



**HAL**  
open science

# Caractérisation biochimique des protéines REC114, MEI4 et IHO1 impliquées dans l'initiation de la recombinaison méiotique chez la souris

Hamida Laroussi

► **To cite this version:**

Hamida Laroussi. Caractérisation biochimique des protéines REC114, MEI4 et IHO1 impliquées dans l'initiation de la recombinaison méiotique chez la souris. Biochimie [q-bio.BM]. Université Grenoble Alpes [2020-..], 2022. Français. NNT : 2022GRALV096 . tel-04086873

**HAL Id: tel-04086873**

**<https://theses.hal.science/tel-04086873v1>**

Submitted on 2 May 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

## THÈSE

Pour obtenir le grade de

### DOCTEUR DE L'UNIVERSITÉ GRENOBLE ALPES

École doctorale : CSV- Chimie et Sciences du Vivant

Spécialité : Chimie Biologie

Unité de recherche : Institut de Biologie Structurale

**Caractérisation biochimique des protéines REC114, MEI4 et IH01  
impliquées dans l'initiation de la recombinaison méiotique chez la souris**

**Biochemical characterization of the REC114, MEI4 and IH01 proteins  
involved in the initiation of meiotic recombination in mice**

**Présentée par :**

**Hamida LAROUSI**

Thèse soutenue publiquement le 7 décembre 2022, devant le jury composé de :

**Jan KADLEC**

CHARGE DE RECHERCHE, INSERM

Directeur de thèse

**Sébastien FRIBOURG**

DIRECTEUR DE RECHERCHE, INSERM

Rapporteur

**Valerie BORDE**

DIRECTEUR DE RECHERCHE, CNRS

Rapporteuse

**Christine MÉZARD**

DIRECTEUR DE RECHERCHE, CNRS

Examinatrice

**Marc JAMIN**

PROFESSEUR DES UNIVERSITÉS, UGA

Président





## **Remerciements**

*Je tiens d'abord à remercier Dr. Christine MÉZARD, Pr. Marc JAMIN, Dr. Sébastien FRIBOURG et Dr. Valerie BORDE de m'avoir fait cet honneur en acceptant de constituer mon jury de thèse et d'évaluer mon travail.*

*Je tiens à exprimer mes profonds remerciements à mon directeur de thèse, Jan Kadlec, pour m'avoir accueillie dans son équipe et encadrer ma thèse. Merci de m'avoir confié ce projet et de me permettre d'apprendre plusieurs techniques. Merci de m'avoir guidé tout au long de ces années. Merci pour tes précieux conseils et suggestions scientifiques, ton aide et ta disponibilité sans limite.*

*Merci à Bernard DE MASSY, Montse SOLER LOPEZ et Nolwenn MIGUET, membres du comité de suivi de ma thèse, qui ont été toujours disponibles pour discuter de mon projet. Vous avez grandement contribué à mon travail par vos recommandations, et je vous en suis vraiment reconnaissante.*

*Merci à l'Agence Nationale de Recherche pour le financement de ma thèse et à l'école doctorale de chimie et sciences de la vie de l'Université Grenoble Alpes.*

*Je veux également remercier chaleureusement les membres des différentes plateformes. Un grand merci à Caroline MAS de la plateforme ISBG de m'avoir appris à réaliser les expériences l'ITC et SEC-MALS, merci pour ta disponibilité et ta sympathie. Merci à Aline LE ROY et Christine EBEL de la plateforme IBS POAL pour avoir mené mes expérimentations AUC et SEC-MALLS. Merci à Élisabetta BOERI ERBA de la plateforme de spectrométrie de masse pour les différentes analyses de masse ainsi qu'à Adrien FAVIER de la plateforme RMN pour l'acquisition de spectres. I would like to thank Michael HONS for your hard work on the electron microscope and for the interest that you have shown in this project. Je tiens également à remercier Daphna FENEL de la plateforme de microscopie électronique de l'IBS et Anne Emmanuelle FOUCHER pour toutes les grilles qu'elles ont préparées et analysées. Many thanks to Cameron MACKERETH for acquiring NMR data, thank you for your availability, help and all the work you have done. Sans vous tous, votre aide et votre disponibilité, je n'aurais pas eu autant de résultats intéressants durant cette thèse.*

*Je retourne encore à Anne Emmanuelle à qui je veux dire que ma thèse n'aurait pas été la même sans toi. Ce fut un véritable plaisir de travailler avec toi. Merci d'être ici pour répondre à toutes mes questions, j'ai tant appris de toi et de ton expérience. Merci pour toutes tes qualités humaines et pour les discussions que nous avons eues tout au long de cette thèse sur des sujets professionnels et personnels.*

*Encore une fois, j'aimerais remercier Jan et Anne Emmanuelle, d'avoir pris le temps de relire ce manuscrit.*

*A special thanks to Ariadna, at the origin of this project, for sharing her knowledge with me and for her daily availability during my beginnings.*

*Un grand merci évidemment pour JB et Étienne, merci d'avoir été là pour moi dans les bons et les mauvais moments de cette thèse. Merci pour nos discussions*

*diverses et variées au quotidien au coin de la paillasse ou autour d'un café. Merci à toi Laura pour tous tes conseils et pour tous les moments que nous avons partagés. Je n'oublierai pas ta spontanéité, ton naturel et ton humour.*

*Je tiens à remercier tous les membres de l'équipe EBEV, Delphine, Nolwenn, Guidenne, Borys, Shaya, Haiyan, Kimi, Christine, Rémy et Cécile, avec qui nous avons partagé le labo pendant plus de deux ans. C'était un réel plaisir de vous côtoyer tous au quotidien. Je n'oublierai jamais tous les moments de convivialité que nous avons passés ensemble marqués par votre sympathie et votre bonne humeur.*

*Mes remerciements vont également à tous les membres du groupe EPIGEN, Carlo, Rose laure, Marjo, Isabale, Dimitrios, Yordan, qui nous ont accueillis dans leur labo avec leur grand cœur. Merci pour votre gentillesse et votre sympathie au quotidien. Merci également pour vos avis scientifiques très pertinents lors de nos groupes meeting. Thanks to Hasan and Ahmed for all the laughter we have shared together and for supporting me during my work on this manuscript.*

*Merci également à toutes les personnes de l'IBS qui ont participé de près ou de loin au bon déroulement de ma thèse.*

*Ce travail est dédié à la mémoire de mon défunt père, qui a toujours cru en moi et qui m'a toujours poussée et motivée durant mon parcours. Puisse Dieu, le Tout Puissant, l'avoir en sa sainte miséricorde ! Mes meilleurs remerciements vont à ma mère, ma source de motivation, qui m'a toujours entourée de son amour, soutenue et encouragée.*

*Je ne saurais oublier de dédier ce travail à toutes les personnes qui me sont chères, en particulier mes frères.*

*Je tiens à exprimer ma profonde reconnaissance à mon mari qui m'a toujours appuyée et m'a poussée à donner le meilleur de moi-même. Cette thèse et ma vie n'auraient pas été les mêmes sans toi. Merci pour tout.*

*Pour finir, je dédie cette thèse au soleil qui brille dans ma vie depuis bientôt sept ans, ma fille Minna. Merci pour le bonheur que tu me procures chaque jour.*

# Table des matières

Table des matières .....	1
Liste des figures et des tableaux .....	5
Résumé .....	9
Summary .....	11
<b>1 Introduction .....</b>	<b>13</b>
1.1 Définition et rôle de la méiose .....	15
1.2 Déroulement de la méiose .....	18
1.2.1 Prophase I, stade de la recombinaison méiotique .....	18
1.2.1.1 Leptotène, stade de la formation des axes chromosomes et l'initiation de la recombinaison .....	18
1.2.1.2 Zygotène, stade de la formation du complexe synaptonémal et de la progression de la recombinaison .....	19
1.2.1.3 Pachytène, stade de la formation des crossing-over .....	21
1.2.1.4 Diplotène et Diacinèse, stade de désynapsis .....	21
1.2.2 Métaphase, anaphase et télophase 1 .....	22
1.2.3 Méiose II .....	22
1.3 Mécanisme moléculaire de la recombinaison méiotique .....	22
1.3.1 Formation et maturation des cassures double brin de l'ADN .....	23
1.3.2 Invasion et recherche d'homologie .....	25
1.3.3 Réparation des DSB .....	26
1.3.3.1 Synthesis-Dependent Strand Annealing (SDSA) .....	26
1.3.3.2 Formation d'une double jonction de Holliday .....	26
1.4 Mécanismes moléculaires de la formation des DSB .....	30
1.4.1 Facteurs protéiques impliqués dans la formation des DSB chez les mammifères .....	30
1.4.1.1 Complexe catalytique SPO11-TOPOVIBL .....	32
1.4.1.2 REC114-MEI4-IHO1 .....	36
1.4.1.3 Autres protéines nécessaires à la formation de DSB .....	38
1.5 Localisation des DSB .....	42
1.6 Régulation de la formation des DSB .....	46
1.6.1 Régulation temporelle de la formation des DSB .....	46
1.6.2 Contrôle du nombre et de la distribution des DSB .....	47
Objectifs de la thèse .....	51
<b>2 Matériels et méthodes .....</b>	<b>53</b>
2.1 Biologie moléculaire .....	55
2.1.1 Clonage .....	55
2.1.2 Mutagenèse dirigée .....	57
2.2 Production et purification .....	59
2.2.1 Expression dans <i>E.coli</i> .....	59
2.2.2 Expression des protéines enrichies aux marquages <sup>15</sup> N et <sup>13</sup> C .....	59
2.2.3 Purification .....	60
2.2.3.1 Lyse des bactéries .....	60
2.2.3.2 Chromatographie d'affinité .....	60
2.2.3.3 Coupure à la protéase TEV .....	61

2.2.3.4	Dialyse et deuxième chromatographie d'affinité .....	61
2.2.3.5	Filtration sur gel .....	61
2.3	Caractérisation du complexe IHO1-REC114-MEI4 .....	62
2.3.1	Reconstitution du complexe IHO1-REC114-MEI4 .....	62
2.3.2	Titration calorimétrique isotherme (ITC) .....	62
2.3.3	Pontage chimique .....	63
2.3.4	Pontage chimique couplé à la spectrométrie de masse .....	64
2.4	Caractérisation du complexe REC114-MEI4, de IHO1 et de REC114 C-ter .....	64
2.4.1	<i>Pull down</i> .....	64
2.4.2	Diffusion de la lumière à plusieurs angles couplée à la chromatographie d'exclusion de taille (SEC-MALLS) .....	65
2.4.3	Ultracentrifugation analytique (AUC) .....	65
2.4.4	Spectrométrie de masse en condition native .....	66
2.5	Caractérisation structurale .....	68
2.5.1	Coloration négative .....	68
2.5.2	Cryo-microscopie électronique (Cryo-EM) .....	68
2.5.3	Résonance magnétique nucléaire (RMN) .....	69
2.5.4	Essais de cristallisation .....	70
2.5.5	AlphaFold 2 .....	71
<b>3</b>	<b>Résultats .....</b>	<b>73</b>
3.1.1	Complexe REC114-MEI4-IHO1 .....	75
3.1.2	Purification du sous-complexe REC114-MEI4 et de la protéine IHO1 .....	77
3.1.3	Reconstitution du complexe REC114-MEI4-IHO1 .....	79
3.1.4	Caractérisation de l'état oligomérique du complexe REC114-MEI4-IHO1 .....	81
3.1.5	Caractérisation de l'interaction entre IHO1 et REC114-MEI4 .....	83
3.1.5.1	Calorimétrie par titrage isotherme (ITC) .....	83
3.1.5.2	Pontage chimique couplé à la spectrométrie de masse .....	84
3.1.5.3	Coloration négative .....	87
3.1.6	Implication de la partie N-terminale d'IHO1 dans l'interaction avec REC114- MEI4 .....	89
3.1.7	Caractérisation de l'interaction entre la partie N-terminale d'IHO1 et REC114 .....	91
3.1.8	Modèle AlphaFold2 du complexe IHO1-REC114 .....	93
3.1.9	Validation du modèle d'interaction entre REC114 et IHO1 .....	95
3.1.9.1	REC114-MEI4-TOPOVIBL et IHO1 .....	95
3.1.9.2	REC114-MEI4-ANKRD31 et IHO1 .....	96
3.1.9.3	Identification des résidus importants pour l'interaction entre REC114-IHO1 .....	98
3.1.10	Essais de cristallisation du complexe IHO1-REC114 .....	101
3.2	Complexe REC114-MEI4 .....	105
3.2.1	Caractérisation par RMN du complexe REC114-MEI4 .....	105
3.2.2	Caractérisation de l'état oligomérique du complexe REC114-MEI4 .....	109
3.2.2.1	Spectrométrie de masse native (MS native) .....	109
3.2.2.2	Ultracentrifugation analytique (AUC) .....	111
3.2.3	Modèle structural par AlphaFold2 du complexe REC114-MEI4 .....	113
3.2.4	Interface d'interaction entre REC114 et MEI4 .....	117
3.3	La partie C-terminale de REC114 .....	123
3.3.1	Caractérisation de l'état oligomérique de la partie C-terminale de REC114 .....	123
3.3.2	Modèle structural par AlphaFold2 pour le domaine C-terminal de REC114 .....	124
3.3.3	Interface de dimérisation de la partie C-terminale de REC114 .....	125

3.4	Caractérisation d'IHO1 .....	129
3.4.1	Purification d'IHO1 <sub>125-281</sub> .....	129
3.4.2	Caractérisation d'IHO1 <sub>125-281</sub> .....	131
3.4.3	Modèle AlphaFold2 d'IHO1 .....	135
3.4.4	Purification d'IHO1 <sub>196-245</sub> .....	136
3.4.5	Caractérisation d'IHO1 <sub>196-245</sub> .....	136
<b>4</b>	<b>Discussions et perspectives .....</b>	<b>139</b>
4.1	Caractérisation du complexe REC114-MEI4-IHO1 .....	141
4.2	Caractérisation du sous-complexe REC114-MEI4 .....	145
4.3	Caractérisation d'IHO1 .....	147
<b>5</b>	<b>Conclusion.....</b>	<b>151</b>
	Références bibliographiques.....	157
	Contribution aux publications .....	172



## Liste des figures et des tableaux

Figure 1. La division méiotique.....	16
Figure 2. Stades de la prophase 1 et structures chromosomiques associées.....	20
Figure 4. Réparation des intermédiaires de recombinaison.....	24
Figure 5. Modèle de la formation des DSB chez les mammifères.....	27
Figure 5. Modèle de la formation des DSB chez les mammifères.....	31
Figure 6. Structure de TopoVI.....	32
Figure 7. Comparaison entre TopoVI et TOPOVIL.....	34
Figure 8. REC114-TOPOVIBL.....	37
Figure 9. REC114-ANKRD31.....	41
Figure 10. Domaine de PRDM9.....	42
Figure 11. Modèle du mode d'action de PRDM9.....	45
Figure 12. Modèle de la régulation de la formation des DSB.....	50
Figure 13. Mutagenèse dirigée.....	58
Figure 14. Représentation de la molécule de pontage BS3.....	63
Figure 15. Représentation schématique des protéines IHO1, REC114 et MEI4.....	76
Figure 16. Purification d'IHO <sub>1-281</sub> et de REC114-MEI4 <sub>1-127</sub> .....	78
Figure 17. Reconstruction du complexe REC114-MEI4-IHO1.....	80
Figure 18. Caractérisation par la technique de SEC-MALLS.....	82
Figure 19. Mesures d'affinité de l'interaction entre REC114-MEI4 <sub>1-127</sub> et MBP-IHO1 <sub>1-281</sub> par ITC.....	84
Figure 20. Reconstitution et pontage chimique du complexe His-SUMO-IHO1 <sub>1-281</sub> -REC114-MEI4 <sub>1-127</sub> .....	86
Figure 21. Visualisation par microscopie électronique en coloration négative du complexe IHO1-REC114-MEI4.....	88
Figure 22. Implication de la partie N-terminale d'IHO1 dans l'interaction avec REC114-MEI4.....	90
Figure 23. Caractérisation de l'interaction entre IHO1 et REC114.....	92
Figure 24. Modèle structural du complexe REC114 et IHO1.....	94
Figure 25. Analyse de l'interaction de IHO1 avec REC114-MEI4 en présence de TOPOVBL.....	96
Figure 26. Analyse de l'interaction de IHO1 avec REC114-MEI4 en présence d'ANKRD31.....	97
Figure 27. Effet de la mutation REC114 L104D sur l'interaction entre REC114 et IHO1.....	98
Figure 28. Effet de la mutation F3E de IHO1 sur l'interaction entre REC114 et IHO1.....	100
Figure 29. Purification et essais de cristallisation du domaine PH de REC114 en fusion avec IHO1 <sub>1-16</sub> .....	103
Figure 30. Expérience HSQC <sup>1</sup> H- <sup>15</sup> N réalisée sur le complexe REC114 <sub>226-254</sub> -MEI4 <sub>1-127</sub> .....	106
Figure 31. Analyse des différentes troncatures de MEI4.....	106
Figure 32. Purification du complexe REC114 <sub>227-259</sub> -MEI4 <sub>1-43</sub> et le spectre RMN associé..	108
Figure 33. Caractérisation de l'état oligomérique du complexe REC114-MEI4 par MS native.....	110
Figure 34. Profils de vitesse de sédimentation du complexe REC114-MEI4 (absorbance 280).....	112
Figure 35. Modèle structural de REC114 et MEI4.....	114
Figure 36. Modèle structural du complexe 2REC114-1MEI4.....	116
Figure 37. Analyse de l'interaction entre les mutants de REC114 et MEI4.....	118
Figure 38. Analyse de l'interaction entre les mutants de MEI4 et REC114.....	120
Figure 39. Caractérisation de l'état oligomérique de la partie C terminale de REC114 par	

SEC-MALLS.....	123
Figure 40. Caractérisation de l'état oligomérique de la partie C-terminale de REC114 par MS native.....	124
Figure 41. Modèle structural du dimère formé par le C-terminus de REC114 murin.....	125
Figure 42. Analyse par la technique SEC-MALLS des mutants de REC114 C-terminal.....	126
Figure 43. Analyse par la technique de SEC-MALLS de la partie C-terminale de REC114.	127
Figure 44. Modèle structural du dimère formé par le C-terminus de REC114 de <i>S. cerevisiae</i> .....	128
Figure 45. Purification de IHO1 <sub>125-281</sub> .....	130
Figure 46. Caractérisation de l'état oligomérique de IHO1 <sub>125-281</sub> .....	131
Figure 47 : Coloration négative de IHO1 <sub>125-281</sub> .....	132
Figure 48. Données de la cryo-microscopie électronique de la protéine IHO1 <sub>125-281</sub> .....	133
Figure 49. Essais de cristallisation d'IHO1 <sub>143-257</sub> .....	134
Figure 50. Modèle structural d'IHO1.....	135
Figure 51. Caractérisation de IHO1 <sub>196-259</sub> et essais de cristallisation.....	138
Figure 52. Modèle du complexe REC114-MEI4-IHO1.....	153
Figure 53. Modèles des complexes REC114-MEI4-TOPOVIBL et REC114-MEI4-ANKRD31.....	154
Figure 54. Partenaires de REC114.....	154
Tableau 1 : Récapitulatifs des constructions utilisées durant cette thèse.....	56
Tableau 2 : récapitulatif des résultats ITC sur les interaction entre REC114-IHO1.....	99



## Liste des abréviations

5Y-CAP	Catabolite activator protein	MLH1 /3	Mut L Homolog 1/3
ADN	Acide désoxyribonucléique	MND1	Meiotic nuclear division protein 1 homologue
ANKRD31	Ankyrin repeat domain-containing 31	MRE11	Meiotic recombination 11 homologue
ATM	Ataxia telangiectasia mutated	MRN	MRE11, RAD50 et NBS1
ATR	Ataxia telangiectasia and Rad3-related protein	MRX	MRE11-RAD50-XRS1 complex
AtSPO11	Arabidopsis thaliana SPO11	MS	Mass spectrometry
AUC	Analytical ultracentrifugation	MS native	spectrométrie de masse native
BME	$\beta$ -mercaptoéthanol	MSCI	méiotique sex chromosome inactivation
BRCA1	Breast cancer proteins BRCA1	MSH4	Mut S homolog 4
BS3	Bis(sulfosuccinimidyl)suberate	MSH5	Mut S homolog 5
CDK-S	S-phase cyclin-dependent kinase	MSUC	Meiotic silencing of unsynapsed chromatin
Cfp1	CXXC finger protein 1	Nano-ESI	Nano-electrospray
CO	Crossing-over	NBS1	Nijmegen breakage syndrome protein 1
COMPASS	Complex Proteins Associated with Set1	NCO	Non-crossover
Cryo-EM	Cryo-microscopie électronique	NDR	Nucleosome-depleted region
CTD	C-terminal domain	PAR	Pseudo-autosomal region
CtIP	CtBP-interacting protein	PCR	Polymerase chain reaction
D-loop	Displacement loop	PEG	Polyethylene glycol
Da	Dalton	PH	Pleckstrin Homology
DDK	Dbf4-dependent kinase	PIKK	Phosphatidylinositol 3-kinase-related kinases
dHJ	double jonction de Holliday	pLDDT	predicted local distance difference test
Dmc1	Disrupted meiotic cDNA 1	PRDM9	PR domain zinc finger protein 9
DSB	Double-Strand Breaks	Rad21L	RAD21 like homolog
EDTA	acide éthylènediaminetétraacétique	RAD50 /51	Radation sensitive protein 50/51
EWSR1	Ewing sarcoma breakpoint region 1	REC114	Recombination 114
EXO1	Exonuclease 1	REC8	Recombination-Deficient protein 8
FBX047	F-box only protein 47	RECA	Recombination protein A
GHKL	gyrase, geat-shock protein 90, histidine kinase et MuK1	RMM	Rec114-Mei4-Mer2
H2TH	helix-two-turn-helix	RMN	résonance magnétique nucléaire
H3K36me3	Histone 3 lysine 36 tri methylation	RPA	Replication protein A
H3K4me3	Histone 3 lysine 4 tri methylation	rpm	Revolutions Per Minute
HOP1 /2	Homolog pairing protein 1/2	SC	Synaptonemal complex
HORMA	Hop1, Rev7 and MAD2	SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
HPLC	High-performance liquid chromatography	SDSA	synthesis-dependent strand annealing
HSQC	Heteronuclear Single Quantum Correlation	SEC	Second-end capture
IHO1	interactor of HORMAD1	SEC-MALLS	Size-exclusion chromatography with multi-angle light scattering
IPTG	isopropyl d-thiogalactopyranoside	SEI	single-end invasion
ITC	Isothermal titration calorimetry	SHOC1	Shortage In Chiasmata 1
KASH	Klarsicht, ANC-1, Syne Homology	SPO11 /16	Sporulation protein 11 /16
KRAB	Krueppel-associated box	SSM	signature sequence motifs
LB	Luria-Bertani	SUN	Sad1p, UNC-84
MBP	Maltose-binding protein	SYCE1 / 2 / 3	Synaptonemal complex central element 1/2/3
MEI1/4	Meiotic double-stranded break formation protein 1/4	SYCP1 / 2/ 3	Synaptonemal complex protein 1
MER2	Meiotic recombination 2	TEV	Tobacco Etch Virus
		TEX12	Testis-expressed protein 12

TLAC	tethered loop-axis complex
TOF	Time of flight
TOPO	Topoisomerase
TOPOVIL	TOPOVI-Like
Toprim	topoisomerase-primase
TRIP13	thyroid receptor-interacting protein 13
TROSY	Transverse relaxation optimized spectroscopy
WDR61	WD Repeat Domain 61
XRCC2/3	X-ray repair cross complementing 2/3
ZCWPW1	Zinc Finger CW-Type And PWWP Domain Containing 1
ZMM	Zip1, 2, 3,4, Msh4, 5, Mer3, Spo16

## Résumé

La méiose est une division cellulaire spécialisée des cellules germinales qui permet la production des gamètes haploïdes en réduisant de moitié le nombre de chromosomes des cellules mères. Ce processus est constitué d'un cycle de réplication suivi de deux cycles de division : méiose I et méiose II. La réduction des chromosomes a lieu pendant la méiose I où les chromosomes homologues d'origine maternelle et paternelle sont séparés. Cette ségrégation réductionnelle nécessite la formation de plusieurs cassures double-brin de l'ADN (DSB) et leur réparation par la recombinaison homologue, appelée également recombinaison méiotique. Ce mécanisme permet la création de liens physiques entre les chromosomes via des échanges réciproques d'informations génétiques entre les homologues, ou crossing-over (CO). La formation d'au moins un CO, dit obligatoire, entre chaque paire d'homologues est requise pour une ségrégation réductionnelle appropriée. De plus, ce brassage génétique permet de générer des diversités génétiques au sein des populations. Cependant, les erreurs de la méiose sont les principales causes d'aneuploïdie, de malformations congénitales et d'infertilité.

Lors de la recombinaison méiotique, la formation des DSB est catalysée par un complexe du type topoisomérase, composé de l'enzyme SPO11 et de son partenaire de liaison TOPOVIBL. Toutefois, l'activité de ce complexe nécessite un ensemble partiellement conservé de protéines régulatrices, qui sont sept chez les mammifères : REC114, MEI4, IHO1, HORMAD1/2, MEI1, ANKRD31. Malgré leur importance lors de la formation des DSB, le rôle exact de ces protéines et le mécanisme moléculaire de leurs actions ne sont pas bien caractérisés. Mon travail s'est porté sur les protéines REC114, MEI4, et IHO1. Le but de mon projet de thèse était de réaliser une étude biophysique et structurale de ces facteurs protéiques afin de comprendre leur organisation et de caractériser les détails moléculaires de leurs interactions.

Les constructions protéiques d'IHO1 et du sous-complexe REC114-MEI4 ont été produites en bactérie puis purifiées séparément. Les échantillons protéiques ont été caractérisés par plusieurs techniques (SEC-MALLS, spectrométrie de masse et AUC) pour déterminer leurs états oligomériques. Les différentes analyses ont établi que la protéine IHO1 est tétramérique alors que le sous-complexe REC114-MEI4 est présent principalement sous forme hétérotrimérique en solution (2REC114/1MEI4). Durant cette

thèse, l'interaction entre le sous-complexe REC114-MEI4 et la protéine IHO1 a été identifiée initialement par la chromatographie d'exclusion de taille, puis elle a été confirmée et son affinité a été caractérisée par des analyses ITC. Ce travail a permis de mettre en évidence le rôle de la partie N-terminale d'IHO1 et du domaine PH de REC114 pour cette interaction. La caractérisation structurale du complexe REC114-MEI4 et de la protéine IHO1 seule ou en complexe a été initiée par plusieurs techniques combinant la cristallographie, la résonance magnétique nucléaire et microscopie électronique. Actuellement, des cristaux sont obtenus et sont en cours d'optimisation pour le fragment structuré d'IHO1, tandis que les caractérisations du complexe REC114-MEI4 par RMN et du complexe REC114-IHO1-MEI4 par la microscopie électronique nécessitent encore des optimisations. Par ailleurs, la modélisation de la structure tridimensionnelle des protéines par AlphaFold2 nous a permis d'identifier les résidus impliqués dans la dimérisation de REC114 et dans l'interaction avec MEI4 et IHO1. Ceux-ci ont été mutés et analysés *in vitro*. Un projet de test *in vivo* de l'effet de ces mutations est en cours en collaboration avec l'équipe du Dr. Bernard De Massy (IGH Montpellier). Les résultats obtenus durant ce travail seront poursuivis par des études structurales et fonctionnelles plus approfondies sur ces protéines méiotiques, seules ou en présence de partenaires supplémentaires.

## Summary

Meiosis is a specialized cell division program that allows the formation of haploid gametes from diploid cells in sexually reproducing organisms. The halving of this genetic content is a result of the cell division process where only one round of DNA replication is followed by two divisions (Meiosis I and Meiosis II). Chromosome reduction occurs during meiosis I where maternal and paternal homologous chromosomes are separated. This process involves the formation of DNA double-strand breaks that are repaired by homologous recombination, also called meiotic recombination. This mechanism allows the formation of physical links, the crossing-overs (CO), between homologues through reciprocal exchanges. The formation of at least one CO per homologue pair, called obligatory CO, is required for faithful meiotic recombination. As such, meiosis generates genetic diversity in a population; however, meiotic errors are also the major causes of aneuploidy, congenital anomalies and infertility.

DSB formation is catalyzed by the SPO11-TOPOVIBL complex and a set of partially conserved DSB formation-regulating proteins. In mammals, there are at least seven proteins that are essential for DSB formation: REC114, MEI4, IHO1, HORMAD1/2, MEI1, ANKRD31, but the exact role of these proteins and the molecular mechanism of their action remain unclear. My work focused on REC114, MEI4, and IHO1 proteins. The aim of my thesis was to perform a biophysical and structural analysis of these protein factors and characterize the molecular details of their interactions.

Constructs of IHO1 and the REC114-MEI4 sub-complex were produced in bacteria and then purified separately. Protein samples were characterized by several techniques (SEC-MALLS, mass spectrometry and AUC) to determine their oligomeric states. The different analyses established that the IHO1 protein was tetrameric while the REC114-MEI4 sub-complex is present mainly in heterotrimeric form (2REC114/1MEI4). During this thesis, the interaction between the REC114-MEI4 sub-complex and the IHO1 protein was initially identified by size exclusion chromatography and then confirmed by ITC analyses. This work revealed the role of the N-terminal part of IHO1 and the PH domain of REC114 in this interaction. The structural characterization of the REC114-MEI4 complex and of the IHO1 protein alone or in a complex with REC114 was initiated by several techniques combining crystallography, nuclear magnetic resonance and electronic

microscopy but still requires optimizations. Nevertheless, the modelling of the three-dimensional structure of the proteins by AlphaFold2 allowed us to obtain structural insight into the dimerization of REC114 and its interactions with MEI4 and IHOI. I identified residues involved in these interactions and their mutagenesis was used to validate the structural models *in vitro*. The *in vivo* characterization of the effect of these mutations is been performed in collaboration with the team of Dr. Bernard De Massy (IGH Montpellier). The results obtained during this work will be followed up with further structural and functional studies on these meiotic proteins, either alone or in the presence of additional partners.

# 1 Introduction



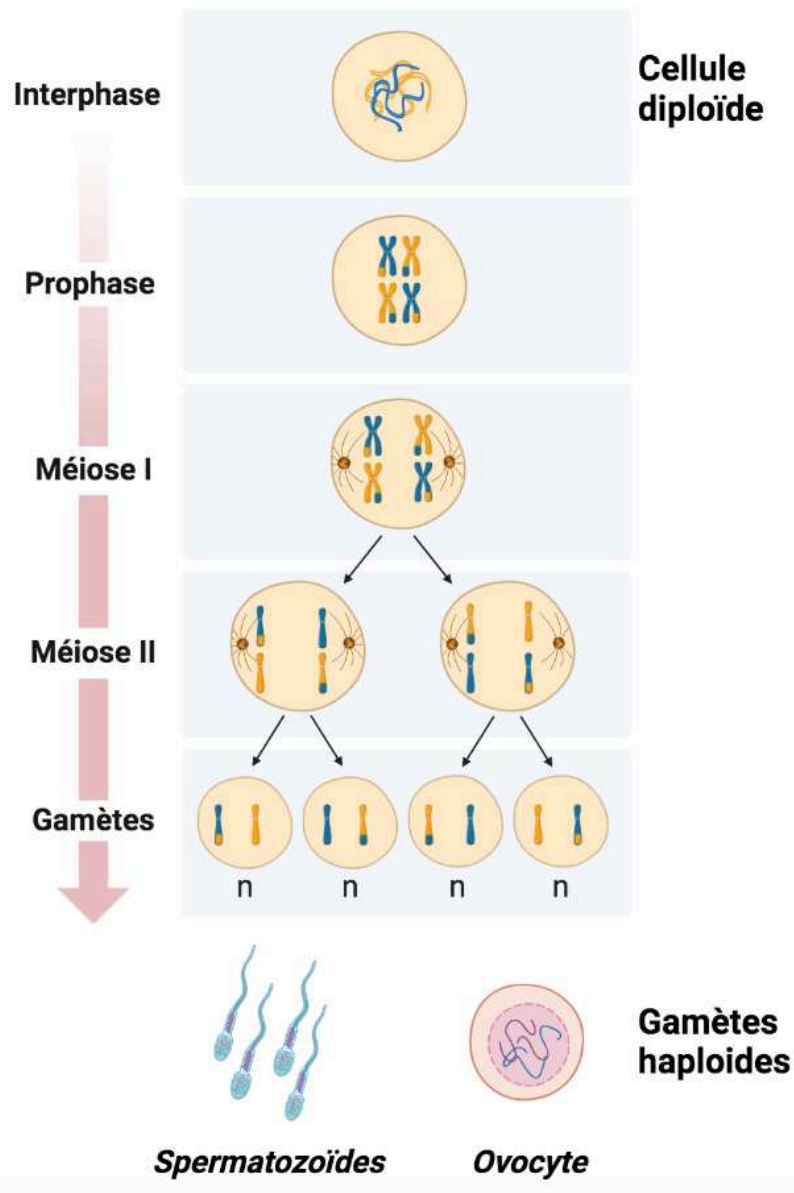


## 1.1 Définition et rôle de la méiose

La méiose est un mécanisme fondamental par lequel les organismes à reproduction sexuée produisent les gamètes mâles et femelles. C'est un mode spécialisé de division cellulaire qui affecte les cellules germinales et permet de générer des cellules haploïdes ( $n$  chromosomes) à partir de cellules mères diploïdes ( $2n$  chromosomes). Le nombre de chromosomes est donc réduit de moitié lors de la formation des gamètes. Cela est dû au fait qu'un seul cycle de réplication du génome est suivi de deux divisions consécutives (méiose I et méiose II). La première division est propre à la méiose et elle sépare les chromosomes homologues d'origine paternelle et maternelle. Par conséquent, elle est à l'origine de la réduction de la ploïdie et elle produit deux cellules filles haploïdes. La méiose II est ensuite initiée directement dans ces cellules filles et sépare les chromatides sœurs de chaque chromosome, comme lors d'une division mitotique, permettant ainsi la production de quatre cellules haploïdes par une division méiotique (Figure 1). La fusion ultérieure du matériel génétique des gamètes haploïdes lors de la fécondation permet de restituer la ploïdie de l'espèce.

Chez les mammifères, la chronologie et le déroulement de la production des gamètes sont différents entre les deux sexes. Chez le mâle, les cellules germinales ne commencent la méiose qu'à la puberté et le stock de cellules souches, spermatogonie, prolifère de façon continue par division mitotique. Cela permet aux mâles de produire des spermatozoïdes tout au long de leur vie. En revanche, toutes les cellules germinales femelles initient la méiose pendant la vie fœtale pour devenir des ovocytes. Après cela, la première prophase de la méiose s'arrête avant la naissance et reprend périodiquement après la puberté. Par conséquent, le nombre d'ovocytes chez les femelles adultes est limité. De plus, contrairement aux cellules germinales mâles, qui produisent quatre spermatozoïdes, les ovocytes ne produisent qu'un seul ovocyte haploïde, et les chromosomes restants sont éliminés lors de la méiose I et II dans des petites structures, appelées globules polaires (Handel and Schimenti, 2010; Herbert et al., 2015).

La ségrégation réductionnelle des chromosomes homologues lors de la méiose I nécessite, chez la majorité des espèces, l'établissement de liens physiques entre les chromosomes. Ces connexions sont établies via des échanges réciproques entre les chromosomes homologues créés par la recombinaison homologue, appelée recombinaison



**Figure 1. La division méiotique.** La cellule germinale diploïde commence la division méiotique par une phase d'interphase où elle réplique son ADN et elle se prépare pour la division méiotique. Ensuite, les chromosomes homologues de la même paire (uniquement, deux paires sont représentées ici) vont s'aligner, s'apparier et s'échanger du matériels génétiques lors de la prophase méiotique I. La méiose I est une division réductionnelle, car elle réduit d'un facteur deux la ploïdie par la séparation des chromosomes homologues. La méiose II est une division équationnelle, comme une division mitotique, et elle sépare les chromatides sœurs. La méiose permet ainsi la production de gamètes haploïdes à partir d'une cellule mère diploïde. Ces gamètes sont de 4 spermatozoïdes pour les mâles et d'un seul ovocyte pour les femelles (Adaptée de 'Non-disjunction in Meiosis', par BioRender.com (2022). Extrait de <https://app.biorender.com/biorender-templates>).

méiotique (De Massy, 2013; Yadav and Claeys Bouuaert, 2021). En plus d'être important pour la ségrégation des chromosomes, l'échange d'informations génétiques, généré par la recombinaison homologue, entraîne un brassage génétique entre les chromosomes maternels et paternels, appelé brassage intrachromosomique. Ensuite, les chromosomes homologues subissent un autre type de brassage, cette fois-ci interchromosomique, suite à une migration aléatoire et de façon indépendante vers les cellules filles. Grâce à ces deux types de brassages lors de la ségrégation des chromosomes, les gamètes produits possèdent de nouvelles combinaisons d'allèles non présentes dans la cellule parentale, aboutissant à des diversités génétiques.

En revanche, l'obtention de gamètes avec un patrimoine génétique complet nécessite une recombinaison méiotique correcte et une ségrégation chromosomique appropriée. Une distribution inégale des chromosomes conduit à la formation de gamètes aneuploïdes avec trop peu ou trop de chromosomes. La fécondation des gamètes avec de tels déséquilibres chromosomiques conduit à des aneuploïdies. Chez l'homme, on estime que 10 à 30% des ovules humains fécondés ont le mauvais nombre de chromosomes (Hassold and Hunt, 2001). Les embryons qui possèdent une seule copie d'un autosome (monosomie autosomique) ne sont pas viables. L'héritage d'une copie supplémentaire d'un autosome (trisomie autosomique) est également associé à de graves anomalies du développement et représente, avec les monosomies, environ un tiers de toutes les fausses couches (Hassold and Hunt, 2001; MacLennan et al., 2015). Certaines trisomies autosomiques et certaines aneuploïdies des chromosomes sexuels sont compatibles avec la naissance chez l'homme et elles en représentent 0,3 % (Hassold and Hunt, 2001). La trisomie du chromosome 21, cause du syndrome de Down, est l'aneuploïdie la plus fréquente (Herbert et al., 2015; MacLennan et al., 2015). Un autre type d'anomalie méiotique correspond à des réarrangements chromosomiques et à des altérations génomiques qui entraînent le gain ou la perte de séquences génomiques lors d'une recombinaison méiotique défectueuse. Ces réarrangements, qui se produisent lors de la méiose, sont transmissibles à la descendance et sont à l'origine de plusieurs maladies (Baudat et al., 2013). Le rôle essentiel de la recombinaison méiotique et les risques qu'elle peut engendrer soulignent l'importance de comprendre son mécanisme moléculaire.

## 1.2 Dérroulement de la méiose

Comme pour la mitose, avant d'entrer dans la méiose, les cellules en division subissent une étape de croissance et de développement appelée interphase. Durant ce temps, la cellule réplique aussi son patrimoine génétique, de sorte que chaque chromosome est constitué de deux chromatides-sœurs qui sont maintenues ensemble par des cohésines.

La cellule entame ensuite les deux cycles successifs de division, méiose I et II, chacune de ces divisions est constituée de quatre étapes ; prophase, métaphase, anaphase et télophase.

### 1.2.1 Prophase I, stade de la recombinaison méiotique

La prophase I est l'étape la plus longue et la plus complexe de la méiose. Elle inclut des événements nécessaires pour la séparation des homologues lors de la première division méiotique. Au cours de cette étape, les paires de chromosomes homologues doivent s'apparier, s'échanger du matériel génétique par recombinaison homologue et s'associer étroitement grâce à la formation du complexe synaptonémal (Kleckner, 1996; Zickler and Kleckner, 2015). Ces mécanismes sont associés à des transitions dans la structure chromosomique, ce qui permet de définir les cinq étapes de la prophase I : leptotène, zygotène, pachytène, diplotène et diacinèse (Figure 2A) (Zickler and Kleckner, 1998).

#### *1.2.1.1 Leptotène, stade de la formation des axes chromosomes et l'initiation de la recombinaison*

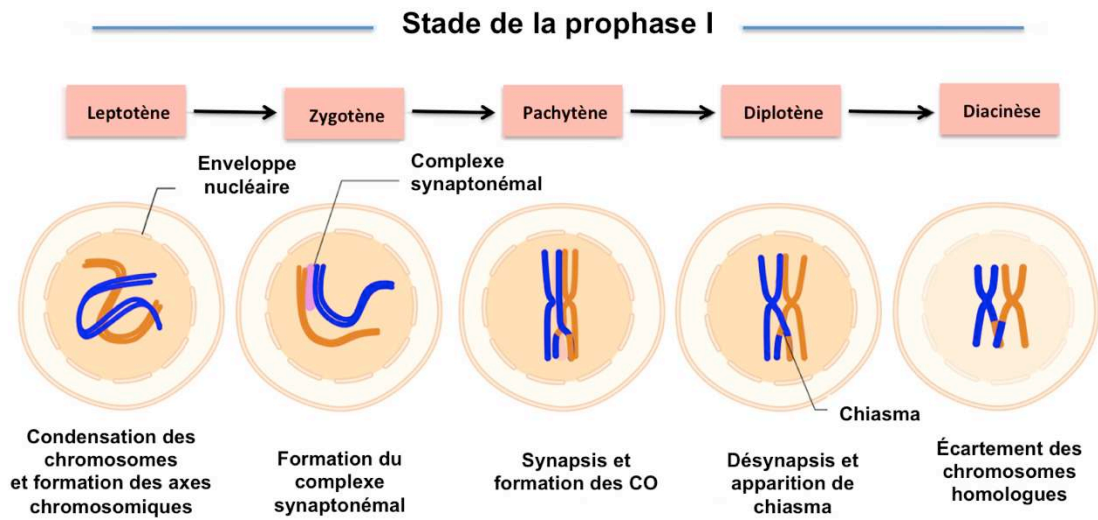
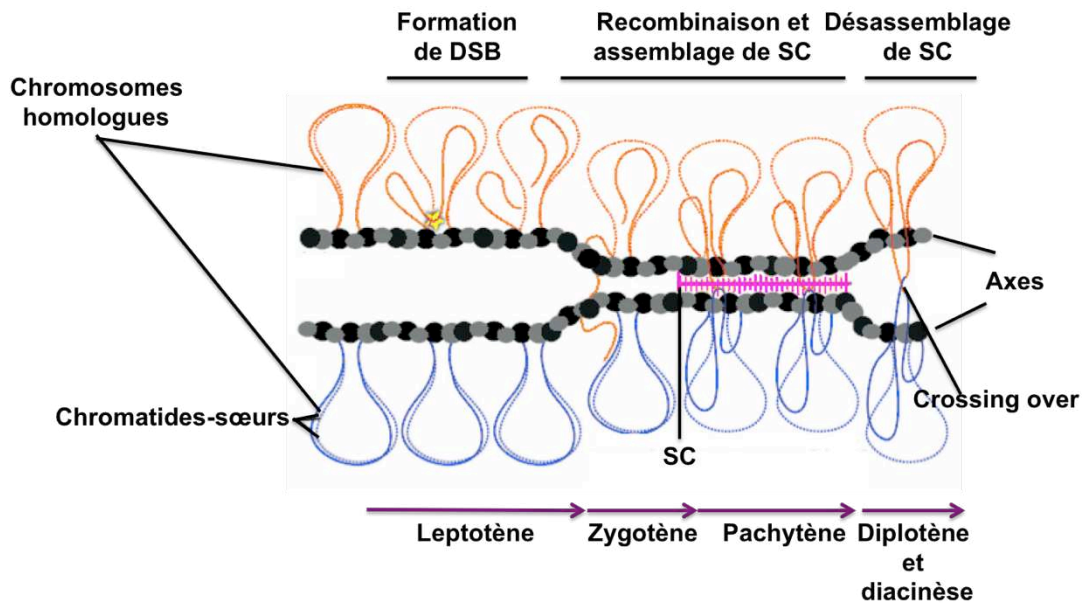
Au cours de ce stade, les chromosomes commencent leur condensation et sont constitués de deux chromatides sœurs qui apparaissent sous la forme de fils minces. La chromatine des deux chromatides sœurs est disposée sous forme de boucles reliées aux axes chromosomiques (Zickler and Kleckner, 1999). Ces axes chromosomiques constituent une structure multi-protéinique appelée élément axial, qui se forme le long de chaque paire de chromatides sœurs (Kleckner, 2006; Zickler and Kleckner, 2015). Chez les mammifères, ces axes sont constitués de complexes de cohésines (notamment le complexe de cohésines méiotiques contenant RAD21L et celui contenant REC8), de protéines du complexe synaptonémal (SYCP3 et SYCP2) et de protéines à domaine HORMAD (HORMAD1) (Grey and de Massy, 2021). Ces axes constitueraient une plateforme pour le recrutement

des protéines nécessaires à l'initiation de la recombinaison par la formation, à ce stade, de plusieurs cassures d'ADN double brin (DSB, pour Double-Strand Breaks) le long du génome (Figure 2B) (Baudat et al., 2013; Zickler and Kleckner, 2015).

Le passage du leptotène au zygotène se caractérise également par un arrangement spécifique des chromosomes qui se définit par l'attachement et le regroupement des télomères à l'enveloppe nucléaire. Ils forment un bouquet chromosomique qui pourrait faciliter l'appariement des chromosomes (Scherthan, 2001). Dans les cellules germinales de mammifères, les protéines à domaine SUN (SUN1), de la membrane nucléaire interne, sont attachées aux télomères et interagissent avec les protéines à domaine KASH, localisées sur la membrane nucléaire externe. Les complexes protéiques SUN-KASH relient les chromosomes au cytosquelette, favorisant les mouvements chromosomiques et l'appariement des homologues pendant la méiose (Ding et al., 2007; Morimoto et al., 2012).

### ***1.2.1.2 Zygotène, stade de la formation du complexe synaptonémal et de la progression de la recombinaison***

Durant cette étape, les chromosomes continuent leurs condensations et ils apparaissent plus courts et plus gros. Les extrémités d'ADN, issues des DSB, s'engagent dans la recherche de séquences complémentaires homologues. L'hybridation entre les deux morceaux d'ADN forme des sites d'interaction physique et permet l'alignement et l'appariement des chromosomes (Page and Hawley, 2004). Ces derniers deviennent ensuite étroitement associés (*synapsis*) par la mise en place d'un complexe protéique appelé complexe synaptonémal (SC pour synaptonemal complex) qui relie progressivement les chromosomes homologues sur toute leur longueur (Page and Hawley, 2004; Zickler and Kleckner, 2015) (Figure 2B). Ce complexe correspond à une structure tripartite formée par le rapprochement étroit des axes des chromosomes homologues, renommés éléments latéraux, et leur connexion par une région centrale (Zickler and Kleckner, 1999). Cette dernière est composée de filaments transversaux et d'un élément central. Les filaments transversaux consistent en un réseau de dimères de la protéine SYCP1 (pour synaptonemal complex protein 1) (Liu et al., 1996), reliant les éléments latéraux aux protéines de l'élément central : SYCE 1, 2, 3 (pour synaptonemal complex central element protein 1, 2 et 3), TEX12 (pour Testis Expressed 12) et SIX6OS1 (Costa et al., 2005; Fraune et al., 2012; Gómez-H et al., 2016; Hamer et al., 2006; Schramm et al., 2011).

**A****B**

**Figure 2. Stade de la prophase I et structures chromosomiques associées.** (A) Résumé des événements, et de la dynamique des chromosomes homologues lors de chaque stade de la prophase I (adaptée de 'Stages of Meiotic Prophase I', par BioRender.com (2022). Extrait de <https://app.biorender.com/biorender-templates>). (B) Relation entre la recombinaison méiotique et la structure des chromosomes homologues lors de la prophase I. Lors du stade leptotène, les chromosomes adoptent une structure en boucles chromatiniennes liées à un axe protéique (noir et gris). Les cassures doubles brins (DSB) sont introduites aux niveaux des boucles d'ADN. Lors du stade zygotène, la progression de la recombinaison rapproche les chromosomes homologues, ce qui initie l'assemblage du complexe synaptonémal (SC) entre les axes chromosomiques (en rose). Le SC atteint sa polymérisation complète au stade pachytène puis il se désassemble au stade diplotène et diacinèse, laissant les chromosomes liés ensemble par les crossing-over (Adaptée de Yadav and Claeys Bouuaert, 2021).

Il existe une relation d'interdépendance entre le SC et la réparation des DSB. Les DSB sont nécessaires pour l'initiation de l'appariement chromosomique (Baudat et al., 2013; Kauppi et al., 2013; Romanienko and Camerini-Otero, 2000), alors que le SC fournit un cadre structural pour la réparation des DSB et la formation d'échanges réciproques entre chromosomes homologues (Page and Hawley, 2004; Zickler and Kleckner, 1998). Cependant, des exceptions ont été observées chez certains organismes, comme *C. elegans* ou de *D. melanogaster*, chez qui l'appariement des homologues se produit indépendamment de la recombinaison (Dernburg et al., 1998; McKim et al., 1998; Page and Hawley, 2004). Chez la souris, un appariement chromosomique précoce est également observé avant la formation des DSB (Boateng et al., 2013).

### ***1.2.1.3 Pachytène, stade de la formation des crossing-over***

Les chromosomes homologues, à cette étape, sont de courtes structures bien condensées et intimement associées par le SC (Zickler and Kleckner, 1998) (Figure 2B). La structure générée est nommée bivalents ou tétrades. Les chromosomes sexuels X et Y des mammifères sont hétéromorphes et s'apparient uniquement dans les portions homologues appelées régions pseudo-autosomiques (PAR pour Pseudo-autosomal region) (Handel, 2004). À ce stade, certains intermédiaires de recombinaison, formés précédemment, sont résolus en générant des échanges réciproques entre chromosomes homologues, appelés enjambements ou crossing-over (CO) (Zickler and Kleckner, 2015).

### ***1.2.1.4 Diplotène et Diacinèse, stade de désynapsis***

À ce stade, le complexe synaptonémal se dégrade (désynapsis). Les chromosomes homologues commencent à s'écarter, mais ils restent connectés par les sites de crossing-over qui deviennent visibles sous forme de chiasmata (Figure 2B) (Zickler and Kleckner, 1998). C'est à cette étape que l'ovocyte dans l'embryon arrête sa méiose en attendant de la reprendre à la puberté (Bowles and Koopman, 2007). Pendant la diacinèse, les chromosomes terminent leur condensation et se détachent de l'enveloppe nucléaire. Ce dernier disparaît et les fibres du fuseau achromatique s'assemblent pour séparer les chromosomes homologues.



## 1.2.2 Métaphase, anaphase et télophase 1

À la métaphase I, les chromosomes homologues sont co-orientés sur la plaque métaphasique et sont connectés aux fibres du fuseau méiotique provenant des pôles de la cellule. L'Anaphase I accomplit la séparation, chaque chromosome migre vers un pôle de la cellule qui est opposé à celui de son homologue. La télophase I couplée à la cytodièrese I produit deux cellules filles, chacune avec un exemplaire de chaque homologue.

## 1.2.3 Méiose II

La méiose II se poursuit dans chaque cellule fille sans l'étape de la réplication du matériel génétique. Les cellules passent par une courte prophase II et elles réassemblent le fuseau méiotique pour séparer les chromatides sœurs cette fois-ci. Pendant la métaphase-II, les chromosomes sont alignés sur la plaque métaphasique et les chromatides sœurs sont connectées aux pôles opposés via le fuseau de microtubules achromatique. Les ovocytes bloquent la méiose II à cette étape et ne la continueront qu'après la fécondation. À l'étape de l'anaphase-II, les chromatides sœurs migrent vers des pôles opposés de la cellule. La télophase II et la cytodièrese II viennent ensuite finir la division et séparer le cytoplasme de la cellule mère en deux cellules filles.

## 1.3 Mécanisme moléculaire de la recombinaison méiotique

La recombinaison méiotique correspond à l'introduction de plusieurs DSB tout le long du génome et leur réparation par recombinaison homologue (De Massy, 2013). Il s'agit d'un mode de réparation semblable à la réparation par recombinaison homologue dans le cycle cellulaire mitotique. Néanmoins, et contrairement aux cellules somatiques, les DSB de la méiose sont dites programmées et elles ne sont pas le résultat des différents agents qui endommagent l'ADN (Andersen and Sekelsky, 2010). Une deuxième différence importante est que la séquence utilisée pour la réparation des DSB méiotique se trouve le plus souvent sur le partenaire homologue, alors que lors de la réparation mitotique de l'ADN, elle se trouve majoritairement sur la chromatide sœur (Haber, 2000; Sung and Klein, 2006). Cette forte préférence pour la recombinaison interhomologue durant la méiose est importante, car elle permet l'appariement des chromosomes et la formation des CO requis pour leurs



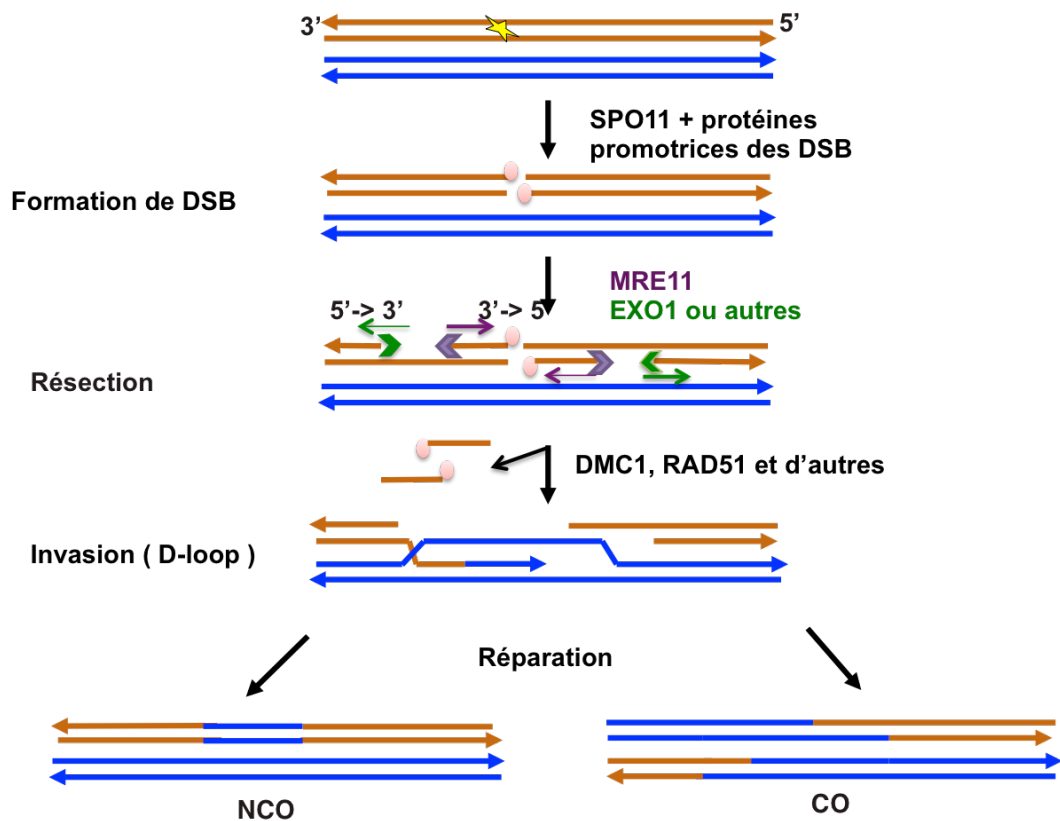
séparations lors de la métaphase I, contrairement à la recombinaison entre chromatides sœurs (Kohl and Sekelsky, 2013).

Les différents facteurs impliqués dans chaque étape de la recombinaison méiotique ont été identifiés, à l'origine, chez la levure mais un bon nombre de ces acteurs moléculaires sont évolutivement conservés chez les mammifères (Baudat et al., 2013).

### 1.3.1 Formation et maturation des cassures double brin de l'ADN

La recombinaison méiotique est initiée par l'introduction de plusieurs cassures d'ADN à travers le génome. Ces DSB sont catalysées par l'endonucléase SPO11, ainsi qu'un ensemble de protéines accessoires (Baudat et al., 2013; Bergerat et al., 1997; Keeney et al., 1997). Les détails moléculaires de cette étape de recombinaison seront abordés dans la section *Mécanismes moléculaires de la formation des DSB*.

Après le clivage de l'ADN, SPO11 reste lié d'une manière covalente aux extrémités 5' des DSB (Keeney et al., 1997). L'étape de maturation permet d'enlever la protéine SPO11 des sites de DSB et de générer des extrémités 3' simple brin. Selon le mécanisme moléculaire de la levure *S. cerevisiae*, c'est le complexe MRX, composé de Mre11, Rad50 et Xrs2 (MRN chez les mammifères-pour MRE11, RAD50 et NBS1), qui détecte les DSB et coopère avec la nucléase Sae2 (CtIP chez l'homme) pour produire une coupure d'ADN sur le brin liant SPO11 via l'activité endonucléase de Mre11 (Cannavo and Cejka, 2014; Ehmsen and Heyer, 2008; Neale et al., 2005). Cette coupure sert comme un point de départ, pour Mre11 d'une part, qui libère SPO11 lié à quelques nucléotides grâce à son activité exonucléase 3'→5', et à l'exonucléase Exo1 d'autre part, qui utilise l'activité 5'→3' pour faire apparaître des extrémités 3' simple brin (Figure 3) (Garcia et al., 2011; Zakharyevich et al., 2010). Chez la souris, l'exonucléase EXO1 n'est pas la déterminante principale de la résection 5'→3'. L'inhibition de l'activité de cette nucléase ne montre qu'une réduction minimale de la résection, suggérant la présence d'exonucléases alternatives pour la résection 5'→3' (Paiano et al., 2020; Yamada et al., 2020). Les extrémités simple brin, générées de cette résection, sont liées à des protéines de liaison à l'ADN simple brin, notamment la protéine RPA, pour être protégées de la dégradation nucléolytique et de la formation de structures secondaires (Ehmsen and Heyer, 2008; Ribeiro et al., 2016).



**Figure 3. Formation et réparation des DSB lors de la recombinaison méiotique.** Représentation des principales étapes de la recombinaison entre les chromosomes homologues. Une chromatide de chaque chromosome est représentée ici. SPO11 (ronds roses) collabore avec plusieurs protéines auxiliaires pour catalyser la formation des DSB (étoile jaune). L'activité endonucléolytique de MRE11 (en violet) permet de libérer SPO11 liées à des oligonucléotides courts. Un traitement ultérieur par une activité exonucléase 3'->5', probablement via l'exonucléase EXO1 (en vert), permet l'exposition de longues extrémités 3' simple brin. Ces dernières sont prises en charge par la protéine RPA (non montrée) puis par DMC1 et RAD51, ainsi que leurs activateurs/cofacteurs, pour envahir le chromosome homologue à la recherche d'une séquence complémentaire pour hybridation. Ces intermédiaires de recombinaison (D-loop) sont ensuite réparés pour donner naissance à des produits qui contiennent ou pas un échange réciproque entre chromosomes homologues (CO ou NCO).

### 1.3.2 Invasion et recherche d'homologie

Chez la plupart des organismes eucaryotes, les extrémités ADN-RPA recrutent deux recombinaisons hautement conservées : RAD51 et DMC1 (Bishop et al., 1992; Shinohara et al., 1992). Ce sont des homologues de RecA, une protéine essentielle pour la réparation d'ADN par recombinaison homologue chez la bactérie (Shinohara and Shinohara, 2004). RAD51 est impliquée à la fois dans la recombinaison mitotique et méiotique alors que DMC1 est uniquement méiotique (Bishop et al., 1992). La coopération entre ces deux protéines est nécessaire pour une recombinaison méiotique efficace (Shinohara and Shinohara, 2004). Toutefois, il semble que l'échange de brins d'ADN entre chromosomes homologues soit favorisé par DMC1, la recombinaison RAD51 jouerait un rôle régulateur essentiel en favorisant l'assemblage de DMC1 sur les sites de la recombinaison (Cloud et al., 2012). La fonction de transfert de brins de Rad51 est d'ailleurs inhibée chez *S. cerevisiae* pendant la méiose par la liaison d'une protéine appelée Hed1 (Busygina et al., 2008).

Les deux recombinaisons se polymérisent sur les extrémités simple brin pour former des structures de filaments nucléoprotéiques (Brown et al., 2015; Shinohara and Shinohara, 2004). Plusieurs cofacteurs / médiateurs ont été impliqués soit dans la favorisation et/ou dans la stabilisation de ces filaments chez les mammifères. Il s'agit notamment du complexe HOP2/MIND1 nécessaire pour promouvoir l'activité de DMC1 (Petukhova et al., 2005; Pezza et al., 2007), de la protéine BRCA2 qui lie DMC1 et RAD51 et assure leurs bonnes localisations (Martinez et al., 2016; Thorslund et al., 2007), ainsi que plusieurs paralogues de RAD51 (RAD51B, RAD51C, RAD51D, XRCC2 et XRCC3) (Zelensky et al., 2014). Ensuite, les complexes recombinaisons-ADN envahissent l'ADN double brin intact du chromosome homologue à la recherche d'une séquence complémentaire dans un processus appelé invasion de brin (Single-End Invasion) (Hunter and Kleckner, 2001). L'hybridation de ce brin avec le brin complémentaire sur le chromosome homologue crée un ADN hybride (hétéroduplex) et déplace le brin complémentaire de la région envahie, formant ainsi une structure en boucle de déplacement (D-loop pour Displacement loop) (Hunter and Kleckner, 2001). Les protéines RAD51 et DMC1 sont ensuite dissociées de l'hétéroduplex et la synthèse d'ADN commence en utilisant l'extrémité 3' comme amorce et la séquence complémentaire de l'homologue envahi comme matrice (Gray and Cohen, 2016) (Figure 3).

### 1.3.3 Réparation des DSB

À ce stade, le mode de réparation des cassures doubles brins est choisi. Les intermédiaires de recombinaison peuvent être réparés soit en CO ou en NCO (Bishop and Zickler, 2004). Chez les mammifères, comme chez la plupart des organismes, le nombre de DSB dépasse largement le nombre de CO résultants, ce qui suggère que la majorité des DSB est réparée en NCO (Baudat and de Massy, 2007; Moens et al., 2002).

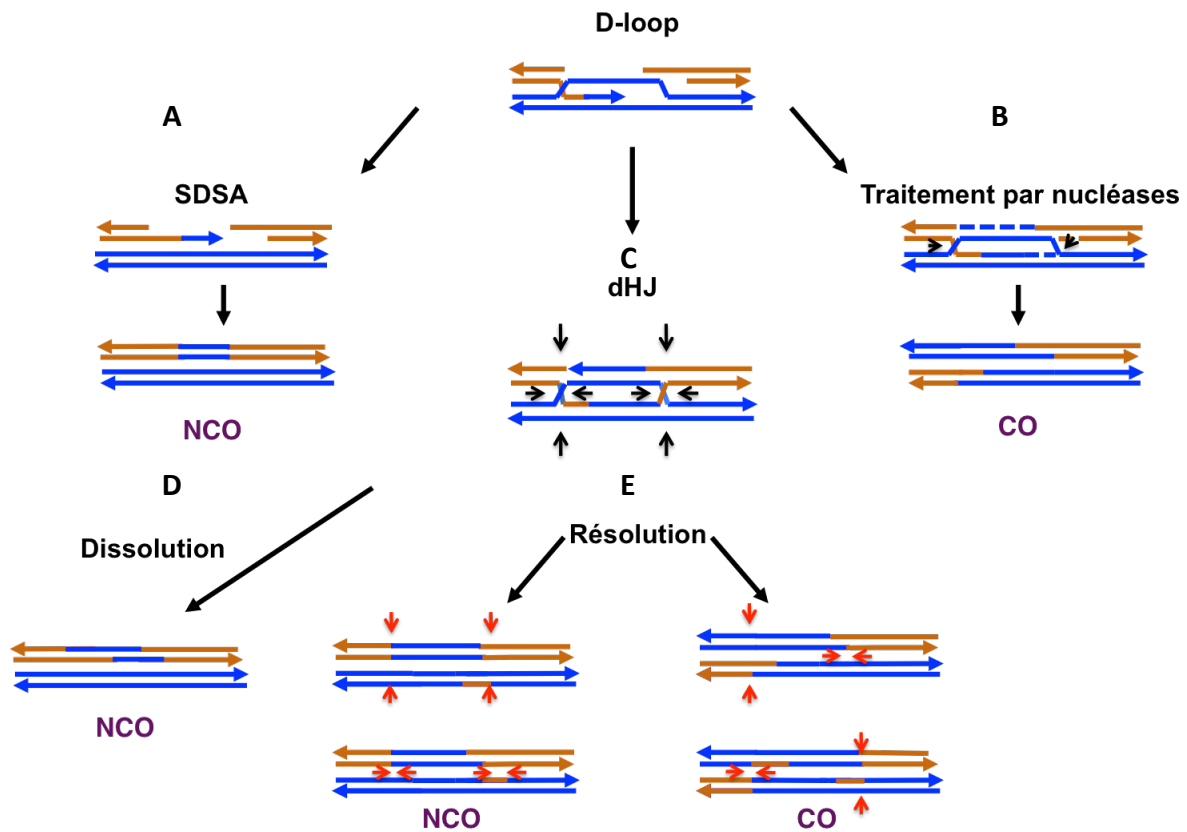
#### 1.3.3.1 *Synthesis-Dependent Strand Annealing (SDSA)*

La principale voie permettant la génération des NCO est la SDSA (pour Synthesis-Dependent Strand Annealing) (McMahill et al., 2007). Dans un premier temps, l'extrémité 3' envahissante est prolongée par la synthèse de l'ADN au sein de la D-loop. Ensuite, un complexe d'hélicase-topoisomérase (complexe BLM-TOPIIIa-RMI1/2 chez les mammifères et Sgs1-Top3-Rmi1 chez *S. cerevisiae*) dissout la D-loop (Fasching et al., 2015; Tang et al., 2015). L'extrémité nouvellement synthétisée se dissocie de sa matrice pour être réalignée avec son partenaire d'origine. Un duplex intact est ensuite reconstitué par la synthèse et la ligation des brèches d'ADN. Cette voie de réparation engendre un transfert unidirectionnel d'informations génétiques autour de la DSB, de la chromatide donneuse vers la chromatide réceptrice, mais ne produit pas un échange réciproque entre les deux chromosomes homologues (Figure 4A) (Baudat and de Massy, 2007; Pâques and Haber, 1999).

#### 1.3.3.2 *Formation d'une double jonction de Holliday*

Dans le cas où la D-loop ne serait pas déstabilisée, la synthèse de l'ADN prolonge davantage le brin envahisseur et permet d'élargir la D-loop jusqu'à l'hybridation du brin déplacé de la D-loop avec la deuxième extrémité de la DSB, c'est le SEC (pour Second-End DNA Capture). À ce stade, ces intermédiaires de recombinaison peuvent être soit traités par des nucléases pour former un CO (Heyer, 2004; Svendsen and Harper, 2010) (Figure 4B), soit poursuivre la synthèse de l'ADN pour réparer les séquences chromosomiques manquantes et générer une double jonction de Holliday (dHJ) (Figure 4C) (Schwacha and Kleckner, 1995). Ces dHJ peuvent être réparées de deux manières différentes, soit par la dissolution, soit par la résolution. Lors de la dissolution, les dHJ sont traitées par des hélicases et donnent naissance à des produits de NCO (Figure 4D) (Bizard and Hickson, 2014; Wu and Hickson, 2003). Tandis que lors de la résolution, les brins d'ADN impliqués

dans la dHJ sont clivés soit d'une façon symétrique pour générer des NCO ou d'une manière asymétriquement pour créer des événements de CO (Szostak et al., 1983) ( Figure 4 E). La résolution de ces dHJ lors de cette dernière étape est fortement biaisée vers la formation de CO (Allers and Lichten, 2001).



**Figure 4. Réparation des intermédiaires de recombinaison.** Les intermédiaires de recombinaison formés par l'invasion de brin peuvent être réparés de plusieurs manières. Une chromatide de chaque chromosome est représentée ici. **(A)** La D-loop peut être déstabilisée et le brin envahissant sera déplacé pour se réhybrider avec le côté opposé du DSB. Cette voie est le Synthesis-dependent strand annealing (SDSA) et son produit ne contient pas de CO. **(B)** Alternativement, la synthèse de l'ADN peut élargir la D-loop jusqu'à la capture de la deuxième extrémité de DSB. Des nucléases peuvent intervenir à cette étape pour traiter ces intermédiaires de recombinaison et former des crossing-over. **(C)** Si les nucléases n'interviennent pas, la ligation des intermédiaires de recombinaison forme une double jonction de Holliday (dHJ). **(D)** Cette dernière peut être dissoute par des hélicases en formant des produits sans CO. **(E)** Alternativement, les dHJ peuvent être résolues par des coupures symétriques, ce qui ne produit pas de CO. Elles peuvent aussi être clivées de manière asymétrique et produire des événements de CO. Les sites de coupure lors de la résolution sont indiqués en rouge (Adaptée de (Gray and Cohen, 2016)).

Les crossing-over formés lors de la réparation de l'ADN peuvent être divisés en deux classes, en fonction de leur distribution le long des chromosomes ainsi que de la machinerie moléculaire impliquée dans leur formation (Gray and Cohen, 2016).

### *Crossing over de classe 1*

La classe I de CO est également appelée la voie dépendante de ZMM (pour Zip1, 2, 3,4, Msh4, 5, Mer3, Spo16) (Pyatnitskaya et al., 2019). C'est une famille de protéines conservées, identifiées initialement chez *S. cerevisiae* puis chez la plupart des eucaryotes, dont la mutation entraîne un retard dans la progression de la méiose et une réduction de la fréquence des CO (Pyatnitskaya et al., 2019). La famille ZMM forme des sous-complexes protéiques fonctionnels parmi lesquels on peut citer le complexe Zip2-Zip4-Spo16 (SHOC1-TEX11-SPO16 chez les mammifères) (Arora and Corbett, 2019; De Muyt et al., 2018; Guiraldelli et al., 2018; Zhang et al., 2019) et le complexe Msh4-Msh5 (MutS $\gamma$ ) (Snowden et al., 2004). Ces derniers reconnaissent, stabilisent et protègent les intermédiaires de recombinaison, tels que les D-loop et dHJ, contre les hélicases qui provoquent leurs désassemblages ou leurs dissolutions (Pyatnitskaya et al., 2019). Le complexe MutS $\gamma$  assure aussi le recrutement des protéines nécessaires à la maturation et/ou la résolution des dHJ en CO, à savoir le complexe d'endonucléases MutL $\gamma$  (Mlh1 et Mlh3) qui agit en coopération avec l'hélicase Sgs1 et la nucléase Exo1 (Marsolier-Kergoat et al., 2018; Santucci-Darmanin et al., 2000; Zakharyevich et al., 2012, 2010).

De plus, certaines protéines de cette famille participent à la connexion des chromosomes en initiant la formation du complexe synaptonémal (Pyatnitskaya et al., 2019). C'est notamment le cas de Zip1 (SYCP1 chez les mammifères), une protéine de l'élément transversal du complexe synaptonémal qui est nécessaire à l'assemblage des autres membres ZMM (Chen et al., 2015; Voelkel-Meiman et al., 2019), de Zip3 (HEI10 et RNF212 chez les mammifères) qui favoriserait également le synapse des chromosomes (Agarwal and Roeder, 2000), et de Zip4 qui coordonnerait la formation des CO avec l'assemblage du complexe synaptonémal (Pyatnitskaya et al., 2021).

Chez la plupart des organismes, les CO de la classe 1 sont majoritaires. De plus, ils ont tendance à suivre une distribution régulièrement espacée sur les axes chromosomiques, c'est le phénomène d'interférence (Berchowitz and Copenhaver, 2010).

## *Crossing over de classe II*

La voie de CO de classe 2 est indépendante des protéines de la famille ZMM et elle représente qu'une fraction minoritaire des CO (10 % chez la souris) (Holloway et al., 2008). Cette voie est également indépendante du phénomène d'interférence (de los Santos et al., 2003; Holloway et al., 2008). Les précurseurs de ces CO seraient des intermédiaires de recombinaison aberrants et complexes qui ont échappé aux traitements des hélicases, comme l'hélicase Sgs1 et deviennent la cible de l'endonucléase Mus81 et son partenaire régulateur Mms4 (MUS81- EME1 chez les mammifères) (Holloway et al., 2011, 2008; Jessop and Lichten, 2008; Oh et al., 2008). Un certain nombre de résolvas ont été aussi évoquées pour leurs capacités à résoudre les intermédiaires de recombinaison, à savoir Slx1-Slx4 et Yen1 (SLX1-BTBD12 et GEN1 chez les mammifères) (De Muyt et al., 2012; Matos et al., 2011; Zakharyevich et al., 2012).

Du fait de leur importance dans la ségrégation des chromosomes et pour la diversité génétique, la formation des CO est hautement régulée pour contrôler à la fois le nombre et la distribution de ces événements (Wang et al., 2015). D'abord, le phénomène d'interférence, observé pour la classe I, permet une structuration spatiale des CO sur les chromosomes qui en subissent plusieurs. En effet, l'apparition d'un CO à une position défavorise l'apparition d'un autre à sa proximité, assurant ainsi une distribution régulière à l'échelle du génome (Hillers, 2004; Jones and Franklin, 2006). De plus, le phénomène de l'assurance permet la formation d'au moins un CO sur chaque chromosome, appelé CO obligatoire, pour garantir une ségrégation précise des homologues lors de la méiose 1 (Hunter, 2015; Wang et al., 2015). Enfin, le phénomène d'homéostasie maintient le nombre de CO stable malgré l'augmentation ou la diminution du nombre des DSB (Martini et al., 2006).

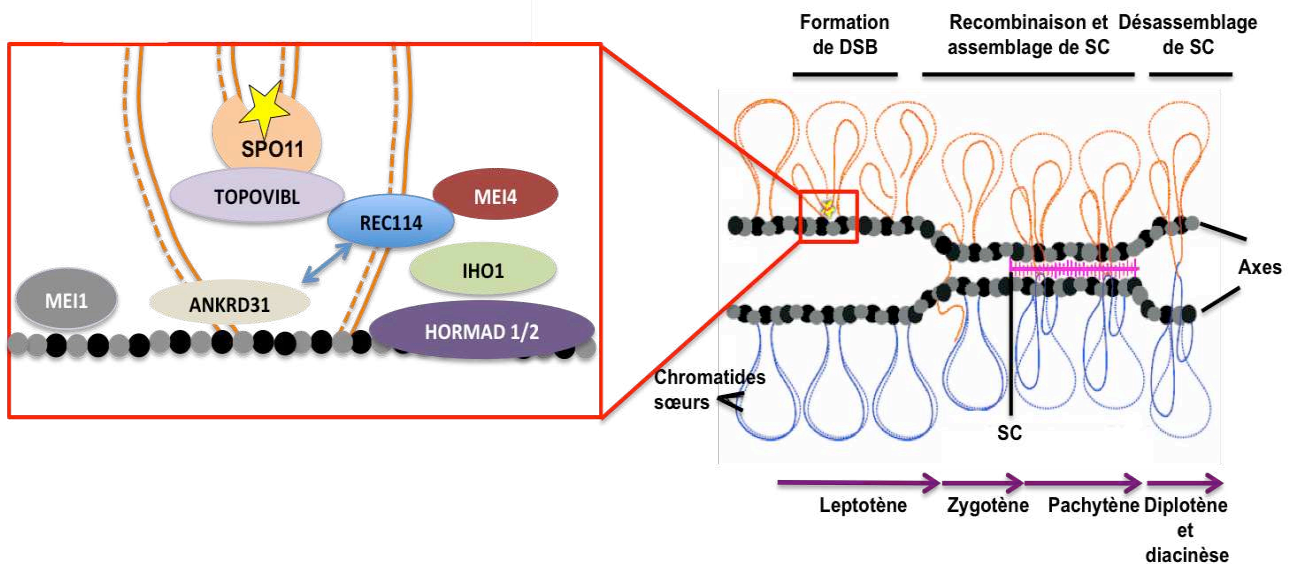
## 1.4 Mécanismes moléculaires de la formation des DSB

### 1.4.1 Facteurs protéiques impliqués dans la formation des DSB chez les mammifères

Les cassures d'ADN se forment dans une structure chromosomique spécifique à la méiose où la chromatine est disposée sous forme de boucles provenant des axes chromosomiques. Les DSB se forment préférentiellement dans l'ADN de ces boucles, mais les facteurs protéiques promoteurs des DSB ne s'assemblent que le long des axes. Il est alors supposé que les DSB ne sont introduites dans les boucles qu'après leur recrutement sur les axes chromosomiques, c'est le modèle TLAC (pour tethered loop axis complexes) (Blat et al., 2002; Borde and de Massy, 2013; Panizza et al., 2011).

Sur le plan moléculaire, la formation de DSB est catalysée par l'endonucléase conservée SPO11 dont l'activité dépend d'un ensemble de protéines partiellement conservées (Lam and Keeney, 2015). Ces acteurs moléculaires forment des sous-groupes fonctionnels qui ont été identifiés dans un premier temps chez *S. cerevisiae* (Yadav and Claeys Bouuaert, 2021). Les recherches menées ensuite chez d'autres organismes ont permis d'identifier les acteurs conservés et de révéler la présence de nouveaux facteurs qui sont propres à chaque espèce. À ce jour, les protéines connues pour leur implication dans la formation de ces coupures d'ADN chez les mammifères sont : SPO11, TOPOVIBL, REC114, MEI4, IHO1, HORMAD1/2, ANKRD31 et MEI1 (Boekhout et al., 2019; Fukuda et al., 2010; Keeney et al., 1999; Kumar et al., 2010, 2015; Papanikos et al., 2019; Reinholdt and Schimenti, 2005; T. Robert et al., 2016; Romanienko and Camerini-Otero, 2000; Stanzione et al., 2016; Wojtasz et al., 2009). Ces facteurs protéiques sont peu caractérisés. La compréhension de leurs mécanismes moléculaires, leurs organisations et leurs fonctions lors de l'initiation de la recombinaison reste très limitée. Cependant, les différentes recherches menées *in vivo* et/ou *in vitro*, surtout chez la souris, ont permis de montrer l'importance de ces facteurs protéiques dans la formation des DSB. Elles ont aussi révélé certains réseaux d'interactions entre ces facteurs protéiques, permettant d'établir un modèle de recrutement et d'assemblage de cette machinerie méiotique sur les sites de la formation des DSB chez les mammifères (Figure 5).

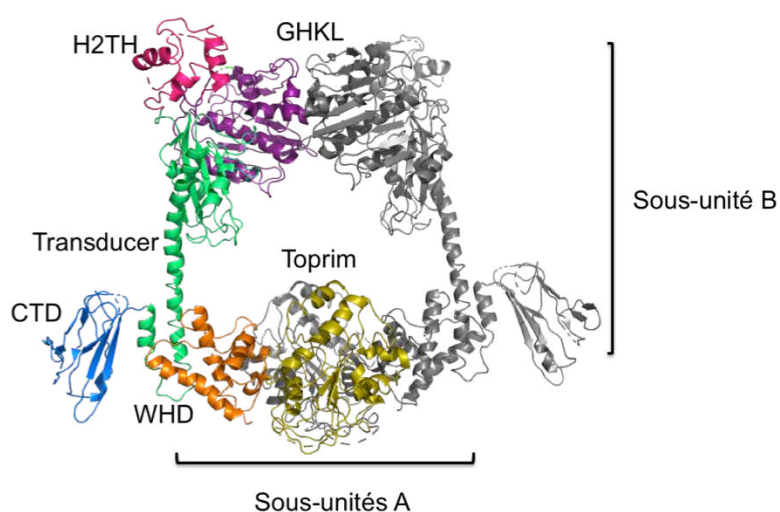




**Figure 5. Modèle de la formation des DSB chez les mammifères.** La formation des DSB se déroule dans la structure de boucles d'ADN liées aux axes protéiques. Les DSB se forment aux niveaux des boucles chromatiniennes tandis que les protéines responsables sont localisées aux niveaux des axes protéiques. Il a été proposé que les DSB se forment après leurs liaisons vers les axes chromosomiques. Les protéines d'axes chromosomiques HORMAD1/2 servent de plateforme pour le recrutement de IHO1 sur les axes, qui recrute à son tour le complexe REC114-MEI4. L'interaction de REC114 avec ANKRD31 semble être essentielle pour le recrutement de REC114 sur les régions pseudo-autosomiques des chromosomes sexuels et sa stabilisation sur le reste du génome. L'assemblage de ces protéines sur les axes chromosomiques entraîne la formation d'un recombinosome promoteur de DSB. Ce complexe recrute ou active le complexe catalytique SPO11-TOPOVIBL, via l'interaction entre REC114 et TOPOVIBL, induisant la formation des DSB (étoile jaune). La protéine MEI1 est également indispensable pour la formation des DSB, mais son rôle exact reste inconnu (Adaptée de (Stanzione et al., 2016; Yadav and Claeys Bouuaert, 2021)).

### 1.4.1.1 Complexe catalytique SPO11-TOPOVIBL

SPO11 est l'entité catalytique de la formation des DSB, c'est une endonucléase hautement conservée et similaire à la sous-unité A de la topoisomérase TopoVI d'archées (Bergerat et al., 1997; Keeney et al., 1997). Les topoisomérases du type II, à qui appartiennent TopoVI, hydrolysent l'ATP pour moduler la topologie de l'ADN. Elles clivent un duplexe d'ADN, passent à travers ces coupures le duplex intact, puis elles referment les coupures par une ligation (Chen et al., 2013). Les TopoVI sont composées de deux sous-unités (sous-unité A et sous-unité B) et agissent sous forme d'un homotétramère  $A_2-B_2$  (Figure 6) (Corbett et al., 2007; Graille et al., 2008). La sous-unité A assure la liaison et le clivage de l'ADN. Elle est constituée de deux domaines : le domaine 5Y-CAP (pour catabolite activator protein), qui contient la tyrosine catalytique, et le domaine Toprim (pour topoisomérase-primase) en C-terminal, qui coordonne les ions divalents importants pour la catalyse (Corbett et al., 2007; Graille et al., 2008; Nichols et al., 1999). La sous-unité B est impliquée dans la liaison et l'hydrolyse de l'ATP qui est nécessaire aux changements conformationnels de l'enzyme. Le module de liaison à l'ATP de cette sous-unité est semblable au domaine des ATPases GHKL (pour gyrase, heat-shock protein, histidine kinase et MuK1), également appelé Bergerat-fold. Ce dernier est relié à la sous-unité A grâce au domaine C-terminal transducteur (transducer). Chez certaines espèces, il existe un domaine C-terminal supplémentaire (CTD) qui adopte un repliement de type immunoglobuline (Corbett et al., 2007). Le domaine central de cette sous-unité, H2TH (pour helix-two-turn-helix), assure la liaison entre GHKL et le domaine transducteur (Corbett et al., 2007; Corbett and Berger, 2005, 2003; Graille et al., 2008).



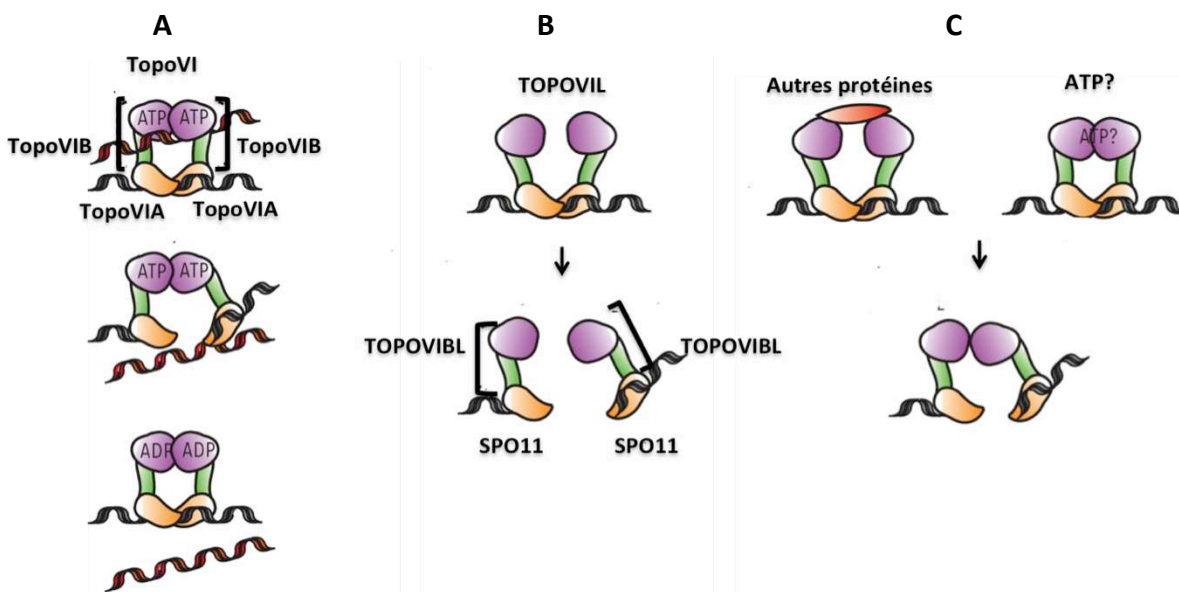
**Figure 6. Structure de TopoVI.** Représentation en ruban de la structure de l'hétérotétramère TopoVI de *M. Mazei* (PDB: 2Q2E). Un dimère A-B est représenté en gris, l'autre est coloré. GHKL: gyrase, HSP90, histidine kinase, domaine MutL. H2TH: helix-two-turn-helix. 5Y-CAP: catabolite activator protein. CTD: C-terminal domain (Corbett et al., 2007).

Pour assurer son action, les sous-unités A de la TopoVI se dimérisent et lient le duplex d'ADN qui sera clivé. Le deuxième duplex d'ADN pénètre dans l'enzyme par une porte formée par les domaines ATPase des deux sous-unités B (Figure 7A). La sous-unité B est également impliquée dans la reconnaissance de l'ADN (Wendorff and Berger, 2018). L'engagement du deuxième duplex active la dimérisation dépendante de l'ATP du domaine GHKL de la sous-unité B (Corbett et al., 2007). La dimérisation du domaine GHKL est transmise à la sous-unité A par le domaine transducteur pour activer le clivage de l'ADN. Les deux tyrosines du site actif réalisent une attaque sur le squelette de l'ADN, générant ainsi des intermédiaires 5'-phosphotyrosyl. L'hydrolyse de l'ATP induit un changement de conformation qui perturbe l'interface dimérique de la sous-unité A et permet le passage du brin non clivé (Corbett and Berger, 2004; Thomas Robert et al., 2016). À la fin, la coupure d'ADN est refermée, l'enzyme et soit relâchée ou engagée dans un second cycle.

D'après les similarités observées entre SPO11 et la sous-unité A de TopoVI, il a été proposé que la formation des DSB méiotiques soit catalysée par un mécanisme du type topoisomérase II. Cependant, la sous-unité A des topoisomérases n'est pas suffisante pour promouvoir la coupure de la molécule d'ADN. L'interaction avec son partenaire B s'avère indispensable pour assurer cette activité. Pour cette raison, des recherches ont été menées afin d'identifier la sous-unité B qui interagirait avec SPO11 pour assurer son activité catalytique et des protéines, équivalentes à la sous-unité B de la topoisomérase VI, ont été identifiées chez la souris (*M.musculus*), les levures (*S. cerevisiae*, *S. pombe*), la plante (*A. thaliana*) et la mouche (*D. melanogaster*) (T. Robert et al., 2016; Vrielynck et al., 2016). La formation des DSB méiotiques chez ces organismes impliquerait donc un complexe du type TopoVI, dénommé TOPOVI-Like (TOPOVIL). Toutefois, à la différence du complexe TopoVI, les protéines SPO11 restent liés d'une manière covalente aux extrémités qu'elles génèrent et elles n'assurent pas la religation des coupures d'ADN (Figure 7 B). Ce sont les étapes de la résection et de la réparation décrites précédemment, qui permettent, respectivement, de libérer SPO11 et de réparer les lésions (Keeney, 2008; Thomas Robert et al., 2016).

Sur le plan évolutif, et contrairement à SPO11 qui présente une forte conservation avec son homologue d'archée, la séquence primaire de la nouvelle sous-unité de TOPOVIL (TOPOVIBL chez la souris) a fortement divergé et certains motifs dans les domaines GHKL de plantes et de souris ne sont pas conservés (Thomas Robert et al., 2016). Par

ailleurs, la comparaison du modèle structural, prédit par AlphaFold2, pour le TOPOVIL de la souris avec la structure cristalline de son homologue d'archée a révélé la présence de similitudes dans l'organisation générale du complexe TOPOVIL, particulièrement au niveau du site catalytique de SPO11 ainsi que l'interface de liaison entre SPO11 et TOPOVIBL. En revanche, la sous-unité B de souris a nettement divergé. D'abord, les éléments structuraux permettant la formation du site de liaison à l'ATP, au niveau du GHKL, adoptent une conformation différente qui interfère avec la formation de ce site dans TOPOVIBL. De plus, le domaine H2TH, impliqué dans la dimérisation de TOPOVIB est absent chez la sous-unité murine, remettant en question sa capacité à fixer l'ATP et à se dimériser (Nore et al., 2022). Ainsi, le complexe TOPOVIL pourrait assurer sa fonction soit sans la dimérisation de la sous-unité B, soit en présence des facteurs protéiques supplémentaires qui stabilisent ou induisent la dimérisation du complexe (Figure 7 B, C) (Robert et al., 2017).



**Figure 7. Comparaison entre TopoVI et TOPOVIL (A)** Mode d'action de la topoisomérase d'archées TopoVI. TopoVI est constituée de deux sous-unités A et B et forme un tétramère  $A_2B_2$ . Les deux sous-unités A (en orange) se dimérisent et fixent le duplex d'ADN à cliver (en noir). Les deux sous-unités B sont liées aux sous-unités A par le domaine « *transducer* » (en vert). Le deuxième duplex d'ADN entre dans l'enzyme par les deux domaines de fixation de l'ATP des deux sous-unités B, domaine GHKL (en mauve). La fixation de l'ATP ferme cette porte par la dimérisation de ces deux domaines GHKL. Les sous-unités A clivent le duplex noir, et passent le duplex rouge à travers la coupure après l'hydrolyse de l'ATP. La coupure de l'ADN, sur le duplex noir, est ensuite refermée. **(B)** Le complexe catalytique de la recombinaison méiotique (TOPOVIL) est constitué de deux sous-unités SPO11 similaires aux sous-unités A (Orange) et de deux sous-unités B appelées, pour la souris, TOPOVIBL (En mauve). SPO11 interagit et clive le duplexe de l'ADN, il reste lié d'une manière covalente aux extrémités et il ne referme pas la coupure d'ADN. **(C)** Étant donné que le domaine de fixation de l'ATP de TOPOVIBL est faiblement conservé, la formation des DSB par TOPOVI est, soit indépendante de la fixation de l'ATP et de la dimérisation de TOPOVIBL, soit dépendante la dimérisation de TOPOVIBL et implique la fixation de l'ATP ou /et la présence de facteurs protéiques stabilisateurs (Robert et al., 2017).

Chez la levure *S. cerevisiae*, l'unité catalytique, Spo11, interagit avec Ski8, Rec102 et Rec104 pour former un complexe présentant des similitudes structurales avec le complexe TopoVI. La sous-unité B de ce complexe serait formée par l'association entre Rec102 et Rec104. Rec102 formerait le domaine transducteur et Rec104 occuperait une position similaire au domaine ATPase du type GHK. Cependant, Rec104 semble être dépourvue du motif nécessaire à la liaison et l'hydrolyse de l'ATP (Claeys Bouuaert et al., 2021b). Ski8 est une protéine contenant des répétitions de motifs WD (pour tryptophane-acide aspartique), elle est connue pour avoir des fonctions cytoplasmiques dans les cellules végétatives et elle est membre du complexe Ski impliqué dans le métabolisme des ARNm (Anderson and Parker, 1998; Brown et al., 2000). Sa relocalisation dans le noyau lui permet d'exercer son activité méiotique qui reste mal définie. Néanmoins, son interaction avec Spo11 paraît essentielle pour la formation des DSB, ainsi que la stabilité du complexe catalytique (Arora et al., 2004; Claeys Bouuaert et al., 2021b). WDR61 est l'orthologue de Ski8 chez les mammifères, il fait aussi partie du complexe Ski et du complexe PAF (impliqué dans l'initiation et l'élongation de la transcription) (Zhu et al., 2005). Par contre, on ignore si WDR61 joue également un rôle dans la formation de DSB méiotique chez les mammifères (Baudat et al., 2013).

Bien qu'il soit nécessaire d'avoir deux sous-unités de Spo11 pour assurer un clivage double brin de l'ADN et malgré le fait que TopoVI d'archée forme un hétérotétramère A<sub>2</sub>B<sub>2</sub>, le complexe Spo11-Rec102-Rec104-Ski8 purifié, *in vitro*, est monomérique et il a une stœchiométrie (1 : 1 : 1 : 1). La dimérisation de ce complexe, *in vivo*, pourrait être un mécanisme de contrôle de la formation des DSB (Claeys Bouuaert et al., 2021b; Yadav and Claeys Bouuaert, 2021).

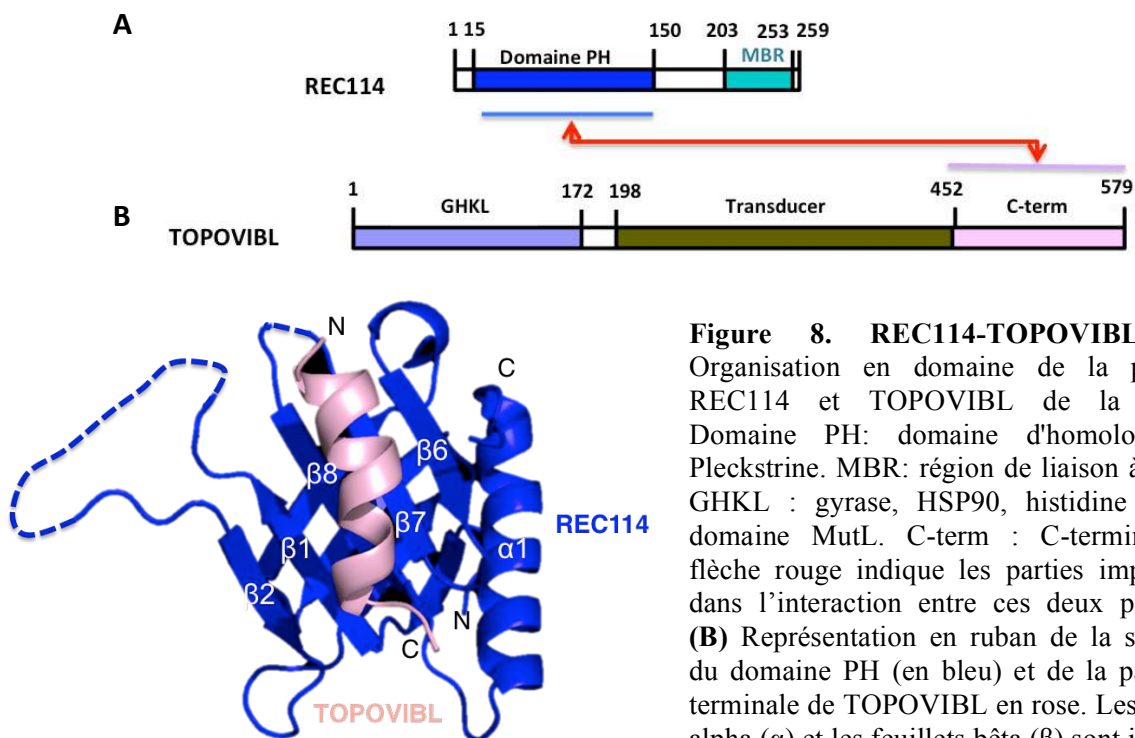
Chez *Arabidopsis thaliana*, il existe trois homologues de SPO11 (AtSPO11-1, AtSPO11-2 et AtSPO11-3) (Hartung and Puchta, 2001, 2000). Néanmoins, seules les protéines AtSPO11-1 et AtSPO11-2 sont essentielles pour l'initiation de la recombinaison méiotique (Stacey et al., 2006). La sous-unité TOPOVIBL de *A. thaliana* (MTOPOVIB) interagit avec les deux orthologues SPO11 (SPO11-1 et SPO11-2) et est nécessaire à la formation de l'hétérodimère SPO11-1/SPO11-2 (Vrielynck et al., 2016).

### 1.4.1.2 REC114-MEI4-IHO1

REC114, MEI4 et IHO1 forment un sous-groupe de protéines promotrices de DSB, initialement décrit chez *S. cerevisiae* (Rec114-Mei4-Mer2 (RMM)) (Arora et al., 2004; Li et al., 2006; Maleki et al., 2007). La séquence de ces protéines a fortement divergé entre les eucaryotes, mais les homologues de ces protéines sont actuellement connus chez de nombreuses espèces (Kumar et al., 2010; Tessé et al., 2017). Les protéines MEI4, REC114 et IHO1 sont importantes pour l'initiation de la recombinaison méiotique. Elles montrent une colocalisation importante sous forme de foci sur les axes chromosomiques. Cette localisation est observée au stade de leptotène méiotique, moment de la formation de DSB méiotique, puis diminue progressivement avec l'appariement des chromosomes (Kumar et al., 2018, 2010; Stanzione et al., 2016). De plus, les souris délétées des gènes codant ces protéines présentent des défauts dans la formation de DSB, la réparation des DSB et le *synapsis* des chromosomes, ce qui induit l'arrêt de la prophase méiotique (Kumar et al., 2018, 2010; Stanzione et al., 2016). Bien que l'existence d'un complexe ternaire entre ces protéines n'ait jamais été confirmée, les analyses par double hybride ainsi que par co-immunoprécipitation suggèrent que ces trois protéines pourraient interagir et former un complexe. L'assemblage de ce dernier sur les axes chromosome avant la formation des DSB promeut l'activité du complexe catalytique et favorise la formation des DSB. Il a été proposé que ce soit IHO1 qui servirait de plateforme pour le recrutement de REC114-MEI4 sur les chromosomes (Kumar et al., 2018; Stanzione et al., 2016).

L'interaction entre REC114 et MEI4, initialement observée par la technique de double hybride (Arora et al., 2004; Kumar et al., 2010; Maleki et al., 2007), a été caractérisée récemment, *in vitro*, chez la souris et la levure (Claeys Bouuaert et al., 2021a; Kumar et al., 2018). La formation d'un complexe stable entre ces deux protéines implique le domaine C-terminal de REC114 et le domaine N-terminal de MEI4 (Claeys Bouuaert et al., 2021a; Kumar et al., 2018). La caractérisation biochimique récente de ce complexe chez la levure a montré que le domaine C-terminal de REC114 se dimérise et son interaction avec MEI4 produit un complexe d'une stœchiométrie 2:1 (Claeys Bouuaert et al., 2021a).

En outre, l'étude structurale de la partie N-terminale de REC114 de la souris a montré la présence d'un domaine PH (pour Pleckstrin Homology), composé de deux ensembles de feuillets  $\beta$  anti-parallèles, disposés perpendiculairement, suivis d'une hélice  $\alpha$  (Kumar et al., 2018). Ce domaine est supposé être impliqué dans les interactions protéine/protéine, ainsi REC114 pourrait être une plateforme pour plusieurs interactions (Kumar et al., 2018). En effet, il a été montré récemment que le domaine PH de REC114 murin interagit avec le peptide C-terminal conservé de TOPOVIBL (du complexe catalytique) et la structure de ce complexe a été résolue par cristallographie (Figure 8 A, B) (Nore et al., 2022). L'effet exact de cette interaction sur l'organisation du complexe catalytique SPO11-TOPOVIBL reste inconnu, mais la dimérisation de REC114, observée pour *S. cerevisiae*, via son domaine C-terminal pourrait stabiliser ou induire la dimérisation du complexe catalytique (Claeys Bouuaert et al., 2021a; Nore et al., 2022). La perte de l'interaction entre REC114 et TOPOVIBL provoque des défauts de formation de DSB méiotiques chez souris mâle et femelle. Contrairement aux ovocytes qui présentent une réduction considérable de la formation des DSB le long du génome, la réduction des DSB n'a été observée qu'aux niveaux des régions subtélomériques pour les spermatocytes. L'activité des DSB a été retardée sur le reste des autosomes mâles (Nore et al., 2022). Ce résultat montre, qu'en fonction des caractéristiques chromosomiques spécifiques, l'interaction entre REC114/TOPOVIBL induit soit la formation catalytique des DSB ou la régulation temporelle de leur formation (Nore et al., 2022).



**Figure 8. REC114-TOPOVIBL.** (A) Organisation en domaine de la protéine REC114 et TOPOVIBL de la souris. Domaine PH: domaine d'homologie de Pleckstrine. MBR: région de liaison à MEI4, GHKL : gyrase, HSP90, histidine kinase, domaine MutL. C-term : C-terminal. La flèche rouge indique les parties impliquées dans l'interaction entre ces deux protéines (B) Représentation en ruban de la structure du domaine PH (en bleu) et de la partie C-terminale de TOPOVIBL en rose. Les hélices alpha ( $\alpha$ ) et les feuillets bêta ( $\beta$ ) sont indiqués (Nore et al., 2022).



Chez la levure, la protéine IHO1, Mer2, adopte une forme tétramérique en solution (Claeys Bouuaert et al., 2021a; Rousová et al., 2021). Sa partie centrale, prédite pour être un *coiled-coil*, formerait deux paires d'hélices parallèles disposées dans une configuration antiparallèle. Il a été également montré que la partie C-terminale de Mer2 et celle de REC114 sont impliquées dans la liaison à l'ADN (Claeys Bouuaert et al., 2021a). La même étude a révélé que Rec114-Mei4 et Mer2 coopèrent fortement, *in vitro*, pour lier l'ADN et former de très grands assemblages nucléoprotéiques, appelés condensat. Ces derniers sont réversibles, peuvent fusionner et sont assemblés via des interactions électrostatiques ADN/protéines et des interactions protéines/protéines. Ces assemblages nucléoprotéiques ressemblent au système de séparation de phase, qui permet de rassembler les biomolécules dédiées à un processus biologique particulier sous forme de larges compartiments intracellulaires (Li et al., 2012). Ces condensats de protéines RMM observés, *in vitro*, pourraient correspondre au foci de protéines RMM visibles *in vivo*. En effet, les mutations qui perturbent la formation de condensat réduisent le nombre de foci protéiques et perturbent également la formation des DSB. Ce qui suggère que la condensation des protéines RMM, induite par l'ADN, est importante pour l'initiation de la recombinaison. Il a été proposé que ces condensats servent de plateforme pour le recrutement du complexe catalytique SPO11 ainsi que d'autres protéines nécessaires à la formation et la régulation de DSB (Claeys Bouuaert et al., 2021a). Le recrutement de SPO11 pourrait être assuré grâce à l'interaction entre le domaine N-terminal de Rec114 et les sous-unités Rec102-Rec104 du complexe catalytique, équivalent de la TOPOVIBL (Claeys Bouuaert et al., 2021a). Il est intéressant de noter qu'*in vitro*, la protéine IHO1 de souris subit également une séparation de phase, ce qui suggère la conservation de ce mécanisme (Tsai et al., 2020).

### ***1.4.1.3 Autres protéines nécessaires à la formation de DSB***

#### ***HORMAD 1 et HORMAD 2***

La protéine HORMAD1 et HORMAD2 sont des protéines méiotiques appartenant à la famille des protéines à domaine HORMA (pour Hop1, Rev7 and MAD2, trois familles de protéines initialement identifiées pour partager ce domaine), qui sont impliquées dans la réparation de l'ADN en agissant comme adaptateurs pour recruter d'autres protéines sur la chromatine (Aravind and Koonin, 1998; Wojtasz et al., 2009).



Ces deux protéines sont préférentiellement associées à l'axe chromosomique au début de la prophase, puis elles y sont délocalisées au fur et à mesure de l'appariement des axes chromosomes par la mise en place du complexe synaptonémal (Fukuda et al., 2010; Wojtasz et al., 2009). La protéine HORMAD1 co-localise, et interagirait directement ou indirectement, avec les cohésines méiotiques (RAD21L et REC8) et les protéines de l'axe chromosomique (SYCP2 et SYCP3) qui assurent respectivement son recrutement et sa stabilité (Fujiwara et al., 2020). Cette protéine est également essentielle pour la gamétogenèse des mammifères. En effet, les souris déficientes en HORMAD1 sont stériles, les niveaux de DSB sont réduits et elles présentent des défauts dans la formation de complexe synaptonémal et l'alignement des chromosomes homologues (Daniel et al., 2011; Shin et al., 2010).

HORMAD1 est également nécessaire à la stabilisation de MEI4 sur les axes chromosomiques dans les spermatocytes de souris (Kumar et al., 2015). En plus, les analyses de double hybride, de l'immunoprécipitation et de la co-localisation ont suggéré qu'IHO1 interagit avec HORMAD1, ce qui permettrait le recrutement et/ou la stabilisation des protéines REC114-MEI4-IHO1 sur les axes chromosomiques (Stanzione et al., 2016). Ce rôle de HORMAD1 semble être conservé. Hop1, l'orthologue de *S. cerevisiae*, est aussi une protéine d'axes chromosomes et assurerait le recrutement des protéines RMM (Panizza et al., 2011). Une étude récente a montré, *in vitro*, la présence d'une interaction entre Mer2 (orthologue d'IHO1) et Hop1 (Rousová et al., 2021).

La protéine HORMAD2 ne semble pas être indispensable pour la formation des DSB ou l'appariement des chromosomes. En effet, les souris mutantes pour cette protéine ne présentent qu'une faible réduction des taux de DSB et des *synapsis* (Kogo et al., 2012; Wojtasz et al., 2012). En revanche, l'absence de cette protéine provoque la stérilité chez les mâles. Il paraît que HORMAD2 est impliquée dans le mécanisme d'inactivation des gènes situés sur les régions des chromosomes sexuels XY qui ne s'apparient pas pendant la prophase I (Kogo et al., 2012; Wojtasz et al., 2012). Ce mécanisme est appelé MSCI (pour méiotique sex chromosome inactivation) et son échec provoque l'arrêt de la méiose et la mort des spermatozoïdes (Royo et al., 2010). Chez les femelles, HORMAD2 coopère avec HORMAD1 et d'autres partenaires pour la mise en place du MSUC (pour meiotic silencing of unsynapsed chromatin) (Daniel et al., 2011; Kogo et al., 2012), un mécanisme de

contrôle permettant d'inactiver la transcription génique dans les régions chromosomiques qui présentent un défaut de *synapsis* (Schimenti, 2005).

### ***Ankyrin repeat domain-containing 31 (ANKRD31)***

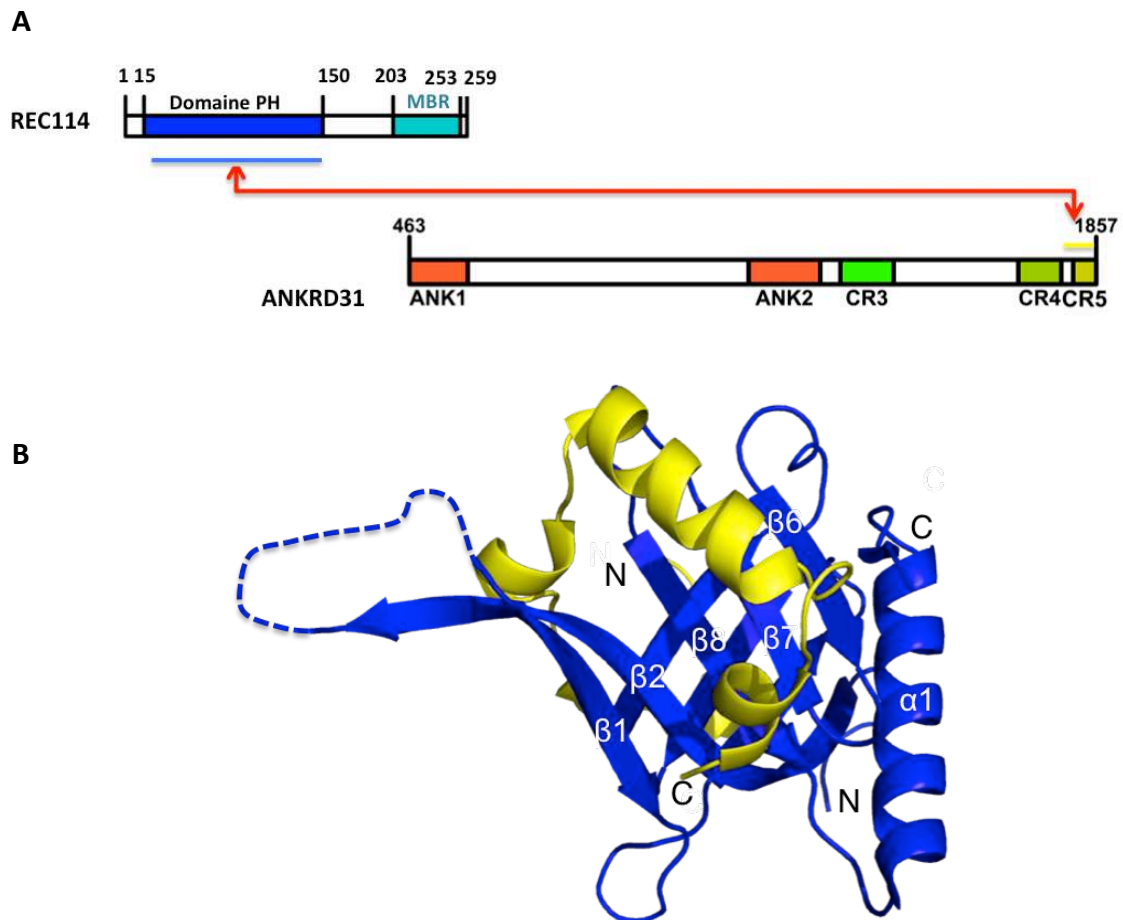
ANKRD31 est une protéine méiotique qui présente la même localisation cytologique que REC114 et MEI4, IHO1 et MEI1 (Acquaviva et al., 2020; Boekhout et al., 2019; Papanikos et al., 2019). Cette protéine joue un rôle majeur dans le contrôle temporel et spatial des DSB méiotiques. En effet, les souris déficientes en cette protéine présentent un retard dans la formation des DSB, un changement dans leurs distributions, ainsi qu'un retard dans l'assemblage du complexe synaptonémal. Ces conséquences sont observées sur les autosomes des deux sexes, mais elles sont plus prononcées dans les régions pseudo-autosomiques des chromosomes sexuels chez les mâles. En effet, l'absence de ANRKD31 provoque la stérilité des souris mâles (Boekhout et al., 2019; Papanikos et al., 2019).

La protéine ANKRD31 contient deux domaines à répétitions ankyrine, connus pour servir de plateforme et assurer des interactions protéines-protéines (Mosavi et al., 2004). C'est une protéine spécifique des vertébrés et sa séquence montre la présence de cinq domaines conservés, dont trois dans la partie C-terminale (Figure 9 A) (Papanikos et al., 2019). Le dernier domaine de la partie C-terminale de ANKRD31 interagit, *in vitro*, avec le PH domaine de REC114 et la structure cristalline de ce complexe a été résolue (Figure 9 B) (Boekhout et al., 2019). Cette interaction serait mutuellement exclusive avec TOPOVIBL qui interagit au même endroit du domaine PH (Nore et al., 2022). L'interaction entre REC114 et ANKRD31 permettrait, *in vivo*, le recrutement de REC114 et d'autres facteurs des DSB sur l'axe chromosomique, en particulier dans les régions pseudo-autosomiques, dans lesquelles l'accumulation des protéines nécessaires à la formation des DSB semble être indépendante de IHO1. Cette protéine méiotique est impliquée dans la localisation, le timing et le nombre de DSB à travers le génome (Boekhout et al., 2019; Papanikos et al., 2019).

### ***MEI1 (meiosis defective 1)***

MEI1 est une protéine essentielle pour la formation du DSB, mais sa fonction reste inconnue (Lam and Keeney, 2015). Les spermatocytes mutants pour *Mei1* présentent de nombreux phénotypes similaires à l'absence de SPO11, indiquant des défauts de formation

de DSB et de l'appariement chromosomique (Libby et al., 2003, 2002). MEI1 joue également un rôle clé dans la localisation de MEI4 au niveau de l'axe chromosomique (Kumar et al., 2015). Elle colocalise avec les autres protéines nécessaires à la formation de DSB, en l'occurrence REC114, MEI4, IHO1 et ANKRD31 (Acquaviva et al., 2020).

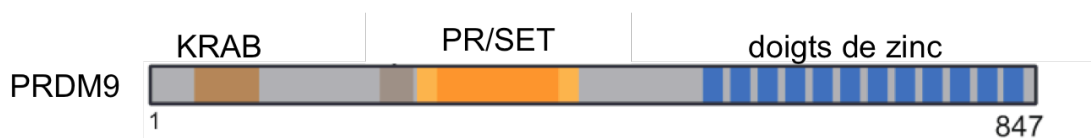


**Figure 9. REC114-ANKRD31.** (A) Organisation en domaine de la protéine REC114 et de la partie 457 à 1857 d'ANKRD31 de la souris. Domaine PH : domaine d'homologie de Pleckstrine, MBR: région de liaison à MEI4, ANK : répétition ankyrine. CR : région conservée. La flèche rouge indique les régions impliquées dans l'interaction entre ces deux protéines. (B) Représentation en ruban de la structure du domaine PH de REC114 (en bleu) et de la partie C-terminale de ANKRD31 en jaune (PDB: 6NXF). Les hélices alpha ( $\alpha$ ) et les feuillets bêta ( $\beta$ ) sont indiqués.

## 1.5 Localisation des DSB

La distribution spatiale des cassures double brin lors de la méiose n'est pas aléatoire, elle est fortement régulée afin d'éviter des réarrangements ou de l'instabilité du génome (Baudat et al., 2013). En effet, la formation des DSB dans des régions très riches en séquence répétitive peut conduire à leurs réparations en utilisant des séquences non-alléliques conduisant à des recombinaisons non alléliques (Sasaki et al., 2010). De plus, la formation des CO dans des régions proches du centromère peut conduire à une séparation des chromatides sœur lors de la méiose I (Lam and Keeney, 2015). Pour éviter ces conséquences néfastes pour la cellule, il existe des emplacements préférentiels pour la formation des DSB, appelés points chauds, et d'autres emplacements qui sont rarement la cible d'événements de recombinaison, appelés points froids (De Massy, 2013; Kauppi et al., 2013; Lam and Keeney, 2015). Chez la plupart des mammifères, les DSB se produisent dans les régions enrichies en tri-méthylations des lysines 4 et 36 de l'histone H3 (H3K4me3 et H3K36me3). Ces méthylations, associées aux points chauds, sont indépendantes de celles fréquemment situées aux niveaux des promoteurs pour activer la transcription des gènes (Buard et al., 2009; Powers et al., 2016; Smagulova et al., 2011).

Un déterminant clé de la répartition des DSB, chez la plupart des mammifères, est la protéine PRDM9 (Baudat et al., 2010; Brick et al., 2012; Diagouraga et al., 2018; Myers et al., 2010; Parvanov et al., 2010). Prdm9 est exprimé spécifiquement dans les méiocytes lors de la prophase I et code pour une protéine présentant principalement un domaine KRAB (Krueppel-associated box) dans la région N-terminale, un domaine PR/SET avec une activité de (H3K4me3 et H36K4me3), et un domaine avec une série de doigts de zinc, permettant à la protéine de fixer des régions spécifiques d'ADN (Figure 10) (Grey et al., 2018).



**Figure 10. Domaine de PRDM9.** Représentation schématique de la protéine PRDM9 de souris montrant le domaine KRAB (Krueppel-associated box), le domaine à activité méthyltransférase (PR / SET) et le domaine à doigts de zinc (Diagouraga et al., 2018).

Les doigts de zinc de PRDM9 sont hautement polymorphes et se caractérisent par une évolution rapide (Oliver et al., 2009). Cette caractéristique permet à la protéine de changer fréquemment sa séquence de la liaison à l'ADN (Buard et al., 2014; Kono et al., 2014; Oliver et al., 2009). Ce changement évolutif rapide peut surmonter le phénomène dit d'érosion de points chauds. Ce phénomène de perte des sites de liaison à PRDM9 peut se produire au cours de l'évolution quand le chromosome utilisé pour la réparation de l'ADN par la recombinaison homologue contient une version allélique de faible affinité à PRDM9. Ce qui induit un changement dans les motifs reconnus par PRDM9 et la transmission de l'allèle qui présente une faible affinité à PRDM9 (Grey et al., 2018). L'évolution rapide de la séquence de liaison à l'ADN de PRDM9 lui permettrait de reconnaître de nouveaux sites d'ADN, ce qui change la distribution des points chauds sur le génome (Baudat et al., 2013).

Pour la formation des DSB, il a été proposé que le réseau de doigts de zinc reconnaisse une séquence spécifique de l'ADN. Une fois le motif d'ADN reconnu, le domaine SET assure les modifications post-traductionnelles (H3K4me3, H3k36me3) du nucléosome et la protéine recrute les facteurs nécessaires à la formation de DSB au voisinage de son site de liaison (Baudat et al., 2010). De plus, par un mécanisme inconnu, la présence de PRDM9 empêche la formation des DSB dans d'autres régions génomiques riches en H3K4me3 telles que les promoteurs de transcription. En effet, en absence de PRDM9, la plupart des recombinaisons sont initiées au niveau du promoteur et dans d'autres régions régulatrices, riches en H3K4me3 (Brick et al., 2012). Les régions pseudo-autosomiques des chromosomes X et Y de souris font exception à ce mode de localisation, car une activité élevée de formation des DSB se produit dans ces régions, indépendamment de PRDM9 (Brick et al., 2012).

