bioRxiv preprint doi: https://doi.org/10.1101/2021.11.08.466737; this version posted November 8, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

# Structural Basis for Mis18 Complex Assembly: Implications for Centromere Maintenance

Reshma Thamkachy[1,7], Bethan Medina-Pritchard[1,7], Sang Ho Park[2,7], Carla G. Chiodi[1], Juan Zou[1], Maria de la Torre-Barranco[1], Kazuma Shimanaka[2], Maria Alba Abad[1], Cristina Gallego Páramo[1], Regina Feederle[3], Emilija Ruksenaite[4], Patrick Heun[1], Owen R. Davies[1], Juri Rappsilber[1,5], Dina Schneidman-Duhovny[6], Uhn-Soo Cho[2], A. Arockia Jeyaprakash[1*]

[1]Wellcome Centre for Cell Biology, University of Edinburgh, Edinburgh, EH9 3BF, UK
[2]Department of Biological Chemistry, University of Michigan, Michigan, 48109, USA.
[3]Monoclonal Antibody Core Facility, Helmholtz Zentrum München, German Research Center for Environmental Health (GmbH), 85764 Neuherberg, Germany
[4]Institute Novo Nordisk Foundation Centre for Protein Research, Copenhagen, Denmark
[5]Institute of Biotechnology, Technische Universität Berlin, 13355 Berlin, Germany
[6]School of Computer Science and Engineering, The Hebrew University of Jerusalem, Jerusalem, Israel.

*Correspondance: jeyaprakash.arulanandam@ed.ac.uk

[7]Equal contribution

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.08.466737; this version posted November 8, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

## 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. During each round of the cell cycle, CENP-A levels undergo DNA replication-mediated dilution. To maintain centromere identity, CENP-A levels must be restored. A central player mediating this process is the Mis18 complex (composed of Mis18α, Mis18β and Mis18BP1), which recruits the CENP-A specific chaperone HJURP to centromeres for CENP-A deposition. Here, using a multipronged approach we provide the structural basis for the assembly of the Mis18 complex. We show that the Mis18α/β hetero-trimer (2 Mis18α:1 Mis18β) is assembled by the formation of a triple helical bundle with a Mis18α/$\beta_{Yippee}$ hetero-dimer and Mis18$\alpha_{Yippee}$ monomer on opposite ends. Two such Mis18α/β hetero-trimers, each bound to a Mis18BP1, assemble as a hetero-octamer via Mis18$\alpha_{Yippee}$ homo-dimerisation. Evaluation of structure-guided separation of function mutants in cells reveal structural determinants essential for Mis18 complex assembly and function.

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.08.466737; this version posted November 8, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

## Introduction

Faithful chromosome segregation during cell division requires 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 [1-4]. 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 [4-7]. CENP-A containing nucleosomes recruit CENP-C and CENP-N, two proteins that are part of the constitutive centromere-associated network (CCAN) and that recruits the rest of the kinetochore components at the centromeric region of the chromosome [8-10].

Whilst canonical histone loading is coupled with DNA replication, CENP-A loading is not [11]. 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 [12, 13]. 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 Mis18α, Mis18β and Mis18BP1) and the CENP-A chaperone, HJURP [12-16]. 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 [17-19]. Once at the centromere, the Mis18 complex has been implicated in facilitating the deposition of CENP-A in several ways. There is evidence that it affects DNA methylation and histone acetylation, which may facilitate CENP-A loading. But one of its most important and well-established roles is the recruitment of HJURP, which binds a single CENP-A/H4 dimer and brings it to the centromere [13, 16, 20]. 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

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.08.466737; this version posted November 8, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

genome, which requires several remodelled factors such as MgcRacGAP, RSF, Ect2, and Cdc42 [21, 22].

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 [23-28]. 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) [24, 25]. Cdk1 also phosphorylates HJURP, which negatively regulates its binding to the Mis18 complex at the centromere [27-29]. 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 Mis18α/β and Mis18BP1, but it has also been proposed to interact with phosphorylated Mis18 complex through its polo-box domain (PBD) [26].

As outlined above, a central event in the process of CENP-A deposition at centromeres is the Mis18 complex assembly. The Mis18 proteins, Mis18α and Mis18β, possess a well-conserved globular domain called the Yippee domain (also known as the MeDiY domain; spanning residues 77-180 in Mis18α and 73-176 in Mis18β) and C-terminal α-helices (residues 196-233 in Mis18α and 191-229 in Mis18β). We and others previously showed that the Yippee domains of Mis18 proteins can form a hetero-dimer, while the C-terminal helices form a hetero-trimer with two Mis18α and one Mis18β. However, the full-length proteins form a hetero-hexameric assembly with 4 Mis18α and 2 Mis18β. This led to a proposed model, where the Mis18α and Mis18β mainly interact via the C-terminal helices to form a hetero-trimer, and two such hetero-trimers interact via the Yippee hetero-dimerisation (Mis18α/Mis18β) or/and homo-dimerisation (Mis18α/Mis18α) to form a hetero-hexameric assembly [24, 25, 30, 31].

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

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.08.466737; this version posted November 8, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

are known to have roles in regulating chromatin remodelling [32-34]. In-between these two domains resides the CENP-C binding domain (CBD) [17, 18]. *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 [18]. We and others have previously shown that the N-terminal 130 amino acids of Mis18BP1 are sufficient for interaction with Mis18α/β through their Yippee domains, and Cdk1 phosphorylation of Mis18BP1 at residues T40 and S110 inhibits its interaction with Mis18α/β [24, 25].

Although the importance of the Mis18 complex assembly and function is well appreciated, our understanding of the contribution of different domains of the Mis18 complex subunits for Mis18 complex assembly is limited. Particularly, the structural basis for the Mis18 complex assembly and function is yet to be identified. Here, we have determined the structural architecture of the Mis18 complex using an integrative structural modelling approach that combines X-ray crystallography, Electron Microscopy (EM), Small Angle X-ray Scattering (SAXS), Cross-Linking Mass Spectrometry (CLMS) 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

### Structure determination of Mis18α/β core assembly

Both Mis18α and Mis18β possess two distinct but conserved structural entities, the Yippee domain and a C-terminal α-helix (**Fig. 1a-c**). Mis18α possesses an additional α-helical domain upstream of the Yippee domain (residues 39-76).

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 homo-dimer [35], and used this to model human Yippee domains. To determine the actual structure of human Mis18 Yippee

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.08.466737; this version posted November 8, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

domains, we purified and crystalised Mis18$\alpha_{Yippee}$ (residues 77-190). The crystals diffracted X-rays to about 3 Å resolution, and the structure was determined using molecular replacement method. The final model was refined to R and $R_{free}$ factors of 20.26% and 25.00%, respectively (**Table 1** and **Fig. 1d**, PDB ID: 7SFZ). The overall fold of the Mis18$\alpha_{Yippee}$ looks remarkably like the previously solved *S. pombe* Mis18$_{Yippee}$ homo-dimer structure [35]. One striking feature is that the crystal packing interactions of Mis18$\alpha_{Yippee}$ are similar to that of pombe Mis18$_{Yippee}$ and both proteins show two unique dimerisation interfaces (Interface I and Interface II) (**Fig 3a**). We had previously shown that mutations in Interface I disrupts homo/hetero-dimerisation of Mis18$\alpha_{Yippee}$ in solution, highlighting the major contribution of Interface I for dimerisation [35]. However, considering the observation that interactions involving Interface II are preserved both in *pombe* Mis18$_{Yippee}$ and human Mis18$\alpha_{Yippee}$, we speculate this interface is also physiologically relevant.

Using the Mis18$\alpha_{Yippee}$ as a template we generated high-confidence structural models for the Mis18$\alpha$ and Mis18$\beta$ Yippee domains (using the homology modelling server Phyre2, www.sbg.bio.ic.ac.uk/phyre2/ [36]). These models were almost identical with those obtained using Raptorx (http://raptorx.uchicago.edu/) and AlphaFold2 [37]; structure prediction programs that employ deep learning approach independent of co-evolution information [38] (**Fig. 1e**).

Previous studies have shown that recombinantly purified C-terminal α-helices of Mis18$\alpha$ and Mis18$\beta$ form a hetero-trimer with 2 Mis18$\alpha$ and 1 Mis18$\beta$ [24, 25]. However, in the absence of high-resolution structural information, how Mis18 $\alpha$-helices interact to form a hetero-trimer and how the structural arrangements of $\alpha$-helices influence the relative orientations of the Yippee domains, and hence the overall architecture of the Mis18$\alpha$/$\beta$ hexamer assembly, remained unclear. We purified recombinantly expressed Mis18$\alpha$ spanning aa residues 191 to 233 and Mis18$\beta$ spanning aa residues 188 and 229 and crystallised the reconstituted complex. The

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.08.466737; this version posted November 8, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

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 1,** PDB ID: 7SFY). The asymmetric unit contained two copies of Mis18α/β hetero trimer. The final model included Mis18α residues 191 to 231 in one copy, Mis18α residues 193 to 230 in the second copy, and Mis18β residues 190 to 223 (**Fig. 1f**). The two Mis18α 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 Mis18α helices (**Fig. 4a & b**). This observation is consistent with the theoretically calculated pI of the Mis18α helix (pI=4.9). In contrast, the pI of the Mis18β helix is 8.32. This charge complementarity appears to facilitate the interaction with Mis18α, as a positively charged surface of the Mis18β helix snug fits in the negatively charged groove of the Mis18α/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 Mis18β helix (**Fig 4b**). The binding free energy calculated based on the buried accessible surface area suggests a nanomolar affinity interaction between the helices of Mis18α and Mis18β.

**Overall architecture of the Mis18 complex**

We and others have previously shown that the N-terminal 130 aa of Mis18BP1 are sufficient to interact with the Mis18α/β complex and that this binding is mediated by the Mis18α/$β_{Yippee}$ hetero-dimers [18, 24, 25]. Recombinantly purified full-length Mis18α/β complex or the Mis18 complex containing the minimal fragment of Mis18BP1 (Mis18α/β/Mis18BP1$_{20-130}$, now on referred as the Mis18$_{core}$ complex) were not amenable for structural characterisation; hence we took an integrative structural approach.

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.08.466737; this version posted November 8, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

SAXS analysis of the Mis18α/β ΔN (Mis18α residues 77-187 and Mis18β residues 56-183), Mis18α/β and Mis18$_{core}$ complexes suggest that all of these complexes possess an elongated shape with flexible features (**Fig. S1, Table S1**). The measured radius of gyration (*Rg*) and cross-sectional radius of gyration (*Rc*) for the Mis18α/β ΔN and Mis18α/β complexes are *Rg* – 53 Å/60 Å and *Rc* – 26 Å/30 Å (**Fig. S1b & c**). The corresponding values for the Mis18$_{core}$ complex show an incremental increase, *Rg* – 63 Å and *Rc* – 31 Å. A similar trend is observed in the calculated maximum interatomic distance values (*Dmax*): 190 Å, 215 Å and 230 Å for the Mis18α/β ΔN, Mis18α/β and Mis18$_{core}$ complexes, respectively (**Fig. S1d**). The positively skewed peaks observed for all complexes suggest their elongated shape. Together, these analyses suggest that Mis18BP1 binding results in a slight increase in the overall shape and bulkiness of the complex.

To gain further insights into the structure of the Mis18$_{core}$ complex, we analysed this sample using negative stain Electron Microscopy (EM). GraFix was used to cross-link the sample [39]. The micrographs revealed a good distribution of particles and were processed using CryoSPARC [40] (**Fig. 2a**). Particle picking, followed by a few rounds of 2D classifications using CryoSPARC, revealed classes with defined structural features (**Fig. 2b**). 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. Hence, we performed *ab-initio* reconstructions with more than one model and refined each model against its respective particle sets with C1 symmetry. This resulted in three models with distinguishable conformational variability (**Fig. 2c**). Overall the dimensions of these models were similar (approximately 220 x 105 x 80 Å) and in agreement with the $D_{max}$ calculated from SAXS analysis. However, the bulkier features that resemble 'ear' and 'mouth' pieces show different relative orientations with respect to 'handle' of the handset. The resolution is approximately 18 Å based on the Fourier shell correlation (FSC) = 0.143 and 21 Å based on FSC = 0.5. At this resolution, the Yippee dimers and the triple helical bundle

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.08.466737; this version posted November 8, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

have similar sizes and shapes, hindering their placement into the models. To gain further insights about the region and the structure, CLMS and computational modelling were performed.

