UCLouvain

AGRO LOUVAIN

Faculté des bioingénieurs

Study of the conservation of the structure and the function of *S. cerevisiae* Rec114-Mei4-Mer2 proteins during evolution

Autrice : Emma De Jonge

Promoteurs : Prof. Corentin Claeys Bouuaert,

Prof. Pierre Morsomme

Lecteurs : Prof. François Chaumont,

Prof. Patrice Soumillion

Année académique 2022-2023

Mémoire de fin d'études présenté en vue de l'obtention du diplôme de Bioingénieur : Chimie et bio-industries

Remerciements

Tout d’abord, je souhaiterais remercier le professeur Corentin Claeys Bouuaert, pour m'avoir accueillie dans son laboratoire et m'avoir offert cette précieuse opportunité de développer mes compétences en recherche. Professeur Claeys Bouuaert, je suis reconnaissante de tout le temps que vous avez consacré à me guider dans mes expériences. Vos précieux conseils, votre disponibilité et votre bienveillance m'ont permis d'avancer dans mes expériences.

Je tiens également à remercier chaleureusement Dima Daccache, mon encadrante. Son expertise, sa générosité et sa disponibilité ont été des atouts majeurs pour mon apprentissage. Grâce à ses conseils avisés et à son soutien constant, j'ai pu surmonter les obstacles et progresser dans mes recherches. Je lui suis profondément reconnaissante pour tout le temps qu'elle m'a accordé. Dima, un grand merci pour tous les moments partagés ensemble !

Je souhaite également exprimer ma reconnaissance envers les membres du laboratoire CCB Lab, à savoir : Pascaline, Cédric, Marita, Karen, David, Hajar, Mahesh, Chi Wai, Priyanka, Chaïmae et Marwane. Leurs conseils éclairés, leur aide, leur expertise et l'ambiance de travail conviviale ont grandement contribué à mon épanouissement professionnel. Leur support et leur enthousiasme ont rendu cette expérience encore plus enrichissante.

J’aimerais aussi adresser un remerciement spécial à Lisa Viatour, ma cokoteuse et amie, pour les repas délicieux qu'elle me préparait lorsque je travaillais tard. Sa générosité et sa gentillesse m'ont donné l'énergie nécessaire pour continuer mes recherches.

Je souhaite également remercier mes parents pour le soutien indéfectible qu'ils m'ont apporté tout au long de mes études. Leur amour, leur encouragement et leur confiance en moi ont été des moteurs essentiels dans la poursuite de mon cursus universitaire.

Enfin, je tiens à remercier Louis Fivet, mon amoureux, pour son soutien émotionnel constant tout au long de cette année. De plus, j’aimerais le remercier pour la réalisation des magnifiques tableaux reprenant l’alignement de séquences protéiques des SMM des orthologues de Rec114, Mei4 et Mer2 que vous pourrez admirer en Figure 16, 17 et 26. Louis, un tout grand merci pour le temps que tu as passé pour que j’aie ces jolies figures dans mon mémoire.

Je suis consciente que cette réussite n'aurait pas été possible sans l’aide de toutes ces personnes et je leur en suis grandement reconnaissante.

Résumé

En prophase I de méiose, a lieu le phénomène de recombinaison génétique qui peut mener à la formation de crossing-over entre deux chromosomes homologues. Ce phénomène est essentiel pour permettre une ségrégation correcte des chromosomes homologues en fin d'anaphase I de méiose.

Le phénomène de recombinaison génétique est initié par une coupure double brin de l'ADN médiée, chez *S. cerevisiae*, par la protéine Spo11. Cependant, ce phénomène requiert également la présence de 9 autres protéines réparties en trois complexes : Le core complexe (qui contient l'unité catalytique Spo11), le complexe MRX et le complexe RMM (composé de Rec114, Mei4 et de Mer2). Il a été montré *in vivo* et *in vitro* que les protéines du complexe RMM forment des condensés en présence d'ADN. Ce phénomène se réfère à de la séparation de phase. Il est supposé que ce complexe jouerait le rôle de plateforme qui recruterait les 7 autres protéines essentielles à l'endroit où se produisent les coupures double brin de l'ADN.

Dans ce mémoire, l'objectif est d'étudier la conservation au sein de l'évolution de la structure et de la fonction de liaison à l'ADN *in vitro* des protéines orthologues au complexe RMM de *S. cerevisiae*. Ces protéines ne pouvant pas se purifier ensemble, l'étude est divisée en deux parties distinctes : Une partie sur les complexes orthologues à Rec114-Mei4 où les complexe de souris (REC114-MEI4), de *S. pombe* (Rec7-Rec24*)*, d'*A. thaliana* (PHS1-PRD2) et du maïs (PHS1-MPS1) seront étudiés et une autre sur les orthologues de Mer2 où la protéine de souris (IHO1), de *S. pombe* (Rec15), d'*A. thaliana* (PRD3), du maïs (PAIR1) et de *S. macrospora* (ASY2) seront étudiées.

En effet, malgré la grande différence des séquences en acides aminés entre ces protéines orthologues, la structure prédite par AlphaFold est assez similaire. Ces modèles du complexe formé par des orthologues de Rec114 et de Mei4, avec une stœchiométrie 2 :1 respectivement, vont être vérifiés par une mutagenèse suivie d'un Pulldown afin d'identifier des acides aminés importants dans les interactions protéiques entre la partie C-terminale de Rec114 et la partie N-terminale de Mei4. Concernant les orthologues de Mer2, seule la structure des parties des protéines formant un « coiled-coil domain » va être étudiée par SEC-MALS, SAXS et XL-MS, afin de voir si la structure homotétramérique et parallèle de Mer2 est conservée.

Afin d'étudier la fonction de liaison à l'ADN de ces protéines, des EMSA sont réalisées sur les complexes minimaux et les « coiled-coil domain » des protéines étudiées. En effet, chez *S. cerevisiae*, ces protéines se lient *in vitro* à l'ADN et la conservation de cette activité est étudiée dans ce mémoire.

Abstract

In prophase I of meiosis, the phenomenon of genetic recombination takes place, which can lead to the formation of crossovers between two homologous chromosomes. This phenomenon is essential for the correct segregation of homologous chromosomes at the end of anaphase I of meiosis.

Genetic recombination is initiated by a DNA double-strand break mediated, in *S. cerevisiae*, by the Spo11 protein. However, this phenomenon also requires the presence of 9 other proteins divided into three complexes: The core complex (which contains the Spo11 catalytic unit), the MRX complex and the RMM complex (composed of Rec114, Mei4 and Mer2). It has been shown *in vivo* and *in vitro* that the proteins of the RMM complex form condensates in the presence of DNA. This phenomenon is referred to as phase separation. It is hypothesized that this complex acts as a platform that recruits the 7 other essential proteins to the site where DNA double-strand breaks occur.

The aim of this master thesis is to study the conservation of the structure and DNA-binding function of the *S. cerevisiae* RMM complex orthologs throughout evolution. As these proteins cannot be purified together, the study is divided into two distinct parts: A part on the Rec114-Mei4 orthologs where the complexes from mice (REC114-MEI4), *S. pombe* (Rec7-Rec24), *A. thaliana* (PHS1-PRD2) and *Z. mays* (PHS1-MPS1) will be studied and another on Mer2 orthologs where the protein from mouse (IHO1), *S. pombe* (Rec15), *A. thaliana* (PRD3), *Z. mays* (PAIR1) and *S. macrospora* (ASY2) will be studied.

Despite the large difference in amino acid sequences between these orthologous proteins, the structure predicted by AlphaFold is quite similar. These models of the complex formed by orthologs of Rec114 and Mei4, with a 2:1 stoichiometry respectively, will be verified by mutagenesis followed by a Pulldown in order to identify the amino acids important in the protein interactions between the C-terminal part of Rec114 and the N-terminal part of Mei4. Concerning the Mer2 orthologs, only the structure of the parts of the proteins forming a coiled-coil domain will be studied by SEC-MALS, SAXS and XL-MS, to see whether the homotetrameric and parallel structure of Mer2 is conserved.

In order to study the DNA-binding function of these proteins, EMSAs were carried out on the minimal complexes and coiled-coil domains of the proteins studied. In *S. cerevisiae*, these proteins bind DNA *in vitro* and the conservation of this activity is studied in this work.

Table of contents

I. Acronyms

aa	Amino Acid
AFM	Atomic Force Microscopy
Amp	Ampicilin
BSA	Bovine Serum Albumin
cc	Coiled-Coil
CDK	Cyclin Dependent Kinase
CO	Crossover
Cryo-EM	Cryogenic Electron Microscopy
DNA	Deoxyribonucleic Acid
dNTP	Deoxynucleotide Triphosphate
OD	Optical Density
O/N	Over Night
DSB	DNA Double-Strand Break
DSS	Disuccinimidyl Suberate
DTT	Dithiothreitol
EDTA	Ethylene Diamine Tetra-acetic Acid
EMSA	Electrophoretic Mobility Shift Assay
FL	Full-Length
HJ	Holliday Junction
IPTG	Isopropyl β-D-1-thiogalactopyranoside
Kan	Kanamycin
LC	Liquid Chromatography
MBP	Maltose binding protein
mRM	$Rec114^{375-428}$ - $Mei4^{1-43}$
MS	Mass Spectrometry
PCR	Polymerase Chain Reaction

RM	Rec114-Mei4
RMM	Rec114-Mei4-Mer2
RNA	Ribonucleic Acid
SAXS	Small Angle X-rays Scattering
SC	Synaptonemal Complex
SDS-PAGE	Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis
SEC-MALS	Size-Exclusion Chromatography – Multi-Angle Light Scattering
SSM	Signature Sequence Motif
SUMO	Small Ubiquitin-like Modifier
Ulp1	Ubl-specific protease 1
XL	Crosslink
XL- MS	Crosslinking-Mass Spectrometry
WT	Wild Type

II. Introduction

In this master thesis, we will investigate the evolutionary conservation of structure and function of *Saccharomyces cerevisiae* proteins involved in DNA double-strand breaks during meiosis prophase I.

1. The main stages of meiosis

To form a diploid (2n chromosomes) individual through fertilization, the production of haploid (n chromosomes) gametes by parental cells is required via a process called meiosis. During meiosis, the genetic material of the parent cell undergoes a phase of DNA replication followed by two successive divisions that will lead to the formation of 4 haploid gametes. Each of these divisions is composed of four different stages: prophase, metaphase, anaphase, and telophase (Figure 1) (Page and Hawley 2003).

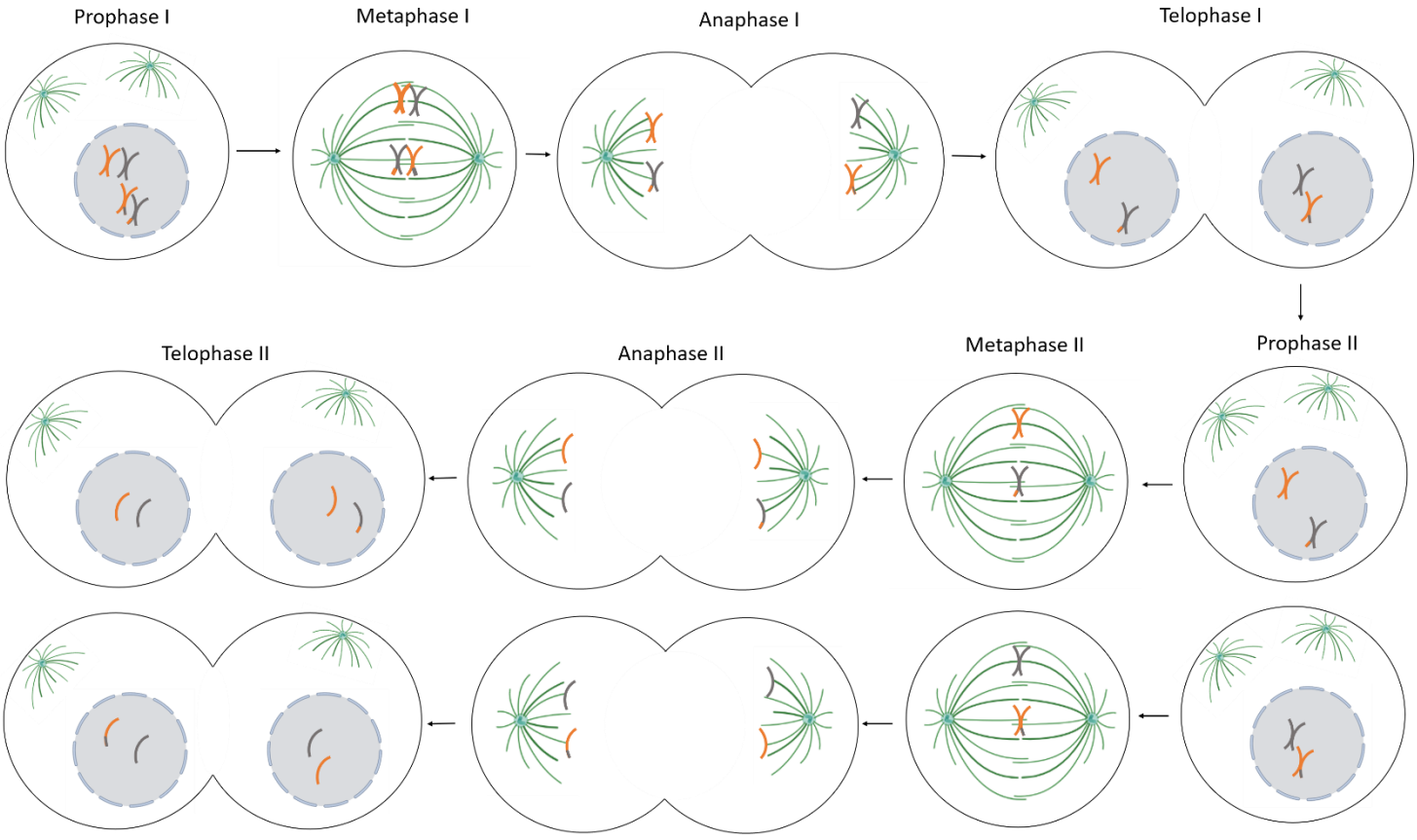

Figure 1: Meiosis phases. *General scheme of meiosis for a eukaryotic cell with two pairs of chromosomes leading to the formation of four gametes with a single copy of each chromosome.*

1.1 Interphase I

Also known as S-Phase, preprophase is a stage during which the cells genetic material will replicate, which will lead to the formation of two identical copies of each chromosome, linked together via a centromere to form a chromosome with two sister chromatids (Zickler and Kleckner 1998).

1.2 First meiotic division

The first meiotic division is called "reductional". Indeed, the number of chromosomes will decrease by half. It is also during this first division that genetic recombination takes place (Kleckner 1996).

1.2.1 Prophase I

This phase of meiosis is the longest and consists of five different chromosome stages: leptotene, zygotene, pachytene, diplotene and diacinese (Figure 2).

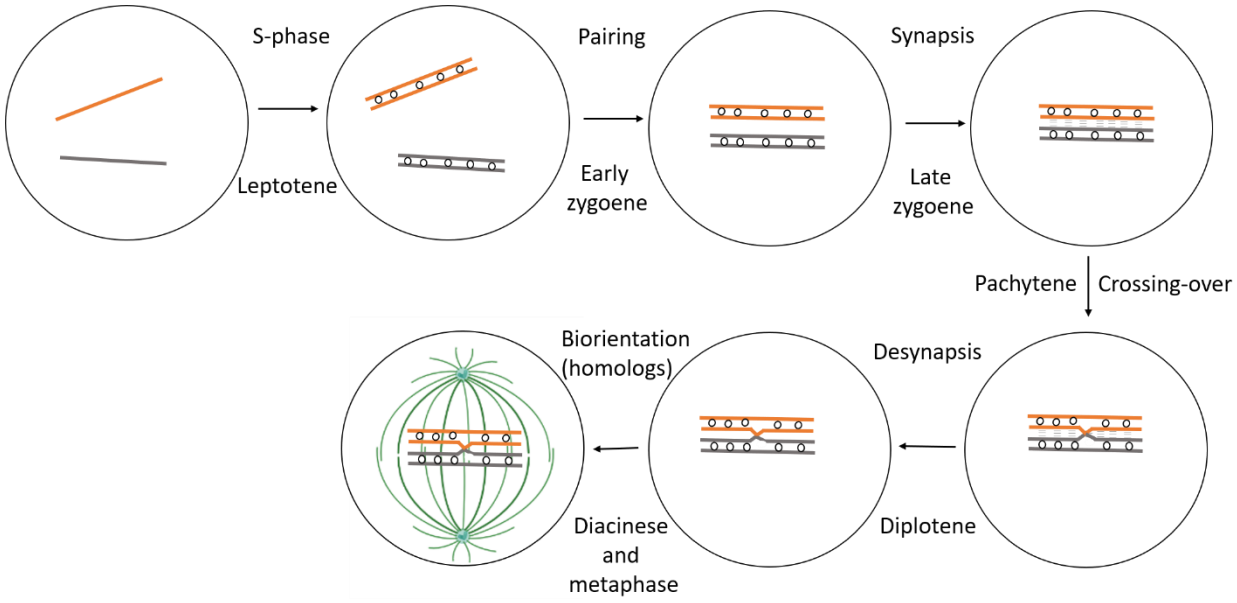

***Figure 2:* Stages of prophase I of meiosis.** *Diagram showing the chromosomal aspects of a single pair of homologous chromosomes (orange and grey lines) of the different stages of prophase I of meiosis including leptotene, zygotene, pachytene, diplotene and diacinese. Black rings indicate cohesion complexes connecting the sister chromatids. The synaptonemal complex is indicated by the triple-dashed line located between the homologs (inspired by Hunter 2015).*

Leptotene

Chromosomes individualize as filaments attached via their ends to the nuclear membrane and begin to condense. In addition, it is in this stage that the programmed DNA double strand breaks occur (DSBs), this is the first step of the genetic recombination (Zickler and Kleckner 2015).

Zygotene

At this stage, each chromosome is searching for its homologue and it will allow them to be paired "gene to gene" to form a bivalent. This pairing is also called synapsis and the synaptonemal complex (SC) is formed. The SC is formed when the two homologous chromosomes are connected by intermediate filaments (Zickler and Kleckner 2015).

Pachytene

This stage marks the end of homologous chromosome pairing. It is during the end of this stage that crossovers (CO) occur between the chromatids of two homologous chromosomes (Hunter 2015). CO are an exchange of genetic material between two homologous chromosomes.

Diplotene

During this stage, the divalents separate and only remain attached to each other through crossovers (chiasms) (Hunter 2015).

Diacinese

Chromosomes are associated end-to-end and they undergo a final stage of condensation. This step marks the end of prophase I (Pawlowski and Cande 2005).

1.2.2 Metaphase I

Pairs of homologous chromosomes are arranged along the equatorial plate and are held together by the mitotic spindle at the kinetochores. The homologous chromosomes are held together by the chiasmata. Each of the bivalent chromosomes is connected to one of the spindle poles (Kleckner 1996; Page and Hawley 2003).

1.2.3 Anaphase I

Chiasms are broken, the homologous chromosomes are separated at the different poles of the cell. It is during this stage that the random distribution of homologous chromosomes occurs so that haploid gametes can be formed (Kleckner 1996; Page and Hawley 2003).

1.2.4 Telophase I

Formation of 2 distinct cells with n chromosomes. In each cell, there is only one chromosome from the homologous chromosome pair.

1.3 Second meiotic division

The second meiotic division is called "equational". This division is similar to a mitotic division except that it is not preceded by a DNA replication phase. At the end of telophase II, four gametes with n chromosomes are produced, each cell has one of the two chromatids of each chromosome present at the end of telophase I (O'Connor 2008).

2. Crossing-over (CO) and genetic recombination

A crossing-over is an exchange of genetic material between two homologous chromosomes. The main purpose of these exchanges is to increase genetic diversity, but also to associate homologous chromosomes in pairs and allow their proper segregation at the end of meiosis I (Zickler and Kleckner 2015; Wilkins and Holliday 2009; Petronczki, Siomos, and Nasmyth 2003).

Crossovers are initiated by DNA double-strand breaks (DSBs) followed by homologous recombination, which consists of copying the genetic information of an intact template. These DSBs are carried out by the conserved topoisomerase-like protein Spo11 (Bergerat *et al.* 1997; Lam and Keeney 2015; Keeney, Giroux, and Kleckner 1997). There are hundreds of DSBs per meiosis in mice and yeast. It should be noted that not all of these DNA double-strand breaks result in crossovers. They may be repaired and lead to non-crossovers (Yadav and Claeys Bouuaert 2021; Martini *et al.* 2006).

In *Saccharomyces cerevisiae*, Spo11 is recruited at the chromosome level and after the DNA cut, Spo11 will remain covalently attached to the 5' ends of the cut DNA. Next, the MRX and Sae2 proteins will release Spo11 and an oligonucleotide by cutting the single-stranded DNA. The remaining 5' end of each strand will be removed by the activity of an exonuclease (Exo1) to leave a single strand of DNA that will be surrounded by RPA proteins that prevents single-stranded DNA from binding to itself. These RPA proteins will be replaced by the Rad51 and Dmc1 proteins which will seek an intact template similar to the damaged one. This template will be found in the homologous chromosome. Depending on how the DNA is copied, crossovers may or may not be observed (Figure 3) (Yadav and Claeys Bouuaert 2021; Bee *et al.* 2013).

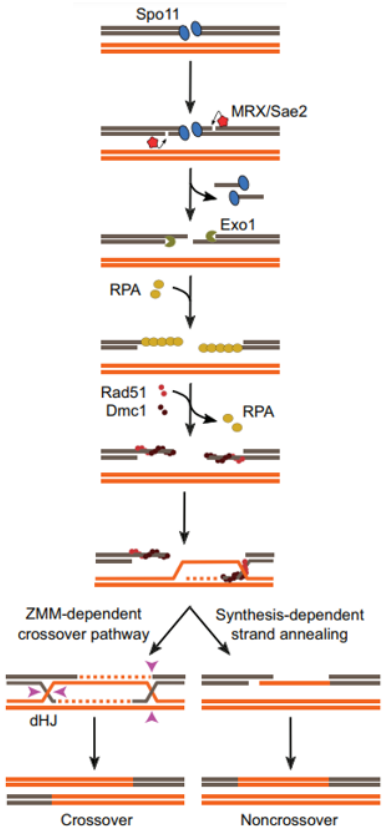

Figure 3: Overview of the mechanism of crossover formation. *Overview of crossover formation or noncrossover by synthesis-dependent strand annealing from a Spo11-catalyzed DNA double-strand break (Yadav and Claeys Bouuaert 2021).*

The frequency of these crossovers depends on several factors such as distance of the sequences from the centromere. A region close to the centromere is less likely to undergo a crossover. It has been shown that COs occur at different positions in the chromosomes in different meiotic nuclei. However, COs tend to be spaced apart from each other. Indeed, if a CO takes up a position on one of the chromosomes, the probability of having a CO close (physically) to it is reduced. This phenomenon is called CO interference (Zickler and Kleckner 2015; Fowler *et al.* 2018).

The appearance of DSBs is highly regulated because they are not randomly distributed in the genome of the organisms. DSBs are concentrated in regions of the genome called hotspots and are separated by regions called cold spots, which have a low frequency of DSBs (Zhu and Keeney 2015). Hotspots are often AT-rich and are flanked by sequences enriched in methylated histone: H3 lysine 4 trimethylation (Borde *et al*. 2009). In addition, a constant number of DSBs were observed in the cells. For example, in *S. cerevisiae*, only 150 to 200 DSBs occur in the entire genome (Zhu and Keeney 2015).

3. The structure of the DNA during prophase I

In prophase I of meiosis, the DNA takes on a particular three-dimensional structure that allow the genetic recombination to take place (Figure 4).

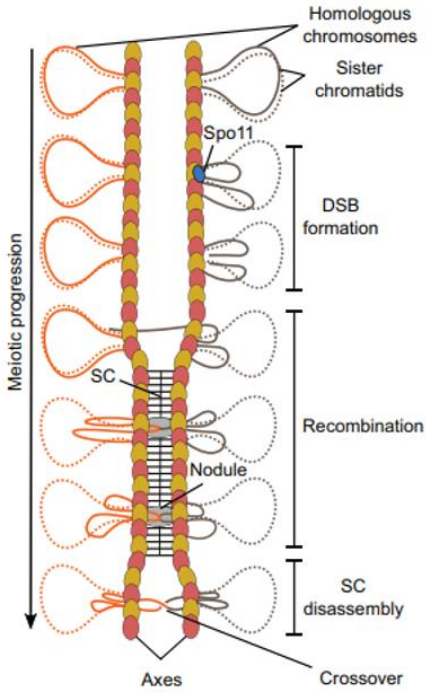

Figure 4: DNA organization in the prophase. *Schema of the DNA organization in prophase I of meiosis. From the beginning of prophase, the chromosomes are organized to create loops anchored in a nucleoprotein axis. DSBs arrive in these loops, which are captured at the axis by certain proteins to allow Spo11, located at the axis, to cut the DNA. Next, the synaptonemal complex is formed by connecting the two homologous chromosomes to allow recombination to take place. The synaptonemal complex then disassembles when the genetic recombinations are complete and the chromosomes remain attached by the crossovers (Yadav and Claeys Bouuaert 2021).*

Indeed, chromosomes are organized into a series of loops that are anchored along their length to an axis formed by structural proteins (Lam and Keeney 2014; Kleckner 2006). The components of the structural protein axis include condensins, cohesins and cohesin-associated proteins and a series of meiosis-specific proteins (Red1, Hop1...) and

its structure is evolutionarily conserved. These axis proteins recruit the machinery responsible for the formation of DSBs, the first step of a genetic recombination. However, the DSBs occur in theses loops. Therefore, it is suggested that the loops are tethered to the nucleoprotein axe to allow the DNA to be cut. During the zygotene stage of meiosis prophase I, the synaptonemal complex (SC) is formed. The SC is formed when the nucleoprotein axes of the two homologous chromosomes are connected by intermediate filaments (Zickler and Kleckner 1999). The SC has several roles that have been identified such as stabilizing the structure of chromosomes around future crossover sites to allow the recombination to take place. When recombinations are complete, the SC disassembles and the homologous chromosomes are held together only by their crossovers until anaphase I (Figure 4) (Yadav and Claeys Bouuaert 2021).

4. DNA Double-Strand Break (DSB) proteins

In *S. cerevisiae*, it has been shown that DNA double-strand breaks are formed by Spo11 and 9 other proteins subdivided in tree complex of proteins: the Core complex, containing Spo11, the MRX complex and the RMM complex. Other proteins are also important but if the expression of one of these 10 proteins is suppressed, no more DSBs are observed which lead to a sporulation defect due to a mis-segregation of chromosomes in meiosis I (Lam and Keeney 2014). However, not all of these proteins in amino acid sequence are well conserved in evolution which made their identification complicated.

4.1 The core complex

The core complex is a complex formed by the proteins Spo11, Ski8, Rec102 and Rec104. These proteins are meiosis specific except Ski8 (Anderson and Parker 1998). These proteins interact with each other to form a stoichiometric complex (1:1:1:1). It is assumed that this complex dimerizes as a dimer of tetramers in presence of DNA , but this has not yet been biochemically proven (Figure 5) (Claeys Bouuaert, Tischfield, *et al.* 2021).

The catalytic unit Spo11 evolved from the catalytic subunit of a type II B topoisomerase, Topo VI. Rec102 and Rec 104 have been shown to play an important role by interacting with the RMM proteins (in particular with the Rec114 protein) which will lead to the recruitment of the core complex to the chromosome axis (Maleki *et al.* 2007)

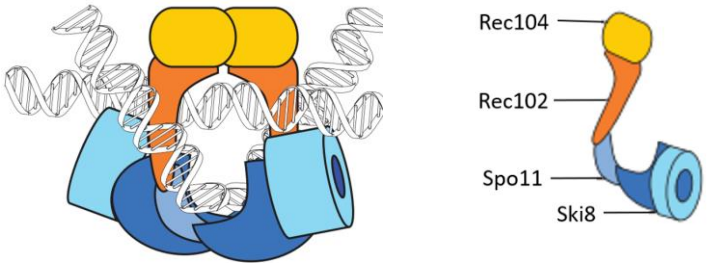

Figure 5: Core complex representation. *Cartoon showing the core complex in S. cerevisiae. This core complex is formed by Rec104, Rec102, Spo11 and Ski8. It has been shown that this complex binds to two DNA molecules in order to cut both strands of one of these molecules and to remain covalently bound to the 5' ends of the DNA molecule (Yadav and Claeys Bouuaert 2021).*

4.2 The MRX complex

The MRX complex consists of the proteins Xrs2, Mre11 and Rad50 and has a stoichiometry 2:2:2. These proteins play a role in DSB formation, in the recognition of DSBs, activation of the DNA-damage checkpoint and telomere maintenance. They also play a role in the repair of double-strand breaks by resecting DNA ends covalently bound to Spo11(Steiner, Kohli, and Ludin 2010; Gnügge and Symington 2017). The Rad50 and Mre11 proteins dimerize to form the complex shown in Figure 6.

Mre11 has 5'-end directed endonuclease and 3'-5' exonuclease activity. Rad50 is an ATPase, which by hydrolysis of ATP, makes the active site of Mre11 available to bind to DNA and cleave it by dissociating the Rad50 dimer (Gobbini *et al.* 2016). This has the effect of, as we have seen in Figure 3, releasing a single-stranded DNA oligo bound to Spo11.

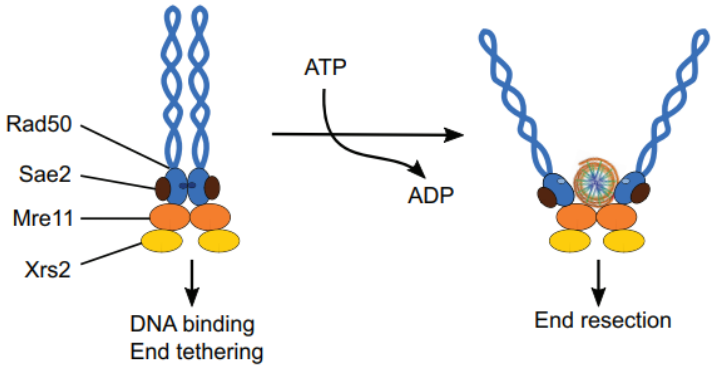

Figure 6: Representation of the MRX complex. *Cartoon of the structural arrangement of the MRX complex before and after ATP hydrolysis by Rad50. Before ATP hydrolysis, the complex can bind to DNA and after hydrolysis, the exonuclease activity of Mre1 can access the DNA (Yadav and Claeys Bouuaert 2021).*

The recruitment of Mre11 to the nucleoprotein axis where chromosomes are anchored requires all other DSBs, suggesting that this complex is the last to be recruited to the axis. One hypothesis is that its recruitment is necessary to activate Spo11 (Borde *et al.* 2004).

4.3 The RMM complex

This complex of proteins comprises two subcomplexes: one formed by Rec114-Mei4 and another formed by Mer2. The proteins Rec114 and Mei4 are specifically expressed during meiosis, while Mer2 is also expressed at a low level in vegetative cells, its expression is different during meiosis due to an intron that is only spliced with the help of meiosis-specific transcription factors (Engebrecht, Voelkel-Meiman, and Roeder 1991).

It was demonstrated *in vivo* by fluorescent signal co-localization experiments that the signals obtained for Rec114, Mei4 and Mer2 overlap. Co-immunoprecipitation experiments using meiotic cells also confirmed the fact that these proteins form a complex (J. Li, Hooker, and Roeder 2006; Maleki *et al.* 2007; Steiner, Kohli, and Ludin 2010). However, it has been shown than these three proteins cannot be purified together *in vitro*. It was demonstrated recently that only Rec114 and Mei4 can be copurified and it has been shown recently to form a 2:1 heterotrimer with two Rec114 and one Mei4. A yeast two hybrid experiment also showed that the N-terminal part of Mei4 is required for interaction with Rec114 (Maleki *et al.* 2007) and a XL-MS experiment showed that the C-terminal part of Rec114 interact with the N-terminal part of Mei4 (Figure 11). Mer2 has been shown to form a homotetrameric complex (Claeys Bouuaert, Tischfield, *et al.* 2021).

In fact, their mutual dependencies are not complete, suggesting that they could exist independently. Mer2 can bind to chromosomes even in the absence of Rec114 and Mei4. Rec114 can bind to chromosomes in the absence of Mei4, but the reverse relationship is not clear (J. Li, Hooker, and Roeder 2006; Maleki *et al.* 2007). Moreover, it has been shown that the phosphorylation of Mer2 is essential for the recruitment of Rec114 on the chromosomes (Panizza *et al.* 2011).

With regard to the binding of these proteins to DNA, *in vitro*, each sub-complexes have been shown to bind to DNA in a highly cooperative manner leading to an assembly of a structure comprising hundreds of proteins, referred to as a condensate. Indeed, atomic force microscopy (AFM) imaging, reveals the formation of clusters of proteins that form cooperatively in the presence of plasmid DNA (Figure 7). *In vitro*, fluorescence microscopy analyses show that Rec114-Mei4 and Mer2 complexes overlap to form mixed condensates (Claeys Bouuaert, Pu, *et al.* 2021).