Le mécanisme moléculaire exact du fonctionnement de PRMD9 n'est pas encore bien compris (Grey et al., 2018). La formation des DSB sur les boucles chromatiniennes après leur recrutement sur les axes chromosomiques, proposée par le modèle TLAC, suggère la présence d'un facteur qui va reconnaître ces marques épigénétiques et amener les sites liés à PRDM9 à l'axe chromosomique contenant les protéines initiatrices de DSB (Figure 11A) (Grey et al., 2018). Le premier candidat proposé pour ce rôle, était la protéine CXXC1 ou Cfp1, membre du complexe COMPASS (complexe associé à la protéine SET1, protéine nécessaire pour la tri-méthylation de H3K4) (Lee and Skalnik, 2005). Des analyses par double hybride ont montré que la protéine CXXC1 interagit d'une part avec PRDM9 au

niveau du domaine KRAB et d'autre part avec la protéine de l'axe chromosomique IHO1 (Imai et al., 2017; Parvanov et al., 2017). Cela suggère que l'interaction de CXXC1 avec PRDM9 et IHO1 permettrait le recrutement des sites liés à PRDM9 à la machinerie promotrice de DSB, située au niveau de l'axe chromosomiques. Le recrutement ou l'activation de l'unité catalytique, SPO11, assurerait le clivage de l'ADN au niveau de ces sites (Figure 11 B) (Grey et al., 2018; Imai et al., 2017; Parvanov et al., 2017).

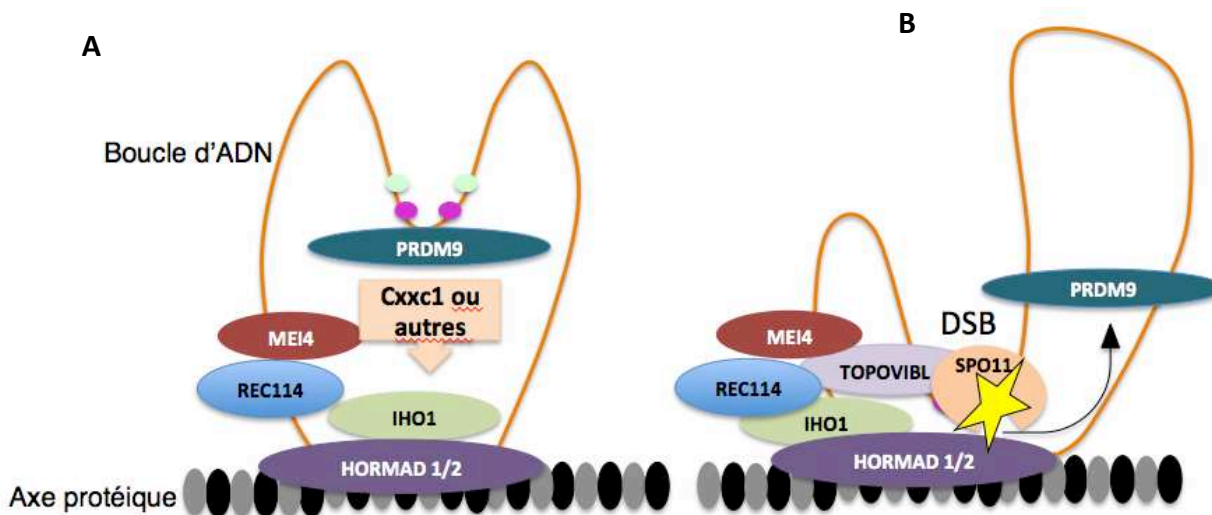
Ce mécanisme de reconnaissance de marque H3K4me3 semble intéressant, d'autant plus qu'il est similaire avec ce qui a été observé chez *S. cerevisiae*. En effet, cette espèce ne possède pas d'homologue de PRMD9, et ses DSB méiotiques se produisant majoritairement dans les régions promotrices (Pan et al., 2011). Les marques H3K4me3 de ces régions sont reconnues par la protéine Spp1, orthologue de CXXC1 et membre du complexe COMPASS (He et al., 2019; Murton et al., 2010). L'interaction de Spp1 avec Mer2 (IHO1 des levures) permettrait le recrutement des boucles d'ADN vers les axes chromosomiques (Acquaviva et al., 2013; Rousová et al., 2021; Sommermeyer et al., 2013).

Cependant, et contrairement à la levure, CXXC1 ne paraît pas être indispensable pour moduler l'activité catalytique de PRDM9 dans l'initiation de la recombinaison méiotique. L'absence de cette protéine chez les souris n'affecte pas la triméthylation des points chauds par PRDM9, la formation ou la réparation des DSB méiotiques (Tian et al., 2018).

Récemment, la protéine EWSR1 a été proposée pour favoriser la méthylation des histones par PRDM9 et de participer à la liaison des sites de DSB aux axes chromosomiques grâce à son interaction avec PRDM9 et la cohésines des axes méiotiques phosphorylée REC8 (Tian et al., 2021). D'autres protéines ont été également proposées comme partenaire de PRDM9, parce que leur évolution (c'est-à-dire leur présence ou leur absence chez les vertébrés) est corrélée à celle de PRDM9. Parmi ces facteurs, on peut citer : ZCWPW1, dont un rôle dans l'aide de réparation de DSB induite par PRDM9 lui a été déjà attribué (Huang et al., 2020; Mahgoub et al., 2020; Wells et al., 2020), son paralogue ZCWPW2, TEX15 et FBX047 (Cavassim et al., 2022).

En plus de localiser le site de DSB, PRDM9 agit aussi sur la structure de la chromatine pour créer des régions plus ouvertes et pauvres en nucléosome appelé NDR

(pour nucleosome-depleted region). Cette réorganisation favorise l'accessibilité des points chauds à la machinerie de DSB et elle est assurée en partie par le recrutement des remodeleurs de la chromatine, notamment HELLS (Baker et al., 2014; Spruce et al., 2020; Yamada et al., 2020). L'action de PRDM9 sur la chromatine est également observée sur le chromosome homologue intact, générant ainsi une structure chromatinienne plus accessible à l'endroit de l'engagement des extrémités simple brin pour la recherche d'homologie (Yamada et al., 2020). Cela est cohérent avec les observations antérieures montrant que la fixation de PRMD9, d'une manière symétrique, sur les chromosomes homologues augmente les chances de réparation des DSB générées à ces endroits par la formation d'un CO (Davies et al., 2016; Hinch et al., 2019).



**Figure 11. Modèle du mode d'action de PRDM9.** (A) La localisation des DSB sur le génome est déterminée, chez la plupart des mammifères, par la protéine PRDM9. Cette dernière se lie, via son domaine à doigt de zinc, à des motifs d'ADN situés dans des boucles de chromatine (une seule chromatide est représentée ici). PRDM9 catalyse la H3K4me3 (points verts) et H3K36me3 (points roses) du nucléosome à proximité de son site de fixation. Ces sites de méthylation correspondent aux points chauds de la formation des DSB. Les protéines promotrices de la formation des DSB telles que HORMAD1, MEI4, REC114 et IHO1, sont situées sur l'axe chromosomique. Selon le modèle de recrutement des boucles d'ADN sur les axes chromosomiques, il doit y avoir un facteur intermédiaire qui assurerait l'interaction de PRDM9 avec la machinerie de formation des DSB sur l'axe chromosomique. Ce facteur peut être CxxC1 qui interagirait avec PRDM9 et IHO1, ou EWSR1 qui est supposé interagir avec PRDM9 et les cohésines de l'axe chromosomique. (B) Après la connexion des boucles d'ADN sur l'axe chromosomique, le complexe catalytique SPO11-TOPOVIBL est recruté ou activé et il assure le clivage de l'ADN autour des sites de liaison de PRDM9 (Adaptée de (Grey et al., 2018)).

## 1.6 Régulation de la formation des DSB

Bien qu'elles soient requises pour l'initiation de la recombinaison méiotique et la ségrégation précise des chromosomes, les cassures d'ADN double brin restent des lésions chromosomiques graves qui, si elles ne sont pas bien réparées, peuvent être létales pour la cellule méiotique ou conduire à plusieurs phénotypes pathologiques chez la descendance. Pour minimiser ces risques délétères, la formation des DSB est étroitement contrôlée afin qu'elles ne se produisent qu'au bon moment et au bon endroit.

### 1.6.1 Régulation temporelle de la formation des DSB

La formation de DSB est limitée à une fenêtre de temps assez étroite pendant la prophase I. Elles doivent se produire après la réplication de l'ADN et avant la ségrégation des chromosomes (Keeney et al., 2014). L'un des mécanismes de contrôle du moment de la formation des DSB consiste à réguler l'expression de la machinerie catalytique. Chez *S. cerevisiae*, les gènes codants les protéines Spo11, Rec102, Rec104, Rec114 et Mei4 ne sont transcrits qu'au moment de la méiose. (Murakami and Keeney, 2008). Un autre niveau de contrôle comprend la coordination entre la réplication et la formation de DSB (Borde et al., 2000). La séparation temporelle de la réplication et de la formation du DSB est importante, car elle garantit que les échanges réciproques se forment uniquement entre les chromosomes dupliqués. Cette coordination semble être assurée par les points de contrôle de la réplication, ainsi que l'action des kinases du cycle cellulaire (Figure 12A, B) (Hunter, 2015).

Les points de contrôle de réplication chez *S. cerevisiae* empêchent le clivage de l'ADN en cas d'un retard de réplication du génome (stress réplcatif). Cela est assuré par l'inhibition de l'expression, de l'activation et/ou de l'accumulation de certains facteurs protéiques essentiels à la formation des DSB (Blitzblau and Hochwagen, 2013). En outre, la progression de la réplication à la formation du DSB est supposée être fortement liée à l'activité des kinases régulatrices du cycle cellulaire CDK-S (S-phase cyclin-dependent kinase) et DDK (Dbf4-dependent kinase) (Murakami and Keeney, 2008). En effet, la réplication de l'ADN est déclenchée à des faibles niveaux d'activité de ces kinases et c'est quand leurs activités atteignent des niveaux plus élevés que la formation des DSB commence (Murakami and Keeney, 2008). De plus, ces deux kinases assurent la phosphorylation de Mer2, protéine cruciale pour la formation des DSB chez *S. cerevisiae*



(l'orthologue de IHO1) (Henderson et al., 2006; Sasanuma et al., 2007; Wan et al., 2008). Mer2 se lie à la chromatine et elle est phosphorylée par CDK sur quelques sites uniquement. Ensuite, la kinase DDK recrutée sur le replisome termine la phosphorylation de Mer2 sur plusieurs sites et elle permet une phosphorylation préférentielle de la protéine Mer2 liée à la chromatine répliquée (Murakami and Keeney, 2014). La phosphorylation de Mer2 favoriserait ensuite le recrutement de Rec114 et d'autres protéines initiateuses des DSB (Henderson et al., 2006; Panizza et al., 2011; Sasanuma et al., 2007). Les connaissances actuelles ne permettent pas de confirmer la conservation de ce mécanisme ainsi que la phosphorylation de IHO1 lors de la formation des DSB chez les mammifères.

## 1.6.2 Contrôle du nombre et de la distribution des DSB

En plus de déterminer la localisation des différents points chauds de la formation de DSB, les cellules méiotiques disposent de plusieurs niveaux de contrôle pour déterminer les points chauds qui vont acquérir des cassures doubles brins parmi tous les points chauds méthylés. Ainsi, elles modulent le nombre et la distribution des DSB pour former uniquement le nombre de DSB nécessaires au bon déroulement de la recombinaison méiotique et l'appariement des chromosomes, tout en évitant des lésions excessives qui peuvent être toxiques pour les cellules.

Chez les mammifères, comme chez la plupart des espèces, l'un des facteurs jouant un rôle central dans la régulation de la formation de DSB méiotique est l'ATM (ataxia telangiectasia mutated) (Keeney et al., 2014; Lukaszewicz et al., 2018). Une kinase Ser / Thr appartenant à la famille des protéines phosphatidylinositol-3-kinase-like (PIKK). L'activation de cette kinase déclenche la signalisation cellulaire nécessaire pour la réparation de l'ADN dans les cellules somatiques (Derheimer and Kastan, 2010; Shiloh and Ziv, 2013). Cette kinase est également importante lors de la méiose, les méiocytes de souris déficientes en ATM présentent des défauts dans l'appariement des chromosomes, des fragmentations des chromosomes et elles sont éliminées par apoptose au stade zygotène/pachytène (Barlow et al., 1998; Xu et al., 1996). L'absence d'ATM induit une augmentation significative du nombre des DSB qui se forment en dehors des points chauds déterminés par PRDM9 (Lange et al., 2011; Paiano et al., 2020). C'est pourquoi, il a été suggéré que la protéine ATM jouerait un rôle important lors de la méiose et elle assurerait la régulation quantitative des DSB en réprimant leur formation répétée dans la même section chromosomique. Cette activité d'ATM apparaît plus importante aux niveaux des

régions non-homologues des chromosomes sexuels (Lange et al., 2016). Le rôle de cette kinase semble être conservé, son orthologue Tell1 de *S. cerevisiae* empêche la formation des DSB rapprochées au sein de la même chromatide (appelée interférence en cis) mais également sur les mêmes loci entre les chromatides sœurs ou les chromosomes homologues (interférence en trans) (Garcia et al., 2015; Zhang et al., 2011).

La façon dont ATM régule négativement la formation DSB méiotique n'est pas encore comprise. La présence des lésions double brin d'ADN active la kinase pour phosphoryler les différentes cibles nécessaires à la réparation de l'ADN (Figure 12 C) (Keeney et al., 2014). Ces cibles ne sont pas encore connues dans la cellule méiotique, mais REC114 et HORMAD1 ont été proposées comme des candidats potentiels (Lukaszewicz et al., 2018). Chez *S. cerevisiae* ces deux protéines sont des cibles connues de Tell1, orthologue d'ATM (Carballo et al., 2013, 2008). De plus, la phosphorylation de REC114 agit négativement sur la formation des DSB (Carballo et al., 2013). Par conséquent, la phosphorylation de ces protéines, ainsi que d'autres cibles ATM, a été suggérée pour empêcher la formation de DSB supplémentaires (Figure 12 D). Cette inhibition peut être localisée en déstabilisant le complexe IHO1-REC114-MEI4 au voisinage immédiat des extrémités coupées. Elle peut également se propager aux boucles de chromatine voisines en raison de la phosphorylation de HORMAD1 et de l'inhibition de l'assemblage du nouveau complexe IHO1-REC114-MEI4 (Lukaszewicz et al., 2018).

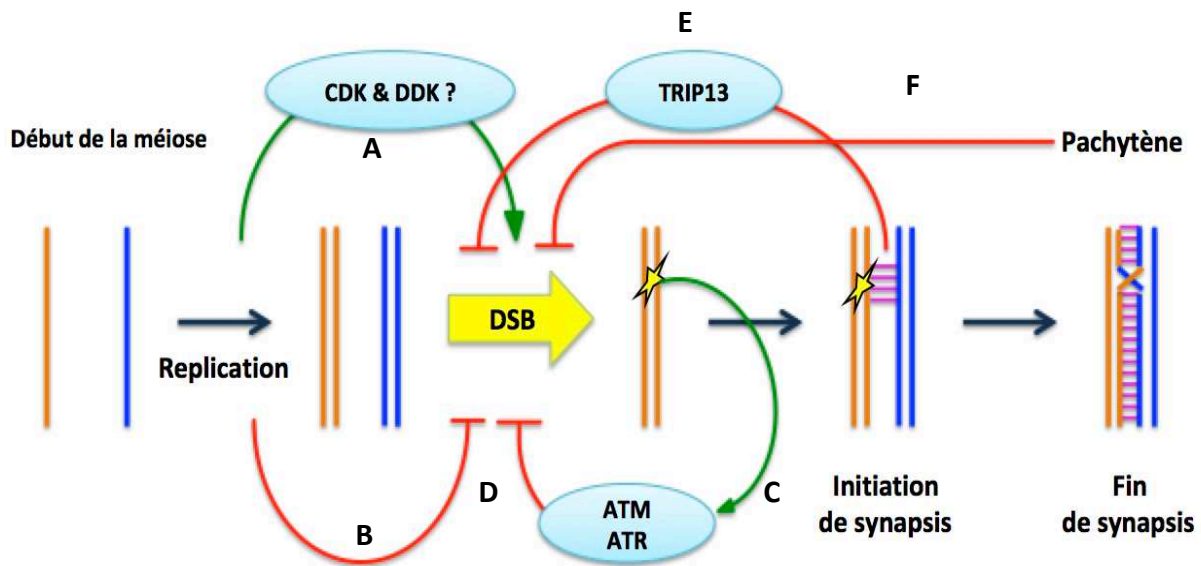
L'activation d'ATM a été également proposée pour promouvoir l'étape de la résection (Paiano et al., 2020; Yamada et al., 2020). De plus, il a été suggéré que les extrémités simple brin générées lors de la résection active l'ATR (pour ataxia telangiectasia and Rad3-related protein), un autre membre de la famille PKK (Figure 12 C) (Dereli et al., 2021; Shiotani and Zou, 2009). Cette kinase favorise l'élimination de IHO1 sur les axes chromosomiques proches de la DSB. Le mécanisme exact de cette voie de régulation reste inconnu, mais elle permettrait de prévenir la formation de nouvelle DSB à proximité des régions clivées et d'assurer leur distribution d'une manière uniforme le long du génome (Figure 12 D) (Dereli et al., 2021).

Le phénomène de condensat observé pour les protéines RMM de levure pourrait par ailleurs jouer un rôle dans la répartition des DSB. En effet, l'assemblage des protéines Rec114, Mei4 et Mer2 avec l'ADN sous forme de condensats pour recruter Spo11 limite

leurs disponibilités locales, ce qui réduit la probabilité de la formation de nouveaux site de DSB à proximité (Claeys Bouuaert et al., 2021a).

Un autre niveau de régulation du nombre de DSB reposerait sur la formation du complexe synaptonémal et l'appariement des chromosomes homologues. Cette voie de régulation éliminerait la machinerie de DSB dans les régions chromosomiques déjà appariées. Elle serait liée en partie à la protéine HORMAD1 (Daniel et al., 2011; Stanzione et al., 2016; Wojtasz et al., 2012). Cette dernière favorise la formation des DSB en jouant un rôle de plateforme sur l'axe chromosomique pour améliorer le recrutement des protéines promotrices des DSB, probablement à travers l'interaction avec IHO1 (Stanzione et al., 2016). Cependant, la formation du complexe synaptonémal enlève HORMAD1 des axes chromosomiques grâce à l'activité de l'ATPase TRIP13 (Wojtasz et al., 2009). L'élimination de HORMAD1 permettrait de réduire l'accumulation des protéines REC114, MEI4 et IHO1 dans les régions déjà associées par le complexe synaptonémal et limiterait l'activité de SPO11 aux sections chromosomiques qui ne sont pas encore appariées (Figure 12 E) (Stanzione et al., 2016).

Au début du pachytène, la protéine IHO1 est éliminée de tous les axes chromosomiques, y compris les chromosomes non appariés, afin d'arrêter la formation des DSB (Figure 12 F). Ce mécanisme empêcherait la formation des cassures d'ADN à ce stade de la méiose, dans lequel les chromosomes homologues autosomes sont censés être associés sur toute leur longueur et les chromosomes sexuels hétérologues aux niveaux des PAR (Dereli et al., 2021).



**Figure 12. Modèle de la régulation de la formation des DSB.** La cellule dispose de plusieurs mécanismes pour réguler le moment, le nombre et la distribution des DSB. Les flèches vertes représentent l'activation et les rouges l'inhibition. La méiose commence par la réplication de l'ADN induisant la duplication des chromosomes, une chromatide de chaque chromosome homologue est représentée pour cette paire d'homologues. **(A)** La formation des DSB est initiée dans les régions répliquées du génome. Chez *S. cerevisiae*, ce sont les kinases de cycle cellulaire qui limitent la phosphorylation de Mer2 aux régions répliquées. **(B)** Un retard dans la réplication, suite à un stress, inhibe l'initiation de la formation des DSB. **(C)** Suite à la progression de la méiose, la formation des DSB (étoile jaune) et leur maturation en extrémités simple brin actif, respectivement, l'ATM et l'ATR. **(D)** Ces dernières inhibent l'accumulation des facteurs promoteurs de DSB et empêchent la formation de nouvelles DSB à proximité de la coupure d'ADN. **(E)** L'appariement des chromosomes et la formation de SC (en rose) empêchent la formation des DSB dans les régions déjà associées, grâce à l'activité de l'ATPase TRIP13 qui élimine HORMAD1. Par ce mécanisme, le SC empêche la formation de nouvelles machineries de DSB et limite l'activité de SPO11 aux régions qui ne sont pas encore appariées. **(F)** Au début du stade du pachytène, toutes les machineries des DSB sont désassemblées par l'élimination d'IHO1 des axes chromosomiques empêchant la formation des DSB au-delà de ce stade (Adaptée de (Dereli et al., 2021; Keeney et al., 2014)).

## Objectifs de la thèse

La formation et la réparation des cassures double brin d'ADN, lors de la recombinaison méiotique, est une étape cruciale pour la ségrégation des chromosomes et la formation des gamètes. Cependant, la formation de telles lésions présente un risque majeur pour l'intégrité génomique. Elles peuvent être létales pour la cellule méiotique ou conduire à des réarrangements génomiques et à des gamètes aneuploïdes. Pour éviter cette génotoxicité, la formation de cassures double-brin d'ADN implique plusieurs niveaux de contrôle et sollicite de nombreux facteurs protéiques pour réguler l'activité catalytique de SPO11.

Mon projet de thèse porte principalement sur les protéines murines REC114, MEI4 et IHO1, un sous groupe de protéines méiotiques proposées pour contrôler l'activité du complexe catalytique SPO11/TOPOVIBL. Il a été proposé qu'IHO1 jouerait un rôle important et permettrait le recrutement de MEI4 et de REC114 sur les axes chromosomiques. Néanmoins, les connaissances actuelles sur l'interaction directe entre IHO1 et REC114-MEI4 sont limitées, et reposent sur les interactions par double hybride, la co-immunoprécipitation et les analyses cytologiques (Kumar et al., 2018; Stanzione et al., 2016). La formation d'un complexe tripartite n'a jamais été établie. De plus, les propriétés biochimiques de ces protéines ne sont pas bien caractérisées. Les données biochimiques disponibles sur ces protéines méiotiques n'ont été publiées que récemment sur les protéines de la levure *S. cerevisiae* (Claeys Bouuaert et al., 2021a).

L'objectif de mon projet de thèse est de caractériser les propriétés biochimiques, d'élucider l'architecture moléculaire et d'analyser les interactions au sein du complexe et/ou des complexes formées par les protéines REC114, MEI4 et IHO1. Ce travail a été mené par différentes approches combinant la biochimie, la biophysique et la biologie structurale. Des analyses fonctionnelles, *in vivo*, sont également en cours en collaboration avec l'équipe de Dr. De Massy à l'Institut de Génétique Humain à Montpellier pour valider les données obtenues *in vitro*.



## **2 Matériels et méthodes**





## 2.1 Biologie moléculaire

Mon travail sur ce projet de thèse s'inscrit dans la continuité des travaux effectués par Ariadna Juarez pendant son post-doc. Elle a mené plusieurs expériences et a préparé plusieurs constructions sur les protéines méiotiques. Par conséquent, plusieurs constructions étaient disponibles au laboratoire à mon arrivée. Il s'agit en l'occurrence de : His-MBP-IHO1<sub>1-128</sub>, His-MBP-IHO1<sub>125-128</sub>, Strep-REC114, His-MEI4<sub>1-127</sub>, Strep-REC114<sub>1-150</sub>, Strep-REC114<sub>1-150</sub> L104D.

### 2.1.1 Clonage

Les protéines étudiées au cours de ce travail ont été exprimées en système bactérien. Les vecteurs d'expression utilisés permettent de cloner les séquences d'intérêt sous la dépendance du promoteur T7. Concernant les constructions que j'ai préparées, le fragment d'ADN contenant le gène d'intérêt a d'abord été amplifié par l'ADN polymérase Phusion (NEB) dans une réaction de polymérase en chaîne (PCR). Les amorces utilisées pour cette amplification sont faites de manière à introduire des sites de restrictions aux deux extrémités pour l'intégration de ces fragments d'ADN dans les vecteurs choisis. Elles permettent aussi, si besoin, d'introduire des étiquettes de purification.

Les produits de la PCR ont été vérifiés sur un gel d'agarose 1 %, puis purifiés par le Kit NucleoSpin Gel (Macherey-nagel). Ils ont ensuite été digérés par les enzymes de restriction appropriées. La même digestion a été effectuée pour le plasmide d'expression. Les ADN, du plasmide et de l'insert issus de la digestion, ont été isolés sur un gel d'agarose, puis purifiés en utilisant le même kit de purification. Ils ont été mélangés par la suite dans un ratio molaire (plasmide/insert) d'environ 1 : 5 pour être fusionnés dans une réaction en utilisant un Mix de ligation (Takara). Ce produit de ligation a été utilisé pour transformer la souche de clonage et de l'amplification de plasmide d'*E. coli*, Top10 (Thermo Fisher). Pour cette transformation, 50 µl de bactéries compétentes ont été incubées avec le mélange de la ligation, dans la glace, pendant 30 minutes. À la fin de cette durée, un choc thermique a été effectué à 42 °C durant 45 secondes. Le mélange a été incubé durant 3 minutes dans la glace, puis additionné avec 250 µl du milieu Luria-Bertani (LB). Le tout a été soumis à une agitation durant une heure à 37 °C avant d'être étalé sur des boîtes de LB-agar contenant l'antibiotique nécessaire, et d'être incubé à 37 °C. Ce protocole de transformation a été également utilisé pour la souche d'expression. Par la

suite, des petites cultures bactériennes ont été lancées à partir des colonies isolées. L'ADN purifié de ces dernières, par le kit NucleoSpin® Plasmid (Macherey-Nagel), a été vérifié par une restriction enzymatique, en utilisant les mêmes enzymes que ceux de la ligation. La présence de la séquence d'intérêt dans les plasmides extraits a été ensuite confirmée par séquençage au niveau de la plateforme de séquençage de GATC Biotech/Eurofins.

Les informations sur les constructions et les plasmides utilisés sont résumées dans le Tableau 1.

Construction	Plasmide	Antibiotique	Tag	Restriction (NEB)
ANKRD31 <sub>1808-1857</sub>	pETM41	Kanamycine	N-ter-His-MBP-TEV	NcoI/NotI
IHO1 <sub>1-24</sub>	pETM41	Kanamycine	N-ter-His-MBP-TEV	NcoI/NotI
IHO1 <sub>1-280</sub>	pETM41	Kanamycine	N-ter-His-MBP-TEV	NcoI/NotI
IHO1 <sub>125-281</sub>	pETM41	Kanamycine	N-ter-His-MBP-TEV	NcoI/NotI
IHO1 <sub>24-281</sub>	pETM41	Kanamycine	N-ter-His-MBP-TEV	NcoI/NotI
IHO1 <sub>133-191</sub>	pETM41	Kanamycine	N-ter-His-MBP-TEV	NcoI/NotI
IHO1 <sub>143-240</sub>	pETM41	Kanamycine	N-ter-His-MBP-TEV	NcoI/NotI
IHO1 <sub>143-257</sub>	pETM41	Kanamycine	N-ter-His-MBP-TEV	NcoI/NotI
IHO1 <sub>175-240</sub>	pETM41	Kanamycine	N-ter-His-MBP-TEV	NcoI/NotI
IHO1 <sub>175-257</sub>	pETM41	Kanamycine	N-ter-His-MBP-TEV	NcoI/NotI
IHO1 <sub>196-245</sub>	pETM41	Kanamycine	N-ter-His-MBP-TEV	NcoI/NotI
MEI4 <sub>1-127</sub>	pProEXHTb	Ampiciline	N-term His, TEV	BspHI/NotI
MEI4 <sub>1-43</sub>	pProEXHTb	Ampiciline	N-term His, TEV	BspHI/NotI
MEI4 <sub>1-43</sub>	pETM41	Kanamycine	N-ter-His-MBP-TEV	BspHI/NotI
REC114 <sub>1-159</sub>	pRSFduet	Kanamycine	N-ter-Strep	NcoI/NotI
REC114 <sub>FL</sub>	pRSFduet	Kanamycine	N-ter-Strep	NcoI/NotI
REC114 <sub>16-150</sub> IHO1 <sub>1-16</sub>	pProEXHTb	Ampiciline	N-term His, TEV	NcoI/NotI
REC114 <sub>226-254</sub>	pETM41	Kanamycine	N-ter-His-MBP-TEV	NcoI/NotI
TOPOVIBL <sub>450-579</sub>	pProEXHTb	Ampiciline	N-term His, TEV	NcoI/NotI
REC114 <sub>222-257</sub>	pETM41	Kanamycine	N-ter-His-MBP-TEV	NcoI/NotI

**Tableau 1. Récapitulatif des constructions utilisées durant cette thèse.** Le nom de la protéine, les limites des constructions, le plasmide utilisé plasmide (pETM41 (EMBL), pRSFduet (Novagen), pProEXHTb (Invitrogen)), la sélection d'antibiotique des enzymes de restriction, l'étiquette de la purification et la présence ou non du site de clivage à la TEV sont indiqués. Les constructions déjà présentes au laboratoire à mon arrivée sont écrites en bleu. N-ter : extrémité N-terminale.

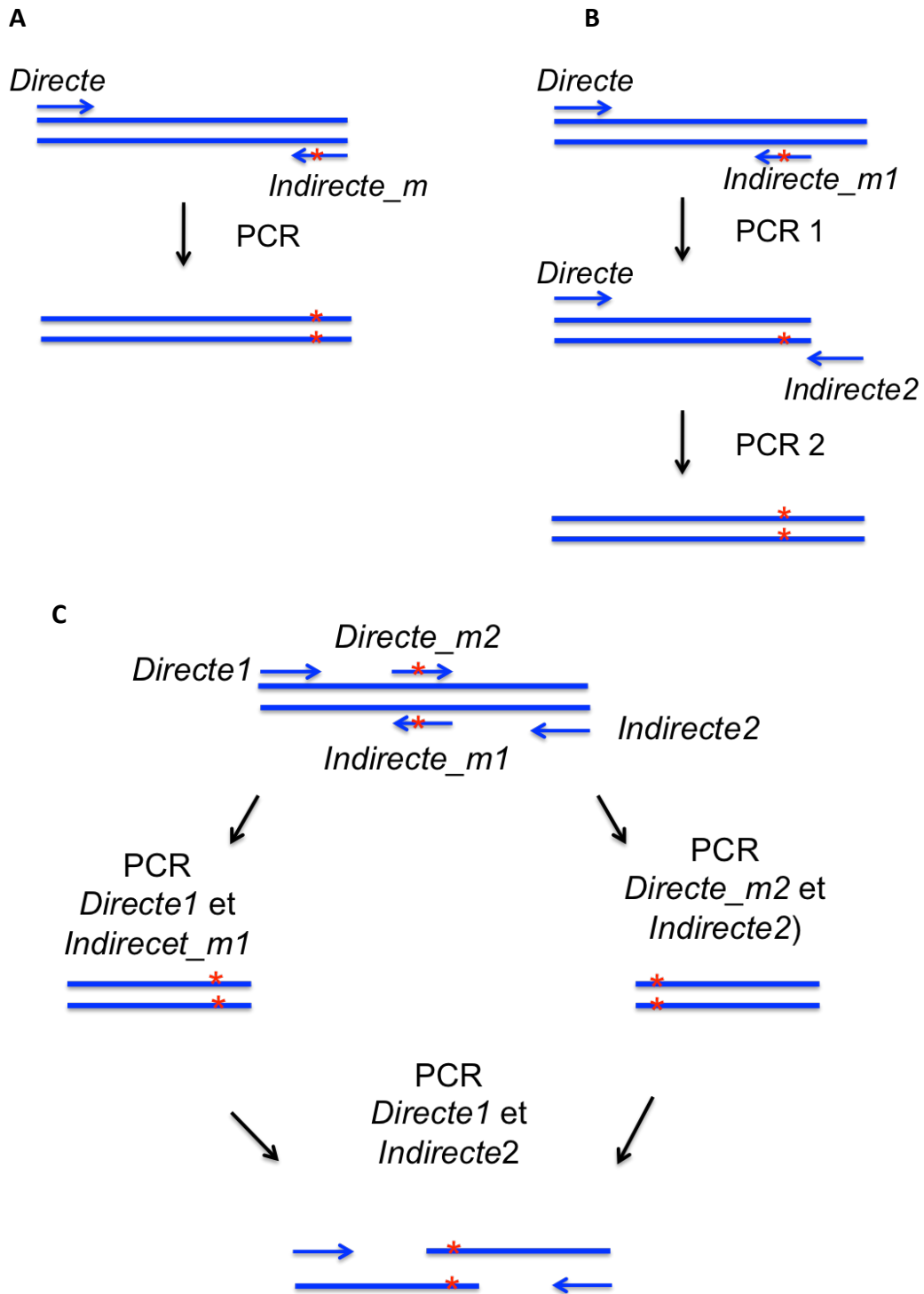
## 2.1.2 Mutagenèse dirigée

Des mutants ont été réalisés sur les protéines REC114, MEI4 et IHO1 pour analyser l'importance de certains résidus dans la formation de complexes entre ces protéines et/ou de valider les modèles structuraux obtenus. L'introduction des mutations a été effectuée par PCR.

Pour les fragments REC114<sub>222-257</sub> et IHO1<sub>1-281</sub> qui ont été mutées dans les extrémités de la séquence d'ADN, les mutations ont été introduites dans les amorces utilisées pour amplifier leurs séquences d'ADN (Figure 13 A).

Pour les mutations de REC114<sub>FL</sub> qui se trouvent proches de l'extrémité C-terminale, une première PCR a été effectuée avec l'amorce directe native (*Directe1*) est une amorce intermédiaire mutée (*Indirecte\_m1*). Cette réaction permet d'introduire la mutation, mais ne produit qu'un fragment du gène, le reste du gène a été complété en utilisant une deuxième amorce *Indirecte2* dans une deuxième réaction d'amplification avec l'amorce *Directe1* (Figure 13 B).

Pour les mutations qui se trouvent au milieu de la séquence d'ADN, comme le cas de MEI4, deux amorces complémentaires et porteuses de la mutation au milieu de la séquence ont été conçues (*Directe\_m2* et *Indirecte\_m1*) (Figure 13 C). Deux réactions de PCR ont été ensuite lancées en parallèle. Dans l'une, l'amorce complémentaire avec l'extrémité 5' a été utilisée avec l'amorce indirecte mutée (*Directe1* et *Indirecte\_m1*), ce qui permet d'avoir la mutation à l'extrémité 3' de la séquence et d'amplifier la séquence avant la mutation. Dans l'autre, c'est l'amorce directe mutée qui est utilisée avec l'amorce indirecte de l'extrémité 3' (*Directe\_m2* et *Indirecte2*), introduisant ainsi la mutation à l'extrémité 5' et amplifiant la séquence de la construction qui se trouve après la mutation. Les produits de ces deux PCR, étant complémentaires à l'endroit de la mutation, ont été utilisés comme matrice pour une troisième PCR dans laquelle on utilise les amorces des extrémités 5' et 3' pour avoir la totalité de la séquence d'ADN. Les amplicons ont été ensuite intégrés dans les plasmides d'expression par le clonage comme décrit précédemment.



**Figure 13. Mutagenèse dirigée.** Représentation schématique des différentes étapes de la mutagenèse dirigée en fonction de la position de la mutation (Etoile rouge). **(A)** La mutation se trouve à l'extrémité de la séquence nucléotidique (les mutations en C-terminale sont présentées ici, mais le même protocole est suivi pour les mutations du côté N-terminal). **(B)** La mutation se trouve proche de l'extrémité, mais assez loin pour pouvoir l'introduire directement dans l'amorce indirecte. **(C)** La mutation se trouve au milieu de la construction. PCR : réaction de polymérase en chaîne. L'amorce porteuse de la mutation est suffixée par -m.

## 2.2 Production et purification

### 2.2.1 Expression dans *E.coli*

Après le clonage, les plasmides, codant les séquences d'intérêts, ont été utilisés pour transformer la souche d'expression BL21 star (DE3, Invitrogen). Après le choc thermique, les transformants ont été cultivés pendant 1 h dans du milieu LB à 37 °C, puis étalés sur des boîtes de LB-agar additionnées de l'antibiotique nécessaire. Les colonies obtenues ont été resuspendues dans un faible volume de LB, supplémenté d'antibiotique, et laissées pousser pour 12 h à 18 h. Ces précultures ont été utilisées par la suite pour inoculer un plus large volume de culture bactérienne. Les bactéries ont été cultivées à 37 °C sous agitation à 180 rotations par minute (rpm) jusqu'à une densité optique à 600 nm de 0,8. L'expression de la protéine a été alors induite avec 1 mM d'isopropyl d-thiogalactopyranoside (IPTG) à 20 °C. Après une nuit d'induction, les cellules ont été récoltées par centrifugation à 4 °C et à 4000 g pendant 10 min. Les culots ont été stockés à -20 °C ou utilisés directement.

### 2.2.2 Expression des protéines enrichies aux marquages <sup>15</sup>N et <sup>13</sup>C

Afin d'analyser le complexe REC114-MEI4 par RMN, le plasmide pETM41-REC114<sub>222-256</sub> et le plasmide pProex-MEI4<sub>1-43</sub> ont été cotransformés dans les cellules compétentes BL21 star (DE3, Invitrogen). Une préculture initiale a été lancée durant 9 h à 37 °C dans 5 ml de LB additionné de 100 mg/ml d'ampicilline et de 50 mg/ml de kanamycine. Une deuxième préculture de 100 ml a été ensuite lancée à partir de la première, dans le milieu minimum M9 contenant 10 g/L Na<sub>2</sub>HPO<sub>4</sub>; 3 g/L K<sub>2</sub>HPO<sub>4</sub>; 0,5 g/L NaCl; 1 g/L <sup>15</sup>NH<sub>4</sub>Cl; 3 g/L glucose <sup>13</sup>C, 1 mM MgSO<sub>4</sub>; 0,1 mM CaCl<sub>2</sub>; 0,01 mM MnCl<sub>2</sub>; 0,05 mM ZnSO<sub>4</sub>; 0,05 mM FeCl<sub>3</sub>; 1mg/L biotine, 1 mg/L thiamine et les antibiotiques. Le culot bactérien obtenu de cette préculture a été utilisé pour inoculer 1 L de culture dans le même milieu M9. La surexpression du complexe a été ensuite poursuivie comme décrite dans la partie précédente.

## 2.2.3 Purification

### 2.2.3.1 Lyse des bactéries

Les culots bactériens ont été re-suspendus dans un tampon de lyse contenant 20 mM Tris-HCl pH 8,0, 125 mM NaCl, 2 % glycérol, 2 mM  $\beta$ -mercaptoéthanol (BME) et un mélange d'inhibiteurs de protéases cOmplete EDTA-free (Roche) (1 pastille par 50 ml). Le volume de tampon utilisé est d'environ 30 mL pour un culot issu d'un litre de culture. Les cellules ont été ensuite lysées par sonication dans la glace, à une amplitude de 70 % durant 3 min (3 secondes de sonication espacées par des pauses de 6 secondes). L'extrait total a été centrifugé pendant 30 minutes à 50 000 g et à 4 °C et l'extrait soluble a été récupéré.

### 2.2.3.2 Chromatographie d'affinité

#### *Chromatographie d'affinité sur la résine nickel*

Cette étape de purification a été utilisée pour isoler les protéines produites en fusion avec une étiquette poly-histidine. L'extrait soluble obtenu de l'étape de la lyse bactérienne a été supplémenté de 10 mM imidazole et chargé sur une colonne de chromatographie contenant 3 ml de la résine Ni<sup>2+</sup> chelating Sepharose (Cytiva), préalablement équilibrée avec le tampon de lyse. Après le passage de tout le lysat par gravité, la résine a été lavée avec le tampon de lyse supplémenté de 20 Mm Imidazole, puis en augmentant la concentration de l'imidazole à 100 mM. La protéine a été ensuite éluée avec le tampon d'éluion contenant 20 mM Tris-HCl pH 8,0, 125 mM NaCl, 2 % glycérol, 2 mM BME et 500 Mm Imidazole. Les fractions d'éluion ont été analysées par SDS-PAGE sur un gel de à 15 %.

#### *Chromatographie d'affinité sur la résine amylose*

La protéine Maltose-binding protein (MBP) a été utilisée pour augmenter la solubilité et le rendement de certaines constructions qui sont indiquées dans le tableau 1. Elle a également servi comme une étiquette de purification. De manière similaire à la purification sur la résine chargée de nickel, l'extrait soluble, contenant des protéines exprimées en fusion avec une étiquette His-MBP, a été passé sur une colonne de résine d'amylose (NEB). Celle-ci a été ensuite lavée avec le tampon de lyse et a été éluée avec le même tampon additionné de 10 mM maltose.

### *Chromatographie d'affinité sur la résine Strep-Tactin*

Le tag Strep est une séquence peptidique de huit résidus (Trp-Ser-His-Pro-Gln-Phe-Glu-Lys) qui est conçue pour se lier spécifiquement à la Strep-tactine (dérivée de la streptavidine) au niveau du site de fixation de la biotine. Ce système de purification est très sélectif, ce qui permet l'obtention de protéines avec une très grande pureté. La résine utilisée pour cette purification est la Strep-Tactin XT (high capacity) (IBA). L'éluion de la protéine retenue par cette dernière a été réalisée avec le tampon de lyse supplémenté de 50 mM biotine, puis analysée par SDS-PAGE.

#### *2.2.3.3 Coupure à la protéase TEV*

Cette étape n'était pas effectuée systématiquement, mais était réservée au cas où le tag His ou le tag His-MBP devaient être retirés. À cet effet, les fractions d'éluion de la chromatographie d'affinité sur la résine amylose ou nickel ont été regroupées et incubées avec de la TEV, en présence de 10 mM BME pendant la nuit à 4 °C (0.01 mg de TEV pour 1 mg de protéine).

#### *2.2.3.4 Dialyse et deuxième chromatographie d'affinité*

Après le clivage à la TEV, les échantillons protéiques ont été dialysés pour enlever l'imidazole, en utilisant une membrane dialysante (rétention > 3500Da), contre un tampon composé de 20 mM Tris-HCl pH 8,0, 125 mM NaCl et 2 % glycérol à 4 °C pendant 3 heures. Les échantillons dialysés ont été ensuite passés sur colonne contenant 3 ml de la résine Ni<sup>2+</sup> chelating Sepharose (GE Healthcare). La protéine qui ne contient plus d'étiquette d'affinité n'est plus retenue et elle passe à travers la résine. La résine a été lavée avec un tampon de 20 mM Tris-HCl pH 8,0, 125 mM NaCl, 2 % glycérol, 2 mM BME, 20 mM Imidazole jusqu'à la récupération de toute la protéine.

#### *2.2.3.5 Filtration sur gel*

Cette technique est basée sur le mécanisme d'exclusion de taille. Les molécules sont séparées de la plus grande à la plus petite en fonction de leur taille moléculaire en solution ainsi que de leur forme. Les échantillons protéiques issus de la purification par affinité ont été concentrés sur une membrane Amicon Ultra (Millipore) jusqu'à un volume de 500 µl. La limite d'exclusion de la membrane a été sélectionnée en fonction du poids moléculaire de la protéine purifiée. L'échantillon concentré a été ensuite chargé sur une

colonne de chromatographie d'exclusion de taille (Superdex 200, Superdex 75 ou Superose 6 (10/300 GL, GE Healthcare)), équilibrée avec un tampon de 50 mM Tris-HCl pH 8,0, 125 mM NaCl, 2 % glycérol, 2 mM BME. Le choix de la colonne dépend de la taille de la protéine ainsi que de la résolution recherchée. Les fractions des pics d'élution ont été analysées par SDS-PAGE.

## **2.3 Caractérisation du complexe IHO1-REC114-MEI4**

### **2.3.1 Reconstitution du complexe IHO1-REC114-MEI4**

Les fractions d'élution de His-MBP-IHO1<sub>1-281</sub> et celles du complexe Strep-REC114-His-MEI4<sub>1-128</sub>, obtenues après la purification par la chromatographie d'affinité ont été concentrées sur la membrane de concentration (Amicon 30 kDa et 50 kDa respectivement). Elles ont été mélangées à un ratio molaire 1 : 2 (IHO1/ REC114-MEI4) dans un volume de 500 µl. Cet échantillon a été ensuite injecté une colonne d'exclusion de taille (Superose 6, GE Healthcare) en utilisant un tampon de 50 mM Tris-HCl pH 8,0, 125 mM NaCl, 2 % glycérol, 2 mM BME. L'élution du complexe a été analysée par SDS-PAGE, et les fractions correspondant à la co-purification des trois protéines ont été regroupées.

### **2.3.2 Titration calorimétrique isotherme (ITC)**

L'ITC est une technique puissante pour caractériser les interactions entre macromolécules biologiques en solution. Cette technique est basée sur la mesure de la chaleur libérée ou absorbée lors d'une interaction. Les expériences d'ITC sont réalisées à l'aide d'un appareil, le microcalorimètre, constitué de deux cellules, une cellule de référence et une cellule qui contient l'échantillon protéique. Des petits volumes d'un ligand ou d'un partenaire protéique sont injectés dans la cellule de l'échantillon, à un intervalle régulier. L'interaction de ligand avec l'échantillon de la cellule induit une émission ou une absorption de la chaleur. L'appareil détecte la différence de température entre les deux cellules et compense cette différence. L'énergie nécessaire pour cela permet de mesurer la chaleur libérée ou absorbée pour chaque ajout du ligand. Au fur et à mesure de la titration du ligand dans la cellule, la saturation des sites de liaison des ligands sur les protéines réduit les variations de température entre deux cellules. La courbe d'interaction est ensuite déduite à partir de ces variations et elle permet de mesurer la stœchiométrie (n), l'enthalpie



( $\Delta H$ ) et la constante d'association  $K_a$ . À partir de ces dernières, on peut déduire l'affinité ( $K_d$ ), les variations d'énergie libre  $\Delta G$  et l'entropie  $\Delta S$ .

Les mesures d'ITC ont été réalisées à la plateforme ISBG avec l'aide de Caroline Mas. Ces mesures ont été utilisées pour caractériser l'interaction entre REC114 et IHO1. Ces expériences ont été réalisées à 25°C en utilisant un microcalorimètre ITC200 (MicroCal). Les deux échantillons protéiques ont été préparés dans un tampon de 20 mM Tris-HCl pH 8,0, 125 mM NaCl, 2 mM BME. La protéine REC114 (ou le complexe REC114-MEI4) a été placée dans la cellule à une concentration de 30  $\mu$ M. 500  $\mu$ M (ou 1mM) d'IHO1<sub>1-281</sub> (ou l'un de ses variant) a été titré dans la cellule par 16 injections (la première injection à 0.5  $\mu$ l puis les restes à 2.5  $\mu$ l). Les données de la première injection ont été supprimées et les données restantes ont été traitées par logiciel Origin, version 7.0 (MicroCal), avec un modèle à un site de liaison.

### 2.3.3 Pontage chimique

Le pontage chimique (cross-link) permet de lier d'une manière covalente des molécules, ou des sous-unités, qui interagissent ou qui sont proches dans l'espace avec des liaisons non-covalentes. Ces liaisons sont formées par l'intermédiaire d'agents réticulants. Ces derniers contiennent des groupes terminaux qui réagissent avec les groupes fonctionnels (amines primaires et les sulfhydryles) des acides aminés. Cette technique permet de stabiliser et de prévenir la dissociation du complexe protéique dont les interactions sont faibles ou transitoires.

Le BS3 (bis(sulfosuccinimidyl)suberate) (Thermo Fisher) a été l'agent de pontage le plus utilisé durant cette thèse. Il contient dans ses deux extrémités un ester de N-hydroxysulfosuccinimide (NHS) réactif aux amines primaires (notamment les groupes  $\epsilon$ -amino des résidus lysine) en formant des liaisons amides stables. Ces deux groupements sont espacés de 8 atomes de carbone, ce qui correspond à une distance de 11.4 Å (Figure 14).

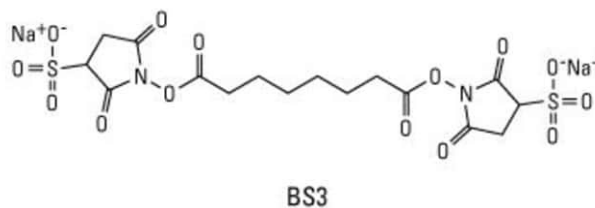


Figure 14 : Représentation de la molécule de pontage BS3.

Le pontage chimique par BS3 a été effectué pour le complexe Strep-REC114-HisMEI4<sub>1-127</sub>-His-SUMO-IHO1<sub>1-281</sub>. La réactivité de BS3 aux amines empêche l'utilisation d'un tampon contenant du Tris. Pour cette raison, la reconstitution du complexe REC114-MEI4-IHO1 par la filtration sur gel a été effectuée dans un tampon contenant 20 mM HEPES pH 8, 125 mM NaCl. Les fractions de la co-elution du complexe ternaire ont été utilisées pour le pontage chimique. 250 µg du complexe dans un volume de 500 µl ont été réticulés en présence de 20 µM de BS3 (Thermo Fisher). Après 30 minutes d'incubation à température ambiante, la réaction a été arrêtée par l'ajout de 25 mM de Tris-HCl, pH 8,0.

Après une vérification par SDS-PAGE, l'échantillon protéique réticulé a été concentré jusqu'à un volume de 500 µL, puis élué sur la colonne d'exclusion stérique (Superose 6) avec 20 mM HEPES pH 8, 125 mM NaCl.

### **2.3.4 Pontage chimique couplé à la spectrométrie de masse**

Le produit du pontage de REC114-MEI4-IHO1 avec le BS3, élué par gel-filtration, a été analysé par SDS-PAGE. La bande correspondant au complexe a été découpée du gel pour être analysée par pontage chimique couplé à la spectrométrie de masse. Cette technique permet de montrer les peptides contenant les résidus qui sont impliqués dans des pontages chimiques avec le BS3. La digestion des protéines par la trypsine, l'extraction des peptides du gel, leurs analyses par la chromatographie liquide couplée à la spectrométrie de masse en tandem (LC-MS/MS), et le traitement des données ont été réalisés par plateforme de spectrométrie de masse (MFPL, Max Perutz Labs) à Vienne en Autriche.

## **2.4 Caractérisation du complexe REC114-MEI4, de IHO1 et de REC114 C-ter**

### **2.4.1 Pull down**

Cette méthode a été utilisée pour analyser l'effet de la mutation de certains résidus sur l'interaction entre REC114-MEI4. Pour analyser les mutants de REC114, le fragment (1-43) de la protéine MEI4, exprimé en fusion avec l'étiquette His-MBP, a été purifié sur colonne amylose, puis incubé pendant 1 heure avec la résine Strep-tactin sur laquelle les différents mutants de Strep-REC114<sub>FL</sub> sont immobilisés. Après plusieurs lavages de la résine avec un tampon de 20 mM Tris-HCl pH 8,0, 125 mM NaCl, 2 % glycérol, 2 mM

BME, les protéines liées ont été éluées, par le même tampon complété de 50 mM de biotine, puis analysées par SDS-PAGE.

### **2.4.2 Diffusion de la lumière à plusieurs angles couplée à la chromatographie d'exclusion de taille (SEC-MALLS)**

Le SEC-MALLS est une technique robuste pour la détermination de la masse moléculaire des protéines et des complexes en solution. Elle est basée sur une séparation des échantillons par une chromatographie d'exclusion de taille (SEC) couplée à un système de détection multiple comprenant notamment un détecteur d'indice de réfraction (RI) et un détecteur de diffusion de lumière à des angles multiples (MALLS). Étant donné que l'intensité de la lumière diffusée par les particules en solution est directement liée au produit de leur masse moléculaire par leur concentration, la détermination de la masse moléculaire par cette technique ne dépend que des lectures obtenues à partir de ces détecteurs.

Les expériences de SEC-MALLS ont été effectuées à la plateforme ISBG par Aline Le Roy sur une chaîne HPLC (Schimadzu, Kyoto, Japan), composée d'un dégazeur DGU-20AD, d'une pompe LC-20AD, d'un passeur d'échantillon SIL20ACHT, d'une interface de communication CBM-20A, d'un détecteur UV-Visible SPD-M20A, d'un four à colonne XL-Therm (WynSep, Sainte Foy d'Aiguille, France), d'un détecteur de diffusion statique de la lumière MiniDAWN Treos, d'un détecteur de diffusion dynamique de la lumière DynaPro Nanostar et d'un détecteur d'indice de réfraction Optilab rEX (Wyatt, Santa-Barbara, USA). Un volume de 100 µl d'échantillon a été injecté à une concentration d'environ 3 mg/ml dans une colonne Superdex 200 10/300 GL (GE Healthcare), préalablement équilibrée avec le tampon de la protéine. L'albumine de sérum Bovine (BSA), à 2 mg/mL, dans du tampon phosphate salin (PBS) a été injectée comme contrôle. Les analyses ont été faites par Aline le Roy à l'aide du logiciel ASTRA, version 5.4.3.20.

### **2.4.3 Ultracentrifugation analytique (AUC)**

L'AUC est une technique robuste pour analyser les caractéristiques des macromolécules biologiques telles que la taille, la forme, la stœchiométrie et les interactions protéiques. Elle est basée sur la séparation des composants, présents dans un échantillon, par une très forte centrifugation, et le suivi de leurs limites de sédimentation grâce à des systèmes optiques d'absorbance et d'interférence. La vitesse de sédimentation

des molécules de l'échantillon, ou le coefficient de sédimentation, dépend de leurs tailles, de leurs masses, leurs formes, et de leurs densités.

Cette technique a été utilisée pour analyser la stœchiométrie et la dynamique du complexe REC114-MEI4. Dans cette étude, les constructions d'ADN codant le fragment C-terminal de REC114 (His-MBP-REC114<sub>226-254</sub>) et la partie N-terminale de MEI4 (His-MEI4<sub>1-43</sub>) ont été co-exprimées en bactéries. Après la co-purification sur la résine amylose et le clivage des étiquettes de purification par la TEV, l'échantillon protéique a été passé sur une résine nickel pour se débarrasser des étiquettes de purification et des protéines éventuellement non clivées. Le complexe a été ensuite purifié sur une colonne de chromatographie d'exclusion de taille et il a été préparé à trois concentrations (0.2 mg/ml, 1 mg/ml et 4.5 mg/ml) dans le tampon de 20 mM Tris pH8, 125 mM NaCl, 2 % de glycérol et 2 mM BME.

L'expérience de la vitesse de sédimentation a été réalisée à plateforme de l'ISBG, par Aline Le Roy. Elles ont été effectuées à 42000 rpm, à 10 °C, sur une ultracentrifugeuse analytique XLI avec le rotor Anti-50 (Beckman Coulter, Brea, USA), et des cellules à doubles canaux de 12 mm communicants et 3 mm de trajet optique équipées de fenêtres de Saphir (Nanolytics, Potsdam, DE). La référence utilisée est le tampon de la protéine. L'acquisition a été faite en absorbance à 280 et en interférence. Les données ont été traitées par Aline Le Roy avec le logiciel REDATE v 1.0.1 et les analyses ont été effectuées avec les logiciels SEDFIT v 16.36 et GUSSE 1.4.2.

#### **2.4.4 Spectrométrie de masse en condition native**

La spectrométrie de masse (MS) est une technique précise pour mesurer la masse moléculaire des composés analysés. Cette technique est basée sur l'ionisation des molécules par la source d'ionisation, puis la séparation des ions générés en fonction de leur rapport masse sur charge ( $m/z$ ) au niveau de l'analyseur. Les ions séparés sont ensuite quantifiés au niveau du détecteur et les signaux obtenus permettent de générer des spectres de masse, où les valeurs  $m/z$  des ions sont tracées en fonction de leurs intensités. Cette technique peut également être appliquée aux complexes biologiques tout en préservant leurs interactions non covalentes. Dans ce cas, on parle de la spectrométrie de masse en condition native (MS native). Elle permet de déterminer la masse précise des complexes intacts, et de fournir des informations précieuses sur leurs stœchiométries, les

interactions entre les sous-unités et les dynamiques d'assemblages de ces dernières. La source d'ionisation la plus utilisée, pour transférer ces complexes de la phase soluble à la phase gazeuse, est le nano-electrospray (Nano-ESI). Lors de l'ESI, les ions moléculaires sont générés en appliquant une haute tension à un échantillon dissout dans une solution volatile puis mis en circulation dans un capillaire étroit. La MS native est souvent effectuée à l'aide d'un tampon volatil tel que l'acétate d'ammonium qui maintient l'intégrité des interactions et qui est compatible avec le processus de l'ionisation. Par conséquent, le tampon d'échantillon doit être remplacé avant l'analyse. Pour analyser les ions générés du complexe protéique, la source ESI est généralement associée à l'analyseur de quadripole à temps de vol (TOF), qui permet l'analyse de larges gammes de masses moléculaires avec une haute précision. Les analyses de MS native réalisées durant cette thèse ont été effectuées au niveau de la plateforme de spectrométrie de masse à l'IBS par Elisabetta Boeri Erba.

Avant d'analyser l'état natif des protéines, une étape de MS en conditions dénaturantes a été effectuée afin de déterminer la masse expérimentale de chaque protéine ou de chaque monomère. Ces mesures ont été réalisées sur un spectromètre LC/ESI-TOF (6210), couplé à un système de chromatographie liquide à pompe binaire (Agilent Technologies). Les données MS ont été traitées par Elisabetta Boeri Erba, avec le logiciel MassHunter (Agilent Technologies).

Les échantillons analysés par MS natives sont : Strep-REC114<sub>FL</sub>-MEI4<sub>1-127</sub>, His-MBP-REC114<sub>226-254</sub> et His-REC114<sub>226-254</sub>-His--MBP-MEI4<sub>1-43</sub>. Le tampon de ces derniers a été changé en utilisant la filtration sur gel. Les échantillons purifiés ont été chargés sur une colonne d'exclusion stérique puis élués avec un tampon d'ammonium acétate à 300 mM. Pour un changement optimal de tampon, deux étapes successives de gel-filtration ont été effectuées. 20 µM de l'échantillon ont ensuite été donnés pour la plateforme de spectrométrie de masse de l'IBS pour effectuer les analyses de MS native. Ces dernières ont été réalisées sur le spectromètre d'ionisation par nano-électrospray-temps de vol quadripolaire (nano-ESI-Q-TOF). De nouveaux spectres ont été acquis par la suite en mode tandem MS (mode MS/MS) sur les espèces qui semblent pertinentes lors des premiers spectres. Lors de cette expérience, une population spécifique d'ions est transmise à la cellule de collision où elle entre en collision avec des molécules d'un gaz inerte. Ce qui provoque la dissociation du complexe et permet la confirmation de la stœchiométrie et

l'identification des sous-unités impliquées. Les données ont été traitées avec le logiciel MassLynx ® 4.0 (Waters) (Boeri Erba et al., 2018) .

## 2.5 Caractérisation structurale

### 2.5.1 Coloration négative

La coloration négative est une méthode de microscopie électronique à transmission qui permet d'observer les échantillons biologiques. Elle peut être utilisée pour contrôler la qualité des échantillons purifiés, obtenir des informations sur l'homogénéité de l'échantillon et vérifier l'assemblage ou le désassemblage des complexes protéiques. Elle est basée sur le dépôt de l'échantillon protéique sur une grille metalo-carbonée qui est ensuite traité avec une solution de sels de métaux lourds (le colorant). L'agent colorant va entourer le volume occupé par l'échantillon. Lorsque le faisceau d'électrons traverse la grille, l'agent contrastant dévie davantage les électrons comparés aux protéines. Ce qui permet de visualiser des particules biologiques claires entourées de zones plus sombres du colorant.

Cette méthode a été utilisée pour visualiser le complexe IHO1-REC114-MEI4 et la protéine IHO1. Certaines grilles ont été préparées et été analysées par Michael Hons au niveau de la plateforme de la microscopie de l'EMBL, d'autres par Anne Emmanuelle Foucher au niveau de la plate-forme de la microscopie électronique de l'IBS.

Les grilles de coloration négative à carbone continu ont été d'abord traitées par une décharge lumineuse pour les rendre plus hydrophiles afin d'assurer une bonne absorption de l'échantillon. Un échantillon protéique de 3  $\mu$ L d'environ 0.2 mg/ml a été pipeté sur la grille et laissé absorber durant 1 minute. L'excès de l'échantillon a été retiré par un papier absorbant. La grille a ensuite été incubée sur une goutte d'acétate d'uranyle à 3% durant 1 minute. L'excès du colorant a été ensuite retiré et les grilles ont été laissées pour séchage à l'air libre avant de les analyser. Les grilles ont été visualisées avec le microscope Tecnai G2 Spirit BT T12 à 120 kV avec caméra 16 M CETA à l'EMBL, ou avec le microscope à 200kV équipée d'une caméra 4kx4k Ceta FEI à l'IBS.

### 2.5.2 Cryo-microscopie électronique (Cryo-EM)

Ces dernières années, la Cryo-EM est devenue une technique puissante pour la détermination des structures protéiques. Elle permet d'analyser les biocomplexes sans

coloration, dans des conditions natives et à de multiples états conformationnels. Cette technique est basée sur la visualisation des échantillons biologiques dans un état vitrifié. Pour cela, et dans un appareil appelé Vitrobot, une faible quantité de l'échantillon est déposée sur une grille recouverte d'une fine couche du carbone perforé, l'excès de l'échantillon est ensuite éliminé et la grille est prolongée dans l'éthane liquide qui permet une congélation de l'échantillon par formation d'une glace vitreuse. L'échantillon est ensuite visualisé par microscope électronique à transmission et il est exposé à un faisceau d'électrons incidents du microscope. Les électrons diffusés par l'échantillon traversent les différentes lentilles de microscope pour créer une image à deux dimensions de l'échantillon au niveau du détecteur. Ces images sont appelées micrographes et elles correspondent à des projections 2D de l'échantillon dans différentes orientations. Ces données sont ensuite traitées pour combiner les différentes projections en une reconstruction 3D de l'objet.

Un jeu de données d'environ 1000 micrographes a été collecté pour la construction IHO1<sub>125-281</sub>, par Michael HONS sur le microscope électronique Glacios équipé d'une caméra à comptage direct d'électrons Falcon 3 en utilisant un voltage de 200 kV (Thermo Fisher). Les données obtenues ont été traitées par le programme cryoSPARC v3.3.2.

### **2.5.3 Résonance magnétique nucléaire (RMN)**

La RMN est une technique polyvalente qui permet d'analyser la structure, la dynamique et les interactions protéiques. Pour obtenir ces informations, l'échantillon protéique est enrichi en isotopes  $^{15}\text{N}$  et/ou  $^{13}\text{C}$  est placé dans un champ magnétique appliqué par le spectromètre puis excité par une impulsion de champs radiofréquence. Le retour à l'équilibre du système, après son excitation, génère des fréquences de résonance qui sont propres à chaque noyau ( $^1\text{H}$ ,  $^{13}\text{C}$  ou  $^{15}\text{N}$ ) et qui dépendent en partie de son environnement moléculaire. L'attribution de ces fréquences ou ces déplacements chimiques aux protons et aux atomes de carbone et d'azote est la première étape pour accéder aux informations structurales. Les spectres 2D et 3D acquis lors des expériences de RMN permettent de corréler séquentiellement les atomes de  $^1\text{H}$ ,  $^{15}\text{N}$  et  $^{13}\text{C}$  sur le squelette peptidique puis sur toutes les chaînes latérales. Ils permettent également de déterminer un certain nombre de contraintes géométriques (angle, distance spatiale) qui vont être utilisées pour le calcul de la structure.

La méthode RMN a été utilisée pour avoir des informations structurales sur le complexe REC114-MEI4. Le complexe minimal REC114<sub>226-254</sub>-MEI4<sub>1-43</sub> a été préparé à 500  $\mu$ M dans 20 mM Tris pH8, 125 mM NaCl et 2mM BME.

Les spectres 2D qui ont été collectés sont HSQC ((Heteronuclear Single Quantum Correlation) et TROSY (Transverse relaxation optimized spectroscopy), corrélant le proton H à celui d'un atome  $^{15}\text{N}$  auquel il est lié.

Les spectres 3D s'interprètent par paires, car dans la même paire, les résonances des atomes  $^1\text{H}$ ,  $^{15}\text{N}$  et  $^{13}\text{C}$  appartiennent au résidu (i-1) dans un spectre, et elles appartiennent au résidu (i) et au résidu précédent (i-1) dans l'autre spectre, ce qui permet de relier le résidu i-1 au résidu i. Les spectres 3D collectés sont :

-HNCO qui lie  $^1\text{H}$ ,  $^{15}\text{N}$  au  $^{13}\text{CO}$  du résidu précédent et HNcaCO qui les corrèle au  $^{13}\text{CO}$  du résidu actuel et du résidu précédent.

- HNcoCA qui lie les résonances  $^1\text{H}$ ,  $^{15}\text{N}$  au Ca du résidu précédent et HNCA qui les corrèle au Ca du résidu actuel et du résidu précédent.

L'acquisition des spectres RMN a été effectuée par Cameron Mackereth, à IECB de Bordeaux. Les expériences ont été réalisées à 298 K sur le spectromètre Bruker Neo à 800 MHz, équipé d'une cryosonde. Les données ont été ensuite analysées par Sparky 3 (T.D. Goddard and D.G. Kneller, University of California). Des acquisitions ont été également réalisées à la plateforme RMN de l'ISBG par Adrien Fabier à 298 K sur un spectromètre Bruker 700 MHz, équipé d'une cryosonde. Ceci en vue d'optimiser la qualité des spectres en testant différentes conditions de tampon et de températures d'acquisition.

#### **2.5.4 Essais de cristallisation**

La cristallographie aux rayons X est une technique permettant de résoudre la structure des macromolécules biologiques à l'échelle atomique. Cette technique repose sur la diffraction des rayons X par tous les électrons des macromolécules empilées de manière périodique et ordonnée dans un cristal. Le profil de diffraction obtenu reflète l'arrangement des atomes dans le cristal et est utilisé pour déterminer la structure originale de la macromolécule. La première étape de cette technique, après l'obtention d'un échantillon pure et homogène, est donc la formation des cristaux avec une qualité de diffraction



optimale. Le criblage des conditions de cristallisation pour les constructions testées pendant cette thèse a été effectué d'une manière automatisée au niveau de la plateforme de cristallisation de l'EMBL. 6 kits commerciaux de 96 conditions (The Classics Suite (Qiagen), The JCSG (MD), The PACT (MD), The PEGs I (Qiagen), Wizard I & II (Rigaku), Salt-Grid\_hampton) ont été testés. Les échantillons protéiques ont été purifiés puis préparés à des concentrations de 5 mg/ml, 10 mg/ml et 15 mg/ml dans un tampon contenant 20 mM Tris pH8, 125 mM NaCl et 2mM BME. Un robot a été utilisé pour la réalisation des gouttes selon la technique de diffusion de vapeur en goutte assises. Pour réaliser ces tests de cristallisation sur nos échantillons, un volume de 100 nl de protéine à 5 mg/ml, 10 mg/ml et 15 mg/ml a été mélangé à 100 nl de solution de cristallisation. Les boîtes ont ensuite été scellées, étiquetées et conservées à 20 ou 4 °C. Des photographies de la goutte ont été prises régulièrement pour suivre leurs évolutions. Les cristaux obtenus ont été congelés instantanément dans de l'azote liquide avec une solution composée de la solution de cristallisation et 30 % de glycérol, puis testés sur les lignes de lumière du synchrotron à l'ESRF.

Des tests de cristallisation ont été également effectués pour IHO1<sub>143-257</sub> en utilisant la technique de la goutte suspendue dans des plaques 24 puits à 20 °C. Les gouttes ont été préparées avec 1 µL de la protéine et 1 µL de tampon de cristallisation et elles ont été suspendues sur un puits de 1 ml de solution de cristallisation.

### 2.5.5 AlphaFold 2

Alphafold est le nouveau programme d'intelligence artificielle qui a révolutionné la prédiction de la structure 3D des protéines. Ce système prend en entrée la séquence d'acide aminé de la protéine recherchée. Cette séquence est utilisée d'une part, pour construire un alignement de séquences multiples (MSA) avec les séquences protéiques présentes dans les bases de données, et d'autre part pour identifier les protéines pouvant avoir des structures similaires dans la banque des structures protéiques. Cela permet de déterminer les parties de la séquence qui ont co-varié au cours de l'évolution et de trouver les résidus susceptibles d'être en contact les uns avec les autres. Ces informations sont ensuite traitées afin de construire un modèle des interactions au sein de la protéine et de générer sa structure 3D. AlphaFold2 combine plusieurs apprentissages en profondeur pour prédire non seulement la structure 3D des protéines seules, mais également celles des complexes et des multimères.

Ce programme a été utilisé pour modéliser la structure des protéines REC114, MEI4 et IHO1 ainsi que les complexes qu'elles forment. Etant donné que AlphaFold2 n'est pas encore installé localement et n'est pas disponible à l'IBS, j'ai utilisé le serveur public ColabFold (<https://github.com/sokrypton/ColabFold>).

## **3 Résultats**



### 3.1.1 Complexe REC114-MEI4-IHO1

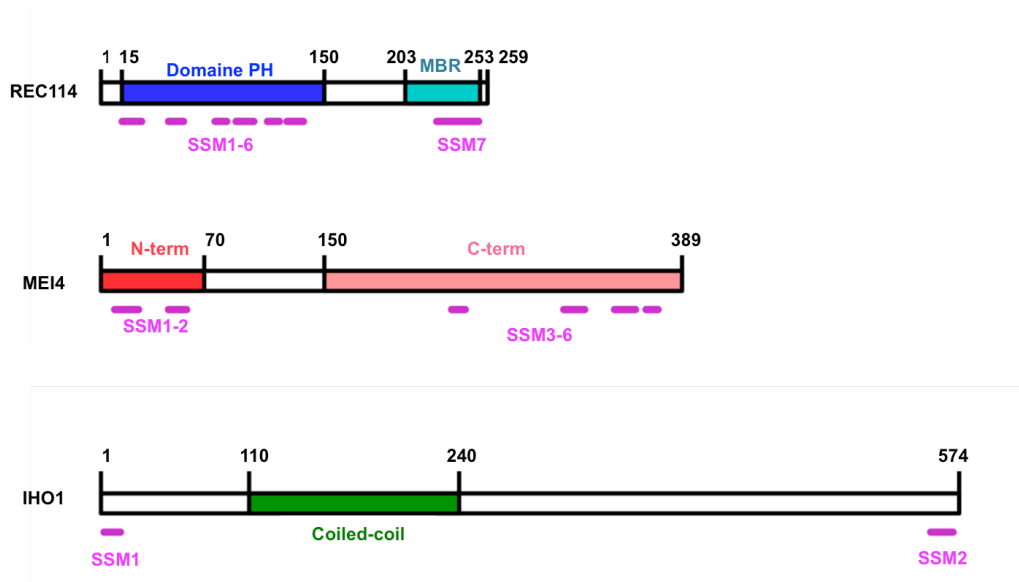
Durant cette thèse, je me suis intéressée à la caractérisation des protéines impliquées dans la formation des cassures de l'ADN lors de la recombinaison méiotique. Mon travail s'est porté principalement sur un sous-ensemble fonctionnel de trois protéines, REC114, MEI4 et IHO1. La co-localisation de ces protéines sur les axes chromosomiques et l'analyse de leurs interactions, réalisée par double hybride et co-immunoprécipitation chez la souris et la levure, suggèrent que ces protéines agissent comme un complexe et favorisent le clivage de l'ADN (Kumar et al., 2018; Li et al., 2006; Maleki et al., 2007; Stanzione et al., 2016). Néanmoins, les propriétés biochimiques, les mécanismes moléculaires de leurs interactions ne sont pas bien élucidés, en particulier chez les mammifères. À cet égard, plusieurs approches ont été menées en parallèle durant ma thèse pour apporter des informations biochimiques et moléculaires sur les protéines de souris REC114, MEI4 et IHO1 seules ou en complexes.

Le gène *rec114* code pour une protéine d'environ 28 kDa. Malgré son importance dans la méiose, cette protéine est peu conservée entre la souris et la levure. Les similarités de REC114 entre les deux espèces ne sont pas faciles à détecter, mais l'analyse bioinformatique a identifié 7 segments faiblement conservés appelés SSM (SSM pour short similarity motifs) (Kumar et al., 2010). Les 6 premiers SSM se trouvent dans la région N-terminale et il a été récemment démontré par une analyse structurale, chez la souris, qu'ils adoptent un repliement semblable au domaine homologue de la pleckstrine (domaine PH) (Kumar et al., 2018). Le dernier SSM se trouve dans la partie C-terminale et interagit avec MEI4 (Kumar et al., 2018, 2010) (Figure 15). Le domaine PH et la région C-terminale sont reliés par une portion probablement désordonnée.

MEI4 est une protéine de 45 kDa qui contient également six motifs SSM partiellement conservés (Kumar et al., 2010). Les deux SSM sont situés dans la partie N-terminale, et les autres sont repartis entre le milieu et l'extrémité C-terminale de la protéine. Les prédictions de structure secondaire indiquent que MEI4 est en grande partie structurée, mais aucun domaine connu ne lui a été attribué. Encore une fois, l'extrémité N-terminale semble être liée au reste de la protéine par un *linker* désordonné (Figure 15).

Enfin, le gène *iho1* code pour une protéine de 64 kDa. Une étude, utilisant la recherche d'homologie couplée à des alignements multiples, a montré que la protéine IHO1

de différentes espèces (champignons, plantes et mammifères) présente deux motifs similaires, l'un du côté N-terminal (SSM1) et l'autre du côté C-terminal (SSM2) (Tessé et al., 2017). La région centrale de cette protéine est prédite pour adopter une structure en *coiled-coil*, mais le reste de la protéine est prédit non-structuré (Figure 15).



**Figure 15. Représentation schématique des protéines REC114, MEI4 et IHO1.** Le gène *rec114* code pour une protéine avec un domaine d'homologie de la pleckstrine (PH) (bleu) situé dans la partie N terminale. La région C-terminale qui interagit avec MEI4 (MBR pour MEI4 Binding Region, Cyan) est indiquée. L'extrémité N-terminale (N-term, rouge) et la partie C-terminale (C-term, saumon) qui sont prédites structurées pour la protéine MEI4 sont représentées. La protéine IHO1 contient une région prédite *coiled-coil* située au centre (vert). Les motifs SSM (pour Signature Sequence Motifs) sont indiqués en rose pour chaque protéine. Les parties représentées en blanc sont probablement désordonnées.

À mon arrivée dans l'équipe, l'expression en bactérie de la protéine REC114 complète (REC114) était déjà maîtrisée. Cependant, les protéines entières MEI4 et IHO1 n'ont pas pu être obtenues en bactérie. L'interaction entre le fragment C-terminal de REC114 et le N-terminal de MEI4 a été déjà mise en évidence par notre équipe (Kumar et al., 2018). Dans le travail qui a précédé ma thèse, la construction MEI4<sub>1-127</sub> a été créée et c'est la plus longue construction qu'on a réussi à produire soluble pour cette protéine (toutes versions plus longues de MEI4 étaient insolubles). Cette construction interagit avec REC114 et elle contient la région N-terminale conservée (SSM1 et SSM2) et le *linker* prédit. La protéine IHO1 n'a pas pu être produite entièrement, mais plutôt sous forme de deux fragments consécutifs. La construction qui couvre la moitié N-terminale (résidus 1 à 281 dont la région prédite comme *coiled-coil*) a montré, dans des expériences

préliminaires, qu'elle peut interagir avec REC114-MEI4, mais la partie C-terminale n'a pas interagi.

Pour effectuer les analyses biochimiques et structurales de ces protéines, j'ai exploré la possibilité de co-exprimer les protéines entières REC114, MEI4 et IHO1 dans les cellules d'insectes. Néanmoins, la production n'a pas été fructueuse, un très faible niveau d'expression a été observé pour REC114, MEI4 et une absence totale pour IHO1. Pour cette raison, je me suis concentrée sur la caractérisation de ces protéines en utilisant les constructions exprimées en système bactérien.

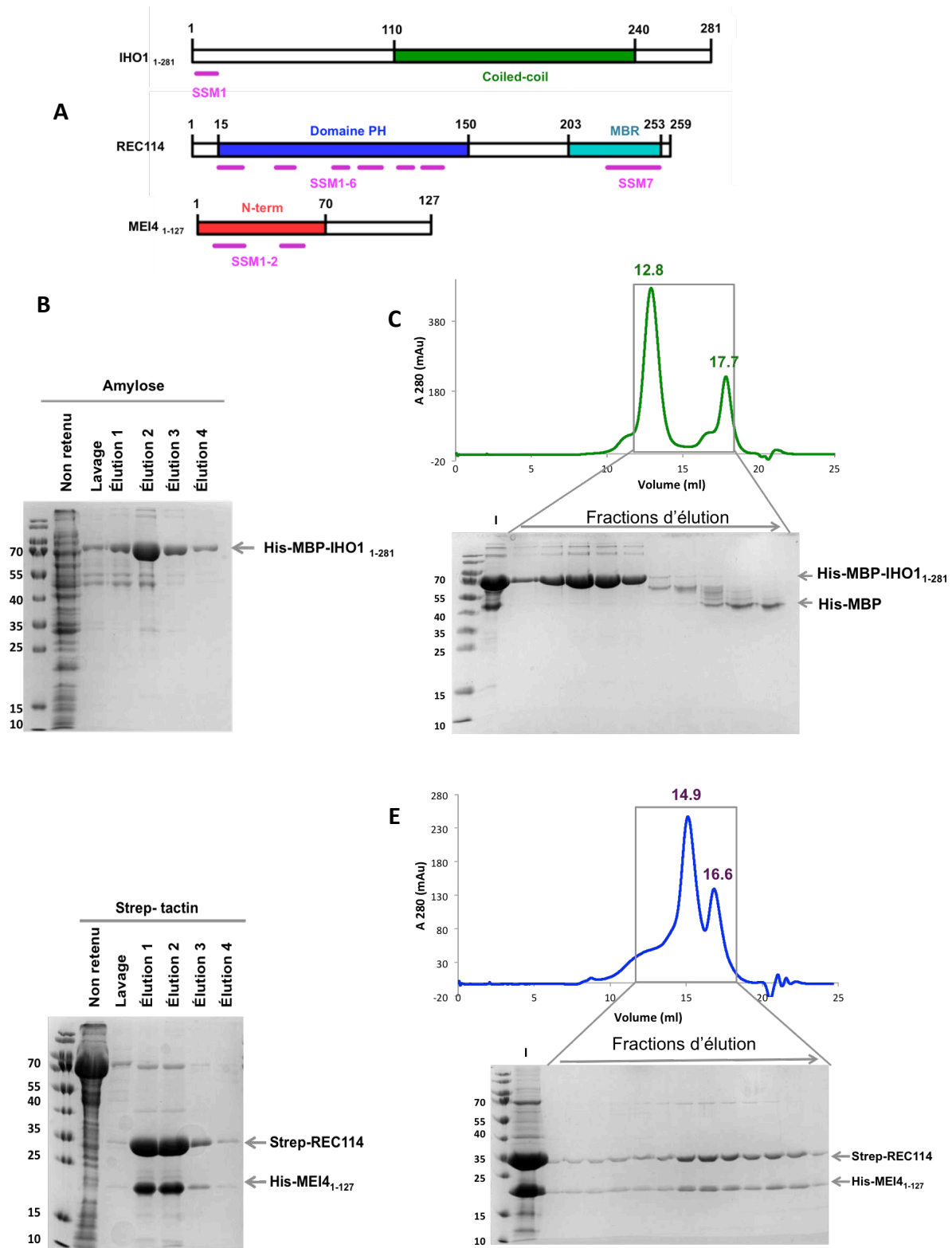
### **3.1.2 Purification du sous-complexe REC114-MEI4 et de la protéine IHO1**

Le fragment codant les 281 premiers acides aminés d'IHO1 (IHO1<sub>1-281</sub>) et les constructions REC114 et MEI4<sub>1-127</sub> ont été utilisés pour caractériser l'interaction entre le sous-complexe REC114-MEI4 et la protéine IHO1 (Figure 16 A). Avant d'analyser cette interaction, la première étape était d'établir l'expression et la purification de ces protéines.

L'étiquette MBP a été utilisée pour améliorer le rendement d'expression et la stabilité d'IHO1. En effet, sans elle, cette construction protéique se dégrade. Pour la purification du complexe REC114-MEI4, j'ai utilisé l'étiquette Strep sur REC114 qui permet habituellement d'obtenir des protéines très pures. MEI4 contenait une étiquette 6XHis N-terminale, qui n'a finalement pas été utilisée pour la purification.

La protéine His-MBP-IHO1<sub>1-281</sub> a été purifiée sur la résine amylose puis injectée en colonne de gel-filtration (Figure 16 B, C). Le profil d'élution de la chromatographie d'exclusion de taille présente un pic principal suivi d'un pic minoritaire. L'analyse par électrophorèse en condition dénaturante montre que la MBP-IHO1 est éluée dans le pic principal. Le pic minoritaire de cet échantillon correspond à la MBP seule (Figure 16 C).

Le sous-complexe Strep-REC114-His-MEI4<sub>1-127</sub> a été purifié sur la résine de Strep-tactin suivie d'une chromatographie d'exclusion de taille (Figure 16 D, E). Le complexe Strep-REC114-His-MEI4<sub>1-127</sub> a été élué en deux pics. L'analyse par SDS-PAGE montre que le sous-complexe REC114-MEI4 est présent dans les deux pics (Figure 16 E), ce qui suggère la présence de différentes espèces oligomériques.



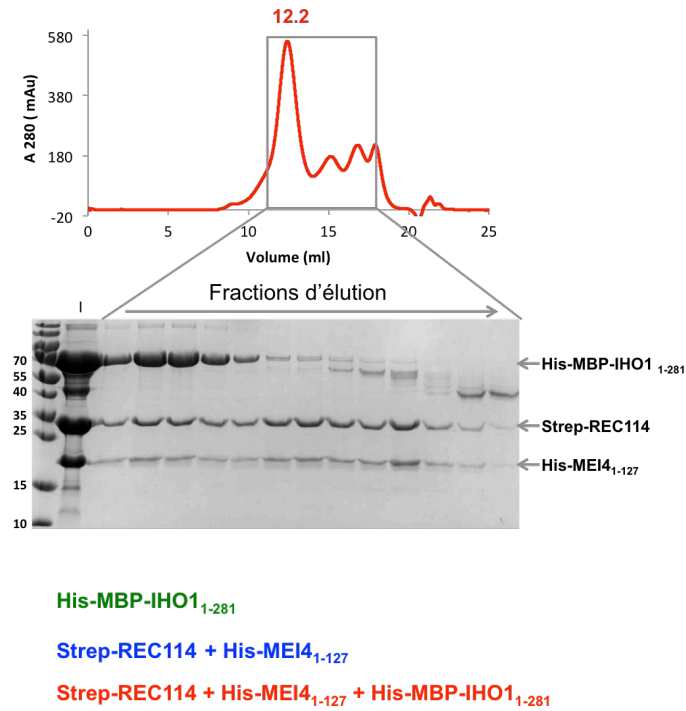
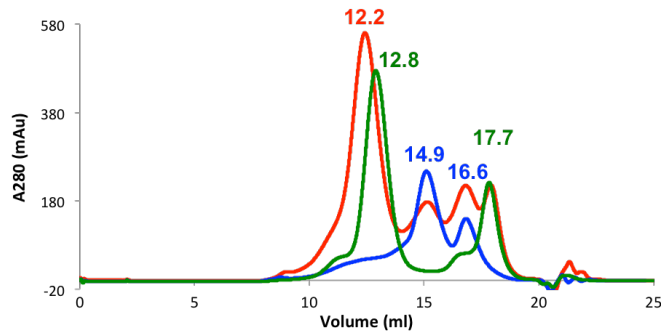
**Figure 16. Purification d'IHO<sub>1-281</sub> et de REC114-MEI4<sub>1-127</sub>** (A) Représentation schématique des constructions utilisées pour la caractérisation du complexe REC114-MEI4-IHO1. Le code couleur des différents domaines est identique à celui de la figure 15. (B) Analyse par SDS-PAGE sur un gel 15 % des fractions de la purification de His-MBP-IHO1<sub>1-281</sub> par chromatographie d'affinité sur la résine amylose. (C) Profil d'élué en chromatographie d'exclusion de taille de His-MBP-IHO1<sub>1-281</sub>. L'échantillon injecté (I) et les fractions d'élué (zone encadrée en noir) ont été analysés par SDS-PAGE. (D) Analyse par SDS-PAGE de la purification de Strep-REC114-His-MEI4<sub>1-127</sub> sur la résine Strep-tactin. (E) Purification par gel-filtration de Strep-REC114-His-MEI4<sub>1-127</sub> et analyse des fractions d'élué par SDS-PAGE. Le volume d'élué (en millilitres) sur la colonne Superose 6 (GE Healthcare) est indiqué en dessus de chaque pic.



Ces résultats montrent que les protéines REC114-MEI4<sub>1-127</sub> et IHO1<sub>1-281</sub> peuvent être produites en bactérie avec un rendement très satisfaisant. Sans la MBP, la protéine IHO1<sub>1-281</sub> se dégrade, donc l'étiquette MBP va être maintenue sur cette construction pour le reste des analyses. Le complexe REC114-MEI4 semble former différents états oligomériques, ce qui n'est pas très favorable pour une éventuelle caractérisation structurale.

### 3.1.3 Reconstitution du complexe REC114-MEI4-IHO1

Pour vérifier si REC114-MEI4 et IHO1 peuvent former un complexe, les fractions d'élution de la chromatographie d'affinité de la protéine His-MBP-IHO1<sub>1-281</sub> et du sous-complexe Strep-REC114-MEI4<sub>1-127</sub> ont été mélangées dans un ratio molaire 1 : 2 et injectées sur la colonne d'exclusion stérique (Figure 17 A). Le chromatogramme obtenu montre la présence un pic majoritaire qui se trouve décalé des pics d'élution des protéines seules (Figure 17 B). Des pics minoritaires sont également observés et sont supposables avec les pics d'élution du complexe REC114-MEI4 et de la MBP seule. L'analyse des fractions d'élution par SDS-PAGE a montré que le pic principal contient les protéines IHO1, REC114 et MEI4 qui ont été éluées ensemble. L'excès de REC114-MEI4, qui n'a pas lié IHO1, et de la MBP ont été élués dans les pics minoritaires (Figure 17 A). La co-elution des trois protéines dans le même pic suggère la présence d'une interaction entre REC114-MEI4<sub>1-127</sub> et MBP-IHO1<sub>1-281</sub>. Une interaction possible entre REC114 et IHO1 a été suggérée précédemment en utilisant la technique du double hybride (Stanzione et al., 2016), mais c'est la première fois qu'un complexe contenant les trois protéines a pu être formé *in vitro*.

**A****B**

**Figure 17. Reconstitution du complexe REC114-MEI4-IHO1.** (A) Les fractions contenant la protéine His-MBP-IHO1<sub>1-281</sub> et Strep-REC114-His-MEI4<sub>1-127</sub> purifiées sur la résine amylose et Strep-tactin respectivement (Figure 16) ont été mélangées et injectées en colonne d'exclusion de taille. L'échantillon injecté (I) et les fractions d'élutions (encadrées en noir) ont été analysés par SDS-PAGE sur un gel 15 %. (B) Superposition de profil d'élution de la protéine His-MBP-IHO1<sub>1-281</sub> (vert, Figure 16), Strep-REC114-His-MEI4<sub>1-127</sub> (bleu, Figure 16) et le mélange de ces deux derniers (rouge). Le volume d'élution (en millilitres) sur la colonne Superose 6 (GE Healthcare) est indiqué au dessus de chaque pic.

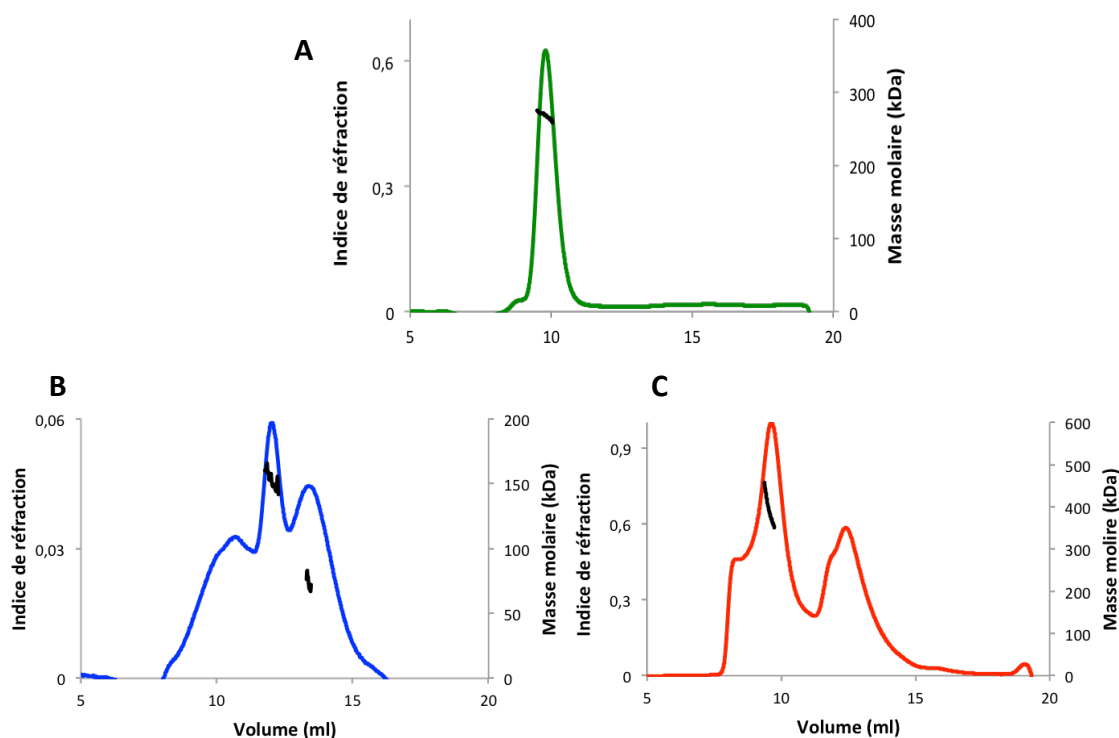
### 3.1.4 Caractérisation de l'état oligomérique du complexe REC114-MEI4-IHO1

Après l'obtention du complexe ternaire par gel-filtration, les premiers efforts se sont portés sur la détermination de son état oligomérique. À cette fin, des mesures de SEC-MALLS ont été réalisées sur la protéine IHO1, le sous-complexe REC114-MEI4 et sur le mélange contenant ces deux échantillons.

Le profil d'élution par la filtration sur gel de la protéine MBP-IHO1<sub>1-281</sub> montre la présence d'un seul pic. Un poids moléculaire d'environ 274 kDa a été calculé à partir de la lumière diffusée et de l'indice de réfraction des espèces présentes dans ce pic (Figure 18 A). Étant donné que la masse théorique pour cette protéine en fusion avec la MBP est de 75 kDa, ce résultat est compatible avec la présence d'un tétramère d'IHO1 en solution.

Comme vu précédemment, le complexe Strep-REC114-MEI4<sub>1-127</sub> est élué dans deux pics majeurs. La masse de l'échantillon élué dans le premier pic est d'environ 150 kDa et la masse des espèces présentes dans le deuxième pic est réduite de moitié, soit d'environ 75 kDa (Figure 18 B). La protéine Strep-REC114 et la protéine MEI4<sub>1-127</sub> ont des masses de 29,5 kDa et 17,5 kDa respectivement. La comparaison des masses mesurées aux valeurs théoriques suggère que le deuxième pic correspondrait à un complexe 2REC114-1MEI4 (masse théorique 76,5 kDa). Le premier pic serait un dimère du complexe élué sous le pic 2, et correspondrait à une stœchiométrie de 4REC114-2MEI4 (152,3 kDa). Ce résultat suggère que le complexe REC114-MEI4<sub>1-127</sub> est un hétérotrimère (2REC114-1MEI4) qui se dimérise. Cette dimérisation est probablement issue de l'auto-association de REC114 via sa partie N-terminale. En effet, il a été observé par notre équipe que le domaine PH de REC114 peut se dimériser et il présente un échange rapide entre la forme monomérique et la forme dimérique (Kumar et al., 2018).

La forme tétramérique d'IHO1 et hétérotrimérique du complexe REC114-MEI4 ont également été observées récemment pour les protéines de la levure *S. cerevisiae* Mer2 et Rec114-Mei4, orthologues des protéines murines respectivement (Claeys Bouuaert et al., 2021a; Rousová et al., 2021). Ceci suggère que les propriétés biochimiques et l'organisation de ces protéines sont conservées.



**Figure 18. Caractérisation par la technique de SEC-MALLS.** Chromatogramme d'élution sur une colonne de chromatographie d'exclusion de taille Superdex 200 Increase. La courbe verte (A), bleue (B) et rouge (C) représente l'indice de réfraction en fonction du volume d'élution de His-MBP-IHO1<sub>1-281</sub>, Strep-REC114-His-MEI4<sub>1-127</sub> et le mélange de ces deux derniers respectivement. La masse moléculaire calculée (en kDa) sous les pics d'élution est représentée en trait noir en fonction du volume d'élution.

Les échantillons His-MBP-IHO1<sub>1-281</sub> et Strep-REC114-MEI4<sub>1-127</sub> ont été mélangés puis injectés en SEC-MALLS pour déterminer l'état oligomérique du complexe formé par ces protéines (Figure 18 C). Cependant, la masse calculée est hétérogène et elle décroît sous le pic principal. L'extrémité avant du pic est constituée de composants qui ont une masse d'environ 450 kDa, ce qui correspondrait à un complexe constitué de 4IHO1 (300 kDa) + 4REC114-2MEI4 (150 kDa). Par contre, la fin du pic contient des protéines de masse molaire de 375 kDa, ce qui correspondrait à un complexe de 4IHO1 (300 kDa) + 2REC114-1MEI4 (75kDa). Ce résultat suggère que l'échantillon protéique du complexe n'est pas très homogène et il présente un mélange de complexes avec différents états oligomériques sous le même pic d'élution. Ceci rend difficile la caractérisation exacte de sa stœchiométrie.

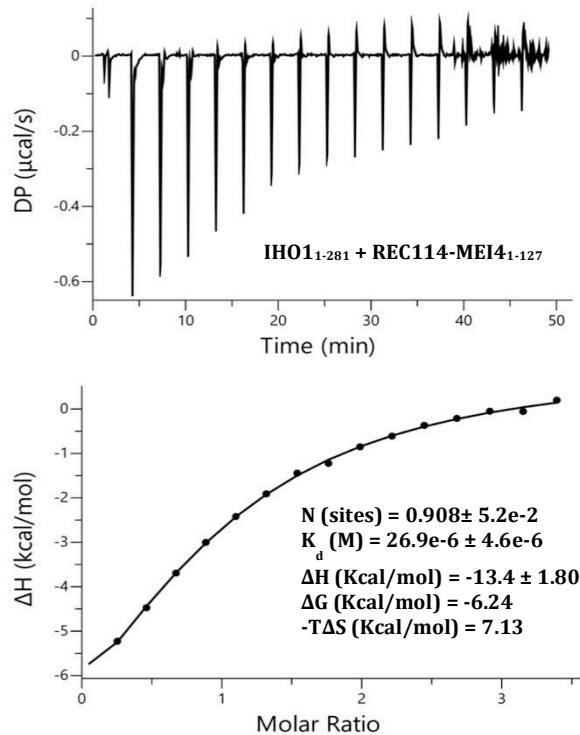
La caractérisation du complexe a été également explorée avec la spectrométrie de masse native, mais les données obtenues n'étaient pas exploitables. Cela est probablement dû à un échange insuffisant du tampon du complexe en acétate d'ammonium, malgré deux étapes de filtration sur gel.

Globalement, ces résultats indiquent que la protéine MBP-IHO1<sub>1-281</sub> forme des tétramères. Comme la MBP est connue pour être monomérique, il est très probable que la tétramérisation observée soit médiée par IHO1. Cette technique confirme l'hétérogénéité oligomérique du sous-complexe REC114-MEI4. Malheureusement, le complexe REC114-MEI4-IHO1 semble être trop hétérogène pour caractériser sa stœchiométrie.

### **3.1.5 Caractérisation de l'interaction entre IHO1 et REC114-MEI4**

#### ***3.1.5.1 Calorimétrie par titrage isotherme (ITC)***

L'étape suivante dans la caractérisation biophysique du complexe était de mesurer l'affinité de l'interaction entre REC114-MEI4 et IHO1. Le complexe REC114-MEI4<sub>1-127</sub>, concentré à 30  $\mu\text{M}$ , a été titré par 500  $\mu\text{M}$  de His-MBP-IHO1<sub>1-281</sub>. Des concentrations plus élevées de cette construction d'IHO1 n'ont pas pu être testées, car cette construction précipite au-dessus de cette concentration. Les données obtenues ont montré la présence d'une interaction entre ces deux échantillons. Bien que le plateau de saturation ne soit pas atteint, l'intégration de ces données brutes a permis de calculer une constante de dissociation ( $K_d$ ) d'environ 26  $\mu\text{M}$  entre MBP-IHO1<sub>1-281</sub> et REC114-MEI4<sub>1-127</sub> (Figure 19). Ce résultat confirme la formation d'un complexe entre ces deux échantillons comme il a été observé par la chromatographie d'exclusion de taille. Néanmoins, cette valeur d'affinité semble indiquer que l'interaction entre la construction d'IHO1<sub>1-281</sub> et REC114-MEI4<sub>1-127</sub> est plutôt faible *in vitro*.



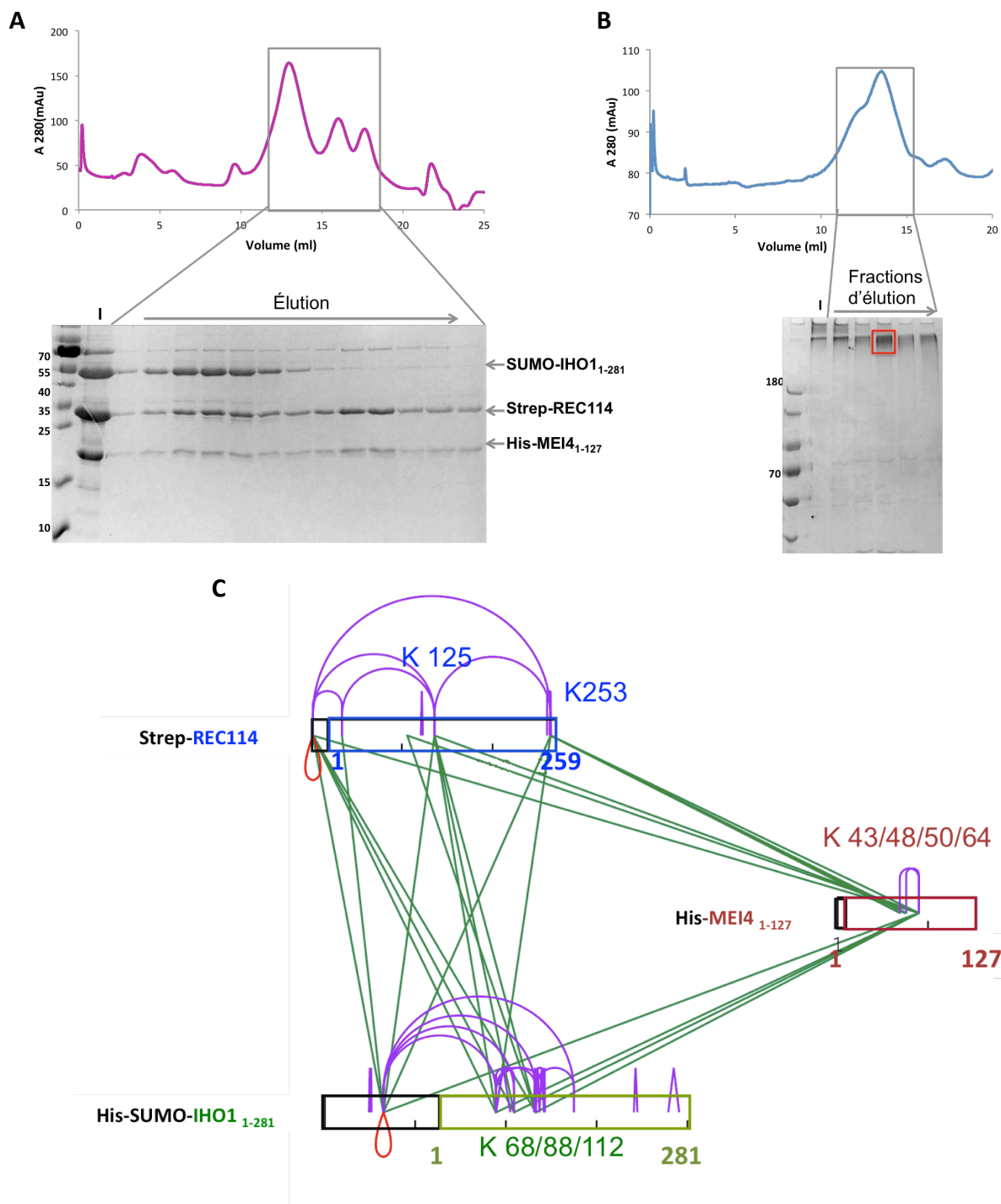
**Figure 19. Mesures d'affinité de l'interaction entre REC114-MEI4<sub>1-127</sub> et MBP-IHO1<sub>1-281</sub> par ITC.** En haut, données brutes montrant les variations de chaleur (puissance électrique différentielle) mesurées entre la cellule d'échantillon et la cellule de référence pour chaque injection de la protéine IHO1 dans le complexe REC114-MEI4 en fonction du temps. En bas, courbe d'interaction après l'intégration des données brutes en fonction du ratio molaire MBP-IHO1<sub>1-281</sub>/REC114-MEI4<sub>1-127</sub>. N : stoechiométrie, K<sub>d</sub> : constante de dissociation, ΔH : enthalpie, ΔG : énergie libre et ΔS : entropie.

### 3.1.5.2 Pontage chimique couplé à la spectrométrie de masse

Pour déterminer l'interface de l'interaction entre REC114-MEI4<sub>1-127</sub> et IHO1<sub>1-281</sub>, une analyse par pontage chimique couplé à la spectrométrie de masse a été réalisée. Les premières expériences de pontage chimique ont été réalisées en utilisant la construction IHO1<sub>1-281</sub> exprimée en fusion avec la MBP. Cependant, le résultat de ces expériences n'était pas satisfaisant, indiquant une forte agrégation du complexe malgré plusieurs tentatives d'optimisation en modifiant le temps de réaction ou la concentration de l'agent réticulant. Ce résultat est probablement dû au poids moléculaire assez important de la MBP qui est susceptible d'apporter de la flexibilité et d'induire des pontages chimiques non spécifiques provoquant l'agrégation du complexe. Une étape de clivage à la protéase TEV devrait être ajoutée à la purification de MBP-IHO1<sub>1-281</sub>. Néanmoins, en absence de la MBP, cette construction de la protéine IHO1 est moins stable et a tendance à se dégrader. Pour pallier ce problème, nous avons décidé d'exprimer IHO1<sub>1-281</sub> en fusion avec la protéine SUMO, connue pour améliorer la solubilité des protéines d'intérêt et ayant un poids moléculaire

moins important que la MBP. La protéine IHO1<sub>1-281</sub> en fusion avec His-SUMO a été exprimée et purifiée sur la résine Ni<sup>2+</sup> puis utilisée pour former le complexe His-SUMO-IHO1<sub>1-281</sub>-REC114-MEI4<sub>1-127</sub> par la chromatographie d'exclusions stérique (Figure 20 A). Les fractions contenant le complexe ternaire ont été pontées chimiquement avec un agent de pontage réactif aux amines, BS3, et passées sur la filtration sur gel (Figure 20 B). Les fractions d'élution du produit de pontage ont été analysées par SDS-PAGE et la bande du gel de la fraction du pic principal a été analysée par la plateforme de spectrométrie de masse (MFPL, Max Perutz Labs, Vienne) pour déterminer les résidus impliqués dans le pontage chimique. Les protéines His-SUMO-IHO1<sub>1-281</sub>, Strep-REC114 et His-MEI4<sub>1-127</sub> contiennent plusieurs lysines réparties sur la longueur de leurs séquences protéiques. Les données obtenues ont détecté la présence de certains pontages liant les trois protéines ensemble (Figure 20 C). C'est notamment le cas des lysines 68/88/112 d'IHO1 qui sont pontées d'une part avec la lysine 65 de MEI4<sub>1-127</sub>, et d'autre part avec la lysine 125 de REC114. Ce résultat indique que les lysines, impliquées dans le pontage, se trouveraient proches dans l'espace avec une distance d'environ 26 à 30 Å au sein du complexe (distance entre les C $\alpha$  de lysines pontées par BS3). De plus, la protéine REC114 est pontée via sa lysine C-terminale (K253) avec les lysines N-terminales 43/48/50/64 de MEI4<sub>1-127</sub>. Ceci est en accord avec les études, réalisées *in vivo* et *in vitro*, qui ont montré l'interaction entre la partie N-terminale de MEI4 et la partie C-terminale de REC114 (Kumar et al., 2018, 2010). Certains pontages ont également été détectés entre les étiquettes de purification et les protéines d'intérêt. Ces pontages seraient non spécifiques et ils peuvent être expliqués par la flexibilité de ces étiquettes.

Bien qu'elle ait montré la présence de plusieurs liens de pontage chimique entre REC114-MEI4 et IHO1, cette expérience ne nous a pas permis de comprendre comment interagissent ces protéines.



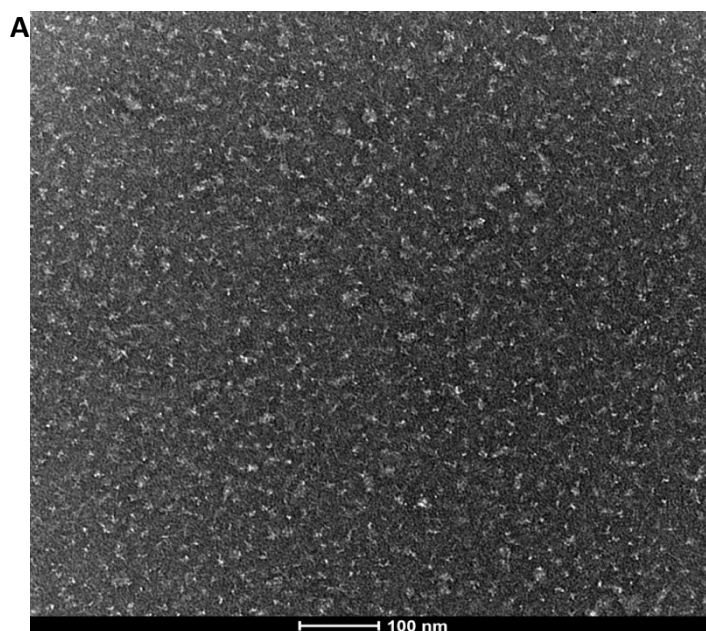
**Figure 20. Reconstitution et pontage chimique du complexe His-SUMO-IHO1<sub>1-281</sub>-REC114-MEI4<sub>1-127</sub>.** (A) Les échantillons protéiques His-SUMO-IHO1<sub>1-281</sub> et Strep-REC114-His-MEI4<sub>1-127</sub> ont été mélangés dans un ratio molaire 1 : 2 et injectés en gel filtration sur colonne Superose 6. Le profil d'élution de ce complexe est représenté en rose, les protéines injectées sur la colonne (I) et les fractions d'élution encadrées en noir ont été analysées par SDS-PAGE. (B) Profil de chromatographie d'exclusion de taille pour le complexe SUMO-IHO1-MEI4-REC114 après le pontage chimique. Les fractions d'élution ont été analysées sur gel polyacrylamide 8%. La fraction encadrée en rouge a été utilisée pour la spectrométrie de masse couplée au pontage chimique. (C) Les trois protéines analysées sont schématisées en rectangles IHO1 (vert), REC114 (bleu) et MEI4 (rouge foncé). La partie représentée en noir pour chaque rectangle correspond à l'étiquette de la purification associée à la protéine. Les lysines les plus impliquées dans la réaction de pontage sont indiquées pour chaque protéine. Les pontages inter-protéines (entre les lysines des protéines différentes) sont représentés en lignes vertes. Les pontages intramoléculeux (lysines au sein de la même protéine ou au sein d'un homomultimère) sont visualisés par des courbes violettes. Les pontages représentés en rouge correspondent à des liens au sein d'un homomultimère.



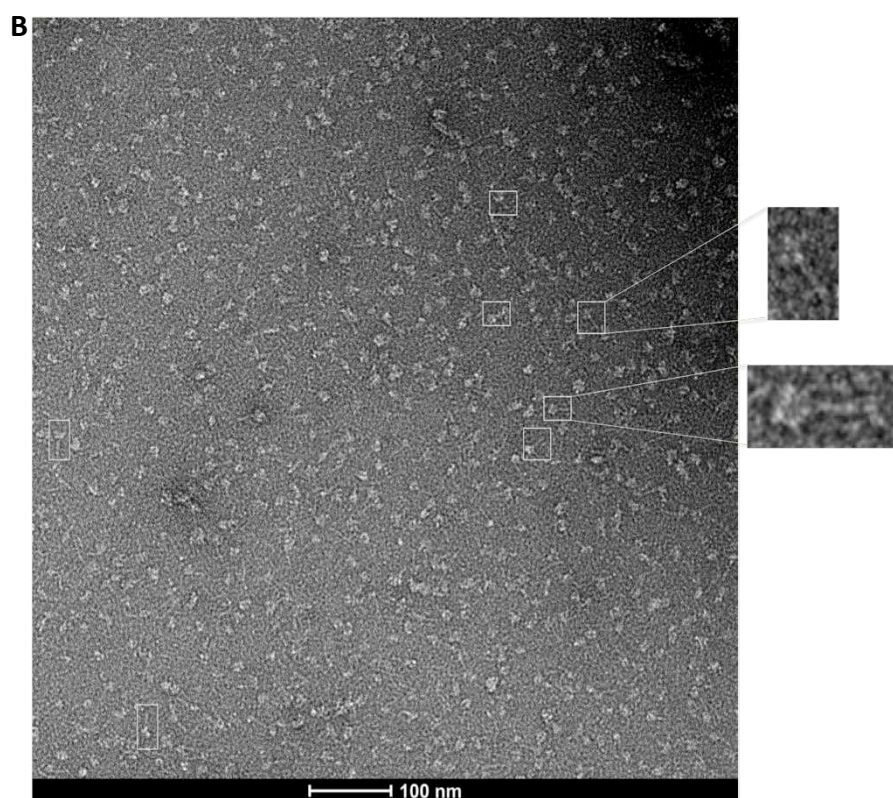
### 3.1.5.3 Coloration négative

Le poids moléculaire d'environ 400 kDa mesuré par la technique SEC-MALLS fait du complexe MBP-IHO1<sub>1-281</sub>-REC114-MEI4<sub>1-127</sub> un échantillon adapté à la caractérisation structurale par la Cryo-EM. Cette dernière est une technique de choix pour l'étude des grands complexes protéiques. Néanmoins, en plus de la taille des macromolécules, l'utilisation d'un échantillon homogène est une condition cruciale pour atteindre la résolution atomique. À cet effet, le complexe ternaire a été analysé d'abord par la coloration négative pour évaluer son homogénéité. La fraction principale du pic du complexe His-MBP-IHO1<sub>1-281</sub>-Strep-REC114-MEI4<sub>1-127</sub> éluée de la filtration sur gel a été préparée pour être observé en microscopie électronique en utilisant l'acétate d'uranyle comme colorant. Cependant, les micrographes obtenus (dont un est présenté dans Figure 21 A) ont montré que l'échantillon est très hétérogène, il présente un mélange d'agrégats et de petites particules. Étant donné la faible affinité de l'interaction du complexe, cette hétérogénéité est probablement provoquée par la dissociation du complexe lors de la préparation des grilles. Pour tester cette hypothèse, plusieurs acquisitions en coloration négative ont été effectuées avec le complexe réticulé His-SUMO-IHO1<sub>1-281</sub>-Strep-REC114-MEI4<sub>1-127</sub>. Les fractions du pic d'éluion après le pontage chimique avec le BS3 ont été ensuite visualisées par microscopie électronique en coloration négative. Les images obtenues montrent que la qualité de l'échantillon s'est améliorée comparée aux précédentes expériences. La présence d'une forme allongée liée à une forme globulaire a été constatée avec certaines particules (sous forme d'une clé) (Figure 21 B), cependant l'échantillon est toujours hétérogène. Cette hétérogénéité est peut-être provoquée par la présence des différents états oligomériques observés par la technique de SEC-MALLS. Dans certains cas, l'hétérogénéité des particules peut aussi être causée par les solutions de coloration utilisées lors de la préparation de l'échantillon pour la coloration négative. Le pH acide de ces colorants peut provoquer la dissociation ou l'agrégation de l'échantillon. Pour cette raison, nous avons essayé de visualiser le complexe REC114-MEI4-IHO1 directement par cryo-microscopie électronique (Cryo-EM) en utilisant le microscope Glacios de l'EMBL à Grenoble, mais l'homogénéité ne s'est pas améliorée. Pour l'instant, la qualité de l'échantillon ne permet pas une étude structurale à haute résolution par cryo-microscopie électronique.

Sans pontage chimique



Avec pontage chimique



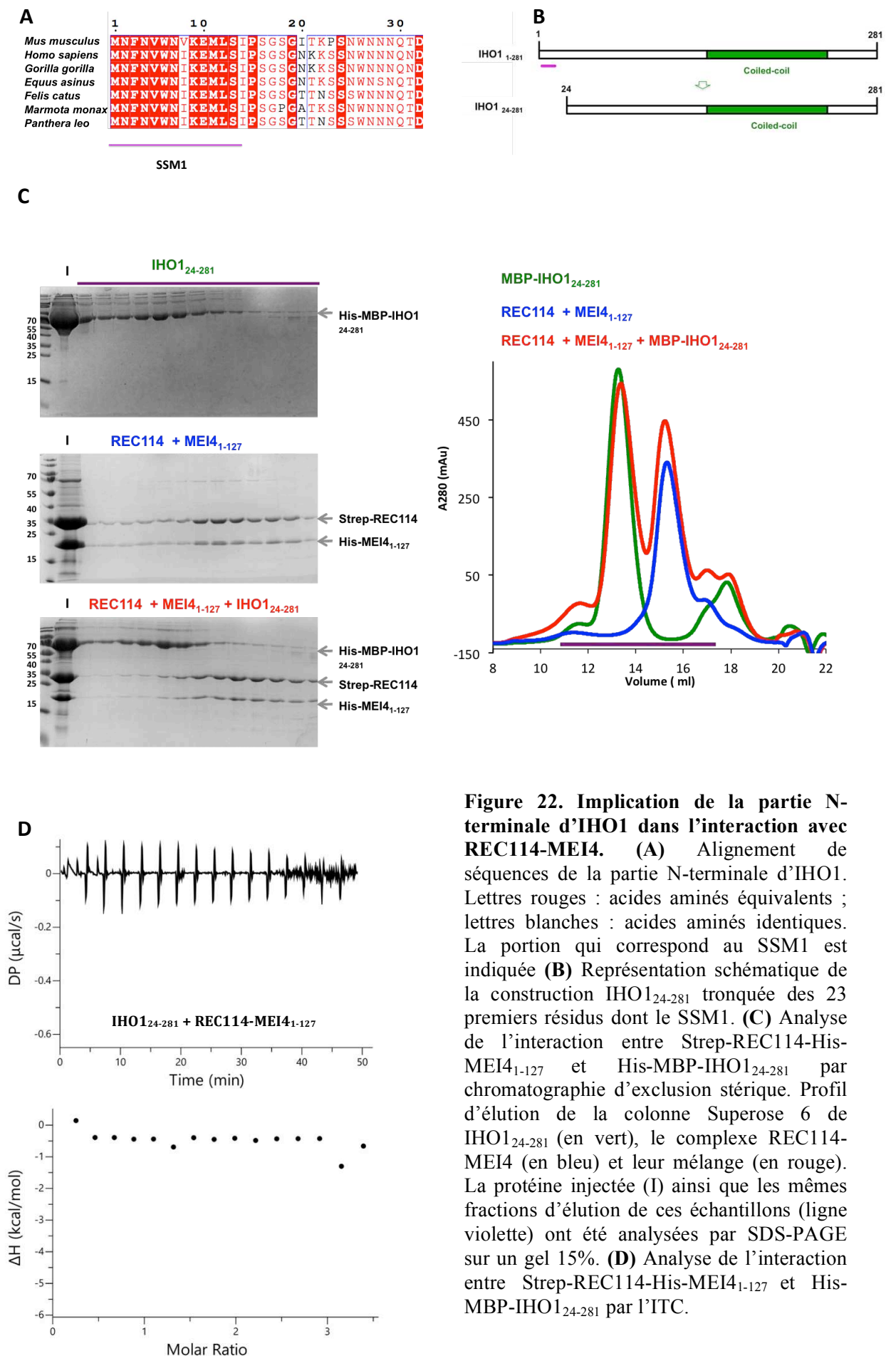
**Figure 21. Visualisation par microscopie électronique en coloration négative du complexe IH01-REC114-MEI4. (A) Micrographe obtenu par la coloration négative du complexe avant le pontage chimique (B) et après le pontage chimique avec BS3. Les particules observées sous forme de clés sont agrandies.**

### 3.1.6 Implication de la partie N-terminale d'IHO1 dans l'interaction avec REC114-MEI4

Comme l'analyse structurale du complexe entier s'est avérée très difficile, et compte tenu des connaissances actuelles limitées sur ces protéines, j'ai décidé de concentrer mes efforts sur les sous-complexes individuels afin d'améliorer notre compréhension de la façon dont ces protéines interagissent. Comme indiqué précédemment, la protéine IHO1 contient deux motifs SSM, l'un du côté N-terminal et l'autre du côté C-terminal. Il a été prédit chez le champignon *Sordaria macrospora* que le motif N-terminal de l'orthologue d'IHO1 est nécessaire pour l'interaction avec l'orthologue de REC114 (Tessé et al., 2017).

