



Structural basis for Mis18 complex assembly and its implications for centromere maintenance

Reshma Thamkachy^{1,8}, Bethan Medina-Pritchard^{ID 1,8}, Sang Ho Park^{ID 2,8}, Carla G Chiodi¹, Juan Zou^{ID 1}, Maria de la Torre-Barranco¹, Kazuma Shimanaka^{ID 2}, Maria Alba Abad¹, Cristina Gallego Páramo^{ID 1}, Regina Feederle^{ID 3}, Emilija Ruksenaite⁴, Patrick Heun¹, Owen R Davies^{ID 1}, Juri Rappsilber^{ID 1,5}, Dina Schneidman-Duhovny⁶, Uhn-Soo Cho^{ID 2} & A Arockia Jeyaprakash^{ID 1,7}✉

Abstract

The centromere, defined by the enrichment of CENP-A (a Histone H3 variant) containing nucleosomes, is a specialised chromosomal locus that acts as a microtubule attachment site. To preserve centromere identity, CENP-A levels must be maintained through active CENP-A loading during the cell cycle. A central player mediating this process is the Mis18 complex (Mis18 α , Mis18 β and Mis18BP1), which recruits the CENP-A-specific chaperone HJURP to centromeres for CENP-A deposition. Here, using a multi-pronged approach, we characterise the structure of the Mis18 complex and show that multiple hetero- and homo-oligomeric interfaces facilitate the hetero-octameric Mis18 complex assembly composed of 4 Mis18 α , 2 Mis18 β and 2 Mis18BP1. Evaluation of structure-guided/separation-of-function mutants reveals structural determinants essential for cell cycle controlled Mis18 complex assembly and centromere maintenance. Our results provide new mechanistic insights on centromere maintenance, highlighting that while Mis18 α can associate with centromeres and deposit CENP-A independently of Mis18 β , the latter is indispensable for the optimal level of CENP-A loading required for preserving the centromere identity.

Keywords Centromere; CENP-A; Mis18 Complex; Centromere Inheritance; Integrative Structural Analysis

Subject Categories Cell Cycle; Chromatin, Transcription & Genomics; Structural Biology

<https://doi.org/10.1038/s44319-024-00183-w>

Received 22 March 2024; Revised 6 May 2024;

Accepted 6 June 2024

Published online: 01 July 2024

Introduction

Faithful chromosome segregation during cell division requires the bi-orientation of chromosomes on the mitotic spindle through the physical attachment of kinetochores to microtubules.

Kinetochores are large multiprotein scaffolds that assemble on a special region of chromosomes known as the centromere (Catania and Allshire, 2014; Cheeseman, 2014; Fukagawa and Earnshaw, 2014; Musacchio and Desai, 2017). Whilst centromeres in some organisms, such as budding yeast, are defined by a specific DNA sequence, in most eukaryotes, centromeres are distinguished by an increased concentration of nucleosomes containing a histone H3 variant called CENP-A (Black et al, 2010; Fukagawa and Earnshaw, 2014; McKinley and Cheeseman, 2016; Stellfox et al, 2013). CENP-A containing nucleosomes recruit CENP-C and CENP-N, two proteins that are part of the constitutive centromere-associated network (CCAN) and that recruit the rest of the kinetochore components at the centromeric region of the chromosome (Carroll et al, 2010; Kato et al, 2013; Weir et al, 2016).

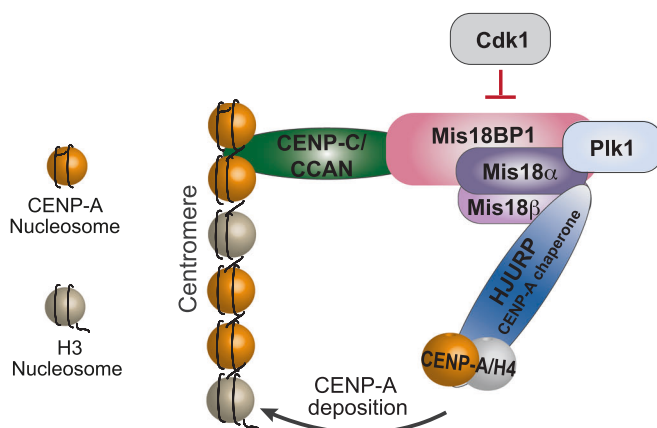
Whilst canonical histone loading is coupled with DNA replication, CENP-A loading is not (Dunleavy et al, 2011). This results in a situation where, after S-phase, the level of CENP-A nucleosomes at the centromere is halved due to the distribution of existing CENP-A to the duplicated DNA (Dunleavy et al, 2009; Jansen et al, 2007). To maintain centromere identity, centromeric CENP-A levels must be restored. This is achieved through active CENP-A loading at centromeres (during G1 in humans) via a pathway that requires the Mis18 complex (consisting of Mis18 α , Mis18 β and Mis18BP1) and the CENP-A chaperone, HJURP (Barnhart et al, 2011; Dunleavy et al, 2009; Foltz et al, 2009; Fujita et al, 2007; Jansen et al, 2007) (Fig. 1A). The Mis18 complex can recognise and localise to the centromere, possibly through its proposed binding to CENP-C and/or other mechanisms which have not yet been identified (Dambacher et al, 2012; Moree et al, 2011; Stellfox et al, 2016). Once at the centromere, the Mis18 complex has been implicated in facilitating the deposition of CENP-A in several ways. There is evidence that the Mis18 complex affects DNA methylation and histone acetylation, which may facilitate CENP-A loading (Hayashi et al, 2004; Kim et al, 2012). But one of the most important and well-established roles of the Mis18 complex is the recruitment of HJURP, which binds a single

¹Wellcome Centre for Cell Biology, University of Edinburgh, Edinburgh EH9 3BF, UK. ²Department of Biological Chemistry, University of Michigan, Ann Arbor, MI 48109, USA.

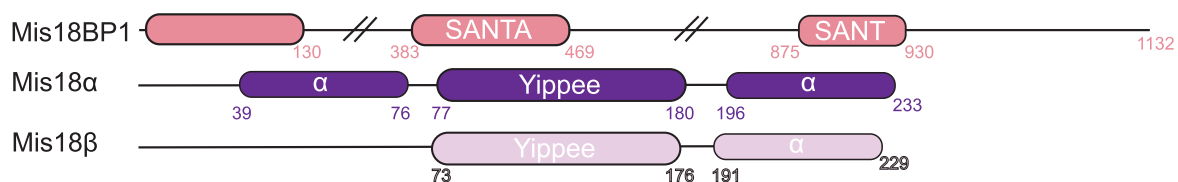
³Monoclonal Antibody Core Facility, Helmholtz Zentrum München, German Research Center for Environmental Health (GmbH), 85764 Neuherberg, Germany. ⁴Institute Novo Nordisk Foundation Centre for Protein Research, Copenhagen, Denmark. ⁵Institute of Biotechnology, Technische Universität Berlin, 13355 Berlin, Germany. ⁶School of Computer Science and Engineering, The Hebrew University of Jerusalem, Jerusalem, Israel. ⁷Gene Center, Department of Biochemistry, Ludwig Maximilians Universität, Munich, Germany.

⁸These authors contributed equally: Reshma Thamkachy, Bethan Medina-Pritchard, Sang Ho Park. ✉E-mail: jeyaprakash.arulnandam@ed.ac.uk

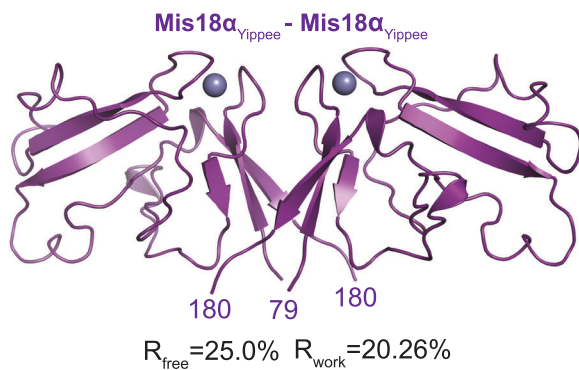
A



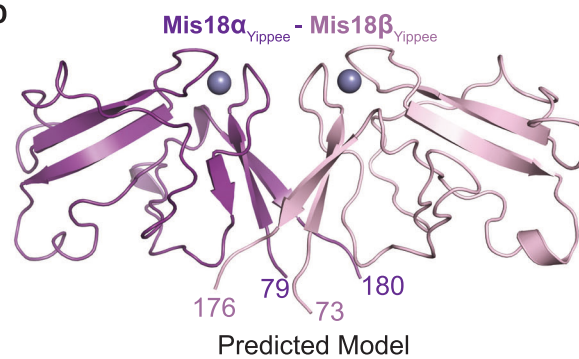
B



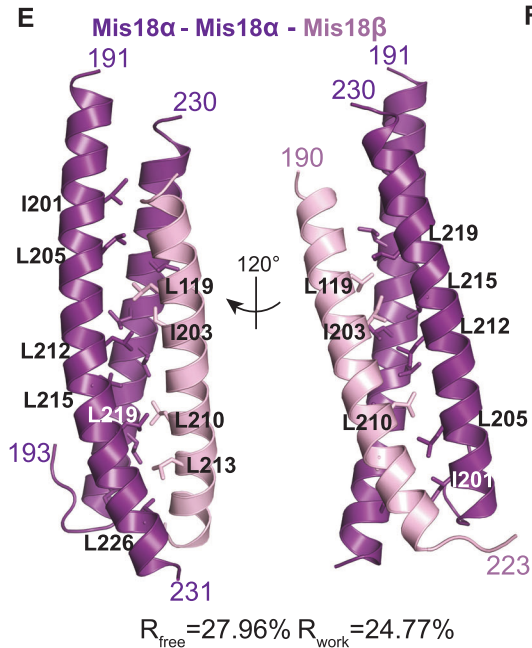
C



D



E



F

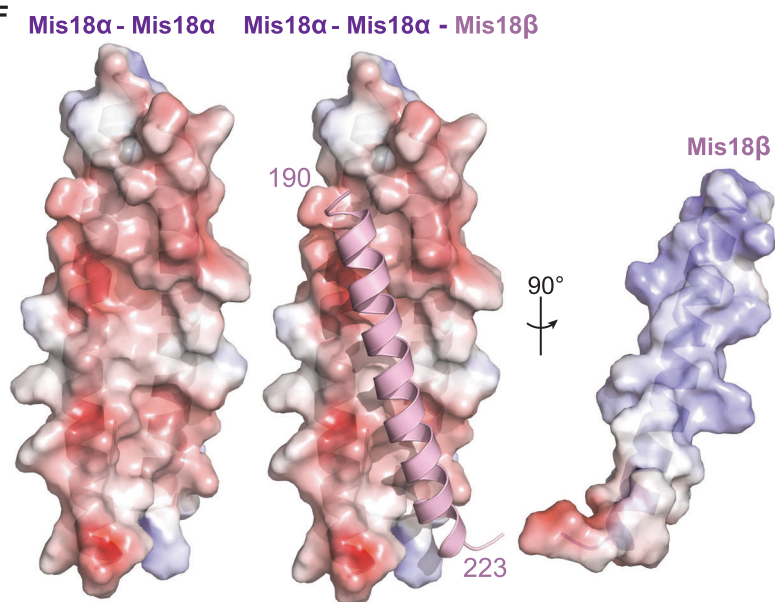


Figure 1. Mis18 α/β contains two independent structural domains that can oligomerise.

(A) Diagram of proteins involved in CENP-A deposition at the centromere. The Mis18 complex (Mis18BP1 (salmon), Mis18 α (purple) and Mis18 β (light pink)) forms once Cdk1 activity is reduced. It interacts with CCAN/CENP-C (green) to localise to the centromere, where Plk1 regulation helps promote the recruitment of HJURP (Blue), a CENP-A chaperone. (B) Schematic representation of structural features of Mis18BP1 (salmon), Mis18 α (purple) and Mis18 β (light pink). Filled boxes represent folded domains. SANTA and SANT domain boundaries as defined in UniProt (Q6PONO). (C) Cartoon representation of the crystal structure of human Mis18 α_{Yippee} homodimer (PDB ID: 7SFZ). (D) Cartoon representation of the human Mis18 α_{Yippee} /Mis18 β_{Yippee} heterodimer modelled by homology to the structure in Fig. 1C. Mis18 α is shown in purple and Mis18 β in light pink (modelled using Phyre2, www.sbg.bio.ic.ac.uk/phyre2/ (Kelley et al, 2015)). (E) Cartoon representation of the crystal structure of Mis18 $\alpha_{\text{C-term}}$ /Mis18 $\beta_{\text{C-term}}$ (PDB ID: 7SFY). Mis18 α is shown in purple and Mis18 β in light pink. (F) Mis18 $\alpha_{\text{C-term}}$ domains are shown in surface representation and coloured based on electrostatic surface potential calculated using APBS (Baker et al, 2001). Mis18 $\beta_{\text{C-term}}$ shown as cartoon.

CENP-A/H4 dimer and brings it to the centromere (Barnhart et al, 2011; Dunleavy et al, 2009; Hu et al, 2011). This then triggers a poorly understood process in which the H3 nucleosomes are removed and replaced with CENP-A nucleosomes. Finally, the new CENP-A nucleosomes are stably integrated into the genome, which requires several remodelling factors such as MgcRacGAP, RSF, Ect2 and Cdc42 (Lagana et al, 2010; Perpelescu et al, 2009).

The timing of CENP-A deposition is tightly regulated, both negatively and positively, by the kinases Cdk1 and Plk1, respectively, in a cell cycle-dependent manner (McKinley and Cheeseman, 2014; Muller et al, 2014; Pan et al, 2017; Silva et al, 2012; Spiller et al, 2017; Stankovic et al, 2017). Previous studies demonstrated that Cdk1 phosphorylation of Mis18BP1 prevents the Mis18 complex assembly and localisation to centromeres until the end of mitosis (when Cdk1 levels are reduced) (Pan et al, 2017; Spiller et al, 2017). Cdk1 also phosphorylates HJURP, which negatively regulates its binding to the Mis18 complex at the centromere (Muller et al, 2014; Stankovic et al, 2017; Wang et al, 2014). In cells, Plk1 is a positive regulator, and its activity is required for G1 centromere localisation of the Mis18 complex and HJURP. Plk1 has been shown to not only phosphorylate Mis18 α/β and Mis18BP1, but it has also been proposed to interact with phosphorylated Mis18 complex through its polo-box domain (PBD) (McKinley and Cheeseman, 2014).

As outlined above, a central event in the process of CENP-A deposition at centromeres is the Mis18 complex assembly. The Mis18 proteins, Mis18 α and Mis18 β , possess a well-conserved globular domain called the Yippee domain (also known as the MeDiY domain; spanning residues 77–180 in Mis18 α and 73–176 in Mis18 β) and C-terminal α -helices (residues 196–233 in Mis18 α and 191–229 in Mis18 β). We and others previously showed that the Yippee domains of Mis18 proteins can form a heterodimer, while the C-terminal helices form a heterotrimer with two Mis18 α and one Mis18 β . However, the full-length proteins form a hetero-hexameric assembly with 4 Mis18 α and 2 Mis18 β . This led to a proposed model, where the Mis18 α and Mis18 β mainly interact via the C-terminal helices to form a heterotrimer, and two such heterotrimers interact via the Yippee heterodimerisation (Mis18 α /Mis18 β) or/and homodimerisation (Mis18 α /Mis18 α) to form a hetero-hexameric assembly (Nardi et al, 2016; Pan et al, 2017; Pan et al, 2019; Spiller et al, 2017).

Mis18BP1, the largest subunit of the Mis18 complex (1132 aa residues), is a multi-domain protein containing SANTA (residues 383–469) and SANT (residues 875–930) domains, which are known to have roles in regulating chromatin remodelling (Aasland et al, 1996; Maddox et al, 2007; Zhang et al, 2006). In between these two domains resides the CENP-C binding domain (CBD) (Dambacher

et al, 2012; Stellfox et al, 2016). In vivo, the CBD alone is not sufficient to recruit Mis18BP1 to the centromere and requires the N-terminus of the protein for proper localisation (Stellfox et al, 2016). We and others have previously shown that the N-terminal 130 amino acids of Mis18BP1 are sufficient for interaction with Mis18 α/β through their Yippee domains, and Cdk1 phosphorylation of Mis18BP1 at residues T40 and S110 inhibits its interaction with Mis18 α/β to form an octamer complex consisting of 2 Mis18BP1, 4 Mis18 α and 2 Mis18 β (Pan et al, 2017; Spiller et al, 2017). Perturbing the Yippee domain-mediated hexameric assembly of Mis18 α/β (that resulted in a Mis18 α/β heterotrimer, 2 Mis18 α and 1 Mis18 β) abolished its ability to bind Mis18BP1 in vitro and in cells (Spiller et al, 2017), emphasising the requirement of maintaining correct stoichiometry of Mis18 α/β subunits. Consistent with this, artificial dimerisation of Mis18BP1, by expressing Mis18BP1 as a GST-tagged protein, enhanced the centromere localisation of Mis18BP1 (Pan et al, 2017).

Although the importance of the Mis18 complex assembly and function is well-appreciated, a structural understanding of the intermolecular interfaces responsible for the Mis18 complex assembly and their functions is yet to be identified. Here, we have characterised the structural basis of the Mis18 complex assembly using an integrative structure modelling approach that combines X-ray crystallography, Electron Microscopy (EM), Small Angle X-ray Scattering (SAXS), Cross-Linking Mass Spectrometry (CLMS), AlphaFold and computational modelling. By evaluating the structure-guided mutations in vitro and in vivo, we provide important insights into the key structural elements responsible for Mis18 complex assembly and centromere maintenance.