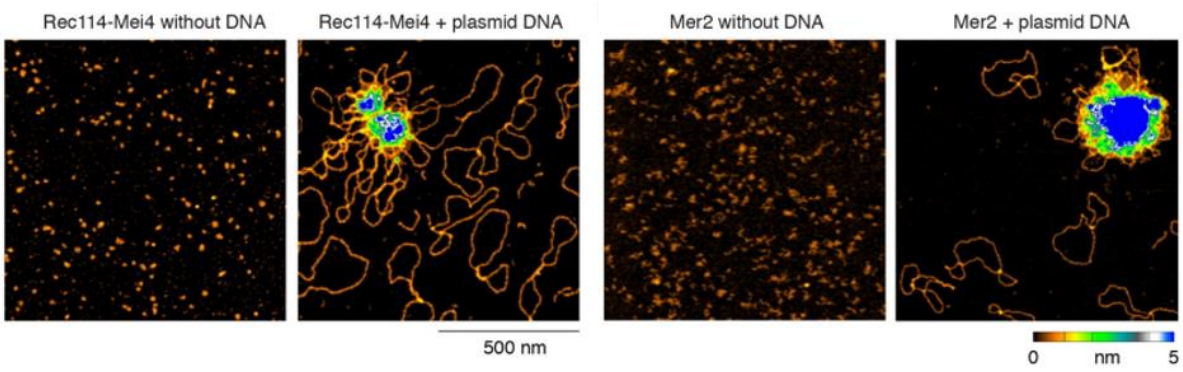

Figure 7: Condensation of RMM proteins. *AFM image showing in vitro condensation in the presence of Rec114-Mei4 complex (12 nM) in the presence of plasmid DNA (1 nM) (left) and in the presence of Mer2 (50 nM) and plasmid DNA (1 nM) (right) (Claeys Bouuaert, Pu, et al. 2021).*

The condensation of RMM proteins is thought to be induced by phase separation. This phenomenon leads to the formation of sub-cellular compartments that facilitate various biochemical reactions through local enrichment of specific biomolecules. Condensates are typically dynamic and can assemble and disassemble in a reversible way. Phase separation is driven by a physical process of supersaturation of a component that leads to the spontaneous formation of two distinct phases. Indeed, the interaction between these proteins is more thermodynamically favorable than the interaction between these proteins and the solvent in which they are found (So, Cheng, and Schuh 2021; Boeynaems *et al.* 2018). These condensates are formed by weak multivalent interactions between proteins and may also involve nucleic acids. These multivalent interactions may be due to folded proteins showing well-defined and/or flexible interaction sites or to the intrinsic disordered parts of the proteins (Boeynaems *et al.* 2018).

Based on this condensation phenomenon, our laboratory proposed that the function of the RMM complex in the formation of DSBs is to assemble the machinery necessary for double strand breaks along the meiotic chromosome axis.

The DSB proteins are connected through each other through an intricate interaction network (Arora *et al.* 2004). The Mer2 protein, when phosphorylated by a cyclin-dependent kinase (CDK) and by a Dbf4-dependent kinase (DDK), interacts with Rec114 and can associate with Spo11 and allow the formation of DSBs (Maleki *et al.* 2007; Henderson *et al.* 2006). It has also been shown *in vivo* that Rec114 and Mei4 interact with Rec102 and Rec104, proteins of the core complex. It is known that it is the N-terminal part of Rec114 (known as the Pleckstrin Homology (PH domain)) interacts with Rec102 and Rec104 and thus linking the RMM complex with the core complex (Claeys Bouuaert, Pu, *et al.* 2021). Mer2 also interacts with Xrs2, a protein belonging to the MRX complex. Moreover, Mer2 is recruited to the nucleoprotein axis by Hop1 and Red1 (composing the

nucleoproteic axis) but it is not known whether this is via a direct physical interaction (Lam and Keeney 2014).

The hotspots where DSBs frequently occur are often associated, in *S. cerevisiae*, with histones methylated regions (H3 lysine 4 trimethylate (H3K4 me 2/3)). However, these H3K4 me2/3 are present in the chromosomal loop formed during prophase I. These histones are captured by the Spp1 protein, which in turn interacts with Mer2, located at the level of the chromosomal axis. The tethered loop axis model suggest that Spp1-Mer2 interaction is a key point in capturing the loop and bringing it close to the axis, near Spo11, the catalytic unit, so that the DNA can be cut (Figure 8 and 9) (Acquaviva *et al.* 2013).

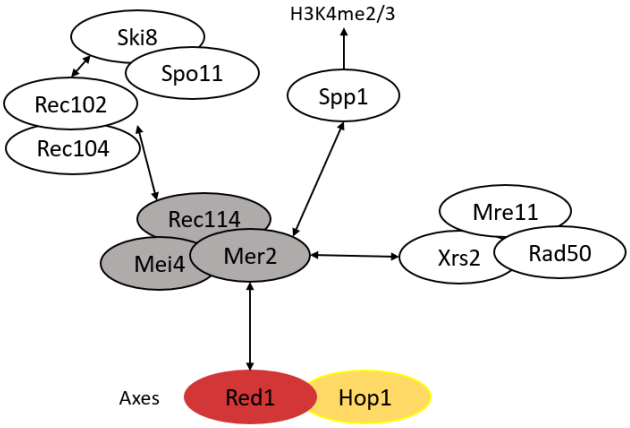

Figure 8: Interaction between the RMM complex and other proteins. *Diagram showing the interactions between proteins of the S. cerevisiae RMM complex and other proteins (inspired from Lam and Keeney 2014).*

This information taken together (Figure 8) leads to the following structural model: The condensate formed by the RMM complex recruits all the other proteins necessary for DSB formation at the nucleoprotein axis. This complex also indirectly tethers the loop to the chromosome axis to allow the DNA to be cut by Spo11 (Figure 9) (Yadav and Claeys Bouuaert 2021).

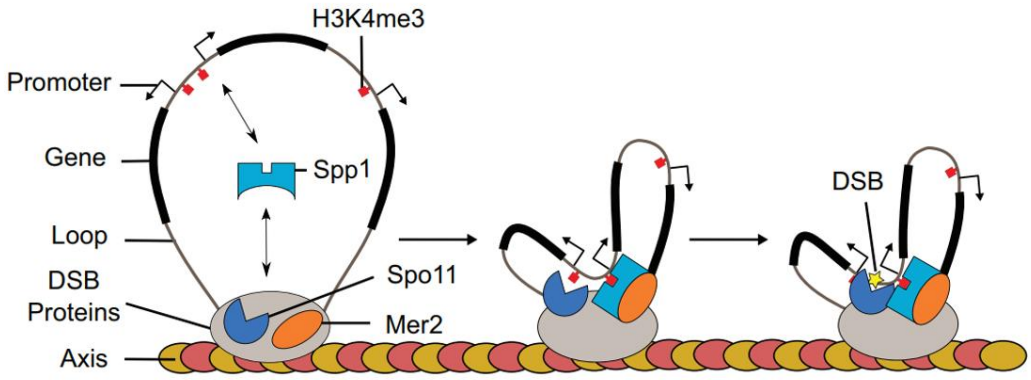

Figure 9: Capture of the loop by Spp1. *Representation of the capture of a chromosomal loop by the Spp1 protein in order to allow the catalytic unit, Spo11, present in the nucleoprotein axis to cut the double-stranded DNA. Spp1 is recruited by Mer2 (Yadav and Claeys Bouuaert 2021).*

4.3.1 Characterization of the structure of the Rec114-Mei4 complex

Rec114 is a 428 amino acid protein (accession number P32841) with a molecular weight of 48.7 kDa and is a highly disordered protein which makes it impossible to perform crystallography on this protein. Rec114 is highly divergent between species, but regions with recognizable sequence similarities (referred to as signature sequence motifs, SSM) can be identified. Rec114 has 6 SSMs in the N-terminal part domain and one in the C-terminal domain (Kumar, Bourbon, and de Massy 2010). The N-terminal part of mouse REC114 was crystallized and showed that it forms a Pleckstrin Homology (PH) domain, with a N-terminal α-helix sandwiched between two antiparallel β-sheets (Kumar *et al.* 2018). The PH domain of mouse REC114 is important for interaction with other proteins (ANKRD31 (protein essential for the formation of DSB in mice), TOPVIBL (SPO11-associated protein), and IHO1 (Mer2 ortholog))(Boekhout *et al.* 2019; Xu *et al.* 2023; Nore *et al.* 2022; Laroussi *et al.* 2023). Based on AlphaFold predictions, this structural fold seems to be conserved in yeast (Daccache *et al.* 2022).

Mei4 is a 408 amino acid protein (accession number P29467) with a molecular weight of 48.2 kDa. This protein is highly ordered and contains 6 SSMs (Kumar, Bourbon, and de Massy 2010). The AlphaFold representation of this protein shows 19 α-helices that form a HEAT-repeat like fold (Daccache *et al.* 2022).

The 2:1 heterotrimeric stoichiometry of Rec114-Mei4 has been confirmed with multiple methods, including Size-Exclusion Chromatography with Multi-Angle Light Scattering (SEC-MALS). This technique allows accurate measurement of molecular weights using a triple light-scattering detector at the outlet of the column. SEC-MALS analysis revealed a molecular weight for the complex of 114 kDa, similar to the theoretical value of 145.6 kDa (Claeys Bouuaert, Pu, *et al.* 2021).

The two Rec114 stoichiometry for one Mei4 can also be confirmed by the intensity of the Coomassie-stained blue on the gel from the untagged complex purification (Figure 10) (Claeys Bouuaert, Pu, *et al.* 2021).

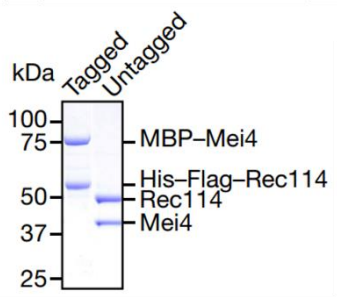

Figure 10 : Rec114-Mei4 complex purification. *SDS-PAGE gel showing purification of the Rec114-Mei4 complex, stained with Coomassie Blue (Claeys Bouuaert, Pu, et al. 2021).*

A spectral count of a mass spectrometry analysis (LC-MS/MS) of the purified and trypsin digested complex also confirmed a 2:1 ratio (Claeys Bouuaert, Pu, *et al.* 2021).

Furthermore, crosslinking followed by mass spectrometry (XL-MS) using a crosslinking agent (DSS) binding spatially close lysines showed crosslinks between the C-terminal part of Rec114 and the N-terminal part of Mei4. In addition, intermolecular self-links (i.e. crosslinks of identical residues) at the C-terminal part of Rec114 suggest that this domain homo-dimerizes (Figure 11).

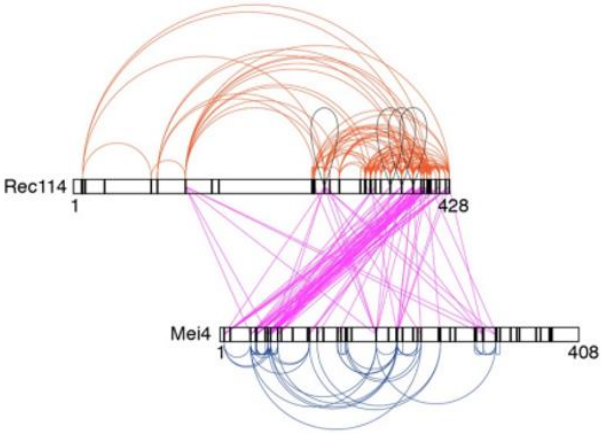

Figure 11: XL-MS results on the Rec114-Mei4 complex. *Representation of the crosslinks observed between Rec114 and Mei4 following XL-MS analysis with in red the crosslink between Rec114 residues, in magenta the crosslink between Rec114 and Mei4, in blue the crosslink between Mei4 residues and in black the crosslink between identical lysines (self-links) (Claeys Bouuaert, Pu, et al. 2021).*

The XL-MS analysis identified the interaction domain between Rec114 and Mei4. Co-expression of the Rec114 C-terminus (residues 375-428) and Mei4 N-terminus (residues 1-43) confirmed that these domains are sufficient to assemble heterotrimeric complex, referred to as the minimal complex. AlphaFold prediction suggests that the minimal complex forms a unique asymmetrical structure with an N-terminal α-helix from Mei4 inserted into a dimeric ring made up of α-helices from the C-terminal part of Rec114 (Figure 12).

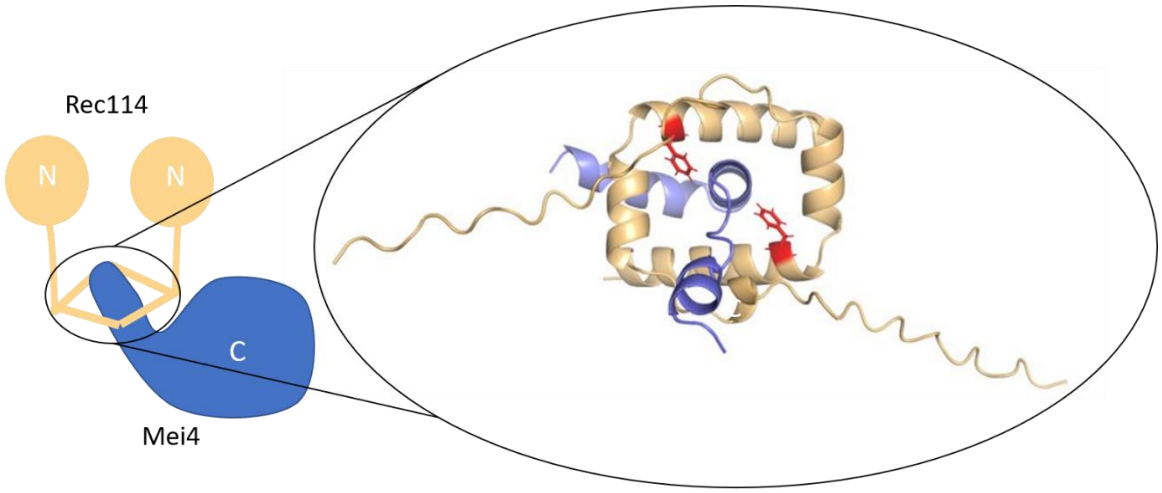

Figure 12: Representation of the Rec114-Mei4 complex. *On the left of the picture, schematic of the complex formed by two Rec114 and one Mei4. The C-terminal parts of Rec114 interacting with the N-terminal part of Mei4. On the right-hand side of the image, the AlphaFold representation of the trimeric Rec114-Mei4 minimal complex, with Mei4 in blue and Rec114 in beige. The amino acid shown in red is phenylalanine F411, an essential amino acid in the protein-protein interaction of this complex.*

Experiments on this minimal part of this complex were carried out (Daccache *et al.* 2022). A thermal-shift analysis showed that Mei4 stabilizes Rec114. In addition, the AlphaFold model was validated by NMR spectroscopy.

In addition, the model was confirmed by previous results showing that by mutating the phenylalanine F411 of Rec114 (shown in red in Figure 12) to alanine, and performing a Pulldown experiment, the protein interaction between Rec114 and MEI4 was abolished. Furthermore, the expression of this protein instead of Rec114 WT inhibited the formation of *in vivo* foci of Rec114 and DSBs, and led to spore death (Claeys Bouuaert, Pu, *et al*. 2021). These experiments validate the model because the phenylalanine F411 of Rec114 points towards Mei4. In addition, amino acids substitutions of K405, E416 of Rec114 and residues E16 and D18 of Mei4 to alanine lowers the aggregation temperature of the complex, indicating that the residues stabilize this complex (Daccache *et al.* 2022).

4.3.2 DNA binding function of the Rec114-Mei4 complex

In order to understand the condensation phenomenon observed with RMM proteins and DNA, several experiments were conducted to characterize the DNA binding properties of this complex *in vitro*.

Electrophoretic mobility shift assays (EMSA) were realized to show a DNA binding function of this complex *in vitro* and that measured a binding activity of the minimal Rec114-Mei4 complex by the appearance of complexes that cannot enter the gel (Figure 13).

It has also been shown that the minimal Rec114-Mei4 complex binds preferentially to a Holliday junction (HJ) type DNA branching structure. This suggests multiple binding sites in the complex that collaborate to form a stable nucleoprotein complex (Figure 13) (Daccache *et al.* 2022). This was expected based on the condensation activity of the complex, which is presumably driven by multivalent protein-DNA interactions (Claeys Bouuaert, Pu, *et al.* 2021).

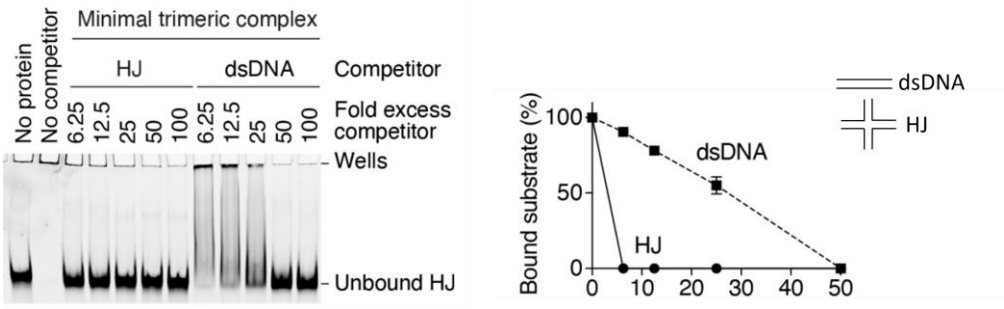

Figure 13: DNA binding by the minimal part of Rec114-Mei4 complex. *Competition assay of the minimal parts of Rec114-Mei4 complex (200 nM) binding to a fluorescent HJ substrate (10 nM) in the presence of unlabeled HJ substrates or dsDNA. Error bars are ranges from 2 independent experiments (most are too small to be visible) (modified from Daccache et al. 2022).*

Mutation of the amino acids R395, K396, K399 and R400 of Rec114 inhibits the interaction of the complex with DNA *in vitro* and inhibit DSB formation *in vivo*. When observing the location of these positively-charged amino acids, they arrange themselves into two separate patches that extend outward from the complex in opposite directions (Figure 14). This incidates that a single B-form DNA molecule could not simultaneously occupy its two binding sites (Daccache *et al.* 2022).

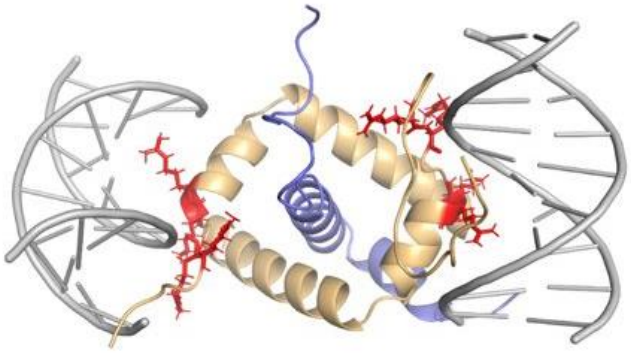

Figure 14: AlphaFold representation of the minimal complex of RM. *AlphaFold representation of the minimal complex formed by two Rec114 in beige and Mei4 in blue with two different DNA molecules. In red are the residues R395, K396, K399 and R400 of Rec114 (Daccache et al. 2022).*

4.3.3 Orthologs of *S. cerevisiae* Rec114-Mei4 complex

In the following sections, the current knowledge of the complex Rec114-Mei4 of *Mus musculus, Schizosaccharomyces pombe, Arabidopsis thaliana* and *Zea mays* will be developed.

By phylogenomic-directed BLAST analysis, orthologs of Mei4 and Rec114 could be identified in most animals, plants and fungi. As a starting point, here is a schematic of the sequences of different orthologs of the Rec114 and Mei4 proteins (Kumar, 2010). These proteins are highly divergent within the eukaryotic kingdom and pair-wise comparisons between these orthologous proteins in different organisms show a sequence identity of less than 10%. However, it can be noticed the presence of evolutionarily conserved SSMs (7 for Rec114 and 6 for Mei4) (Figure 15 and 16).

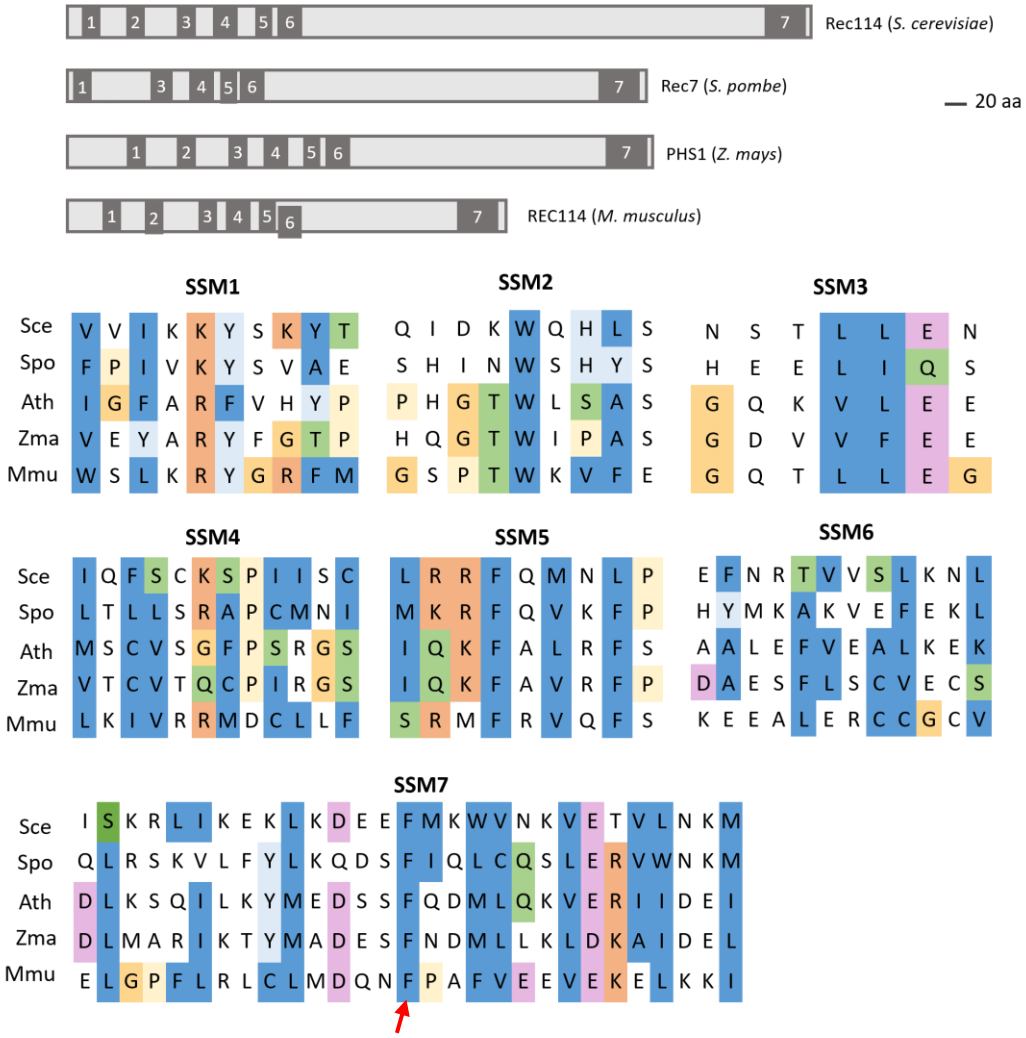

Figure 15: ***Conservation of Rec114. (Top)*** *Representation of the protein sequences orthologous to Rec114 with the relative position of the SSMs represented in dark grey.* ***(Bottom)*** *Sequence alignment of Rec114 orthologous proteins for the different SSMs found. Amino acids highlighted in color are those that are conserved. In SSM 7, we can see the conservation of phenylalanine, which plays a potential role in the interaction between Rec114 and Mei4, pointed by the red arrow. With Sce = Saccharomyces cerevisiae; Spo = Schizosaccharomyces pombe; Ath = Arabidopsis thaliana; Zma = Zea mays; Mmu = Mus musculus (inspired from Kumar, 2010).*

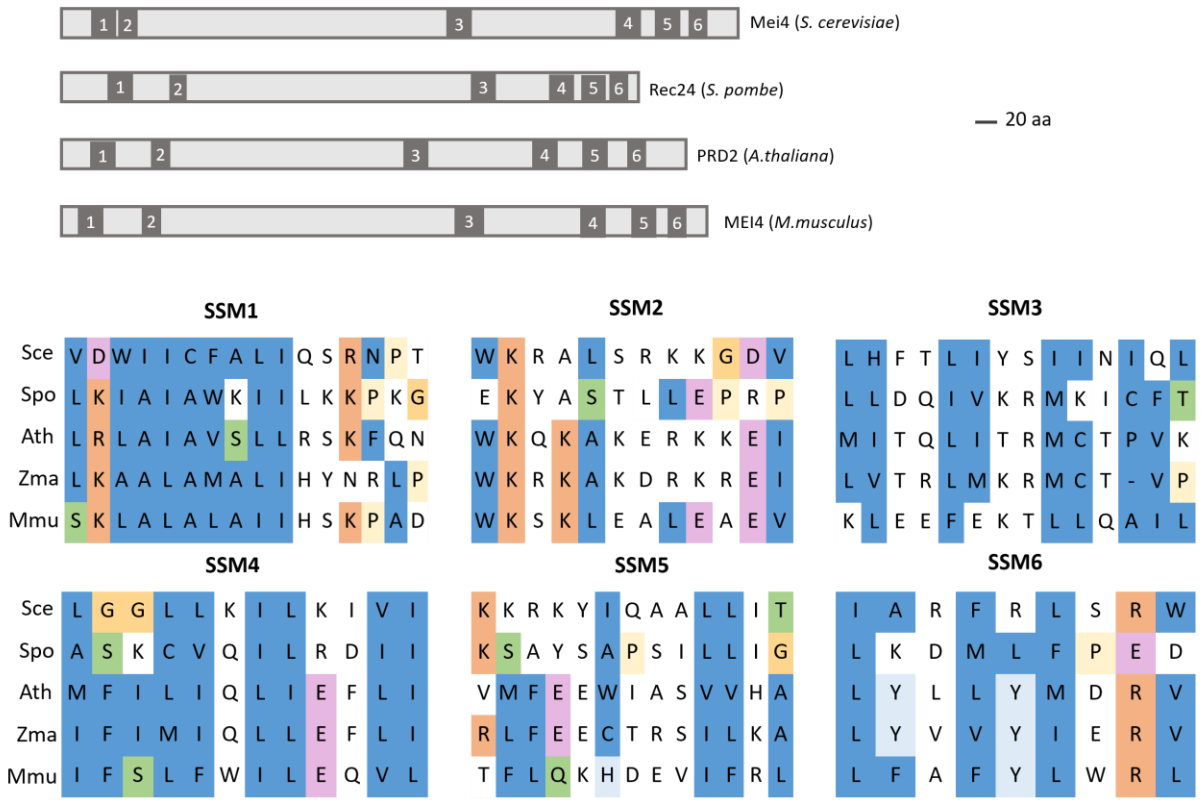

Figure 16: Conservation of Mei4. (Top) *Representation of the Mei4 orthologous protein sequences with the relative position of the SSMs represented in dark grey.* ***(Bottom)*** *Sequence alignment of Mei4 orthologous proteins for the different SSMs found. Amino acids highlighted in color are those that are conserved. With Sce = Saccharomyces cerevisiae; Spo = Schizosaccharomyces pombe; Ath = Arabidopsis thaliana; Zma = Zea mays; Mmu = Mus musculus (inspired from Kumar, 2010).*

Although the amino acid sequences are quite different between species, the structures of these proteins share similarities according to the AlphaFold prediction for all the orthologs of Rec114 discussed below (Figure 17A). The N-terminal part of these proteins is thought to share a conserved PH domain (A PH domain consist of two perpendicular anti-parallel β-sheets and a C-terminal amphipathic helix (Scheffzek and Welti 2012)) and the C-terminal part is composed of two to four α-helices. For the Mei4 orthologs, 4 HEAT-repeats, which comprise consecutive arrangements of helix A-turn-helix B motifs forming an α-solenoid structure with an approximate angle of 15° between each repeat (Andrade *et al.* 2001), are predicted as well as an N-terminal part composed of 3 to 5 helices (Figure 17B) (Daccache *et al.* 2022).

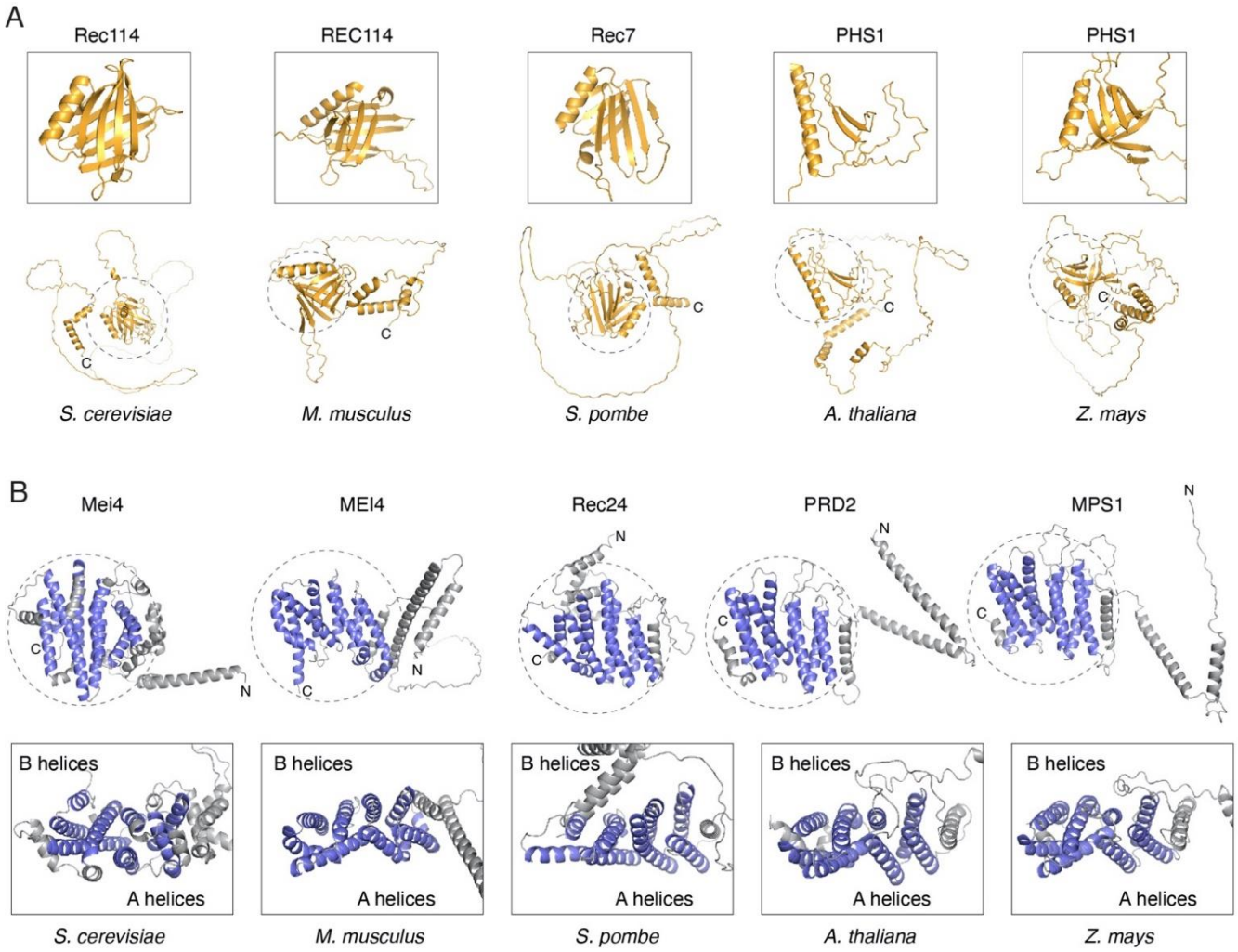

Figure 17: Conservation of Rec114 and Mei4 tridimensional structures. (A) *AlphaFold representation of the Rec114 orthologs shows that they all have a N-terminal PH domain (zoom in on the top of the image), a central IDR and C-terminal α-helix.* ***(B)*** *AlphaFold representation of Mei4 orthologs with a zoom on the C-terminal part below the image (Daccache et al. 2022).*

For all these orthologs, AlphaFold representations of the minimal complex showed that these complexes have a rather similar structure composed of a ring of two α-helices from the C-terminal part of the Rec114 ortholog into which an α-helix from the N-terminal part of the Mei4 ortholog is inserted (Figure 18). This is very similar to the three-dimensional structure of the minimal complex of *S. cerevisiae* (Figure 12). Indeed, the phenylalanine F411 of Rec114, whose mutation to alanine completely inhibits the *in vitro* interaction between Rec114 and Mei4 of *S. cerevisiae*, is probably conserved during evolution because, according to the AlphaFold models, it is always found at the interaction surface between the two different proteins of the complex (they are represented in pink in Figure 18). We will see if this amino acid is important for the interaction between the two proteins of these complexes in the section result and discussion of this master thesis.

