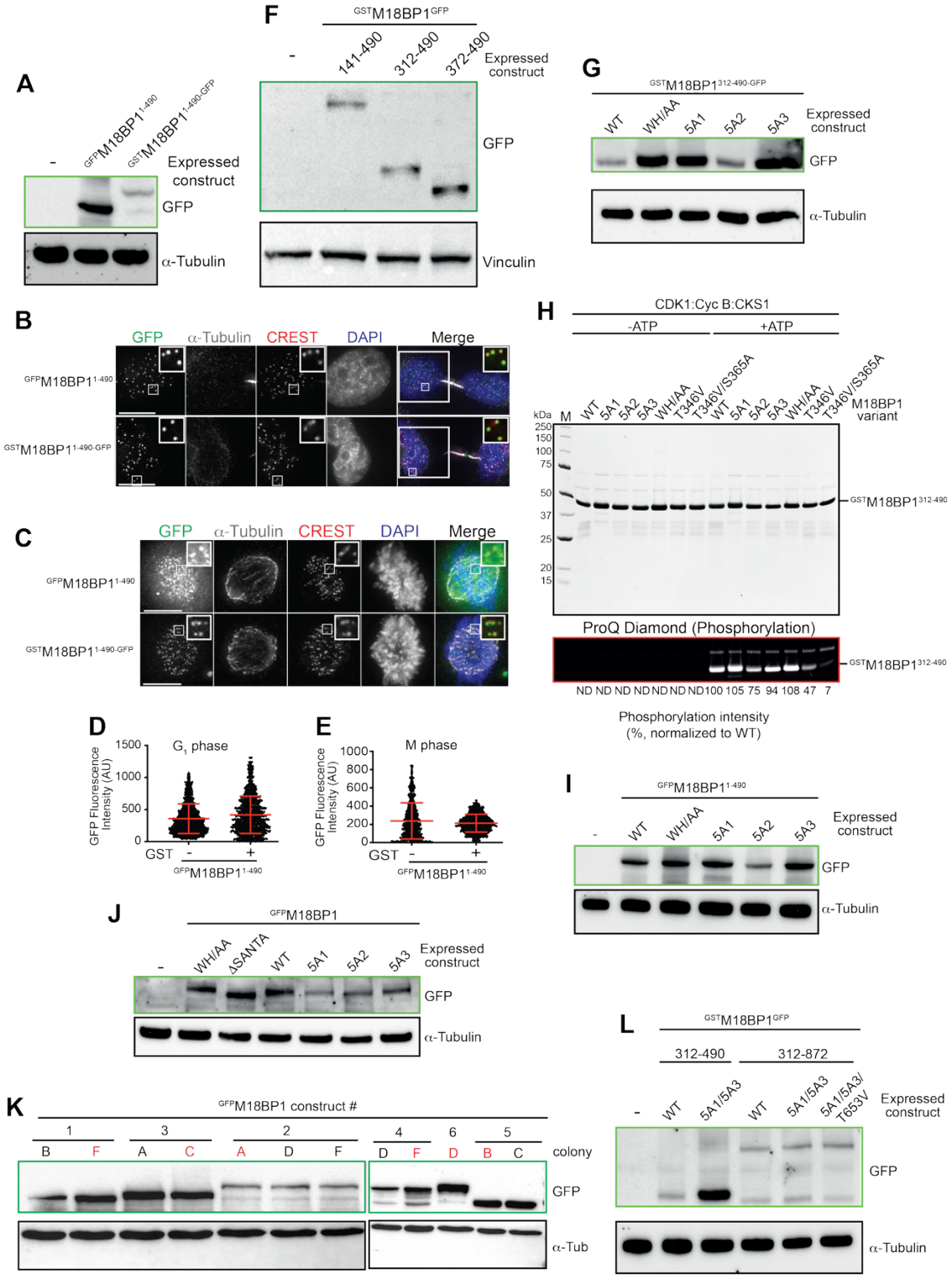
M18BP1 valency and a distributed interaction footprint determine epigenetic centromere specification in humans

by

Kai Walstein, Louisa Hill, Doro Vogt, Lina Oberste-Lehn, Petra Janning, Ingrid R. Vetter, Dongqing Pan & Andrea Musacchio

