosen markers to illustrate S-phase and G2 phase. The displayed cells are also displayed in Figure 2H. Scale bars = 5 μm.

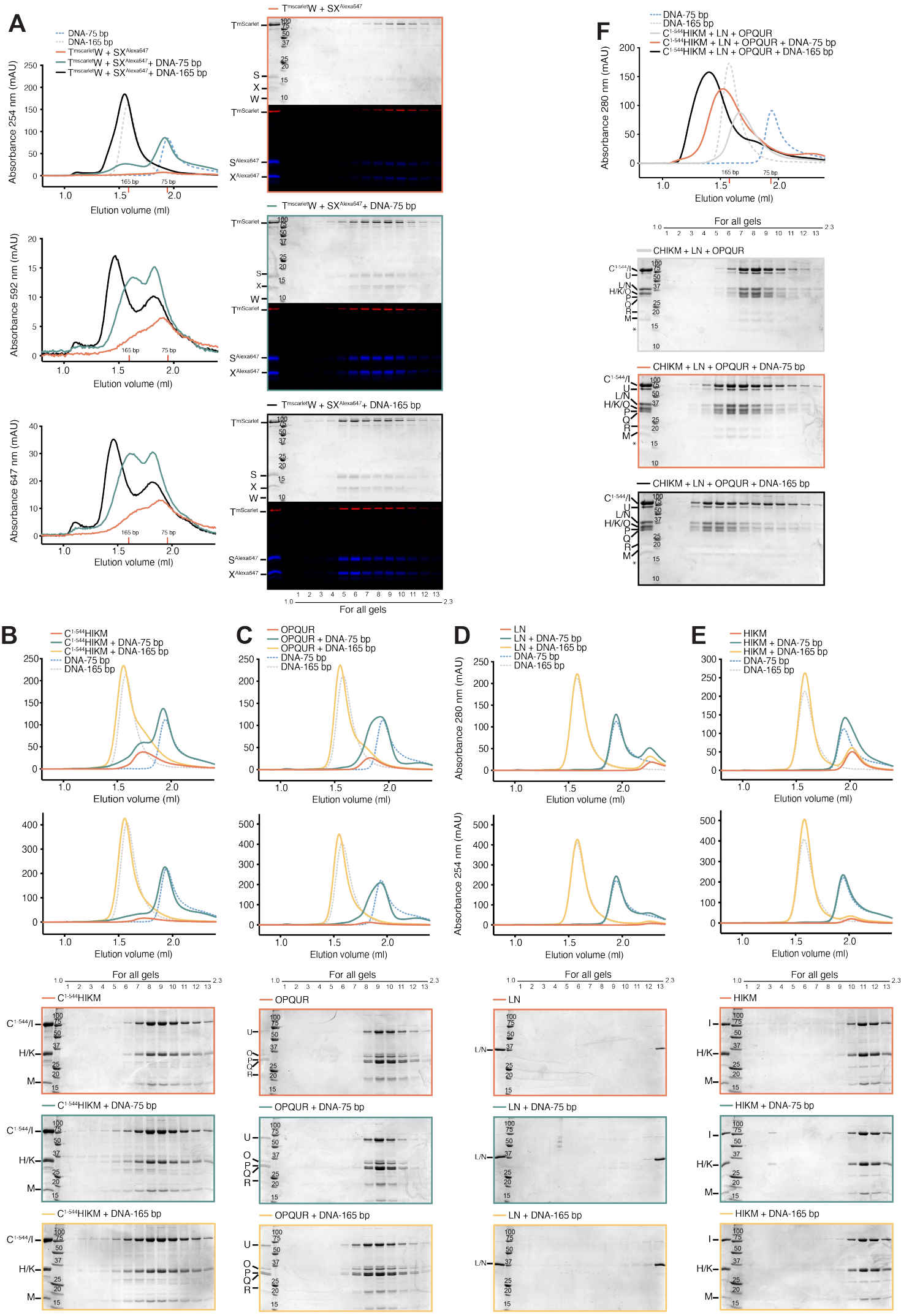


Figure S12 – Related to Figure 3

Figure S12 (related to Figure 3). DNA binding by CCAN subunits

Size-exclusion chromatography on a Superose 6 5/150 column of the indicated samples to monitor protein:DNA binding using 75-bps and 165-bps DNA. **A**) Binding of 75- and 165-bp DNA to fluorescently labelled CENP-TWSX complex. Elution profiles showing absorbance at the indicated wavelengths are shown on the left. Coomassie stained SDS-PAGE of the indicated fractions (100 µl fractions collected between 1.0 and 2.3 ml) are shown on the right together with fluorescence readout from CENP-T^{mScarlet} and Alexa647-labelled CENP-SX complex. **B**) Weak CENP-C$^{1-544}$HIKM binding to DNA is demonstrated by a modest left shift upon incubation of protein complex with DNA. The control chromatograms and gels in panels B-D are also shown in Figure 1A-C. **C**) CENP-OPQUR also binds weakly to DNA. **D**) CENP-LN does not bind to DNA. **E**) CENP-HIKM does not bind to DNA. **F**) The combined CCAN subunits in CENP-12 bound DNA with substantial binding affinity.