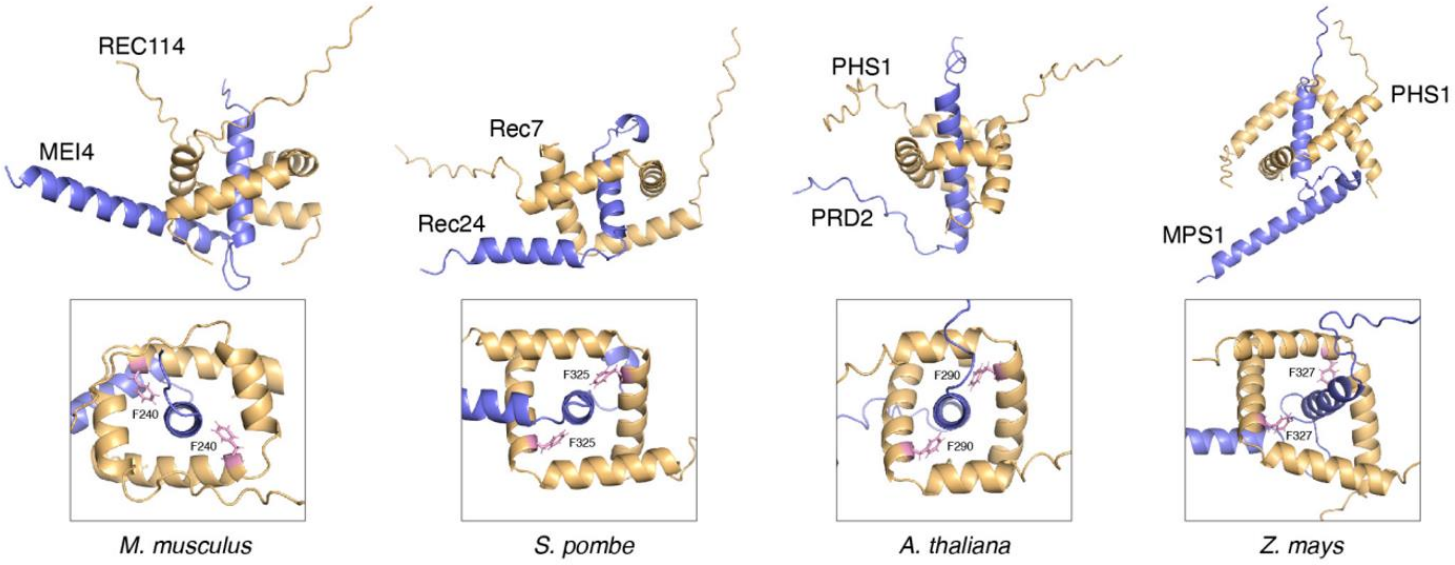

Figure 18: AlphaFold representation of the minimal RM complex of orthologs. *This AlphaFold representation of the minimal complex of other organisms shows that the phenylalanine important in the interaction between Rec114 and Mei4 in S. cerevisiae appears to be conserved. It is shown in pink in this figure (Daccache et al. 2022).*

Like the *S. cerevisiae* complex, these orthologs also have positively charged amino acid patches pointing outwards from the complex, which could explain a possible multivalent binding to DNA that we will test. This can be seen in Figure 19 below where these amino acids (lysines and arginines) are shown in red.

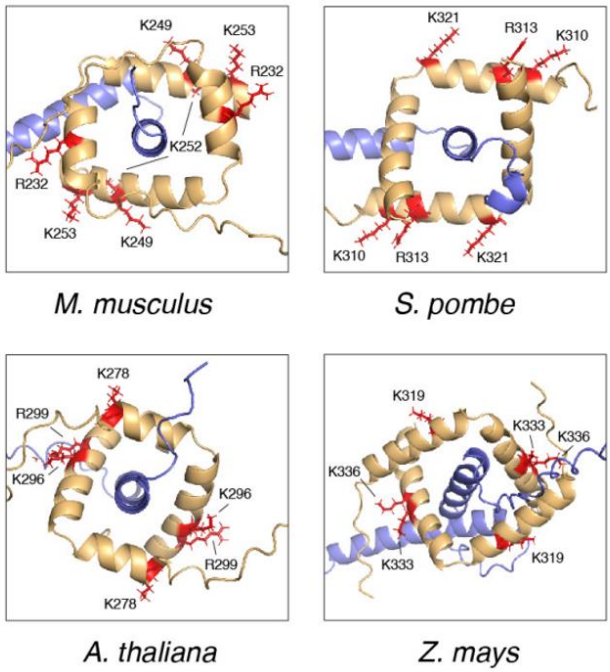

Figure 19: Positively charged amino acids that point away from the complex. *AlphaFold representation of the minimal complexes of the Rec114-Mei4 orthologues. In red are positively charged amino acids (lysine, arginine) that point outwards from the complex and may be involved in the multivalent binding of the complex to DNA (Daccache et al. 2022).*

4.3.3.1 Mouse REC114-MEI4 complex

REC114 is a 259 aa protein (accession number: NP_082874.1) and MEI4 is a 389 aa protein (accession number NP_780422.1) (Laroussi *et al.* 2023).

Regarding the structure of this complex, it was shown *in vitro* by co-purification that a direct interaction between the C-terminal part of REC114 (residues 226-254) and the N-terminal part of MEI4 (residues 1-43). The stoichiometry of the complex has been shown to be 2:1 with 2 REC114 and one MEI4 (Laroussi *et al.* 2023).

Concerning the function of this complex, it has been shown *in vivo* that REC114 and MEI4 co-localize with IHO1, Mer2 homolog in mouse, at the chromosomal axes during meiosis prophase (Nore *et al.* 2022). It was also demonstrated that IHO1 binds the REC114-MEI4 complex *in vitro* (Laroussi *et al.* 2023).

4.3.3.2 *S. pombe* Rec7-Rec24 complex

Rec7, the ortholog of Rec114 is a 339 amino acid protein (accession number: AAA35333.2) and Rec24 is the ortholog of Mei4 and consists of 350 amino acids (accession number: NP_594817.1). Rec7 shares 9% amino acid identity with Rec114 and Mei4 shares 7% amino acid identity with Rec24. The low sequence conservation observed even between yeast shows a rapid evolution of these proteins (Bonfils *et al*. 2011).

Rec7 localizes to chromosomes and it has been shown to interact with Rec24 *in vivo* by yeast two hybrid, and more specifically that the C-terminal part of Rec7 interacts with Rec24. A conserved amino acid of Rec7, phenylalanine 325 (shown in pink in Figure 19) is essential to maintain this interaction *in vivo* and to maintain the formation of DSBs (Steiner, Kohli, and Ludin 2010).

No *in vitro* studies of the structural biochemistry and DNA binding function have yet been performed.

4.3.3.3 The Rec1114-Mei4 complex in plants

No *in vitro* biochemical studies have yet been performed on these proteins: no structural studies and their DNA binding function has not been shown *in vitro*.

***Arabidopsis thaliana* PHS1- PRD2 proteins**

PHS1 is a 310 amino acid protein and is the homolog of Rec114 (accession number: OAP13975.1). This protein is not essential for DSBs formation but is essential

for DNA double-strand break repair and homologous pairing (de Massy 2013). This suggests that the conservation of PHS1 function has been lost in dicotyledonous plants (divergence of about 150 Myr from monocotyledons such as mays) (Vrielynck *et al.* 2021). But it has been shown *in vivo* that PHS1 interacts directly with PRD2 (Vrielynck *et al.* 2021).

PRD2 is a 202 amino acid protein and is the homolog of Mei4 (accession number: OAO95537.1). This protein does not show a functional domain meaning that the structural domain was not predicted before but is still essential for DSBs formation (Y. Li *et al.* 2021). PRD2 interacts with the chromosomal axis.

In contrast to the organisms discussed so far, the RMM protein family in *A. thaliana* appears to be formed by the PHS1-PRD2 complex and the DFO protein, which has no ortholog in non-plant organisms. No direct interaction was detected by yeast two hybrid between PRD3 and the PHS1-PRD2 complex (Thangavel *et al.* 2023; Vrielynck *et al.* 2021).

***Zea mays* PHS1- MPS1 proteins**

PHS1, or poor homologous synapsis 1 protein, is the ortholog of Rec114, has 347 aa and is an essential protein for DSB formation in *Zea mays* (accession number: AAQ84720.1) (Vrielynck *et al.* 2021).

MPS1, or Multipolar Spindle 1, the ortholog of Mei4, has 391 residues (accession number: PWZ29589.1)(Kumar, Bourbon, and Massy 2010). No study has yet shown its role in the formation of DSBs, or its interaction with PHS1.

4.3.4 Characterization of the Mer2 structure

Mer2 is a protein consisting of 314 amino acids (accession number: CAA60944.1) and has a molecular weight of 35.712 kDa. The protein contains two SSMs and has a central coiled-coil domain that tetramerizes and is flanked by a disordered N- and C-terminal portion (Figure 20).

Figure 20: Domain organization of Mer2 *with an intrinsically-disordered regions in white and a central part that tetramerizes forming a coiled coil in dark grey.*

An AlphaFold representation of these homotetrameric coiled-coils domain predicts that they are four α-helices arranged parallel to each other (Figure 21) (Daccache *et al.* 2022).

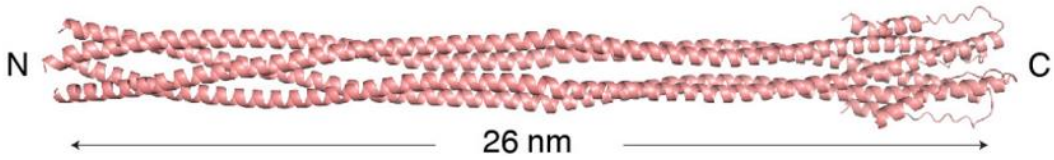

Figure 21: Tetrameric coiled-coil domain. *AlphaFold model of the coiled-coil domain formed by 4 coiled-coil domains of Mer2. The size of this coiled coil is predicted to be 26 nm (Daccache et al. 2022).*

However, the confidence score in the center of this coiled-coil is low and the AlphaFold prediction of the full-length protein suggests that a disordered region may exist in the center of the coiled coil involving residues from 110 to 160, which could give the coiled coil some flexibility (Figure 22) (Daccache *et al.* 2022).

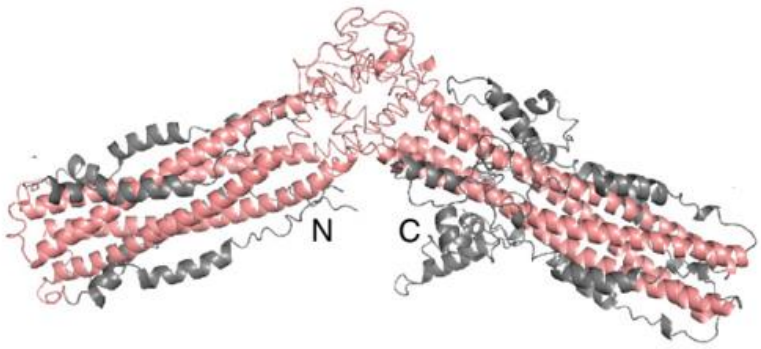

Figure 22: Mer2 full length*. AlphaFold model of the Mer2 full length protein. The coiled-coil part is shown in pink and the N- and C-terminal regions are shown in grey. A disordered region appears in the center of the coiled coil (modified from Daccache et al. 2022).*

4.3.5 DNA binding of Mer2

As with the Rec114-Mei4 complex, it was demonstrated via an EMSA that Mer2 binds to DNA and that, again, a branched HJ substrate is preferred (Figure 23). This suggests multiple binding sites in the complex that collaborate to form a stable nucleoprotein complex (Daccache *et al.* 2022).

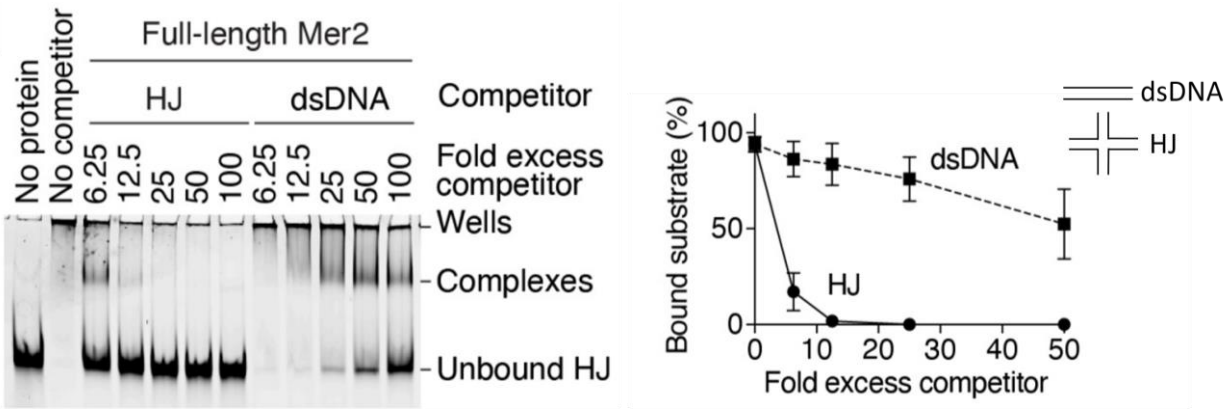

Figure 23: DNA binding by Mer2. *Competition assay on Mer2 (200 nM) binding to a fluorescent HJ substrate (10 nM) in the presence of unlabeled HJ substrates or dsDNA. Error bars are ranges from 2 independent experiments (most are too small to be visible) (modified from Daccache et al. 2022).*

The coiled-coil domain is necessary and sufficient for this DNA binding. However, the most important part of Mer2 for interaction with DNA is located outside the coiled-coil, at the C-terminal part of the protein (Daccache *et al.* 2022).

4.3.6 Orthologs of *S. cerevisiae* Mer2

Here is a diagram of the sequences of the different Mer2 orthologs that will be studied in this thesis (Figure 24) (Tessé *et al.* 2017).

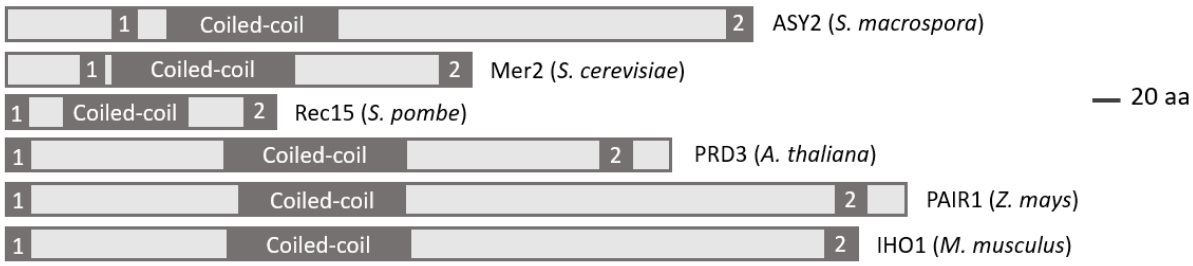

Figure 24: Mer2 orthologs*. Sequence diagram of Mer2 orthologous proteins showing the relative position of the coiled-coil domain (shown in dark grey), SSM1 in blue and SSM2 in dark grey (inspired from Tessé et al. 2017) .*

Following a bioinformatics analysis and as represented on the diagram above, you can see the presence of two conserved signature sequence motifs (SSMs) on either side of the coiled-coil domain of each protein. These proteins have been shown to be required for meiotic DSB formation in their respective organism (Figure 25) (Tessé *et al*. 2017).

Figure 25: Conservation of Mer2. *Sequence alignments of SSM1 and SSM2 of different Mer2 orthologous proteins with Sce =Saccharomyces cerevisiae; Spo = Schizosaccharomyces pombe; Smac = Sordaria macrospora, Ath = Arabidopsis thaliana; Mmu = Musculus (inspired from Tessé et al. 2017).*

AlphaFold predicts that these coiled-coils form parallel homotetrameric complexes, like Mer2 (Figure 26). The confidence score of these predictions are high (Daccache *et al.* 2022).

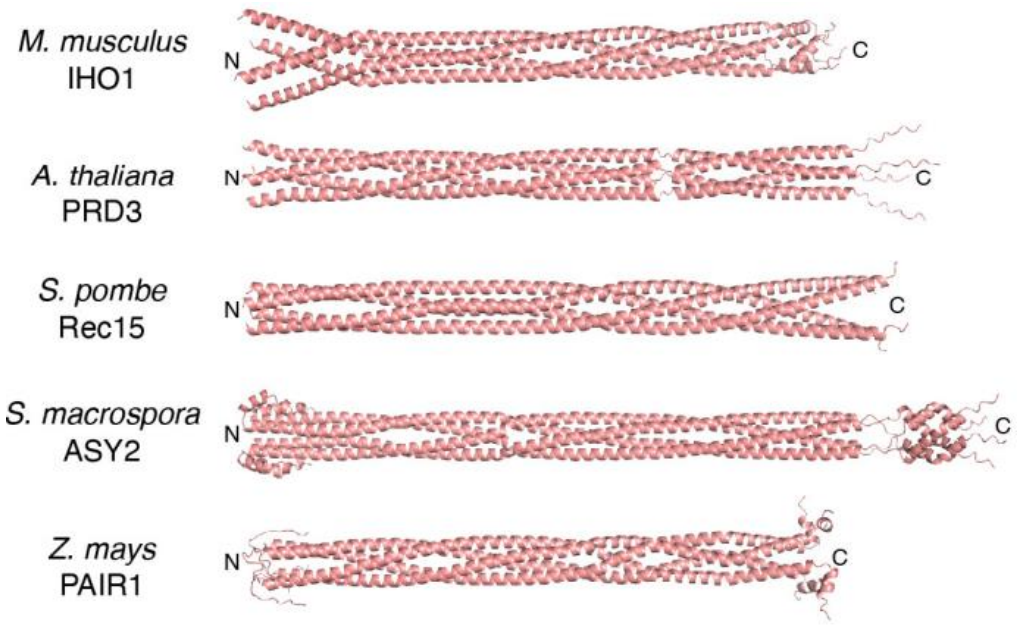

Figure 26: Coiled-coil domains of Mer2 orthologs. *AlphaFold models of homotetrameric coiled-coil domains of M. musculus IHO1 (residues 109-267, length 2.15 nm), A. thaliana PRD3 (residues 120-270, length 2.25 nm), S. pombe Rec15 (residues 1-160, length 2.3 nm), S. macrospora ASY2 (residues 55-275, length 2.85 nm), and Z. mays PAIR1 (residues 140-310, length 2.15 nm) (Daccache et al. 2022).*

4.3.6.1 Mouse IHO1 protein

IHO1 (accession number: NP_001366587.1), or interactor of HORMAD1 protein 1, is a 574-residues protein and like Mer2, this protein is highly disordered, except for a central region where a coiled-coil domain is observable.

It has been shown *in vivo* that IHO1 is located at the chromosomal axis in prophase I (Stanzione *et al.* 2016).

No *in vitro* DNA binding studies have yet been performed.

4.3.6.2 *S. pombe* Rec15 protein

Rec15 has 180 amino acids (accession number: NP_595887.1) (Doll *et al.* 2005). No structural analysis has been carried out, nor has its DNA binding function been studied *in vitro*.

4.3.6.3 S. *macrospora* ASY2 protein

ASY2 (protein SMAC_02785) is a protein of 504 amino acids (accession number: XP_003348288.1). *In vivo*, it has been shown that this protein localizes all along the chromosomes during prophase I of meiosis (Tessé *et al.* 2017).

No structural analysis has been carried out, nor has its DNA binding function been studied *in vitro*.

4.3.6.4 Plants Mer2 orthologs

No *in vitro* studies of the structure and DNA binding function of these proteins has yet been performed.

A. thaliana PDR3 protein

PRD3 (accession number: OAP15284.1), a protein of 449 residues, does not interact directly with the Rec114-Mei4 complex and it has been shown *in vivo* that PRD3 forms clusters along chromosomes (Vrielynck *et al.* 2021).

Z. mays PAIR1 protein

PAIR1 is a protein of 590 residues (accession number: XP_008649842.1).

III Objectives

In order to study the evolutionary conservation of structure and the DNA-binding activity of the RMM complex of proteins, the aims of my master thesis are divided into two main parts.

The first part consists in studying the structure and the function of proteins orthologous to *S. cerevisiae* Rec114-Mei4. In this first part, several experiments will be carried out on the minimal parts of the proteins of *Mus musculus* ($REC114^{210-259}$- $MEI4^{1-58}$), *Schizosaccharomyces pombe* ($Rec7^{289-339}$- $Rec24^{1-50}$), *Zea mays* ($PSH1^{297-347}$- $MPS1^{1-87}$) and *Arabidopsis thaliana* ($PSH1^{260-310}$- $PRD2^{1-50}$). In order to do that, we will focus on two different sub-aims. On the one hand, the analysis of the structure of these complexes by analyzing interaction domains between the two proteins of the complex and by identifying the amino acids that are important for the interaction between the C-terminal part of the orthologs of Rec114 and the N-terminal part of the orthologs of Mei4 in order to test AlphaFold prediction of these complex shown in Figure 18. On the other hand, the DNA binding capacity *in vitro* of these minimal parts will be studied.

The second aim is to characterize the structure and the function of Mer2 orthologs including: IHO1 (*Mus musculus*), Rec 15 (*Schizosaccharomyces pombe*), PAIR1 (*Zea mays*), PRD3 (*Arabidopsis thaliania*) and ASY2 (*Sordaria macrospora*). This aim will be divided in two parts. The first concerns the study of the conservation of the parallel homotetrameric structure predicted by AlphaFold (Figure 26) and the second concerns the characterization of the DNA binding function *in vitro* and identifying the amino acids involved in the protein-DNA interaction. For these proteins, only the coiled-coil domain will be purified and studied.

IV Results

1. Orthologs of the Rec114-Mei4 complex

1.1 Conservation of the structure of minimal complex of RM

In order to partially validate the structural models proposed by AlphaFold for the minimal complexes orthologous to the Rec114-Mei4 complex of *S. cerevisiae*, we have identified amino acids that could play a role in the interaction between the two proteins. Indeed, in Figures 27 below, hydrophobic amino acids from the Rec114 orthologous protein pointing inwards from the complex to the Mei4 orthologous protein are represented in pink. For the mouse complex, the amino acids identified as potentially important in the protein-protein interaction were identified are as follows: F230, F240 and F243. For the Rec7-Rec24 complex of *S. pombe*, the amino acids selected are: Y320, F325 and L328. For the *A. thaliana* complex formed by PHS1-PRD2, the residues are Y284, F290 and L294. And finally, for the *Z. mays* complex formed by PHS1-MPS1: F327, L334 and I338. The phenylalanines marked in magenta above are evolutionarily conserved and correspond to phenylalanine F411 of Rec114 of *S. cerevisiae*, which by mutating it, totally inhibits the interaction between the proteins of the complex (Claeys Bouuaert, Tischfield, *et al.* 2021).

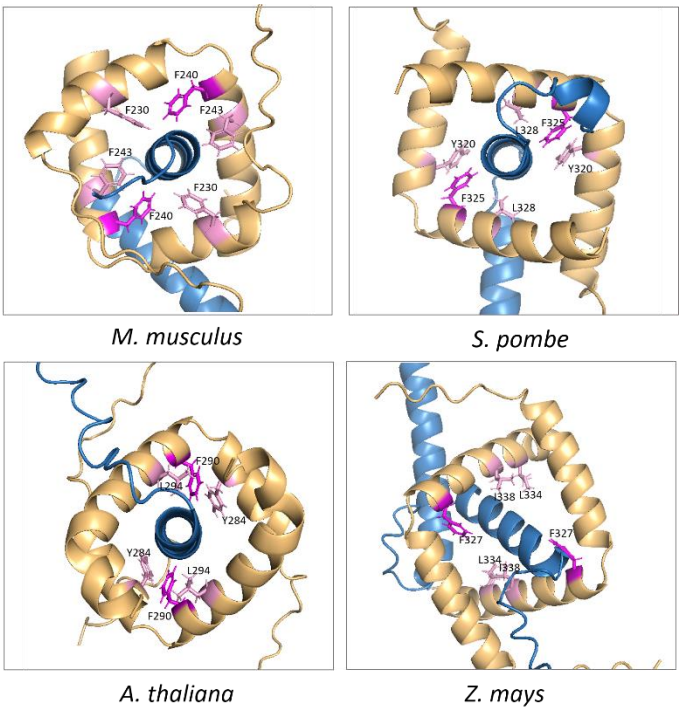

Figure 27: Interaction domains between Rec114 and Mei4 orthologous. *AlphaFold representation of the minimal domains of the 2:1 stoichiometry complex formed by the orthologs of Rec114-Mei4 including mouse REC114-MEI4 complex, the S. pombe Rec7-Rec24, the A. thaliana PHS1-PRD2 and the Z. mays PHS1-MPS1 complexes. In magenta is a phenylalanine conserved in all sequences (F411 in S. cerevisiae). In pink are represented other hydrophobic amino acids pointing towards the interior of the complex.*

In order to test the AlphaFold model of the mouse minimal complex, we carried out a Pulldown on tagged proteins where we mutated a single amino acid at a time: F230, F240 or F243 in alanine. The impact of the mutation on the interaction between the two proteins can be measured by comparing the ratio of the intensity of the MEI4 band to the intensity of the REC114 band visible on the SDS-PAGE gel of the Ni-NTA column elution between the WT and the mutant. If this ratio decreases, this means that the interaction has been disrupted by the mutation. As you can see from the gel shown in Figure 28, only the F240 phenylalanine mutation had a major negative impact on the interaction between REC114 and MEI4.

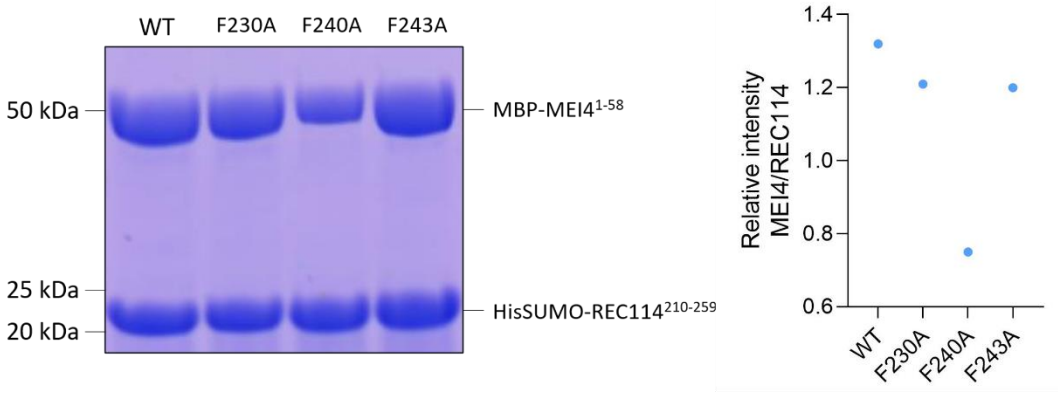

Figure 28: Pulldown of single mutations on the mouse REC114-MEI4 minimal complex. *SDS-PAGE gel of Pulldown elution on a Ni-NTA column. WT REC114$^{210-259}$ and MEI$^{1-58}$ proteins were tested, where REC114 F230, F240 or F243 were mutated to alanine.*

By making triplicates, it is shown that for the *M. musculus* complex, mutation of the phenylalanine F240 to alanine inhibits the REC114-MEI4 interaction by 25% (Figure 30). The interaction between proteins can be quite strong, so a simple mutation may not be able to completely inhibit it. As a result, mutations in the three amino acids tested would have a greater impact on protein interaction, we tested a Pulldown on proteins where F230, F240 and F243 are mutated to alanine. And in fact, the triple mutation inhibits the protein-protein interaction by 70% (Figure 30A). While this work was underway, another study has shown in a Pulldown experiment that by mutating the following amino acids of REC114 to glutamate, the interaction between the two proteins is inhibited: F240, V244 and V247 as well as the A16 mutation in MEI4 (Figure 29) (Laroussi *et al.* 2023). These mutagenesis validated that these amino acids are involved in the interaction between proteins and that they point well inside the complex as predicted by the model.

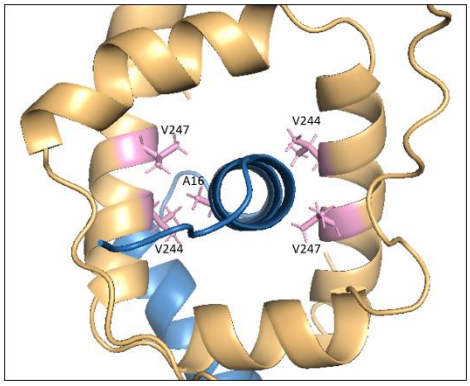

Figure 29: AlphaFold representation of mouse mRM. *In pink are represented the amino acids V244 and V247 of REC114 and the amino acid A16 of Mei4, whose mutation has been shown to inhibit protein interaction by Laroussi et al (2023). These amino acids point towards the interior of the complex.*

In the *S. pombe* Rec24-Rec7 complex, the F235 phenylalanine mutation is not sufficient to significantly decrease the interaction between the proteins of the complex (Figure 31B). However, *in vivo*, it has been shown that mutation of this amino acid completely inhibits DSB formation and interaction between Rec114 and Mei4 (Steiner, Kohli, and Ludin 2010). But the triple mutation of amino acids Y320, F235 and L328 inhibits the interaction by 80% (Figure 30B). The same effect is observed in the *A. thaliana* complex: The phenylalanine F290 mutation does not significantly impact the interaction between the two proteins of the complex, whereas the triple mutation of amino acids Y284, F290 and L294 inhibits the interaction between PHS1 and PRD2 by 80% (Figure 30C). The reduced interaction between proteins orthologous to Rec114 and orthologous to Mei4 caused by this mutagenesis directed towards hydrophobic amino acids pointing towards the interior of the complex supports the three-dimensional structure model predicted by AlphaFold.

For the *Z. mays* complex, the phenylalanine F327 mutation surprisingly increases the interaction between PHS1 and MPS1 by 32%. The triple mutation shows no significant difference in the interaction between the two proteins of the complex compared to the WT proteins (Figure 30D). Due to this mutagenesis, it has not been shown that these residues are involved in the interaction between PHS1 and MPS1. However, the confidence score for the three-dimensional structure of this complex given by AlphaFold is lower than that of the other complexes tested which tells us that we can have less confidence in it.

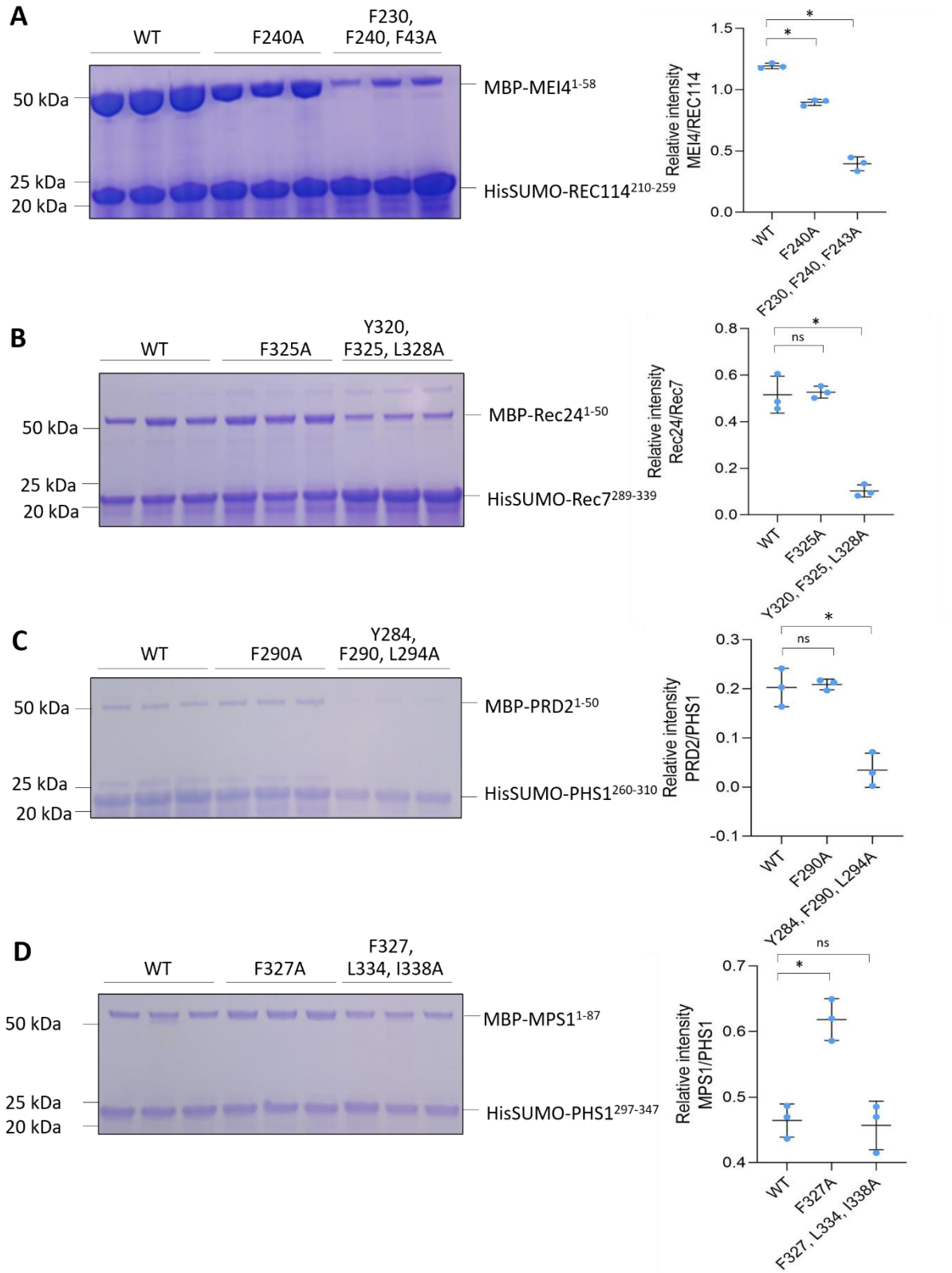

Figure 30: Pulldown result. *On the left of the image, pictures of the SDS-PAGE gels of the Ni-NTA column elution allowing to measure the impact of mutation brought to the Rec114 orthologue on the interaction between the two proteins. On the right, a graph reporting these effects. Error bars are ranges from three replicates and with * = significant difference with an α = 0.01, ns = not significant defined by a Student test. It is shown that we cannot reject the fact that these data follow a normal distribution (via a Shapiro-Wilk test) and that they have equal variances (via a Fisher test). Pulldown on* ***(A)*** *$REC114^{210-259}$-$MEI4^{1-58}$ (mouse),* ***(B)*** *$Rec7^{289-339}$-$Rec24^{1-50}$ (S. pombe),* ***(C)*** *$PHS1^{260-310}$-$PRD2^{1-50}$ (A. thaliana)* ***(D)*** *$PHS1^{297-347}$-$MPS1^{1-87}$ (Z. mays).*

1.2 Conservation of DNA binding of the RM complex

In order to test the DNA binding of the minimal RM complex orthologs, EMSAs were performed on purified tagged complexes which were overexpressed in *E. coli* (Figure 31).