Results

Structural basis for the assembly of Mis18 α/β core modules

Mis18 α and Mis18 β possess two distinct but conserved structural entities, a Yippee domain and a C-terminal α -helix (Figs. 1B and EV1A,B). Mis18 α possesses an additional α -helical domain upstream of the Yippee domain (residues 39–76) as compared with Mis18 β . Previous studies have shown that Mis18 α Yippee domain can form a homodimer or a heterodimer with Mis18 β Yippee domain whereas Mis18 α/β C-terminal helices form a robust 2:1 heterotrimer (Pan et al, 2017; Spiller et al, 2017; Subramanian et al, 2016). Disrupting Yippee homo- or heterodimerisation in full-length proteins, while did not abolish their ability to form a complex, did perturb the dimerisation of Mis18 α/β heterotrimer (Spiller et al, 2017). Contrarily, intermolecular interactions involving the

C-terminal helices of Mis18 α and Mis18 β are essential for Mis18 α/β complex assembly (Nardi et al, 2016). Overall, the available biochemical data suggest the presence of at least three independent structural core modules within the Mis18 α/β complex: the Mis18 α Yippee homodimer, the Mis18 α/β Yippee heterodimer and the Mis18 α/β C-terminal helical assembly. Here, we structurally characterised these modules individually and together as a holo-complex.

Mis18 α Yippee homodimer

We previously determined a crystal structure of the Yippee domain in the only homologue of Mis18 in *S. pombe* (PDB: 5HJ0), showing that it forms a homodimer (Subramanian et al, 2016). To determine the structure of human Mis18 Yippee domains, we purified and crystallised Mis18 α_{Yippee} (residues 77–190). The crystals diffracted X-rays to about 3 Å resolution, and the structure was determined using the molecular replacement method. The final model was refined to R and R_{free} factors of 20.26% and 25.00%, respectively (Table EV1; Fig. 1C, PDB ID: 7SFZ). The overall fold of the Mis18 α_{Yippee} is remarkably similar to the previously determined *S. pombe* Mis18 α_{Yippee} homodimer structure with a RMSD of 0.92 Å (Subramanian et al, 2016). In brief, the monomeric Mis18 α_{Yippee} is formed by two antiparallel β -sheets that are held together by a Zn²⁺ ion coordinated via loops containing C-X-X-C motifs. The Mis18 α_{Yippee} dimerisation is mediated via a back-to-back arrangement of a ‘three-stranded’ β -sheet from each monomer.

Mis18 α/β Yippee heterodimer

As repeated efforts to crystallise the Mis18 $\alpha/\beta_{\text{Yippee}}$ heterodimer were not successful, using the Mis18 α_{Yippee} as a template we generated high-confidence structural models for the Mis18 $\alpha/\beta_{\text{Yippee}}$ heterodimer using RaptorX (<http://raptorx6.uchicago.edu/>) (Källberg et al, 2012) and AlphaFold2 (Jumper et al, 2021) (Figs. 1D and EV1C,D). As observed for Mis18 α_{Yippee} homodimer, the Mis18 $\alpha/\beta_{\text{Yippee}}$ heterodimerisation is also mediated via the back-to-back arrangement of the three-stranded beta sheets of Mis18 α and Mis18 β Yippee domains.

Mis18 α/β C-terminal helical assembly

Previous studies have shown that recombinantly purified C-terminal α -helices of Mis18 α and Mis18 β form a heterotrimer with two copies of Mis18 α and one copy of Mis18 β (Pan et al, 2017; Spiller et al, 2017). However, in the absence of high-resolution structural information, how Mis18 C-terminal helices interact to form a heterotrimer and how the structural arrangements of α -helices influence the relative orientations of the Yippee domains, and hence the overall architecture of the Mis18 α/β hexamer assembly, remained unclear. We purified Mis18 α spanning aa residues 191 to 233 and Mis18 β spanning aa residues 188 and 229 (Figs. 1B and EV1A,B) and crystallised the reconstituted complex. The crystals diffracted X-rays to about 2.5 Å resolution. The structure was determined using single-wavelength anomalous dispersion method. After iterative cycles of refinement and model building, the final model was refined to R and R_{free} factors of 24.77% and 27.96%, respectively (Table EV1, PDB ID: 7SFY). The asymmetric unit contained two copies of Mis18 α/β heterotrimer. The final model included Mis18 α residues 191–231 in one copy, Mis18 α residues 193 to 230 in the second copy, and Mis18 β residues 190–223 (Fig. 1E). The two Mis18 α helices interact in an antiparallel orientation, and one helix is stabilised in a slightly

curved conformation. This arrangement results in a predominantly negatively charged groove that runs diagonally on the surface formed by the Mis18 α helices (Fig. 1E,F). In contrast, the pI of the Mis18 β helix is 8.32. This charge complementarity appears to facilitate the interaction with Mis18 α , as a positively charged surface of the Mis18 β helix snugly fits in the negatively charged groove of the Mis18 α/α interface. A closer look at the intermolecular interactions reveals tight hydrophobic interactions along the ‘spine’ of the binding groove with electrostatic interactions ‘zipping-up’ both sides of the Mis18 β helix (Fig. 1F). The binding free energy calculated based on the buried accessible surface area suggests a nanomolar affinity interaction between the helices of Mis18 α and Mis18 β . It should be noted that the crystal structure presented here differs from the previously predicted models in terms of either the subunit stoichiometry (Nardi et al, 2016) or the directional arrangement of individual subunits (Mis18 α and Mis18 β in parallel orientation with the 2nd Mis18 α in an antiparallel orientation (this work) vs all parallel (Pan et al, 2019)). Although the Pan et al, 2019 model presented the 2nd Mis18 α in a parallel orientation, they did not rule out the possibility of this assembling in an antiparallel orientation within the Mis18 α/β C-terminal helical assembly (Pan et al, 2019).

Multiple surfaces of Mis18 α/β Yippee heterodimers contribute to the overall oligomeric assembly of the Mis18 complex

Full-length Mis18 α/β complex or the Mis18_{core} complex (Mis18 α –Mis18 β –Mis18BP1_{20–130}) were not amenable for structural characterisation using X-ray crystallography possibly due to their intrinsic flexibility. Consistent with this notion, the SAXS profiles collected for the Mis18 α/β Δ N (Mis18 α residues 77–End and Mis18 β residues 56–End), Mis18 α/β and Mis18_{core} complexes suggest that these complexes possess an elongated shape with flexible features (Fig. EV2; Table EV2). Hence, to understand the overall assembly of the Mis18 complex, we took an integrative structure modelling approach, combining the crystal structures of Mis18 α_{Yippee} dimer and Mis18 α/β C-terminal hetero-trimeric helical assembly together with the homology/AlphaFold modelling of Mis18 $\alpha_{\text{Yippee}}/\text{Mis18}\beta_{\text{Yippee}}$ heterodimer, negative staining EM, SAXS and CLMS analysis of the Mis18_{core} complex.

The negative staining electron micrographs of the Mis18_{core} complex cross-linked using GraFix (Kastner et al, 2008) revealed a good distribution of particles (Fig. EV3A). Particle picking, followed by a few rounds of 2D classifications, revealed classes with defined structural features (Fig. 2D). Some of the 2D projections resembled the shape of a ‘handset’ of a telephone with bulkier ‘ear’ and ‘mouth’ pieces. Differences in the relative orientation of bulkier features of the 2D projection suggested conformational heterogeneity (Figs. 2E and EV3B). The three-dimensional volumes calculated for the particles were similar ($\sim 220 \times 105 \times 80$ Å) and in agreement with the D_{max} calculated from SAXS analysis (Fig. EV2D). Consistent with these models, when we compared the theoretical SAXS scattering curve with the experimentally measured one for Mis18 α/β Δ N, we observed a good match with χ^2 value of 1.36 (Fig. 2F).

We attempted to assemble the whole Mis18 complex using AlphaFold-multimer (AF2M), with full-length Mis18 α (in purple), Mis18 β (in pink) and two small region of Mis18BP1 (20–51 and

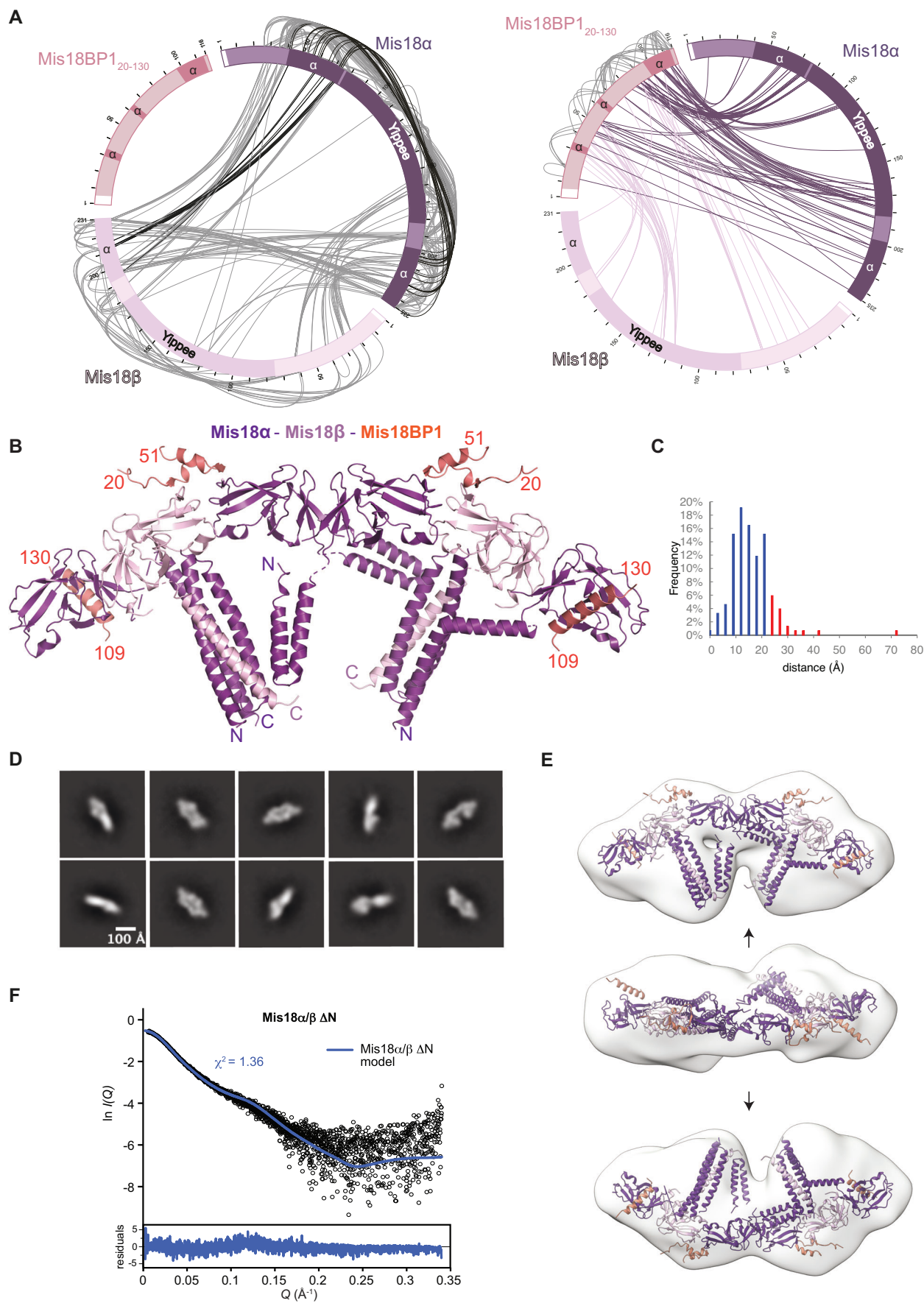


Figure 2. Mis18 complex oligomeric assembly requires multiple surfaces.

(A) Linkage map showing the sequence position and cross-linked residue pairs between the different Mis18_{core} complex subunits, Mis18 α , Mis18 β and Mis18BP1₂₀₋₁₃₀. Left panel highlights cross-linked residues between Mis18 α and Mis18 β . Black lines highlight cross-links between N-terminal α -helix of Mis18 α and C-terminal helical regions of proteins. Right panel highlights cross-links observed between (i) Mis18BP1₂₀₋₁₃₀ and Mis18 α (purple), (ii) Mis18BP1₂₀₋₁₃₀ and Mis18 β (light pink), (iii) Mis18BP1₂₀₋₁₃₀ self cross-links (light grey). White boxes represent residual residues left over from tag cleavage. Dark boxes show Yippee domains and regions of α -helices. (B) Model of the Mis18_{core} complex generated using partial structures determined using X-ray crystallography and AlphaFold2 (Jumper et al, 2021) and cross-linking restrained molecular docking in EM maps. Mis18BP1 shown in salmon, Mis18 α in purple and Mis18 β in light pink. (C) Histograms show the percentage of satisfied or violated cross-links for structures modelled using MODELLER (Sali and Blundell, 1993). (D) Representative images of 2D classes from Mis18_{core} particles picked using CryoSPARC (Punjani et al, 2017). Scale bar shows 100 Å. (E) Model (Class I) generated for Mis18_{core} from negative staining EM analysis. This shows that the overall shapes of the Mis18_{core} resemble a telephone handset with 'ear' and 'mouth' pieces. Arrows denote the different orientations shown. (F) Theoretical SAXS scattering curves of Mis18 α / β Δ N model compared to experimental data.

109–130; in salmon) (preprint: Evans et al, 2022). The AF2M converged towards a structure with six Yippee domains stacked in a line-like arrangement in the Mis18 α _{Yippee}-Mis18 β _{Yippee}-Mis18 α _{Yippee}-Mis18 α _{Yippee}-Mis18 β _{Yippee}-Mis18 α _{Yippee} order and two triple helix bundles, each formed by C-terminal α -helices of 2 copies of Mis18 α and 1 copy of Mis18 β . However, the modelled two helical bundles had all three helices in a parallel orientation that is not supported by our crystal structures (Fig. 1E) and cross-links (Fig. 2A). We modified the relative orientation of the helices to match the crystal structure by superposing the latter on the AF2M model (Figs. 2B,C and EV3B). Using cross-links and docking we added the N-terminal helices of the Mis18 α . Cross-linking data indicates that these helices have multiple orientations with respect to the rest of the structure, contacting both Yippee domains and triple helix bundles. The linker between the Yippee domain and the C-terminal helix is the shortest in Mis18 β (Fig. 1B), further supporting the arrangement of the Yippee domains within the assembly. The integrative model of the Mis18 complex fits well in the EM map. Interestingly, the serial arrangement of the Yippee domains utilises the second Yippee dimerisation interface observed in the crystal packing of both human Mis18 α _{Yippee} and *S. pombe* Mis18_{Yippee} (Fig. EV3C, highlighted by zoom in view). Accordingly, disrupting this interface by mutating Mis18 α residues C154 and D160 (Fig. EV3C) perturbed Mis18 oligomerisation as evidenced by SEC analysis (Fig. EV3D).

Mis18 α oligomerisation via the C-terminal helical bundle assembly is essential for Mis18 α / β centromere localisation and new CENP-A loading

Although the subunit stoichiometry and the arrangement of Mis18 α / β C-terminal helices within the helical bundle proposed by Nardi et al, 2016 are different from the data presented here, the Mis18 α residues (I201, L205, L212, L215 and L219) that were predicted by them to stabilise the helical bundle do indeed form the 'spine' of the hydrophobic core running along the triple-helical bundle (Fig. 1E,F). Mutating these residues perturbed the ability of Mis18 α tethered at an ectopic LacO site to facilitate CENP-A deposition at the tethering site (Nardi et al, 2016). However, how these Mis18 α mutants perturb the oligomeric structure of the Mis18 α / β C-terminal helical bundle and how this structural perturbation affects CENP-A loading at endogenous centromeres remain as open questions.

To address these questions, we first tested these mutants using in vitro amylose pull-down assays by mixing recombinantly purified WT and mutant His-MBP-Mis18 β ₁₈₈₋₂₂₉ and His-SUMO-Mis18 α ₁₉₁₋₂₃₃ proteins. Mutating these residues to Ala

(Mis18 α _{I201A/L205A} and Mis18 α _{L212A/L215A/L219A}) or Asp (Mis18 α _{I201D/L205D}) abolished the ability of Mis18 α α -helix to interact with Mis18 β ₁₈₈₋₂₂₉ (Fig. EV4A). Co-immunoprecipitation (Co-IP) assays using an anti-Mis18 α antibody were performed on cells where endogenous Mis18 α was depleted, and Mis18 α -mCherry was co-expressed with Mis18 β -GFP to check for complex formation (Fig. EV4B). In line with our in vitro pull-downs, the Co-IPs using a Mis18 α antibody revealed that Mis18 α _{WT}-mCherry interacted with Mis18 β -GFP while Mis18 α _{I201A/L205A} and Mis18 α _{L212A/L215A/L219A} mutants did not (Fig. EV4B). SEC-MALS analysis of His-SUMO tagged Mis18 α ₁₈₈₋₂₃₃ showed that on its own, Mis18 α WT protein can form a dimer, whilst introducing I201A/L205A or L212A/L215A/L219A results in both proteins forming a monomer (Fig. EV4C). To evaluate the role of C-terminal helical bundle assembly, mediated via the Mis18 α oligomerisation, on centromere localisation of Mis18 α and Mis18 β and CENP-A deposition, these mutants were further tested in HeLa cells.