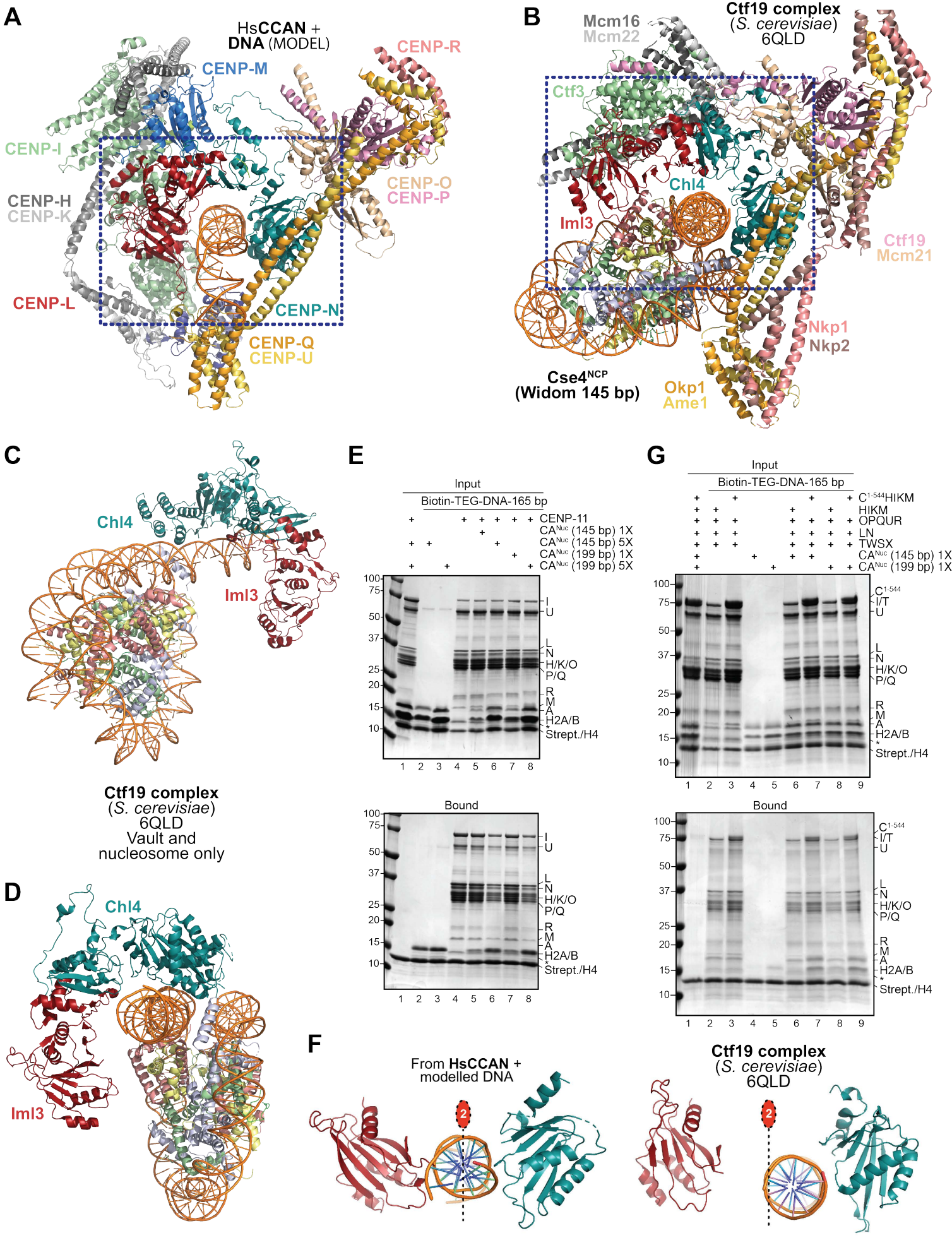


Figure S13 – Related to Figure 4

Figure S13 (related to Figure 4). 2-fold pseudosymmetry and DNA binding mode of the human vault

A) Cartoon model of human CCAN bound to DNA. The DNA fragment was extracted from 6C0W (CENP-N N-terminal region bound to CENP-A nucleosome (Pentakota et al., 2017)) using the DNA gyre facing CENP-N after superposition of CENP-N. **B**) View of the *S. cerevisiae*'s Ctf19^{CCAN}-Cse4^{CENP-A} nucleosome complex (6QLD (Yan et al., 2019)) with the same orientation as the human complex in panel A. **C**) View of 6QLD where only the vault and the nucleosome are displayed. **D**) As in panel C, but from a different orientation to emphasize lack of contacts of Iml3^{CENP-L} with the DNA and nucleosome. **E**) Biotin-TEG DNA immobilized on Straptavidin beads was incubated with preys as indicated. Octameric CENP-A nucleosomes created with 145- or 199-bps of DNA did not bind CENP-11 (CENP-HIKM, CENP-LN, CENP-OPQUR), and only caused modest release of the complex when used at 5x concentration. Samples were visualized by SDS-PAGE and Coomassie staining. Three technical repeats were performed. Note that the preparation of nucleosome had an excess of H2A:H2B that bound DNA also in absence of CENP-11. Lack of nucleosome binding is testified by absence of CENP-A in bound fractions. **F**) The 2-fold pseudosymmetry axis linking the two LNHDs of CENP-L and CENP-N and the modelled DNA in human CCAN is offset in the yeast structure (6QLD). The LNHD of Chl4^{CENP-N} is displayed with the same orientation used for the human complex. **G**) Addition of CENP-C$^{1-544}$ to HIKM allows direct comparison of CENP-A nucleosome binding by CENP-15 and CENP-16. CENP-A bound only in presence of CENP-C$^{1-544}$. Three technical repeats were performed.