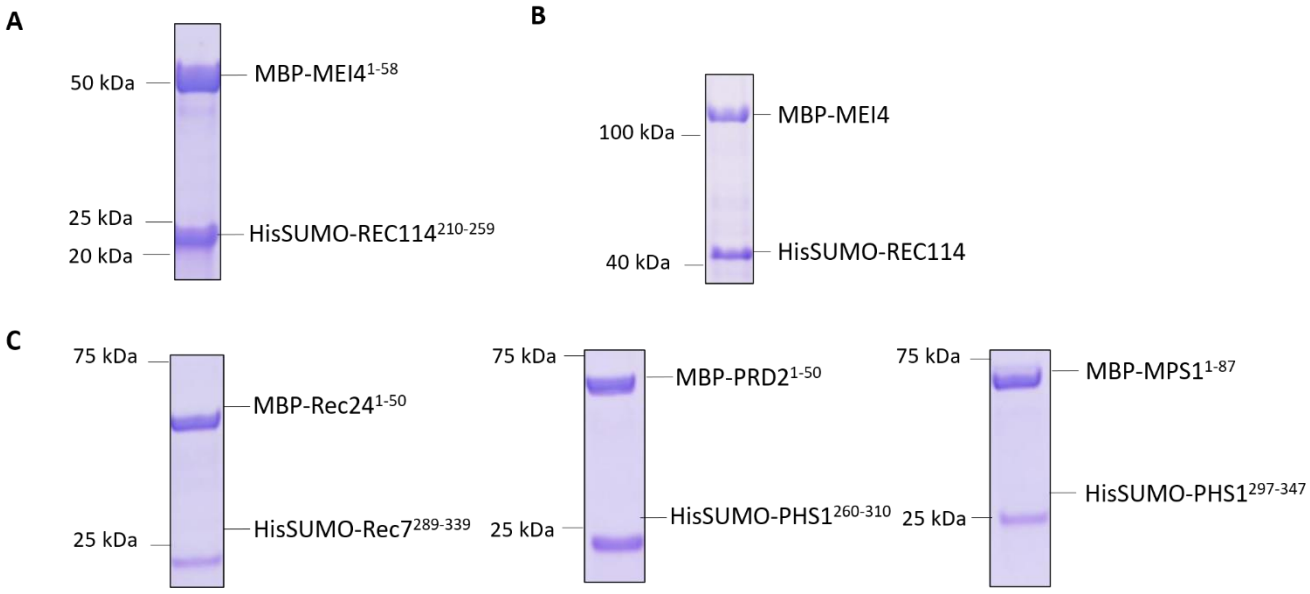

Figure 31: SDS-PAGE gel of protein purification of RM orthologs. (A) *Purification gel of the minimal REC114-MEI4 complex from M. musculus.* ***(B)*** *Purification of the REC114-MEI4 full-length complex from M. musculus.* ***(C)*** *Purification of the minimal complexes from S. pombe (Rec7-Rec24), A. thaliana (PHS1-PRD2) and Z. mays (PHS1-MPS1).*

In order to test the binding of the Rec114-Mei4 orthologous complex of *S. cerevisiae* to DNA *in vitro*, EMSAs were performed. An EMSA consists of migrating a mix of fluorophore-labelled DNA and proteins in a polyacrylamide gel. As the DNA is negatively charged, it will migrate through the gel more or less quickly depending on its size or the size of the complex it forms with proteins. Indeed, the larger the complex formed, the less the labelled DNA will migrate. After the gel is revealed, it is possible to measure whether there is an interaction between the labelled DNA and the proteins tested.

To measure the preference of proteins for a particular DNA substrate, competition EMSA were realized by, to the starting reaction mix, adding a fluorophore-labelled DNA substrate (HJ) and a large excess of an unlabelled "competitor" substrate (HJ or dsDNA). If the proteins prefer the competitor substrate to the initial one, the labelled one will be able to migrate freely and far in the gel and the protein will form a complex with the competitor and then migrate more slowly and less far in the gel. After gel revelation, if complexes are visible, this means that the proteins have a preference for the fluorophore-labelled substrate rather than for the competitor.

Dima Daccache has shown that the full-length REC114-MEI4 proteins, the *M. musculus* complex, bind to DNA *in vitro* and that this complex binds to DNA and has a >20-fold preference for a cross-shaped HJ substrate rather than a linear substrate (Figure 32) (Daccache *et al*. 2022). That suggests that this complex contains multiple DNA-binding sites, like the Rec114-Mei4 complex of *S. cerevisiae*.

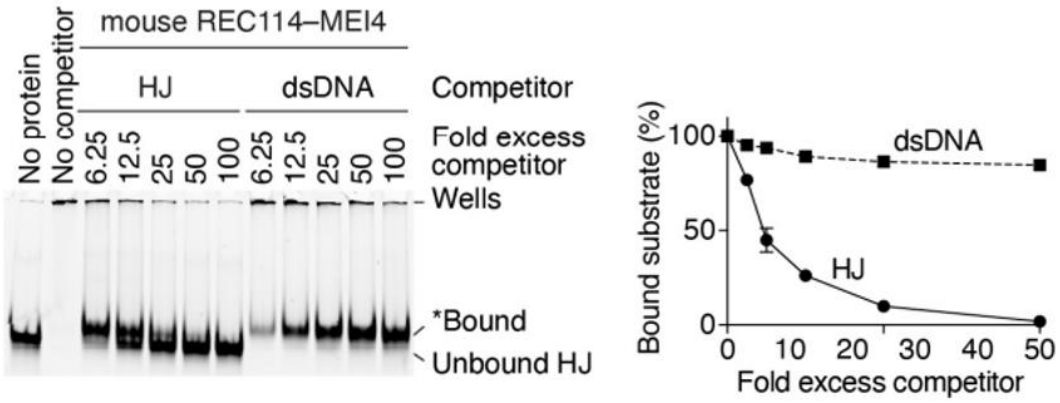

Figure 32: DNA binding by REC114-MEI4. *Competition assay of tagged FL mouse REC114-MEI4 complex (1000 nM) binding to a fluorescent HJ substrate (10 nM) in the presence of unlabeled HJ substrates or dsDNA. The band labeled "*Bound" migrates close to the position of the unbound substrate. It is likely due to the rapid dissociation of REC114-MEI4 from the substrate at the start of the electrophoresis. Error bars are ranges from 2 independent experiments (most are too small to be visible) (Daccache et al. 2022).*

Concerning the binding of the minimal tagged complexes of RM orthologs to HJ, it can be noted that, for EMSA performed in a TAE buffer, only the minimal complex of *A. thaliana* (PHS1-PRD2) is sufficient to bind DNA as complexes are visible in the gel and in the wells. Indeed, in presence of 800 nM of PHS1-PRD2, up to 50% of DNA was bound (Figure 33B). Therefore, the DNA-binding properties of the minimal complex Rec114-Mei4 of *S. cerevisiae* is conserved only for the *A. thaliana* minimal complex (Figure 33). Indeed, for the other complexes tested (*M. musculus* REC114-MEI4, *S. pombe* Rec7-Rec24 and *Z. mays PHS1-MPS1*), no DNA binding was detected on the gel migrated with a TAE buffer (Figure 33).

However, according to AlphaFold predictions, positively charged residues form patches that point outwards from the proteins, showing that they can potentially bind to DNA (Figure 19). We know that EMSA is a technique in which the DNA-protein interaction must survive migration in a gel. We therefore tested different techniques to try to show their binding to DNA *in vitro* trying to destabilize the possible protein-DNA interaction as little as possible.

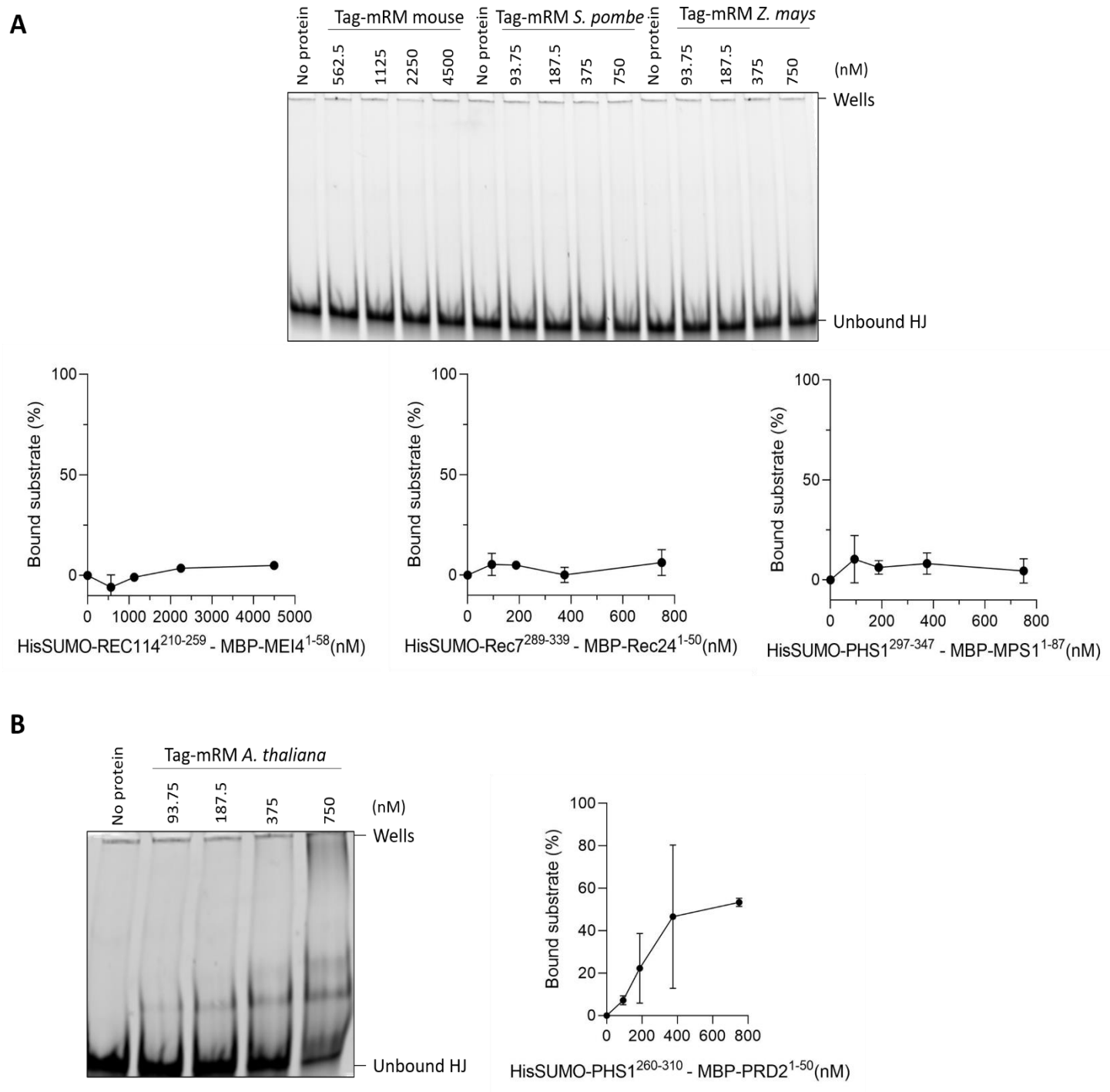

Figure 33: DNA-binding activity of the orthologs of RM. (A) *Gel-shift assay of different concentrations of tagged RM minimal complexes of the mouse (REC114-MEI4), S. pombe (Rec7-Rec24) and Z. mays (PHS1-MPS1) binding to a fluorescent HJ substrate (10 nM). Error bars are ranges from two independent experiments (some are too small to be visible).* ***(B)*** *Gel-shift assay of different concentration of the minimal tagged complex RM of A. thaliana (PHS1-PRD2). Error bars are ranges from 2 independent experiments (some are too small to be visible).*

Firstly, we tested adding 5 mM Mg^{2+} to the samples, 1.5 mM in the gel and in the electrophoresis buffer, as Mg^{2+} is known to improve some protein-DNA interactions (Guéroult *et al.* 2012). Given that DNA is negatively charged and Mg^{2+} is positively charged, it can be at the protein-DNA interface and help the proteins to bind the DNA. As you can see, these minimal complexes still failed to bind to DNA (Figure 34A). We then tested another TAE buffer in order to lower the ionic strength of the medium. Its composition is the following: 0.1 mM EDTA, 4 mM Tris at pH 8.3, 4 mM NaOAc, 0.2 mM MgOAc. The gel migrated for 4 h at 100 V. However, still no DNA binding was detected (Figure 34B).

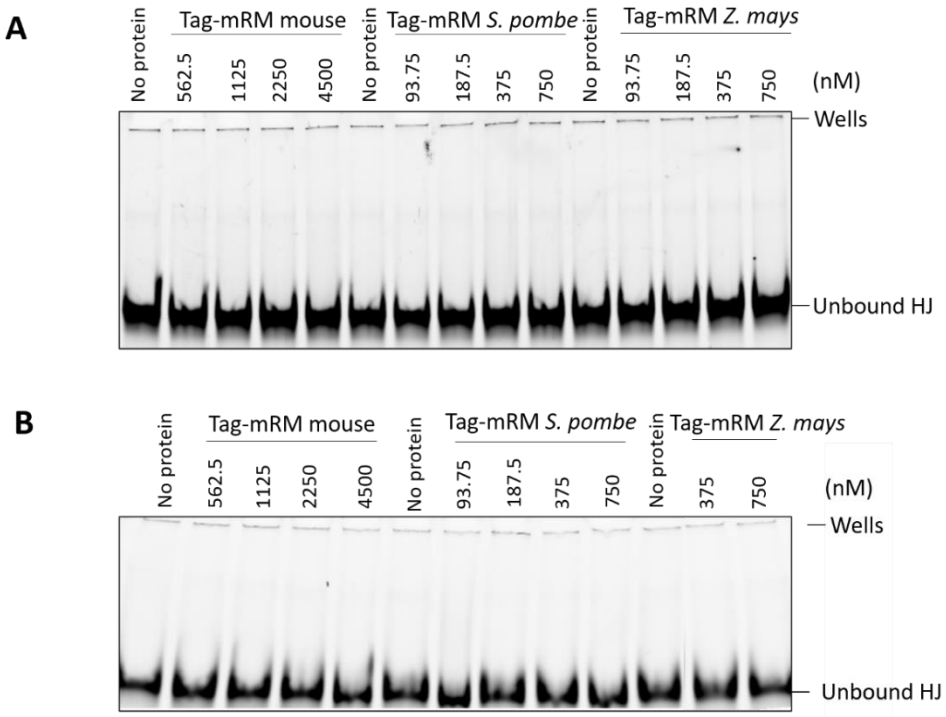

Figure 34: DNA-binding property of mRM of mouse, S. pombe and Z. mays. *Gel-shift assay of different concentrations of minimal RM complexes tagged of the mouse (REC114-MEI4), S. pombe (Rec7-Rec24) and Z. mays (PHS1-MPS1) binding to a fluorescent HJ substrate (10 nM).* ***(A)*** *Test with 5 mM Mg^{2+} in the samples, 1.5 mM in the buffers of electrophoresis and the polyacrylamide gel.* ***(B)*** *Test with another TAE (0.1 mM EDTA, 4 mM Tris at pH 8.3, 4 mM NaOAc, 0.2 mM MgOAc).*

There are several hypotheses as to the lack of DNA-binding activities of these complexes. Firstly, it could be that the main DNA-binding site of these proteins is not located in the minimal part. Secondly, it could be that the tags used (HisSUMO and MBP) are a steric hindrance to the binding of these complexes to DNA. Thirdly, it could be that this interaction is weak and therefore the complex is immediately destabilized when electrophoresis is initiated. Indeed, as shown by Dima Daccache in Figure 32, the full-length proteins of the mouse complex manage to bind to DNA but this association is rapidly dissociated at the start of electrophoresis, indicating that the protein-DNA interaction is fairly weak and does not resist EMSA for long, which may explain the fact that for the minimal mouse complex, no binding to DNA is measured by EMSA, as this interaction may be even weaker than for the full-length complex. However, a recent as

yet unpublished article has shown that the minimal complex formed by amino acids 203 to 259 of REC114 and residues 1 to 42 of MEI4 manages to bind to DNA, as demonstrated by an EMSA. However, the affinity of this complex for DNA is less than that of the *S. cerevisiae* complex. The concentration tested was 30 µM, 6.67 times higher than our tested concentration. Moreover, only 30% of binging is measured at the concentration of proteins of 30 µM. At a concentration of 15 µM, no binding was measured, so our results could be explained by the fact that the protein concentrations tested were not high enough. In addition, the DNA substrate tested was also larger (150 bp) and linear (Liu *et al.* 2023). In addition, AFM experiments by Karen Mechleb, a member of our laboratory, have shown that the mouse REC114-MEI4 minimal complex binds to DNA (Figure 35) (Daccache *et al*. 2022).

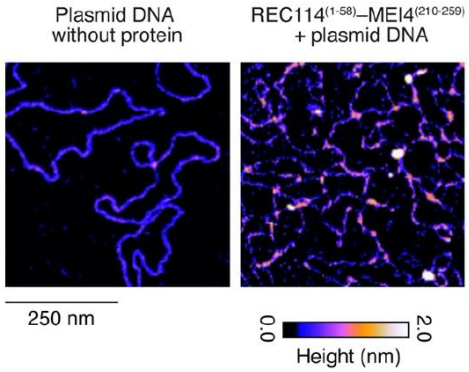

Figure 35: AFM imaging on mouse mRM. *AFM imaging of 1 nM plasmid DNA in the presence of minimal untagged mouse REC114-MEI4 (100 nM) (Daccache et al. 2022).*

Therefore, mode of DNA binding is probably conserved but quantitatively very different from *S. cerevisiae*.

2. Orthologs of the coiled-coil domain of Mer2

Tagged and untagged coiled-coils orthologs to Mer2 were overexpressed in *E. coli* and then purified (Figure 36).

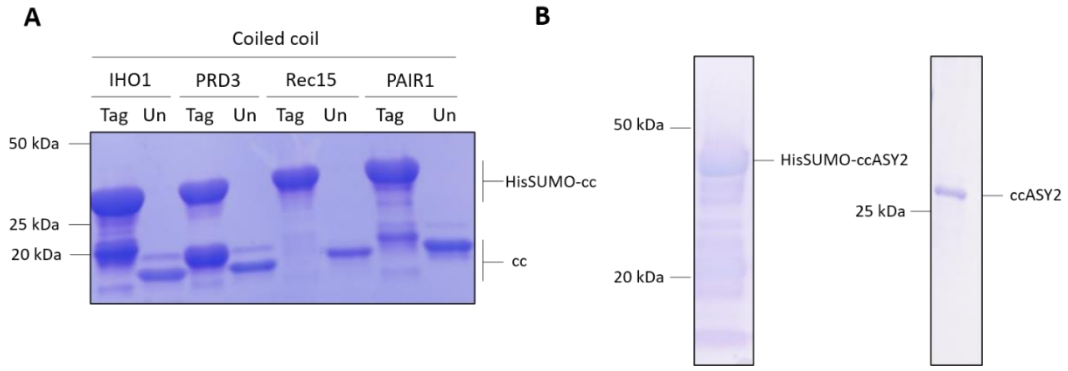

Figure 36: Purification of the coiled coils. *SDS-PAGE gel of the purification* ***(A)*** *of tagged and untagged coiled coil from mouse (IHO1), A. thaliana (PRD3), S. pombe (Rec15) and Z. mays (PAIR1)* ***(B)*** *of the coiled coil tagged on the left and untagged on the right of S. macrospora (ASY2).*

2.1 Conservation of the structure of the cc of Mer2

SEC-MALS analyses of HisSUMO-tagged coiled-coils were performed and showed that all the coiled-coils tested had a homotetrameric structure. Indeed, the measured molecular weights by the MALS are close to those expected in case of a homotetrameric structure. The results obtained are shown in table 1.

Table 1 : SEC-MALS result. *Table of SEC-MALS results obtained on coiled-coils tagged with a HisSUMO tag in comparison to the theoretical weights of a monomer and a tetramer.*

Organisms	Protein	Mw expected monomer (kDa)	Mw expected tetramer (kDa)	Mw obtained by SEC-MALS (kDa)
M. musculus	HisSUMO-IHO1	31.75	127.0	113.0
A. Thaliana	HisSUMO-PRD3	32.19	128.76	135.2
S. pombe	HisSUMO-Rec15	32.10	128.40	132.7
Z. mays	HisSUMO-PAIR1	32.95	131.80	121.2
S. macrospora	HisSUMO-ASY2	38.75	155.00	142.2

These results are indeed derived from the main peak detected by the UV detector at the exit of the SEC of the purified protein samples. The results as a function of the elution volume of the SEC of a purified protein sample are shown in Figure 37.

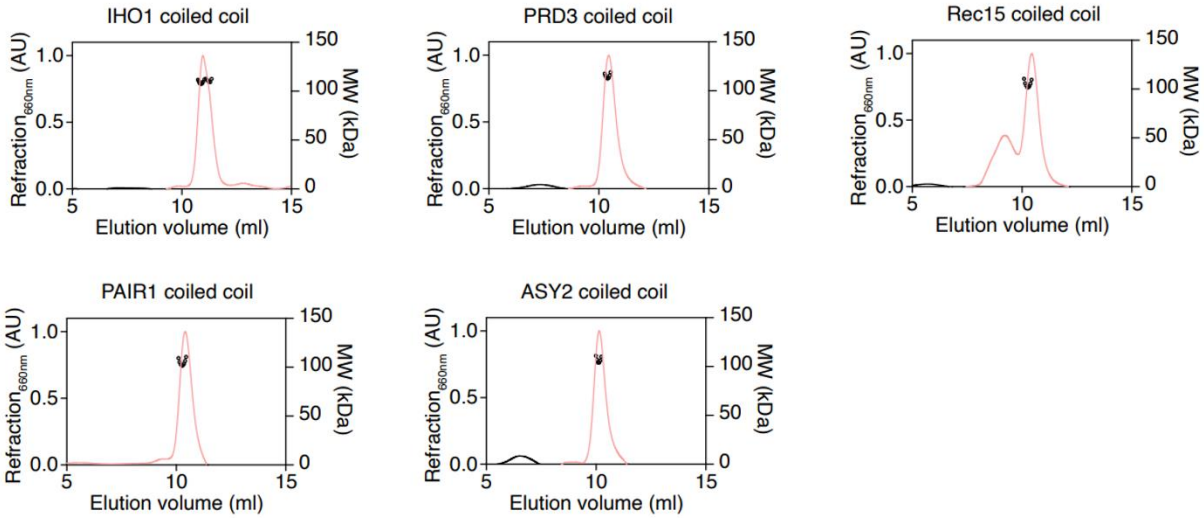

Figure 37: SEC-MALS on the coiled-coil domains*. SEC-MALS analysis of tagged coiled-coil domains of IHO1, PRD3, Rec15, PAIR1 and AYS2. The traces show differential refraction at 660 nm (arbitrary units) and circles are molar mass measurements across the peak in function of the elution volume of the SEC.*

By using Small Angle X-ray Scattering (SAXS) technique, it is possible to distinguish between the parallel and antiparallel topology of the proteins that make up the coiled-coil domain. Indeed, if these proteins are arranged in a parallel direction, all the tags HisSUMO fused to the N-terminal part of the proteins will be on the same side, whereas in the case of an antiparallel arrangement, the tags will be equally distributed on both sides of the proteins (Figure 38).

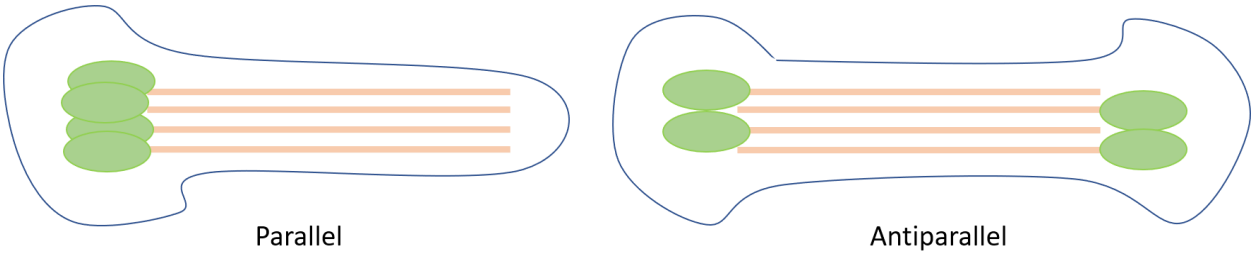

Figure 38: Representation of the orientation of a coiled-coil domain. *In pink are the α-helices. Green represents the HisSUMO tag fused to the N-terminal part of the coiled-coil domain. If the α-helices are arranged in parallel as shown on the left of the figure, all the HisSUMO tags are on the same side. The electron density, shown in blue, would therefore be much greater on their side. If the orientation is antiparallel, the tags are on both sides of the helices and the electron density is distributed on both sides of the coiled coil.*

Consequently, the shape represented by the electron density measured by SAXS will not be the same in the case of parallel or antiparallel orientation (Figure 38).

For this experiment, we carried out the coiled coil purifications and the SAXS analyses were performed by Yann Sterckx and his team. In Figure 39, the α-helices in a parallel or antiparallel orientation are shown in red. In grey is the electron density measured by SAXS. The results of SAXS (*ab initio* particle reconstitutions) on all the coiled-coil domains orthologous to that of Mer2 show a density excess on one side of the tetrameric complex, which validates the AlphaFold model and confirms the parallel organization of the α-helices of the coiled-coil domain (Figure 39).

Looking at Figure 40, we can see that the parallel helix model is encompassed by the excess density measured, which is not the case for the antiparallel model (measured by the D_{max}, which is the minimum the more closely the model resembles the measurement).

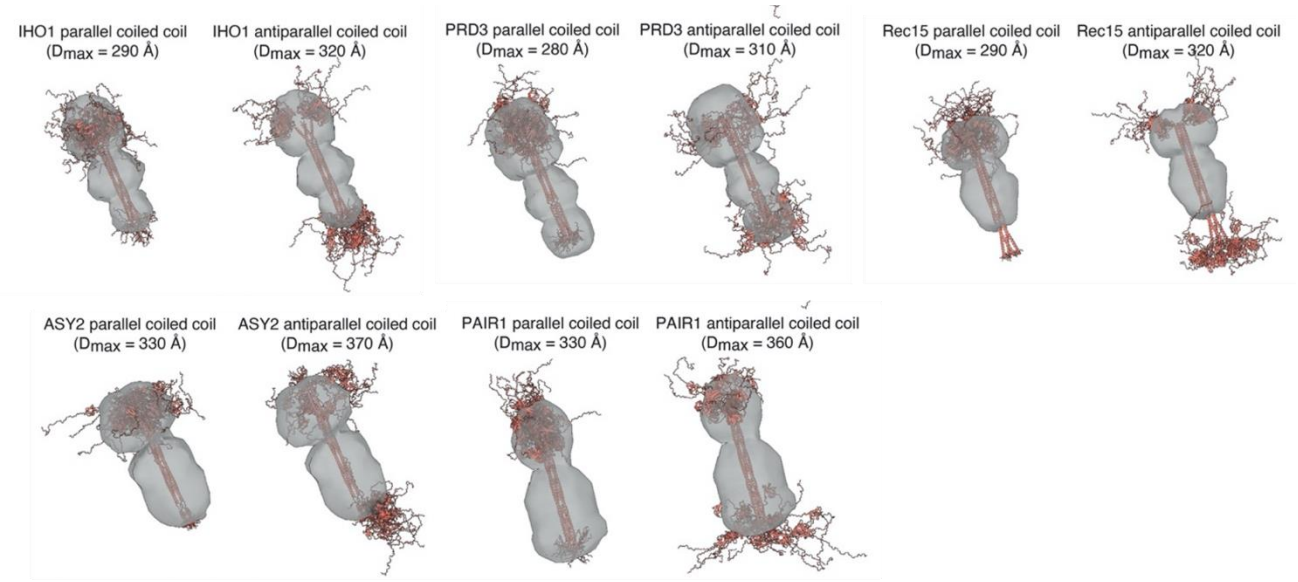

Figure 39: Parallel structure of the coiled-coil domains tested by SAXS. *Overlay of parallel model of coiled coil tagged with a HisSUMO in their N-terminal part with the ab initio reconstructed shape obtained from SEC-SAXS analyses in comparison with the antiparallel model (Daccache et al. 2022).*

Now that we have verified the parallel and homotetrameric structure of the coiled-coils, we wanted to go further in validating the model predicted by AlphaFold using a completely independent technique. To do this, we performed a XL-MS experiment. We first had to determine what concentration of DSS, the crosslinking agent, was necessary to observe sufficient XL between the lysines of the coiled coil. We performed a small-scale crosslink test by varying the DSS concentration on the IHO1 coiled coil. The crosslinking was verified by an SDS-PAGE gel shown in Figure 40. Without DSS, one major band is observed, corresponding to monomeric IHO1. However, as soon as 0.5 mM DSS is added, lower-migrating bands are visible, corresponding to dimers and multimers.

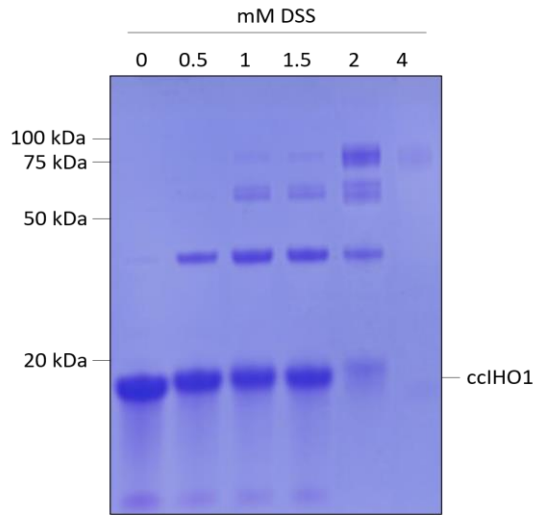

Figure 40: XL test on the coiled coil of IHO1. *Testing the effect of DSS concentration on the coiled coil of IHO1 (7.55 µg = 5µl of protein solution).*

At a concentration of 2 mM DSS, most of the proteins has been crosslinked. At 4 mM almost nothing was observed, presumably because the protein was not able to penetrate the gel due to the large number of XL. We decided to perform the XL-MS experiment with a DSS concentration of 2 mM to have sufficient levels of XL (Figure 40).

XL-MS experiments were performed using 60µg of untagged coiled coils. The SDS-PAGE gels below show that the proteins have been crosslinked by DSS. Multimers are visible on the SDS-PAGE gels shown in Figure 41A. After analysis, the crosslink obtained are represented in Figure 41B.