HeLa Mis18 β -GFP CENP-A-SNAP cells (McKinley and Cheeseman, 2014) were depleted of endogenous Mis18 α by siRNA (Fig. EV4D) and simultaneously rescued with either WT or mutant Mis18 α -mCherry (Fig. EV4E), then visualised by immunofluorescence along with ACA. Unlike Mis18 α _{WT}, the Mis18 α mutants (Mis18 α _{I201A/L205A}, Mis18 α _{I201D/L205D} and Mis18 α _{L212A/L215A/L219A}) all failed to localise to centromeres (Fig. 3A). As expected, Mis18 β -GFP co-expression showed co-localisation between Mis18 β _{WT} with Mis18 α _{WT}. However, in cells expressing Mis18 α _{I201A/L205A}, Mis18 α _{I201D/L205D} and Mis18 α _{L212A/L215A/L219A}, Mis18 β could no longer co-localise with Mis18 α at the centromere. Together, this confirms that Mis18 β depends on its interaction with Mis18 α and the formation of the C-terminal triple helical assembly to localise at centromeres.

We then evaluated the impact of Mis18 α mutants not capable of forming the C-terminal helical bundle on new CENP-A deposition. We did this by performing a Quench-Chase-Pulse CENP-A-SNAP Assay according to Jansen et al (Jansen et al, 2007) (Fig. 3B). HeLa CENP-A-SNAP cells were depleted of endogenous Mis18 α and rescued with either Mis18 α _{WT} or Mis18 α mutants (Mis18 α _{I201A/L205A}, Mis18 α _{I201D/L205D} and Mis18 α _{L212A/L215A/L219A}). The existing CENP-A was blocked with a non-fluorescent substrate of the SNAP, and the new CENP-A deposition in the early G1 phase was visualised by staining with the fluorescent substrate of the SNAP. Mis18 α _{WT} rescued new CENP-A deposition to levels compared to that of control siRNA (Fig. 3C). However, Mis18 α _{I201A/L205A}, Mis18 α _{I201D/L205D} and Mis18 α _{L212A/L215A/L219A} abolished new CENP-A loading almost completely, indicating that the formation of the Mis18 triple-helical bundle is essential for CENP-A deposition (Fig. 3C).

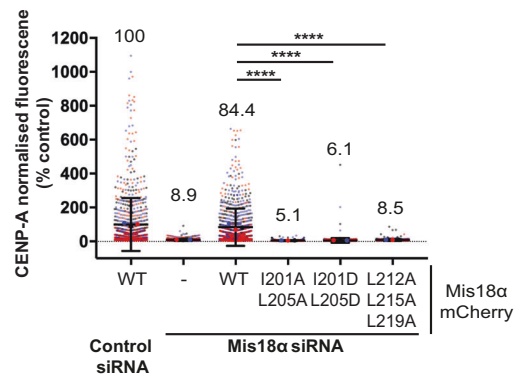
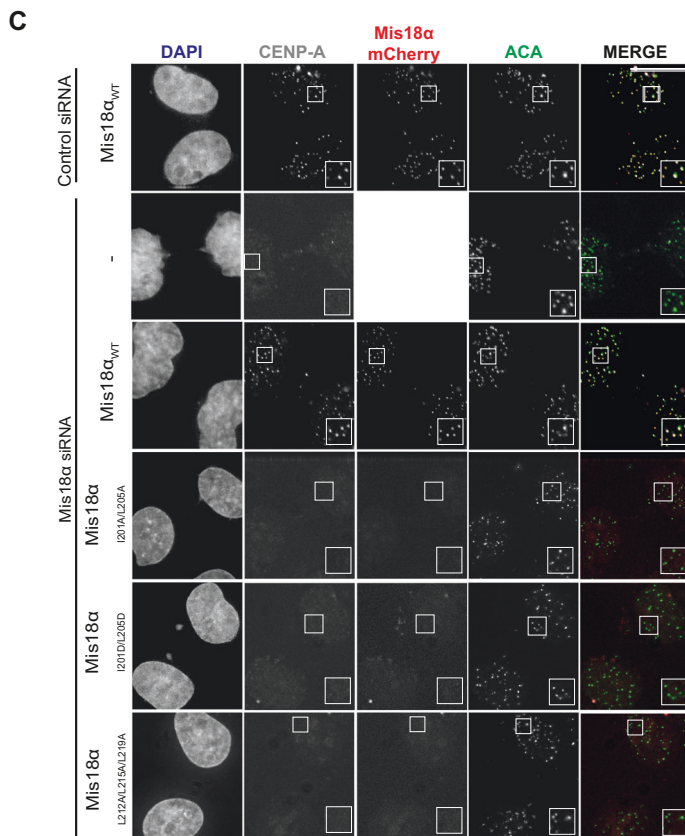
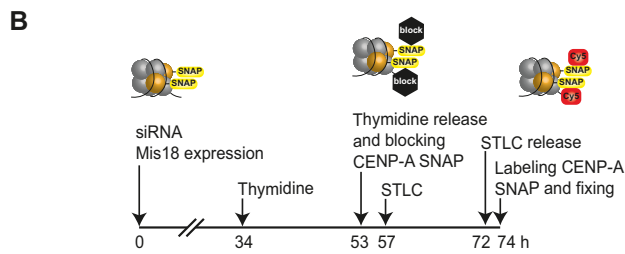
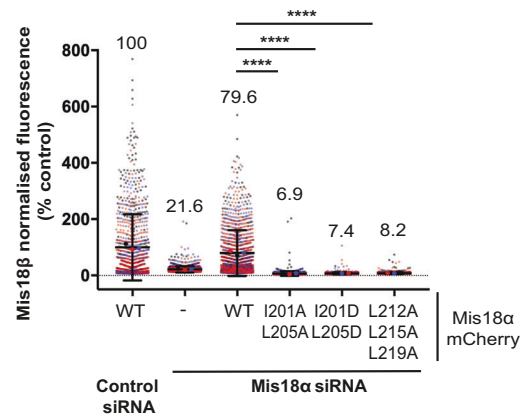
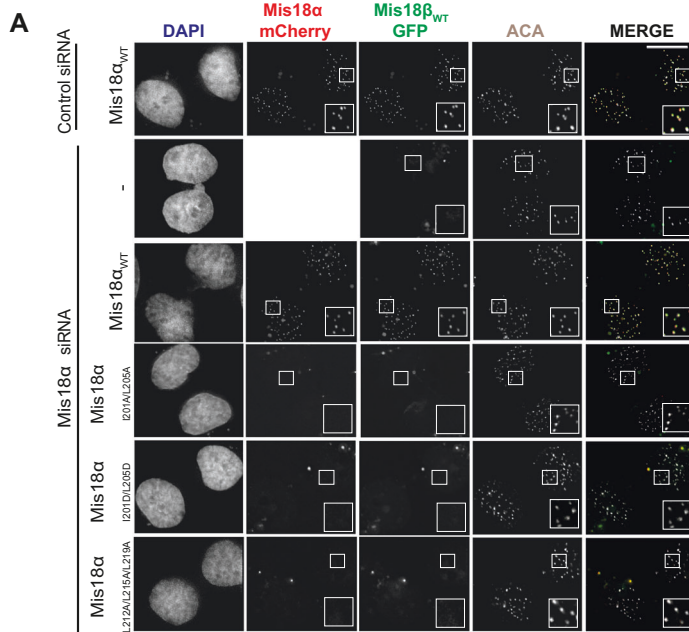


Figure 3. Mis18 α mutations disrupting the Mis18 α / β triple helical assembly result in loss of Mis18 α / β centromere localisation and CENP-A deposition.

(A) Representative fluorescence images (left panel) and quantification (right panel) assessing the ability of Mis18 α _{WT}-mCherry ($n = 1236$), Mis18 α _{L201A/L205A}-mCherry ($n = 1368$), Mis18 α _{L201D/L205D}-mCherry ($n = 1373$) and Mis18 α _{L212A/L215A/L219A}-mCherry ($n = 1383$) to co-localise with Mis18 β -GFP at endogenous centromeres in HeLa (Mann-Whitney U test; **** $P \leq 0.0001$). Cells were co-transfected with either control ($n = 944$) or Mis18 α siRNA (with no transfected Mis18 α -mCherry $n = 1572$), as stated, in three independent experiments shown in black, blue and red. Error bars show mean \pm SD. Scale bars, 10 μ m. All conditions have been normalised to control conditions: cells transfected with control siRNA and Mis18 α _{WT}-mCherry. (B) Schematic representation of the experimental set-up used to evaluate the effect of Mis18 α and Mis18 β mutants on new CENP-A-SNAP loading. (C) Representative fluorescence images (left panel) and quantification (right panel) assessing the ability of Mis18 α _{WT}-mCherry ($n = 896$), Mis18 α _{L201A/L205A}-mCherry ($n = 886$), Mis18 α _{L201D/L205D}-mCherry ($n = 1434$) and Mis18 α _{L212A/L215A/L219A}-mCherry ($n = 1188$) to deposit new CENP-A-SNAP at endogenous centromeres (Mann-Whitney U test; **** $P \leq 0.0001$). Cells were co-transfected with either control ($n = 852$) or Mis18 α siRNA (with no transfected Mis18 α -mCherry $n = 1736$), as stated, in three independent experiments shown in black, blue and red. Error bars show mean \pm SD. Scale bars, 10 μ m. All conditions have been normalised to control conditions: cells transfected with control siRNA and Mis18 α _{WT}-mCherry. Source data are available online for this figure.

Mis18 α associates with the centromere independently of Mis18 β and can deposit CENP-A, but efficient CENP-A loading requires Mis18 β

We again performed amylose in vitro pull-down assays, using His-SUMO-Mis18 α _{L191-233} WT and mutant His-MBP-Mis18 β _{L188-229} proteins, to assess the ability of Mis18 β mutant to form a triple-helical bundle with Mis18 α . Based on our X-ray crystal structure (Fig. 1E), we identified one cluster (L199/I203) in Mis18 β and observed that mutating these residues to either Ala (Mis18 β _{L199A/I203A}) or Asp (Mis18 β _{L199D/I203D}) either reduced or abolished its ability to interact with Mis18 α _{L191-233} (Fig. 4A, left panel). Co-IP analysis using an anti-Mis18 α antibody was performed on cells where endogenous Mis18 β was depleted, and Mis18 β -GFP was expressed along Mis18 α -mCherry to check for complex formation. Western blot analysis showed that Mis18 β _{WT} could interact with Mis18 α -mCherry and that the ability of Mis18 β _{L199D/I203D} to interact with Mis18 α was reduced (Fig. 4A, right panel).

To assess the contribution of Mis18 β for the centromere association and function of Mis18 α , we evaluated the Mis18 β mutant (Mis18 β _{L199D/I203D}), which cannot form the triple-helical assembly with Mis18 α , in siRNA rescue assays by expressing Mis18 β -GFP-tagged proteins in a mCherry-Mis18 α cell line (McKinley and Cheeseman, 2014). Depletion of endogenous Mis18 β and simultaneous transient expression of Mis18 β _{WT}-GFP led to co-localisation of Mis18 β with Mis18 α at centromeres (Figs. 4B and EV4D,E). Under these conditions, Mis18 β _{WT}-GFP levels at centromeres were comparable to that of the control siRNA. Whereas Mis18 β _{L199D/I203D} failed to localise at the centromeres. Strikingly, Mis18 β _{L199D/I203D} perturbed centromere association of Mis18 α only moderately (Fig. 4B). This suggests that Mis18 α can associate with centromeres in a Mis18 β independent manner.

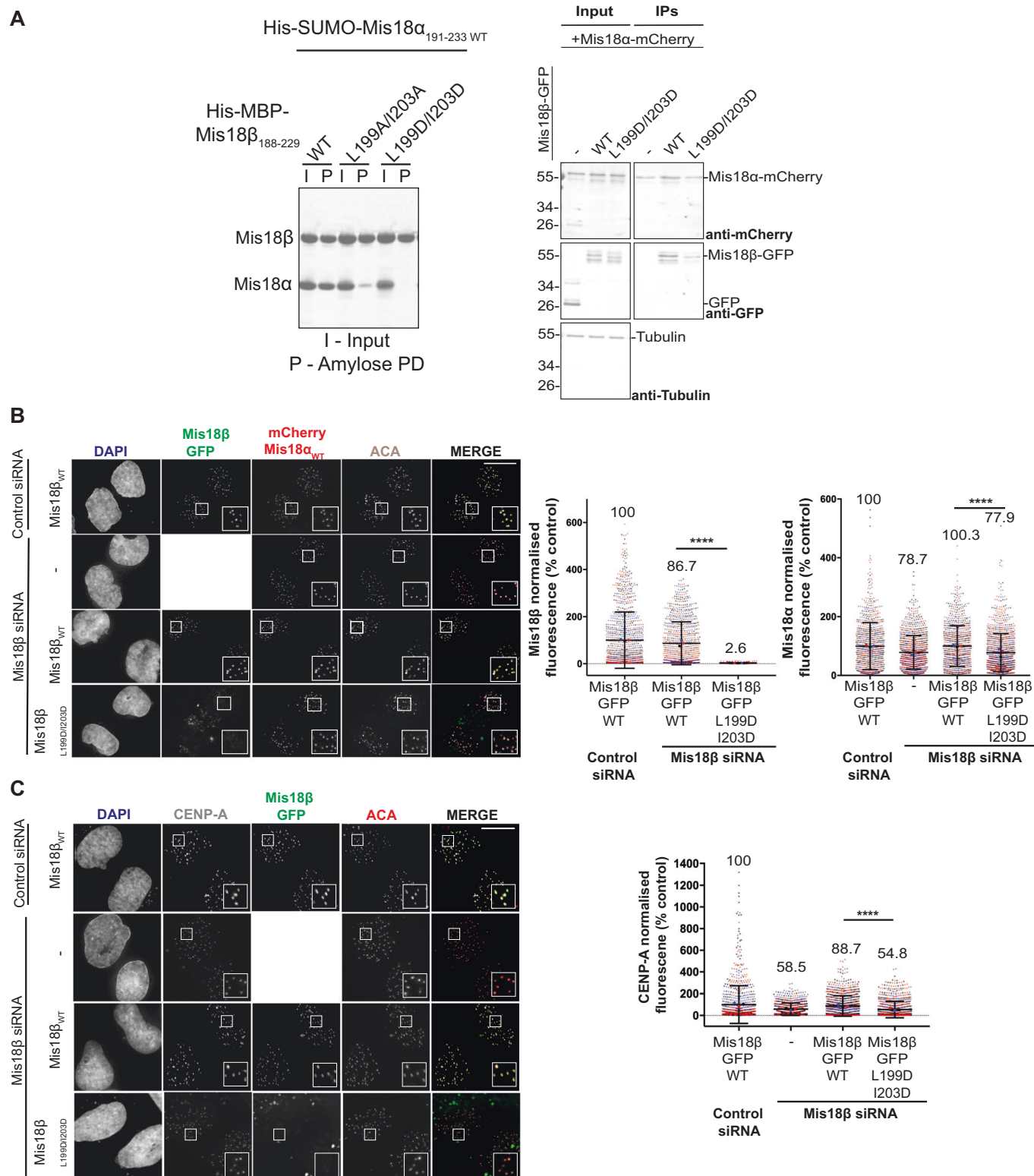
Next, we assessed the contribution of Mis18 β for CENP-A deposition in the Quench-Chase-Pulse CENP-A-SNAP assay described above. Endogenous Mis18 β was depleted using siRNA, and Mis18 β _{WT} and Mis18 β _{L199D/I203D} were transiently expressed as GFP-tagged proteins in HeLa cells expressing CENP-A-SNAP. Mis18 β _{WT} rescued new CENP-A deposition to comparable levels to the ones observed in the control siRNA-Mis18 β WT condition (Fig. 4C). Interestingly, unlike the Mis18 α mutants (Mis18 α _{L201A/L205A}, Mis18 α _{L201D/L205D} and Mis18 α _{L212A/L215A/L219A}), Mis18 β _{L199D/I203D} did not abolish new CENP-A loading but reduced the levels only moderately.

Together, these analyses demonstrate that Mis18 α can associate with centromeres and deposit new CENP-A independently of Mis18 β . However, efficient CENP-A loading requires Mis18 β .