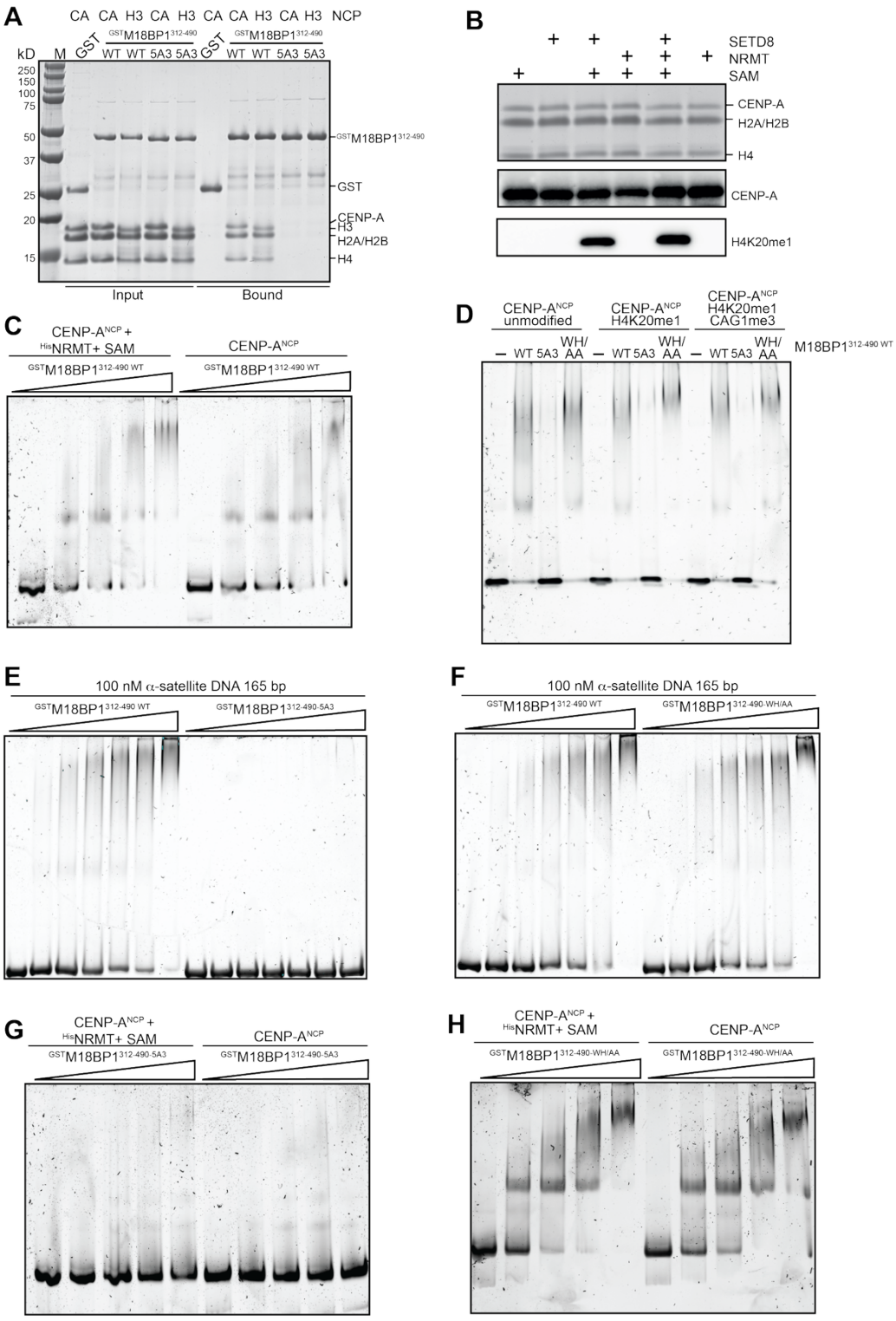
Table of Content of Appendix Material



Appenndix Figure S1

Appendix Figure S1 ***M18BP1 construct expression levels***

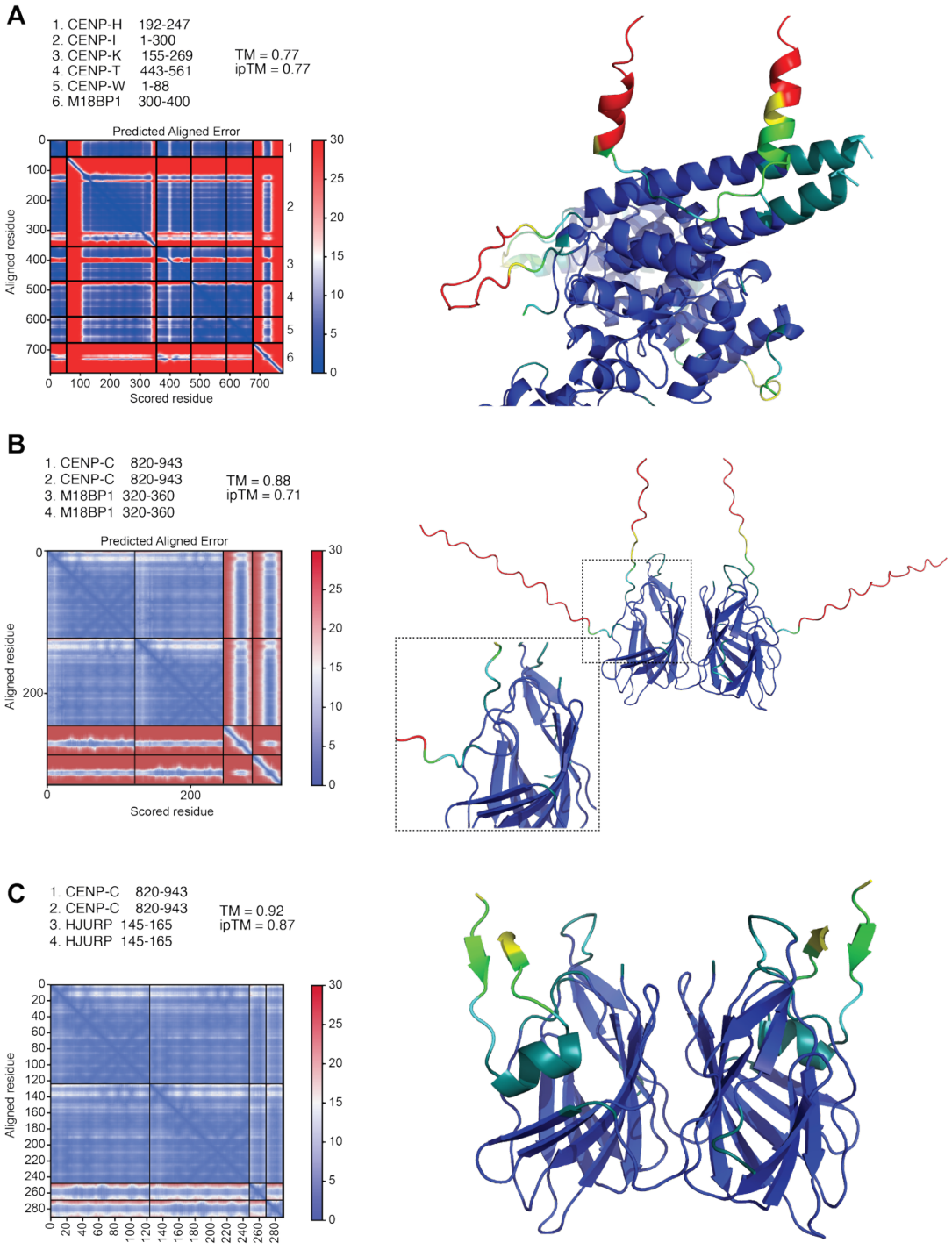
(**A**) Anti-GFP and anti-α-Tubulin immunoblots of HeLa cell lysates expressing the indicated GFP tagged M18BP1 constructs. (**B**) and (**C**) Representative images of fixed HeLa cells showing GFP fluorescence of indicated stably expressed M18BP1 variants and immuno-stained α-tubulin in G1 and mitosis, respectively. Centromeres were visualized by CREST sera, and DNA was stained by DAPI. White scale bars indicate 10 μm. (**D**) and (**E**) Quantification of the centromeric GFP-fluorescence intensities of the indicated expressed M18BP1 variants in G1 and mitosis, respectively. Data shown as mean with SD. Number of quantified centromeres in (D) was 1568 (-GST) and 1373 (+GST), and in (E) 1147 (-GST) and 1082 (+GST). (**F**) Anti-GFP and anti-α-Vinculin immunoblots of HeLa cell lysates expressing the indicated ^{GST}M18BP1^{GFP} constructs. (**G**) Anti-GFP and anti-Tubulin immunoblots of HeLa cell lysates expressing the indicated ^{GST}M18BP1$^{312\text{-}490\text{-}GFP}$ variants. (**H**) SDS-PAGE and ProQ staining showing the phosphorylation intensities of the indicated ^{GST}M18BP1$^{312\text{-}490}$ variants. All samples were incubated with CCC kinase, ATP was added as indicated. (**I**) and (**J**) Anti-GFP and anti-α-Tubulin immunoblots of HeLa cell lysates expressing the indicated GFP tagged M18BP1 constructs. (**K**) Anti-GFP and anti-α-Tubulin immunoblots of HeLa cell lysates from different colonies to compare the expression levels of the indicated GFP tagged M18BP1 constructs. The colonies marked in red were used for the experiment shown in Figure 5I-L. (**L**) Anti-GFP and anti-α-Tubulin immunoblots of HeLa cell lysates expressing the indicated GFP-tagged exogenous M18BP1 constructs.