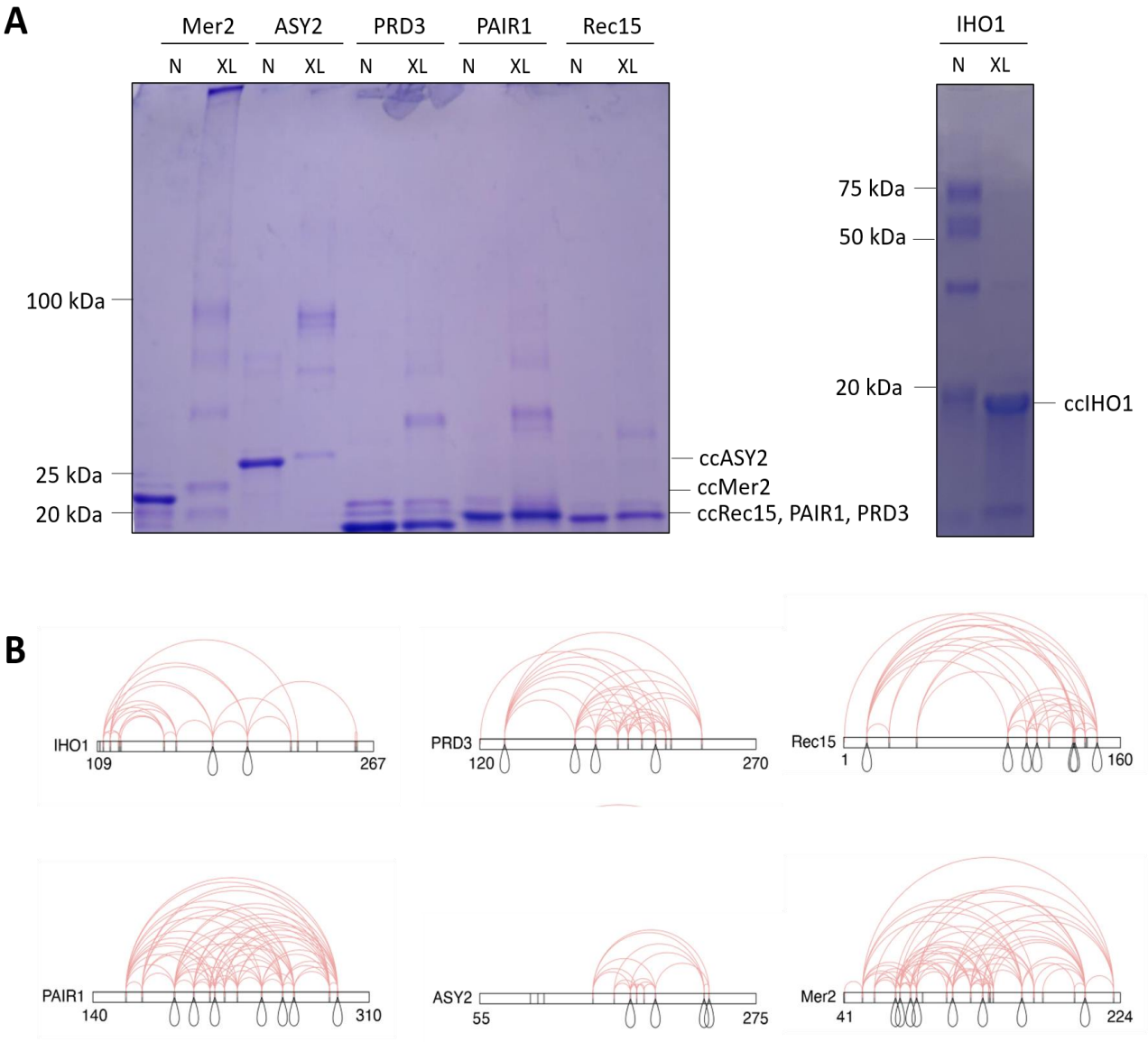

Figure 41: XL-MS on the coiled-coil domains. (A) *SDS-PAGE gel of coiled-coil domains (from Mer2, ASY2, PRD3, PAIR1, Rec15 and IHO1) of proteins untreated with DSS (named N) and proteins crosslinked with DSS (named XL).* ***(B)*** *Representation of the crosslink obtained after XL-MS on the coiled-coils. The black loops represent the bonds between a lysine and itself while the pink curves represent crosslinks observed between two different lysines. The protein concentrations used were: 0.95 µg/µl for IHO1, 0.76 µg/µl for PRD3 and PAIR1, 0.63 µg/µl for Rec15, 0.56 µg/µl for ASY2 and 0.54 µg/µl for Mer2.*

In all these coiled coils, lysines have been bound to identical lysines (curve in black in Figure 41B), indicating that the proteins form a multimer (consistent with SEC-MALS and SAXS). However, the distance of the majority of the crosslinked lysine (shown in pink in Figure 41B) observed and projected on the AlphaFold representation of the coiled coils exceeds the threshold value of 27 Å for distances measured within the same coiled-coil domain (Figure 42).

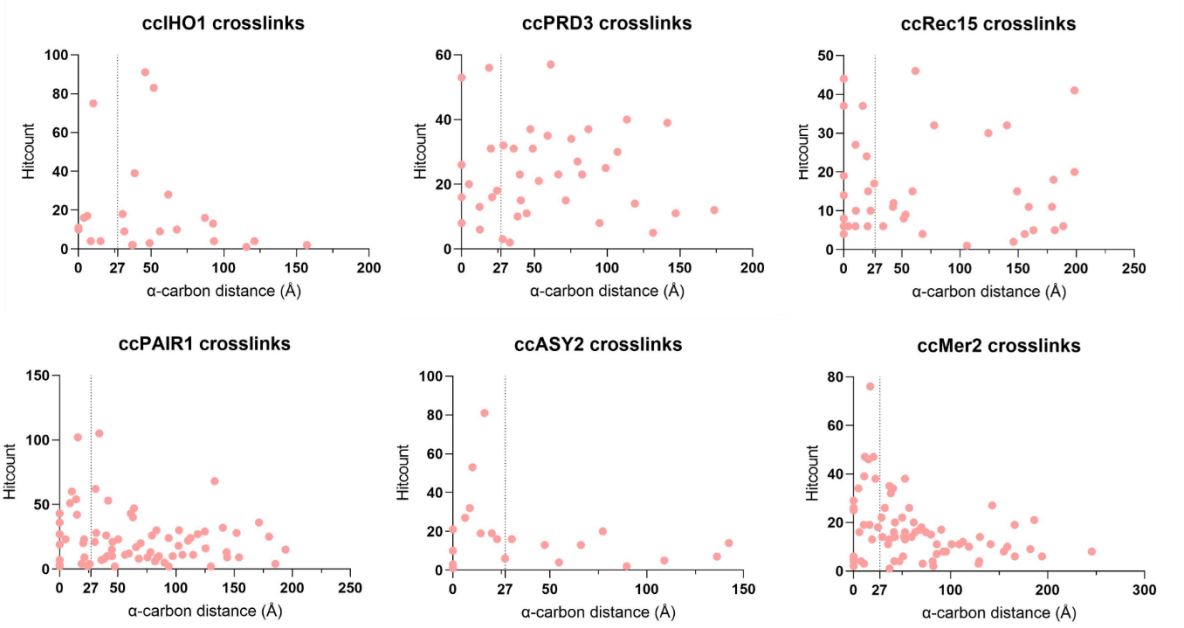

Figure 42: Hitcount and distances of XL. *Graph showing the number of times the same crosslink was detected by MS (Hitcount) as a function of the distance between the α-carbons of the crosslinked lysines in the same coiled-coil domain. The threshold of validity of the 27 Å measurements is represented by a vertical dashed line. The corresponding data can be seen in Appendix 1.*

Indeed, any value exceeding this distance is inconsistent with the structural model proposed by AlphaFold because the spacer linking the two DSS reaction groups measures 11.4 Å (Figure 43) when fully extended, thus allowing two α-carbons of crosslinked lysines to be approximately 24 Å apart. However, an allowance of 3 Å is made because experience shows that many probable XLs are at a distance of 27 Å, giving a maximum distance of 27Å (Calabrese and Radford 2018; Merkley *et al.* 2014).

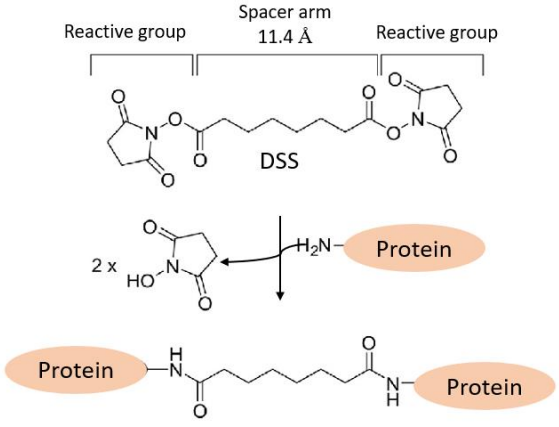

Figure 43: Size of XL carried out by the DSS. *Representation of the DSS with its two reaction groups and its 11.4 Å spacer arm. DSS reacts with lysines to bind two different proteins or two lysines from the same protein (modified from Calabrese and Radford 2018).*

Certain long-range XL could be explained by the fact that the coiled coils are flexible but the confidence score of these coiled coils are high for the orthologs of Mer2 and AlphaFold don't predict any flexible region in these coiled-coil. Furthermore, not all long-range XLs can be explained by this hypothesis. Some of these long-range XLs can also be explained by an antiparallel orientation of the coiled coils, but this hypothesis is not consistent with the SAXS analysis. Another explanation could be that inter-tetramer crosslinks occurred, in addition to intra-tetrameric (intra or intermolecular) crosslinks, because the proteins were too concentrated during the crosslinking experiment. A representation of the various possible crosslinks can be seen in Figure 44.

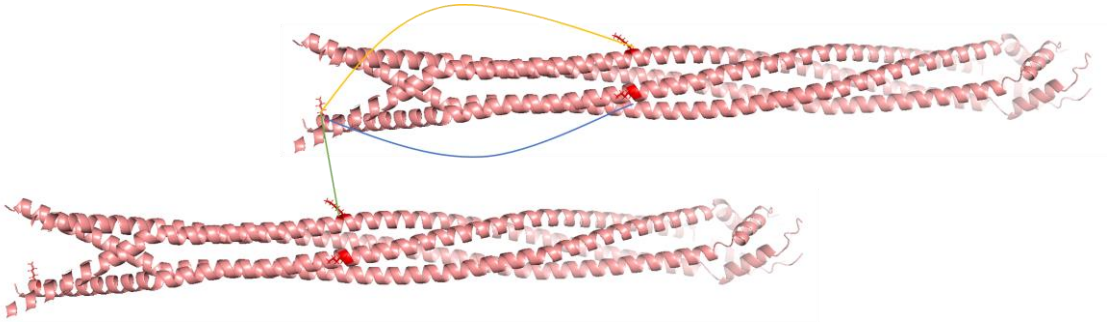

Figure 44: Representation of possible crosslinks for the IHO1 coiled-coil. *In XL-MS, crosslinks were detected between residue 4 and 67 of the IHO1 coiled-coil. These crosslinks can be intramolecular (shown in blue), intermolecular within the same coiled-coil (shown in orange) or they can be inter-coiled-coil (shown in green).*

In order to test the hypothesis that long-range XLs exceeding the 27 Å threshold are in fact inter-coiled coil XLs, we carried out the same experiment but in dilute protein concentrations. By reducing the protein concentration, we expect to increase proportion of intra-tetrameric crosslink vs. inter-tetrameric events. This experiment was carried out with the Mer2 coiled coil (Figure 45B).

In fact, as shown in Figure 45, far fewer long-range XL are detected. For the first condition tested, 59.3% of the XLs exceeded the threshold value of 27 Å and therefore violated the model predicted by AlphaFold, whereas in the condition where the protein was diluted 7 times, this percentage fell by more than half to 22.1% (Figure 45C). Therefore, by diluting the protein during the XL reaction, we significantly decreased inter-coiled coil XL.

Therefore, we know that a lot of XL shown in Figure 45B are actually inter-coiled coil XL, which makes the results obtained on orthologs coiled coil uninterpretable.

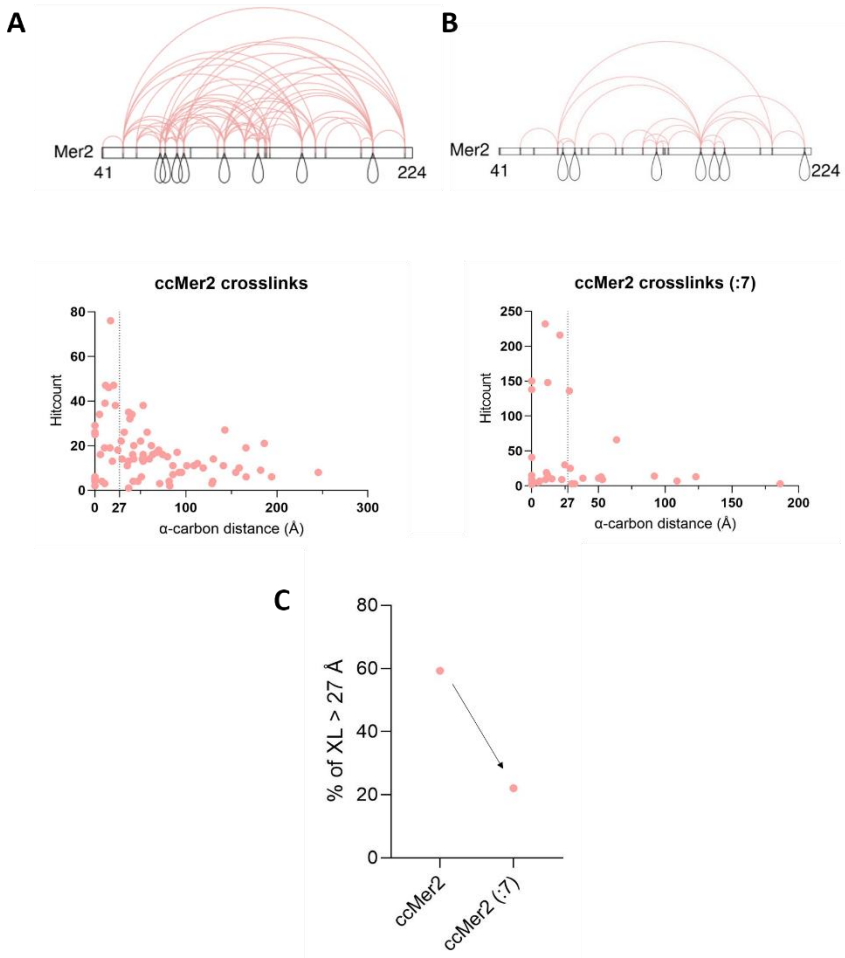

Figure 45: XL-MS results on Mer2 coiled-coil under diluted conditions*. **(Top)** Representation of the crosslinking obtained after XL-MS on Mer2 coiled-coil. The black loops represent the bonds between a lysine and itself, while the pink curves represent the cross-links observed between two different lysines. **(Bottom)** Graph showing the number of times the same cross-linking was detected by MS (Hitcount) as a function of the distance between the α-carbons of the cross-linked lysines. The validity threshold for measurements at 27 Å is represented by a vertical dashed line. The corresponding data are shown in Appendix 2. **(A)** XL-MS results on the cc of Mer2 (0.54 µg/µl) and **(B)** XL-MS results on the cc of Mer2 in dilute condition (0.08 µg/µl). **(C)** Graph showing the percentage of XL with a distance of more than 27 Å (projected onto the model predicted by AlphaFold) for the first condition where the proteins are concentrated and for the condition where the proteins are diluted almost 7 times.*

The remaining long-range XL (Figure 45B) can be explained by the fact that XL inter-coiled coil remains or by the fact that the coiled coil has a flexible region at its center (however, this hypothesis is not confirmed by a SAXS experiment (Daccache *et al.* 2022)). The only result that can be drawn from this experiment is the fact that the dilution experiment indicates that the concentration-dependent long-range crosslinks are evidence of inter-tetrameric interactions. Since SEC-MALS and SAXS indicate that the protein forms mostly tetramers, these intermolecular interactions are likely weak and transient, a property that presumably participates in Mer2 condensation.

2.2 Conservation of DNA binding of the coiled-coil domain

It has been shown that the coiled coil of IHO1, similarly to Mer2, is sufficient to bind the DNA and that a branched substrate is preferred >20-fold to a linear substrate (Figure 46A). This indicates that the tetrameric IHO1 coiled-coil domain undergo multivalent protein-DNA interactions (Daccache *et al.* 2022).

By mutating the residues K121, K122, R123 and R124 (mutant KKRR) positively charged that point outward the coiled coil (Figure 46C) to alanine, the DNA binding by the coiled-coil domain of IHO1 is inhibited (Figure 46B). These amino acids are therefore responsible for the DNA-binding activity of this protein.

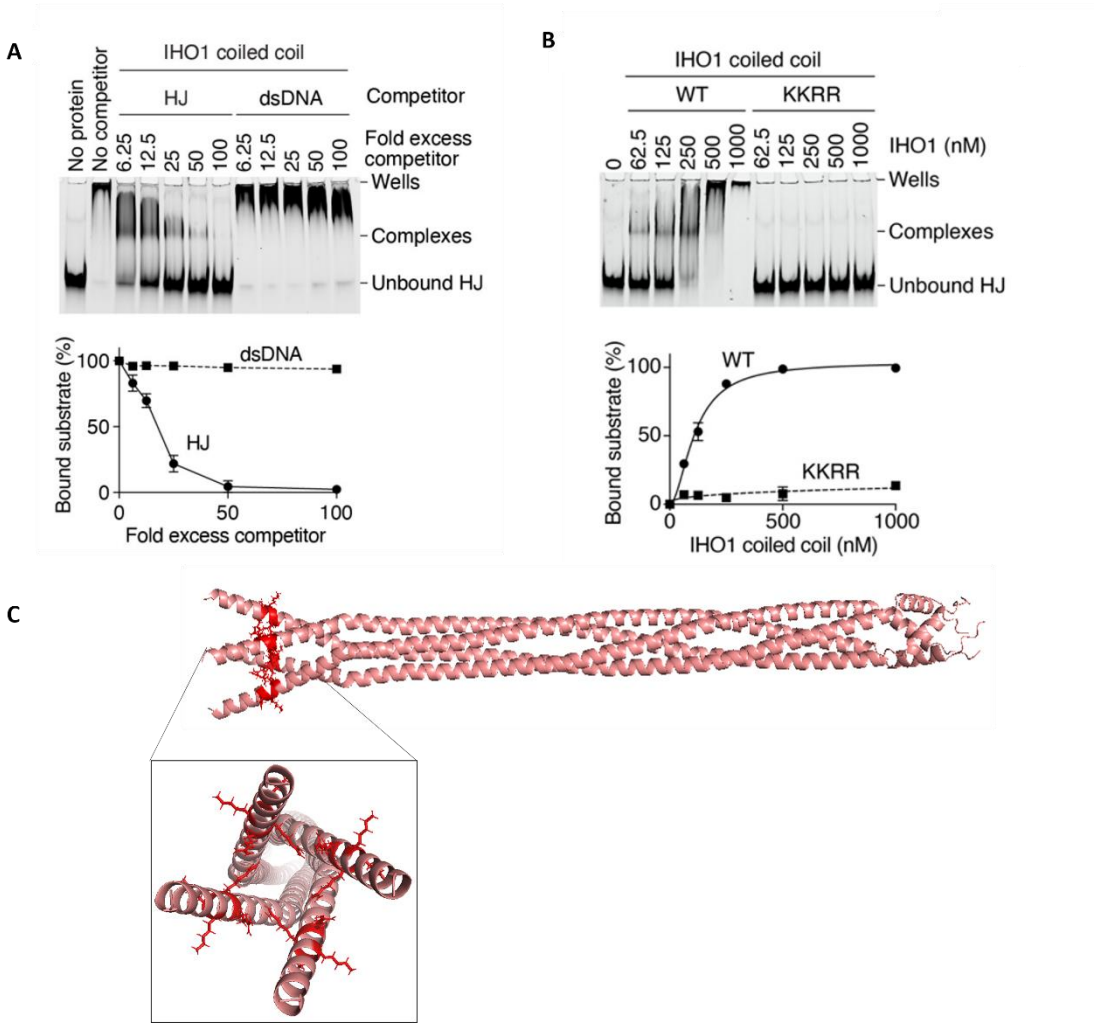

Figure 46: DNA binding properties of the coiled-coil of IHO1. (A) *Competition assay of IHO1 coiled-coil domain (600 nM) binding to a fluorescent HJ substrate (10 nM) in the presence of unlabeled HJ or dsDNA substrates. Error bars are ranges from two independent experiments (some are too small to be visible) (Daccache et al. 2022).* ***(B)*** *Gel-shift assay of different concentration of IHO1 coiled-coil domain WT and the coiled-coil domain of IHO1 with the following mutation: K121, K122, R123, R124A binding to fluorescently-labeled HJ substrate. Error bars are ranges from two independent experiments (some are too small to be visible) (Daccache et al. 2022).* ***(C)*** *AlphaFold representation of the coiled-coil tetrameric domain of IHO1with the positive amino acids K121, K122, R123 and R124 pointing outwards from the complex highlighted in red.*

We are also able to detect that the coiled coil of Rec15 is sufficient to bind DNA by a gel-shift assay. Furthermore, by substrate competition EMSA, this protein has a ~17-fold preference for the branched substrate type HJ over a linear dsDNA substrate, meaning that this tetrameric coiled-coil domain mediates multivalent DNA-protein interactions (Figure 47A).

By mutating amino acids K134, K135, K141, K142, R143 (mutant 5KR) to alanine, we are able to completely inhibits Rec15 binding to the DNA showing that these residues are responsible for the DNA binding property of this protein (Figure 47B). These amino acids are positively charged pointing outward of the coiled coil (Figure 47C).

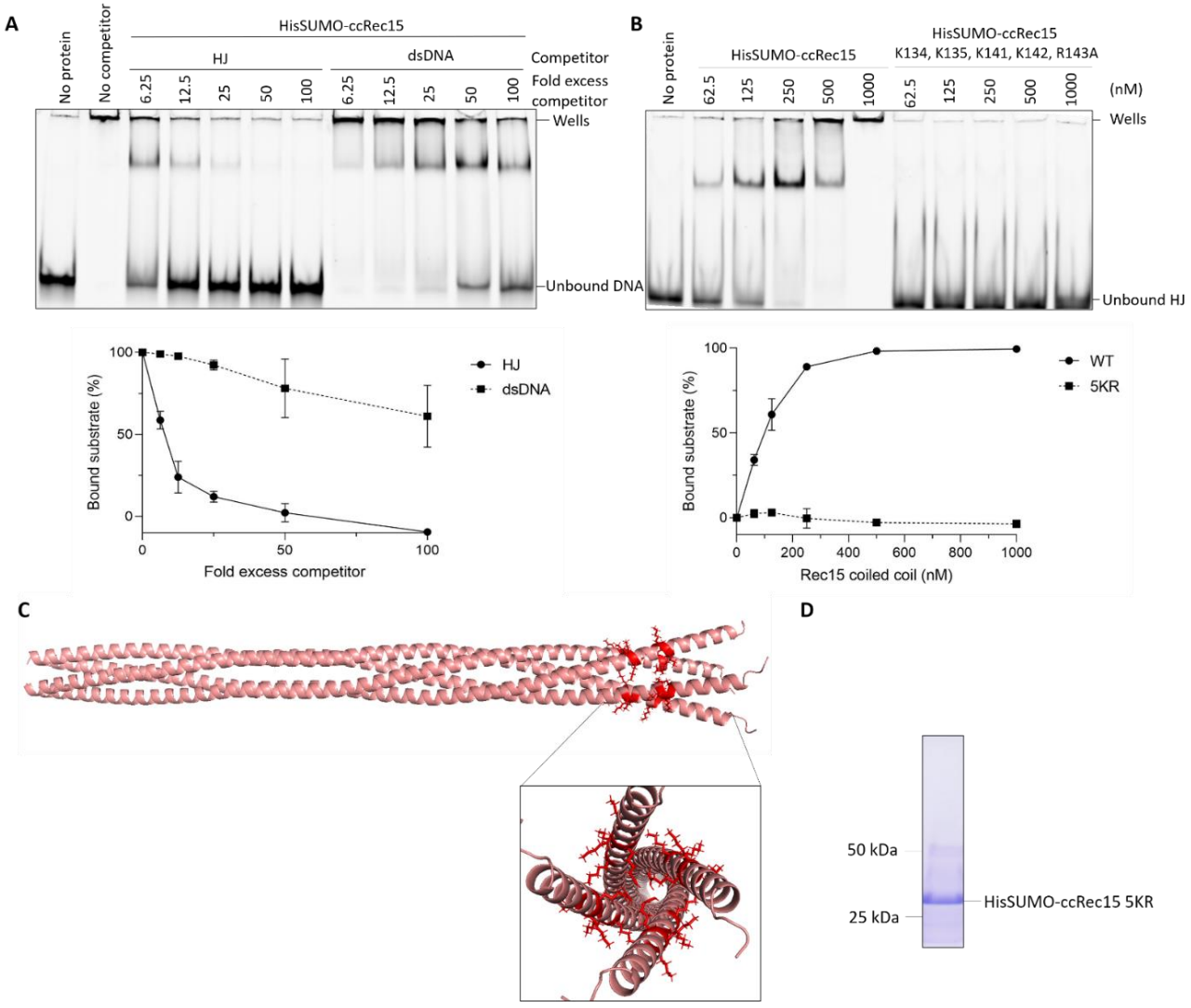

Figure 47: DNA-binding properties of the coiled coil of Rec15. (A) *Competition assay of Rec15 coiled-coil domain (600 nM) binding to a fluorescent HJ substrate (10 nM) in the presence of unlabeled HJ or dsDNA substrates. Error bars are ranges from two independent experiments (some are too small to be visible).* ***(B)*** *Gel-shift assay of different concentration of Rec15 coiled-coil domain WT and the coiled-coil domain of Rec15 with the following mutation: K134, K135, K141, K142, R143A (5KR) binding to fluorescently-labeled HJ substrate. Error bars are ranges from two independent experiments (some are too small to be visible).* ***(C)*** *AlphaFold representation of the coiled-coil tetrameric domain of Rec15 with the positive amino acids K134, K135, K141, K142 and R143 pointing outwards from the complex highlighted in red.* ***(D)*** *Purification of the tagged coiled-coil domain of the mutant 5KR of Rec15.*

An EMSA performed on the others coiled coil tagged showed that the coiled coil of PAIR1 is sufficient to bind with a substrate of HJ type DNA because complexes in the gels and in the wells can be observed. This EMSA was realized with the special TAE: 0.1 mM EDTA, 4 mM Tris at pH 8.3, 4 mM NaOAc, 0.2 mM MgOAc (Figure 48A).

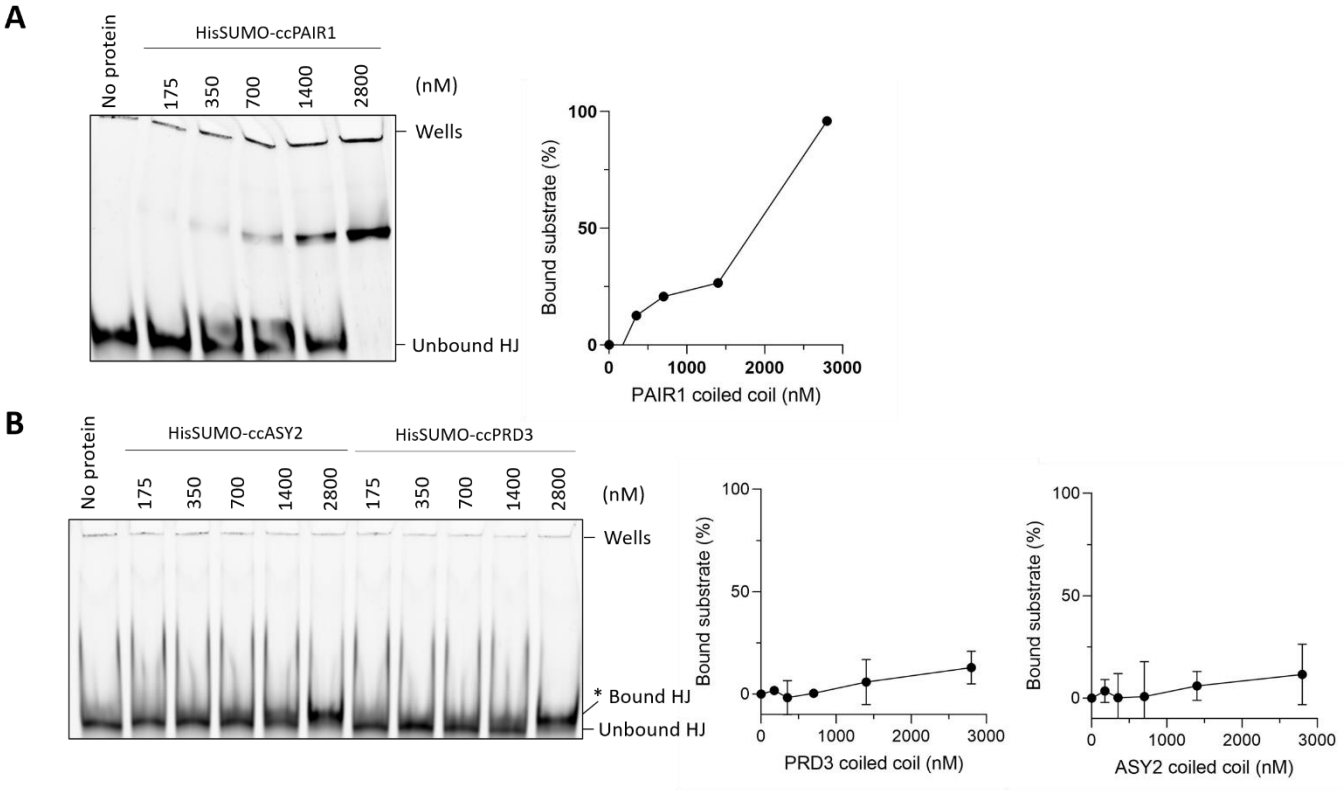

Figure 48: DNA binding properties of the coiled-coil of PAIR1, PRD3 and ASY2. (A) *Gel-shift assay of different concentration of PAIR1 tagged coiled-coil domain. This gel shift was realised using a special TAE (0.1 mM EDTA, 4 mM Tris at pH 8.3, 4 mM NaOAc, 0.2 mM MgOAc and the gel migrate 4h at 100V).* ***(B)*** *Gel-shift assay of different concentration of ASY2 and PRD3 tagged coiled-coil domain. Error bars are ranges from two independent experiments (some are too small to be visible).*

The coiled coils of PRD3 and ASY2 do not show a DNA binding for concentrations below 2800 nM. However, a complex located slightly above the unbound HJ band is visible for a concentration of 2800 nM. This can be explained by the fact that the complex formed by DNA and proteins is rapidly dissociated at the beginning of electrophoresis due to lack of strong interaction. This can also be explained by the presence of a possible contaminant that can bind the DNA, because our non-tagged proteins have not been run on size exclusion chromatography and, as can be seen on the purification gels (Figure 35), there are other proteins or degraded proteins present. It could be that small contaminating proteins are bound to the DNA and that, as a result, a change in the progression of the DNA is visible Figure 48B.

2.3 Stability of the coiled coils in different buffers

To validate the structure of these coiled coils with a high resolution, it might be possible to carry out crystallography of these coiled coils. However, before we can do this, we need to test the stability of these coiled coils in different solutions containing different salt concentrations and at room temperature. A protein will not be stable if it precipitates. The proteins were then dialyzed O/N in solutions with different salt concentrations and then centrifuged. If the proteins remained soluble, they would be found in the supernatant, whereas if precipitates were obtained, they would be found in the pellet.

For HisSUMO-ccIHO1, the protein does not appear to be stable in a buffer with a salt concentration of 0 mM because most of the protein is found in the sample pellet, indicating that it has precipitated. However, it is stable under the other conditions tested. For the other coiled coils, the majority of the protein was found in the supernatant of the samples, regardless of the buffer tested, indicating a certain stability of the protein as it did not precipitate. To conclude, all tagged coiled coil are stable at a salt concentration of 50 mM (Figure 49).