Structural basis for centromere recruitment of Mis18 α / β by Mis18BP1

Previous studies have shown that the first 130 amino acids of Mis18BP1 are required to bind Mis18 α / β _{Yippeee} domains (Spiller et al, 2017). However, how Mis18 α / β _{Yippeee} domains recognise Mis18BP1 is not clear. Our structural analysis based on the AlphaFold model suggests that two Mis18BP1 fragments, a short helical segment spanning aa residues 109–130 (Mis18BP1₁₀₉₋₁₃₀) and a region spanning aa residues 20–51 (Mis18BP1₂₀₋₅₁) interact with Mis18 α _{Yippeee} domain and with an interface formed between Mis18 α / β _{Yippeee} heterodimers, respectively (Fig. 5A). Mis18BP1₁₀₉₋₁₃₀ binds at a hydrophobic pocket of the Mis18 α _{Yippeee} domain formed by amino acids L83, F85, W100, I110, V172 and I175. This hydrophobic pocket is surrounded by hydrophilic amino acids E103, D104, T105, S169 E171 facilitating additional electrostatic interactions with Mis18BP1₁₀₉₋₁₃₀ (Fig. 5A). Mis18BP1₂₀₋₅₁ contains two short β strands that interact at Mis18 α / β _{Yippeee} interface extending the six-stranded- β sheets of both Mis18 α and Mis18 β Yippeee domains. This provides the structural rationale for why Yippeee domains-mediated Mis18 α / β hetero-hexamersation is crucial for Mis18BP1 binding (Spiller et al, 2017). Notably, the two Cdk1 phosphorylation sites on Mis18BP1 (T40 and S110) that we and others have shown to disrupt Mis18 complex assembly (Pan et al, 2017; Spiller et al, 2017) lie directly within the Mis18 α / β binding interface predicted by this model providing the structural basis for Cdk1 mediated regulation of Mis18 complex assembly. Consistent with this model, several cross-links observed between Mis18BP1 and Mis18 α and Mis18 β map to these residues. Mutating the negatively charged amino acid cluster of Mis18 α (E103, D104 and T105) that is juxtaposed to Mis18BP1₁₀₉₋₁₃₀ in a TetR-eYFP-Mis18 α vector (TetR-eYFP-Mis18 α _{E103R/D104R/T105R}) transfected in HeLa cells with an ectopic synthetic alphoid^{test} array integrated in a chromosome arm significantly perturbed Mis18 α 's ability to recruit Mis18BP1₂₀₋₁₃₀-mCherry to the tethering site as compared to Mis18 α _{WT} (Fig. 5B).

Furthermore, we probed the effects of perturbing Mis18 α -Mis18BP1 interaction on endogenous centromeres. We depleted Mis18 α in a cell line that stably expresses CENP-A-SNAP and allows inducible expression of GFP-Mis18BP1 (McKinley and Cheeseman, 2014). We then assessed the ability of transfected Mis18 α -mCherry to co-localise with Mis18BP1 at centromeres. Depletion of Mis18 α and simultaneous expression of either Mis18 α _{WT}-mCherry or Mis18 α _{E103R/D104R/T105A}-mCherry revealed that, unlike Mis18 α _{WT}, Mis18 α _{E103R/D104R/T105A} failed to localise at endogenous centromeres (Fig. 5C, middle panel). We also observed



a slight decrease in the levels of GFP-Mis18BP1 at the centromere when Mis18 $\alpha_{E103R/D104R/T105A}$ was expressed as compared to Mis18 α_{WT} (Fig. 5C, right panel). Consistent with the observation of reduced centromeric Mis18 α , when Mis18 $\alpha_{E103R/D104R/T105A}$

mCherry is expressed, the quantification of new CENP-A deposition in HeLa cell expressing CENP-A-SNAP showed a significant reduction of new CENP-A deposition at the centromere indicating that the interaction of Mis18 α with Mis18BP1 is essential for

Figure 4. Mis18 α associates with centromeres in a Mis18 β -independent manner but requires Mis18 β for efficient CENP-A loading.

(A) Left panel shows SDS-PAGE analysis of cobalt and amylose pull-down of His-MBP-Mis18 β ₁₈₈₋₂₂₉ WT and mutants with His-SUMO-Mis18 α ₁₉₁₋₂₃₃. SDS-PAGE shows protein bound to nickel resin as input (I) and protein bound to amylose resin to assess interaction (P). Right panel shows Western blot analysis of co-immunoprecipitation (Co-IP) experiments using Mis18 α antibody to test interaction of Mis18 α -mCherry and Mis18 β -GFP with and without mutations in the C-terminal α -helices or GFP as a control. Top panel shows blot against mCherry, middle panel shows blot against GFP, and bottom panel shows blot against tubulin as loading control. (B) Representative fluorescence images (left panel) and quantification (right panel) used to evaluate the ability of Mis18 β _{WT}-GFP ($n = 963, 927$) and Mis18 β _{L199D/1203D}-GFP ($n = 1312, 1221$) to co-localise with mCherry-Mis18 α at endogenous centromeres. Middle panel, quantification of Mis18 β signal. Right panel, quantification of Mis18 α signal (Mann-Whitney U test; **** $P \leq 0.0001$). Cells were co-transfected with either control ($n = 1131, 935$) or Mis18 β siRNA (with no transfected Mis18 β -GFP $n = 1170$), in three independent experiments shown in black, blue and red. Error bars show mean \pm SD. (C) Representative fluorescence images (left panel) and quantification (right panel) used to evaluate the ability of Mis18 β _{WT}-GFP ($n = 1036$) and Mis18 β _{L199D/1203D} GFP ($n = 947$) to deposit new CENP-A-SNAP at endogenous centromeres (Mann-Whitney U test; **** $P \leq 0.0001$). Cells were co-transfected with either control ($n = 840$) or Mis18 β siRNA (with no transfected Mis18 β -GFP $n = 824$), as stated, in three independent experiments shown in black, blue and red. Error bars show mean \pm SD. Scale bars, 10 μ m. All conditions have been normalised to control conditions: cells transfected with control siRNA and Mis18 β _{WT}-GFP. Source data are available online for this figure.

centromeric recruitment of the Mis18 complex and for CENP-A loading (Fig. 5D).

Discussion

Mis18 complex assembly is a central process essential for the recruitment of CENP-A/H4 bound HJURP and the subsequent CENP-A deposition at centromeres (Dunleavy et al, 2009; Fujita et al, 2007; Jansen et al, 2007). Thus far, several studies, predominantly biochemical and cellular, have characterised interactions and functions mediated by the two distinct structural domains of the Mis18 proteins, the Yippee and C-terminal α -helical domains of Mis18 α and Mis18 β (Nardi et al, 2016; Pan et al, 2017; Spiller et al, 2017; Stellfox et al, 2016). Some of the key conclusions of these studies include: (1) Mis18 α/β is a hetero-hexamers made of 4 Mis18 α and 2 Mis18 β ; (2) The Yippee domains and C-terminal α -helices of Mis18 α and Mis18 β have the intrinsic ability to homo- or hetero-oligomerise, and form three distinct oligomeric modules in different copy numbers—a Mis18 α _{Yippee} homodimer, two copies of Mis18 α/β _{Yippee} heterodimers and two heterotrimers made of Mis18 α/β C-terminal helices (2 Mis18 α and 1 Mis18 β); (3) the two copies of Mis18 α/β _{Yippee} heterodimers each bind one Mis18BP1₂₀₋₁₃₀ and form a hetero-octameric Mis18_{core} complex (Mis18 α /Mis18 β /Mis18BP1₂₀₋₁₃₀: a Mis18 α/β hetero-hexamers bound to 2 copies of Mis18BP1₂₀₋₁₃₀). However, no experimentally determined structural information is available for the human Mis18 complex. This is crucial to identify the amino acid residues essential for the assembly of Mis18 α/β and the holo-Mis18 complexes and to determine the specific interactions that are essential for the localisation of Mis18 complex to centromeres and its function.

Here, we have taken an integrative structural approach that combines X-ray crystallography, electron microscopy and homology modelling with cross-linking mass spectrometry to characterise the structure of the Mis18 complex. Our analysis shows that Mis18 α/β heterotrimer is stabilised by the formation of a triple-helical bundle with a Mis18 α/β _{Yippee} heterodimer and Mis18 α _{Yippee} monomer arranged as a linear array. Two such Mis18 α/β heterotrimers assemble as a hetero-hexamers via the homodimerisation of the Mis18 α _{Yippee} domains. The crystal structure of Mis18 α/β _{C-term} triple-helical structure allowed us to design several separation-of-function Mis18 α and Mis18 β mutants. These mutations specifically perturb the ability of Mis18 α or Mis18 β to assemble into the helical bundle, while retaining their other

functions, if there are any. Functional evaluation of these mutants in cells has provided important new insights into the molecular interdependencies of the Mis18 complex subunits. Particularly, the observations that: (1) Mis18 α can associate with centromeres and deposit CENP-A independently of Mis18 β , and (2) depletion of Mis18 β or disrupting the incorporation of Mis18 β into the Mis18 complex, while does not abolish CENP-A loading, reduces the CENP-A deposition amounts, questions the consensus view that Mis18 α and Mis18 β always function as a single structural entity to exert their function to maintain centromere maintenance.

Whilst proteins involved in CENP-A loading have been well-established, the mechanism by which the correct levels of CENP-A are controlled is yet to be thoroughly explored and characterised. The data presented here suggest that Mis18 β mainly contributes to the quantitative control of centromere maintenance by ensuring the right amount of CENP-A deposition at centromeres. We also note that the Mis18 β mutant, which cannot interact with Mis18 α , moderately reduced Mis18 α levels at centromeres, and hence, it is possible that Mis18 β ensures the correct level of CENP-A deposition by facilitating optimal Mis18 α centromere recruitment. Future studies will focus on dissecting the mechanisms underlying the Mis18 β -mediated control of CENP-A loading amounts along with any other mechanisms involved.

Previously published work identified amino acid sequence similarity between the N-terminal region of Mis18 α and R1 and R2 repeats of the HJURP that mediates Mis18 α/β interaction (Pan et al, 2019). Deletion of the Mis18 α N-terminal region enhanced HJURP interaction with the Mis18 complex (Pan et al, 2019). Here, we show that the N-terminal helical region of Mis18 α makes extensive contact with the C-terminal helices of Mis18 α and Mis18 β , which had previously been shown to mediate HJURP binding by Pan et al, 2019. Collectively these observations suggest that the N-terminal region of Mis18 α might directly interfere with HJURP– Mis18 complex interaction. Two independent recent studies (preprint: Conti et al, 2024; preprint: Parashara et al, 2024) reveal that this is indeed the case and a Plk1-mediated phosphorylation cascade involving several phosphorylation and binding events of the Mis18 complex subunits relieve the intramolecular interactions between the Mis18 α N-terminal helical region and the HJURP binding surface of the Mis18 α/β C-terminal helical bundle. This facilitates robust HJURP–Mis18 α/β interaction in vitro and efficient HJURP centromere recruitment and CENP-A loading in cells. Overall, these studies also highlight the importance of the critical structural insights into the Mis18 complex we report here.

One of the key outstanding questions in the field is how does the Mis18 complex associate with the centromere. Previous studies identified

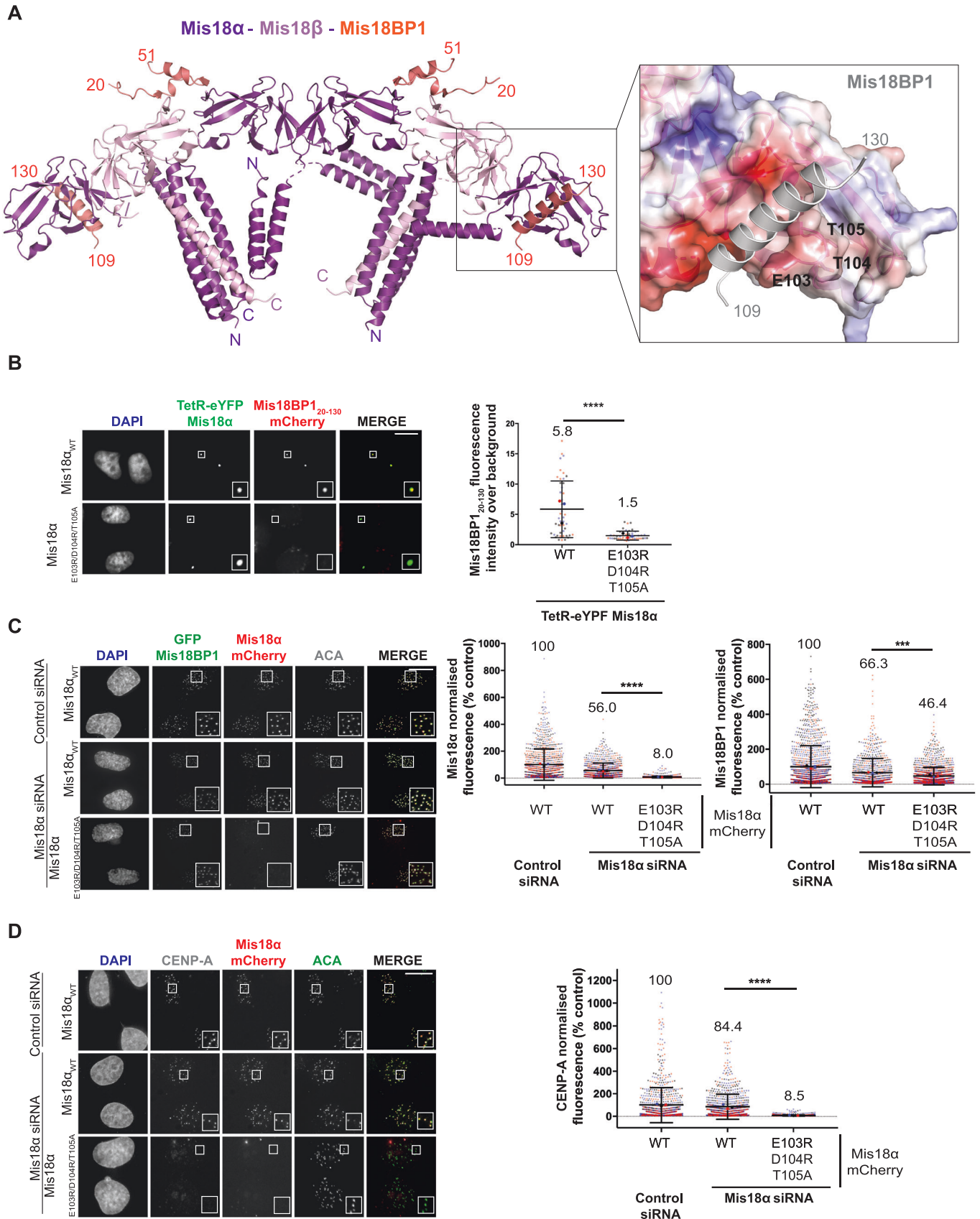


Figure 5. Disrupting the Mis18BP1 binding interface of Mis18 α prevents its centromere localisation and CENP-A deposition.

(A) Mis18 α /Mis18 β model and its surface representation coloured based on electrostatic surface potential (zoom panel), highlighting the residues proposed to be involved in Mis18BP1 binding. Mis18 α shown in purple, Mis18 β shown in light pink and Mis18BP1 shown in salmon or grey for the zoom panel for clarity. (B) Representative images and quantification showing the recruitment of either Mis18BP1₂₀₋₁₃₀-mCherry by different Mis18 α constructs (WT and mutant) tethered to the aliphoid^{test} array in HeLa 3-8. Tethering of TetR-eYFP-Mis18 α _{WT} ($n = 45$) and TetR-eYFP-Mis18 α _{E103R/D104R/T105A} ($n = 46$) testing recruitment of Mis18BP1₂₀₋₁₃₀ mCherry (Mann-Whitney U test; **** $P \leq 0.0001$). Data from three independent experiments shown in black, blue and red. Error bars show mean \pm SD. Scale bars, 10 μ m. (C) Representative fluorescence images (left panel) and quantifications (right panel) evaluating the ability of Mis18 α _{WT}-mCherry ($n = 985, 856$) and Mis18 α _{E103R/D104R/T105A} ($n = 1497, 1511$) to co-localise with GFP-Mis18BP1 at endogenous centromeres. Middle panel, quantification of Mis18 α signal and right panel, quantification of Mis18BP1 signal (Mann-Whitney U test; *** $P \leq 0.001$, **** $P \leq 0.0001$). Cells were co-transfected with either control ($n = 1016, 1403$) or Mis18 α siRNA, as stated, in three independent experiments shown in black, blue and red. Error bars show mean \pm SD. Scale bars, 10 μ m. (D) Representative fluorescence images (left panel) and quantifications (right panel) evaluating the ability of Mis18 α _{WT}-mCherry ($n = 896$) and Mis18 α _{E103R/D104R/T105A} ($n = 1430$) to deposit new CENP-A-SNAP at endogenous centromeres (Mann-Whitney U test; **** $P \leq 0.0001$). Cells were co-transfected with either control ($n = 852$) or Mis18 α siRNA, as stated, in three independent experiments shown in black, blue and red. Error bars show mean \pm SD. Scale bars, 10 μ m. All conditions have been normalised to control conditions: cells transfected with control siRNA and Mis18 α _{WT}-mCherry. Source data are available online for this figure.

CCAN subunits CENP-C and CENP-I as major players mediating the centromere localisation of the Mis18 complex mainly via Mis18BP1 (Dambacher et al, 2012; Moree et al, 2011; Shono et al, 2015), although Mis18 β subunit has also been suggested to interact with CENP-C (Stellfox et al, 2016). Within the Mis18 complex, we and others have shown that the Mis18 α / β _{Yippe} heterodimers can directly interact with Mis18BP1. Here our structural analysis allowed us to map the interaction interface mediating Mis18 α / β -Mis18BP1 binding. Perturbing this interface on Mis18 α completely abolished Mis18 α centromere localisation and reduced Mis18BP1 centromere levels. These observations show that Mis18 α associates with the centromere mainly via Mis18BP1, and assembly of the Mis18 complex itself is crucial for its efficient centromere association, as previously suggested. Future work aimed at characterising the intermolecular contact points between the subunits of the Mis18 complex, centromeric chromatin and CCAN components and understanding if the Mis18 complex undergoes any conformational and/or compositional variations upon centromere association and/or during CENP-A deposition process, will be crucial to delineate the mechanisms underpinning the centromere maintenance.

Methods

Plasmids

For crystallisation, a polycistronic expression vector for the C-terminal coiled-coil domains of Mis18 α (residues 191–233, Mis18 α _{C-term}) and Mis18 β (residues 188–229, Mis18 β _{C-term}) were produced with the N-terminal 6His-SUMO- (His-SUMO) and 6His-MBP-tags (His-MBP), respectively. Mis18 α _{Yippe} (residues 77–190) was cloned into the pET3a vector with the N-terminal 6His-tag.