Appendix Figure S2

Appendix Figure S2 ***M18BP1 interacts with centromeric DNA and nucleosomes***

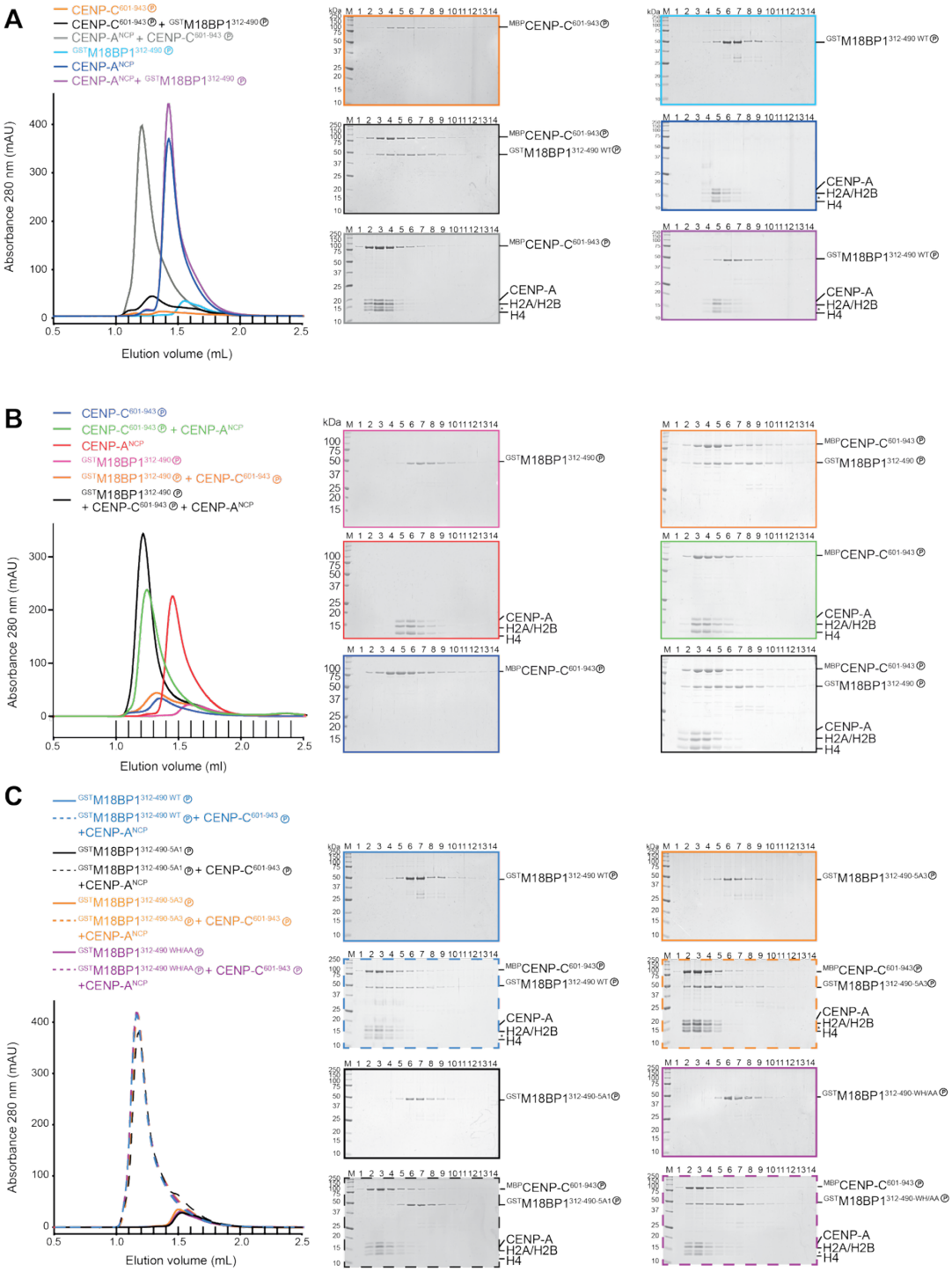
(**A**) GSH-resin pull-down assay with GST (negative control) and ^{GST}M18BP1$^{312-490}$ as baits testing the binding to CENP-A^{NCP} ("CA") and H3^{NCP}. The NCP variant used as prey is indicated above each lane. (**B**) SDS-PAGE, anti-CENP-A and anti-H4K20me1 immunoblots of CENP-A^{NCP} treated with SETD8, NRMT and SAM as indicated above each lane. (**C**) EMSA experiment showing the interaction between ^{GST}M18BP1$^{312-490}$ and CENP-A^{NCP}. The NCPs were added at 100 nM concentration, the M18BP1 concentration was varied from 0 mM (lowest concentration) to 1 mM (highest concentration). (**D**) EMSA experiment showing the interaction between ^{GST}M18BP1$^{312-490}$ variants and differently modified CENP-A^{NCP}. The NCPs were added at 100 nM concentration, the indicated M18BP1 variants were added at 1 mM concentration. (**E**) and (**F**) EMSA experiments showing the interactions between a-satellite DNA and different ^{GST}M18BP1$^{312-490}$ variants. The DNA was added at 100 nM concentration, the indicated M18BP variant concentrations were varied from 0 mM (lowest concentration) to 2 mM (highest concentration). (**G**) and (**H**) EMSA experiments showing the interactions between ^{GST}M18BP1$^{312-490}$ variants and CENP-A^{NCP}. The NCPs were added at 100 nM concentration, the M18BP1 concentration was varied from 0 mM (lowest concentration) to 1 mM (highest concentration).