Le motif N-terminal d'IHO1 correspond aux 14 premiers résidus (Figure 22 A) (Tessé et al., 2017). Pour déterminer chez la souris l'implication de ce motif dans l'interaction entre IHO1 et REC114-MEI4, une construction tronquée des 23 premiers résidus a été créée (MBP-IHO1<sub>24-281</sub>) (Figure 22 B). L'impact de cette délétion sur la formation du complexe ternaire IHO1-REC114-MEI4 a été analysé par la chromatographie d'exclusion stérique. La protéine MBP-IHO1<sub>24-281</sub> et le complexe Strep-REC114-His-MEI4<sub>1-127</sub> ont été purifiés par la chromatographie d'affinité, puis mélangés dans un ratio molaire 1 : 2 et injectés en gel-filtration. Contrairement au mélange contenant la protéine IHO1<sub>1-281</sub>, le chromatogramme obtenu pour le mélange MBP-IHO1<sub>24-281</sub> et Strep-REC114-His-MEI4<sub>1-127</sub> est parfaitement superposable avec les chromatogrammes des protéines individuelles. L'analyse des fractions d'élution par SDS-PAGE n'a montré aucune co-purification du complexe ternaire, les protéines ont été éluées individuellement dans des pics séparés (Figure 22 C). Ce résultat indique que la protéine IHO1 tronquée des 23 premiers résidus est incapable de se lier au complexe REC114-MEI4.

Pour le confirmer, des mesures ITC ont été réalisées. La protéine IHO1<sub>24-281</sub> a été injectée dans le sous-complexe REC114-MEI4<sub>1-127</sub>, de la même manière que pour la construction IHO1<sub>1-281</sub>. Les données ITC mesurées n'ont révélé aucune interaction entre les deux échantillons. Cela indique que le motif N-terminal de la protéine IHO1 de souris est requis pour l'interaction avec REC114-MEI4 (Figure 22 D).



**Figure 22. Implication de la partie N-terminale d'IHO1 dans l'interaction avec REC114-MEI4.** (A) Alignement de séquences de la partie N-terminale d'IHO1. Lettres rouges : acides aminés équivalents ; lettres blanches : acides aminés identiques. La portion qui correspond au SSM1 est indiquée (B) Représentation schématique de la construction IHO1<sub>24-281</sub> tronquée des 23 premiers résidus dont le SSM1. (C) Analyse de l'interaction entre Strep-REC114-His-MEI4<sub>1-127</sub> et His-MBP-IHO1<sub>24-281</sub> par chromatographie d'exclusion stérique. Profil d'éluion de la colonne Superose 6 de IHO1<sub>24-281</sub> (en vert), le complexe REC114-MEI4 (en bleu) et leur mélange (en rouge). La protéine injectée (I) ainsi que les mêmes fractions d'éluion de ces échantillons (ligne violette) ont été analysées par SDS-PAGE sur un gel 15%. (D) Analyse de l'interaction entre Strep-REC114-His-MEI4<sub>1-127</sub> et His-MBP-IHO1<sub>24-281</sub> par l'ITC.

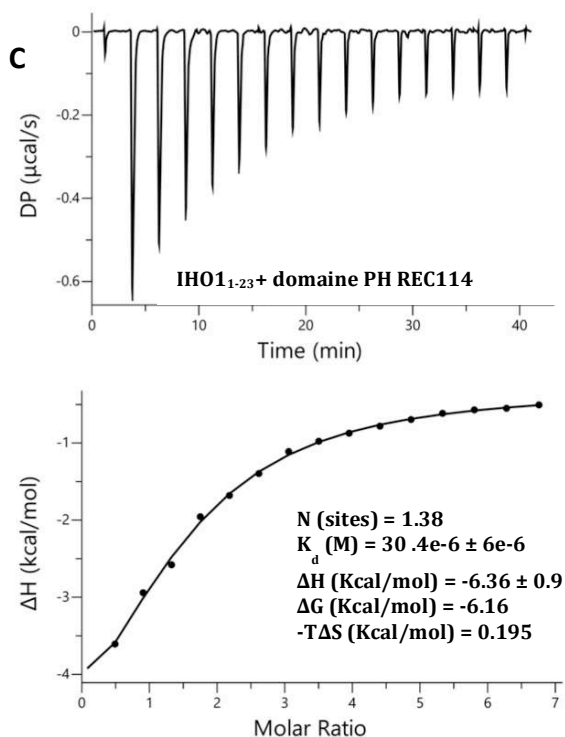
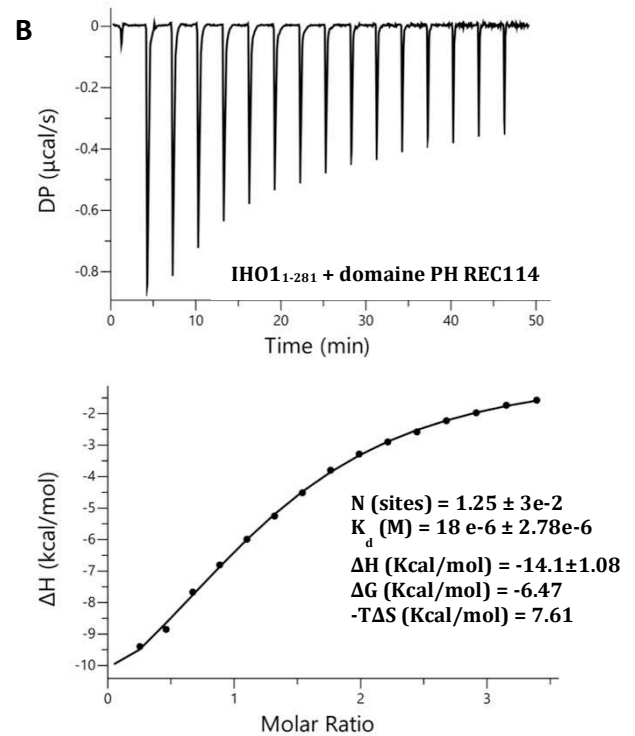
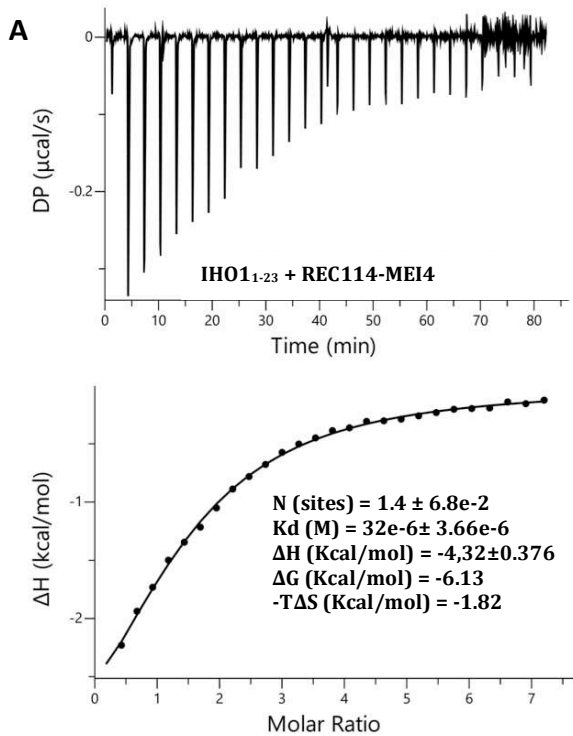
### 3.1.7 Caractérisation de l'interaction entre la partie N-terminale d'IHO1 et REC114

Ensuite, pour déterminer si la partie N-terminale d'IHO1 est suffisante pour lier le complexe REC114-MEI4, une construction couvrant les 24 premiers résidus d'IHO1 a été clonée et exprimée en fusion avec la MBP. L'interaction entre ce fragment d'IHO1 et le sous-complexe REC114-MEI4 a été évaluée par des mesures d'ITC. La construction His-MBP-IHO1<sub>1-23</sub> a pu être concentrée jusqu'à 1 mM et elle a été titrée dans 30  $\mu$ M du complexe REC114-MEI4. Les données obtenues ont montré que cette construction lie le complexe REC114-MEI4 avec une constante de dissociation ( $K_d$ ) d'environ 30  $\mu$ M (Figure 23 A). Cette valeur d'affinité est comparable à celle d'IHO1<sub>1-281</sub> avec le complexe REC114-MEI4 (26  $\mu$ M). Ce qui montre que l'interaction avec REC114-MEI4 est maintenue en présence du fragment N-terminal d'IHO1 uniquement.

L'investigation s'est portée ensuite sur la partie N-terminale de REC114 pour déterminer son rôle dans l'interaction avec IHO1. Des mesures ITC ont été réalisées, en absence de MEI4 et de la partie C-terminale de REC114, entre la construction codante le domaine PH de la protéine REC114 (résidus 1 à 159) et les deux constructions de MBP-IHO1, la construction la plus longue IHO1<sub>1-281</sub> ou celle contenant uniquement les 24 premiers résidus (IHO1<sub>1-23</sub>). Les résultats de ces analyses ont permis de montrer que le domaine PH se lie de manière équivalente à ces deux fragments d'IHO1. Des constantes de dissociation d'environ 20  $\mu$ M et 30  $\mu$ M, ont été mesurées pour l'interaction entre le domaine PH et le grand ou le petit fragment d'IHO1 respectivement (Figure 23 B, C). Ces valeurs d'affinités sont également comparables à celle, en présence de MEI4 et de la partie C-terminale REC114, entre IHO1<sub>1-281</sub> et REC114-MEI4.

L'ensemble de ces analyses montre que l'absence de MEI4 n'affecte pas l'interaction entre REC114 et IHO1. Elles prouvent également que la partie N-terminale d'IHO1 et le domaine PH de REC114 sont les principaux acteurs de l'interaction entre IHO1 et REC114-MEI4. L'interaction de REC114, via sa partie C-terminale, avec la partie N-terminale de MEI4 permettrait la formation du complexe ternaire REC114-MEI4-IHO1.





**Figure 23. Caractérisation de l'interaction entre IHO1 et REC114.**

(A) Titrage de 30  $\mu\text{M}$  Strep-REC114-His-MEI4<sub>1-127</sub> par 1mM de His-MBP-IHO1<sub>1-23</sub>.

(B) Titrage de 30  $\mu\text{M}$  Strep REC114<sub>1-159</sub> (domaine PH) par 500  $\mu\text{M}$  de His-MBP-IHO1<sub>1-281</sub>.

(C) Titrage de 30  $\mu\text{M}$  Strep REC114<sub>1-159</sub> (domaine PH) par 1 mM de His-MBP-IHO1<sub>1-23</sub>.

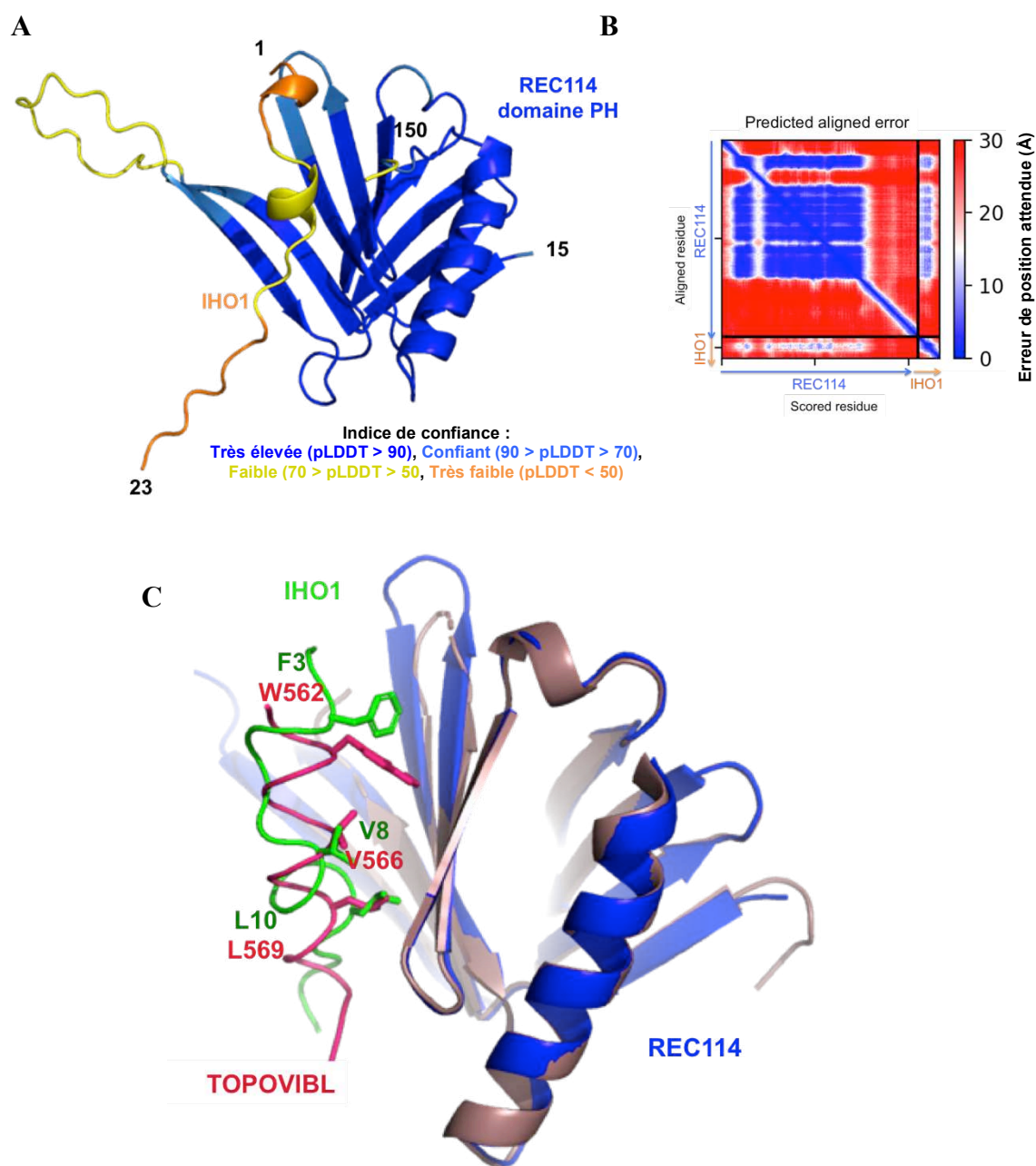
En haut, données brutes montrant les variations de chaleur (puissance électrique différentielle) mesurées entre la cellule d'échantillon et la cellule de référence pour chaque injection d'IHO1 dans le complexe REC114-MEI4 ou dans le domaine PH en fonction temps. En bas, courbe d'interaction après l'intégration des données brutes en fonction du ratio molaire His-MBP-IHO1<sub>1-23</sub> (ou His-MBP-IHO1<sub>1-23</sub>) et REC114<sub>1-159</sub> (ou Strep-REC114-His-MEI4<sub>1-127</sub>).

N : stœchiométrie,  $K_d$  : constante de dissociation,  $\Delta H$  : enthalpie,  $\Delta G$  : énergie libre et  $\Delta S$  : entropie.

### 3.1.8 Modèle AlphaFold2 du complexe IHO1-REC114

Au moment où je travaillais sur l'analyse structurale du complexe REC114-IHO1, le logiciel d'intelligence artificielle AlphaFold2 (Jumper et al., 2021) a été développé pour permettre la prédiction des structures 3D des protéines ainsi que des complexes protéiques à partir de la séquence en acides aminés. La prédiction de la structure 3D de nombreuses protéines, y compris celles impliquées dans l'initiation de la recombinaison méiotique, sont accessibles dans la base de données d'AlphaFold2 chez EBI (European Bioinformatics Institute, <https://alphafold.ebi.ac.uk/>). Ces modèles prédits confirment les prédictions précédentes que la protéine MEI4 (Q8BRM6) est principalement structurée, tandis qu'IHO1 (Q6PDM4) ne présente qu'une partie centrale structurée. Le modèle AlphaFold2 suggère qu'en plus du domaine N-terminal de REC114 (Q9CWH4) (domaine PH dont la structure est déjà résolue), la partie C-terminale de REC114, qui interagit avec MEI4, est également structurée. Les modèles structuraux des protéines entières seront présentés dans les prochaines sections.

Grâce à cette avancée majeure, l'interaction entre REC114 et IHO1 a pu être modélisée. Le modèle AlphaFold2 prédit, avec un faible indice de confiance, la structure du peptide IHO1 (Figure 24 A). Ce modèle suggère, avec une probabilité d'erreur relativement moyenne, qu'IHO1 se positionne sur la surface hydrophobe du domaine PH (Figure 24 A, B). Cet emplacement est similaire à celui du peptide de TOPOVIBL et la partie C-terminale du peptide ANKRD31 (Figure 8 et Figure 9). En effet, le domaine PH de REC114 est la cible de deux autres protéines méiotiques, la protéine TOPOVIBL et ANKRD31 qui interagiraient de manière mutuellement exclusive avec REC114 (Boekhout et al., 2019; Nore et al., 2022). La comparaison de la structure du complexe REC114-TOPOVIBL résolue par notre équipe au modèle AlphaFold2 du complexe IHO1-REC114 montre que les résidus F3, V8 et L10 d'IHO1 sont situés dans des positions similaires aux résidus W562, V566 et L596 de TOPOVIBL (Figure 24 B). Ces derniers sont importants pour l'interaction entre TOPOVIBL et REC114 et leurs mutations empêchent la formation de ce complexe. Toutefois, le résidu W572 de TOPOVIBL est important pour lier REC114, mais il ne semble pas être conservé dans IHO1 (Nore et al., 2022).



**Figure 24. Modèle structural du complexe REC114 et IHO1.** (A) Modélisation de l'interaction entre le domaine PH de REC114 et le peptide N-terminal d'IHO1 avec AlphaFold2 (Jumper et al., 2021). Les deux protéines sont colorées en fonction de la valeur pLDDT (pour predicted local distance difference test) qui correspond au score de confiance de la prédiction de chaque résidu. AlphaFold2 considère que le degré de confiance est bon pour une valeur pDDLT > 70 et faible pour une valeur de pDDLT < 70. (B) Graphique indiquant l'erreur sur la distance attendue (en Ångströms) dans la prédiction de la position du résidu X par rapport au résidu Y, une couleur bleue à la position (x, y) indique que la probabilité d'erreur dans la prédiction est faible tandis qu'une couleur rouge traduit une probabilité élevée. (C) Superposition de la structure du complexe REC114<sub>16-148</sub>-TOPOVIBL<sub>559-576</sub> (marron et magenta respectivement (Nore et al., 2022)) avec le modèle AlphaFold2 du complexe REC114-IHO1 (bleu et vert respectivement). Les résidus importants à l'interaction entre REC114 et TOPOVIBL qui occupent des positions similaires sur le peptide IHO1 sont montrés. Le résidu W572 de TOPOVIBL n'est pas montré.

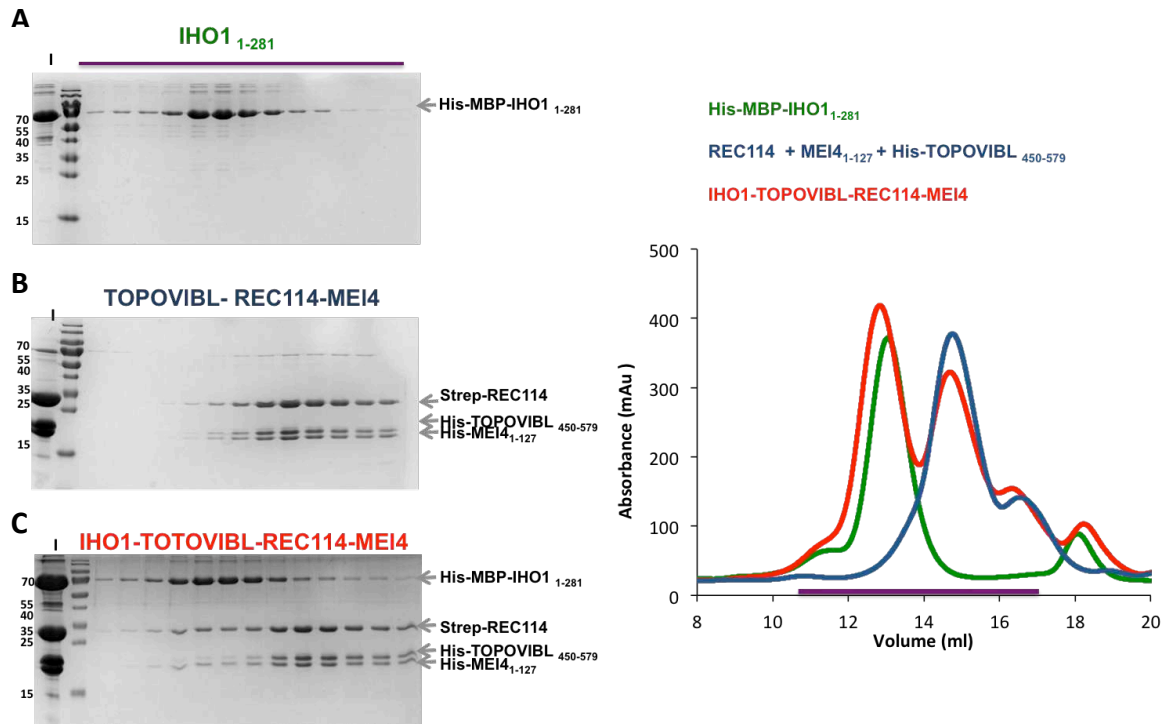


### 3.1.9 Validation du modèle d'interaction entre REC114 et IHO1

Pour confirmer l'interface d'interaction entre IHO1 et REC114 déterminée par AlphaFold2, nous avons procédé à la vérification du modèle structural par des expériences *in vitro*. À cette fin, nous avons décidé d'analyser l'effet de la liaison de la protéine TOPOVIBL ou ANKRD31 avec REC114 sur son interaction avec IHO1.

#### 3.1.9.1 REC114-MEI4-TOPOVIBL et IHO1

L'effet de la protéine TOPOVIBL sur l'interaction entre REC114 et IHO1 a été analysé par chromatographie d'exclusion stérique. D'une part, la protéine His-MBP-IHO1<sub>1-281</sub>, a été purifiée sur une colonne d'amylose puis injectée en gel-filtration (Figure 25 A). D'autre part, le fragment TOPOVIBL<sub>452-579</sub>, qui interagit avec REC114, a été purifié sur la résine Ni<sup>2+</sup>, puis mélangé dans un ratio équimolaire avec le complexe binaire Strep-REC114-MEI4<sub>1-127</sub>. Le mélange a été injecté sur la colonne Superose 6 afin de reconstituer le complexe ternaire REC114-MEI4-TOPOVIBL (Figure 25 B). Ensuite, les fractions du pic correspondant à la co-élution des trois protéines (REC114-MEI4-TOPOVIBL) ont été mélangées avec His-MBP-IHO1<sub>1-281</sub>. L'échantillon a été réinjecté sur la colonne de SEC pour analyser l'interaction entre REC114-MEI4-TOPOVIBL et IHO1 (Figure 25 C). Le chromatogramme obtenu et l'analyse sur gel des fractions ne montrent pas la formation d'un complexe tétramérique IHO1-REC114-MEI4-TOPOVIBL (Figure 25 C). Ce résultat indique que la protéine REC114 impliquée dans l'interaction avec TOPOVIBL n'interagit plus avec IHO1. Toutefois, une faible quantité de REC114-MEI4 a été éluée avec MBP-IHO1<sub>1-281</sub>. Cela peut être expliqué par la présence d'un excès de REC114-MEI4 lors de la reconstruction du complexe REC114-MEI4-TOPOVIBL, cet excès se retrouverait libre et interagirait avec MBP-IHO1<sub>1-281</sub>. Cela pourrait également s'expliquer par le fait que TOPOVIBL a été partiellement déplacée du sous-complexe REC114-MEI4 par IHO1.

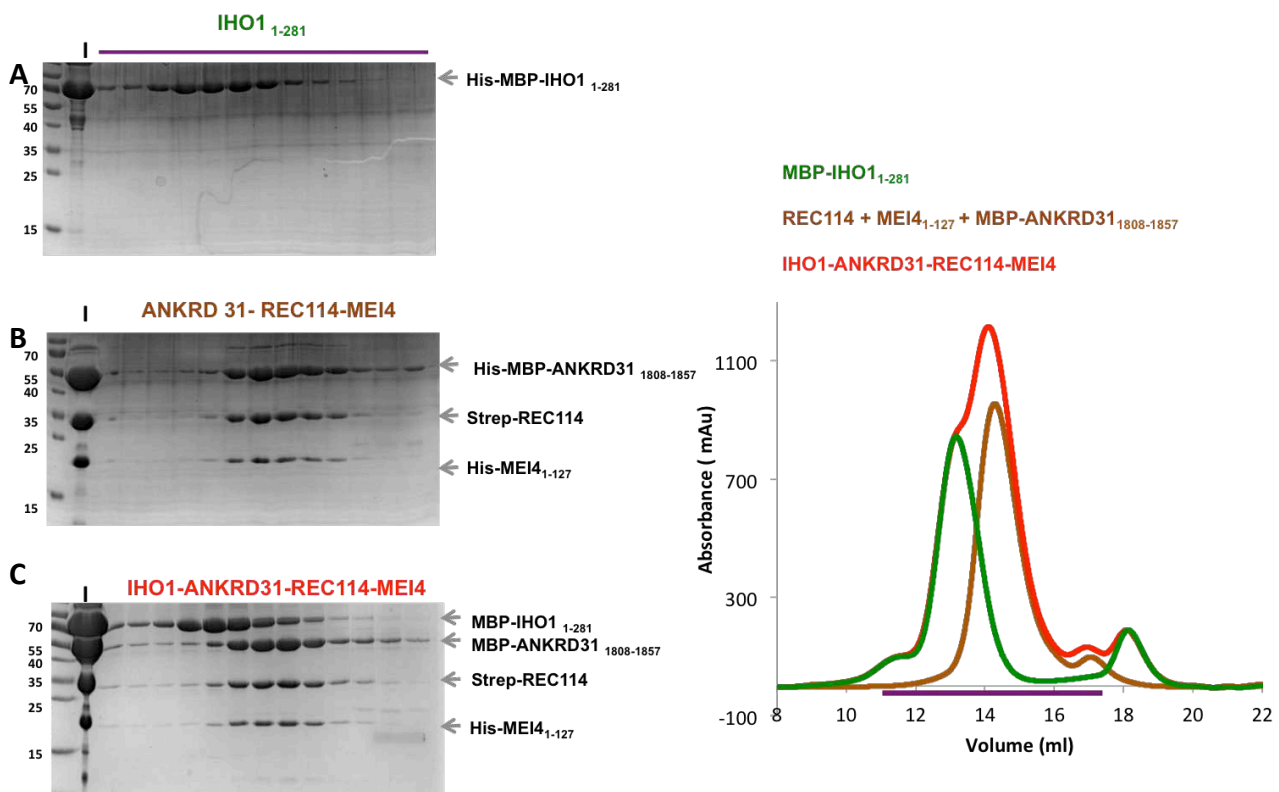


**Figure 25. Analyse de l'interaction de IHO1 avec REC114-MEI4 en présence de TOPOVBL.** (A) La protéine His-MBP-IHO1<sub>1-281</sub> purifiée par chromatographie d'affinité a été injectée sur colonne d'exclusion de taille Superose 6 (profil d'élution vert). (B) La protéine His-TOPOVBL<sub>452-579</sub> a été mélangée avec strep-REC114-His-MEI4<sub>1-127</sub> et injectée sur la filtration sur gel (profil d'élution bleu). (C) Les fractions d'élution de IHO1 et du complexe TOPOVBL-REC114-MEI4 ont été ensuite mélangées dans un ratio molaire 1 : 2 et injectées sur colonne d'exclusion de taille (profil d'élution rouge). La protéine injectée (I) ainsi que les mêmes fractions d'élution de ces échantillons (ligne violette) ont été analysées par SDS-PAGE sur un gel polyacrylamide à 15%.

### 3.1.9.2 REC114-MEI4-ANKRD31 et IHO1

La même expérience a été reproduite en présence de MBP-ANKRD31<sub>1808-1857</sub> comme partenaire du complexe REC114-MEI4. Cette fois, le complexe ternaire REC114-MEI4-ANKRD31 a été obtenu par co-lyse et co-purification de Strep-REC114-His-MEI4<sub>1-127</sub> et MBP-ANKRD31<sub>1808-1857</sub>. Ce résultat montre que REC114 peut interagir avec ANKRD31 en présence de MEI4, formant un complexe ternaire REC114-MEI4-ANKRD31. Ensuite, la protéine MBP-IHO1<sub>1-281</sub> et le complexe REC114-MEI4-ANKRD31 purifiés par la chromatographie d'exclusion de taille (Figure 26 A, B) ont été mélangés puis réinjectés sur la filtration sur gel. Le mélange de la protéine IHO1 et du complexe REC114-MEI4-ANKRD31 a été élué au même volume d'élution des protéines seules. L'analyse des pics obtenus ne montre aucune co-élution d'IHO1<sub>1-281</sub> avec le complexe REC114-MEI4-ANKRD31 (Figure 26 C).

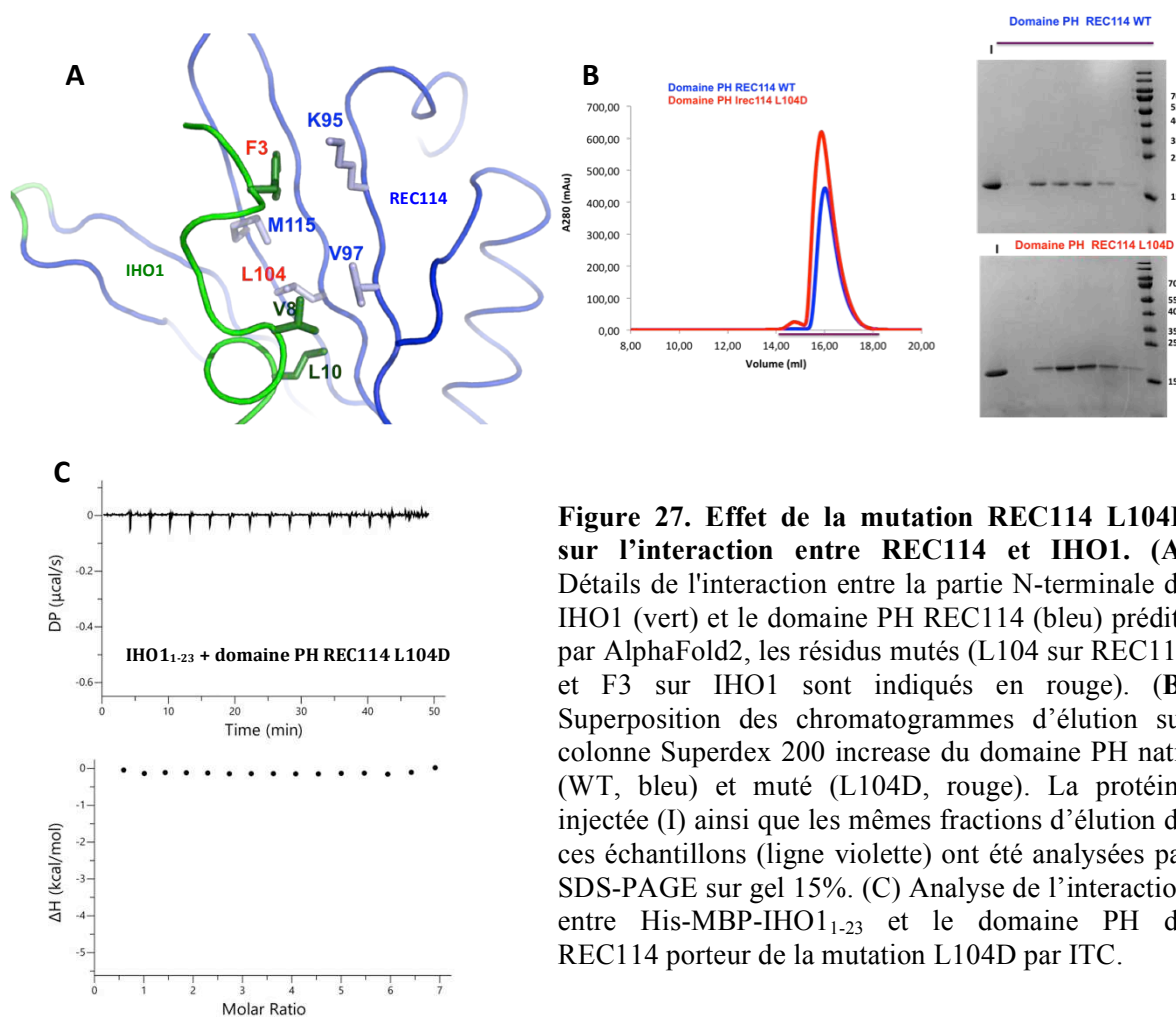
Les résultats de ces deux expériences montrent que la liaison des peptides de TOPOVIBL ou ANKRD31 à la protéine REC114 interfère avec son interaction avec la protéine IHO1. Ceci suggère que ces protéines (TOPOVIBL, ANKRD31 et IHO1) partagent l'interface de liaison à REC114. La dernière étude réalisée en partie par l'équipe a montré que l'affinité du fragment TOPOVIBL<sub>450-579</sub> pour la protéine REC114 est de 1,2  $\mu$ M. Il a été également proposé que l'affinité d'ANKRD31 serait supérieure, car la pré-liaison de cette protéine à REC114 empêche la fixation de TOPOVIBL (Nore et al., 2022). Ces résultats indiquent que l'affinité d'IHO1 pour REC114, *in vitro*, est inférieure à celle de TOPOVIBL<sub>450-579</sub> et ANKRD31, bien qu'ils partagent le même site de liaison sur REC114.



**Figure 26. Analyse de l'interaction de IHO1 avec REC114-MEI4 en présence d'ANKRD31.** (A) La protéine His-MBP-IHO1<sub>1-281</sub> et (B) le complexe Strep-REC114.His-MBP-ANKRD31<sub>1808-1857</sub>-His-MEI4<sub>1-127</sub> purifiés par chromatographie d'affinité ont été injectés sur colonne d'exclusion de taille Superose 6 (profil d'élution vert et marron respectivement). (C) Les fractions d'élution de la protéine IHO1 et celles du complexe ANKRD31-REC114-MEI4 ont été ensuite mélangées dans un ratio molaire 1 : 2 et injectées sur colonne d'exclusion de taille (profil d'élution rouge). La protéine injectée (I) ainsi que les mêmes fractions d'élution de ces échantillons (ligne violette) ont été analysées par SDS-PAGE sur un gel 15%.

### 3.1.9.3 Identification des résidus importants pour l'interaction entre REC114-IHO1

Récemment, il a été montré par Nore *et al* (2022) que la mutation du résidu hydrophobe L104D au niveau du feuillet  $\beta$  de REC114 aboli l'interaction avec TOPOVIBL. Nous avons souhaité vérifier l'effet de cette mutation sur la liaison d'IHO1 (Figure 27A). La construction codant pour le domaine PH de REC114 porteur de la mutation (Strep-REC114<sub>1-150</sub> L104D) a été exprimée et purifiée sur la résine de Strep-Tactin puis par gel-filtration. Le chromatogramme d'élution de ce mutant est comparable au profil d'élution de la protéine native (Figure 27 B), ce qui indique que cette mutation n'affecte pas le repliement de la protéine. L'échantillon purifié a été utilisé pour des mesures d'ITC avec le peptide N-terminal de la construction IHO1<sub>1-24</sub> exprimée en fusion avec la MBP. Comme avec la protéine TOPOVIBL, les données d'ITC montrent que la mutation du résidu L104 de la protéine REC114 entraîne la perte de l'interaction avec IHO1 (Figure 27 C, Tableau 3). Ce résultat confirme davantage le mode d'interaction d'IHO1 avec REC114, qui selon AlphaFold2 est similaire à celui de TOPOVIBL.



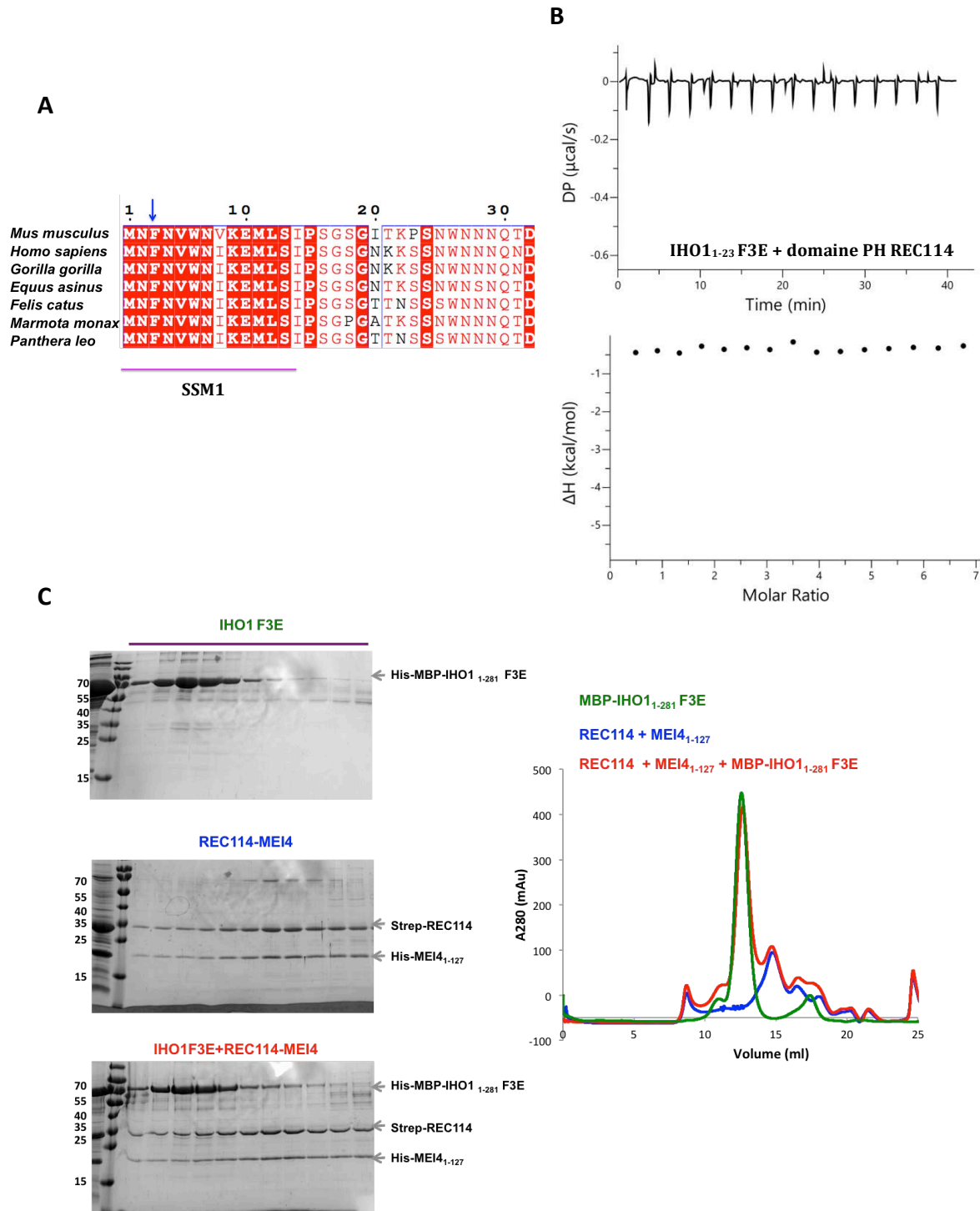
**Figure 27. Effet de la mutation REC114 L104D sur l'interaction entre REC114 et IHO1.** (A) Détails de l'interaction entre la partie N-terminale de IHO1 (vert) et le domaine PH REC114 (bleu) prédits par AlphaFold2, les résidus mutés (L104 sur REC114 et F3 sur IHO1) sont indiqués en rouge). (B) Superposition des chromatogrammes d'élution sur colonne Superdex 200 increase du domaine PH natif (WT, bleu) et muté (L104D, rouge). La protéine injectée (I) ainsi que les mêmes fractions d'élution de ces échantillons (ligne violette) ont été analysées par SDS-PAGE sur gel 15%. (C) Analyse de l'interaction entre His-MBP-IHO1<sub>1-23</sub> et le domaine PH de REC114 porteur de la mutation L104D par ITC.

Les mutations W562A, sur TOPOVIBL, et W1842A, sur ANKRD31, perturbent l'interaction des deux protéines avec REC114 (Boekhout et al., 2019; Nore et al., 2022). Selon le modèle AlphaFold2, le résidu F3 occuperait la même position que ces résidus dans la protéine IHO1 (Figure 24 B). Pour tester son importance, le résidu conservé F3 (Figure 27 A, Figure 28 A) d'IHO1<sub>1-24</sub> a été muté en acide glutamique et exprimé en fusion avec MBP. L'interaction entre ce peptide muté et le PH domaine a été testée par l'ITC. Le résultat obtenu a mis en évidence que le peptide muté ne s'est pas lié à REC114 (Figure 28 B, Tableau 3). Cette mutation a également été introduite dans la construction la plus longue d'IHO1<sub>1-281</sub> et son effet sur la liaison de REC114-MEI4 a été analysé par gel-filtration. La protéine His-MBP-IHO1<sub>1-281</sub> mutée (F3E) et le complexe Strep-REC114<sub>FL</sub>-MEI4<sub>1-127</sub> ont été mélangés puis injectés sur une colonne de chromatographie stérique. Contrairement au complexe impliquant la protéine IHO1 native (Figure 17), le chromatogramme du mélange formé avec la protéine IHO1 mutée est parfaitement superposable aux protéines individuelles, aucun décalage dans le volume d'éluion n'a été observé. La protéine IHO1<sub>1-281</sub>F3E et le complexe REC114-MEI4 ont été élués à leurs volumes d'éluions séparément (Figure 28 C).

	<b>MBP-IHO1<sub>1-281</sub></b>	<b>MBP-IHO1<sub>24-281</sub></b>	<b>MBP-IHO1<sub>1-23</sub></b>	<b>MBP-IHO1<sub>1-23</sub>F3E</b>
<b>REC114-MEI4</b>	Interaction (K <sub>d</sub> = 26 μM)	Pas d'interaction	Interaction (K <sub>d</sub> = 30 μM)	Pas d'interaction
<b>REC114 (domaine PH)</b>	Interaction (K <sub>d</sub> = 20 μM)	NA	Interaction (K <sub>d</sub> = 30 μM)	NA
<b>REC114 (domaine PH) L104D</b>	NA	NA	Pas d'interaction	NA

**Tableau 2. Récapitulatif des résultats ITC sur les interactions entre REC114 et IHO1.**  
NA : Interaction non analysée.

Ces résultats d'analyse d'interaction par ITC et la chromatographie d'exclusion de taille indiquent que le résidu F3 d'IHO1 est important pour l'interaction d'IHO1 avec REC114, tout comme les résidus W1842 et W562 d'ANKRD31 et de TOPOVIBL respectivement. Ces résultats, obtenus par des analyses *in vitro*, sont cohérents et supportent le modèle AlphaFold2.



**Figure 28. Effet de la mutation F3E de IHO1 sur l'interaction entre REC114 et IHO1. (A)** Alignement de séquences de l'extrémité N-terminale d'IHO1 chez les mammifères. Lettres rouges : acides aminés équivalents ; lettres blanches : acides aminés identiques. Le résidu F3 muté est montré par une flèche. **(B)** analyse de l'interaction entre la His-MBP-IHO1<sub>1-23</sub> muté en F3E et le domaine PH de REC114 par ITC. **(C)** La protéine His-MBP-IHO1<sub>1-281</sub> F3E et le complexe Strep-REC114-His-MEI4<sub>1-127</sub> purifiés par chromatographie d'affinité ont été injectés sur colonne Superose 6 (profil d'élution vert et bleu respectivement). Les fractions d'élution de ces deux échantillons ont été ensuite mélangées dans un ratio molaire 1 : 2 et injectées sur colonne d'exclusion de taille (profil d'élution rouge). La protéine injectée (I) ainsi que les mêmes fractions d'élution de ces échantillons (ligne violette) ont été analysées par SDS-PAGE sur un gel à 15%.

### 3.1.10 Essais de cristallisation du complexe IHO1-REC114

Compte tenu de la difficulté d'optimiser la qualité des échantillons du complexe IHO1-REC11-MEI4 pour la caractérisation structurale par Cryo-EM, et du fait que la structure du domaine PH de REC114 seul ou en complexe avec le C-terminus de TOPOVIBL, ou le C-terminus d'ANKRD31, a déjà été déterminée par cristallographie, nous avons décidé de nous orienter vers cette technique dans le but d'élucider la structure du domaine PH de REC114 en interaction avec le peptide N-terminal d'IHO1. Cependant, et comparée au complexe REC114-TOPOVIBL ou REC114-ANKRD31, la faible affinité et la nature transitoire de l'interaction entre REC114 et IHO1 pourraient rendre difficile la formation des cristaux contenant les deux partenaires en interaction. Pour pallier ce problème, nous avons procédé au clonage du peptide codant les 16 premiers résidus de la protéine IHO1 fusionné à la partie C-terminale de la séquence codant le domaine PH via un *linker* de 7 glycines, en se basant sur le modèle AlphaFold2. Cette approche permettrait de lier de manière covalente les deux partenaires et d'augmenter la concentration locale du peptide IHO1<sub>1-16</sub>, augmentant ainsi la chance d'obtenir des cristaux qui captent les deux protéines en interaction.

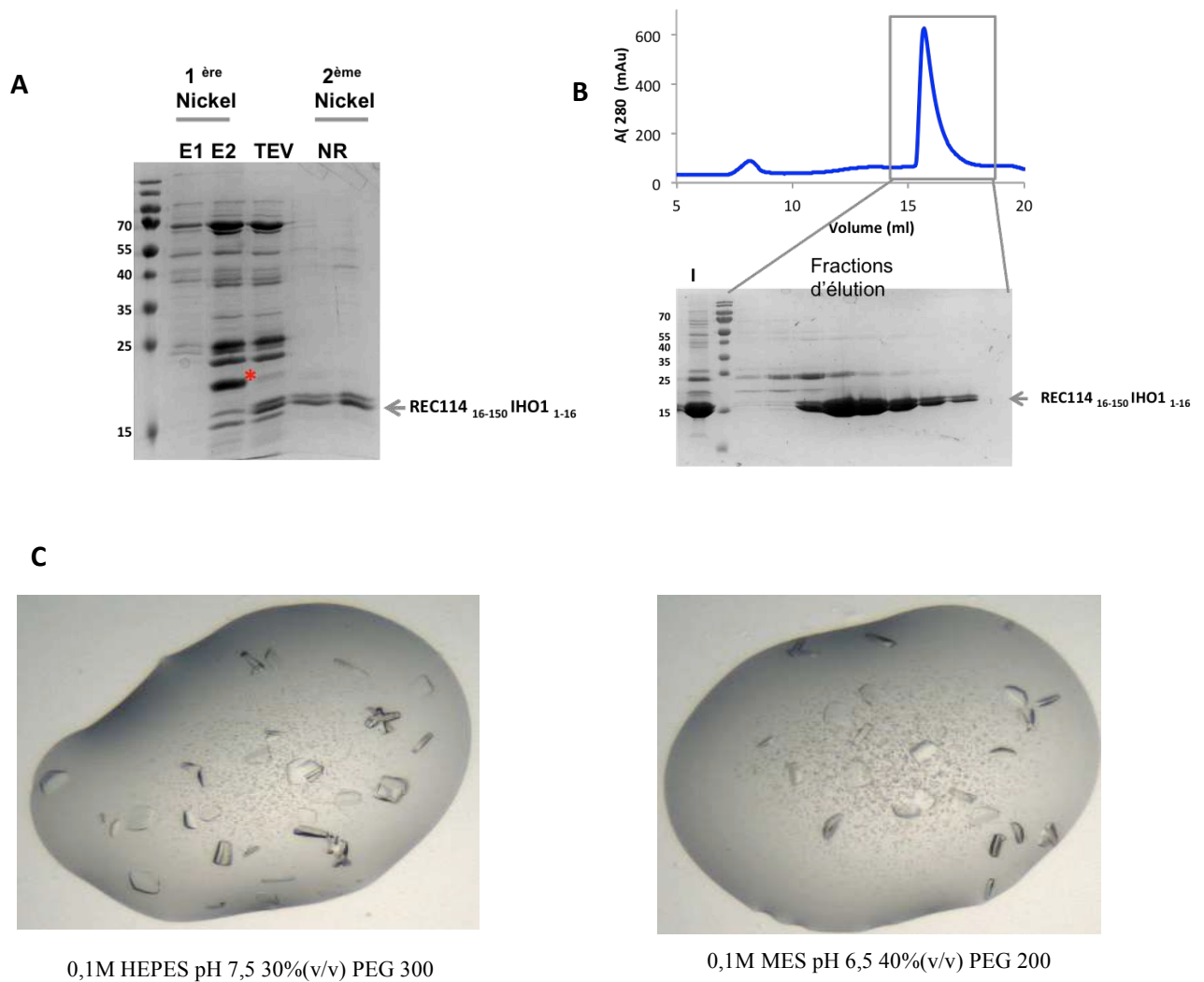
Après la purification de la fusion His-REC114<sub>16-150</sub> IHO1<sub>1-16</sub> sur la résine nickel, l'étiquette de 6xHis de l'extrémité N-terminale de REC114 a été clivée par la TEV. L'échantillon a été ensuite repassé sur une colonne de chromatographie d'affinité pour se débarrasser de cette étiquette et récupérer la protéine (Figure 29 A). Cependant, une bande de dégradation de cette fusion a été constatée à cette étape. Cette bande correspondrait à une dégradation C-terminale d'une partie ou de la totalité du peptide IHO1, car la purification du domaine His-REC114<sub>16-150</sub> ne présente pas cette dégradation (Figure 27 B). La bande de dégradation est toujours présente après l'étape de gel-filtration et elle est co-éluee avec la fusion entière (Figure 29 B). Toutefois, des tests de cristallisation ont été effectués sur cet échantillon et des cristaux ont été obtenus dans des conditions différentes de celles du domaine PH seul ou en interaction avec TOPOVIBL ou ANKRD31 (Figure 29 B). Le traitement des données de diffraction de ces derniers (résolution à 2,5 Å) ne montre pas la présence du peptide d'IHO1, mais uniquement le domaine PH de REC114. Plusieurs pistes ont été explorées pour essayer de résoudre ce problème. Dans un premier temps, la durée d'expression de cette fusion a été réduite à 3 h et toutes les étapes de la purification ont été réalisées à 4 °C pour limiter cette dégradation, mais cela n'a pas apporté

d'amélioration et la dégradation a persisté. Ensuite, une étiquette Strep a été ajoutée à l'extrémité C-terminale du peptide d'IHO1<sub>1-16</sub> pour essayer de la protéger contre cette dégradation, mais aussi pour ne récupérer que l'échantillon qui contient le peptide entier lors de la purification sur la résine Strep-tactin. Cette méthode permet actuellement de prévenir la dégradation de cette fusion, mais le rendement de cette construction était très faible et insuffisant pour les tests de cristallisation.

Nous avons également tenté de déterminer la structure de ce complexe en utilisant un peptide synthétique d'IHO1 couvrant les résidus 1-16, mais en raison des problèmes de solubilité, il n'a finalement pas pu être utilisé.

Bien qu'une structure cristalline expérimentale soit la meilleure façon de valider le modèle structural du complexe IHO1-REC114, cela n'ait pas pu être réalisé pour le moment. Toutefois, malgré que les détails du modèle AlphaFold2 puissent ne pas être exacts, les études de mutagenèses et les analyses biochimiques décrites ci-dessus montrent clairement que l'extrémité N-terminale de IHO1 interagit avec le même feuillet  $\beta$  de REC114 que TOPOVIBL et ANKRD31.





**Figure 29. Purification et essais de cristallisation du domaine PH de REC114 en fusion avec IHO1<sub>1-16</sub>.** (A) Analyse par SDS-PAGE des fractions des différentes étapes de purification de la construction de la fusion entre domaine PH (REC11<sub>15-150</sub>) et IHO1<sub>1-16</sub> par chromatographie d'affinité. E1-E2 : élution sur une résine chargée de nickel. TEV : coupure à la protéase TEV. NR : échantillon non retenu sur la seconde colonne d'affinité sur résine Ni<sup>2+</sup>. (B) Profil d'élution de la fusion entre domaine PH (REC11<sub>15-150</sub>) et IHO1<sub>1-16</sub> en chromatographie d'exclusion de taille. La protéine injectée sur la colonne (I) et les fractions d'élution encadrées en noir ont été analysées par SDS-PAGE. (C) Les cristaux obtenus pour le domaine PH en fusion avec IHO1<sub>1-16</sub>. Les conditions de cristallisation sont indiquées.

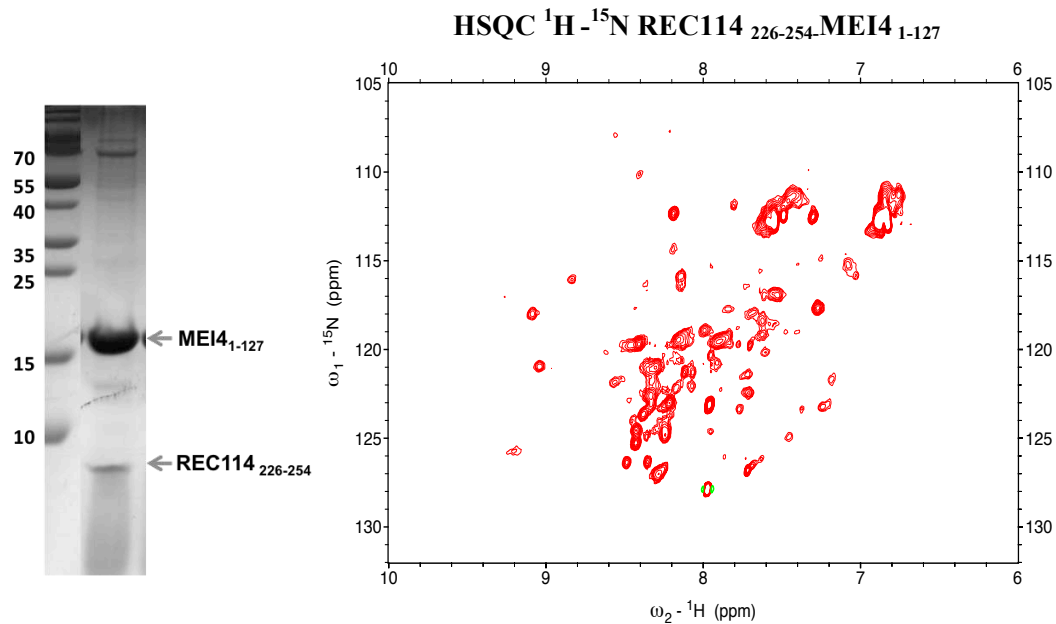


## 3.2 Complexe REC114-MEI4

L'autre composant du complexe REC114-MEI4-IHO1 est le sous-complexe REC114-MEI4, dont les travaux précédents de l'équipe ont démontré, *in vitro*, que la partie C-terminale de REC114 interagit avec le fragment N-terminal de MEI4 (Kumar et al., 2018). Dans le but d'obtenir les détails moléculaires de cette interaction, plusieurs essais de cristallisation ont été effectués, avant ma thèse, sur le complexe formé entre REC114<sub>226-254</sub> et MEI4<sub>1-127</sub>. Malheureusement, aucun cristal n'a pu être obtenu. Comme on l'a vu dans la partie précédente, la caractérisation du complexe REC114-MEI4<sub>1-127</sub> par la technique de SEC-MALLS a montré qu'il est constitué d'un hétérotrimère 2REC114-1MEI4<sub>1-271</sub> qui se dimérise, probablement à travers le domaine PH de REC114 (Figure 18 B). La partie REC114<sub>227-259</sub> est dépourvue du domaine PH, elle formerait ainsi avec MEI4 un complexe hétérotrimérique d'environ 22 kDa (2REC114<sub>226-254</sub> (4 kDa) et 1MEI4<sub>1-127</sub> (14 kDa)). Avec un tel poids moléculaire, le complexe REC114<sub>226-254</sub>-MEI4<sub>1-127</sub> serait un objet adéquat pour l'analyse par la technique de la résonance magnétique nucléaire (RMN). Par conséquent, il a été décidé de s'orienter vers cette technique pour déterminer la structure de ce complexe, ou d'identifier des régions flexibles présentes dans ces protéines et qui sont susceptibles d'interférer avec sa cristallisation.

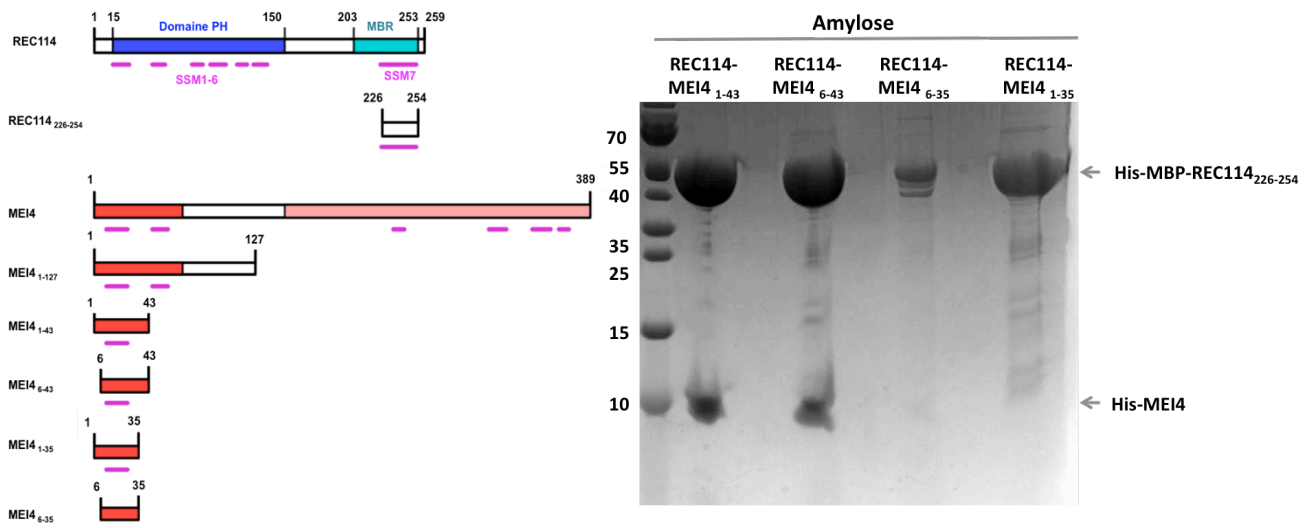
### 3.2.1 Caractérisation par RMN du complexe REC114-MEI4

Les premières données RMN collectées sur le complexe doublement marqué [<sup>13</sup>C, <sup>15</sup>N] REC114<sub>226-254</sub> et MEI4<sub>1-127</sub> n'étaient pas exploitables. Le spectre 2D collecté, corrélant le proton <sup>1</sup>H à celui d'un atome <sup>15</sup>N auquel est lié (<sup>1</sup>H-<sup>15</sup>N HSQC), contient beaucoup moins de taches de corrélation que le nombre de résidus des deux protéines. Ceci suggère que seule une partie du complexe est visible sur le spectre. Les taches sont également larges, de faibles intensités et ils se chevauchent (Figure 30). Ce qui rend l'attribution de ces résonances aux résidus associés très difficile. Ces propriétés de spectres sont dues soit à la présence d'états oligomériques de hauts poids moléculaires, soit à des processus d'échanges entre les différentes stœchiométries du complexe. De plus, ce complexe n'était pas stable et il se dégradait avec le temps. Par conséquent, il a fallu rechercher de nouvelles constructions plus stables et plus courtes pour améliorer et simplifier les spectres RMN.



**Figure 30.** Expérience HSQC  $^1\text{H}$ - $^{15}\text{N}$  réalisée sur le complexe REC114<sub>226-254</sub>-MEI4<sub>1-127</sub>, doublement marqué  $^{15}\text{N}$ - $^{13}\text{C}$ , dont une fraction analysée par SDS-PAGE est représentée.

La construction de REC114<sub>226-254</sub> correspond au fragment minimal requis pour l'interaction avec MEI4. Dans les travaux précédents de l'équipe, de courtes tronctions de chaque côté de REC114<sub>226-254</sub> interféraient avec la liaison de MEI4. Pour cette raison, je me suis concentrée sur la protéine MEI4, j'ai testé diverses constructions et j'ai finalement trouvé que les 43 premiers résidus de MEI4 étaient suffisants pour interagir avec REC114. Par contre, la construction couvrant les résidus 1 à 35 perd sa capacité à lier REC114 (Figure 31).

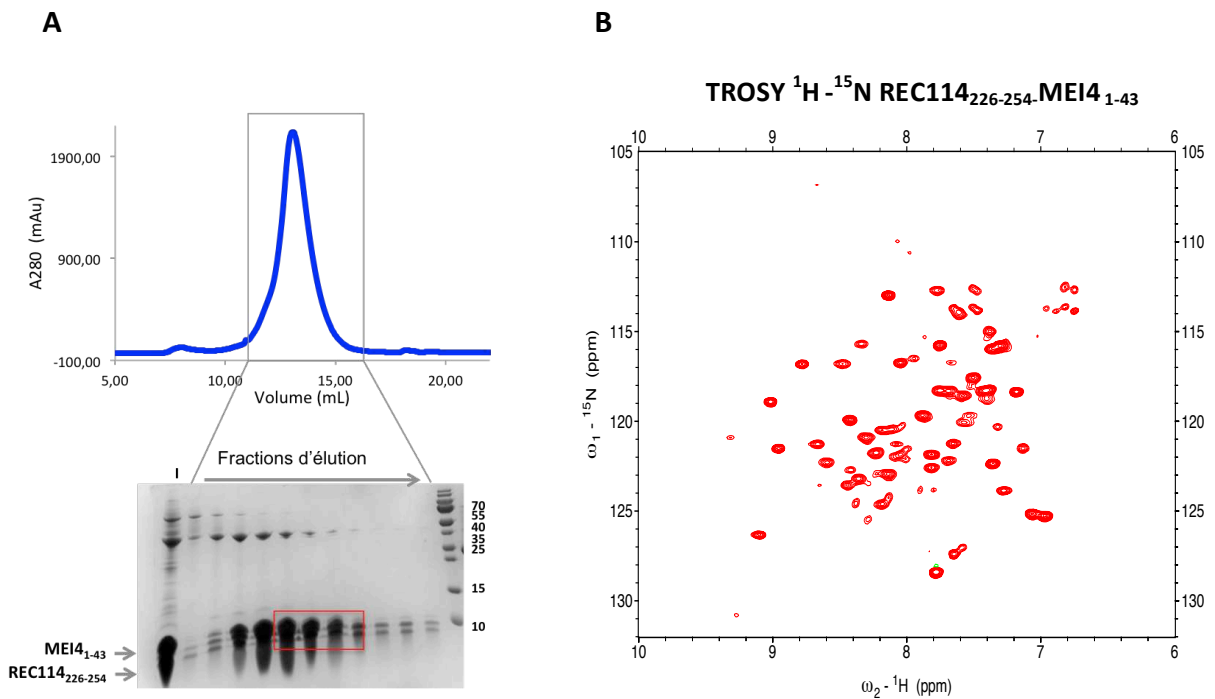


**Figure 31.** Analyse des différentes tronctions de MEI4. À gauche, schématisation de la construction REC114<sub>226-254</sub> qui correspond au fragment C-terminal nécessaire pour l'interaction avec MEI4 et les différentes tronctions de MEI4 analysées. À droite, Analyse par SDS-PAGE des éluions sur la résine amylose de la co-purification des constructions de MEI4 (1-43), (6-43), (1-35) et (6-35) co-exprimées avec His-MBP-REC114<sub>226-254</sub>. La tronction (1-127) n'est pas montrée sur ce gel et elle est représentée dans la Figure 30.

Pour effectuer de nouvelles acquisitions RMN, la construction d'ADN codant pour le fragment C-terminal de REC114 (His-MBP-REC114<sub>226-254</sub>) a été co-exprimée avec le plasmide contenant la partie N-terminale de MEI4 (His-MEI4<sub>1-43</sub>) dans un milieu enrichi en <sup>15</sup>N et <sup>13</sup>C. Après la co-purification sur la résine amylose et le clivage des étiquettes de purification par la TEV, l'échantillon protéique a été rechargé sur la résine nickel pour se débarrasser des étiquettes de purification et de la protéine éventuellement non clivée. Le complexe a été ensuite purifié sur la colonne de chromatographie d'exclusion de taille. L'analyse par SDS-PAGE montre la présence d'une contamination à la His-MBP, malgré le passage de l'échantillon clivé sur la chromatographie d'affinité sur la résine d'affinité au nickel (Figure 32 A). Les fractions d'élution, les moins contaminées en His-MBP, ont été concentrées à 150 µM pour l'acquisition de nouveaux spectres RMN. Pour avoir des données de meilleures qualités, un spectre 2D du type TROSY (Transverse Relaxation Optimized Spectroscopy) a été collecté (Figure 32 B). C'est une technique RMN développée pour les complexes de tailles importantes. Elle permet de supprimer les données des régions très mobiles et d'obtenir un spectre avec des taches de résonances bien définies. Bien qu'il n'y ait toujours pas la totalité des pics, la qualité des données obtenues avec cette méthode semble être suffisante pour collecter des mesures 3D, indiqués dans la section *Matériels et Méthodes*. Ces derniers permettent la corrélation de l'atome d'azote et le proton amide d'un résidu avec l'atome de carbone (<sup>13</sup>C) appartenant soit au même résidu, soit au résidu précédent, permettant ainsi de connecter les résidus consécutifs dans la chaîne protéique.

Malheureusement, plusieurs signaux de résonances sont absents des spectres RMN, ce qui rompt souvent le chemin d'attribution des acides aminés. Pour cette raison, nous n'avons pas réussi à faire l'attribution complète du squelette peptidique, qui est la première étape pour accéder aux données structurales par RMN. Dans le but d'améliorer la qualité de ces données, des acquisitions de spectres HSQC ont été effectuées à différents pH et à différentes concentrations salines, mais aucune amélioration n'a été constatée.

Ces résultats ont permis de définir les régions minimales pour la formation du complexe REC114-MEI4 (REC114<sub>226-254</sub> et MEI4<sub>1-43</sub>), mais l'échantillon purifié n'était pas adapté à la détermination de la structure par RMN. De plus, aucun cristal n'a pu être obtenu pour une éventuelle caractérisation de ce complexe par la cristallographie aux rayons X.



**Figure 32. Purification du complexe REC114<sub>226-254</sub>-MEI4<sub>1-43</sub> et le spectre RMN associé.** (A) Profil d'éluion du complexe REC114<sub>226-254</sub>-MEI4<sub>1-43</sub> en chromatographie d'exclusion de taille. L'échantillon injecté (I, purifié sur colonne amylose, coupé à la TEV et passé ensuite sur la résine chargée de nickel) et les fractions d'éluion encadrées en noir ont été analysées par SDS-PAGE. (B) Les fractions encadrées en rouge ont été utilisées pour l'acquisition d'un spectre de type TROSY corrélant le proton H à celui d'un atome  $^{15}\text{N}$  auquel il est lié.

### 3.2.2 Caractérisation de l'état oligomérique du complexe REC114-MEI4

Compte tenu des résultats de la technique SEC-MALLS décrits précédemment suggérant deux états oligomériques différents du complexe REC114-MEI4 (2 :1 et 4 :2), et de la difficulté d'obtenir des spectres RMN exploitables, j'ai décidé d'approfondir la caractérisation de la stœchiométrie du complexe REC114-MEI4.

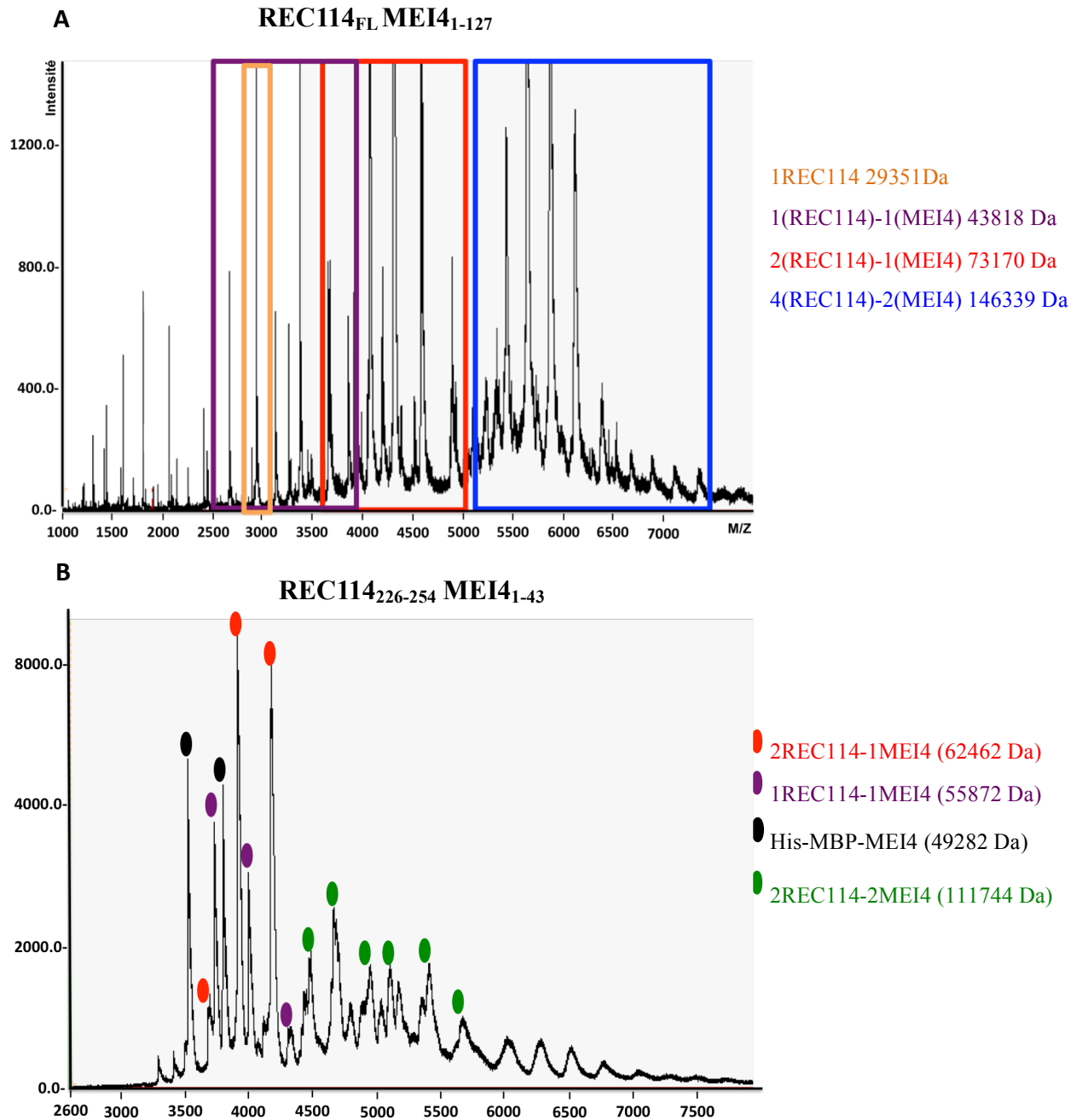
#### 3.2.2.1 Spectrométrie de masse native (MS native)

La technique de MS native a été utilisée pour confirmer l'état oligomérique, observé par SEC-MALLS, du complexe formé entre la protéine complète REC114<sub>FL</sub> et la plus longue construction de MEI4, MEI4<sub>1-127</sub>. Après les différentes étapes de purification et d'échange de tampon à l'acétate d'ammonium, décrites dans la section *Matériels et Méthodes*, l'échantillon protéique a été préparé à une concentration d'environ 10 µM et analysé par la plateforme de la spectrométrie de masse à l'IBS. Le spectre obtenu était très complexe et l'échantillon était très hétérogène. L'analyse des principaux signaux montre la présence de plusieurs populations dont un hétérodimère 1REC114-1MEI4 (44 kDa), un hétérotrimère 2REC114-1MEI4 (78 kDa), et un dimère de l'hétérotrimère 2(2REC114-1MEI4) (146 kDa) (Figure 33 A).

Par ailleurs, cette technique a été utilisée pour caractériser la stœchiométrie du complexe formé par la partie C-terminale de REC114, dépourvu du domaine PH, en interaction avec le plus petit fragment N-terminale de MEI4 (His-REC114<sub>226-254</sub>-His-MBP-MEI4<sub>1-43</sub>). Les données obtenues sont également complexes et elles indiquent la présence d'un mélange de populations oligomériques. Néanmoins, les principaux signaux détectés par cette technique correspondent à un hétérotrimère 2REC114-1MBP-His-MEI4 (62.5 kDa) (Figure 33 B). Cette stœchiométrie a été également confirmée par la MS/MS en tandem.

Les résultats de cette technique montrent la présence de plusieurs formes oligomériques pour le complexe REC114-MEI4, suggérant une dynamique d'association et de dissociation entre les espèces de ce complexe. Selon ces résultats, cette dynamique est observée non seulement dans le complexe formé avec REC114<sub>FL</sub>, mais aussi dans le complexe contenant uniquement la partie C-terminale de REC114. Ce qui suggère que la

dynamique d'association et de dissociation du complexe REC114-MEI4 ne se limite pas au domaine PH, qui a tendance à se dimériser.



**Figure 33. Caractérisation de l'état oligomérique du complexe REC114-MEI4 par MS native.** Spectre de la spectrométrie de masse en condition native pour le complexe REC114<sub>FL</sub>-MEI4<sub>1-127</sub> (A) et le complexe formé entre le fragment C-terminal de (REC114<sub>226-254</sub>) et la partie N-terminale de MEI4 (MEI4<sub>1-43</sub>). (B) Les différentes espèces attribuées aux pics de chaque spectre sont indiquées.



### 3.2.2.2 Ultracentrifugation analytique (AUC)

La technique AUC a été utilisée pour mieux caractériser le comportement du complexe REC114-MEI4 et analyser l'effet de la concentration sur son état oligomérique. Dans cette technique, les particules sont soumises à la force centrifuge et séparées selon une vitesse de sédimentation qui est étroitement liée à leurs paramètres physico-chimiques (taille, forme, densité). L'échantillon protéique du complexe REC114<sub>226-254</sub>-MEI4<sub>1-43</sub>, purifié de manière similaire à l'échantillon utilisé en RMN, a été préparé à trois concentrations 0.2 mg/mL, 1 mg/mL et 4.5 mg/mL et analysé au niveau de la plateforme ISBG.

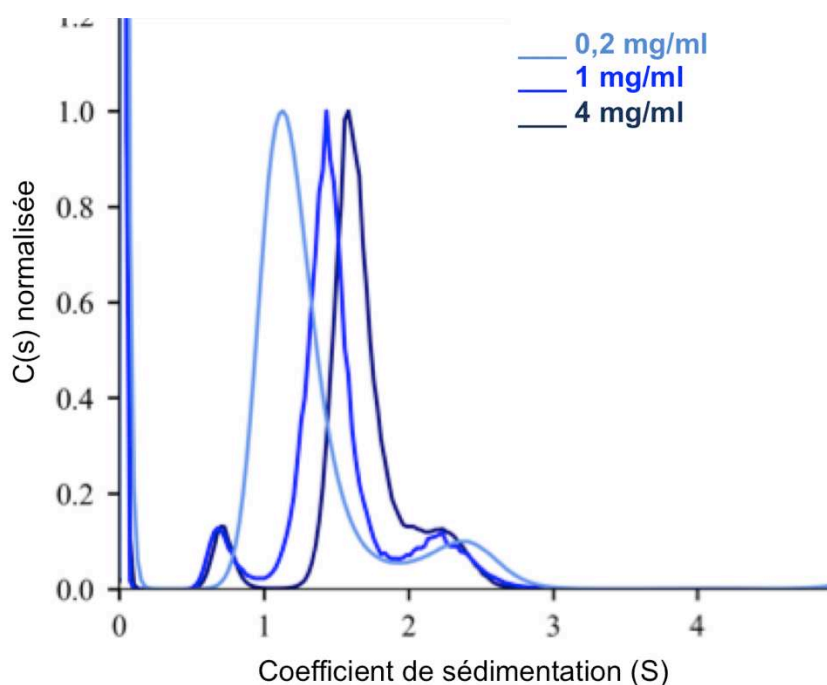
Les données de sédimentation obtenues montrent qu'il existe deux pics minoritaires et un pic majeur. Les coefficients de sédimentation des deux pics minoritaires sont constants, mais celui du pic majoritaire varie avec la concentration.

Les deux pics minoritaires sont observés pour les trois concentrations (Figure 34, données d'absorbance uniquement). Le premier à 0.72 S pourrait correspondre au complexe 1REC114-1MEI4. Alors que le deuxième pic de 2,3 S représente une espèce de haut poids moléculaire et peut correspondre à un complexe de stoechiométrie 8REC114-4MEI4.

La majeure partie du complexe (84 % du signal total) a été sédimentée dans le pic principal dont le coefficient de sédimentation augmente avec l'augmentation de la concentration. Il est de 1.2 S pour 0.2 mg/mL, de 1.44 S pour 1 mg/mL et de 1.64 S pour 4.5 mg/mL (Figure 34). Cette augmentation du coefficient de sédimentation avec l'augmentation de concentration suggère un équilibre d'association-dissociation rapide entre les espèces de ce complexe. Les complexes 2REC114-1MEI4 et 4REC114-2MEI4 devraient avoir des coefficients de sédimentation de 1,13 S et 1,58 S, respectivement. Ces valeurs sont telles que la concentration de 0,2 mg/mL (1,2 S) est majoritairement peuplée par le complexe 2REC114-1MEI4, alors que la concentration de 4 mg/mL (1,64 S) correspondrait à un complexe 4REC114-2MEI4. Le pic intermédiaire observé avec l'échantillon concentré à 1 mg/mL serait un mélange formé par l'échange rapide entre les deux états oligomériques, 2REC114-1MEI4, 4REC114-2MEI4. Les expériences RMN ont été réalisées sur un échantillon dans la même gamme de concentration que l'échantillon qui sédimente dans le pic intermédiaire et qui présente un mélange de différentes espèces oligomériques. La dynamique de l'équilibre entre les différentes formes oligomériques est défavorable aux études RMN (Tong et al., 2005). En effet, l'échange entre le monomère et

l'oligomère conduit à un élargissement, voire une disparition des résonances pour les résidus qui connaissent des environnements chimiques différents entre le monomère et l'oligomère. Ce qui expliquerait la qualité limitée des spectres RMN du complexe REC114-MEI4 et la difficulté d'attribuer séquentiellement les résidus de la chaîne principale.

Combinés aux données de la spectrométrie en condition native, les résultats obtenus par cette approche indiquent que le complexe REC114-MEI4 est composé principalement de deux REC114 et d'une MEI4, mais il s'oligomérisse de manière concentration-dépendante. La dynamique d'échange entre les différentes espèces de ce complexe serait rapide.



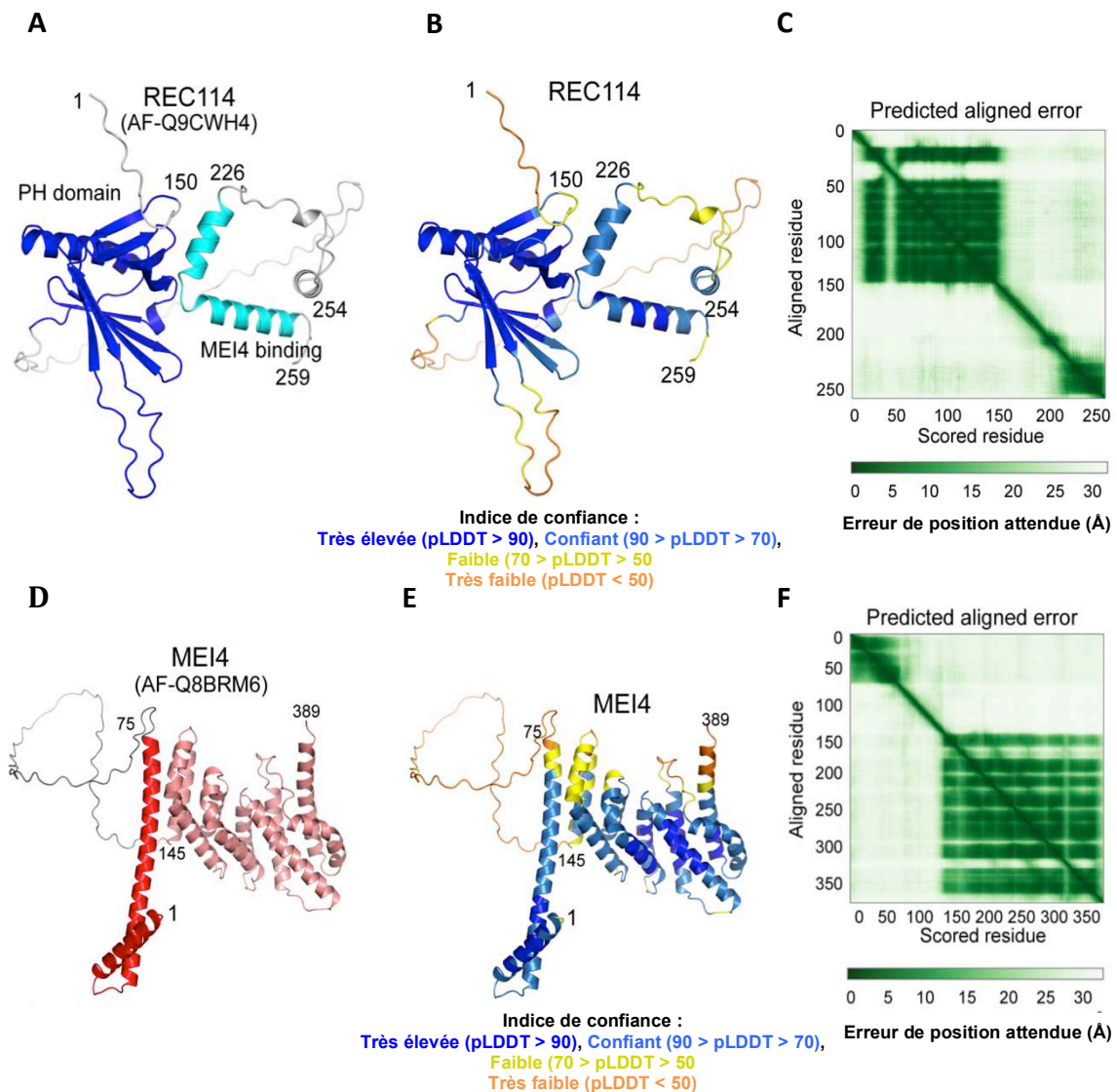
**Figure 34. Profils de vitesse de sédimentation du complexe REC114-MEI4 (absorbance 280).** Les expériences d'AUC ont été effectuées avec trois concentrations (0.2 mg/mL, 1 mg/mL et 4.5 mg/mL).

### 3.2.3 Modèle structural par AlphaFold2 du complexe REC114-MEI4

Il y a un an, une autre façon d'obtenir les informations structurales sur le complexe REC114-MEI4 est devenue possible. Le programme de prédiction de structure par intelligence artificielle AlphaFold2 a pu prédire la structure du complexe sans être affecté par les problèmes d'oligomérisation.

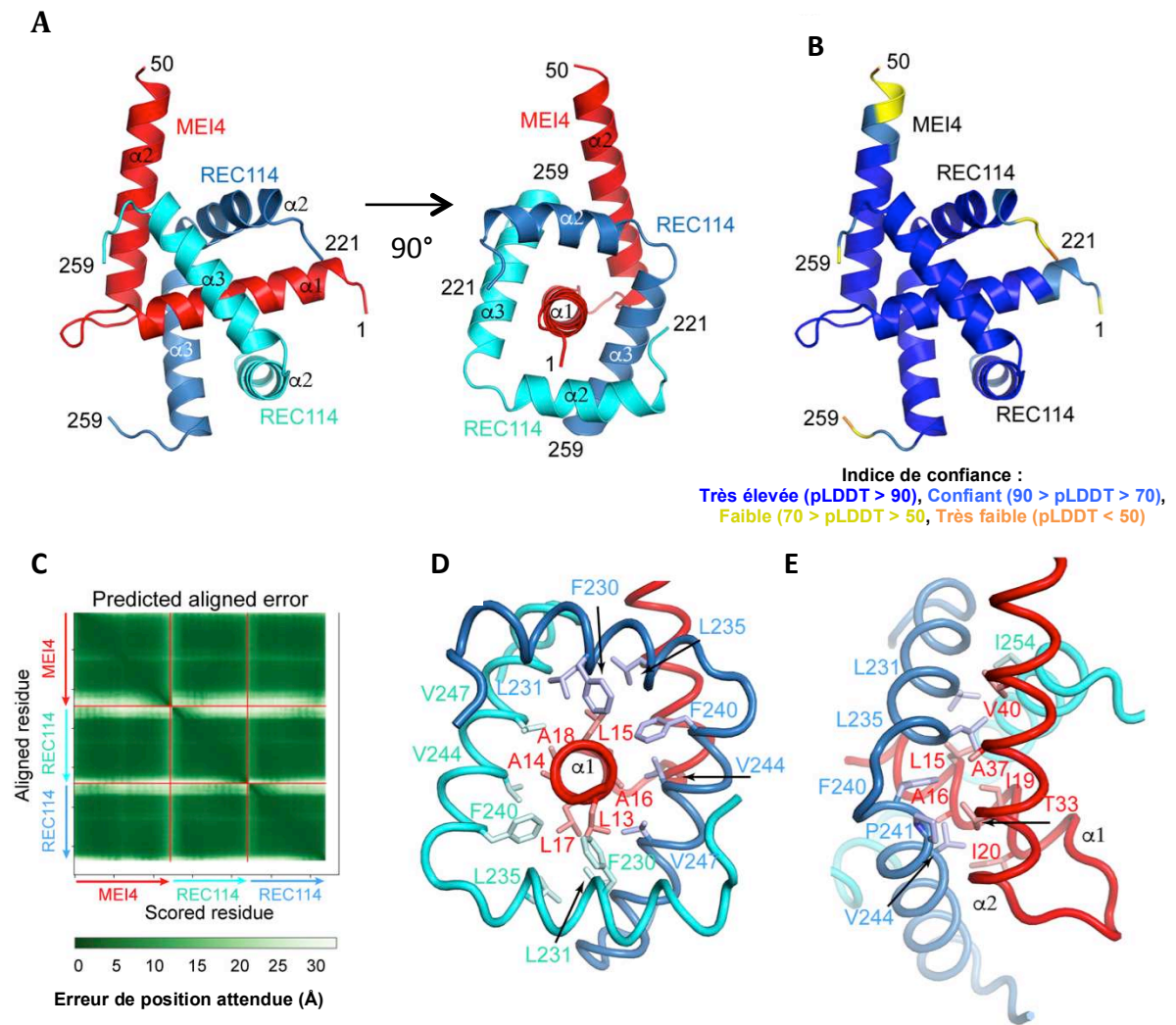
Les prédictions AlphaFold2 disponibles sur le site de l'EBI comprennent les modèles des protéines REC114 et MEI4 de souris (Figure 35) qui sont prédits avec un très bon score de confiance. Les deux modèles confirment notre analyse bioinformatique initiale. REC114 est constituée d'un domaine PH N-terminal et d'une courte partie C-terminale hélicoïdale. Ces deux derniers sont prédits avec un indice de confiance élevé et ils sont liés par un *linker* désordonné (Figure 35 A, B). Les positions indépendantes des deux domaines sont confirmées par le graphique indiquant la probabilité d'erreur sur la distance attendue dans la prédiction de la position de chaque paire d'acides aminés (Figure 35 C). De même, MEI4 consiste en une partie N-terminale hélicoïdale et un grand domaine hélicoïdal également relié par une séquence peptidique désordonnée (Figure 35 D, E). Bien que le modèle suggère un contact possible entre les deux parties structurées, le graphique de la prédiction des erreurs d'alignement indique qu'elles sont indépendantes (Figure 35 F).

Grâce à AlphaFold2 (Jumper et al., 2021), la structure 3D du complexe REC114-MEI4 a pu être modélisée. Nous avons utilisé la stœchiométrie minimale 2REC114-1MEI4, montrée pour le complexe REC114-MEI4 de la levure *S. cerevisiae* (Claeys Bouuaert et al., 2021a), et observée par les différentes techniques pour les protéines de la souris, pour modéliser l'interaction entre ces deux protéines. Nous avons essayé de modéliser le complexe incluant les protéines entières, mais les seuls contacts observés concernent la partie C-terminale de REC114 et la partie-N terminale de MEI4, caractérisées dans la partie précédente (Figure 36 A).



**Figure 35. Modèle structural de REC114 et MEI4.** (A, B) Modèle AlphaFold2 (Jumper et al., 2021) de la protéine REC114. En (A) le modèle est coloré en suivant le code couleur de la Figure 30 et en (B) selon la valeur de pLDDT (pour predicted local distance difference test) qui correspond au score de confiance de la prédiction de chaque résidu. (C) Graphique indiquant l'erreur sur distance (en Ångströms) attendue dans la prédiction de la position du résidu X par rapport au résidu Y est, une couleur foncée à la position (x, y) indique que la probabilité d'erreur est faible tandis qu'une couleur claire traduit une probabilité d'erreur élevée. (D, E) modélisation de la protéine MEI4 par AlphaFold2 (Jumper et al., 2021). En (D) le modèle est coloré en suivant le code couleur de la figure 30 et en (E) selon la valeur de pLDDT. (F) Graphique indiquant l'erreur de distance (en Ångströms) attendue dans la prédiction de chaque paire de résidus (résidu en X et résidu en Y).

Le modèle AlphaFold2 du complexe REC114-MEI4 a été prédit avec un très bon score de confiance (Figure 36 B, C). La région de REC114 qui interagit avec MEI4 est  $\alpha$ -hélicoïdale. Elle est constituée de deux hélices  $\alpha 2$  et  $\alpha 3$  séparées par une boucle. Les deux monomères interagissent de façon antiparallèle pour former le dimère REC114. L'extrémité de l'hélice  $\alpha 2$  d'un monomère interagit avec l'hélice  $\alpha 3$  de l'autre monomère et vice versa, en adoptant une disposition rectangulaire (Figure 36 A). La surface interne de cette configuration est constituée de résidus conservés hydrophobes des deux monomères de REC114, à savoir F230, L231, F240, V244 et V247 (Figure 36 D). Ces derniers seraient impliqués dans l'interaction avec MEI4. La partie N-terminale de MEI4 est également formée par deux hélices (Figure 36 A, B), une courte ( $\alpha 1$  résidus 5-22) et une plus longue ( $\alpha 2$  résidus 26-76). La première hélice de MEI4 interagit de manière assez symétrique avec le dimère formé par REC114 via la séquence peptidique (13-20) LALALAI. Les résidus, L13, A14, L17 et A18 sont orientés vers un monomère de REC114 et les résidus L15, A16, I19 et I20 vers l'autre monomère (Figure 36 D, E). Le début de la deuxième hélice ( $\alpha 2$  jusqu'au résidu 43) de MEI4 fournit des contacts supplémentaires entre les deux protéines, en interagissant avec un monomère de REC114 (Figure 36 E). Ces contacts semblent cruciaux pour la formation du complexe. En effet, la construction de MEI4<sub>1-35</sub> dépourvue de cette partie (résidus 36 à 43) d'hélice perd la capacité d'interagir avec REC114 (Figure 31).



**Figure 36. Modèle structural du complexe 2REC114-1MEI4.** (A, B) Modélisation du complexe 2REC114<sub>221-251</sub>-1MEI4<sub>1-50</sub> avec AlphaFold2 (Jumper et al., 2021). Le modèle représenté en (A) est coloré en suivant le code couleur de la Figure 30 et celui en (B) selon la valeur de pLDDT. (C) Graphique indiquant l'erreur sur distance (en Ångströms) attendue dans la prédiction de la position du résidu X par rapport au résidu Y, une couleur foncée à la position (x, y) indique que l'erreur est faible tandis qu'une couleur claire traduit une erreur élevée. (D) Interface d'interaction entre le C-terminal de REC114 et le N-terminal de MEI4. (E) Détails de l'interaction entre l'hélice  $\alpha 2$  de MEI4 et le monomère de REC114.

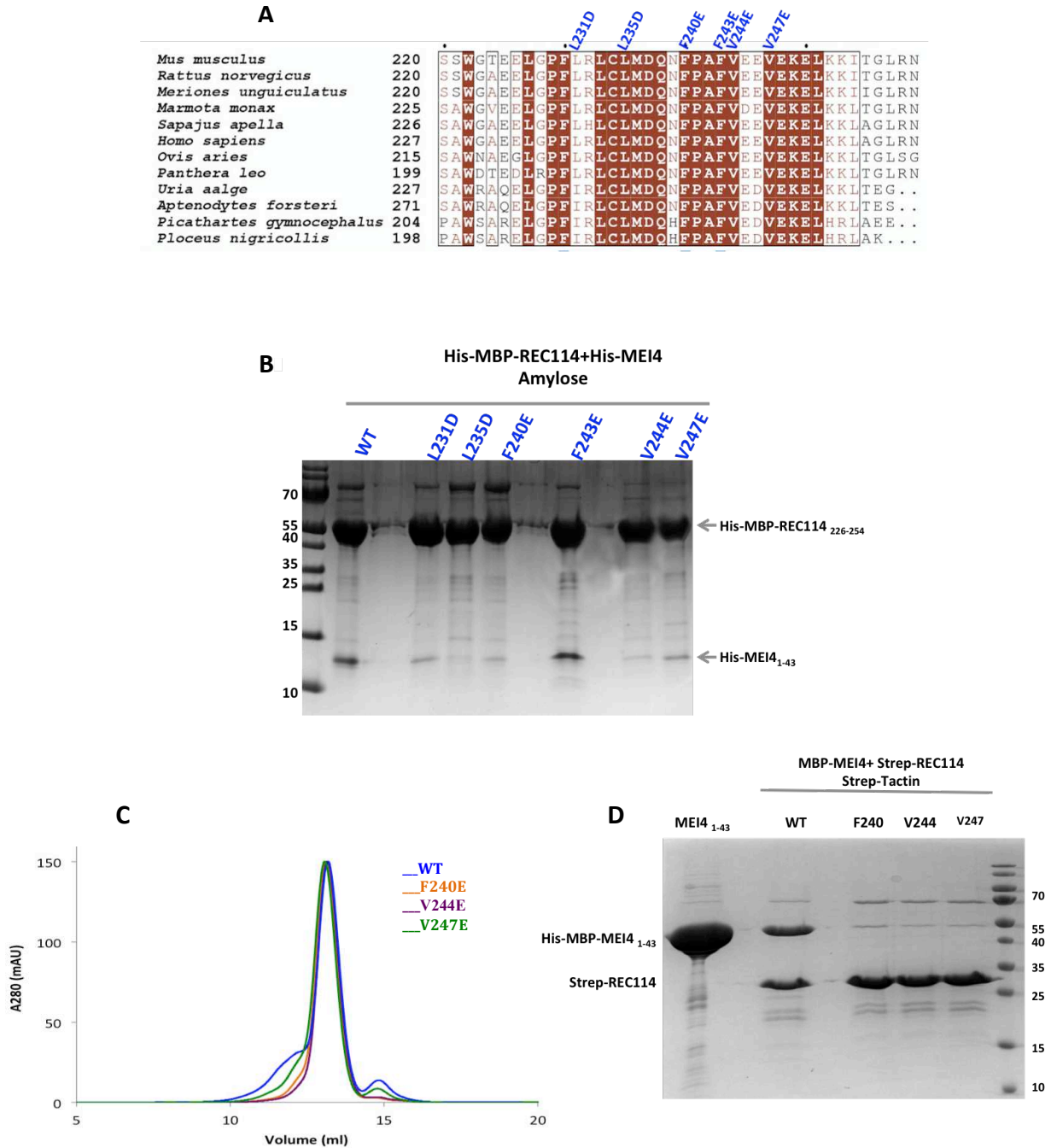
### 3.2.4 Interface d'interaction entre REC114 et MEI4

L'étape suivante consistait à valider le modèle AlphaFold2 décrit ci-dessus et à identifier les résidus clés dont la mutation pourrait perturber l'interaction entre REC114 et MEI4 et qui pourraient à l'avenir être utilisés dans une analyse fonctionnelle *in vivo*. Pour cette raison, plusieurs résidus ont été mutés sur la base de la conservation des séquences protéiques et du modèle structural AlphaFold2.

Du côté REC114, nous avons décidé de muter les résidus conservés L231, L235, F240, V244 et V247 (Figure 37 A). Ceux-ci semblent être importants pour la liaison de MEI4. Le résidu F243 a également été muté comme un contrôle, car il est supposé ne pas perturber l'interaction entre les deux protéines. De simples mutations de ces résidus en acide glutamique ou aspartique ont été introduites dans la construction His-MBP-REC114<sub>226-254</sub>. Ces mutants ont été co-exprimés avec la construction His-MEI4<sub>1-43</sub>. L'échantillon protéique a été co-purifié sur la résine amylose. L'analyse par électrophorèse des fractions de la co-elution des échantillons montre que la présence de la mutation : L231D, L235D, F240E, V244E ou V247E réduit significativement la quantité de la protéine MEI4 co-éluée avec REC114 (Figure 37 B). Ce qui indique que ces résidus sont importants pour la formation du complexe binaire. On a également vérifié par la filtration sur gel, pour les mutants F240E, V244E, et V247E, si la perte de l'interaction n'est pas provoquée par la perturbation du repliement de la protéine. Le chromatogramme obtenu pour les protéines mutées est similaire à celui de la protéine native (Figure 37 C), indiquant que les mutations n'affectent pas le repliement de la protéine. Comme attendue, la mutation du résidu F243 (Figure 37 B), qui a été utilisé comme un contrôle, n'affecte pas l'interaction entre les deux protéines. Ces résultats, *in vitro*, sont conformes à la prédiction du modèle AlphaFold2.

Pour confirmer leurs rôles dans le contexte de la protéine entière, les résidus F240, V244, et V247 ont été mutés séparément dans la construction d'ADN codant la protéine entière REC114 (Strep-REC114<sub>FL</sub>). L'effet de ces mutations dans cette construction a été testé par *pull down*. Pour cela, la protéine His-MBP-MEI4 a été purifiée sur la résine amylose, puis incubée avec des résines Strep-tactin sur lesquelles sont immobilisés les





**Figure 37. Analyse de l'interaction entre les mutants de REC114 et MEI4. (A)** Alignement de séquences de l'extrémité C-terminale de REC114. Les résidus en brun sont des acides aminés équivalents et ceux en blanc sont des acides aminés identiques. Les résidus mutés sont indiqués en bleu. **(B)** Analyse par SDS-PAGE des éluions de la co-purification sur la colonne d'amylose de la protéine His-MBP-REC114<sub>226-254</sub> native (WT), et celles mutées, avec His-MEI4<sub>1-43</sub>. **(C)** Superposition des profils d'éluion en chromatographie d'exclusion stérique de la protéine His-MBP- REC114<sub>226-254</sub> WT (bleu), et la protéine mutées (F240E : orange, V244E: violet, V247E : vert). **(D)** Analyse par SDS-PAGE des éluions de *pull down* entre les mutants de REC114<sub>FL</sub> et MEI4. La protéine MBP-MEI4<sub>1-43</sub> a été purifiée sur une colonne amylose (ligne 1) puis incubée avec de la résine Strep-Tactin sur laquelle est liée la protéine REC114 WT ou mutée.



différents mutants de Strep-REC114. L'analyse par SDS-PAGE des fractions d'élution montre que toutes les mutations de REC114<sub>FL</sub> abolissent l'interaction avec MEI4, confirmant l'importance de ces résidus dans la formation du complexe binaire (Figure 37 D). La mutation du résidu équivalent à F240 dans REC114 de *S. cerevisiae* (F411A) a également perturbé l'interaction avec son partenaire MEI4 (Claeys Bouuaert et al., 2021a), suggérant l'importance de ces résidus dans la formation du complexe binaire chez les deux espèces.

Nous nous sommes ensuite intéressés à l'identification des mutations dans la protéine MEI4 qui pourraient perturber son interaction avec REC114. Pour cela, les résidus L13, A14, L15 et A16 (Figure 38 A) ont été mutés séparément en acides glutamiques dans la construction His-MBP-MEI4<sub>1-43</sub> et co-exprimés avec la construction de His-REC114<sub>226-254</sub>. Après co-purification sur la résine amylose, l'analyse par SDS-PAGE des fractions de co-élution montre que la protéine MEI4 n'est plus co-purifiée avec REC114 en présence des mutations A16E et L15E (Figure 38 B). Ce qui indique que ces mutations abolissent l'interaction entre les deux protéines et reflète l'importance des résidus L15 et A16 pour la formation du complexe REC114-MEI4. Nous avons également confirmé par la filtration sur gel que la mutation A16E ne perturbe pas la structure 3D de MEI4 (Figure 38 C). Selon le modèle AlphaFold2, les résidus L15 et A16, situés sur l'hélice  $\alpha$ 1, interagissent avec le même monomère de REC114 que la deuxième hélice de MEI4 ( $\alpha$ 2) (Figure 36 E). La mutation de ces résidus déstabiliserait également les contacts importants apportés par cette hélice, conduisant à la perte d'interaction entre le dimère de REC114 et MEI4.

La mutation des résidus L13 ou A14 qui interagiraient avec le deuxième monomère de REC114 ne montre pas d'effet sur l'interaction entre les deux protéines par cette technique. L'explication la plus probable est que ces mutations ne font que réduire la liaison de ce monomère, et ce qu'on observe sur le gel est un complexe REC114-MEI4 d'une stoechiométrie 1 : 1. La différence entre la stoechiométrie 1 : 1 et 2 : 1 du complexe REC114<sub>226-254</sub>-MEI4<sub>1-43</sub> serait difficile à distinguer sur le gel, surtout compte tenu de la petite taille du fragment REC114. Un mutant avec la double mutation de ces deux résidus, L13EA14E, a été généré aussi et co-exprimé par His-REC114<sub>226-254</sub>. La co-élution de cet échantillon semble montrer une réduction de His-REC114<sub>226-254</sub> co-purifiée avec MEI4 doublement mutée comparée à la protéine MEI4 native (Figure 38 D). Ce résultat devrait



être confirmé par une approche plus quantitative pour identifier l'effet de ces mutations sur le complexe REC114-MEI4.

Dans l'ensemble, ces résultats fournissent les premières informations structurales sur le complexe REC114-MEI4. Le modèle AlphaFold2 est prédit avec une grande confiance et est confirmé par nos analyses de mutagenèses. Il n'est toujours pas clair si les deux mutants de MEI4 (L13E, A14E) n'ont pas d'effet sur l'interaction avec REC114. Il serait possible qu'ils perturbent la liaison d'une seule molécule REC114, mais que le complexe 1REC114-1MEI4 reste stable. De plus, la modélisation AlphaFold2 ne nous a pas permis jusqu'à présent de modéliser de manière convaincante le complexe 4REC114-2MEI4.

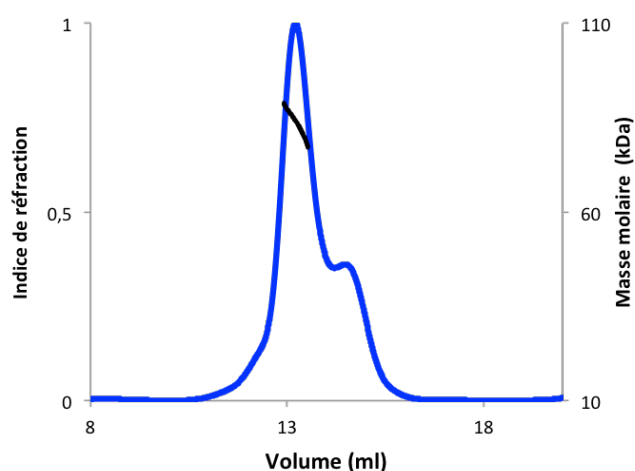


### 3.3 La partie C-terminale de REC114

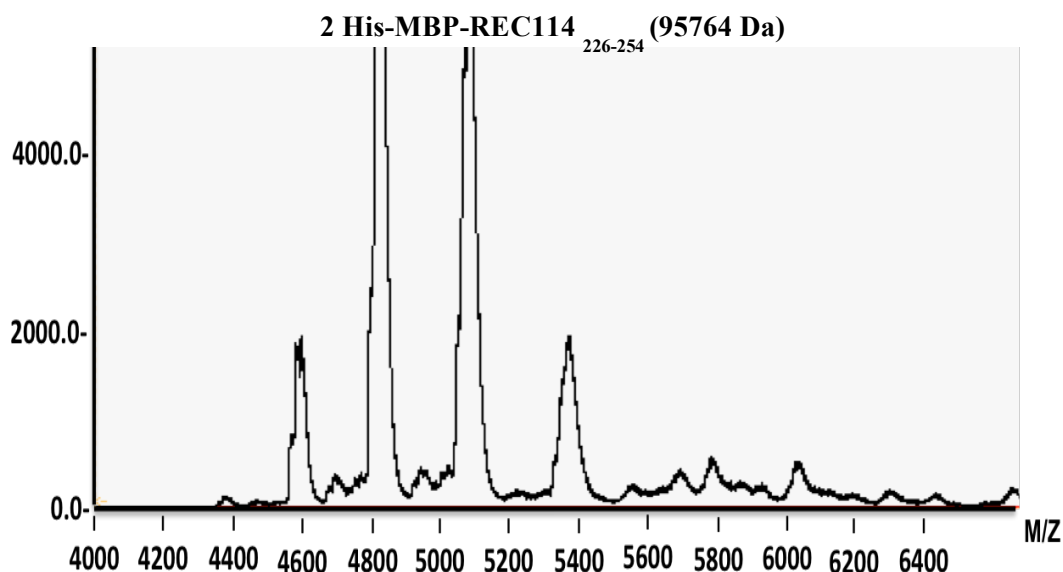
#### 3.3.1 Caractérisation de l'état oligomérique de la partie C-terminale de REC114

Comme nous l'avons vu par différentes méthodes dans la section précédente, le complexe REC114-MEI4 est principalement composé d'un dimère de REC114 lié à un monomère de MEI4. Nous avons alors cherché à comprendre si cette dimérisation de la partie C-terminale de REC114 dépendait de la présence de la partie N-terminale de MEI4. Pour cela, l'échantillon protéique His-MBP-REC114<sub>226-254</sub>, qui correspond aux deux hélices C-terminales de REC114, a été élué sur la résine amylose puis analysé par la technique de SEC-MALLS et de MS native pour contrôler son état oligomérique.

La mesure SEC-MALLS montre la présence d'un pic principal de 92 kDa (Figure 39). La comparaison de cette dernière avec la masse théorique de His-MBP-REC114<sub>226-254</sub> (48 kDa) suggère une forme dimérique pour cet échantillon en solution. Les données obtenues par la MS à l'état natif ont confirmé ce qui avait été observé avec la technique de SEC-MALLS. La masse mesurée par cette méthode est de 95,7 kDa (Figure 40), ce qui est compatible avec la présence de dimères en solution. Cela signifie que la présence de MEI4 n'est pas nécessaire pour la dimérisation de la partie C-terminale de REC114. De même, la dimérisation du fragment C-terminal de Rec114 de la levure *S.cerevisiae* n'a pas nécessité Mei4 (Claeys Bouuaert et al., 2021a). Cela traduit la conservation des caractéristiques des deux protéines chez les deux espèces.



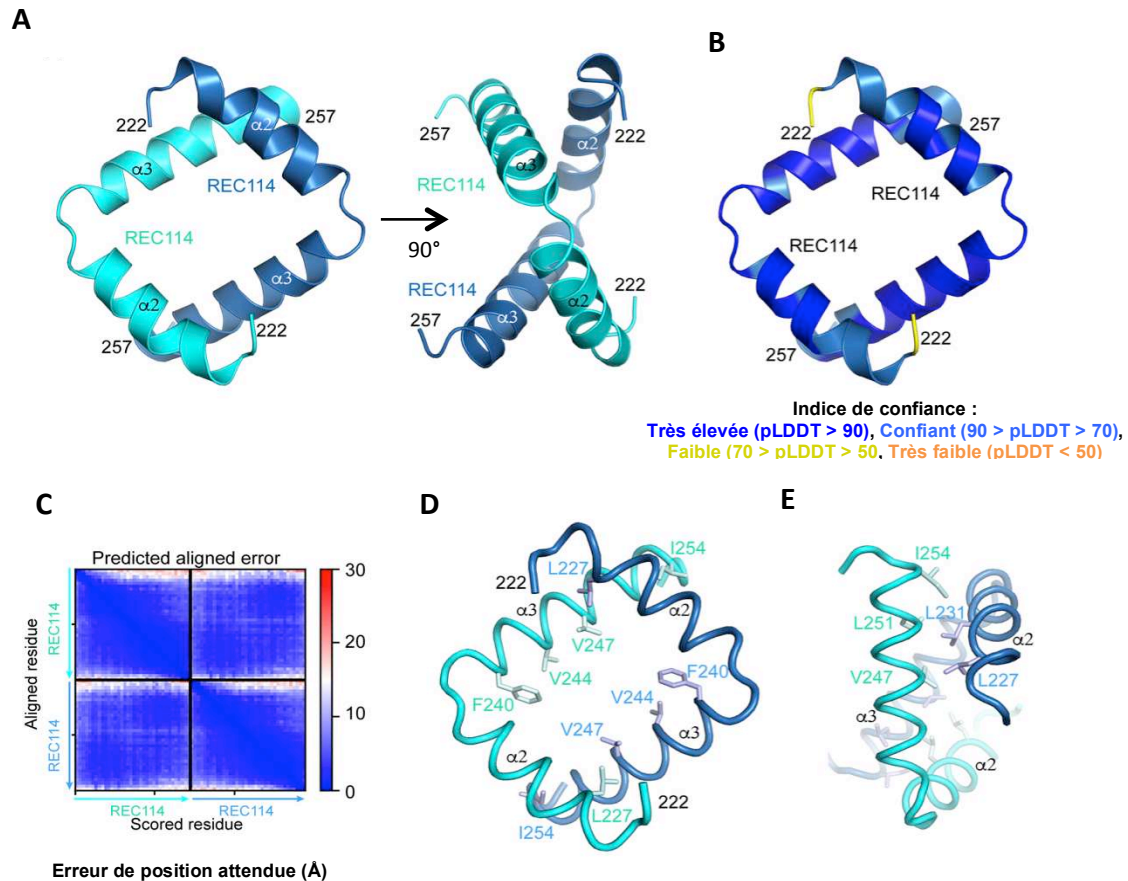
**Figure 39. Caractérisation de l'état oligomérique de la partie C terminale de REC114 par SEC-MALLS.** Analyse de la protéine His-MBP-REC114<sub>226-254</sub> par la technique de SEC-MALLS. La courbe bleue représente l'indice de réfraction en fonction du volume d'élué. La masse moléculaire calculée (en kDa) sous le pic d'élué est représentée en trait noir en fonction du volume d'élué en ml.



**Figure 40. Caractérisation de l'état oligomérique de la partie C-terminale de REC114 par MS native.** Analyse de la protéine His-MBP-REC114<sub>226-254</sub> par la spectrométrie de masse en condition non-dénaturante.

### 3.3.2 Modèle structural par AlphaFold2 pour le domaine C-terminal de REC114

Pour obtenir des informations structurales sur le dimère REC114<sub>226-254</sub>, nous avons modélisé sa structure avec AlphaFold2 (Figure 41 A). Le modèle obtenu a été prédit avec un très bon score de confiance (Figure 41 B, C) et il montre que chaque monomère est constitué de deux hélices de longueurs équivalentes séparées par une boucle (Figure 41 A). Les deux monomères adoptent une configuration antiparallèle comparable à celle observée en présence de MEI4. Selon ce modèle structural, les résidus F240, V244, V247 nécessaires à l'interaction avec MEI4 semblent éloigner de l'interface de dimérisation (Figure 41 D). Les contacts entre les deux monomères sont assurés par les résidus situés aux extrémités des hélices (Figure 41 E).



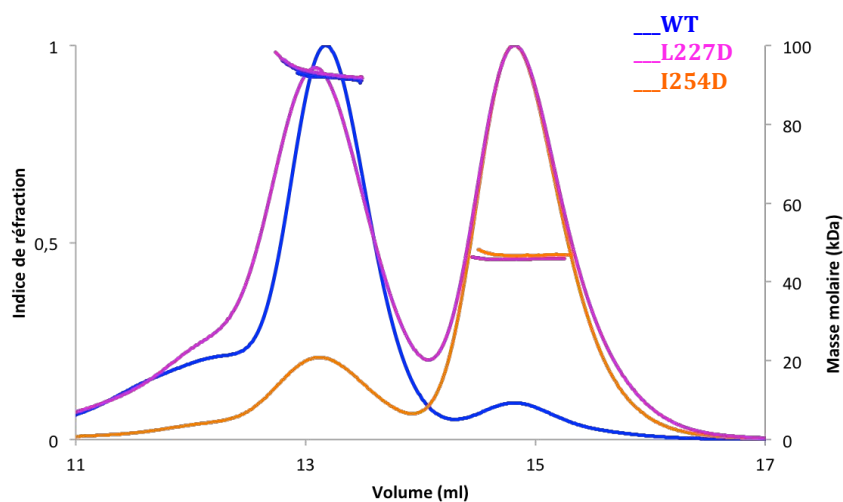
**Figure 41. Modèle structural du dimère formé par le C-terminus de REC114 murin.** Modélisation du dimère REC114<sub>222-257</sub> avec AlphaFold2 (Jumper et al., 2021). Le modèle représenté en (A) est coloré en suivant le code couleur de la Figure 30 et celui en (B) selon de la valeur de pLDDT (score de confiance de la prédiction de chaque résidu). (C) Graphique indiquant l'erreur sur distance (en Ångströms) attendue dans la prédiction de la position du résidu X par rapport au résidu Y, une couleur foncée à la position (x, y) indique que l'erreur est faible tandis qu'une couleur claire traduit une erreur élevée. (D) Détails du dimère REC114<sub>222-257</sub> mettant en évidence les résidus impliqués dans l'interaction avec MEI4 (F240, V244 et V247) et ceux qui semblent importants pour la dimérisation à savoir (L227 et I254). (E) Interface d'interaction entre deux hélices de REC114<sub>222-257</sub>.

### 3.3.3 Interface de dimérisation de la partie C-terminale de REC114

Pour caractériser l'interface de dimérisation, la construction de la partie C-terminale de REC114 a été étendue aux résidus (222-256) car le modèle suggère, avec une faible confiance, des contacts supplémentaires entre des résidus additionnels. Les résidus qui semblent être impliqués dans l'interface de dimérisation (L227 et I254) ont été mutés, en acides aspartiques, et exprimés en fusion avec His-MBP. Les résidus prédits pour être en dehors du site de dimérisation, et identifiés précédemment pour empêcher la fixation de MEI4, ont été utilisés comme témoins (F240, V244, V247). L'échantillon His-MBP-

REC114<sub>227-259</sub> WT et ceux des protéines mutées ont été préparés à une concentration de 3 mg/ml et analysés par la technique de SEC-MALLS.

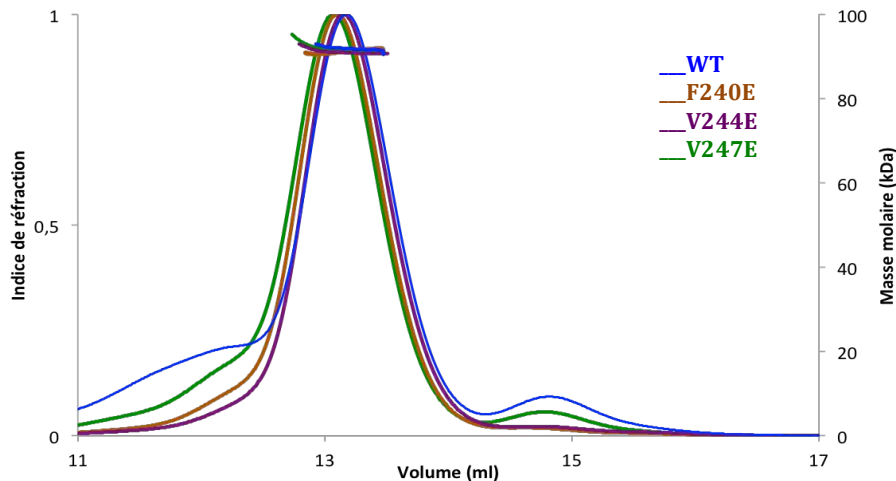
Concernant le mutant L227D, le résultat obtenu par la technique de SEC-MALLS montre que ce mutant s'élue en deux pics d'intensités équivalentes. Le premier pic a été observé à 13,1 ml et correspond à une masse de 93,5 +/- 0,9 kDa, indiquant l'état dimérique de l'échantillon à ce pic. La deuxième élution de 14,9 ml correspond à une population de 46,4 kDa et est compatible avec une forme monomérique (Figure 42). Ce résultat traduit une perte partielle de la dimérisation de la partie C-terminale de REC114 en présence de cette mutation. Le mutant I254D a également été analysé par SEC-MALLS. Le profil d'élution par filtration sur gel montre la présence d'un pic majoritaire à un volume d'élution de 14,9 ml. Le poids moléculaire calculé pour les espèces du pic principal est de 46 kDa (Figure 42), ce qui est compatible avec la présence de monomères de REC114 C-terminal porteur de la mutation I254D. D'après ces résultats, les mutations conçues sur la base du modèle AlphaFold2 ont affecté la dimérisation de la partie C-terminale de REC114. Par conséquent, ce modèle nous a permis d'identifier des résidus importants dans la formation d'un dimère entre les parties C-terminales de REC114.