We performed EDC chemical cross-linking. EDC is a zero-length cross-linker that covalently links Asp or Glu residues with Lys, and to a lesser extent Ser, Thr and Tyr. Purified untagged $Mis18_{core}$ complex was dialysed into PBS, and a cross-linking titration series was performed with EDC (10 µg - 16 µg) and the corresponding amount of Sulfo-NHS (22 µg - 35.2 µg). Analysis of the cross-linking reactions on an SDS-PAGE identified a cross-linked species that migrated as expected for an intact $Mis18_{core}$ complex (4 Mis18α, 2 Mis18β and 2 $Mis18BP1_{20-130}$ (178.2 kDa)) (**Fig. S2a**). This condition was subsequently replicated and analysed by MS.

CLMS analysis revealed several cross-links between and within the subunits of the Mis18 complex (**Fig. 2d**). Particularly, we made four key observations: (1) Consistent with the crystal structure of the $Mis18\alpha/\beta_{C\text{-}term}$ helical assembly (**Fig. 1f**), several residues of Mis18α and Mis18β spanning the C-terminal helices are involved in cross-linking; (2) cross-links were observed between residues of the Yippee domains and C-terminal helices of Mis18α/β; (3) N-terminal helical region of Mis18α makes several cross-links with C-terminal helices of Mis18α and Mis18β (highlighted in black); and (4) in agreement with previous studies [24, 25], Mis18BP1 residues cross-link with $Mis18\alpha/\beta_{Yippee}$ domains.

To understand the architecture of the $Mis18_{core}$ complex, we first assembled $Mis18\alpha/\beta_{Yippee}$ hetero-dimer with the triple helical bundle using the CLMS data as restraints for docking. The Yippee hetero-dimer connects to the two parallel C-terminal helices, while the second anti-parallel Mis18α helix can freely connect to the Yippee domain on the other side of the bundle (**Fig. 3b**). We have also attempted to fold the structure using AlphaFold2 [37] with two Mis18α and one Mis18β sequences as an input. While most of the runs produced parallel triple helix

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.08.466737; this version posted November 8, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

bundles, a few reproduced the anti-parallel triple helix and had consistent Mis18$\alpha$/$\beta_{Yippee}$ hetero-dimer orientation. The linkers between the triple helix bundle and the Yippee dimer enable some flexibility in the orientation of the domains, consistent with EM data (**Fig. 2c**). The initial hexamer was obtained using the Mis18$\alpha_{Yippee}$ homo-dimer formed via Interface II (Dimer II) (**Fig. 3**). Then the two remaining Mis18$\alpha_{Yippee}$ domains as well as the N-terminal helical region (residues 37-55 and 60-76) of Mis18$\alpha$ were docked to this initial hexamer using the CLMS data as restraints. We explored three different options for the two Mis18$\alpha_{Yippee}$ domains (i) homo-dimer with the interface different from hetero-dimer; (ii) homo-dimer with the interface identical to hetero-dimer; and (iii) no dimerisation (**Fig. 3b**). The first two options were consistent with class I and II EM density maps (**Fig. 2c**), while the third one fitted best into the class III density map (**Fig. 2c**).

Finally, the Mis18BP1 was docked to the hexamer. The overall structure is disordered; however, a helical structure was predicted for residues 21-33, 42-50, and 90-111 by AlphaFold (**Fig. 3b**). These helices were docked to the Mis18$\alpha$/$\beta_{Yippee}$ hetero-dimer using the CLMS data as restraints. A highly similar model was obtained using AlphaFold2 with the Mis18$\alpha$/$\beta_{Yippee}$ hetero-dimer and Mis18BP1 sequences. The remaining disordered regions were added using MODELLER [41] to enable calculation of cross-link satisfaction. Overall, the best models satisfied ~60% of the 584 EDC cross links (**Fig. S3**). Most of the violations have distances under 40 Å and corresponded to cross-links from disordered regions that can move significantly. In the structured regions there were 75% satisfied cross-links (64 violated cross-links out of 250). Overall, there were only 15 cross links with distances larger than 40 Å that were not satisfied by the 3 models simultaneously (**Fig. S3**). The models were also validated using the sulfoSDA cross-links that were not used in modelling. Similarly, ~80% of the cross-links were satisfied by the structured regions and ~50% when disordered regions were added.

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.08.466737; this version posted November 8, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Overall, our integrative modelling shows that multiple hetero-oligomeric interfaces stabilise the hetero-octameric Mis18 complex assembly. Particularly, the Mis18α/β hetero-trimers, each bound to a single copy of Mis18BP1, dimerise mainly via the homo-dimerisation of Mis18α/β hetero-dimers and the Yippee domains of anti-parallel Mis18α from each are flexibly connected to the helical bundles and can assume different conformations (**Fig. 3b**).

**Mis18α mutants disrupting the C-terminal helical bundle assembly fail to localise to the centromere and abolish new CENP-A loading at centromeres.**

To evaluate the contribution of the triple helical bundle, formed by the C-terminal helices of Mis18α and Mis18β, for Mis18 complex assembly and function, we designed several mutants based on the crystal structure of the Mis18α/$\beta_{C\text{-term}}$ helical assembly (**Fig. 1f** and **4b**). We first tested these mutants using *in vitro* pull-down assays by mixing recombinantly purified WT and mutant His-MBP-Mis18$\beta_{188\text{-}229}$ and His-SUMO-Mis18$\alpha_{191\text{-}233}$ proteins. Pull-downs using cobalt resin represent inputs, while amylose pull-downs assess intermolecular interactions (**Fig. 4c & d**). We identified two hydrophobic clusters (I201/L205 and L212/L215/L219) in Mis18α that form the 'spine' of the hydrophobic core running along the triple helical bundle. Mutating these residues to Ala (Mis18$\alpha_{I201A/L205A}$ and Mis18$\alpha_{L212A/L215A/L219A}$) or Asp (Mis18$\alpha_{I201D/L205D}$) abolished its ability to interact with Mis18β (**Fig. 4c**). Co-immunoprecipitation (Co-IP) assays using an anti-Mis18α antibody were performed on cells where endogenous Mis18α was depleted, and Mis18α mCherry was co-expressed with Mis18β GFP to check for complex formation. In line with our *in vitro* pull-downs (**Fig 4c**), Co-IPs revealed that Mis18$\alpha_{WT}$ mCherry interacted with Mis18β GFP while Mis18$\alpha_{I201A/L205A}$ and Mis18$\alpha_{L212A/L215A/L219A}$ mutants did not (**Fig. S4a**, left panel). These mutants were further tested *in vivo* to evaluate their effect on centromere localisation of Mis18α and Mis18β and CENP-A deposition.

HeLa CENP-A-SNAP cells [26] were depleted of the endogenous Mis18α by siRNA (**Fig. S4b**) and simultaneously rescued with either WT or mutant Mis18α mCherry, then visualised by

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.08.466737; this version posted November 8, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

immunofluorescence along with ACA. Mis18β GFP was also co-expressed with the Mis18α mCherry (**Fig. S4c**). Unlike Mis18$\alpha_{WT}$, the Mis18α mutants (Mis18$\alpha_{I201A/L205A}$, Mis18$\alpha_{I201D/L205D}$ and Mis18$\alpha_{L212A/L215A/L219A}$) all failed to localise to centromeres (**Fig. 5a**). Analysis of the GFP signal revealed co-localisation of Mis18$\beta_{WT}$ with Mis18$\alpha_{WT}$ as expected. However, in cells expressing Mis18$\alpha_{I20A1/L205A}$, Mis18$\alpha_{I201D/L205D}$ and Mis18$\alpha_{L212A/L215A/L219A}$, Mis18β could no longer co-localise with Mis18α at the centromere. Together, this confirms that Mis18β depends on Mis18α to localise at centromeres.

We then evaluated the impact of Mis18α 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.* [12] (**Fig. 5b**). HeLa CENP-A-SNAP cells were depleted of the endogenous Mis18α and rescued with either Mis18$\alpha_{WT}$ or Mis18α mutants (Mis18$\alpha_{I20A1/L205A}$, Mis18$\alpha_{I201D/L205D}$ and Mis18$\alpha_{L212A/L215A/L219A}$). 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$\alpha_{WT}$ rescued new CENP-A deposition to levels compared to that of control siRNA (**Fig. 5c**). However, Mis18$\alpha_{I20A1/L205A}$, Mis18$\alpha_{I201D/L205D}$ and Mis18$\alpha_{L212A/L215A/L219A}$ abolished new CENP-A loading almost completely, indicating that the formation of the Mis18 triple helical bundle is essential for CENP-A deposition.

**Mis18α can associate with the centromere and facilitate CENP-A deposition, independently of Mis18β**

We again perform cobalt resin and amylose *in vitro* pull-down assays, using His-SUMO-Mis18$\alpha_{191\text{-}233\ WT}$ and mutant His-MBP-Mis18$\beta_{188\text{-}229}$ proteins, to assess the ability of a structure guided Mis18β mutant to form the triple-helical bundle with Mis18α. We identified one cluster (L199/I203) in Mis18β and observed that mutating these residues to either Ala

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.08.466737; this version posted November 8, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

(Mis18$\beta_{L199A/I203A}$) or Asp (Mis18$\beta_{L199D/I203D}$) either reduced or abolished its ability to interact with Mis18$\alpha_{191-233}$ (**Fig. 4d**). Co-IP analysis using an anti-Mis18α antibody was performed on cells where endogenous Mis18β was depleted, and Mis18β GFP was expressed along Mis18α mCherry to check for complex formation. Western blot analysis showed that Mis18$\beta_{WT}$ could interact with Mis18α mCherry and that the ability of Mis18$\beta_{L199D/I203D}$ to interact with Mis18α was reduced (**Fig. S4a,** right panel).

To assess the contribution of Mis18β for the centromere association and function of Mis18α, we evaluated the Mis18β mutant (Mis18$\beta_{L199D/I203D}$), that cannot form the triple helical assembly with Mis18α, in siRNA rescue assays by expressing Mis18β GFP tagged proteins in a mCherry Mis18α cell line [26]. Depletion of endogenous Mis18β and simultaneous transient expression of Mis18$\beta_{WT}$ GFP led to co-localisation of Mis18β with Mis18α at centromeres (**Fig. S4b, S4c & 6a**). Under these conditions, Mis18$\beta_{WT}$ GFP levels at centromeres were comparable to that of the control siRNA. Whereas Mis18$\beta_{L199D/I203D}$ failed to localise at the centromeres. Strikingly, Mis18$\beta_{L199D/I203D}$ perturbed centromere association of Mis18α only moderately (**Fig 6a**, right panel). This suggests that Mis18α can associate with centromere in a Mis18β independent manner.

Next, we assessed the contribution of Mis18β for CENP-A deposition in the Quench-Chase-Pulse CENP-A-SNAP assay described above. Endogenous Mis18β was depleted using siRNA, and Mis18$\beta_{WT}$ and Mis18$\beta_{L199D/I203D}$ were transiently expressed as GFP-tagged proteins in HeLa cells expressing CENP-A-SNAP. Mis18$\beta_{WT}$ rescued new CENP-A deposition to comparable levels observed in the control experiment (**Fig 6b**). Interestingly, unlike the Mis18α mutants (Mis18$\alpha_{I20A1/L205A}$, Mis18$\alpha_{I201D/L205D}$ and Mis18$\alpha_{L212A/L215A/L219A}$), Mis18$\beta_{L199D/I203D}$ did not abolish new CENP-A loading but reduced the levels only moderately.

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.08.466737; this version posted November 8, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Together, these analyses demonstrate that Mis18α can associate with centromeres and deposit new CENP-A independently of Mis18β. However, efficient CENP-A loading requires Mis18β.