For all other recombinant proteins, codon-optimised sequences (GeneArt) for Mis18 α and Mis18 β were cloned into pET His6 TEV or pET His6 msfGFP TEV (9B Addgene plasmid #48284, 9GFP Addgene plasmid #48287, a kind gift from Scott Gradia), respectively. They were combined to make a single polycistronic plasmid. The boundaries of ΔN for Mis18 α and Mis18 β were 77–187 and 56–183 Mis18BP1₂₀₋₁₃₀ was cloned in pEC-K-3C-His-GST and pET His6 MBP TEV (9C Addgene plasmid #48286).

Non-codon-optimised sequences were amplified from a human cDNA library (MegaMan human transcription library, Agilent). Mis18 α , Mis18 β and Mis18BP1₂₀₋₁₃₀ were cloned into pcDNA3 mCherry LIC vector, pcDNA3 GFP LIC vector (6B Addgene

plasmid #30125, 6D Addgene plasmid #30127, a kind gift from Scott Gradia) and TetR-eYFP-IRES-Puro vector as stated. All mutations were generated following QuikChange site-directed mutagenesis protocol (Stratagene), using primers in Table EV3.

Expression and purification of recombinant proteins

For crystallisation, both Mis18 α / β _{C-term} domains and Mis18 α _{Yippe} were transformed and expressed in *Escherichia coli* BL21 (DE3) using the auto-inducible expression system (Studier, 2005). The cells were harvested and resuspended in the lysis buffer containing 30 mM Tris-HCl pH 7.5, 500 mM NaCl and 5 mM β -mercaptoethanol with protease inhibitor cocktails. The resuspended cells were lysed using the ultra-sonication method and centrifuged at 20,000 $\times g$ for 50 min at 4 $^{\circ}$ C to remove the cell debris. After 0.45 μ m filtration of the supernatant, the lysate was loaded into the cobalt affinity column (New England Biolabs) and eluted with a buffer containing 30 mM Tris-HCl pH 7.5, 500 mM NaCl, 5 mM β -mercaptoethanol, and 300 mM imidazole. The eluate was loaded into the amylose affinity column (New England Biolabs) and washed with a buffer containing 30 mM Tris-HCl pH 7.5, 500 mM NaCl and 5 mM β -mercaptoethanol. To cleave the His-MBP tag, on-column cleavage was performed by adding Tobacco Etch Virus (TEV) protease (1:100 ratio) into the resuspended amylose resin and incubated overnight at 4 $^{\circ}$ C. The TEV cleavage released the untagged Mis18 α / β _{C-term} domains in solution, and the flow-through fraction was collected and concentrated using a Centricon (Millipore). The protein was loaded onto a HiLoad[™] 16/600 Superdex[™] 200 column (GE Healthcare) equilibrated with a buffer containing 30 mM Tris-HCl pH 7.5, 100 mM NaCl and 1 mM TCEP. To further remove the contaminated MBP tag, the sample was re-applied into the amylose affinity column, and the flow-through fraction was collected and concentrated to 20 mg/ml for the crystallisation trial. SeMet (selenomethionine) incorporated Mis18 α / β _{C-term} domains were expressed with PASM-5052 auto-inducible media (Studier, 2005). The SeMet-substituted Mis18 α / β _{C-term} domains were purified using the same procedure described above.

The purification of His tagged Mis18 α _{Yippe} employed the same purification method used for Mis18 α / β _{C-term} domains except for the amylose affinity chromatography step. The purified Mis18 α _{Yippe} from the HiLoad[™] 16/600 Superdex[™] 200 chromatography was concentrated to 13.7 mg/ml with the buffer containing 30 mM Tris-HCl pH 7.5, 100 mM NaCl and 1 mM TCEP.

All other proteins were expressed in *Escherichia coli* BL21 (DE3) Gold cells using LB. After reaching an O.D. ~0.6 at 37 °C, cultures were cooled to 18 °C and induced with 0.35 mM IPTG overnight. The His-Mis18 α /His-GFP-Mis18 β complex was purified by resuspending the pellet in a lysis buffer containing 20 mM Tris-HCl pH 8.0 at 4 °C, 250 mM NaCl, 35 mM imidazole pH 8.0 and 2 mM β -mercaptoethanol supplemented with 10 μ g/ml DNase, 1 mM PMSF and cComplete™ EDTA-free (Sigma). After sonication, clarified lysates were applied to a 5 ml HisTrap™ HP column (GE Healthcare) and washed with lysis buffer followed by a buffer containing 20 mM Tris-HCl pH 8.0 at 4 °C, 1 M NaCl, 35 mM imidazole pH 8.0, 50 mM KCl, 10 mM MgCl₂, 2 mM ATP and 2 mM β -mercaptoethanol and then finally washed with lysis buffer. The complex was then eluted with 20 mM Tris-HCl pH 8.0 at 4 °C, 250 mM NaCl, 500 mM imidazole pH 8.0 and 2 mM β -mercaptoethanol. Fractions containing proteins were pooled, and TEV was added (if needed) whilst performing overnight dialyses against 20 mM Tris-HCl pH 8.0 at 4 °C, 150 mM NaCl and 2 mM DTT.

His-GST-Mis18BP1₂₀₋₁₃₀ was purified in the same manner as above with the following modifications: the lysis and elution buffers contained 500 mM NaCl, whilst the dialysis buffer contained 75 mM NaCl. His-MBP-Mis18BP1₂₀₋₁₃₀ was purified using the same lysis buffer containing 500 mM NaCl and purified using amylose resin (New England Biolabs). Proteins were then eluted by an elution buffer containing 10 mM maltose.

If needed, proteins were subjected to anion exchange chromatography using the HiTrap™ Q column (GE Healthcare) using the ÄKTA™ start system (GE Healthcare). Concentrated fractions were then injected onto either Superdex™ 75 increase 10/300 or Superdex™ 200 increase 10/300 columns equilibrated with 20 mM Tris-HCl pH 8.0 at 4 °C, 100–250 mM NaCl and 2 mM DTT using the ÄKTA™ Pure 25 system (GE Healthcare).

Interaction trials

Pull-down assays used to test the interaction between the C-terminus of Mis18 α and Mis1 β were performed by initially purifying the proteins through the cobalt affinity chromatography, as described for wild-type proteins, and the eluted fractions were loaded into the amylose affinity resin, pre-equilibrated with a binding buffer consisting of 30 mM Tris-HCl pH 7.5, 500 mM NaCl and 5 mM β -mercaptoethanol. Amylose resins were washed with the binding buffer, and the proteins were eluted with a binding buffer containing 20 mM maltose. The fractions were subjected to SDS-PAGE analysis.

Pull-down assay using the amylose resin to test interactions between Mis18 α / β and Mis18BP1₂₀₋₁₃₀ were done as described previously (Pan et al, 2017). Briefly, purified proteins were diluted to 10 μ M in 40 μ l binding buffer, 50 mM HEPES pH 7.5, 1 M NaCl, 1 mM TCEP, 0.01% Tween® 20. One-third of the mixture was taken as input, and the remaining fraction was incubated with 40 μ l amylose resin for 1 h at 4 °C. The bound protein was separated by washing with binding buffer three times, and the input and bound fractions were analysed by SDS-PAGE.

Crystallisation, data collection, and structure determination

Purified Mis18 α / β _{C-term} domains and Mis18 α _{Yippeee} were screened and crystallised using the hanging-drop vapour diffusion method at

room temperature with a mixture of 0.2 μ l of the protein and 0.2 μ l of crystallisation screening solutions. The crystals of Mis18 α / β _{C-term} domains were grown within a week with a solution containing 0.2 M magnesium acetate and 20% (w/v) PEG 3350. SeMet-substituted Mis18 α / β _{C-term} domains crystals were grown by the micro-seeding method with a solution containing 0.025 M magnesium acetate and 14% (w/v) PEG 3350. The crystals of SeMet-substituted Mis18 α / β _{C-term} domains were further optimised by mixing 1 μ l of the protein and 1 μ l of the optimised crystallisation solution containing 0.15 M magnesium acetate and 20% (w/v) PEG 3350. The crystals of Mis18 α _{Yippeee} were obtained in 2 M ammonium sulfate, 2% (w/v) PEG 400, and 100 mM HEPES at pH 7.5. The crystals of Mis18 α / β _{C-term} domains and Mis18 α _{Yippeee} were cryoprotected with the crystallisation solutions containing 20% and 25% glycerol, respectively. The cryoprotected crystals were flash-frozen in liquid nitrogen. Diffraction datasets were collected at the beamline LS-CAT 21 ID-G and ID-D of Advanced Photon Source (Chicago, USA). The data set were processed and scaled using the DIALS (Winter et al, 2018) via Xia2 (Winter et al, 2013). The initial model of Mis18 α / β _{C-term} domains was obtained using the SAD method with SeMet-derived data using the Autosol program (Terwilliger, 2000). The molecular replacement of the initial model as a search model against native diffraction data was performed using the Phaser programme within the PHENIX programme suite (Liebschner et al, 2019). The initial model of Mis18 α _{Yippeee} was calculated by molecular replacement method (Phaser) using yeast Mis18 Yippeee-like domain structure (PDB ID: 5HJ0) (Subramanian et al, 2016) as a search model. The final structures were manually fitted using the Coot programme (Emsley and Cowtan, 2004) and the refinement was carried out using REFMAC5 (Afonine et al, 2010). The quality of the final structures was validated with the MolProbity programme (Chen et al, 2010).

SEC-MALS

Size-exclusion chromatography (ÄKTA-Micro™, GE Healthcare) coupled to UV, static light scattering and refractive index detection (Viscotek SEC-MALS 20 and Viscotek RI Detector VE3580; Malvern Instruments) was used to determine the molecular mass of protein and protein complexes in solution. Injections of 100 μ l of 2–6 mg/ml material were used.

His-SUMO-Mis18 α ₁₈₈₋₂₃₃ ($\partial A_{280\text{nm}}/\partial c = 0.43 \text{ AU.ml.mg}^{-1}$) WT and mutants were run on a Superdex 75 increase 10/300 GL size-exclusion column pre-equilibrated in 50 mM HEPES pH 8.0, 150 mM NaCl and 1 mM TCEP at 22 °C with a flow rate of 1.0 ml/min. Light scattering, refractive index (RI) and $A_{280\text{nm}}$ were analysed by a homo-polymer model (OmniSEC software, v5.02; Malvern Instruments) using the parameters stated for the protein, $\partial n/\partial c = 0.185 \text{ ml.g}^{-1}$ and buffer RI value of 1.335. The mean standard error in the mass accuracy determined for a range of protein-protein complexes spanning the mass range of 6–600 kDa is $\pm 1.9\%$.

SAXS

SEC-SAXS experiments were performed at beamline B21 of the Diamond Light Source synchrotron facility (Oxfordshire, UK). Protein samples at concentrations >5 mg/ml were loaded onto a Superdex™ 200 Increase 10/300 GL size-exclusion chromatography column (GE

Healthcare) in 20 mM Tris pH 8.0, 150 mM KCl at 0.5 ml/min using an Agilent 1200 HPLC system. The column outlet was fed into the experimental cell, and SAXS data were recorded at 12.4 keV, detector distance 4.014 m, in 3.0 s frames. Data were subtracted, averaged and analysed for Guinier region R_g and cross-sectional R_g (R_c) using ScÅtter 3.0 (ScÅtter), and $P(r)$ distributions were fitted using PRIMUS (Konarev et al, 2003). Ab initio modelling was performed using DAMMIN (Svergun, 1999), in which 30 independent runs were performed in P1 or P2 symmetry and averaged.

Gradient fixation (GraFix)

Fractions from the gel filtration peak were concentrated to 1 mg/mL using a Vivaspin® Turbo (Sartorius) centrifugal filter, and the buffer exchanged into 20 mM HEPES pH 8.0, 150 mM NaCl, and 2 mM DTT for GraFix (Kastner et al, 2008; Stark, 2010). A gradient was formed with buffers A, 20 mM HEPES pH 8.0, 150 mM NaCl, 2 mM DTT, and 5% sucrose and B, 20 mM HEPES pH 8.0, 150 mM NaCl, 2 mM DTT, 25% sucrose, and 0.1% glutaraldehyde using the Gradient Master (BioComp Instruments). In total, 500 μ l of the sample was applied on top of the gradient, and the tubes were centrifuged at 40,000 rpm at 4 °C using a Beckman SW40 rotor for 16 h. The gradient was fractionated in 500- μ l fractions from top to bottom, and the fractions were analysed by SDS-PAGE with Coomassie blue staining and negative staining EM.

Negative staining sample preparation, data collection and processing

Copper grids, 300 mesh, with continuous carbon layer (TAAB) were glow-discharged using the PELCO easiGlow™ system (Ted Pella). GraFix fractions with and without dialysis were used. Dialysed fractions were diluted to 0.02 mg/ml. In all, 4 μ l of sample were adsorbed for 2 min onto the carbon side of the glow-discharged grids, then the excess was side blotted with filter paper. The grids were washed in two 15- μ l drops of buffer and one 15 μ l drop of 2% uranyl acetate, blotting the excess between each drop, and then incubated with a 15 μ l drop of 2% uranyl acetate for 2 min. The excess was blotted by capillary action using a filter paper, as previously described (Scarff et al, 2018).

The grids were loaded into a Tecnai F20 (Thermo Fisher Scientific) electron microscope, operated at 200 kV, field emission gun (FEG), with pixel size of 1.48 Å. Micrographs were recorded using an 8k \times 8k CMOS F816 camera (TVIPS) at a defocus range of -0.8 to -2 μ m. For Mis18 α / β /Mis18BP1₂₀₋₁₃₀ (Mis18_{core}), 163 micrographs were recorded and analysed using CryoSPARC 3.1.0 (Punjani et al, 2017). The contrast transfer function (CTF) was estimated using Gctf (Zhang, 2016). Approximately 750 particles were manually picked and submitted to 2D classification. The class averages served as templates for automated particle picking. Several rounds of 2D classification were employed to remove bad particles and assess the data, reducing the 14,840 particles to 5540. These were used to generate three ab initio models followed by homogeneous refinement with the respective particle sets.

CLMS

Cross-linking was performed on gel-filtered complexes dialysed into PBS. In total, 16 μ g EDC and 35.2 μ g sulfo-NHS were used to

cross-link 10 μ g of Mis18 α / β with Mis18BP1₂₀₋₁₃₀ (Mis18_{core}) for 1.5 h at RT. The reactions were quenched with final concentration 100 mM Tris-HCl before separation on Bolt™ 4–12% Bis-Tris Plus gels (Invitrogen). Sulfo-SDA (sulfosuccinimidyl 4,4'-azipentanoate) (Thermo Scientific Pierce) cross-linking reaction was a two-step process. First, sulfo-SDA mixed with Mis18 α / β (0.39 μ g/ μ l) at different ratio (w/w) of 1:0.07, 1:0.13, 1:0.19, 1:0.38, 1:0.5, 1:0.75, 1:1 and 1:1.4 (Mis18 α / β :Sulfo-SDA) was allowed to incubate 30 min at room temperature to initiate incomplete lysine reaction with the sulfo-NHS ester component of the cross-linker. The diazirine group was then photoactivated for 20 min using UV irradiation from a UVP CL-1000 UV Cross-linker (UVP Inc.) at 365 nm (40 W). The reactions were quenched with 2 μ l of 2.7 M ammonium bicarbonate before loading on Bolt™ 4–12% Bis-Tris Plus gels (Invitrogen) for separation. Following previously established protocol (Maiolica et al, 2007), either the whole sample or specific bands were excised, and proteins were digested with 13 ng/ μ l trypsin (Pierce) overnight at 37 °C after being reduced and alkylated. The digested peptides were loaded onto C18-Stage-tips (Rappsilber et al, 2007) for LC-MS/MS analysis.

LC-MS/MS analysis was performed using an Orbitrap Fusion Lumos (Thermo Fisher Scientific) coupled online with an Ultimate 3000 RSLCnano system (Thermo Fisher Scientific) with a “high/high” acquisition strategy. The peptide separation was carried out on a 50-cm EASY-Spray column (Thermo Fisher Scientific). Mobile phase A consisted of water and 0.1% v/v formic acid. Mobile phase B consisted of 80% v/v acetonitrile and 0.1% v/v formic acid. Peptides were loaded at a flow rate of 0.3 μ l/min and eluted at 0.2 μ l/min or 0.25 μ l/min using a linear gradient going from 2% mobile phase B to 40% mobile phase B over 109 or 79 min, followed by a linear increase from 40% to 95% mobile phase B in 11 min. The eluted peptides were directly introduced into the mass spectrometer. MS data were acquired in the data-dependent mode with a 3 s acquisition cycle. Precursor spectra were recorded in the Orbitrap with a resolution of 120,000. The ions with a precursor charge state between 3+ and 8+ were isolated with a window size of 1.6 m/z and fragmented using high-energy collision dissociation (HCD) with a collision energy of 30. The fragmentation spectra were recorded in the Orbitrap with a resolution of 15,000. Dynamic exclusion was enabled with a single repeat count and 60-s exclusion duration. The mass spectrometric raw files were processed into peak lists using ProteoWizard (version 3.0.20388) (Kessner et al, 2008), and cross-linked peptides were matched to spectra using Xi software (version 1.7.6.3) (Mendes et al, 2019) (<https://github.com/Rappsilber-Laboratory/XiSearch>) with in-search assignment of monoisotopic peaks (Lenz et al, 2018). Search parameters were MS accuracy, 3 ppm; MS/MS accuracy, 10 ppm; enzyme, trypsin; cross-linker, EDC; max missed cleavages, 4; missing monoisotopic peaks, 2. For EDC search cross-linker, EDC; fixed modification, carbamidomethylation on cysteine; variable modifications, oxidation on methionine. For sulfo-SDA search: fixed modifications, none; variable modifications, carbamidomethylation on cysteine, oxidation on methionine, SDA-loop SDA cross-link within a peptide that is also cross-linked to a separate peptide. Fragments b and y type ions (HCD) or b, c, y, and z type ions (EThcD) with loss of H₂O, NH₃ and CH₃SOH. 5% on link level False discovery rate (FDR) was estimated based on the number of decoy identification using XiFDR (Fischer and Rappsilber, 2017).

