Evolutionary conservation of the structure and function of meiotic Rec114–Mei4 and Mer2 complexes

Dima Daccache[1], Pascaline Liloku[1], Emma De Jonge[1], Alexander N. Volkov[2,3,*], and Corentin Claeys Bouuaert[1,*]

[1] Louvain Institute of Biomolecular Science and Technology, Université catholique de Louvain, 1348 Louvain-La-Neuve, Belgium.
[2] VIB-VUB Center for Structural Biology, VIB, Pleinlaan 2, Brussels, Belgium.
[3] Jean Jeener NMR Centre, Vrije Universiteit Brussel (VUB), Pleinlaan 2, Brussels, Belgium.
* Correspondence to corentin.claeys@uclouvain.be (C.C.B.), Oleksandr.Volkov@vub.be (A.N.V.).

Meiosis-specific Rec114–Mei4 and Mer2 complexes are thought to enable Spo11-mediated DNA double-strand-break (DSB) formation through a mechanism that involves DNA-dependent condensation. However, the structure, molecular properties, and evolutionary conservation of Rec114–Mei4 and Mer2 are unclear. Here, we present structures of Rec114–Mei4 and Mer2 complexes, supported by AlphaFold modeling, nuclear magnetic resonance (NMR) spectroscopy, crosslinking-mass spectrometry (XL-MS), and mutagenesis. We show that dimers composed of the Rec114 C-terminus form α-helical chains that cup an N-terminal Mei4 α-helix, and that Mer2 forms a parallel homotetrameric coiled coil. Both Rec114–Mei4 and Mer2 bind preferentially to branched DNA substrates, indicative of multivalent protein-DNA interactions. Indeed, the Rec114–Mei4 interaction domain contains two independent DNA-binding sites that point in opposite directions and likely drive condensation. In addition, Mer2 binds efficiently to nucleosomes *in vitro*, while Rec114–Mei4 does not. Finally, we show that the structure and properties of Rec114–Mei4 and Mer2 are conserved across eukaryotes. This work provides insights into the molecular mechanism whereby Rec114–Mei4 and Mer2 complexes promote the assembly of the meiotic DSB machinery.

Introduction

Throughout eukaryotes, the formation of haploid gametes requires Spo11-dependent catalysis of DNA double-strand breaks (DSBs) to initiate meiotic recombination, which is essential for the accurate segregation of homologous chromosomes (Hunter, 2015). Spo11 activity relies on a higher-order assembly that is tied to chromosome structure and subject to overlapping regulatory pathways that control the timing, number and distribution of DSBs (Lam and Keeney, 2015; Yadav and Claeys Bouuaert, 2021). However, the molecular assemblies required for meiotic DSB formation are not well characterized, and the degree to which they are conserved is unclear.

From fungi to plants and animals, Spo11 activity depends on a cohort of accessory factors. While Spo11 itself is ubiquitous and well-conserved at the sequence level, the auxiliary proteins that constitute the DSB machinery vary more broadly between organisms, and functional homologs tend to be highly divergent (de Massy, 2013; Keeney, 2008; Lam and Keeney, 2015). Nevertheless, some of the key partners are found throughout eukaryotes, including a sub-group referred to as RMM (Rec114, Mei4 and Mer2) in *S. cerevisiae* (Arora et al., 2004; Kumar et al., 2010; Li et al., 2006; Maleki et al., 2007; Stanzione et al., 2016; Tesse et al., 2017; Vrielynck et al., 2021).

We recently showed that the RMM proteins constitute two distinct sub-complexes, a Rec114–Mei4 heterotrimer and a Mer2 homotetramer (Claeys Bouuaert et al., 2021). *In vitro*, both complexes undergo DNA-driven condensation independently, and can mingle together to form mixed condensates. Condensation is a fundamental property of RMM proteins, and mutations that compromise DNA binding reduce condensation *in vitro* and *in vivo* and abolish meiotic DSB formation, suggesting that this activity is important for their biological function (Claeys Bouuaert et al., 2021). Hence, we proposed that RMM condensation organizes discrete chromatin sub-compartments within which DSB formation takes place. Nevertheless, little was known regarding the structures of Rec114–Mei4 and Mer2 complexes, how they relate to their biological functions, and whether their structural and molecular properties are conserved.

Here, we address this using AlphaFold structural modeling, supported by biochemical characterizations. We show that Rec114–Mei4 and Mer2 complexes show similar architectures throughout eukaryotes, and reveal mechanistic insights into their multivalent interactions with DNA that underlie the assembly of the DSB machinery by DNA-driven condensation.

Results

Structure of a minimal Rec114–Mei4 complex

Rec114 and Mei4 form a heterotrimeric complex with a 2:1 stoichiometry where the C-terminus of Rec114 homodimerizes and interacts with the N-terminus of Mei4 (**Figure 1A**) (Claeys Bouuaert et al., 2021). We purified the Rec114 dimerization domain (residues 375-428) and a minimal trimeric Rec114–Mei4 complex (that includes Mei4 residues 1-43). Thermal shift analyses revealed melting temperatures of 74.5 ± 1.0 °C and 81.5 ± 0.9 °C for Rec114 and the Rec114–Mei4 complex, respectively, indicating that the presence of Mei4 stabilizes Rec114 (**Figure 1B**).

We used AlphaFold to predict the structure of the minimal Rec114–Mei4 complex (Jumper et al., 2021). This yielded a high-quality model showing two Rec114 α-helical chains (residues 399 to 426) cupping a Mei4 α-helix (residues 16 to 29) (**Figure 1C**). Quality assessment indicated high-confidence predictions of the folded regions and the relative orientations of the interacting domains (**Supplementary Figure 1**).

To gain experimental evidence to support this structural model, we purified isotopically labeled U-[^{13}C, ^{15}N] Rec114 and Rec114–Mei4 complexes, and studied the proteins by nuclear magnetic resonance (NMR) spectroscopy. Although NMR analyses did not allow us to determine the structures of the complexes *de novo* (**Supplementary Figure 2A**), the data strongly support the AlphaFold models.

First, AlphaFold predicts that Rec114 and Mei4 peptides feature two α-helices preceded by N-terminal unstructured tails. This topology is confirmed by NMR chemical shift index analysis that revealed a good agreement between the NMR and AlphaFold α-helical regions (**Supplementary Figure 2B)**.

Second, the presence of Mei4 breaks the symmetry of the Rec114 dimer. This is clearly seen in the HSQC spectrum of the Rec114–Mei4 complex, where multiple Rec114 residues give rise to two backbone amide resonances (**Figure 1D, E**). This indicates that the same Rec114 amino acid experiences different chemical environments in the two protein chains. Mapping the differences in chemical shifts of the double Rec114 HSQC peaks shows that the largest effects are observed for the residues in the first α-helix (residues 399-407) and the C-terminal part of the protein (residues 423-428) (**Supplementary Figure 2C**). Indeed, these are the Rec114 regions with the highest dissimilarity in the predicted structure of the complex, where residues in one protein chain

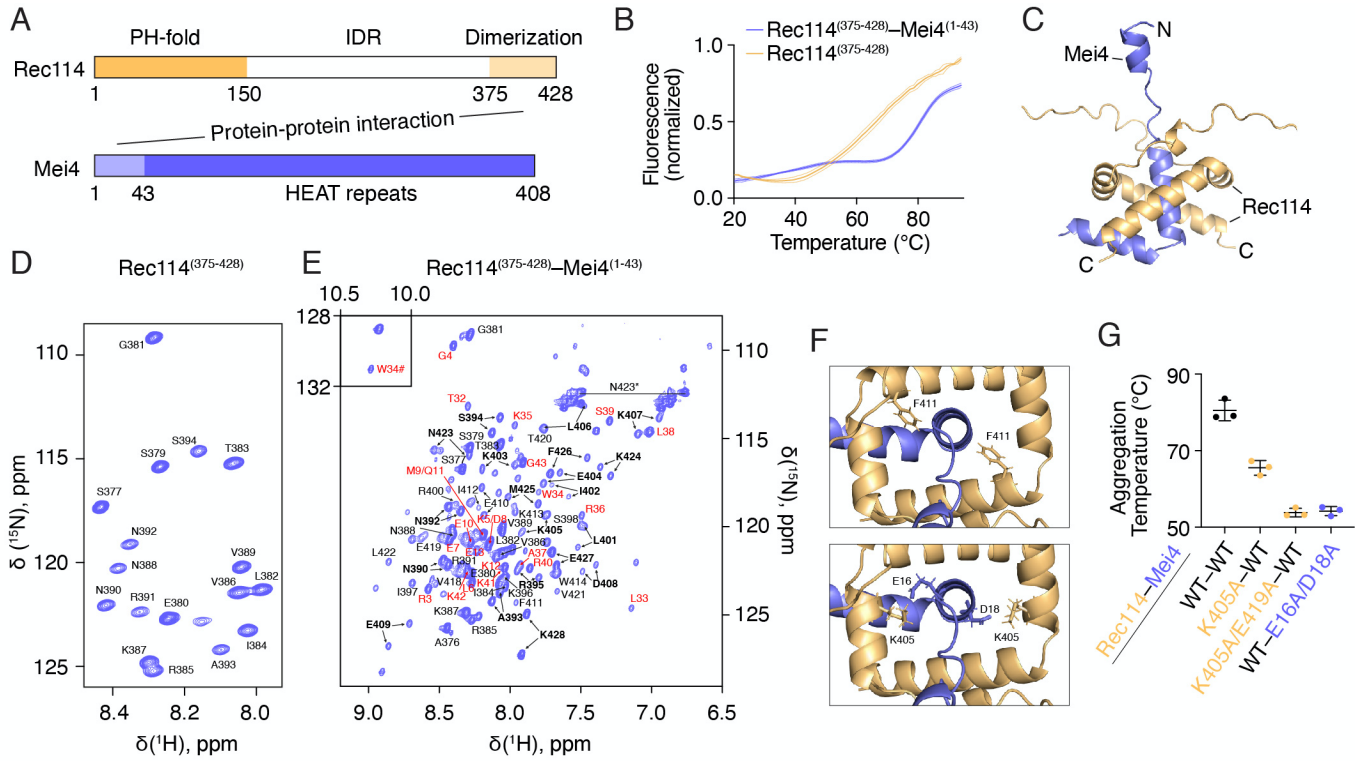


Figure 1: Structure of a minimal Rec114–Mei4 complex.
(A) Domain organization of Rec114 and Mei4 proteins. PH domain, Pleckstrin Homology; IDR, intrinsically-disordered region. (B) Thermal shift analyses of Rec114 C-terminal domain and the minimal trimeric Rec114–Mei4 complex. Error bars show mean ± SD of three replicates. (C) AlphaFold structure of a minimal Rec114–Mei4 complex. (D) [^{1}H,^{15}N] HSQC spectrum of the Rec114 C-terminal domain annotated with its backbone amide assignments. (E) [^{1}H,^{15}N] HSQC spectrum of the minimal trimeric Rec114–Mei4 complex. Black and red labels show backbone amide resonance assignments of Rec114 and Mei4, respectively. The Rec114 residues exhibiting two sets of peaks are in bold. The side-chain NH_2 peaks of N423 are indicated by an asterisk and joined by a horizontal line. The indole amide resonance of Mei4 W34 is labelled by a hash symbol and shown in the inset. (F) Protein-protein interaction interface between Rec114 and Mei4. Rec114-F411 was previously shown to be important for the interaction with Mei4 and for DSB formation (Claeys Bouuaert et al., 2021). Mei4 residues E16 and D18 make strong hydrogen bonds with Rec114 K405 in both dimer chains. (G) Thermal shift analysis of wild type and mutant Rec114–Mei4 minimal complexes. Error bars show mean ± SD of three replicates.

interact with the C-terminal α-helix of Mei4 (residues 32-43), while the same groups in the other chain do not.

Third, we recorded and analyzed nuclear Overhauser effect spectra (NOESY), which allow to detect pairs of ^{1}H atoms lying in close spatial proximity (typically < 5Å (Wüthrich, 1986)). We successfully assigned a number of sidechain methyl and aromatic groups of both Rec114 and Mei4, which are buried in the core of the protein complex and, thus, can be used to detect specific residue-residue contacts. A set of well-resolved methyl resonances of Leu, Ile, Val, and Met, as well as aromatic protons of Trp and Phe, were inspected in NOESY spectra, which allowed identification of several key ^{1}H-^{1}H interactions (**Supplementary Figure 2D**). In total, 13 unambiguous NOEs were detected (**Supplementary Table 1**). In particular, we observed contacts between residues in the same protein chain, interchain Rec114–Rec114 interactions, and several intermolecular Rec114–Mei4 contacts. Overall, the observed NOEs are fully consistent with the predicted structure and validate the AlphaFold model of the minimal heterotrimeric Rec114–Mei4 complex.