**Figure 42. Analyse par la technique SEC-MALLS des mutants de REC114 C-terminal.** (A) Résultat de l'analyse par SEC-MALLS de la protéine His-MBP-REC114<sub>226-254</sub> WT (bleu), ou mutée sur les résidus qui semblent importants pour la dimérisation (L227D : rose et I254 : orange). Ces expériences ont été effectuées à 3 mg/ml. Le chromatogramme de l'indice de réfraction et de la masse moléculaire (kDa) en fonction du volume d'élution en ml est représenté en suivant le même code couleur.



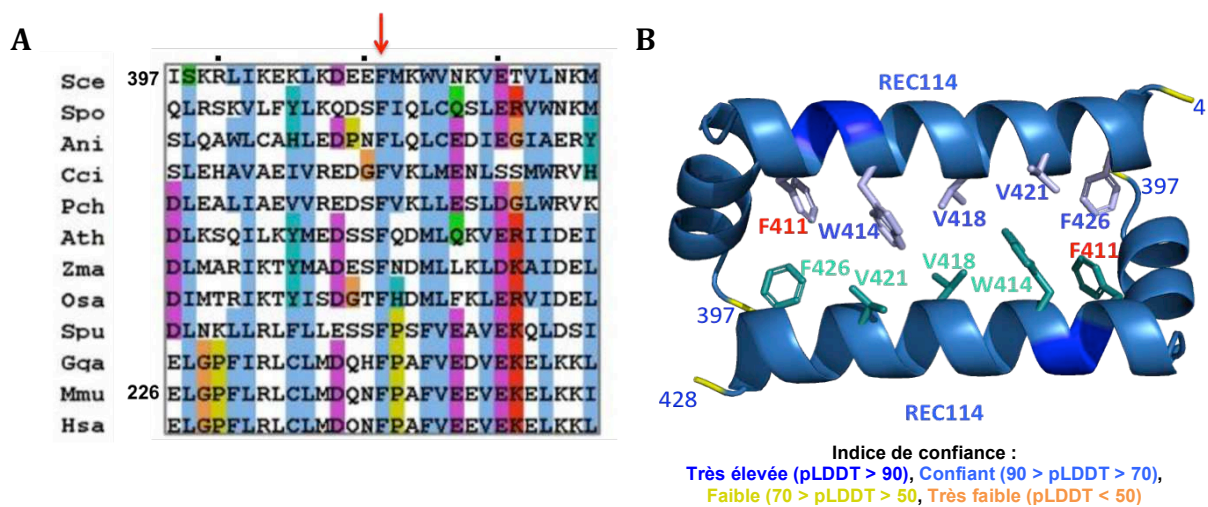
D'autre part, la chromatographie d'exclusion stérique a montré une élution de la protéine WT des mutants F240E, V244E et V247E dans le même volume d'élution. La masse calculée en dessous de ces pics d'élution, de 95 kDa, correspond à la masse d'un dimère (Figure 43). Ce résultat indique que ces mutations n'affectent pas la dimérisation de REC114 C-terminale, ce qui est cohérent avec l'emplacement de ces résidus en dehors de la région de dimérisation prédite par AlphaFold2.



**Figure 43. Analyse par la technique de SEC-MALLS de la partie C-terminale de REC114** Superposition des résultats de la technique SEC-MALLS pour la protéine His-MBP- REC114<sub>226-254</sub> WT (bleu), ou mutée sur les résidus qui ne sont pas censés abolir la dimérisation (F240E : marron, V244E : violet, V247E : vert). Ces expériences ont été effectuées à 3 mg/ml. Le chromatogramme de l'indice de réfraction et de la masse moléculaire (kDa) en fonction du volume d'élution en ml est représenté en suivant le même code couleur.

Cependant, la mutation du résidu équivalent à F240 chez *S. cerevisiae* (F411A, Figure 44 A) a suffi pour abolir la dimérisation de REC114 chez cette espèce (Claeys Bouuaert et al., 2021a). Il faut noter que cette similarité est basée sur l'alignement, publié par Kumar *et al.*, 2010, qui montre une faible conservation entre les séquences de souris et de levures. Pour comprendre cette différence entre les deux espèces, nous avons modélisé la structure de Rec114 C-ter de *S. cerevisiae* par AlphaFold2. Le résultat obtenu montre que la partie structurée C-terminale de Rec114 de *S. cerevisiae* est également constituée de deux hélices. Selon le modèle AlphaFold2, la dimérisation du fragment C-terminal de Rec114 de levure fait interagir les premières hélices de chaque monomère sur leur longueur, et le résidu F411A se trouverait dans cette interface d'interaction (Figure 44 B), alors que le modèle de REC114 murin montre que le résidu F240 est éloigné de la région

de dimérisation formée par des contacts entre les extrémités des hélices de chaque monomère. Les différences conformationnelles observées dans le modèle structural de la partie C-terminale de REC114 entre les deux espèces pourraient expliquer les différences observées dans l'effet des mutations de ces résidus équivalents. Néanmoins, si on considère que le modèle du dimère Rec114 de levure est correct, il semble peu probable que la mutation de F411 en Alanine soit suffisante pour perturber complètement l'interface de dimérisation.



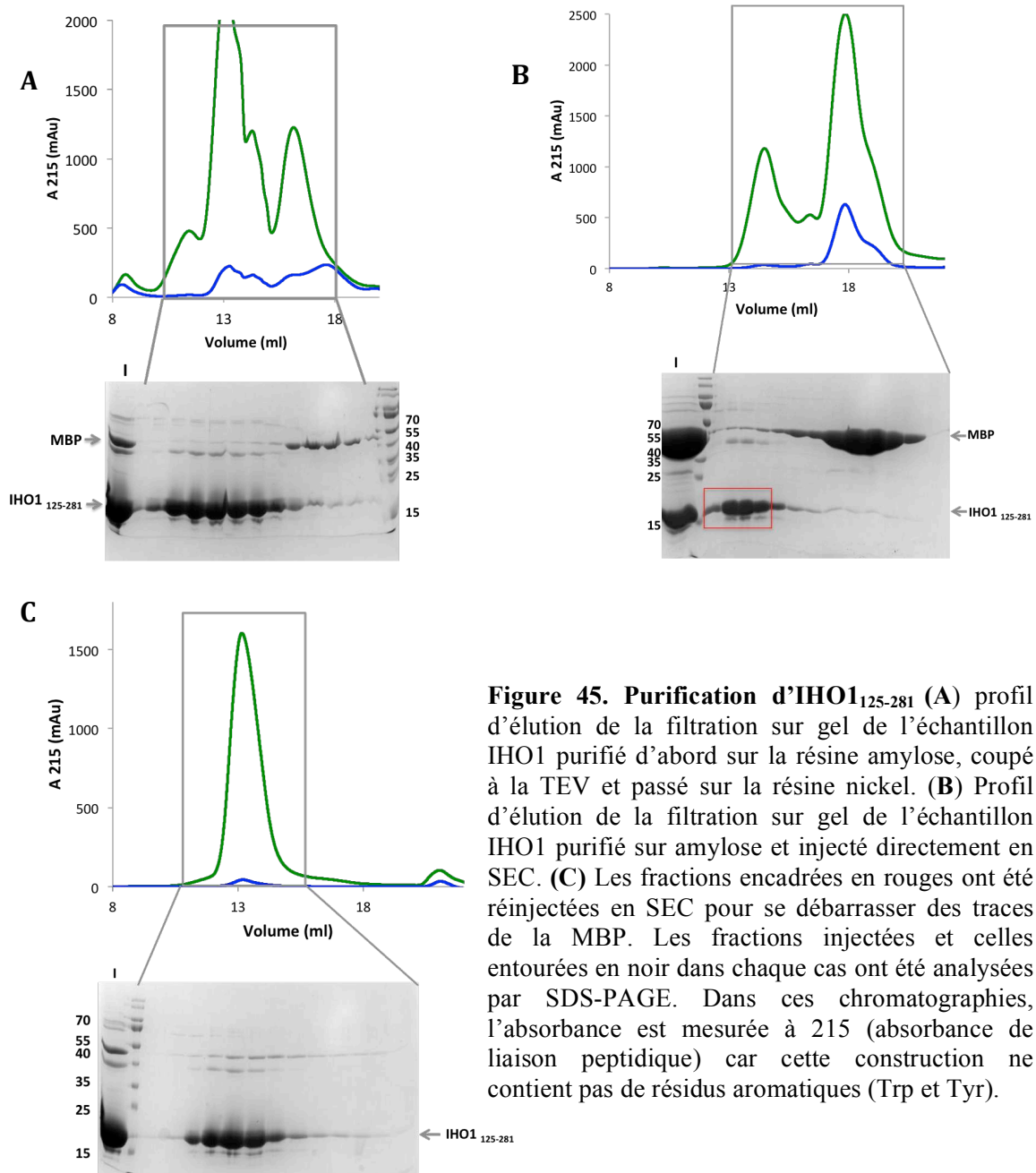
**Figure 44. Modèle structural du dimère formé par le C-terminus de REC114 de *S. cerevisiae*.** (A) Alignement de la partie C-terminale de REC114 entre les différentes espèces eucaryotes (extrait de Kumar et al., 2010). (Sce) *Saccharomyces cerevisiae*; (Spo) *Schizosaccharomyces pombe*; (Ani) *Aspergillus nidulans*; (Cci) *Coprinopsis cinerea*; (Pch) *Phanerochaete chrysosporium*; (Ath) *Arabidopsis thaliana*; (Zma) *Zea mays*; (Osa) *Oryza sativa*; (Spu) *Strongylocentrotus purpuratus*; (Gga) *Gallus gallus*; (Mmu) *Mus musculus*; (Hsa) *Homo sapiens*. Le résidu F411 de *S. cerevisiae* qui est l'équivalent du résidu F240 chez la souris est indiqué par flèche rouge. (B) Modélisation du dimère REC114<sub>397-428</sub> avec AlphaFold2 (Jumper et al., 2021). Les deux monomères de REC114 sont colorés en fonction de la valeur de pLDDT. Les résidus qui semblent importants pour la dimérisation de cette protéine sont indiqués. Le résidu dont la mutation empêche la dimérisation de cette protéine (F411) est indiqué en rouge.

## 3.4 Caractérisation d'IHO1

### 3.4.1 Purification d'IHO1<sub>125-281</sub>

Le dernier composant du complexe REC114-MEI4-IHO1 que j'ai tenté de caractériser est la région *coiled-coil* d'IHO1. La caractérisation de la partie centrale d'IHO1, IHO1<sub>125-281</sub>, a été initiée avant mon arrivée dans l'équipe. Ce fragment a été identifié en utilisant une protéolyse limitée avec la trypsine, et couvre la partie prédite comme *coiled-coil*. Plusieurs essais de cristallisation ont été effectués dans le but de déterminer la structure de ce fragment, mais aucun cristal n'a pu être obtenu.

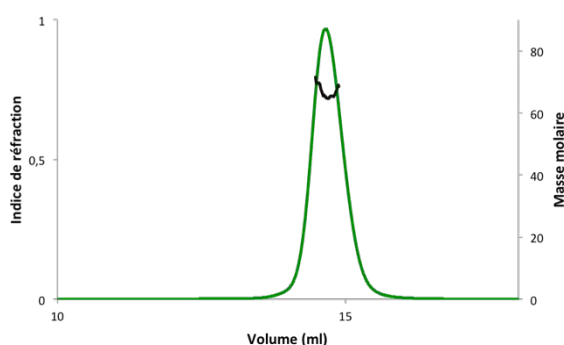
Pour continuer la caractérisation de cette protéine, j'ai suivi dans un premier temps le protocole de purification établi dans l'équipe. Après une purification sur la résine amylose et clivage de la His-MBP par la TEV, l'échantillon protéique a été rechargé sur la résine nickel pour se débarrasser de l'étiquette His-MBP et de la protéine éventuellement non clivée. La protéine a été ensuite injectée sur une colonne de SEC. Néanmoins, le chromatogramme obtenu présente plusieurs pics très larges suggérant une hétérogénéité de l'échantillon (Figure 45 A). Ce comportement de la protéine n'est pas observé avec l'échantillon injecté en SEC en présence de la MBP. Cela suggère que le clivage de la MBP ou le passage sur la résine de nickel n'est pas adapté à cette construction. Pour comprendre l'origine de cet élargissement des pics d'éluion, l'échantillon protéique a été purifié par gel-filtration directement après le clivage de la MBP, sans l'étape de la deuxième chromatographie d'affinité sur résine nickel. Le résultat obtenu montre que la protéine est éluée dans un seul pic, mais elle présente une contamination à la MBP (Figure 45 B). Une deuxième étape chromatographie d'exclusion stérique a été effectuée sur cet éluât, ce qui a permis de se débarrasser de la majorité de la contamination par la MBP et d'obtenir un échantillon homogène avec une pureté satisfaisante (Figure 45 C). La suppression de la deuxième chromatographie d'affinité sur la résine nickel a permis de récupérer un échantillon protéique dans un seul pic symétrique en sortie de la filtration sur gel. Une des explications possibles à ce comportement est que la protéine fixe le nickel de manière non-spécifique, ce qui peut conduire à une hétérogénéité conformationnelle, qui peut entraîner des pics très larges lors de la chromatographie d'exclusion de taille. Ce protocole de purification a permis de récupérer un échantillon, plus stable en absence de la MBP comparé à IHO1<sub>1-281</sub>, qui semble homogène pour les analyses structurales.



**Figure 45. Purification d'IHO1<sub>125-281</sub>** (A) profil d'éluion de la filtration sur gel de l'échantillon IHO1 purifié d'abord sur la résine amylose, coupé à la TEV et passé sur la résine nickel. (B) Profil d'éluion de la filtration sur gel de l'échantillon IHO1 purifié sur amylose et injecté directement en SEC. (C) Les fractions encadrées en rouges ont été réinjectées en SEC pour se débarrasser des traces de la MBP. Les fractions injectées et celles entourées en noir dans chaque cas ont été analysées par SDS-PAGE. Dans ces chromatographies, l'absorbance est mesurée à 215 (absorbance de liaison peptidique) car cette construction ne contient pas de résidus aromatiques (Trp et Tyr).

### 3.4.2 Caractérisation d'IHO1<sub>125-281</sub>

Comme nous l'avons vu précédemment par la technique de SEC-MALLS, la protéine IHO<sub>1-281</sub> est tétramérique en solution, nous avons voulu confirmer si la construction tronquée de la partie N-terminale IHO<sub>125-281</sub>, conserve cet état oligomérique. Des mesures par SEC-MALLS ont été effectuées pour cette construction. La masse calculée de 70 kDa (Figure 46), est compatible avec une forme tétramérique. Ce qui suggère que c'est la partie centrale qui est responsable de la tétramérisation de cette



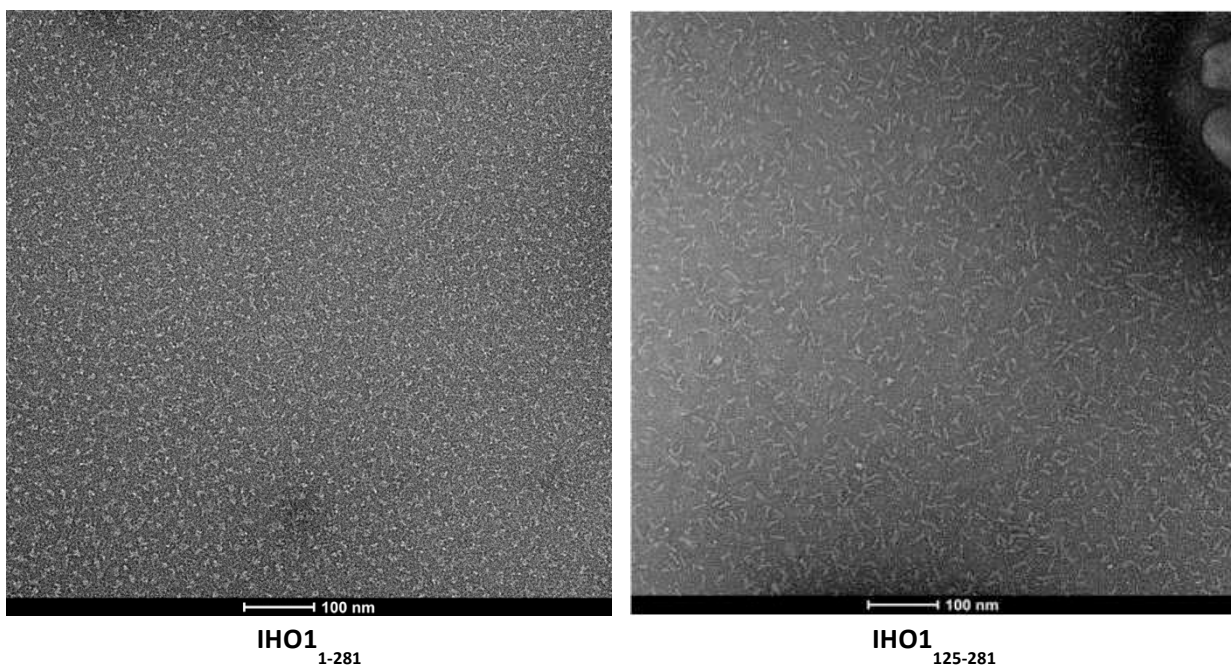
**Figure 46. Caractérisation de l'état oligomérique d'IHO1<sub>125-281</sub>.** La partie centrale d'IHO1 a été analysée par SEC-MALLS. La réfractométrie et la masse calculée sont représentées en fonction du volume d'éluion.

protéine.

Notre première approche pour obtenir les détails structuraux du tétramère d'IHO1 a été d'utiliser la cristallographie aux rayons X. L'échantillon protéique purifié a été préparé à différentes concentrations (5 mg/ml, 10 mg/ml et 15 mg/ml) et utilisé pour cribler les conditions de cristallisation à l'aide des kits commerciaux (sur la plateforme de cristallisation haut débit HTX de l'EMBL). Ceux-ci permettent de tester un grand nombre de combinaisons entre agents précipitants, pH, additifs et tampons. Ces tests de cristallisation ont été réalisés à 20 °C et 4 °C. L'évolution des gouttes a été régulièrement suivie, mais ces tests de cristallisation n'ont montré aucun cristal ou précipité microcristallin.

Puisqu'il est si difficile d'obtenir des cristaux, nous nous sommes tournés vers la microscopie électronique pour obtenir les caractéristiques structurales de la protéine IHO1. Un échantillon protéique de la construction la plus longue de IHO1 (IHO<sub>1-281</sub> 128 kDa pour un tétramère) a été préparé et visualisé par la coloration négative. Les micrographes

obtenus montrent la présence de particules allongées, de taille et de forme très hétérogènes (Figure 47). En revanche, quand l'échantillon protéique de la partie centrale (IHO1<sub>125-281</sub>) a été visualisé par la coloration négative, des particules assez homogènes avec une structure allongée et en fibres ont été observées (Figure 47). Cette différence entre les deux échantillons est probablement provoquée par la partie N-terminale qui serait trop flexible, gênant ainsi la visualisation de fibres homogènes. Ces fibres sont comparables aux structures allongées, observées dans les images de coloration négatives du complexe

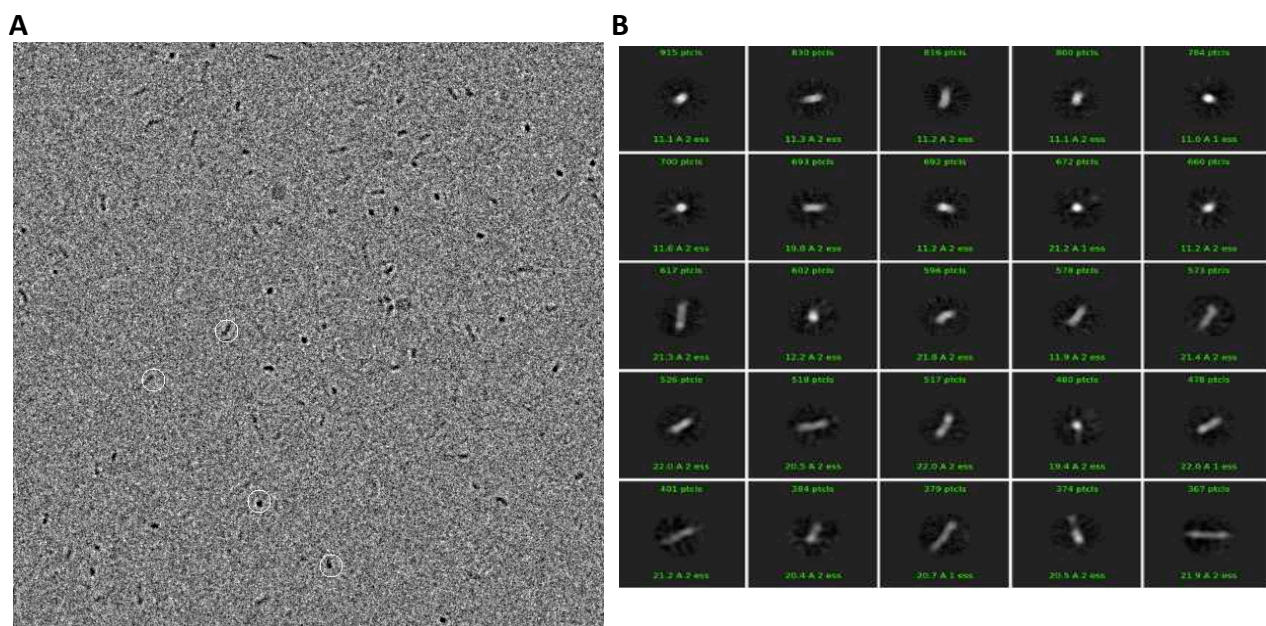


**Figure 47 : Coloration négative d'IHO1<sub>125-281</sub>.** Micrographes obtenus par coloration négative pour la construction IHO1<sub>1-281</sub> et la construction dépourvue de la partie N-terminale IHO1<sub>125-281</sub> REC114-MEI4-IHO1 (Figure 21).

Les premières images de la coloration négative de la partie IHO1<sub>125-281</sub> étaient très encourageantes et ouvraient la perspective pour l'analyse structurale de la protéine par cryo-microscopie électronique (Cryo-EM). Cependant, en raison de la petite taille de cette protéine (70 kDa), cette approche en elle-même est un défi. En effet, la caractérisation à haute résolution des structures de petites particules (moins de 100 kDa) est beaucoup plus difficile en raison du faible contraste et du faible rapport signal sur bruit. Néanmoins, grâce à diverses avancées dans la préparation des échantillons, l'acquisition et le traitement d'images au cours des dernières années, des structures de plus en plus petites de molécules biologiques ont été rapportées.

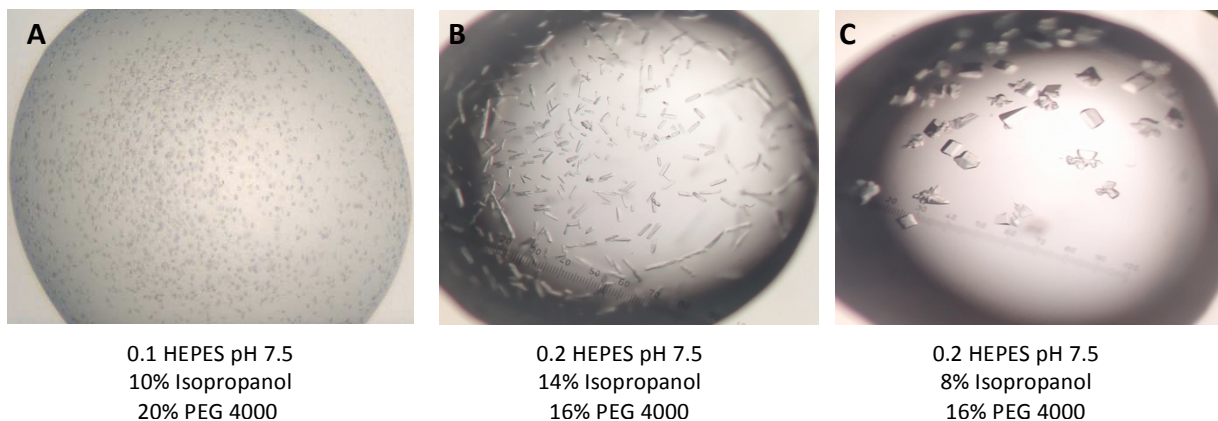


Un échantillon protéique de la construction IHO1<sub>125-281</sub> a été préparé pour collecter des données de Cryo-EM. Un jeu de données d'environ 1000 micrographes a été collecté au microscope Glacios de l'EMBL à Grenoble. Contrairement aux images de la coloration négative, les particules obtenues sur les micrographes de cryo-EM sont de tailles et de formes hétérogènes (Figure 48 A). Les particules semblent ne pas survivre au processus de la vitrification d'échantillon. Les classes 2D obtenues à partir de cet ensemble de données sont de qualité très limitée et montrent la présence de particules de différentes tailles (Figure 48 B). Les modèles 3D *ab-initio* qui ont été calculés par cryoSPARC sont très limités en résolution. Nous n'avons pas poursuivi la caractérisation par Cryo-EM, car ces premiers résultats n'étaient pas très prometteurs, mais aussi parce que notre équipe n'avait, à ce moment là, ni l'expertise ni l'accès suffisant aux microscopes Cryo-EM.



**Figure 48. Données de la cryo-microscopie électronique de la protéine IHO1<sub>125-281</sub>.** (A) exemple de micrographie de la protéine IHO1<sub>125-281</sub> en Cryo-EM. (B) exemples de 2D obtenues par cryoSPARC après le traitement des données.

Je suis donc revenue à la cristallographie aux rayons X pour essayer de déterminer la structure d'IHO1. J'ai décidé de tester plusieurs constructions qui resteraient tétramérique et ne contiendraient que la région *coiled-coil*, qui était cependant difficile à prédire de façon précise. Par conséquent, j'ai produit quatre autres constructions: IHO<sub>143-240</sub>, IHO<sub>143-257</sub>, IHO<sub>175-240</sub>, IHO<sub>175-257</sub>, et le meilleur échantillon a été obtenu avec la construction His-MBP-IHO<sub>143-257</sub>. Cette dernière a été purifiée avec le protocole établi pour la construction IHO<sub>125-281</sub>. Après avoir obtenu un échantillon d'une pureté satisfaisante, elle a été utilisée pour des essais de cristallisation. Plusieurs petits cristaux ont été obtenus dans différentes conditions, notamment dans 100 mM HEPES pH 7.5, 10 % Isopropanol et 20 % PEG 4000 (Figure 49 A). Ces derniers ont été reproduits en plus grandes tailles lors de l'optimisation manuelle dans deux conditions. Dans la condition de 200 mM HEPES pH 7.5, 14% Isopropanol et 16 % PEG 4000, les cristaux apparaissent sous forme de petites baguettes de taille hétérogènes avec plusieurs événements de nucléation conduisant à la formation de plusieurs microcristaux (Figure 49 B). Les cristaux formés dans la condition de 200 mM HEPES pH 7.5, 8 % Isopropanol et 16 % PEG 4000 sont plus gros, mais leurs formes est très irrégulières (Figure 49 C). Ces cristaux ont été testés, mais aucune diffraction n'a été observée. La cristallisation a été testée à 4 °C pour ralentir la croissance cristalline et améliorer la qualité des cristaux, sans résultat. Plusieurs conditions de cristallisation ont été testées et des additifs ont été ajoutés, mais la qualité des cristaux n'a pas pu être améliorée.

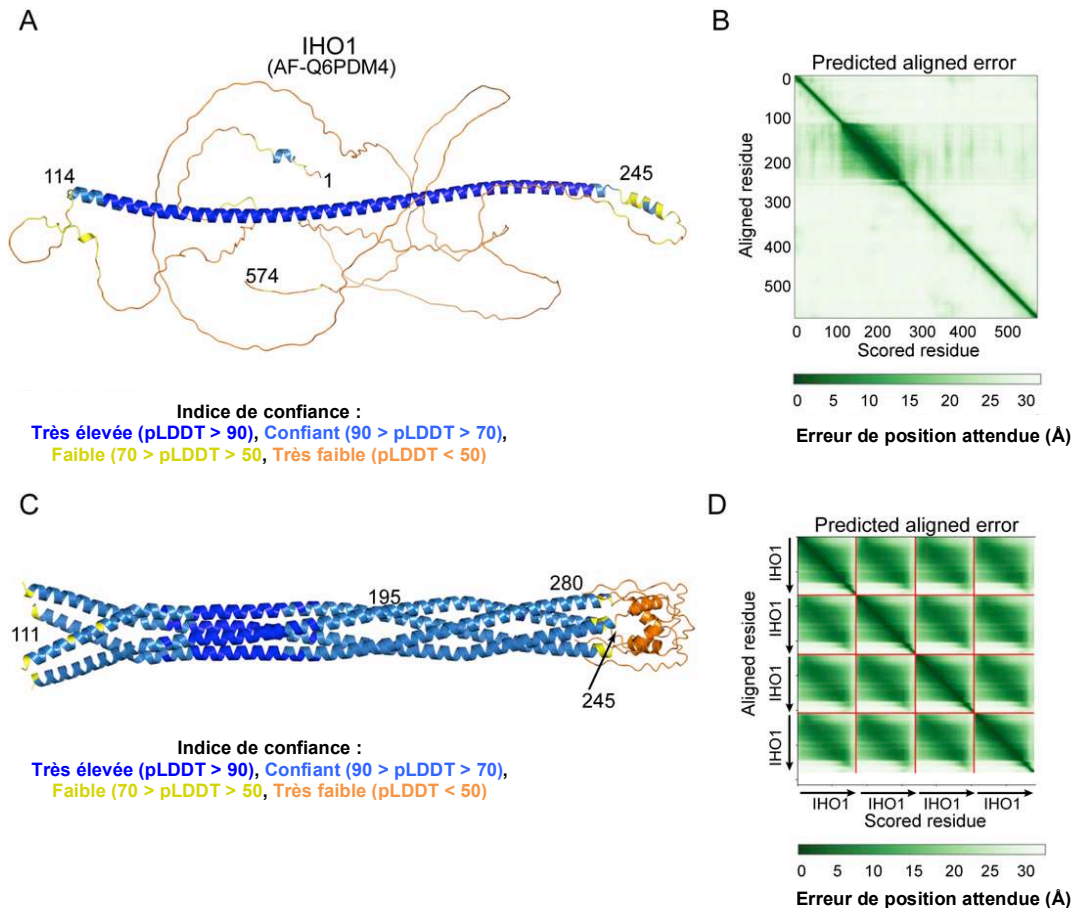


**Figure 49. Essais de cristallisation d'IHO<sub>143-257</sub>.** (A) Les cristaux obtenus lors des essais de cristallisation en utilisant des kits commerciaux par la plateforme de cristallisation. (B, C) Les cristaux obtenus après une optimisation manuelle des conditions. La condition de cristallisation est indiquée pour chaque goutte.



### 3.4.3 Modèle AlphaFold2 d'IHO1

Avec le développement d'AlphaFold2 (Jumper et al., 2021), le modèle structural d'IHO1<sub>FL</sub> est devenu disponible dans la base de données associée (Q6PDM4). Le modèle prédit, avec un très haut score de confiance, la présence d'une partie centrale structurée couvrant les résidus 114 à 245 (Figure 50 A, B). Ce modèle confirme également le défi que représente l'analyse structurale de cette protéine, car à part l'hélice centrale, la protéine est essentiellement désordonnée. La caractérisation de l'état oligomérique effectuée précédemment pour IHO1<sub>1-281</sub> et IHO1<sub>125-281</sub> a démontré la présence d'une forme tétramérique de cette protéine en solution. Une prédiction de cet assemblage a donc été effectuée par AlphaFold2 et le résultat obtenu suggère que cette partie d'IHO1 formerait un tétramère d'hélices alpha disposées dans une configuration parallèle (Figure 50 C, D).



**Figure 50. Modèle structural d'IHO1.** (A) Modèle AlphaFold2 (Jumper et al., 2021) de la protéine IHO1 coloré en fonction de la valeur de pLDDT qui correspond au score de confiance de la prédiction de chaque résidu (B) Graphique indiquant l'erreur de distance (en Ångströms) attendue dans la prédiction de la position du résidu X par rapport au résidu Y, une couleur foncée à la position (x, y) indique que l'erreur est faible tandis qu'une couleur claire traduit une erreur élevée. (D, E) Modélisation de la tétramérisation de la protéine IHO1<sub>1-281</sub> par AlphaFold2 (Jumper et al., 2021). (D) Les hélices sont colorées en fonction du degré de confiance de la prédiction pLDDT. (E) Le graphique mesurant la confiance de la prédiction de la distance pour chaque paire d'acides aminés est représenté.

### 3.4.4 Purification d'IHO1<sub>196-245</sub>

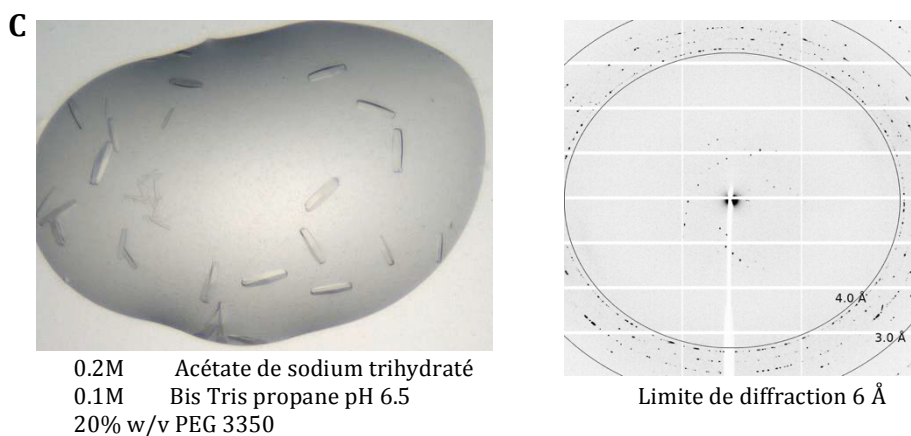
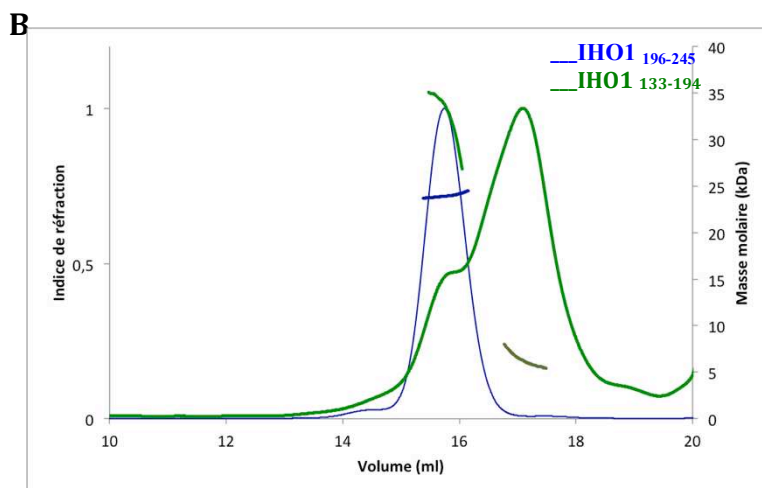
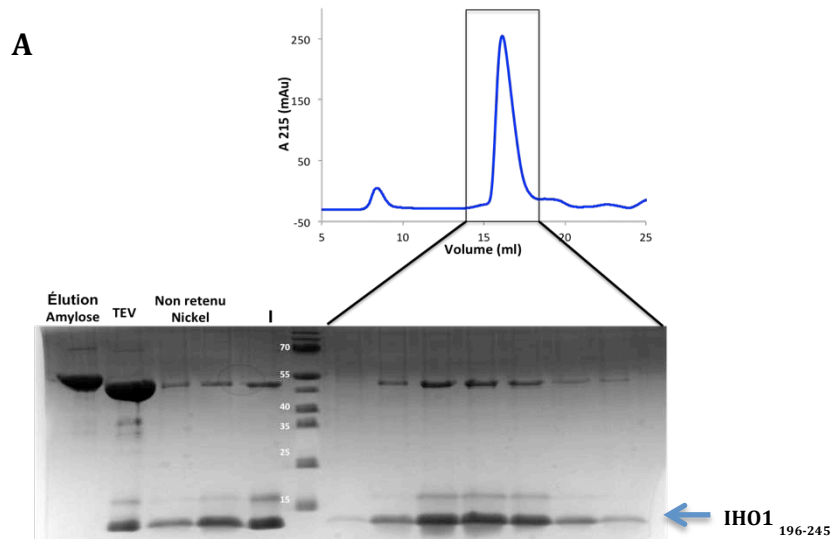
Dans la section précédente, j'ai décrit qu'IHO1<sub>143-259</sub> peut être cristallisée mais aucune diffraction n'a été obtenue. Une des raisons pour laquelle la protéine forme des cristaux qui ne diffractent pas pourrait provenir de la flexibilité du long tétramère hélicoïdale.

J'ai donc décidé de générer deux nouvelles constructions qui correspondent aux deux moitiés de la partie structurée prédite par AlphaFold2 : IHO1<sub>133-194</sub> et IHO1<sub>196-245</sub>. Comme les autres constructions de IHO1, ces deux dernières ont été exprimées en fusion avec une étiquette His-MBP et purifiées sur la résine amylose. Après le clivage des étiquettes de purification, la séparation par filtration sur gel des échantillons protéiques de la MBP n'a pas pu être utilisée, car le volume d'élution de ces deux fragments protéiques, qui sont plus petits que la construction centrale, est comparable à celui de la MBP. Il a fallu donc trouver un autre moyen pour séparer la protéine d'intérêt de la His-MBP. À cet égard, l'échantillon protéique clivé a été repassé sur une chromatographie d'affinité sur la résine nickel pour séparer la protéine de la His-MBP en ajoutant dans le tampon de la protéine 1mM EDTA (acide éthylènediaminetétraacétique). L'EDTA est un agent chélateur des ions métalliques divalents, il permettrait alors d'éliminer les cations divalents contaminants qui perturberaient la conformation de la protéine. Cependant, la concentration de l'EDTA doit être assez basse pour ne pas chélater le nickel présent sur la résine. Après cette étape de la purification, la protéine a été purifiée par gel-filtration. Le chromatogramme obtenu montre que la protéine est éluée sous forme d'un seul pic assez symétrique. Ce protocole de purification a permis d'obtenir un échantillon d'une pureté acceptable pour la suite des analyses (Figure 51 A).

### 3.4.5 Caractérisation d'IHO1<sub>196-245</sub>

Les échantillons purifiés ont été ensuite analysés par la technique de SEC-MALLS pour déterminer si ces deux constructions étaient encore présentes dans la solution sous forme de tétramères. Les résultats de cette expérience montrent qu'un pic majeur de 24 kDa a été observé dans l'échantillon IHO1<sub>196-245</sub>, ce qui correspond à un tétramère. Par contre, l'échantillon IHO1<sub>133-194</sub> est élué dans un pic majoritaire d'une masse de 6,4 kDa et cela correspond à un monomère (Figure 51 B). Ce résultat indique que les contacts requis pour la tétramérisation d'IHO1 se situent dans la région entre les résidus 196 à 245. La

construction tétramérique de IHO1 a été ensuite utilisée pour des essais de cristallisation. Plusieurs criblages ont été réalisés et des cristaux ont été obtenus dans la condition de 200 mM acétate de sodium trihydraté, 100 mM Bis-Tris propane pH 6.5 et 20% PEG 3350 (Figure 51 C). Les cristaux obtenus pour cette construction sont des baguettes et leur aspect est plus régulier, comparé à ceux obtenus pour IHO1<sub>143-257</sub>. Cependant, le pouvoir diffractif de ces cristaux était faible. Les cristaux testés appartiennent au groupe d'espace P4<sub>1</sub>2<sub>1</sub>2 ou P4<sub>3</sub>2<sub>1</sub>2 et avec des dimensions de la maille de a=70, b=70, c=142 Å. La meilleure diffraction obtenue était d'environ 6 Å et n'a pas permis pour le moment d'élucider la structure d'IHO1. Une meilleure diffraction pourrait être obtenue en améliorant la qualité cristalline. L'obtention de toute information structurale sur cette région centrale d'IHO1 serait utile pour confirmer le modèle AlphaFold2 et établir la topologie correcte de ce tétramère.



**Figure 51. Caractérisation d'IHO1<sub>196-259</sub> et essais de cristallisation** (A) Analyse par SDS-PAGE des différentes étapes de purification de IHO1<sub>196-254</sub>. La protéine a été purifiée d'abord sur la résine amylose, coupé à la TEV et passé ensuite sur la résine pour se débarrasser de His-MBP. La fraction non retenue sur la résine Ni<sup>2+</sup> a été concentrée et injectée (I) sur la filtration sur gel. (B) Caractérisation de l'état oligomérique de la construction IHO1<sub>133-194</sub> et IHO1<sub>196-254</sub> par SEC-MALLS. (C) Cristaux de la construction IHO1<sub>196-254</sub> et le cliché de diffraction obtenu.

## **4 Discussions et perspectives**

La formation de cassures d'ADN par le complexe catalytique SPO11/TOPOVIBL et leur réparation par recombinaison méiotique sont des mécanismes importants non seulement pour la ségrégation précise des chromosomes homologues mais aussi pour la diversité génomique. Néanmoins, les cassures double brin peuvent être dangereuses et peuvent entraîner des réarrangements chromosomiques indésirables si elles ne sont pas introduites au bon moment et réparées correctement. Par conséquent, l'activité du complexe SPO11-TOPOVIBL est étroitement contrôlée et dépend de plusieurs facteurs protéiques tels que les protéines MEI4, REC114 et IHO1, qui s'assemblent sur les chromosomes autour des sites des DSB. Malgré l'importance de ce mécanisme au cours de la méiose, peu d'informations sont disponibles sur les rôles exacts de ces facteurs protéiques, leur organisation et leur réseau d'interaction. Bien que ce mécanisme, qui initie la recombinaison méiotique, fasse l'objet d'étude chez plusieurs organismes modèles, la plupart de nos connaissances actuelles provient d'études réalisées chez la levure. C'est dans le but d'apporter de nouvelles connaissances structurales et fonctionnelles sur la formation et la régulation des DSB méiotique chez les mammifères que la collaboration entre les équipes de Dr. Jan Kadlec (IBS Grenoble) et Dr. Bernard De Massy (IGH Montpellier) a été mise en place. Les travaux menés par l'équipe ont permis de caractériser l'interaction de la protéine méiotique REC114 avec la protéine TOPOVIBL, membre du complexe catalytique, d'une part et avec la protéine MEI4 d'autre part. Ils ont également caractérisé avec succès la structure du domaine PH de REC114 seul ou en interaction avec TOPOVIBL (Kumar et al., 2018; Nore et al., 2022).

Pour ma part, je me suis intéressée à la protéine IHO1 et au complexe REC114-MEI4. Ce sont trois acteurs clés pour la formation des cassures d'ADN méiotique. Il a été suggéré que ces protéines agissent ensemble et régulent l'activité ou le recrutement du complexe catalytique. Cependant, l'interaction entre REC114-MEI4 et IHO1 n'a été analysée que par des tests d'interaction en double hybride (Stanzione et al., 2016). L'existence d'interaction directe et le mécanisme moléculaire la régissant n'ont pas été établis. Mon travail de thèse s'est porté sur la caractérisation biophysique du sous-complexe REC114-MEI4 et de la protéine IHO1, et sur l'analyse des interactions entre ces facteurs méiotiques. Le but de ce travail est de fournir de nouvelles informations sur l'organisation de ces protéines pour faire progresser notre compréhension sur le mode de la formation et de la régulation des DSB méiotiques. La comparaison des données obtenues pendant cette thèse sur les protéines murines avec les données biochimiques récemment

publiées sur les protéines de *S. cerevisiae* permettrait d'analyser les conservations et les différences entre ces acteurs de DSB chez les deux espèces.

## 4.1 Caractérisation du complexe REC114-MEI4-IHO1

Dans un premier temps, puisque les protéines MEI4 et IHO1 entières ne sont pas exprimées dans les bactéries, nous avons testé l'expression d'IHO1, REC114 et MEI4 entières dans les cellules d'insectes. La production dans les cellules eucaryotes est plus adaptée à l'expression de protéines des mammifères et favorise le repliement correct des protéines. Néanmoins, le niveau d'expression de ces protéines n'était pas satisfaisant.

En parallèle, j'ai entrepris la caractérisation de l'interaction entre le complexe REC114-MEI4 et la protéine IHO1 en utilisant des constructions tronquées. Comme observé par la filtration sur gel, le mélange de la moitié N-terminale d'IHO1 (IHO1<sub>1-281</sub>) avec le sous-complexe REC114.MEI4<sub>1-127</sub> a entraîné la formation d'un complexe contenant les trois protéines. C'est donc la première indication d'une interaction directe entre REC114-MEI4 et IHO1. Ensuite, les mesures d'ITC entre IHO1 et REC114-MEI4 ont révélé une affinité dans la gamme de 20 à 30  $\mu$ M, confirmant ainsi l'existence d'un complexe ternaire IHO1-REC114-MEI4. En outre, les différentes mesures d'ITC effectuées ont montré que l'absence de MEI4 n'affectait pas l'interaction entre REC114 et IHO1. Elles nous ont également permis d'identifier que la région N-terminale d'IHO1 (1-23) et la partie N terminale de REC114 (domaine PH) sont les acteurs impliqués dans l'interaction entre ces deux protéines.

Ces résultats sont importants, car l'existence de ce complexe a récemment été remise en question chez la levure. Il a été démontré que REC114, MEI4 et IHO1 se co-localisent sur les axes chromosomiques et co-immunoprécipitent (Kumar et al., 2018; Stanzione et al., 2016). Il a également été démontré, chez la souris et la levure par la technique du double hybride, que la protéine REC114 se lie à IHO1 (Arora et al., 2004; Maleki et al., 2007; Stanzione et al., 2016). Cependant, une étude récente chez *S. cerevisiae* suggère que Rec114-Mei4 et Mer2 (l'orthologue d'IHO1) sont des entités biochimiquement distinctes qui se lient de manière coopérative à l'ADN en formant des condensats de manière ADN-dépendants. La même étude a aussi montré par *pull down* que Mer2 se lie faiblement au complexe Rec114-Mei4 (Claeys Bouuaert et al., 2021a). Nos

résultats montrent maintenant que les protéines REC114 et IHO1 de souris interagissent directement *in vitro* et identifient les régions protéiques impliquées dans cette interaction.

De plus, l'interaction entre REC114 et IHO1 a été modélisée par AlphaFold2. Le peptide d'IHO1 se positionnerait contre le feuillet  $\beta$  du domaine PH de REC114. Ce modèle a été prédit avec une faible confiance et n'est probablement pas complètement correct. Néanmoins, la mutagenèse des résidus susceptibles d'être impliqués dans cette interaction a permis d'identifier que les résidus F3 d'IHO1 et L104 du domaine PH sont indispensables pour l'interaction entre les deux protéines. Ces résultats apportent des informations moléculaires sur l'interaction entre REC114 et IHO1. La mutation du résidu F3 sera utilisée pour analyser, *in vivo*, l'effet de la perte d'interaction entre REC114 et IHO1. Il faut noter qu'il n'existe actuellement pas de modèles cellulaires pour la formation des DSB méiotiques chez la souris et que l'analyse fonctionnelle des mutations identifiées dans cette étude nécessite la préparation de souris génétiquement modifiées.

Le modèle AlphaFold2 a montré qu'IHO1 interagit avec REC114 sur la même surface du domaine PH de REC114 que les deux autres partenaires de REC114 (TOPOVIBL et ANKRD31) (Boekhout et al., 2019; Nore et al., 2022). Les expériences réalisées pendant cette thèse par la chromatographie d'exclusion stérique ont montré également que l'interaction d'IHO1 sur le complexe REC114-MEI4 est inhibée par la présence du domaine C-terminal d'ANKRD31 ou de TOPOVIBL. Cette observation est particulièrement intéressante pour comprendre l'activité et la dynamique de ces facteurs protéiques lors de l'initiation de la recombinaison méiotique, mais son implication exacte n'est pas claire. Il est intéressant de noter que la formation des foci des protéines REC114, MEI4 sur l'axe chromosomique dépend de la protéine IHO1 et de son interaction avec HORMAD1 et est favorisée par ANKRD31, mais n'en dépend pas. En revanche, la localisation de ces protéines sur les PAR et les extrémités distales de certains chromosomes est indépendante d'IHO1 mais nécessite ANKRD31 (Papanikos et al., 2019; Stanzione et al., 2016). D'autres analyses, y compris des travaux *in vivo*, seront nécessaires pour comprendre une éventuelle mutuelle exclusivité dans la liaison de ces protéines à REC114.

L'affinité de liaison d'IHO1 à REC114 mesurée durant ce travail est relativement faible comparée à celles des complexes TOPOVIBL-REC114 et ANKRD31-REC114. Une des explications possibles est que l'interaction entre REC114 et IHO1 nécessiterait la



présence de modifications post-traductionnelles de la protéine IHO1 qui, comme son homologue de levure *S. cerevisiae*, peut être régulée par la phosphorylation. En effet, des études antérieures chez la levure ont montré que la phosphorylation de la région N-terminale de Mre2 (l'orthologue d'IHO1) promeut le recrutement de REC114 aux sites des DSB (Sasanuma et al., 2007) et elle semble nécessaire pour l'interaction des deux protéines analysée par double hybride (Henderson et al., 2006). De plus, comme nous avons pu montrer que la protéine IHO1 forme des tétramères et que le complexe REC114-MEI4 est composé d'oligomères 2:1 et 4:2, l'affinité entre ces partenaires pourrait donc être plus élevée au sein de ces assemblages multimériques. D'autre part, nos analyses ont été réalisées avec des constructions tronquées, les régions manquantes de ces protéines pourraient fournir des contacts supplémentaires, rendant le complexe REC114-MEI4-IHO1 plus stable.

Le complexe Rec114-Mei4 et la protéine Mer2 chez *S. cerevisiae* coopèrent pour lier l'ADN, formant un grand assemblage nucléoprotéique composé d'ADN, du complexe Rec114-Mei4 et de la protéine Mer2. Au sein de ces *clusters* de nucléoprotéiques, des interactions entre les complexes Mer2 et Rec114-Mei4 existeraient et participeraient au maintien et à l'assemblage de ces condensats (Claeys Bouuaert et al., 2021a). Un mécanisme similaire à celui proposé pour la levure pourrait exister chez les mammifères, d'autant plus que la protéine IHO1 s'est avérée capable d'induire de la séparation de phases *in vitro* (Tsai et al., 2020). Pour valider la formation de ces assemblages nucléoprotéiques chez la souris, des tests d'interaction de REC114-MEI4 et de la protéine IHO1 avec de l'ADN devront d'abord être effectués. Il est à noter qu'aucune contamination par l'ADN n'a été observée lors de la purification du complexe REC114-MEI4 ou de la construction IHO1<sub>1-281</sub> dans nos conditions expérimentales, bien que ce phénomène soit courant lors de la purification des protéines de liaison à l'ADN. Cela indique que nos constructions n'interagissent probablement pas avec les acides nucléiques. Par conséquent, ces interactions protéines-ADN, si elles existent, nécessiteraient des conditions expérimentales différentes et affecteraient probablement la partie C-terminale d'IHO1.

Pour continuer la caractérisation du complexe ternaire, il serait intéressant d'optimiser l'expression des protéines exprimées en cellules d'insecte afin de réaliser des caractérisations biochimiques et structurales sur des protéines entières. Il va falloir essayer différentes étiquettes d'expression qui peuvent améliorer le rendement et la solubilité de

ces protéines ou de tester d'autres systèmes d'expression comme les cellules de mammifères.

L'état oligomérique des protéines individuelles et du complexe qu'elles forment a été caractérisé pendant ma thèse par la technique du SEC-MALLS. Bien que les données de SEC-MALLS n'aient pas permis de déterminer la stœchiométrie exacte du complexe REC114-MEI-MBP-IHO1, elles ont montré qu'il s'agit d'un échantillon à haut poids moléculaire (400 kDa en présence de la MBP) dont la taille est adaptée à la caractérisation structurale par la Cryo-EM. Une analyse par coloration négative a été donc effectuée et une étape de pontage chimique a été optimisée et ajoutée au protocole de préparation du complexe pour prévenir sa dissociation. À l'heure actuelle, la qualité des échantillons n'est toujours pas satisfaisante pour procéder à des études structurales à haute résolution par la cryo-microscopie électronique. Ainsi, nous devrions continuer à optimiser ce complexe, peut-être en supprimant certaines régions flexibles et désordonnées de ces protéines et en testant différents agents de pontage pour prévenir sa dissociation. La technique du Grafix pourrait également être utilisée (Stark, 2010). L'échantillon serait chargé au-dessus d'un gradient de saccharose ou de glycérol contenant une concentration croissante d'agent de pontage et soumis à une ultracentrifugation. Cette technique permettrait de séparer les différentes populations du complexe, augmentant ainsi son homogénéité.

Dans le même contexte, la technique de diffusion des rayons X aux petits angles (SAXS) peut s'avérer être une bonne méthode pour évaluer l'état oligomérique du complexe formé entre REC114-MEI4 et IHO1. De plus, cette méthode génère une grande quantité d'informations structurales à partir du profil de diffusion résultant, y compris le rayon de giration ( $R_g$ ), le poids moléculaire, la flexibilité, la forme tridimensionnelle globale, la polydispersité et la distance intraparticule maximale ( $D_{max}$ ).

Par ailleurs, l'analyse structurale du complexe minimal IHO1-REC114 a été initiée par des essais de cristallisation sur les 16 premiers résidus d'IHO1 exprimés en fusion avec le domaine PH de REC114. Ces tests ont donné place à des cristaux, mais qui ne contenaient que le domaine PH de REC114. Il faut alors optimiser le protocole d'expression et de purification de cette construction afin de prévenir la dégradation et la perte du peptide d'IHO1. Le peptide codant cette partie N-terminale d'IHO1 a été également synthétisé, mais des problèmes dans la solubilité ont été rencontrés. Il est donc

nécessaire de surmonter ce problème avant de pouvoir lancer des tests de co-cristallisation avec le domaine PH.

En parallèle, pour valider davantage le modèle structural du complexe REC114-IHO1, il sera utile de tester l'effet de la mutation de résidus conservés supplémentaires à l'extrémité N-terminale d'IHO1 sur son interaction avec REC114.

En outre, le complexe REC114-IHO-MEI4 devrait être caractérisé dans un contexte plus large en incluant la protéine HORMAD1. En effet, il a été suggéré que la protéine HORMAD1 permet le recrutement d'IHO1 sur les axes chromosomiques. De plus, il a été démontré que la protéine Mer2 est capable, *in vitro*, de lier la protéine Hop1, orthologue de HORMAD1 chez la levure (Rousová et al., 2021). Chez la souris, l'interaction entre HORMAD1 et IHO1 a été observée par des tests d'interaction en double hybride (Stanzione et al., 2016). Cette interaction devrait être analysée sur le plan moléculaire pour caractériser davantage l'organisation de la machinerie médiatrice des DSB méiotiques.

## 4.2 Caractérisation du sous-complexe REC114-MEI4

En plus du complexe ternaire qui réunit IHO1-REC114-MEI4, nous avons étudié le sous-complexe REC114-MEI4 pour caractériser son organisation et les détails moléculaires de son interaction. L'analyse des différentes troncations a permis de montrer que les résidus 6 à 43 de MEI4 sont suffisants pour l'interaction avec la partie C-terminale de REC114. L'analyse du complexe minimal REC114<sub>226-254</sub> en interaction avec MEI4<sub>1-43</sub> a été tentée par RMN. Les spectres obtenus ont suggéré l'oligomérisation du complexe ou la présence d'un phénomène d'échange entre les différentes formes oligomériques, ce qui limite l'acquisition des données sur l'ensemble des résidus. La caractérisation de l'état oligomérique du complexe a été menée par différentes techniques (la SEC-MALLS, MS native et AUC) et elle a permis de montrer que la stœchiométrie du complexe REC114-MEI4 n'est pas stable, ce qui pourrait expliquer par ailleurs l'échec des études structurales par RMN ou par cristallographie. L'espèce prédominante en solution de ce complexe semble être le complexe 2REC114-1MEI4, bien qu'il puisse former des oligomères plus gros en fonction de la concentration. Cette observation est intéressante pour comprendre l'activité de ce complexe. En effet, ce phénomène d'oligomérisation pourrait être un moyen pour la formation et la croissance de gros assemblages, tels que les condensats de

séparation de phase induits par l'ADN observés, *in vitro*, chez la levure ou les foci observés *in vivo*. De même, la caractérisation, *in vitro*, du complexe REC114-MEI4 de levure a révélé qu'il est également constitué de deux sous-unités Rec114 et d'une sous-unité Mei4 (qui interagissent par des régions équivalentes au complexe de souris) et il s'assemble en condensats nucléoprotéiques (Claeys Bouuaert et al., 2021a). Pour caractériser davantage le complexe REC114-MEI4 murin, il serait judicieux d'analyser son état oligomérique en présence de la protéine MEI4 entière, car la partie manquante de MEI4 pourrait stabiliser la stœchiométrie du complexe. Ensuite, la formation de condensats nucléoprotéiques devrait être testé pour confirmer la conservation évolutive de ce mécanisme chez la souris.

L'arrivée d'AlphaFold2, l'année dernière, a ouvert de nouvelles voies pour l'avancée de mon projet. Le modèle tridimensionnel obtenu pour le complexe REC114-MEI4 a été exploré et validé, *in vitro*, pour obtenir des informations précises sur le mode d'interaction entre les deux protéines. Les résultats d'analyse d'interaction entre REC114 et MEI4 ont montré que les simples mutations des résidus F240E, V244E et V247E sur REC114 ainsi que les mutations L15E et A16E sur MEI4 empêchent la formation du complexe binaire REC114-MEI4. L'organisation du complexe REC114-MEI4 semble être conservée au cours de l'évolution. Le système AlphaFold2 prédit des structures similaires pour le complexe REC114-MEI4 de la levure *S. cerevisiae* et celui de la souris. De plus, la mutation de résidus équivalents entre la souris, la levure *S. cerevisiae* et la levure *S. pombe* perturbe la fixation de la protéine MEI4 (Claeys Bouuaert et al., 2021a; Steiner et al., 2010).

Dans un second temps, les expériences réalisées durant cette thèse par la technique de SEC-MALLS et de MS native ont montré également que la partie C-terminale de REC114 adopte une forme dimérique en solution. Ce qui suggère que la dimérisation de REC114 est indépendante du recrutement de MEI4. Le modèle AlphaFold2 de ce dimère a été validé expérimentalement par une série de mutations et il nous a permis d'identifier les résidus impliqués dans cette dimérisation (I227, I244). Bien que les modèles structuraux du complexe Rec114-Mei4 paraissent conservés entre la levure et la souris, la structure du dimère REC114 semble être différente. La dimérisation de REC114 de levure *S. cerevisiae* n'implique pas la même surface d'interaction que celle de son homologue murin, de sorte que la mutation de résidu équivalent entre les deux espèces n'a pas le même impact sur la dimérisation. La fixation de MEI4 pourrait induire des changements de conformation dans

le C-terminus de REC114, ce qui lui permettrait d'adopter une structure similaire à celle du complexe REC114-MEI4 murin.

Les mutations identifiées, *in vitro*, devraient être testées, *in vivo*, pour comprendre l'effet spécifique de la perturbation de l'interaction entre REC114 et MEI4 ou de la dimérisation de REC114 sur la formation des cassures double brin méiotique ainsi que la progression de la méiose. Un projet de test, *in vivo*, de l'effet des mutations identifiées durant cette thèse (notamment le A16 de MEI4) est en cours, en collaboration avec l'équipe du Dr. Bernard De Massy.

Le problème majeur que j'ai rencontré durant la caractérisation du complexe REC114-MEI4 était l'échange rapide observé entre les différents états oligomériques, ce qui a compromis tous mes efforts pour obtenir expérimentalement la structure de ce complexe. Pour atteindre ce but, les modèles AlphaFold2 du complexe REC114-MEI4 et de la partie C-terminale de REC114 peuvent être utilisés pour la détermination des résidus potentiellement impliqués dans l'oligomérisation du complexe. La mutation de ces résidus peut interférer avec l'échange entre les différents états oligomériques et améliorerait la qualité des données RMN pour pouvoir résoudre la structure de ce complexe par cette technique, ou éventuellement par cristallisation. Alternativement, les conditions de concentration et de tampons qui stabilisent le complexe REC114<sub>FL</sub>-MEI4 dans un état oligomérique à haut poids moléculaire peuvent être recherchées pour accéder à la structure de ce multimère par Cryo-EM. Dans ce but, la technique de photométrie de masse pourrait être une excellente méthode pour la recherche des conditions optimales.

### 4.3 Caractérisation d'IHO1

D'autre part, la protéine IHO1<sub>1-281</sub> a été caractérisée et, comme l'orthologue Mer2, s'est avérée tétramérique en solution (Claeys Bouuaert et al., 2021a; Rousová et al., 2021). Bien que cette protéine n'ait pas pu être cristallisée, les images de la coloration négative de l'échantillon IHO1<sub>125-281</sub> étaient très encourageantes, indiquant la présence de particules de structure en fibres. Des données en cryo-microscopie électronique ont été collectées, mais la taille limitée et l'hétérogénéité des particules sur les grilles de Cryo-EM étaient un obstacle lors du traitement des données. En raison de la taille relativement faible du complexe sur les micrographes, le processus de sélection de particules a confondu des

artefacts et des impuretés avec les particules de l'échantillon étudié, ce qui a troublé leurs alignements et a empêché l'obtention de classes 2D satisfaisantes.

Toujours dans le but de résoudre la structure de la partie centrale d'IHO1, de nouvelles constructions d'IHO1 ont été analysées et des cristaux ont été ensuite obtenus, mais malheureusement leurs pouvoirs diffractifs étaient très faibles. Le tétramère de la partie centrale d'IHO1 a été modélisé à l'aide d'AlphaFold2. Il montre la présence d'un *coiled-coil* tétramérique adoptant une forme parallèle. Cette organisation est différente avec celle proposée pour l'orthologue de la levure (Mer2). En effet, la combinaison des données de pontage chimique couplé à la spectrométrie de masse aux données de SEC-MALLS effectuée sur la fusion de deux monomères de *coiled-coil* de Mer2 par un *linker* trop court pour une disposition parallèle, a suggéré la présence d'un assemblage de deux paires d'hélices parallèles disposées de manière antiparallèle (Claeys Bouuaert et al., 2021a). L'obtention de la structure de la partie centrale d'IHO1 serait d'une très grande aide pour comprendre l'organisation de cet homotétramère chez la souris.

Dans le but d'augmenter la diffraction des cristaux de la partie centrale de la protéine IHO1, le modèle AlphaFold2 a été utilisé pour faire deux autres constructions divisant la partie structurée d'IHO1 en deux parties. La caractérisation de l'état oligomérique de ces deux constructions a montré que la région responsable de la tétramérisation de cette protéine se trouverait dans la partie entre le résidu 196 et le résidu 259. Les cristaux obtenus avec la construction IHO1<sub>196-259</sub> sont de meilleures qualités, mais leur pouvoir diffractif, limité à 6 Å, ne permet pas de résoudre la structure de cette protéine. Il est donc nécessaire d'améliorer la diffraction de ces cristaux pour résoudre la structure d'IHO1<sub>196-259</sub> par cristallographie à résolution atomique et déterminer l'organisation tétramérique d'IHO1.

Pour avoir la structure de la partie centrale d'IHO1 par cristallographie, il faut essayer d'améliorer la diffraction des cristaux obtenus. Pour cela, des additifs peuvent être ajoutés à la solution de cristallisation et pourront améliorer la qualité des cristaux. L'ensemencement de ces cristaux dans les conditions des kits commerciaux pourrait également nous aider à trouver de nouvelles conditions qui amélioreraient la diffraction des cristaux. Les bornes de la séquence protéique peuvent être modifiées en réalisant de nouvelles constructions qui seront probablement plus homogènes et auront un meilleur

pouvoir diffractif. En parallèle, si ces efforts n'aboutissent pas, les résidus qui semblent importants pour la tétramérisation d'IHO1 dans le modèle Alphafold2 peuvent être mutés et analysés afin de confirmer la prédiction de ce tétramère.

Une autre voie consiste à consacrer plus d'effort à la caractérisation de la structure de cette protéine par Cryo-EM. Grâce à l'évolution rapide de la technique de Cryo-EM, plusieurs pistes pourraient être explorées pour résoudre à haute résolution la structure du tétramère du fragment IHO1<sub>125-281</sub> et déterminer les propriétés structurales qui régissent sa tétramérisation. Pour améliorer le contraste des images, la technologie de volta phase plate (VPP) pourrait être utilisée, car elle permet l'acquisition de données proches de la focale en introduisant un décalage de phase entre les électrons diffusés par l'échantillon, ce qui va permettre d'augmenter significativement le contraste des images tout en maintenant les informations à haute résolution. Cette technique est adaptée aux protéines de petits poids moléculaires et elle a permis récemment de résoudre la structure de la streptavidine (52kDa) et de l'hémoglobine humaine (64 kDa) à une résolution de 3,2 Å (Fan et al., 2019; Khoshouei et al., 2017). La structure de cette protéine pourrait être également caractérisée par Cryo-EM en la fusionnant avec une autre protéine structurée et homogène afin d'augmenter la taille des particules. Néanmoins, cette technique nécessite beaucoup d'optimisation et la fusion doit être assez rigide pour ne pas introduire de la flexibilité dans l'échantillon et troubler les alignements des particules, ce qui provoquerait la perte du signal de la protéine cible lors de la reconstruction et du traitement d'images. Une nouvelle approche a été également développée et elle est appelée Megabodies. Elle consiste en une molécule chimérique formée par la fusion d'un anticorps spécifique à la protéine cible (nanobodies notamment) avec des protéines plus grandes pour produire des chimères stables, raisonnablement rigides et se repliant efficacement. Cette stratégie a permis récemment des reconstructions 3D par Cryo-EM pour une protéine membranaire d'environ 57 kDa (Uchański et al., 2021).

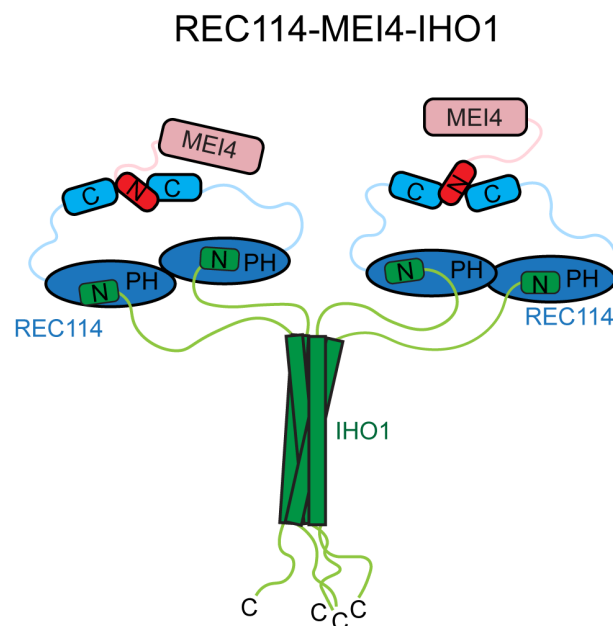




## **5 Conclusion**

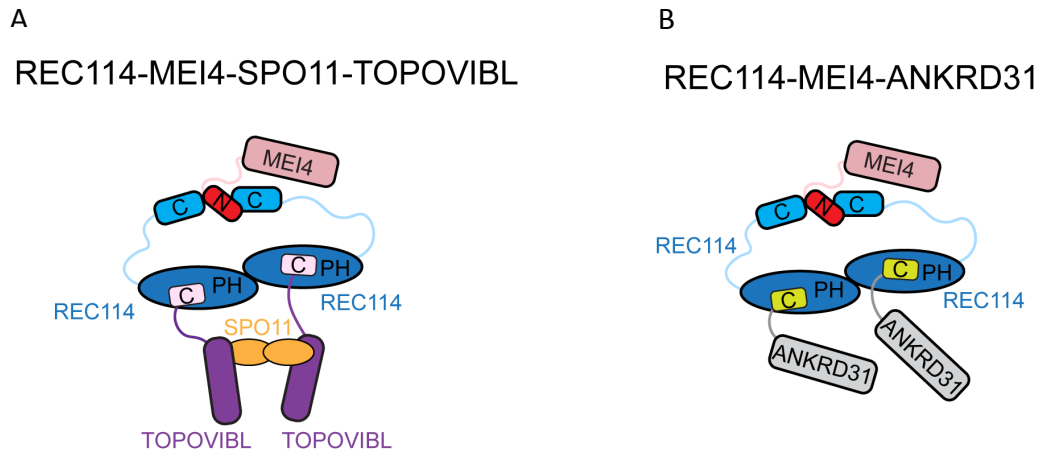


Bien que la structure expérimentale du complexe IHO1-REC114-MEI4 ou des protéines seules n'ait pas pu être obtenue durant cette thèse, ce travail apporte des informations précieuses sur l'organisation de ces protéines méiotiques et peut servir de base pour des caractérisations plus poussées de la machinerie de DSB. Nos résultats démontrent l'existence d'une interaction directe entre REC114 et IHO1. Cette interaction a été proposée *in vivo* pour assurer le recrutement ou la stabilité de REC114-MEI4 sur l'axe chromosomique lié par la protéine IHO1. De plus, les différentes analyses effectuées ont élucidé les domaines nécessaires à cette interaction. L'ensemble de nos résultats permettent d'éclaircir l'état oligomérique des protéines REC114-MEI4 et IHO1 et de suggérer que le complexe ternaire serait composé d'un tétramère d'IHO1 en interaction avec un ou deux hétérotrimères du complexe REC114-MEI4 (2REC114-MEI4 ou 4REC114-2MEI4) (Figure 52). Cette association est probablement régulée, *in vivo*, par la phosphorylation et/ou est contrôlée par les systèmes des séparations de phase.



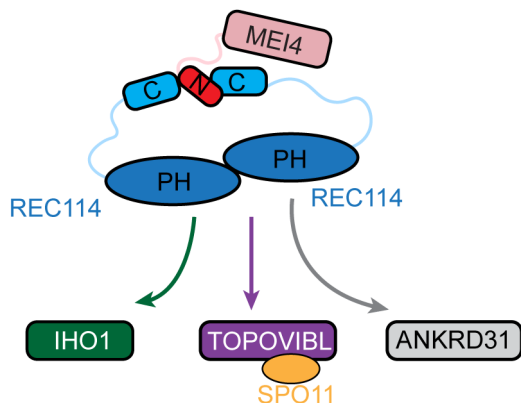
**Figure 52. Modèle du complexe REC114-MEI4-IHO1.** Schématisation du complexe IHO1-REC114-MEI4 montrant le tétramère d'IHO1 (vert) qui interagit, via son N-terminus, avec le domaine PH de REC114 (bleu), la partie C-terminale des deux REC114 (cyan) est en interaction avec le N-terminus de MEI4 (rouge).

Combiné avec les données de Nore *et al* (2022) montrant l'interaction entre le domaine PH de REC114 et TOPOVIBL, la forme dimérique de REC114 déterminée au cours de ce travail pourrait stabiliser la dimérisation du complexe catalytique SPO11/TOPOVIBL (Figure 53 A), dont les éléments nécessaires à sa dimérisation dans TOPOVIBL semblent dégénérés. De la même façon l'interaction d'ANKRD31 avec le complexe REC114-MEI4 pourrait induire sa dimérisation (Figure 53 B).



**Figure 53. Modèles des complexes REC114-MEI4-TOPOVIBL et REC114-MEI4-ANKRD31.** (A) Schématisation de la dimérisation du complexe catalytique SPO11-TOPOVIBL par son interaction avec l'hétérotrimère 2REC114-1MEI4. (B) Schématisation de l'interaction entre le complexe 2REC114-MEI4 et ANKRD31.

L'interaction d'IHO1 sur la même région de REC114 que celle liée par TOPOVIBL et ANKRD31 suggère des interactions mutuellement exclusives, et soulève des questions sur les relations entre ces différents partenaires et la coordination de leur liaison avec REC114 (Figure 54).



**Figure 54. Partenaires de REC114.** Modèle résumant les partenaires de REC114. Le C-terminus est impliqué dans une interaction avec MEI4 et le domaine PH dans interaction probablement mutuellement exclusive avec IHO1, TOPOVIBL et ANKRD31.

Enfin, et grâce à AlphaFold2, les résidus essentiels à la formation des complexes REC114-MEI4 et IHO1-REC114, ainsi qu'à la dimérisation de REC114 identifiés durant ce travail vont servir pour de nouvelles études de mutagenèses, *in vivo*. Ceci permettrait de mieux comprendre ces réseaux d'interaction et leur importance *in vivo*, et de fournir de nouvelles informations sur la régulation et la formation des DSB méiotiques.



## Références bibliographiques

- Acquaviva, L., Boekhout, M., Karasu, M.E., Brick, K., Pratto, F., Li, T., van Overbeek, M., Kauppi, L., Camerini-Otero, R.D., Jasin, M., Keeney, S., 2020. Ensuring meiotic DNA break formation in the mouse pseudoautosomal region. *Nature* 582, 426–431. <https://doi.org/10.1038/s41586-020-2327-4>
- Acquaviva, L., Székvölgyi, L., Dichtl, B., Dichtl, B.S., de La Roche Saint André, C., Nicolas, A., Géli, V., 2013. The COMPASS subunit Spp1 links histone methylation to initiation of meiotic recombination. *Science* 339, 215–218. <https://doi.org/10.1126/science.1225739>
- Agarwal, S., Roeder, G.S., 2000. Zip3 Provides a Link between Recombination Enzymes and Synaptonemal Complex Proteins. *Cell* 102, 245–255. [https://doi.org/10.1016/S0092-8674\(00\)00029-5](https://doi.org/10.1016/S0092-8674(00)00029-5)
- Allers, T., Lichten, M., 2001. Differential timing and control of noncrossover and crossover recombination during meiosis. *Cell* 106, 47–57. [https://doi.org/10.1016/s0092-8674\(01\)00416-0](https://doi.org/10.1016/s0092-8674(01)00416-0)
- Andersen, S.L., Sekelsky, J., 2010. Meiotic versus Mitotic Recombination: Two Different Routes for Double-Strand Break Repair. *BioEssays News Rev. Mol. Cell. Dev. Biol.* 32, 1058–1066. <https://doi.org/10.1002/bies.201000087>
- Anderson, J.S., Parker, R.P., 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. *EMBO J.* 17, 1497–1506. <https://doi.org/10.1093/emboj/17.5.1497>
- Aravind, L., Koonin, E.V., 1998. The HORMA domain: a common structural denominator in mitotic checkpoints, chromosome synapsis and DNA repair. *Trends Biochem. Sci.* 23, 284–286. [https://doi.org/10.1016/s0968-0004\(98\)01257-2](https://doi.org/10.1016/s0968-0004(98)01257-2)
- Arora, C., Kee, K., Maleki, S., Keeney, S., 2004. Antiviral protein Ski8 is a direct partner of Spo11 in meiotic DNA break formation, independent of its cytoplasmic role in RNA metabolism. *Mol. Cell* 13, 549–559. [https://doi.org/10.1016/s1097-2765\(04\)00063-2](https://doi.org/10.1016/s1097-2765(04)00063-2)
- Arora, K., Corbett, K.D., 2019. The conserved XPF:ERCC1-like Zip2:Spo16 complex controls meiotic crossover formation through structure-specific DNA binding. *Nucleic Acids Res.* 47, 2365–2376. <https://doi.org/10.1093/nar/gky1273>
- Baker, C.L., Walker, M., Kajita, S., Petkov, P.M., Paigen, K., 2014. PRDM9 binding organizes hotspot nucleosomes and limits Holliday junction migration. *Genome Res.* 24, 724–732. <https://doi.org/10.1101/gr.170167.113>
- Barlow, C., Liyanage, M., Moens, P.B., Tarsounas, M., Nagashima, K., Brown, K., Rottinghaus, S., Jackson, S.P., Tagle, D., Ried, T., Wynshaw-Boris, A., 1998. Atm deficiency results in severe meiotic disruption as early as leptotema of prophase I. *Dev. Camb. Engl.* 125, 4007–4017. <https://doi.org/10.1242/dev.125.20.4007>
- Baudat, F., Buard, J., Grey, C., Fledel-Alon, A., Ober, C., Przeworski, M., Coop, G., de Massy, B., 2010. PRDM9 is a Major Determinant of Meiotic Recombination Hotspots in humans and mice. *Science* 327, 836–840. <https://doi.org/10.1126/science.1183439>
- Baudat, F., de Massy, B., 2007. Regulating double-stranded DNA break repair towards crossover or non-crossover during mammalian meiosis. *Chromosome Res. Int. J. Mol. Supramol. Evol. Asp. Chromosome Biol.* 15, 565–577. <https://doi.org/10.1007/s10577-007-1140-3>

- Baudat, F., Imai, Y., de Massy, B., 2013. Meiotic recombination in mamMALLS: localization and regulation. *Nat. Rev. Genet.* 14, 794–806. <https://doi.org/10.1038/nrg3573>
- Berchowitz, L.E., Copenhaver, G.P., 2010. Genetic interference: don't stand so close to me. *Curr. Genomics* 11, 91–102. <https://doi.org/10.2174/138920210790886835>
- Bergerat, A., de Massy, B., Gadelle, D., Varoutas, P.C., Nicolas, A., Forterre, P., 1997. An atypical topoisomerase II from Archaea with implications for meiotic recombination. *Nature* 386, 414–417. <https://doi.org/10.1038/386414a0>
- Bishop, D.K., Park, D., Xu, L., Kleckner, N., 1992. DMC1: a meiosis-specific yeast homolog of *E. coli* recA required for recombination, synaptonemal complex formation, and cell cycle progression. *Cell* 69, 439–456. [https://doi.org/10.1016/0092-8674\(92\)90446-j](https://doi.org/10.1016/0092-8674(92)90446-j)
- Bishop, D.K., Zickler, D., 2004. Early decision; meiotic crossover interference prior to stable strand exchange and synapsis. *Cell* 117, 9–15. [https://doi.org/10.1016/s0092-8674\(04\)00297-1](https://doi.org/10.1016/s0092-8674(04)00297-1)
- Bizard, A.H., Hickson, I.D., 2014. The Dissolution of Double Holliday Junctions. *Cold Spring Harb. Perspect. Biol.* 6, a016477. <https://doi.org/10.1101/cshperspect.a016477>
- Blat, Y., Protacio, R.U., Hunter, N., Kleckner, N., 2002. Physical and functional interactions among basic chromosome organizational features govern early steps of meiotic chiasma formation. *Cell* 111, 791–802. [https://doi.org/10.1016/s0092-8674\(02\)01167-4](https://doi.org/10.1016/s0092-8674(02)01167-4)
- Blitzblau, H.G., Hochwagen, A., 2013. ATR/Mec1 prevents lethal meiotic recombination initiation on partially replicated chromosomes in budding yeast. *eLife* 2. <https://doi.org/10.7554/eLife.00844>
- Boateng, K.A., Bellani, M.A., Gregoretti, I.V., Pratto, F., Camerini-Otero, R.D., 2013. Homologous pairing preceding SPO11-mediated double-strand breaks in mice. *Dev. Cell* 24, 196–205. <https://doi.org/10.1016/j.devcel.2012.12.002>
- Boekhout, M., Karasu, M.E., Wang, J., Acquaviva, L., Pratto, F., Brick, K., Eng, D.Y., Xu, J., Camerini-Otero, R.D., Patel, D.J., Keeney, S., 2019. REC114 Partner ANKRD31 Controls Number, Timing, and Location of Meiotic DNA Breaks. *Mol. Cell* 74, 1053–1068.e8. <https://doi.org/10.1016/j.molcel.2019.03.023>
- Boeri Erba, E., Signor, L., Oliva, M.F., Hans, F., Petosa, C., 2018. Characterizing Intact Macromolecular Complexes Using Native Mass Spectrometry. *Methods Mol. Biol.* Clifton NJ 1764, 133–151. [https://doi.org/10.1007/978-1-4939-7759-8\\_9](https://doi.org/10.1007/978-1-4939-7759-8_9)
- Borde, V., de Massy, B., 2013. Programmed induction of DNA double strand breaks during meiosis: setting up communication between DNA and the chromosome structure. *Curr. Opin. Genet. Dev., Genome architecture and expression* 23, 147–155. <https://doi.org/10.1016/j.gde.2012.12.002>
- Borde, V., Goldman, A.S., Lichten, M., 2000. Direct coupling between meiotic DNA replication and recombination initiation. *Science* 290, 806–809. <https://doi.org/10.1126/science.290.5492.806>
- Bowles, J., Koopman, P., 2007. Retinoic acid, meiosis and germ cell fate in mamMALLS. *Development* 134, 3401–3411. <https://doi.org/10.1242/dev.001107>
- Brick, K., Smagulova, F., Khil, P., Camerini-Otero, R.D., Petukhova, G.V., 2012. Genetic recombination is directed away from functional genomic elements in mice. *Nature* 485, 642–645. <https://doi.org/10.1038/nature11089>
- Brown, J.T., Bai, X., Johnson, A.W., 2000. The yeast antiviral proteins Ski2p, Ski3p, and Ski8p exist as a complex in vivo. *RNA* 6, 449–457.



- Brown, M.S., Grubb, J., Zhang, A., Rust, M.J., Bishop, D.K., 2015. Small Rad51 and Dmc1 Complexes Often Co-occupy Both Ends of a Meiotic DNA Double Strand Break. *PLOS Genet.* 11, e1005653. <https://doi.org/10.1371/journal.pgen.1005653>
- Buard, J., Barthès, P., Grey, C., de Massy, B., 2009. Distinct histone modifications define initiation and repair of meiotic recombination in the mouse. *EMBO J.* 28, 2616–2624. <https://doi.org/10.1038/emboj.2009.207>
- Buard, J., Rivals, E., Dunoyer de Segonzac, D., Garres, C., Caminade, P., de Massy, B., Boursot, P., 2014. Diversity of Prdm9 Zinc Finger Array in Wild Mice Unravels New Facets of the Evolutionary Turnover of this Coding Minisatellite. *PLoS ONE* 9, e85021. <https://doi.org/10.1371/journal.pone.0085021>
- Busygina, V., Sehorn, M.G., Shi, I.Y., Tsubouchi, H., Roeder, G.S., Sung, P., 2008. Hed1 regulates Rad51-mediated recombination via a novel mechanism. *Genes Dev.* 22, 786–795. <https://doi.org/10.1101/gad.1638708>
- Cannavo, E., Cejka, P., 2014. Sae2 promotes dsDNA endonuclease activity within Mre11-Rad50-Xrs2 to resect DNA breaks. *Nature* 514, 122–125. <https://doi.org/10.1038/nature13771>
- Carballo, J.A., Johnson, A.L., Sedgwick, S.G., Cha, R.S., 2008. Phosphorylation of the axial element protein Hop1 by Mec1/Tel1 ensures meiotic interhomolog recombination. *Cell* 132, 758–770. <https://doi.org/10.1016/j.cell.2008.01.035>
- Carballo, J.A., Panizza, S., Serrentino, M.E., Johnson, A.L., Geymonat, M., Borde, V., Klein, F., Cha, R.S., 2013. Budding yeast ATM/ATR control meiotic double-strand break (DSB) levels by down-regulating Rec114, an essential component of the DSB-machinery. *PLoS Genet.* 9, e1003545. <https://doi.org/10.1371/journal.pgen.1003545>
- Cavassim, M.I.A., Baker, Z., Hoge, C., Schierup, M.H., Schumer, M., Przeworski, M., 2022. PRDM9 losses in vertebrates are coupled to those of paralogs ZCWPW1 and ZCWPW2. *Proc. Natl. Acad. Sci. U. S. A.* 119, e2114401119. <https://doi.org/10.1073/pnas.2114401119>
- Chen, S.H., Chan, N.-L., Hsieh, T., 2013. New mechanistic and functional insights into DNA topoisomerases. *Annu. Rev. Biochem.* 82, 139–170. <https://doi.org/10.1146/annurev-biochem-061809-100002>
- Chen, X., Suhandynata, R.T., Sandhu, R., Rockmill, B., Mohibullah, N., Niu, H., Liang, J., Lo, H.-C., Miller, D.E., Zhou, H., Börner, G.V., Hollingsworth, N.M., 2015. Phosphorylation of the Synaptonemal Complex Protein Zip1 Regulates the Crossover/Noncrossover Decision during Yeast Meiosis. *PLOS Biol.* 13, e1002329. <https://doi.org/10.1371/journal.pbio.1002329>
- Claeys Bouuaert, C., Pu, S., Wang, J., Oger, C., Daccache, D., Xie, W., Patel, D.J., Keeney, S., 2021a. DNA-driven condensation assembles the meiotic DNA break machinery. *Nature* 592, 144–149. <https://doi.org/10.1038/s41586-021-03374-w>
- Claeys Bouuaert, C., Tischfield, S.E., Pu, S., Mimitou, E.P., Arias-Palomo, E., Berger, J.M., Keeney, S., 2021b. Structural and functional characterization of the Spo11 core complex. *Nat. Struct. Mol. Biol.* 28, 92–102. <https://doi.org/10.1038/s41594-020-00534-w>
- Cloud, V., Chan, Y.-L., Grubb, J., Budke, B., Bishop, D.K., 2012. Dmc1 catalyzes interhomolog joint molecule formation in meiosis with Rad51 and Mei5-Sae3 as accessory factors. *Science* 337, 1222–1225. <https://doi.org/10.1126/science.1219379>
- Corbett, K.D., Benedetti, P., Berger, J.M., 2007. Holoenzyme assembly and ATP-mediated conformational dynamics of topoisomerase VI. *Nat. Struct. Mol. Biol.* 14, 611–619. <https://doi.org/10.1038/nsmb1264>

- Corbett, K.D., Berger, J.M., 2005. Structural dissection of ATP turnover in the prototypical GHF ATPase TopoVI. *Struct. Lond. Engl.* 1993 13, 873–882.  
<https://doi.org/10.1016/j.str.2005.03.013>
- Corbett, K.D., Berger, J.M., 2004. Structure, molecular mechanisms, and evolutionary relationships in DNA topoisomerases. *Annu. Rev. Biophys. Biomol. Struct.* 33, 95–118. <https://doi.org/10.1146/annurev.biophys.33.110502.140357>
- Corbett, K.D., Berger, J.M., 2003. Structure of the topoisomerase VI-B subunit: implications for type II topoisomerase mechanism and evolution. *EMBO J.* 22, 151–163. <https://doi.org/10.1093/emboj/cdg008>
- Costa, Y., Speed, R., Öllinger, R., Alsheimer, M., Semple, C.A., Gautier, P., Maratou, K., Novak, I., Höög, C., Benavente, R., Cooke, H.J., 2005. Two novel proteins recruited by synaptonemal complex protein 1 (SYCP1) are at the centre of meiosis. *J. Cell Sci.* 118, 2755–2762. <https://doi.org/10.1242/jcs.02402>
- Daniel, K., Lange, J., Hached, K., Fu, J., Anastassiadis, K., Roig, I., Cooke, H.J., Stewart, A.F., Wassmann, K., Jasin, M., Keeney, S., Tóth, A., 2011. Meiotic homologous chromosome alignment and its surveillance are controlled by mouse HORMAD1. *Nat. Cell Biol.* 13, 599–610. <https://doi.org/10.1038/ncb2213>
- Davies, B., Hatton, E., Altemose, N., Hussin, J.G., Pratto, F., Zhang, G., Hinch, A.G., Moralli, D., Biggs, D., Diaz, R., Preece, C., Li, R., Bitoun, E., Brick, K., Green, C.M., Camerini-Otero, R.D., Myers, S.R., Donnelly, P., 2016. Re-engineering the zinc fingers of PRDM9 reverses hybrid sterility in mice. *Nature* 530, 171–176. <https://doi.org/10.1038/nature16931>
- de los Santos, T., Hunter, N., Lee, C., Larkin, B., Loidl, J., Hollingsworth, N.M., 2003. The Mus81/Mms4 endonuclease acts independently of double-Holliday junction resolution to promote a distinct subset of crossovers during meiosis in budding yeast. *Genetics* 164, 81–94.
- De Massy, B., 2013. Initiation of Meiotic Recombination: How and Where? Conversation and Specificities Among Eukaryotes. *Annu. Rev. Genet.* 47, 563–599.
- De Muyt, A., Jessop, L., Kolar, E., Sourirajan, A., Chen, J., Dayani, Y., Lichten, M., 2012. BLM helicase ortholog Sgs1 is a central regulator of meiotic recombination intermediate metabolism. *Mol. Cell* 46, 43–53. <https://doi.org/10.1016/j.molcel.2012.02.020>
- De Muyt, A., Pyatnitskaya, A., Andréani, J., Ranjha, L., Ramus, C., Laureau, R., Fernandez-Vega, A., Holoch, D., Girard, E., Govin, J., Margueron, R., Couté, Y., Cejka, P., Guérois, R., Borde, V., 2018. A meiotic XPF–ERCC1-like complex recognizes joint molecule recombination intermediates to promote crossover formation. *Genes Dev.* 32, 283–296. <https://doi.org/10.1101/gad.308510.117>
- Dereli, I., Stanzione, M., Olmeda, F., Papanikos, F., Baumann, M., Demir, S., Carofiglio, F., Lange, J., de Massy, B., Baarends, W.M., Turner, J., Rulands, S., Tóth, A., 2021. Four-pronged negative feedback of DSB machinery in meiotic DNA-break control in mice. *Nucleic Acids Res.* 49, 2609–2628. <https://doi.org/10.1093/nar/gkab082>
- Derheimer, F.A., Kastan, M.B., 2010. Multiple roles of ATM in monitoring and maintaining DNA integrity. *FEBS Lett.* 584, 3675–3681. <https://doi.org/10.1016/j.febslet.2010.05.031>
- Dernburg, A.F., McDonald, K., Moulder, G., Barstead, R., Dresser, M., Villeneuve, A.M., 1998. Meiotic recombination in *C. elegans* initiates by a conserved mechanism and is dispensable for homologous chromosome synapsis. *Cell* 94, 387–398. [https://doi.org/10.1016/s0092-8674\(00\)81481-6](https://doi.org/10.1016/s0092-8674(00)81481-6)

- Diagouraga, B., Clément, J.A.J., Duret, L., Kadlec, J., de Massy, B., Baudat, F., 2018. PRDM9 Methyltransferase Activity Is Essential for Meiotic DNA Double-Strand Break Formation at Its Binding Sites. *Mol. Cell* 69, 853-865.e6. <https://doi.org/10.1016/j.molcel.2018.01.033>
- Ding, X., Xu, R., Yu, J., Xu, T., Zhuang, Y., Han, M., 2007. SUN1 Is Required for Telomere Attachment to Nuclear Envelope and Gametogenesis in Mice. *Dev. Cell* 12, 863–872. <https://doi.org/10.1016/j.devcel.2007.03.018>
- Ehmsen, K.T., Heyer, W.-D., 2008. Biochemistry of Meiotic Recombination: Formation, Processing, and Resolution of Recombination Intermediates. *Genome Dyn. Stab.* 3, 91. [https://doi.org/10.1007/7050\\_2008\\_039](https://doi.org/10.1007/7050_2008_039)
- Fan, X., Wang, J., Zhang, X., Yang, Z., Zhang, J.-C., Zhao, L., Peng, H.-L., Lei, J., Wang, H.-W., 2019. Single particle cryo-EM reconstruction of 52 kDa streptavidin at 3.2 Angstrom resolution. *Nat. Commun.* 10, 2386. <https://doi.org/10.1038/s41467-019-10368-w>
- Fasching, C.L., Cejka, P., Kowalczykowski, S.C., Heyer, W.-D., 2015. Top3-Rmi1 dissolve Rad51-mediated D loops by a topoisomerase-based mechanism. *Mol. Cell* 57, 595–606. <https://doi.org/10.1016/j.molcel.2015.01.022>
- Fraune, J., Schramm, S., Alsheimer, M., Benavente, R., 2012. The mammalian synaptonemal complex: Protein components, assembly and role in meiotic recombination. *Exp. Cell Res., Experimental Cell Research Special Review Issue: Chromosome Biology*, 2012 318, 1340–1346. <https://doi.org/10.1016/j.yexcr.2012.02.018>
- Fujiwara, Y., Horisawa-Takada, Y., Inoue, E., Tani, N., Shibuya, H., Fujimura, S., Kariyazono, R., Sakata, T., Ohta, K., Araki, K., Okada, Y., Ishiguro, K., 2020. Meiotic cohesins mediate initial loading of HORMAD1 to the chromosomes and coordinate SC formation during meiotic prophase. *PLOS Genet.* 16, e1009048. <https://doi.org/10.1371/journal.pgen.1009048>
- Fukuda, T., Daniel, K., Wojtasz, L., Toth, A., Höög, C., 2010. A novel mammalian HORMA domain-containing protein, HORMAD1, preferentially associates with unsynapsed meiotic chromosomes. *Exp. Cell Res.* 316, 158–171. <https://doi.org/10.1016/j.yexcr.2009.08.007>
- Garcia, V., Gray, S., Allison, R.M., Cooper, T.J., Neale, M.J., 2015. Tel1(ATM)-mediated interference suppresses clustered meiotic double-strand-break formation. *Nature* 520, 114–118. <https://doi.org/10.1038/nature13993>
- Garcia, V., Phelps, S.E.L., Gray, S., Neale, M.J., 2011. Bidirectional resection of DNA double-strand breaks by Mre11 and Exo1. *Nature* 479, 241–244. <https://doi.org/10.1038/nature10515>
- Gómez-H, L., Felipe-Medina, N., Sánchez-Martín, M., Davies, O.R., Ramos, I., García-Tuñón, I., de Rooij, D.G., Dereli, I., Tóth, A., Barbero, J.L., Benavente, R., Llano, E., Pendas, A.M., 2016. C14ORF39/SIX6OS1 is a constituent of the synaptonemal complex and is essential for mouse fertility. *Nat. Commun.* 7, 13298. <https://doi.org/10.1038/ncomms13298>
- Graille, M., Cladière, L., Durand, D., Lecoite, F., Gadelle, D., Quevillon-Cheruel, S., Vachette, P., Forterre, P., van Tilbeurgh, H., 2008. Crystal Structure of an Intact Type II DNA Topoisomerase: Insights into DNA Transfer Mechanisms. *Structure* 16, 360–370. <https://doi.org/10.1016/j.str.2007.12.020>
- Gray, S., Cohen, P.E., 2016. Control of Meiotic Crossovers: From Double-Strand Break Formation to Designation. *Annu. Rev. Genet.* 50, 175–210. <https://doi.org/10.1146/annurev-genet-120215-035111>

- Grey, C., Baudat, F., de Massy, B., 2018. PRDM9, a driver of the genetic map. *PLoS Genet.* 14, e1007479. <https://doi.org/10.1371/journal.pgen.1007479>
- Grey, C., de Massy, B., 2021. Chromosome Organization in Early Meiotic Prophase. *Front. Cell Dev. Biol.* 9, 688878. <https://doi.org/10.3389/fcell.2021.688878>
- Guiraldelli, M.F., Felberg, A., Almeida, L.P., Parikh, A., de Castro, R.O., Pezza, R.J., 2018. SHOC1 is a ERCC4-(HhH)2-like protein, integral to the formation of crossover recombination intermediates during mammalian meiosis. *PLoS Genet.* 14, e1007381. <https://doi.org/10.1371/journal.pgen.1007381>
- Haber, J.E., 2000. Partners and pathways. *Trends Genet.* 16, 259–264. [https://doi.org/10.1016/S0168-9525\(00\)02022-9](https://doi.org/10.1016/S0168-9525(00)02022-9)
- Hamer, G., Gell, K., Kouznetsova, A., Novak, I., Benavente, R., Höög, C., 2006. Characterization of a novel meiosis-specific protein within the central element of the synaptonemal complex. *J. Cell Sci.* 119, 4025–4032. <https://doi.org/10.1242/jcs.03182>
- Handel, M.A., 2004. The XY body: a specialized meiotic chromatin domain. *Exp. Cell Res.* 296, 57–63. <https://doi.org/10.1016/j.yexcr.2004.03.008>
- Handel, M.A., Schimenti, J.C., 2010. Genetics of mammalian meiosis: regulation, dynamics and impact on fertility. *Nat. Rev. Genet.* 11, 124–136. <https://doi.org/10.1038/nrg2723>
- Hartung, F., Puchta, H., 2001. Molecular characterization of homologues of both subunits A (SPO11) and B of the archaeobacterial topoisomerase 6 in plants. *Gene* 271, 81–86. [https://doi.org/10.1016/s0378-1119\(01\)00496-6](https://doi.org/10.1016/s0378-1119(01)00496-6)
- Hartung, F., Puchta, H., 2000. Molecular characterisation of two paralogous SPO11 homologues in *Arabidopsis thaliana*. *Nucleic Acids Res.* 28, 1548–1554.
- Hassold, T., Hunt, P., 2001. To err (meiotically) is human: the genesis of human aneuploidy. *Nat. Rev. Genet.* 2, 280–291. <https://doi.org/10.1038/35066065>
- He, C., Liu, N., Xie, D., Liu, Y., Xiao, Y., Li, F., 2019. Structural basis for histone H3K4me3 recognition by the N-terminal domain of the PHD finger protein Spp1. *Biochem. J.* 476, 1957–1973. <https://doi.org/10.1042/BCJ20190091>
- Henderson, K.A., Kee, K., Maleki, S., Santini, P., Keeney, S., 2006. Cyclin-dependent kinase directly regulates initiation of meiotic recombination. *Cell* 125, 1321–1332. <https://doi.org/10.1016/j.cell.2006.04.039>
- Herbert, M., Kalleas, D., Cooney, D., Lamb, M., Lister, L., 2015. Meiosis and Maternal Aging: Insights from Aneuploid Oocytes and Trisomy Births. *Cold Spring Harb. Perspect. Biol.* 7, a017970. <https://doi.org/10.1101/cshperspect.a017970>
- Heyer, W.-D., 2004. Recombination: Holliday Junction Resolution and Crossover Formation. *Curr. Biol.* 14, R56–R58. <https://doi.org/10.1016/j.cub.2003.12.043>
- Hillers, K.J., 2004. Crossover interference. *Curr. Biol.* 14, R1036–R1037. <https://doi.org/10.1016/j.cub.2004.11.038>
- Hinch, A.G., Zhang, G., Becker, P.W., Moralli, D., Hinch, R., Davies, B., Bowden, R., Donnelly, P., 2019. Factors Influencing Meiotic Recombination Revealed by Whole Genome Sequencing of Single Sperm. *Science* 363, eaau8861. <https://doi.org/10.1126/science.aau8861>
- Holloway, J.K., Booth, J., Edelman, W., McGowan, C.H., Cohen, P.E., 2008. MUS81 Generates a Subset of MLH1-MLH3-Independent Crossovers in Mammalian Meiosis. *PLoS Genet.* 4. <https://doi.org/10.1371/journal.pgen.1000186>
- Holloway, J.K., Mohan, S., Balmus, G., Sun, X., Modzelewski, A., Borst, P.L., Freire, R., Weiss, R.S., Cohen, P.E., 2011. Mammalian BTBD12 (SLX4) protects against genomic instability during mammalian spermatogenesis. *PLoS Genet.* 7, e1002094. <https://doi.org/10.1371/journal.pgen.1002094>

- Huang, T., Yuan, S., Gao, L., Li, M., Yu, X., Zhan, J., Yin, Y., Liu, C., Zhang, C., Lu, G., Li, W., Liu, J., Chen, Z.-J., Liu, H., 2020. The histone modification reader ZCWPW1 links histone methylation to PRDM9-induced double-strand break repair. *eLife* 9, e53459. <https://doi.org/10.7554/eLife.53459>
- Hunter, N., 2015. Meiotic Recombination: The Essence of Heredity. *Cold Spring Harb. Perspect. Biol.* 7. <https://doi.org/10.1101/cshperspect.a016618>
- Hunter, N., Kleckner, N., 2001. The Single-End Invasion: An Asymmetric Intermediate at the Double-Strand Break to Double-Holliday Junction Transition of Meiotic Recombination. *Cell* 106, 59–70. [https://doi.org/10.1016/S0092-8674\(01\)00430-5](https://doi.org/10.1016/S0092-8674(01)00430-5)
- Imai, Y., Baudat, F., Taillepierre, M., Stanzione, M., Toth, A., de Massy, B., 2017. The PRDM9 KRAB domain is required for meiosis and involved in protein interactions. *Chromosoma* 126, 681–695. <https://doi.org/10.1007/s00412-017-0631-z>
- Jessop, L., Lichten, M., 2008. Mus81/Mms4 endonuclease and Sgs1 helicase collaborate to ensure proper recombination intermediate metabolism during meiosis. *Mol. Cell* 31, 313–323. <https://doi.org/10.1016/j.molcel.2008.05.021>
- Jones, G.H., Franklin, F.C.H., 2006. Meiotic crossing-over: obligation and interference. *Cell* 126, 246–248. <https://doi.org/10.1016/j.cell.2006.07.010>
- Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Židek, A., Potapenko, A., Bridgland, A., Meyer, C., Kohl, S.A.A., Ballard, A.J., Cowie, A., Romera-Paredes, B., Nikolov, S., Jain, R., Adler, J., Back, T., Petersen, S., Reiman, D., Clancy, E., Zielinski, M., Steinegger, M., Pacholska, M., Berghammer, T., Bodenstein, S., Silver, D., Vinyals, O., Senior, A.W., Kavukcuoglu, K., Kohli, P., Hassabis, D., 2021. Highly accurate protein structure prediction with AlphaFold. *Nature* 596, 583–589. <https://doi.org/10.1038/s41586-021-03819-2>
- Kauppi, L., Barchi, M., Lange, J., Baudat, F., Jasin, M., Keeney, S., 2013. Numerical constraints and feedback control of double-strand breaks in mouse meiosis. *Genes Dev.* 27, 873–886. <https://doi.org/10.1101/gad.213652.113>
- Keeney, S., 2008. Spo11 and the Formation of DNA Double-Strand Breaks in Meiosis. *Genome Dyn. Stab.* 2, 81–123. [https://doi.org/10.1007/7050\\_2007\\_026](https://doi.org/10.1007/7050_2007_026)
- Keeney, S., Baudat, F., Angeles, M., Zhou, Z.-H., Copeland, N.G., Jenkins, N.A., Manova, K., Jasin, M., 1999. A Mouse Homolog of the *Saccharomyces cerevisiae* Meiotic Recombination DNA Transesterase Spo11p. *Genomics* 61, 170–182. <https://doi.org/10.1006/geno.1999.5956>
- Keeney, S., Giroux, C.N., Kleckner, N., 1997. Meiosis-Specific DNA Double-Strand Breaks Are Catalyzed by Spo11, a Member of a Widely Conserved Protein Family. *Cell* 88, 375–384. [https://doi.org/10.1016/S0092-8674\(00\)81876-0](https://doi.org/10.1016/S0092-8674(00)81876-0)
- Keeney, S., Lange, J., Mohibullah, N., 2014. Self-organization of meiotic recombination initiation: general principles and molecular pathways. *Annu. Rev. Genet.* 48, 187–214. <https://doi.org/10.1146/annurev-genet-120213-092304>
- Khoshouei, M., Radjainia, M., Baumeister, W., Danev, R., 2017. Cryo-EM structure of haemoglobin at 3.2 Å determined with the Volta phase plate. *Nat. Commun.* 8, 16099. <https://doi.org/10.1038/ncomms16099>
- Kleckner, N., 2006. Chiasma formation: chromatin/axis interplay and the role(s) of the synaptonemal complex. *Chromosoma* 115, 175–194. <https://doi.org/10.1007/s00412-006-0055-7>
- Kleckner, N., 1996. Meiosis: how could it work? *Proc. Natl. Acad. Sci.* 93, 8167–8174. <https://doi.org/10.1073/pnas.93.16.8167>
- Kogo, H., Tsutsumi, M., Inagaki, H., Ohye, T., Kiyonari, H., Kurahashi, H., 2012. HORMAD2 is essential for synapsis surveillance during meiotic prophase via the

- recruitment of ATR activity. *Genes Cells* 17, 897–912.  
<https://doi.org/10.1111/gtc.12005>
- Kohl, K.P., Sekelsky, J., 2013. Meiotic and Mitotic Recombination in Meiosis. *Genetics* 194, 327–334. <https://doi.org/10.1534/genetics.113.150581>
- Kono, H., Tamura, M., Osada, N., Suzuki, H., Abe, K., Moriwaki, K., Ohta, K., Shiroishi, T., 2014. Prdm9 Polymorphism Unveils Mouse Evolutionary Tracks. *DNA Res. Int. J. Rapid Publ. Rep. Genes Genomes* 21, 315–326.  
<https://doi.org/10.1093/dnares/dst059>
- Kumar, R., Bourbon, H.-M., de Massy, B., 2010. Functional conservation of Mei4 for meiotic DNA double-strand break formation from yeasts to mice. *Genes Dev.* 24, 1266–1280. <https://doi.org/10.1101/gad.571710>
- Kumar, R., Ghyselinck, N., Ishiguro, K., Watanabe, Y., Kouznetsova, A., Höög, C., Strong, E., Schimenti, J., Daniel, K., Toth, A., de Massy, B., 2015. MEI4 – a central player in the regulation of meiotic DNA double-strand break formation in the mouse. *J. Cell Sci.* 128, 1800–1811. <https://doi.org/10.1242/jcs.165464>
- Kumar, R., Oliver, C., Brun, C., Juarez-Martinez, A.B., Tarabay, Y., Kadlec, J., de Massy, B., 2018. Mouse REC114 is essential for meiotic DNA double-strand break formation and forms a complex with MEI4. *Life Sci. Alliance* 1, e201800259.  
<https://doi.org/10.26508/lsa.201800259>
- Lam, I., Keeney, S., 2015. Mechanism and Regulation of Meiotic Recombination Initiation. *Cold Spring Harb. Perspect. Biol.* 7, a016634.  
<https://doi.org/10.1101/cshperspect.a016634>
- Lange, J., Pan, J., Cole, F., Thelen, M.P., Jasin, M., Keeney, S., 2011. ATM controls meiotic double-strand-break formation. *Nature* 479, 237–240.  
<https://doi.org/10.1038/nature10508>
- Lange, J., Yamada, S., Tischfield, S.E., Pan, J., Kim, S., Zhu, X., Socci, N.D., Jasin, M., Keeney, S., 2016. The Landscape of Mouse Meiotic Double-Strand Break Formation, Processing, and Repair. *Cell* 167, 695-708.e16.  
<https://doi.org/10.1016/j.cell.2016.09.035>
- Lee, J.-H., Skalnik, D.G., 2005. CpG-binding Protein (CXXC Finger Protein 1) Is a Component of the Mammalian Set1 Histone H3-Lys4 Methyltransferase Complex, the Analogue of the Yeast Set1/COMPASS Complex \*. *J. Biol. Chem.* 280, 41725–41731. <https://doi.org/10.1074/jbc.M508312200>
- Li, J., Hooker, G.W., Roeder, G.S., 2006. *Saccharomyces cerevisiae* Mer2, Mei4 and Rec114 form a complex required for meiotic double-strand break formation. *Genetics* 173, 1969–1981. <https://doi.org/10.1534/genetics.106.058768>
- Li, P., Banjade, S., Cheng, H.-C., Kim, S., Chen, B., Guo, L., Llaguno, M., Hollingsworth, J.V., King, D.S., Banani, S.F., Russo, P.S., Jiang, Q.-X., Nixon, B.T., Rosen, M.K., 2012. Phase transitions in the assembly of multivalent signalling proteins. *Nature* 483, 336–340. <https://doi.org/10.1038/nature10879>
- Libby, B.J., De La Fuente, R., O'Brien, M.J., Wigglesworth, K., Cobb, J., Inselman, A., Eaker, S., Handel, M.A., Eppig, J.J., Schimenti, J.C., 2002. The Mouse Meiotic Mutation *mei1* Disrupts Chromosome Synapsis with Sexually Dimorphic Consequences for Meiotic Progression. *Dev. Biol.* 242, 174–187.  
<https://doi.org/10.1006/dbio.2001.0535>
- Libby, B.J., Reinholdt, L.G., Schimenti, J.C., 2003. Positional cloning and characterization of *Mei1*, a vertebrate-specific gene required for normal meiotic chromosome synapsis in mice. *Proc. Natl. Acad. Sci.* 100, 15706–15711.  
<https://doi.org/10.1073/pnas.2432067100>

- Liu, J.-G., Yuan, L., Brundell, E., Björkroth, B., Daneholt, B., Höög, C., 1996. Localization of the N-terminus of SCP1 to the Central Element of the Synaptonemal Complex and Evidence for Direct Interactions between the N-termini of SCP1 Molecules Organized Head-to-Head. *Exp. Cell Res.* 226, 11–19. <https://doi.org/10.1006/excr.1996.0197>
- Lukaszewicz, A., Lange, J., Keeney, S., Jasin, M., 2018. Control of meiotic double-strand-break formation by ATM: local and global views. *Cell Cycle* 17, 1155–1172. <https://doi.org/10.1080/15384101.2018.1464847>
- MacLennan, M., Crichton, J.H., Playfoot, C.J., Adams, I.R., 2015. Oocyte development, meiosis and aneuploidy. *Semin. Cell Dev. Biol.* 45, 68–76. <https://doi.org/10.1016/j.semdb.2015.10.005>
- Mahgoub, M., Paiano, J., Bruno, M., Wu, W., Pathuri, S., Zhang, X., Ralls, S., Cheng, X., Nussenzweig, A., Macfarlan, T.S., n.d. Dual histone methyl reader ZCWPW1 facilitates repair of meiotic double strand breaks in male mice. *eLife* 9, e53360. <https://doi.org/10.7554/eLife.53360>
- Maleki, S., Neale, M.J., Arora, C., Henderson, K.A., Keeney, S., 2007. Interactions between Mei4, Rec114, and other proteins required for meiotic DNA double-strand break formation in *Saccharomyces cerevisiae*. *Chromosoma* 116, 471–486. <https://doi.org/10.1007/s00412-007-0111-y>
- Marsolier-Kergoat, M.-C., Khan, M.M., Schott, J., Zhu, X., Llorente, B., 2018. Mechanistic View and Genetic Control of DNA Recombination during Meiosis. *Mol. Cell* 70, 9–20.e6. <https://doi.org/10.1016/j.molcel.2018.02.032>
- Martinez, J.S., von Nicolai, C., Kim, T., Ehlén, Å., Mazin, A.V., Kowalczykowski, S.C., Carreira, A., 2016. BRCA2 regulates DMC1-mediated recombination through the BRC repeats. *Proc. Natl. Acad. Sci. U. S. A.* 113, 3515–3520. <https://doi.org/10.1073/pnas.1601691113>
- Martini, E., Diaz, R.L., Hunter, N., Keeney, S., 2006. Crossover homeostasis in yeast meiosis. *Cell* 126, 285–295. <https://doi.org/10.1016/j.cell.2006.05.044>
- Matos, J., Blanco, M.G., Maslen, S., Skehel, J.M., West, S.C., 2011. Regulatory control of the resolution of DNA recombination intermediates during meiosis and mitosis. *Cell* 147, 158–172. <https://doi.org/10.1016/j.cell.2011.08.032>
- McKim, K.S., Green-Marroquin, B.L., Sekelsky, J.J., Chin, G., Steinberg, C., Khodosh, R., Hawley, R.S., 1998. Meiotic Synapsis in the Absence of Recombination. *Science* 279, 876–878. <https://doi.org/10.1126/science.279.5352.876>
- Mimitou, E.P., Yamada, S., Keeney, S., 2017. A global view of meiotic double-strand break end resection. *Science* 355, 40–45. <https://doi.org/10.1126/science.aak9704>
- Moens, P.B., Kolas, N.K., Tarsounas, M., Marcon, E., Cohen, P.E., Spyropoulos, B., 2002. The time course and chromosomal localization of recombination-related proteins at meiosis in the mouse are compatible with models that can resolve the early DNA-DNA interactions without reciprocal recombination. *J. Cell Sci.* 115, 1611–1622. <https://doi.org/10.1242/jcs.115.8.1611>
- Morimoto, A., Shibuya, H., Zhu, X., Kim, J., Ishiguro, K., Han, M., Watanabe, Y., 2012. A conserved KASH domain protein associates with telomeres, SUN1, and dynactin during mammalian meiosis. *J. Cell Biol.* 198, 165–172. <https://doi.org/10.1083/jcb.201204085>
- Mosavi, L.K., Cammett, T.J., Desrosiers, D.C., Peng, Z.-Y., 2004. The ankyrin repeat as molecular architecture for protein recognition. *Protein Sci. Publ. Protein Soc.* 13, 1435–1448. <https://doi.org/10.1110/ps.03554604>

- Murakami, H., Keeney, S., 2014. Temporospatial coordination of meiotic DNA replication and recombination via DDK recruitment to replisomes. *Cell* 158, 861–873. <https://doi.org/10.1016/j.cell.2014.06.028>
- Murakami, H., Keeney, S., 2008. Regulating the formation of DNA double-strand breaks in meiosis. *Genes Dev.* 22, 286–292. <https://doi.org/10.1101/gad.1642308>
- Murton, B.L., Chin, W.L., Ponting, C.P., Itzhaki, L.S., 2010. Characterising the binding specificities of the subunits associated with the KMT2/Set1 histone lysine methyltransferase. *J. Mol. Biol.* 398, 481–488. <https://doi.org/10.1016/j.jmb.2010.03.036>
- Myers, S., Bowden, R., Tumian, A., Bontrop, R.E., Freeman, C., MacFie, T.S., McVean, G., Donnelly, P., 2010. Drive Against Hotspot Motifs in Primates Implicates the PRDM9 gene in Meiotic Recombination. *Science* 327. <https://doi.org/10.1126/science.1182363>
- Neale, M.J., Pan, J., Keeney, S., 2005. Endonucleolytic processing of covalent protein-linked DNA double-strand breaks. *Nature* 436, 1053–1057. <https://doi.org/10.1038/nature03872>
- Nichols, M.D., DeAngelis, K., Keck, J.L., Berger, J.M., 1999. Structure and function of an archaeal topoisomerase VI subunit with homology to the meiotic recombination factor Spo11. *EMBO J.* 18, 6177–6188. <https://doi.org/10.1093/emboj/18.21.6177>
- Nore, A., Juarez-Martinez, A.B., Clément, J., Brun, C., Diagouraga, B., Grey, C., Bourbon, H.M., Kadlec, J., Robert, T., de Massy, B., 2022. TOPOVIBL-REC114 interaction regulates meiotic DNA double-strand breaks. *bioRxiv* 2021.11.30.470517. <https://doi.org/10.1101/2021.11.30.470517>
- Oh, S.D., Lao, J.P., Taylor, A.F., Smith, G.R., Hunter, N., 2008. RecQ Helicase, Sgs1, and XPF-Family Endonuclease, Mus81-Mms4, Resolve Aberrant Joint Molecules During Meiotic Recombination. *Mol. Cell* 31, 324–336. <https://doi.org/10.1016/j.molcel.2008.07.006>
- Oliver, P.L., Goodstadt, L., Bayes, J.J., Birtle, Z., Roach, K.C., Phadnis, N., Beatson, S.A., Lunter, G., Malik, H.S., Ponting, C.P., 2009. Accelerated evolution of the Prdm9 speciation gene across diverse metazoan taxa. *PLoS Genet.* 5, e1000753. <https://doi.org/10.1371/journal.pgen.1000753>
- Page, S.L., Hawley, R.S., 2004. THE GENETICS AND MOLECULAR BIOLOGY OF THE SYNAPTONEMAL COMPLEX. *Annu. Rev. Cell Dev. Biol.* 20, 525–558. <https://doi.org/10.1146/annurev.cellbio.19.111301.155141>
- Paiano, J., Wu, W., Yamada, S., Sciascia, N., Callen, E., Paola Cotrim, A., Deshpande, R.A., Maman, Y., Day, A., Paull, T.T., Nussenzweig, A., 2020. ATM and PRDM9 regulate SPO11-bound recombination intermediates during meiosis. *Nat. Commun.* 11. <https://doi.org/10.1038/s41467-020-14654-w>
- Pan, J., Sasaki, M., Kniewel, R., Murakami, H., Blitzblau, H.G., Tischfield, S.E., Zhu, X., Neale, M.J., Jasin, M., Socci, N.D., Hochwagen, A., Keeney, S., 2011. A hierarchical combination of factors shapes the genome-wide topography of yeast meiotic recombination initiation. *Cell* 144, 719–731. <https://doi.org/10.1016/j.cell.2011.02.009>
- Panizza, S., Mendoza, M.A., Berlinger, M., Huang, L., Nicolas, A., Shirahige, K., Klein, F., 2011. Spo11-accessory proteins link double-strand break sites to the chromosome axis in early meiotic recombination. *Cell* 146, 372–383. <https://doi.org/10.1016/j.cell.2011.07.003>
- Papanikos, F., Clément, J.A.J., Testa, E., Ravindranathan, R., Grey, C., Dereli, I., Bondarieva, A., Valerio-Cabrera, S., Stanzone, M., Schleiffer, A., Jansa, P., Lustyk, D., Fei, J.-F., Adams, I.R., Forejt, J., Barchi, M., de Massy, B., Toth, A.,