Appendix Figure S3

Appendix Figure S3 ***Quality indicators of AlphaFold predictions***

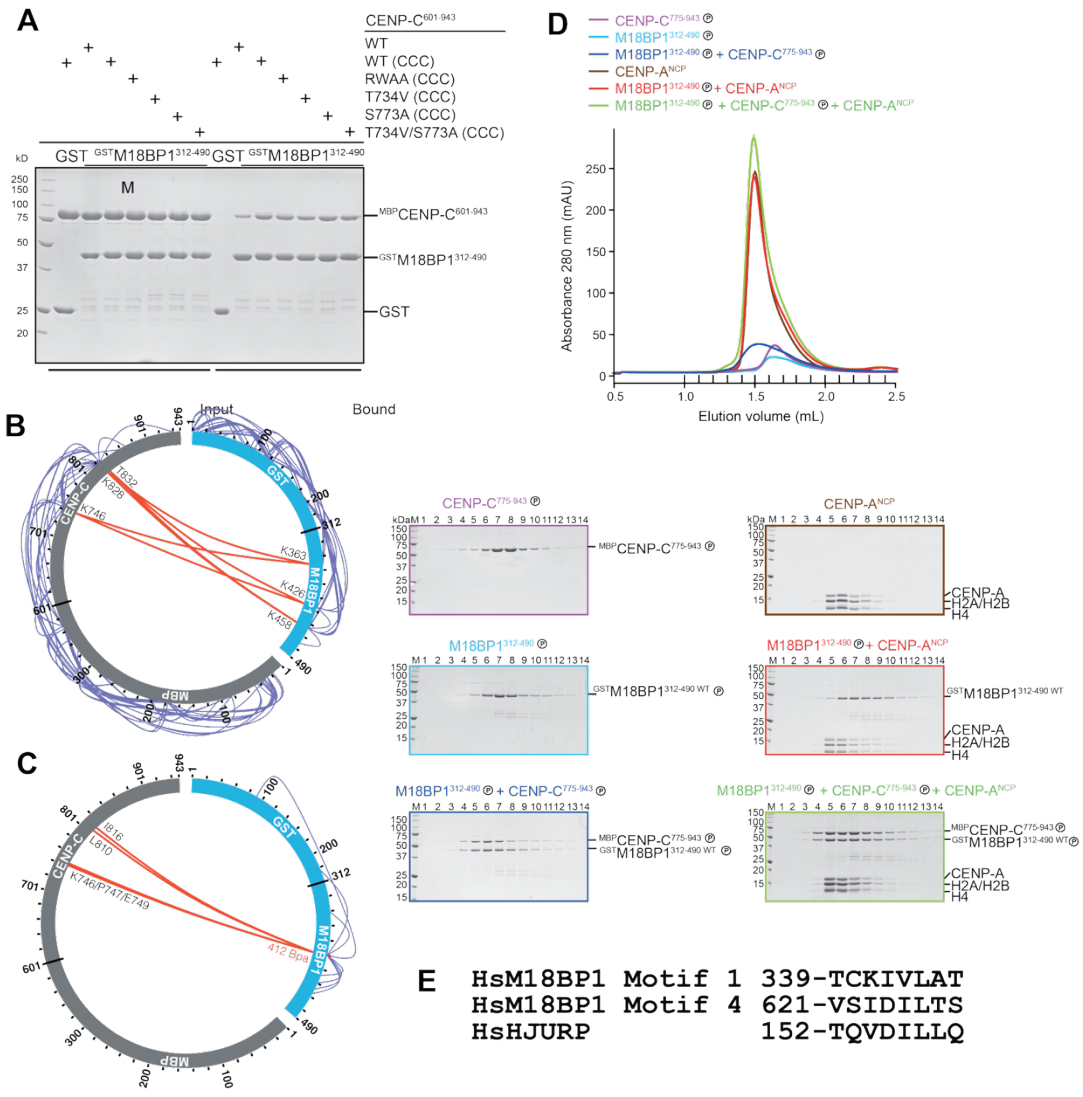
(**A**) AlphaFold2 prediction of the interaction of M18BP1 residues 300-400 with the CENP-HIKM complex. PAE plot, domain boundaries, and quality indicators are shown. (**B**) AlphaFold3 prediction of the interaction of the CENP-C$^{820\text{-}943}$ Cupin dimer with M18BP1$^{320\text{-}360}$. PAE plot, domain boundaries, and quality indicators are shown. (**C**) AlphaFold3 prediction of the interaction of the CENP-C$^{820\text{-}943}$ Cupin dimer with HJURP$^{145\text{-}165}$, encompassing the motif shown in Appendix Figure S5E. PAE plot, domain boundaries, and quality indicators are shown.