Integrative structure modelling

Input subunits

Using the Mis18 α _{Yippee} as a template, we generated high-confidence structural models for the Mis18 α and Mis18 β Yippee domains (using the homology modelling server Phyre2, www.sbg.bio.ic.ac.uk/phyre2/ (Kelley et al, 2015)). These models were almost identical with those obtained using Raptorx (<http://raptorx6.uchicago.edu/>) and AlphaFold2 (Jumper et al, 2021); structure prediction programmes that employ deep learning approach independent of co-evolution information (Källberg et al, 2012) (Fig. 1E).

Scoring function for CLMS

A cross-link was considered satisfied if the Calpha-Calpha distance was less than 22 Å. The final score was the fraction of satisfied cross-links.

Sampling

To determine the structure of the Mis18 complex, we used XlinkAssembler, an algorithm for multi-subunit assembly based on combinatorial docking approach (Inbar et al, 2005; Schneidman-Duhovny and Wolfson, 2020). The input to XlinkAssembler is N subunit structures and a list of cross-links. First, all subunit pairs are docked using cross-links as distance restraints (Schneidman-Duhovny et al, 2005). Pairwise docking generates multiple docked configurations for each pair of subunits that satisfy a large fraction of cross-links (>70%). Second, the combinatorial assembler hierarchically enumerates pairwise docking configurations to generate larger assemblies that are consistent with the CLMS data.

XlinkAssembler was used with 11 subunits to generate a model for Mis18 α/β : initial hexamer structure based on AlphaFold (Jumper et al, 2021), two Mis18 α _{Yippee} domains as well as four copies of the two helices in the Mis18 α N-terminal helical region (residues 37–55 and 60–76). For docking Mis18BP1 helices, XlinkAssembler was used with 4 subunits: the Mis18 α/β _{Yippee} domains heterodimer and the three Mis18BP1 helices predicted by AlphaFold (residues 21–33, 42–50 and 90–111).

Cell culture and transfection

The cell line HeLa Kyoto, HeLa 3–8 (having an alphoid^{tetO} array integrated into one of its chromosome arms) (Ohzeki et al, 2012), as well as HeLa CENP-A-SNAP, GFP-Mis18BP1 inducible CENP-A-SNAP, Mis18 β -GFP CENP-A-SNAP and mCherry-Mis18 α CENP-A-SNAP (kind gift from Iain Cheeseman (McKinley and Cheeseman, 2014)) were maintained in DMEM (Gibco) containing 10% FBS (Biowest) and 1 \times Penicillin/Streptomycin antibiotic mixture (Gibco). The cells were incubated at 37°C in a CO₂ incubator in humid condition containing 5% CO₂. GFP-Mis18BP1 was induced with 10 μ g/ml doxycycline for 18 h. siRNAs (AllStars Negative Control siRNA 1027280. Mis18 α : ID s28851, Mis18 β : ID s22367; Silencer[®] Select, Thermo Fisher Scientific) were used in the rescue assays by transfecting the cells using jetPRIME[®] (Polyplus transfection[®]) reagent according to the manufacturer's instructions. Briefly, HeLa CENP-A-SNAP, Mis18 β -GFP CENP-A-SNAP, GFP-Mis18BP1 inducible CENP-A-SNAP and mCherry-Mis18 α CENP-A-SNAP cells were seeded in 12-well plates and incubated overnight. siRNAs (50 pmol), vectors (200 ng) and the jetPRIME[®] reagent were diluted in the jetPRIME[®] buffer, vortexed and spun

down. The transfection mixture was incubated for 15 min before adding to the cells in a drop-by-drop manner. The cells were then incubated for 48 h.

The TetR-eYFP tagged proteins were transfected using the XtremeGene-9 (Roche) transfection reagent according to the manufacturer's protocol. The HeLa 3–8 cells attached onto the coverslip in a 12-well plate were transfected with the corresponding vectors (500 ng) and the transfection reagent diluted in Opti-MEM (Invitrogen) followed by incubation for 36–48 h.

Generation of monoclonal antibodies against Mis18 α /Mis18 β

Lou/c rats and C57BL/6J mice were immunised with 60 μ g purified recombinant human Mis18 α/β protein complex, 5 nmol CpG (TIB MOLBIOL, Berlin, Germany), and an equal volume of Incomplete Freund's adjuvant (IFA; Sigma, St. Louis, USA). A boost injection without IFA was given 6 weeks later and 3 days before fusion of immune spleen cells with P3X63Ag8.653 myeloma cells using standard procedures. Hybridoma supernatants were screened for specific binding to Mis18 α/β protein complex and also for binding to purified GST-Mis18 β protein in ELISA assays. Positive supernatants were further validated by western blot analyses on purified recombinant human Mis18 α/β complex, on cell lysates from *Drosophila* S2 cells overexpressing human Mis18 α and on HEK293 cell lysates. Hybridoma cells from selected supernatants were subcloned at least twice by limiting dilution to obtain stable monoclonal cell lines. Experiments in this work were performed with hybridoma supernatants mouse anti-Mis18 α (clone 25G8, mouse IgG2b/k) and rat anti-Mis18 β (clone 24C8; rat IgG2a/k).

Western blot

To study the efficiency of DNA and siRNA transfected, HeLa cells were transfected as stated above. Protein was extracted with RIPA buffer and analysed by SDS-PAGE followed by wet transfer using a Mini Trans-Blot[®] Cell (BioRad). Antibodies used for Western blots were: mouse Mis18 α (25G8), rat Mis18 β (24C8) (1:100, Helmholtz Zentrum München), Mis18BP1 (1:500, PA5-46777, Thermo Fisher Scientific or 1 μ g/ml, ab89265, Abcam), GFP (1:5000, ab290, Abcam), mCherry (1:1000, ab167453, Abcam) and tubulin (1:2000, T5168, Sigma). Secondary antibodies used were ECL Rabbit IgG, ECL Mouse IgG and ECL Rat IgG (1:5000, NA934, NA931, NA935, GE Healthcare), and immunoblots were imaged using NuGlow ECL (Alpha Diagnostics). For imaging with the Odyssey[®] CLx system, goat anti-mouse 680 and donkey anti-rabbit 800 secondary antibodies were used (1:5000, 926-68070, 926-32213, LI-COR).

Co-immunoprecipitation

HeLa Kyoto cells were seeded in 100-mm dishes. The cells were depleted of the endogenous Mis18 α or Mis18 β by siRNA transfection with jetPRIME[®] (Polyplus transfection[®]) and simultaneously rescued with siRNA-resistant versions of WT or mutant Mis18 α -mCherry and Mis18 β -GFP. The cells were harvested after 48 h and lysed by resuspending in immunoprecipitation buffer, 75 mM HEPES pH 7.5, 1.5 mM EGTA, 1.5 mM MgCl₂, 150 mM NaCl, 10% glycerol, 0.1% NP40, 1 mM PMSF, 10 mM NaF, 0.3 mM Na-vanadate and cOmplete™ Mini Protease Inhibitor; adapted from (Pan et al, 2017). Cells were

incubated with mixing for 30 min at 4 °C before sonicating with a Bioruptor® Pico (Diagenode). Lysates were then spun for 10 min at 15,000 × g. The protein concentrations were determined and adjusted to the same concentration. Protein was taken for inputs, and the rest was incubated with Protein G Mag Sepharose® (GE Healthcare), previously coupled to Mis18α antibody, for 1 h at 4 °C. Next, the bound fraction was separated from unbound by bind beads to the magnet and washing three times with the IP buffer with either 150 mM or 300 mM NaCl. The protein was extracted from the beads by boiling with SDS-PAGE loading dye for 5 min and were analysed by SDS-PAGE followed by western blotting with anti-mCherry, GFP and tubulin antibodies.

Immunofluorescence and quantification

The transfected cells were washed with PBS and fixed in 4% paraformaldehyde for 10 min, followed by permeabilisation in PBS with 0.5% Triton™ X-100 (Sigma) for 5 min. The cells were then blocked in 3% BSA containing 0.1% Triton™ X-100 for 1 h at 37 °C. The blocked cells were subsequently stained with the indicated primary antibodies for 1 h at 37 °C followed by secondary antibody staining under similar conditions. The following primary antibodies were used for immunofluorescence: anti-ACA (1:300; 15-235; Antibodies Inc.) and anti-CENP-A (1:100, MA 1-20832, Thermo Fisher Scientific). The secondary antibodies used were Alexa Fluor® 488 AffiniPure donkey anti-human IgG, Cy5-conjugated AffiniPure donkey anti-human, and TRITC-conjugated AffiniPure donkey anti-mouse (1:300; 709-546-149, 709-175-149, 715-025-150, Jackson ImmunoResearch). Vector shield with DAPI (Vector Laboratories) was used for DNA staining.

Micrographs were acquired at the Centre Optical Instrumentation Laboratory on a DeltaVision Elite™ system (Applied Precision) or Nikon Ti2 inverted microscope. Z stacks were obtained at a distance of 0.2 μm and were deconvolved using SoftWoRx, or AutoQuant software, respectively, followed by analysis using Image J software. The intensity at the tethering site was obtained using a custom-made plugin. Briefly, the Mis18BP1₂₀₋₁₃₀-mCherry signal at the tethering site (eYFP) was found for every z-section within a 7-square pixel box. The mean signal intensity obtained was subtracted from the minimum intensities within the section. The values were obtained from a minimum of three biological repeats. Prism 9.1.2 was used to establish if data was normally distributed, before the statistical significance of the difference between normalised intensities at the centromere and tethering region was established by a Mann–Whitney *U* two-tailed test.

SNAP-CENP-A assay and quantification

SNAP-CENP-A quench pulse labelling was done as described previously (Jansen et al, 2007). Briefly, the existing CENP-A was quenched by 10 μM SNAP-Cell® Block BTP (S9106S, New England Biolabs). The cells were treated with 1 μM STLC for 15 h for enriching the mitotic cell population, and the newly formed CENP-A was pulse labelled with 3 μM SNAP-Cell® 647-SiR (S90102S, New England Biolabs), 2 h after release from the STLC block (early G1). After pulse labelling, the cells were washed, fixed and processed for immunofluorescence. Images were obtained using DeltaVision Elite™ system (Applied Precision), deconvolved by SoftwoRx and processed by Image J. The average centromere intensities were obtained using a previously described macro CraQ (Bodor et al,

2012). Briefly, the centromeres were defined by a 7 × 7 pixel box using a reference channel, and the corresponding mean signalling intensity at the data channel was obtained by subtracting the minimum intensities within the selection. The values plotted were obtained from a minimum of three independent experiments. After testing is the data was normally distributed, the statistical significance of the difference between normalised intensities at the centromere region was established by a Mann–Whitney *U* test using Prism 9.1.2.

Data availability

PDB ID: 7SFY for Mis18α/β_{C-term}: <https://www.rcsb.org/structure/7SFY>. PDB ID: 7SFZ for Mis18α_{Y100E}: <https://www.rcsb.org/structure/7SFZ>. The mass spectrometry proteomics data have been deposited in the ProteomeXchange Consortium via the PRIDE (Perez-Riverol et al, 2019) partner repository: identifier PXD047345. Access codes for the EM density maps deposited in EMDB: EMD-50218, EMD-50219, EMD-50220. Access code for the integrative structure model deposited in PDB-Dev: PDBDEV_00000380. Access code for the original microscopy images deposited in BioImage Archive: S-BIAD1181. Plugin for analysing intensities at tethering site deposited in Zenodo: <https://doi.org/10.5281/zenodo.5708337> (<https://zenodo.org/records/5708337>).

The source data of this paper are collected in the following database record: [biostudies:S-SCDT-10_1038-S44319-024-00183-w](https://www.ebi.ac.uk/biostudies/studies/S-SCDT-10_1038-S44319-024-00183-w).

Expanded view data, supplementary information, appendices are available for this paper at <https://doi.org/10.1038/s44319-024-00183-w>.

Peer review information

A peer review file is available at <https://doi.org/10.1038/s44319-024-00183-w>

References

- Aasland R, Stewart AF, Gibson T (1996) The SANT domain: a putative DNA-binding domain in the SWI-SNF and ADA complexes, the transcriptional co-repressor N-CoR and TFIIB. *Trends Biochem Sci* 21:87–88
- Afonine PV, Mustyakimov M, Grosse-Kunstleve RW, Moriarty NW, Langan P, Adams PD (2010) Joint X-ray and neutron refinement with phenix.refine. *Acta Crystallogr D Biol Crystallogr* 66:1153–1163
- Baker NA, Sept D, Joseph S, Holst MJ, McCammon JA (2001) Electrostatics of nanosystems: application to microtubules and the ribosome. *Proc Natl Acad Sci USA* 98:10037–10041
- Barnhart MC, Kuich PH, Stellfox ME, Ward JA, Bassett EA, Black BE, Foltz DR (2011) HJURP is a CENP-A chromatin assembly factor sufficient to form a functional de novo kinetochore. *J Cell Biol* 194:229–243
- Black BE, Jansen LE, Foltz DR, Cleveland DW (2010) Centromere identity, function, and epigenetic propagation across cell divisions. *Cold Spring Harb Symp Quant Biol* 75:403–418
- Bodor DL, Rodriguez MG, Moreno N, Jansen LE (2012) Analysis of protein turnover by quantitative SNAP-based pulse-chase imaging. *Curr Protoc Cell Biol* Chapter 8:Unit8 8
- Bond CS, Schüttelkopf AW (2009) ALINE: a WYSIWYG protein-sequence alignment editor for publication-quality alignments. *Acta Crystallogr D Biol Crystallogr* 65:510–512