Next, we sought to test the model by mutagenesis. We previously showed that a Rec114-F411A mutation abolishes the interaction with Mei4 in yeast-two-hybrid and pulldown assays, and abolishes meiotic DSB formation (Claeys Bouuaert et al., 2021). The AlphaFold model indicates that the F411 residue points directly towards Mei4 (**Figure 1F**, top), explaining these results. Based on the model, we selected for mutagenesis other residues located within the predicted interaction surface. Rec114 residues K405 and E419, and Mei4 residues E16 and D18 make strong intermolecular hydrogen bonds (**Figure 1F**, bottom). Alanine substitutions did not completely abolish complex formation but strongly lowered the aggregation temperature of the complex in a thermal shift assay, indicating that these residues indeed stabilize the complex (**Figure 1G**).

Model of full-length Mei4 bound to Rec114

Next, we examined the architecture of full-length Rec114–Mei4 complexes. Rec114 is predicted to have a long central intrinsically-disordered region preceded by a structured N-terminal domain (Claeys Bouuaert et al., 2021). The N-terminal domain of *M. musculus* REC114 has been crystallized and shows a Pleckstrin Homology (PH)-like fold composed of an α-helix sandwiched between two anti-parallel β-sheets (Boekhout et al., 2019; Kumar et al., 2018). AlphaFold predicted with high confidence a similar structural fold for the yeast Rec114 N-terminal domain (residues 1-140) (**Supplementary Figure 5A**).

In addition, AlphaFold prediction of full-length Mei4 revealed a structure composed of 19 α-helices, with the two N-terminal Rec114-binding helices pointing out of this ordered structure (**Figure 2A**). Analysis of the Mei4 structural fold using the DALI server (Holm, 2022) revealed similarities with HEAT-repeat proteins, including NOT1 (PDB 5FU7 (Raisch et al., 2016) with a Cα RMSD of 3.7 Å over 202 residues) and CAND1 (PDB 4A0C (Fischer et al., 2011) with a Cα RMSD of 3.5

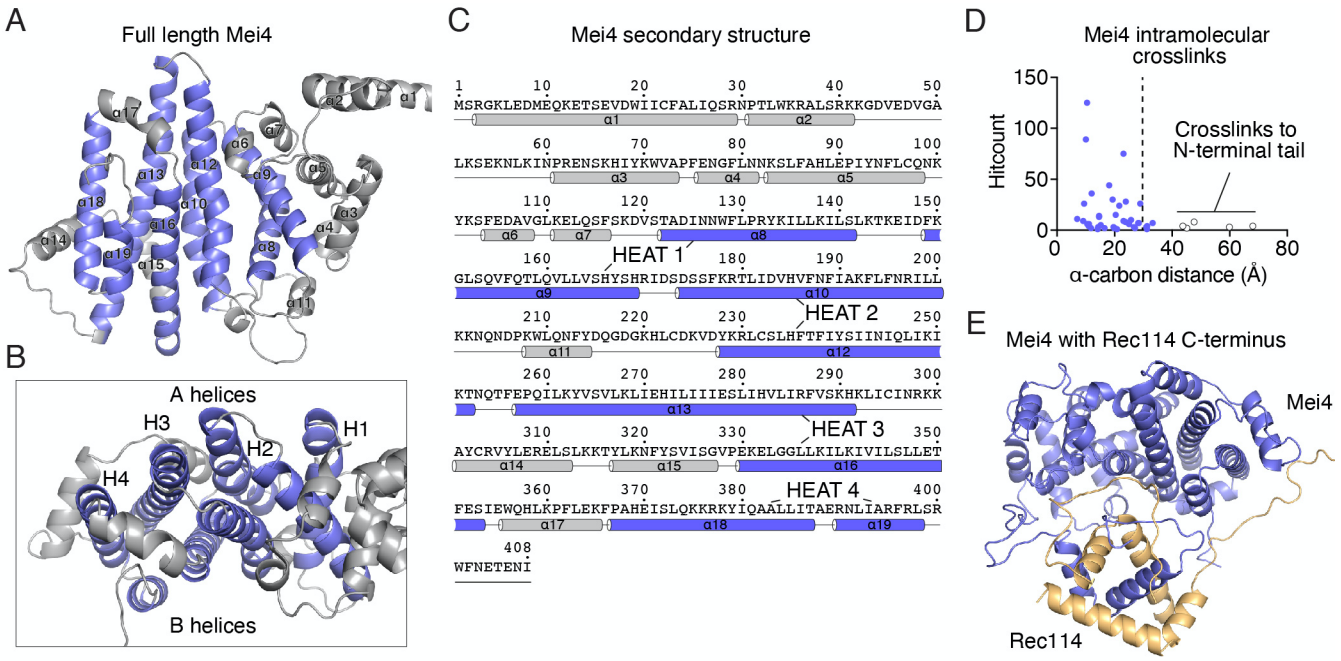


Figure 2: Structural model of full-length Mei4 with Rec114 C-terminus.
(A) AlphaFold predicted structure of full-length Mei4 (AF-P29467). Helices that make up the HEAT-repeat core are in blue, other helices are grey. (B) Mei4 has a concave side (top) made up of HEAT-repeat A helices and a convex side (bottom) made up of B helices. (C) Secondary structures of Mei4 based on AlphaFold model. Helices that constitute the HEAT repeats are annotated. (D) Analysis of Mei4 intramolecular crosslinks (690 crosslinks, 44 distinct pairs) (Claeys Bouuaert et al., 2021). The histogram shows the distances separating α-carbons of crosslinked lysines. The crosslinkable limit (dashed line) is 27.4 Å. Model-clashing crosslinks (open circles) involve the flexible N-terminal tail of Mei4. (E) Lowest energy structure of the dimeric Rec114 C-terminus (orange) in complex with full-length Mei4 (blue).

Å over 172 residues). Typical HEAT repeats consist of helix A-turn-helix B motifs arranged in tandem with a ~15° angle between repeats and form an α-solenoid (Andrade et al., 2001). Indeed, the predicted Mei4 structure consists of a core of four pairs of anti-parallel helices stacked against each other yielding a convex surface made of A helices and a concave surface made of B helices, similar to other HEAT-repeat structures (**Figure 2B**). However, these are not assembled from canonical helix-turn-helix motifs, since two out of four pairs of antiparallel helices are interrupted by additional helices (**Figure 2C**). Hence, Mei4 has an atypical HEAT-repeat structure.

We used our published crosslinking coupled to mass spectrometry (XL-MS) data of Rec114–Mei4 complexes (Claeys Bouuaert et al., 2021) to test the predicted Mei4 structure. The dataset contains 690 intramolecular Mei4 crosslinks (44 distinct pairs). Most of the α-carbons of the crosslinked lysines were closer to each other than the crosslinkable limit of 27,4 Å (**Figure 2D**). The five pairs of crosslinked residues that were significantly farther apart involved the N-terminal Rec114-interacting tail, which is presumably flexible. Hence, intramolecular Mei4 crosslinks strongly support the predicted structure.

We sought to model the interaction of the Rec114 C-terminal domain with full-length Mei4. To do this, we used the Rec114–Mei4 XL-MS data as distance restraints to calculate the lowest-energy structure the complex (**Figure 2E**). This revealed a binding geometry where the Rec114 C-terminal dimer is docked on the concave side of Mei4, formed by HEAT-repeat B helices (α9, α12, α16, α19). The refined structure satisfies most of the abundant crosslinks (>10 hits) of the XL-MS data (**Supplementary Figure 3, Supplementary Table 2**).

Rec114–Mei4 heterotrimers have two duplex-DNA binding sites

We previously showed that Rec114–Mei4 binds duplex DNA substrates *in vitro* (Claeys Bouuaert et al., 2021). However, the substrate specificity of the complex had remained unexplored. We used a gel shift assay to investigate the binding of Rec114–Mei4 to different DNA structures, assembled with synthetic oligonucleotides. We found that Rec114–Mei4 binds preferentially to branched DNA substrates, compared to duplex DNA (**Figure 3A, B**). This was confirmed in a competition assay, where a DNA substrate that mimics a Holliday Junction (HJ) was a ~10-fold better competitor than duplex DNA (**Figure 3C**). In addition, binding to duplex DNA leads to well-shifts, while binding to the branched DNA substrates lead to complexes that can enter the gel. The preference of Rec114–Mei4 for branched DNA suggests the presence of multiple DNA-binding sites that collaborate to stabilize the nucleoprotein complexes. Indeed, multivalent interactions between Rec114–Mei4 and DNA likely underlie their DNA-driven condensation activity (Claeys Bouuaert et al., 2021).

To gain further insights into the DNA-binding properties of Rec114–Mei4, we examined whether the minimal trimeric complex also interacts preferentially with branched DNA substrates. Indeed, the minimal complex showed ~10-fold higher affinity for a HJ substrate compared to a duplex DNA in a competition experiment, similar to full-length Rec114–Mei4 (**Figure 3D**).

Alanine mutations of Rec114 residues R395, K396, K399 and R400 compromise the DNA-binding and condensation activities of Rec114–Mei4 *in vitro*, and abolish DSB formation *in vivo* (Claeys Bouuaert et al., 2021). These residues form two patches of positively-charged amino acids that face in opposite directions on the structural model, revealing two independent DNA-binding sites within the complex (**Figure 3E**). This geometrical arrangement indicates that the two binding sites could not be occupied simultaneously on the same B-form DNA substrate, which explains the preference of Rec114–Mei4 for branched structures.

Finally, we assessed whether Rec114–Mei4 complexes bind reconstituted nucleosomes *in vitro*. However, we found no evidence for preferential binding of the complex to nucleosomes, compared to a cognate 147 bp duplex DNA substrate (**Figure 3F**).

Structural model of Mer2 homotetramers

Mer2 has a central coiled-coil domain that tetramerizes, flanked by N- and C-terminal disordered regions (**Figure 4A**). We previously proposed that the coiled coil is assembled as pairs of parallel α-helices arranged in an antiparallel configuration (Claeys Bouuaert et al., 2021). To gain further insight, we used AlphaFold to predict the structure of a tetrameric Mer2 coiled-coil domain (residues 41-224). Surprisingly, AlphaFold generated with high-confidence a parallel tetrameric model (**Figure 4B, Supplementary Figure 4A-C**). This parallel structure was largely consistent with our published XL-MS data of Mer2, except for a few long-range crosslinks that cannot be accommodated by the model (**Figure 4C**) (Claeys Bouuaert et al., 2021).