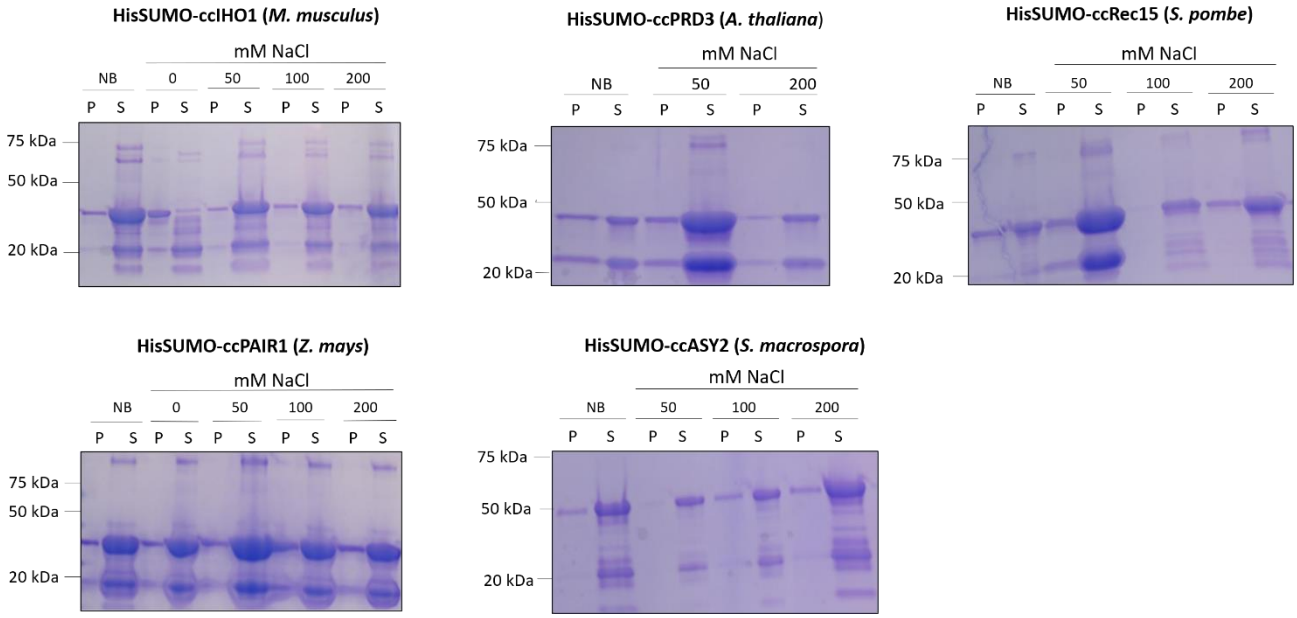

Figure 49: Stability test. *Stability test for HisSUMO-tagged coiled-coils of IHO1, PRD3, Rec15, PAIR1 and ASY2 performed with 10µl of protein in different buffer containing different salt concentrations at room temperature and with 20mM Tris. Where NB = "normal buffer", being the base buffer in which the proteins are stored (Here, Buffer C), P = pellet and S = supernatant.*

V Discussion and perspectives

1. Structural conservation of Rec114-Mei4 and Mer2

The orthologs of the *S. cerevisiae* RMM complex have evolved rapidly and have not retained their amino acid sequence. However, despite the difference in sequences, AlphaFold proposes a highly conserved structure for these proteins.

In fact, the minimal RM complex of the different orthologs predicted by AlphaFold shares a structure that is highly similar, composed of two α-helices belonging to two different Rec114 which form a ring in which a Mei4 α-helix is inserted. Several ways to validate these structures have been tested. Rec114-Mei4 interaction domains are conserved in *M. musculus*, *S. pombe, A. thaliana*, and *Z. mays* because we can copurify all of these minimal domains together. This is particularly important for the minimal complex of PHS1 and MPS1 from *Z. mays*, as the role of MPS1 in DSB formation has not been tested *in vivo*, telling us that MPS1 is likely a Mei4 ortholog.

Concerning the validity of the structural models proposed by AlphaFold of the *M. musculus*, *S. pombe* and *A. thaliana* complex, it was shown by a Pulldown that certain amino acids of the Rec114 ortholog, identified as pointing towards the Mei4 ortholog on the model, play an essential role in the protein-protein interaction of this complex. This confirms the significance of the three-dimensional representations of these complexes. As a reminder, these amino acids are as follows: F230, F240 and F243 for REC114 (*M. musculus*), Y320, F325 and L328 for Rec7 (*S. pombe*) and Y284, F290 and L294 for PHS1 (*A. thaliana*). All these amino acids are found in SSM7, identified in the C-terminal part of the orthologs of Rec114 (Figure 15).

The phenylalanines occupying a position similar to that of F411 in the *S. cerevisiae* complex (that is required to maintain the interaction between Rec114 and Mei4) appear to retain its structural role, since mutating this residue reduces the protein-protein interaction of the orthologs of Rec114 and Mei4. In mouse, the effect of the single mutation of this phenylalanine F240 is sufficient to notice a significant decrease in the interaction, unlike the *S. pombe* and *A. thaliana* complex where more mutations are required to observe a negative impact on the interaction between RM.

However, for the Rec7-Rec24 complex of *S. pombe*, it has been shown *in vivo* that the simple mutation of phenylalanine F325 is sufficient to inhibit the formation of DSBs and the interaction between the two proteins (Steiner, Kohli, and Ludin 2010). This can be explained by the fact that the stringency of the protein interaction is stronger *in vivo* than *in vitro*, despite the presence of detergent (0.5% Triton x100), a high salt concentration (500 mM NaCl) and numerous column washes, which was still insufficient

to show, *in vitro,* the inhibition of interaction between Rec7 and Rec24 by the F325 mutation of Rec7. In addition, it should be remembered that this experiment was carried out using proteins overexpressed in *E. coli*, so the concentration of proteins is much higher than *in vivo*. By mutating the F325 amino acid to alanine, we showed, *in vitro*, that the interaction between the two proteins in the complex was reduced. We can therefore assume that the dissociation constant (K_d) of this complex has increased compared with that of the WT complex (which has a lower K_d). Therefore, it is possible that the *in vivo* concentration of these proteins is below the threshold K_d value of the mutant and not of the WT, so that the complex for the mutant cannot associate, whereas it can associate for the WT. *In vitro*, however, we are working with much higher protein concentrations, which can exceed both K_d values allowing the two different complexes to be formed.

Concerning the Pulldown assay of the *Z. mays* complex, we can see that the mutations of residues F327, L334 and I338 do not have a negative effect on the interaction between the two proteins in the complex, although the residues point towards the inside of the complex represented by AlphaFold. In fact, the simple mutation of the phenylalanine F327, which is evolutionarily conserved, impacts positively this interaction. We can therefore assume that this model is not supported by this mutagenesis. If we take a look at the accuracy of the prediction of the structural model, we can see that the pLDDT score from AlphaFold of the PHS1-MPS1 minimal complex is weak (Figure 50), this complex has a low confidence score.

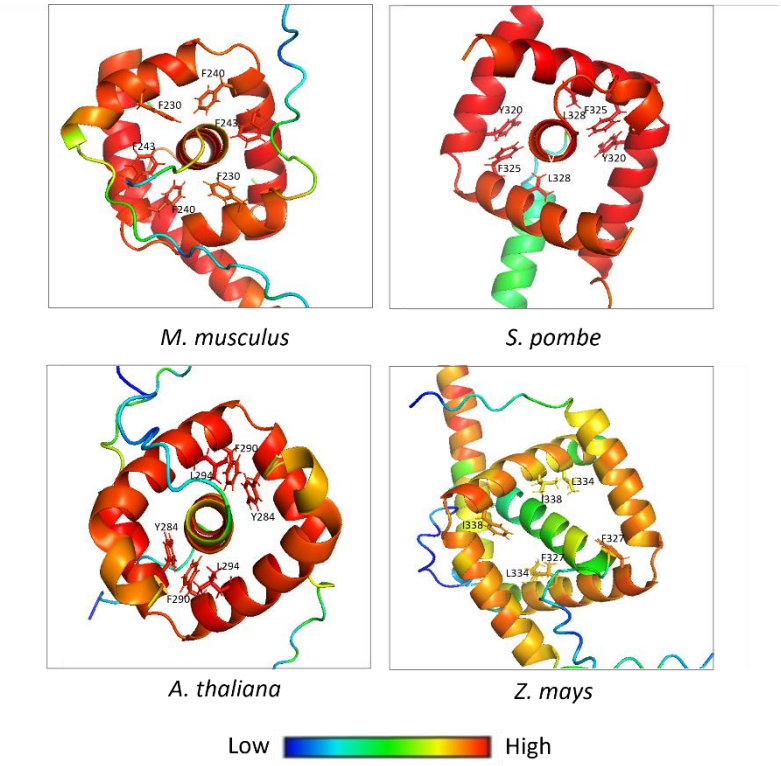

Figure 50: Confidence score of the AlphaFold representations of the RM complexes*. Color-coded AlphaFold representation of the minimal complexes of the RM orthologs highlighting the hydrophobic amino acids identified as pointing towards the interior of the protein complex that were mutated for the Pulldown experiment. Red represents a high confidence score; blue indicates a low confidence score.*

For the other complexes, the predicted structures all have a higher confidence score, shown in red. The value of the prediction is therefore higher and it was validated by the Pulldown. Another way of verifying these AlphaFold models would be to analyze them using NMR spectroscopy, as was done for the *S. cerevisiae* complex (Daccache *et al.* 2022). However, in order to obtain a higher resolution and more accurate representation of these minimal complexes, crystallography or Cryogenic electron microscopy (Cryo-EM) could be envisaged.

AlphaFold predictions of Mer2 orthologs coiled coil domains are validated by SEC-MALS and SAXS experiments. That showed that the tetrameric and parallel structure of the Mer2 coiled coil is well conserved for the orthologous coiled-coil domains tested (that of IHO1 (*M. musculus*), Rec15 (*S. pombe*), PRD3 (*A. thaliana*), PAIR1 (*Z. mays*) and ASY2 (*S. macrospora*)). Crystallography or Cryo-EM could be used to provide a higher resolution and more accurate representation of these coiled-coil domains.

2. DNA-binding properties conservation of RMM

Only the full-length REC114-MEI4 complex from mouse was tested and showed to maintain the DNA-binding function shown in *S. cerevisiae*. It also has a preference for a branch DNA substrate, meaning that this complex of proteins mediates multivalent DNA-protein interactions, like its *S. cerevisiae* counterpart. But in our experiments, the minimal complex of REC114-MEI4 is not sufficient to bind DNA. However, another study was able to show the binding of this minimal complex to DNA *in vitro* by EMSA (Liu *et al.* 2023). Their protein concentration tested is much higher than the one we tested, the size, the shape of the substrate and the migration buffer were different from ours. In addition, Karen Mechleb, from the laboratory where I did my master thesis, carried out an AFM experiment on this untagged minimal trimeric complex and demonstrated DNA-binding activity (Figure 35) (Daccache *et al.* 2022). This indicates that the mouse minimal complex can bind to DNA.

In our experiments, it has been shown, that only the *A. thaliana* minimal complex retains the DNA-binding function of the *S. cerevisiae* minimal complex. However, proteins do have the potential for bivalent interaction with DNA due to their positively-charged amino acid patches, which point outwards from the complex predicted by AlphaFold (Figure 19). This means that the DNA-binding activity of these minimal complexes is less robust than the DNA-binding activity of the *S. cerevisiae* minimal complex.

Failure to detect DNA binding for minimum complexes orthologous to Rec114-Mei4 in our EMSAs can be partly explained by the fact that the electrostatic surface potential of the orthologous complexes is significatively more electronegative than the *S. cerevisiae* complex (Figure 51). Given that DNA is negatively charged, so the affinity for

these proteins is lower than for the *S. cerevisiae* complex which is more electropositive. It could mean that the protein-DNA complexes are linked only by weak interactions that are immediately broken on initiation of electrophoresis and migration in the gel, which may explain the fact that we do not measure any binding. In fact, evidence for this hypothesis is provided by the EMSA realized on the mouse FL proteins shown in Figure 32, which shows that the protein-DNA interaction is immediately interrupted at the start of the EMSA. In addition, AFM experiments performed on mouse proteins show that these proteins bind to DNA (Figure 35).

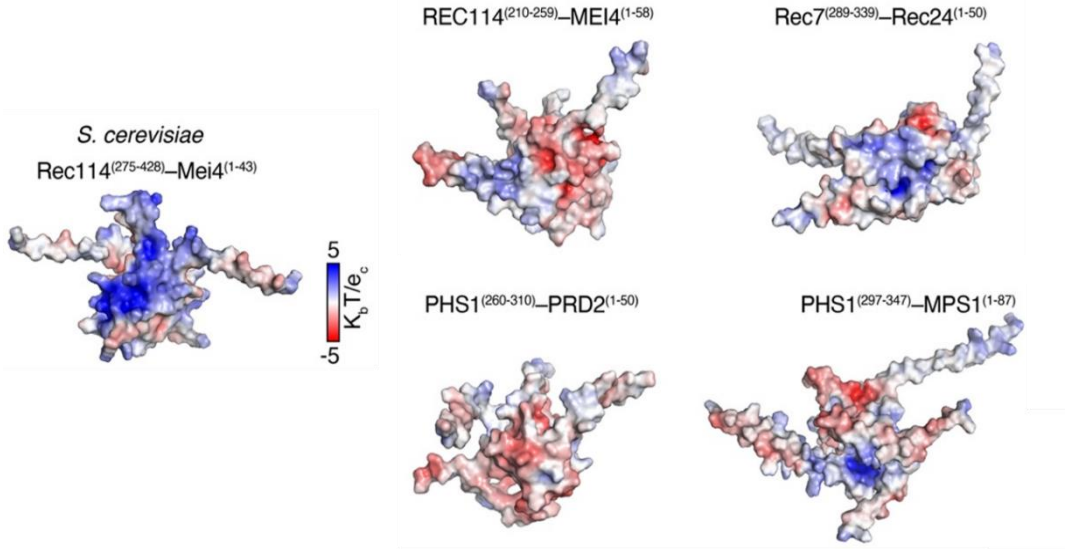

Figure 51: Electrostatic potential of orthologous complexes at RM. *electrostatic potential of orthologous complexes at RM. Surface electrostatic potential of the minimal complex of Rec114-Mei4 from S. cerevisiae and orthologs from mouse, S. pombe, A. thaliana and Z. mays. Blue represents an electropositive surface potential while red represents an electronegative surface potential (Daccache et al. 2022).*

Other techniques that destabilize protein-DNA interactions less than EMSA could be tested to demonstrate this binding, such as filter binding assay or an ELISA-type microtiter method consisting by fixing DNA and quantifying the proteins attached to it after several washes using antibodies specific to a tag to which the proteins are fused. In addition, the same experiment as for the mouse mRM in AFM could be carried out on the other minimal complexes.

Another possible explanation for the low affinity of RM complexes to DNA is that it has been shown *in vivo* in *S. cerevisiae* that Mer2 binds to chromosomes independently of the RM complex. This is not the case for Rec114-Mei4, which needs Mer2 to bind to DNA. We can therefore assume that the DNA-protein interactions of the RM complex are less important than the protein-protein interactions that would attract them to the nucleoprotein axis where Mer2 is bound the chromosomes (J. Li, Hooker, and Roeder 2006).

Moreover, since the full-length proteins were not tested, it cannot be concluded that the function of the Rec114-Mei4 proteins is not evolutionarily conserved, since the

parts essential for DNA binding could also be outside the minimal complex tested. EMSAs could therefore be carried out on these full-length proteins.

Concerning the DNA-binding of coiled coils, only the coiled-coil domains of IHO1, Rec15 and PAIR1 showed DNA binding under the tested conditions, as did the coiled-coil domain of Mer2. The DNA binding function therefore appears to be conserved for these Mer2 orthologous proteins. The coiled coils of IHO1 and Rec15 both show a preference for branched DNA substrates meaning that these tetrameric coiled-coil domains mediate multivalent DNA-protein interactions, similar to Mer2. We are able to identify IHO1 and Rec15 amino acids that are responsible for the DNA-binging (K121, K122, R123 and R124 for IHO1 and K134, K135, K141, K142 and R143 for Rec15). The position of these amino acids is not conserved but the DNA-binding activity well. The binding of these two coiled coils to DNA was also demonstrated by AFM experiments (carried out by Karen Mechleb). In addition, she showed that the Rec15 coiled coil formed condensates in the presence of DNA (Figure 52).

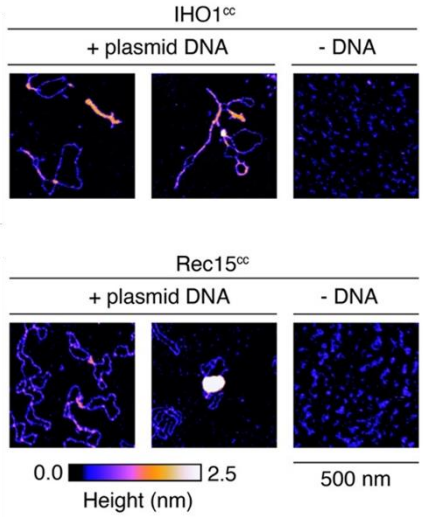

Figure 52: AFM imaging of IHO1 and Rec15 coiled coil. *AFM imaging of 1 nM plasmid DNA* ***(top)*** *and coiled coil of IHO1 (100 nM),* ***(bottom)*** *and coiled coil of* Rec15 (50 nM left and 100 nM right) *(Daccache et al. 2022).*

Concerning the coiled coil of PRD3 and ASY2, there is the assumption that these cc bind to DNA but with a very low binding intensity because these complexes could have disassociated very early on during EMSA. To test this, experiments like AFM, filter binding assay or ELISA could be tested

In Mer2, residues of the C-terminal part of the protein, outside the coiled coil, bind to DNA and are shown to be important but not essential for this binding (Claeys Bouuaert, Pu, *et al.* 2021). Other EMSAs could be carried out on the full-length proteins. Therefore, no hasty conclusions concerning the DNA binding function of *Sordaria macrospora* (ASY2) and *Arabidopsis thaliana* (PRD3) can be drawn because it could be that the element essential for DNA binding is located outside the coiled coil. No non-conservation of function can therefore be deduced for the ASY2 and PRD3 proteins compared with Mer2.

VI Conclusion

The aim of this master thesis was to study the conservation of structure and function of *S. cerevisiae* RMM proteins during evolution, despite their great divergence in the amino acid sequence. Given that the RMM complex cannot be purified together, the study was separated into two distinct parts: one on the Rec114-Mei4 complex and the other on Mer2.

To study the conservation of the structure of the minimal Rec114-Mei4 complex composed of the C-terminal part of Rec114 and the N-terminal part of Mei4, structural models were proposed by AlphaFold. All these models resemble each other (two α-helices of two Rec114 forming a ring in which one α-helix of Mei4 slides) and the aim was to test the interaction domains between the proteins and eventually identify the amino acids important in this protein-protein interaction based on the predicted structures. The following minimal complexes were studied: REC114$^{210\text{-}259}$-MEI4$^{1\text{-}58}$ (*M. musculus*), Rec7$^{289\text{-}339}$-Rec24$^{1\text{-}50}$(*S. pombe*), PHS1$^{260\text{-}310}$-PRD2$^{1\text{-}50}$ (*A. thaliana*) and PHS1$^{297\text{-}347}$-MPS1$^{1\text{-}87}$ (*Z. mays*). The structural models were partially confirmed by mutagenesis for the mouse, *S. pombe* and *A. thaliana* complexes, where the structural function of the F411 amino acid in *S. cerevisiae* appears to be preserved, whereas the model proposed by AlphaFold was not confirmed by our mutagenesis for the *Z. mays* complex. However, it is important to note that the PHS1 and MPS1 proteins of *Z. mays* were able to be copurified whereas no study has yet shown their interaction, which means that the interaction domains are still conserved and that MPS1 can be the Mei4 ortholog. The homotetrameric and parallel structure of the Mer2 coiled-coil domain appears to be conserved for the orthologs tested: IHO1 (*M. musculus*), Rec15 (*S. pombe*), PRD3 (*A. thaliana*), PAIR1 (*Z. mays*) and ASY2 (*S. macrospora*).

Concerning the conservation of the binding function of the RMM complex to DNA, it was shown that the mouse RMM complex conserves this function and have a preference for branched DNA substrates which highlight on the multivalent side of the interactions between DNA and proteins. The residues essential for DNA binding of the Mer2 ortholog, IHO1, are located in the N-terminal part of the coiled coil. Concerning the *S. pombe* RMM proteins, this study showed that the Mer2 ortholog, Rec15, was bound to DNA and that it prefers branched DNA. We also show that de C-terminal part of the coiled coil contains the essential residues for this binding., whereas the Rec7-Rec24 minimal complex was not sufficient to draw any conclusions. In the *A. thaliana* complex, the coiled-coil domain of PRD3 did not show clear binding to DNA. However, the minimal parts of the PHS1-PRD2 complex were sufficient to bind DNA, so the DNA-binding function was retained. For the *Z. mays* RMM complex, the coiled-coil domain of PAIR1 is sufficient to bind DNA whereas no DNA binding could be detected for the RM (PHS1-MPS1) complex. Finally,

only the Mer2 ortholog of *S. macrospora* was studied and the coiled-coil domain of ASY2 did not show DNA binding. No orthologs to the Rec114-Mei4 complex were found in *S. macrospora*.

VII Material and method

1. Solutions

Table 2: Table of solutions used in the experiments of this master thesis

Solution	Composition
Buffer A40	25 mM HEPES-NaOH pH 7.5, 500 mM NaCl (or 300 mM NaCl if specified), 10% glycerol, 0.1 mM DTT, 40 mM imidazole, sterile water to bring to desired volume
Buffer A40 + PMSF	Buffer A40 + 0.4 mM PMSF fresh before use
Buffer A40 + Triton	Buffer A40 (with 300 mM NaCl) + 0.5% Triton x100
Buffer Amylose	25 mM HEPES-NaOH pH 7.5, 500 mM NaCl, 10% glycerol, 0.1 mM DTT, sterile water to bring to desired volume
Buffer Amylose Elution	Buffer Amylose with 10mM maltose
Buffer C	25 mM HEPES-NaOH pH7.5, 500 mM NaCl, 10% glycerol, 1 mM DTT, 5mM EDTA, sterile water to bring to desired volume
Buffer B500	Buffer A40 with 500 mM imidazole final
Coomassie Blue Staining Solution	0.1% Coomassie R250, 10% glacial acetic acid, 40% methanol, sterile water to bring to desired volume
Destain	50% sterile water, 40% methanol, 10% glacial acetic acid

5x Ficoll Loading Buffer	0.01 M Tris-HCl pH 8, 0.001M EDTA, 0.2 g/ml Ficoll, 2 mg/ml Bromophenol Blue, 2 mg/ml Xylene Cyanol FF, 0.1% SDS, sterile water to bring to desired volume
Laemmli Buffer	50 mM Tris-HCl pH 6.8, 2% SDS, 0.1 % Bromophenol Blue, 40% glycerol, 50 µl/ml β-mercaptoethanol
LB	10 g/l Bacto-tryptone, 5g/l Bacto-yeast extract, 10 g/l NaCl, sterile water to bring to desired volume
Lysis Buffer (Insect Cells)	50mM HEPES pH 7.5, 1 :500 PI green, 1:500 PI red, 1 mM PMSF, 0.17 mM DTT, 65 mM imidazole, sterile water to bring to desired volume
MES Buffer	50 mM MES, 50 mM Tris, 1 mM EDTA, 0.1% (w/v) SDS
SOC	2% Bacto-tryptone, 0.5 % Bacto-yeast extract, 10mM NaCl, 2.5 mM KCl, 10 mM $MgCl_2$, 10 mM $MgSO_4$, 20 mM Glucose (Dextrose), sterile water to bring to desired volume
Special TAE	0.1 mM EDTA, 4 mM Tris at pH 8.3, 4 mM NaOAc, 0.2 mM MgOAc
TBE	108 g/l Tris, 55 g/l Boric acid, 0.04 M Na_2EDTA (pH 8.0), sterile water to bring to desired volume
TE	10 M Tris-HCl pH 8, 1 mM EDTA, sterile water to bring to desired volume
Tris-Glycine	(4 x concentrated) 12 g/l Tris, 1.5 g/l Na_2EDTA, 4 g/l SDS, 57.6 g/l glycine, sterile water to bring to desired volume

2. PCR primers

Table 3: *Table of primers used to create various plasmids*

Primer name	Sequence	Direction	Use	Organism
CCB1486	ggattggaagtacaggttttc	Reverse	Open pDD092 and pDD094 without MEI4	*Mus musculus* and *A. thaliana* or *Z. mays*
CCB1500	tccaccaatctgttctctg	Reverse	Open pDD066 without REC114	*Mus musculus*
CCB1501	attcgagctcggcgc	Forward	Open pDD066 without REC114	*Mus musculus*
CCB1504	ctcgagtctggtaaagaaac	Forward	Open pDD092 and pDD094 without MEI4	*Mus musculus* and *A. thaliana* or *Z. mays*
DD127	cagcagctgtggaagaggtg	Forward	REC114-F243A	*Mus musculus*
DD128	ggaagttttgatccatgag	Reverse	REC114-F243A	*Mus musculus*
DD129	caaaacgccccagcatttg	Forward	REC114-F240A	*Mus musculus*
DD130	atccatgagacataagcg	Reverse	REC114-F240A	*Mus musculus*
EDJ12	gcagtagaggcaattatcgacgaaattg	Forward	PHS1-K296A, K299A	*A. thaliana*
EDJ13	ctgcagcatatcttgaaatgaagaatc	Reverse	PHS1-K296A, K299A	*A. thaliana*
EDJ14	gcgtcgcagatactgaag	Forward	PHS1-K278A	*A. thaliana*
EDJ15	tagatcaggattctgc	Reverse	PHS1-K278A	*A. thaliana*
EDJ18	gccatccaagcatgtcaaagtttggaac	Forward	Rec7-Y320A, L328A	*S. pombe*
EDJ19	gctgtcttgttttagagcgaataataccttg	Reverse	Rec7- Y320A, L328A	*S. pombe*
EDJ20	cttcagctcaagatatgctgcag	Forward	PHS1-F290A	*A. thaliana*
EDJ21	aatcttccatgtacttcagtatc	Reverse	PHS1-F290A	*A. thaliana*
EDJ22	gaagaatcttccatggccttcagtatctgcg	Reverse	PHS1- Y284A, L294A	*A. thaliana*

EDJ23	agctcaagatatggcgcagaaag tagagag	Forward	PHS1-Y284A, L294A	*A. thaliana*
EDJ24	aacatatcattagcggattcatca gccatatatg	Reverse	PHS1-F327A	*Z. mays*
EDJ25	gctcaagcttgataaagccattg	Forward	PHS1-F327A	*Z. mays*
EDJ26	gctcaaggctgataaagccgctga tgaacttggtg	Forward	PHS1-L334A, I338A	*Z. mays*

3. Plasmids

Table 4: *Table of the different plasmids used in this master thesis. All protein sequences are in Appendix 3.*

Plasmids	**Description**	**Antibiotic resistance marker**	**Organism**
pCCB982	HisSUMO-IHO1$^{109-267}$	Kan	*Mus Musculus*
pCCB983	HisSUMO-REC114, MBP-MEI4 (full-length)	Amp	*Mus musculus*
pCCB990	HisSUMO-PRD3$^{120-270}$	Kan	*Arabidopsis thaliana*
pCCB991	HisSUMO-Rec15$^{1-160}$	Kan	*Schizosaccharomyces pombe*
pCCB992	HisSUMO-ASY2$^{55-275}$	Kan	*Sordaria macrospora*
pCCB993	HisSUMO-PAIR1$^{140-310}$	Kan	*Zea mays*
pDD066	HisSUMO-REC114$^{210-259}$, MBP-MEI4$^{1-58}$	Amp	*Mus musculus*
pDD085	HisSUMO-Rec7$^{289-339}$, MBP- Rec24$^{1-50}$	Amp	*Schizosaccharomyces pombe*
pDD086	HisSUMO-Rec15$^{1-160}$ (K134, K135, K141, K142, R143A)	Kan	*Schizosaccharomyces pombe*
pDD091	HisSUMO-Rec7$^{289-339}$ (F325A), MBP-Rec24$^{1-50}$	Amp	*Schizosaccharomyces pombe*

pDD092	HisSUMO-PSH1$^{260-310}$, MBP-MEI4$^{1-58}$	Amp	*Arabidopsis thaliana* / *Mus musculus*
pDD093	HisSUMO-PSH1$^{260-310}$, MBP-PRD2$^{1-50}$	Amp	*Arabidopsis thaliana*
pDD094	HisSUMO-PSH1$^{297-347}$, MBP-MEI4$^{1-58}$	Amp	*Zea mays* / *Mus musculus*
pDD095	HisSUMO-PSH1$^{297-347}$, MBP-MPS1$^{1-87}$	Amp	*Zea mays*
pEDJ07	HisSUMO-PSH1$^{260-310}$ (K296, K299A), MBP-PRD2$^{1-50}$	Amp	*Arabidopsis thaliana*
pEDJ09	HisSUMO-PSH1$^{260-310}$ (K278, K296, K299A), MBP-PRD2$^{1-50}$	Amp	*Arabidopsis thaliana*
pEDJ10	HisSUMO-Rec7$^{289-339}$ (Y320, F325, L328A), MBP- Rec24$^{1-50}$	Amp	*Schizosaccharomyces pombe*
pEDJ11	HisSUMO-PSH1$^{260-310}$ (F290A), MBP-PRD2$^{1-50}$	Amp	*Arabidopsis thaliana*
pEDJ12	HisSUMO-PSH1$^{297-347}$ (F327A), MBP-MPS1$^{1-87}$	Amp	*Zea mays*
pEDJ13	HisSUMO-PSH1$^{260-310}$ (Y284, F290, L294A), MBP-PRD2$^{1-50}$	Amp	*Arabidopsis thaliana*
pEDJ14	HisSUMO-PSH1$^{297-347}$ (F327, L334, I338A), MBP-MPS1$^{1-87}$	Amp	*Zea mays*

3.1 Tags fused protein

Proteins were fused to different tags at their N-terminal part. Either the HisSUMO (Small Ubiquitin-like Modifier) tag (for Rec114 and Mer2 orthologs) or the MBP (Maltose binding protein) tag (for Mei4 orthologs). Indeed, these tags are amino acid sequences, which are fused to the amino acid sequence of the proteins of interest and which aim to facilitate their purification. Indeed, the HisSUMO tag is retained on a Nickel column and can be eluted by saturating the Nickel column with imidazole. The MBP tag is retained on an Amylose column and can be eluted with a maltose rich solution.

The HisSUMO tag consists of 122 amino acids and has a molecular weight of 13.8 kDa. It has the characteristic of having 6 histidines, allowing for better affinity and specificity to a Nickel column. HisSUMO allows better expression and solubilization of the protein. The HisSUMO tag can be cleaved by a SUMO protease such as Ulp1 (Ubl-specific

protease 1)(« SUMO Tag and SUMO Tag Purification | Sino Biological » s. d.; « SUMO Protease Small Ubiquitin-Like Modifier Protease Sigma » s. d.) .

The MBP tag consists of 393 amino acids and has a molecular weight of 43 kDa. This tag is also known to increase the solubility of proteins to which it is fused (« MBP Affinity Tag | NEB » s. d.). To this tag is fused a TEV cleavage site, which is a sequence recognized by the TEV enzyme, which will come and cleave the protein at this point. This allows the TAG to be removed. This sequence contains 7 amino acids.

3.2 Use of the operon LacO

For each of these plasmids, the gene for the proteins of interest is preceded by the gene for the lactose operon LacO:

Sequence: 5'-GGAATTGTGAGCGGATAACAATT- 3'

This makes protein expression inducible by adding IPTG to the culture medium.

4. Preparation of expression vectors

All plasmids whose synthesis is not detailed below have been synthesized by Dima Daccache.

The REC114 C-terminal domain from pDD066 was replaced by the C-terminal part of PHS1 of *A. thaliana* and of *Z. mays* using primers CCB1500 and CCB1501 and cloned into pDD066 vector by Gibson assembly to yield pDD092 and pDD094 respectively. The gene coding for MEI4 was replaced in pDD092 by the gene coding for PRD3 (*A. Thaliana)* to obtain plasmid pDD093 and by the gene coding for MPSI (*Z. mays*) in pDD094 to obtain plasmid pDD095 and it was carried out by the primers CCB15004 and CCB1486 and by a Gibson assembly.

Other plasmids were prepared by a reverse Polymerase Chain Reaction. In a PCR tube, 1 µl of dNTP, 0.5 µl of template, 10 µl of Q5 reaction buffer®, 10 µl of enhancer buffer®, 1.5 µl of each primer, 26.5 µl of sterile water and 0.5 µl of Q5 enzyme were mixed or 2 µl of dNTP, 2.5 µl of each primer, 10 µl of Phusion® Buffer, 0.3 µl of template and 31.7 µl of sterile water were mixed with 1µl of Phusion polymerase depending on the plasmid (Table 5).

Each sample was put at a denaturation temperature of 98°C for 30 s. Then, 35 cycles of denaturation at 98°C for 20 s, hybridization at 56°C for 30 s and polymerization at 72°C for 4 min. Finally, the elongation was at 72°C for 7 minutes.