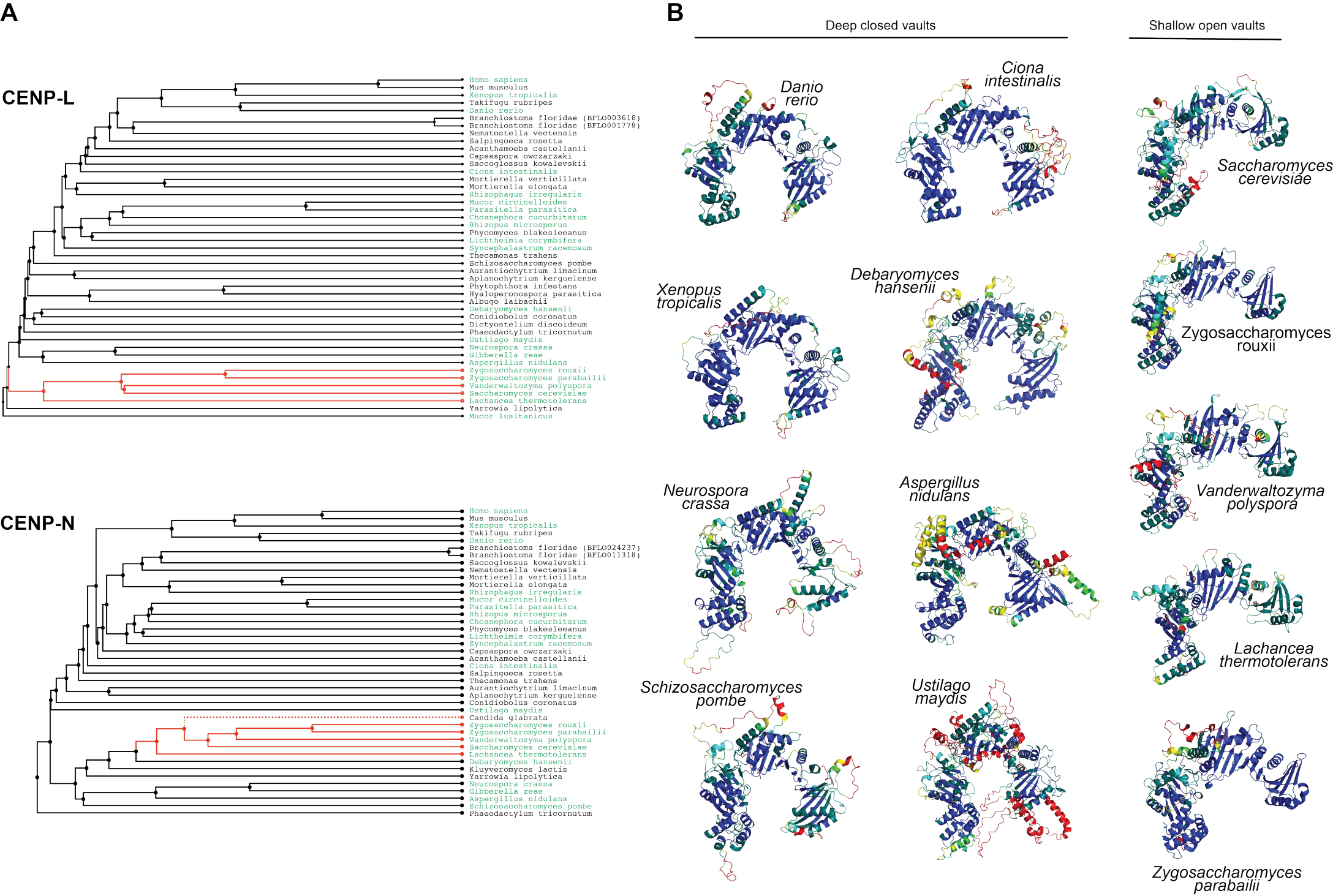


Figure S14 – Related to Figure 5

Figure S14 (related to Figure 5). Phylogenetic analysis of open and closed vaults

A) Cladograms of the indicated CENP-L and CENP-N species were calculated from sequence alignments with https://mafft.cbrc.jp/alignment/server/phylogeny.html. Structures of CENP-LN dimers predicted by AF2 are indicated by green species names. The part of the branch indicated with red lines reflects presence of open vaults (either experimentally ascertained, like *Saccharomyces cerevisiae*'s or predicted by AF2. No CENP-L homolog was found in C. glabrata (dashed line). **B**) Gallery of a representative subset of AF2 predictions (from those indicated in green in panel A) of vertebrate and yeast CENP-LN orthologs from the indicated species, represented in pLDDT score (blue, high confidence prediction; red, low confidence prediction).