To test this structural arrangement, we purified tagged and untagged truncations of Mer2, and determined their molar masses (MW) by size exclusion chromatography followed by multi-angle light scattering (SEC-MALS) (**Figure 4D**). The Mer2 coiled coil and the C-terminal half alone (residues 138-224) form tetramers, which is not compatible with a parallel-antiparallel model. However, further truncating the C-terminal part of the coiled coil (residues 161-224) gives an apparent molecular weight that corresponds to a trimer, suggesting that the tetramer is unstable and partially dissociates during the chromatography. In addition, the N-terminal half of the coiled coil (residues 41-136) yields a molecular weight that corresponds to a dimer, while further truncating the N-

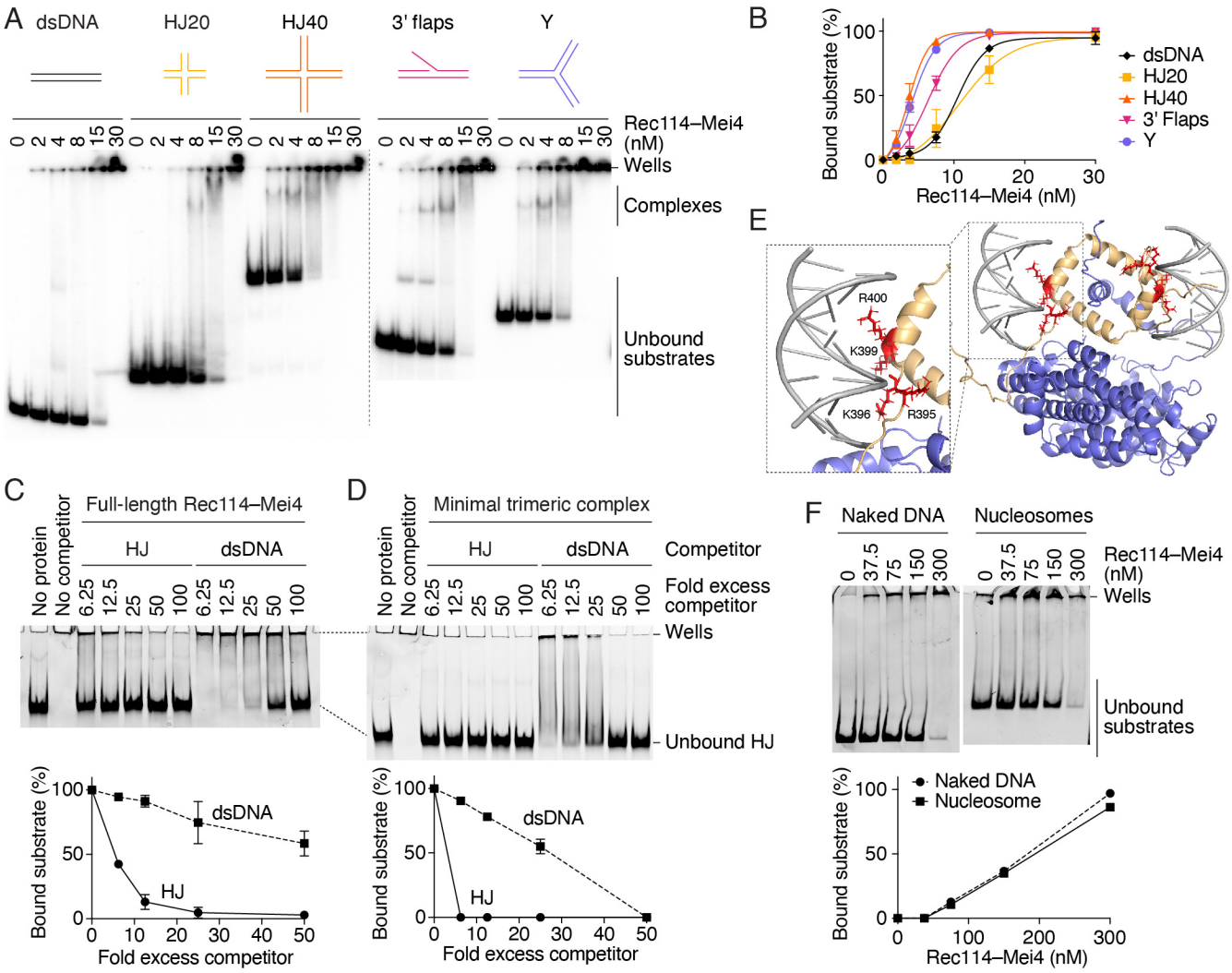
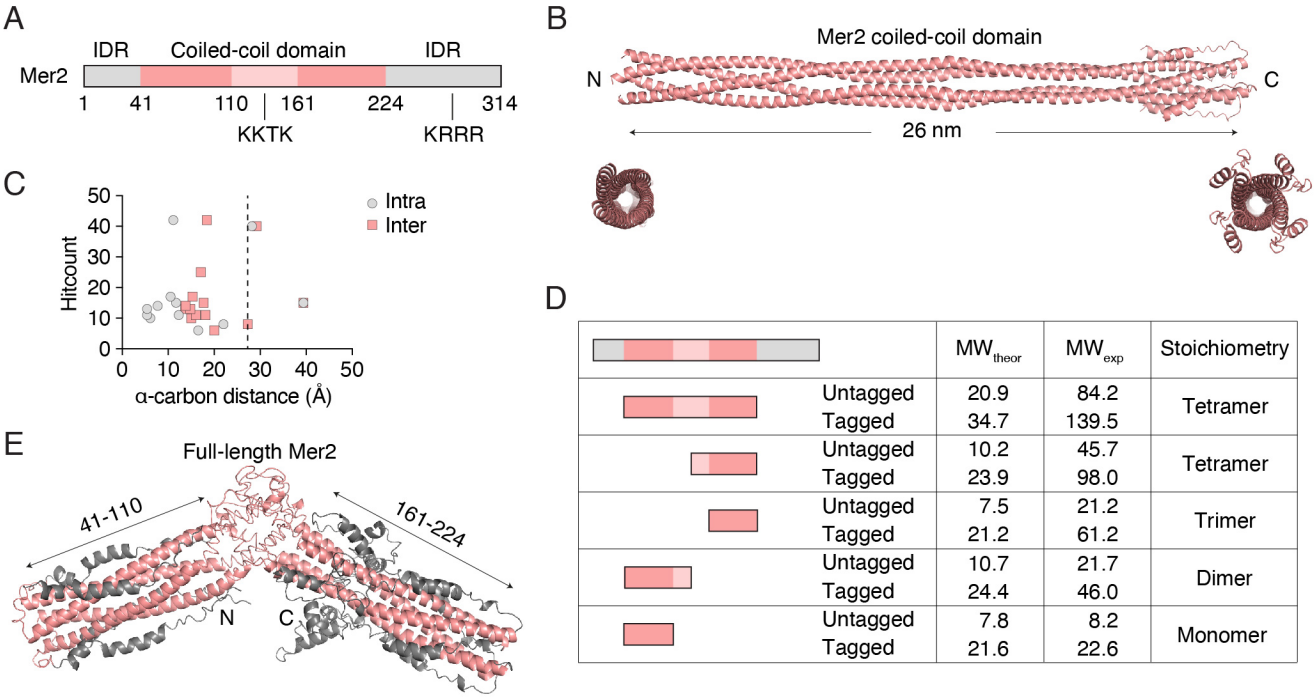


Figure 3: Rec114–Mei4 has two distinct DNA-binding sites.
(A) Gel-shift assay of tagged Rec114–Mei4 binding to ^{32}P-labeled 80-bp duplex DNA (dsDNA), Holliday Junctions with 20-bp (HJ20) or 40-bp (HJ40) arms, substrates with a 40-nt 3' single-stranded flap (3' flaps), or branched structures with three 40-bp arms (Y). Here and elsewhere, concentrations refer to a 2:1 Rec114–Mei4 heterotrimer. (B) Quantification of the gel-shift assay in panel A. Error bars are ranges from two independent experiments. Lines are sigmoidal curves fit to the data. The apparent affinities of Rec114–Mei4 for the DNA substrates are: 10.5 ± 1.4 nM (dsDNA, mean and range), 10.8 ± 3.0 nM (HJ20), 3.5 ± 0.7 nM (HJ40), 6.3 ± 1.0 nM (3' flaps), 4.1 ± 0.4 nM (Y). (C, D) Competition assays of full-length Rec114–Mei4 (180 nM) (C) and minimal trimeric Rec114–Mei4 (200 nM) (D) complexes binding to a fluorescent HJ substrate (10 nM) in the presence of unlabeled dsDNA or HJ40 substrates. Error bars are ranges from two independent experiments. (E) Model of Rec114–Mei4 complex bound to two DNA duplexes. The insert shows the position of Rec114 residues R395, K396, K399 and R400, previously implicated in DNA binding (Claeys Bouuaert et al., 2021). (F) Gel-shift assay of Rec114–Mei4 binding to fluorescently-labeled nucleosomes, compared to a cognate naked DNA substrate.



		MW_{theor}	MW_{exp}	Stoichiometry
	Untagged	20.9	84.2	Tetramer
	Tagged	34.7	139.5	
	Untagged	10.2	45.7	Tetramer
	Tagged	23.9	98.0	
	Untagged	7.5	21.2	Trimer
	Tagged	21.2	61.2	
	Untagged	10.7	21.7	Dimer
	Tagged	24.4	46.0	
	Untagged	7.8	8.2	Monomer
	Tagged	21.6	22.6	

Figure 4: Structural models of Mer2 homotetramers.
(A) Domain organization of Mer2. The positions of sequence motifs analyzed in this work are shown. IDR, intrinsically-disordered region. The central part of the coiled coil (light pink) may be unstable. (B) AlphaFold predicted structure of the tetrameric Mer2 coiled-coil domain. (C) XL-MS analysis of Mer2 crosslinks that map to the coiled-coil domain (Claeys Bouuaert et al., 2021). Distances separating α-carbons of crosslinked lysines are shown based on whether the crosslinks were considered as intramolecular or intermolecular. The crosslinkable limit (dashed line) is 27.4 Å. Only crosslinks that were found >5 times were counted, representing 240 individual crosslinks (14 distinct pairs). (D) SEC-MALS analysis of untagged and HisSUMO-tagged Mer2 truncations. (E) AlphaFold prediction of full-length Mer2.

terminal yields a monomer, indicating that the packaging of the coiled-coil N-terminus is less stable than that of the C-terminus.

Analysis of the predicted structure using Twister (Strelkov and Burkhard, 2002) identified about 20 heptad repeats interrupted by six stutters (insertion of 4 amino acids) and one skip (+1) located between residues 80 and 168 (**Supplementary Figure 9A**). The AlphaFold model predicts geometrical distortions of the coiled coil in the vicinity of the stutters, which are mostly compensated by local unwinding of the coiled coil, with a local shift to a right-handed geometry. In addition, the model shows a wide coiled-coil radius around residues 120-150, which is indicative of sub-optimal hydrophobic packaging of the coiled coil, resulting in a less stable structure (**Supplementary Figure 9B**).

The center of the coiled coil contains a skip at residue 138 (**Supplementary Figure 9A**). The confidence score of the model is lowest around this sequence (**Supplementary Figure 4C**). Interestingly, AlphaFold modeling of a full-length Mer2 tetramer predicted that the coiled-coil domain is separated in two, with a disordered region spanning residues ~110 to 160 (**Figure 4E, Supplementary Figure 4D-F**). This raises the possibility that the center of the coiled coil is flexible, leading to different co-existing configurations of tetrameric complexes. The flexibility of the coiled coil could account for the presence of long-range crosslinks in the Mer2 XL-MS data (**Figure 4C**) (Claeys Bouuaert et al., 2021).

Mer2 engages in multivalent protein-DNA interactions

Mer2 was previously shown to bind double-stranded DNA *in vitro*, but binding preferences for distinct DNA structures had not been investigated (Claeys Bouuaert et al., 2021; Rousová et al., 2021; Tsai et al., 2020). Similar to Rec114–Mei4, we found that Mer2 binds preferentially to branched substrates (**Figure 5A, B**). This was confirmed in a competition assay, which showed ~20-fold increased affinity for a HJ substrate compared to duplex DNA (**Figure 5C**).

The Mer2 coiled-coil domain, which is necessary and sufficient for DNA binding (Claeys Bouuaert et al., 2021), also binds preferentially to branched DNA substrates (**Figure 5D**). This indicates that the coiled-coil domain engages in multivalent protein-DNA interactions. A KKTK motif located at the center of the coiled coil (**Figure 4A, 5E**) was previously implicated in DNA binding (Tsai et al., 2020). However, in our hands, mutating the three lysines of this motif to alanine did not affect the DNA-binding activity of Mer2 (**Figure 5F**). Hence, the DNA-binding residues within the Mer2 coiled-coil domain remain unknown. On the other hand, we previously showed that the KRRR motif located within the C-terminal intrinsically-disordered region is important for DNA binding, condensation, and DSB formation (Claeys Bouuaert et al., 2021). As expected, the Mer2-KRRR mutant also shows decreased binding affinity for a branched substrate (**Figure 5F**).

Finally, in agreement with a previous report (Rousová et al., 2021), we found that Mer2 binds preferentially to nucleosomes compared to a 147 bp cognate DNA substrate (**Figure 5G**), in contrast to Rec114–Mei4 complexes. Since Mer2 probably binds DNA exposed at the surface of the nucleosome (Rousová et al., 2021), we propose that nucleosomes present DNA duplexes with the appropriate geometry to allow multivalent interactions between Mer2 tetramers and DNA. However, this geometry is not favorable for multivalent interactions between Rec114–Mei4 complexes and DNA.