2019. Mouse ANKRD31 Regulates Spatiotemporal Patterning of Meiotic Recombination Initiation and Ensures Recombination between X and Y Sex Chromosomes. *Mol. Cell* 74, 1069-1085.e11. <https://doi.org/10.1016/j.molcel.2019.03.022>
- Pâques, F., Haber, J.E., 1999. Multiple Pathways of Recombination Induced by Double-Strand Breaks in *Saccharomyces cerevisiae*. *Microbiol. Mol. Biol. Rev.* 63, 349–404.
- Parvanov, E.D., Petkov, P.M., Paigen, K., 2010. Prdm9 controls activation of mammalian recombination hotspots. *Science* 327, 835. <https://doi.org/10.1126/science.1181495>
- Parvanov, E.D., Tian, H., Billings, T., Saxl, R.L., Spruce, C., Aithal, R., Krejci, L., Paigen, K., Petkov, P.M., 2017. PRDM9 interactions with other proteins provide a link between recombination hotspots and the chromosomal axis in meiosis. *Mol. Biol. Cell* 28, 488–499. <https://doi.org/10.1091/mbc.E16-09-0686>
- Petukhova, G.V., Pezza, R.J., Vanevski, F., Ploquin, M., Masson, J.-Y., Camerini-Otero, R.D., 2005. The Hop2 and Mnd1 proteins act in concert with Rad51 and Dmc1 in meiotic recombination. *Nat. Struct. Mol. Biol.* 12, 449–453. <https://doi.org/10.1038/nsmb923>
- Pezza, R.J., Voloshin, O.N., Vanevski, F., Camerini-Otero, R.D., 2007. Hop2/Mnd1 acts on two critical steps in Dmc1-promoted homologous pairing. *Genes Dev.* 21, 1758–1766. <https://doi.org/10.1101/gad.1562907>
- Powers, N.R., Parvanov, E.D., Baker, C.L., Walker, M., Petkov, P.M., Paigen, K., 2016. The Meiotic Recombination Activator PRDM9 Trimethylates Both H3K36 and H3K4 at Recombination Hotspots In Vivo. *PLoS Genet.* 12, e1006146. <https://doi.org/10.1371/journal.pgen.1006146>
- Pyatnitskaya, A., Andreani, J., Guérois, R., Muylt, A.D., Borde, V., 2021. The Zip4 protein directly couples meiotic crossover formation to synaptonemal complex assembly. *Genes Dev.* <https://doi.org/10.1101/gad.348973.121>
- Pyatnitskaya, A., Borde, V., De Muylt, A., 2019. Crossing and zipping: molecular duties of the ZMM proteins in meiosis. *Chromosoma* 128, 181–198. <https://doi.org/10.1007/s00412-019-00714-8>
- Reinholdt, L.G., Schimenti, J.C., 2005. Mei1 is epistatic to Dmc1 during mouse meiosis. *Chromosoma* 114, 127–134. <https://doi.org/10.1007/s00412-005-0346-4>
- Ribeiro, J., Abby, E., Livera, G., Martini, E., 2016. RPA homologs and ssDNA processing during meiotic recombination. *Chromosoma* 125, 265–276. <https://doi.org/10.1007/s00412-015-0552-7>
- Robert, T., Massy, B. de, Grelon, M., 2017. TopoVIL - Un ciseau moléculaire essentiel à la reproduction. *médecine/sciences* 33, 512–518. <https://doi.org/10.1051/medsci/20173305015>
- Robert, T., Nore, A., Brun, C., Maffre, C., Crimi, B., Bourbon, H.-M., de Massy, B., 2016. The TopoVIB-Like protein family is required for meiotic DNA double-strand break formation. *Science* 351, 943–949. <https://doi.org/10.1126/science.aad5309>
- Robert, Thomas, Vrielynck, N., Mézard, C., de Massy, B., Grelon, M., 2016. A new light on the meiotic DSB catalytic complex. *Semin. Cell Dev. Biol., Mechanisms of cancer cachexia* 54, 165–176. <https://doi.org/10.1016/j.semcd.2016.02.025>
- Romanienko, P.J., Camerini-Otero, R.D., 2000. The Mouse Spo11 Gene Is Required for Meiotic Chromosome Synapsis. *Mol. Cell* 6, 975–987. [https://doi.org/10.1016/S1097-2765\(00\)00097-6](https://doi.org/10.1016/S1097-2765(00)00097-6)
- Rousová, D., Nivsarkar, V., Altmannova, V., Raina, V.B., Funk, S.K., Liedtke, D., Janning, P., Müller, F., Reichle, H., Vader, G., Weir, J.R., 2021. Novel mechanistic insights

- into the role of Mer2 as the keystone of meiotic DNA break formation. *eLife* 10, e72330. <https://doi.org/10.7554/eLife.72330>
- Royo, H., Polikiewicz, G., Mahadevaiah, S.K., Prosser, H., Mitchell, M., Bradley, A., de Rooij, D.G., Burgoyne, P.S., Turner, J.M.A., 2010. Evidence that meiotic sex chromosome inactivation is essential for male fertility. *Curr. Biol.* CB 20, 2117–2123. <https://doi.org/10.1016/j.cub.2010.11.010>
- Santucci-Darmanin, S., Walpita, D., Lespinasse, F., Desnuelle, C., Ashley, T., Paquis-Flucklinger, V., 2000. MSH4 acts in conjunction with MLH1 during mammalian meiosis. *FASEB J.* 14, 1539–1547. <https://doi.org/10.1096/fj.99-0851com>
- Sartori, A.A., Lukas, C., Coates, J., Mistrik, M., Fu, S., Bartek, J., Baer, R., Lukas, J., Jackson, S.P., 2007. Human CtIP promotes DNA end resection. *Nature* 450, 509–514. <https://doi.org/10.1038/nature06337>
- Sasaki, M., Lange, J., Keeney, S., 2010. Genome destabilization by homologous recombination in the germline. *Nat. Rev. Mol. Cell Biol.* 11, 182–195. <https://doi.org/10.1038/nrm2849>
- Sasanuma, H., Murakami, H., Fukuda, T., Shibata, T., Nicolas, A., Ohta, K., 2007. Meiotic association between Spo11 regulated by Rec102, Rec104 and Rec114. *Nucleic Acids Res.* 35, 1119–1133. <https://doi.org/10.1093/nar/gkl1162>
- Scherthan, H., 2001. A bouquet makes ends meet. *Nat. Rev. Mol. Cell Biol.* 2, 621–627. <https://doi.org/10.1038/35085086>
- Schimenti, J., 2005. Synapsis or silence. *Nat. Genet.* 37, 11–13. <https://doi.org/10.1038/ng0105-11>
- Schramm, S., Fraune, J., Naumann, R., Hernandez-Hernandez, A., Höög, C., Cooke, H.J., Alsheimer, M., Benavente, R., 2011. A Novel Mouse Synaptonemal Complex Protein Is Essential for Loading of Central Element Proteins, Recombination, and Fertility. *PLOS Genet.* 7, e1002088. <https://doi.org/10.1371/journal.pgen.1002088>
- Schwacha, A., Kleckner, N., 1995. Identification of double Holliday junctions as intermediates in meiotic recombination. *Cell* 83, 783–791. [https://doi.org/10.1016/0092-8674\(95\)90191-4](https://doi.org/10.1016/0092-8674(95)90191-4)
- Shiloh, Y., Ziv, Y., 2013. The ATM protein kinase: regulating the cellular response to genotoxic stress, and more. *Nat. Rev. Mol. Cell Biol.* 14, 197–210. <https://doi.org/10.1038/nrm3546>
- Shin, Y.-H., Choi, Y., Erdin, S.U., Yatsenko, S.A., Kloc, M., Yang, F., Wang, P.J., Meistrich, M.L., Rajkovic, A., 2010. Hormad1 mutation disrupts synaptonemal complex formation, recombination, and chromosome segregation in mammalian meiosis. *PLoS Genet.* 6, e1001190. <https://doi.org/10.1371/journal.pgen.1001190>
- Shinohara, A., Ogawa, H., Ogawa, T., 1992. Rad51 protein involved in repair and recombination in *S. cerevisiae* is a RecA-like protein. *Cell* 69, 457–470. [https://doi.org/10.1016/0092-8674\(92\)90447-k](https://doi.org/10.1016/0092-8674(92)90447-k)
- Shinohara, A., Shinohara, M., 2004. Roles of RecA homologues Rad51 and Dmc1 during meiotic recombination. *Cytogenet. Genome Res.* 107, 201–207. <https://doi.org/10.1159/000080598>
- Shiotani, B., Zou, L., 2009. Single-Stranded DNA Orchestrates an ATM-to-ATR Switch at DNA Breaks. *Mol. Cell* 33, 547–558. <https://doi.org/10.1016/j.molcel.2009.01.024>
- Smagulova, F., Gregoret, I.V., Brick, K., Khil, P., Camerini-Otero, R.D., Petukhova, G.V., 2011. Genome-Wide Analysis Reveals Novel Molecular Features of Mouse Recombination Hotspots. *Nature* 472, 375–378. <https://doi.org/10.1038/nature09869>
- Snowden, T., Acharya, S., Butz, C., Berardini, M., Fishel, R., 2004. hMSH4-hMSH5 recognizes Holliday Junctions and forms a meiosis-specific sliding clamp that

- embraces homologous chromosomes. *Mol. Cell* 15, 437–451.  
<https://doi.org/10.1016/j.molcel.2004.06.040>
- Sommermeier, V., Béneut, C., Chaplais, E., Serrentino, M.E., Borde, V., 2013. Spp1, a member of the Set1 Complex, promotes meiotic DSB formation in promoters by tethering histone H3K4 methylation sites to chromosome axes. *Mol. Cell* 49, 43–54.  
<https://doi.org/10.1016/j.molcel.2012.11.008>
- Spruce, C., Dlamini, S., Ananda, G., Bronkema, N., Tian, H., Paigen, K., Carter, G.W., Baker, C.L., 2020. HELLS and PRDM9 form a pioneer complex to open chromatin at meiotic recombination hot spots. *Genes Dev.* 34, 398–412.  
<https://doi.org/10.1101/gad.333542.119>
- Stacey, N.J., Kuromori, T., Azumi, Y., Roberts, G., Breuer, C., Wada, T., Maxwell, A., Roberts, K., Sugimoto-Shirasu, K., 2006. Arabidopsis SPO11-2 functions with SPO11-1 in meiotic recombination. *Plant J. Cell Mol. Biol.* 48, 206–216.  
<https://doi.org/10.1111/j.1365-313X.2006.02867.x>
- Stanzione, M., Baumann, M., Papanikos, F., Dereli, I., Lange, J., Ramlal, A., Tränkner, D., Shibuya, H., Massy, B. de, Watanabe, Y., Jasin, M., Keeney, S., Tóth, A., 2016. Meiotic DNA break formation requires the unsynapsed chromosome axis-binding protein IHO1 (CCDC36) in mice. *Nat. Cell Biol.* 18, 1208–1220.  
<https://doi.org/10.1038/ncb3417>
- Stark, H., 2010. GraFix: stabilization of fragile macromolecular complexes for single particle cryo-EM. *Methods Enzymol.* 481, 109–126. [https://doi.org/10.1016/S0076-6879\(10\)81005-5](https://doi.org/10.1016/S0076-6879(10)81005-5)
- Steiner, S., Kohli, J., Ludin, K., 2010. Functional interactions among members of the meiotic initiation complex in fission yeast. *Curr. Genet.* 56, 237–249.  
<https://doi.org/10.1007/s00294-010-0296-0>
- Sung, P., Klein, H., 2006. Mechanism of homologous recombination: mediators and helicases take on regulatory functions. *Nat. Rev. Mol. Cell Biol.* 7, 739–750.  
<https://doi.org/10.1038/nrm2008>
- Svendsen, J.M., Harper, J.W., 2010. GEN1/Yen1 and the SLX4 complex: solutions to the problem of Holliday junction resolution. *Genes Dev.* 24, 521–536.  
<https://doi.org/10.1101/gad.1903510>
- Szostak, J.W., Orr-Weaver, T.L., Rothstein, R.J., Stahl, F.W., 1983. The double-strand-break repair model for recombination. *Cell* 33, 25–35. [https://doi.org/10.1016/0092-8674\(83\)90331-8](https://doi.org/10.1016/0092-8674(83)90331-8)
- Tang, S., Wu, M.K.Y., Zhang, R., Hunter, N., 2015. Pervasive and essential roles of the Top3-Rmi1 decatenase orchestrate recombination and facilitate chromosome segregation in meiosis. *Mol. Cell* 57, 607–621.  
<https://doi.org/10.1016/j.molcel.2015.01.021>
- Tessé, S., Bourbon, H.-M., Debuchy, R., Budin, K., Dubois, E., Liangran, Z., Antoine, R., Piolot, T., Kleckner, N., Zickler, D., Espagne, E., 2017. Asy2/Mer2: an evolutionarily conserved mediator of meiotic recombination, pairing, and global chromosome compaction. *Genes Dev.* 31, 1880–1893.  
<https://doi.org/10.1101/gad.304543.117>
- Thorslund, T., Esashi, F., West, S.C., 2007. Interactions between human BRCA2 protein and the meiosis-specific recombinase DMC1. *EMBO J.* 26, 2915–2922.  
<https://doi.org/10.1038/sj.emboj.7601739>
- Tian, H., Billings, T., Petkov, P.M., 2021. EWSR1 affects PRDM9-dependent histone 3 methylation and provides a link between recombination hotspots and the chromosome axis protein REC8. *Mol. Biol. Cell* 32, 1–14.  
<https://doi.org/10.1091/mbc.E20-09-0604>

- Tian, H., Billings, T., Petkov, P.M., 2018. CXXC1 is not essential for normal DNA double-strand break formation and meiotic recombination in mouse. *PLoS Genet.* 14, e1007657. <https://doi.org/10.1371/journal.pgen.1007657>
- Tong, Y., Hughes, D., Placanica, L., Buck, M., 2005. When monomers are preferred: A strategy for the identification and disruption of weakly oligomerized proteins. *Struct. Lond. Engl.* 1993 13, 7–15. <https://doi.org/10.1016/j.str.2004.10.018>
- Tsai, B., Liu, W., Dong, D., Shi, K., Chen, L., Gao, N., 2020. Phase separation of Mer2 organizes the meiotic loop-axis structure of chromatin during meiosis I (preprint). *Biophysics*. <https://doi.org/10.1101/2020.12.15.422856>
- Uchański, T., Masiulis, S., Fischer, B., Kalichuk, V., López-Sánchez, U., Zarkadas, E., Weckener, M., Sente, A., Ward, P., Wohlkönig, A., Zögg, T., Remaut, H., Naismith, J.H., Nury, H., Vranken, W., Aricescu, A.R., Pardon, E., Steyaert, J., 2021. Megabodies expand the nanobody toolkit for protein structure determination by single-particle cryo-EM. *Nat. Methods* 18, 60–68. <https://doi.org/10.1038/s41592-020-01001-6>
- Voelkel-Meiman, K., Cheng, S.-Y., Parziale, M., Morehouse, S.J., Feil, A., Davies, O.R., Muylt, A. de, Borde, V., MacQueen, A.J., 2019. Crossover recombination and synapsis are linked by adjacent regions within the N terminus of the Zip1 synaptonemal complex protein. *PLOS Genet.* 15, e1008201. <https://doi.org/10.1371/journal.pgen.1008201>
- Vrielynck, N., Chambon, A., Vezon, D., Pereira, L., Chelysheva, L., De Muylt, A., Mézard, C., Mayer, C., Grelon, M., 2016. A DNA topoisomerase VI-like complex initiates meiotic recombination. *Science* 351, 939–943. <https://doi.org/10.1126/science.aad5196>
- Vries, F.A.T. de, Boer, E. de, Bosch, M. van den, Baarends, W.M., Ooms, M., Yuan, L., Liu, J.-G., Zeeland, A.A. van, Heyting, C., Pastink, A., 2005. Mouse Sycp1 functions in synaptonemal complex assembly, meiotic recombination, and XY body formation. *Genes Dev.* 19, 1376–1389. <https://doi.org/10.1101/gad.329705>
- Wan, L., Niu, H., Futcher, B., Zhang, C., Shokat, K.M., Boulton, S.J., Hollingsworth, N.M., 2008. Cdc28–C1b5 (CDK-S) and Cdc7–Dbf4 (DDK) collaborate to initiate meiotic recombination in yeast. *Genes Dev.* 22, 386–397. <https://doi.org/10.1101/gad.1626408>
- Wang, S., Zickler, D., Kleckner, N., Zhang, L., 2015. Meiotic crossover patterns: Obligatory crossover, interference and homeostasis in a single process. *Cell Cycle* 14, 305–314. <https://doi.org/10.4161/15384101.2014.991185>
- Wells, D., Bitoun, E., Moralli, D., Zhang, G., Hinch, A., Jankowska, J., Donnelly, P., Green, C., Myers, S.R., 2020. ZCWPW1 is recruited to recombination hotspots by PRDM9 and is essential for meiotic double strand break repair. *eLife* 9, e53392. <https://doi.org/10.7554/eLife.53392>
- Wendorff, T.J., Berger, J.M., 2018. Topoisomerase VI senses and exploits both DNA crossings and bends to facilitate strand passage. *eLife* 7, e31724. <https://doi.org/10.7554/eLife.31724>
- Wojtasz, L., Cloutier, J.M., Baumann, M., Daniel, K., Varga, J., Fu, J., Anastassiadis, K., Stewart, A.F., Reményi, A., Turner, J.M.A., Tóth, A., 2012. Meiotic DNA double-strand breaks and chromosome asynapsis in mice are monitored by distinct HORMAD2-independent and -dependent mechanisms. *Genes Dev.* 26, 958–973. <https://doi.org/10.1101/gad.187559.112>
- Wojtasz, L., Daniel, K., Roig, I., Bolcun-Filas, E., Xu, H., Boonsanay, V., Eckmann, C.R., Cooke, H.J., Jasin, M., Keeney, S., McKay, M.J., Toth, A., 2009. Mouse HORMAD1 and HORMAD2, two conserved meiotic chromosomal proteins, are

- depleted from synapsed chromosome axes with the help of TRIP13 AAA-ATPase. *PLoS Genet.* 5, e1000702. <https://doi.org/10.1371/journal.pgen.1000702>
- Wu, L., Hickson, I.D., 2003. The Bloom's syndrome helicase suppresses crossing over during homologous recombination. *Nature* 426, 870–874. <https://doi.org/10.1038/nature02253>
- Xu, Y., Ashley, T., Brainerd, E.E., Bronson, R.T., Meyn, M.S., Baltimore, D., 1996. Targeted disruption of ATM leads to growth retardation, chromosomal fragmentation during meiosis, immune defects, and thymic lymphoma. *Genes Dev.* 10, 2411–2422. <https://doi.org/10.1101/gad.10.19.2411>
- Yadav, V.K., Claeys Bouuaert, C., 2021. Mechanism and Control of Meiotic DNA Double-Strand Break Formation in *S. cerevisiae*. *Front. Cell Dev. Biol.* 9.
- Yamada, S., Hinch, A.G., Kamido, H., Zhang, Y., Edelman, W., Keeney, S., 2020. Molecular structures and mechanisms of DNA break processing in mouse meiosis. *Genes Dev.* 34, 806–818. <https://doi.org/10.1101/gad.336032.119>
- Zakharyevich, K., Ma, Y., Tang, S., Hwang, P.Y.-H., Boiteux, S., Hunter, N., 2010. Temporally and biochemically distinct activities of Exo1 during meiosis: double-strand-break resection and resolution of double-Holliday Junctions. *Mol. Cell* 40, 1001–1015. <https://doi.org/10.1016/j.molcel.2010.11.032>
- Zakharyevich, K., Tang, S., Ma, Y., Hunter, N., 2012. Delineation of joint molecule resolution pathways in meiosis identifies a crossover-specific resolvase. *Cell* 149, 334–347. <https://doi.org/10.1016/j.cell.2012.03.023>
- Zelensky, A., Kanaar, R., Wyman, C., 2014. Mediators of Homologous DNA Pairing. *Cold Spring Harb. Perspect. Biol.* 6, a016451. <https://doi.org/10.1101/cshperspect.a016451>
- Zhang, B., Tang, Z., Li, L., Lu, L.-Y., 2020. NBS1 is required for SPO11-linked DNA double-strand break repair in male meiosis. *Cell Death Differ.* 27, 2176–2190. <https://doi.org/10.1038/s41418-020-0493-4>
- Zhang, L., Kim, K.P., Kleckner, N.E., Storlazzi, A., 2011. Meiotic double-strand breaks occur once per pair of (sister) chromatids and, via Mec1/ATR and Tel1/ATM, once per quartet of chromatids. *Proc. Natl. Acad. Sci. U. S. A.* 108, 20036–20041. <https://doi.org/10.1073/pnas.1117937108>
- Zhang, Q., Ji, S.-Y., Busayavalasa, K., Yu, C., 2019. SPO16 binds SHOC1 to promote homologous recombination and crossing-over in meiotic prophase I. *Sci. Adv.* 5, eaau9780. <https://doi.org/10.1126/sciadv.aau9780>
- Zhu, B., Mandal, S.S., Pham, A.-D., Zheng, Y., Erdjument-Bromage, H., Batra, S.K., Tempst, P., Reinberg, D., 2005. The human PAF complex coordinates transcription with events downstream of RNA synthesis. *Genes Dev.* 19, 1668–1673. <https://doi.org/10.1101/gad.1292105>
- Zickler, D., Kleckner, N., 2015. Recombination, Pairing, and Synapsis of Homologs during Meiosis. *Cold Spring Harb. Perspect. Biol.* 7. <https://doi.org/10.1101/cshperspect.a016626>
- Zickler, D., Kleckner, N., 1999. Meiotic chromosomes: integrating structure and function. *Annu. Rev. Genet.* 33, 603–754. <https://doi.org/10.1146/annurev.genet.33.1.603>
- Zickler, D., Kleckner, N., 1998. THE LEPTOTENE-ZYGOTENE TRANSITION OF MEIOSIS. *Annu. Rev. Genet.* 32, 619–697. <https://doi.org/10.1146/annurev.genet.32.1.619>

## Contribution aux publications

### **TOPOVIBL-REC114 interaction regulates meiotic DNA double-strand breaks**

Pour cet article j'ai réalisé certaines caractérisations biochimiques avant la première soumission, notamment l'analyse par gel-filtration de l'interaction entre REC114-MEI4 et TOPOVIBL. Ensuite, pendant la révision, j'ai participé à la réalisation de plusieurs expériences de caractérisations biochimiques.

### **Structural analysis of Red1 as a conserved scaffold of the RNA-targeting MTREC/PAXT complex**


En dehors de mon projet de thèse, j'ai contribué au travail de Foucher *et al* (2022) sur l'analyse structurale et fonctionnelle des protéines Red1, Iss10 et Ars2, membres du complexe MTREC/PAXT. Ce dernier est impliqué dans la reconnaissance des ARN non pertinents pour les acheminer vers les machineries de dégradation. J'ai participé à l'attribution des données RMN de la chaîne principale du complexe Red1-Iss10. J'ai également analysé l'interaction entre les mutants de Red1 et Iss10 par *pull down*, j'ai préparé l'échantillon pour l'analyse SEC-MALLS du complexe Red1-Iss10 et j'ai réalisé le clonage des mutants d'Ars2.

# TOPOVIBL-REC114 interaction regulates meiotic DNA double-strand breaks

Received: 27 September 2022

Accepted: 8 November 2022

Published online: 17 November 2022

 Check for updates

Alexandre Nore<sup>1,5</sup>, Ariadna B. Juarez-Martinez<sup>2,5</sup>, Julie Clément<sup>1</sup>, Christine Brun<sup>1</sup>, Boubou Diagouraga<sup>3</sup>, Hamida Laroussi<sup>2</sup>, Corinne Grey<sup>1</sup>, Henri Marc Bourbon<sup>4</sup>, Jan Kadlec<sup>2</sup>✉, Thomas Robert<sup>3</sup>✉ & Bernard de Massy<sup>1</sup>✉

Meiosis requires the formation of programmed DNA double strand breaks (DSBs), essential for fertility and for generating genetic diversity. DSBs are induced by the catalytic activity of the TOPOVIL complex formed by SPO11 and TOPOVIBL. To ensure genomic integrity, DNA cleavage activity is tightly regulated, and several accessory factors (REC114, MEI4, IHO1, and MEI1) are needed for DSB formation in mice. How and when these proteins act is not understood. Here, we show that REC114 is a direct partner of TOPOVIBL, and identify their conserved interacting domains by structural analysis. We then analyse the role of this interaction by monitoring meiotic DSBs in female and male mice carrying point mutations in TOPOVIBL that decrease or disrupt its binding to REC114. In these mutants, DSB activity is strongly reduced genome-wide in oocytes, and only in sub-telomeric regions in spermatocytes. In addition, in mutant spermatocytes, DSB activity is delayed in autosomes. These results suggest that REC114 is a key member of the TOPOVIL catalytic complex, and that the REC114/TOPOVIBL interaction ensures the efficiency and timing of DSB activity.

Sexual reproduction relies on the specialized cell division of meiosis to generate haploid cells that eventually differentiate into gametes. In most taxa, proper segregation of the homologous chromosomes (homologs) depends on homologous recombination, which physically connects homologs through at least one reciprocal exchange (crossover) between each homolog pair<sup>1</sup>. Meiotic recombination initiates at the onset of prophase I by the formation of hundreds of programmed DNA double-strand breaks (DSBs) at preferred DNA sites, named hotspots<sup>2</sup>. DSBs are formed by the collective action of a conserved set of proteins: the TOPOVIL complex and its accessory partners<sup>3–6</sup>.

The TOPOVIL complex is evolutionarily related to the TopoVI type IIB topoisomerase<sup>7</sup>, and is composed of two conserved subunits: the SPO11 catalytic subunit and TOPOVIBL. TOPOVIBL shares partial homology with the GHKL-ATPase domain (thus named GHKL-like),

with the central transducer domain and to a lesser extent, with the regulatory C-terminal domain (CTD) of the TopoVIB subunit of archaeal TopoVI<sup>5,6,8–11</sup>. The TOPOVIL meiosis-specific activity is finely regulated in time (i.e., must be turned on and off at precise time windows) and in space (i.e., active at specific chromosomal locations) during meiotic prophase. However, the molecular mechanisms underlying this complex regulation remain largely elusive.

In *Saccharomyces cerevisiae*, Rec102, the identified TOPOVIBL homolog, shares only partial similarity with TopoVIB, with the conserved transducer domain but not with the GHKL-like domain<sup>5</sup>. Interaction and biochemical studies suggest that Rec104, which is also essential for DSB formation in yeast, could replace the GHKL domain and that the Rec102/Rec104 complex fulfils the function of TOPOVIBL<sup>12–14</sup>. The *S. cerevisiae* core complex defined biochemically is

<sup>1</sup>Institut de Génétique Humaine (IGH), Centre National de la Recherche Scientifique, Univ Montpellier, Montpellier, France. <sup>2</sup>Univ. Grenoble Alpes, CNRS, CEA, IBS, F-38000 Grenoble, France. <sup>3</sup>CBS, Univ Montpellier, CNRS, INSERM, Montpellier, France. <sup>4</sup>Centre de Biologie Intégrative, CNRS, Université de Toulouse, Toulouse, France. <sup>5</sup>These authors contributed equally: Alexandre Nore, Ariadna B. Juarez-Martinez. ✉ e-mail: [jan.kadlec@ibs.fr](mailto:jan.kadlec@ibs.fr); [thomas.robert@cbs.cnrs.fr](mailto:thomas.robert@cbs.cnrs.fr); [bernard.de-massy@igh.cnrs.fr](mailto:bernard.de-massy@igh.cnrs.fr)

composed of Spo11, Rec102, Rec104 and Ski8, a protein directly interacting with Spo11<sup>13</sup>. In addition, to the core complex, several accessory proteins are required for DSB formation. In *S. cerevisiae*, these accessory partners are Rec114, Mei4 and Mer2, which interact with DNA that mediates the formation of condensate-like RMM protein clusters<sup>15</sup>. These condensates might recruit the Spo11/Rec102/Rec104/Ski8 complex to DNA through an interaction of Rec114 with Rec102/Rec104<sup>12,15,16</sup>. The finding that the Rec114 residues involved in this interaction are essential for DSB formation supports this hypothesis<sup>15</sup>.

In *M. musculus*, SPO11 and TOPOVIBL are evolutionarily conserved<sup>5,17,18</sup> and the DSB sites are determined by PRDM9 that recognizes specific DNA motifs and modifies chromatin upon binding to these sites<sup>19–21</sup>. In the mouse, the accessory proteins REC114, MEI4, IHO1 (orthologues of *S. cerevisiae* Rec114, Mei4 and Mer2, respectively), and MEI1 are essential for DSB formation, and localize as foci on chromosome axes at meiotic prophase onset. It has been proposed that IHO1, MEI4 and REC114 directly control TOPOVIL through its recruitment or activation<sup>22–26</sup>. MEI4 and REC114 form a stable complex, and structural analyses revealed that REC114 N-terminus forms a Pleckstrin homology (PH) domain<sup>25,27</sup> harbouring exposed conserved residues potentially involved in protein-protein interactions. This suggests that REC114 acts as a regulatory platform. In line with this hypothesis, it was recently reported that ANKRD31 directly interacts with REC114 PH domain and is involved in regulating DSB number and localization<sup>27,28</sup>. However, how these accessory proteins participate in TOPOVIL activity remains to be determined.

Here, using structural analysis, we found that the mouse REC114 PH domain directly interacts with a conserved C-terminal peptide of TOPOVIBL, identifying its CTD as a predicted regulatory unit of the TOPOVIL complex. Accordingly, in mice where TOPOVIBL interaction with REC114 was disrupted, we observed meiotic DSB formation defects in both sexes associated with reduced fertility. In females, DSB formation was drastically reduced. In males, DSB formation was delayed, but DSB levels were not reduced except for sub-telomeric regions. On the X and Y chromosomes, which recombine specifically in the distal sub-telomeric pseudo-autosomal region (PAR) region, DSB activity reduction led to chromosome synapsis defects. These results suggest that REC114 is part of the TOPOVIL catalytic complex and acts to regulate the level and timing of DSB formation.

## Results

### TOPOVIBL forms a stable complex with REC114

Using yeast two-hybrid assays (Y2H) we identified a specific interaction between mouse TOPOVIBL and REC114. The deletion analysis showed that the N-terminal PH domain of REC114 was required for this interaction because TOPOVIBL did not bind to REC114 lacking the first 39 amino acids (40–259) (Fig. 1a, b and Supplementary Fig. 1a). This interaction also required TOPOVIBL C-terminus as indicated by the absence of binding upon deletion of its last 29 amino acids (TOPOVIBL<sup>1–550</sup>) (Fig. 1a, c and Supplementary Fig. 1b). TOPOVIBL<sup>1–550</sup> could still interact with SPO11 $\beta$ , indicating that the deletion of the 29 amino acids did not disrupt TOPOVIBL folding (Fig. 1c and Supplementary Fig. 1b). We obtained similar results in vitro using Strep-tag pull-down assays (Fig. 1d). While it was difficult to produce full-length TOPOVIBL, we could express and purify a 6xHis-SUMO fusion of TOPOVIBL that lacks part of its transducer domain (construct 1–385; Fig. 1d, lane 2) and also the 6xHis-TOPOVIBL C-terminal domain (CTD; 452–579) (Fig. 1d, lane 3). REC114 clearly interacted with TOPOVIBL<sup>452–579</sup> (Fig. 1d, lane 12), but not with TOPOVIBL<sup>1–385</sup> (Fig. 1d, lane 11). The REC114 N-terminal domain (REC114<sup>1–159</sup>), which includes the PH domain, was sufficient for the interaction with TOPOVIBL<sup>452–579</sup> (Fig. 1d, lane 14). In agreement, the REC114 PH domain and TOPOVIBL<sup>452–579</sup> co-eluted in a single peak during size exclusion chromatography (Fig. 1e and Supplementary Fig. 2a).

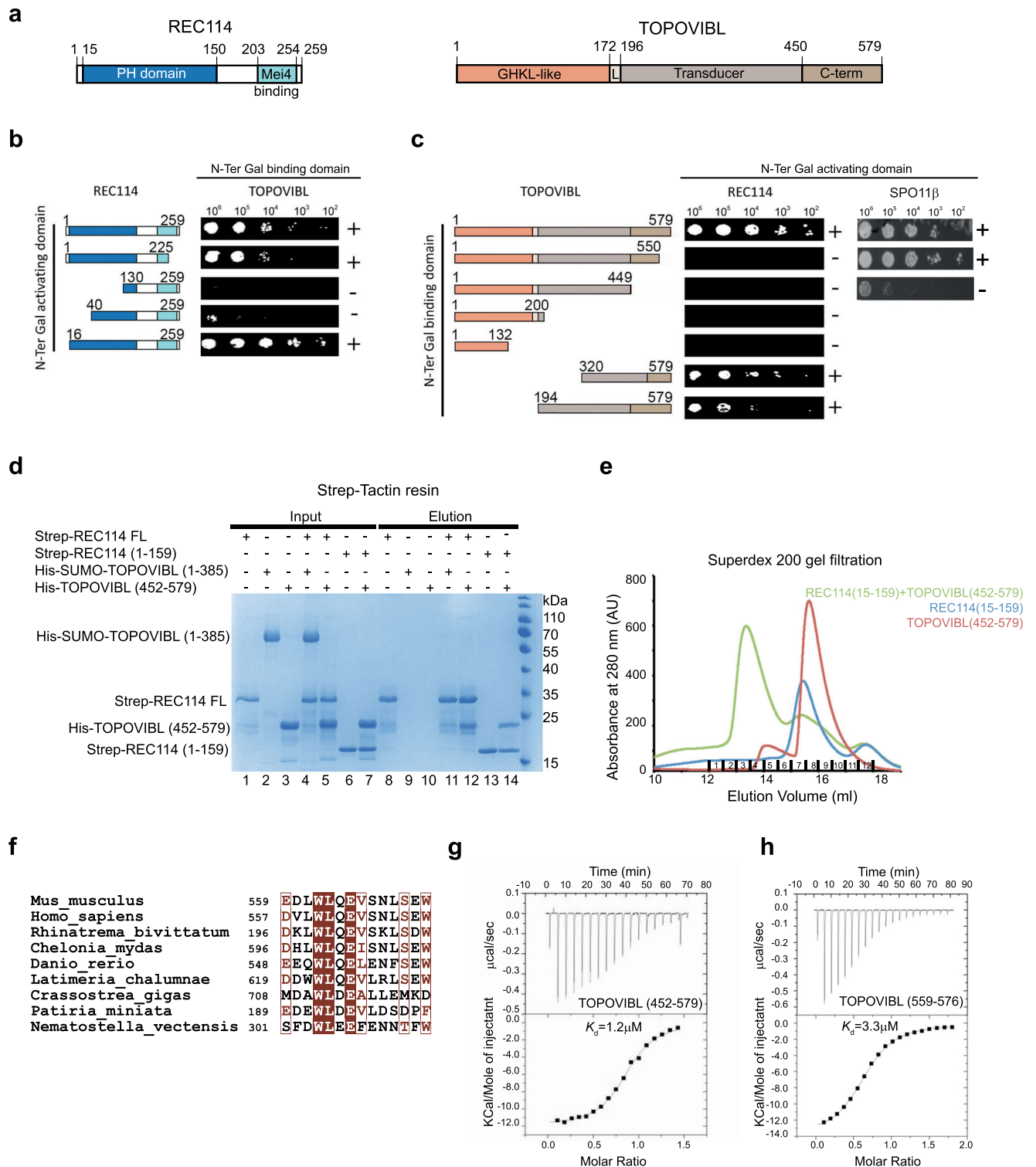
The structure prediction analysis and AlphaFold2 model (AF-J3QMY9) suggested that TOPOVIBL CTD does not contain any known globular domain and was partially intrinsically disordered. However, it indicated the presence of a putative helix at its C-terminus. Interestingly, within the last C-terminal residues, we detected a high conservation among metazoan species (Fig. 1f). An extensive analysis of TOPOVIBL phylogeny indeed identified that the three helices of the transducer domain and this predicted C-terminal helix are the main conserved regions of TOPOVIBL in metazoans<sup>7</sup>. As our yeast two-hybrid assays suggested that the corresponding C-terminal 29 residues of TOPOVIBL (residues 551 to 579) were important for the interaction with REC114 (Fig. 1c), we hypothesized that this predicted conserved helix (residues 559–572, Fig. 1f) might represent the REC114 binding region. Using isothermal titration calorimetry (ITC), we showed that REC114 PH domain bound to TOPOVIBL CTD (452–579) with a dissociation constant ( $K_d$ ) of 1.2  $\mu$ M (Fig. 1g). Moreover, the interaction between a TOPOVIBL peptide spanning residues 559–576 and REC114 PH domain was in the same range ( $K_d$  = 3.3  $\mu$ M) (Fig. 1h). These results demonstrate that this highly conserved motif in TOPOVIBL CTD (559–576) is sufficient for interaction with REC114. Whether residues in the 452–559 region of TOPOVIBL contribute to the interaction with REC114 cannot be excluded, but no conserved motif could be identified in this interval (Supplementary Fig. 3). This question is also addressed by the mutational analysis presented below.

### Crystal structure of the TOPOVIBL-REC114 complex

We then determined the crystal structure of the complex formed by the REC114 PH domain (residues 15–159) and the TOPOVIBL<sup>559–576</sup> peptide by X-ray crystallography (2.3 Å resolution,  $R_{\text{free}}$  of 24.9%, and R-factor of 22.6%) (Supplementary Table 1, Supplementary Fig. 2b). The structure of the TOPOVIBL-bound REC114 PH domain (two perpendicular antiparallel  $\beta$ -sheets followed by a C-terminal helix) was essentially the same as in its unbound form<sup>25</sup>. The TOPOVIBL C-terminal peptide folds into a single helix that interacts with the REC114 PH domain  $\beta$ -sheet formed of strands  $\beta$ 1,  $\beta$ 2 and  $\beta$ 6– $\beta$ 8, burying 731 Å<sup>2</sup> of surface area (Fig. 2a, b). The interaction surface on REC114 is formed of highly conserved surface residues (Fig. 2c and Supplementary Fig. 2c). In the N-terminal part of the TOPOVIBL peptide, L561, W562 and V566 pack against a hydrophobic surface of REC114 (Fig. 2d). Specifically, W562 is located in a hydrophobic pocket formed by aliphatic side chains of K95, V97, L104 and M115 of REC114. In the central part of the TOPOVIBL helix, L569 binds to a hydrophobic groove in REC114 formed by V97, R99, C102, L104, and R117 (Fig. 2e). The well-conserved R99 and R117 residues form several hydrogen bonds with main-chain carbonyls in the TOPOVIBL helix, and R117 forms a salt bridge with E571. The C-terminal part of the TOPOVIBL peptide forms a  $3_{10}$  helix where W572 inserts into another hydrophobic pocket of REC114 made of R24, V53, C102 and R117 (Fig. 2f). W572 also forms cation- $\pi$  interactions with the guanidinium groups of the two arginine residues and main-chain hydrogen bonds with R24 and Q119. Finally, L573 packs against a hydrophobic surface around M100 (Fig. 2f).

Most of the TOPOVIBL residues involved in the interaction with REC114 are highly conserved across metazoan species (Fig. 1f). We showed the interaction conservation by modelling the complex from different vertebrate species using AlphaFold2<sup>29</sup> (Supplementary Fig. 4a). To identify the TOPOVIBL residues required for the interaction with REC114, we mutated several candidate residues (W562, V566, L569 and W572) based on structure and conservation. In pull-down assays with the Strep-tagged REC114 PH domain, mutations W562A, W562E, V566R, L569R and W572L led to undetectable interaction with TOPOVIBL (Fig. 2g, lanes 9–13). We also mutated two hydrophobic residues of the REC114  $\beta$ -sheet (V97D, L104D). Both mutations lead to undetectable interaction with TOPOVIBL (Supplementary Fig. 4b, lanes 9 and 10). The mutations introduced in TOPOVIBL and REC114 did not alter the structure of the proteins as judged by gel filtration





analysis (Supplementary Figs. 4c, d, 5a, b). We confirmed the role of TOPOVIBL W562 in the interaction with REC114 by ITC measurement and showed that both the TOPOVIBL<sup>559-576</sup> peptide and the TOPOVIBL<sup>452-579</sup> purified protein containing the W562A mutation did not bind to REC114 PH domain (Supplementary Fig. 4e, f). This indicates an essential role for W562 in the interaction with the REC114 PH domain and the absence of any significant interaction in the adjacent 452–559 region of TOPOVIBL. Moreover, in yeast two-hybrid assays using full-length proteins we showed that W562A and W562G led to undetectable interaction between TOPOVIBL and REC114 (Fig. 2h and Supplementary Fig. 1c). We conclude that the TOPOVIBL W562A

mutation specifically disrupts the interaction with REC114 *in vitro*, and we selected this mutation for the *in vivo* studies (see below).

### REC114 partners

Mouse REC114 directly interacts also with MEI4<sup>25</sup> and ANKRD31<sup>27</sup>. The REC114/MEI4 interaction does not involve the REC114 PH domain but the C-terminal domain of REC114 (203–254)<sup>25</sup>. In agreement, we could show by size exclusion chromatography the simultaneous interaction of REC114 with MEI4 and TOPOVIBL (Fig. 2i and Supplementary Fig. 5c). Conversely, the REC114/ANKRD31 interaction involves the REC114 PH domain. The crystal structure of the ANKRD31-REC114 complex<sup>27</sup>

**Fig. 1 | The C-terminal region of TOPOVIBL interacts with the N-terminal domain of REC114.** **a** Schematic representation of the domain structure of mouse REC114 (left) and TOPOVIBL (right). PH: Pleckstrin homology domain, MEI4 binding region, GHKL-like: Gyrase, HSP90, Histidine Kinase, MutL domain, L: linker, C-term: C-terminal. **b** Yeast two-hybrid assays showed that REC114 N-terminal domain is required for the interaction with TOPOVIBL. Full length and truncated REC114 proteins were tested for interaction with full length TOPOVIBL. Growth (+ or -) was assayed on medium lacking leucine, tryptophane and histidine and with 5 mM 3-amino-1,2,4-triazole. **c** Yeast two-hybrid assays indicated that TOPOVIBL last 29 residues are required for the interaction with REC114 but not with SPO11 $\beta$ . Full length and truncated TOPOVIBL proteins were tested for interaction with full length REC114 and SPO11 $\beta$ . Growth (+ or -) was assayed on medium lacking leucine, tryptophane and histidine. **d** Pull-down experiments of Strep-tagged REC114 with TOPOVIBL domains. All proteins were first purified by affinity chromatography and

gel filtration. Proteins were mixed as indicated above the lanes. A total of 0.8% of the input (lanes 1–7) and 1.2% of the eluates (lanes 8–14) were analyzed on 12% SDS-PAGE gels stained with coomassie brilliant blue. TOPOVIBL (1-385) is not retained by REC114 (compare lanes 4 and 11), whereas TOPOVIBL C-terminal region (452-579) is sufficient for the interaction with FL-REC114 (compare lanes 5 and 12) as well as REC114<sup>1-159</sup> (compare lanes 7 and 14). Source data are provided as a Source Data file. **e** Overlay of Superdex 200 gel filtration elution profiles of REC114<sup>15-159</sup>, TOPOVIBL<sup>452-579</sup>, and their complex. SDS-PAGE gels with eluted fractions are shown in Supplementary Fig. 2a. **f** Sequence alignment of the 14-aa conserved motif at the TOPOVIBL C-terminus in metazoans. Brown letters: equivalent amino acids; white letters: identical amino acids. **g** ITC measurement of the interaction affinity between REC114<sup>15-159</sup> and TOPOVIBL<sup>452-579</sup>. **h** ITC measurement of the interaction affinity between REC114<sup>15-159</sup> and TOPOVIBL<sup>559-576</sup>.

shows that the ANKRD31 interacting fragment (45 residues) is significantly longer than that of TOPOVIBL and covers a larger surface on REC114 (~1800 Å<sup>2</sup>) packing against both of its  $\beta$ -sheets (Supplementary Fig. 5d, e). The C-terminal part of the ANKRD31 peptide forms two helices and packs against the same surface as TOPOVIBL, and both peptides interact with equivalent REC114 residues (Fig. 2j). This structural comparison suggests mutually exclusive binding of these two proteins to REC114, likely with higher affinity for ANKRD31. Mutation of REC114 L104 reduced the binding to both ANKRD31<sup>27</sup> and to TOPOVIBL (Supplementary Fig. 4b, lane 10). Similarly, the ANKRD31 W1842A mutation disrupted the interaction with REC114<sup>27</sup>, as did the corresponding W562A mutation in TOPOVIBL (Fig. 2g, lane 9). As in our hands SUMO or MBP fusions of ANKRD31<sup>1808-1857</sup> aggregated unless bound to REC114, we could not determine its  $K_d$  for REC114. To test for the mutually exclusive nature of the ANKRD31/TOPOVIBL binding to REC114, we performed ITC measurements of the interaction of REC114 with TOPOVIBL<sup>452-579</sup> in the absence or presence of saturating amounts of ANKRD31<sup>1808-1857</sup>. While TOPOVIBL<sup>452-579</sup> normally interacted with REC114 with a  $K_d$  of 1.2  $\mu$ M, we did not observe any binding when REC114 was pre-saturated with ANKRD31<sup>1808-1857</sup> (Supplementary Fig. 5f). Similarly, in pull-down assays, the interaction of REC114 with TOPOVIBL<sup>452-479</sup> was prevented when REC114 was first bound to ANKRD31<sup>1808-1857</sup> (Fig. 2k, lanes 8-11). These results strongly suggest that ANKRD31 prevents TOPOVIBL binding to the REC114 PH domain.

### In vivo analysis of the TOPOVIBL-REC114 interaction

To evaluate the biological significance of the interaction between the C-terminal region of TOPOVIBL and REC114, we generated mice that express mutant alleles of *Top6bl*, which are predicted to alter the interaction in vivo: i) *Top6bl*<sup>W562A</sup> harbouring the point mutation W562A that disrupt the interaction between TOPOVIBL and REC114 in vitro (Fig. 2); and ii) *Top6bl* <sup>$\Delta$ 17Ct</sup> harbouring a truncation of the C-terminal helix of TOPOVIBL that interacts with REC114 by replacing the last 17 amino acids with 9 unrelated residues as a consequence of a frameshift immediately after W562 (Supplementary Fig. 6a–c). We verified by size exclusion chromatography of the C-terminal domain of TOPOVIBL (452-579) that the truncation of the last 17 amino acids did not significantly change the protein elution profile and thus did not induce major changes in its structure (Supplementary Fig. 6d). We also verified that both mutant proteins were expressed in mouse testes. As the TOPOVIBL signal is weak and only detected after immunoprecipitation, a quantitative assessment of protein levels in vivo is not possible (Supplementary Fig. 6e). In ovaries, *Top6bl* expression was detected by RT-PCR in both mutant mice (Supplementary Fig. 6f). Homozygous mutant mice were viable.

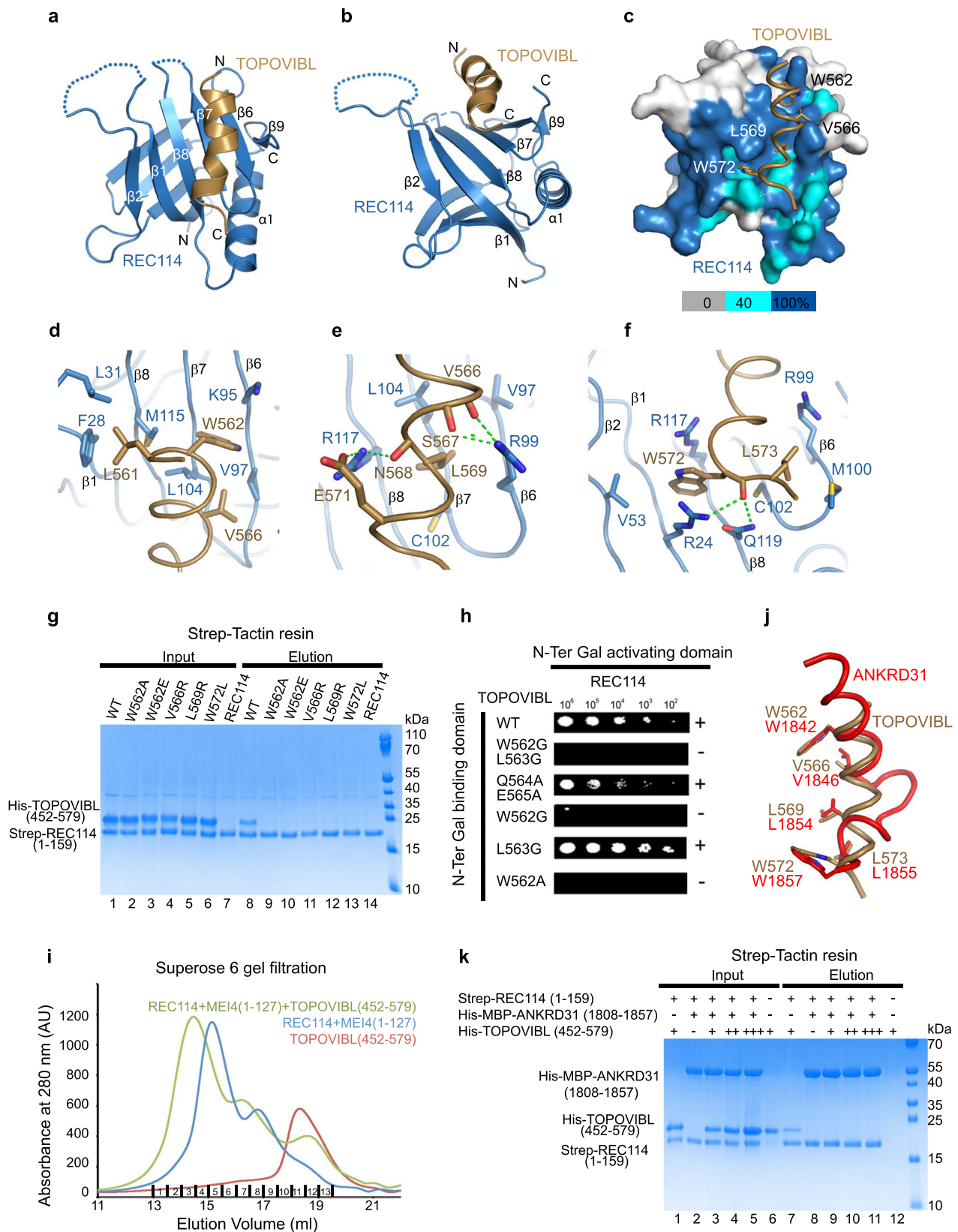
### *Top6bl* mutations lead to a DSB activity reduction in oocytes

We first investigated DSB formation during meiotic prophase of oocytes from embryonic ovaries (16 days post-coitum, dpc), when leptoneuma and zygonema are predominant<sup>30</sup>. To follow DSB

formation, we analysed the phosphorylated form of H2AX ( $\gamma$ H2AX) that appears at chromatin domains around DSB sites upon DSB formation<sup>31</sup>. In wild-type oocytes,  $\gamma$ H2AX was present over large chromatin domains at leptoneuma and zygonema and mostly disappeared at pachynema (Fig. 3a). In both *Top6bl* mutants,  $\gamma$ H2AX intensity was strongly reduced at leptoneuma and zygonema: by 5.2- and 2.5-fold, respectively, in *Top6bl*<sup>W562A/W562A</sup>, and by 8-fold and 7.6-fold, respectively, in *Top6bl* <sup>$\Delta$ 17Ct/ $\Delta$ 17Ct</sup> oocytes, compared with wild-type oocytes (Fig. 3a, b). To determine whether this could be due to delayed DSB formation, we monitored  $\gamma$ H2AX levels at a later developmental stage (18dpc) in *Top6bl* <sup>$\Delta$ 17Ct/ $\Delta$ 17Ct</sup> mice and found a reduction by 7.8-fold (Supplementary Fig. 7a). These findings suggest that the two *Top6bl* mutations lead to a decrease of DSB activity.

We confirmed this hypothesis by quantifying DSB repair through the detection of DMC1 and RPA2. DMC1 binds to resected DSB ends and catalyses homologous strand exchange for DSB repair. RPA is recruited to resected DSB ends before and also after strand exchange during second-end capture for repair<sup>2,32</sup>. In wild-type oocytes, we detected DMC1 and RPA foci that colocalized with the chromosome axis at leptoneuma and zygonema. The number of DMC1 foci was decreased in both *Top6bl* mutants at leptoneuma (3.1-fold in *Top6bl*<sup>W562A/W562A</sup> and 18.5-fold in *Top6bl* <sup>$\Delta$ 17Ct/ $\Delta$ 17Ct</sup> mice) (Fig. 3c, d). DMC1 level in *Top6bl* <sup>$\Delta$ 17Ct/ $\Delta$ 17Ct</sup> mice was close to the background level because DMC1 foci at leptoneuma and zygonema were reduced by 22.9-fold in *Top6bl*<sup>-/-</sup> mice where DSB formation is abolished (Fig. 3d). The reduction of DMC1 foci was also observed at zygonema (Fig. 3d). We obtained similar results for RPA2 foci (Supplementary Fig. 7b, c). In wild-type oocytes, DSB repair promotes interactions between homologues that are stabilized by the recruitment and assembly of several proteins, including SYCP1, to form the synaptonemal complex (Fig. 3a)<sup>33</sup>. In both *Top6bl*<sup>W562A/W562A</sup> and *Top6bl* <sup>$\Delta$ 17Ct/ $\Delta$ 17Ct</sup> mice, we observed only short stretches of synapsis and very few nuclei with full synapses. In 16dpc wild-type ovaries, 42.5%, 32.5% and 20% of oocytes were in leptoneuma, zygonema and pachynema ( $n = 315$ ), respectively, compared with 61.9%, 37.7% and 0.4% ( $n = 496$ ) in *Top6bl*<sup>W562A/W562A</sup>, and 57.6%, 42.4% and 0% ( $n = 151$ ) in *Top6bl* <sup>$\Delta$ 17Ct/ $\Delta$ 17Ct</sup> ovaries. Overall, these results are consistent with a reduced DSB activity that affects synapsis formation between homologues.

As the formation of meiotic DSBs depends on the pre-DSB proteins IH01, REC114, MEI4, ANKRD31 and MEI1<sup>23,25-28,34</sup>, and because the two *Top6bl* mutations disrupt the interaction interface with REC114, we assessed REC114 cytological localization in *Top6bl*<sup>W562A/W562A</sup> and *Top6bl* <sup>$\Delta$ 17Ct/ $\Delta$ 17Ct</sup> mice. In wild-type oocytes, REC114, IH01, MEI4 and ANKRD31 form several hundred foci on chromosome axes at leptoneuma and they progressively disappear as DSBs form<sup>24-28</sup>. The number of REC114 axis-associated foci was significantly higher in *Top6bl*<sup>W562A/W562A</sup> and particularly in *Top6bl* <sup>$\Delta$ 17Ct/ $\Delta$ 17Ct</sup> mice at leptoneuma and especially at zygonema compared with wild-type oocytes (Fig. 3e, f). This higher number of foci can be explained by a reduction of DSB activity. Indeed, it was previously shown that REC114 (and



MEI4) foci are displaced from chromosome axes upon DSB formation in wild-type meocytes, likely as a regulation to turn-off DSB activity, and accumulate in DSB-defective mutant spermatocytes such as *Spo11*<sup>-/-</sup><sup>24,25</sup>. This accumulation of REC114 at zygonema in the absence of meiotic DSB activity also applies to oocytes (Supplementary Fig. 7f)<sup>25</sup>. We obtained similar results for MEI4 and ANKRD31 foci

which were readily detectable on axes at both leptonema and zygonema but with an increased number mostly at zygonema in *Top6bl*<sup>W562A/W562A</sup> and *Top6bl*<sup>Δ17Ct/Δ17Ct</sup> oocytes (Supplementary Figs. 7d, e, 8a, b). Overall, these analyses demonstrate that the loading of REC114, MEI4 and ANKRD31 is not altered by the *Top6bl*<sup>W562A</sup> and *Top6bl*<sup>Δ17Ct</sup> mutations.



**Fig. 2 | Structure of the REC114-TOPOVIBL complex.** **a** Ribbon representation of the overall structure of the REC114-TOPOVIBL complex. The REC114 PH domain is in blue, and the TOPOVIBL peptide in brown. Alpha helices ( $\alpha$ ) and beta sheets ( $\beta$ ) are labelled. **b** REC114-TOPOVIBL structure rotated 60° around the horizontal axis, compared to panel **a**. **c** Surface representation of REC114 to highlight the conserved surface residues. Sequence conservation is represented from grey to blue according to the colour scale bar below. The sequence alignment used for conservation calculation is shown in Supplementary Fig. 2c. TOPOVIBL is shown as a cartoon and the key interacting residues as sticks. **d** Details of the interaction between the N-terminal part of the TOPOVIBL helix (brown) and REC114 (blue). TOPOVIBL W562 inserts into a hydrophobic pocket on the  $\beta$ -sheet formed by strands  $\beta 1$  and  $\beta 6$ - $\beta 8$ . **e** The central part the TOPOVIBL helix (brown) forms several hydrogen bonds (green dotted lines) with conserved REC114 residues. L569 interacts with another hydrophobic cavity formed by  $\beta 6$ - $\beta 8$ . **f** The C-terminal W572 residue of TOPOVIBL forms hydrophobic and charged interactions with REC114. **g** Pull-down experiments of Strep-tagged REC114<sup>1-159</sup> with TOPOVIBL<sup>452-579</sup> mutants

indicated above the lanes. **h** Essential role of TOPOVIBL W562 in the interaction with REC114 shown by yeast two-hybrid assays. Full length TOPOVIBL (WT or with the indicated point mutations) was tested for interaction with full length REC114. Growth (+ or -) was assayed on medium lacking leucine, tryptophane and histidine and with 5 mM 3-amino-1,2,4-triazole. **i** Overlay of Superose 6 gel filtration elution profiles of Strep-REC114-ME14<sup>1-127</sup>, TOPOVIBL<sup>452-579</sup> and their complex. SDS-PAGE gels with eluted fractions are shown in Supplementary Fig. 5c. **j** The key TOPOVIBL and ANKRD31 residues that interact with the REC114  $\beta$ -sheet ( $\beta 1$ ,  $\beta 2$ ,  $\beta 6$ - $\beta 8$ ) are in similar positions. ANKRD31 structure PDB code- 6NXF. **k** ANKRD31 prevents TOPOVIBL binding to REC114. His-TOPOVIBL<sup>452-579</sup>, Strep-REC114<sup>1-159</sup> and Strep-REC114<sup>1-159</sup> bound to His-MBP-ANKRD31<sup>1808-1857</sup> were first purified by affinity chromatography and gel filtration. Proteins were mixed as indicated above the lanes. While TOPOVIBL<sup>452-579</sup> co-purifies with REC114<sup>1-159</sup> (lane 7), it does not co-purify even when used at two (++) or four (+++) times higher concentrations when REC114<sup>1-159</sup> is previously bound to His-MBP-ANKRD31 (lanes 10,11). Source data are provided as a Source Data file.

These meiotic prophase defects affected oogenesis: follicle number was strongly reduced in ovaries of *Top6bl*<sup>W562A/W562A</sup> and particularly *Top6bl* <sup>$\Delta 17Ct/\Delta 17Ct$</sup>  mice (Supplementary Fig. 9a, b, d). Indeed, *Top6bl* <sup>$\Delta 17Ct/\Delta 17Ct$</sup>  female mice were sterile, whereas *Top6bl*<sup>W562A/W562A</sup> female mice were sub-fertile (Supplementary Fig. 9c). We conclude that the TOPOVIBL-REC114 interaction is important for the normal level of meiotic DSB formation in oocytes, and that the truncation of the last 17 amino acids of TOPOVIBL leads to a stronger decrease of DSB activity as compared to the W562A mutation.

### In *Top6bl* spermatocytes, DSBs are delayed genome-wide and decreased in sub-telomeric regions

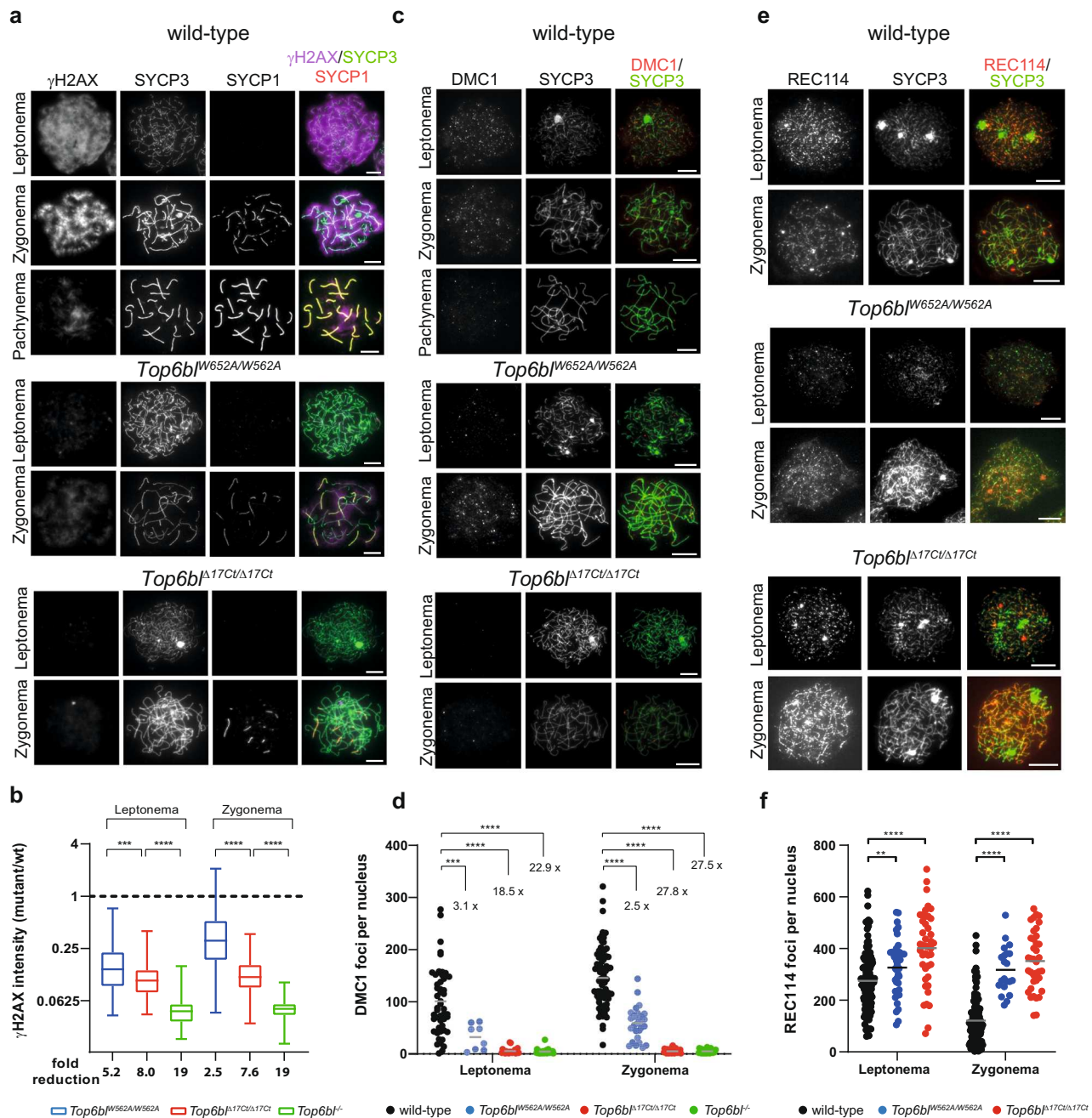
Unlike female meiosis, DSB activity was efficient in spermatocytes from both *Top6bl* mutants, as indicated by the detection and quantification of  $\gamma$ H2AX at late leptotema and zygonema (Fig. 4a–c). Conversely, at early/mid leptotema,  $\gamma$ H2AX levels were lower in *Top6bl*<sup>W562A/W562A</sup> and *Top6bl* <sup>$\Delta 17Ct/\Delta 17Ct$</sup>  than wild-type spermatocytes, suggesting a delay in DSB formation (Fig. 4c).

DMC1 and RPA2 foci also appeared later in both mutants compared with wild-type spermatocytes (Fig. 4d–f and Supplementary Fig. 10b–d). The number of DMC1 and RPA2 foci was lower in *Top6bl*<sup>W562A/W562A</sup> and *Top6bl* <sup>$\Delta 17Ct/\Delta 17Ct$</sup>  than in wild-type spermatocytes, particularly at early/mid leptotema (8.7- and 3.7-fold reduction of DMC1 foci and 11.7- and 7.4-fold reduction of RPA2 foci in *Top6bl*<sup>W562A/W562A</sup> and *Top6bl* <sup>$\Delta 17Ct/\Delta 17Ct$</sup> , respectively, compared with wild-type spermatocytes). Conversely, at zygonema and pachynema, the number of DMC1 and RPA2 foci was similar in wild-type and mutant spermatocytes (Fig. 4e and Supplementary Fig. 10b). These findings suggest an efficient but delayed formation of DMC1 and RPA2 foci, and efficient DSB repair in *Top6bl*<sup>W562A/W562A</sup> and *Top6bl* <sup>$\Delta 17Ct/\Delta 17Ct$</sup>  spermatocytes.

The localization of the DSB axis-associated proteins REC114, ANKRD31 and ME14 at early/mid leptotema and the number of foci were similar or higher than in wild-type spermatocytes (Fig. 5a, b, d and Supplementary Fig. 11a–c). The number of foci gradually decreased from early/mid leptotema in wild-type spermatocytes, but only after late leptotema (ANKRD31) or after early/mid leptotema (REC114) and with a slower kinetic in the *Top6bl* mutants (Fig. 5c, e). These kinetic alterations are compatible with the observed delayed DSB activity because these axis-associated proteins disassemble from the axis upon DSB formation<sup>25,27,28</sup>.

To directly evaluate DSB activity and to map DSB sites, we monitored DMC1 enrichment by DMC1 chromatin-immunoprecipitation (ChIP), followed by ssDNA enrichment (DMC1-Single Strand DNA Sequencing, SSSS)<sup>35</sup> in *Top6bl* <sup>$\Delta 17Ct/\Delta 17Ct$</sup>  mice. We identified 16780 DSB hotspots in *Top6bl* <sup>$\Delta 17Ct/\Delta 17Ct$</sup> , among which 13261 (79%) overlapped with wild-type hotspots (Fig. 6a). This indicated efficient PRDM9-dependent DSB localization in the mutant, as confirmed also by the

absence of significant signal at PRDM9-independent hotspots<sup>36</sup> (Supplementary Fig. 12). The hotspot intensity was similar in *Top6bl* <sup>$\Delta 17Ct/\Delta 17Ct$</sup>  and wild-type samples for most hotspots (Fig. 6b). 3519 peaks were however specific to *Top6bl* <sup>$\Delta 17Ct/\Delta 17Ct$</sup>  mice (Fig. 6a). The average intensity of the DMC1-SSDS signal of the 3519 *Top6bl* <sup>$\Delta 17Ct/\Delta 17Ct$</sup> -specific peaks was weak (Supplementary Fig. 13a upper panel). But interestingly these regions also show some DSB activity in wild-type, although with a 1.7-fold lower average DMC1-SSDS signal as compared to *Top6bl* <sup>$\Delta 17Ct/\Delta 17Ct$</sup>  mice (Supplementary Fig. 13a upper panel). At these sites, the H3K4me3 enrichment detected in wild-type mice of the same *Prdm9* genotype (*Prdm9*<sup>Dom2</sup> from B6 strain) but not from mice expressing the *Prdm9*<sup>Cst</sup> variant (Rj2 strain) confirms that they are *Prdm9*-dependent hotspots (Supplementary Fig. 13a lower panel). This indicates that these 3519 peaks are not new hotspots specific to the *Top6bl* <sup>$\Delta 17Ct/\Delta 17Ct$</sup>  mutant context but that their detection in this mutant is due to the increased DMC1-SSDS signal. We asked if we could also detect any increase of DMC1-SSDS signal in *Top6bl* <sup>$\Delta 17Ct/\Delta 17Ct$</sup>  among the subset of the weakest hotspots identified in wild-type, but this was not the case (bin 1 from Supplementary Fig. 13b). It is therefore possible that these 3519 sites have some specific feature leading to their differential activity in *Top6bl* <sup>$\Delta 17Ct/\Delta 17Ct$</sup>  vs wild-type. A differential analysis of DMC1-SSDS signal intensity at common hotspots (using the DESeq2 R package, see Methods) showed that in *Top6bl* <sup>$\Delta 17Ct/\Delta 17Ct$</sup>  mice, most of the hotspots (86%) were of similar intensity compared to wild-type, and that the signal was decreased (by 1.2- to 24-fold) in 12% of hotspots, and increased (by 1.45- to 3.3-fold) in 2% of hotspots (green and red dots respectively, Fig. 6c). Visual analysis highlighted that in *Top6bl* <sup>$\Delta 17Ct/\Delta 17Ct$</sup>  mice, the DMC1-SSDS signal was decreased at hotspots located near the q-arm telomeres, the telomeres distal to centromeres (Fig. 6d and Supplementary Figs. 14, 15). We quantified this sub-telomeric phenotype by different approaches. First, by quantifying the ratio of the mutant/wild-type signal along the chromosome arms, we observed a  $\geq 2$ -fold decrease in the last few megabases proximal to the q-arm telomeres (Fig. 6e). Second, in the ten hotspots closest to the q-arm telomeres, located within about 4 Mb of the telomere, DMC1-SSDS signal intensity at most hotspots for each chromosome was lower in *Top6bl* <sup>$\Delta 17Ct/\Delta 17Ct$</sup>  than wild-type samples (blue for autosomes, red for the X chromosome) (Fig. 6b). Third, quantification of this effect at each chromosome showed a reduced DMC1-SSDS signal within the ten q-arm telomeric adjacent hotspots in *Top6bl* <sup>$\Delta 17Ct/\Delta 17Ct$</sup>  mice at most chromosomes (Fig. 6f and Supplementary Fig. 13c). Fourth, using DESeq, we evaluated the effect of the distance from telomeres, and found that the effect was most pronounced at hotspots located within 3 Mb from telomeres and decreased when we tested larger intervals (Fig. 6g and Supplementary Fig. 13d). In this analysis, the statistical significance also depended on the hotspot number in the tested region for each chromosome (see, for instance, chromosome 17, Supplementary Fig. 13e). We conclude that in *Top6bl* <sup>$\Delta 17Ct/\Delta 17Ct$</sup>  mice, the

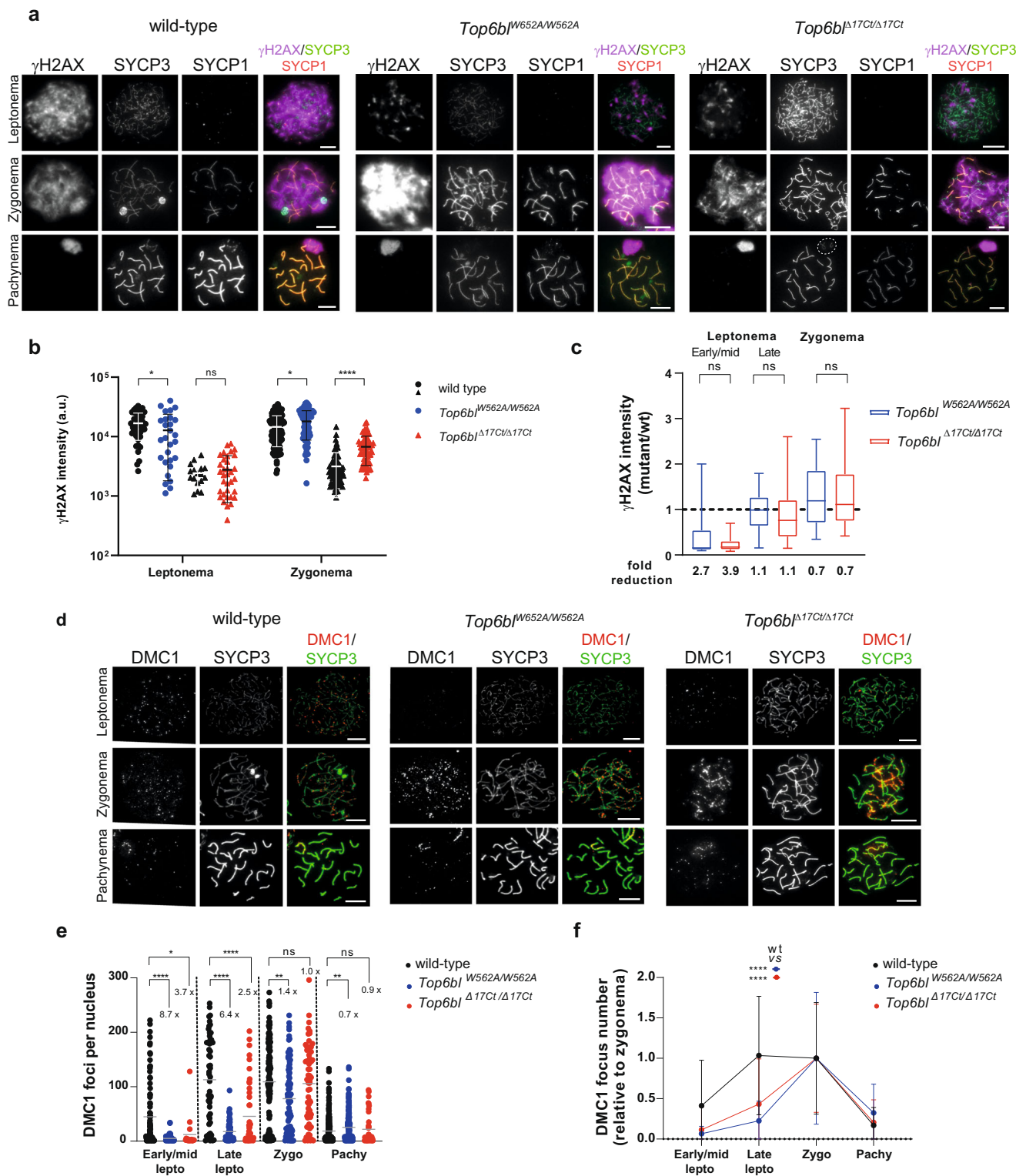


**Fig. 3 | In *Top6bl*<sup>W562A/W562A</sup> and *Top6bl*<sup>Δ17Ct/Δ17Ct</sup> mice, meiotic DSB activity is decreased in oocytes. **a** Immunostaining of γH2AX, SYCP3 and SYCP1 in oocytes from 16 dpc wild-type (+/+), *Top6bl*<sup>W562A/W562A</sup>, and *Top6bl*<sup>Δ17Ct/Δ17Ct</sup> ovaries. Scale bar, 10 μm. **b** Quantification of γH2AX signal intensity in leptotene and zygotene (or zygotene-like) nuclei of oocytes from 16 dpc wild-type (+/+ or *Top6bl*<sup>Δ17Ct/Δ17Ct</sup>), *Top6bl*<sup>W562A/W562A</sup>, *Top6bl*<sup>Δ17Ct/Δ17Ct</sup>, and *Top6bl*<sup>-/-</sup> mice ( $n = 5, 2, 2,$  and  $1$  mouse/genotype). Number of nuclei at leptonema: 126, 123, 73, and 23; number of nuclei at zygonema: 241, 132, 64, and 33 for each genotype. Ratios of the integrated intensity between the mean values in mutant nuclei and in wild-type nuclei are plotted (box plot as defined in Methods).  $P$  values were determined using the two-tailed unpaired Mann-Whitney test. The fold reduction is the ratio of the wild-type to mutant mean values. **c** Immunostaining of DMC1 and SYCP3 in oocytes from 16 dpc or 17 dpc wild-type (+/+), *Top6bl*<sup>W562A/W562A</sup> and *Top6bl*<sup>Δ17Ct/Δ17Ct</sup> ovaries. Scale bar, 10 μm. **d** Quantification of DMC1 foci. DMC1 axis-associated foci were counted in**

leptotene and zygotene nuclei of oocytes from 16 and 17 dpc wild-type (+/+ or *Top6bl*<sup>W562A/W562A</sup>), *Top6bl*<sup>W562A/W562A</sup>, *Top6bl*<sup>Δ17Ct/Δ17Ct</sup>, and *Top6bl*<sup>-/-</sup> mice. Number of nuclei at leptonema: 50, 8, 22, and 28; number of nuclei at zygonema: 79, 28, 29, and 27 for each genotype, respectively. Grey bars show the mean values.  $P$  values were determined using the two-tailed unpaired Mann-Whitney test. The fold difference relative to wild-type is shown. **e** Immunostaining of REC114 and SYCP3 in oocytes from 16dpc wild-type (+/+), *Top6bl*<sup>W562A/W562A</sup>, and *Top6bl*<sup>Δ17Ct/Δ17Ct</sup> ovaries. Scale bar, 10 μm. **f** Quantification of axis-associated REC114 foci in leptotene and zygotene oocytes from 15 dpc wild-type (+/+), *Top6bl*<sup>W562A/W562A</sup>, and *Top6bl*<sup>Δ17Ct/Δ17Ct</sup> mice ( $n = 1$  mouse/genotype). Number of nuclei at leptonema: 51 and 43; number of nuclei at zygonema: 58 and 39 in wild-type and *Top6bl*<sup>Δ17Ct/Δ17Ct</sup> oocytes, respectively. Grey bars show the mean values.  $P$  values were determined using the two-tailed unpaired Mann-Whitney test. Source data are provided as a Source Data file.

DMC1-SSDS signal is specifically reduced in the 3 Mb sub-telomeric region of the q-arm of most chromosomes. This finding could indicate a DSB decrease or a higher DMC1 turnover. This region-specific alteration in *Top6bl*<sup>Δ17Ct/Δ17Ct</sup> mice may not affect homologous

interactions between autosomes because these interactions should be ensured by DSB sites along the chromosome arms which level of activity is overall similar in *Top6bl*<sup>Δ17Ct/Δ17Ct</sup> and wild-type mice. However, on the X and Y chromosomes, which depend on recombination in



the PAR, and which is located proximal to the telomere, the decreased DSB activity in this region in *Top6bl*<sup>Δ17Ct/Δ17Ct</sup> mice (Fig. 6d) could influence their homologous interaction.

We tested this possibility by monitoring synapsis and bivalent formation on autosomes and on the X and Y chromosomes. Synapsis formation was normal on autosomes at pachynema in both mutants (Supplementary Table 4). However, the X and Y chromosomes were frequently unsynapsed in pachytene nuclei (41% of *Top6bl*<sup>W562A/W562A</sup>, 72% of *Top6bl*<sup>Δ17Ct/Δ17Ct</sup> vs 18% in wild-type) (Fig. 7a). At metaphase, an increased frequency of nuclei with 21 DAPI staining bodies (Fig. 7b, d)

and a decreased number of XY bivalent (Fig. 7c, e) was detected in both *Top6bl* mutants. We conclude that the decreased DMC1 signal on the X chromosome sub-telomeric region is compatible with decreased DSB activity rather than with rapid DMC1 turnover. We propose that the TOPOVIBL-REC114 interaction is specifically required at sub-telomeric regions for full DSB activity and therefore, is essential for X/Y chromosome synapsis and segregation.

These molecular and cytological phenotypes should have specific consequences on meiotic prophase and downstream events, because unsynapsed X/Y chromosomes induce metaphase I arrest<sup>37,38</sup>. Indeed,



**Fig. 4 | Efficient but delayed DSB formation in *Top6bl* mutant spermatocytes.** **a** Immunostaining of  $\gamma$ H2AX, SYCP3 and SYCP1 in spermatocytes from 12 dpp wild-type (+/+), *Top6bl*<sup>W562A/W562A</sup>, and *Top6bl* <sup>$\Delta$ 17C $\Delta$ 17C $\Delta$</sup>  mice. A white dotted circle highlights the unsynapsed X and Y chromosomes at pachynema (SYCP3 staining) in *Top6bl* <sup>$\Delta$ 17C $\Delta$ 17C $\Delta$</sup>  mice. Scale bar, 10  $\mu$ m. **b** Quantification of  $\gamma$ H2AX intensity (mean  $\pm$  SD; a.u., arbitrary units) in leptotene and zygotene spermatocytes from 12 dpp wild-type, *Top6bl*<sup>W562A/W562A</sup>, and *Top6bl* <sup>$\Delta$ 17C $\Delta$ 17C $\Delta$</sup>  mice ( $n = 1$  mouse per genotype). Each mutant had a wild-type control tested in parallel (labelled with black spheres or triangles). Number of nuclei: 44, 28, 17, and 34 at leptotene; 132, 61, 77, and 66 at zygonema.  $P$  values were determined using the two-tailed unpaired Mann-Whitney test. **c** Quantification of  $\gamma$ H2AX intensity in *Top6bl*<sup>W562A/W562A</sup> and *Top6bl* <sup>$\Delta$ 17C $\Delta$ 17C $\Delta$</sup>  spermatocytes relative to wild-type at early/mid, late leptotene, and zygotene (box plot as defined in Methods). Stages were defined as described in Methods (see Supplementary Fig. 10a). Number of nuclei: 19, 201 at early/mid, 33, 290 at late leptotene, and 88, 333 at zygotene in *Top6bl*<sup>W562A/W562A</sup> and *Top6bl* <sup>$\Delta$ 17C $\Delta$ 17C $\Delta$</sup>  mice, respectively. The fold reduction is the ratio of the mean values in wild-type and mutant samples.  $P$  values were determined using the two-tailed unpaired Mann-

Whitney test. **d** Immunostaining of DMCI and SYCP3 in spermatocytes from 14 dpp wild-type (+/+), *Top6bl*<sup>W562A/W562A</sup>, and *Top6bl* <sup>$\Delta$ 17C $\Delta$ 17C $\Delta$</sup>  mice. Scale bar, 10  $\mu$ m.

**e** Quantification of DMCI foci. Axis-associated DMCI foci were counted in leptotene (early/mid and late), zygotene, and pachytene nuclei of spermatocytes from wild-type (+/+), *Top6bl*<sup>W562A/W562A</sup>, and *Top6bl* <sup>$\Delta$ 17C $\Delta$ 17C $\Delta$</sup>  mice ( $n = 3$  wild-type, and  $n = 2$  mice per mutant genotype). Number of nuclei: 95, 50 and 18 at early/mid leptotene, 55, 63 and 45 at late leptotene, 120, 92 and 66 at zygotene, and 261, 207 and 36 at pachytene for each genotype, respectively. Grey bars show the mean values.  $P$  values were determined using the two-tailed unpaired Mann-Whitney test.

**f** Variation of DMCI focus number during prophase. For each genotype (wild-type, *Top6bl*<sup>W562A/W562A</sup>, and *Top6bl* <sup>$\Delta$ 17C $\Delta$ 17C $\Delta$</sup> ), the number of DMCI foci at the indicated stages (dataset as in panel e) was normalized to the mean number at zygonema (set to 1). Mean values  $\pm$  SD are shown. Statistical significance between wild-type and each mutant (blue *Top6bl*<sup>W562A/W562A</sup>; red *Top6bl* <sup>$\Delta$ 17C $\Delta$ 17C $\Delta$</sup> ) was tested at late leptotene using the two-tailed unpaired Mann-Whitney test. Source data are provided as a Source Data file.

*Top6bl*<sup>W562A/W562A</sup> and *Top6bl* <sup>$\Delta$ 17C $\Delta$ 17C $\Delta$</sup>  spermatocytes proceed through prophase like wild-type spermatocytes, but they arrested in metaphase and many cells were apoptotic, particularly in the *Top6bl* <sup>$\Delta$ 17C $\Delta$</sup>  mutant. This metaphase I arrest was correlated with reduced sperm production in both mutants and fertility loss in *Top6bl* <sup>$\Delta$ 17C $\Delta$</sup>  mice (Supplementary Fig. 9e–k).

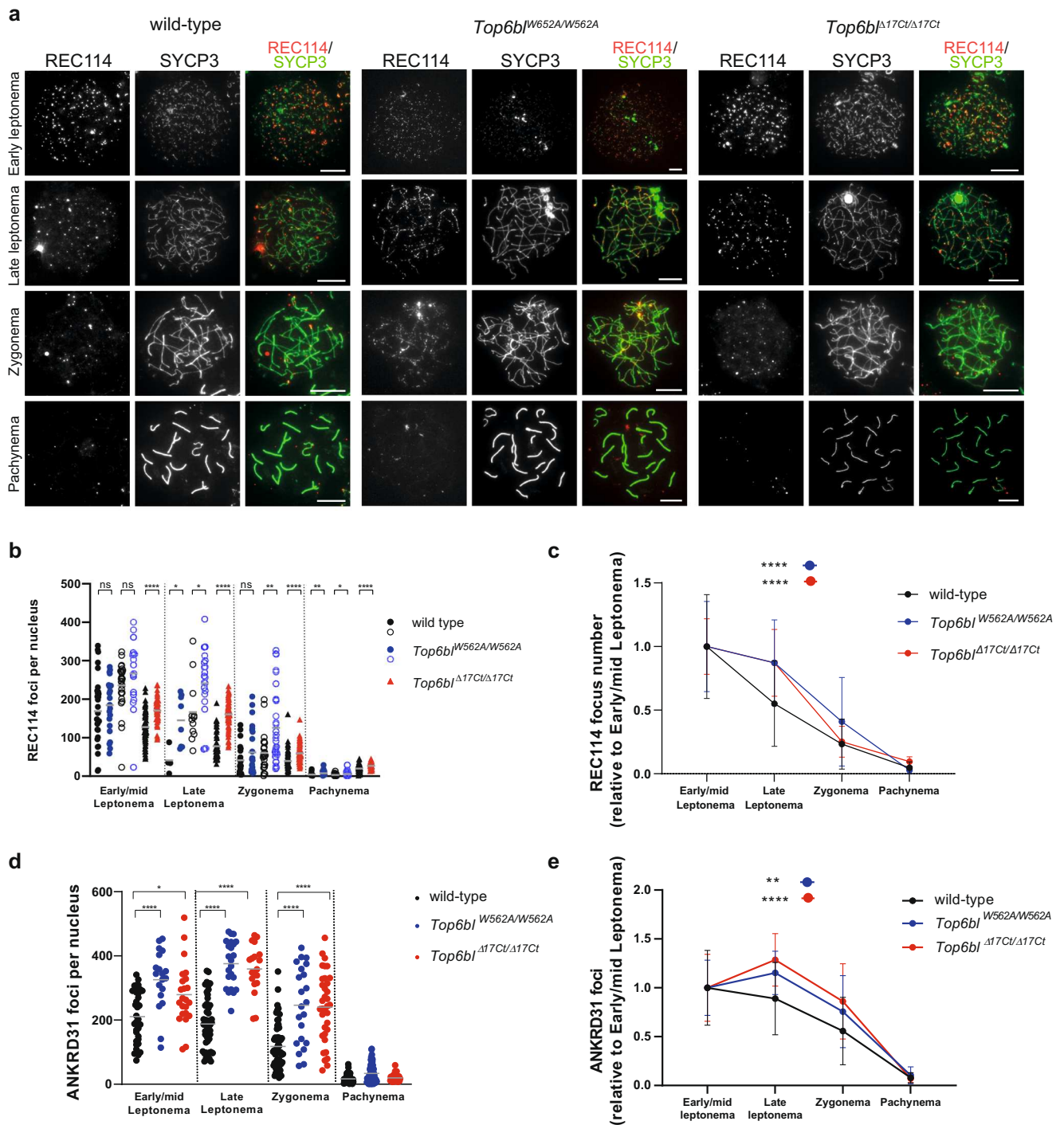
## Discussion

A central question after the identification of the axis-associated proteins essential for meiotic DSB formation is to understand their function. Here, we found that TOPOVIBL CTD directly interacts with the REC114 PH domain, and identified residues required for this interaction *in vitro*.

We propose that TOPOVIBL can bind simultaneously to SPO11 through its transducer domain and to REC114 through the CTD. Both mouse TOPOVIBL and SPO11 structures have been modelled by the AlphaFold2 structure prediction program<sup>29</sup> with high predicted accuracy (AlphaFold2 Protein Structure Database code- AF-J3QMY9 and AF-Q9WTK8). We further used AlphaFold2<sup>29</sup> to model the SPO11/TOPOVIBL complex and combined this model with the crystal structure of the REC114-PH/TOPOVIBL<sup>559-576</sup> complex determined in this study (Fig. 8a and Supplementary Fig. 16d, e). The model gave very high or confident pLDDT scores for most residues in the GHKL-like and transducer domains of TOPOVIBL as well as in 5Y-CAP and Toprim domains of SPO11 (Supplementary Fig. 16a, b, e, g–l, respectively). The predicted aligned errors are low except for the relative positions of the transducer domains of each monomer (Supplementary Fig. 16c, f). Compared to the TopoVIB structure from archaea<sup>39</sup>, the putative mouse TOPOVIBL ATP binding site is predicted to be degenerated with the lack of the ATP-lid structure and its replacement by a helical insertion of the transducer domain (Supplementary Fig. 8b). Indeed, TOPOVIBL contains mutations within the conserved Bergerat fold motifs<sup>5,7</sup> essential for ATP binding and hydrolysis<sup>40</sup>. TOPOVIBL is thus predicted not to bind ATP. TOPOVIBL also lacks both the N-terminal “strap” region involved in dimerization and the H2TH domain. Thus, the ATP mediated dimerization, observed for archaeal TopoVIB, is unlikely to occur in the case of TOPOVIBL. SPO11 however can be modelled as homodimer and its catalytic site, formed by the two protomers is predicted to be equivalent to the one described for TopoVIA<sup>41</sup> (Fig. 8d, e and Supplementary Fig. 16h, i). Finally, the modelled SPO11-TOPOVIBL interface involving the three conserved helices from TOPOVIBL transducer domain and part of the 5Y-CAP domain of SPO11 starting with the second helix (44–63), resembles the archaeal TopoVI complex but with an additional helix of the transducer domain (Fig. 8c and Supplementary Fig. 16g). The REC114-TOPOVIBL complex, as characterized in this study, is connected by the flexible region of the TOPOVIBL CTD to the SPO11/TOPOVIBL core (Fig. 8a).

Although the exact REC114 effect on SPO11/TOPOVIBL organization/conformation is unknown, REC114 binding to TOPOVIBL could stabilize the SPO11/TOPOVIBL complex dimerization. This hypothesis is supported by the stoichiometry (2:1) of the Rec114/Mei4 complex in *S. cerevisiae*<sup>15</sup> and by a predicted heterotrimer from *C. elegans* DSB-1/DSB-2/DSB-3 complex<sup>42</sup>, where DSB-1 and –2 are predicted REC114 orthologs<sup>43,44</sup> and DSB-3, the MEI4 ortholog<sup>45</sup>. Our observation that REC114 could bind simultaneously to TOPOVIBL and MEI4 (Fig. 2j) is compatible with a model where MEI4/REC114 would act as clamp to hold the TOPOVIBL/SPO11 complex. Nevertheless, in mice, REC114 interactions are certainly more complex because its PH domain also interacts with ANKRD31<sup>27</sup>, thus likely competing with TOPOVIBL (Fig. 2k and Supplementary Fig. 5f). Moreover, REC114 interacts with IHO1 in yeast-two hybrid assays<sup>26</sup>. The interplay between these different interactions remains to be determined and disrupting the interaction between TOPOVIBL and REC114 may have additional consequences on other interactions.

The interaction between the C-terminal motif of TOPOVIBL and the PH domain of REC114 is predicted to be conserved in many metazoans based on the presence of the motif and the PH domain (Fig. 1<sup>27,23,25,27</sup>). TOPOVIBL is found outside from metazoans but is highly divergent and no eukaryotic consensus motif can be identified at its C-terminal end<sup>7</sup>. With the contribution of genetics, both TOPOVIBL and REC114 homologs have also been identified in fungi and green plants<sup>5,23</sup>. In *S. cerevisiae*, TOPOVIBL appears to be split into two proteins Rec102 and Rec104. Phylogenetic and biochemical analysis identified Rec102 as homologous to the transducer domain of TOPOVIBL, while Rec104 is predicted to be at the position of the GHKL domain<sup>5,15</sup>. *S. cerevisiae* Rec114 has the conserved PH domain interacting with both, Rec102 and Rec104 in yeast-two hybrid assays and the REC114 residues involved in the interaction are essential for DSB formation<sup>15,16</sup>. According to the AlphaFold2 model of *S. cerevisiae* Rec114 (AF-A0A6V8S448), many of the residues of the mouse REC114 PH domain required for the interaction with mouse TOPOVIBL are not conserved in *S. cerevisiae*. These include mouse REC114 V97 and L104, shown in this study as essential for binding to TOPOVIBL (Supplementary Fig. 4b). In addition, the hydrophobic pocket of REC114 accommodating TOPOVIBL W562 (Fig. 2d) seems absent in *S. cerevisiae* Rec114. Thus, we conclude that a potentially equivalent regulation of DSB activity could be mediated by *S. cerevisiae* Rec114 but with distinct molecular interactions. The regulatory function of *S. cerevisiae* Rec114 for DSB activity is also illustrated by the Mec1/Tel1 (ATR/ATM) dependent phosphorylation of Rec114 and its consequence on down-regulating DSB activity<sup>46</sup>, a modification and regulation also observed for *C. elegans* DSB-1<sup>42</sup>. In *A. thaliana*, both the TOPOVIBL ortholog, MTPOVIB<sup>9</sup>, and REC114<sup>23</sup> have been identified. Although the PH domain of REC114 is conserved<sup>27</sup>, the metazoan C-terminal motif is

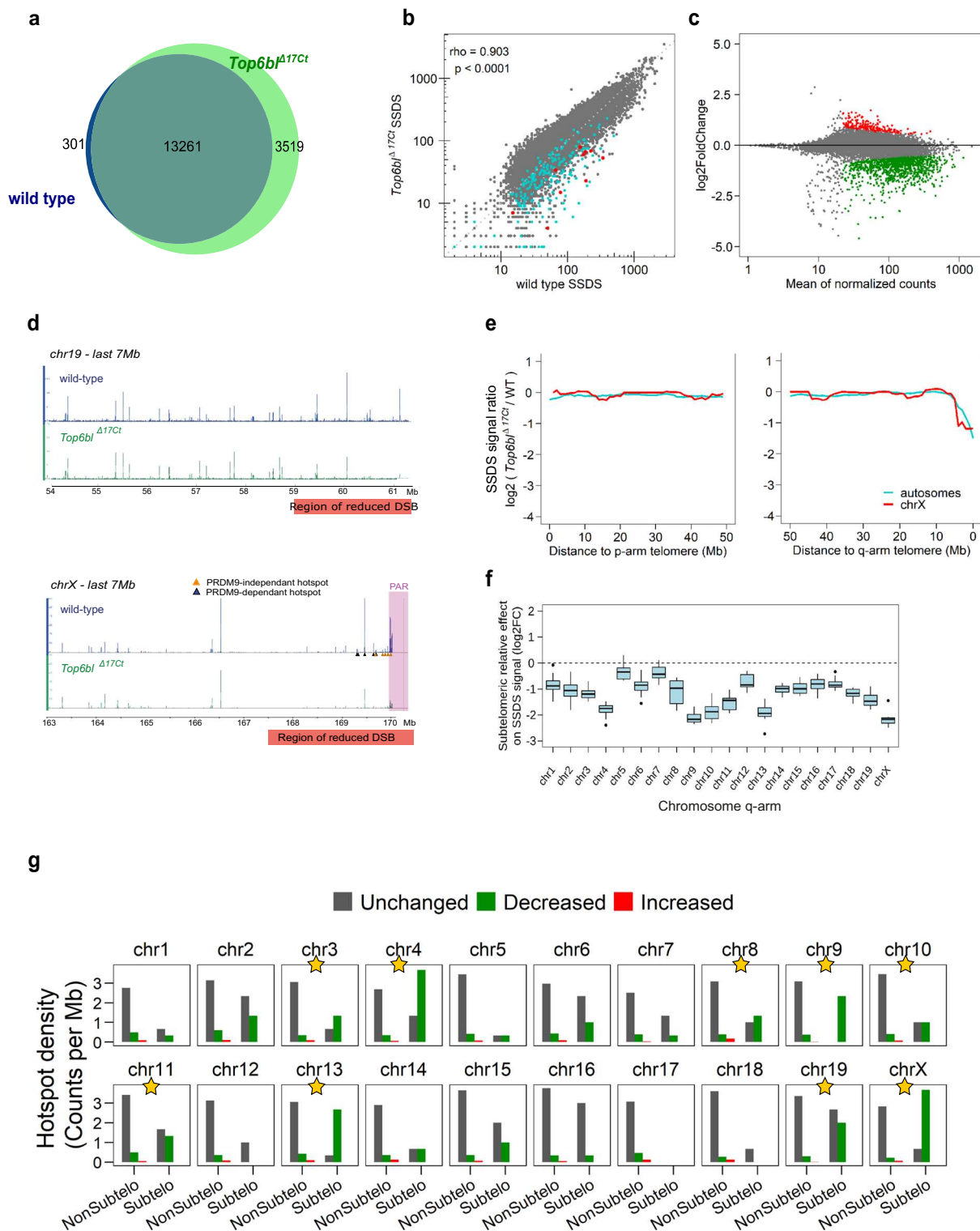


**Fig. 5 | Efficient REC114 and ANKRD31 loading in *Top6bl* mutant spermatocytes.**

**a** Immunostaining of REC114 and SYCP3 in spermatocytes from 12–14dpp wild-type (+/+), *Top6bl*<sup>W562A/W562A</sup>, and *Top6bl*<sup>Δ17C1/Δ17C1</sup> mice. Scale bar, 10 μm. **b** Quantification of axis-associated REC114 foci in early/mid and late leptotene, zygotene and pachytene spermatocytes from 12–14dpp wild-type (+/+ or *Top6bl*<sup>Δ17C1</sup>; n = 4), *Top6bl*<sup>W562A/W562A</sup> (n = 2), and *Top6bl*<sup>Δ17C1/Δ17C1</sup> (n = 2) mice. Each mutant is compared with a wild-type control in parallel to control for variations between experiments. Results from *Top6bl*<sup>W562A/W562A</sup> are from two independent experiments (labelled with filled and open circles respectively). Number of nuclei: wild-type: 31 early/mid L, 4 late L, 26 Z, 51 P; *Top6bl*<sup>W562A/W562A</sup>: 22 early/mid L, 8 late L, 26 Z, 24 P; wild-type 28 early/mid L, 12 late L, 24 Z, 46 P; *Top6bl*<sup>Δ17C1/Δ17C1</sup>: 18 early/mid L, 21 late L, 32 Z, 38 P; wild-type: 84 early/mid L, 59 late L, 84 Z, 75 P; *Top6bl*<sup>Δ17C1/Δ17C1</sup>: 76 early/mid L, 68 late L, 79 Z, 61 P. Grey bars show the mean values. P values were determined using the two-tailed unpaired Mann-Whitney test. **c** Variation of REC114 foci during prophase. The number of foci at early/mid leptotene, late leptotene, zygotene and pachynema (dataset as in panel b) relative to the mean number at early/mid

leptotene was plotted for wild-type, *Top6bl*<sup>W562A/W562A</sup>, and *Top6bl*<sup>Δ17C1/Δ17C1</sup> mice. The mean values ± SD are shown. Statistical significance was tested at late leptotene using the two-tailed unpaired Mann-Whitney test. **d** Quantification of axis-associated ANKRD31 foci in early/mid, late leptotene, zygotene and pachytene spermatocytes from 12–14 dpp wild-type (+/+ or *Top6bl*<sup>Δ17C1</sup>; n = 2), *Top6bl*<sup>W562A/W562A</sup> (n = 1), and *Top6bl*<sup>Δ17C1/Δ17C1</sup> (n = 1) mice. Mean number of nuclei per genotype at early/mid leptotene (37, 19, and 23), late leptotene (51, 23, and 19), zygotene (61, 22, and 35), and pachytene (100, 61, and 31). Grey bars show the mean values. P values were determined using the two-tailed unpaired Mann-Whitney test. **e** Variation of ANKRD31 foci during prophase. The number of foci at early/mid leptotene, late leptotene, zygotene, and pachynema (dataset as in panel d) relative to the mean number at early/mid leptotene (set at 1) was plotted for wild-type, *Top6bl*<sup>W562A/W562A</sup>, and *Top6bl*<sup>Δ17C1/Δ17C1</sup> mice. The mean values ± SD are shown. Statistical significance was tested at late leptotene using the two-tailed unpaired Mann-Whitney test. Source data are provided as a Source Data file.



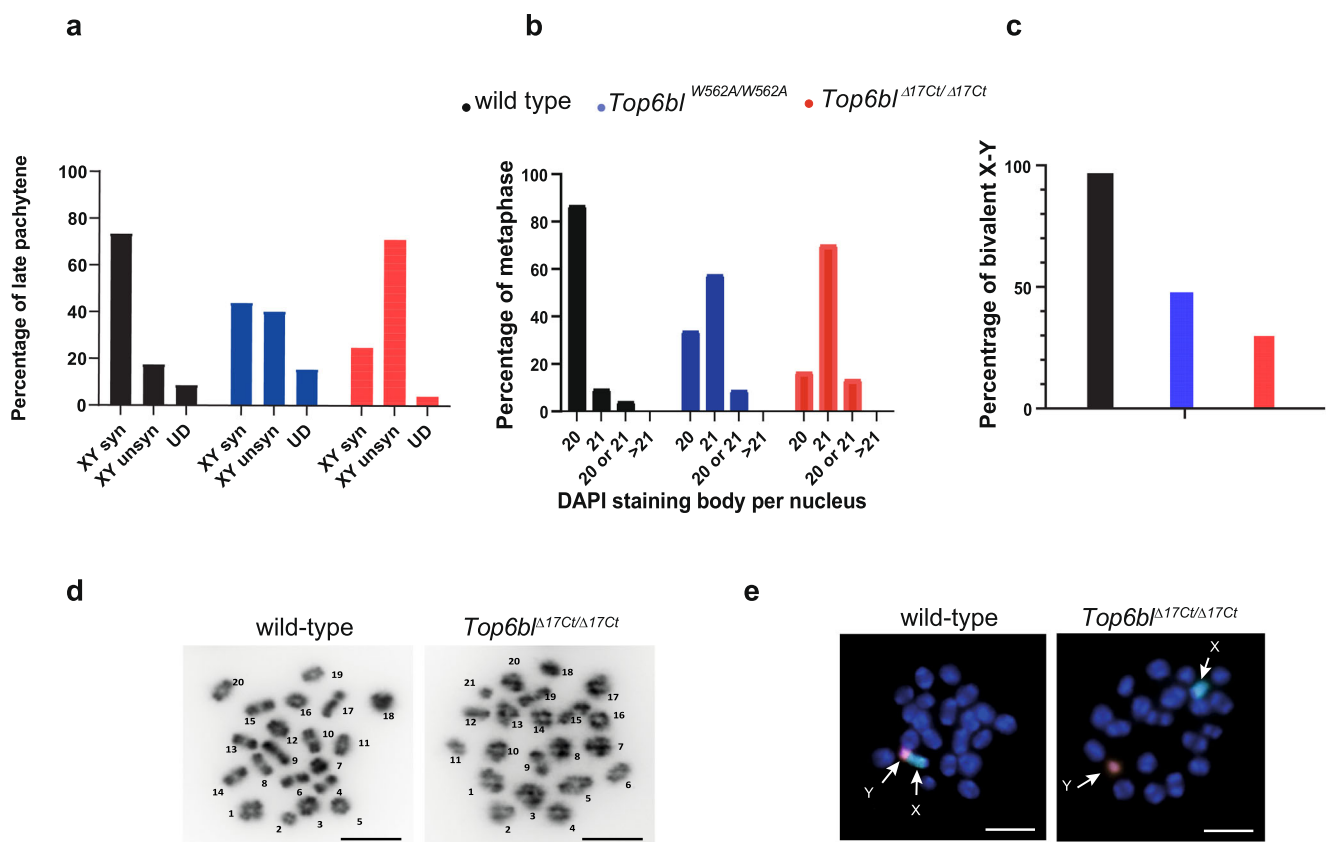


absent in MTOPVIB. It is therefore possible that the function of these proteins is mediated by distinct molecular interactions as suggested by interaction and protein localization assays in *A. thaliana*<sup>47</sup>. These divergences are correlated with a distinct role for *A. thaliana* REC114 which is dispensable for DSB formation<sup>47</sup>. AtREC114 may however regulate DSB activity at a distinct level, and it would be interesting to know if it plays a role in controlling the timing of DSB activity, since this is one of the REC114 roles we have uncovered in male mice.

The observation of meiotic DSB defects in female and male mice harbouring *Top6bl* mutations indicates that in vivo REC114 acts by directly interacting with the SPO11/TOPOVIBL complex. This interaction requires W562, a highly conserved residue of TOPOVIBL in metazoans (Fig. 1f)<sup>7</sup>. The stronger phenotype of mice carrying the *Top6bl*<sup>Δ17Ct</sup> mutation (compared with *Top6bl*<sup>W562A</sup>) could be explained by a lack of interaction between TOPOVIBL Δ17Ct and REC114 and a remaining weak interaction between TOPOVIBL W562A and REC114.

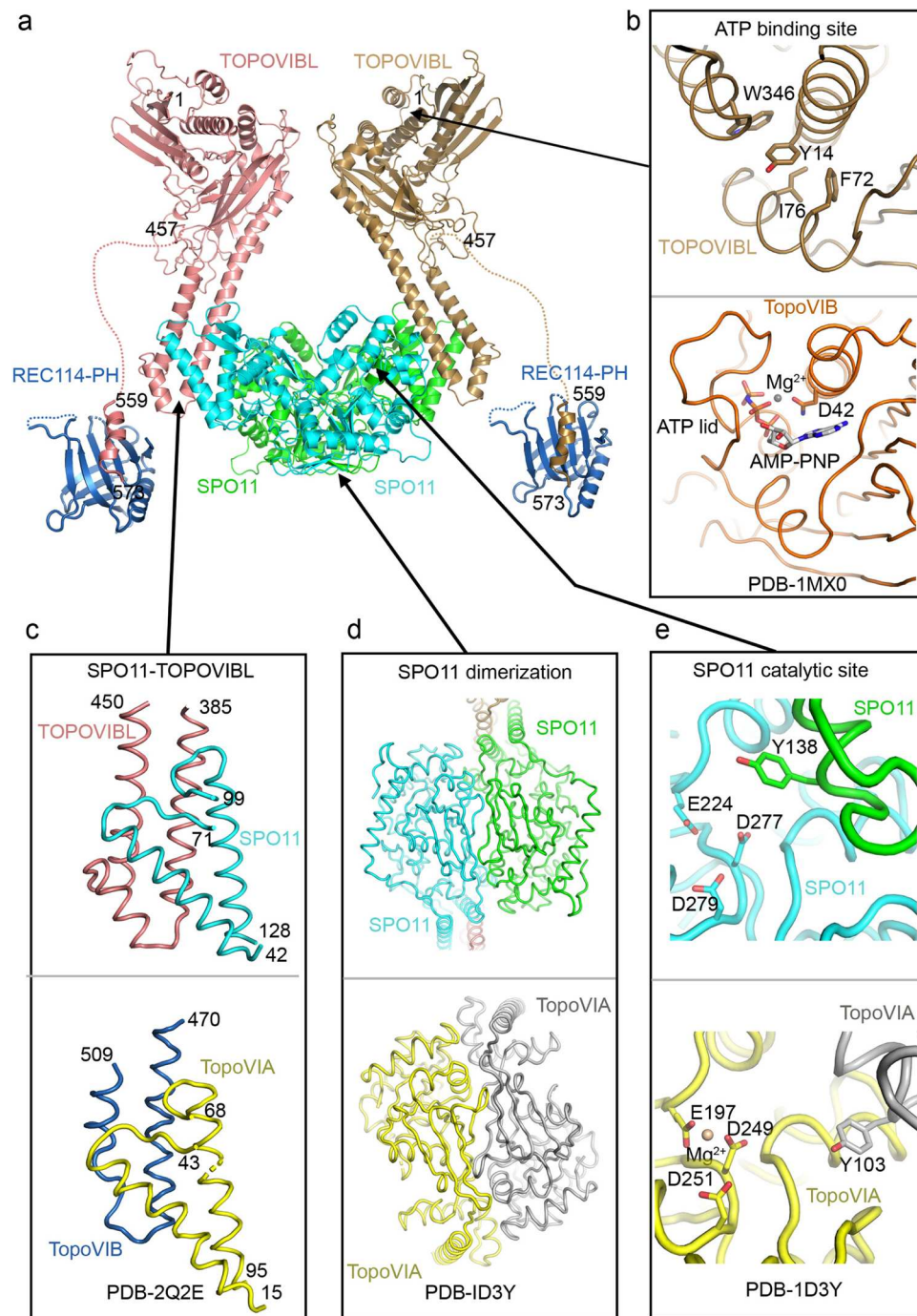
**Fig. 6 | Proper DMCI-SSDS signal localisation and intensity, excepted for q-arm sub-telomeric regions in *Top6bl*<sup>Δ17Cq/Δ17Cq</sup> spermatocytes.** **a** In *Top6bl*<sup>Δ17Cq/Δ17Cq</sup> spermatocytes, almost all wild-type hotspots and 20% of new hotspots are detected. **b** DMCI-SSDS signal correlation between wild-type and *Top6bl*<sup>Δ17Cq/Δ17Cq</sup> mice at wild-type hotspots. The Spearman rho and associated p-value (two-sided) are shown. Ten telomere-proximal hotspots are highlighted for each autosome (blue) and for the X chromosome (red). In the PAR, only part of the DMCI-SSDS signal, which covers a large domain (see panel d), is included within hotspots. **c** MA plot of the DMCI-SSDS signal in *Top6bl*<sup>Δ17Cq/Δ17Cq</sup> compared with wild-type samples. The log<sub>2</sub> of the ratios was calculated using DESeq2. Hotspot with significantly increased or decreased signal are highlighted in red ( $n = 214$ ) and green ( $n = 1098$ ), respectively (adjusted  $p$  value  $< 0.1$ ). Unchanged hotspots are in grey ( $n = 8043$ ). The mean normalized count corresponds to the baseMean value from the DESeq2 analysis. **d** DSB maps of the 7 Mb telomere-proximal regions of the chromosomes 19 and X in wild-type (blue) and *Top6bl*<sup>Δ17Cq/Δ17Cq</sup> (green) mice. On chromosome X PRDM9-independent and -dependent hotspots as defined by analysis in *Prdm9*<sup>-/-</sup> mice<sup>36</sup> are identified by orange and blue triangles, respectively. The PAR is highlighted with a pink rectangle. All chromosome ends are shown in Supplementary Figs. 14, 15.

**e** DMCI-SSDS signal intensity decreases in *Top6bl*<sup>Δ17Cq/Δ17Cq</sup> samples relative to wild-type samples in the q-arm sub-telomeric region (right panel). The DMCI-SSDS signal ratio within hotspots (log<sub>2</sub>-fold change estimated by DESeq2) was computed over 5Mb-windows with a 1Mb-step. The same analysis was performed in the 50 Mb adjacent to p-arm telomeres (left panel). **f** The q-arm sub-telomeric region effect in *Top6bl*<sup>Δ17Cq/Δ17Cq</sup> mice. The averaged DMCI-SSDS signal ratio (*Top6bl*<sup>Δ17Cq/Δ17Cq</sup>/wild-type) of the last ten hotspots of a given chromosome was compared to the averaged DMCI-SSDS signal ratio of ten randomly chosen, non-telomeric consecutive hotspots in the same chromosome. Box plots (as defined in Methods) represents the log<sub>2</sub> fold-change (FC) between these values for ten randomizations. The control is shown in Supplementary Fig. 13c. **g** Decreased (green), increased (red), or unchanged (grey) hotspot density within the 3 Mb sub-telomeric (Subtelo) region relative to the nonsub-telomeric regions (NonSubtelo). Decreased, increased and unchanged hotspots were determined from the DESeq2 analysis, as shown in panel c. The densities of decreased and unchanged hotspots in the two studied regions were compared using the Pearson's Chi-square test, and p-values were adjusted for multiple testing using the Benjamini & Yekutieli method. Yellow stars indicate  $p$  value  $< 0.05$ . The Chi-square test results are provided in Supplementary Table 2.



**Fig. 7 | Defective XY chromosome synapsis in *Top6bl* mutants.** **a** Synapsis quantification between the X and Y chromosomes at pachynema. Synapsis formation was monitored on spreads from late pachytene spermatocytes from adult wild-type (+/+), *Top6bl*<sup>W562A/W562A</sup> or *Top6bl*<sup>Δ17Cq/Δ17Cq</sup>, *Top6bl*<sup>W562A/W562A</sup>, and *Top6bl*<sup>Δ17Cq/Δ17Cq</sup> mice, identified by staining with  $\gamma$ H2AX, SYCP3 and SYCP1. Number of nuclei: wild-type (227), *Top6bl*<sup>W562A/W562A</sup> (108), *Top6bl*<sup>Δ17Cq/Δ17Cq</sup> (171). syn: synapsed; unsyn: unsynapsed; UD: undefined. The synapsed/unsynapsed ratios were significantly different between wild-type and *Top6bl*<sup>W562A/W562A</sup> and *Top6bl*<sup>Δ17Cq/Δ17Cq</sup> spermatocytes (Pearson's Chi-Square, 26.31 and 115.15, respectively). **b** Quantification of bivalent formation at metaphase I. Percentage of metaphases with 20, 21, 20 or 21, or  $>21$  DAPI-stained bodies per nucleus from adult wild-type (*Top6bl*<sup>+/+</sup>), *Top6bl*<sup>W562A/W562A</sup>, and *Top6bl*<sup>Δ17Cq/Δ17Cq</sup> mice ( $n = 2$  mice per genotype). Number of nuclei: 75, 80, and 97, respectively. The number of nuclei with 20 and 21 bivalents was significantly different between wild-type and *Top6bl*<sup>W562A/W562A</sup> and between wild-type and *Top6bl*<sup>Δ17Cq/Δ17Cq</sup> metaphases (Pearson's Chi-Square: 44.4 and 78.8

respectively). **c** Quantification of X and Y bivalents at metaphase I. The X and Y chromosomes were detected by FISH in metaphase spreads from adult wild-type (*Top6bl*<sup>+/+</sup>), *Top6bl*<sup>W562A/W562A</sup>, and *Top6bl*<sup>Δ17Cq/Δ17Cq</sup> mice ( $n = 2$  mice per genotype). Number of nuclei: 84, 144, and 164 for wild-type, *Top6bl*<sup>W562A/W562A</sup> and *Top6bl*<sup>Δ17Cq/Δ17Cq</sup>, respectively. The percentages of metaphase spreads with XY bivalents were significantly different between wild-type and *Top6bl*<sup>W562A/W562A</sup>, and between wild-type and *Top6bl*<sup>Δ17Cq/Δ17Cq</sup> (Chi-Square Pearson: 54.2 and 96.9, respectively). **d** Representative images of DAPI-stained wild-type and *Top6bl*<sup>Δ17Cq/Δ17Cq</sup> metaphase spreads. Scale bar, 10  $\mu$ m. DAPI-stained bodies are numbered (arbitrarily): 20 are observed in wild-type, and 21 in *Top6bl*<sup>Δ17Cq/Δ17Cq</sup> metaphase spreads (samples from panel b). **e** Representative images of a wild-type spermatocyte nucleus with the X and Y chromosomes forming a bivalent (left) and of a *Top6bl*<sup>Δ17Cq/Δ17Cq</sup> spermatocyte nucleus with separated X and Y chromosomes (right) (samples from panel c). Blue, nuclei (DAPI staining); green, X chromosome probe; red, Y chromosome probe. Scale bar, 10  $\mu$ m.



**Fig. 8 | Model of the SPO11-TOPOVIBL-REC114 complex.** **a** The structure of the complex of SPO11 and the C-terminal part of the TOPOVIBL transducer domain (residues 389-457) was modelled by AlphaFold2<sup>29</sup>. Full-length TOPOVIBL AlphaFold2-modelled structure (AF-J3QMY9) was then superimposed on the transducer domains of the SPO11-TOPOVIBL complex. Crystal structure of the REC114 PH domain bound the C-terminus of TOPOVIBL determined in this study is shown linked to the SPO11-TOPOVIBL complex via a long-disordered linker. No contact between REC114 PH-TOPOVIBL C-ter and the SPO11-TOPOVIBL complex could be modelled. The modelled complex structure coloured according to the AlphaFold2 per-residue estimate of confidence (pLDDT) and predicted aligned error plots are shown in Supplementary Fig. 16. **b** The AlphaFold2 model (AF-J3QMY9) suggests that the putative ATP-binding site of TOPOVIBL is degenerated. Secondary structure elements forming the ATP-binding site in *S. shibatae* (PDB - 1MX0, lower panel), are predicted to be organised differently in mouse TOPOVIBL.