**Identification and characterisation of Mis18α residues critical for Mis18BP1 binding and subsequent centromere association and function**

CLMS analysis revealed several contacts between residues of Mis18BP1$_{20-130}$ and Yippee domains of Mis18α and Mis18β (**Fig. 2d**). We mapped these contacts on the three-dimensional model of the Mis18α/β$_{Yippee}$ hetero-dimer. This together with the analysis of the electrostatic potential on the surface of the Mis18α/β$_{Yippee}$ hetero-dimer revealed three negatively charged amino acid clusters as potential Mis18BP1 interacting residues. These are Mis18α S169, E171, Mis18α E103, D104 and T105 and Mis18β E116 and E124 (**Fig. 7a**). We assessed the contribution of these amino acid clusters for Mis18BP1 binding using recombinant proteins in pull-down assays. Recombinantly purified untagged wild type and mutant Mis18α/β complexes (Mis18α$_{E103R/D104R/T105A}$/Mis18β$_{WT}$, Mis18α$_{S169A/E171R}$/Mis18β$_{WT}$ and Mis18α$_{WT}$/Mis18β$_{E116R/E124R}$) were mixed with MBP-Mis18BP1$_{20-130}$ and interactions were tested by performing MBP pull-downs using amylose resins followed by the analysis of the pull-down eluants in Coomassie-stained SDS-PAGE (**Fig. 7b**). Assessment of Mis18α/β band intensities in the pull-downs revealed that the Mis18α$_{E103R/D104R/T105A}$/Mis18β$_{WT}$ complex bind to MBP-Mis18BP1$_{20-130}$ weakly as compared to other Mis18α/β complexes tested suggesting the direct contribution of Mis18α residues E103, D104 and T105 for Mis18BP1 binding. These residues are part of an acidic protrusion that borders two surface grooves on the Mis18α$_{Yippee}$ domain.

To strengthen the conclusions of the *in vitro* pull-down assays, we tested the Mis18α mutants using a TetO array-based tethering assay in HeLa 3-8 cells [42], where a synthetic alphoid$^{tetO}$ array was integrated in a chromosome arm. We expressed Mis18BP1$_{20-130}$ mCherry

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.08.466737; this version posted November 8, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

along with TetR-eYFP Mis18$\alpha_{WT}$, TetR-eYFP Mis18$\alpha_{S169A/E171R}$ or TetR-eYFP Mis18$\alpha_{E103R/D104R/T105A}$ and assessed their ability to recruit Mis18BP1$_{20-130}$ to the alphoid$^{tetO}$ array [24] (**Fig. S4c & S5a**). Consistent with the *in vitro* pull-down assay, Mis18$\alpha_{E103R/D104R/T105A}$ recruited significantly less Mis18BP1$_{20-130}$ to the alphoid$^{tetO}$ array as compared to Mis18$\alpha_{S160A/E171R}$ and Mis18$\alpha_{WT}$.

Furthermore, we probed the effects of these mutants on endogenous centromeres. We depleted Mis18$\alpha$ in a cell line that stably expresses CENP-A-SNAP and allows inducible expression of GFP Mis18BP1 [26]. We then assessed the ability of transfected Mis18$\alpha$ mCherry to co-localise with Mis18BP1 at centromeres. Depletion of Mis18$\alpha$ and simultaneous expression of either Mis18$\alpha_{WT}$ mCherry or Mis18$\alpha_{E103R/D104R/T105A}$ mCherry revealed that, unlike Mis18$\alpha_{WT,}$ Mis18$\alpha_{E103R/D104R/T105A}$ failed to localise at endogenous centromeres (**Fig. 7c**, middle panel, **& S4c**). 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$\alpha_{WT}$ (**Fig. 7c,** right panel). Consistent with the observation of reduced centromeric Mis18$\alpha$, 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 (**Fig. 7d**).

Overall, these observations demonstrate a direct contribution of the acidic surface protrusion of Mis18$\alpha_{Yippee}$ domain, formed by E103, D104 and T105, for its interaction with Mis18BP1 and subsequent centromere association and function of the Mis18 complex.

**N-terminal $\alpha$-helical region of Mis18$\alpha$ modulates HJURP binding by directly interacting with the Mis18$\alpha/\beta$ C-terminal $\alpha$-helical assembly**

Previous studies have established that HJURP binding of the Mis18 complex is mainly mediated by the C-terminal domains of Mis18$\alpha/\beta$ complex and removing these abolished

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.08.466737; this version posted November 8, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

HJURP interaction *in vitro* and *in vivo* [30, 31]. Interestingly, the Mis18α/β complex lacking the N-terminal α-helical region of Mis18α bound HJURP more efficiently than the full-length Mis18α/β complex [31]. Consistent with these observations, the Mis18α mutants lacking the N-terminal helical regions (TetR-eYFP-Mis18$\alpha_{54\text{-End}}$ and TetR-eYFP-Mis18$\alpha_{77\text{-End}}$), when tethered to the alphoid$^{tetO}$ array integrated in a chromosome arm, recruited more HJURP and deposited more CENP-A at the tethering site as compared to that of Mis18$\alpha_{WT}$ (**Fig. S5b & c**). However, the structural basis for these observations is not clear yet.

Our CLMS analysis reported here revealed several cross-links between the Mis18α N-terminal α-helical region (spanning aa 39-76) and C-terminal α-helical regions of Mis18α and Mis18β (**Fig. 2d**). The Yippee domains of Mis18α and Mis18β fold in a way that would orient their N- and C-terminal ends in the same direction and, as consequence, the N-terminal helical region of Mis18α will be in close proximity to the triple helical bundle formed by the C-terminal helices of Mis18α and Mis18β required for HJURP binding. This provides a structural basis for how the N-terminal helical domain of Mis18α influences HJURP binding of the Mis18 complex.

## 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 [12-14]. 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 Mis18α and Mis18β [18, 24, 25, 30]. Some of the key conclusions of these studies include: (1) Mis18α/β is a hetero-hexamer made of 4 Mis18α and 2 Mis18β; (2) The Yippee domains and C-terminal α-helices of Mis18α and Mis18β have the intrinsic ability to homo- or hetero-oligomerise, and form three distinct oligomeric modules in different copy numbers – a Mis18$\alpha_{Yippee}$ homo-dimer, two copies of Mis18α/$\beta_{Yippee}$ hetero-dimers and two hetero-trimers made of Mis18α/β C-terminal helices (2 Mis18α and 1 Mis18β);

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.08.466737; this version posted November 8, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

(3) the two copies of Mis18α/$\beta_{Yippee}$ hetero-dimers each bind one Mis18BP1$_{20-130}$ and form a hetero-octameric Mis18$_{core}$ complex (Mis18α/Mis18β/Mis18BP1$_{20-130}$: a Mis18α/β hetero-hexamer bound to 2 copies of Mis18BP1$_{20-130}$). 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 Mis18α/β and the holo-Mis18 complexes and to determine the specific interactions that are essential for Mis18 complex 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 Mis18α/β hetero-trimer is stabilised by the formation of a triple helical bundle with a Mis18α/$\beta_{Yippee}$ hetero-dimer on one end and Mis18$\alpha_{Yippee}$ monomer on the other. Two such Mis18α/β hetero-trimers assemble as hetero-hexamer via the homo-dimerisation of Mis18$\alpha_{Yippee}$ domain. The crystal structure of Mis18α/$\beta_{C\text{-}term}$ triple helical structure allowed us to design several separation of function Mis18α and Mis18β mutants. These mutations specifically perturb the ability of Mis18α or Mis18β 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) Mis18α can associate with centromeres and deposit CENP-A independently of Mis18β, and (2) depletion of Mis18β or disrupting the incorporation of Mis18β into the Mis18 complex, while does not abolish CENP-A loading, reduces the CENP-A deposition amounts, questions the consensus view that Mis18α and Mis18β always function as a single structural entity to exert their function to maintain centromere maintenance. The data presented here suggest that Mis18β mainly contributes to the quantitative control of centromere maintenance – by ensuring the right amounts of CENP-A deposition at centromeres. Future studies will focus on dissecting the mechanisms underlying the Mis18β-mediated control of CENP-A loading amounts.

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.08.466737; this version posted November 8, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Mis18α mutants that disrupt Mis18BP1 binding shows that Mis18BP1 can associate with centromeres independently of Mis18α [14, 27], but efficient centromere association requires its association with Mis18α. The separation of function Mis18α mutant characterised here shows that disrupting Mis18α-Mis18BP1 interaction completely abolishes Mis18α's ability to associate with centromeres and new CENP-A loading [14]. This highlights that Mis18BP1-mediated centromere targeting is the major centromere recruitment pathway for the Mis18α/β complex.

Previously published work identified amino acid sequence similarity between the N-terminal region of Mis18α and R1 and R2 repeats of the HJURP that mediates Mis18α/β interaction [31]. Deletion of the Mis18α N-terminal region enhanced HJURP interaction with the Mis18 complex. This led to speculation that the N-terminal region of Mis18α might directly interact with the HJURP binding site of the Mis18 complex and thereby modulating HJURP binding. Our work presented here strengthens this speculation and provides the structural justification. We show that the N-terminal helical region of Mis18α makes extensive contacts with the C-terminal helices of Mis18α and Mis18β that mediate HJURP binding and tethering Mis81α lacking the N-terminal region to an ectopic site in cells recruited more HJURP and deposited more CENP-A at the tethering site. In the future, it will be important to address how and when the interference caused by the N-terminal region of Mis18α is relieved for efficient HJURP binding by the Mis18 complex.

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.08.466737; this version posted November 8, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

## Material and Methods

### Plasmids

For crystallisation, a polycistronic expression vector for the C-terminal coiled-coil domains of Mis18α (residues 191-233, Mis18$\alpha_{C\text{-term}}$) and Mis18β (residues 188-229, Mis18$\beta_{C\text{-term}}$) were produced with the N-terminal 6His-SUMO- (His-SUMO) and 6His-MBP-tags (His-MBP), respectively. Mis18$\alpha_{Yippee}$ (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 Mis18α and Mis18β 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 Mis18α and Mis18β were 77-187 and 56-183. Mis18BP1$_{20\text{-}130}$ 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). Mis18α, Mis18β and Mis18BP1$_{20\text{-}130}$ 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).

### Expression and purification of recombinant proteins

For crystallisation, both Mis18α/$\beta_{C\text{-term}}$ domains and Mis18$\alpha_{Yippee}$ were transformed and expressed in *Escherichia coli* BL21 (DE3) using the auto-inducible expression system [43]. The cells were harvested and resuspended in the lysis buffer containing 30 mM Tris-HCl pH7.5, 500 mM NaCl, and 5 mM β-mercaptoethanol with protease inhibitor cocktails. The

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.08.466737; this version posted November 8, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

resuspended cells were lysed using the ultra-sonication method and centrifuged at 20,000 x g for 50 min at 4°C to remove the cell debris. After 0.45 μm 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 pH7.5, 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 pH7.5, 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°C. The TEV cleavage released the untagged Mis18$\alpha/\beta_{C\text{-term}}$ 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 pH7.5, 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$\alpha/\beta_{C\text{-term}}$ domains were expressed with PASM-5052 auto-inducible media [43]. The SeMet-substituted Mis18$\alpha/\beta_{C\text{-term}}$ domains were purified using the same procedure described above.

The purification of His tagged Mis18$\alpha_{Yippee}$ employed the same purification method used for Mis18$\alpha/\beta_{C\text{-term}}$ domains except for the amylose affinity chromatography step. The purified Mis18$\alpha_{Yippee}$ from the HiLoad™ 16/600 Superdex™ 200 chromatography was concentrated to 13.7 mg/ml with the buffer containing 30 mM Tris-HCl pH7.5, 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-Mis18α/His-GFP-Mis18β 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

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.08.466737; this version posted November 8, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

8.0 and 2 mM β-mercaptoethanol supplemented with 10 μg/ml DNase, 1mM PMSF and cOmplete™ 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 $MgCl_2$, 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-Mis18BP1$_{20-130}$ 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-Mis18BP1$_{20-130}$ was purified using the same lysis buffer containing 500 mM NaCl and purified using amylose resin (NEB). 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 Mis18α and Mis1β 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 pH7.5, 500

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.08.466737; this version posted November 8, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

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 Mis18α/β and Mis18BP1$_{20-130}$ were done as described previously [25]. Briefly, purified proteins were diluted to 10 μM in 40 μl 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 μl 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 Mis18α/β$_{C\text{-}term}$ domains and Mis18α$_{Yippee}$ were screened and crystallised using the hanging-drop vapour diffusion method at room temperature with a mixture of 0.2 μl of the protein and 0.2 μl of crystallisation screening solutions. The crystals of Mis18α/β$_{C\text{-}term}$ domains were grown within a week with a solution containing 0.2 M magnesium acetate and 20% (w/v) PEG 3350. SeMet-substituted Mis18α/β$_{C\text{-}term}$ 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 Mis18α/β$_{C\text{-}term}$ domains were further optimised by mixing 1 μl of the protein and 1 μl of the optimised crystallisation solution containing 0.15 M magnesium acetate and 20% (w/v) PEG 3350. The crystals of Mis18α$_{Yippee}$ were obtained in 2 M ammonium sulfate, 2% (w/v) PEG 400, and 100 mM HEPES at pH 7.5. The crystals of Mis18α/β$_{C\text{-}term}$ domains and Mis18α$_{Yippee}$ 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

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.08.466737; this version posted November 8, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

the DIALS [44] via Xia2 [45]. The initial model of Mis18$\alpha/\beta_{C\text{-term}}$ domains was obtained using the SAD method with SeMet-derived data using the Autosol program [46]. The molecular replacement of the initial model as a search model against native diffraction data was performed using the Phaser program within the PHENIX program suite [47]. The initial model of Mis18$\alpha_{Yippee}$ was calculated by molecular replacement method (Phaser) using yeast Mis18 Yippee-like domain structure (PDB ID: 5HJ0) [35] as a search model. The final structures were manually fitted using the Coot program [48] and the refinement was carried out using REFMAC5 [49]. The quality of the final structures was validated with the MolProbity program [50].