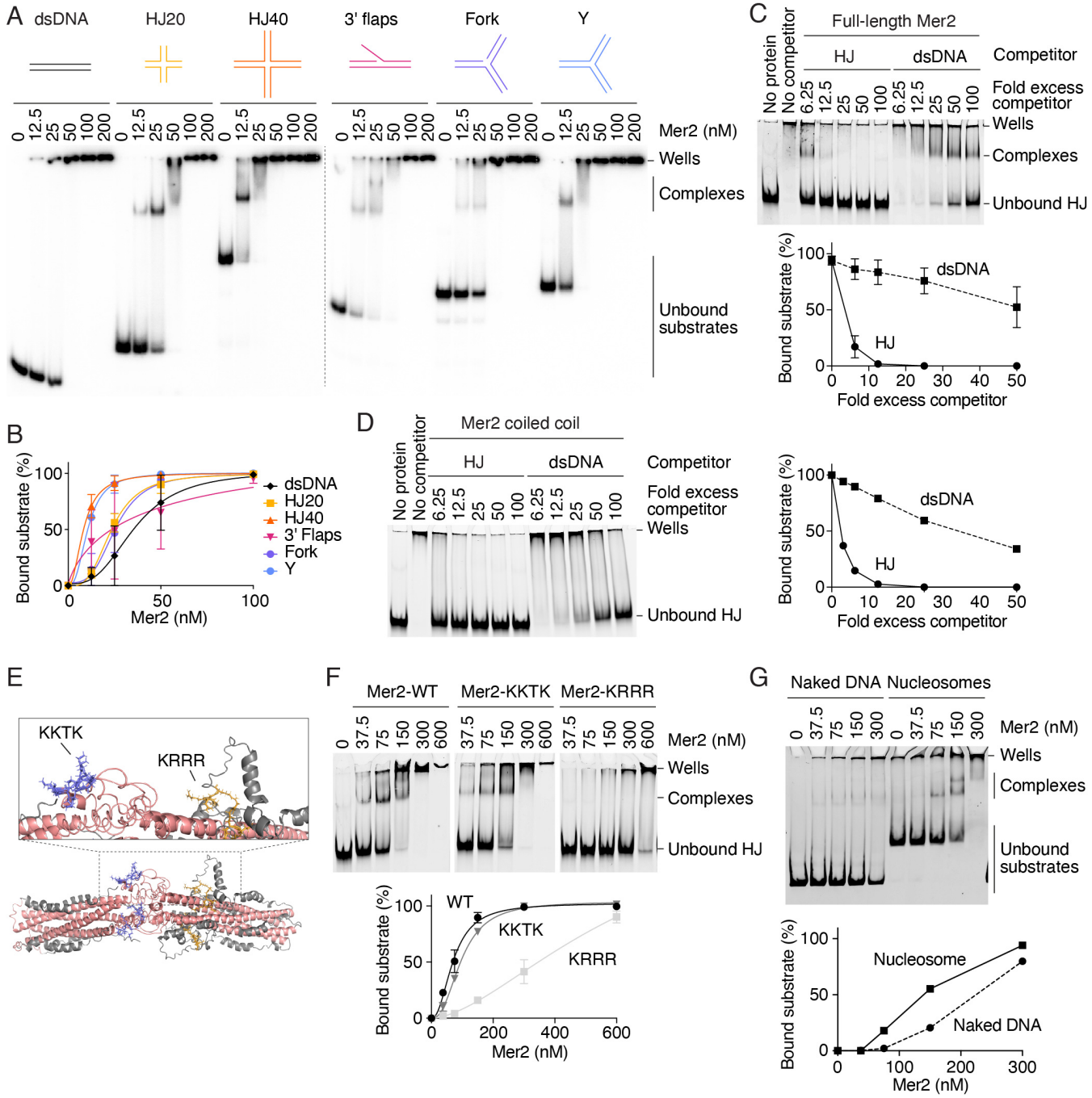


Figure 5: DNA-binding activities of Mer2.
(A) Gel-shift assay of Mer2 binding to ^{32}P-labeled 80-bp duplex DNA (dsDNA), Holliday Junctions with 20-bp (HJ20) or 40-bp (HJ40) arms, substrates with a 40-nt 3' single-stranded flap (3' flaps), or branched structures with three 40-bp arms without (Y) and with a central nick (Fork). Concentrations refer to Mer2 monomers. (B) Quantification of the gel-shift assay in panel A. Error bars are ranges from two independent experiments. Lines are sigmoidal curves fit to the data. The apparent affinities of Mer2 for the DNA substrates are: 36.8 ± 13.3 nM (dsDNA, mean and range), 23.9 ± 2.9 nM (HJ20), 9.2 ± 1.0 nM (HJ40), 34.4 ± 25.2 nM (3' flaps), 24.4 ± 5.5 nM (Fork), 10.6 ± 1.2 nM (Y). (C, D) Competition assays of full-length Mer2 (200 nM) (C) or the coiled-coil domain (750 nM) (D) binding to a fluorescent HJ substrate (10 nM) in the presence of unlabeled dsDNA or HJ40 substrates. Error bars are ranges from two independent experiments. Error bars in panel D are too small to be visible. (E) Positions of the KKTK and KRRR motifs on the AlphaFold model of full-length Mer2. (F) Gel-shift assay of wild-type and mutant Mer2 binding to a fluorescent HJ substrate. In the mutants, the lysines and arginines of the respective motifs are mutated to alanine. Error bars are ranges from two independent experiments. (G) Gel-shift assay of Mer2 binding to fluorescently-labeled nucleosomes or 147 bp cognate DNA.

Conservation of Rec114–Mei4 structure and DNA-binding properties

The RMM proteins are highly diverged across the eukaryotic kingdom, and pair-wise comparisons between homologs of distantly related species typically show sequence identities well below 10% (Kumar et al., 2010; Stanzione et al., 2016; Tesse et al., 2017; Wang et al., 2019). To gain insights into their structural conservation, we used AlphaFold to model the architecture of Rec114 and Mei4 orthologs from *M. musculus* (REC114, MEI4), *S. pombe* (Rec7, Rec24), *A. thaliana* (PHS1, PRD2) and *Z. mays* (PHS1, MPS1) (Bonfils et al., 2011; De Muyt et al., 2009; Hinman et al., 2021; Kumar et al., 2010; Molnar et al., 2001; Pawlowski et al., 2004; Rosu et al., 2013; Stamper et al., 2013; Steiner et al., 2010; Vrielynck et al., 2021).

As expected, Rec114 orthologs showed a long central intrinsically-disordered region flanked by an N-terminal PH domain, and 2-4 C-terminal α-helices (**Supplementary Figure 5A**). While the PH-fold was well defined for the *M. musculus*, *S. pombe*, and *Z. mays* homologs, the PH-fold of *A. thaliana* PHS1 was incomplete. Interestingly, PHS1 was recently shown to be dispensable for meiotic DSB formation (Vrielynck et al., 2021), in contrast to other Rec114 orthologs, including maize PHS1 (Kumar et al., 2018; Molnar et al., 2001; Pawlowski et al., 2004; Rosu et al., 2013; Stamper et al., 2013).

AlphaFold models of Mei4 orthologs revealed HEAT-repeat domains flanked by 3-5 flexibly-connected N-terminal helices (**Supplementary Figures 5B & 6**). For each ortholog, four well-defined HEAT repeats could be identified (**Supplementary Figure 6**). The helix A-turn-helix B motifs were more canonical than for yeast Mei4, and consecutive HEAT repeats were arranged with a 5-15° angle, creating slightly curved structures (**Supplementary Figure 5B**).

Next, we modeled the interaction domains of Rec114 and Mei4 orthologs. In all cases, AlphaFold predicted similar heterotrimeric complexes with an N-terminal α-helix of the Mei4 ortholog inserted within a dimeric ring composed of C-terminal α-helices of Rec114

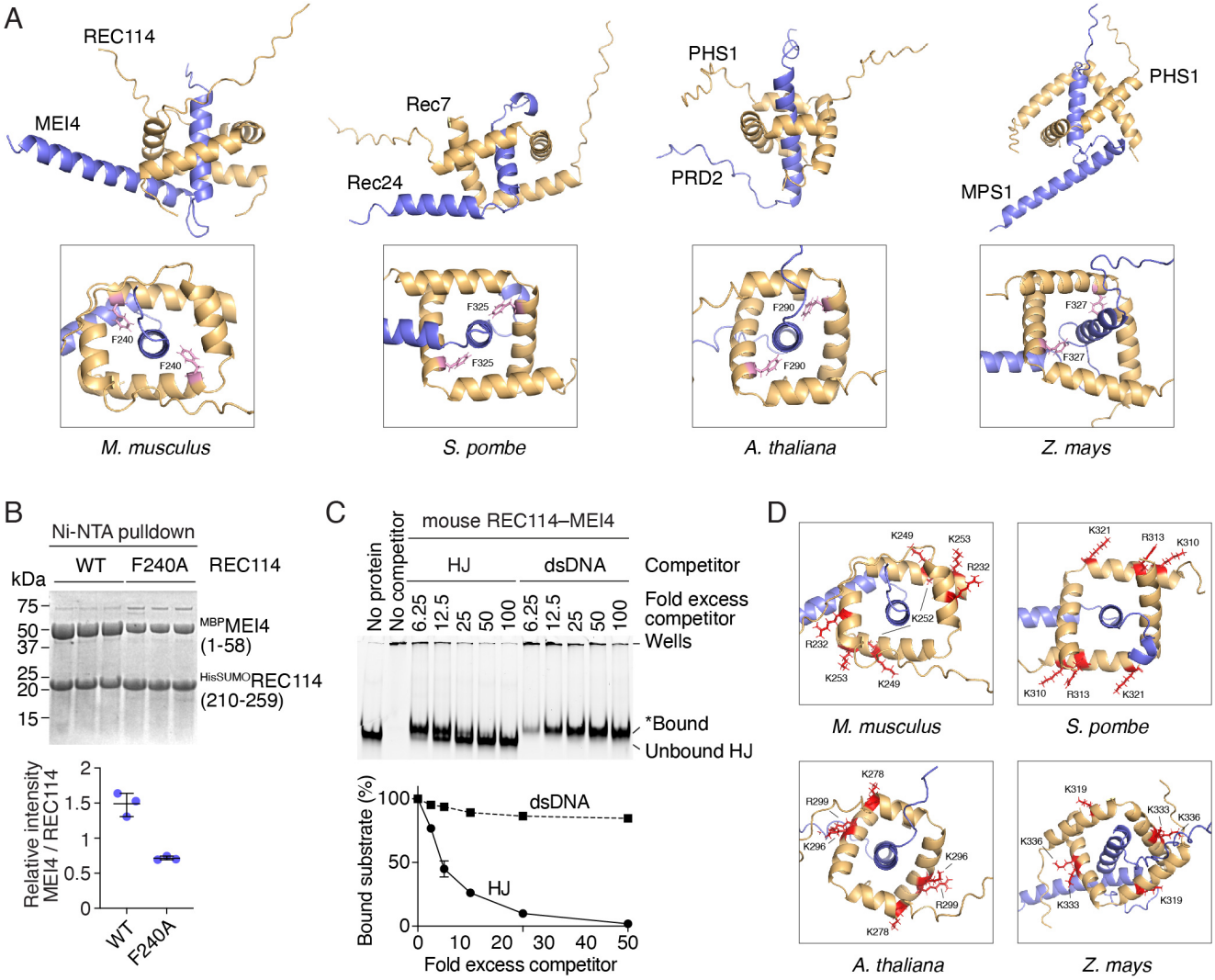


Figure 6: Conservation of Rec114–Mei4 structure and DNA-binding properties.
(A) AlphaFold models of minimal trimeric complexes for Rec114–Mei4 orthologs in *M. musculus* (REC114–MEI4), *S. pombe* (Rec7–Rec24), *A. thaliana* (PHS1–PRD2), and *Z. mays* (PHS1–MPS1). The position of the conserved phenylalanine at the interface between Rec114 and Mei4 orthologs is shown. (B) Pulldown analysis of the effect of *M. musculus* REC114-F240A mutation on the interaction with MEI4. Error bars are ranges from three replicates. (C) Competition assay of tagged mouse REC114–MEI4 complex (1000 nM) binding to a fluorescent HJ substrate (10 nM) in the presence of unlabeled dsDNA or HJ substrates. The band labeled '*Bound' migrates close to the position of the unbound substrate. It is likely due to the rapid dissociation of REC114–MEI4 from the substrate at the start of the electrophoresis. Error bars are ranges from two independent experiments (most are too small to be visible). (D) Putative DNA-binding sites within *M. musculus*, *S. pombe*, *A. thaliana* and *Z. mays* complexes.

orthologs (**Figure 6A, Supplementary Figure 7**). The *S. pombe* Rec7-F325 residue, previously shown to be important for the interaction with Rec24 (Steiner et al., 2010) points directly towards Rec24 in the model, similarly to *S. cerevisiae* Rec114-F411. In addition, the equivalent phenylalanine residue in *M. musculus* (F240), *A. thaliana* (F290), and *Z. mays* (F327) (Kumar et al., 2010), occupy a similar position inside the dimeric ring (**Figure 6A**, bottom).