In particular, the “ATP-lid” possessing key residues ATP binding residues is missing and is replaced by a helical insertion in the transducer domain. In consequence, a corresponding ATP binding site is not formed, and is area is filled with aromatic and hydrophobic residues of TOPOVIBL. TOPOVIBL also lacks the N-terminal “strap” region and the H2TH domain, both involved in the TopoVIB dimerization, indicating that TOPOVIBL dimerization, if occurs, should differ from TopoVIB<sup>39</sup>. **c** The modelled SPO11-TOPOVIBL interface is similar to that of the archaeal TopoVI complex (PDB - 2Q2E, lower panel), with the short additional helix of the TOPOVIBL transducer domain<sup>7</sup>. **d** The modelled SPO11 dimerization interface resembles the one described for archaeal TopoVIA, including the formation of a pseudo-continuous  $\beta$ -sheet formed by the two protomers<sup>41</sup>. **e** The modelled SPO11 catalytic site is equivalent to the one previously described for TopoVIA (PDB - 1D3Y, lower panel) where the two protomers contribute three negatively charged residues that co-ordinate the magnesium atom and the catalytic tyrosine<sup>41</sup>.



Such differential effect of the two mutations on the interaction is however not supported by in vitro analysis since no interaction between TOPOVIBL W562A and REC114 PH domain was detected in vitro. It remains possible that in vivo TOPOVIBL W562A can interact weakly with REC114. In vivo interaction assays would be required to test this hypothesis. Alternatively, the stronger phenotype observed in *Top6bl<sup>Δ17Ct</sup>* mice may be due to additional interactions or properties mediated by the last 17 amino acids of TOPOVIBL. The different phenotypes of *Top6bl* mutants, show that REC114 is a regulatory subunit of the activity, and not just an accessory factor of the TOPOVIBL complex. The reduced DSB activity in oocytes fits exactly the simple interpretation that REC114 binding to TOPOVIBL is required for the catalytic activity. The male phenotype highlights REC114 double role in fine tuning DSB activity by regulating the timing of DSB formation genome-wide and DSB formation at sub-telomeric regions. It is remarkable that the delayed DSB formation with respect to axis formation in *Top6bl<sup>Δ17Ct/Δ17Ct</sup>* mice does not alter homologous synapsis and seems to allow weak DSB sites to be more active than in wild-type. This could be explained if DSB sites can be turned off upon synapsis, as shown in *S. cerevisiae*<sup>48</sup>. Delaying synapsis would thus allow weak DSB sites to be active. The sub-telomeric effect in *Top6bl<sup>Δ17Ct/Δ17Ct</sup>* mice potentially applies to all chromosomes and is not restricted to the four chromosomes (chr. 4, 9, 13 and X) shown to accumulate REC114, MEI4, IHO1 and ANKRD31 aggregates<sup>27,28</sup>. The property we detected is therefore not strictly linked to those aggregates, which is also consistent with the observation that *Top6bl* mutants do not have any detectable defect in the formation of those aggregates (Supplementary Fig. 11a). We do notice inter-chromosomal quantitative differences of the extent of decreased DSB activity in *Top6bl<sup>Δ17Ct/Δ17Ct</sup>* mice that may at least in part be due to inter-chromosomal differences in hotspot activity within sub-telomeric regions (Supplementary Fig. 13e). Sub-telomeric regions display specific features (i.e., nuclear organization during meiosis<sup>49,50</sup>, architecture, organization and other epigenetic properties) that may influence the control of SPO11/TOPOVIBL activity. These regions are known to behave differently from chromosome arms for meiotic DSBs, based on their sensitivity to the expression of a GAL4BD-SPO11 fusion protein on DSB activity<sup>51</sup>, to a lower DMCI1/SPO11-oligonucleotides ratio<sup>52</sup> and to differential activities in male and female meiosis<sup>53</sup>. It is also possible that mice expressing SPO11β-only (the long isoform of SPO11) and shown to have decreased DSB activity in the PAR, would a specific decrease of DSB activity near autosome ends<sup>38</sup>. This would imply a specific regulation of SPO11 near chromosome ends, as the phenotypes due to the *Top6bl<sup>Δ17Ct</sup>* mutation suggest. As SPO11α (the short isoform of SPO11) is lacking the interaction domain with TOPOVIBL, it is not expected to be catalytically active. SPO11α may rather repress an inhibitor of TOPOVIBL. One possibility is that this inhibitor interferes with the REC114-TOPOVIBL interaction. Such scenario, although certainly others can be envisioned, would fit with the potential similarity of phenotypes between *Spo11β*-only and *Top6bl<sup>Δ17Ct/Δ17Ct</sup>* male mice, including the XY synapsis defect. Overall, these observations imply that i) the TOPOVIBL-REC114 interaction is not essential for DSB activity in all genomic contexts; ii) REC114 activity senses or responds to specific chromosomal features. Several studies have highlighted differences of recombination and/or chromosome organization between sexes as well along chromosomes but which links to DSB activity remain to be determined<sup>53,54</sup>.

Programming hundreds of DSBs in the genome is a challenge for the cell, and the current knowledge that the catalytic complex of SPO11/TOPOVIBL requires several other proteins is coherent with the need to regulate these events. Here, we described the central role of REC114 in SPO11/TOPOVIBL activity through its direct interaction with TOPOVIBL, thus highlighting a first level of this regulation. Moreover, the DSB program must be executed in two very different cell types

(oocytes and spermatocytes), and our findings show that the REC114-TOPOVIBL interaction is sensitive to these differences. Other components of the meiotic DSB machinery and their potential multiple interactions should contribute to SPO11/TOPOVIBL regulation. Additional directed-mutagenesis studies will unravel them and will identify the complex(es) active in vivo.

## Methods

### Mouse strains

Mice were in the C57BL/6J background. Mice carrying the homozygous mutant alleles *Top6bl*<em1(W562A)BdM> and *Top6bl*<em2(delta17)BdM> were named *Top6bl<sup>W562A/W562A</sup>* and *Top6bl<sup>Δ17Ct/Δ17Ct</sup>*, respectively. *Top6bl<sup>-/-</sup>* mice carry the Gm960<sup>em2Arte</sup> allele, a null allele due to a 5 bp deletion in *Top6bl*<sup>6</sup>. Mice housing conditions were: temperature 22 °C, humidity 55%, dark/light cycle 12/12 corresponding to 8 am/8 pm in summer and 7 am/7 pm in winter. All experiments were carried out according to the CNRS guidelines and were approved by the ethics committee on live animals (Comité National de Réflexion Ethique sur l'Expérimentation Animale; project CE-LR-0812 and 1295).

### Generation of mutant mice by CRISPR/Cas9

Mutant mice were created at the Jackson Laboratory using the CRISPR-Cas9 technology with three different guides and two different donor oligos (Supplementary Table 3). Guides were selected to minimize off-target effects. The donor oligos were designed to change the W562 codon TGG (W) to GCG (A), and to introduce a silent mutation (A to G) to generate a *Pst*I restriction site. The *Top6bl*<em1(W562A)BdM> allele, named *Top6bl<sup>W562A</sup>*, is the result of homologous recombination. The *Top6bl*<em2(delta17)BdM> allele, named *Top6bl<sup>Δ17Ct</sup>*, is the result of non-homologous repair and has a 4 bp deletion (Supplementary Fig. 6a). Founders were backcrossed with C57BL/6J animals to obtain heterozygous animals. The predicted TOPOVIBL protein expressed from each mutant allele and the genotyping strategies are shown in Supplementary Fig. 6.

### Yeast two-hybrid assays: clones, assays, western blotting

All plasmids used in yeast two-hybrid assays were cloned with the Gateway® Gene Cloning Technology (Invitrogen) and transformed in the AH109 and Y187 haploid yeast strains (Clontech). AH109 and Y187 cells were transformed with Gal4 DNA binding domain (GBD) fusion plasmids derived from pAS2 and Gal4 activation domain (GAD) fusion plasmids that were obtained from pGAD. Purified colonies of diploid strains were streaked on synthetic medium (SD) plates lacking leucine and tryptophan (-LW), or leucine, tryptophan and histidine (-LWH), or leucine, tryptophan and histidine with 5 mM aminotriazole (-LWH + 3AT), or leucine, tryptophan, histidine and adenine (-LWHA). Dilution assays were performed by spotting cells on -LW, -LWH, -LWH + 3AT and -LWHA plates that were incubated at 30 °C for 3 days. For verification of protein expression, protein extracts were prepared and analysed by western blotting, as previously described<sup>5</sup>, with anti-GAD (1:3000; UPSTATE-06-283) and anti-GBD (1:1000; SIGMA; G3042) antibodies.

### Protein expression, purification and crystallization

His-tagged mouse REC114 15-159 was expressed in *E. coli* BL21-Gold (DE3) cells (Agilent) from the pProEXHTb expression vector (Invitrogen). The protein was first purified by affinity chromatography using the Ni<sup>2+</sup> resin (Chelating Sepharose, GE Healthcare). After His-tag cleavage with the TEV protease, it was purified through a second Ni<sup>2+</sup> column and size-exclusion chromatography on Superdex 200 (GE Healthcare). The pure protein was concentrated to 20 mg ml<sup>-1</sup> in a buffer containing 20 mM Tris, pH 8.0, 200 mM NaCl and 5 mM β-mercaptoethanol, and supplemented with a three-fold molar excess of TOPOVIBL peptide (559-EDLWLQEVSNLSEWLNPG-576). The complex

was crystallized using the hanging drop vapor diffusion method at 20 °C. The best diffracting crystals grew within seven days in a solution containing 1.6 M MgSO<sub>4</sub>, 100 mM MES (pH 6.5), and 10% (v/v) dioxane. For data collection at 100 K, crystals were snap-frozen in liquid nitrogen with a solution containing mother liquor and 25% (vol/vol) glycerol.

### Data collection and structure determination

Crystals of the mouse REC114-TOPOVIBL complex belong to the space group *P*<sub>6</sub><sub>1</sub>22 with the unit cell dimensions *a*, *b* = 108.7 Å and *c* = 83. Å. The asymmetric unit contains one REC114-TOPOVIBL dimer and has a solvent content of 68%. A complete native dataset was collected to a 2.5 Å resolution, partially extending to 2.26 Å on the ESRF beamline ID30B using the MXcuBE3 software (ESRF). The data were processed using autoPROC<sup>55</sup>. Phases were obtained by molecular replacement using PHASER<sup>56</sup> with the crystal structure of the REC114 PH domain (PDB code: 6HFG) as search model. The initial map was improved using the prime-and-switch density modification option of RESOLVE<sup>57</sup>. After manual model rebuilding with COOT<sup>58</sup>, the structure was refined using Refmac5<sup>59</sup> to a final *R*-factor of 22.6% and *R*<sub>free</sub> of 24.9% (Supplementary Table 1) with all residues in the allowed (95.2% in favored) regions of the Ramachandran plot, as analysed by MOLPROBITY<sup>60</sup>. A representative part of the 2*F*<sub>o</sub> - *F*<sub>c</sub> electron density map covering the TOPOVIBL-REC114 interface is shown in Supplementary Fig. 2b.

### Pull-down assays

Full-length REC114 and its variants were cloned as Strep-tag fusions (using a single Strep-tag: WSHPQFEK) into pRSFDuet-1 (Novagen). TOPOVIBL<sup>1-385</sup> was cloned as a 6xHis-SUMO fusion protein in pETMLL and TOPOVIBL<sup>452-579</sup> and its mutated versions as 6xHis fusion proteins in pProExHTb. Proteins were expressed individually in *E. coli* BL21Gold (DE3) cells. All proteins were first purified by affinity chromatography (Strep-Tactin XT resin (IBA), Ni-Chelating Sepharose, GE Healthcare) and gel filtration on Superdex 200 (GE Healthcare). Mixtures of REC114 and TOPOVIBL were loaded onto Strep-Tactin XT (IBA) resin columns. Columns were then extensively washed with a buffer containing 100 mM Tris pH 8, 100 mM NaCl, 5 mM β-mercaptoethanol. Bound proteins were eluted by addition of 50 mM of D-Biotin, and analysed on SDS-PAGE. ANKRD31<sup>1808-1857</sup> was cloned as 6xHis-MBP fusion in pETM41. Strep-REC114<sup>1-159</sup> and ANKRD31<sup>1808-1857</sup> were individually expressed in *E. coli* BL21Gold (DE3) cells. Following cell disruption, supernatants containing soluble Strep-REC114<sup>1-159</sup> and ANKRD31<sup>1808-1857</sup> were mixed. The complex was further purified by Superdex 200 (GE Healthcare) gel filtration column. The pull-down with increasing amounts of TOPOVIBL<sup>452-579</sup> on Strep-Tactin XT (IBA) resin column was performed as above. For pull-down presented in Fig. 2g, a total of 0.8% of the input (lanes 1–6) and 1.2% of the eluates (lanes 7–14) were analyzed on 12% SDS-PAGE gels stained with coomassie brilliant blue. Control lanes 7 and 14 show REC114 alone. For pull-down presented in Fig. 2k, a total of 0.3% of the input (lanes 1–6) and 1.2% of the eluates (lanes 7–14) were analyzed on 12% SDS-PAGE gels stained with coomassie brilliant blue.

### Isothermal Titration Calorimetry (ITC)

ITC experiments were performed at 25 °C using an ITC200 microcalorimeter (MicroCal). Experiments included one 0.5 μl injection and 18–20 injections of 1.5–2 μl of 0.3–1.8 mM TOPOVIBL (TOPOVIBL<sup>452-579</sup>, TOPOVIBL<sup>452-562,WS62A,563-579</sup>, TOPOVIBL<sup>559-576</sup> or TOPOVIBL<sup>559-562,WS62A,563-576</sup>) into the sample cell that contained 30–40 μM REC114<sup>15-159</sup> or Strep-REC114<sup>1-159</sup> in 20 mM Tris (pH 8.0), 100 mM NaCl, 5% glycerol and 5 mM β-mercaptoethanol. The initial data point was deleted from the data sets. Binding isotherms were fitted with a one-site binding model by nonlinear regression using the Origin software, version 7.0 (MicroCal).

### Preparation of mouse protein extracts, immunoprecipitation and western blotting

Whole cell protein extracts were prepared from eight frozen testes collected at 14 dpp for each genotype. After protein extraction by homogenizing cells with a Dounce homogenizer in HNTG buffer (150 mM NaCl, 20 mM HEPES pH7.5, 1% Triton X100, 10% glycerol, 1 mM MgCl, Complete protease Inhibitor (Roche 11873580001)), followed by sonication, benzonase (250 U) was added at 4 °C for 1 h. After centrifugation (16000 g, 4 °C, 10 min) to remove debris, immunoprecipitation was performed with 5 μg of homemade anti-TOPOVIBL antibody. For each immunoprecipitation, 3.5 mg of whole cell protein extract and 50 μl of Protein A Dynabeads (Invitrogen 10001D) were used. Then, immunoprecipitates were resuspended in 40 μl of Laemmli buffer and TOPOVIBL immunoprecipitation was assessed by western blotting with a homemade affinity-purified anti-TOPOVIBL (1/1000) antibody followed by an anti-rabbit LC mouse monoclonal secondary antibody (Jackson ImmunoResearch 211-032-171, 1/3000).

### RT-PCR assays

Total RNA was extracted with the miRNeasy Mini Kit (Qiagen) according to the manufacturer's instructions. For RT-PCR, first-strand DNA was synthesized using oligo d(T)<sub>18</sub> (Ambion), SuperScriptIII (Invitrogen), and total RNA (1–2 μg) from 16dpc ovaries.

The open reading frames of *Top6bl* and *Spo11* were amplified using standard PCR conditions and the primer pairs Oli63/Oli70 and Spo11:116U22/Spo11:655L22, respectively (Supplementary Table 3). PCR cycling conditions were: 3 min at 94 °C, 35 cycles of 30 sec at 94 °C, 30 sec at 54 °C, and 2 min or 30 sec at 72 °C, followed by 5 min at 72 °C. *Top6bl* ORF was then digested with the *EcoI* enzyme.

### Histological analysis of paraffin sections and TUNEL assay

Mouse testes or ovaries were fixed in Bouin's solution for Periodic Acid Schiff (PAS) staining of testes and haematoxylin eosin staining of ovaries. Fixation was in 4% paraformaldehyde/1X PBS for immunostaining and TUNEL assay. Testes and ovaries were embedded in paraffin and cut in 3 μm-thick sections. Sections were scanned using the automated tissue slide-scanning tool of a Hamamatsu NanoZoomer Digital Pathology system. TUNEL assay was performed with the DeadEnd Fluorometric TUNEL System (Promega), according to the manufacturer's protocol.

### Spermatozoid counting

After dissection of the epididymis caudal part from adult testes (2 month-old), spermatozooids were extracted from the epididymis by smashing or crushing the tissue in PBS. After homogenization by pipetting, 10 μl of the soluble part was diluted in 1 mL of water, and spermatozooids were counted.

### Immunocytology

Spread from spermatocytes and oocytes were prepared with the dry down technique, as described<sup>61</sup>. Briefly, a suspension of testis cells was prepared in PBS, and then incubated in a hypotonic solution for 8 min at room temperature. Cells were centrifuged, resuspended in a solution of 66 mM sucrose and spread on slides with 1% paraformaldehyde, 0.05% Triton. Nuclei were dried for 1 to 2 h in a humid chamber. Immunostaining was performed using a milk-based blocking buffer (5% milk, 5% donkey serum in PBS)<sup>62</sup>. Primary antibodies were incubated overnight at room temperature. Secondary antibodies were incubated at 37 °C for 1.5 h. Nuclei were stained with DAPI (4'-6-Diamidino-2-phenylindole, 2 μg/ml) during the final washing step.

### Antibodies

Guinea pig anti-SYCP3<sup>62</sup> (1/500), rabbit anti-SYCP1 (Abcam, 15090, 1/400), rabbit anti-DMC1 (Santa Cruz, H100, 1/200) anti-RPA2 (Abcam, ab76420 clone name EPR2877Y, 1/200), anti-MEI4<sup>23</sup> (1/100),

anti-REC114<sup>25</sup> (1/50), anti-REC114 (gift from S. Keeney, 1/1000), anti-ANKRD31<sup>28</sup> (1/400), anti-IHO1<sup>26</sup> (1/2000) and mouse monoclonal anti-phospho-histone H2AX (Ser139) ( $\gamma$ H2AX) (Millipore, 05-636, 1/10000) antibodies were used for immunostaining. Homemade affinity purified anti-TOPOVIBL antibody: rabbits were injected with full-length mouse His-TOPOVIBL protein prepared from *E. coli* inclusion bodies. Rabbit serum was purified by affinity using His-TOPOVIBL purified from inclusion bodies.

### Metaphase spread preparation

Tubules from decapsulated testes were pulled apart in 1% trisodium citrate and transferred into a 15 ml tube. After pipette homogenization and 1 min sedimentation, the cell-containing supernatant was transferred in a new 15 ml tube. Following the same procedure, the tubule pieces were rinsed twice with 3 ml of 1% trisodium citrate. The cell solution was centrifuged at 180 g for 10 min, and the pellet resuspended in 100  $\mu$ l of supernatant. Then, 3 ml of methanol: acetic acid: chloroform (3:1:0.05) solution was added drop by drop to the cell solution (rolling the first drops down the sides of the tube while flicking the tube). Cells were then centrifuged at 180 g for 10 min and resuspended in 100  $\mu$ l of supernatant, and 3 ml of methanol: acetic acid (3:1) was added to the tube. After 10 min of incubation at room temperature, cells were centrifuged again (180 g for 10 min) and resuspended in -1 ml methanol: acetic acid (3:1). To prepare the slides, 40  $\mu$ l of the cell suspension was dropped from a height of ~40 cm onto a slide that was held tilted at 45°. Slides were dried in a humid chamber.

### FISH for chromosome painting

X (D-1420-050-FI; D-1420-050-OR) and Y (D-1421-050-FI; D-1421-050-OR) chromosome-specific probes were used according to the manufacturer (Metasystems Probes). 10  $\mu$ l of probe mixture were added onto slides with metaphase spreads, covered with coverslips and sealed. The samples were then denatured, hybridized, washed and stained with DAPI as recommended by the manufacturer (Metasystems Probes).

### Image analysis

For focus quantification, all images were deconvoluted using the Huygens software. Image J was used to quantify foci that colocalized with the chromosome axis defined by SYCP3 staining.

For  $\gamma$ H2AX quantification in oocytes, signal intensity was obtained using Cell Profiler on non-deconvoluted images. Integrated intensity was used for the analysis. For spermatocytes, both Cell Profiler and Image J quantifications were performed and gave similar results. The output of Image J integrated intensity is presented.

Staging criteria were as follows. Pre-leptotene nuclei had weak SYCP3 nuclear signal and no or very weak  $\gamma$ H2AX signal; early leptotene nuclei were  $\gamma$ H2AX-positive and with only short SYCP3 fragments; mid leptotene nuclei were  $\gamma$ H2AX-positive and with short and long SYCP3 fragments; late leptotene were  $\gamma$ H2AX-positive and with only long SYCP3 fragments; zygotene nuclei had partially synapsed homologs; and pachytene cells had all 19 autosomes fully synapsed (Supplementary Fig. 10a).

### DMC1-SSDS analysis

**Library preparation and sequencing.** DMC1 ChIP-seq was performed as described in<sup>63</sup> using a goat anti-DMC1 antibody (0.5 mg/ml; Santa Cruz, C-20). Six *Top6bl* <sup>$\Delta$ 17Ct/ $\Delta$ 17Ct</sup> testes and two wild-type testes from 12 to 25-week-old mice were used for each replicate. Sequencing was performed on a HiSeq 2500 instrument in paired-end mode (2x150bp).

**DMC1-SSDS mapping and hotspot identification.** After quality control and read trimming to remove adapter sequences and low-quality reads, DMC1 ChIP-SSDS reads were mapped to the UCSC mouse genome assembly build GRCm38/mm10. The previously published

method<sup>35</sup> was used for DMC1-SSDS read mapping (i.e., the BWA modified algorithm and a customized script that were specifically developed to align and recover ssDNA fragments). A filtering step was performed on the aligned reads to keep only non-duplicated and high-quality uniquely mapped reads with no more than one mismatch per read. To identify meiotic hotspots from biologically replicated samples in DMC1-SSDS, the Irreproducible Discovery Rate (IDR) method was used, as done in our previous studies. This method was developed for ChIP-seq analysis and extensively used by the ENCODE and modENCODE projects<sup>64</sup>. The framework developed by Qunhua Li and Peter Bickel's group (<https://sites.google.com/site/anshulkundaje/projects/idr>) was followed. Briefly, this method allows testing the reproducibility within and between replicates by using IDR statistics. Following their pipeline, peak calling was performed using MACS version 2.0.10 with relaxed conditions ( $-p$ -value = 0.1-bw1000-nomodel-shift400) for each of the two replicates, the pooled dataset, and pseudo-replicates that were artificially generated by randomly sampling half of the reads twice for each replicate and the pooled dataset. Then IDR analyses were performed, and reproducibility was checked. Final peak sets were built by selecting the top N peaks from pooled datasets (ranked by increasing p values), with N defined as the highest value between N1 (the number of overlapping peaks with an IDR below 0.01, when comparing pseudo-replicates from pooled datasets) and N2 (the number of overlapping peaks with an IDR below 0.05 when comparing the true replicates, as recommended for the mouse genome). Hotspot centring and strength calculation were computed following the method described by Khil et al.<sup>35</sup>. All read distributions and signal intensities presented in this work were calculated after pooling reads from both replicates, if not otherwise stated. When DSB maps were compared between mouse genotypes, the 1bp-overlaps were restricted to the central 400 bp of hotspots (+/- 200 bp around the peak centre). For correlation plots, the type 1 single-strand DNA signal was library-normalized (fragment per million).

**Differential analysis of hotspot strength.** To compare hotspot usage between *Top6bl* <sup>$\Delta$ 17Ct/ $\Delta$ 17Ct</sup> and wild-type mice, DMC1-SSDS signal intensity was compared at the 13562 wild-type hotspots using DESeq2<sup>65</sup>. Among these hotspots, 4207 hotspots (31%) were filtered out with the default independent filtering option using the mean of normalized counts as filter statistic. The aim was to remove sites with too low counts (mean count below 22) that have zero or low chance of showing significant differences to increase the detection power for the other sites. For the 9355 tested hotspots, log fold change shrinkage was performed to correct data dispersion using the *apeglm* method<sup>66</sup>. p-value were adjusted for multiple testing within DESeq2 using the procedure of Benjamini and Hochberg. The hotspots with increased or decreased DMC1-SSDS signal intensity were then determined using an adjusted p-value threshold of 0.1 and a log2 fold change value below or above zero, respectively. This led to the identification of 214 increased (1.2% of total hotspots, and 2% of the tested ones), 1098 decreased (8% of total hotspots, and 12% of the tested ones), and 8043 unchanged hotspots in *Top6bl* <sup>$\Delta$ 17Ct/ $\Delta$ 17Ct</sup> mice (Fig. 6c).

**Analysis of hotspot distribution at sub-telomeric regions.** Visual inspection of DMC1-SSDS signal intensity along chromosomes and the localization of hotspots with decreased DMC1-SSDS signal intensity in *Top6bl* <sup>$\Delta$ 17Ct/ $\Delta$ 17Ct</sup> suggested that much of the DMC1 signal decrease was located near the q-arm telomeres (telomeres distant from centromeres). Note that as the genomic DNA sequence of p-arms has not been assembled, it is immediately flanked by centromeric and q-arm DNA sequences, where the signal can be mapped and quantified. To test whether this biased distribution was significant, each hotspot was annotated as sub-telomeric when within the sub-telomeric region defined with a variable size from 1 to 10 Mb. For each sub-telomeric region, the numbers of unchanged, decreased and increased hotspots



in the sub-telomeric versus the non-sub-telomeric region (i.e., the rest of the chromosome) were counted. Pearson's Chi-square tests were computed (by taking into account or not the increased hotspots) and *p*-values were adjusted for multiple testing using the Benjamini & Yekutieli method. Megabase-normalized counts (hotspot density) were measured and plotted (Fig. 6g for sub-telomeric regions of 3 Mb). For each chromosome, the sub-telomeric over non-sub-telomeric ratio of hotspot density was calculated for each unchanged, decreased or increased hotspot category (Supplementary Fig. 13d). Alternatively, to evaluate hotspot activity without a fixed distance from the telomere, for each chromosome, the *Top6bl*<sup>Δ17C/Δ17C'</sup>/wild-type signal ratio was averaged over the ten most q-arm telomeric hotspots, and then compared to the averaged signal ratio measured over another set of ten consecutive hotspots randomly chosen along the chromosome (excluding the last tens). Then, the ratio of these two mean values was computed. The procedure was repeated 10 times, each time with a different random set of 10 non-sub-telomeric hotspots. The sub-telomeric effects are presented in Fig. 6f as the distribution of these ratios. As control, these ratios were computed not between the last ten and ten non-sub-telomeric hotspots, but between two non-sub-telomeric random hotspot sets (procedure repeated 10 times) (Supplementary Fig. 13c).

### Statistical analysis

The statistical analyses of cytological observations were done with GraphPad Prism 9. The nonparametric Mann-Whitney test was used to compare the number of foci, and the Pearson's Chi square test to compare distributions, as indicated in the figure legends. The Chi square tests were performed at <http://vassarstats.net/> and <https://www.quantitativeskills.com/sisa/index.htm>. Box plots (25–75 percentiles) show the median and 5–95 percentiles. Statistical tests for DMCI-SSDS data were done using R version 4.0.3. All tests and *p* values (n.s., not significant, \**P* < 0.05, \*\**P* < 0.01, \*\*\**P* < 0.001, \*\*\*\**P* < 0.0001) are provided in the corresponding legends and/or figures.

### Statistics and reproducibility

The pull-down assays shown in Figs. 1d, 2g, were performed twice. The pull-down assays shown Fig. 2k and Supplementary Fig. 4b were performed three times. The gel filtrations shown in Supplementary Figs. 2a, 4d, 5b, c, 6d were performed at least twice. IP from Supplementary Fig. 6e was reproduced three times. RT-PCR from Supplementary Fig. 6f were performed twice per genotype.

### Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

### Data availability

The DMCI-SSDS raw and processed data for this study have been deposited in the European Nucleotide Archive (ENA) at EMBL-EBI and are available through the project identifier [PRJEB43730](https://www.ebi.ac.uk/pdbe/). The atomic coordinates and structure factors of the mouse REC114-TOPOVIBL complex determined in this study have been deposited at the Protein Data Bank (<http://www.ebi.ac.uk/pdbe/>) under the PDB accession code [7QWV](https://www.ebi.ac.uk/pdbe/). Additional information, resources and reagents are available from the corresponding author upon reasonable request. Source data are provided with this paper.

### References

- Hunter, N. Meiotic Recombination: The Essence of Heredity. *Cold Spring Harb. Perspect. Biol.* **7**, a016618 (2015).
- Baudat, F., Imai, Y. & de Massy, B. Meiotic recombination in mammals: localization and regulation. *Nat. Rev. Genet.* **14**, 794–806 (2013).
- de Massy, B. Initiation of meiotic recombination: how and where? Conservation and specificities among eukaryotes. *Annu. Rev. Genet.* **47**, 563–599 (2013).
- Lam, I. & Keeney, S. Mechanism and regulation of meiotic recombination initiation. *Cold Spring Harb. Perspect. Biol.* **7**, a016634 (2014).
- Robert, T. et al. The TopoVIB-Like protein family is required for meiotic DNA double-strand break formation. *Science* **351**, 943–949 (2016).
- Vrielynck, N. et al. A DNA topoisomerase VI-like complex initiates meiotic recombination. *Science* **351**, 939–943 (2016).
- Brinkmeier, J., Coelho, S. M., de Massy, B. & Bourbon, H.-M. Evolution and diversity of the TopoVI and TopoVI-like subunits with extensive divergence of the TOPOVIBL subunit. *Mol. Biol. Evol.* **39**, msac227 (2022).
- Bergerat, A. et al. An atypical topoisomerase II from Archaea with implications for meiotic recombination [see comments]. *Nature* **386**, 414–417 (1997).
- Corbett, K. D., Benedetti, P. & Berger, J. M. Holoenzyme assembly and ATP-mediated conformational dynamics of topoisomerase VI. *Nat. Struct. Mol. Biol.* **14**, 611–619 (2007).
- Graille, M. et al. Crystal Structure of an Intact Type II DNA Topoisomerase: Insights into DNA Transfer Mechanisms. *Structure* **16**, 360–370 (2008).
- Keeney, S., Giroux, C. N. & Kleckner, N. Meiosis-specific DNA double-strand breaks are catalyzed by Spo11, a member of a widely conserved protein family. *Cell* **88**, 375–384 (1997).
- Arora, C., Kee, K., Maleki, S. & Keeney, S. Antiviral protein ski8 is a direct partner of spo11 in meiotic DNA break formation, independent of its cytoplasmic role in RNA metabolism. *Mol. Cell* **13**, 549–559 (2004).
- Claeys Bouuaert, C. et al. Structural and functional characterization of the Spo11 core complex. *Nat. Struct. Mol. Biol.* **28**, 92–102 (2021).
- Jiao, K., Salem, L. & Malone, R. Support for a meiotic recombination initiation complex: interactions among Rec102p, Rec104p, and Spo11p. *Mol. Cell Biol.* **23**, 5928–5938 (2003).
- Claeys Bouuaert, C. et al. DNA-driven condensation assembles the meiotic DNA break machinery. *Nature* **592**, 144–149 (2021).
- Maleki, S., Neale, M. J., Arora, C., Henderson, K. A. & Keeney, S. Interactions between Mei4, Rec114, and other proteins required for meiotic DNA double-strand break formation in *Saccharomyces cerevisiae*. *Chromosoma* **116**, 471–486 (2007).
- Baudat, F., Manova, K., Yuen, J. P., Jasin, M. & Keeney, S. Chromosome synapsis defects and sexually dimorphic meiotic progression in mice lacking spo11. *Mol. Cell* **6**, 989–998 (2000).
- Romanienko, P. J. & Camerini-Otero, R. D. The mouse spo11 gene is required for meiotic chromosome synapsis. *Mol. Cell* **6**, 975–987 (2000).
- Grey, C., Baudat, F. & de Massy, B. PRDM9, a driver of the genetic map. *PLoS Genet.* **14**, e1007479 (2018).
- Imai, Y. et al. PRDM9 activity depends on HELLS and promotes local 5-hydroxymethylcytosine enrichment. *Elife* **9**, e57117 (2020).
- Spruce, C. et al. HELLS and PRDM9 form a pioneer complex to open chromatin at meiotic recombination hot spots. *Genes Dev.* **34**, 398–412 (2020).
- Dereli, I. et al. Four-pronged negative feedback of DSB machinery in meiotic DNA-break control in mice. *Nucleic Acids Res.* **49**, 2609–2628 (2021).
- Kumar, R., Bourbon, H. M. & de Massy, B. Functional conservation of Mei4 for meiotic DNA double-strand break formation from yeasts to mice. *Genes Dev.* **24**, 1266–1280 (2010).
- Kumar, R. et al. MEI4 - a central player in the regulation of meiotic DNA double-strand break formation in the mouse. *J. Cell Sci.* **128**, 1800–1811 (2015).

25. Kumar, R. et al. Mouse REC114 is essential for meiotic DNA double-strand break formation and forms a complex with MEI4. *Life Sci Alliance* **1**, e201800259 (2018).
26. Stanzione, M. et al. Meiotic DNA break formation requires the unsynapsed chromosome axis-binding protein IHO1 (CCDC36) in mice. *Nat. Cell Biol.* **18**, 1208–1220 (2016).
27. Boekhout, M. et al. REC114 Partner ANKRD31 Controls Number, Timing, and Location of Meiotic DNA Breaks. *Mol. Cell* **74**, 1053–1068 e8 (2019).
28. Papanikos, F. et al. Mouse ANKRD31 Regulates Spatiotemporal Patterning of Meiotic Recombination Initiation and Ensures Recombination between X and Y Sex Chromosomes. *Mol. Cell* **74**, 1069–1085 e11 (2019).
29. Jumper, J. et al. Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589 (2021).
30. Cohen, P. E., Pollack, S. E. & Pollard, J. W. Genetic analysis of chromosome pairing, recombination and cell cycle control during first meiotic prophase in mammals. *Endocr. Rev.* **27**, 398–426 (2006).
31. Mahadevaiah, S. K. et al. Recombinational DNA double-strand breaks in mice precede synapsis. *Nat. Genet.* **27**, 271–276 (2001).
32. Hinch, A. G. et al. The Configuration of RPA, RAD51, and DMC1 Binding in Meiosis Reveals the Nature of Critical Recombination Intermediates. *Mol. Cell* **79**, 689–701 e10 (2020).
33. Fraune, J., Schramm, S., Alsheimer, M. & Benavente, R. The mammalian synaptonemal complex: protein components, assembly and role in meiotic recombination. *Exp. Cell Res* **318**, 1340–1346 (2012).
34. Libby, B. J., Reinholdt, L. G. & Schimenti, J. C. Positional cloning and characterization of Mei1, a vertebrate-specific gene required for normal meiotic chromosome synapsis in mice. *Proc. Natl Acad. Sci. USA* **100**, 15706–15711 (2003).
35. Khil, P. P., Smagulova, F., Brick, K. M., Camerini-Otero, R. D. & Petukhova, G. V. Sensitive mapping of recombination hotspots using sequencing-based detection of ssDNA. *Genome Res* **22**, 957–965 (2012).
36. Brick, K., Smagulova, F., Khil, P., Camerini-Otero, R. D. & Petukhova, G. V. Genetic recombination is directed away from functional genomic elements in mice. *Nature* **485**, 642–645 (2012).
37. Odorisio, T., Rodriguez, T. A., Evans, E. P., Clarke, A. R. & Burgoyne, P. S. The meiotic checkpoint monitoring synapsis eliminates spermatocytes via p53-independent apoptosis. *Nat. Genet.* **18**, 257–261 (1998).
38. Kauppi, L. et al. Distinct Properties of the XY Pseudoautosomal Region Crucial for Male Meiosis. *Science* **331**, 916–920 (2011).
39. Corbett, K. D. & Berger, J. M. Structure of the topoisomerase VI-B subunit: implications for type II topoisomerase mechanism and evolution. *EMBO J.* **22**, 151–163 (2003).
40. Dutta, R. & Inouye, M. GHKL, an emergent ATPase/kinase superfamily. *Trends Biochem Sci.* **25**, 24–28 (2000).
41. Nichols, M. D., DeAngelis, K., Keck, J. L. & Berger, J. M. Structure and function of an archaeal topoisomerase VI subunit with homology to the meiotic recombination factor Spo11. *Embo J.* **18**, 6177–6188 (1999).
42. Guo, H. et al. Phosphoregulation of DSB-1 mediates control of meiotic double-strand break activity. *Elife* **11**, e77956 (2022).
43. Rosu, S. et al. The *C. elegans* DSB-2 Protein Reveals a Regulatory Network that Controls Competence for Meiotic DSB Formation and Promotes Crossover Assurance. *PLoS Genet* **9**, e1003674 (2013).
44. Stamper, E. L. et al. Identification of DSB-1, a Protein Required for Initiation of Meiotic Recombination in *Caenorhabditis elegans*, Illuminates a Crossover Assurance Checkpoint. *PLoS Genet* **9**, e1003679 (2013).
45. Hinman, A.W. et al. *Caenorhabditis elegans* DSB-3 reveals conservation and divergence among protein complexes promoting meiotic double-strand breaks. *Proc Natl Acad Sci USA* **118**, e2109306118 (2021).
46. Carballo, J. A. et al. Budding Yeast ATM/ATR Control Meiotic Double-Strand Break (DSB) Levels by Down-Regulating Rec114, an Essential Component of the DSB-machinery. *PLoS Genet* **9**, e1003545 (2013).
47. Vrielynck, N. et al. Conservation and divergence of meiotic DNA double strand break forming mechanisms in *Arabidopsis thaliana*. *Nucleic Acids Res* **49**, 9821–9835 (2021).
48. Thacker, D., Mohibullah, N., Zhu, X. & Keeney, S. Homologue engagement controls meiotic DNA break number and distribution. *Nature* **510**, 241–246 (2014).
49. Scherthan, H. Telomere attachment and clustering during meiosis. *Cell Mol. Life Sci.* **64**, 117–124 (2007).
50. Shibuya, H. et al. MAJIN Links Telomeric DNA to the Nuclear Membrane by Exchanging Telomere Cap. *Cell* **163**, 1252–1266 (2015).
51. Smagulova, F. et al. Suppression of genetic recombination in the pseudoautosomal region and at subtelomeres in mice with a hypomorphic Spo11 allele. *BMC Genomics* **14**, 493 (2013).
52. Yamada, S. et al. Genomic and chromatin features shaping meiotic double-strand break formation and repair in mice. *Cell Cycle*, 0 (2017).
53. Brick, K. et al. Extensive sex differences at the initiation of genetic recombination. *Nature* **561**, 338–342 (2018).
54. Gruhn, J. R., Rubio, C., Broman, K. W., Hunt, P. A. & Hassold, T. Cytological studies of human meiosis: sex-specific differences in recombination originate at, or prior to, establishment of double-strand breaks. *PLoS One* **8**, e85075 (2013).
55. Vonrhein, C. et al. Data processing and analysis with the autoPROC toolbox. *Acta Crystallogr D. Biol. Crystallogr* **67**, 293–302 (2011).
56. McCoy, A. J. et al. Phaser crystallographic software. *J. Appl Crystallogr* **40**, 658–674 (2007).
57. Terwilliger, T. C. Maximum-likelihood density modification. *Acta Crystallogr D. Biol. Crystallogr* **56**, 965–972 (2000).
58. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr D. Biol. Crystallogr* **66**, 486–501 (2010).
59. Murshudov, G. N., Vagin, A. A. & Dodson, E. J. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D. Biol. Crystallogr* **53**, 240–255 (1997).
60. Chen, V. B. et al. MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr D. Biol. Crystallogr* **66**, 12–21 (2010).
61. Peters, A. H., Plug, A. W., van Vugt, M. J. & de Boer, P. A drying-down technique for the spreading of mammalian meiocytes from the male and female germline. *Chromosome Res* **5**, 66–68 (1997).
62. Grey, C., Baudat, F. & de Massy, B. Genome-Wide Control of the Distribution of Meiotic Recombination. *PLoS Biol.* **7**, e35 (2009).
63. Grey, C. et al. In vivo binding of PRDM9 reveals interactions with noncanonical genomic sites. *Genome Res* **27**, 580–590 (2017).
64. Landt, S. G. et al. ChIP-seq guidelines and practices of the ENCODE and modENCODE consortia. *Genome Res* **22**, 1813–1831 (2012).
65. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).
66. Zhu, A., Ibrahim, J. G. & Love, M. I. Heavy-tailed prior distributions for sequence count data: removing the noise and preserving large differences. *Bioinformatics* **35**, 2084–2092 (2019).

## Acknowledgements

We thank the following Biocampus facilities from Montpellier for their service: the Réseau des Animaleries de Montpellier (RAM) for animal care, the management of our mouse strains by Manon Leportier, the Réseau d'Histologie Expérimentale de Montpellier (RHEM) for histology,



and Montpellier Resources Imagerie (MRI) for microscopy. We thank all lab members for discussion and Pauline Auffret for support in bioinformatic analysis. We thank Scott Keeney for the anti-REC114 antibody, Attila Toth for the anti-ANKRD31 and anti-IHO1 antibodies. We thank Caroline Mas for assistance and access to the biophysics platform, the staff of the ESRF-EMBL (European Synchrotron Radiation Facility-European Molecular Biology Laboratory) Joint Structural Biology Group, particularly Andrew McCarthy and Matthew Bowler, for access to and help with the ESRF beamlines. We thank the EMBL high-throughput crystallization facility (HTX). BdM and JK were funded by ANR Topobreaks (ANR-18-CE11-0024-01). BdM was also funded by Prize Coups d'Élan for French Research from the Fondation Bettencourt-Schueller, ERC (European Research Council (ERC) Executive Agency under the European Community's Seventh Framework Programme (FP7/2007-2013 Grant Agreement no. 322788) and under the European Union's Horizon 2020 research and innovation programme (Grant Agreement no. 883605)) and CNRS. We acknowledge also the CNRS-Plan Cancer ATIP-Avenir program funding to JK, and the CNRS INSERM ATIP-Avenir 2017 program funding to TR. Ariadna B. Juarez-Martinez was supported by the Labex GRAL (Grenoble Alliance for Integrated Structural Cell Biology) (ANR-10-LABX-49-01) and the People Program (Marie Curie Actions) of the European Union's Seventh Framework Program (FP7/2007–2013) under REA grant agreement PCOFUND-GA-2013-609102, through the PRESTIGE program coordinated by Campus France. Financial support from the Center National de la Recherche Scientifique (IR-RMN-THC Fr3050) is gratefully acknowledged. IBS acknowledges integration into the Interdisciplinary Research Institute of Grenoble (IRIG, CEA). This work used the platforms of the Grenoble Instruct-ERIC centre (ISBG; UAR 3518 CNRS-CEA-UGA-EMBL) within the Grenoble Partnership for Structural Biology (PSB), supported by FRISBI (ANR-10-INBS-0005-02) and GRAL, financed within the University Grenoble Alpes graduate school (Ecoles Universitaires de Recherche) CBH-EURGS (ANR-17-EURE-0003).

### Author contributions

J.K., T.R. and B.dM. designed the project. Acquisition and analysis of the structural and biochemical studies were by A.J.-M. and J.K. with some complementary experiments by H.L. A.N. performed most Y2H and mouse experiments. T.R. and B.dM. performed some image acquisition and analysis and interpreted the data. C.B. performed immunocytochemistry. B.D. performed immunoprecipitations experiments. C.G.

performed DMC1 Chip-Seq. DMC1 SSDS was analyzed by J.C. H.M.B. analyzed the conservation of TOPOVIBL and identified the conserved motif. H.M.B. and J.K. generated the Alfold2 models. B.dM. wrote the first draft of the manuscript; B.dM., T.R. and J.K. revised the manuscript.

### Competing interests

The authors declare no competing interests.

### Additional information

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41467-022-34799-0>.

**Correspondence** and requests for materials should be addressed to Jan Kadlec, Thomas Robert or Bernard de Massy.

**Peer review information** *Nature Communications* thanks John Weir and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

**Reprints and permissions information** is available at <http://www.nature.com/reprints>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

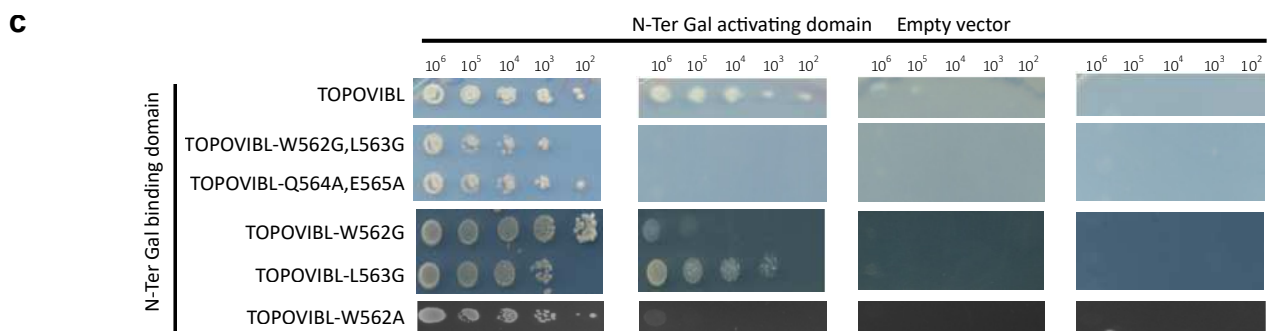
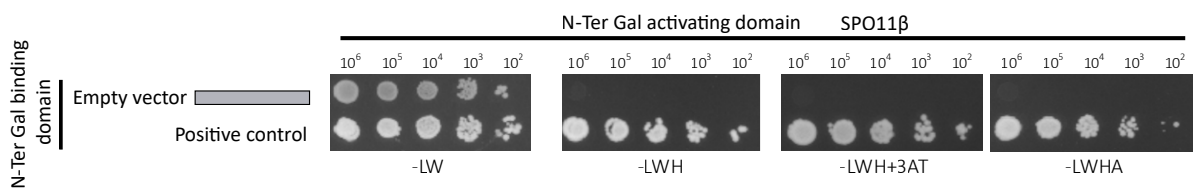
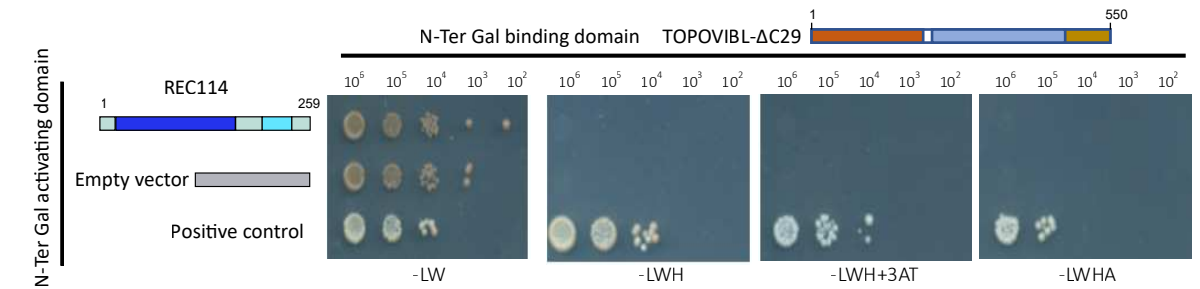
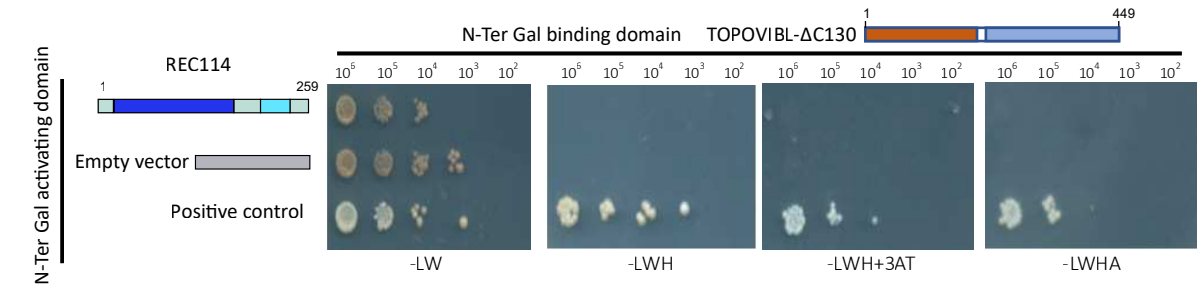
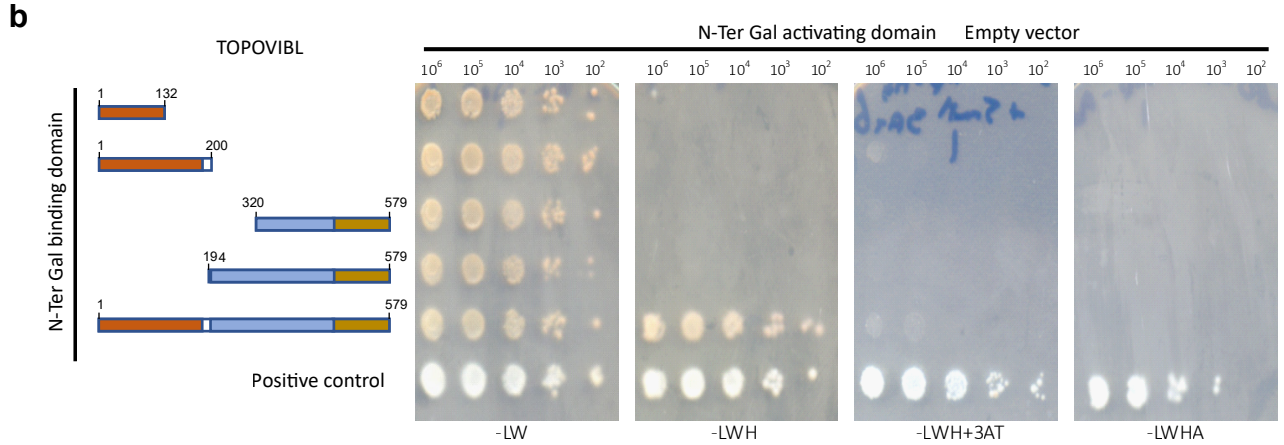
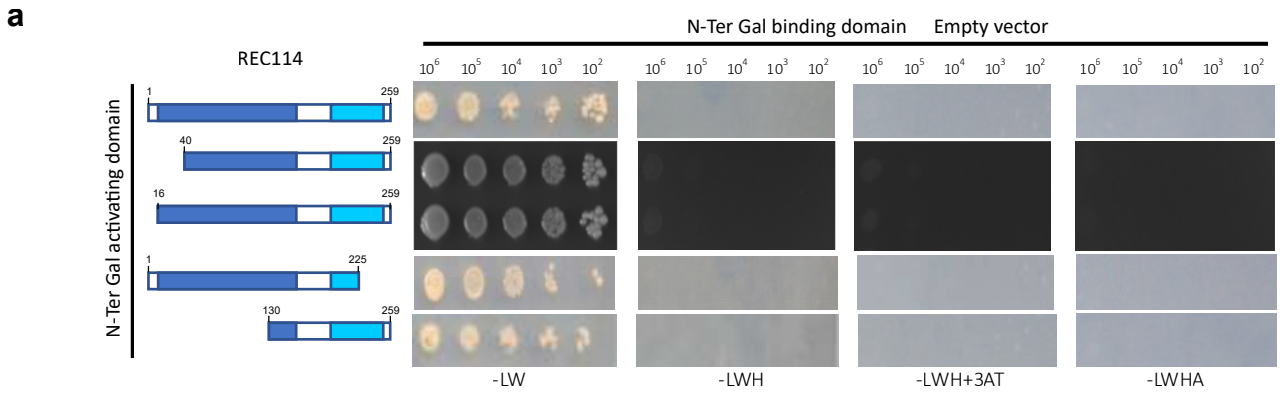
**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022

**Supplementary data**

**Supplementary Figures 1-16**

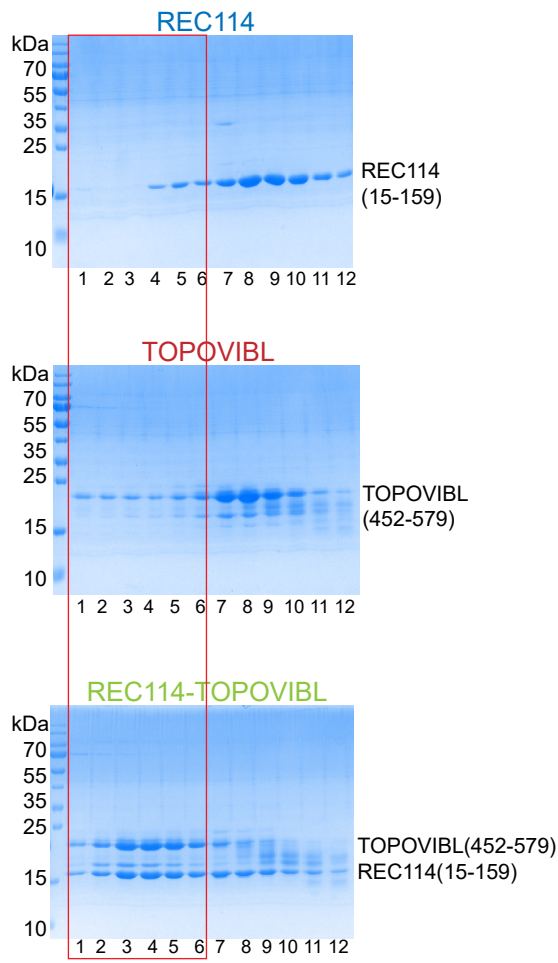
**Supplementary Tables 1-4**



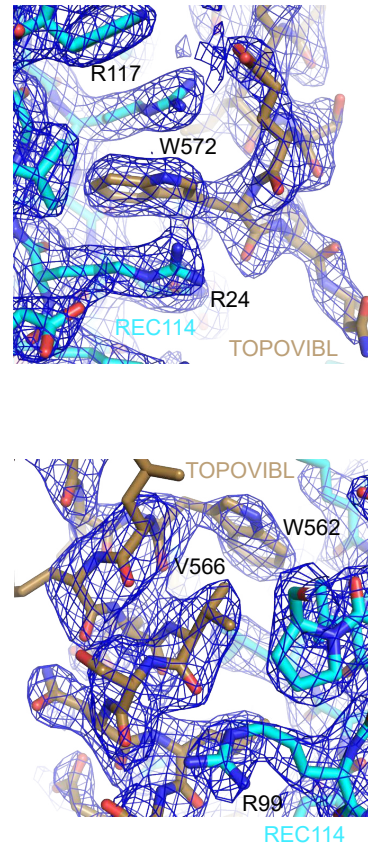
**Supplementary Figure 1. Controls for Y2H**

- a.** Negative controls for Y2H assays for Fig. 1b
- b.** Negative controls for Y2H assays for Fig. 1c
- c.** Negative controls for Y2H assays for Fig. 2h.

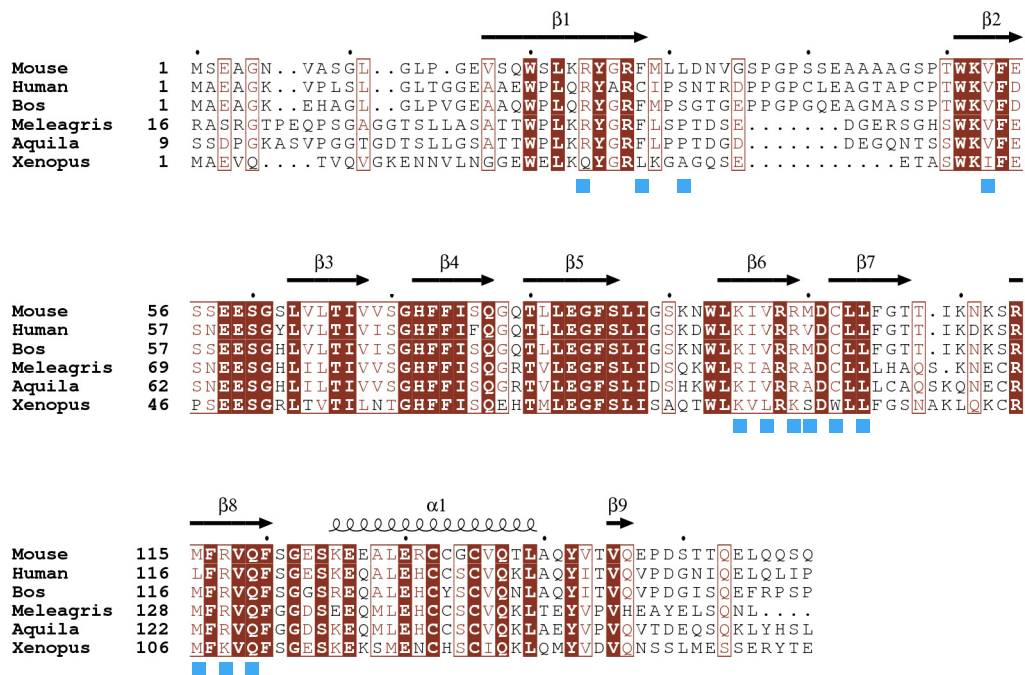
a



b



c



## Supplementary Figure 2. Properties of the REC114-TOPOVIBL complex

- a.** SDS-PAGE analysis of fractions 1-12 of the Superdex 200 gel filtration elution profiles of REC114<sup>15-</sup><sup>159</sup>, TOPOVIBL<sup>452-579</sup> and their complex shown in Fig. 1e. The red rectangle highlights fractions containing the REC114-TOPOVIBL complex. Source data are provided as a Source Data file.
- b.** Representative part of the *2Fo-Fc* electron density map covering the TOPOVIBL-REC114 interface contoured at 1.0  $\sigma$ .
- c.** Sequence alignment of the REC114 PH domain from *M. musculus*, *H. sapiens*, *B. mutus*, *Meleagris gallopavo*, *Aquila chrysaetos canadensis* and *Xenopus laevis* used to highlight the amino-acid conservation on the structure (Fig. 2c). Brown letters: equivalent amino acids; white letters: identical amino acids. The residues interacting with TOPOVIBL are shown as blue squares.

*Mus\_musculus/396-579*      <sup>a5</sup>      <sup>a6</sup>      <sup>a7</sup>

1      10      20      30      40      50      60

*Mus\_musculus/396-579*      **L**SVAVNS**L**MSV**L**T**G**S**T**RS**S****R**K**T**C**L**Q**A**L**E**A**D**T**Q**E**F**G**V**K**L**H**R**I**F**Y**D**T**T**Q**H**Q**F**L**K**H**C**S**D**T

*HomO\_sapiens/395-577*      **L**SVAVNS**L**MS**L**T**G**S**T**RS**S****R**K**M**C**L**Q**T**L**Q**A**A**D**D**Q**E**F**R**T**K**L**H**V**F**R**E**T**T**Q**H**Q**F**L**H**H**C**S**C**E**V**

*Chelonia\_mydas/398-611*      **R**P**V**L**G**A**L**S**S**V**V**S**S**T**N**S**Q**E**R**R**E**C**L**Q**S**L**Q**V**S**D**T**F**E**L**M**A**A**L**Q**T**A**F**E**N**V**T**R**K**R**L**M**P**N**G**T**C**S**I

*Rhinatrema\_bivittatum/70-211*      **Q**S**L**V**T**D**C**I**C**S**I**V**T**S**S**A**D**S**G**E**R**E**R**C**L**S**D**M**Q**V**T**D**T**Q**G**F**R**T**V**Q**D**S**L**H**R**T**A**L**N**R**L**L**S**R**T**C**R**N**L**

*Danio\_rerio/391-562*      **I**P**I**V**L**S**S**V**N**S**I**A**G**S**S**N**G**E**R**T**A**C**V**D**S**M**R**V**K**S**T**C**E**L**Q**T**S**L**R**Q**S**L**S**I**V**D**G**R**F**T**P**R**T**K**R**C**R**K

*Latimeria\_chalumnae/391-634*      **I**P**V**M**I**S**S****L**I**T**V**V**S**S**T**N**R**D**E**R**K**K**C**L**Q**D**F**Q**V**G**D**T**Q**E**L**M**E**A**V**G**N**S**L**Q**R**V**T**Q**N**R**F**L**T**S**R**T**C**G**I

*Aplysia\_californica/466-956*      **I**P**D**I**A**N**S****L**T**D**M**F**I**R****N**T**E**R**I**N**V**G**H**L**F**E**N**R**A**T**M**S**K**H**E**K**L**W**N**L**G**.....

*Patiria\_minitata/30-206*      **V**A**N**T**E**A**N****A**T**V**S**S****T**N**D**E**R**Q**K**F**S**L**K****V**P**S**S**G**L**V**S**S**L**P**S**Q**L**H**S**L**A**C**R**H**M**G**V**T**L**C**A**E**S

*Nematostella\_vectensis/21-315*      **V**P**A**I**S**M**A**L**A**S**I**V**S****S**K**N**K**E**Q**Q**R**C**L**E**I**Q**E**T**R**V**D**R**I**E**K**L**N**T**L**K**W**L**A**R**Q**K**F**G**I**S**.....

*Mus\_musculus/396-579*      70      80      90

*Mus\_musculus/396-579*      E**Q**H**L**T**F**E**K**N**I**S**A**Q**N**T**K**D**Q**H.....**K**N**I**A**Q**E**F**F**E**S.....

*HomO\_sapiens/395-577*      K**Q**O**L**T**L**E**K**K**D**S**A**Q**T**E**D**A**P**.....**D**N**S**L**L**E**L**L**A**D**T**.....

*Chelonia\_mydas/398-611*      T**K**O**V**P**E**D**A**S**P**A**W**L**A**A**C**D**L**P**T**A**G**K**L**S**Q**E**E**A**S**D**V**.....

*Rhinatrema\_bivittatum/70-211*      R**K**E**L**L**E**K**F**N**V**N**Q**L**T**P**C**T**Y**E**A**N**R**S**Q**S**Q**G**N**A**S**C**I**.....

*Danio\_rerio/391-562*      E**K**K**V**T**Q**A**V**C**S**Q**K**R**S**P**D**I**L**E**A**F**E**N**P**L**L**S**S**K**R**I.....

*Latimeria\_chalumnae/391-634*      R**T**S**F**V**D**E**N**K**Q**T**Q**S**T**D**S**Q.....**D**R**L**M**E**D**S**F**E**T**I**C**T**S**I**S**L**R**K**H**G**L**S**L**P**M**D**S**S**K**E**D.....

*Aplysia\_californica/466-956*      .....**C**H**E**S**N**A**D**E**D**T**D**A**K**P**L**G**N**Q**D**E**K**E**I**C**V**D**E**.....**W**S**D**S**Q**F**S**Q**E**R**H**E**D**R**R**N**D**

*Patiria\_minitata/30-206*      Q**V**O**E**F**S**K**S**S**E**K**A**M**V**T**S**Q**P**A**S**Q**C**L**Q**H**I**F**A**S**Q**.....**P**F**G**A**D**D**S**L**E**D**E**F**D**

*Nematostella\_vectensis/21-315*      ..**Q**E**A**S**E**G**E**Q**V**A**P**E**I**S**Q**A**S**V**G**S**D**T**G**K**T**P**T**Q.....**T**S**I**S**S**S**Y**F**D**L**S**D.....

*Mus\_musculus/396-579*      100      110

*Mus\_musculus/396-579*      .....**I**G**Q**A**E****N**R**K**R**K**R**G**S**P**N**H**G.....**R**E**S**R**V**L**G**S**A**R.....

*HomO\_sapiens/395-577*      .....**S**Q**Q**A**E****N**K**R**L**K**R**G**S**F**R.....**T**E**M**R**A**L**R**S**A**R.....

*Chelonia\_mydas/398-611*      .....**S**I**S**S**L****N**S**E**P**A**.....**S**L**L**D**R**V**D**L**G**W**E**I**S**E**T**R**R**R**T**A.....

*Rhinatrema\_bivittatum/70-211*      .....**I**V**E**S**D**T**G**E**F**V**V**P**S**D**N**R**E**I**R**T**A**A**P**K**G**Q**P**D**S**T**E**Y**I**Q**N**Y**A**E**T**G**D**P**C**R**R**E.....

*Danio\_rerio/391-562*      F**Q**F**G**M**L**G**L**P**N**E**N**E**Q**K**G****N**Q**Y**S**R**.....**H**Y**L**Q**N**D**F**Q**N**D**S**H**S**A**Q**N**L**I**N**D**V**G**G**Q**F**

*Latimeria\_chalumnae/391-634*      F**S**Q**L**D**F**I**D**G**E**G**V**T**G****N**L**S**S.....**Q**L**F**D**E**A**K**S**A**G**T**A**T**S**G**M**E**.....

*Aplysia\_californica/466-956*      .....**Y**E**A**D**E**I**C**N**M**S**L**Y**D**S**Q**E**F**R.....**Q**Y**T**Q**S**Q**A**T**Q**D**K**L**P**L**N**T**Q**E.....**K**N**T**

*Patiria\_minitata/30-206*      .....**Y**E**A**D**E**I**C**N**M**S**L**Y**D**S**Q**E**F**R.....**Q**Y**T**Q**S**Q**A**T**Q**D**K**L**P**L**N**T**Q**E.....**K**N**T**

*Nematostella\_vectensis/21-315*      .....**Y**E**A**D**E**I**C**N**M**S**L**Y**D**S**Q**E**F**R.....**Q**Y**T**Q**S**Q**A**T**Q**D**K**L**P**L**N**T**Q**E.....**K**N**T**

*Mus\_musculus/396-579*      120      130      140

*Mus\_musculus/396-579*      .....**D**R**S**F**P**K**S**A**T**R**D**R.....**E**L**T**E**V**S**L**T**A**R**G**S**Q**.....

*HomO\_sapiens/395-577*      .....**T**I**E**N**N**A**G**E**A**L**E**Q**E**P**R**G**L**Q**E**R**T**D**N**L**M****D**F**G**S**Q**L**S**.....**F**C**L**W**S**P.....

*Chelonia\_mydas/398-611*      .....**R**E**D**.....**L**S**D**N**G**S**T**.....

*Rhinatrema\_bivittatum/70-211*      .....**H**C**D**L**E**S**S**F**I**S**E**K**R**Q**C**T**E**P**F**S**E**D**S**D**T**N**E**T**S**D**F**Q**K**S**P**.....**F**S**L**V**H**L.....

*Danio\_rerio/391-562*      .....**E**A**R**S**G**S**G**A**K**R**E**R.....**R**G**D**S**F**P.....**L**P**V**C**R**R.....

*Latimeria\_chalumnae/391-634*      C**Q**L**S**E**Q**N**E**Q**Q**P**P**V**E**D**P**K**D**D**S**E**L**D**V**L**S**L**Q**S**D**V**O**S**E**D**E**R**L**P**I**D**S**P**F**S**P**D**S**L**C**V**S**S**P**C**Y**E**D**T**E**

*Aplysia\_californica/466-956*      C**I**T**S**E**P**S**E**L**G**I**D**I**P**L**P**N**N**T**T**Q**R**K**T**Q**E**V**N**S**G**I**T**S**E**.....**L**G**D**K**P**.....

*Patiria\_minitata/30-206*      C**I**T**S**E**P**S**E**L**G**I**D**I**P**L**P**N**N**T**T**Q**R**K**T**Q**E**V**N**S**G**I**T**S**E**.....**L**G**D**K**P**.....

*Nematostella\_vectensis/21-315*      C**I**T**S**E**P**S**E**L**G**I**D**I**P**L**P**N**N**T**T**Q**R**K**T**Q**E**V**N**S**G**I**T**S**E**.....**L**G**D**K**P**.....

*Mus\_musculus/396-579*      T**Q**A**A**H**G**.....

*Mus\_musculus/396-579*      .....**H**R**A**H**G**.....

*HomO\_sapiens/395-577*      .....**M**F**L**S**E**R**E**G**F**.....

*Chelonia\_mydas/398-611*      .....**S**N**S**I**H**M**S**Q**I**.....

*Rhinatrema\_bivittatum/70-211*      .....**P**Q**I**Q**A**R.....**M**.....

*Danio\_rerio/391-562*      .....**H**R**D**R**E**A**A**.....**Q**P**H**A**E**A**G**S**C**R**A**R**T**.....

*Latimeria\_chalumnae/391-634*      L**C**D**D**P**L**I**E**L**S**V**V**E**S**N**S**H**V**H**M**R**N**S**P**A**E**T**D**A**G**L**E**P**F**V**L**P**H**E**T**A**D**A**G**K**A**L**L**S**L**F**D**N**S**L**T**Q

*Aplysia\_californica/466-956*      .....**L**R**N**N.....

*Patiria\_minitata/30-206*      .....**I**P**Q**N**T**Q**E**N**T**C**F**S.....

*Nematostella\_vectensis/21-315*      .....**I**P**Q**N**T**Q**E**N**T**C**F**S.....

*Mus\_musculus/396-579*      A**P**.....

*Mus\_musculus/396-579*      H**T**L**Q**N**D**H**K**E**T**L**R**Q**T**N**V**A**D**Q**S**I**E**I**S**D**S**E**D**D**S**P**V**T**K**A**K**N**V**P**L**E**P**H**V**S**I**I**K**T**C**K**L**D**Q**P**A**H**F**

*HomO\_sapiens/395-577*      .....**E**P.....**G**L**M**G**E**C.....**F**

*Chelonia\_mydas/398-611*      .....**E**P.....**G**L**M**G**E**C.....**F**

*Rhinatrema\_bivittatum/70-211*      .....**E**P.....**G**L**M**G**E**C.....**F**

*Danio\_rerio/391-562*      .....**E**P.....**G**L**M**G**E**C.....**F**

*Latimeria\_chalumnae/391-634*      .....**E**P.....**G**L**M**G**E**C.....**F**

*Aplysia\_californica/466-956*      M**S**D**R**D**Q**C**T**E**P**L**Q**S**E**Q**E**K**N**S**D**E**H**C**E**T**P**K**S**L**S**V**V**C**T**D**F**E**K**R**C**P**F**R**K**S**S**E**Y**I.....

*Patiria\_minitata/30-206*      G**S**.....**Q**E**T**E**D**.....**F**L**S**E**V**V**L**E**F**.....**S**E**Y**I**S**K**N**S**Q**

*Nematostella\_vectensis/21-315*      G**S**.....**Q**E**T**E**D**.....**F**L**S**E**V**V**L**E**F**.....**S**E**Y**I**S**K**N**S**Q**

*Mus\_musculus/396-579*      150      160      000000000

*Mus\_musculus/396-579*      .....**R**A**Q**A**E**A**A**S**P**A**G**.....**G**L**E**D**L**.....

*HomO\_sapiens/395-577*      .....**R**A**L**A**P**G**R**A**S**L**G**S.....**R**L**E**D**V**L.....

*Chelonia\_mydas/398-611*      .....**T**A**V**S**E**D**G**A**R**S**S**E.....**N**L**W**Q**G**H**L**.....

*Rhinatrema\_bivittatum/70-211*      .....**Y**D**F**F**D**G**L**Q**Q**.....**Q**E**K**L.....

*Danio\_rerio/391-562*      .....**T**E**D**K**P**S**V**K**L**I**N**E.....**Q**V**E**E**Q**.....

*Latimeria\_chalumnae/391-634*      .....**C**A**A**G**G**S**R**S**D**H**G**S.....**V**H**Q**E**D**D**W**.....

*Aplysia\_californica/466-956*      .....**D**D**I**T**D**V**M**R**S**C**S**V**E**S**T**S**S**E**I**C**I**D**A**K**R**F**K**P**R**G**E**K**S**P**D**L**V**C**N**D**V**L**D**E**L**Q**F**S**Q**D**S**L**N**F**S**E**D**

*Patiria\_minitata/30-206*      .....**S**V**K**S**F**E**V**S**Q**S**K**S.....**M**L**E**D**E**.....

*Nematostella\_vectensis/21-315*      A**D**P.....**T**G**L**E**F**H**T**E**K**T**D**E.....**N**L**L**K**E**D**D**L**E**.....**T**K**D**A**L**H.....

*Mus\_musculus/396-579*      170      180

*Mus\_musculus/396-579*      .....**W**L**Q**E**V**S**N**L**S**E**W**L**N**F**G**H**R**S.....

*HomO\_sapiens/395-577*      .....**W**L**Q**E**V**S**N**L**S**E**W**L**S**P**S**P**G**P.....

*Chelonia\_mydas/398-611*      .....**W**L**Q**E**V**S**N**L**S**E**W**L**S**M**S**.....

*Rhinatrema\_bivittatum/70-211*      .....**W**L**Q**E**V**S**K**L**S**D**H**T**C**.....

*Danio\_rerio/391-562*      .....**W**L**Q**E**N**F**S**E**L**.....**D**.....

*Latimeria\_chalumnae/391-634*      .....**W**L**Q**E**V**L**R**L**S**E**W**T**T**.....

*Aplysia\_californica/466-956*      G**E**E**D**V**L**K**H**W**Q**R**L**S**P**L**S**G**Q**L**E**L**N**A**E**H**R**S**K**T**S**Q**L**D**S**E**D**W**L**N**E**E**F**S**K**W**Q**A**N**.....**M**T

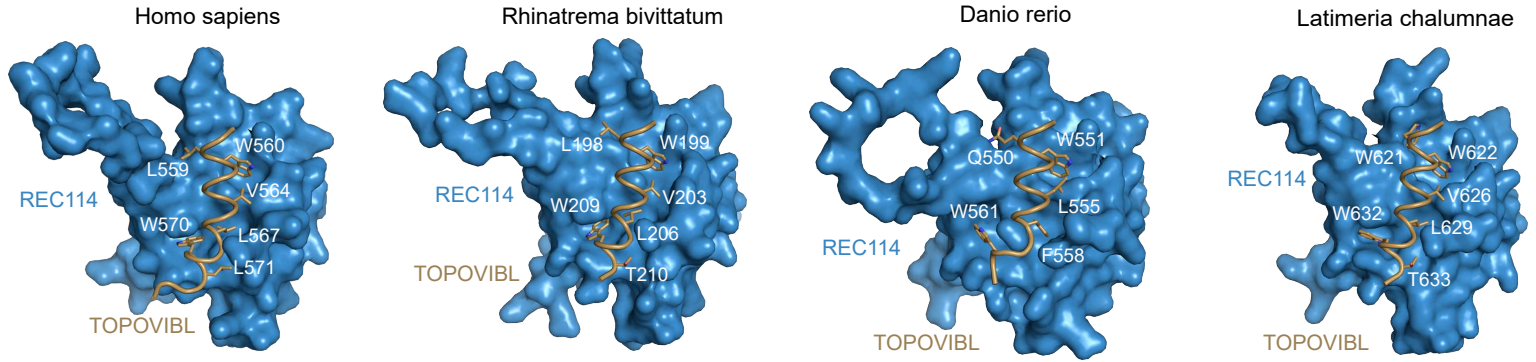
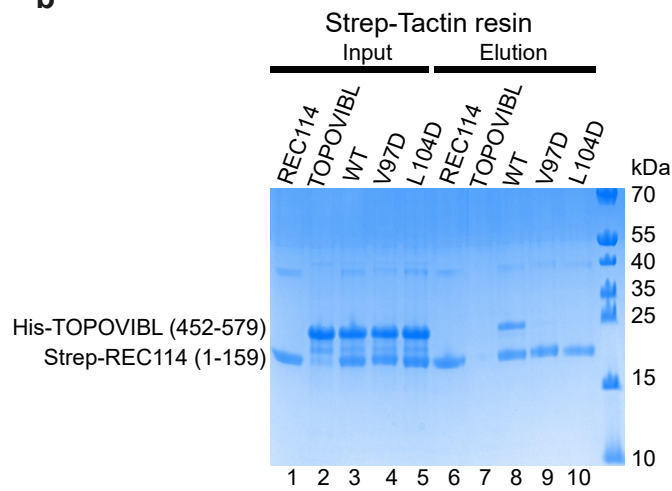
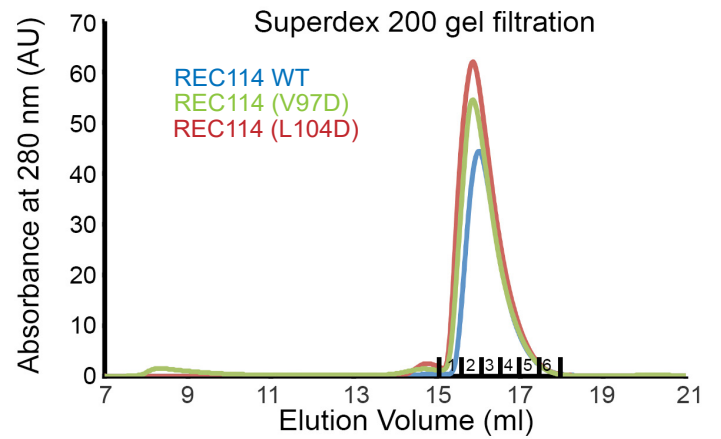
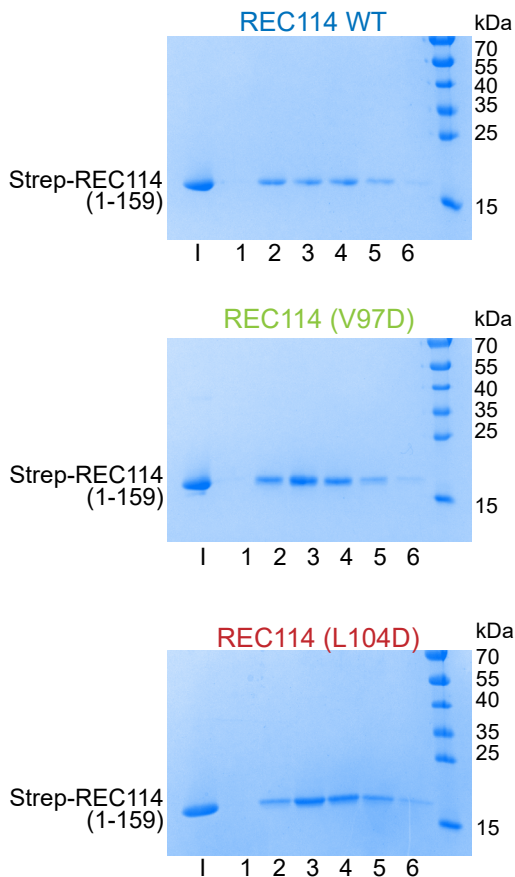
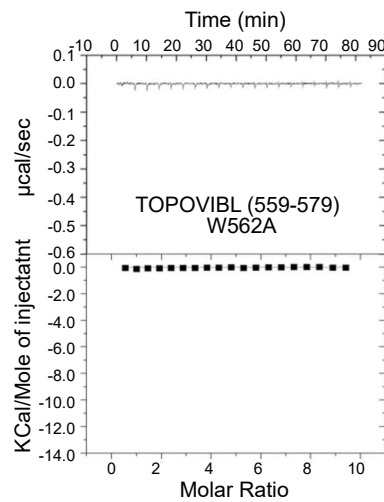
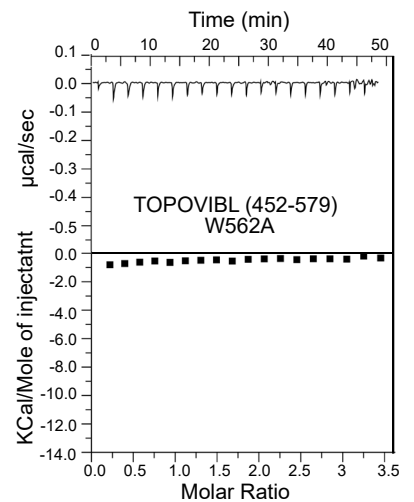
*Patiria\_minitata/30-206*      .....**W**L**D**E**V**L**D**S**D**P**F**.....**S**.....

*Nematostella\_vectensis/21-315*      .....**K**D**V**L**N**.....**E**E**R**G**I**E**K**E.....**N**S**F**D**W**L**E**F**E**N**N****F****A**V**N**.....

### **Supplementary Figure 3. Conservation of TOPOVIBL C-terminal region**

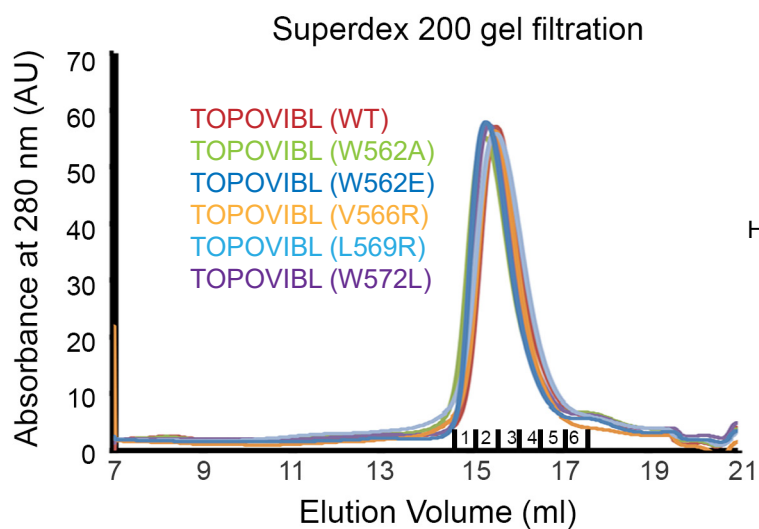
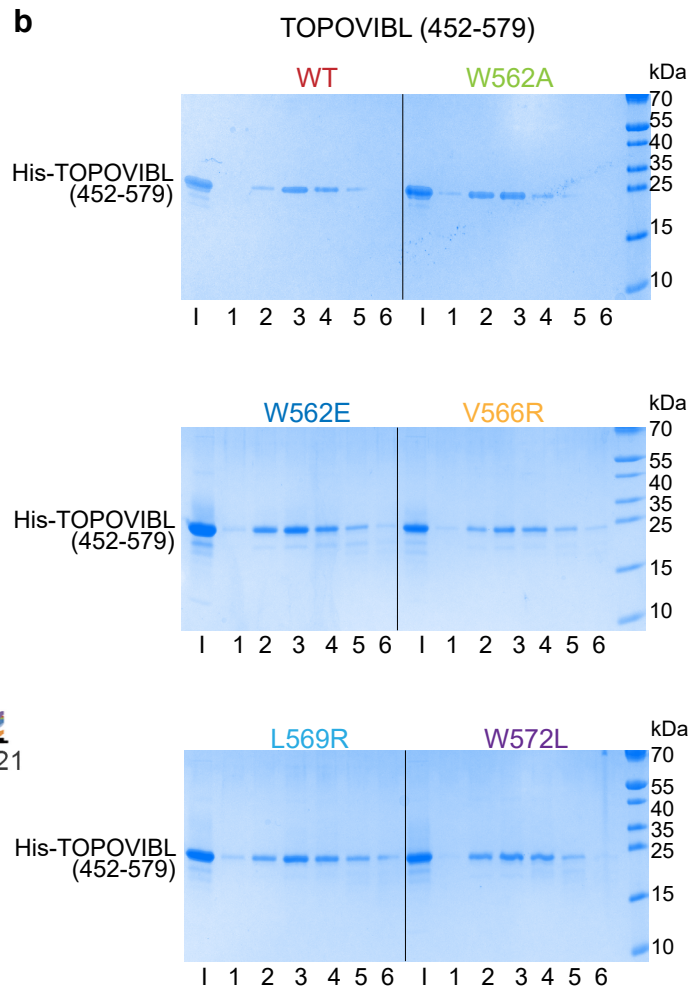
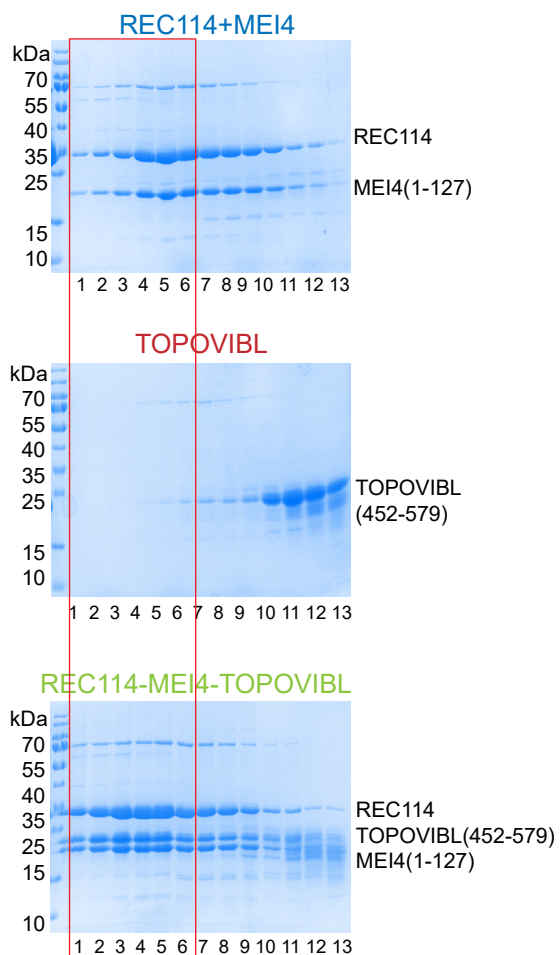
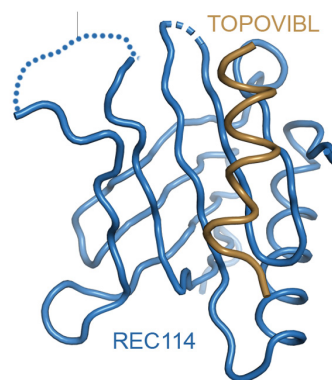
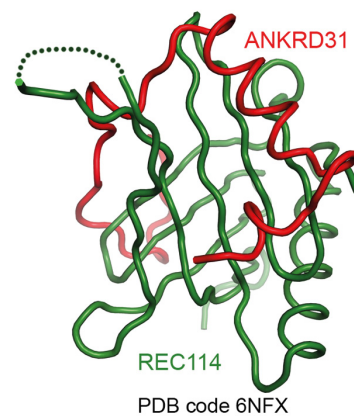
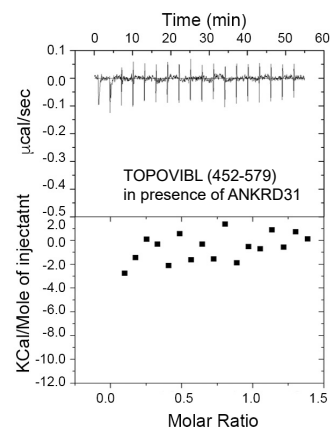
Alignment of TOPOVIBL C-terminal region from nine metazoan species (as in Fig. 1f). The region aligned starts from the three conserved helices of the transducer domain ( $\alpha 5$ ,  $\alpha 6$  and  $\alpha 7$ ) up to the C-terminal end. The position of S452 in *M. musculus* TOPOVIBL which defines the N-terminal boarder of the protein used in biochemical assays (452-579) is shown by an arrow. Brown letters: equivalent amino acids; white letters: identical amino acids.



**a****b****c****d****e****f**

#### Supplementary Figure 4. Properties of the TOPOVIBL-REC114 complex and mutants

- a.** The TOPOVIBL-REC114 complex models predicted by AlphaFold2<sup>29</sup>. The structure of the REC114 PH domain is shown as surface in blue. Key interacting residues of TOPOVIBL are shown as sticks. Most of the interactions observed in the mouse crystal structure are conserved in the predicted complexes of other species. In *D. rerio*, the residue corresponding to mouse Val566 is Leu555. This mutation seems to be compensated by the valine to glutamine mutation at position 92 in REC114.
- b.** Pull-down experiments of Strep-tagged REC114<sup>1-159</sup> mutants indicated above the lanes with TOPOVIBL<sup>452-579</sup>. All proteins were first purified by affinity chromatography and gel filtration. A total of 1.2% of the input (lanes 1–5) and 1.2% of the eluates (lanes 6–10) were analyzed on 12% SDS-PAGE gels stained with coomassie brilliant blue. Control lanes 1-2 and 6-7 show inputs and elutions of REC114 and TOPOVIBL alone, respectively.
- c.** Superdex 200 gel filtration elution profiles of WT and mutant Strep-REC114<sup>1-159</sup>. The elution profiles are similar, indicating that the REC114 mutations do not significantly affect the overall structure of this REC114 fragment.
- d.** SDS-PAGE analysis of fractions 1-6 of the Superdex 200 gel filtration elution profiles shown in **c**. I indicates input samples loaded onto the column.
- e.** ITC measurement of the interaction affinity between REC114<sup>15-159</sup> and the TOPOVIBL<sup>559-576</sup> peptide harbouring the W562A mutation (the wild-type control is in Fig. 1g).
- f.** ITC measurement of the interaction affinity between REC114<sup>1-159</sup> and TOPOVIBL<sup>452-579</sup> (W562A) (the wild-type control is in Fig. 1g). Source data are provided as a Source Data file.

**a****b****c****d****e****f**

### **Supplementary Figure 5. Properties of the TOPOVIBL-REC114-MEI4 complex and mutants**

- a.** Superdex 200 gel filtration elution profiles of WT and mutant TOPOVIBL<sup>452-579</sup>. The elution profiles are similar, indicating that the TOPOVIBL mutations do not significantly affect the overall structure of this TOPOVIBL fragment.
- b.** SDS-PAGE analysis of fractions 1-6 of the Superdex 200 gel filtration elution profiles shown in **a**. I indicates input samples loaded onto the column.
- c.** SDS-PAGE analysis of fractions 1-13 of the Superose 6 gel filtration elution profiles shown in Fig. 2i. The red rectangle highlights fractions of containing the TOPOVIBL-REC114-MEI4 complex.
- d. and e.** Comparison of the REC114-TOPOVIBL (d) and the REC114-ANKRD31 (e) structures.
- f.** ITC measurement of the interaction between REC114<sup>15-159</sup> and TOPOVIBL<sup>452-579</sup> shows the absence of binding when REC114 is pre-saturated with ANKRD31<sup>1808-1857</sup>. Source data are provided as a Source Data file.

**a**

CTGGAGGACTTGTGGCTACAGGAGGTATCCAATCTGTCCGAGTGGCTGAATCCGGGACACAGGTCC TAA



HR, allele *Top6b*<sup>W562A</sup>

4 bp deletion,  
allele *Top6b*<sup>Δ17Ct</sup>

CTGGAGGACTTGTgCCTgCAGGAGGTATCCAATCTGTCCGAGTGGCTGAATCCGGGACACAGGTCC TAA

CTGGAGGACTTGTGGCTACAGGAGGTATCCAATCTGTCCGAGTGGCTGAATCCGGGACACAGGTCC TAA

**b**

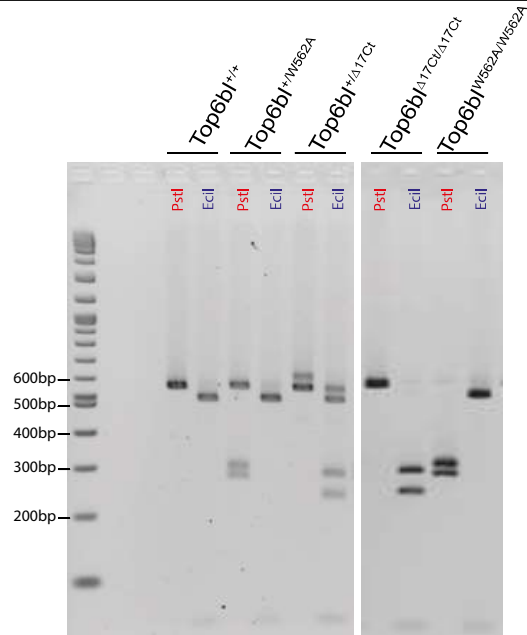
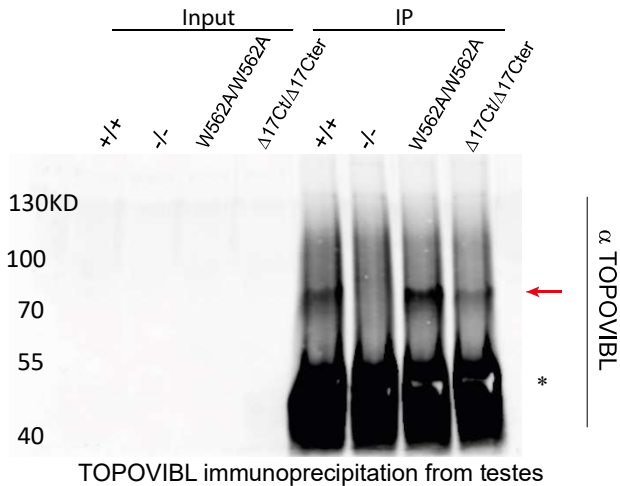
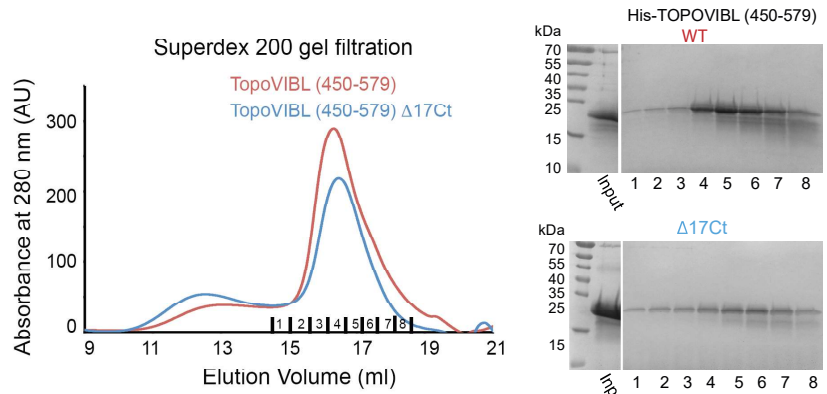
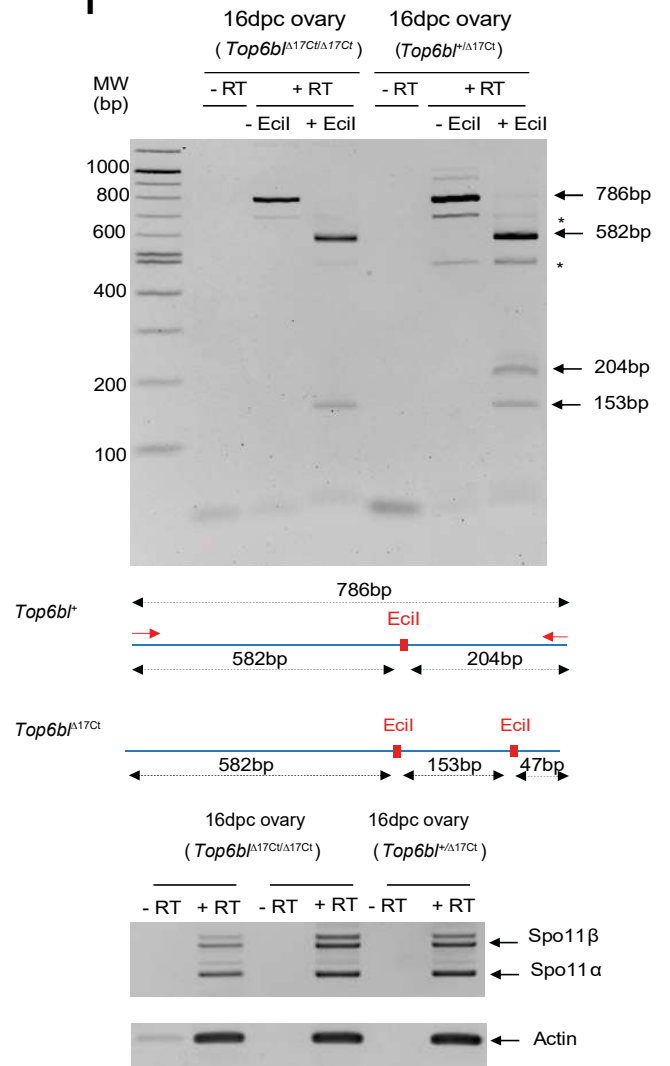
TOPOVIBL ... LEDLWLQEVSNLSEWLNPGHRS\*

TOPOVIBL-W562A ... LEDLALQEVSNLSEWLNPGHRS\*

TOPOVIBL-Δ17Ct ... LEDLWRRYPICPSG\*

**c**

	<i>Top6bl</i>	<i>Top6bl</i> <sup>W562A</sup>	<i>Top6bl</i> <sup>Δ17Ct</sup>
PstI	569 bp	294 + 275 bp	569 bp
EciI	517 + 52 bp	517 + 52 bp	280 + 233 + 52 bp

**e****d****f**

**Supplementary Figure 6. Generation and characterization of mice carrying *Top6bl*<sup>W562A</sup> and *Top6bl*<sup>Δ17Ct</sup> alleles.**

**a.** *Top6bl*, *Top6bl*<sup>W562A</sup> and *Top6bl*<sup>Δ17Ct</sup> cDNA sequences. The top sequence corresponds to the 65 last nucleotides of the *Top6bl* open reading frame. Orange and blue letters mark exons 15 and 16, respectively. The TGG codon in green corresponds to the W562 residue in the protein sequence. Red arrows mark the position of the predicted CRISPR-Cas9 cleavage sites, corresponding to the three guide RNAs used. TAA in black: original stop codon.

The middle sequence corresponds to the repair product by homologous recombination (HR) following CRISPR-Cas9-mediated DSB formation. The TGG codon is replaced by the GCG codon (red letters). The G residue in purple corresponds to a silent substitution to create a *PstI* restriction site for genotyping. TAA in black: original stop codon.

The bottom sequence corresponds to the product of the deletion of four nucleotides in the sequence following DSB formation and repair (grey crossed letters). TGA in black: new stop codon due to frameshift.

**b.** Protein sequences of the C-terminal part of TOPOVIBL, TOPOVIBL-W562A and TOPOVIBL-Δ17Ct. Blue letters correspond to the original residues, bold red A to the mutant substitution of W562 in the TOPOVIBL-W562A sequence, and red letters to the nine-residue substitution in the TOPOVIBL-Δ17Ct sequence. \*: Stop.

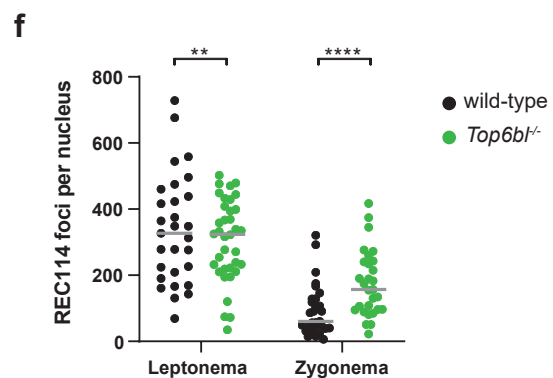
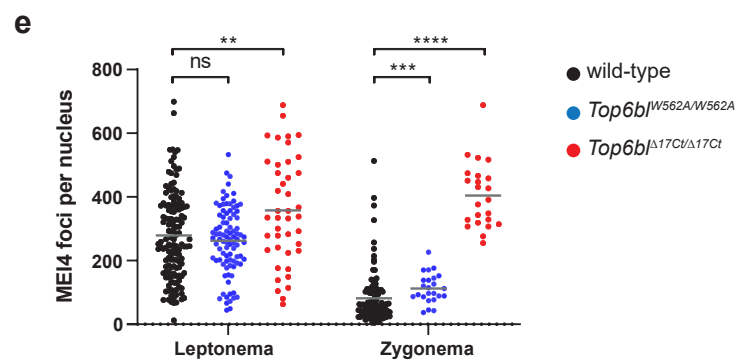
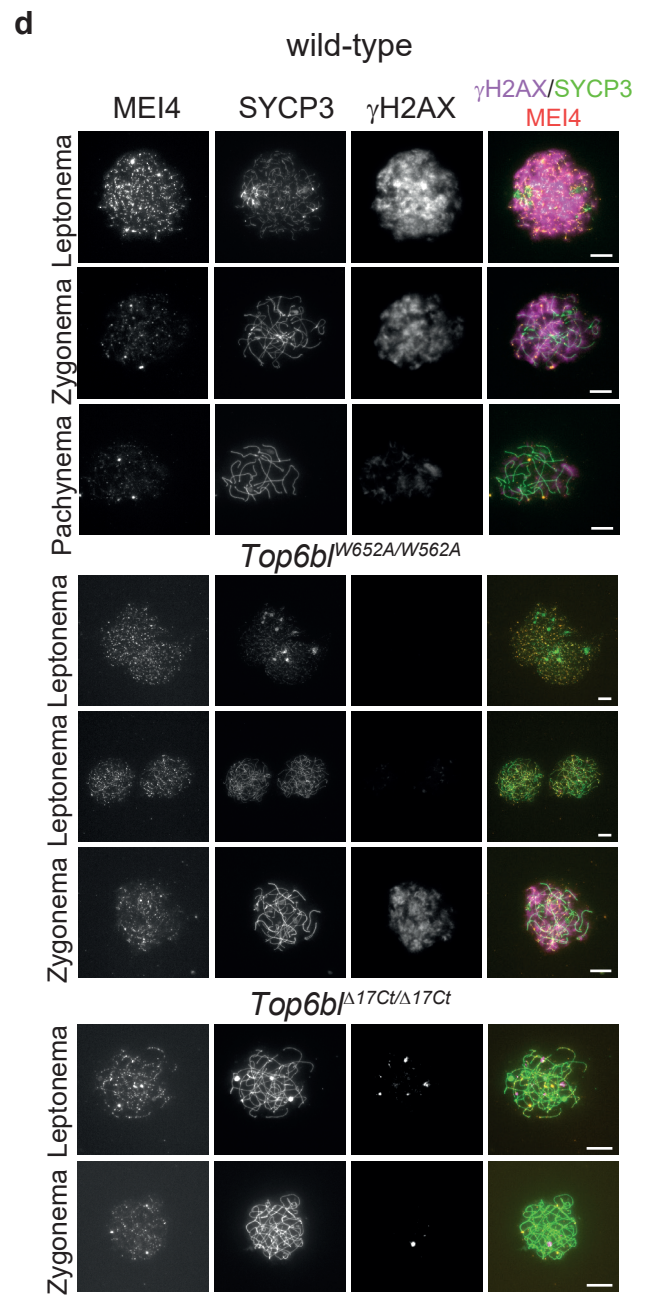
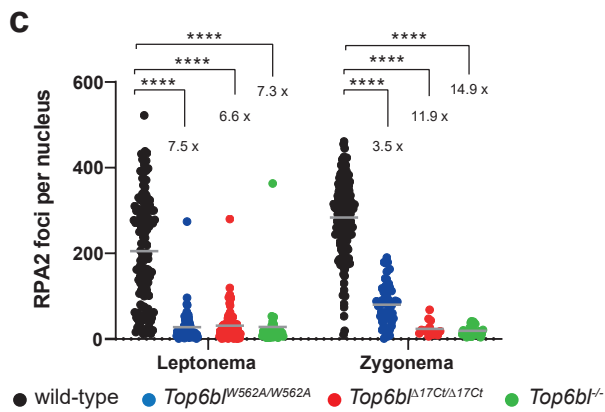
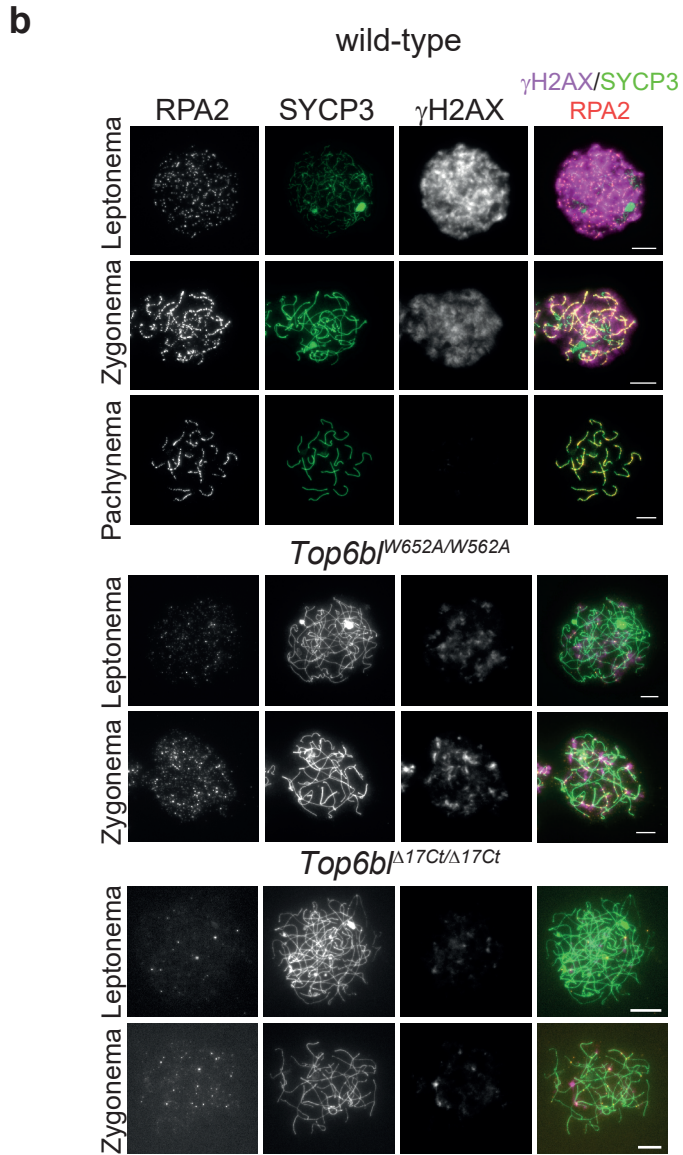
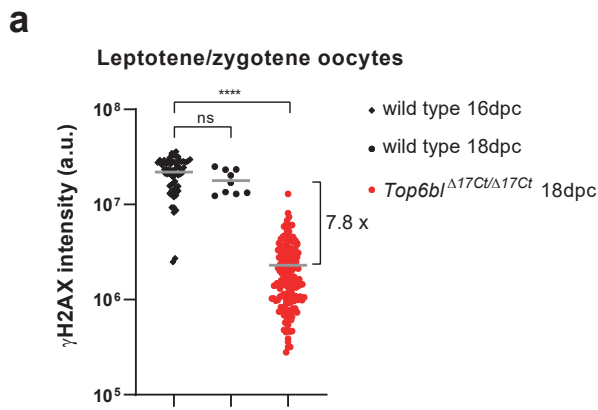
**c.** Genotyping strategy. Following DNA extraction from tails of *Top6bl*<sup>+/+</sup>, *Top6bl*<sup>+/W562A</sup>, *Top6bl*<sup>+/Δ17Ct</sup>, *Top6bl*<sup>W562A/W562A</sup> and *Top6bl*<sup>Δ17Ct/Δ17Ct</sup> mice, PCR amplification with the primer pair specific for a 569 bp region containing the mutations was performed. PCR products were digested with *PstI* or *EciI*. The top table presents the expected size following digestion of the different alleles (red correspond to *PstI* digestion, and dark blue to *EciI* digestion). Bottom panel: DNA fragments separated on agarose gels.

**d.** Left panel: Overlay of the Superdex 200 gel filtration elution profiles of wild-type TOPOVIBL<sup>452-579</sup> and TOPOVIBL<sup>52-579</sup> Δ17Ct. Both profiles are similar, indicating that the Δ17Ct mutation does not significantly affect the overall structure of this TOPOVIBL fragment. Right panel: SDS-PAGE analysis of the input and fractions 1-8 of the Superdex 200 gel filtration elution profiles. Gels were stained with coomassie brilliant blue.

e. TOPOVIBL immunoprecipitation. Following extraction of total protein from 14 dpp mouse testes obtained from *Top6bl*<sup>+/+</sup>, *Top6bl*<sup>W562A/W562A</sup>, *Top6bl*<sup>Δ17Ct/Δ17Ct</sup> and *Top6bl*<sup>-/-</sup> mice, TOPOVIBL was immunoprecipitated using the homemade anti-TOPOVIBL antibody. Six testes of each genotype were used for one immunoprecipitation. The immunoblot was revealed using the homemade anti-TOPOVIBL antibody. The red arrow shows the TOPOVIBL expected band. \*corresponds to a non-specific band (immunoglobulin heavy chain).

f. Expression of *Top6bl* assessed by RT-PCR. RNA extracted from ovaries of 16 dpc *Top6bl*<sup>Δ17Ct/Δ17Ct</sup> or *Top6bl*<sup>+/Δ17Ct</sup> mice was incubated with or without reverse transcriptase (RT), amplified by PCR and incubated with or without *EciI* restriction enzyme. The size of the fragments after *EciI* digestion are shown for the *Top6bl*<sup>+</sup> and *Top6bl*<sup>Δ17Ct</sup> alleles. Stars indicate non-specific amplification products. Control PCR assays for monitoring the expression of the two *Spo11* splice variants (*Spo11α* and *β*) and of actin are shown. Source data are provided as a Source Data file.







## Supplementary Figure 7. Immunocytochemistry on oocytes

**a.** Quantification of  $\gamma$ H2AX signal intensity in leptotene (L)/zygotene (Z) oocytes from 16 and 18 dpc wild-type and 18dpc *Top6bl* <sup>$\Delta$ 17Ct/ $\Delta$ 17Ct</sup> mice. Number of nuclei: 64, 9, and 167, respectively. Grey bars show the mean values. P values were determined using the two-tailed unpaired Mann-Whitney test. Progression into prophase at 18dpc was: L 46%, Z 54%, Pachynema (P) 0% in *Top6bl* <sup>$\Delta$ 17Ct/ $\Delta$ 17Ct</sup>, and L 2%, Z 8%, P 90% in wild-type oocytes.

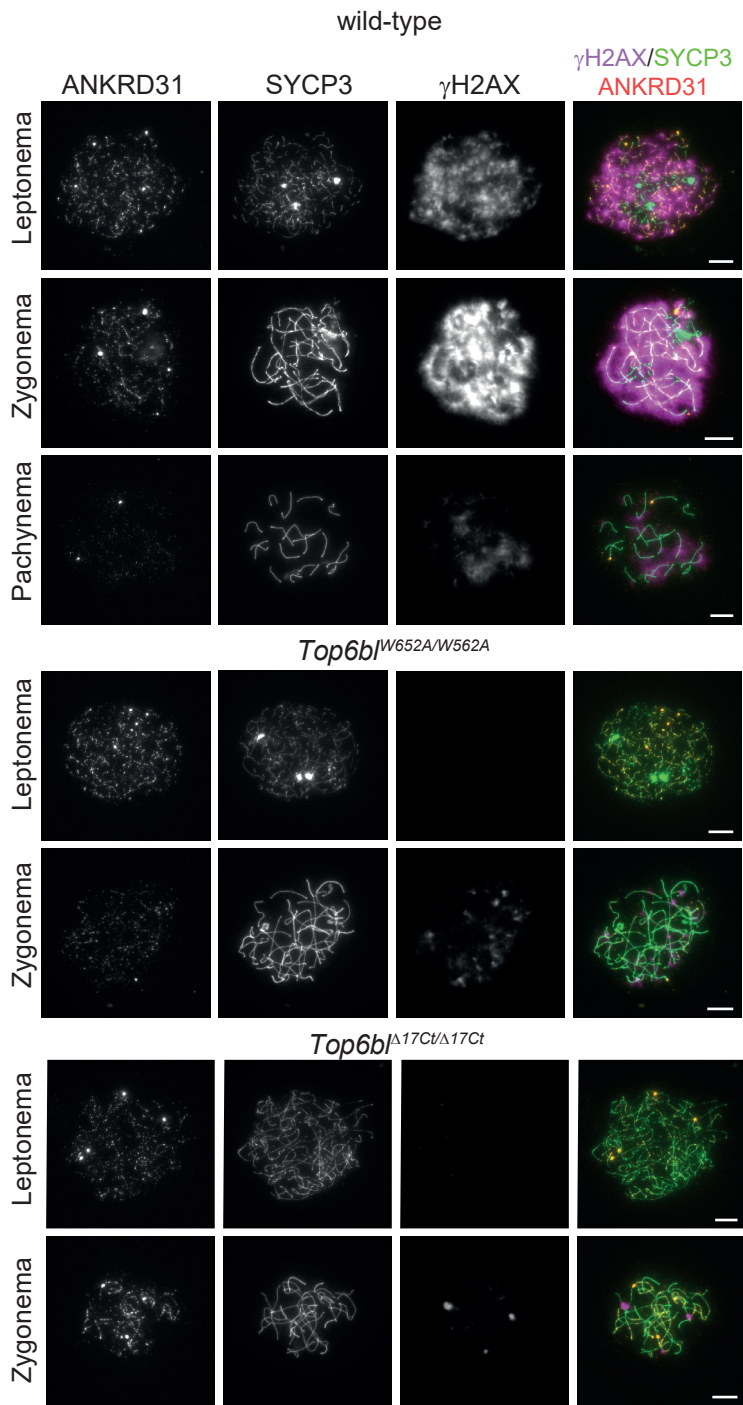
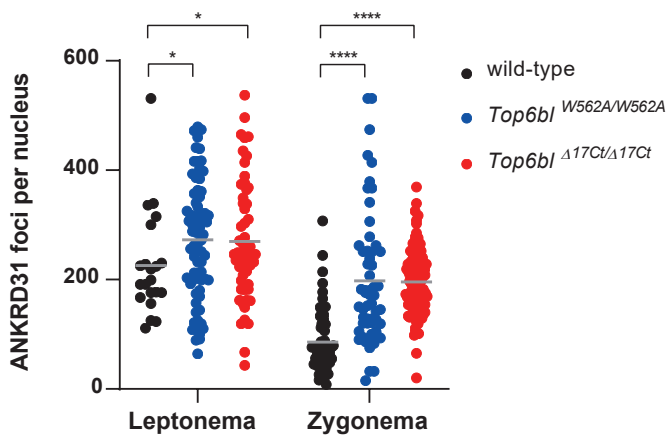
**b.** Immunostaining of RPA2, SYCP3 and  $\gamma$ H2AX in oocytes from 16 dpc wild-type (+/+), *Top6bl*<sup>W562A/W562A</sup>, and *Top6bl* <sup>$\Delta$ 17Ct/ $\Delta$ 17Ct</sup> ovaries. Scale bar, 10 $\mu$ m.

**c.** Quantification of RPA foci per nucleus in leptotene and zygotene oocytes from 16 dpc wild-type (+/+ or *Top6bl*<sup>+/ $\Delta$ 17Ct</sup>), *Top6bl*<sup>W562A/W562A</sup>, *Top6bl* <sup>$\Delta$ 17Ct/ $\Delta$ 17Ct</sup>, and *Top6bl*<sup>-/-</sup> mice (n=4 wild-type and n=2 mice per mutant genotype). Number of nuclei per genotype at leptotene (157, 85, 14, and 29) and at zygotene (165, 55, 105, and 28). Grey bars show the mean value. P values were determined using the two-tailed unpaired Mann-Whitney test. Ratios of the wild-type and mutant mean values are shown.

**d.** Immunostaining of MEI4, SYCP3 and  $\gamma$ H2AX in oocytes from 16 dpc wild-type (+/+), *Top6bl*<sup>W562A/W562A</sup>, and *Top6bl* <sup>$\Delta$ 17Ct/ $\Delta$ 17Ct</sup> ovaries. Scale bar, 10 $\mu$ m.

**e.** Quantification of axis-associated MEI4 foci in leptotene and zygotene oocytes from 15 and 16 dpc wild-type (+/+ or *Top6bl*<sup>+/ $\Delta$ 17Ct</sup>), *Top6bl*<sup>W562A/W562A</sup>, and *Top6bl* <sup>$\Delta$ 17Ct/ $\Delta$ 17Ct</sup> mice (n=3 wild-type, n=1 *Top6bl*<sup>W562A/W562A</sup>, and n=2 *Top6bl* <sup>$\Delta$ 17Ct/ $\Delta$ 17Ct</sup> mice). Number of nuclei per genotype at leptotene (139, 91, and 42) and at zygotene (106, 24, and 23). Grey bars show the mean values. P values were determined using the two-tailed unpaired Mann-Whitney test.