**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 *Rg* and cross-sectional *Rg* (*Rc*) using ScÅtter 3.0 (http://www.bioisis.net), and *P(r)* distributions were fitted using *PRIMUS* [51]. *Ab-initio* modelling was performed using *DAMMIN* [52], 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 pH8.0, 150 mM NaCl, and 2 mM DTT for GraFix [39, 53]. 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

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.08.466737; this version posted November 8, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

NaCl, 2 mM DTT, 25% sucrose, and 0.1% glutaraldehyde using the Gradient Master (BioComp Instruments). 500 μl of sample was applied on top of the gradient, and the tubes centrifuged at 40,000 rpm at 4ºC using a Beckman SW40 rotor for 16 h. The gradient was fractionated in 500 μl 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. 4 μl 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 μl drops of buffer and one 15 μl drop of 2% uranyl acetate, blotting the excess between each drop, and then incubated with a 15 μl drop of 2% uranyl acetate for 2 min. The excess was blotted by capillary action using a filter paper, as previously described [54].

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 x 8k CMOS F816 camera (TVIPS) at a defocus range of -0.8 to -2 μm. For Mis18α/β/Mis18BP1$_{20-130}$ (Mis18$_{core}$), 163 micrographs were recorded and analysed using CryoSPARC 3.1.0 [40]. The contrast transfer function (CTF) was estimated using Gctf [55]. 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 5,540. These were used to generate three *ab-initio* models followed by homogeneous refinement with the respective particle sets.

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.08.466737; this version posted November 8, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

**CLMS**

Cross-linking was performed on gel filtered complexes dialysed into PBS. 16 μg EDC and 35.2 μg sulpho-NHS were used to cross-link 10 μg of Mis18α/β with Mis18BP1$_{20-130}$ (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 Mis18α/β (0.39 μg/μl) 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 (Mis18α/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 mins using UV irradiation from a UVP CL-1000 UV Cross-linker (UVP Inc.) at 365 nm (40 W). The reactions were quenched with 2 μl of 2.7 M ammonium bicarbonate before loading on Bolt™ 4–12% Bis-Tris Plus gels (Invitrogen) for separation. Following previously established protocol [38], either the whole sample or specific bands were excised, and proteins were digested with 13 ng/μl trypsin (Pierce) overnight at 37°C after being reduced and alkylated. The digested peptides were loaded onto C18-Stage-tips [39] for LC-MS/MS analysis.

LC-MS/MS analysis was performed using an Orbitrap Fusion Lumos (Thermo Fisher Scientific) coupled on-line with an Ultimate 3000 RSLCnano system (Thermo Fisher Scientific) with a "high/high" acquisition strategy. The peptide separation was carried out on a 50cm 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 μl/min and eluted at 0.2 μl/min or 0.25 μl/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

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.08.466737; this version posted November 8, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

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 single repeat count and 60 s exclusion duration. The mass spectrometric raw files were processed into peak lists using ProteoWizard (version 3.0.20388) [56], and cross-linked peptides were matched to spectra using Xi software (version 1.7.6.3) [57] (https://github.com/Rappsilber-Laboratory/XiSearch) with in-search assignment of monoisotopic peaks [58]. Search parameters were MS accuracy, 3 ppm; MS/MS accuracy, 10ppm; enzyme, trypsin; cross-linker, EDC; max missed cleavages, 4; missing mono-isotopic 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 $H_2O$, $NH_3$ and $CH_3SOH$. 5% on link level False discovery rate (FDR) was estimated based on the number of decoy identification using XiFDR [59].

**Integrative structure modelling**

To determine the structure of the complexes we used XlinkAssembler, an algorithm for multi-subunit assembly based on combinatorial docking approach [60, 61]. 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 [62]. 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.

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.08.466737; this version posted November 8, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

XlinkAssembler was used with 11 subunits to generate a model for Mis18α/β: initial hexamer structure based on AlphaFold [37], two Mis18α$_{Yippee}$ domains as well as four copies of the two helices in the Mis18α N-terminal helical region (residues 37-55 and 60-76). For docking Mis18BP1 helices, XlinkAssembler was used with 4 subunits: the Mis18α/β$_{Yippee}$ domains hetero-dimer 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), as well as HeLa CENP-A-SNAP, GFP Mis18BP1 inducible CENP-A-SNAP and mCherry Mis18α CENP-A-SNAP (kind gift from Iain Cheeseman [26]) were maintained in DMEM (Gibco) containing 10% FBS (Biowest) and 1X Penicillin/Streptomycin antibiotic mixture (Gibco). The cells were incubated at 37°C in a $CO_2$ incubator in humid condition containing 5% $CO_2$. GFP Mis18BP1 was induced with 10 μg/ml doxycycline for 18 h. siRNAs (AllStars Negative Control siRNA 1027280. Mis18α: ID s28851, Mis18β: ID s22367, ThermoFisher) were used in the rescue assays by transfecting the cells using jetPRIME® (Polyplus transfection®) reagent according to manufacturer's instructions. Briefly, HeLa CENP-A-SNAP, GFP Mis18BP1 inducible CENP-A-SNAP and mCherry Mis18α 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 on to the coverslip in a 12-well plate were transfected with the corresponding vectors (500 ng)

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.08.466737; this version posted November 8, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

and the transfection reagent diluted in Opti-MEM (Invitrogen) followed by incubation for 36-48 h.

**Generation of monoclonal antibodies against Mis18α/Mis18β**

Lou/c rats and C57BL/6J mice were immunized with 60 μg purified recombinant human Mis18α/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 three days before fusion of immune spleen cells with P3X63Ag8.653 myeloma cells using standard procedures. Hybridoma supernatants were screened for specific binding to Mis18α/β protein complex and also for binding to purified GST-Mis18β protein in ELISA assays. Positive supernatants were further validated by Western blot analyses on purified recombinant human Mis18α/β complex, on cell lysates from *Drosophila* S2 cells overexpressing human Mis18α 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-Mis18α (clone 25G8, mouse IgG2b/k) and rat anti-Mis18β (clone 24C8; rat IgG2a/k). These antibodies are not commercially available.

**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 Mis18α (25G8), rat Mis18β (24C8) (1:100, Helmholtz Zentrum München), Mis18BP1 (1:500, PA5-46777, Thermo Fisher or 1ug/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

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.08.466737; this version posted November 8, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

(Alpha Diagnostics). For imaging with the Odyssey® CLx system, F secondary antibodies were used.

**Co-Immunoprecipitation**

HeLa Kyoto cells were seeded in 100 mm dishes. The cells were depleted of the endogenous Mis18α or Mis18β by siRNA transfection with jetPRIME® (Polyplus transfection®) and simultaneously rescued with siRNA resistant versions of WT or mutant Mis18α mCherry and Mis18β GFP. The cells were harvested after 48 h and lysed by resuspending in immunoprecipitation buffer, 75 mM HEPES pH 7.5, 1.5mM EGTA, 1.5mM $MgCl_2$, 150mM NaCl, 10% glycerol, 0.1 % NP40, 1mM PMSF, 10 mM NaF, 0.3 mM Na-vanadate and cOmplete™ Mini Protease Inhibitor; adapted from [25]. 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 150mM or 300mM 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-

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.08.466737; this version posted November 8, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

A (1:100, MA 1-20832, Thermofisher). 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; 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 ImageJ software. The intensity at the tethering site was obtained using a custom-made plugin. Briefly, the CENP-A signal at the tethering site (eYFP) was found for every z-section within a 7-square pixel box. The mean signal intensity thus obtained was subtracted from the minimum intensities within the section, which was then normalised with the average CENP-A intensities of the endogenous centromeres. The values were obtained from a minimum of three biological repeats. Statistical significance of the difference between normalised intensities at the centromere and tethering region was established by a Mann–Whitney U two tailed test using Prism 9.1.2.

**SNAP-CENP-A assay and quantification**

SNAP-CENP-A quench pulse labelling was done as described previously [12]. Briefly, the existing CENP-A was quenched by 10 μM SNAP Cell® Block BTP (S9106S, NEB). 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, NEB), 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 [63]. Briefly, the centromeres were defined by a 7x7 pixel box using a reference channel, and the

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.08.466737; this version posted November 8, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

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. 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 Mis18$\alpha/\beta_{C\text{-term}}$

PDB ID: 7SFZ for Mis18$\alpha_{Yippee}$

The MS proteomics data will be deposited to the ProteomeXchange Consortium via the PRIDE [64]) partner repository.

**Code availability**

Plugin for analysing intensities at tethering site will be deposited in Zenodo.

**Acknowledgements**

We would like to thank David Kelly from the Centre Optical Instrumentation Laboratory as well as Marcus Wilson and Maarten Tuijtel of the Cryo-Electron Microscopy Facility for their help. We also thank Diamond Light Source and the staff of beamline B21 (proposal sm23510), as well as Advanced Photon Source and the staff at the 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 A.A. Jeyaprakash (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. The work of D. Schneidman is supported by ISF 1466/18 and Israeli Ministry of Science and Technology.

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.08.466737; this version posted November 8, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

The authors declare no competing financial interests.

Author contributions: A.A. Jeyaprakash and U.S. Cho conceived the project. R. Thamkachy, B. Medina-Pritchard, S.H. Park, M.A. Abad and A.A. Jeyaprakash designed the experiments. S.H. Park and K. Shimanaka preformed crystal structure characterisation. C.G. Chiodi and M. de la Torre-Barranco preformed EM characterisation. B. Medina-Pritchard and J. Zou preformed cross-linking analysis. C. Gallego Páramo and O.R. Davies analysed SAXS data. D. Schneidman-Duhovny and A.A. Jeyaprakash preformed modelling. R. Thamkachy and S.H. Park preformed biochemical characterisation. R. Thamkachy and B. Medina-Pritchard preformed *in vivo* assays and analysis. R. Feederle and E. Ruksentaite generated and validated anti-bodies. J. Rappsilber and P. Heun provided resources, expertise, and feedback. B. Medina-Pritchard, C.G. Chiodi, R. Thamkachy, D. Scheidman-Duhovny and A.A. Jeyaprakash wrote the manuscript.