To test these models, we purified a complex containing the C-terminus of mouse REC114 (residues 210-259) bound to the N-terminus of MEI4 (residues 1-58). As expected, mutating REC114 residue F240 to alanine compromised the interaction with MEI4 (**Figure 6B**).

To address whether the DNA-binding properties of Rec114–Mei4 are conserved, we performed gel-shift analyses with full-length mouse REC114–MEI4 in the presence of a HJ substrate. Similar to the yeast complex, REC114–MEI4 binds to a HJ substrate with >20-fold higher affinity than to duplex DNA (**Figure 6C**), suggesting that the complex contains multiple DNA-binding sites. Indeed, the C-terminus of Rec114 orthologs of *M. musculus*, *S. pombe*, *A. thaliana*, and *Z. mays* all have basic residues that form two positively-charged patches that point in opposite directions on the structural models (**Figure 6D**), suggesting a conserved mechanism for multivalent protein-DNA interactions.

Conservation of Mer2 structure and DNA-binding properties

To investigate the structural conservation of Mer2 orthologs, we generated AlphaFold predictions of tetrameric coiled-coil domains of *M. musculus* IHO1, *A. thaliana* PRD3, *S. pombe* Rec15, *S. macrospora* ASY2, and *Z. mays* PAIR1 (De Muyt et al., 2009; Miyoshi et al., 2012; Nonomura et al., 2004; Stanzione et al., 2016; Tesse et al., 2017). AlphaFold modeling yielded high-confidence predictions of the different coiled coils, all arranged in a parallel configuration (**Figure 7A, Supplementary Figure 8**).

To test these models, we purified the coiled-coil domain of IHO1 (residues 109-267). SEC-MALS analyses of HisSUMO-tagged and untagged complexes revealed experimental molecular masses of 113 kDa

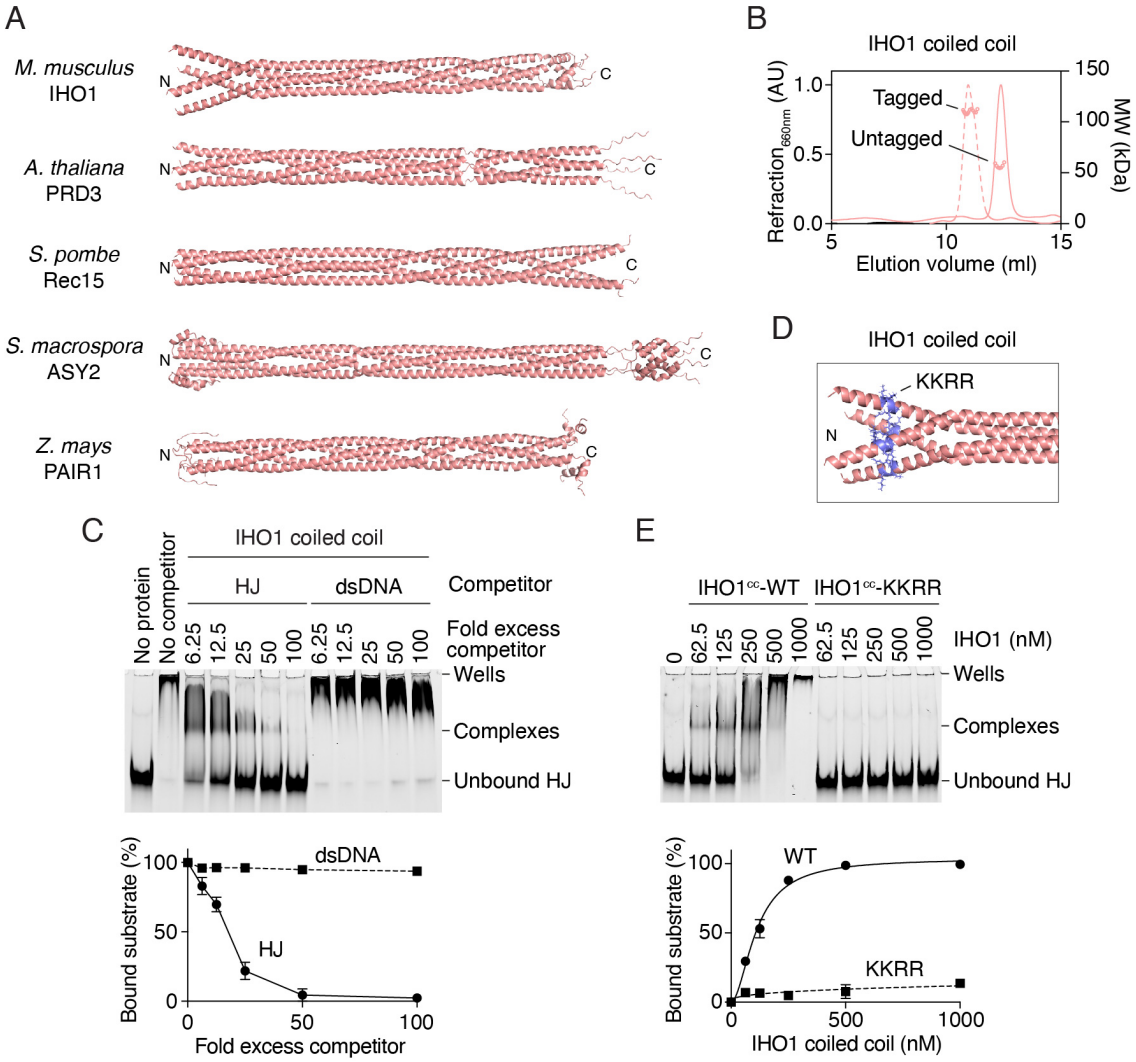


Figure 7: Conservation of Mer2 structure and DNA-binding properties.
(A) AlphaFold models of homotetrameric coiled-coil domains of *M. musculus* IHO1 (residues 109-267, length 2.15 nm), *A. thaliana* PRD3 (residues 120-270, length 2.25 nm), *S. pombe* Rec15 (residues 1-160, length 2.3 nm), *S. macrospora* ASY2 (residues 55-275, length 2.85 nm), and *Z. mays* PAIR1 (residues 140-310, length 2.15 nm). (B) SEC-MALS analysis of tagged and untagged IHO1 coiled-coil domains. The traces show differential refraction at 660 nm (arbitrary units) and circles are molar mass measurements across the peak. (C) Competition assay of IHO1 coiled-coil domain (600 nM) binding to a fluorescent HJ substrate (10 nM) in the presence of unlabeled dsDNA or HJ substrates. Error bars are ranges from two independent experiments. (D) Position of the KKRR motif on the structural model of the IHO1 coiled-coil domain. (E) Gel-shift assay of wild type and mutant IHO1 coiled-coil domain binding to a fluorescently-labeled HJ substrate. In the mutant, the KKRR residues are mutated to alanine.

and 64 kDa, respectively, consistent with homotetramers (expected sizes 127 kDa and 72 kDa, respectively) (**Figure 7B**).

Similar to *S. cerevisiae* Mer2, Twister analyses (Strelkov and Burkhard, 2002) of the predicted coiled-coil structures identified heptad repeats and revealed the presence of discontinuities (stutters, skips and stammers), which are locally compensated by distortions of the coiled coils, primarily through increased radius of the coiled coils and local switches to right-handed helices (**Supplementary Figures 9-11**). Hence the coiled-coil structure of Mer2 orthologs is largely conserved.

The coiled-coil domain of IHO1 was sufficient to bind DNA and showed a >20-fold binding preference for a HJ substrate compared to duplex DNA in a competition assay (**Figure 7C**). This indicates that the tetrameric IHO1 coiled-coil domain mediates multivalent protein-DNA interactions. On the structural model, the helices dissociate at the N-terminus over a ~25 amino-acid sequence that contains a KKRR motif (**Figure 7D**). Mutating these residues to alanine abolished the DNA-binding activity of the coiled coil (**Figure 7E**). Hence, IHO1 and Mer2 share DNA-binding properties, although the position of the DNA-interacting residues may differ.

Discussion

We presented structural models of *S. cerevisiae* Rec114–Mei4 and Mer2 complexes, supported by NMR, XL-MS, and mutagenesis, and showed that their architecture is conserved in higher eukaryotes. In addition, we showed that Rec114–Mei4 and Mer2 complexes and their orthologs bind preferentially to branched DNA structures, which results from the combined action of multiple DNA-binding sites within the complexes. These multivalent interactions are likely key to their condensation activity that drives the assembly of the DSB machinery.

Structural conservation of Rec114–Mei4 and Mer2 complexes

Rec114 has an N-terminal PH domain, followed by a ~250 amino-acid intrinsically-disordered region. The C-terminal ~30 amino acids form two α-helices that homodimerize to form a ring in which the N-terminus of Mei4 is inserted, yielding an asymmetric 2:1 complex. Mei4 presents an atypical HEAT-repeat like fold made up of four HEAT repeats that create a curved structure, with the C-terminal Rec114 dimer lodged along the concave side. The Rec114 PH domain, the Mei4 HEAT repeat structure, and the 2:1 Rec114–Mei4 interaction domain are generally conserved in *S. pombe*, *M. musculus*, *A. thaliana*, and *Z. mays*. In addition, AlphaFold modeling revealed a similar trimeric structure for the *C. elegans* homologs (DSB-1, DSB-2 and DSB-3) (Guo et al., 2022).

While the yeast Mei4 structure presents non-canonical HEAT repeats, the four HEAT repeats that compose the core structure of Mei4 orthologs displayed more typical helix A-turn-helix B motifs. The *Z. mays* Mei4 homolog, previously identified as Multipolar Spindle 1 (MPS1) (Kumar et al., 2010), also presents a similar HEAT-repeat architecture and N-terminal Rec114 (PHS1)-interacting helices, although, to our knowledge, its meiotic function has not been established.

In contrast to most Rec114 orthologs, AlphaFold modeling of *A. thaliana* PHS1 suggested that its PH-domain is incomplete. This is surprising because PHS1 contains signature sequence motifs (SSMs) within its N-terminal domain that are shared with other Rec114 homologs (Kumar et al., 2010), and these SSMs constitute the core secondary structures of the PH fold (Boekhout et al., 2019; Kumar et al., 2018). Interestingly, *A. thaliana* PHS1 was recently shown to be dispensable for meiotic DSB formation (Vrielynck et al., 2021). Nevertheless, this does not seem to be shared with other plant species, since the predicted maize PHS1 structure shows a well-folded PH-domain, and maize PHS1 is required for meiotic DSB formation (Pawlowski et al., 2004).

Mer2 forms a homotetramer with a ~200 amino-acid central coiled coil that folds as a ~25 nm parallel α-helical bundle. This parallel tetrameric configuration is conserved in *S. pombe*, *S. macrospora*, *M. musculus*, *Z. mays*, and *A. thaliana*. We previously proposed a parallel-antiparallel arrangement for the Mer2 tetrameric coiled coil (Claeys Bouuaert et al., 2021). This was based on two lines of evidence. First, XL-MS experiments revealed long-range crosslinks within the coiled-coil domain, which cannot easily be accounted for assuming a parallel arrangement of four uninterrupted helices. Second, an engineered single-chain dimer with two copies of the coiled coil connected with a short flexible linker behaved as a tetramer in SEC-MALS analyses. Since the linker was too short to allow parallel folding of the coiled coil, we concluded that Mer2 tetramers most likely consisted of two pairs of parallel α-helices arranged in an antiparallel configuration. However, both observations can be explained by a parallel homotetrameric structure, provided that the coiled coil is flexible. Indeed, the Alphafold model of full-length Mer2 suggested that the center of the coiled coil may be unstructured. This is consistent with the presence of multiple stutters and a skip that span the central region of the heptad repeats and likely destabilize the coiled coil.