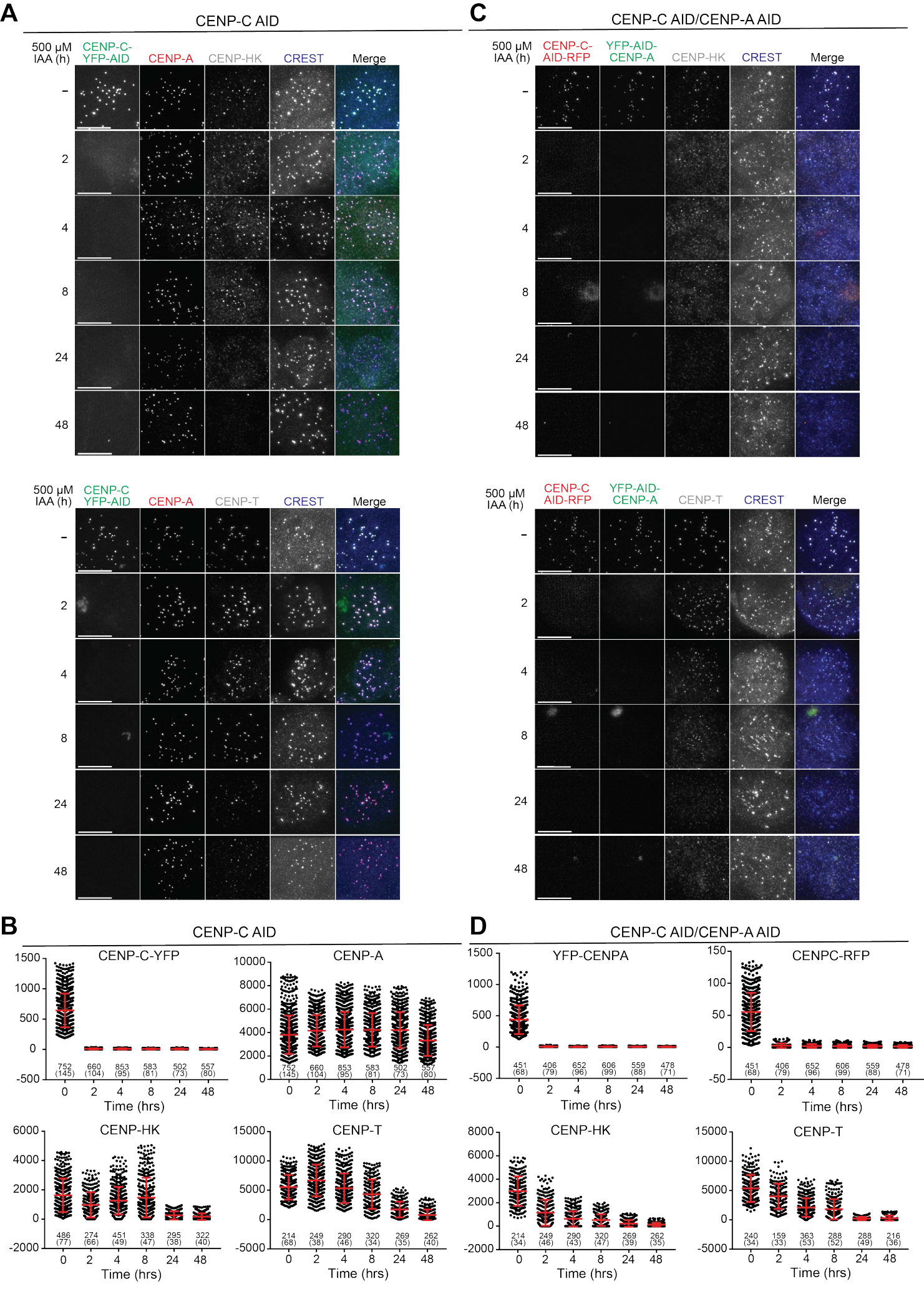


Figure S15 – Related to Figure 6

Figure S15 (related to Figure 6). Time course of CENP-C and CENP-A depletion

A) Degradation of CENP-C endogenously tagged with a EYFP-AID cassette (Fachinetti et al., 2015) was induced by addition of indole acetic acid (IAA, Auxin). The full time-course is shown here. Levels of CENP-A, CENP-HK, CENP-T, and CREST were also monitored. Scale bar (here and in panel C) = 10 μm. **B**) Quantification of the experiments in panels A. Number of kinetochores and number of cells (in parentheses) are indicated. Levels of CENP-A and CENP-$C^{YFP\text{-}AID}$ from both experiments were quantified together. Levels of CENP-T and CENP-HK were quantified from the individual experiments in panel A. Red bars represent average fluorescence intensities and standard deviation of quantified centromere foci. **C**) Degradation of CENP-A endogenously tagged with a EYFP-AID cassette and of CENP-C endogenously tagged with an RFP-AID cassette (Hoffmann et al., 2020) at the indicated time points after addition of IAA. Levels of CENP-HK, CENP-T, and CREST were also monitored. **D**) Quantification of the experiments in panel C was performed and displayed precisely as described for panel A.