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.08.466737; this version posted November 8, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

## References


1. Musacchio, A. and A. Desai, *A Molecular View of Kinetochore Assembly and Function.* Biology (Basel), 2017. **6**(1).
2. Cheeseman, I.M., *The kinetochore.* Cold Spring Harb Perspect Biol, 2014. **6**(7): p. a015826.
3. Catania, S. and R.C. Allshire, *Anarchic centromeres: deciphering order from apparent chaos.* Curr Opin Cell Biol, 2014. **26**: p. 41-50.
4. Fukagawa, T. and W.C. Earnshaw, *The centromere: chromatin foundation for the kinetochore machinery.* Dev Cell, 2014. **30**(5): p. 496-508.
5. McKinley, K.L. and I.M. Cheeseman, *The molecular basis for centromere identity and function.* Nat Rev Mol Cell Biol, 2016. **17**(1): p. 16-29.
6. Stellfox, M.E., A.O. Bailey, and D.R. Foltz, *Putting CENP-A in its place.* Cell Mol Life Sci, 2013. **70**(3): p. 387-406.
7. Black, B.E., et al., *Centromere identity, function, and epigenetic propagation across cell divisions.* Cold Spring Harb Symp Quant Biol, 2010. **75**: p. 403-18.
8. Carroll, C.W., K.J. Milks, and A.F. Straight, *Dual recognition of CENP-A nucleosomes is required for centromere assembly.* J Cell Biol, 2010. **189**(7): p. 1143-55.
9. Kato, H., et al., *A conserved mechanism for centromeric nucleosome recognition by centromere protein CENP-C.* Science, 2013. **340**(6136): p. 1110-3.
10. Weir, J.R., et al., *Insights from biochemical reconstitution into the architecture of human kinetochores.* Nature, 2016. **537**(7619): p. 249-253.
11. Dunleavy, E.M., G. Almouzni, and G.H. Karpen, *H3.3 is deposited at centromeres in S phase as a placeholder for newly assembled CENP-A in G(1) phase.* Nucleus, 2011. **2**(2): p. 146-57.
12. Jansen, L.E., et al., *Propagation of centromeric chromatin requires exit from mitosis.* J Cell Biol, 2007. **176**(6): p. 795-805.
13. Dunleavy, E.M., et al., *HJURP is a cell-cycle-dependent maintenance and deposition factor of CENP-A at centromeres.* Cell, 2009. **137**(3): p. 485-97.
14. Fujita, Y., et al., *Priming of centromere for CENP-A recruitment by human hMis18alpha, hMis18beta, and M18BP1.* Dev Cell, 2007. **12**(1): p. 17-30.
15. Foltz, D.R., et al., *Centromere-specific assembly of CENP-a nucleosomes is mediated by HJURP.* Cell, 2009. **137**(3): p. 472-84.
16. Barnhart, M.C., et al., *HJURP is a CENP-A chromatin assembly factor sufficient to form a functional de novo kinetochore.* J Cell Biol, 2011. **194**(2): p. 229-43.
17. Dambacher, S., et al., *CENP-C facilitates the recruitment of M18BP1 to centromeric chromatin.* Nucleus, 2012. **3**(1): p. 101-10.
18. Stellfox, M.E., et al., *Differential Binding Partners of the Mis18alpha/beta YIPPEE Domains Regulate Mis18 Complex Recruitment to Centromeres.* Cell Rep, 2016. **15**(10): p. 2127-2135.
19. Moree, B., et al., *CENP-C recruits M18BP1 to centromeres to promote CENP-A chromatin assembly.* J Cell Biol, 2011. **194**(6): p. 855-71.
20. Hu, H., et al., *Structure of a CENP-A-histone H4 heterodimer in complex with chaperone HJURP.* Genes Dev, 2011. **25**(9): p. 901-6.
21. Lagana, A., et al., *A small GTPase molecular switch regulates epigenetic centromere maintenance by stabilizing newly incorporated CENP-A.* Nat Cell Biol, 2010. **12**(12): p. 1186-93.
22. Perpelescu, M., et al., *Active establishment of centromeric CENP-A chromatin by RSF complex.* J Cell Biol, 2009. **185**(3): p. 397-407.

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.08.466737; this version posted November 8, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.


23. Silva, M.C., et al., *Cdk activity couples epigenetic centromere inheritance to cell cycle progression.* Dev Cell, 2012. **22**(1): p. 52-63.
24. Spiller, F., et al., *Molecular basis for Cdk1-regulated timing of Mis18 complex assembly and CENP-A deposition.* EMBO Rep, 2017. **18**(6): p. 894-905.
25. Pan, D., et al., *CDK-regulated dimerization of M18BP1 on a Mis18 hexamer is necessary for CENP-A loading.* Elife, 2017. **6**.
26. McKinley, K.L. and I.M. Cheeseman, *Polo-like kinase 1 licenses CENP-A deposition at centromeres.* Cell, 2014. **158**(2): p. 397-411.
27. Stankovic, A., et al., *A Dual Inhibitory Mechanism Sufficient to Maintain Cell-Cycle-Restricted CENP-A Assembly.* Mol Cell, 2017. **65**(2): p. 231-246.
28. Muller, S., et al., *Phosphorylation and DNA binding of HJURP determine its centromeric recruitment and function in CenH3(CENP-A) loading.* Cell Rep, 2014. **8**(1): p. 190-203.
29. Wang, J., et al., *Mitotic regulator Mis18beta interacts with and specifies the centromeric assembly of molecular chaperone holliday junction recognition protein (HJURP).* J Biol Chem, 2014. **289**(12): p. 8326-36.
30. Nardi, I.K., et al., *Licensing of Centromeric Chromatin Assembly through the Mis18alpha-Mis18beta Heterotetramer.* Mol Cell, 2016. **61**(5): p. 774-787.
31. Pan, D., et al., *Mechanism of centromere recruitment of the CENP-A chaperone HJURP and its implications for centromere licensing.* Nat Commun, 2019. **10**(1): p. 4046.
32. Zhang, D., C.J. Martyniuk, and V.L. Trudeau, *SANTA domain: a novel conserved protein module in Eukaryota with potential involvement in chromatin regulation.* Bioinformatics, 2006. **22**(20): p. 2459-62.
33. Aasland, R., A.F. Stewart, and T. Gibson, *The SANT domain: a putative DNA-binding domain in the SWI-SNF and ADA complexes, the transcriptional co-repressor N-CoR and TFIIIB.* Trends Biochem Sci, 1996. **21**(3): p. 87-8.
34. Maddox , P.S., et al., *Functional genomics identifies a Myb domain–containing protein family required for assembly of CENP-A chromatin.* Journal of Cell Biology, 2007. **176**(6): p. 757-763.
35. Subramanian, L., et al., *Centromere localization and function of Mis18 requires Yippee-like domain-mediated oligomerization.* EMBO Rep, 2016. **17**(4): p. 496-507.
36. Kelley, L.A., et al., *The Phyre2 web portal for protein modeling, prediction and analysis.* Nat Protoc, 2015. **10**(6): p. 845-58.
37. Jumper, J., et al., *Highly accurate protein structure prediction with AlphaFold.* Nature, 2021. **596**(7873): p. 583-589.
38. Källberg, M., et al., *Template-based protein structure modeling using the RaptorX web server.* Nat Protoc, 2012. **7**(8): p. 1511-22.
39. Kastner, B., et al., *GraFix: sample preparation for single-particle electron cryomicroscopy.* Nat Methods, 2008. **5**(1): p. 53-5.
40. Punjani, A., et al., *cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination.* Nat Methods, 2017. **14**(3): p. 290-296.
41. Sali, A. and T.L. Blundell, *Comparative protein modelling by satisfaction of spatial restraints.* J Mol Biol, 1993. **234**(3): p. 779-815.
42. Ohzeki, J., et al., *Breaking the HAC Barrier: histone H3K9 acetyl/methyl balance regulates CENP-A assembly.* Embo j, 2012. **31**(10): p. 2391-402.
43. Studier, F.W., *Protein production by auto-induction in high density shaking cultures.* Protein Expr Purif, 2005. **41**(1): p. 207-34.
44. Winter, G., et al., *DIALS: implementation and evaluation of a new integration package.* Acta Crystallogr D Struct Biol, 2018. **74**(Pt 2): p. 85-97.

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.08.466737; this version posted November 8, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.


45. Winter, G., C.M. Lobley, and S.M. Prince, *Decision making in xia2.* Acta Crystallogr D Biol Crystallogr, 2013. **69**(Pt 7): p. 1260-73.
46. Terwilliger, T.C., *Maximum-likelihood density modification.* Acta Crystallogr D Biol Crystallogr, 2000. **56**(Pt 8): p. 965-72.
47. Liebschner, D., et al., *Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix.* Acta Crystallogr D Struct Biol, 2019. **75**(Pt 10): p. 861-877.
48. Emsley, P. and K. Cowtan, *Coot: model-building tools for molecular graphics.* Acta Crystallogr D Biol Crystallogr, 2004. **60**(Pt 12 Pt 1): p. 2126-32.
49. Afonine, P.V., et al., *Joint X-ray and neutron refinement with phenix.refine.* Acta Crystallogr D Biol Crystallogr, 2010. **66**(Pt 11): p. 1153-63.
50. Chen, V.B., et al., *MolProbity: all-atom structure validation for macromolecular crystallography.* Acta Crystallogr D Biol Crystallogr, 2010. **66**(Pt 1): p. 12-21.
51. Konarev, P.V., et al., *PRIMUS: a Windows PC-based system for small-angle scattering data analysis.* Journal of Applied Crystallography, 2003. **36**(5): p. 1277-1282.
52. Svergun, D.I., *Restoring low resolution structure of biological macromolecules from solution scattering using simulated annealing.* Biophys J, 1999. **76**(6): p. 2879-86.
53. Stark, H., *GraFix: stabilization of fragile macromolecular complexes for single particle cryo-EM.* Methods Enzymol, 2010. **481**: p. 109-26.
54. Scarff, C.A., et al., *Variations on Negative Stain Electron Microscopy Methods: Tools for Tackling Challenging Systems.* J Vis Exp, 2018(132).
55. Zhang, K., *Gctf: Real-time CTF determination and correction.* J Struct Biol, 2016. **193**(1): p. 1-12.
56. Kessner, D., et al., *ProteoWizard: open source software for rapid proteomics tools development.* Bioinformatics, 2008. **24**(21): p. 2534-6.
57. Mendes, M.L., et al., *An integrated workflow for cross-linking/mass spectrometry.* bioRxiv, 2018.
58. Lenz, S., et al., *In-Search Assignment of Monoisotopic Peaks Improves the Identification of Cross-Linked Peptides.* J Proteome Res, 2018. **17**(11): p. 3923-3931.
59. Fischer, L. and J. Rappsilber, *Quirks of Error Estimation in Cross-Linking/Mass Spectrometry.* Anal Chem, 2017. **89**(7): p. 3829-3833.
60. Schneidman-Duhovny, D. and H.J. Wolfson, *Modeling of Multimolecular Complexes.* Methods Mol Biol, 2020. **2112**: p. 163-174.
61. Inbar, Y., et al., *Prediction of multimolecular assemblies by multiple docking.* J Mol Biol, 2005. **349**(2): p. 435-47.
62. Schneidman-Duhovny, D., et al., *PatchDock and SymmDock: servers for rigid and symmetric docking.* Nucleic Acids Res, 2005. **33**(Web Server issue): p. W363-7.
63. Bodor, D.L., et al., *Analysis of protein turnover by quantitative SNAP-based pulse-chase imaging.* Curr Protoc Cell Biol, 2012. **Chapter 8**: p. Unit8.8.
64. Perez-Riverol, Y., et al., *The PRIDE database and related tools and resources in 2019: improving support for quantification data.* Nucleic Acids Res, 2019. **47**(D1): p. D442-d450.
65. Madeira, F., et al., *The EMBL-EBI search and sequence analysis tools APIs in 2019.* Nucleic Acids Res, 2019. **47**(W1): p. W636-W641.
66. Bond, C.S. and A.W. Schüttelkopf, *ALINE: a WYSIWYG protein-sequence alignment editor for publication-quality alignments.* Acta Crystallogr D Biol Crystallogr, 2009. **65**(Pt 5): p. 510-2.
67. Baker, N.A., et al., *Electrostatics of nanosystems: application to microtubules and the ribosome.* Proc Natl Acad Sci U S A, 2001. **98**(18): p. 10037-41.