Multivalent protein-DNA interactions drive RMM condensation

DNA-dependent condensation by Rec114–Mei4 and Mer2 require protein-DNA interactions, protein-protein interactions, and multivalency. Theoretically, multivalency may arise from protein-protein interactions, protein-DNA interactions, or both. We previously identified DNA-binding residues within Rec114–Mei4 and Mer2 complexes (Claeys Bouuaert et al., 2021). However, insights into the nature of the multivalency remained limited. The preference of Rec114–Mei4 and Mer2 for branched DNA structures and the formation of defined stoichiometric complexes with these substrates indicates that they mediate multivalent protein-DNA interactions.

This is explained structurally for Rec114–Mei4 since the AlphaFold model of the heterotrimeric protein-interaction domain, which is also responsible for DNA binding and the preference for branched structures, shows two duplex DNA binding sites that point in opposite directions. Since those cannot be occupied simultaneously by a single DNA duplex (under the persistence length of ~150 bp), the complex binds more stably to a branched substrate presenting multiple flexibly-connected duplexes, which can be contacted simultaneously.

In addition, the AlphaFold structures of Rec114–Mei4 orthologs from *M. musculus*, *S. pombe*, *A. thaliana* and *Z. mays*, also present two independent putative DNA-binding sites, suggesting a conserved mechanism for multivalent binding to DNA duplexes. This conclusion was confirmed with the mouse REC114-MEI4 complex that showed preference for branched DNA.

Similarly, Mer2 also binds preferentially to branched DNA, and forms well-defined complexes, particularly with four-way (HJ) structures. Hence, multivalency underlying Mer2 condensation also likely arises from independent DNA-binding sites within Mer2 tetramers. In this case, however, the structural details are elusive. The Mer2 coiled coil is necessary and sufficient for DNA binding, but residues located within the C-terminal flexible domain are also important for the interaction with DNA. The KKTK motif within the coiled-coil domain, previously implicated in Mer2 DNA binding (Tsai et al., 2020), was not important for DNA binding in our assays.

The preference for branched DNA substrates was conserved in IHO1 and required only the coiled-coil domain, further indicating that the assembly of the DSB machinery through the combined action of multivalent protein-DNA interactions is conserved across eukaryotes. In addition, we identified a KKRR motif that is essential for the DNA-binding activity of the coiled coil. Whether other residues outside of the coiled coil also contribute to IHO1 DNA binding is unclear.

Impact of chromatin on RMM condensation

Despite fundamentally different architectures, Rec114–Mei4 and Mer2 complexes share many biochemical properties, including binding preferences for branched DNA substrates and the propensity to undergo

DNA-driven condensation. However, the complexes differ in one important way: Mer2 binds efficiently to nucleosomes *in vitro*, while Rec114–Mei4 does not. This suggests that chromatin impacts differently Rec114–Mei4 and Mer2 condensation. Nevertheless, it is likely that Mer2 does not directly contact nucleosomes, and that instead nucleosomes present DNA duplexes with a favorable geometry to allow multivalent Mer2 interactions, but that this configuration is refractory to Rec114–Mei4 DNA binding.

In vivo, Mer2 can bind chromatin in the absence of Rec114 and Mei4, but chromatin binding by Rec114 requires Mei4 and Mer2 (Panizza et al., 2011). In addition, Rec114 chromatin association depends on CDK-dependent Mer2 phosphorylation at S30, which promotes protein-protein interactions between Mer2 and Rec114 (Henderson et al., 2006; Panizza et al., 2011). This suggests a model where chromatin inhibits condensation of Rec114–Mei4, but not Mer2, and this inhibition is overcome by phosphorylation-dependent interactions between Mer2 and Rec114. However, whether this is indeed the case remains to be established.

In summary, our results yield insights into the structure of Rec114–Mei4 and Mer2 complexes and the multivalent protein-DNA interactions that drive their DNA-dependent condensation activity, and reveal that these structural and functional properties are conserved throughout eukaryotes.

Materials and Methods

Preparation of expression vectors

The sequences of the oligos are listed in **Supplementary Table 3**. Plasmids are listed in **Supplementary Table 4**. The vectors used to express ^{HisFlag}Rec114 (pCCB789), ^{MBP}Mei4 (pCCB791) from Sf9 cells and ^{HisSUMO}Mer2 (pCCB750) from *E. coli* were previously described (Claeys Bouuaert et al., 2021).

The Rec114 C-terminal domain and Mei4 N-terminal domain were amplified from pCCB649 and pCCB652 using primers dd015 and dd016, and dd017 and dd018, and cloned into pCCB206 vector by Gibson assembly to yield pDD003 and pDD004 respectively. The Rec114 C-terminal domain was amplified from pDD003 using dd027 and dd028 and cloned into pETDuet-1 vector by Gibson assembly to yield pDD006. N-terminal ^{HisSUMO}Mei4 was amplified from pDD004 using primers dd025 and dd026 and cloned into pDD006 vector by Gibson assembly to yield pDD009.

Expression vectors for minimal trimeric Rec114–Mei4 mutant complexes were obtained using pCCB825 as a template by inverse PCR and self-ligation using dd084 and dd085, dd095 and dd114 to yield, respectively, pDD044 (^{HisSUMO}Rec114-K405A, Mei4-WT) and pDD045 (^{HisSUMO}Rec114-WT, Mei4-E16A/D18A). Plasmid pDD051(^{HisSUMO}Rec114-K405A/E419A, Mei4-WT) was obtained using pDD044 as a template and dd080 and dd118 as primers. Expression vectors for Mer2 truncations were obtained by inverse PCR and self-ligation using pCCB750 as a template to generate pCCB973 (^{HisSUMO}Mer2(41-314)), pCCB975 (^{HisSUMO}Mer2(161-314)) and pCCB978 (^{HisSUMO}Mer2(1-110)) using cb1346 and cb1497, cb1346 and cb1495, and cb1342 and cb1492 primers, respectively. Plasmids pCCB973, pCCB978 and pCCB975 were further used as templates to generate pCCB981 (^{HisSUMO}Mer2(41-224)), pCCB979 (^{HisSUMO}Mer2(41-110)) and pCCB980 (^{HisSUMO}Mer2(161-224)) using primers cb1342 and dd121, cb1346 and cb1497, and cb1342 and dd121, respectively. Plasmid pDD078 (^{HisSUMO}Mer2(138-224)) and pDD079 (^{HisSUMO}Mer2(41-136)) were obtained by using, respectively, cb1346 and dd146, or cb1342 and dd147 as primers and pCCB981 as a template. The expression vector for the ^{HisSUMO}Mer2-KKTK mutant (pDD015) was obtained by inverse PCR and self-ligation by using dd060 and dd067 as primers and pCCB750 as a template.

Expression vectors for *M. musculus* REC114, MEI4 and IHO1 were generated by PCR-amplification of mouse testes cDNA (using primers cb1315 and cb1316, cb1317 and cb1318, and cb1322 and cb1327, respectively), and cloned by Gibson assembly into vectors pFastBac1-Flag (REC114), pFastBac1-MBP (MEI4), or pSMT3 (IHO1) to yield pCCB805, pCCB806, and pCCB808. The coiled-coil domain of IHO1 was amplified from pCCB808 using primers cb1498 and cb1499 and cloned into pSMT3 by Gibson assembly to yield pCCB982. The expression vector for the ^{HisSUMO}IHO1(coiled-coil)-KKRR mutant (pDD081) was obtained by inverse PCR and self-ligation with dd150 and dd151 as primers and pCCB982 as a template.

Expression and purification of recombinant proteins

Viruses were produced using a Bac-to-Bac Baculovirus Expression System (Invitrogen) according to the manufacturer's instructions. We infected $2x10^9$ *Spodoptera frugiperda* Sf9 cells (Gibco, Thermo Fisher) with combinations of viruses at a multiplicity of infection (MOI) of 2.5 each. Expression of $^{HisFlag-TEV}$Rec114, $^{MBP-TEV}$Mei4 used viruses generated from pCCB789 and pCCB791, and mouse $^{HisFlag-TEV}$REC114, $^{MBP-TEV}$MEI4 used viruses generated from pCCB805 and pCCB806. After 72h infection, cells were collected, washed with phosphate buffer saline (PBS), frozen in dry ice and kept at -80 °C until use. All purification steps were carried out at 0-4 °C. Cell pellets were resuspended in 3 volumes of lysis buffer (50 mM HEPES-NaOH pH 7.5, 0.17 mM DTT, 65 mM imidazole, 4.6 µM leupeptin, 0.3 µM aprotinin, 3.3 µM antipain, 2.9 µM pepstatin, 3.3 µM chymostatin and 1 mM phenylmethanesulfonyl fluoride (PMSF)) and then pooled in a beaker and mixed slowly with a stir bar for 20 minutes. 10% of ice-cold glycerol and 335 mM NaCl were added to the cell lysate that was then centrifuged at 43,000 g for 30 min. The cleared extract was loaded onto 1 ml pre-equilibrated Ni-NTA resin (Thermo Scientific). The column was washed extensively with nickel buffer (25 mM HEPES-NaOH pH 7.5, 500 mM NaCl, 10% glycerol, 0.1 mM DTT, 20 mM imidazole, 0.1 mM PMSF). The tagged complexes were then eluted in nickel buffer containing 500 mM imidazole. The complexes were further purified on amylose resin (NEB). Fractions containing protein were pooled and diluted in 3 volumes of amylose buffer (25 mM HEPES-NaOH pH 7.5, 500 mM NaCl, 10% glycerol, 2 mM DTT, 5 mM EDTA). Next, the complexes were bound to 1 ml of the amylose resin in a disposable chromatography column (Thermo Scientific) and the resin was washed extensively. Complexes were eluted from amylose resin with buffer containing 10 mM maltose. Fractions containing protein were concentrated in 50-kDa cutoff Amicon centrifugal filters (Millipore). Aliquots were flash frozen in liquid nitrogen and stored at −80 °C.

For expression of recombinant proteins in *E. coli*, expression vectors were transformed in BL21 DE3 cells and plated on LB plates containing the appropriate antibiotic. For the expression of IHO1, the expression vector was transformed in *E. coli* C41 cells. Cells were then cultured in liquid medium at 37 °C to an optical density (OD_{600}) of 0.6. Expression was carried out at 30 °C for 3 hours with 1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG). Cells were resuspended in nickel buffer (25 mM HEPES-NaOH pH 7.5, 500 mM NaCl, 10% glycerol, 0.1 mM DTT, 20 mM imidazole, 0.1 mM PMSF), frozen dropwise in liquid nitrogen and kept at −80 °C until use. All the purification steps were carried out at 0–4 °C. Cells were lysed by sonication and centrifuged at 43,000 g for 30 minutes. The cleared extract was loaded onto 1 ml pre-equilibrated Ni-NTA resin (Thermo Scientific). The column was washed extensively with nickel buffer then eluted in buffer containing 500 mM imidazole. The 6His–SUMO tag was cleaved with Ulp1 during overnight dialysis in nickel buffer. The sample was then loaded on a second nickel column to remove 6His–SUMO and Ulp1. The flow-through was then loaded on a Superose 6 Increase 10/300 GL column preequilibrated with gel filtration buffer (25 mM HEPES-NaOH pH 7.5, 300 mM NaCl, 10% glycerol, 40 mM imidazole, 1 mM DTT, 5 mM EDTA). After gel filtration, fractions containing protein were concentrated in 10-kDa cutoff Amicon centrifugal filters (Millipore). Aliquots were flash frozen in liquid nitrogen and stored at −80 °C.

For the production of the doubly-labeled U-[^{13}C,^{15}N] protein, the minimal medium contained M9 salts (6.8 g/L Na_2HPO_4, 3 g/L KH_2PO_4, and 1 g/L NaCl), 2 mM $MgSO_4$, 0.2 mM $CaCl_2$, trace elements (60 mg/L $FeSO_4 \cdot 7H_2O$, 12 mg/L $MnCl_2 \cdot 4H_2O$, 8 mg/L $CoCl_2 \cdot 6H_2O$, 7 mg/L $ZnSO_4 \cdot 7H_2O$, 3 mg/L $CuCl_2 \cdot 2H_2O$, 0.2 mg/L H_3BO_3, and 50 mg/L EDTA), BME vitamin mix (Sigma), and 1 g/L $^{15}NH_4Cl$ and 2 g/L [$^{13}C_6$]glucose (CortecNet) as the sole nitrogen and carbon sources, respectively. Expression was carried out at 30 °C for 6 h with 1 mM IPTG. After affinity chromatography on Ni-NTA resin, the sample was loaded to Superdex 75 Increase 10/300 GL column. Proteins were eluted in a buffer containing 20 mM sodium phosphate (NaP) and 30 mM NaCl (pH 6.0).