Appendix Figure S4

Appendix Figure S4 ***The interaction between M18BP1 and CENP-A is mediated by CENP-C***

(**A**) to (**C**) Chromatograms and SDS PAGEs of analytical size exclusion chromatography (SEC) experiments. ^{MBP}CENP-C$^{601-943}$ and ^{GST}M18BP1$^{312-490}$ variants were incubated as indicated at 10 μM, CENP-A^{NCP} at 5 μM for 30 min. The individual proteins or incubated mixtures were loaded on an analytical Superdex 200 column. The 14 indicated fractions between 1 mL and 2.5 mL elution volume were subsequently analyzed by SDS PAGE. Components pre-phosphorylated by CCC kinase are marked by an encircled P.



Appendix Figure S5

Appendix Figure S5 ***The interaction between M18BP1 and CENP-C***

(**A**) GSH-resin pull-down assay with GST (negative control) and ^{GST}M18BP1$^{312-490}$ variants as baits testing the binding to different ^{MBP}CENP-C$^{601-943}$ constructs carrying mutations in conserved residues of the CENP-C motif, known to bind CENP-A nucleosomes. The M18BP1 variants used as baits and the addition of CDK1/Cyclin B/CKS1 (CCC) kinase complex are indicated above each lane. (**B**) Schematic showing the results of a DSBU crosslinking experiment followed by mass spectrometry analysis. ^{GST}M18BP1$^{312-490WT}$ was incubated with ^{MBP}CENP-C$^{601-943}$. Intermolecular crosslinks are highlighted in red, intramolecular crosslinks are highlighted in blue. Crosslinks involving the GST or MBP tags were excluded for clarity. (**C**) Schematic showing the results of UV-induced crosslinking experiments followed by mass spectrometry analysis. Ultraviolet (UV) light was used to induce crosslinking of CENP-C$^{601-943}$ to ^{GST}M18BP1$^{312-490_412BPA}$ (Benzoyl-phenylalanine introduced at residue 412). Intermolecular crosslinks are highlighted in red, intramolecular crosslinks are highlighted in blue. (**D**) Chromatograms and SDS PAGEs of SEC experiments. ^{MBP}CENP-C$^{775-943}$ and ^{GST}M18BP1$^{312-490}$ were incubated as indicated at 10 μM, CENP-A^{NCP} at 5 μM for 30 min. The individual proteins or incubated mixtures were loaded on an analytical Superdex 200 column. The 14 indicated fractions between 1 mL and 2.5 mL elution volume were subsequently analyzed by SDS PAGE. Components pre-phosphorylated by CCC kinase are marked by an encircled P. (**E**) Similarity of motifs in M18BP1 (Motifs 1 and 4) and HJURP.

Appendix Table S1 ***X-ray data collection and refinement statistics***

Resolution range (Å)	45.04-2.42 (2.51-2.42)*
Space group	C2
a,b, c (Å)	143.2 120.28 77.36
α, β, γ (°)	90.00 121.28 90.00
Molecules/asymmetric unit	2
Wavelength	1.0332
Number of crystals	2
synchrotron	PETRA
date	08. July 2019
Detector	Pilatus6MF
Dataset statistics:	
Total reflections	574151 (43031)
Unique reflections	42292 (3857)
Multiplicity	13.6 (11.1)
Completeness (%)	98.97 (90.67)
Mean $I/\sigma I$	10.37 (1.10)
CC1/2 (%)	99.7 (47.1)
R-merge (%)	18.49 (207.7)
R-meas (%)	19.21 (217.7)
R-pim (%)	5.18 (63.61)
Wilson B-factor	56.74
Refinement:	
Reflections in refinement	42265 (3857)
R_{work} / R_{free}(%)	22.91/26.19
Protein atoms	7320
Ligand atoms	187
Solvent	94
average B protein/ligand/solvent (Å^2)	64.19/64.63/50.93
Clashscore	2.48
Rotamer outliers (%)	2.0
R.m.s. deviations:	
Bond lengths (Å)	0.007
Bond angles (°)	0.99
Ramachandran:	
Favored/allowed regions (%)	97.68/2.32
Outliers (%)	0.0

***** Statistics for highest resolution shell are shown in parentheses.