- Carroll CW, Milks KJ, Straight AF (2010) Dual recognition of CENP-A nucleosomes is required for centromere assembly. *J Cell Biol* 189:1143–1155
- Catania S, Allshire RC (2014) Anarchic centromeres: deciphering order from apparent chaos. *Curr Opin Cell Biol* 26:41–50
- Cheeseman IM (2014) The kinetochore. *Cold Spring Harb Perspect Biol* 6:a015826
- Chen VB, Arendall 3rd WB, Headd JJ, Keedy DA, Immormino RM, Kapral GJ, Murray LW, Richardson JS, Richardson DC (2010) MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr* 66:12–21
- Conti D, Verza AE, Pesenti ME, Cmentowski V, Vetter IR, Pan D, Musacchio A (2024) Role of PLK1 in the epigenetic maintenance of centromeres. Preprint at <https://www.biorxiv.org/content/10.1101/2024.02.23.581696v1>
- Dambacher S, Deng W, Hahn M, Sadic D, Frohlich J, Nuber A, Hoischen C, Diekmann S, Leonhardt H, Schotta G (2012) CENP-C facilitates the recruitment of M18BP1 to centromeric chromatin. *Nucleus* 3:101–110
- Dunleavy EM, Almouzni G, Karpen GH (2011) H3.3 is deposited at centromeres in S phase as a placeholder for newly assembled CENP-A in G(1) phase. *Nucleus* 2:146–157
- Dunleavy EM, Roche D, Tagami H, Lacoste N, Ray-Gallet D, Nakamura Y, Daigo Y, Nakatani Y, Almouzni-Pettinotti G (2009) HJURP is a cell-cycle-dependent maintenance and deposition factor of CENP-A at centromeres. *Cell* 137:485–497
- Emsley P, Cowtan K (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* 60:2126–2132
- Evans R, O'Neill M, Pritzel A, Antropova N, Senior A, Green T, Židek A, Bates R, Blackwell S, Yim J et al (2022) Protein complex prediction with AlphaFold-Multimer. Preprint at <https://www.biorxiv.org/content/10.1101/2021.10.04.463034v1>
- Fischer L, Rappsilber J (2017) Quirks of error estimation in cross-linking/mass spectrometry. *Anal Chem* 89:3829–3833
- Foltz DR, Jansen LE, Bailey AO, Yates 3rd JR, Bassett EA, Wood S, Black BE, Cleveland DW (2009) Centromere-specific assembly of CENP-a nucleosomes is mediated by HJURP. *Cell* 137:472–484
- Fujita Y, Hayashi T, Kiyomitsu T, Toyoda Y, Kokubu A, Obuse C, Yanagida M (2007) Priming of centromere for CENP-A recruitment by human hMis18alpha, hMis18beta, and M18BP1. *Dev Cell* 12:17–30
- Fukagawa T, Earnshaw WC (2014) The centromere: chromatin foundation for the kinetochore machinery. *Dev Cell* 30:496–508
- Hayashi T, Fujita Y, Iwasaki O, Adachi Y, Takahashi K, Yanagida M (2004) Mis16 and Mis18 are required for CENP-A loading and histone deacetylation at centromeres. *Cell* 118:715–729
- Hu H, Liu Y, Wang M, Fang J, Huang H, Yang N, Li Y, Wang J, Yao X, Shi Y et al (2011) Structure of a CENP-A-histone H4 heterodimer in complex with chaperone HJURP. *Genes Dev* 25:901–906
- Inbar Y, Benyamini H, Nussinov R, Wolfson HJ (2005) Prediction of multimolecular assemblies by multiple docking. *J Mol Biol* 349:435–447
- Jansen LE, Black BE, Foltz DR, Cleveland DW (2007) Propagation of centromeric chromatin requires exit from mitosis. *J Cell Biol* 176:795–805
- Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, Tunyasuvunakool K, Bates R, Zidek A, Potapenko A et al (2021) Highly accurate protein structure prediction with AlphaFold. *Nature* 596:583–589
- Källberg M, Wang H, Wang S, Peng J, Wang Z, Lu H, Xu J (2012) Template-based protein structure modeling using the RaptorX web server. *Nat Protoc* 7:1511–1522
- Kastner B, Fischer N, Golas MM, Sander B, Dube P, Boehringer D, Hartmuth K, Deckert J, Hauer F, Wolf E et al (2008) GraFix: sample preparation for single-particle electron cryomicroscopy. *Nat Methods* 5:53–55
- Kato H, Jiang J, Zhou BR, Rozendaal M, Feng H, Ghirlando R, Xiao TS, Straight AF, Bai Y (2013) A conserved mechanism for centromeric nucleosome recognition by centromere protein CENP-C. *Science* 340:1110–1113
- Kelley LA, Mezulis S, Yates CM, Wass MN, Sternberg MJ (2015) The Phyre2 web portal for protein modeling, prediction and analysis. *Nat Protoc* 10:845–858
- Kessner D, Chambers M, Burke R, Agus D, Mallick P (2008) ProteoWizard: open source software for rapid proteomics tools development. *Bioinformatics* 24:2534–2536
- Kim IS, Lee M, Park KC, Jeon Y, Park JH, Hwang EJ, Jeon TI, Ko S, Lee H, Baek SH et al (2012) Roles of Mis18α in epigenetic regulation of centromeric chromatin and CENP-A loading. *Mol Cell* 46:260–273
- Konarev PV, Volkov VV, Sokolova AV, Koch MHJ, Svergun DI (2003) PRIMUS: a Windows PC-based system for small-angle scattering data analysis. *J Appl Crystallogr* 36:1277–1282
- Lagana A, Dorn JF, De Rop V, Ladouceur AM, Maddox AS, Maddox PS (2010) A small GTPase molecular switch regulates epigenetic centromere maintenance by stabilizing newly incorporated CENP-A. *Nat Cell Biol* 12:1186–1193
- Lenz S, Giese SH, Fischer L, Rappsilber J (2018) In-search assignment of monoisotopic peaks improves the identification of cross-linked peptides. *J Proteome Res* 17:3923–3931
- Liebschner D, Afonine PV, Baker ML, Bunkóczi G, Chen VB, Croll TI, Hintze B, Hung LW, Jain S, McCoy AJ et al (2019) Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. *Acta Crystallogr D Struct Biol* 75:861–877
- Maddox PS, Hyndman F, Monen J, Oegema K, Desai A (2007) Functional genomics identifies a Myb domain-containing protein family required for assembly of CENP-A chromatin. *J Cell Biol* 176:757–763
- Madeira F, Park YM, Lee J, Buso N, Gur T, Madhusoodanan N, Basutkar P, Tivey ARN, Potter SC, Finn RD et al (2019) The EMBL-EBI search and sequence analysis tools APIs in 2019. *Nucleic Acids Res* 47:W636–W641
- Maiolica A, Cittaro D, Borsotti D, Sennels L, Ciferri C, Tarricone C, Musacchio A, Rappsilber J (2007) Structural analysis of multiprotein complexes by cross-linking, mass spectrometry, and database searching. *Mol Cell Proteomics* 6:2200–2211
- McKinley KL, Cheeseman IM (2014) Polo-like kinase 1 licenses CENP-A deposition at centromeres. *Cell* 158:397–411
- McKinley KL, Cheeseman IM (2016) The molecular basis for centromere identity and function. *Nat Rev Mol Cell Biol* 17:16–29
- Mendes ML, Fischer L, Chen ZA, Barbon M, O'Reilly FJ, Giese SH, Bohlke-Schneider M, Belsom A, Dau T, Combe CW et al (2019) An integrated workflow for crosslinking mass spectrometry. *Mol Syst Biol* 15:e8994
- Moree B, Meyer CB, Fuller CJ, Straight AF (2011) CENP-C recruits M18BP1 to centromeres to promote CENP-A chromatin assembly. *J Cell Biol* 194:855–871
- Muller S, Montes de Oca R, Lacoste N, Dingli F, Loew D, Almouzni G (2014) Phosphorylation and DNA binding of HJURP determine its centromeric recruitment and function in CenH3(CENP-A) loading. *Cell Rep* 8:190–203
- Musacchio A, Desai A (2017) A molecular view of kinetochore assembly and function. *Biology* 6:5
- Nardi IK, Zasadzinska E, Stellfox ME, Knippler CM, Foltz DR (2016) Licensing of centromeric chromatin assembly through the Mis18alpha-Mis18beta heterotetramer. *Mol Cell* 61:774–787
- Ohzeki J, Bergmann JH, Kouprina N, Noskov VN, Nakano M, Kimura H, Earnshaw WC, Larionov V, Masumoto H (2012) Breaking the HAC Barrier: histone H3K9 acetyl/methyl balance regulates CENP-A assembly. *EMBO J* 31:2391–2402
- Pan D, Klare K, Petrovic A, Take A, Walstein K, Singh P, Rondelet A, Bird AW, Musacchio A (2017) CDK-regulated dimerization of M18BP1 on a Mis18 hexamer is necessary for CENP-A loading. *eLife* 6:e23352
- Pan D, Walstein K, Take A, Bier D, Kaiser N, Musacchio A (2019) Mechanism of centromere recruitment of the CENP-A chaperone HJURP and its implications for centromere licensing. *Nat Commun* 10:4046

- Parashara P, Medina-Pritchard B, Abad MA, Sotelo-Parrilla P, Thamkachy R, Grundei D, Zou J, Das V, Yan Z, Kelly DA et al (2024) PLK1-mediated phosphorylation cascade activates the Mis18 complex to ensure centromere inheritance. Preprint at <https://www.biorxiv.org/content/10.1101/2024.02.23.581399v1.full.pdf>
- Perez-Riverol Y, Csordas A, Bai J, Bernal-Llinares M, Hewapathirana S, Kundu DJ, Inuganti A, Griss J, Mayer G, Eisenacher M et al (2019) The PRIDE database and related tools and resources in 2019: improving support for quantification data. *Nucleic Acids Res* 47:D442–d450
- Perpelescu M, Nozaki N, Obuse C, Yang H, Yoda K (2009) Active establishment of centromeric CENP-A chromatin by RSF complex. *J Cell Biol* 185:397–407
- Punjani A, Rubinstein JL, Fleet DJ, Brubaker MA (2017) cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nat Methods* 14:290–296
- Rappsilber J, Mann M, Ishihama Y (2007) Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. *Nat Protoc* 2:1896–1906
- Sali A, Blundell TL (1993) Comparative protein modelling by satisfaction of spatial restraints. *J Mol Biol* 234:779–815
- Scarff CA, Fuller MJG, Thompson RF, Iadanza MG (2018) Variations on negative stain electron microscopy methods: tools for tackling challenging systems. *J Vis Exp* 132:e57199
- Schneidman-Duhovny D, Inbar Y, Nussinov R, Wolfson HJ (2005) PatchDock and SymmDock: servers for rigid and symmetric docking. *Nucleic Acids Res* 33:W363–367
- Schneidman-Duhovny D, Wolfson HJ (2020) Modeling of multimolecular complexes. *Methods Mol Biol* 2112:163–174
- Shono N, Ohzeki J, Otake K, Martins NM, Nagase T, Kimura H, Larionov V, Earnshaw WC, Masumoto H (2015) CENP-C and CENP-I are key connecting factors for kinetochore and CENP-A assembly. *J Cell Sci* 128:4572–4587
- Silva MC, Bodor DL, Stellfox ME, Martins NM, Hochegger H, Foltz DR, Jansen LE (2012) Cdk activity couples epigenetic centromere inheritance to cell cycle progression. *Dev Cell* 22:52–63
- Spiller F, Medina-Pritchard B, Abad MA, Wear MA, Molina O, Earnshaw WC, Jeyapakash AA (2017) Molecular basis for Cdk1-regulated timing of Mis18 complex assembly and CENP-A deposition. *EMBO Rep* 18:894–905
- Stankovic A, Guo LY, Mata JF, Bodor DL, Cao XJ, Bailey AO, Shabanowitz J, Hunt DF, Garcia BA, Black BE et al (2017) A dual inhibitory mechanism sufficient to maintain cell-cycle-restricted CENP-A assembly. *Mol Cell* 65:231–246
- Stark H (2010) GraFix: stabilization of fragile macromolecular complexes for single particle cryo-EM. *Methods Enzymol* 481:109–126
- Stellfox ME, Bailey AO, Foltz DR (2013) Putting CENP-A in its place. *Cell Mol Life Sci* 70:387–406
- Stellfox ME, Nardi IK, Knippler CM, Foltz DR (2016) Differential binding partners of the Mis18alpha/beta YIPPEE domains regulate Mis18 complex recruitment to centromeres. *Cell Rep* 15:2127–2135
- Studier FW (2005) Protein production by auto-induction in high density shaking cultures. *Protein Expr Purif* 41:207–234
- Subramanian L, Medina-Pritchard B, Barton R, Spiller F, Kulasegaran-Shylini R, Radaviciute G, Allshire RC, Arockia Jeyapakash A (2016) Centromere localization and function of Mis18 requires Yippee-like domain-mediated oligomerization. *EMBO Rep* 17:496–507
- Svergun DI (1999) Restoring low resolution structure of biological macromolecules from solution scattering using simulated annealing. *Biophys J* 76:2879–2886
- Terwilliger TC (2000) Maximum-likelihood density modification. *Acta Crystallogr D Biol Crystallogr* 56:965–972
- Wang J, Liu X, Dou Z, Chen L, Jiang H, Fu C, Fu G, Liu D, Zhang J, Zhu T et al (2014) Mitotic regulator Mis18beta interacts with and specifies the centromeric assembly of molecular chaperone holliday junction recognition protein (HJURP). *J Biol Chem* 289:8326–8336
- Weir JR, Faesen AC, Klare K, Petrovic A, Basilico F, Fischbock J, Pentakota S, Keller J, Pesenti ME, Pan D et al (2016) Insights from biochemical reconstitution into the architecture of human kinetochores. *Nature* 537:249–253
- Winter G, Lobley CM, Prince SM (2013) Decision making in xia2. *Acta Crystallogr D Biol Crystallogr* 69:1260–1273
- Winter G, Waterman DG, Parkhurst JM, Brewster AS, Gildea RJ, Gerstel M, Fuentes-Montero L, Vollmar M, Michels-Clark T, Young ID et al (2018) DIALS: implementation and evaluation of a new integration package. *Acta Crystallogr D Struct Biol* 74:85–97
- Zhang D, Martyniuk CJ, Trudeau VL (2006) SANTA domain: a novel conserved protein module in Eukaryota with potential involvement in chromatin regulation. *Bioinformatics* 22:2459–2462
- Zhang K (2016) Gctf: real-time CTF determination and correction. *J Struct Biol* 193:1–12

Acknowledgements

The authors would like to thank David Kelly from the Centre Optical Instrumentation Laboratory, Martin Wear from the Edinburgh Protein Production Facility as well as Marcus Wilson and Maarten Tuijtel of the Cryo-Electron Microscopy Facility for their help. The authors also thank Diamond Light Source and the staff of beamline B21 (proposal sm23510), as well as Advanced Photon Source and the staff at beamlines LS-CAT 21 ID-G and ID-D. Thanks also to Iain Cheeseman for the kind gift of cell lines. The Wellcome Trust generously supported this work through Senior Research Fellowships to AA Jeyapakash (202811), J Rappsilber (084229), O Davies (219413/Z/19/Z) and P Heun (103897/Z/14/Z), a Centre Core Grant (092076 and 203149) and an instrument grant (108504) to the Wellcome Trust Centre for Cell Biology. AA Jeyapakash and his team are also funded by the European Union (ERC Advanced Grant, CHROMSEG, 101054950) and the Medical Research Council (MRC, United Kingdom; MR/X0012451/1). Views and opinions expressed are, however, those of the authors only and do not necessarily reflect those of the European Union or the European Research Council. Neither the European Union nor the granting authority can be held responsible for them. The NIH supported the work of U Cho and SH Park (NIH/NIDDK; R01 DK111465). The work of D Schneidman-Duhovny is supported by ISF 1466/18 and the Israeli Ministry of Science and Technology.

Author contributions

Reshma Thamkachy: Formal analysis; Investigation; Writing—review and editing. **Bethan Medina-Pritchard:** Data curation; Formal analysis; Investigation; Writing—original draft; Writing—review and editing. **Sang Ho Park:** Investigation. **Carla G Chiodi:** Formal analysis; Investigation. **Juan Zou:** Formal analysis; Investigation. **Maria de la Torre-Barranco:** Formal analysis; Investigation. **Kazuma Shimanaka:** Investigation. **Maria Alba Abad:** Investigation. **Cristina Gallego Páramo:** Formal analysis. **Regina Feederle:** Resources. **Emilija Ruksenaite:** Investigation. **Patrick Heun:** Resources; Supervision. **Owen R Davies:** Formal analysis. **Juri Rappsilber:** Resources. **Dina Schneidman-Duhovny:** Data curation; Formal analysis; Writing—original draft. **Uhn-Soo Cho:** Conceptualisation; Resources; Supervision. **A Arockia Jeyapakash:** Conceptualisation; Resources; Supervision; Writing—original draft; Writing—review and editing.

Source data underlying figure panels in this paper may have individual authorship assigned. Where available, figure panel/source data authorship is listed in the following database record: [biostudies:S-SCDT-10_1038-544319-024-00183-w](https://www.ebi.ac.uk/biostudies/studies/S-SCDT-10_1038-544319-024-00183-w).

Disclosure and competing interests statement

The authors declare no competing interests.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. Creative Commons Public Domain Dedication waiver <http://creativecommons.org/public-domain/zero/1.0/> applies to the data associated with this article, unless otherwise stated in a credit line to the data, but does not extend to the graphical or creative elements of illustrations, charts, or figures. This waiver removes legal barriers to the re-use and mining of research data. According to standard scholarly practice, it is recommended to provide appropriate citation and attribution whenever technically possible.

© The Author(s) 2024

Expanded View Figures

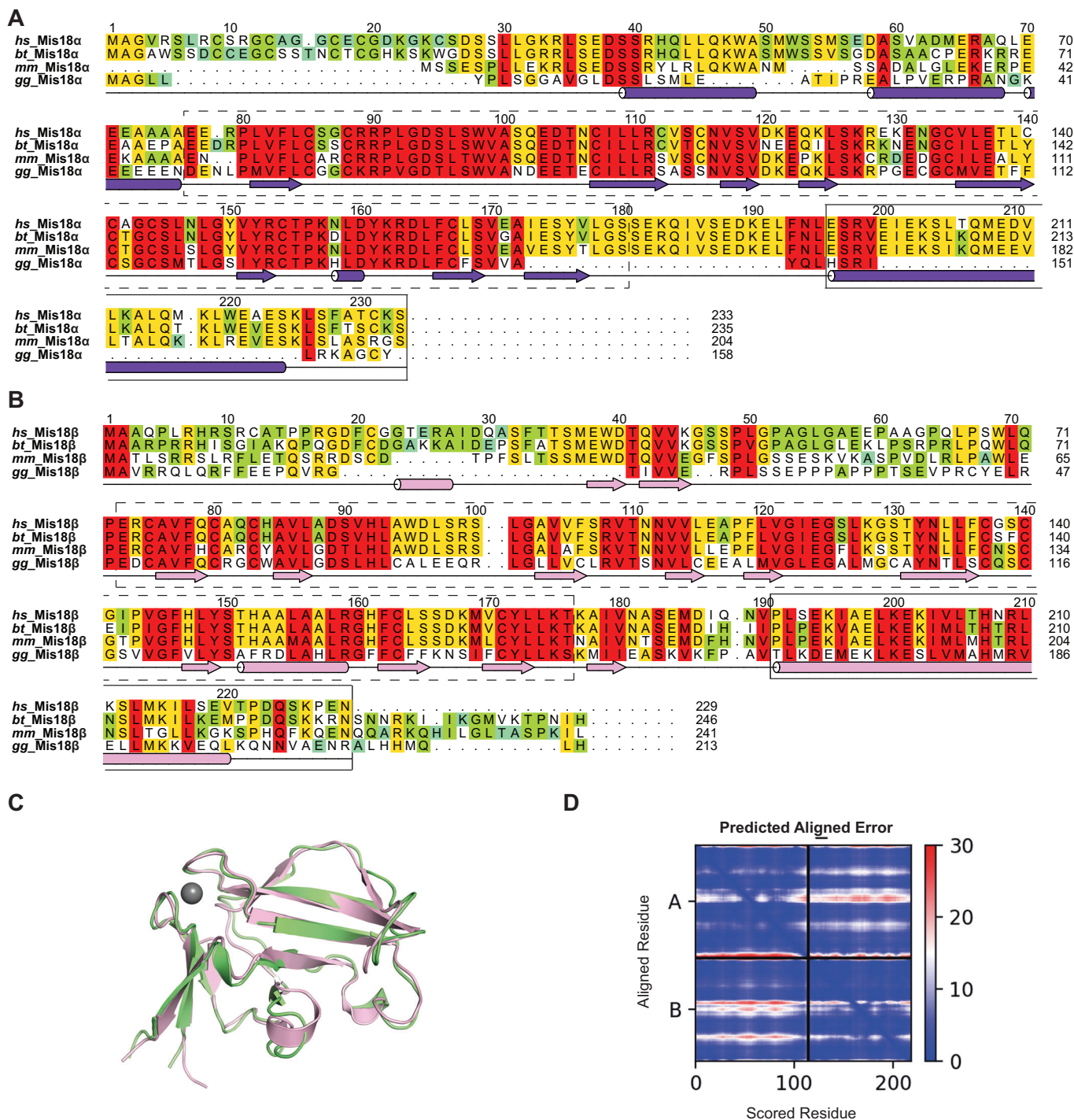


Figure EV1. Mis18 α and Mis18 β contain two domains capable of oligomerising.