Histone octamers were produced by co-expression in *E. coli* BL21 cells using a polycistronic vector (pET29a-YS14 from Addgene) (Shim et al., 2012). After induction, cells were harvested, resuspended in 20 mM Tris-HCl pH 8.0, 2 M NaCl, 1 mM PMSF, 0.1 mM DTT, and lysed by sonication. The lysate was centrifuged and the supernatant incubated with Ni-NTA resin. The resin was subsequently washed in buffer containing 20 mM Tris-HCl pH 8.0, 2.0 M NaCl, 0,1 mM DTT, 50 mM imidazole, and eluted with 150 mM imidazole. Histidine tags were cleaved by overnight digestion with thrombin, and histone octamers further purified by size exclusion chromatography using a Superdex 200 column in buffer containing 20 mM Tris-HCl pH 8, 2 M NaCl, 1 mM DTT. Aliquots were flash frozen in liquid nitrogen and stored at −80 °C.

Nuclear Magnetic Resonance (NMR) spectroscopy

The samples contained 0.8 mM of U-[^{13}C,^{15}N] Rec114 dimer in 20 mM sodium phosphate 20 mM NaCl pH 6.0, 0.02% NaN_3 and 10% D_2O or 1.4 mM of U-[^{13}C,^{15}N] 2:1 Rec114–Mei4 complex in 20 mM sodium phosphate 30 mM NaCl pH 6.0, 5 mM DTT, 0.02 % NaN_3 and 10% D_2O. All NMR spectra were acquired at 298 °K on a Bruker Avance III HD 800 MHz spectrometer, equipped with a TCI cryoprobe. The experimental set comprised 2D [^{1}H,^{15}N] HSQC, [^{1}H,^{13}C] HSQC, and constant-time [^{1}H,^{13}C] HSQC for the aromatic region; 3D ^{15}N edited NOESY-HSQC and ^{13}C-edited NOESY-HSQC for aliphatic and aromatic regions (all recorded with the mixing time of 120 ms); and triple-resonance BEST-HNCACB, BESTHN(CO)CACB, BEST-HNCO, BEST-HN(CA)CO, HBHA(CO)NH, (H)CCH-TOCSY, and H(C)CH-TOCSY spectra. The NMR data were processed in TopSpin 3.6 (Bruker) or NMRPipe (Delaglio et al., 1995), and analyzed in CCPNMR (Vranken et al., 2005). Semi-automatic assignment of the protein backbone was performed in CCPNMR (Vranken et al., 2005). The assignments of N, NH, Hα, Hβ, CO, Cα, and Cβ atoms were obtained from the identification of intra- and inter-residue connectivities in HNCACB, HN(CO)CACB, HNCO, HN(CA)CO, and HBHA(CO)HN experiments at the ^{1}H,^{15}N frequencies of every peak in the [^{1}H,^{15}N] HSQC spectrum. Assignments were extended to the side chain signals using correlations within (H)CCH-TOCSY and H(C)CH-TOCSY experiments. Aromatic ^{1}H and ^{13}C assignments were obtained from constant-time [^{1}H,^{13}C] HSQC and ^{13}C-edited NOESY-HSQC spectra focused on the aromatic region. Remaining aliphatic and aromatic side-chains were assigned from 3D ^{15}N- and ^{13}C-edited NOESY-HSQC spectra. The ^{1}H, ^{13}C and ^{15}N chemical shifts for the 2:1 Rec114–Mei4 complex were deposited in the Biological Magnetic Resonance Bank under the accession number 26335.

The average chemical shift difference ($\Delta\delta_{avg}$) reported in **Supplementary Figure 2C** was calculated as $\Delta\delta_{avg} = (\Delta\delta_N^2/50 + \Delta\delta_H^2/2)^{0.5}$, where $\Delta\delta_N$ and $\Delta\delta_H$ are the chemical shift differences of the backbone amide nitrogen and proton, respectively, for the double HSQC resonances of a given Rec114 residue, The secondary structure of Rec114 and Mei4 in their 2:1 complex (**Supplementary Figure 2B**) was predicted from the backbone chemical shifts using the chemical shift index function and the DANGLE module (Cheung et al., 2010) in CCPNMR (Vranken et al., 2005).

Thermal shift assays

Aggregation (T_{agg}) and melting (T_m) temperatures were obtained from analysis of the static light scattering (SLS) and intrinsic tryptophan fluorescence, respectively, measured simultaneously in UNcle UNI (Unchained Labs, CA, USA). Series of samples containing 2-3 mg/ml of ternary Rec114-Mei4 complexes or Rec114 dimer in 25 mM HEPES-NaOH pH 7.5, 10% Glycerol, 300 mM NaCl, 2 mM DTT, 5 mM EDTA were prepared in triplicate and loaded into the UNcle quartz cells (9 µl per cell). The tryptophan fluorescence spectra and SLS at 266 or 434 nm were measured during a linear temperature gradient of 1 °C/min from 20 to 95 °C. To maximize the frequency of measurements, a holding time was not used. The T_{agg} and T_m values – defined as the inflection points of the corresponding thermal curves – were obtained, respectively, from the analysis of the SLS absorption at 266 nm or the barycentric mean given by the equation 1:

$$\lambda_{BCM} = \sum_{\lambda} \lambda I(\lambda) \Big/ \sum_{\lambda} I(\lambda) \qquad (1),$$

where λ and I(λ) are the wavelength and the corresponding intensity in the fluorescence spectrum, while the summation covers the 300-430 nm region. The thermal curves were analyzed with a two-state transition model given by the equation 2 (van Nuland et al., 1998):

$$y = \frac{y_N + a_N T}{1 + \exp\left(\Delta H_T/RT_T - \Delta H_T/RT\right)} + \frac{(y_U + a_U T)\exp\left(\Delta H_T/RT_T - \Delta H_T/RT\right)}{1 + \exp\left(\Delta H_T/RT_T - \Delta H_T/RT\right)} \qquad (2),$$

where y is the signal observed at temperature T, R is the absolute gas constant, $y_N + a_N T$ and $y_U + a_U T$ are the linear slopes of the pre- and post-transitional regions of the thermal curves, respectively, and ΔH_T is the change in enthalpy at the transition temperature T_T. Thus, T_T values obtained from the non-linear fit of the SLS and tryptophan fluorescence thermal scan curves correspond to T_{agg} and T_m, respectively. The obtained values agreed well with those determined by the UNcle Analysis software.

DNA substrates and gel shift assays

Short double-strand DNA substrates were generated by annealing complementary oligos. The substrates were the following (with oligo names in parentheses): dsDNA (cb95 and cb100), HJ20 (cb922, cb923, cb924 and cb925), HJ40 (cb095, cb096, cb097, cb098), 3'-Flaps (cb095, cb098 and cb122), Fork (cb095, cb098, cb122 and cb120), Y (cb095, cb098 and cb101). The 40-nt and 80-nt oligos were first purified on 10% polyacrylamide-urea gels. Oligos were subsequently mixed in equimolar concentrations (10 µM) in STE (100 mM NaCl, 10 mM Tris-HCl pH 8, 1 mM EDTA), heated and slowly cooled on a PCR thermocycler (98 °C for 3 min, 75 °C for 1 h, 65 °C for 1 h, 37 °C for 30 min, 25 °C for 10 min). For radioactive labeling, 1/20th of the annealed substrates were 5′-end-labelled with [γ-^{32}P]-ATP and T4 polynucleotide kinase (New England Biolabs). For fluorescently labeled HJ40 substrates, oligo cb095 was replaced by 5'-6FAM modified version (dd077). Labeled and unlabeled substrates were purified by native polyacrylamide gel electrophoresis.

For nucleosome assembly, a 147 bp DNA fragment (601 Widom sequence) was amplified by PCR, incubated with histones octamers in a molar ratio of 1.1:1 (octamer:DNA) in buffer containing 10 mM Tris-HCl pH 8, 2 M NaCl, 1 mM EDTA and the salt concentration was decreased step-wise to 0 mM NaCl by dialysis.

Binding reactions (20 µl) were carried out in 25 mM Tris-HCl pH 7.5, 7.5% glycerol, 100 mM NaCl, 2 mM DTT, 1 mM EDTA, and 1 mg/ml BSA. Unless stated otherwise, reactions contained 1 nM radiolabeled substrate or 10 nM fluorescently labeled substrate and the indicated concentration of protein. For Mer2 and IHO1 complexes, concentrations are expressed as monomers. For Rec114–Mei4 complexes, concentrations are expressed as 2:1 heterotrimers. Complexes were assembled for 30 minutes at 30 °C and separated by gel electrophoresis. Binding reactions were separated on 5% TAE-polyacrylamide gels at 150 V for 2 h 30 min, and fluorescent gels were visualized using a Typhoon scanner (Cytiva), while radioactive gels were dried and imaged by autoradiography.

SEC-MALS

Light scattering data were collected using a Superdex 200 increase 10/300 GL Size Exclusion Chromatography (SEC) column connected to a AKTA Pure Chromatography System (Cytiva). The elution from SEC was monitored by a differential refractometer (Optilab, Wyatt), and a static and dynamic, multiangle laser light scattering (LS) detector (miniDAWN, Wyatt). The SEC–UV/LS/RI system was equilibrated in buffer 25 mM HEPES-NaOH pH 7.5, 500 mM NaCl, 10% glycerol, 2 mM EDTA at a flow rate of 0.3 ml/min. The weight average molecular masses were determined across the entire elution profile at intervals of 0.5 s from static LS measurement using ASTRA software.

Structural refinement of the Rec114–Mei4 complex

All simulations were performed in Xplor-NIH v 2.49 (Schwieters et al., 2003), starting from the AlphaFold models of the minimal Rec114–Mei4 complex and the full-length Mei4 obtained in this work. The intermolecular XL-MS data were converted into pairwise distance restraints between lysine sidechains as described elsewhere (Gong et al., 2020). In all refinement runs, the position of the core Rec114–Mei4 minimal complex (residues 399-426 of Rec114 and 16-42 of Mei4) was kept fixed; Mei4 globular domain (residues 65-401) treated as a rigid body group; while the intervening linkers, N- and C-terminal tails of both proteins, and sidechains of crosslinked lysines given full torsional degree of freedom. The computational protocol comprised an initial simulated annealing step followed by the side-chain energy minimization. The total minimized energy function consisted of the standard geometric (bonds, angles, dihedrals, and impropers) and steric (van der Waals) terms, a knowledge-based dihedral angle potential (Schwieters et al., 2003) and the experimental XL-MS restraints term (Gong et al., 2020). In each refinement run, 100 structures were calculated and 10 lowest-energy solutions – representing the best agreement with the experimental data – retained for the subsequent analysis. To model the DNA-bound Rec114-Mei4 complex, we generated the double helical structure of the canonical Watson-Crick paired B-DNA (20-bp oligomer 5'-GAGATGTCCATGGACATCTC-3'), and docked two copies of the DNA duplex to the best structure of the XL-MS refined Rec114-Mei4 complex. The model was obtained by minimizing the distance between the DNA-binding Rec114 residues R395, K396, K399, and R400 and the central grooves of the double-stranded DNA oligomer, while avoiding steric clashes between protein sidechains and the DNA. The resulting model shows the DNA-bound Rec114-Mei4 complex, where Rec114 residues R395, K396, K399, and R400 make a number of intermolecular contacts with the DNA phosphate backbone and nucleotide bases.

Twister analysis of predicted coiled-coil structures

AlphaFold predicted coiled-coil structures were analyzed using Twister (Strelkov and Burkhard, 2002). A coiled coil can be described by the radius and

pitch (i.e. distance along the axis that corresponds to a full turn) of the superhelix, and the radius and pitch of the α-helices. In addition, the pitch of the coiled coil and the α-helices can be expressed per residue (phase yield). Analyses of the coiled-coil and the α-helical phase yields reveal whether the coiled coil and/or the individual helices are locally distorted. Analyses of coiled-coil left and right pitch parameters reveal whether the normally left-handed coiled coil locally switches to a right-handed superhelix. Finally, Crick's angle describes the position of a specific residue relative to the axis of the coiled coil. Variations in Crick's angles between residues that occupy equivalent positions (e.g. a or d) further indicates geometrical distortions of the coiled coil.