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.08.466737; this version posted November 8, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

**Table 1. Crystallographic Data collection and refinement statistics**

| | **SeMet-Mis18α/β$_{C_term}$** | **Mis18α/β$_{C\text{-}term}$** | **Mis18α$_{Yippee}$** |
|---|---|---|---|
| Space group | $C222_1$ | $C222_1$ | $P2_12_12$ |
| Unit cell parameters (Å) | a=77.73<br>b=101.91<br>c=88.51<br>α=β=γ=90° | a=77.591<br>b=101.815<br>c=88.563<br>α=β=γ=90° | a=110.725<br>b=114.864<br>c=116.279<br>α=β=γ=90° |
| Wavelength (Å) | 0.97917 | 0.97856 | 1.12713 |
| **Data collection statistics** | | | |
| Resolution range (Å) | 50.00-2.75 (2.75-2.80) | 50.00-2.50 (2.50-2.54) | 50.00-3.00 (3.05-3.00) |
| Number of unique reflections | 9402 (469) | 12467 (600) | 30063 (1479) |
| Completeness (%) | 99.4 (98.3) | 99.0 (96.0) | 99.3 (99.7) |
| $R_{merge}$ | 0.387 (5.327) | 0.078 (0.919) | 0.105 (0.804) |
| $R_{pim}$ | 0.112 (1.456) | 0.040 (0.447) | 0.055 (0.410) |
| Redundancy | 13.4 (12.7) | 5.2 (4.7) | 4.4 (4.5) |
| Mean I/σ | 8.7 (1.5) | 9.2 (1.8) | 16.5 (1.4) |
| **Refinement statistics** | | | |
| Resolution range (Å) | | 44.29-2.50 | 49.99-3.00 |
| $R_{work}/R_{free}$ (%) | | 24.77/27.96 | 20.26/25.00 |
| RMSD bonds (Å) | | 0.008 | 0.012 |
| RMSD angles (deg) | | 1.069 | 1.294 |
| Average B factor (Å$^2$) | | 91.18 | 83.84 |
| Number of water molecules | | 17 | 8 |
| Ramachandran favoured (%) | | 97.89 | 95.51 |
| allowed (%) | | 2.11 | 4.49 |
| not allowed (%) | | 0 | 0 |

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.08.466737; this version posted November 8, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

## Figures

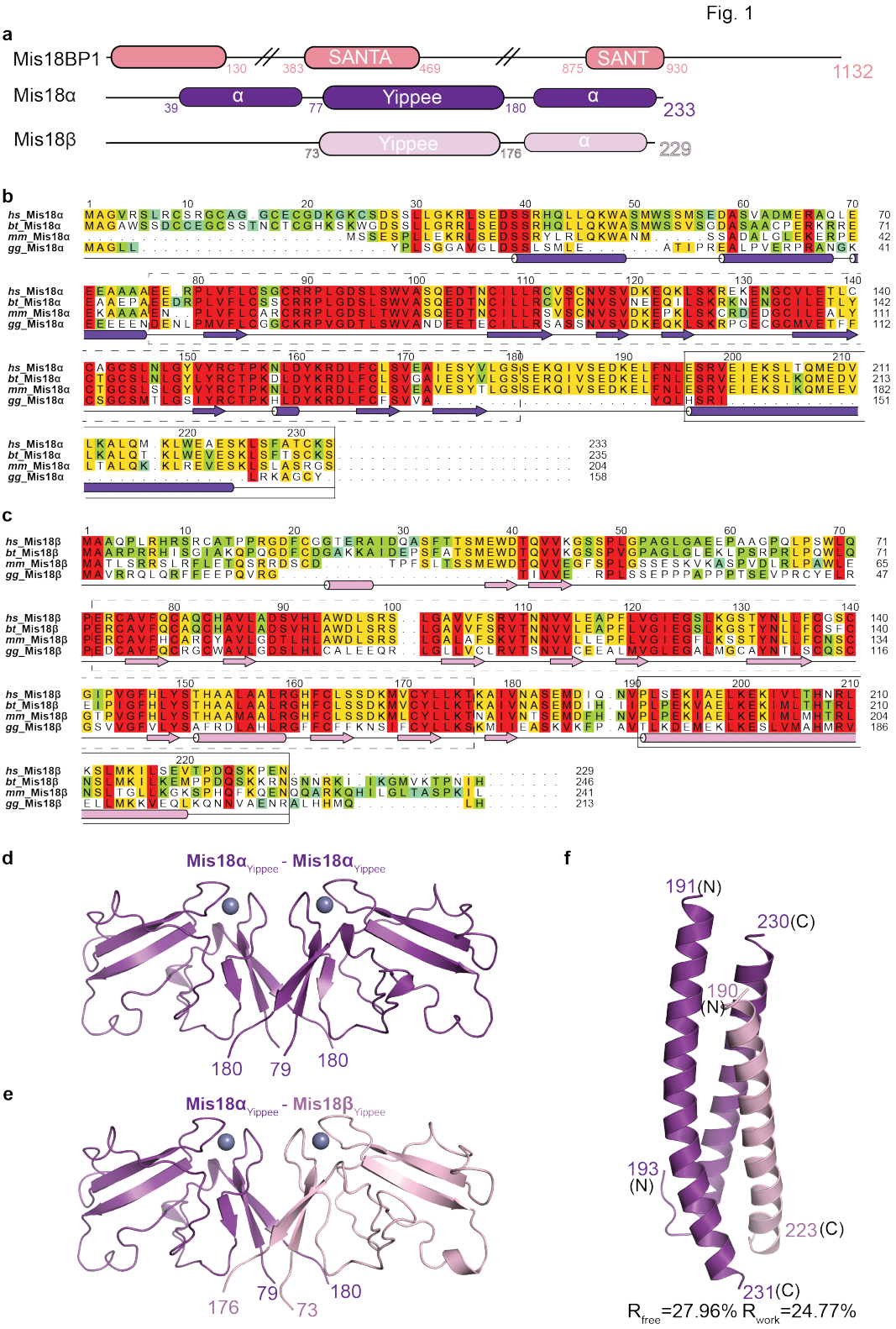


Figure 1: Mis18α/β Contains Two Independent Structural Domains that can Oligomerise.

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.08.466737; this version posted November 8, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

**a**) Schematic representation of structural features of Mis18BP1 (salmon), Mis18α (purple) and Mis18β (light pink). Filled boxes represent folded domains. SANTA and SANT domain boundaries as defined in UniProt (Q6P0N0).

**b** & **c**) Domain architecture and amino acid conservation of (**b**) Mis18α and (**c**) Mis18β. Alignments include *Homo sapiens* (*hs*), *Bos taurus* (*bt*), *Mus musculus* (*mm*) and *Gallus gallus* (*gg*). Conservation score is mapped from red to cyan, where red corresponds to highly conserved and cyan to poorly conserved. Secondary structures as annotated/predicted by Conserved Domain Database [CDD] and PsiPred, http://bioinf.cs.ucl.ac.uk/psipred. Multiple sequence alignments were performed with MUSCLE [65] and edited with Aline [66]. Dashed boxes highlight Yippee domains whilst solid boxes highlight C-terminus α-helices.

**d**) Cartoon representation of the crystal structure of human Mis18$\alpha_{Yippee}$ homo-dimer (PDB ID: 7SFZ).

**e**) Cartoon representation of the homology-modelled human Mis18$\alpha_{Yippee}$/Mis18$\beta_{Yippee}$ hetero-dimer based on the structure in **Fig. 1d**. Mis18α is shown in purple and Mis18β in light pink (modelled using Phyre2, www.sbg.bio.ic.ac.uk/phyre2/ [36]).

**f**) Cartoon representation of the crystal structure of Mis18$\alpha_{C\text{-}term}$/Mis18$\beta_{C\text{-}term}$ (PDB ID: 7SFY). Mis18α is shown in purple and Mis18β in light pink.

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.08.466737; this version posted November 8, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Fig. 2

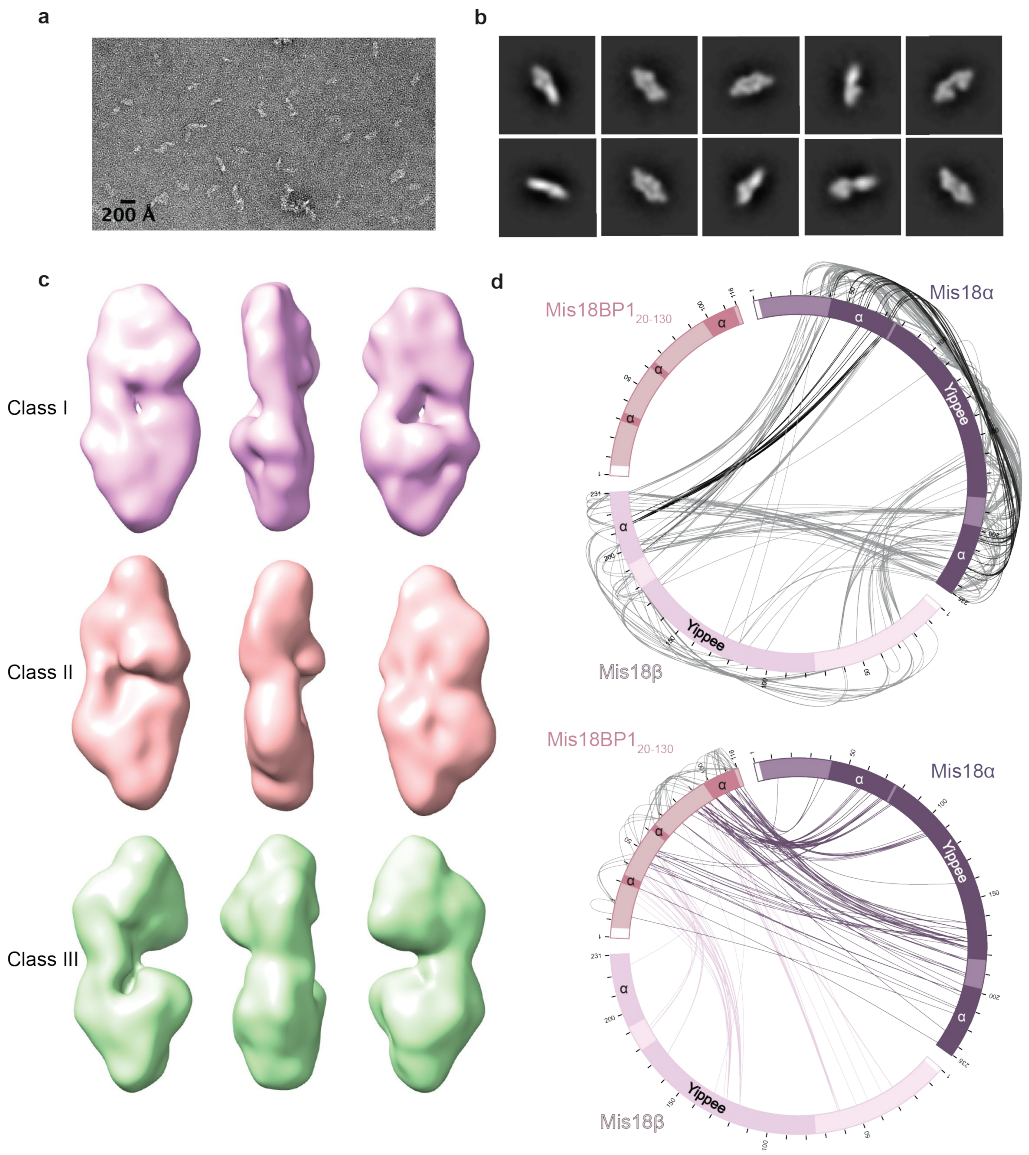


Figure 2: Structural Characterisation of the $Mis18_{core}$ Complex

**a**) Representative micrograph of negative staining EM of the Mis18α/Mis18β/$Mis18BP1_{20-130}$ ($Mis18_{core}$) complex cross-linked using GraFix [39, 53].

**b**) Representative images of 2D classes from $Mis18_{core}$ particles picked using CryoSPARC [40].

**c**) Three models (Class I-III) generated for $Mis18_{core}$ from negative staining EM analysis. All three show that the overall shapes of the $Mis18_{core}$ resemble a telephone handset with 'ear' and 'mouth' pieces assuming different relative orientations.

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.08.466737; this version posted November 8, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

**d**) Linkage map showing the sequence position and cross-linked residue pairs between the different $Mis18_{core}$ complex subunits, Mis18α, Mis18β and $Mis18BP1_{20-130}$. Top panel highlights cross-linked residues between Mis18α and Mis18β. Black lines highlight cross-links between N- and C- terminal helical regions of Mis18α. Bottom panel highlights cross-links observed between i) $Mis18BP1_{20-130}$ and Mis18α (purple) ii) $Mis18BP1_{20-130}$ and Mis18β (light pink) iii) $Mis18BP1_{20-130}$ 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.