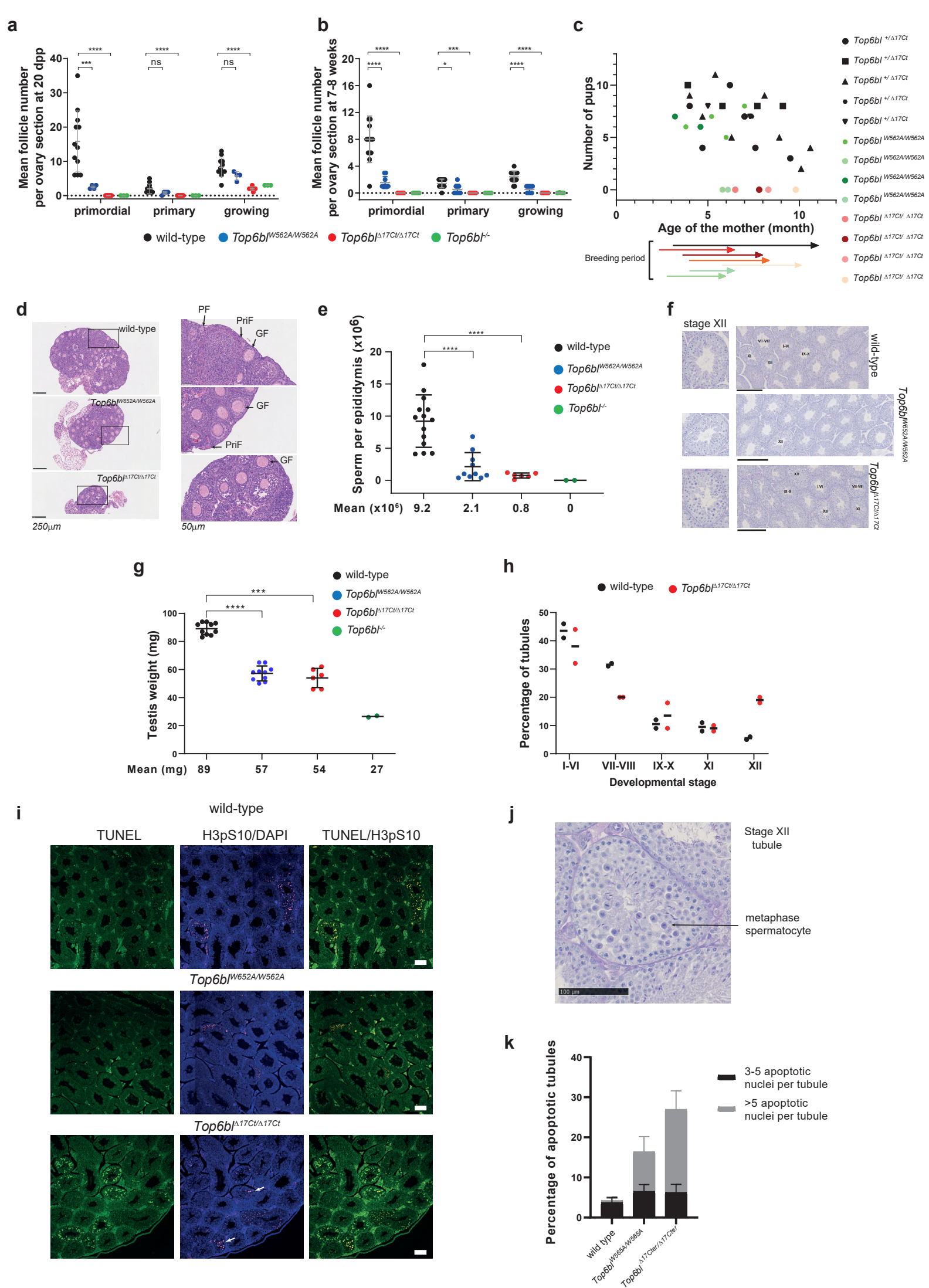
**f.** Quantification of axis-associated REC114 foci in leptotene and zygotene oocytes from 16 dpc wild-type (+/+) and *Top6bl*<sup>-/-</sup> mice (n=1 mouse/genotype). Number of nuclei: 29 and 36 L, and 31 and 29 Z, respectively. Grey bars show the mean values. P values were determined using the two-tailed unpaired Mann-Whitney test. Source data are provided as a Source Data file.

**a****b**

**Supplementary Figure 8. Immunocytochemistry of ANKRD31 on oocytes**

**a.** Immunostaining of ANKRD31, SYCP3 and  $\gamma$ H2AX in oocytes from 16 dpc wild-type (+/+), *Top6bl*<sup>W562A/W562A</sup>, and *Top6bl* <sup>$\Delta$ 17Ct/ $\Delta$ 17Ct</sup> ovaries. Scale bar, 10 $\mu$ m.

**b.** Quantification of axis-associated ANKRD31 foci in leptotene and zygotene oocytes from 16 and 18 dpc wild-type (+/+ or *Top6bl*<sup>+/ $\Delta$ 17Ct</sup>), *Top6bl*<sup>W562A/W562A</sup>, and *Top6bl* <sup>$\Delta$ 17Ct/ $\Delta$ 17Ct</sup> mice (n=3 wild-type, n=1 *Top6bl*<sup>W562A/W562A</sup>, and n=2 *Top6bl* <sup>$\Delta$ 17Ct/ $\Delta$ 17Ct</sup> mice). Number of nuclei per genotype at leptoneuma (21, 74, and 56) and at zygonema (56, 57, and 108). Grey bars show the mean values. P values were determined using the two-tailed unpaired Mann-Whitney test. Source data are provided as a Source Data file.



## Supplementary Figure 9. Histology on ovaries and testis

a. Quantification of primordial, primary, and growing follicles in ovaries from 20 dpp wild-type homozygous or heterozygous (+/+ , *Top6bl*<sup>+/ $\Delta$ 17Ct</sup> , *Top6bl*<sup>+/ $W^{562A}$</sup>  , +/ *Top6bl*), *Top6bl* <sup>$W^{562A}/W^{562A}$</sup>  , *Top6bl* <sup>$\Delta$ 17Ct/ $\Delta$ 17Ct</sup> , and *Top6bl*<sup>-/-</sup> mice. Each data point is the mean  $\pm$  SD of at least 10 sections from one ovary. Number of mice: 6, 2, 4, and 2, respectively. P values were determined using the two-tailed unpaired Mann-Whitney test.

b. Quantification of primordial, primary, and growing follicles in ovaries from 7-8-week-old wild-type (+/+ or *Top6bl*<sup>+/ $W^{562A}$</sup> ), *Top6bl* <sup>$W^{562A}/W^{562A}$</sup>  , *Top6bl* <sup>$\Delta$ 17Ct/ $\Delta$ 17Ct</sup> , and *Top6bl*<sup>-/-</sup> mice. Each data point is the mean  $\pm$  SD of 20 sections from one ovary. P values were determined using the two-tailed unpaired Mann-Whitney tests. Number of mice: 7, 5, 3, and 2 respectively.

c. Fertility tests. *Top6bl*<sup>+/ $\Delta$ 17Ct</sup> , *Top6bl* <sup>$W^{562A}/W^{562A}$</sup>  and *Top6bl* <sup>$\Delta$ 17Ct/ $\Delta$ 17Ct</sup> females (n= 5, 4 and 4 respectively), from 2.3 to 10.6 months of age, were mated with heterozygous males (*Top6bl*<sup>+/ $\Delta$ 17Ct</sup> or *Top6bl*<sup>+/ $W^{562A}$</sup> ). Each symbol refers to a female. Only two of the four *Top6bl* <sup>$W^{562A}/W^{562A}$</sup>  females were fertile (large circle, dark green, and small circle), and after the last progeny, they were not mated any longer. Breeding periods of non-productive females and for *Top6bl*<sup>+/ $\Delta$ 17Ct</sup> females are shown.

d. Haematoxylin and eosin-stained sections of ovaries from 20 dpp wild-type (+/+), *Top6bl* <sup>$W^{562A}/W^{562A}$</sup>  and *Top6bl* <sup>$\Delta$ 17Ct/ $\Delta$ 17Ct</sup> mice (samples from panel a). Scale bars, 250 and 50  $\mu$ m in the left and right panel, respectively. Primordial (PriF), primary (PF), and growing (GF) follicles are indicated.

e. Sperm count in *Top6bl* mutants. Sperm in epididymis was counted in 2-month-old mice. Each data point is one epididymis (number of mice= 7, 5, 3 and 1 for wild-type (+/+ or +/-), *Top6bl* <sup>$W^{562A}/W^{562A}$</sup>  , *Top6bl* <sup>$\Delta$ 17Ct/ $\Delta$ 17Ct</sup> , and *Top6bl*<sup>-/-</sup> , respectively). The mean  $\pm$  SD is shown. P values were determined using the two-tailed unpaired Mann-Whitney test.

f. Periodic acid-Schiff staining of testis sections from 2 month-old wild-type (+/+), *Top6bl* <sup>$W^{562A}/W^{562A}$</sup>  , and *Top6bl* <sup>$\Delta$ 17Ct/ $\Delta$ 17Ct</sup> mice. Scale bar, 250 $\mu$ m. Zoomed images of stage XII tubules are shown on the right.

g. Testis weight in *Top6bl* mutants. Testis weight was measured in 2 month-old mice. Each data point is one testis (number of mice= 5, 5, 3, and 1 for wild-type (+/+), *Top6bl* <sup>$W^{562A}/W^{562A}$</sup>  , *Top6bl* <sup>$\Delta$ 17Ct/ $\Delta$ 17Ct</sup> , and

*Top6bl<sup>-/-</sup>*, respectively). The mean value  $\pm$  SD is shown. P values were determined using the two-tailed unpaired Mann-Whitney test.

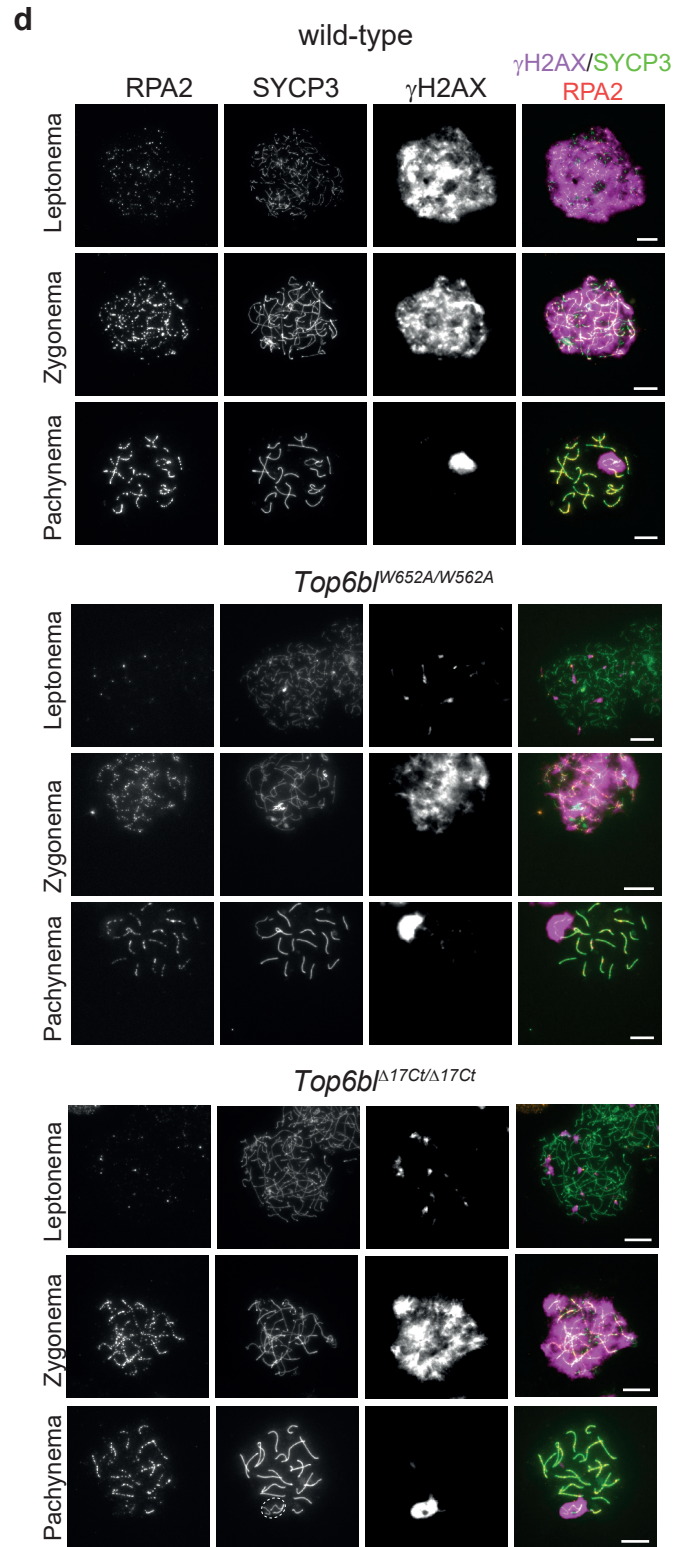
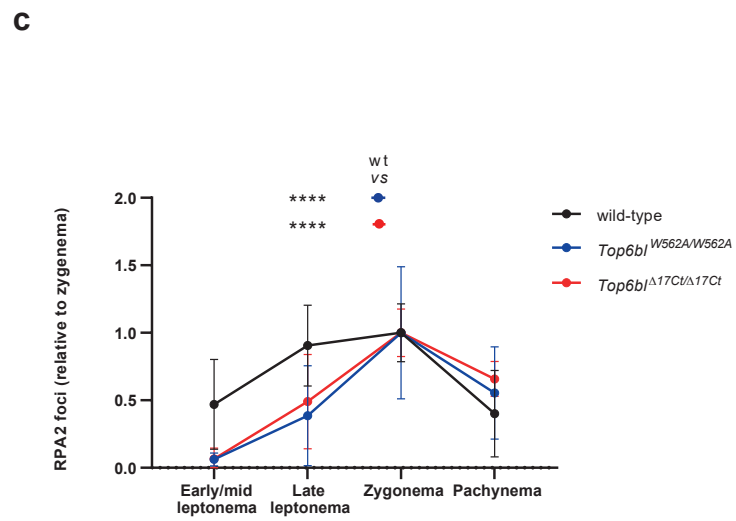
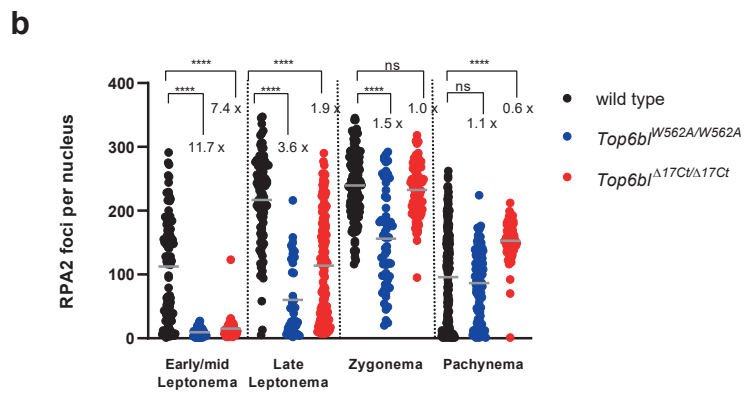
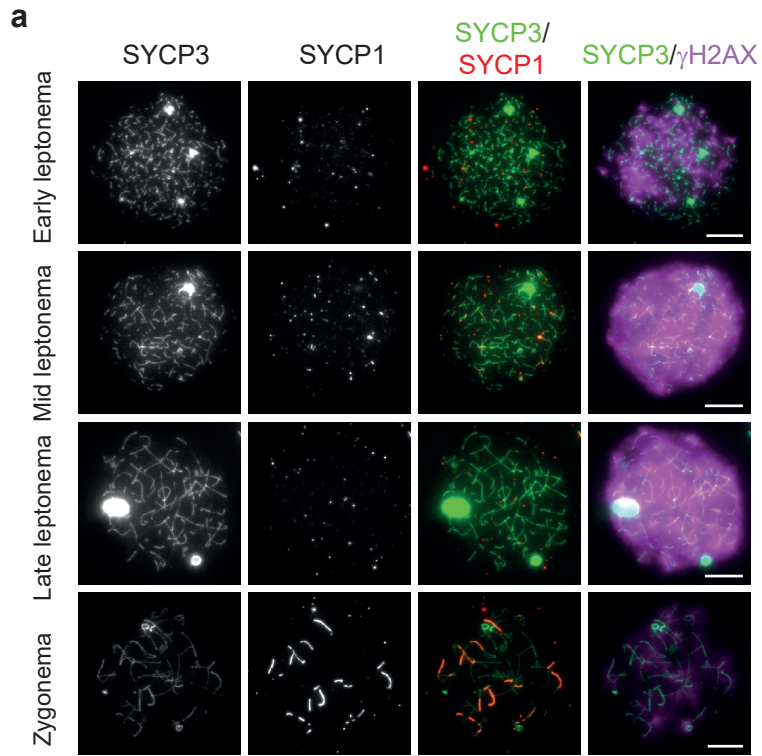
**h.** Staging of tubules in testis sections from 2-month-old wild-type and *Top6bl <sup>$\Delta$ 17Ct/ $\Delta$ 17Ct</sup>* mice (n=2 mice/genotype). Tubules were divided in five groups based on several criteria. I-VI: round spermatid with acrosomes, 0 to ~25% of the perimeter of the spermatid, elongated spermatids on the lumen; VII-VIII: round spermatids with acrosomes, ~50% of the perimeter; IX-X: elongated spermatids with acrosomes not hook-shaped; XI: elongated spermatids with a hook; XII: presence of metaphases. Number of staged tubules: 310 in wild-type and 366 in *Top6bl <sup>$\Delta$ 17Ct/ $\Delta$ 17Ct</sup>* mice. The horizontal bars show the mean value. The distributions are statistically different (Pearson's Chi-Square: 36.028).

**i.** Detection of apoptosis and metaphase in testis sections. Sections were labelled with TUNEL (green) and H3pS10 (red), and stained with DAPI (blue). Scale bar, 100 $\mu$ m. White arrowheads indicate tubules positive for H3pS10 and TUNEL.

**j.** A stage XII tubule with metaphase spermatocytes (from *Top6bl <sup>$\Delta$ 17Ct/ $\Delta$ 17Ct</sup>* mice). Scale bar, 100 $\mu$ m. Several metaphases with condensed chromosomes are seen, and one (arrow) shows relatively more intense nuclear staining, suggesting apoptosis.

**k.** Detection of apoptosis by TUNEL assay. Tubules were classified in three categories: <3, 3 to 5, and >5 apoptotic nuclei. The mean values  $\pm$  SD of the 3 to 5 and >5 categories are shown. Number of 2 month-old mice: 4 wild-type (+/+ and *Top6bl<sup>+/ $\Delta$ 17Ct</sup>*) and 2 mice per mutant. The percentage of tubules with <3, 3 to 5, and >5 apoptotic nuclei was significantly different between wild-type and *Top6bl<sup>W562A/W562A</sup>* and *Top6bl <sup>$\Delta$ 17Ct/ $\Delta$ 17Ct</sup>* mice (Chi-square Pearson: 354.03 and 811.63, respectively). Source data are provided as a Source Data file.





### Supplementary Figure 10. Staging and immunocytochemistry on spermatocytes

**a.** Identification of spermatocytes at early, mid, late leptotene and zygonema. Examples of wild-type nuclei at the different stages by staining with SYCP3 (green), SYCP1 (red) and  $\gamma$ H2AX (magenta). Scale bar, 10 $\mu$ m.

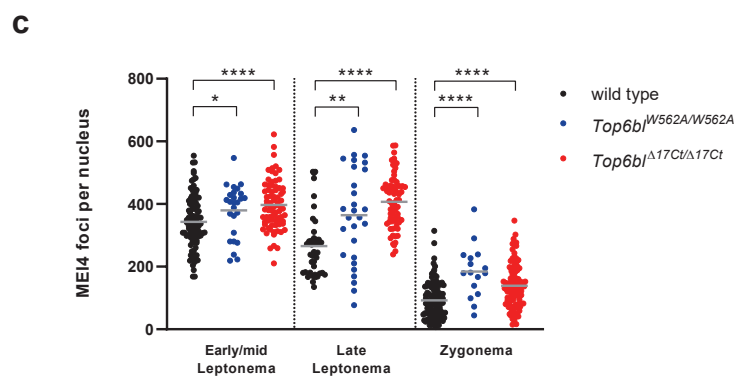
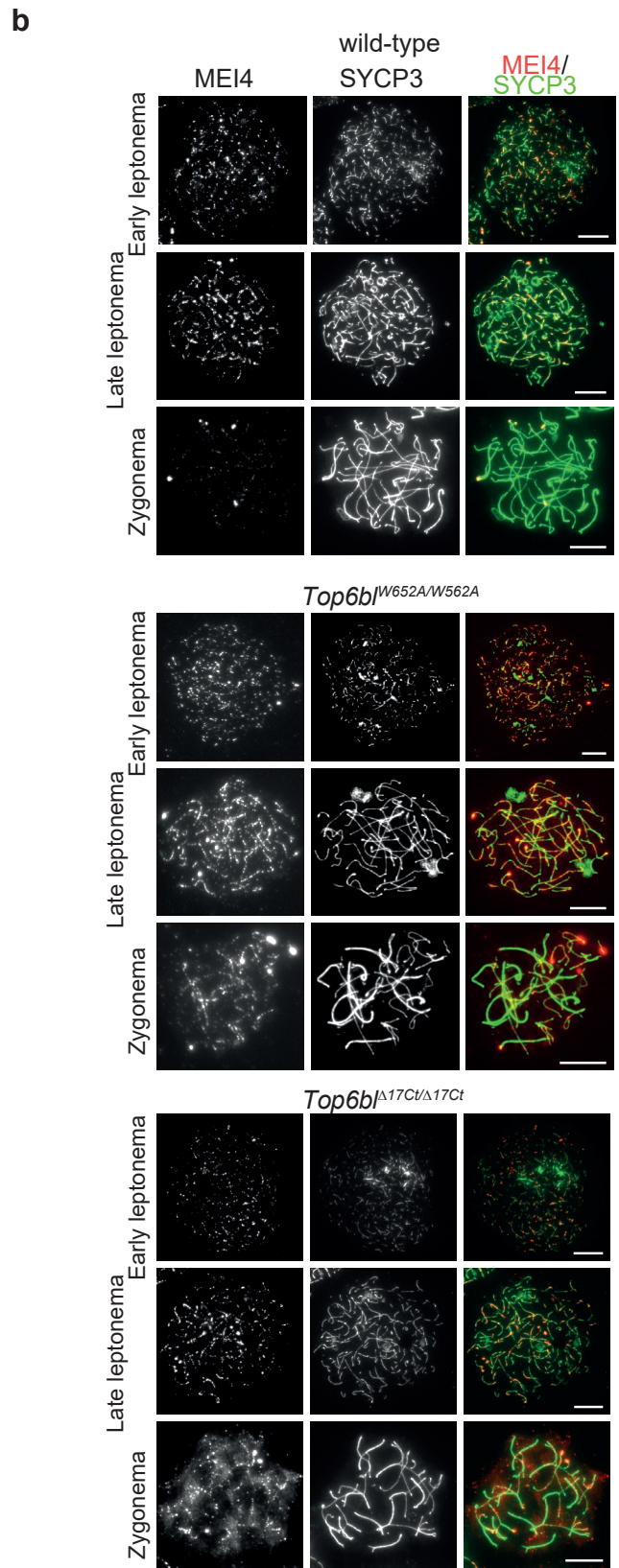
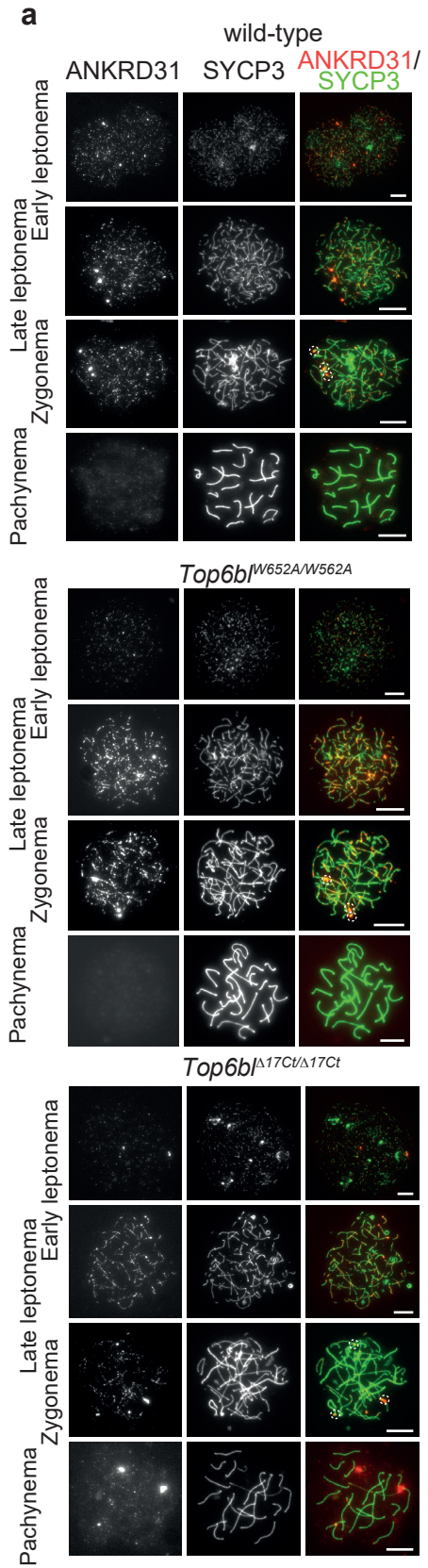
**b.** Quantification of axis-associated RPA2 foci in early/mid and late leptotene, zygotene and pachytene spermatocytes from 12-14dpp wild-type (+/+ or *Top6bl*<sup>+/ $\Delta$ 17Ct</sup>; n=3), *Top6bl*<sup>W562A/W562A</sup> (n=2), and *Top6bl* <sup>$\Delta$ 17Ct/ $\Delta$ 17Ct</sup> (n=2) mice. Number of nuclei per genotype at early/mid leptotene (105, 36, and 41), late leptotene (91, 38, and 131), zygotene (135, 62, and 85), and pachytene (178, 119, and 80). Grey bars show the mean values. P values were determined using the two-tailed unpaired Mann-Whitney test.

**c.** RPA2 focus variation during prophase. For each genotype (wild-type, *Top6bl*<sup>W562A/W562A</sup>, and *Top6bl* <sup>$\Delta$ 17Ct/ $\Delta$ 17Ct</sup>), the number of RPA2 foci at the different stages (dataset as in panel b) was normalized to the mean number at zygonema (set to 1). The mean value  $\pm$  SD is shown. Statistical significance between wild-type and each mutant (*Top6bl*<sup>W562A/W562A</sup>; *Top6bl* <sup>$\Delta$ 17Ct/ $\Delta$ 17Ct</sup>) was tested at late leptotene using the two-tailed unpaired Mann-Whitney test.

**d.** Immunostaining of RPA2, SYCP3 and  $\gamma$ H2AX in spermatocytes from 12-14dpp wild-type (+/+), *Top6bl*<sup>W562A/W562A</sup>, and *Top6bl* <sup>$\Delta$ 17Ct/ $\Delta$ 17Ct</sup> spermatocytes. A white dotted circle highlights the unsynapsed X and Y chromosomes at pachynema (SYCP3 staining) in *Top6bl* <sup>$\Delta$ 17Ct/ $\Delta$ 17Ct</sup> mice. Scale bar, 10 $\mu$ m.

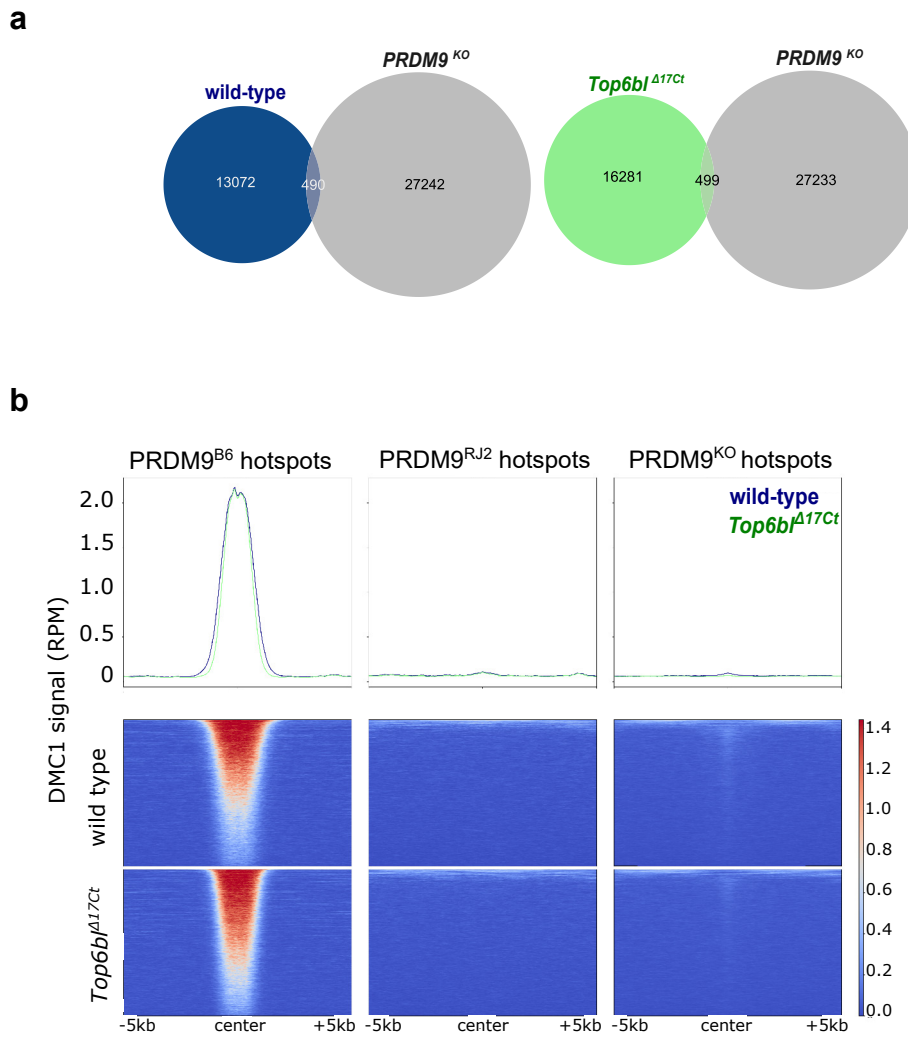
Source data are provided as a Source Data file.





### Supplementary Figure 11. Immunocytochemistry on spermatocytes

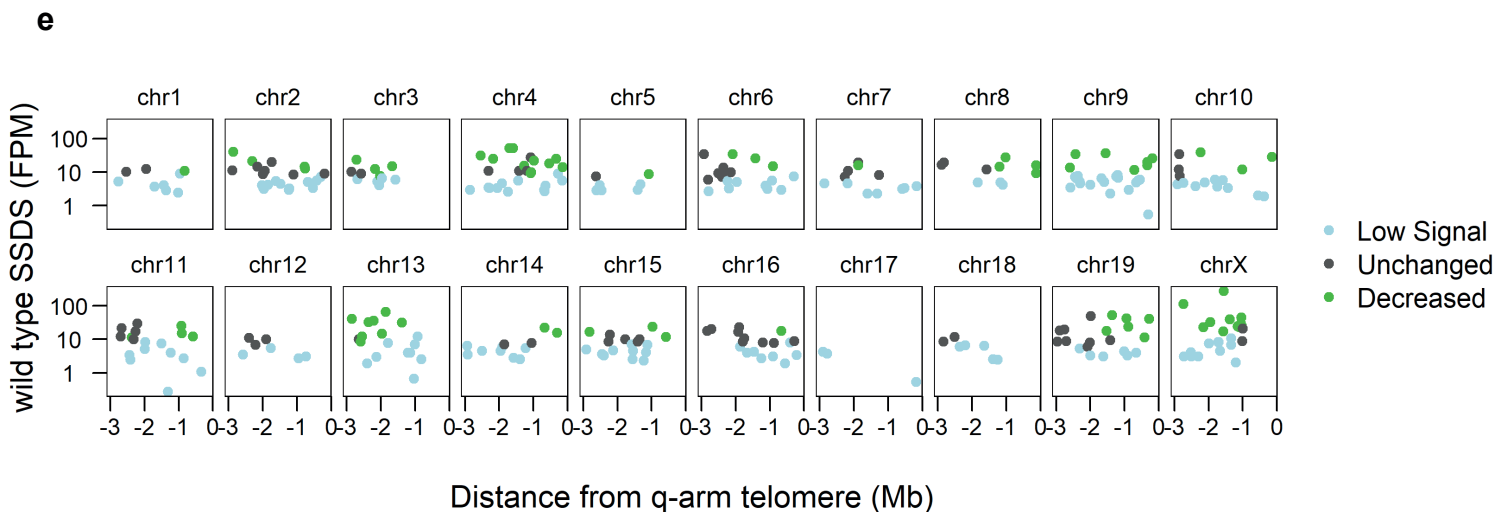
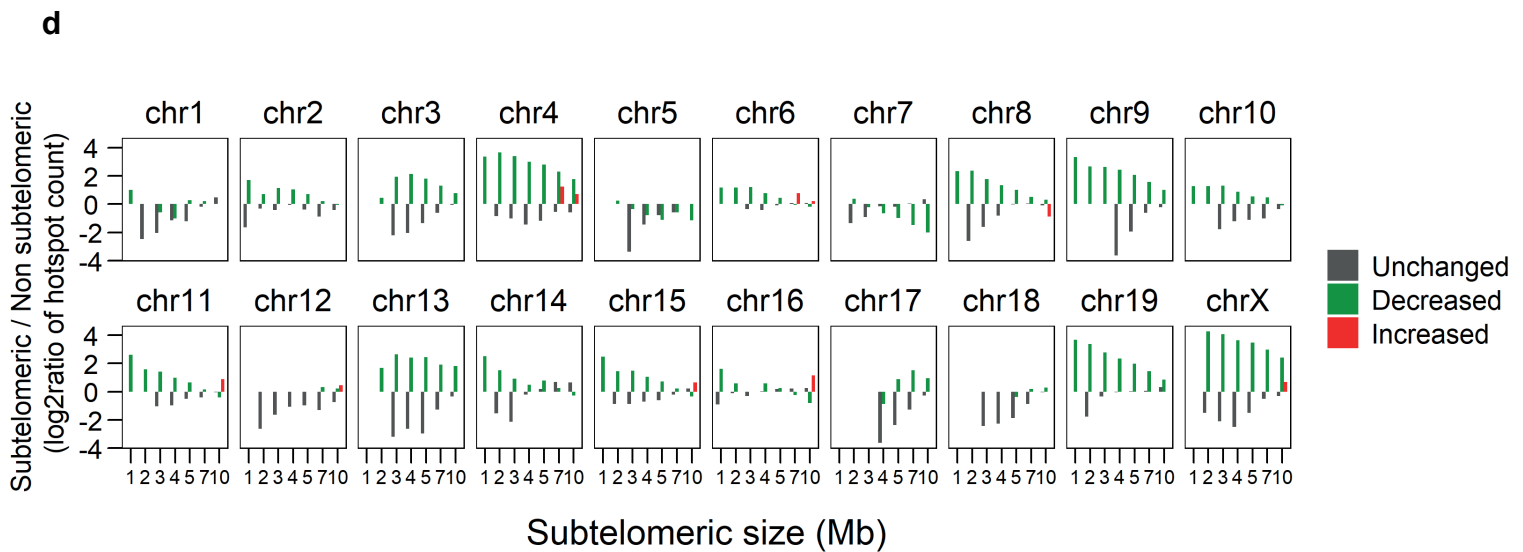
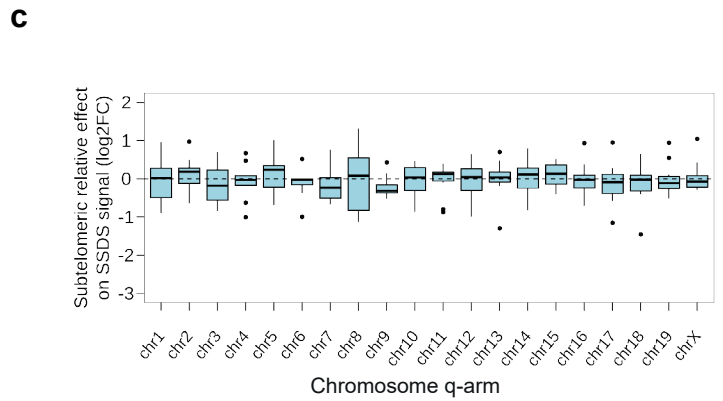
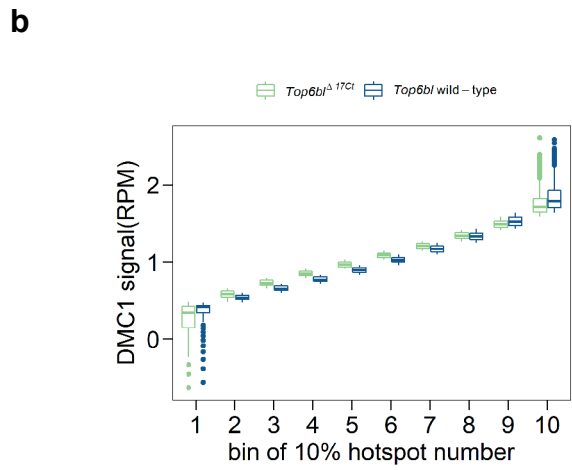
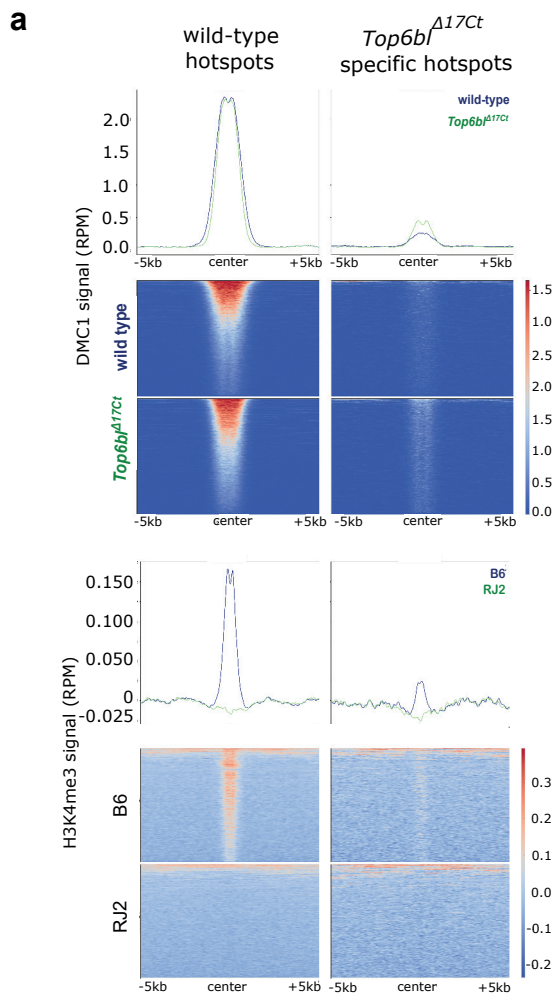
- a.** Immunostaining of ANKRD31 and SYCP3 in spermatocytes from 12-14dpp wild-type (+/+), *Top6bl*<sup>W562A/W562A</sup>, and *Top6bl*<sup>Δ17Ct/Δ17Ct</sup> mice. At zygonema, dotted white circles (merge panels) highlight ANKRD31 aggregates at the ends of some chromosome axes. Scale bar, 10μm.
- b.** Immunostaining of MEI4 and SYCP3 in spermatocytes from 12-14 dpp wild-type (+/+), *Top6bl*<sup>W562A/W562A</sup>, and *Top6bl*<sup>Δ17Ct/Δ17Ct</sup> mice. Scale bar, 10μm.
- c.** Quantification of axis-associated MEI4 foci in early/mid, late leptotene and zygotene spermatocytes from 12-14 dpp wild-type (+/+ or *Top6bl*<sup>+ /Δ17Ct</sup>; n=2), *Top6bl*<sup>W562A/W562A</sup> (n=1), and *Top6bl*<sup>Δ17Ct /Δ17Ct</sup> (n=2) mice. Number of nuclei: 97 early/mid L, 43 late L, 114 Z in wild-type; 26 early/mid L, 28 late L and 16 Z in *Top6bl*<sup>W562A/W562A</sup>; 77 early/mid L, 68 late L and 103 Z in *Top6bl*<sup>Δ17Ct /Δ17Ct</sup>. Grey bars show the mean values. P values were determined using the two-tailed unpaired Mann-Whitney test. Source data are provided as a Source Data file.



### Supplementary Figure 12. Hotspot overlap analysis

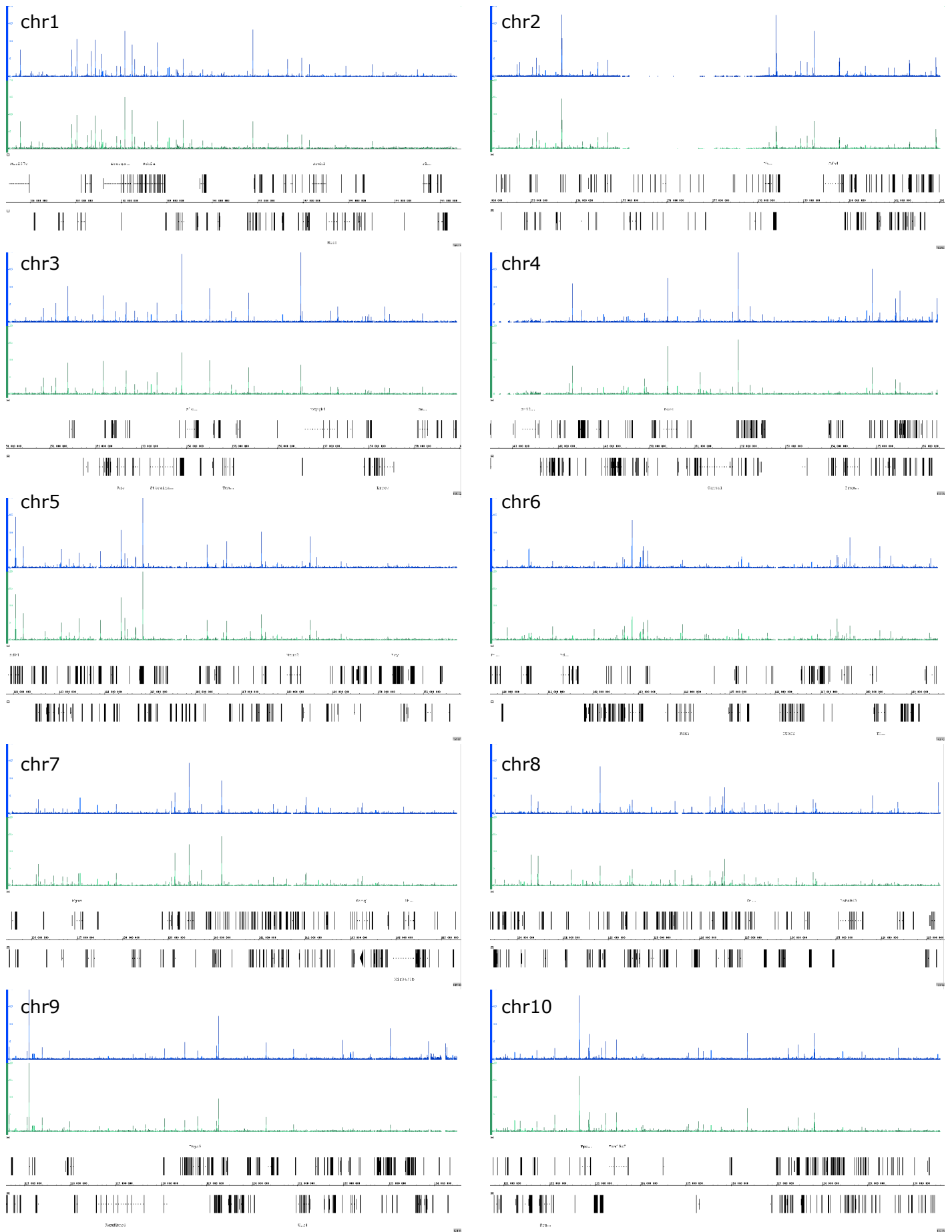
**a.** Overlap between wild-type and *Top6bl*<sup>Δ17Ct/Δ17Ct</sup> hotspots and the default hotspots used in the *PRDM9*<sup>KO</sup> mouse strain. The overlaps were restricted to the central 400bp around the estimated hotspot centre.

**b.** DMC1-SSDS signal (heatmap and average plots) from wild-type and *Top6bl*<sup>Δ17Ct/Δ17Ct</sup> mice at *PRDM9*<sup>B6</sup>-, *PRDM9*<sup>RJ2</sup>- and *PRDM9*<sup>KO</sup>-defined hotspots.



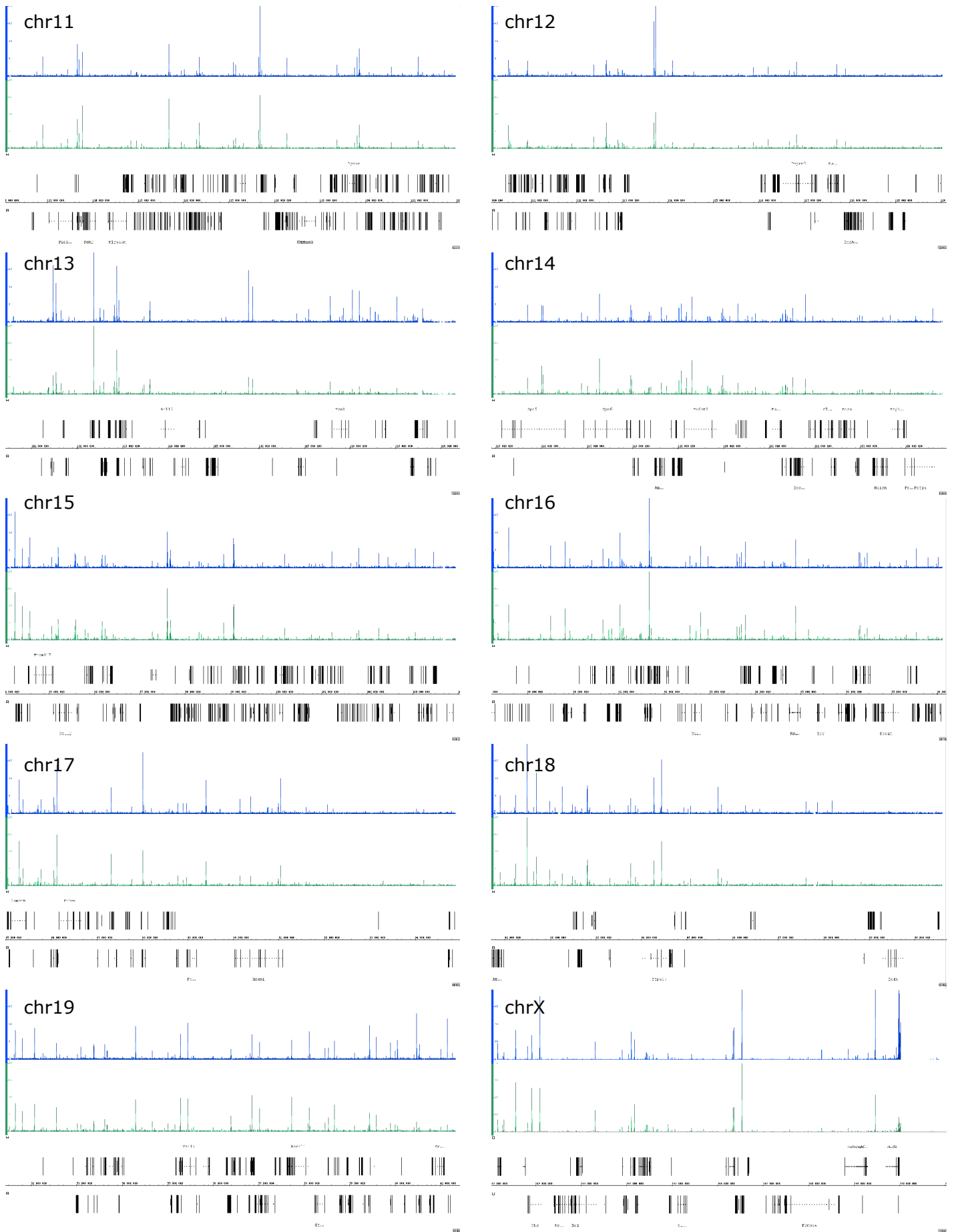
### Supplementary Figure 13. Hotspots signal quantification analysis

- a.** DMC1-SSDS signal from *Top6bl* <sup>$\Delta 17Ct/\Delta 17Ct$</sup>  and wild-type mice at wild-type hotspots (13562) and at hotspots specifically activated in *Top6bl* <sup>$\Delta 17Ct/\Delta 17Ct$</sup>  (3519), and H3K4me3 signal (heatmap and average plots) from B6 and RJ2 mice. Peaks are ranked as a function of the DMC1 or H3K4me3 signal intensity.
- b.** DMC1-SSDS signal in *Top6bl* <sup>$\Delta 17Ct/\Delta 17Ct$</sup> , *Ankrd31*<sup>-/-</sup> and their respective wild-type control mice. Signal is represented as box plots (as defined in Methods) computed for hotspots (n= 13562) ranked according to increased DMC1-SSDS signal intensity and grouped into 10 bins (1 bin = 10% of the total hotspot number; bin 1 = weakest signal, bin 10= strongest signal).
- c.** Control for the sub-telomeric effect. Quantification was done as in Fig. 6f, but with ratios calculated between two random pools of 10 non-sub-telomeric consecutive hotspots (box plot as defined in Methods).
- d.** Sub-telomeric/Non-sub-telomeric hotspot density ratios for unchanged, decreased, and increased hotspots, as defined by DESeq2 analysis, for sub-telomeric regions defined using various distances from the telomeric annotated ends (1-2-3-4-5-7-10Mb). The log<sub>2</sub> ratio was plotted and showed excess (positive values) or lack (negative values) of hotspots at sub-telomeric regions compared with non-sub-telomeric regions.
- e.** Distribution of hotspots within 3 Mb from annotated telomeric ends and according to their DESeq2 category (unchanged, decreased, or low signal). Note the absence of increased hotspots within this sub-telomeric window size. The normalized wild-type DMC1-SSDS signal is shown on the y-axis.



**Supplementary Figure 14. Browser windows of DMC1-SSDS**

Browser windows of DMC1-SSDS in wild-type (blue) and *Top6bl*<sup>Δ17Ct/Δ17Ct</sup> (green) mice within the 10Mb q-arm telomere-proximal regions of chromosomes 1 to 10.

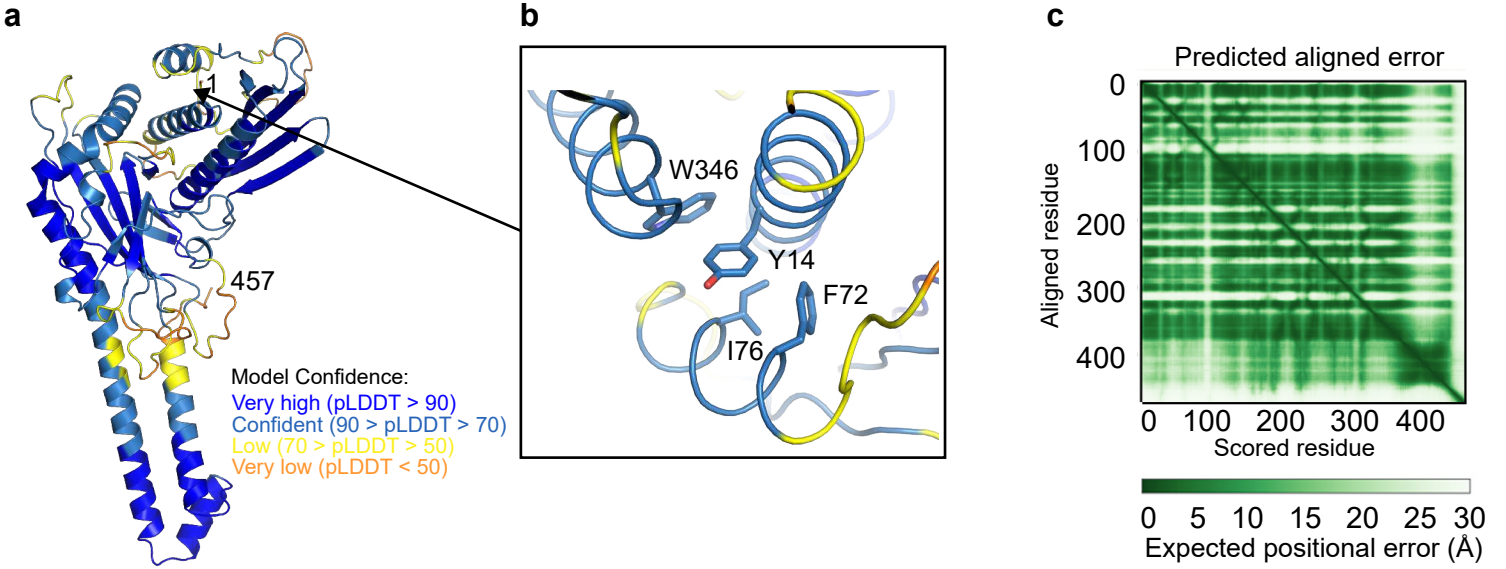


### Supplementary Figure 15. Browser windows of DMC1-SSDS

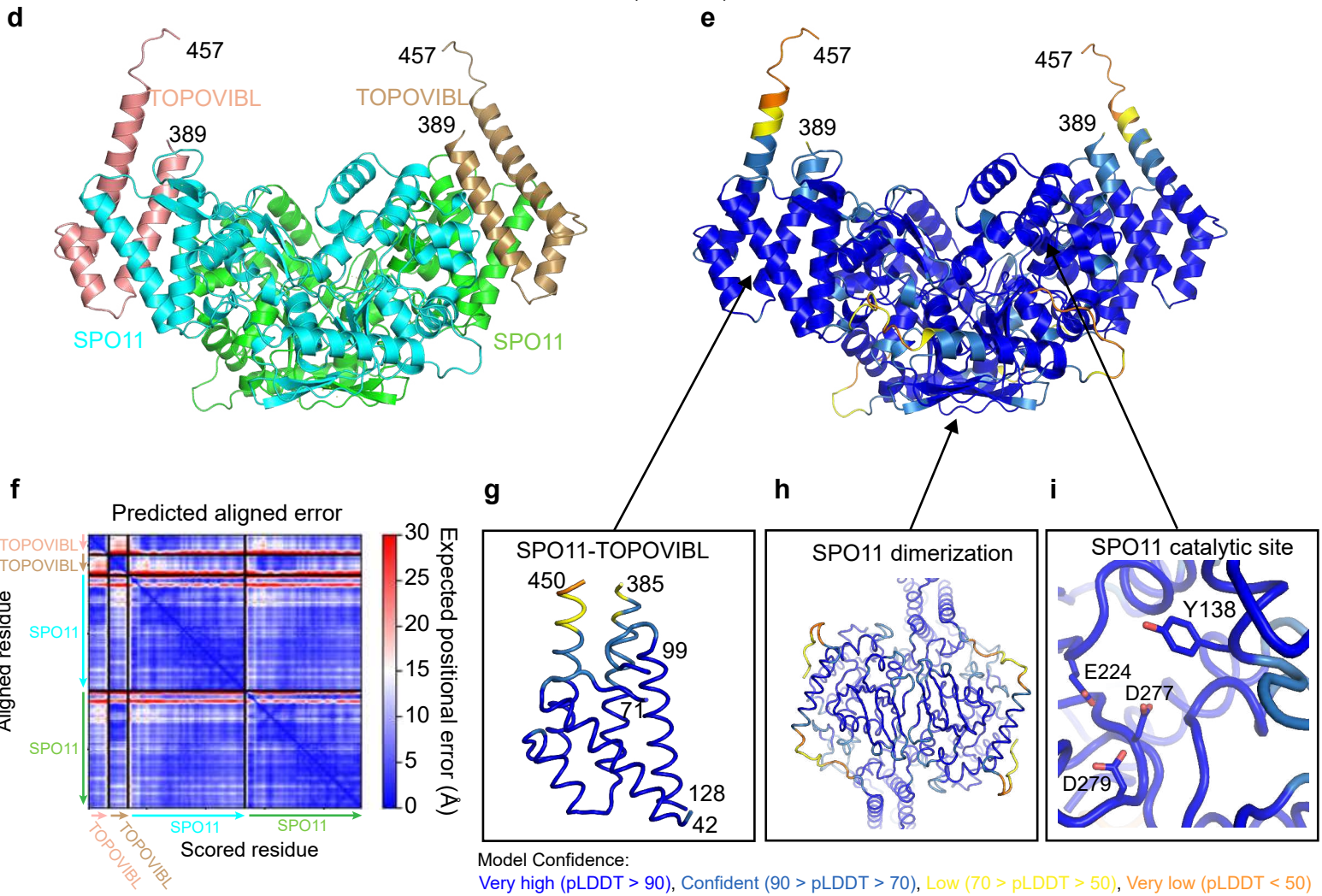
Browser windows of DMC1-SSDS in wild-type (blue) and *Top6bl*<sup>Δ17C1/Δ17C1</sup> (green) mice within the 10Mb q-arm telomere-proximal regions of chromosomes 11 to X.



TOPOVIBL (1-457) AF model -AF-J3QMY9



SPO11-TOPOVIBL (389-457) AF model





### **Supplementary Figure 16. Quantitative evaluations of AlphaFold2 models**

- a.** AlphaFold2 model of the mouse TOPOVIBL<sup>1-457</sup>, including the GHKL-like, linker and transducer domains coloured according to the AlphaFold2 per-residue estimate of confidence (pLDDT).
- b.** Detail of the TOPOVIBL GHKL-like domain shown in Fig. 8b and coloured according to pLDDT.
- c.** Predicted aligned error plot for the model shown in **a**.
- d.** AlphaFold2 model of the SPO11-TOPOVIBL<sup>389-457</sup> complex coloured by protein chain.
- e.** AlphaFold2 model of the SPO11-TOPOVIBL<sup>389-457</sup> complex coloured according to pLDDT.
- f.** Predicted aligned error plot for the model shown in **d**.
- g-i** Details of the SPO11-TOPOVIBL complex shown in Fig. 8c-e coloured according to pLDDT.

	<b>TOPVIBL-REC114</b>
<b>Data collection</b>	
Space group	<i>P</i> 6 <sub>1</sub> 22
Cell dimensions	
<i>a</i> , <i>b</i> , <i>c</i> (Å)	108.7, 108.7, 83
$\alpha$ $\beta$ $\gamma$ (°)	90, 90, 120
Resolution (Å)	62.3-2.26 (2.35-2.26) <sup>a</sup>
<i>R</i> <sub>merge</sub> (%)	6.7(169)
I / $\sigma$ I	23.9 (1.4)
<i>CC</i> <sub>1/2</sub> (%)	99.9 (57.8)
Completeness (%)	84.4 (38.9)
Redundancy	18 (12.2)
<b>Refinement</b>	
Resolution (Å)	94.03-2.26
No. reflections	11285
<i>R</i> <sub>work</sub> / <i>R</i> <sub>free</sub>	0.226/0.249
<i>B</i> -factors	70
R.m.s. deviations	
Bond lengths (Å)	0.006
Bond angles (°)	1.412

<sup>a</sup> Values in parentheses are for highest resolution shell.

### Supplementary Table 1

Data collection and refinement statistics.

	Unchanged-Decreased-Increased									
	Subtel 1M	Subtel 2M	Subtel 3M	Subtel 4M	Subtel 5M	Subtel 1M	Subtel 2M	Subtel 3M	Subtel 4M	Subtel 5M
chr1	0.19688	1	1	1	1	0.337954	0.707152	0.95226	1	0.30121
chr2	0.19688	1	0.9538	1	1	0.118562	0.728138	0.170249	0.26614	0.26614
chr3	NA	0.09331	0.00059	0	0.00062	NA	0.209354	0.000178	1.00E-06	8.70E-05
chr4	2.00E-05	0	0	0	0	1.20E-05	0	0	0	0
chr5	NA	0.09472	1	1	1	NA	0.214951	0.530377	1	1
chr6	0.1378	0.00765	1	1	1	0.275433	0.00955	0.263484	0.426418	0.848232
chr7	NA	1	1	1	1	NA	0.867484	1	1	0.92624
chr8	0.00204	9.00E-05	0.00548	0.2626	1	0.005766	5.50E-05	0.001712	0.039004	0.37051
chr9	2.00E-05	0	0	0	0	5.00E-06	0	0	0	0
chr10	0.07514	0.00201	0.06632	0.68638	1	0.213463	0.003835	0.016055	0.106033	0.211257
chr11	0.00041	0.00046	0.13714	0.58528	1	0.000331	0.000331	0.022116	0.06639	0.301525
chr12	NA	1	1	1	1	NA	1	1	0.877914	0.708196
chr13	NA	0.00056	0	0	0	NA	0.000478	0	0	0
chr14	0.00156	0.20581	0.25	1	1	0.004174	0.094238	0.094238	0.886754	0.784146
chr15	0.00041	0.58244	0.25	1	1	0.001482	0.189291	0.0572	0.170896	0.329774
chr16	0.85779	1	1	1	1	0.596137	1	1	0.932596	1
chr17	NA	NA	NA	1	0.02142	NA	NA	NA	0.622188	0.003701
chr18	NA	NA	1	1	1	NA	NA	1	0.644324	0.875619
chr19	0	0	0.00162	0.05963	0.26777	0	0	0.00019	0.004521	0.020866
chrX	NA	0	0	0	0	NA	0	0	0	0

### Supplementary Table 2.

Statistics to test the biased distribution of hotspot densities at sub-telomeric vs. non-sub-telomeric regions in *Top6bl*<sup>Δ17Ct</sup> samples: Pearson's Chi square adjusted p-value for tests considering two (unchanged/decreased) or three (unchanged/decreased/increased) hotspot categories for different sizes of sub-telomeric regions. NA: not applicable due to the absence of hotspot with quantifiable signal.

Guides for Cirspr	6509_Top6bl_sgRNA1	CCCGCAGGACTTGTGGCTAC
	6510_Top6bl_sgRNA2	GCAGGACTTGTGGCTACAGG
	6511_Top6bl_sgRNA3	TCCTGTAGCCACAAGTCCTG
Donor for CrispR	6512_Top6bl_donor1	TTCTTGGCTTGCTGGGCGTCTCAGAGAGAATGCTTAGGAC CTGTGTCCCGGATTTCAGCCACTCGGACAGATTGGATACCTC CTGCAGCGCCAAGTCCTGCGGGAGCAGAGGAGGCGGTGGT GAGGA
	6513_Top6bl_donor2	GCGGCCTGGAGGTACCGCAGCGAGAGCAGGGCGGGTGCG GGCCTGAGCTACCACCGCCGGGTCCTCACCACCGCCTCCTC TGCTCCCGCAGGACTTGCGCTGCAGGAGGTATCCAATCT GTCCGAG
Genotyping	oliAN 320	CAACCATGGCAGGGAAGAATC
	oliAN 3201	TACCGACAGTAAACGCAGCC
RT-PCR <i>Top6bl</i>	Oli63	GGACCAGCTCTGCTACTTTT
	Oli70	CAGAGAGAATGCTTAGGACC
RT-PCR <i>Spo11</i>	Spo11:116U22	CTGTTGGCCATGGTGAAGAGAG
	Spo11:655L22	TGTGCCGACTAACATTCAAGGA

### Supplementary Table 3

Oligonucleotides

Genotype	Pachytene with full autosomal synapsis (%)	Nuclei (n)
+/+	76.6	77
+/ <i>Top6bl</i> <sup>W562A</sup>	86.7	30
wild-type	83.2	107
<i>Top6bl</i> <sup>W562A</sup> / <i>Top6bl</i> <sup>W562A</sup>	86.4	81
<i>Top6bl</i> <sup>W562A</sup> / <i>Top6bl</i> <sup>W562A</sup>	96.3	27
<i>Top6bl</i> <sup>W562A</sup> mutant	88.9	108
+/+	91.0	79
+/ <i>Top6bl</i> <sup>Δ17Ct</sup>	72.2	54
wild-type	88	133
<i>Top6bl</i> <sup>Δ17Ct</sup> / <i>Top6bl</i> <sup>Δ17Ct</sup>	84.4	96
<i>Top6bl</i> <sup>Δ17Ct</sup> / <i>Top6bl</i> <sup>Δ17Ct</sup>	85.2	108
<i>Top6bl</i> <sup>Δ17Ct</sup> mutant	85	204

#### Supplementary Table 4

Percentage of pachytene spermatocytes with full synapsis




# Structural analysis of Red1 as a conserved scaffold of the RNA-targeting MTREC/PAXT complex

Received: 18 January 2022

Accepted: 4 August 2022

Published online: 24 August 2022

 Check for updates

Anne-Emmanuelle Foucher<sup>1,6</sup>, Leila Touat-Todeschini<sup>2,6</sup>,  
Ariadna B. Juarez-Martinez<sup>1,6</sup>, Auriane Rakitch<sup>2</sup>, Hamida Laroussi<sup>1</sup>,  
Claire Karczewski<sup>2</sup>, Samira Acajjaoui<sup>3</sup>, Montserrat Soler-López<sup>3</sup>,  
Stephen Cusack<sup>4</sup>, Cameron D. Mackereth<sup>5</sup> , André Verdel<sup>1,2</sup>  &  
Jan Kadlec<sup>1</sup> 

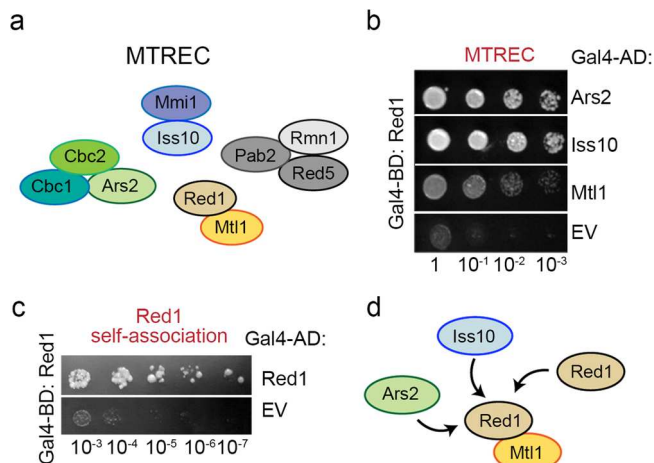
To eliminate specific or aberrant transcripts, eukaryotes use nuclear RNA-targeting complexes that deliver them to the exosome for degradation. *S. pombe* MTREC, and its human counterpart PAXT, are key players in this mechanism but inner workings of these complexes are not understood in sufficient detail. Here, we present an NMR structure of an MTREC scaffold protein Red1 helix-turn-helix domain bound to the Iss10 N-terminus and show this interaction is required for proper cellular growth and meiotic mRNA degradation. We also report a crystal structure of a Red1-Ars2 complex explaining mutually exclusive interactions of hARS2 with various ED/EGEI/L motif-possessing RNA regulators, including hZFC3H1 of PAXT, hFLASH or hNCBP3. Finally, we show that both Red1 and hZFC3H1 homo-dimerize via their coiled-coil regions indicating that MTREC and PAXT likely function as dimers. Our results, combining structures of three Red1 interfaces with in vivo studies, provide mechanistic insights into conserved features of MTREC/PAXT architecture.

An astonishing amount of RNA, including a sizeable portion of aberrant transcripts, is constantly being synthesized within eukaryotic nuclei<sup>1–3</sup>. To detect and eliminate specific or aberrant transcripts, cells rely on highly regulated RNA degradation machineries. Unwanted transcripts are recognized by specialized RNA-targeting complexes that deliver them to the nuclear RNA exosome for degradation<sup>4,5</sup>.

In the fission yeast *Schizosaccharomyces pombe*, MTREC (Mtl1–Red1 core) or NURS (nuclear RNA silencing) is such an RNA-targeting complex and is responsible for recognition of cryptic unstable transcripts (CUTs), meiotic mRNAs and unspliced pre-mRNAs<sup>6–8</sup>. The MTREC core module comprises the zinc-finger

protein Red1 and the RNA helicase Mtl1. In addition, there are four associated sub-modules - the nuclear cap-binding complex (Cbc-Ars2), the Iss10-Mmi1 complex that targets meiotic RNAs, the Pab2-Rmn1-Red5 complex likely recognizing poly(A) sequences and the more loosely associated poly(A) polymerase Pla1<sup>6–8</sup>. The subunit conservation indicates that MTREC is the fission yeast counterpart to the human PAXT (pA-tail-exosome-targeting) complex that targets long and polyadenylated nuclear RNAs<sup>7,9–11</sup>. Indeed, PAXT contains a large zinc-finger protein ZFC3H1 and the RNA helicase MTR4, human homologs of MTREC Red1 and Mtl1, as well as the cap-binding and poly(A)-binding sub-modules<sup>9–11</sup>. The well-studied human NEXT (nuclear

<sup>1</sup>Univ. Grenoble Alpes, CNRS, CEA, IBS, F-38000 Grenoble, France. <sup>2</sup>Institut for Advanced Biosciences, UMR Inserm U1209/CNRS 5309/University Grenoble Alpes, La Tronche, France. <sup>3</sup>Structural Biology Group, European Synchrotron Radiation Facility (ESRF), CS 40220, 38043 Grenoble, France. <sup>4</sup>European Molecular Biology Laboratory, 71 Avenue des Martyrs, CS 90181, Grenoble Cedex 9 38042, France. <sup>5</sup>Univ. Bordeaux, Inserm U1212, CNRS UMR 5320, ARNA Laboratory, Institut Européen de Chimie et Biologie, 33607 Pessac, France. <sup>6</sup>These authors contributed equally: Anne-Emmanuelle Foucher, Leila Touat-Todeschini, Ariadna B. Juarez-Martinez. ✉e-mail: [cameron.mackereth@inserm.fr](mailto:cameron.mackereth@inserm.fr); [andre.verdel@univ-grenoble-alpes.fr](mailto:andre.verdel@univ-grenoble-alpes.fr); [jan.kadlec@ibs.fr](mailto:jan.kadlec@ibs.fr)



**Fig. 1 | Red1 forms multiple interactions within the MTREC complex.** **a** Schematic representation of the MTREC complex. Red1 and Mtl1 form the MTREC complex core<sup>7</sup>. **b** Y2H assays revealed that Red1 interacts with Ars2 and Iss10. Mtl1 was used as a positive control. Red1 was used as bait. Growth assayed on selective medium lacking histidine and with 50 mM 3-amino-1,2,4-triazole (3AT) histidine inhibitor is shown. EV- empty vector, BD- binding domain, AD- activating domain. **c** Y2H analysis of Red1 self-association. Growth was assayed on medium lacking histidine and with 30 mM 3AT. **d** Schematic representation of the interactions characterized in this study.

exosome-targeting) complex, on the other hand, is composed of the MTR4 helicase, zinc-knuckle protein ZCCHC8 and the RNA-binding subunit RBM7<sup>12</sup>. NEXT is involved in degradation of short-lived RNAs, such as promoter upstream transcripts (PROMPTs) and enhancer RNAs (eRNAs)<sup>12,13</sup> but its relation to MTREC is less clear.

The subunit composition of the *S. pombe* MTREC and human PAXT complexes has been characterized by proteomics<sup>6–9,11</sup>. The Red1 protein was shown to form a scaffold of the MTREC complex, interacting with multiple subunits and linking individual sub-modules together<sup>14</sup>. Red1 consists of 712 residues (compared to the 1989 residues of its ZFC3H1 counterpart in human PAXT) and is predicted to be mostly disordered. A crystal structure of a minimal Red1-Mtl1 complex, revealed how Red1 interacts with the Mtl1 Arch domain via its zinc-finger containing domain at its C-terminus<sup>14</sup>. In human and *S. cerevisiae*, Mtl1 counterpart MTR4 has been structurally characterized<sup>15</sup> revealing molecular details of its interactions with its partners including the NEXT subunit ZCCHC8<sup>16</sup> and the RNA exosome to whom MTR4 presents the targeted RNA for degradation<sup>17,18</sup>. Recently, using yeast two-hybrid (Y2H) and pull-down assays Dobrev et al. defined the minimal interacting regions involved in the interactions of Red1 with Iss10 of the Iss10-Mmi1 sub-module and with Ars2 of the cap-binding complex<sup>14</sup>. However, there is no 3D structural information available on the interactions that the Red1-Mtl1 core forms with the individual MTREC sub-modules. Similarly, the function of these modules in MTREC/PAXT activity still remains unclear.

The MTREC complex is involved in specific elimination of meiotic transcripts and heterochromatin formation in *S. pombe* and this activity involves the Iss10/Mmi1 sub-module<sup>6–8,19,20</sup>. Indeed, in *S. pombe*, expression of a number of meiotic genes is strictly repressed during the vegetative growth by selective RNA degradation, since their aberrant expression is highly deleterious to cell viability<sup>21</sup>. Meiotic transcripts and non-coding RNAs that contain repeats of a hexanucleotide motif called determinant of selective removal (DSR) are specifically recognized by the YTH family RNA-binding protein Mmi1, which in turn recruits MTREC to mediate their degradation<sup>20–24</sup>. In vegetatively growing cells, Mmi1 forms nuclear foci<sup>21</sup>, where most of its partners, including MTREC, nuclear exosome and the target RNA, colocalize<sup>6,8,25</sup>. The Iss10 subunit is also directly involved in the selective

elimination of meiotic DSR-containing transcripts and is important for the MTREC sub-nuclear localization<sup>26</sup>. In addition, Iss10 and the MTREC-dependent RNA degradation has been connected to the mTOR cellular energy-sensing signaling<sup>27</sup>. Under vegetative growth, Iss10 is phosphorylated by TORC1 (TOR complex 1), which prevents its degradation by the proteasome. However, under conditions of nitrogen depletion, TORC1 phosphorylation of Iss10 is blocked, and Iss10, in turn, eliminated<sup>27</sup>. Lack of Iss10 comes with a loss of MTREC colocalization with Mmi1 in nuclear foci, and accumulation of DSR-containing mRNAs and meiotic proteins, thereby allowing the cells to proceed with meiosis<sup>26,27</sup>. Thus, Iss10 is a key regulatory element of MTREC-dependent degradation of meiotic mRNAs, that most likely links the MTREC core complex to the Mmi1 bound RNAs<sup>26</sup>.

The Red1-Mtl1 core also recruits the Cbc-Ars2 sub-module<sup>6,7</sup>, which has been more extensively studied in human. CBC (nuclear cap-binding complex) is a key RNA biogenesis regulator that co-transcriptionally binds the m<sup>7</sup>G cap at the 5' end of RNA polymerase II transcripts<sup>28,29</sup>. The CBC and its partner ARS2 are implicated in fate determination of Pol II transcripts, including their processing, transport or degradation<sup>30</sup>. CBC-ARS2 interacts in a mutually exclusive way with positive RNA biogenesis factors (such as PHAX, FLASH and NCBP3), or with the RNA destructive MTREC/PAXT and NEXT assemblies<sup>30–32</sup>. While the connection to PAXT and NEXT involves ZC3H18<sup>9</sup>, no such protein has been identified in the MTREC complex<sup>6,7</sup>. The atomic details underlying how the CBC-ARS2 complex is recruited to the exosome-linked RNA targeting complexes and what role this module plays are poorly understood.

Here, we report a biochemical and structural analysis of two important interactions of Red1 with MTREC sub-modules, combined with *in vivo* studies. We determined an NMR structure of the Red1-Iss10 complex, which reveals how Red1 recognizes Iss10 via its helix-turn-helix motif. We also report a crystal structure of the Red1-Ars2 complex, which shows how the Ars2 C-terminal zinc-finger domain interacts with the conserved EDGE1 motif of Red1. In addition, we demonstrate equivalent interaction between ZFC3H1 and hARS2 within the PAXT complex, providing structural insights into how the Cbc-Ars2 complex is integrated into the RNA-targeting complexes. Functional *in vivo* experiments using Red1 structure-driven mutants showed that both Iss10 and Ars2 interactions are important for MTREC-mediated cellular activities. Finally, we demonstrate that Red1 dimerizes via its coiled-coil region, strongly suggesting that the MTREC complex acts as a dimeric assembly and this feature is conserved in human PAXT and NEXT complexes.

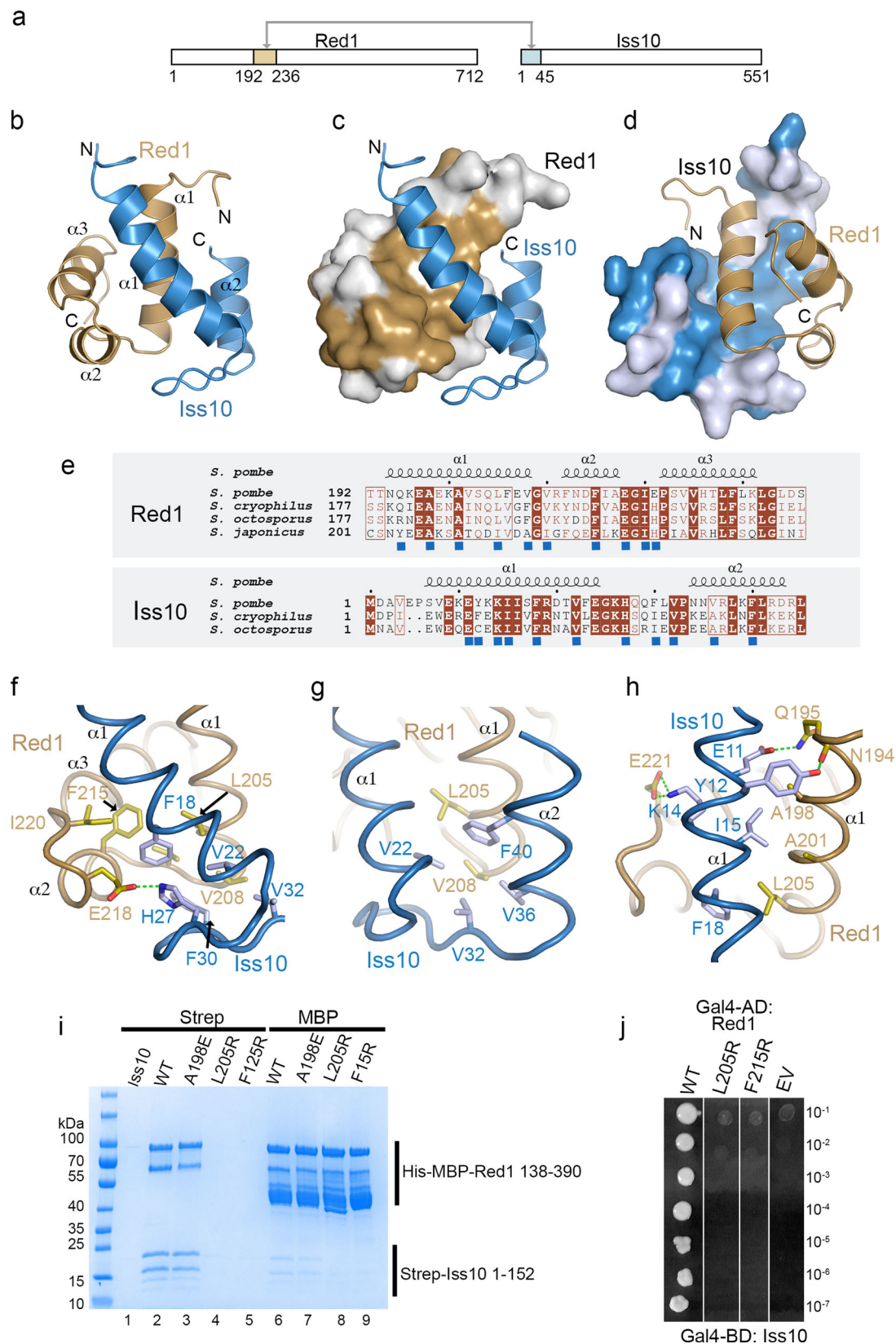
## Results

### Yeast two-hybrid analysis of Red1 interactions within MTREC

Red1, together with the Mtl1 helicase, has been shown to form the MTREC core<sup>6,14</sup> (Fig. 1a), but the structural basis of its interactions with the MTREC sub-modules remains elusive. To understand how Red1 associates with its partners, we first performed Y2H assays to analyze its interactions with several MTREC members. In agreement with the recent Y2H analysis of the MTREC complex interactions<sup>14</sup>, we detected strong interactions with Ars2, Iss10 and Mtl1 as well as between Iss10 and Mmi1 (Fig. 1b and Supplementary Fig. 1a, b). In addition, our Y2H assay revealed an unexpected Red1 self-association (Fig. 1c). To obtain mechanistic insights into the MTREC/PAXT complex architecture, in this study, we performed structural and functional characterization of the Red1 interactions with Iss10 and Ars2 as well as of its self-association (Fig. 1d).

### NMR structure of the Red1-Iss10 complex

A Y2H and GST pull-down analysis recently revealed that Red1<sup>187–236</sup> and Iss10<sup>1–51</sup> represent the interacting regions of the complex Red1 forms with Iss10<sup>14</sup>. Our characterization of the complex using Y2H, Strep-tag pull-down, and limited proteolysis experiments identified equivalent



fragments Red1<sup>192-236</sup> and Iss10<sup>1-45</sup> (Supplementary Fig. 1c–e). Multi angle laser light scattering (MALLS) showed that these constructs form a homogenous complex sample with a 1:1 stoichiometry (Supplementary Fig. 1f).

The Red1<sup>192-236</sup>-Iss10<sup>1-45</sup> complex was then subject to structural analysis (Fig. 2a). While the complex resisted crystallization, we

obtained a high quality <sup>15</sup>N-HSQC NMR spectrum (Supplementary Fig. 1g) that prompted us to determine its structure by NMR. The resulting structural ensemble (Supplementary Table 1; Supplementary Fig. 2a, b) revealed that Red1 residues 192–236 and Iss10 residues 1–44 comprise the structured core of the interaction complex. Red1<sup>192-236</sup> folds into three short helices arranged in a way reminiscent



**Fig. 2 | Red1 interacts with Iss10 with its HTH domain.** **a** A schematic diagram highlighting the Red1 and Iss10 interacting regions. **b** Ribbon representation of the lowest energy NMR-based structure of the Red1-Iss10 complex. The Red1 HTH domain is shown in brown and the Iss10 N-terminus is in blue. The secondary structures are labeled. **c** Surface representation of Red1 with conserved surface residues shown in brown. **d** Surface representation of Iss10 with conserved surface residues highlighted in blue. **e** Sequence alignment of the Red1 HTH domain (upper panel) and Iss10 (lower panel) in different *Schizosaccharomyces* species. Identical residues are in brown boxes. Blue squares indicate residues involved in the interaction. **f** Details of the interaction between the Iss10 helix  $\alpha 1$  and the following loop

with Red1 helices  $\alpha 1$  and  $\alpha 3$ . **g** Details of the Red1 helix  $\alpha 1$  packing against the Iss10 surface made of helices  $\alpha 1$  and  $\alpha 2$ . **h** Coiled-coil interaction and hydrogen bonding between Red1 and Iss10. **i** Strep-tag pull-down experiments with Strep-Iss10<sup>1-45</sup> and His-MBP-Red1<sup>192-236</sup> mutants indicated above the lanes. Iss10<sup>1-45</sup> and Red1<sup>192-236</sup> were co-expressed, the cultures were split into two halves and the proteins purified on Strep-Tactin or amylose resins. In absence of the interaction with Red1, Strep-Iss10<sup>1-45</sup> is unstable. Source data are provided as a Source Data file. **j** Essential role of Red1 L205 and F215 in the interaction with Iss10 shown by Y2H assay. Growth was assayed on medium lacking histidine and with 30 mM 3AT.

to the classical helix-turn-helix (HTH) motif (Fig. 2b–d and Supplementary Fig. 2c). Indeed, a search of the Protein Data Bank using the PDBeFold server at the EBI revealed that Red1<sup>192-236</sup> is highly similar to other HTH motifs, such as the DNA binding domain of the *B. subtilis* DeoR protein<sup>33</sup> (Supplementary Fig. 2c–e). Nevertheless, the surface of Red1 helix  $\alpha 3$ , that in classical HTH motifs frequently mediates dsDNA binding, carries mostly hydrophobic residues (Supplementary Fig. 2f). While Red1 DNA binding has not been reported and no unspecific DNA binding was observed during Red1 purification, specific DNA interaction cannot be excluded. The Iss10<sup>1-45</sup> fragment consists of two helices, packing tightly against Red1 (Fig. 2b–d). Mutual interaction surfaces on both proteins are made of residues that are highly conserved among evolutionary distant *Schizosaccharomyces* species underlining the importance of this interaction (Fig. 2c–e).

### The Red1-Iss10 interface

The interface between Red1 and Iss10 is dominated by hydrophobic interactions. Iss10 helix  $\alpha 1$  and the following loop form numerous contacts with the groove between Red1 helices  $\alpha 1$  and  $\alpha 3$  (Fig. 2f). Iss10 F18, V22, F30, and V32 interact with the Red1 hydrophobic surface composed of L205, V208, V210, F215 and I220. E218 also makes a salt bridge with H27 (Fig. 2f). In parallel, Red1 helix  $\alpha 1$  packs against the Iss10 surface made of helices  $\alpha 1$  and  $\alpha 2$ . Red1 L205 and V208 interact with Iss10 residues around V22, V32, V36 and F40 (Fig. 2g). Finally, Iss10 and Red1 establish a short coiled-coil interaction, including contacts between Red1 A198, A201 and L205 and Iss10 I15 and F18 (Fig. 2h). These interactions are further strengthened by several hydrogen bonds of Iss10 E11, Y12 and K14 and Red1 N194, Q195 and E221 (Fig. 2h).

In order to assess the role of the Red1-Iss10 interaction in vivo, we were then interested in identifying Red1 mutations that could disrupt its interaction with Iss10. To this end, we generated an A198E mutation in the center of the coiled-coil formed with Iss10 and two L205R and F215R mutations that we hypothesized should directly disrupt the interaction with Iss10 helix  $\alpha 1$  (Fig. 2f–h). These mutations were introduced into the His-MBP-Red1<sup>138-390</sup> and tested for their interaction with Iss10<sup>1-152</sup> by using Strep- and MBP- pull-down assays (Fig. 2i). While A198E retained the WT level of binding (Fig. 2i, lanes 3 and 7), the L205R and F215R mutants no longer bound Iss10 (Fig. 2i, lanes 4, 5, 8 and 9). Since co-expression with Red1 is required for Iss10<sup>1-152</sup> stability, Iss10<sup>1-152</sup> produced with these Red1 mutants becomes unstable and does not purify employing the Strep-Tactin resin. To assess the possible impact of the L205R and F215R mutations on the Red1 HTH domain, we produced Red1<sup>192-236</sup>, but 1D NMR analysis revealed that in absence of Iss10, the WT and mutated variants are unfolded (Supplementary Fig. 2g). Together with size exclusion chromatography, the 1D NMR spectra confirmed that both L205R and F215R mutants are similar in behavior to WT Red1<sup>192-236</sup> being soluble with no sign of aggregation or higher order oligomerization (Supplementary Fig. 2g, h). We then confirmed the essential role of L205 and F215 for the interaction with Iss10 in the context of full-length proteins using an Y2H assay (Fig. 2j). Ultimately, the L205R mutation was selected for in vivo studies (see below).

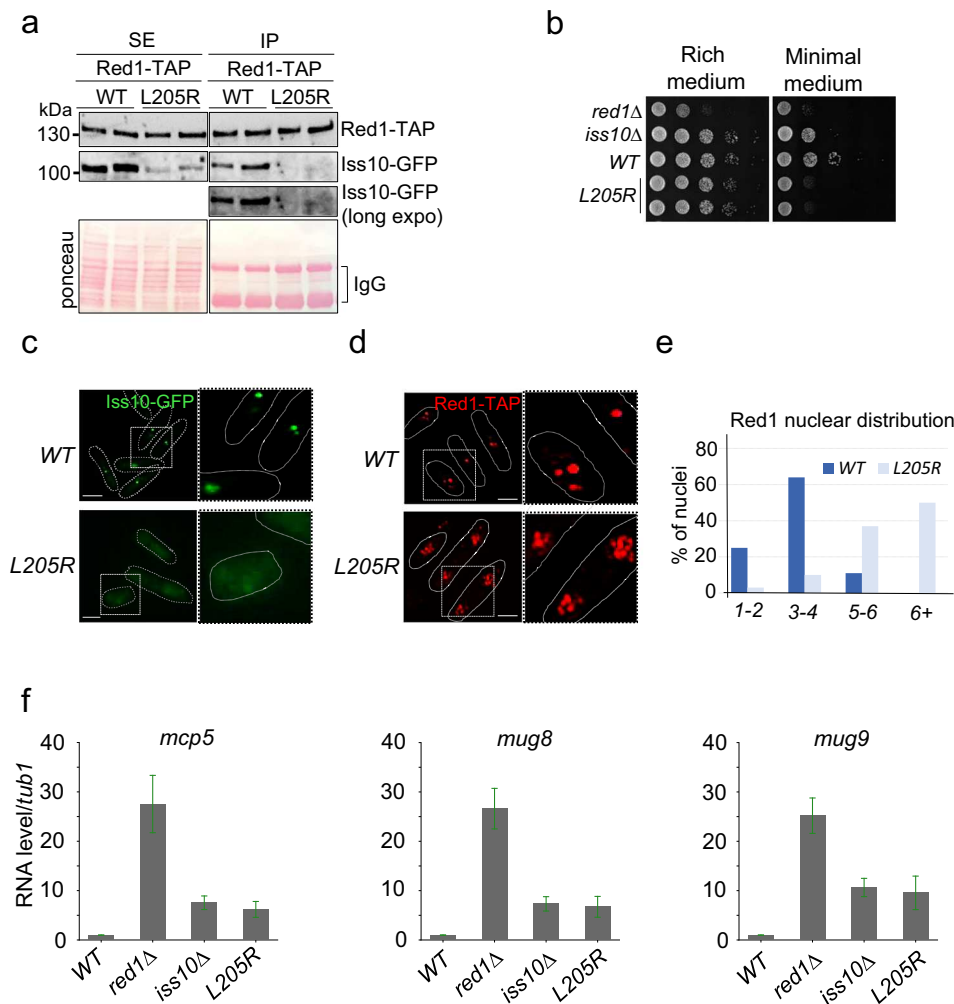
### Red1-Iss10 binding interface is required for proper degradation of meiotic DSR-containing mRNAs and clustering of Red1 foci in vivo

In order to evaluate in vivo the function of the Red1-Iss10 binding interface, recombinant Red1-TAP or Red1-L205R-TAP proteins in combination with Iss10-GFP protein were expressed in *S. pombe* cells in place of the endogenous proteins from the respective endogenous promoters. Using these cells, reciprocal co-immunoprecipitation experiments showed that the Red1-L205R mutation also compromised Red1 association with Iss10 in *S. pombe* (Fig. 3a and Supplementary Fig. 3a). Next, we examined the growth of *red1-L205R* mutant cells. While *red1Δ* cells showed only a moderate growth defect on solid rich medium at 30 °C, this defect was more pronounced at lower temperature of 25 or 18 °C (Supplementary Fig. 3b), as published previously<sup>14,34</sup> or when the cells were grown on minimal medium (Supplementary Fig. 3c). Importantly, *red1-L205R* cells showed a similar growth defect, when grown on minimal medium, suggesting that Red1-Iss10 interaction can be important for Red1-dependent cell growth function depending on the environmental conditions (Fig. 3b). We then examined the subcellular localization of Red1-L205R-TAP mutant and Iss10-GFP. Both Red1 and Iss10 have been previously reported to co-localize within nuclear foci<sup>23,26,34</sup>. We observed that the *red1-L205R* mutation had an effect on both Iss10 and Red1 localization by, respectively, inducing the disappearance of Iss10 nuclear foci and the increase in the number of Red1 foci per nucleus (Fig. 3c–e). These findings indicate that the interaction of Iss10 with Red1 is required for Iss10 recruitment to Red1 nuclear foci and suggest that it may promote the clustering of Red1 nuclear foci. Finally, we examined the level of meiotic DSR-containing mRNAs which are degraded by the MTREC machinery<sup>7</sup>. Expression of the Red1-L205R mutant, instead of WT Red1, led to a defect comparable to *iss10Δ* mutation, and it was more pronounced when cells were grown on minimal medium (Fig. 3f and Supplementary Fig. 3d). Taken together, we conclude that in vivo, the Red1-Iss10 binding interface plays a role in mediating functions of Red1, and its importance is determined by the environmental growth conditions.

### Crystal structure of the Ars2-Red1 complex

The interaction between Ars2 and Red1 has recently been shown, by Y2H and pull-down assays, to be mediated by a short conserved “EDGE1” motif of Red1<sup>14</sup>. Using Y2H and gel filtration chromatography we confirmed the implication of the EDGE1 motif in the interaction with full-length Ars2, identifying a slightly longer construct Red1<sup>20-40</sup> (Fig. 4a and Supplementary Fig. 4). To further characterize the interaction between Ars2 and this Red1 motif, we used isothermal titration calorimetry (ITC). Both Red1<sup>1-54</sup> and Red1<sup>20-40</sup> (with and without MBP tag) bound Ars2 with an equivalent dissociation constant ( $K_d$ ) of about 5  $\mu$ M (Fig. 4b and Supplementary Fig. 5a, b), confirming that additional sequences surrounding EDGE1 motif do not enhance the binding affinity for Ars2.

*S. pombe* Ars2 consists of 609 residues and in analogy to hARS2 is predicted to contain intrinsically disordered N- and C-termini and a central structured region. Ars2 residues 450–516 were shown to bind Red1 by Y2H<sup>14</sup>, but in our hands short constructs encompassing this



**Fig. 3 | Red1<sup>L205R</sup> mutation affects cell growth, degradation of DSR-containing mRNAs and clustering of Red1 foci.** **a** Co-immunoprecipitation experiments, analyzed by western blot, showing that Red1<sup>L205R</sup> mutation compromises Red1-Iss10 interaction in vivo. Soluble cell extracts (SE) were prepared from exponentially growing cells expressing Red1-TAP and Iss10-GFP in liquid minimal medium. **b** Plating assays monitoring the growth of *red1-L205R* compared to *WT*, *iss10Δ* and *red1Δ* cells. Cells were grown on solid rich or minimal media for 5 days at 25 °C. **c** Images of live fluorescence microscopy showing the localization Iss10-GFP in *WT* and *red1-L205R* cells. Scale bar = 5 μm. **d** Images of immuno-fluorescence

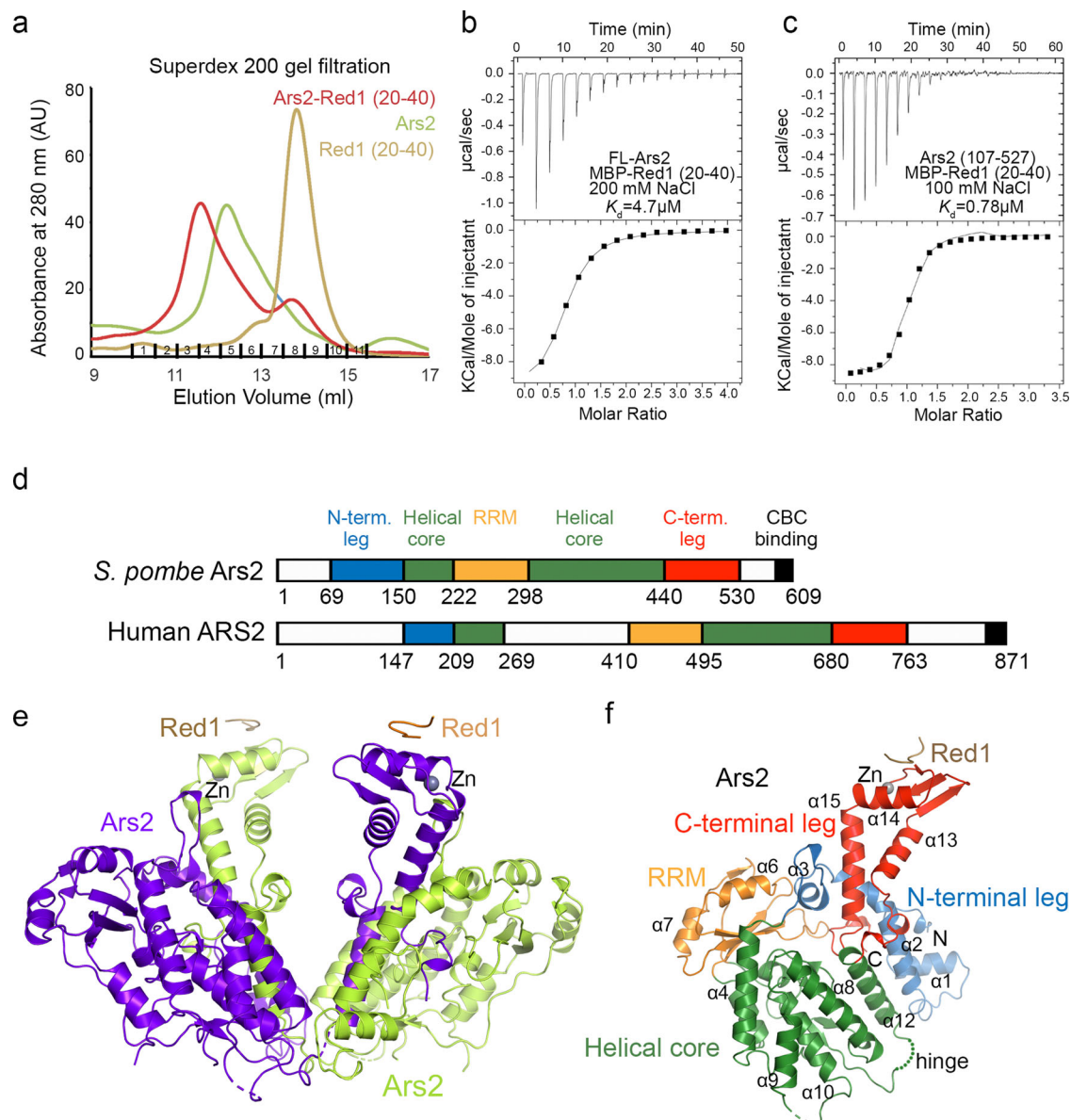
microscopy showing Red1-TAP localization in *WT* and *red1-L205R* cells. Scale bar = 5 μm. **e** Graph showing the number of Red1 foci per nucleus in *WT* and *red1-L205R* cells. A minimum of 100 nuclei from two different biological isolates were counted. **f** Quantitative RT-PCR experiments showing the level of *mcp5*, *mug8* and *mug9*, three meiotic and DSR-containing mRNAs normalized to *tub1* mRNA in *WT*, *red1Δ*, *iss10Δ* and *red1-L205R* cells. Total RNAs were purified from cells growing exponentially in liquid minimal medium. Error bars represent standard deviation from three independent biological replicates. Source data are provided as a Source Data file.

region of hARS2 did not produce stable protein. Based on deletion mutagenesis, limited proteolysis and sequence comparisons to hARS2, we generated the Ars2<sup>107–527</sup> construct lacking both disordered termini, which turned to be more stable than the full-length protein and retained its interaction with Red1. The strength of the interaction depends on salt concentration, since when NaCl concentration was reduced from 200 to 100 mM, the  $K_d$  measured by ITC for Ars2<sup>107–527</sup> and Red1<sup>20–40</sup> increased from 5 to 0.8 μM (Fig. 4c).

To gain molecular insights into the Ars2-Red1 interaction, we set out to determine the atomic structure of the complex by X-ray crystallography. First, we obtained crystals of the complex between Ars2<sup>107–527</sup> and Red1<sup>20–40</sup>, that however only diffracted to 5 Å. Inspired by the recently available AlphaFold2 model of *S. pombe* Ars2 (AF-094326)<sup>35</sup>, we then extended both the N- and C-terminal limits of the Ars2<sup>107–527</sup> construct and removed one predicted mobile loop (residues 184–205). This new construct was produced by co-expression of its two fragments - Ars2<sup>68–183</sup> and Ars2<sup>206–531</sup>. Two separate elution peaks of Ars2 were obtained by size exclusion chromatography (Supplementary Fig 5c, d) and MALLS analysis identified them as a homodimer and

a monomer, respectively (Supplementary Fig. 5e, f). Only the minor dimeric version of Ars2 in complex with Red1<sup>20–40</sup> yielded single crystals, which diffracted to a 2.8 Å resolution. The structure was determined by molecular replacement using the AlphaFold2 Ars2 model (AF-094326) and refined to an  $R_{\text{free}}$  of 30.2% and an  $R_{\text{work}}$  of 24.5% (Supplementary Table 2).

The *S. pombe* Ars2 structure consists of four defined domains, including a helical core with an inserted RNA recognition motif (RRM) and is flanked by two extended, mostly helical structures. These extensions are referred to as the N- and C-terminal leg, in similarity to its human counterpart<sup>32</sup> (Fig. 4d–f). Ars2<sup>68–183,206–531</sup> crystallized as a domain-swapped dimer, having the C-terminal leg segment exchanged between the two protomers (Fig. 4e). The hinge region, allowing the rotation of the C-terminal leg, corresponds to residues 420–421 (Supplementary Fig. 7a). Analysis of the extensive dimer interface with PDBePISA server<sup>36</sup> at the EBI revealed that it buries 4946 Å<sup>2</sup> on each Ars2 protomer, including 37 hydrogen bond and 18 salt bridge contacts between the two molecules and has a Complex Formation Significance Score (CSS) of 1, indicating that the interface plays an



**Fig. 4 | Crystal structure of the Ars2-Red1.** **a** Overlay of Superdex 200 gel filtration elution profiles of full-length (FL) His-MBP-Ars2, His-MBP-Red1<sup>20-40</sup> and their complex. The proteins were expressed individually as MBP fusions and purified on amylose resin. Both proteins were first purified on Superdex 200. **b** ITC measurement of the interaction affinity between FL Ars2 and His-MBP-Red1<sup>20-40</sup> in presence of 200 mM NaCl. **c** ITC measurement of the interaction affinity between Ars2<sup>107-527</sup> and His-MBP-Red1<sup>20-40</sup> in presence of 100 mM NaCl. **d** Schematic representation of the domain structure of *S. pombe* and human Ars2, as defined in this study and in Schulze et al.<sup>32</sup>. The CBC binding site was characterized in<sup>43</sup>. **e** Ribbon

representation of the overall structure of the Ars2<sup>68-183,206-531</sup> dimer. The C-terminal region (420–531) corresponding to the C-terminal leg is swapped between the protomers. The interacting peptide of Red1 is shown in brown. **f** Crystal structure of the Ars2 monomer with C-terminal leg not swapped in complex with Red1. The four defined domains are highlighted in different colors. Red1 peptide is in brown. The hinge region, allowing for the mobility of the C-terminal domain is shown. Annotated secondary-structure elements correspond to those shown in the alignment of Supplementary Fig. 6.

essential role in the dimeric complex formation. The domain swap occurred likely already during the overexpression in bacteria as it could not be reproduced even with high concentrations (up to 40 mg/ml) of the monomeric Ars2. Since it is not known, whether such an Ars2 dimer has a physiological significance, in the following text, the monomeric Ars2 structure, without the domain swap, will be considered.

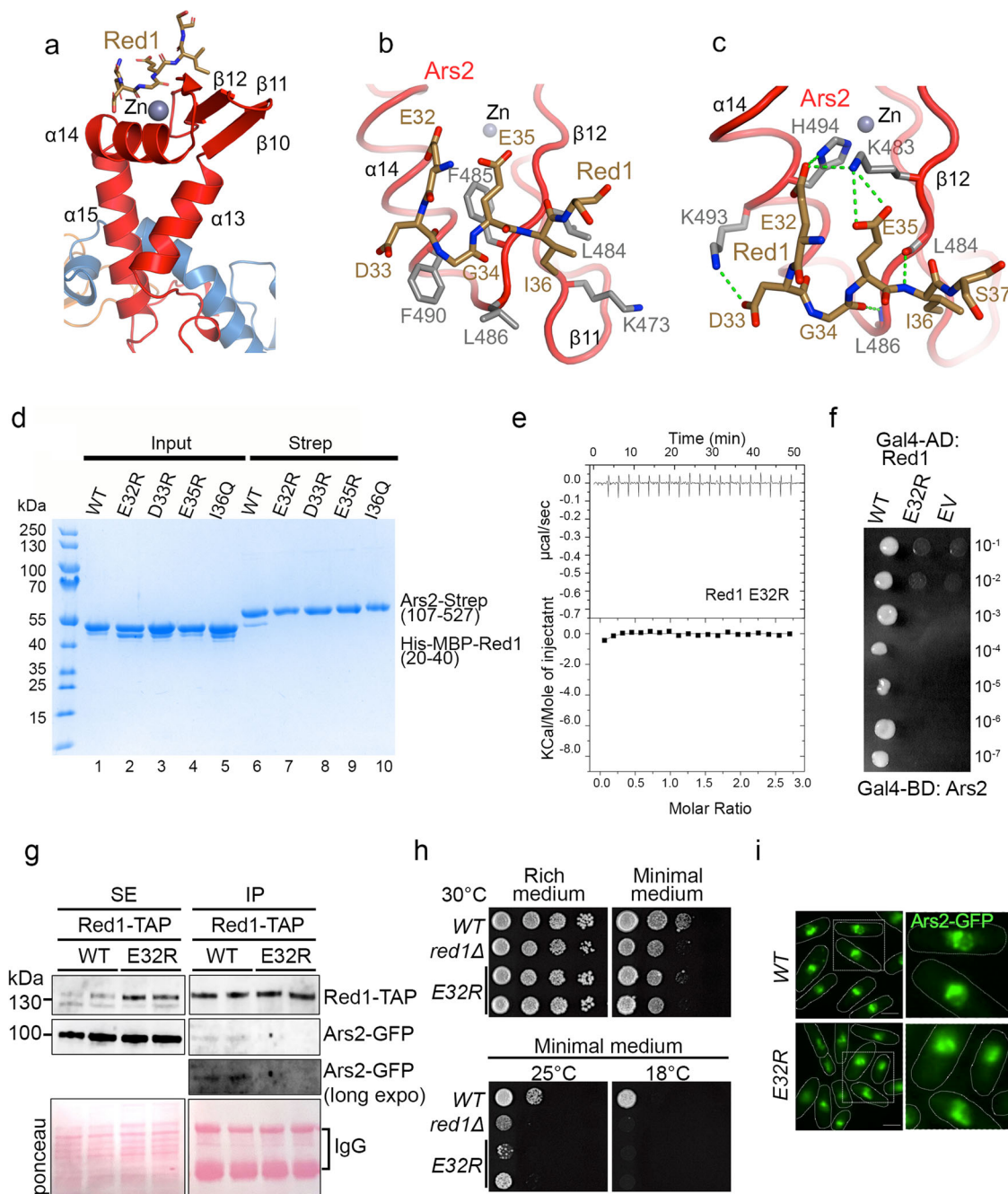
### The Ars2-Red1 interface

Red1 interacts with the Zn-finger domain of the Ars2 C-terminal leg, particularly with the surface formed by  $\alpha 14$ ,  $\beta 11$  and  $\beta 12$  (Fig. 5a). In our Ars2-Red1 complex structure, electron density is visible for six out of the twenty Red1 peptide residues (Supplementary Fig. 7b). The peptide

adopts a U-shape structure bent around the central G34 and establishes hydrophobic and charged contacts with Ars2, burying 300 Å<sup>2</sup> of its surface. Red1 G34 and I36 together with aliphatic portions of D33 and E35 side chains pack against the Ars2 hydrophobic surface generated by K473, L484, F485, L486 and F490 (Fig. 5b). In addition, the three negatively charged residues E32, D33 and E35 make several hydrogen bonds and salt bridge interactions with Ars2 K483, K493, and H494 (Fig. 5c). The electrostatic nature of these interactions is consistent with the variation of the  $K_d$  with salt concentration (Fig. 4b, c). Finally, Red1 I36 and G34 form main chain interactions with Ars2 L484 and L486 (Fig. 5c).

In order to identify the key Ars2 and Red1 residues for stabilizing the interaction, we mutated Ars2 K483, which forms salt bridge





**Fig. 5 | Red1-Ars2 interaction is necessary for proper cell proliferation and Ars2 nuclear localization into Red1 foci.** **a** The Red1 peptide (in brown) interacts with the Zn-finger domain of the Ars2 C-terminal leg (in red). **b** Details of the interaction between the Red1 EDGEI motif and the Ars2 surface formed by  $\alpha 14$ ,  $\beta 11$  and  $\beta 12$ . Hydrophobic contacts between Red1 and the Ars2 residues are shown. **c** Residues of the Red1 EDGEI motif form several side and main chain hydrogen bonds and salt bridge interactions with the Ars2 C-terminus. **d** Strep-tag pull-down experiments with Strep-Ars2<sup>107-527</sup> and His-MBP-Red1<sup>20-40</sup> mutants indicated above the lanes. Ars2 and Red1 proteins were first purified using Ni<sup>2+</sup> resin (lanes 1–5). Ars2 with either WT or mutant Red1 were then co-purified on Strep-tactin resin (lanes 6–10). All mutations abolished the interaction. **e** ITC measurement of the interaction affinity

between Strep-Ars2<sup>107-527</sup> and the His-MBP-Red1<sup>20-40</sup> E32R mutant in presence of 100 mM NaCl. **f** Essential role of Red1 E32 in the interaction with Ars2 as shown by Y2H. Growth was assayed on medium lacking histidine and with 30 mM 3AT. **g** Co-immunoprecipitation experiments of Red1-TAP and Ars2-GFP in WT and *red1-E32R* cells, analyzed by western blot. Soluble cell (SE) extracts were prepared from exponentially growing cells in minimal liquid medium. **h** Plating assays monitoring the growth of WT, *red1Δ* and *red1-E32R* cells on rich and minimal media at different temperatures. **i** Images of live cell fluorescence microscopy showing Ars2-GFP localization in WT and *red1-E32R* cells. Scale bar = 5  $\mu$ m. Source data are provided as a Source Data file.

interactions with Red1, as well as F490, which is central to the hydrophobic contacts between the two proteins to aspartates (Fig. 5b, c). The integrity of the Ars2 mutants was verified by gel filtration, with an elution profile similar to the WT protein (Supplementary Fig. 7c). In Strep-tag pull down assays, both mutations (K483D, F490D) essentially disrupted the Red1 binding (Supplementary Fig. 7d, lanes 3, 4). We also

mutated key Red1 residues including Glu32, Asp33, Glu35 and Ile36 within the Red1 peptide (20–40). In pull-down assays with Strep-tagged Ars2<sup>107-527</sup>, individual mutations of all four residues abolished the binding (Fig. 5d, lanes 7–10), confirming the importance of these residues for the interaction. In addition, the role of E32R was supported by ITC measurements, since no binding to Ars2<sup>107-527</sup> was

observed for the peptide containing the E32R mutation (Fig. 5e). In a previous study, an E32A, D33A, E35A triple mutant was required to disrupt the binding between full length Red1 and Ars2<sup>14</sup>. Using Y2H assays, we could show that the Red1 E32R mutation is sufficient to prevent the interaction with Ars2 in the context of full-length proteins (Fig. 5f). The structure-guided mutagenesis thus validated the key residues observed in the structure for the Red1-Ars2 interaction and identified the Red1 E32R mutation as a suitable candidate for functional analysis.

### In vivo, *red1-E32R* mutation affects Red1-Ars2 interaction, cell growth and Ars2 localization in nuclear foci

We then examined the function of Red1-Ars2 binding interface in vivo. We assessed the impact of the *red1-E32R* point mutation on Red1-Ars2 interaction by co-immunoprecipitation experiments. Recombinant Red1-TAP, Red1-E32R-TAP and Ars2-GFP proteins were expressed instead of endogenous proteins and from respective endogenous promoters. Importantly, while Red1-TAP and Ars2-GFP interact, as expected, this interaction was lost in cells expressing Red1-E32R-TAP (Fig. 5g). Next, we conducted growth assays on different solid media and at different temperatures. As observed previously, the *red1Δ* cells experience a growth defect, which is exacerbated on minimal medium (Fig. 5h and Supplementary Fig. 3b, c) or at lower temperature (Supplementary Fig. 8a; refs. 14, 34). A similar growth defect was visible for the *red1-E32R-TAP* mutant cells grown on minimal medium, suggesting that Red1-Ars2 interaction, depending on the growth conditions, can be important for the normal growth of the cell population (Fig. 5h). As MTREC degrades a specific population of RNAs such as meiotic DSR-containing mRNAs and PROMPTS/CUTS<sup>7</sup>, we asked whether the levels of these transcripts changed in *red1-E32R* mutant cells. However, no significant changes were observed for several meiotic mRNAs and PROMPTS/CUTS MTREC targets (Supplementary Fig. 8b, c). It has been demonstrated that Ars2 regulates *pho1* and *byr2* protein-coding genes<sup>37</sup>. We examined the possible effect of *red1-E32R* mutation on *pho1* and *byr2* mRNAs. However, no significant changes in *pho1* and *byr2* mRNA levels were observed between wild-type and *red1-E32R* mutant cells (Supplementary Fig. 8d). We also assessed the relevance of Red1-Ars2 interaction for their respective localization in nuclear foci. In cells expressing the Red1-E32R-TAP mutant protein, Ars2-GFP signal also formed nuclear foci like in wild-type cells but they tended to be more difficult to distinguish from the more even nuclear GFP signal when compared to wild-type cells (Fig. 5i). Interestingly, quantification of the intensity of the signal within each nucleus showed a reduction of Ars2-GFP nuclear signal in *red1-E32R* mutant cells compared to wild-type cells, suggesting that some of Ars2-GFP proteins diffuse to the cytoplasm (Supplementary Fig. 8e). The localization of Red1-TAP in nuclear foci did not significantly change (Supplementary Fig. 8f, g). Thus, in vivo, Red1-Ars2 binding interface is critical for Red1-Ars2 interaction, for optimal cell growth and nuclear localization of Ars2 where Red1 and other subunits of MTREC localize.

### Comparison of *S. pombe* and human Ars2

Compared to human ARS2, the mutual orientations of the domains in *S. pombe* Ars2 are considerably different. When the two proteins are aligned via their helical domains, large differences in positions of the N- and C-terminal legs and to a lesser extent of the RRM domain are apparent (Fig. 6a, b). Similar differences are also observed when compared to the structure of the *A. thaliana* ARS2 ortholog known as SERRATE<sup>38</sup>. The folds of individual domains, however, are rather well conserved between *S. pombe* and human (Supplementary Fig. 9a–d). Thanks to extending the N- and C-terminal limits of the construct used, the *S. pombe* Ars2 structure shows a more elaborate N-terminal leg, where the extreme N terminus (residues 69–80) packs against helices  $\alpha 1$  and  $\alpha 2$  on one side and interacts with C-terminal residues 527–530 on the other one (Supplementary Fig. 9e). The remaining part of the

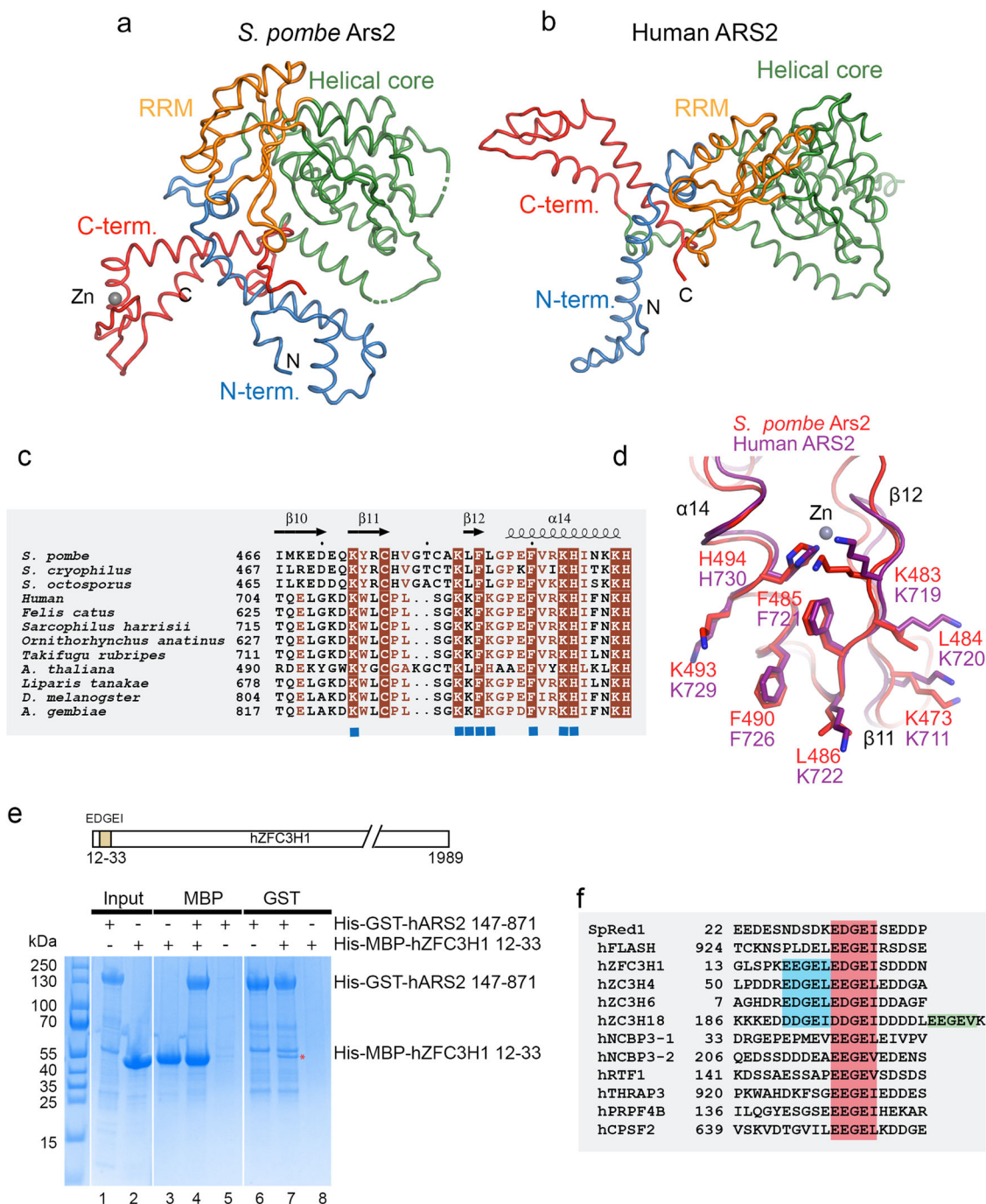
N-terminal leg resembles the one of human ARS2 (Supplementary Fig. 9a). The helical core of Ars2 has a similar topology to its human counterpart (Supplementary Fig. 9b) and has the RRM domain inserted after helix  $\alpha 5$ . While the RRM domain  $\beta$ -sheet surfaces frequently carry exposed aromatic residues in the two RNPI and RNP2 motifs that are involved in the binding of ssRNA sequences<sup>39</sup>, the Ars2 RRM possesses only one W267 located on  $\beta 3$  (RNPI) (Supplementary Fig. 9c). In addition, the  $\beta$ -sheet and W267 are packed against  $\alpha 11$  of the helical core, making it unavailable for interaction with RNA. The C-terminal leg can also be well superimposed onto the human C-terminus (Supplementary Fig. 9d). However, an important difference is the presence of a Zn atom coordinated by C476, C481, H494, and H499 in *S. pombe*, which is lacking in human ARS2, due to replacement of C481 by a serine (Fig. 6c). The Zn co-ordination is preserved in *A. thaliana* ARS2/SERRATE<sup>38</sup>. The mutual orientation of Ars2 domains is unlikely to be affected by the C-terminal leg swap, as observed in our low resolution structure of Ars2<sup>107–527</sup>, where no domain swap occurs, while the Zn-finger region and the N-terminal helix  $\alpha 2$  is partially disordered (Supplementary Fig. 9f).