End Matter

Author Contributions and Notes

D.D. and C.C.B designed research, D.D., P.L., E.D.J. and A.N.V. performed research, D.D., A.N.V. and C.C.B analyzed data; and D.D., A.N.V. and C.C.B wrote the paper.
The authors declare no conflict of interest.

Acknowledgments

We thank members of the CCB lab for discussions and comments on the manuscript, particularly Cédric Oger and David Álvarez Melo. We thank Marita Haddad and Joseph Nader for experimental support, and Gholamreza Hassanzadeh Ghassabeh and Ema Romao of the VIB Nanobody Core Facility for providing access to the UNcle setup and help with the measurements. This work was supported by the European Research Council under the European Union's Horizon 2020 research and innovation program (ERC grant agreement 802525 to CCB), and the Fonds National de la Recherche Scientifique (MIS-Ulysse grant F.6002.20 to CCB). DD was funded by an Aspirant fellowship of the FNRS. CCB is a FNRS Research Associate.

References

Andrade, M.A., Petosa, C., O'Donoghue, S.I., Müller, C.W., and Bork, P. (2001). Comparison of ARM and HEAT protein repeats. J Mol Biol 309, 1-18.

Arora, C., Kee, K., Maleki, S., and Keeney, S. (2004). Antiviral protein Ski8 is a direct partner of Spo11 in meiotic DNA break formation, independent of its cytoplasmic role in RNA metabolism. Mol Cell 13, 549-559.

Boekhout, M., Karasu, M.E., Wang, J., Acquaviva, L., Pratto, F., Brick, K., Eng, D.Y., Xu, J., Camerini-Otero, R.D., Patel, D.J., et al. (2019). REC114 Partner ANKRD31 Controls Number, Timing, and Location of Meiotic DNA Breaks. Mol Cell 74, 1053-1068.e1058.

Bonfils, S., Rozalen, A.E., Smith, G.R., Moreno, S., and Martin-Castellanos, C. (2011). Functional interactions of Rec24, the fission yeast ortholog of mouse Mei4, with the meiotic recombination-initiation complex. J Cell Sci 124, 1328-1338.

Cheung, M.S., Maguire, M.L., Stevens, T.J., and Broadhurst, R.W. (2010). DANGLE: A Bayesian inferential method for predicting protein backbone dihedral angles and secondary structure. J Magn Reson 202, 223-233.

Claeys Bouuaert, C., Pu, S., Wang, J., Oger, C., Daccache, D., Xie, W., Patel, D.J., and Keeney, S. (2021). DNA-driven condensation assembles the meiotic DNA break machinery. Nature.

de Massy, B. (2013). Initiation of meiotic recombination: how and where? Conservation and specificities among eukaryotes. Annu Rev Genet 47, 563-599.

De Muyt, A., Pereira, L., Vezon, D., Chelysheva, L., Gendrot, G., Chambon, A., Laine-Choinard, S., Pelletier, G., Mercier, R., Nogue, F., et al. (2009). A high throughput genetic screen identifies new early meiotic recombination functions in Arabidopsis thaliana. PLoS Genet 5, e1000654.

Delaglio, F., Grzesiek, S., Vuister, G.W., Zhu, G., Pfeifer, J., and Bax, A. (1995). NMRPipe: a multidimensional spectral processing system based on UNIX pipes. J Biomol NMR 6, 277-293.

Fischer, E.S., Scrima, A., Böhm, K., Matsumoto, S., Lingaraju, G.M., Faty, M., Yasuda, T., Cavadini, S., Wakasugi, M., Hanaoka, F., et al. (2011). The molecular basis of CRL4DDB2/CSA ubiquitin ligase architecture, targeting, and activation. Cell 147, 1024-1039.

Gong, Z., Ye, S.X., and Tang, C. (2020). Tightening the Crosslinking Distance Restraints for Better Resolution of Protein Structure and Dynamics. Structure 28, 1160-1167.e1163.

Guo, H., Stamper, E.L., Sato-Carlton, A., Shimazoe, M.A., Li, X., Zhang, L., Stevens, L., Tam, K.C.J., Dernburg, A.F., and Carlton, P.M. (2022). Phosphoregulation of DSB-1 mediates control of meiotic double-strand break activity. Elife 11.

Henderson, K.A., Kee, K., Maleki, S., Santini, P.A., and Keeney, S. (2006). Cyclin-dependent kinase directly regulates initiation of meiotic recombination. Cell 125, 1321-1332.

Hinman, A.W., Yeh, H.Y., Roelens, B., Yamaya, K., Woglar, A., Bourbon, H.G., Chi, P., and Villeneuve, A.M. (2021). Caenorhabditis elegans DSB-3 reveals conservation and divergence among protein complexes promoting meiotic double-strand breaks. Proc Natl Acad Sci U S A 118.

Holm, L. (2022). Dali server: structural unification of protein families. Nucleic Acids Res 50, W210-215.

Hunter, N. (2015). Meiotic Recombination: The Essence of Heredity. Cold Spring Harb Perspect Biol 7.

Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A., et al. (2021). Highly accurate protein structure prediction with AlphaFold. Nature 596, 583-589.

Keeney, S. (2008). Spo11 and the Formation of DNA Double-Strand Breaks in Meiosis. Genome Dyn Stab 2, 81-123.

Kumar, R., Bourbon, H.M., and de Massy, B. (2010). Functional conservation of Mei4 for meiotic DNA double-strand break formation from yeasts to mice. Genes Dev 24, 1266-1280.

Kumar, R., Oliver, C., Brun, C., Juarez-Martinez, A.B., Tarabay, Y., Kadlec, J., and de Massy, B. (2018). Mouse REC114 is essential for meiotic DNA double-strand break formation and forms a complex with MEI4. Life science alliance 1, e201800259.

Lam, I., and Keeney, S. (2015). Mechanism and regulation of meiotic recombination initiation. Cold Spring Harb Perspect Biol 7, a016634.

Li, J., Hooker, G.W., and Roeder, G.S. (2006). Saccharomyces cerevisiae Mer2, Mei4 and Rec114 form a complex required for meiotic double-strand break formation. Genetics 173, 1969-1981.

Maleki, S., Neale, M.J., Arora, C., Henderson, K.A., and Keeney, S. (2007). Interactions between Mei4, Rec114, and other proteins required for meiotic DNA double-strand break formation in Saccharomyces cerevisiae. Chromosoma 116, 471-486.

Miyoshi, T., Ito, M., Kugou, K., Yamada, S., Furuichi, M., Oda, A., Yamada, T., Hirota, K., Masai, H., and Ohta, K. (2012). A central coupler for recombination initiation linking chromosome architecture to S phase checkpoint. Mol Cell 47, 722-733.

Molnar, M., Parisi, S., Kakihara, Y., Nojima, H., Yamamoto, A., Hiraoka, Y., Bozsik, A., Sipiczki, M., and Kohli, J. (2001). Characterization of rec7, an early meiotic recombination gene in Schizosaccharomyces pombe. Genetics 157, 519-532.

Nonomura, K., Nakano, M., Fukuda, T., Eiguchi, M., Miyao, A., Hirochika, H., and Kurata, N. (2004). The novel gene HOMOLOGOUS PAIRING ABERRATION IN RICE MEIOSIS1 of rice encodes a putative coiled-coil protein required for homologous chromosome pairing in meiosis. Plant Cell 16, 1008-1020.

Panizza, S., Mendoza, M.A., Berlinger, M., Huang, L., Nicolas, A., Shirahige, K., and Klein, F. (2011). Spo11-accessory proteins link double-strand break sites to the chromosome axis in early meiotic recombination. Cell 146, 372-383.

Parry, D.A., Fraser, R.D., and Squire, J.M. (2008). Fifty years of coiled-coils and alpha-helical bundles: a close relationship between sequence and structure. J Struct Biol 163, 258-269.

Pawlowski, W.P., Golubovskaya, I.N., Timofejeva, L., Meeley, R.B., Sheridan, W.F., and Cande, W.Z. (2004). Coordination of meiotic recombination, pairing, and synapsis by PHS1. Science 303, 89-92.

Raisch, T., Bhandari, D., Sabath, K., Helms, S., Valkov, E., Weichenrieder, O.,

and Izaurralde, E. (2016). Distinct modes of recruitment of the CCR4-NOT complex by Drosophila and vertebrate Nanos. Embo j 35, 974-990.

Rosu, S., Zawadzki, K.A., Stamper, E.L., Libuda, D.E., Reese, A.L., Dernburg, A.F., and Villeneuve, A.M. (2013). The C. elegans DSB-2 protein reveals a regulatory network that controls competence for meiotic DSB formation and promotes crossover assurance. PLoS Genet 9, e1003674.

Rousová, D., Nivsarkar, V., Altmannova, V., Raina, V.B., Funk, S.K., Liedtke, D., Janning, P., Müller, F., Reichle, H., Vader, G., et al. (2021). Novel mechanistic insights into the role of Mer2 as the keystone of meiotic DNA break formation. Elife 10.

Schwieters, C.D., Kuszewski, J.J., Tjandra, N., and Clore, G.M. (2003). The Xplor-NIH NMR molecular structure determination package. J Magn Reson 160, 65-73.

Shim, Y., Duan, M.R., Chen, X., Smerdon, M.J., and Min, J.H. (2012). Polycistronic coexpression and nondenaturing purification of histone octamers. Anal Biochem 427, 190-192.

Stamper, E.L., Rodenbusch, S.E., Rosu, S., Ahringer, J., Villeneuve, A.M., and Dernburg, A.F. (2013). Identification of DSB-1, a protein required for initiation of meiotic recombination in Caenorhabditis elegans, illuminates a crossover assurance checkpoint. PLoS Genet 9, e1003679.

Stanzione, M., Baumann, M., Papanikos, F., Dereli, I., Lange, J., Ramlal, A., Trankner, D., Shibuya, H., de Massy, B., Watanabe, Y., et al. (2016). Meiotic DNA break formation requires the unsynapsed chromosome axis-binding protein IHO1 (CCDC36) in mice. Nat Cell Biol 18, 1208-1220.

Steiner, S., Kohli, J., and Ludin, K. (2010). Functional interactions among members of the meiotic initiation complex in fission yeast. Curr Genet 56, 237-249.

Strelkov, S.V., and Burkhard, P. (2002). Analysis of alpha-helical coiled coils with the program TWISTER reveals a structural mechanism for stutter compensation. J Struct Biol 137, 54-64.

Tesse, S., Bourbon, H.M., Debuchy, R., Budin, K., Dubois, E., Liangran, Z., Antoine, R., Piolot, T., Kleckner, N., Zickler, D., et al. (2017). Asy2/Mer2: an evolutionarily conserved mediator of meiotic recombination, pairing, and global chromosome compaction. Genes Dev 31, 1880-1893.

Tsai, B., Liu, W., Dong, D., Shi, K., Chen, L., and Gao, N. (2020). Phase separation of Mer2 organizes the meiotic loop-axis structure of chromatin during meiosis I. bioRxiv, 2020.2012.2015.422856.

van Nuland, N.A., Meijberg, W., Warner, J., Forge, V., Scheek, R.M., Robillard, G.T., and Dobson, C.M. (1998). Slow cooperative folding of a small globular protein HPr. Biochemistry 37, 622-637.

Vranken, W.F., Boucher, W., Stevens, T.J., Fogh, R.H., Pajon, A., Llinas, M., Ulrich, E.L., Markley, J.L., Ionides, J., and Laue, E.D. (2005). The CCPN data model for NMR spectroscopy: development of a software pipeline. Proteins 59, 687-696.

Vrielynck, N., Schneider, K., Rodriguez, M., Sims, J., Chambon, A., Hurel, A., De Muyt, A., Ronceret, A., Krsicka, O., Mézard, C., et al. (2021). Conservation and divergence of meiotic DNA double strand break forming mechanisms in Arabidopsis thaliana. Nucleic Acids Res 49, 9821-9835.

Wang, W., Dong, J., Chen, B., Du, J., Kuang, Y., Sun, X., Fu, J., Li, B., Mu, J., Zhang, Z., et al. (2019). Homozygous mutations in REC114 cause female infertility characterised by multiple pronuclei formation and early embryonic arrest. Journal of medical genetics.

Wüthrich, K. (1986). NMR of proteins and nucleic acids (Wiley).

Yadav, V.K., and Claeys Bouuaert, C. (2021). Mechanism and control of meiotic DNA double-strand break formation in S. cerevisiae. Frontiers in cell and developmental biology.