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.08.466737; this version posted November 8, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Fig. 3

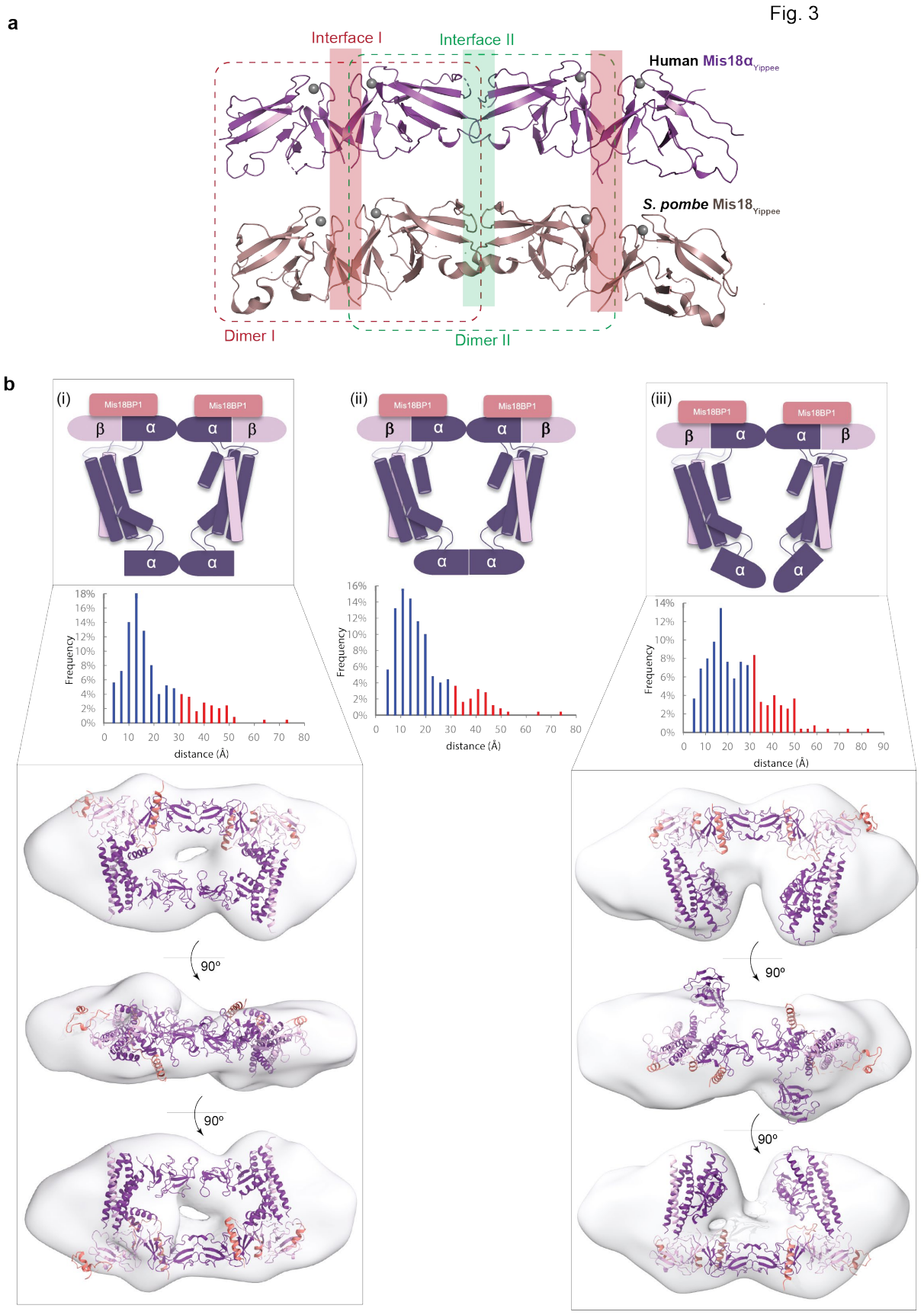


Figure 3: Overall architecture of the $Mis18_{core}$ complex

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.08.466737; this version posted November 8, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

**a**) Cartoon representation of human Mis18$\alpha_{Yippee}$ (top panel) and *S. pombe* Mis18$_{Yippee}$ (bottom panel) oligomers as observed in their respective crystal structures. Crystal packing interactions of human Mis18$\alpha_{Yippee}$ and *S. pombe* Mis18$_{Yippee}$ are strikingly similar and involve two different Yippee domain interfaces for homo-oligomerisation.

**b**) Models of the Mis18$_{core}$ complex generated using partial structures determined using X-ray crystallography and AlphaFold2 [37] and cross-linking restrained molecular docking in EM maps. Cartoon diagrams (top panel) describe the overall architecture and relative orientations of domains for the generated models. Histograms underneath the models show percentage of satisfied (blue) and violated (red) cross-links. Bottom panel shows 3D volumes obtained from EM analysis and corresponding best fitting models.

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.08.466737; this version posted November 8, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

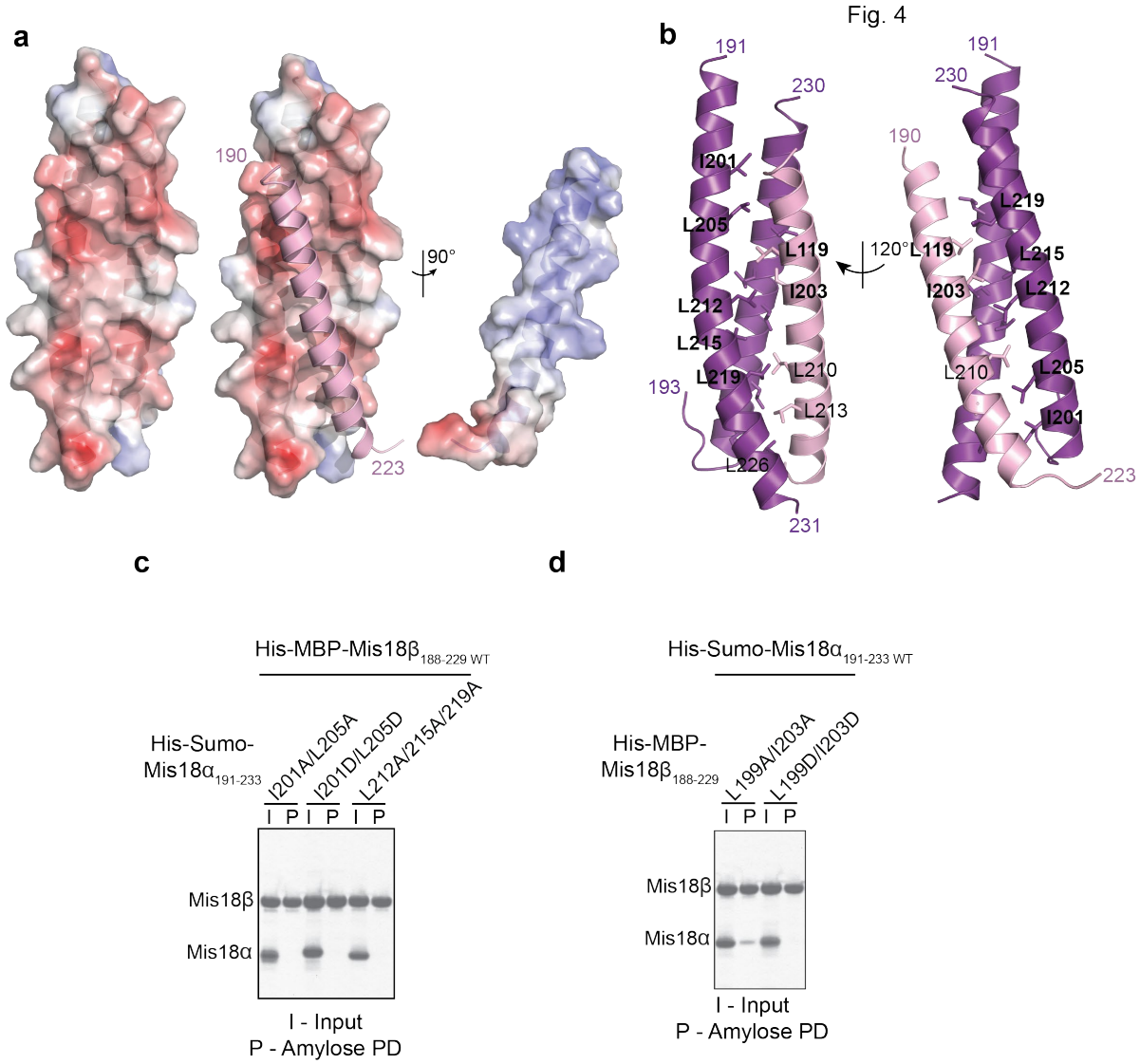


Figure 4: Hydrophobic Interactions Stabilise Mis18α/β C-terminal Helical Bundle Assembly

**a**) Mis18$\alpha_{C\text{-term}}$ domains are shown in surface representation and coloured based on electrostatic surface potential calculated using APBS [67]. Mis18$\beta_{C\text{-term}}$ shown as cartoon.

**b**) Mis18$\alpha_{C\text{-term}}$/Mis18$\beta_{C\text{-term}}$ domains are shown in cartoon representation in two different orientations. Potential residues involved in the interaction are highlighted. Mis18α (purple) and Mis18β (light pink).

**c** & **d**) SDS-PAGE analysis of cobalt and amylose pull-down of His-MBP-Mis18$\beta_{188-229\ WT}$ with His-SUMO-Mis18$\alpha_{191-233}$ mutants and His-SUMO-Mis18$\alpha_{191-233\ WT}$ with His-

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.08.466737; this version posted November 8, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

MBP-Mis18$\beta_{188-229}$ mutants, respectively. SDS-PAGE shows protein bound to nickel resin as input (I) and protein bound to amylose resin to assess interaction (P).

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.08.466737; this version posted November 8, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

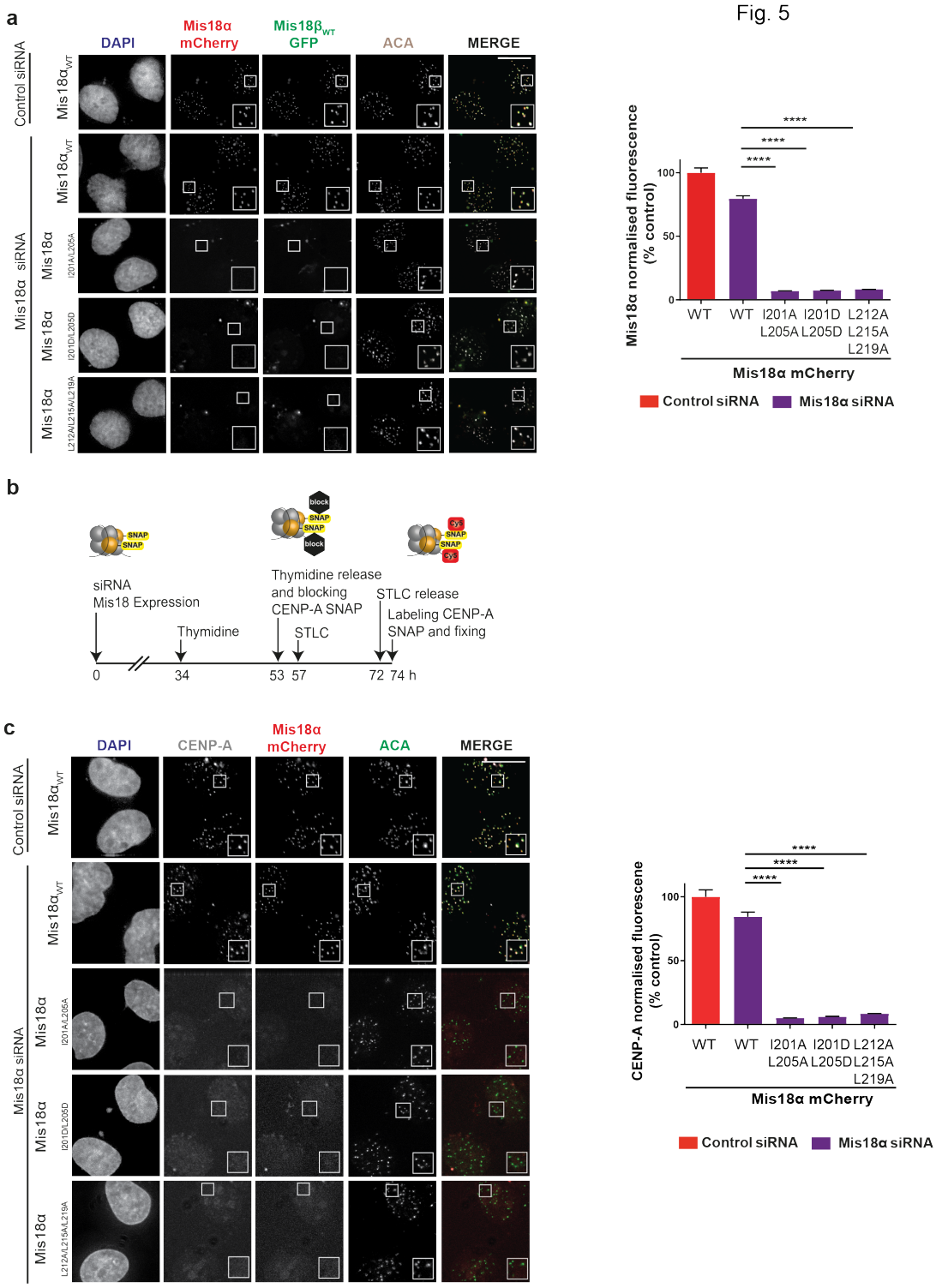


Figure 5: Mis18α Mutations Disrupting the Mis18α/β Triple Helical Assembly Result in Loss of Mis18α/β Centromere Localisation and CENP-A Deposition.