Table 5: *PCR template, primers and polymerase used to create plasmids where a mutation in the ortholog of Rec114 is added.*

Desired plasmid	Polymerase	Template	Forward primer	Reverse primer	Added mutation in Rec114 ortholog	Organisms of protein
pCCB988	Q5	pCCB984	DD129	DD130	F240A	*M. musculus*
pCCB989	Q5	pCCB984	DD127	DD128	F243A	*M. musculus*
pEDJ10	Phusion	pDD091	EDJ18	EDJ19	Y320A, L328A	*S. pombe*
pEDJ11	Phusion	pDD093	EDJ20	EDJ21	F290A	*A. thaliana*
pEDJ12	Phusion	pDD095	EDJ25	EDJ24	F327A	*Z. mays*
pEDJ13	Phusion	pEDJ11	EDJ23	EDJ22	Y284A, L294A	*A. thaliana*
pEDJ14	Phusion	pEDJ12	EDJ26	EDJ24	L334A, L338A	*Z. mays*

After the PCR: The template was digested with DpnI (30 µl of PCR product with 9.5 µl of sterile water, 4.5 µl of Cutsmart® and 1 µl of DpnI) and incubated for 1 h at 37°C. 9 µl of 5x Ficoll loading buffer was added to the sample and the sample was migrated in a 1% Agarose gel with 20000-fold diluted Midori Green and TBE with 100 V current for 40 min. Then the correct size DNA band was extracted from the gel by a gel extraction (Protocol in the Appendix 4). Once done, the ends of the amplified DNA fragment were phosphorylated by adding 3 µl of ligase buffer (10 mM ATP) and 0.5 µl of PNK enzyme. After a 45-min incubation at 37°C, the enzyme is deactivated by incubating the solution at 65°C for 20 min. Finally, a ligation of the phosphorylated DNA ends is performed by adding 0.3 µl of PNK Buffer and 0.5 µl of ligase and incubating at RT for 1 hour. The plasmid created is transformed into 100 µl of thermocompetent *E. coli* cells (strain TOP10) following the protocol in point VII.5. A pre-culture is performed the next day following point VII.6. and the day after, a miniprep is performed to isolate the plasmid DNA (protocol in Appendix 5).

5. Transformation of plasmids into competent cells

0.5 to 1 µl of solution containing the plasmid of interest is added to 100 µl of thermocompetent *E. coli* cells (strain BL21). The tubes are left on ice for 15 min. After a heat shock at 42 °C for exactly 45 s, the cells were leaved on ice for 2 min. 950 µl of SOC medium and the solution are transferred to a 14 ml round-bottom tube. After an incubation of 1 h at 37°C, 100 µl of solution is spread on a petri dish with LB and the appropriate antibiotic followed by an incubation at 37°C O/N.

6. Preculture

One of the colonies isolated from the petri dish is inoculated into 5 mL of liquid LB with 100 µg/ml ampicillin or 50 µg/ml kanamycin. Incubate at 37°C O/N with shaking at 120 rpm.

7. Induction of protein expression

3 ml of the preculture is added to a 5 liters Erlenmeyer flask containing 1 liter of LB with 100 µg/ml ampicillin or 50 µg/l kanamycin and antifoam for protein purification and 150 µl of preculture is added to a 50 ml Erlenmeyer for the Pulldown experiment. Incubation at 37°C with 120 rpm agitation is maintained until an optical density (OD) measurement of 0.6 is obtained. Once this stage is reached, protein expression is induced with IPTG 1mM. Induction is stopped after an incubation time of at least 3 h at 30°C with an agitation of 120 rpm. The medium containing the bacteria is centrifuged at 5000 rpm for 10 min at 4°C and the pellet is resuspended in 40 ml of Buffer A40 or in 10ml of Buffer A40 + Triton for the Pulldown experiment. The whole is poured into a 50 ml flask and centrifuged again at 4200 rpm for 10 min at 4°C. Cell pellets were flash frozen in liquid nitrogen and stored at -80°C.

8. Purification of mRM and cc of Mer2 orthologs

8.1 Tagged mRM purification (pDD085, pDD093, pDD095)

First column: The bacterial pellets are resuspended in 30 ml of Amylose Buffer + PMSF and sonicated on ice + water (5 s pulse, with a magnitude of 50 W/cm^2) for 4 min.

The lysed cells are placed in a Ti45 ultracentrifuge tube and centrifuged for 25 min at 30000 rpm at 4°C. During this time, 3 ml of well-mixed Amylose resin is taken (this is equivalent to approximately 1 ml of resin which allows the purification of 40 mg of protein) and is transferred to a 50 ml Falcon and spin for 2 min at 1300 rpm. The ethanol on the surface is aspirated and the resin is then washed with 45 ml of Amylose Buffer + PMSF (spin for 2 min at 1300 rpm and the supernatant is aspirated). The resin washing step is repeated once. Once the ultracentrifuge is complete, the supernatant is added to the Falcon containing the resin and the Falcon is incubated for 30 min at 4°C on the hot dog roller. The Falcon is then centrifuged for 2 min at 1300 rpm and the supernatant is collected as a control. The resin is then washed with 40 ml of Amylose Buffer (incubated for 3 min at 4°C on the hot dog roller, spin for 2 min at 1300 rpm and the supernatant is aspirated). The resin is transferred to a 10 ml disposable chromatography column. Once the resin is sedimented by gravity, 3 washes of the column are performed with Amylose Buffer. Protein elution is performed by adding 4 ml of Amylose Elution Buffer (fractionated into 500 µl). The presence of proteins is verified by an SDS-PAGE gel using a Mini-PROTEAN TGX Gel 4-15% (with Tris-Glycine buffer) with 30 min of migration with a current of 160V. The gel is revealed by Coomassie Blue staining solution and decolorizing in Destain solution. Fractions are then selected and put on a second column.

Second column: 3 ml of Ni-NTA agarose resin (Qiagen) is pipetted and transferred to a 50 ml Falcon and washed with Buffer A40 twice in the same way as for the first column. The Amylose elution is added to the falcon and incubated for 30 min at 4°C on the Hot-dog roller. The falcon is then centrifuged at 1300 rpm for 2 min at 4°C and the supernatant is collected as a control. The resin is then washed once with A40 buffer and transferred to a 20 ml disposable chromatography column. The resin is sedimented by gravity and then washed 3 times with A40 buffer. Elution takes place using 3 ml of Buffer B500, fractionated into 500 µl fractions. Again, protein purification is verified by an SDS-PAGE gel. After a 1 h dialysis in Buffer C, proteins were concentrated with a 10 kDa cutoff Amicon centrifugal filters (Millipore) and proteins were flash frozen in liquid nitrogen and stored at -80°C.

8.2 Tagged and untagged cc of Mer2 orthologs purification (pCCB990, pCCB991, pCCB992, pCCB993, pDD086)

First Ni column: Bacterial pellets are resuspended in 30 ml Buffer A40 + PMSF and sonicated on ice + water (5 s pulse, with a magnitude of 50 W/cm^2) for 4 min. The lysed cells are placed in a Ti45 ultracentrifuge tube and centrifuged for 25 min at 30,000 rpm at 4°C. During this time, 3 ml of well-mixed Ni-NTA agarose resin (QIAgen) is taken. This resin is transferred to a 50 ml Falcon and spin for 2 min at 1300 rpm. The ethanol on the surface is aspirated and the resin is then washed with 45 ml of Buffer A40 + PMSF

(spin for 2min at 1300 rpm and the supernatant is aspirated). The resin washing step is repeated once. Once the ultracentrifuge is complete, the supernatant is added to the Falcon containing the resin and the Falcon is incubated for 30 min at 4°C on the hot dog roller. The Falcon is then centrifuged for 2 min at 1300 rpm and the supernatant is collected in another Falcon and kept as a control. The resin is then washed with 40 ml of Buffer A40 (incubated for 3 min at 4°C on the Hot dog roller, spun for 2 min at 1300 rpm and the supernatant is aspirated). The resin is transferred to a 10 ml disposable chromatography column. Once the resin is sedimented by gravity, 3 washes of the column are performed with A40 buffer. Protein elution is performed by adding 4 ml of Amylose Elution Buffer (fractionated into 500 µl fractions).

The presence of proteins is verified by an SDS-PAGE gel using a Mini-PROTEAN TGX Gel 4-15% (with Tris-Glycine buffer) with 30 min of migration with an electric current of 160V. The gel is revealed by Coomassie Blue staining solution and decolorizing in Destain solution.

If the gel shows the expected proteins, dialysis is performed with a 10 kDa cassette in Buffer A40 over night to decrease the imidazole concentration.

Second Ni column (reverse Ni): The next day, the 4 ml of the elution from the first Ni column is separated into 2 (one half will be the proteins retaining the HisSUMO tag and for the other half, the tag will be cleared). For the half where the tag is to be removed, 5 µl of UlpI enzyme is added and the solution is incubated for 30 min at 4°C on the hot dog roller. A reverse Nickel chromatography is prepared by pipetting 3ml of well-mixed Ni-NTA agarose resin (QIAgen) into a 15 ml Falcon. The resin is washed once with Buffer A40. Once the incubation with UlpI is complete, the sample is incubated on the resin for 30 min at 4°C on the hot dog roller. The solution is transferred to a 10 ml disposable chromatography column. The flow through is collected and 600 µl of Buffer A40 (representing the resin void) is added. The cleavage of the tag is checked via a Mini-PROTEAN TGX Gel 4-15%. Finally, a 1 h dialysis in Buffer C (for both tagged and untagged proteins) is performed and proteins were flash frozen in liquid nitrogen and stored at -80°C.

8.3 Purification of the tagged cc for the SAXS experiment

Only the first column step of the coiled coil purification (with Buffer A40 with only 5% glycerol) is performed and followed by dialysis in Buffer C for one hour. The 4 ml of elution is then concentrated to 1 ml with a 10 kDa cutoff Amicon centrifugal filters (Millipore). This sample is passed through a Costar® to get rid of aggregates, if any, so as not to block the column and followed by a size exclusion chromatography. The machine

used is an AKTA pure with a Superose 6 Increase 10/300 GL column and a flow rate of 0.450 ml/min and the results were analyzed with a chromatogram and an SDS-PAGE gel.

After analysis of the chromatography peaks, the protein-containing fractions are separated and the samples are concentrated to 1ml and placed in liquid nitrogen and stored at -80°C.

The SAXS experiment was realized by Yann Sterckx and his team.

9. Purification of full-length mouse HisSUMO-REC114-MBP-MEI4

At the time of my arrival in the lab, the insect cell pellets were already stored at -80°C.

The pellet is resuspended in 26 ml of lysis buffer and thawed on a wheel or in a beaker by gently mixing with a magnetic bar. 30 ml of 50% glycerol (cold) is added followed by 10 ml of 5 M NaCl (cold) while mixing the sample. After 30 min incubation on ice with low agitation, the solution is centrifuged at 48000 g for 30 min. Meanwhile, the Ni-NTA resin (3 ml) is prepared in a 50 ml Falcon and washed with 40 ml A40 buffer by centrifuging 2 min at 1000 rpm and discarding the supernatant. The sample supernatant is incubated on the resin for 30 min at 4°C on the hot dog roller. 100 µl of PMSF is added into each tube and the tubes are centrifuged at 1000 rpm for 2 min and the supernatant is discarded. The resin is washed once with 40 ml buffer A40 + PMSF, PI green and PI red. The resin is transferred to a 10 ml disposable chromatography column. The Falcon is rinsed with A40 buffer and this is used to wash the column. 6 times 1 ml of Buffer B500 + PMSF are used to elute proteins and 15 µl of each fraction and 5 µl of Laemmli are taken for an SDS-page gel. During Ni column wash, 2 ml of Amylose resin is transferred to a disposable 10 ml chromatography column. The column is washed twice with 10 ml of Amylose Buffer. The elution of the Ni column is loaded on the Amylose column and the flow through is allowed to pass. The resin is washed 3 times with 10 ml of Amylose Buffer and the proteins are eluted with 6 ml of Amylose Elution Buffer. The purified proteins are analyzed via a Mini PROTEAN TGX Gel 4-15% with a migration of 30 min at 160 V.

10. Pulldown of mRM orthologs proteins

To show the interaction between two proteins, a Pulldown experiment is performed. Indeed, this technique makes it possible to co-purify two or more proteins that form a complex, which interact together.

The pellet is resuspended in 1.3 ml of Buffer A40 + Triton and transferred to a 2 ml Eppendorf tube. Bacteria are sonicated on ice with water at a 5 s pulse of magnitude 50 W/cm^2 for 2 min. The tubes are centrifuged for 20 min at 15000 rpm at 4°C. During this time, the resin is prepared by pipetting 200 µl of Ni-NTA resin into a disposable Bio-Spin column. The resin sedimented by gravity and the resin washed once by filling the column with Buffer A40 + Triton. The sample supernatant is incubated on the resin for 30 min at 4°C on a wheel. 5 washes of the column are performed with Buffer A40 + Triton. The remaining proteins on the resin are eluted with 400 µl of Buffer B500. The protein elution is verified by an SDS-PAGE gel.

11. EMSA

EMSA, electrophoresis mobility shift assay, is a technique for analyzing the binding of a protein to DNA by migrating a solution through a polyacrylamide gel.

The composition of the gel is as follows: 5% polyacrylamide, TAE 1x, 0.66 % APS and 1/1000 (v/v) TEMED. The gel is run in a TAE 1x buffer.

The reaction (20µl) is composed of 10 nM HJ40 labelled with a fluorophore, 25 mM Tris, 1 mg/ml BSA, 7.5% glycerol, 100 mM NaCl, 2mM DTT and the indicated protein concentration. For the different protein dilutions tested, the proteins are always diluted in Buffer C. The tubes are incubated for 30 min at 30°C with finger vortexing every 10 min. The gel is loaded and run in a cold room at 4°C and the gel migrates for 150 min at 150 V. Fluorescent gels are then visualized using a Typhoon scanner (Cytiva).

11.1 Competition EMSA

The EMSA testing the competition between two DNA substrates follows the same principle as a classical EMSA with the exception that the reaction is composed of 10 nM HJ40 labelled with a fluorophore, 25 mM Tris, 1 mg/ml BSA, 7.5% glycerol, 100 mM NaCl, 2mM DTT, the indicated protein concentration and the indicate non-fluorophore labelled competitor substrate concentration (HJ40 or 80 bp dsDNA).

12. Crosslinking-Mass spectrometry (XL-MS)

In order to test the structure of the proteins, a Crosslinking-Mass spectrometry experiment was performed. In order to crosslink the proteins, 60 µg of proteins were used. DTT was added to give 2 mM and the solution was incubated for 10 min at RT. DSS,

the crosslinking agent, was then added to reach 1mM and the whole was incubated for 15 min at 30°C. To quench the unreacted DSS, Tris at pH 7.5 was added and allowed to react for 10-15 min. The sample was then stored at -20°C. In order to verify that the XL had worked, an SDS-PAGE gel of this sample compared to the non-XL protein, was run.

The proteins were digested with trypsin and analyzed by mass spectrometry by Hervé Degand (Figure 53).

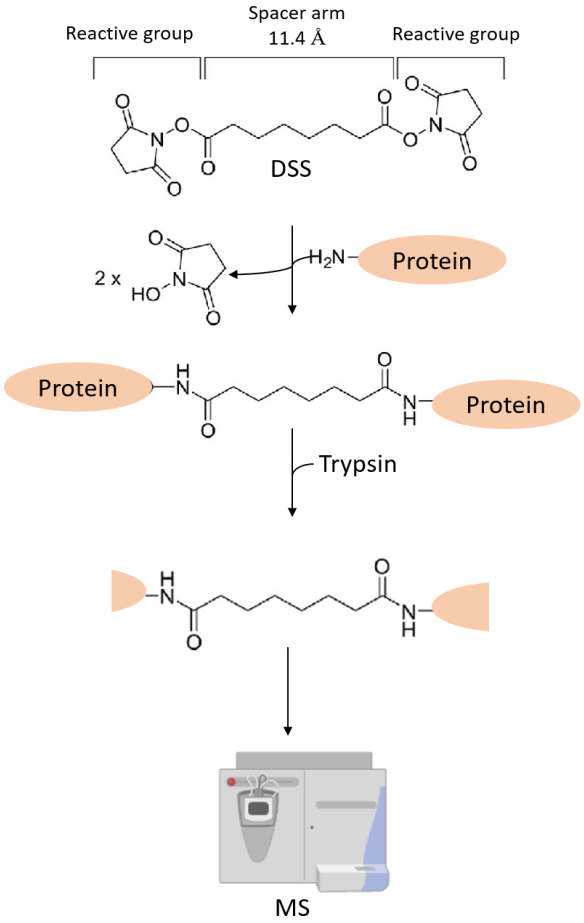

Figure 53: XL-MS experiment synthesis. *Diagram showing crosslink using the crosslinking agent DSS, composed of two reaction groups and a 11.4 Å spacer. The cross-linking agent will react with two lysines of different proteins or of the same protein and then the proteins will be digested into peptides by trypsin. The different fragments are then analyzed by mass spectrometry (modified from Calabrese and Radford 2018).*

12.1 XL-MS in dilute condition

The same reaction conditions as explained above are used except that the 60 µg of protein is diluted 10-fold.

Then, in order to concentrate the sample, acetone precipitation was carried out: The sample containing the cross-linked protein is separated into several 1.5 ml Eppendorf tubes. 4 times as much cold acetone (-20°C) as protein solution was added.

The tubes were vortexed and incubated for 60 min at -20°C before being centrifuged for 10 min at 13000 rpm. The supernatant was discarded and the tubes were left open at RT for a few moments to allow the remaining acetone to evaporate. The pellets were resuspended to obtain sufficient volume to run the sample on the MS.

13. Coiled coil stability test in different buffers

Dialysis O/N at 4°C of 10 µl of solutions containing the coiled coils in 1.9 ml buffer containing the following concentrations

1) 0 mM NaCl, 20 mM Tris, sterile water to bring to desired volume
2) 50 mM NaCL, 20 mM Tris, sterile water to bring to desired volume
3) 100 mM NaCl, 20 mM Tris, sterile water to bring to desired volume
4) 200 mM NaCl, 20 mM Tris, sterile water to bring to desired volume

The dialysis buffer is renewed in the morning and left for at least 3h at RT. The samples containing the coiled coils are recovered and centrifuged for 10 min at 15000 rpm. The supernatant is pipetted and the pellet is resuspended in 10 µl of dialysis buffer. The samples are analyzed on an SDS-PAGE gel.

14. Quantification of bands on gel

ImageJ software was used to quantify the intensity of bands formed by proteins on an SDS-PAGE gel for Pulldowns or bands formed by fluorophore-labelled DNA on a polyacrylamide gel for EMSAs.

VIII Appendix

1. Distances measured of XL on the cc

Table 6: *Table showing the XL distances measured considering intramolecular XL on the coiled-coil domains orthologs of Mer2 and the number of times a XL is detected (Hitcount). c*

Chain	Residue 1	Residue 2	Hitcount	Intramolécular Distance (Å)	Threshold (27Å) satisfied ?
IHO1	4	8	17	6,2	S
IHO1	4	14	4	15,3	S
IHO1	4	39	83	51,8	V
IHO1	4	46	28	61,8	V
IHO1	4	67	13	92,6	V
IHO1	4	87	4	121	V
IHO1	8	13	4	8,5	S
IHO1	8	39	91	45,9	V
IHO1	8	46	9	56	V
IHO1	8	67	16	87	V
IHO1	8	87	1	115,6	V
IHO1	8	116	2	157,2	V
IHO1	13	39	39	38,7	V
IHO1	13	46	3	48,9	V
IHO1	14	39	2	36,9	V
IHO1	39	46	75	10,3	S
IHO1	46	67	9	31,6	V
IHO1	67	67	11	0	S
IHO1	67	87	18	30,4	V
IHO1	67	112	10	67,7	V
IHO1	87	87	10	0	S
IHO1	87	112	2	37,4	V

IHO1	87	150	4	93,3	V
IHO1	149	150	16	3,8	S
PRD3	1	95	39	141,4	V
PRD3	15	15	53	0	S
PRD3	15	56	57	61,3	V
PRD3	15	68	27	79,7	V
PRD3	15	81	25	99,1	V
PRD3	15	87	30	107,2	V
PRD3	15	95	14	119	V
PRD3	15	103	5	131,6	V
PRD3	15	112	11	147,1	V
PRD3	15	130	12	173,6	V
PRD3	56	56	26	0	S
PRD3	56	68	56	18,9	S
PRD3	56	81	10	38,5	V
PRD3	56	87	37	47,3	V
PRD3	56	95	35	59,2	V
PRD3	56	103	15	71,6	V
PRD3	56	109	23	82,8	V
PRD3	56	112	37	87,2	V
PRD3	56	130	40	113,6	V
PRD3	68	68	16	0	S
PRD3	68	81	31	20,2	S
PRD3	68	87	32	28,8	V
PRD3	68	95	15	40,7	V
PRD3	68	103	21	53	V
PRD3	68	130	8	94,8	V
PRD3	81	95	16	21	S
PRD3	81	103	2	33,2	V
PRD3	81	109	11	44,7	V
PRD3	81	112	31	48,9	V

PRD3	81	130	34	75,4	V
PRD3	87	103	18	24,4	S
PRD3	87	109	31	35,8	V
PRD3	87	112	23	40	V
PRD3	87	130	23	66,5	V
PRD3	95	103	6	12,6	S
PRD3	95	112	3	28,3	V
PRD3	103	103	8	0	S
PRD3	103	109	13	12,4	S
PRD3	109	112	20	5	S
Rec15	1	134	41	198,6	V
Rec15	14	14	44	0	S
Rec15	14	27	24	19,8	S
Rec15	14	96	30	124,5	V
Rec15	14	107	32	140,5	V
Rec15	14	113	15	149,2	V
Rec15	14	120	11	159,2	V
Rec15	14	134	11	179,2	V
Rec15	14	135	5	181,6	V
Rec15	14	141	6	188,8	V
Rec15	14	148	20	198,6	V
Rec15	27	135	5	163,2	V
Rec15	27	148	18	180,5	V
Rec15	43	113	1	106	V
Rec15	43	141	2	146	V
Rec15	43	148	4	155,8	V
Rec15	96	96	37	0	S
Rec15	96	107	37	16,5	S
Rec15	96	113	17	26,3	S
Rec15	96	135	15	59,3	V
Rec15	96	141	4	67,8	V

Rec15	96	148	32	77,9	V
Rec15	107	107	8	0	S
Rec15	107	113	27	10,2	S
Rec15	107	120	6	20,6	S
Rec15	107	135	12	42,9	V
Rec15	107	141	8	51,6	V
Rec15	107	148	46	61,7	V
Rec15	113	113	4	0	S
Rec15	113	135	6	33,9	V
Rec15	113	148	9	53,2	V
Rec15	120	135	10	23,1	S
Rec15	120	148	11	42,4	V
Rec15	134	134	14	0	S
Rec15	134	135	6	3,9	S
Rec15	134	141	10	10,5	S
Rec15	134	148	15	21	S
Rec15	135	135	6	0	S
Rec15	135	141	6	10	S
Rec15	141	148	10	10,5	S
Rec15	148	148	19	0	S
PAIR1	21	31	42	14,9	S
PAIR1	21	51	15	44,9	V
PAIR1	21	63	40	63	V
PAIR1	21	73	13	78,1	V
PAIR1	21	76	6	82	V
PAIR1	21	82	5	90,1	V
PAIR1	21	90	18	102,3	V
PAIR1	21	105	16	125,5	V
PAIR1	21	118	9	144	V
PAIR1	21	125	9	154,3	V
PAIR1	21	147	4	185,5	V

PAIR1	21	152	15	194,3	V
PAIR1	31	51	62	30,9	V
PAIR1	31	73	47	64	V
PAIR1	31	76	8	68,1	V
PAIR1	31	118	2	130	V
PAIR1	31	125	32	140,3	V
PAIR1	31	147	36	171,4	V
PAIR1	31	152	25	180,1	V
PAIR1	51	51	43	0	S
PAIR1	51	63	4	18,9	S
PAIR1	51	73	105	34,1	V
PAIR1	51	76	8	38,5	V
PAIR1	51	82	2	47,4	V
PAIR1	51	90	12	59,5	V
PAIR1	51	105	30	83,2	V
PAIR1	51	118	30	102,7	V
PAIR1	51	125	24	112,6	V
PAIR1	51	147	13	143,6	V
PAIR1	51	152	28	152,2	V
PAIR1	63	63	7	0	S
PAIR1	63	73	102	15,6	S
PAIR1	63	76	20	20,3	S
PAIR1	63	90	10	40,9	V
PAIR1	63	118	9	83,9	V
PAIR1	63	125	24	93,8	V
PAIR1	63	147	29	124,8	V
PAIR1	63	152	68	133,3	V
PAIR1	73	76	23	5,1	S
PAIR1	73	82	54	14,3	S
PAIR1	73	90	4	25,7	S
PAIR1	73	125	26	79,4	V

PAIR1	73	147	22	110,3	V
PAIR1	73	152	27	118,7	V
PAIR1	76	76	2	0	S
PAIR1	76	90	9	21,3	S
PAIR1	76	105	21	45,5	V
PAIR1	76	118	17	65,8	V
PAIR1	76	125	9	75,5	V
PAIR1	76	152	11	114,8	V
PAIR1	82	105	7	36,1	V
PAIR1	82	118	11	56,2	V
PAIR1	82	147	10	97,1	V
PAIR1	82	152	11	105,6	V
PAIR1	90	105	2	24,4	S
PAIR1	90	118	10	45,1	V
PAIR1	90	147	10	85,4	V
PAIR1	90	152	2	93,8	V
PAIR1	105	105	3	0	S
PAIR1	105	118	23	21,1	S
PAIR1	105	125	21	30,2	V
PAIR1	105	147	43	61,2	V
PAIR1	105	152	20	69,6	V
PAIR1	118	118	19	0	S
PAIR1	118	125	60	10,5	S
PAIR1	118	147	53	41,7	V
PAIR1	118	152	23	50,4	V
PAIR1	125	125	36	0	S
PAIR1	125	147	28	31,4	V
PAIR1	125	152	26	40	V
PAIR1	147	152	51	8,9	S
PAIR1	152	152	27	0	S
ASY2	91	121	13	47,4	V

ASY2	91	126	4	54,7	V
ASY2	91	141	20	77,4	V
ASY2	91	180	7	136,2	V
ASY2	91	184	14	142,4	V
ASY2	108	121	19	20	S
ASY2	108	126	6	26,9	S
ASY2	108	180	5	109	V
ASY2	121	121	3	0	S
ASY2	121	126	32	8,7	S
ASY2	121	132	81	16,3	S
ASY2	121	141	16	30,4	V
ASY2	121	180	2	89,6	V
ASY2	126	132	53	10,1	S
ASY2	126	141	16	22,8	S
ASY2	132	141	19	14,4	S
ASY2	141	141	1	0	S
ASY2	141	184	13	66	V
ASY2	180	180	10	0	S
ASY2	180	184	27	6,3	S
ASY2	184	184	21	0	S
Mer2	1	13	76	17,3	S
Mer2	13	35	26	32,3	V
Mer2	13	38	13	36,8	V
Mer2	13	45	4	47,2	V
Mer2	13	49	13	53,1	V
Mer2	13	100	3	128,8	V
Mer2	13	119	8	154,9	V
Mer2	13	127	6	166,3	V
Mer2	13	180	8	245,6	V
Mer2	21	35	47	20,6	S
Mer2	21	38	18	25,3	S

Mer2	21	45	11	35,6	V
Mer2	21	49	16	41,5	V
Mer2	21	119	27	142,9	V
Mer2	35	35	5	0	S
Mer2	35	38	34	5,1	S
Mer2	35	45	46	15,2	S
Mer2	35	69	6	51	V
Mer2	35	73	26	57,4	V
Mer2	35	85	16	74,4	V
Mer2	35	93	11	85,8	V
Mer2	35	161	21	186,2	V
Mer2	38	38	26	0	S
Mer2	38	45	19	10,5	S
Mer2	38	49	19	16,5	S
Mer2	38	73	16	53,1	V
Mer2	38	85	18	70,1	V
Mer2	38	93	4	81,6	V
Mer2	38	100	8	92,6	V
Mer2	38	119	10	119	V
Mer2	38	161	9	182,2	V
Mer2	45	45	2	0	S
Mer2	45	49	16	6,1	S
Mer2	45	73	14	42,8	V
Mer2	45	85	14	59,8	V
Mer2	45	93	3	71,3	V
Mer2	45	100	2	82,2	V
Mer2	45	119	11	108,9	V
Mer2	49	49	4	0	S
Mer2	49	73	35	37	V
Mer2	49	85	13	53,8	V
Mer2	49	161	19	166	V

Mer2	49	180	6	194,2	V
Mer2	69	97	4	42,1	V
Mer2	69	127	7	85,8	V
Mer2	73	73	5	0	S
Mer2	73	93	14	29,9	V
Mer2	73	98	1	37	V
Mer2	73	100	34	40,7	V
Mer2	73	119	17	68	V
Mer2	73	127	15	79,8	V
Mer2	73	161	14	130,3	V
Mer2	73	180	10	158,6	V
Mer2	85	93	47	11,8	S
Mer2	85	98	13	19,3	S
Mer2	85	100	38	22,6	S
Mer2	85	119	22	50,2	V
Mer2	85	127	20	62,1	V
Mer2	85	161	12	112,7	V
Mer2	85	180	11	141,2	V
Mer2	93	93	29	0	S
Mer2	93	98	4	7,7	S
Mer2	93	100	39	11,1	S
Mer2	93	119	32	38,4	V
Mer2	93	161	11	101,1	V
Mer2	93	180	4	129,5	V
Mer2	97	161	8	95,3	V
Mer2	100	161	17	90,4	V
Mer2	119	119	6	0	S
Mer2	119	154	38	53,1	V
Mer2	119	161	16	63,6	V
Mer2	127	161	14	52,1	V
Mer2	133	161	20	42,7	V

Mer2	154	161	3	10,7	S
Mer2	161	161	25	0	S
Mer2	161	180	22	28,9	V

2. Distances measured of the XL for the cc of Mer2 in dilute conditions

Table 7: *Table showing the XL distances measured considering intramolecular XL on the coiled-coils domain of Mer2 and the number of times a XL is detected (Hitcount). V= violated and S=satisfied.*

Chain	Residue 1	Residue 2	Hitcount	Intramolécular Distance (Å)	Threshold (27Å) satisfied?
Mer2	13	35	3	32,3	V
Mer2	35	38	5	5,1	S
Mer2	35	45	10	15,2	S
Mer2	35	119	13	123	V
Mer2	35	161	3	186,2	V
Mer2	38	38	8	0	S
Mer2	38	45	9	10,5	S
Mer2	45	45	9	0	S
Mer2	45	119	7	108,9	V
Mer2	53	69	30	24,8	S
Mer2	73	93	3	29,9	V
Mer2	85	93	16	11,8	S
Mer2	85	100	9	22,6	S
Mer2	85	119	11	50,2	V
Mer2	93	93	15	0	S
Mer2	93	97	7	6	S
Mer2	93	100	19	11,1	S
Mer2	93	119	11	38,4	V
Mer2	100	119	136	28,3	V
Mer2	119	119	41	0	S
Mer2	119	127	148	12,1	S
Mer2	119	133	216	21,2	S
Mer2	119	154	9	53,1	V
Mer2	119	161	66	63,6	V
Mer2	119	180	14	91,9	V
Mer2	127	127	138	0	S
Mer2	127	133	232	10,2	S

Mer2	127	161	13	52,1	V
Mer2	133	133	2	0	S
Mer2	161	180	25	28,9	V
Mer2	180	180	150	0	S

3. Protein sequences

3.1 Orthologs of RM

The minimal domain is highlighted in grey.