(A, B) Domain architecture and amino acid conservation of (A) Mis18 α and (B) Mis18 β . Alignments include *Homo sapiens* (hs), *Bos taurus* (bt), *Mus musculus* (mm) and *Gallus gallus* (gg). The conservation score is mapped from red to cyan, where red corresponds to highly conserved and cyan to poorly conserved. Secondary structures are annotated/predicted by Conserved Domain Database [CDD] and PsiPred, <http://bioinf.cs.ucl.ac.uk/psipred>. Multiple sequence alignments were performed with MUSCLE (Madeira et al, 2019) and edited with Aline (Bond and Schüttelkopf, 2009). Dashed boxes highlight Yippee domains whilst solid boxes highlight C-terminus α -helices. (C) Superposition of Mis18 β _{Yippee} structures predicted by AlphaFold (light pink) and RaptorX (green). RaptorX generated five models and the model with the lowest estimated error (1.9 Å) is shown here. The AlphaFold and RaptorX models superpose well with an RMSD of 0.95 Å. (D) The PAE plot corresponding to the Mis18 α / β _{Yippee} AlphaFold model shown in Fig. 1D.

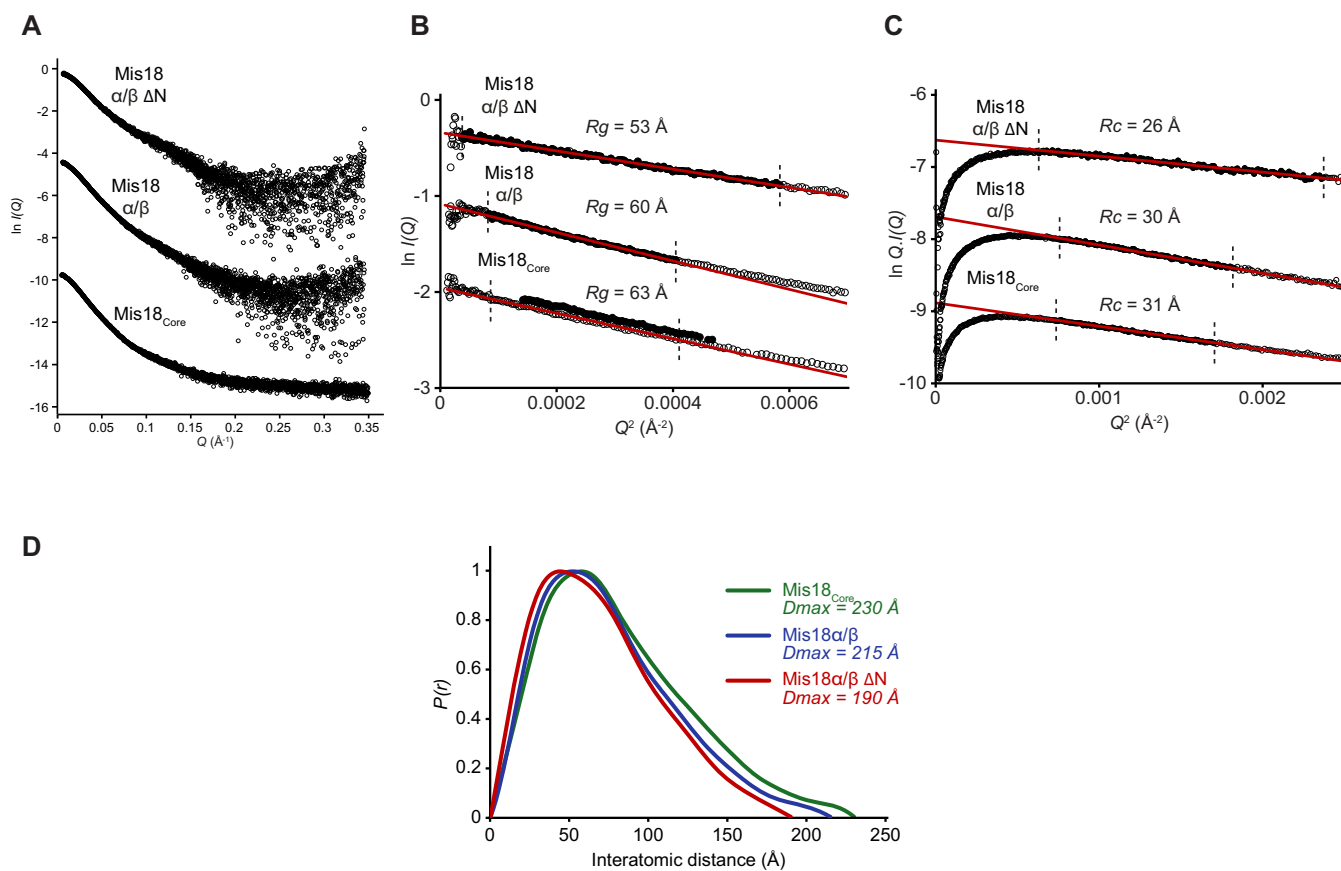
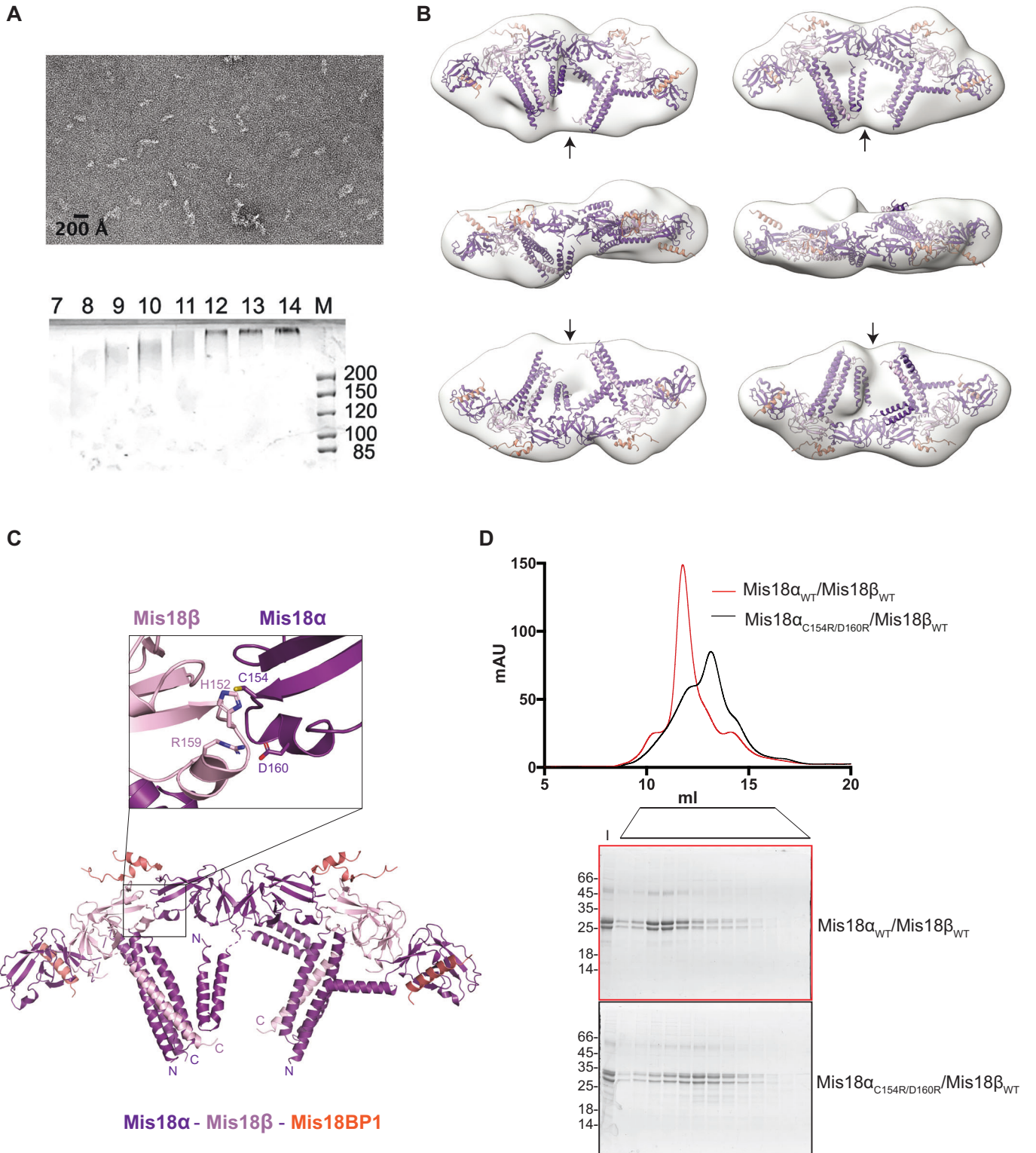


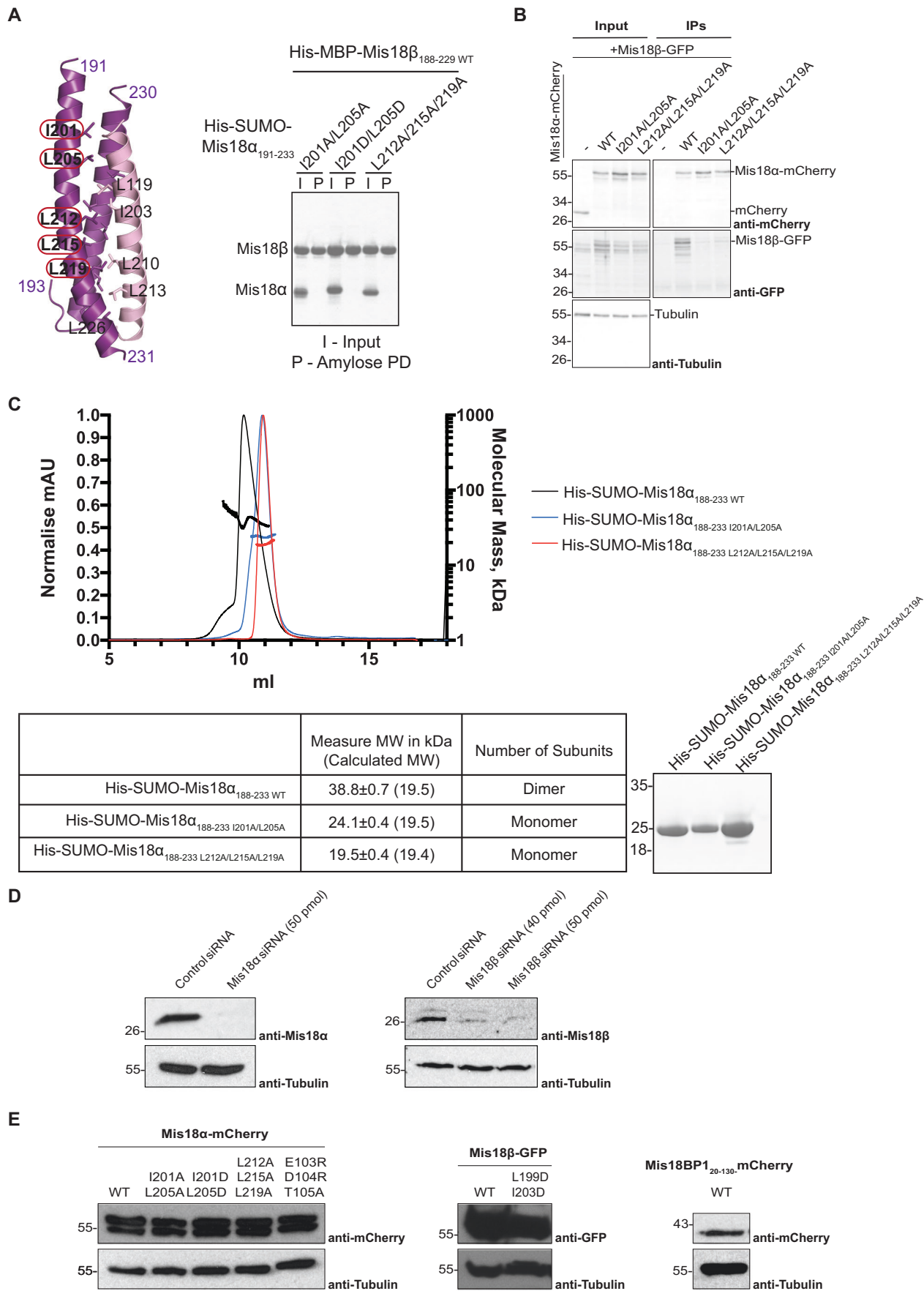
Figure EV2. SAXS analysis of Mis18 α/β Δ N, Mis18 α/β and Mis18_{core}.

(A) SAXS scattering curves of Mis18 α/β Δ N, Mis18 α/β and Mis18_{core}. (B) Guinier Plot showing R_g of 53 \AA , 60 \AA , and 63 \AA for Mis18 α/β Δ N, Mis18 α/β and Mis18_{core} respectively. (C) Modified Guinier Plot showing R_c of 26 \AA , 30 \AA , and 31 \AA for Mis18 α/β Δ N, Mis18 α/β and Mis18_{core} respectively. (D) SAXS $P(r)$ distributions showing maximum dimensions of 190 \AA , 215 \AA , and 230 \AA for Mis18 α/β Δ N, Mis18 α/β and Mis18_{core} respectively.



◀ Figure EV3. Structural characterisation of the Mis18_{core} complex.

(A) Representative micrograph of negative staining EM of the Mis18 α /Mis18 β /Mis18BP1₂₀₋₁₃₀ (Mis18_{core}) complex cross-linked using GraFix (Kastner et al, 2008; Stark, 2010). Beneath is the corresponding SDS-PAGE analysis of fractions from GraFix, fractions 8 and 9 were used to make grids. (B) Two models (Class II-III) generated for Mis18_{core} from negative staining EM analysis. All show that the overall shapes of the Mis18_{core} resemble a telephone handset with 'ear' and 'mouth' pieces assuming different relative orientations. (C) Cartoon representation of the model of Mis18_{core} complex generated in Fig. 2B. Zoomed in panel shows interaction between Mis18 α and Mis18 β Yippee domains using the second interface. Important residues for this interaction highlighted in pink and purple. (D) SEC profile of Mis18 α _{WT}/Mis18 β _{WT} (red) and Mis18 α _{C154R/D160R}/Mis18 β _{WT} (black) and corresponding SDS-PAGE analysis of the fractions. Samples were analysed using Superdex 200 increase 10/300 in 20 mM Tris-HCl pH 8.0, 250 mM NaCl and 2 mM DTT.



◀ **Figure EV4. Structural and biochemical characterisation of Mis18 α C-terminal helix.**

(A) Cartoon representation of the crystal structure of Mis18 α _{C-term}/Mis18 β _{C-term} (PDB ID: 7SFY). Mis18 α is shown in purple and Mis18 β in light pink. Potential residues involved in the interaction are highlighted. Mis18 α (purple) and Mis18 β (light pink). Right panel shows SDS-PAGE analysis of cobalt and amylose pull-down of His-MBP-Mis18 β ₁₈₈₋₂₂₉ WT with His-SUMO-Mis18 α ₁₉₁₋₂₃₃ mutants. SDS-PAGE shows protein bound to nickel resin as input (I) and protein-bound to amylose resin to assess interaction (P). Control with WT proteins shown in Fig. 4A. (B) Western blot analysis of co-immunoprecipitation (Co-IP) experiments using Mis18 α antibody to test interaction of mCherry as a control, Mis18 α -mCherry with and without mutations in the C-terminal α -helices and Mis18 β -GFP. Top panel shows blot against mCherry, middle panel shows blot against GFP, and bottom panel shows blot against tubulin as loading control. (C) SEC-MALS of His-SUMO-Mis18 α ₁₈₈₋₂₃₃ WT, His-SUMO-Mis18 α ₁₈₈₋₂₃₃ I201A/L205A and His-SUMO-Mis18 α ₁₈₈₋₂₃₃ L212A/L215A/L219A. Normalised absorption at 280 nm (mAU, left y-axis) and molecular mass (kDa, right y-axis) are plotted against elution volume (ml, x-axis). Measured molecular weight (MW) and the calculated subunit stoichiometry based on the predicted MW. Samples were analysed using a Superdex 75 increase in 50 mM HEPES pH 8.0, 150 mM NaCl and 1 mM TCEP. (D) Representative immunoblots showing expression levels of endogenous proteins after treatment with siRNA. (E) Representative immunoblots showing expression levels of transiently expressed tagged proteins after transfection.