**a**) Representative fluorescence images (left panel) and quantification (right panel) assessing the ability of Mis18$\alpha_{WT}$ mCherry, Mis18$\alpha_{I201A/L205A}$ mCherry, Mis18$\alpha_{I201D/L205D}$ mCherry and Mis18$\alpha_{L212A/L215A/L219A}$ mCherry to co-localise with Mis18β GFP at

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.08.466737; this version posted November 8, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

endogenous centromeres in HeLa (Mann-Whitney U test; ****$P < 0.0001$, $n \geq 1236$). Cells were co-transfected with either control or Mis18α siRNA as stated, 3 independent experiments. Error bars show ±SEM. Scale bars, 10 μm. All conditions have been normalised to control conditions: cells transfected with control siRNA and Mis18$\alpha_{WT}$ mCherry.

**b**) Schematic representation of the experimental set-up used to evaluate the effect of Mis18α and Mis18β mutants on new CENP-A-SNAP loading.

**c**) Representative fluorescence images (left panel) and quantification (right panel) assessing the ability of Mis18$\alpha_{WT}$ mCherry, Mis18$\alpha_{I201A/L205A}$ mCherry, Mis18$\alpha_{I201D/L205D}$ mCherry and Mis18$\alpha_{L212A/L215A/L219A}$ mCherry to deposit new CENP-A-SNAP at endogenous centromeres (Mann-Whitney U test; ****$P < 0.0001$, $n \geq 886$). Cells were co-transfected with either control or Mis18α siRNA as stated, 3 independent experiments. Error bars show ±SEM. Scale bars, 10 μm. All conditions have been normalised to control conditions: cells transfected with control siRNA and Mis18$\alpha_{WT}$ mCherry.

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.08.466737; this version posted November 8, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Fig. 6

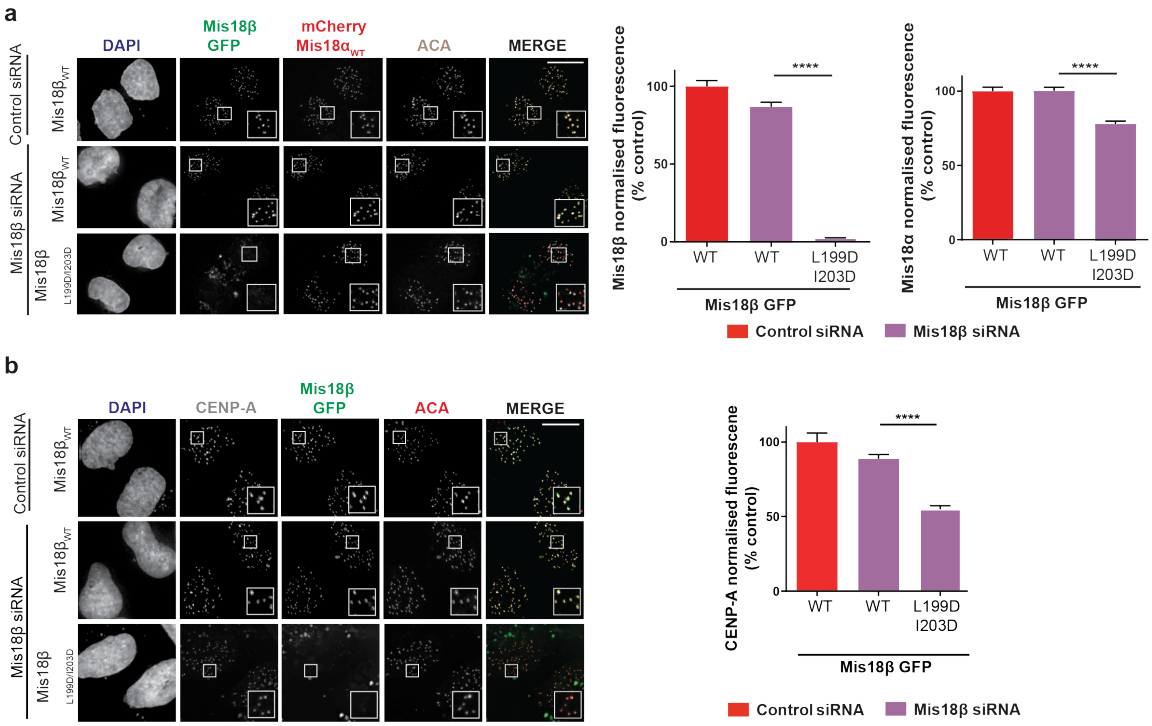


Figure 6: Mis18α Associates with Centromeres in a Mis18β-Independent Manner but Requires Mis18β for Efficient CENP-A Loading.

**a** & **b**) Representative fluorescence images (left panel) and quantification (right panel) used to evaluate the ability of Mis18β$_{WT}$ GFP and Mis18β$_{L199D/I203D}$ GFP to (**a**) co-localise with mCherry Mis18α at endogenous centromeres. Middle panel, quantification of Mis18β signal. Right panel, quantification of Mis18α signal (Mann-Whitney U test; ****$P$ < 0.0001, $n \geq 927$), (**b**) deposit new CENP-A-SNAP at endogenous centromeres (Mann-Whitney U test; ****$P$ < 0.0001, $n \geq 947$). Cells were co-transfected with either control or Mis18β siRNA as stated, 3 independent experiments. Error bars show ±SEM. Scale bars, 10 μm. All conditions have been normalised to control conditions: cells transfected with control siRNA and Mis18β$_{WT}$ GFP.

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.08.466737; this version posted November 8, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Fig. 7

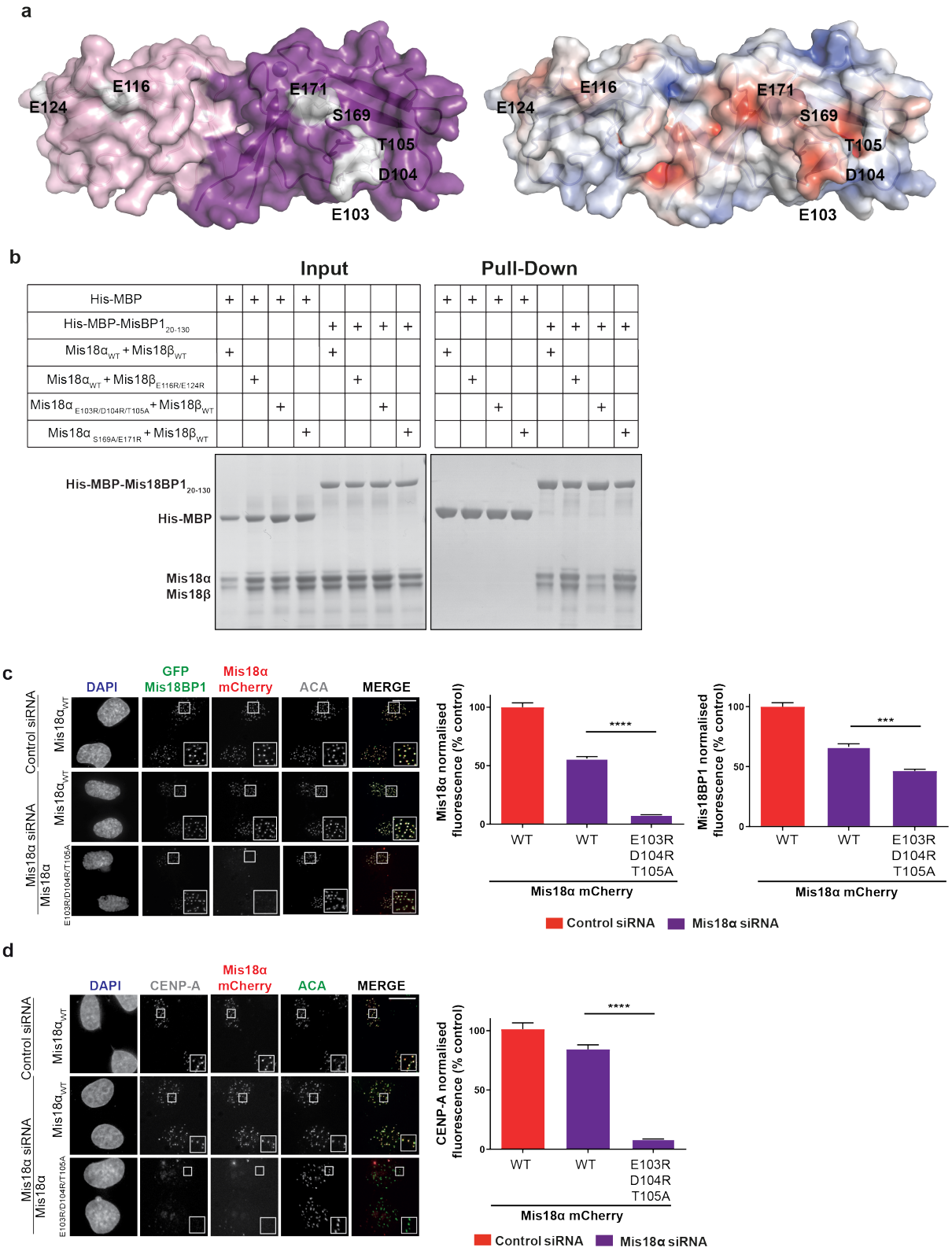


Figure 7: Disrupting the Mis18BP1 Binding Interface of Mis18α Prevents its Centromere Localisation and CENP-A Deposition.

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.08.466737; this version posted November 8, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

**a**) Mis18α/Mis18β$_{Yippee}$ model (left panel) and its surface representation coloured based on electrostatic surface potential (right panel) highlighting the residues proposed to be involved in Mis18BP1 binding. Mis18α shown in purple and Mis18β shown in light pink. **b**) SDS-PAGE analysis of amylose pull-down of His-MBP and His-MBP-Mis18BP1$_{20-130}$ with Mis18α/Mis18β$_{WT}$ complex and mutant complexes. SDS-PAGE shows input and protein bound to amylose resin to assess interaction.

**c** & **d**) Representative fluorescence images (left panel) and quantifications (right panel) assessing the ability of Mis18α$_{WT}$ mCherry and Mis18α$_{E103R/D104R/T105A}$ to (**c**) co-localise with GFP Mis18BP1 at endogenous centromeres. Middle panel, quantification of Mis18α signal and right panel, quantification of Mis18BP1 signal (Mann-Whitney U test; ****$P$ < 0.0001, $n$ ≥ 856), (**d**) deposit new CENP-A-SNAP at endogenous centromeres (Mann-Whitney U test; ***$P$ = 0.0001, ****$P$ < 0.0001, $n$ ≥ 896). Cells were co-transfected with either control or Mis18α siRNA as stated, 3 independent experiments. Error bars show ±SEM. Scale bars, 10 μm. All conditions have been normalised to control conditions: cells transfected with control siRNA and Mis18α$_{WT}$ mCherry.