Rec24: NP_594817.1

MNGTNTEDNSKQILIQTMYTYDSSGETLKIAIAWKIILKKPKGKNIKDYIEALRKGIEDQEHCEKYASTL
LEPRPKTKKDVVLKNSNVTECVALKAKPFSKKIEDMDIFLLTNVHENLQEKRQTSGSLAHLDIEYTFNGL
FRFLKCTADIKLKQTKVYEGADFLRIKTLFEEIFMFLKRDCKSPLVLTRLVELGDYVLDLIIITQSIMQN
NANNGTGVISRAKFLEFYVFLEQLIFNKLSFASVEQLEKLLDQIVKRMKICFTYCKNDNPSIRLLYSECF
FSYAEIYFPCLHSFDAQLSSAASKCVQILRDIITNEELQTDKQELSKSAYSAPSILLIGLKDMLFPEDIS

Rec7: AAA35333.2

MNIFPIVKYSVAEDSYTTDSSHINWSHYSDGGFTLSLTSGLLQIRQHEELIQSINLLDLWHQLIPGTKEK
CLTLLSRAPCMNIRAFTNNVMKRFQVKFPSDVHYMKAKVEFEKLGLVFKDAKSSSEKKQFNNSQSQSNNS
QELSLMNNAYNKSSAQQPNLLLQPSYIPMTQTATTAVNNSTNYVNPAPLQHVMPNAEIFSNTPPLKRFRG
DAGMTQMPLRSDTSIESITASQQPTWDENVVITSSPFNPNRNAYSYGANSQYPIIAATPLNSQTQASWVA
QPENQAYANLIPSPPTTSQILPTELTEEEKQLRSKVLFYLKQDSFIQLCQSLERVWNKM

REC114: NP_082874.1

MSEAGNVASGLGLPGEVSQWSLKRYGRFMLLDNVGSPGPSSEAAAAGSPTWKVFESSEESGSLVLTIVVS
GHFFISQGQTLLEGFSLIGSKNWLKIVRRMDCLLFGTTIKNKSRMFRVQFSGESKEEALERCCGCVQTLA
QYVTVQEPDSTTQELQQSQGPREAGESQGKDPLQQGPSLTLEQHVCMAAGAGVLQERTSVTHRAQSILAP
EKLTLAYEGSSWGTEELGPFLRLCLMDQNFPAFVEEVEKELKKITGLRN

MEI4: NP_780422.1

MDIQPWYLKTSKLALALAIIHSKPADRSSREYTEYLASLVTQKESTWKSKLEALEAEVLQLRQKLLLSRI
SSGLFKNGPDVLPTLSDQEPTSSENTLTLMDDSGCVLSNEQRNEPAELSQHFVESTDPPLLPLPLEKRPR
TTLENPLSSHMQFFQHLLELKKWTESSSLKVYLTHFEKDSSTVSDSVSQLLDALITFYRNPKLPFSSFWT
EAVGTLARLASDFNLSNHIFKRCSKKLEEFEKTLLQAILENNSINRFQVQRYVSQSLVTLGSCSLLRKSI
ISLLLSEVNSFVDDLGAIDQDQGIYDVTRYENIFSLFWILEQVLQQAPQGDRTAHMDHSIPEMQTFLQKH
DEVIFRLSDAFPLFAFYLWRLGVLLNSAEMETVKNESLP

PHS1 (*A. thaliana*): OAP13975.1

MAGSLTASNRRRNAEDSSEIYRWTIGFARFVHYPSSPSPHPVLKPLGKREQYHSPHGTWLSASSSTVSLH
IVDELNRSDVILSVKLGQKVLIQKFALRFSTCDAALEFVEALKEKIKGLKEASTQNQKNKTRCDVSFQSD
YNPSDAIIPRATQKEPNMVRPLNSYVPEMLPRIVYEAQYQKSETRSEVSFQSDYNPSIEIFPRATEEEPN
MVRFFDSSVPEVLPRPEYEAGQALYPSQSTLNQIPSLPPSFTTLLSGCFPDSTLDAGQTTVKQNPDLKSQ
ILKYMEDSSFQDMLQKVERIIDEIGGNWIT

PRD2: OAO95537.1

MSSSVAEANHTEKEESLRLAIAVSLLRSKFHNHQSSSSTSRCYVSSESDALRWKQKAKERKKEIIRLQED
LKDAESSFHRDLFPANASCKCYFFDNLGVFSGRRIGEASESRFNDVLRRRFLRLARRRSRRKLTRSSQRL
QPSEPDYEEEAEHLRISIDFLLELSEADSNDSNFSNWSHQAVDFIFGNIIQSSRVANKLFKS

PHS1 (*Z. mays*): AAQ84720.1

MADAADSSMALVHSSLADSVLTSPRTLRQGQKWEVEYARYFGTPRRDPTAAPPSGLRYIMRGVHRHQGTW
IPASCPASLCVCHPSLPSAVPVLTISIGDVVFEEHFVSILNFSWPQVTCVTQCPIRGSRVVFVSFCDKFK
QIQKFAVRFPQPCDAESFLSCVECSCGSSGTMDIIPFGSDYVCEDSSASEYIVSNGLHHRLDDASNLEEQ
CFDHTIDEPPMNYHEETDQHVLEPLSASNTSNNSAFPPSFNQMLKSCSIDYDQEEPCPLAASNHVLQEVY
VLDTSHDVANEERTAGKGMDAAEGVDASILTYDLMARIKTYMADESFNDMLLKLDKAIDELGGDMSL

MPS1: PWZ29589.1

MALPKPRPPTPTASAATGTSSSRIDSPSLKAALAMALIHYNRLPGKANATAGTSPPSLLHWKRKAKDRKR
EILRLREELKVLQDGVRGEEMEPPVASCRCHFFDGCRDLRPQQGGGGGEHWVDEVLRRRFLRLVRWKEKR
RRVDRSLPSSSLIDFNSEDEMQQLSMSTDFLVELSDGIFAKSEAGHSFATFSHQAVDFILATLKNILSSE
REKDLVGEIIDSLVTRLMKRMCTVPEKLVTSDSGSTGCSDAQFSVQHLFRKLGNDEFIGQRVILVVSQKI
SNVSERLFLADPFADAFPDMHDNIFIMIQLLEFLISDYMKVWLCCEHINKRLFEECTRSILKARNDLQIL
ENMNGLYVVYIERVVGRLARDVAPAAHQGKLDLEVFSKLLC

3.2 Orthologs of Mer2

The coiled-coil domain is highlighted in grey

IHO1: NP_001366587.1

MNFNVWNVKEMLSIPSGSGITKPSNWNNNQTDCSLSDSQFLFGSQFCPENSETLLPSLDAGACLRHPKQT
QQNSVDSEPSIFIKYQAKPQLLGGDTKDESLFSLPLPVGKSKGLSKQFEEKKRRATDQSDSETLHSFVSH
FPEVINKLQTSVEKTEENLSSRSQSILDSVETIAKTFQETARVQHDLMVESVRDKGSMEQAILEIQRTCA
ARQAEFMEMKSTLKNLEVLVVEQTKNLQQFCDNLSQLIVPGILEELKKFTSVPQVAGHLKDSTSQTSPSL
TQSLHFTRQEKHPSEEPATWQAQEAPAGNPSTSSQRPGECGVWDEGAESGVFQKAALPTDGLHRGDGHVK
NKTVPTYCKNWVMTTRSVSNHFSNLPSQRAGNGQGLMAQGASQRDVSKFEARVKNACPEYGPQSMCSFDS
LEQSATEQKGRPCRKRRRGKKQQPQRSKRGGLLDRKQGQTSKAACAFIARHHCPQSPVCDPQGPLICWLT
PRSSTKSTCHILGGTGETSQTARAAQGNLVQHSQRSSTDSSSQGDQQINWFSDLSLENLEPPQCKKGGTN
LLCDPDFDSSDDNF

PRD3: OAP15284.1

MKMNINKACDLKSISVFPPNLRRSAEPQASQQLRSQQSQQSFSQGPSSSQRGCGGFSQMTQSSIDELLIN
DQRFSSQERDLSLKKVSSCLPPINHKREDSQLVASRSSSGLSRRWSSASIGESKSQISEELEQRFGMMET
SLSRFGMMLDSIQSDIMQANRGTKEVFLETERIQQKLTLQDTSLQQLRKEQADSKASLDGGVKFILEEFS
KDPNQEKLQKILQMLTTIPEQVETALQKIQREICHTFTREIQVLASLRTPEPRVRVPTAPQVKAKENLPE
QRGQAAKVLTSLKMPEPRVQVPAAPQAKENFPEQRGPVAKSNSFCNTTLKTKQPQFPRNPNDASARAVKP
YLSPKIQVGCWKTVKPEKSNFKKRATRKPVKSESTRTQFEQCSVVIDSDEEDIDGGFSCLINENTRGTNF
EWDAEKETERILRTARRTKRKFGNPIIIN

ASY2: XP_003348288.1

MASASRLTSPFFNDHSWSQETASTNGSSQSSLFSQPSQHSTQPSRQSSLASSPLEPDLEADLDQEQLDEI
DEHSPEVWASYLSAQYQTHIVETLKQTVFEKLDQLKASITSLQQNYVRQEEASTASSSELRATHDQIFTF
LRDLKPIPGRISALESDVRQNKEQAGRALSDLDDKLSSLKTDLVVKHDGAVRGLKSVRDQNQSLLETLVA
LRGEVMNLQLQQTQTDGRLSTFQKELEKTLAPRPELSNEAMDFLRQLLSRRADLMGLMEPPTTATSGTNV
PLIEAAPQLDLPWRSPQDTVDSTNSNLPVQAAPKTRAAHKRKALPSALDLTGPEPQAEKRANPKRTKFDL

PPVSSSRQLHESRKQLPVAKAKKALSAPETRPSATTKQTGTTKQKAPVAGMGTAPAGLSQQPPVRPKVQD
MKLLFKKFNDEYKAKRPKFDYLFIWKFIGQIESSEKLIHLQVRLLELLPDHVSQRRASHRLGDKFIQISK
TVTWAQFATALAEV

Rec15: NP_595887.1

MSYSVSAAQLWSRKLAMQAEDMQQHQKSQSNQIASCLAEMNTKQEVVNQTIGQLGRSISEVQQQNSQLVL
QSLNQINMSMQQVALGIQDYASRINKLEQTMSDMNLKFEALQKEQNSNTKTLADCTSQMTIITKKLDAEL
KKRYMTTKQTRTVQNQTMPRSNTTTKKRVLAIDFLADDDY

PAIR1: XP_008649842.1

MTGAGMKLKINKACDLSSISVLPPRRTGGSGGAGSSAAAAAVAAGSQQRSQPLSQQSFSQGAVGSGGASS
LLHSQSQLSQASLDENLLTLHLASPARDQGFRLHDDSSKKMTSLPVTSASCVREESQLQLAKTSSNPVHR
WNPSLPDSRCVVPAEDVERKFQHMASSVHKVGMVLDSVQNDTMQLNRAMKEAALDSGSIQQKIVVLDTSM
QKNLKGQGDLKVLVESNTKNIADQLAVLNSHSNKLVEISSALSVLPKQIERDLKQQQSDIFRIFRNNMEE
IARAVKSLNSKIDAIQMPTEQRCTTNGRPLMNQLPVDRNERPQVNQTPEVTRVSQTTVASLVNQAPAANG
RHLMSKTPAANGKSVMRQTPAENGRHLVSQTPAANGKTLKRQTSVTNGISLMSQVPAANGKPVVIQIPAP
KGRPMMRQKTAESGRTEMNEIPVASGRPHTNKIPIAEVHPAPLVCPAKQVATDPKPKVDEEKLKALPQKL
TSSRSRVAPKQEEAANTKVISRAAATEKVVIFIDDSDDDNDVRASCVILRPSGSGSLGERECDLMKVGAE
ESQEIPRRARKRRRREMQATVGVACMPLPA

4. Gel extraction

In order to extract the DNA of interest from an Agarose gel, the gel is first cut and only the strip containing the DNA remains. This strip is placed in a 1.5 ml Eppendorf tube and the gel is placed. The QIAquick® Gel Extraction Kit from QIAgen was used. 3 volumes Buffer QG to 1 volume gel (100 mg gel = 100 µl) are added. The tube is incubated at 50°C for 10 min and the tube is shaken every 2 to 3 min to help the gel dissolve. 1 gel volume of 100% isopropanol is added and the solution is mixed. The solution is transferred to a QIAquick spin column in a provided 2 ml collection tube. After centrifuging for 1 min at 13000 rpm, the flow-through is discarded. The column is washed with 750 µl PE buffer and centrifuged for 1 min at 13000 rpm. The centrifugation step is repeated to remove the residual solvent. The column is placed in a 1.5 ml Eppendorf tube. 27 µl of TE is added and allowed to rest for a few minutes before centrifuging a final time at 13000 rpm for 1 minute. The flow-through is retained.

5. Miniprep

In order to get rid of the genomic DNA of a cell, and to keep only its plasmid DNA, it is necessary to carry out what is more commonly called a "Miniprep". To do this, use the QIAprep® Spin Miniprep Kit (250) from QIAGEN and carry out the following protocol: the bacterial pre-culture is centrifuged at 4200 rpm for 3 min at RT. The bacterial pellet

is resuspended in 250 µl of Buffer P1 and transferred to a 1.5 ml Eppendorf tube. 250 µl of buffer P2 is added. The lysis reaction is not allowed to proceed for more than 5 min and the tube is mixed by inverting it 4 to 5 times. 350 µl of buffer N3 is added and the tube is mixed by inverting 4-5 times. After a 10 min centrifugation at 13 000 rpm, the supernatant is recollected and placed in a QIAprep 2.0 spin column. After a 1 min centrifugation at 13000 rpm, the flow-through is discarded and the column is washed with 750 µl of PE buffer. A further 1 min centrifugation is performed at 13000 rpm and the flow-through is discarded. The centrifugation step is repeated to remove the residual buffer. The column is placed in a 1.5 ml Eppendorf tube. 30 µl of TE is added and allowed to rest for a few minutes before centrifuging a final time at 13000 rpm for 1 min. The flow-through is retained.

IX Bibliography

Acquaviva, Laurent, Lóránt Székvölgyi, Bernhard Dichtl, Beatriz Solange Dichtl, Christophe de La Roche Saint André, Alain Nicolas, and Vincent Géli. 2013. « The COMPASS Subunit Spp1 Links Histone Methylation to Initiation of Meiotic Recombination ». *Science* 339 (6116): 215-18. https://doi.org/10.1126/science.1225739.

Anderson, J S, and R P Parker. 1998. « The 3' to 5' degradation of yeast mRNAs is a general mechanism for mRNA turnover that requires the SKI2 DEVH box protein and 3' to 5' exonucleases of the exosome complex. » *The EMBO Journal* 17 (5): 1497-1506. https://doi.org/10.1093/emboj/17.5.1497.

Andrade, M. A., C. Petosa, S. I. O'Donoghue, C. W. Müller, and P. Bork. 2001. « Comparison of ARM and HEAT Protein Repeats ». *Journal of Molecular Biology* 309 (1): 1-18. https://doi.org/10.1006/jmbi.2001.4624.

Arora, Charanjit, Kehkooi Kee, Shohreh Maleki, and Scott Keeney. 2004. « Antiviral Protein Ski8 Is a Direct Partner of Spo11 in Meiotic DNA Break Formation, Independent of Its Cytoplasmic Role in RNA Metabolism ». *Molecular Cell* 13 (4): 549-59. https://doi.org/10.1016/s1097-2765(04)00063-2.

Bee, Leonardo, Sonia Fabris, Roberto Cherubini, Maddalena Mognato, and Lucia Celotti. 2013. « The Efficiency of Homologous Recombination and Non-Homologous End Joining Systems in Repairing Double-Strand Breaks during Cell Cycle Progression ». *PLOS ONE* 8 (7): e69061. https://doi.org/10.1371/journal.pone.0069061.

Bergerat, Agnès, Bernard de Massy, Danielle Gadelle, Paul-Christophe Varoutas, Alain Nicolas, and Patrick Forterre. 1997. « An Atypical Topoisomerase II from Archaea with Implications for Meiotic Recombination ». *Nature* 386 (6623): 414-17. https://doi.org/10.1038/386414a0.

Boekhout, Michiel, Mehmet E. Karasu, Juncheng Wang, Laurent Acquaviva, Florencia Pratto, Kevin Brick, Diana Y. Eng, et al. 2019. « REC114 Partner ANKRD31 Controls Number, Timing, and Location of Meiotic DNA Breaks ». *Molecular Cell* 74 (5): 1053-1068.e8. https://doi.org/10.1016/j.molcel.2019.03.023.

Boeynaems, Steven, Simon Alberti, Nicolas L. Fawzi, Tanja Mittag, Magdalini Polymenidou, Frederic Rousseau, Joost Schymkowitz, et al. 2018. « Protein Phase Separation: A New Phase in Cell Biology ». *Trends in Cell Biology* 28 (6): 420-35. https://doi.org/10.1016/j.tcb.2018.02.004.

Bonfils, Sandrine, Ana E. Rozalén, Gerald R. Smith, Sergio Moreno, and Cristina Martín-Castellanos. 2011. « Functional interactions of Rec24, the fission yeast ortholog of mouse Mei4, with the meiotic recombination–initiation complex ». *Journal of Cell Science* 124 (8): 1328-38. https://doi.org/10.1242/jcs.079194.

Borde, Valérie, Waka Lin, Eugene Novikov, John H. Petrini, Michael Lichten, and Alain Nicolas. 2004. « Association of Mre11p with Double-Strand Break Sites during Yeast Meiosis ». *Molecular Cell* 13 (3): 389-401. https://doi.org/10.1016/S1097-2765(04)00034-6.

Borde, Valérie, Nicolas Robine, Waka Lin, Sandrine Bonfils, Vincent Géli, and Alain Nicolas. 2009. « Histone H3 Lysine 4 Trimethylation Marks Meiotic Recombination Initiation Sites ». *The EMBO Journal* 28 (2): 99-111. https://doi.org/10.1038/emboj.2008.257.

Calabrese, Antonio, and Sheena Radford. 2018. « Mass spectrometry-enabled structural biology of membrane proteins ». *Methods* 147 (février). https://doi.org/10.1016/j.ymeth.2018.02.020.

Claeys Bouuaert, Corentin, Stephen Pu, Juncheng Wang, Cédric Oger, Dima Daccache, Wei Xie, Dinshaw J. Patel, and Scott Keeney. 2021. « DNA-Driven Condensation Assembles the Meiotic DNA Break Machinery ». *Nature* 592 (7852): 144-49. https://doi.org/10.1038/s41586-021-03374-w.

Claeys Bouuaert, Corentin, Sam E. Tischfield, Stephen Pu, Eleni P. Mimitou, Ernesto Arias-Palomo, James M. Berger, and Scott Keeney. 2021. « Structural and Functional Characterization of the Spo11 Core Complex ». *Nature Structural & Molecular Biology* 28 (1): 92-102. https://doi.org/10.1038/s41594-020-00534-w.

Daccache, Dima, Pascaline Liloku, Emma De Jonge, Alexander N. Volkov, and Corentin Claeys Bouuaert. 2022. « Evolutionary Conservation of the Structure and Function of Meiotic Rec114–Mei4 and Mer2 Complexes ». bioRxiv. https://doi.org/10.1101/2022.12.16.520760.

Doll, Eveline, Monika Molnar, Yasushi Hiraoka, and Jürg Kohli. 2005. « Characterization of Rec15, an Early Meiotic Recombination Gene in Schizosaccharomyces Pombe ». *Current Genetics* 48 (5): 323-33. https://doi.org/10.1007/s00294-005-0024-3.

Engebrecht, J. A., K. Voelkel-Meiman, and G. S. Roeder. 1991. « Meiosis-Specific RNA Splicing in Yeast ». *Cell* 66 (6): 1257-68. https://doi.org/10.1016/0092-8674(91)90047-3.

Fowler, Kyle R., Randy W. Hyppa, Gareth A. Cromie, et Gerald R. Smith. 2018. « Physical Basis for Long-Distance Communication along Meiotic Chromosomes ». *Proceedings of the National Academy of Sciences of the United States of America* 115 (40): E9333-42. https://doi.org/10.1073/pnas.1801920115.

Gnügge, Robert, and Lorraine S. Symington. 2017. « Keeping It Real: MRX-Sae2 Clipping of Natural Substrates ». *Genes & Development* 31 (23-24): 2311-12. https://doi.org/10.1101/gad.310771.117.

Gobbini, Elisa, Corinne Cassani, Matteo Villa, Diego Bonetti, et Maria Pia Longhese. 2016. « Functions and Regulation of the MRX Complex at DNA Double-Strand Breaks ». *Microbial Cell* 3 (8): 329-37. https://doi.org/10.15698/mic2016.08.517.

Guéroult, Marc, Olivier Boittin, Oliver Mauffret, Catherine Etchebest, and Brigitte Hartmann. 2012. « Mg2+ in the Major Groove Modulates B-DNA Structure and Dynamics ». *PLOS ONE* 7 (7): e41704. https://doi.org/10.1371/journal.pone.0041704.

Henderson, Kiersten A., Kehkooi Kee, Shohreh Maleki, Paul A. Santini, and Scott Keeney. 2006. « Cyclin-Dependent Kinase Directly Regulates Initiation of Meiotic Recombination ». *Cell* 125 (7): 1321-32. https://doi.org/10.1016/j.cell.2006.04.039.

Hunter, Neil. 2015. « Meiotic Recombination: The Essence of Heredity ». *Cold Spring Harbor Perspectives in Biology* 7 (12): a016618. https://doi.org/10.1101/cshperspect.a016618.

Keeney, S., C. N. Giroux, and N. Kleckner. 1997. « Meiosis-Specific DNA Double-Strand Breaks Are Catalyzed by Spo11, a Member of a Widely Conserved Protein Family ». *Cell* 88 (3): 375-84. https://doi.org/10.1016/s0092-8674(00)81876-0.

Kleckner, N. 1996. « Meiosis: how could it work? » *Proceedings of the National Academy of Sciences* 93 (16): 8167-74. https://doi.org/10.1073/pnas.93.16.8167.

Kleckner, Nancy. 2006. « Chiasma Formation: Chromatin/Axis Interplay and the Role(s) of the Synaptonemal Complex ». *Chromosoma* 115 (3): 175-94. https://doi.org/10.1007/s00412-006-0055-7.

Kumar, Rajeev, Henri-Marc Bourbon, and Bernard de Massy. 2010. « Functional Conservation of Mei4 for Meiotic DNA Double-Strand Break Formation from Yeasts to Mice ». *Genes & Development* 24 (12): 1266-80. https://doi.org/10.1101/gad.571710.

Kumar, Rajeev, Cecilia Oliver, Christine Brun, Ariadna B. Juarez-Martinez, Yara Tarabay, Jan Kadlec, and Bernard de Massy. 2018. « Mouse REC114 Is Essential for Meiotic DNA Double-Strand Break Formation and Forms a Complex with MEI4 ». *Life Science Alliance* 1 (6). https://doi.org/10.26508/lsa.201800259.

Lam, Isabel, and Scott Keeney. 2014. « Mechanism and Regulation of Meiotic Recombination Initiation ». *Cold Spring Harbor Perspectives in Biology* 7 (1): a016634. https://doi.org/10.1101/cshperspect.a016634.

Laroussi, Hamida, Ariadna B. Juarez-Martinez, Aline Le Roy, Elisabetta Boeri Erba, Bernard de Massy, and Jan Kadlec. 2023. « Characterization of the REC114-MEI4-IHO1 Complex Regulating Meiotic DNA Double-Strand Break Formation ». bioRxiv. https://doi.org/10.1101/2023.01.11.523614.

Li, Jing, Gillian W. Hooker, and G. Shirleen Roeder. 2006. « Saccharomyces Cerevisiae Mer2, Mei4 and Rec114 Form a Complex Required for Meiotic Double-Strand Break Formation ». *Genetics* 173 (4): 1969-81. https://doi.org/10.1534/genetics.106.058768.

Li, Yang, Yu-Fan Wu, Han-Wei Jiang, Ranjha Khan, Qi-Qi Han, Furhan Iqbal, Xiao-Hua Jiang, and Qing-Hua Shi. 2021. « The molecular control of meiotic double-strand break (DSB) formation and its significance in human infertility ». *Asian Journal of Andrology* 23 (6): 555-61. https://doi.org/10.4103/aja.aja_5_21.

Liu, Kaixian, Emily M. Grasso, Stephen Pu, Shixin Liu, David Eliezer, and Scott Keeney. 2023. « Structure and DNA Bridging Activity of the Essential Rec114–Mei4 Trimer Interface ». bioRxiv. https://doi.org/10.1101/2023.01.18.524603.

Maleki, Shohreh, Matthew J. Neale, Charanjit Arora, Kiersten A. Henderson, and Scott Keeney. 2007. « Interactions between Mei4, Rec114, and Other Proteins Required for Meiotic DNA Double-Strand Break Formation in Saccharomyces Cerevisiae ». *Chromosoma* 116 (5): 471-86. https://doi.org/10.1007/s00412-007-0111-y.

Martini, Emmanuelle, Robert L. Diaz, Neil Hunter, and Scott Keeney. 2006. « Crossover Homeostasis in Yeast Meiosis ». *Cell* 126 (2): 285-95. https://doi.org/10.1016/j.cell.2006.05.044.

de Massy, Bernard 2013. « Initiation of Meiotic Recombination: How and Where? Conservation and Specificities among Eukaryotes ». *Annual Review of Genetics* 47: 563-99. https://doi.org/10.1146/annurev-genet-110711-155423.

« MBP Affinity Tag | NEB ». s. d. Consulted on 9 May 2023. https://international.neb.com/applications/protein-purification/affinity-purification-and-expression-tags/mbp-affinity-tag.

Merkley, Eric D, Steven Rysavy, Abdullah Kahraman, Ryan P Hafen, Valerie Daggett, and Joshua N Adkins. 2014. « Distance restraints from crosslinking mass spectrometry: Mining a molecular dynamics simulation database to evaluate lysine–lysine distances ». *Protein Science : A Publication of the Protein Society* 23 (6): 747-59. https://doi.org/10.1002/pro.2458.

Nore, Alexandre, Ariadna B. Juarez-Martinez, Julie Clément, Christine Brun, Boubou Diagouraga, Hamida Laroussi, Corinne Grey, et al. 2022. « TOPOVIBL-REC114 Interaction Regulates Meiotic DNA Double-Strand Breaks ». *Nature Communications* 13 (1): 7048. https://doi.org/10.1038/s41467-022-34799-0.

O'Connor, Clare. 2008. « Stages of Meiosis and Sexual Reproduction | Learn Science at Scitable ». 2008. http://www.nature.com/scitable/topicpage/meiosis-genetic-recombination-and-sexual-reproduction-210.

Page, Scott L., and R. Scott Hawley. 2003. « Chromosome Choreography: The Meiotic Ballet ». *Science* 301 (5634): 785-89. https://doi.org/10.1126/science.1086605.

Panizza, Silvia, Marco A. Mendoza, Marc Berlinger, Lingzhi Huang, Alain Nicolas, Katsuhiko Shirahige, and Franz Klein. 2011. « Spo11-Accessory Proteins Link Double-Strand Break Sites to the Chromosome Axis in Early Meiotic Recombination ». *Cell* 146 (3): 372-83. https://doi.org/10.1016/j.cell.2011.07.003.

Pawlowski, Wojciech P., and W. Zacheus Cande. 2005. « Coordinating the Events of the Meiotic Prophase ». *Trends in Cell Biology* 15 (12): 674-81. https://doi.org/10.1016/j.tcb.2005.10.005.

Petronczki, Mark, Maria F. Siomos, and Kim Nasmyth. 2003. « Un Ménage à Quatre: The Molecular Biology of Chromosome Segregation in Meiosis ». *Cell* 112 (4): 423-40. https://doi.org/10.1016/s0092-8674(03)00083-7.

Scheffzek, Klaus, and Stefan Welti. 2012. « Pleckstrin Homology (PH) like Domains - Versatile Modules in Protein-Protein Interaction Platforms ». *FEBS Letters* 586 (17): 2662-73. https://doi.org/10.1016/j.febslet.2012.06.006.

So, Chun, Shiya Cheng, and Melina Schuh. 2021. « Phase Separation during Germline Development ». *Trends in Cell Biology* 31 (4): 254-68. https://doi.org/10.1016/j.tcb.2020.12.004.

Stanzione, Marcello, Marek Baumann, Frantzeskos Papanikos, Ihsan Dereli, Julian Lange, Angelique Ramlal, Daniel Tränkner, et al. 2016. « Meiotic DNA Break Formation Requires the Unsynapsed Chromosome Axis-Binding Protein IHO1 (CCDC36) in Mice ». *Nature Cell Biology* 18 (11): 1208-20. https://doi.org/10.1038/ncb3417.

Steiner, Silvia, Jürg Kohli, and Katja Ludin. 2010. « Functional Interactions among Members of the Meiotic Initiation Complex in Fission Yeast ». *Current Genetics* 56 (3): 237-49. https://doi.org/10.1007/s00294-010-0296-0.

« SUMO Protease Small Ubiquitin-Like Modifier Protease Sigma ». s. d. Consulted on 9 May 2023. http://www.sigmaaldrich.com/.

« SUMO Tag and SUMO Tag Purification | Sino Biological ». s. d. Consulted on 9 May 2023. https://www.sinobiological.com/resource/protein-review/sumo-tag-purification.

Tessé, Sophie, Henri-Marc Bourbon, Robert Debuchy, Karine Budin, Emeline Dubois, Zhang Liangran, Romain Antoine, et al. 2017. « Asy2/Mer2: an evolutionarily conserved mediator of meiotic recombination, pairing, and global chromosome compaction ». *Genes & Development* 31 (18): 1880-93. https://doi.org/10.1101/gad.304543.117.

Thangavel, Gokilavani, Paulo G. Hofstatter, Raphaël Mercier, and André Marques. 2023. « Tracing the Evolution of the Plant Meiotic Molecular Machinery ». *Plant Reproduction* 36 (1): 73-95. https://doi.org/10.1007/s00497-022-00456-1.

Vrielynck, Nathalie, Katja Schneider, Marion Rodriguez, Jason Sims, Aurélie Chambon, Aurélie Hurel, Arnaud De Muyt, et al. 2021. « Conservation and Divergence of Meiotic DNA Double Strand Break Forming Mechanisms in Arabidopsis Thaliana ». *Nucleic Acids Research* 49 (17): 9821-35. https://doi.org/10.1093/nar/gkab715.

Wilkins, Adam S, and Robin Holliday. 2009. « The Evolution of Meiosis From Mitosis ». *Genetics* 181 (1): 3-12. https://doi.org/10.1534/genetics.108.099762.

Xu, Jiaqi, Tao Li, Soonjoung Kim, Michiel Boekhout, and Scott Keeney. 2023. « Essential Roles of the ANKRD31-REC114 Interaction in Meiotic Recombination and Mouse Spermatogenesis ». bioRxiv. https://doi.org/10.1101/2023.04.27.538541.

Yadav, Vikash Kumar, and Corentin Claeys Bouuaert. 2021. « Mechanism and Control of Meiotic DNA Double-Strand Break Formation in S. cerevisiae ». *Frontiers in Cell and Developmental Biology* 9. https://www.frontiersin.org/articles/10.3389/fcell.2021.642737.

Zhu, Xuan, and Scott Keeney. 2015. « High-Resolution Global Analysis of the Influences of Bas1 and Ino4 Transcription Factors on Meiotic DNA Break Distributions in Saccharomyces cerevisiae ». *Genetics* 201 (2): 525-42. https://doi.org/10.1534/genetics.115.178293.

Zickler, D., and N. Kleckner. 1998. « The Leptotene-Zygotene Transition of Meiosis ». *Annual Review of Genetics* 32 (1): 619-97. https://doi.org/10.1146/annurev.genet.32.1.619.

Zickler, D., and N. Kleckner. 1999. « Meiotic Chromosomes: Integrating Structure and Function ». *Annual Review of Genetics* 33: 603-754. https://doi.org/10.1146/annurev.genet.33.1.603.

Zickler, Denise, and Nancy Kleckner. 2015. « Recombination, Pairing, and Synapsis of Homologs during Meiosis ». *Cold Spring Harbor Perspectives in Biology* 7 (6): a016626. https://doi.org/10.1101/cshperspect.a016626.

Study of the conservation of the structure and the function of *S. cerevisiae* Rec114-Mei4-Mer2 proteins during evolution

Emma De Jonge

In prophase I of meiosis, the phenomenon of genetic recombination takes place, which can lead to the formation of crossovers between two homologous chromosomes. This phenomenon is essential for the correct segregation of homologous chromosomes at the end of anaphase I of meiosis.

Genetic recombination is initiated by a DNA double-strand break mediated, in *S. cerevisiae*, by the Spo11 protein. However, this phenomenon also requires the presence of 9 other proteins divided into three complexes: The core complex (which contains the Spo11 catalytic unit), the MRX complex and the RMM complex (composed of Rec114, Mei4 and Mer2). It has been shown *in vivo* and *in vitro* that the proteins of the RMM complex form condensates in the presence of DNA. This phenomenon is referred to as phase separation. It is hypothesized that this complex acts as a platform that recruits the 7 other essential proteins to the site where DNA double-strand breaks occur.

The aim of this master thesis is to study the conservation of the structure and DNA-binding function of the *S. cerevisiae* RMM complex orthologs throughout evolution. As these proteins cannot be purified together, the study is divided into two distinct parts: A part on the Rec114-Mei4 orthologs where the complexes from *M. musculus* (REC114-MEI4), *S. pombe* (Rec7-Rec24), *A. thaliana* (PHS1-PRD2) and *Z. mays* (PHS1-MPS1) will be studied and another on Mer2 orthologs where the protein from *M. musculus* (IHO1), *S. pombe* (Rec15), *A. thaliana* (PRD3), *Z. mays* (PAIR1) and *S. macrospora* (ASY2) will be studied.

Despite the large difference in amino acid sequences between these orthologous proteins, the structure predicted by AlphaFold is quite similar. These models of the complex formed by orthologs of Rec114 and Mei4, with a 2:1 stoichiometry respectively, will be verified by mutagenesis followed by a Pulldown in order to identify the amino acids important in the protein interactions between the C-terminal part of Rec114 and the N-terminal part of Mei4. Concerning the Mer2 orthologs, only the structure of the parts of the proteins forming a coiled-coil domain will be studied by SEC-MALS, SAXS and XL-MS, to see whether the homotetrameric and parallel structure of Mer2 is conserved.

In order to study the DNA-binding function of these proteins, EMSAs were carried out on RM complexes and coiled-coil domains of the proteins studied. In *S. cerevisiae*, these proteins bind DNA *in vitro* and the conservation of this activity is studied in this work.

UNIVERSITÉ CATHOLIQUE DE LOUVAIN
Faculté des bioingénieurs
Croix du Sud, 2 bte L7.05.01, 1348 Louvain-la-Neuve, Belgique | **www.uclouvain.be/agro**