### The Red1-Ars2 interface is conserved in human

It has been reported that the EDGE1 motif is conserved in the human PAXT subunit ZFC3H1, a putative human counterpart of Red1, implying that an analogous interaction between ARS2 and ZFC3H1 might exist within the PAXT complex<sup>14</sup>. Correspondingly, the Ars2 residues that interact with the Red1 EDGE1 motif are generally very well conserved across species (Fig. 6c, d) suggesting that the interaction of Ars2 with this motif could be universal. The only exceptions are *S. pombe* L484 and L486, which are both replaced by lysine in most other species that likely retain hydrophobic contacts with I36 (Fig. 6c, d and Supplementary Fig. 10a). To assess the conservation of the interface in human PAXT, we used pull-down experiments with GST-tagged human ARS2<sup>147–871</sup> and MBP-tagged ZFC3H1<sup>12–33</sup> which revealed a direct interaction between ARS2 and ZFC3H1 mediated by the conserved ZFC3H1<sup>12–33</sup> motif (Fig. 6e, lanes 4 and 7). Importantly, the E23R mutation analogous to the above-described E32R mutation in Red1 (Fig. 5d–f) severely reduced ZFC3H1<sup>12–33</sup> binding to hARS2 in an MBP-pull down assay (Supplementary Fig. 10b, lane 3). An I27L mutation did not perturb the binding (Supplementary Fig. 10b, lane 4), indicating that isoleucine in position 27 is not strictly required and can be replaced by a leucine. We could also show that a short peptide containing only 12 residues (ZFC3H1<sup>20–31</sup>) is sufficient for the interaction with ARS2 (Supplementary Fig. 10b, lane 2). To further confirm the conservation of the Ars2-Red1 interface, we used MBP-pull down assays, where the GST-tagged human ARS2 (residues 147–871) bound equally well to MBP-Red1 or MBP-ZFC3H1 “EDGE1” motif-containing peptides (Supplementary Fig. 10c).

In addition to ZFC3H1, human ARS2 also interacts via a similar “EEGE1” motif with FLASH, a large protein involved in histone mRNA biogenesis<sup>32,40</sup>. FLASH<sup>931–943</sup> binds ARS2<sup>147–871</sup> with a  $K_d$  of 0.5–5  $\mu$ M depending on salt concentration<sup>32</sup>, reminiscent of the Red1-Ars2 binding. Accordingly, in the human system, a K719A, K722, K734A triple mutant located at the EDGE1 motif-binding surface of ARS2 (Supplementary Fig. 10a) abolished the hARS2-FLASH interaction<sup>32</sup>.

A search for other human proteins containing possible ARS2-binding motifs with SlimSearch<sup>41</sup>, identified several RNA biogenesis factors (Fig. 6f). These include three other Zn-finger proteins - ZC3H4, ZC3H6 and ZC3H18. Similarly to ZFC3H1, they contain two or three tandem copies of the motif. The strong effect of the single E23R mutation in one of the ZFC3H1 motifs (Supplementary Fig. 10b, lane 3) might indicate that individual motifs interact with different affinities depending on surrounding residues. NCBP3, linked to mRNA export<sup>42</sup> also possesses two separated copies of the motif. Notably, NCBP3 was shown to directly bind ARS2, and the C-terminal leg of ARS2 was



**Fig. 6 | The ARS2-Red1 interface is conserved in human. a** Comparison of the *S. pombe* and **b** human ARS2 structures. The proteins were superimposed using the helical core domain (in green) with r.m.s.d. = 2.8 Å for 153 C $\alpha$  atoms. **c** Sequence alignment of ARS2 proteins. Only the Zn-finger domain of the C-terminal leg is shown. Identical residues are in brown boxes. Blue squares indicate residues involved in the interaction with Red1. **d** The Red1 binding surface of *S. pombe* ARS2 is conserved in human ARS2. The secondary structure element labeling corresponds to *S. pombe* ARS2. **e** MBP- and GST-tag pull-down experiments with His-GST-tagged human ARS2<sup>147-871</sup> and His-MBP-tagged ZFC3H1<sup>12-33</sup> showing,

respectively, a mutual interaction (lanes 4 and 7). The proteins were first purified using Ni<sup>2+</sup> resin (lanes 1, 2) and then co-purified on either Amylose or GST resins. The position of the EDGEI motif within hZFC3H1 is shown. Source data are provided as a Source Data file. **f** Sequence comparison of Red1 with human proteins containing possible ARS2-binding motifs. The ARS2 binding motif is highlighted in red. Additional copies of the motif are shown in blue and green. The SlimSearch<sup>41</sup> sequence analysis was done with a more relaxed sequence of the motif [ED][ED]GE[ILV], allowing also an aspartate in position 1 and a valine in position 5, likely compatible with the crystal structure of the Red1-ARS2 complex.

required for this interaction<sup>32</sup>. Finally, THRAP3, PRPF4B, RTF1, and CPSF2 contain one motif. In all these proteins, the conserved ARS2-binding motif is located in disordered regions as judged using AlphaFold2 predicted models. Importantly, ZFC3H1, ZC3H4, NCBP3, THRAP3, PRPF4B and CPSF2 were all found enriched in affinity purification/mass spectrometry (AP-MS) analysis with native hARS2

compared to the hARS2 K719A, K722A, K734A triple mutant in the ARS2 C-terminal leg<sup>32</sup>.

Together, these results show that the *S. pombe* ARS2-Red1 interface is well conserved among species. From these results, we also hypothesize, that in addition to hZFC3H1 and FLASH, the ARS2 ED/EGE/L binding surface is involved in connecting CBC-ARS2 bound-



RNAs to several other regulators of RNA metabolism, likely in a mutually exclusive way.

### Dimerization of Red1 C-terminus

To shed light onto the observed Red1 propensity to self-associate, we assessed which part of Red1 is required for its eventual oligomerization. In Y2H assays, the C-terminal Red1 construct (residues 390–712) interacted strongly with full-length Red1 (Fig. 7a). In addition, the C-terminal region was sufficient for the interaction when tested against the three Red1 fragments (Fig. 7a). We used AlphaFold2<sup>35</sup> to model the dimeric structure of the Red1 C-terminus. While most of the structure is predicted with low confidence, AlphaFold2 predicts an antiparallel dimeric coiled-coil corresponding to residues 470–524 with very high per-residue confidence score (pLDDT) and the predicted aligned error plot indicates high confidence of the mutual positions of the two protomers (Fig. 7b, Supplementary Fig. 11). Equivalent predictions were obtained for *S. octosporus* and *S. cryophilus* Red1 (Supplementary Fig. 11). The predicted structure revealed six complete heptad repeats engaged in classical coiled-coil packing (Fig. 7b–d). To test this prediction, a His-MBP-Red1<sup>452–524</sup> fusion protein was analyzed by size exclusion chromatography, where two elution peaks were observed (Supplementary Fig. 12a, b). MALLS revealed that the two peaks correspond to dimeric and monomeric Red1<sup>452–524</sup>, respectively (Fig. 7e and Supplementary Fig. 12c). The dimeric Red1<sup>452–524</sup> remained dimeric upon reinjection on the gel filtration column (Fig. 7e and Supplementary Fig. 12d). Importantly, an L481R, I485R double mutation, based on the structure model (Fig. 7d), is able to abolish the Red1 dimerization (Supplementary Fig. 12e, f). Using Y2H we could show that the L481R, I485R double mutation essentially disrupted the dimerization of full-length Red1, when one or both protomers are mutated (Fig. 7f), but did not impact the binding of Ars2 nor Iss10 (Supplementary Fig. 13a).

We then investigated whether Red1 dimerizes in vivo. We constructed strains co-expressing the proteins Red1-Myc<sub>13</sub> and Red1-HA<sub>3</sub>. Co-IP experiments showed that indeed Red1 dimerizes within *S. pombe* (Fig. 7g). Co-expression of Red1-Myc<sub>13</sub> together with Red1-ΔCC-HA<sub>3</sub> (Δcoiled-coil corresponds to the deletion of residues 452–524) or Red1-L481R-I485R-HA<sub>3</sub> mutants showed that both mutations compromise Red1 dimerization but do not completely disrupt the interaction (Fig. 7g). We made other strains co-expressing Red1-HA<sub>3</sub> constructions together with Red1-TAP and obtained similar results (Supplementary Fig. 13b). We next explored whether the unique expression of either Red1 dimerization mutants resulted in Red1 loss of function. However, cell growth assays showed that *red1-ΔCC* or *red1-L481R-I485R* mutant cells growth is similar to wild-type cells (Supplementary Fig. 13c). In addition, examination of the RNA levels of transcripts known to be degraded by MTREC showed no significant changes of their level (Supplementary Fig. 13d). Thus, these results show that Red1 dimerizes in vivo and that the Red1 coiled-coil domain contributes to this dimerization in *S. pombe*.

Together, these results provide evidence for a coiled coil-mediated homo-dimerization of Red1. Given that Iss10, Ars2 and Mtl1 each interact with Red1 in a 1:1 or 2:2 stoichiometry, it is thus possible that these subunits are present in two copies within the MTREC complex (Fig. 8a). Importantly, AlphaFold2 also predicts, with high accuracy, coiled coil-based homodimers for the PAXT and NEXT complex scaffolding proteins hZFC3H1 and ZCCHC8, respectively (Supplementary Fig. 14). The dimeric architecture might thus be a conserved feature among all these RNA targeting complexes (Fig. 8b and Supplementary Fig. 15).

### Discussion

The *S. pombe* MTREC complex and its human counterpart PAXT target specific and aberrant nuclear transcripts for degradation by the nuclear RNA exosome<sup>4,5</sup>. While the core RNA Mtl1/MTR4 helicase has

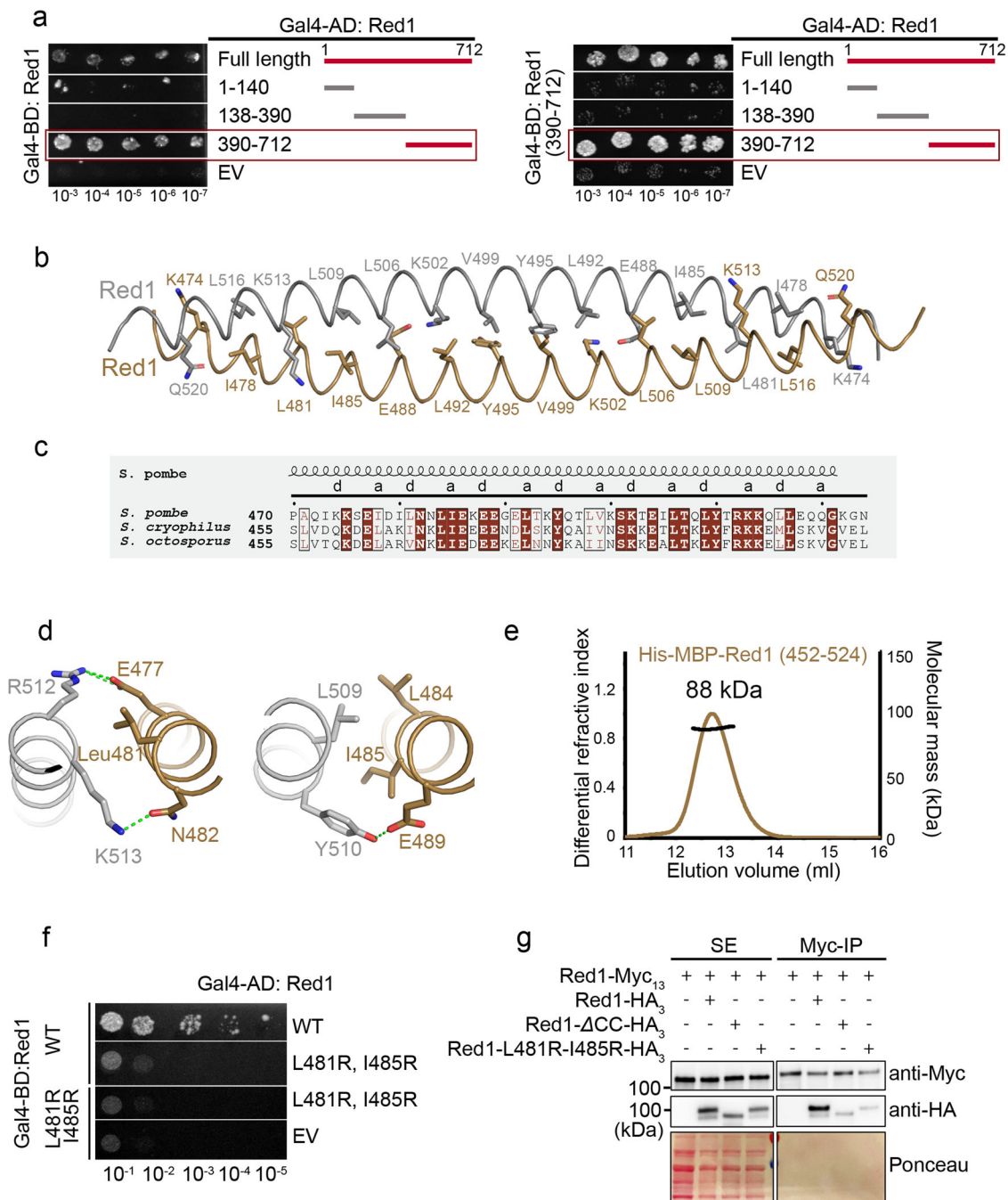
been extensively analyzed<sup>15</sup>, the exact roles of the remaining subunits remain unclear. In this study, we describe atomic details of the Red1 subunit interactions with Iss10 of the Iss10/Mmi1 sub-module and Ars2 of the cap-binding sub-module (Cbc1, Cb2, and Ars2). In addition, we report a coiled-coil mediated homo-dimerization of Red1. Together with the recently reported crystal structure of the complex between Red1 and Mtl1<sup>14</sup>, these results provide structural insights into the Red1 function as a scaffold of the MTREC complex. We also show that these MTREC features are conserved in the human analog PAXT complex.

Our NMR structure of the minimal Red1-Iss10 interaction complex revealed that Red1 uses a HTH motif to bind the Iss10 N-terminal helices. We identified a Red1 L205R mutant that disrupts the Red1-Iss10 interaction and analyzed its impact on MTREC functions in *S. pombe* cells. This mutation was sufficient to compromise the Red1-Iss10 binding in vivo, resulting in an Iss10 protein level decrease, indicating that the Red1 interaction is required for Iss10 stabilization. Accordingly, the amount of Iss10 was previously shown to be drastically reduced in *red1Δ* cells<sup>26</sup>. Noticeably, *red1-L205R* cells show a growth defect similar to *red1Δ* cells on minimal medium and at lower temperatures that is stronger than in *iss10Δ* cells, but the reason for this is currently unclear.

Red1 is crucial for nuclear exosome foci formation<sup>23</sup>. Specifically, this function has been attributed to region 196–245<sup>23</sup>, which we now show forms a HTH domain required for Iss10 binding. Proper Red1 and exosome nuclear localization is also dependent on Iss10<sup>6,23</sup>. Accordingly, we show that the *red1-L205R* mutation affects both Iss10 and Red1 localization. The interaction of Iss10 with Red1 HTH domain is required for Iss10 recruitment to Red1 nuclear foci, may promote the clustering of Red1 nuclear foci and enable recruitment of the nuclear exosome. We note that the mislocalization of Red1 in *red1-L205R* mutant cells, which causes Red1 to localize in an increased number of foci within the nucleus, differs from Red1 mislocalization in *iss10Δ* cells, in which Red1 shows even distribution in the nucleus<sup>6,26</sup>. The reason for this difference is unclear but in *red1-L205R* cells the fact that Iss10 is still expressed may influence Red1 localization either due to residual direct interaction between Red1 and Iss10 or to the existence of an indirect interaction. In favor of the second possibility, affinity purifications of Red1 have suggested that Iss10-Mmi1 submodule may be connected to Red1 by several means<sup>6–8</sup>.

Expression of the Red1-L205R mutant instead of WT Red1 leads to accumulation of meiotic mRNAs that is comparable to *iss10Δ* mutation. This phenotype is further strengthened when cells are grown in minimal medium, however, it still remains lower than in *red1Δ* cells. The Iss10-Red1 interaction is thus dispensable for most of the MTREC-based degradation of meiotic transcripts. Our Y2H assays also show that Iss10 interacts with Mmi1, known to directly bind to meiotic transcripts<sup>21</sup>, indicating that Iss10 links Red1 to Mmi1 and hence to target transcripts. The same Y2H interaction was recently reported in ref. 14. In *iss10Δ* cells, the interaction between Red1 and Mmi1 is severely compromised but not completely lost<sup>7,26</sup>. It is thus possible that while Iss10-mediated interaction of Mmi1 with Red1 is important for MTREC activity, particularly in sub-optimal conditions, Mmi1 might also be recruited into MTREC by alternative interactions as proposed in ref. 14. The fact that the impact of the Iss10-Red1 interaction varies according to the environmental growth conditions might be linked to the Iss10 regulation by TORC1<sup>27</sup>. Further analysis of the Mmi1 interaction with MTREC and Iss10 regulation by TORC1 will help explain the mechanism of MTREC-mediated degradation of meiotic transcripts in *S. pombe*.

We also report the crystal structure of the complex between *S. pombe* Ars2 and a Red1 N-terminal peptide. The structure of Ars2 reveals four structured domains that resemble their counterparts in hARS2 but their mutual orientations are very different, highlighting considerable plasticity of the Ars2 architecture in different species. While ssRNA binding has been reported for the hARS2 RRM domain<sup>32</sup>,



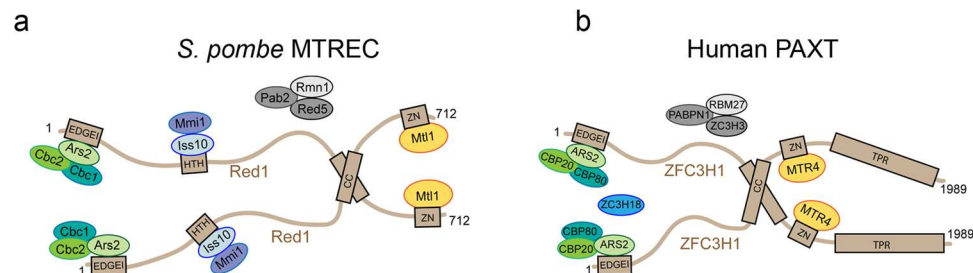
**Fig. 7 | Red1 homo-dimerization.** **a** Y2H mapping of the Red1 regions involved in its self-association. Growth was assayed on medium lacking histidine and with 30 and 50 mM 3AT. EV- empty vector. FL-Red1 (left panel) or Red1<sup>390-712</sup> (right panel) were used as a bait. **b** Ribbon representation of the Red1 dimeric coiled coil as modeled by AlphaFold<sup>25</sup>. Residues in heptad positions indicated in **c** are shown as sticks. **c** Sequence alignment of Red1 proteins comparing different *Schizosaccharomyces* species. Only the sequence of the coiled-coil region is shown. Identical residues are in brown boxes, and conserved residues are highlighted in brown. Heptad positions are indicated. **d** Details of the Red1 dimeric coiled-coil interactions centered on L481 (right panel) and I485 (left panel). **e** Molecular mass

determination of the His-MBP-Red1<sup>452-524</sup> by MALLS. The measured molecular mass of 88 kDa corresponds to a dimer. Calculated molecular mass of a monomer is 53 kDa. The sample was injected at 9 mg/ml. **f** Y2H assay showing that the L481R, I485R double mutant essentially disrupts the full-length Red1 dimerization when one or both protomers are mutated. Growth was assayed on medium lacking histidine and with 50 mM 3AT. EV- empty vector. **g** Co-immunoprecipitation experiments of Red1-Myc<sub>13</sub> and Red1-HA<sub>3</sub>, Red1-ΔCC-HA<sub>3</sub> or Red1-L481R-I485R-HA<sub>3</sub>, analyzed by western blot. Soluble extracts (SE) were prepared from exponentially growing cells in minimal liquid medium. Source data are provided as a Source Data file.

the *S. pombe* Ars2 RRM is tightly packed against the helical core, and its putative RNA binding residues are buried in the interface. hARS2 is involved in numerous interactions that are crucial for the determination of the target transcript fate<sup>30</sup>. Only the interaction of the flexible, extreme ARS2 C-terminus with CBC has so far been characterized in detail<sup>43</sup>. In the first structure of the Ars2 core with a

partner protein, we show how the Zn-finger domain of the Ars2 C-terminal leg interacts with the Red1 conserved EDGEI motif. We generated structure-based mutations in the Zn-finger domain and Red1 peptide that abolish the binding in vitro and showed that Red1 E32R mutation is sufficient to disrupt binding with Ars2 in vivo. The *Red1-E32R-TAP* mutant cells experience growth defect resembling the





**Fig. 8 | MTREC/PAXT complex model. a, b** Schematic models of the MTREC and PAXT complexes, summarizing the structural and biochemical information obtained in this study and in ref. 14. The crystal structure between Mtr4 and Red1 was determined using *C. thermophilum* proteins and the key interacting residues are conserved in *S. pombe* Red1 and Mtl1 as well as in human MTR4 and ZFC3H1<sup>14</sup>.

*red1Δ* cells, suggesting that Red1-Ars2 interaction is important for an optimal cellular growth. In addition, in mutant cells, Ars2 appears to be more diffused and less concentrated in nuclear dots. On the other hand, we did not observe any significant changes in the levels of several meiotic mRNAs and PROMPTS/CUTs known to be targets of MTREC, suggesting the interaction is not critical for the degradation of these targets by MTREC.

The Red1 EDGE1 motif is conserved in human hZFC3H1<sup>14</sup>, a subunit of the PAXT complex. We show that the structure of the C-terminal leg interacting surface is conserved between *S. pombe* and human ARS2. Correspondingly, we also show that hARS2 and hZFC3H1 interact, and the E23R mutation in hZFC3H1 (equivalent to E32R in Red1) is sufficient to disrupt the complex. This is consistent with the demonstration that the hARS2 construct corresponding to the C-terminal leg (628–876) is a major determinant in triggering RNA decay of endogenous PROMPTS transcripts and in tethering assays<sup>44</sup>. Indeed, in tethering assays, the ARS2 activity was reduced upon siRNA silencing of hZFC3H1, but not of ZC3H18 nor ZCCHC8<sup>44</sup>. Moreover, either a K719A, K722A, K734A triple mutation, or the single K719A mutation, in the C-terminal leg of ARS2 inhibited its RNA decay triggering activity<sup>44</sup>. Here, we show that residues K719 and K722 are directly involved in recognizing the Red1/ZFC3H1 EDGE1 motif. It seems highly plausible that this activity of human ARS2 thus depends on its EDGE1 motif-mediated interaction with hZFC3H1. In agreement, the same ARS2 triple mutation prevented interaction of hARS2 with PAXT and NEXT components in cell-based assays<sup>32</sup>.

Human ARS2 has previously been shown to interact via its C-terminal leg with the EEGL motif-containing FARB region of FLASH, a protein important for processing of replication-dependent histone mRNAs, and this interaction was abolished by the K719A, K722A, K734A triple mutation<sup>32</sup>. Our structure now provides an explanation for this interaction and implies a mutually exclusive binding of hARS2 to hZFC3H1/PAXT and FLASH. In addition to hZFC3H1 and FLASH, similar motifs are found in other important RNA regulators including ZC3H4, recently shown to restrict non-coding transcription<sup>45</sup>, poorly characterized ZC3H6, the NEXT/PAXT subunit ZC3H18, NCBP3 connected to mRNA export<sup>42</sup> splicing factors THRAP3 and PRPF4B<sup>46,47</sup>, RNA polymerase-associated protein RTF1, and the CPSF complex subunit CPSF2<sup>48</sup>. Interestingly, NCBP3<sup>1–182</sup> interacts with hARS2 in pull-down assays, which can be abrogated by the ARS2 C-terminal leg K719A, K722A, K734A triple mutation<sup>32</sup>. In addition, ZFC3H1, ZC3H4, NCBP3, THRAP3, PRP4B and CPSF2 all co-purify in AP-MS analysis with native hARS2 but not with the hARS2 K719A, K722A, K734A triple mutant<sup>32</sup>. While the exact nature of these interactions remains to be characterized, we speculate that the conserved binding surface on the ARS2 C-terminal leg accommodates multiple mutually exclusive interactions with competing partner proteins that exhibit the ARS2 binding motif 'EE/DGEI/L', and that this interaction is important for fate determination of targeted RNAs.

AlphaFold2<sup>35</sup> structure prediction of the human complex structure (Supplementary Fig. 15) indicates that that MTR4-binding region of ZFC3H1 might pack against a small helical domain. It is thus likely, that the MTR4-binding domain of ZFC3H1 covers residues 1140–1292. Similarly, in the AlphaFold2 prediction of Red1 (AF-Q9UTR8), this domain spans residues 575–703.

Finally, we also demonstrate the ability of Red1 to dimerize via its conserved coiled-coil region both in vitro and in *S. pombe*. Given that Iss10, Ars2, and Mtl1 can each bind to a single Red1, it is possible that the MTREC complex might exert its function as a dimeric assembly, where at least these subunits are present in pairs. This would be consistent with the reported dimerization of Mmi1<sup>49</sup>, which is a direct binder of meiotic transcripts in *S. pombe*<sup>20–22,50</sup>. Mmi1 is dimerized via its interaction with its dimeric partner Erh1 and this is important for meiotic transcript degradation by MTREC<sup>49</sup>. Functional analyses of Red1 mutants that lost the capacity to homodimerize in vitro and in Y2H assays did not reveal any significant loss of function for the *S. pombe* cells expressing these Red1 mutants. Importantly, our co-immunoprecipitation experiments showed that in *S. pombe* these Red1 mutants can still partially dimerize, indicating that in vivo Red1 can dimerize by at least one other mechanism. Better understanding of the role of Red1 dimerization within MTREC will thus require further characterization. Interestingly, AlphaFold2 modeling predicts with high confidence dimeric coiled-coils also for hZFC3H1 of the PAXT and ZCCHC8 of the NEXT complex. Dimeric architectures might thus be a conserved feature of these RNA targeting complexes.

In conclusion, our structural and in vivo analyses of the MTREC/PAXT scaffolding proteins provide important insights towards elucidation of the molecular architecture and function of these essential RNA regulators.

## Methods

### Yeast two-hybrid (Y2H) Assay

Genes coding for MTREC subunits and their variants were cloned into pDONR221 vector using BP clonase recombination (Life Technologies). The inserts were then introduced into Y2H destination vectors (pDEST32 for bait and pDEST22 for prey) using Gateway cloning system (ProQuest-Life Technologies). Bait and prey plasmids were co-transformed into MaV203 yeast strain and plated on Nitrogen-base (NB) agar medium containing 2% glucose, Histidine, Adenine and Uracil and incubated for 48 h at 30 °C. Co-transformed colonies were replica-plated on the same medium and incubated for two more days at 30 °C. Co-transformants were then assayed for *HIS3* reporter gene activation, by plating them on NB agar medium (2% glucose, Ade, Ura) supplemented with different concentrations (5–50 mM) of 3-aminotriazole (3AT; Sigma-Aldrich) histidine biosynthesis inhibitor. Plates were incubated at 30 °C and growth was monitored. Control plasmids from ProQuest™ two-hybrid system (Life Technologies). Co-transformation with empty prey vector was used as negative control. In the initial screen, Krev/RalGDS interaction was used as a positive control, and RalGDS prey vector was also used as a mock control.

### Pull-down assay

All variants of Red1 were cloned as His-MBP fusions into the pETM41 vector. Iss10 constructs were cloned as N-terminal Strep-tag fusions

into pACYCDuet (Novagen). Combinations of Red1 and Iss10 constructs were co-expressed in *E. coli* BL21Star (DE3, Invitrogen). Following cell disruption, the Red1/Iss10 proteins-containing supernatants were loaded onto a Strep-Tactin XT resin (IBA) that was then extensively washed with a buffer containing 20 mM Tris pH 8, 100 mM NaCl and 5 mM  $\beta$ -mercaptoethanol. Bound proteins were eluted by the addition of 50 mM of D-Biotin, and analyzed on 15% SDS-PAGE. For analysis of the impact of Red1 mutants on the interaction with Iss10, half of each culture was used for Strep-tag pull-down assays, as described above, and the other half for MBP pull-down assays. Supernatants were applied on Amylose resin (NEB) that was extensively washed and the bound proteins were eluted with the addition on 10 mM maltose.

Ars2-Red1 complexes were analyzed in Strep-tag pull-down assays. Ars2<sup>107–527</sup> and its variants were cloned to contain an N-terminal His-MBP- and C-terminal Strep-tag into pETM41. His-MBP fusions of Red1 proteins and Ars2 variants were expressed individually in *E. coli* BL21Star (DE3) and purified using a Ni<sup>2+</sup> chelating Sepharose (GE Healthcare). Strep-Ars2 proteins were applied on Strep-Tactin XT columns and washed extensively. Equal amounts of Ni-purified His-MBP Red1 proteins were added. Columns were extensively washed with a buffer containing 20 mM Tris pH 8 and 100 mM NaCl and bound proteins were eluted with the addition of 50 mM of D-Biotin and analyzed by 12–15% SDS-PAGE.

For analysis of the interaction between human ARS2 and ZFC3H1, hARS2<sup>47–871</sup> was cloned into pETM30 as a His-GST fusion and variants of ZFC3H1 as His-MBP fusions into pETM41. The proteins were expressed individually in *E. coli* BL21Star (DE3) and purified on a Ni<sup>2+</sup> chelating Sepharose (GE Healthcare). His-GST-ARS2 was applied onto a GST resin (GE Healthcare) and His-MBP-ZFC3H1 onto an Amylose resin (NEB). After extensive washing, His-GST-ARS2 was added onto the ZFC3H1-bound MBP column, while His-MBP-ZFC3H1 was added onto the ARS2-bound GST resin. Columns were washed with a buffer containing 20 mM Tris pH 8 and 100 mM NaCl, the proteins were eluted by the addition of 10 mM reduced glutathione or 10 mM maltose, respectively, and analyzed on 15% SDS-PAGE.

### Isothermal titration calorimetry (ITC)

ITC experiments were performed at 25 °C using an ITC200 micro-calorimeter (MicroCal). Experiments included one 0.5  $\mu$ l injection and 15–20 injections of 1.5–2  $\mu$ l of 0.4–0.8 mM His-MBP-Red1 (1–53, 20–40 or 20–40 E32R) or the Red1<sup>20–40</sup> peptide into the sample cell that contained 30–40  $\mu$ M Ars2 (FL, 58–609 or 107–527) in 20 mM Tris pH 8.0 and 100–200 mM NaCl. The initial data point was deleted from the data sets. Binding isotherms were fitted with a one-site binding model by nonlinear regression using the Origin software, version 7.0 (MicroCal).

### Multi angle laser light scattering

Size exclusion chromatography (SEC)-Light scattering (LS) experiments were conducted at 4 °C on an HPLC chromatography system consisting of a degasser DGU-20AD, a LC-20AD pump, an autosampler SIL20-ACHT, a communication interface CBM-20A and a UV-Vis detector SPD-M20A (Schimadzu, Kyoto, Japan), a column oven XL-Therm (WynSep, Sainte Foy d'Aigrefeuille, France), a static light scattering miniDawn Treos, a dynamic light scattering DynaPro Nanostar and a refractive index Optilab rEX detectors (Wyatt, Santa-Barbara, USA). The analysis was carried out with the software ASTRA, v5.4.3.20 (Wyatt, Santa-Barbara, USA). Samples of 20  $\mu$ l were injected at 0.5 mL min<sup>-1</sup> on a Superdex 200 10/300 GL (GE Healthcare), equilibrated with 20 mM Tris-HCl pH8, 100 mM NaCl, 5 mM  $\beta$ -mercaptoethanol. Bovine Serum Albumin, at 2 mg mL<sup>-1</sup>, in PBS buffer was injected as a control.

### Protein Expression and Purification

His-MBP-Red1<sup>192–236</sup> and Strep-Iss10<sup>1–45</sup> were co-expressed in *E. coli* BL21Star (DE3) cells. After affinity chromatography using the Ni<sup>2+</sup> chelating resin (GE Healthcare) in a buffer containing 20 mM Tris pH 8, 200 mM NaCl and 5 mM  $\beta$ -mercaptoethanol, the His-MBP-tag on Red1 was cleaved off by TEV protease. The complex was further purified on Strep-Tactin XT resin and Superdex 200 size exclusion chromatography.

Ars2<sup>68–183</sup> was cloned into pProEXHTb (Invitrogen) as a His-tag fusion and Ars2<sup>206–531</sup> into pRSFduet vector (Novagen) as a Strep-tag fusion. The two fragments were co-expressed in *E. coli* BL21Star (DE3, Invitrogen) and the complex purified by affinity chromatography using the chelating Ni<sup>2+</sup> Sepharose (GE Healthcare). After His-tag cleavage with the TEV protease, the two fragments were co-purified on Strep-Tactin XT resin (IBA). The final purification step was a Superdex 200 size-exclusion chromatography using a buffer containing 20 mM Tris pH 8, 100 mM NaCl and 5 mM  $\beta$ -mercaptoethanol.

### NMR Spectroscopy

NMR spectra were recorded at 298 K using a Bruker Neo spectrometer at 700 MHz or 800 MHz, equipped with a standard triple resonance gradient probe or cryoprobe, respectively. Bruker TopSpin version 4.0 (Bruker BioSpin) was used to collect data. NMR data were processed with NMR Pipe/Draw<sup>51</sup> and analyzed with Sparky 3 (T.D. Goddard and D.G. Kneller, University of California).

### Chemical shift assignment

Backbone <sup>1</sup>H<sup>N</sup>, <sup>1</sup>H <sup>$\alpha$</sup> , <sup>13</sup>C <sup>$\alpha$</sup> , <sup>13</sup>C <sup>$\beta$</sup> , <sup>13</sup>C <sup>$\gamma$</sup>  and <sup>15</sup>N<sup>H</sup> chemical shifts for the minimal Iss10-Red1 heterodimer were assigned from a sample of 500  $\mu$ M [<sup>13</sup>C,<sup>15</sup>N]Red1<sup>192–236</sup>-Iss10<sup>1–45</sup> using 2D <sup>1</sup>H,<sup>15</sup>N-HSQC, 3D HNCO, 3D HNCACO, 3D HNCA, 3D HNCACB, 3D CBCACONH, 3D HNHA and 3D HACACONH spectra. Using the same sample, aliphatic side chain protons were assigned based on 3D H(C)(CO)NH-TOCSY (40 ms mixing time), and 3D (H)C(CO)NH-TOCSY (40 ms mixing time) spectra. This was complemented by 2D <sup>1</sup>H,<sup>13</sup>C-HSQC, 3D (H)CCH-TOCSY (16 ms mixing time), and 3D (H)CCH-TOCSY (16 ms mixing time) spectra collected on the sample after exchange to 99% (v/v) D<sub>2</sub>O. Assignment of sidechain asparagine  $\delta$ 2 amides and glutamine  $\epsilon$ 2 amides used a 3D <sup>15</sup>N-HSQC-NOESY (120 ms mixing time). Stereospecific <sup>1</sup>H-<sup>13</sup>C assignment of leucine and valine methyl groups used a 2D constant-time <sup>13</sup>C-HSQC spectrum with a 150  $\mu$ M sample of 10% <sup>13</sup>C-labeled Red1<sup>192–236</sup>-Iss10<sup>1–45</sup> in 99% D<sub>2</sub>O. Aromatic <sup>1</sup>H chemical shifts were assigned using the same sample with 2D TOCSY (60 ms mixing time) and 2D DQF-COSY spectra.

### NMR structure calculation

Structure ensembles were calculated using Aria 2.3/CNS1.2<sup>52,53</sup>. The <sup>1</sup>H distances were obtained using NOE crosspeaks from a sample of 500  $\mu$ M [<sup>13</sup>C,<sup>15</sup>N]Red1<sup>192–236</sup>-Iss10<sup>1–45</sup> using a 3D <sup>15</sup>N-HSQC-NOESY (120 ms mixing time) and 3D <sup>13</sup>C-HSQC-NOESY (120 ms mixing time). Distance restraints were also obtained from the 150  $\mu$ M sample of Red1<sup>192–236</sup>-Iss10<sup>1–45</sup> in 99% D<sub>2</sub>O from a 2D <sup>1</sup>H,<sup>1</sup>H-NOESY (120 ms mixing time) spectrum. Protein dihedral angles were obtained by using TALOS-N<sup>54</sup> and Sider<sup>55,56</sup>. Hydrogen bond restraints (two per hydrogen bond) were introduced following an initial structural calculation and include only protected amides in the center of helices. Starting at iteration four in the structure calculations, residual dipolar coupling (RDC) values were included by measuring interleaved spin state-selective TROSY experiments. The aligned sample contained 17 mg/mL Pf1 phage (Asla Biotech). RDC-based intervector projection angle restraints used  $D_a$  and  $R$  values of 8 and 0.33, respectively. Final ensembles were refined in explicit water and consisted of the 20 lowest energy structures from a total of 100 calculated models. Complete refinement statistics are presented in Supplementary Table 1.

## X-ray structure determination

Pure Ars2<sup>68-183,206-531</sup> was concentrated to 8 mg ml<sup>-1</sup> in a buffer containing 20 mM Tris pH 8, 100 mM NaCl, and 5 mM  $\beta$ -mercaptoethanol using Amicon ULTRA concentrator (Millipore). The pure protein was supplemented with a three-fold molar excess of Red1 peptide (20-KNEEDESNDSDKEDGEISEDD-40) (PSL GMBH). The complex was crystallized using hanging drop vapor diffusion method at 5 °C. The best diffracting grew within three days in a solution containing 200 mM Ammonium Formate and 20% (w/v) PEG3350. For data collection at 100 K, crystals were snap-frozen in liquid nitrogen with solution containing mother liquor and 30% glycerol. Crystals of the *S. pombe* Ars2<sup>68-183,206-531</sup> - Red1<sup>20-40</sup> complex belong to the space group  $P2_1$  with the unit cell dimensions  $a = 64.670$ ,  $b = 128.281$ ,  $c = 89.3$  and  $\beta = 108^\circ$ . The asymmetric unit contains a 2:2 Ars2/Red1 hetero-tetramer with a swapped C-terminus and has a solvent content of 63%. A complete native dataset was collected to a 3.1 Å resolution, partially extending to 2.8 Å, on the ESRF beamline ID30A-3 using the MXcuBE3 software (ESRF). The data were processed using autoPROC<sup>57</sup>. Phases were obtained by molecular replacement using PHASER<sup>58</sup> with an adjusted AlphaFold2 model of *S. pombe* Ars2 (AF-094326) as a search model. The initial map was improved using the prime-and-switch density modification option of RESOLVE<sup>59</sup>. After manual model rebuilding with COOT<sup>60</sup>, the structure was refined using Refmac5<sup>61</sup> with NCS restraints to a final  $R_{\text{free}}$  of 30.2% and an  $R_{\text{work}}$  of 24.5% (Supplementary Table 2) with 99.8% residues in allowed (94% in favored) regions of the Ramachandran plot, as analyzed by MOLPROBITY<sup>62</sup>. A representative part of the  $2F_o - F_c$  and omit  $F_o - F_c$  electron density omit maps calculated using the refined model with and without the Red1 peptide, respectively, are shown in Supplementary Fig. 7b.

## Standard yeast genetics

Strains were constructed by standard genetic techniques (Moreno et al., 1991), cultured at 30 °C and under agitation at 180 rpm, in YEA (Rich medium) or Edinburgh minimal medium with appropriate supplements (Minimal medium). Genotypes of strains used in this study are listed in Supplementary Table 3.

Red1-TAP, Ars2-GFP and Iss10-GFP *S. pombe* cells were obtained using the PCR-based gene targeting method<sup>63</sup>. Positive transformants were selected by growth on YEA medium containing the appropriate antibiotic and confirmed by genomic PCR. Cells expressing point mutants of Red1 protein were generated in two steps from parental cells expressing *red1-TAP*. First, *ura4* gene was integrated into *red1* locus in order to generate *red1-Δ30-207*. Positive transformants were selected by growth on medium deprived of uracil and validated by genomic PCR. Then, *red1Δ30-207::ura4+* cells were transformed with synthesized DNA fragments of *red1* (GeneArt) containing either the E32R or L205R mutations. Positive transformants were selected by growth on YEA-5FOA and validated by genomic PCR and DNA sequencing. Oligos used for the different strain generation are listed in Supplementary Table 4.

Cells co-expressing one of the different versions of Red1-TAP (WT, E32R or L205R) together with Ars2-GFP or Iss10-GFP were obtained by crossing cells with appropriate genotypes using random spore analysis and selection on the appropriate media. *S. pombe red1-TAP iss10Δ* cells were obtained by crossing *iss10Δ* strain with *red1-TAP* strain and growth on the appropriate selective medium after random spore analysis.

To obtain plasmids expressing the Red1 mutants, the pJRXL-Pnmt81 plasmid expressing *red1-HA* (WT) was subjected for mutagenesis using Q5<sup>R</sup> Site-Direct mutagenesis Kit (NEB) using the oligo GTAGCTGCGGTTGACAACAA to generate *Pnmt81-red1Δ452-524* plasmid, and the oligo GAGATTGACATTAGAAATAACTTAAGA to generate *Pnmt81-red1-L481R-I485R* plasmid.

## Cell growth spotting assay

Spotting assays were done from 10<sup>7</sup> cells of exponentially growing cultures (OD<sub>600nm</sub> < 2) in minimal media. A 1/10th serial dilution of cells was dropped on rich or minimal media and grown at different temperatures (30, 25, and 18 °C). Pictures of the plates were acquired (Syngene bioimaging camera) after 3 days for the plate grown at 30 °C, 5 days for the plates grown at 25 °C and 10 days of grown at 18 °C.

## Live cell imaging

Red1-TAP (WT), Red1-TAP-E32R (E32R) and Red1-TAP-L205R (L205R) *S. pombe* cells co-expressing Ars2-GFP or Iss10-GFP were grown on liquid minimal medium. Live microscopy analysis was performed when cells reached an OD<sub>600</sub> of 0.5–0.8. Images were acquired with a Zeiss Apotome microscope (Carl Zeiss MicroImaging) and using a 63× oil immersion objective with a numerical aperture of 1.4 (WD 190, DICIII and GFP filters). Raw images were analyzed using the AxioVision software (Carl Zeiss MicroImaging), and processed using ImageJ.

## Immunofluorescent microscopy

The same strains used for the live cell imaging were used to assess the Red1-TAP localization by detecting its TAP tag using immunofluorescence experiments. Briefly, 20 ml of *S. pombe* cell cultures in logarithmic phase (OD<sub>600</sub> = 0.5–0.8) and grown in minimal medium were fixed with 3.8% paraformaldehyde at room temperature for 30 min, then washed with 10 ml PEM (100 mM Pipes, 1 mM EGTA, 1 mM MgSO<sub>4</sub> pH6.9). The cells were then transferred to 1.5 ml Eppendorf tubes with 1 ml of PEM and washed 2 more times. Cells were resuspended in 1 ml PEMS (PEM, 1 M D-Sorbitol) with 0.25 mg Novozym (Sigma) and 0.25 mg Zymolyase (Sigma) and incubated at RT (ideal @30 °C) until >90% cells are digested, then washed 3 times in 1 ml PEMS and spin 30 s at 500 g. Next, the cells were permeabilized 2 min with PEMT (PEM 1% Triton X-100), washed once with 1 ml of PEM and then blocked with PEMBAL (PEM, 1% BSA, 100mM L-Lysine) for 30 min. Detection of Red1-TAP was obtained by incubating the cells for 2 h at RT with the anti-TAP primary antibody (Thermo scientific #CAB1001) (1:100) in PEMBAL on a rotating wheel, the cells were washed 3 times 10 min with PEMBAL. The cells were then incubated over night at 4 °C on rotating wheel (or 2 h at RT) with a secondary antibody coupled to a fluorescent dye (DyLight<sup>®</sup> 549 (VECTOR lab # DI 1549) (1:400) dilution in PEMBAL, cells were washed 3 times 10 min with 1 ml of PEMBAL then resuspended in 100  $\mu$ l PEMBAL and mounted on poly L-lysine coated coverslips. Fluorescence microscopy imaging were done with a Zeiss Apotome microscope (Carl Zeiss MicroImaging) and using a 63× oil immersion objective with a numerical aperture of 1.4 (WD 190, DICIII and GFP filters). Raw images were acquired using the AxioVision software (Carl Zeiss MicroImaging), and processed using ImageJ.

## Protein Co-immunoprecipitation

For the Red1-TAP/Iss10-GFP and Red1-TAP-Ars2 GFP CoIPs, 20 ml of *S. pombe* cells were grown to mid-log phase in minimal medium at 30 °C, harvested, and flash-frozen in liquid nitrogen prior to protein extract preparations. Extracts prepared from cells expressing epitope-tagged proteins under the control of native gene promoters. Red1-TAP WT or mutated versions (E32R, L205R) were used for immunoprecipitations to assess the interaction between Red1-TAP (WT or E32R point mutant) with Ars2-GFP, and between Red1-TAP (WT and L205R point mutant) and Iss10-GFP. Total protein extracts were prepared from cells lysated mechanically using glass beads (2× 60 sec) in IP-Lysis Buffer (50 mM HEPES-NaOH pH = 7.5, 5 mM MgCl<sub>2</sub>, 20 mM  $\beta$ -glycerophosphate, 10% glycerol, 1 mM EDTA, 1 mM EGTA, 50 mM NaF, 0.1 mM Na<sub>3</sub>Vo<sub>4</sub>, 0.2% NP-40, 150 mM NaCl, 1  $\mu$ g LABP, 1 mM Benzamidine, 1 mM PMSF). Soluble extract cleared of cellular debris was incubated with IgG sepharose (GE Healthcare) for 2 h at 4 °C on a rotating wheel. Afterwards, beads were washed three times with the Lysis Buffer. Protein elution was performed using 2× Laemmli buffer (50 mM Tris-HCl, pH



6.8, 10%  $\beta$ -mercaptoethanol, 1% SDS, 10% glycerol and 0.1% bromophenol blue) and by heating the samples 10 min at 70 °C. The IP samples and 5% of the input were then separated by SDS-PAGE, transferred to nitrocellulose membranes and analyzed using appropriate primary (anti-Myc -Thermo scientific #CAB 1001, anti-GFP -Sigma # I1814460001) and secondary antibodies (HRP goat anti-Mouse -Dako #P0447, HRP goat anti-Rabbit -Dako #0448). Antibodies were diluted in TBS (0.1% Tween20- 1% fat dry milk; 1:1000 for the primary and 1:10,000 for the secondary), and membranes were washed in TBS- 0.1% Tween20.

For Iss10-GFP/Red1-CBP CoIPs, extracts were prepared from cells expressing the epitope tagged proteins under the control of native gene promoters at their endogenous locus and expressing Red1-TAP Red1-L205R-TAP. Whole cell extracts were prepared as described above with exception that the cell lysates were incubated at 4 °C 1 h on a rotating wheel with 20 U/ml of Arctic-TEV protease (2279 Protean) to remove the two ProteinA tag and then incubated with 20  $\mu$ l of GFP-Trap<sup>®</sup> magnetic beads (gtd-20 Chromotek) for another hour at 4 °C on a rotating wheel. For Red1-Myc/Red1-HA CoIPs, extracts prepared from cells expressing Red1-Myc or Red1-TAP epitope tagged proteins *red1* at endogenous locus, and WT, *red1 $\Delta$ 452-524* or *red1-L481R-I485R* Red1-HA proteins from a plasmid under *pnmt-81* promoter.

Red1-TAP/Red1-HA CoIPs were done as described above. Red1-Myc/Red1-HA CoIPs were done as described above except for the whole cell protein extracts that were incubated with 10  $\mu$ g of mouse anti-Myc 9E10 (Covance) at 4 °C 1 h on a rotating wheel before incubation with 10  $\mu$ l of Dynabeads<sup>™</sup> Protein-G magnetic beads (ThermoFisher) for another hour at 4 °C on a rotating wheel. The washes, elution and detection are done as described above using rabbit anti-HA (ab9110 Abcam) and anti-Myc for the detection of Red1-HA and Red1-Myc, respectively.

Enhanced chemiluminescence (ECL) detection was performed using reagents from Biorad (Clarity Western ECL kit) and revealed using ChemiDocMPS system (BioRad) or the Fusion FX (Vilber) camera. Signals were quantified with the corresponding software, the imageLab software (Biorad) or the Fusion-Capt software (Vilber Lourmat).

### Quantitative RT-PCR

Total RNA was extracted from *S. pombe* cells using hot phenol-acid. Briefly, 25 ml of logarithmic phase ( $OD_{600} = 0.5-0.8$ ) cells cultivated in minimal medium were collected by centrifugation 5 min at 1500  $\times$  g. Cells were washed with 1 ml of dH<sub>2</sub>O, then lysed by the addition 750  $\mu$ l of TES (10 mM Tris pH 7.5; 10 mM EDTA pH 8; 0.5% SDS) and 750  $\mu$ l cold acidic phenol-chloroform (Sigma P-1944), incubated at 65 °C for 15 min, and vortexed 10 s every 3 min during the heat incubation. Samples were then placed on ice for 1 min, vortexed 20 s and centrifuged for 15 min at 12,000  $\times$  g at 4 °C. A total of 700  $\mu$ l of the upper phase was recovered into a phase-lock (heavy) tubes (Eppendorf) and a second phenolic treatment was performed by the addition of 700  $\mu$ l of acidic phenol-chloroform and mixing thoroughly by inverting the tubes. Samples were then centrifuged 5 min at 12,000  $\times$  g at 4 °C. The upper phase was transferred to new phase-lock tubes, and 700  $\mu$ l of chloroform (Sigma) was added to the samples, mixed thoroughly again by inverting the tubes, and centrifuged 5 min at 12,000 g at 4 °C. The aqueous phase (500  $\mu$ l) was carefully transferred to a 2 ml Eppendorf tubes containing 1.5 ml of 100% cold EtOH (-20 °C) and 50  $\mu$ l of 3 M NaAc pH5.2, the mix was vortexed and incubated 15 min at -80 °C then centrifuged for 10 min at 12,000  $\times$  g, 4 °C. The pellet was washed in 75% (v/v) EtOH and centrifuged at 12,000  $\times$  g for 5 min, 4 °C. The pellet was then air-dried, dissolved in 100  $\mu$ l RNase-free water, incubated at 60 °C for 10 min and stored at -20 °C. RNA concentration and the A260/280 ratio were determined using a NanoDrop ND-1000 UV spectrophotometer. RNA integrity was evaluated by electrophoresis on 1 % agarose gel using the GelGreen Nucleic Acid Stain (Biotium) dye following the manufacturer's instructions.

Reverse transcription and PCR were performed on 2  $\mu$ g of purified RNA using the Transcriptor Reverse Transcriptase (Roche). RNA was first subjected to DNase (Roche) treatment in the presence of 5X DNase buffer (100 mM Tris pH 8, 10 mM MgCl<sub>2</sub>), 100 mM DTT (Invitrogen) and RNase inhibitor (Thermo Scientific) for 20 min at 37 °C. Next 1.9  $\mu$ l of 50 mM EDTA were added to the RNA-DNase mixture and incubated 10 min at 70 °C. RNA was then hybridized with 50  $\mu$ M of random hexamer primer (Invitrogen) in the presence of 50 mM MgCl<sub>2</sub> for 10 min at 65 °C. The cDNA synthesis was performed by adding 10 mM dNTP (Thermo Scientific), RNase Inhibitor (Thermo Scientific), RT buffer (Roche) and Reverse Transcriptase (Roche) and subsequent incubation for 10 min at 25 °C, 40 min at 55 °C and 5 min at 85 °C, according to the manufacturer guidelines.

qPCR was performed on a LightCycler480 machine (Roche). 4  $\mu$ l of the diluted (1:4) cDNAs were used with 0.1  $\mu$ M of each primer (forward and reverse) and 2X MESA Blue qPCR mix (Eurogentec) in a 20  $\mu$ l final qPCR reaction. DNA amplification was done using the following program: 15 min incubation at 95 °C, followed by 40 cycles of 95 °C for 15 s, 60 °C for 15 s and 72 °C for 15 s. Oligos used for the qRT-PCRs are listed in Supplementary Table 4.

### Statistics and reproducibility

The pull-down experiments in Figs. 2i, 6e and Supplementary Figs. 1e, 7d, 10b, c were performed twice. The pull-down experiment in Fig. 5d was performed three times. Gel filtration experiments in Supplementary Fig. 2h, 12b, f were done twice. The gel filtration experiments in Supplementary Fig 4b-e and Supplementary Fig. 5d were done at least three times. The co-immunoprecipitation experiments in Fig. 3a, Supplementary Fig. 3a, Fig. 5g, Fig. 7g and Supplementary Fig. 13b were performed two times. The RT-qPCR experiments in Fig. 3f, Supplementary Figs. 3d, 8b, c, and 13d were performed three times. The RT-PCRs experiments in Supplementary Fig. 8d were performed three times. The immunofluorescence experiments in Figs. 3c, d, 5i, and Supplementary Fig. 8f were performed three times. The spotting assays in Fig. 3b, Supplementary Fig. 3b, c, Fig. 5h, Supplementary Figs. 8a and 13c were performed three times.

### Reporting summary

Further information on research design is available in the Nature Research Reporting Summary linked to this article.

### Data availability

The data that support this study are available from the corresponding author upon request. The structure ensemble of the Red1<sup>192-236</sup>-Iss10<sup>1-45</sup> heterodimer generated in this study has been deposited at the Protein Data Bank (<http://www.ebi.ac.uk/pdbe/>) with accession ID 7QUU. Corresponding chemical shift assignments have been deposited in the Biological Magnetic Resonance Data Bank (<http://bmr.bwisc.edu/>) under BMRB accession number 34702. The atomic coordinates and structure factors of the *S. pombe* Ars2-Red1 complex determined in this study have been deposited under the PDB accession codes 7QY5. Source data are provided with this paper.

### References

1. Carninci, P. et al. Molecular biology: The transcriptional landscape of the mammalian genome. *Science* **309**, 1559–1563 (2005).
2. Djebali, S. et al. Landscape of transcription in human cells. *Nature* **489**, 101–108 (2012).
3. Jensen, T. H., Jacquier, A. & Libri, D. Dealing with pervasive transcription. *Mol. Cell* **52**, 473–484 (2013).
4. Zinder, J. C. & Lima, C. D. Targeting RNA for processing or destruction by the eukaryotic RNA exosome and its cofactors. *Genes Dev.* **31**, 88–100 (2017).
5. Schmid, M. & Jensen, T. H. Controlling nuclear RNA levels. *Nat. Rev. Genet.* **19**, 518–529 (2018).

6. Egan, E. D., Braun, C. R., Gygi, S. P. & Moazed, D. Post-transcriptional regulation of meiotic genes by a nuclear RNA silencing complex. *RNA* **20**, 867–881 (2014).
7. Zhou, Y. et al. The fission yeast MTREC complex targets CUTs and unspliced pre-mRNAs to the nuclear exosome. *Nat. Commun.* **6**, 7050 (2015).
8. Lee, N. N. et al. Mtr4-like protein coordinates nuclear RNA processing for heterochromatin assembly and for telomere maintenance. *Cell* **155**, 1061–1074 (2013).
9. Meola, N. et al. Identification of a nuclear exosome decay pathway for processed transcripts. *Mol. Cell* **64**, 520–533 (2016).
10. Ogami, K. et al. An Mtr4/ZFC3H1 complex facilitates turnover of unstable nuclear RNAs to prevent their cytoplasmic transport and global translational repression. *Genes Dev.* **31**, 1257–1271 (2017).
11. Silla, T. et al. The human ZC3H3 and RBM26/27 proteins are critical for PAXT-mediated nuclear RNA decay. *Nucleic Acids Res.* **48**, 2518–2530 (2020).
12. Lubas, M. et al. The human nuclear exosome targeting complex is loaded onto newly synthesized RNA to direct early ribonucleolysis. *Cell Rep.* **10**, 178–192 (2015).
13. Lubas, M. et al. Interaction profiling identifies the human nuclear exosome targeting complex. *Mol. Cell* **43**, 624–637 (2011).
14. Dobrev, N. et al. The zinc-finger protein Red1 orchestrates MTREC submodules and binds the Mtl1 helicase arch domain. *Nat. Commun.* **12**, 3456 (2021).
15. Weick, E. M. & Lima, C. D. RNA helicases are hubs that orchestrate exosome-dependent 3'–5' decay. *Curr. Opin. Struct. Biol.* **67**, 86–94 (2021).
16. Puno, M. R. & Lima, C. D. Structural basis for MTR4-ZCCHC8 interactions that stimulate the MTR4 helicase in the nuclear exosome-targeting complex. *Proc. Natl Acad. Sci. USA* **115**, E5506–E5515 (2018).
17. Schuller, J. M., Falk, S., Fromm, L., Hurt, E. & Conti, E. Structure of the nuclear exosome captured on a maturing preribosome. *Science* **360**, 219–222 (2018).
18. Weick, E. M. et al. Helicase-dependent RNA decay illuminated by a cryo-EM structure of a human nuclear RNA exosome-MTR4 complex. *Cell* **173**, 1663–1677 (2018).
19. Zofall, M. et al. RNA elimination machinery targeting meiotic mRNAs promotes facultative heterochromatin formation. *Science* **335**, 96–100 (2012).
20. Hiriart, E. et al. Mmi1 RNA surveillance machinery directs RNAi complex RITS to specific meiotic genes in fission yeast. *EMBO J.* **31**, 2296–2308 (2012).
21. Harigaya, Y. et al. Selective elimination of messenger RNA prevents an incidence of untimely meiosis. *Nature* **442**, 45–50 (2006).
22. Yamashita, A. et al. Hexanucleotide motifs mediate recruitment of the RNA elimination machinery to silent meiotic genes. *Open Biol.* **2**, 120014 (2012).
23. Shichino, Y., Otsubo, Y., Yamamoto, M. & Yamashita, A. Meiotic gene silencing complex MTREC/NURS recruits the nuclear exosome to YTH-RNA-binding protein Mmi1. *PLoS Genet.* **16**, e1008598 (2020).
24. Touat-Todeschini, L. et al. Selective termination of lncRNA transcription promotes heterochromatin silencing and cell differentiation. *EMBO J.* **36**, 2626–2641 (2017).
25. Shichino, Y., Otsubo, Y., Kimori, Y., Yamamoto, M. & Yamashita, A. YTH-RNA-binding protein prevents deleterious expression of meiotic proteins by tethering their mRNAs to nuclear foci. *Elife* **7**, e32155 (2018).
26. Yamashita, A., Takayama, T., Iwata, R. & Yamamoto, M. A novel factor Iss10 regulates Mmi1-mediated selective elimination of meiotic transcripts. *Nucleic Acids Res.* **41**, 9680–9687 (2013).
27. Wei, Y. et al. TOR targets an RNA processing network to regulate facultative heterochromatin, developmental gene expression and cell proliferation. *Nat. Cell Biol.* **23**, 243–256 (2021).
28. Izaurralde, E. et al. A nuclear cap binding protein complex involved in pre-mRNA splicing. *Cell* **78**, 657–668 (1994).
29. Mazza, C., Segref, A., Mattaj, I. W. & Cusack, S. Large-scale induced fit recognition of an m7GpppG cap analogue by the human nuclear cap-binding complex. *EMBO J.* **21**, 5548–5557 (2002).
30. Lykke-Andersen, S., Rouvière, J. O. & Jensen, T. H. ARS2/SRRT: At the nexus of RNA polymerase II transcription, transcript maturation and quality control. *Biochem. Soc. Trans.* **49**, 1325–1336 (2021).
31. Winczura, K. et al. Characterizing ZC3H18, a multi-domain protein at the interface of RNA production and destruction decisions. *Cell Rep.* **22**, 44–58 (2018).
32. Schulze, W. M., Stein, F., Rettel, M., Nanao, M. & Cusack, S. Structural analysis of human ARS2 as a platform for co-transcriptional RNA sorting. *Nat. Commun.* **9**, 1701 (2018).
33. Soltysova, M. et al. Structural insight into DNA recognition by bacterial transcriptional regulators of the SorC/DeoR family. *Acta Crystallogr. Sect. D. Struct. Biol.* **77**, 1411–1424 (2021).
34. Sugiyama, T. & Sugioka-Sugiyama, R. Red1 promotes the elimination of meiosis-specific mRNAs in vegetatively growing fission yeast. *EMBO J.* **30**, 1027–1039 (2011).
35. Jumper, J. et al. Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–689 (2021).
36. Krissinel, E. & Henrick, K. Inference of macromolecular assemblies from crystalline state. *J. Mol. Biol.* **372**, 774–797 (2007).
37. Thillainadesan, G. et al. Conserved protein Pir2ARS2 mediates gene repression through cryptic introns in lncRNAs. *Nat. Commun.* **11**, 2412 (2020).
38. Machida, S., Chen, H. Y. & Adam Yuan, Y. Molecular insights into miRNA processing by Arabidopsis thaliana SERRATE. *Nucleic Acids Res.* **39**, 7828–7836 (2011).
39. Cléry, A., Blatter, M. & Allain, F. H. T. RNA recognition motifs: Boring? Not quite. *Curr. Opin. Struct. Biol.* **18**, 290–298 (2008).
40. Kiriya, M., Kobayashi, Y., Saito, M., Ishikawa, F. & Yonehara, S. Interaction of FLASH with arsenite resistance protein 2 is involved in cell cycle progression at S phase. *Mol. Cell Biol.* **29**, 4729–4741 (2009).
41. Krystkowiak, I. & Davey, N. E. SLIMSearch: A framework for proteome-wide discovery and annotation of functional modules in intrinsically disordered regions. *Nucleic Acids Res.* **45**, W464–W469 (2017).
42. Dou, Y. et al. NCBP3 positively impacts mRNA biogenesis. *Nucleic Acids Res.* **48**, 10413–10427 (2020).
43. Schulze, W. M. & Cusack, S. Structural basis for mutually exclusive co-transcriptional nuclear cap-binding complexes with either NELF-E or ARS2. *Nat. Commun.* **8**, 1302 (2017).
44. Melko, M. et al. Mapping domains of ARS2 critical for its RNA decay capacity. *Nucleic Acids Res.* **48**, 6943–6953 (2020).
45. Estell, C., Davidson, L., Steketee, P. C., Monier, A. & West, S. Zc3h4 restricts non-coding transcription in human cells. *Elife* **10**, e67305 (2021).
46. Lee, K. M., Hsu, I. W. & Tarn, W. Y. TRAP150 activates pre-mRNA splicing and promotes nuclear mRNA degradation. *Nucleic Acids Res.* **38**, 3340–3350 (2010).
47. Schneider, M. et al. Human PRP4 kinase is required for stable tri-snRNP association during spliceosomal B complex formation. *Nat. Struct. Mol. Biol.* **17**, 216–221 (2010).
48. Kolev, N. G., Yario, T. A., Benson, E. & Steitz, J. A. Conserved motifs in both CPSF73 and CPSF100 are required to assemble the active endonuclease for histone mRNA 3'-end maturation. *EMBO Rep.* **9**, 1013–1018 (2008).

49. Xie, G. et al. A conserved dimer interface connects ERH and YTH family proteins to promote gene silencing. *Nat. Commun.* **10**, 251 (2019).
50. Wang, C. et al. A novel RNA-binding mode of the YTH domain reveals the mechanism for recognition of determinant of selective removal by Mmi1. *Nucleic Acids Res.* **44**, 969–982 (2016).
51. Delaglio, F. et al. NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR* **6**, 277–293 (1995).
52. Rieping, W. et al. ARIA2: Automated NOE assignment and data integration in NMR structure calculation. *Bioinformatics* **23**, 381–382 (2007).
53. Brunger, A. T. Version 1.2 of the crystallography and nmr system. *Nat. Protoc.* **2**, 2728–2733 (2007).
54. Shen, Y. & Bax, A. Protein structural information derived from nmr chemical shift with the neural network program talos-n. *Methods Mol. Biol.* **1260**, 17–32 (2015).
55. Hansen, D. F., Neudecker, P., Vallurupalli, P., Mulder, F. A. A. & Kay, L. E. Determination of Leu side-chain conformations in excited protein states by NMR relaxation dispersion. *J. Am. Chem. Soc.* **132**, 42–43 (2010).
56. Hansen, D. F., Neudecker, P. & Kay, L. E. Determination of isoleucine side-chain conformations in ground and excited states of proteins from chemical shifts. *J. Am. Chem. Soc.* **132**, 7589–7591 (2010).
57. Vonrhein, C. et al. Data processing and analysis with the autoPROC toolbox. *Acta Crystallogr. Sect. D. Biol. Crystallogr.* **67**, 293–302 (2011).
58. McCoy, A. J. et al. Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674 (2007).
59. Terwilliger, T. C. Maximum-likelihood density modification. *Acta Crystallogr. Sect. D. Biol. Crystallogr.* **56**, 965–972 (2000).
60. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. Sect. D. Biol. Crystallogr.* **66**, 486–501 (2010).
61. Murshudov, G. N., Vagin, A. A. & Dodson, E. J. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. Sect. D. Biol. Crystallogr.* **53**, 240–255 (1997).
62. Chen, V. B. et al. MolProbity: All-atom structure validation for macromolecular crystallography. *Acta Crystallogr. Sect. D. Biol. Crystallogr.* **66**, 12–21 (2010).
63. Bähler, J. et al. Heterologous modules for efficient and versatile PCR-based gene targeting in *Schizosaccharomyces pombe*. *Yeast* **14**, 943–951 (1998).

## Acknowledgements

J.K. was funded by the CNRS ATIP-Avenir program. A.V., S.C., and J.K. were funded by ANR MTREC (ANR-21-CE11-0021) and ANR RNAGermSilence (ANR-13-BSV2-0012) to A.V. Ariadna B. Juarez-Martinez was supported by the Labex GRAL (Grenoble Alliance for Integrated Structural Cell Biology) (ANR-10-LABX-49-01) and the People Program (Marie Curie Actions) of the European Union's Seventh Framework Program (FP7/2007–2013) under REA grant agreement PCOFUND-GA-2013-609102, through the PRESTIGE program coordinated by Campus France. Financial support from the Center National de la Recherche Scientifique (IR-RMN-THC Fr3050) is gratefully acknowledged. IBS acknowledges integration into the Interdisciplinary Research Institute of Grenoble (IRIG, CEA). This work used the platforms of the Grenoble Instruct-ERIC centre (ISBG; UAR 3518 CNRS-CEA-UGA-EMBL) within the Grenoble Partnership for Structural Biology (PSB), supported by FRISBI (ANR-10-INBS-0005-02) and GRAL, financed within the University Grenoble Alpes graduate school (Ecoles Universitaires de Recherche) CBH-EUR-GS (ANR-17-EURE-0003). We thank Caroline Mas for assistance and access to the

biophysics platform, Luca Signor for mass spectrometry analysis, Aline Le Roy for assistance and access to the Protein Analysis On Line (PAOL) platform, the staff of the ESRF-EMBL (European Synchrotron Radiation Facility-European Molecular Biology Laboratory) Joint Structural Biology Group, particularly Andrew McCarthy and Matthew Bowler, for access to and help with the ESRF beamlines, Mylène Pezet from the IAB microscopy platform (MicroCell) for help with microscopy analyses as well as Daniel Perazza for the boxplot, statistical analyses and helpful comments. We thank the EMBL high-throughput crystallization facility (HTX). We thank Axelle Grélard, Estelle Morvan, and the structural biology platform at the Institut Européen de Chimie et Biologie (UMS 3033) for access to NMR spectrometers, equipment, and technical assistance. We thank Rachel Fitzgerald for help with initial validation of Red1-Ars2 interaction and Wiebke Schulze for providing hARS2 expression plasmids. We thank Ekaterina Flin for her help with statistical analyses.

## Author contributions

A.-E.F., S.A., and M.S. performed yeast two-hybrid experiments. A.-E.F., A.J.-M., and H.L. did all the in-vitro characterization. L.T.-T., A.R., and C.K. performed all the in-vivo experiments. H.L., J.K., and C.M. determined the NMR structure of Red1-Iss10 complex. A.-E.F. and J.K. determined the crystal structure of Red1-Ars2 complex. S.C. provided ARS2 expertise. M.S., S.C., C.M., A.V., and J.K. wrote the manuscript.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41467-022-32542-3>.

**Correspondence** and requests for materials should be addressed to Cameron D. Mackereth, André Verdel or Jan Kadlec.

**Peer review information** *Nature Communications* thanks the anonymous reviewers for their contribution to the peer review of this work. Peer reviewer reports are available.

**Reprints and permission information** is available at <http://www.nature.com/reprints>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022

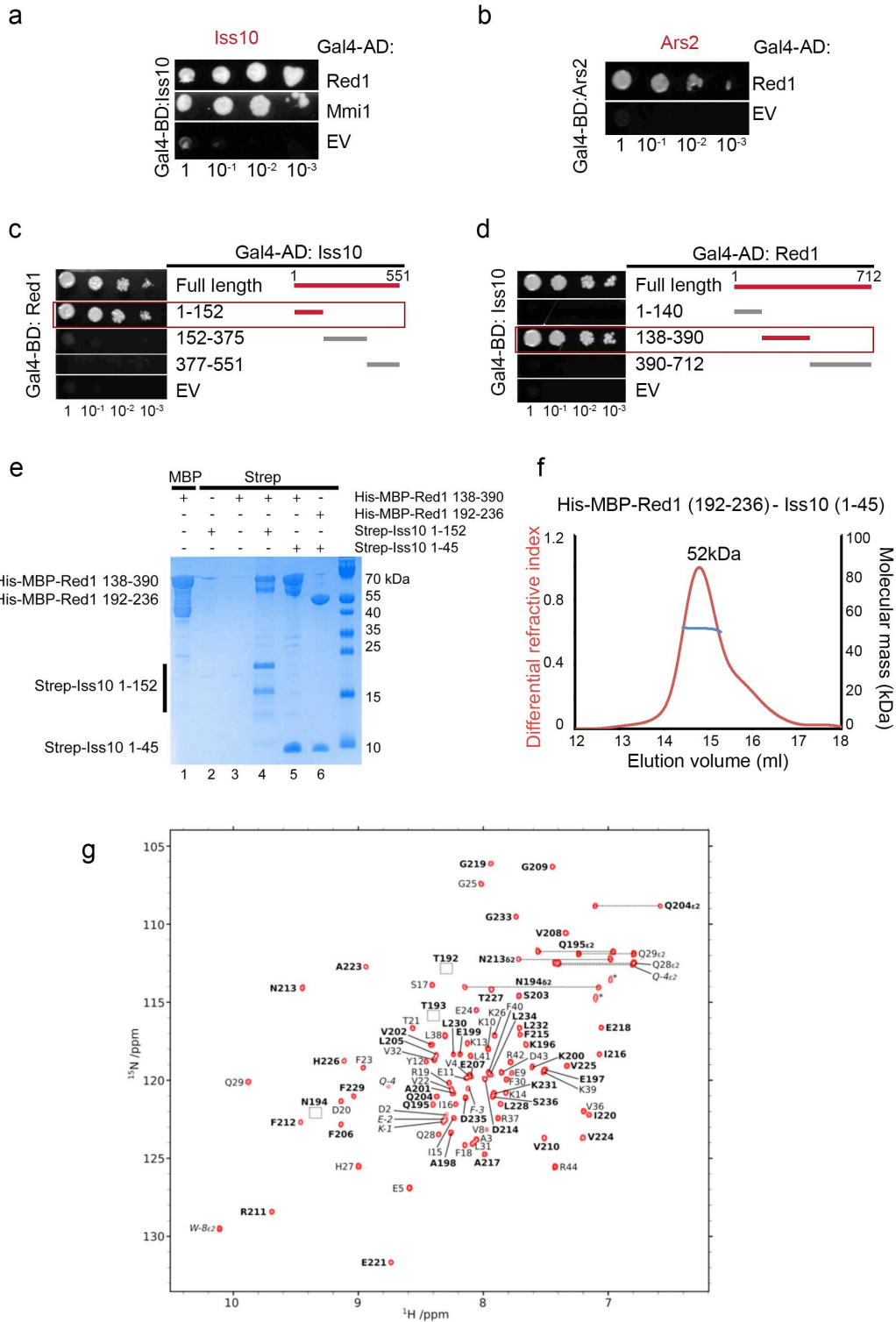
## **SUPPLEMENTARY INFORMATION**

### **Structural analysis of Red1 as a conserved scaffold of the RNA-targeting MTREC/PAXT complex**

Anne-Emmanuelle Foucher<sup>1,6</sup>, Leila Touat-Todeschini<sup>2,6</sup>, Ariadna B. Juarez-Martinez<sup>1,6</sup>, Auriane Rakitch<sup>2</sup>, Hamida Laroussi<sup>1</sup>, Claire Karczewski<sup>2</sup>, Samira Acajjaoui<sup>3</sup>, Montserrat Soler-López<sup>3</sup>, Stephen Cusack<sup>4</sup>, Cameron D. Mackereth<sup>5\*</sup>, André Verdel<sup>2\*</sup> and Jan Kadlec<sup>1\*</sup>

Contains 15 Supplementary Figures, 4 Supplementary Tables and Supplementary references.

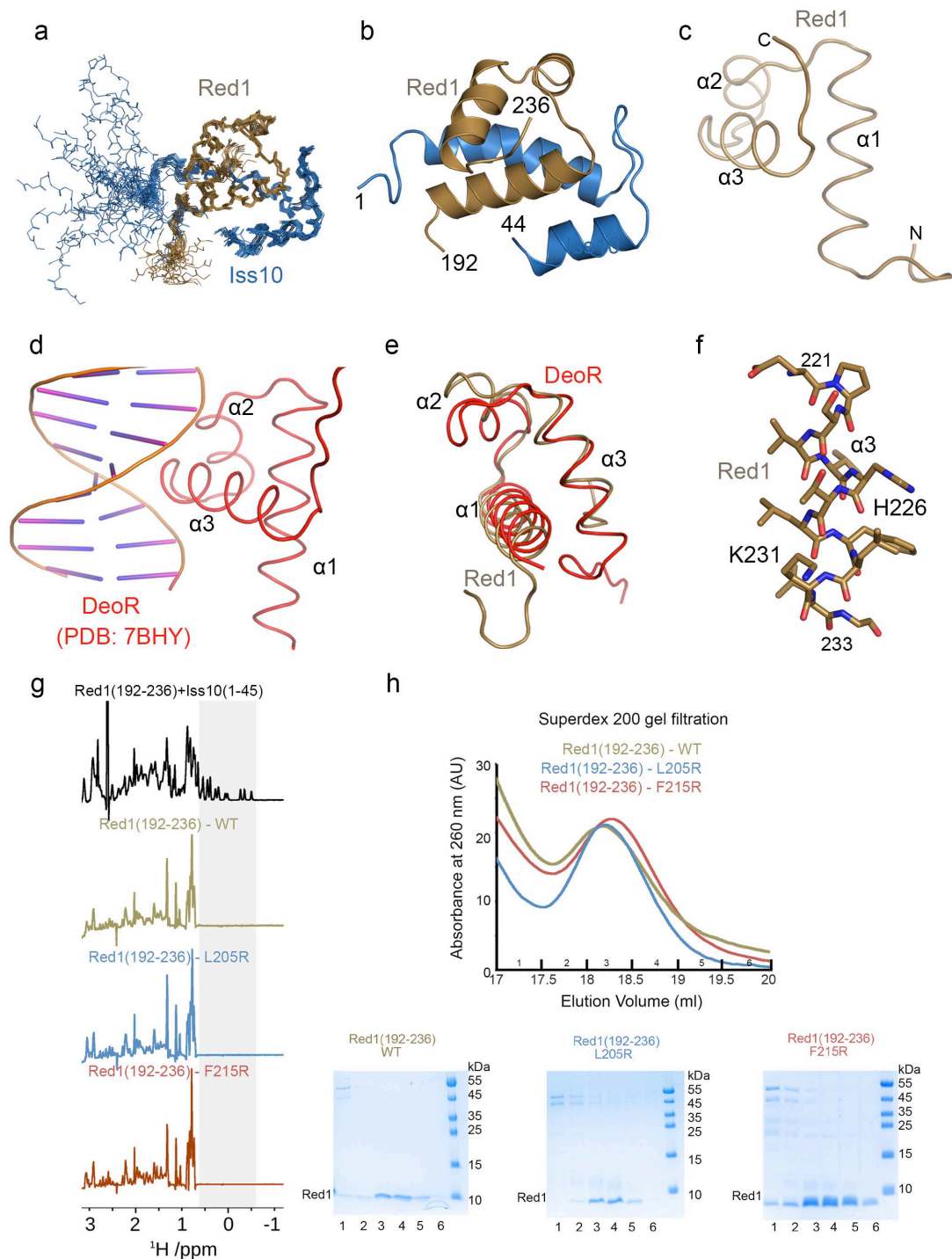




### Supplementary Figure 1

**a** Y2H assays with Red1 as a prey showing that Iiss10 interacts with Red1. Iiss10 interacts also with Mmi1. Growth was assayed on medium lacking histidine and with 20mM 3AT. **b** Y2H assays with Red1 as a prey showing that Ars2 interacts with Red1. Growth was assayed on medium lacking histidine and with 20mM 3AT. **c** Mapping of the Iiss10 interacting regions with Red1 by Y2H assays, revealing the Iiss10 N-terminus to be sufficient for the interaction. Growth was assayed on medium lacking histidine and with 50mM 3AT. EV – empty vector. **d** Mapping of the Red1 interacting regions with Iiss10 by Y2H assays, showing that the Red1 middle segment mediates the interaction. **e** SDS-PAGE analysis of Strep-tag pull-down experiments with Red1 and Iiss10. The indicated combinations

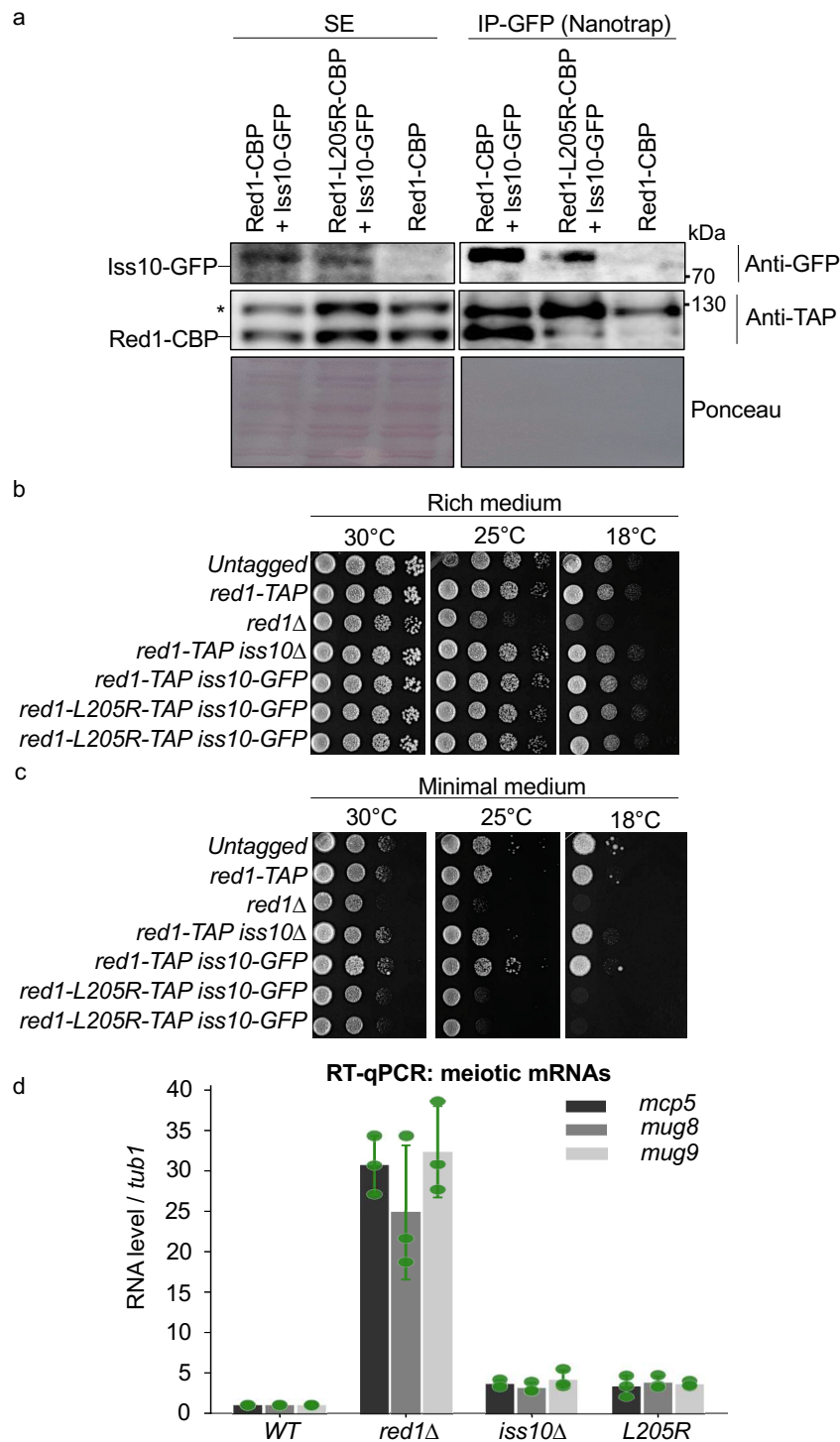
of His-MBP-Red1 and Strep-tagged Iss10 constructs were co-expressed and then purified on Strep-Tactin columns. His-MBP-Red1<sup>138-390</sup> purified alone on Amylose resin is shown as a control in lane 1. Source data are provided as a Source Data file. **f** Molecular mass determination of the His-MBP-Red1<sup>192-236</sup> - Strep-Iss10<sup>1-45</sup> complex by MALLS. The measured molecular mass of 52 kDa corresponds to a 1:1 stoichiometry. Calculated Mw of the complex is 56 kDa (49.4 + 6.7 kDa). The complex was injected at 2 mg/ml. **g** Annotated <sup>15</sup>N-HSQC spectrum of the Iss10-Red1 heterodimer. The spectrum was collected at 298 K at a field strength of 800 MHz on a sample of 150 μM [<sup>15</sup>N]Iss10-Red1 in 20 mM Tris (pH 7.5), 150 mM NaCl, 5 mM beta-mercaptoethanol and 10% (v/v) D<sub>2</sub>O. Amide <sup>1</sup>H, <sup>15</sup>N crosspeaks are annotated by residue type and number. Residues in normal type are from Iss10, and in bold type for Red1. Annotations in italics are for residues in the expression tag, and unassigned side chain amides are indicated with an asterisk. Squares indicate the location of amide crosspeaks with low intensity due to line-broadening.



### Supplementary Figure 2

**a** Superposition of 20 energy-minimized conformers representing the 3D NMR structure of the complex between Red1<sup>192-236</sup> and Iss10<sup>1-45</sup>. Red1 is shown in brown and Iss10 in blue. **b** Cartoon representation of the structured part of the complex including Red1 residues 192-236 and Iss10 residues 1-44. **c** The HTH domain of Red1. **d** The HTH domain of *B. subtilis* DeoR in complex with dsDNA (PDB code-7BHY - <http://doi.org/10.2210/pdb7BHY/pdb>)<sup>1</sup>. The structure was superimposed on Red1 shown in **c** with r.m.s.d. of 2.3 Å for 74 Ca atoms. **e** Superposition of HTH motifs of Red1 and DeoR. **f** Detail view of side chains of the Red1 helix  $\alpha$ 3. **g** Methyl region of the Red1-Iss10 complex (black, top spectrum), as well as the isolated Red1<sup>192-236</sup> wild-type, Red1<sup>192-236</sup> L205R, and Red1<sup>192-236</sup> F215R constructs. Methyl peaks with chemical shifts less than 0.7

ppm indicate the formation of a stable hydrophobic core within the Red1-Iss10 complex. The wild-type and mutant Red1 constructs all lack evidence of a stable fold, with no methyl peaks having a chemical shift less than 0.7 ppm. In addition, all three isolated Red1 constructs are soluble and show no sign of aggregation or higher order oligomerization as evident by the narrow peak linewidths. Protein concentrations are 20-40  $\mu$ M in 20mM Tris pH 8, 100mM NaCl and 5mM b-mercaptoethanol, with NMR measured at 298 K at a field strength of 700 MHz. **h** Superdex 200 gel filtration elution profiles of WT, L205R and F215R Red1<sup>192-236</sup> and SDS-PAGE analysis of fractions 1-6 for each protein. Red1<sup>192-236</sup> does not contain aromatic residues absorbing well at 280nm, but its elution could be followed by a weak absorbance at 260nm. Source data are provided as a Source Data file.



### Supplementary Figure 3

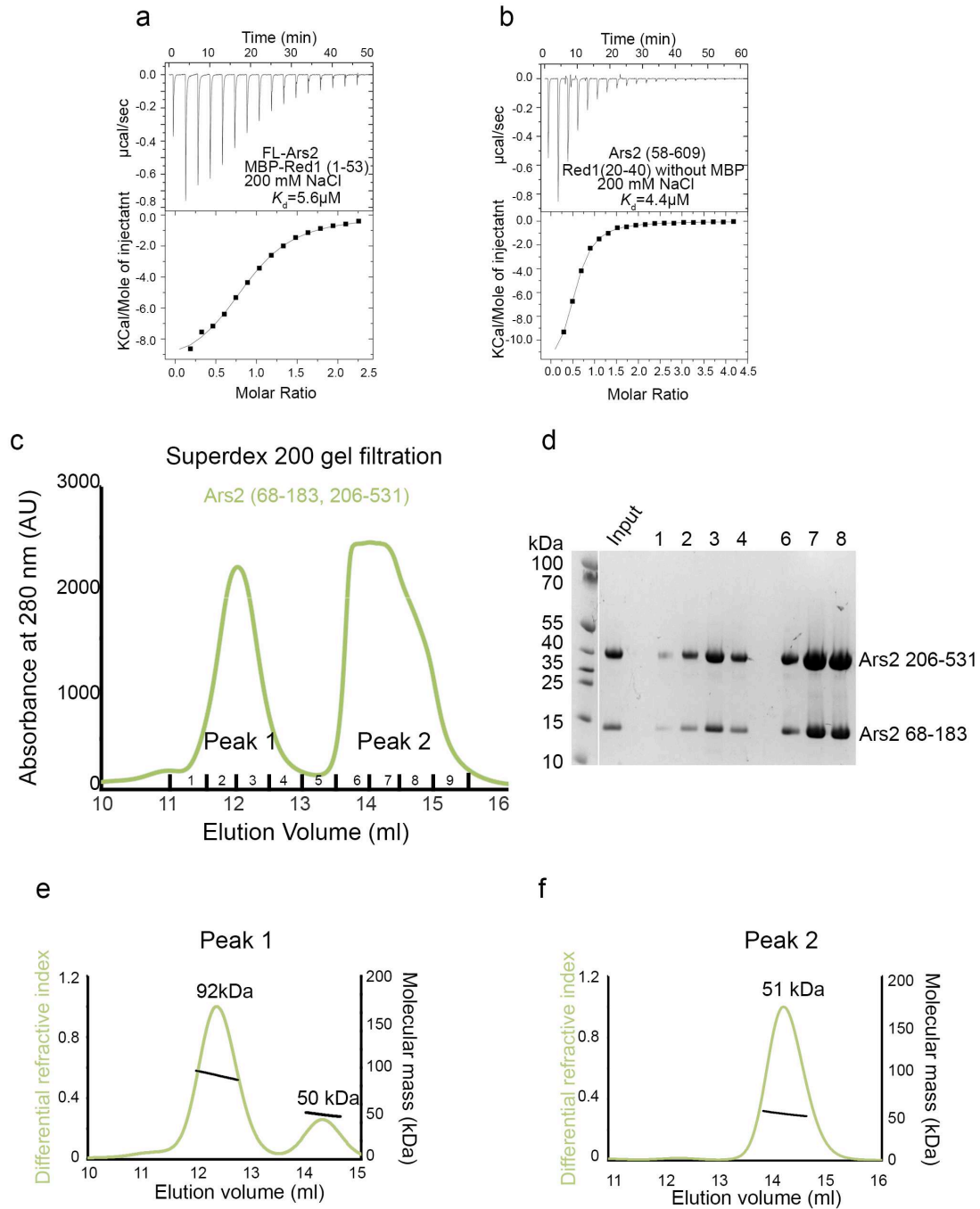
**a** Co-immunoprecipitation experiments of Red1-CBP and Iss10-GFP analysed by western blot. Soluble cell (SE) extracts were prepared from exponentially growing cells in minimal liquid medium. Asterisk indicates a background signal generated from the remaining Red1-TAP proteins after TEV cleavage (see Method for further details). **b,c** Growth assays comparing the proliferation of *red1-L205R-TAP* mutant cells to the parental cells, expressing the endogenous and untagged Red1 (*Untagged*), to cells expressing the wild-type Red1 fused to the TAP tag (*red1-TAP* or *WT*), to cells

deleted for *red1* coding sequence (*red1*Δ), and to cells deleted for *iss10* coding sequence (*iss10*Δ). Cells were grown on rich **b** or minimal **c** medium, at temperatures of 30, 25 or 18°C. **d** Quantitative RT-PCR showing the level of *mcp5*, *mug8* and *mug9* mRNAs normalized to *tub1* mRNA in *WT* and *red1-L205R* mutant cells. Total RNAs were purified from exponentially growing cells in liquid rich medium. Error bars represent standard deviation from three independent biological replicates. Source data are provided as a Source Data file.



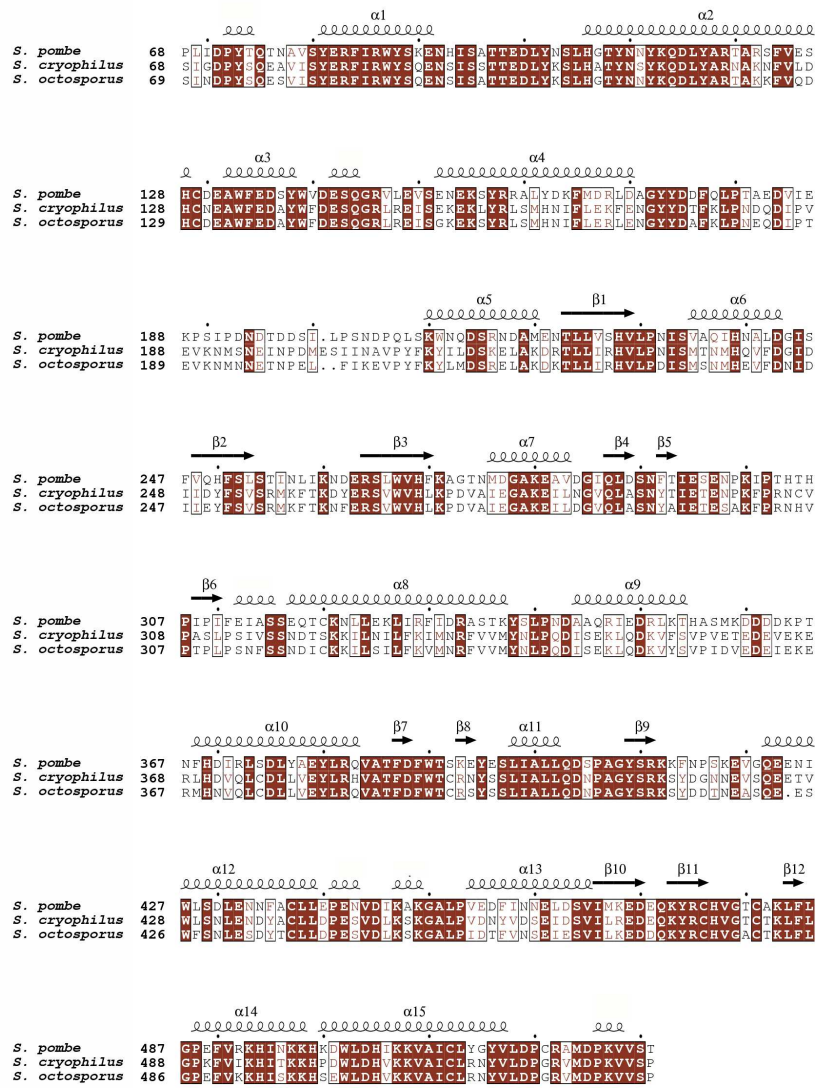


fractions as described above, of His-MBP-Red1<sup>1-53</sup> and its complex with FL His-MBP-Ars2. **e** Superdex 200 gel filtration profiles and SDS-PAGE analysis of eluted fractions as described above, of His-MBP-Red1<sup>20-40</sup> and its complex with FL His-MBP-Ars2. **f** Sequence alignment of the Red1 proteins from different *Schizosaccharomyces* species. Only the sequence of residues 1-53 covering the N-terminal fragment involved in the interaction with Ars2 is shown. The minimal binding segment identified above is highlighted with a blue rectangle. Identical residues are in brown boxes. Blue squares indicate residues whose mutation impacts the interaction as shown in Fig. 5d-f. Source data are provided as a Source Data file.



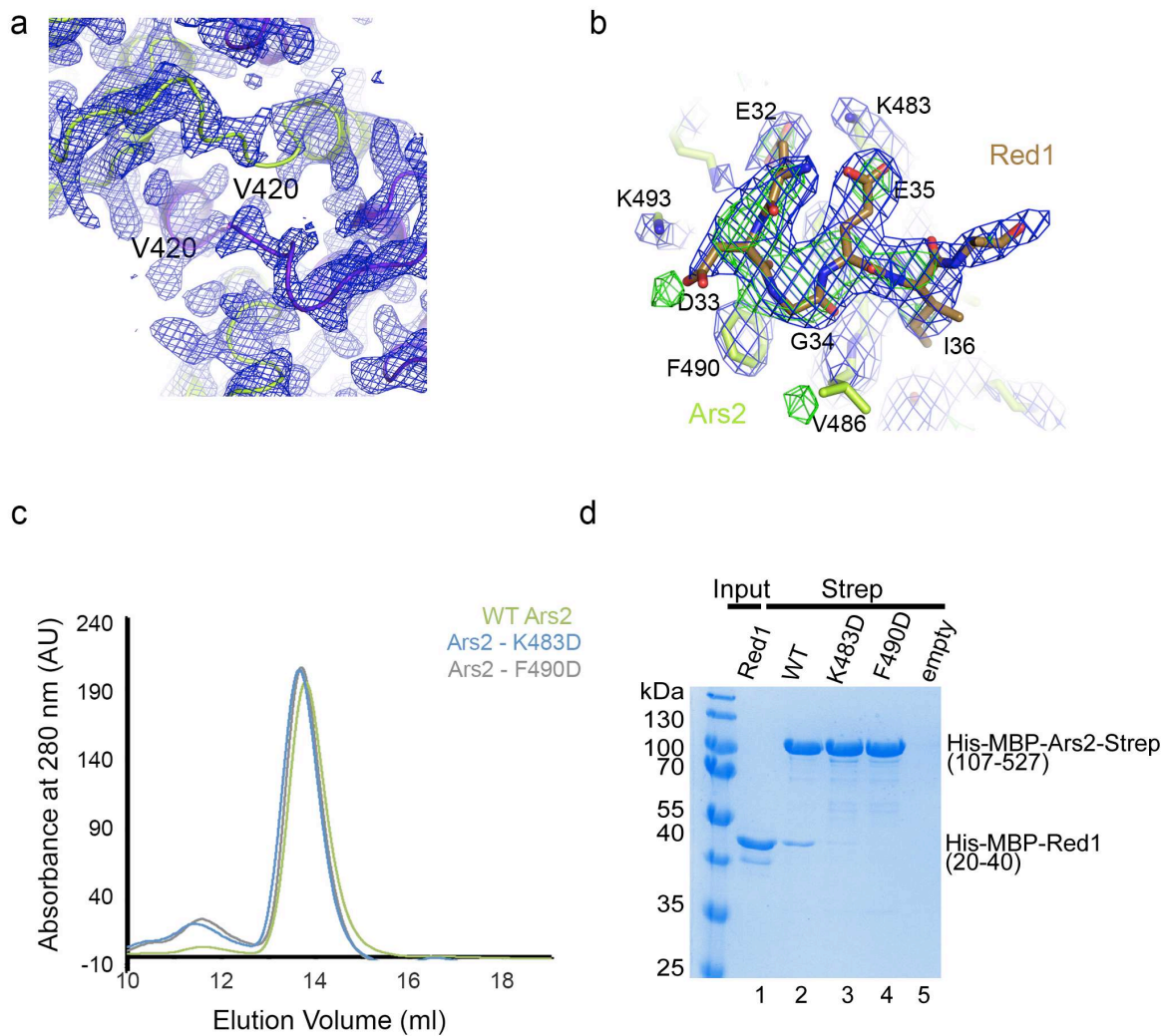
### Supplementary Figure 5

**a** ITC measurement of the interaction affinity between FL Ars2 and His-MBP-Red1<sup>1-53</sup> in presence of 200mM NaCl. **b** ITC measurement of the interaction affinity between Ars2<sup>58-609</sup> and Red1<sup>20-40</sup> peptide in presence of 200mM NaCl. **c** Superdex 200 gel filtration elution profile of Ars2<sup>68-183,206-531</sup> revealing two elution peaks. **d** SDS-PAGE analysis of fractions 1-11 of the Superdex 200 gel filtration elution profile of Ars2<sup>68-183,206-531</sup> shown in **c**. Source data are provided as a Source Data file. **e** Molecular mass determination of the *S. pombe* Ars2<sup>68-183,206-531</sup> originating from the first elution peak shown in **c** by MALLS. The measured molecular masses of 92 and 50 kDa correspond to its dimer and monomer, respectively. The protein was injected at 38 mg/ml. **f** MALLS measurement of the molecular mass of the *S. pombe* Ars2<sup>68-183,206-531</sup> eluting in the second peak shown in **c**. The measured molecular mass of 51 kDa corresponds to a monomer. The protein was injected at 4 mg/ml.



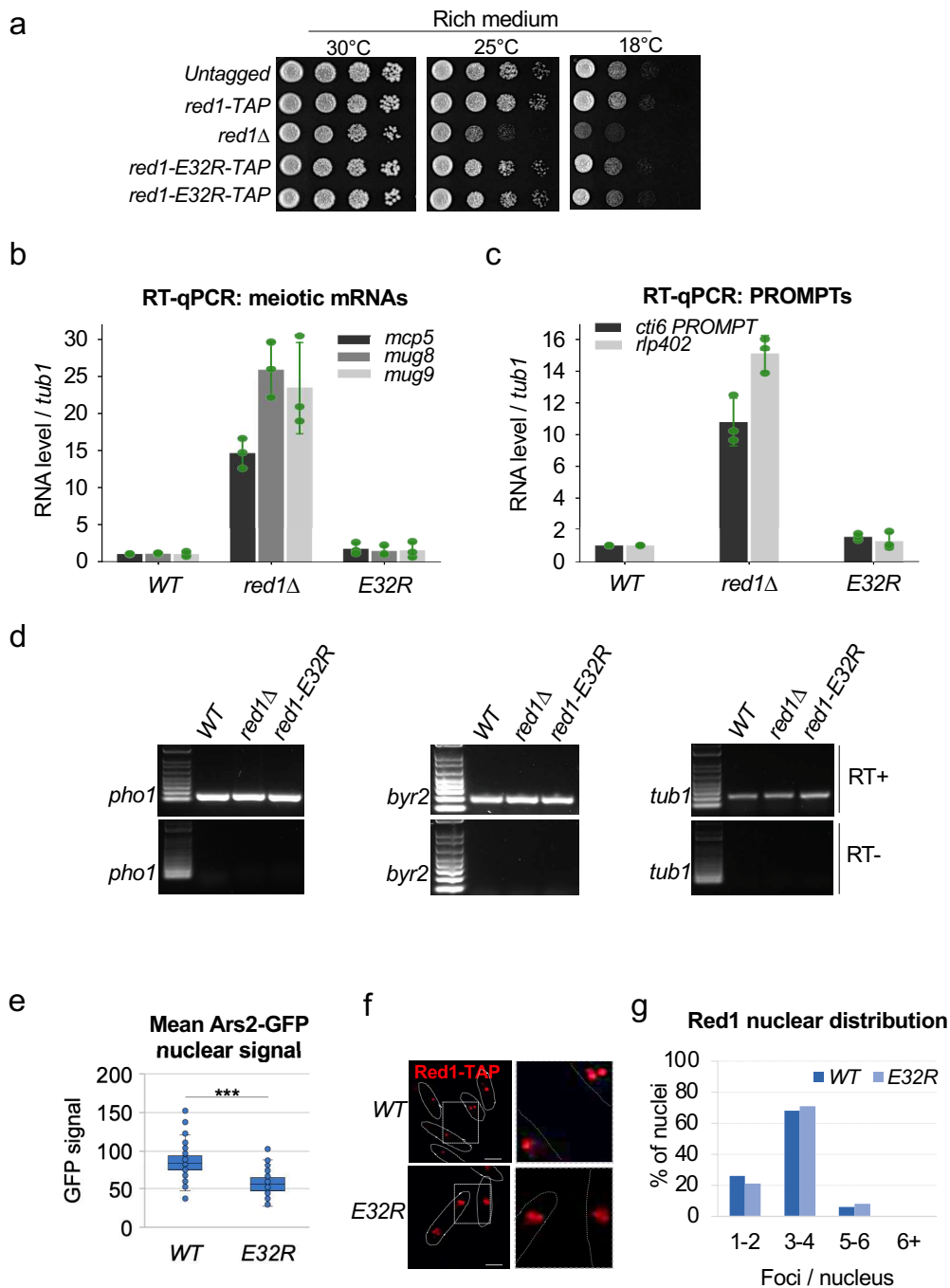
### Supplementary Figure 6

Sequence alignment of Ars2 proteins from *Schizosaccharomyces* species. Identical residues are in brown boxes.



### Supplementary Figure 7

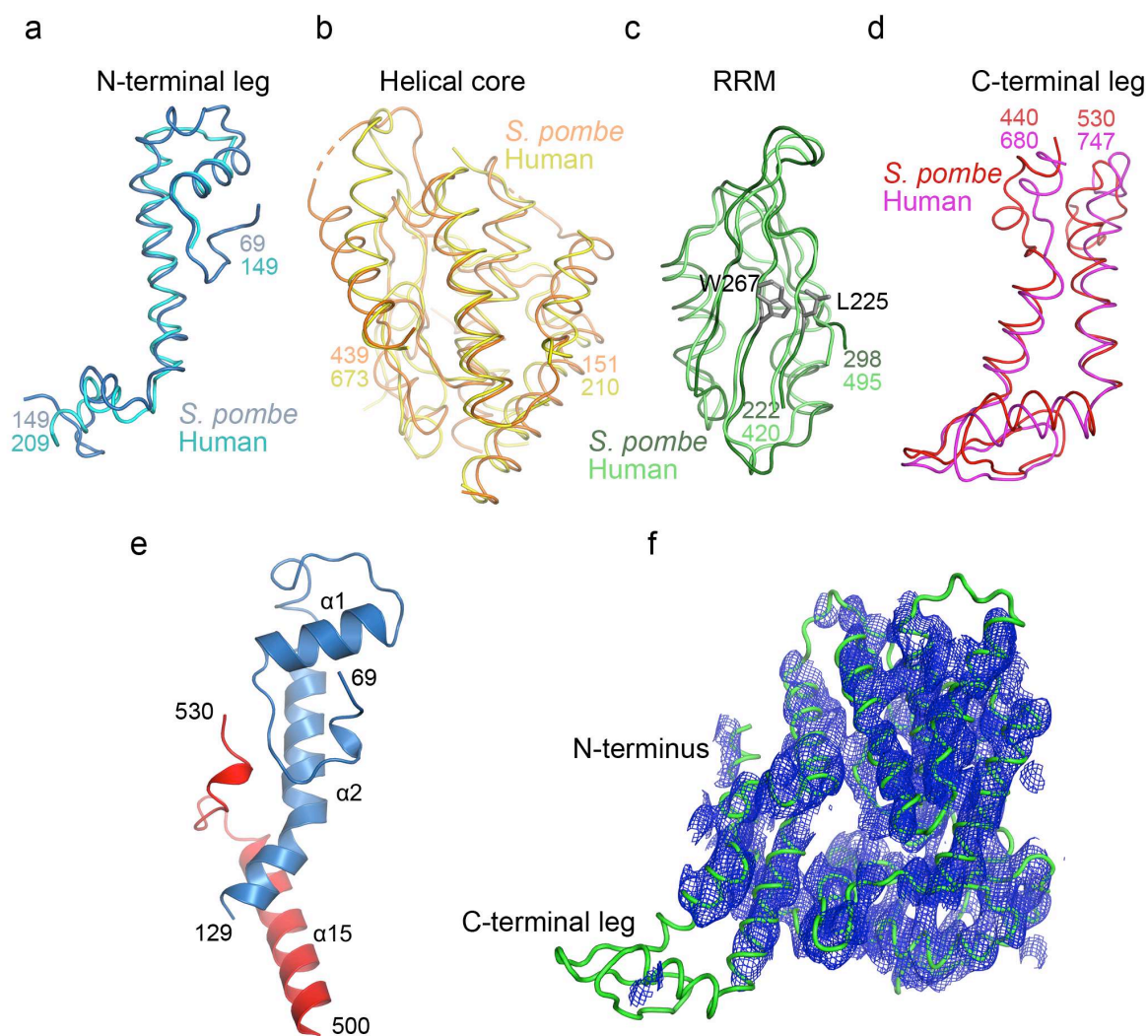
**a** A  $2F_o-F_c$  electron density map covering the two hinge regions around V420 mediating the swap of the C-terminal domain between the two Ars2 protomers contoured at  $1.0 \sigma$ . **b**  $2F_o-F_c$  (blue) and omit  $F_o-F_c$  (green) electron density maps covering the Red1 residues interacting with Ars2 contoured at  $1.0$  and  $2.5 \sigma$ , respectively. **c** Superdex 200 gel filtration elution profiles of WT and mutant ARS2 constructs used in panel **d**. **d** Strep-tag pull-down experiments with His-MBP-Red1<sup>20-40</sup> and Strep-Ars2<sup>107-527</sup> mutants indicated above the lanes. Ars2 and Red1 proteins were first purified using Ni<sup>2+</sup> resin (lane 1). Red1 with Ars2 variants were then co-purified on Strep-Tactin resin (lanes 2-4). Source data are provided as a Source Data file.



### Supplementary Figure 8

**a** Growth assays comparing the growth of *red1-E32R-TAP* mutant cells to the parental cells, which express endogenous and untagged Red1 (*Untagged*), to cells expressing wild-type Red1 fused to the TAP tag (*red1-TAP* or *WT*), to cells deleted for *red1* coding sequence (*red1Δ*). Cells were grown on rich medium, at temperatures of 30, 25 or 18°C. **b** Quantitative RT-PCR monitoring the level of meiotic DSR-containing mRNAs (*ssm4*, *spo5*, *rec8*) and **c** of PROMPTS-CUTs (*cti6* and *rlp402* PROMPTS) in *WT* and *red1-L205R* mutants. Total RNAs were purified from exponentially growing cells in liquid minimal medium. Variation of RNA levels is expressed in fold - change relative to the

*WT* levels and normalized to *tub1* mRNA. Error bars represent standard deviation from three independent biological replicates. **d** Agarose gel pictures showing RT-PCR amplification products to monitor the level of *pho1* and *byr2* mRNAs in *WT*, *red1Δ* and *red1-E32R* mutant cells. **e** Boxplots showing the mean intensity of the nuclear Ars2-GFP signal in *WT* and *red1-E32R* mutant cells. The analyses were performed on a minimum of 100 nuclei from two independent biological replicates. Wild-type box plot has a minima of 37.3, maxima of 152.2 and a centre value of 84.1. *red1-E32R* boxplot has a minima of 29.9, maxima of 102.6 and a centre value of 56.3. Each boxplot is divided in four 25% quartiles. The mean difference was examined with a student t-Test. \*\*\*= a p-value of  $3.915 \times 10^{-50}$ , determined from the two-tailed test. **f** Images of immuno-fluorescence microscopy showing Red1-TAP localization in *WT* and *red1-E32R* cells. Scale bar = 5  $\mu$ m. **g** Graph representing the foci number per nucleus of Red1-TAP in *WT* and *red1-E32R* cells. Count performed on 100 nuclei from two different biological isolates. Source data are provided as a Source Data file.



### Supplementary Figure 9

**a** Comparison of *S. pombe* and human Ars2 N-terminal leg (r.m.s.d.- 2.2 Å for 52 Ca atoms).

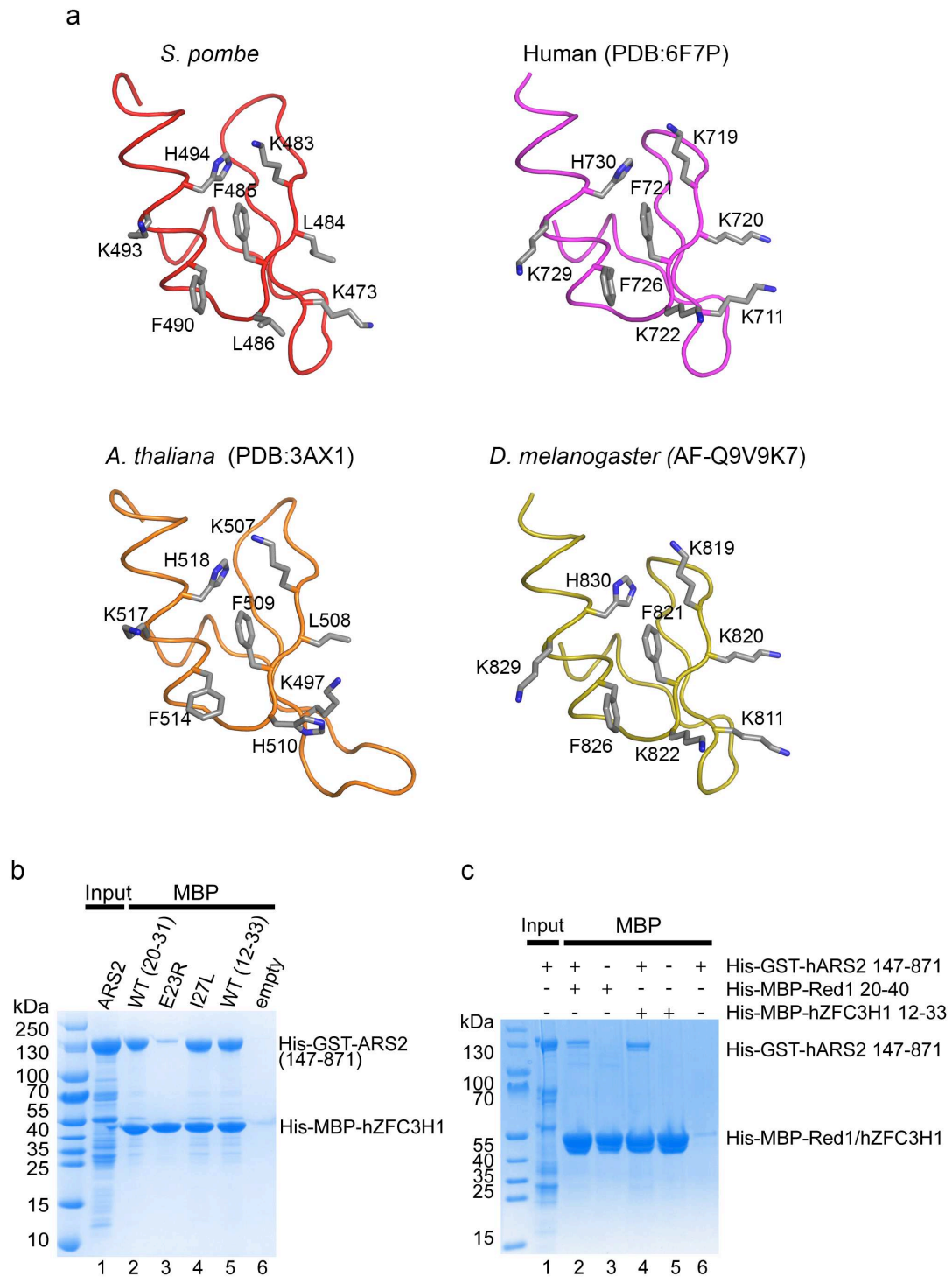
**b** Comparison of *S. pombe* and human Ars2 helical core (r.m.s.d.- 2.8 Å for 153 Ca atoms).

**c** Comparison of *S. pombe* and human Ars2 RRM domain (r.m.s.d.- 1.6 Å for 74 Ca atoms).

**d** Comparison of *S. pombe* and human Ars2 C-terminal leg (r.m.s.d.- 4.2 Å for 60 Ca atoms). **e** The N-terminal segment (residues 69-80) packs against helices  $\alpha 1$  and  $\alpha 2$  and interact with C-terminal residues 527-530.

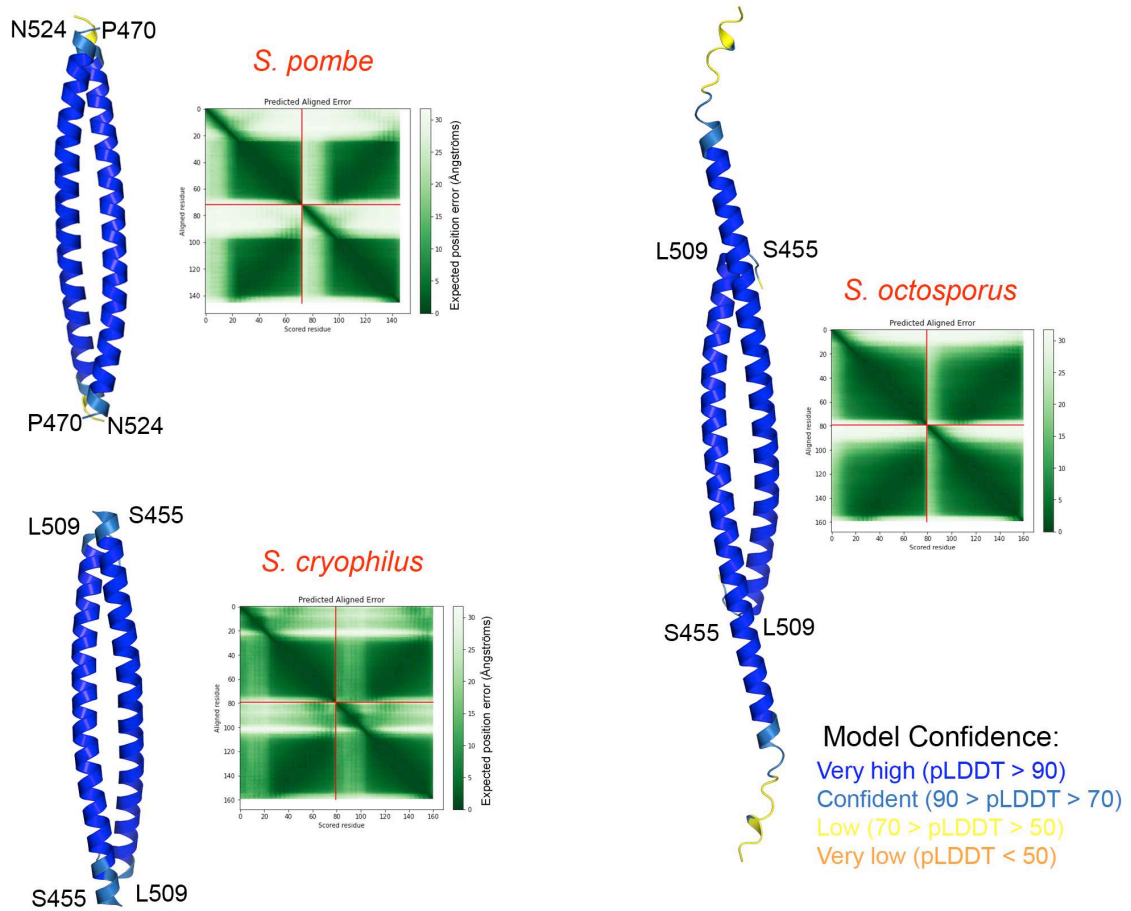
**f** A 5 Å resolution  $2F_o - F_c$  electron density map covering a monomer of Ars2<sup>107-527</sup>. The N-terminal helix and Zn-finger domain of the C-terminal leg are disordered. The structure was determined by molecular replacement using the AlphaFold *S. pombe* Ars2 model (AF-094326). The model bias was reduced using the Prime-and-switch phasing in Resolve<sup>2</sup>.





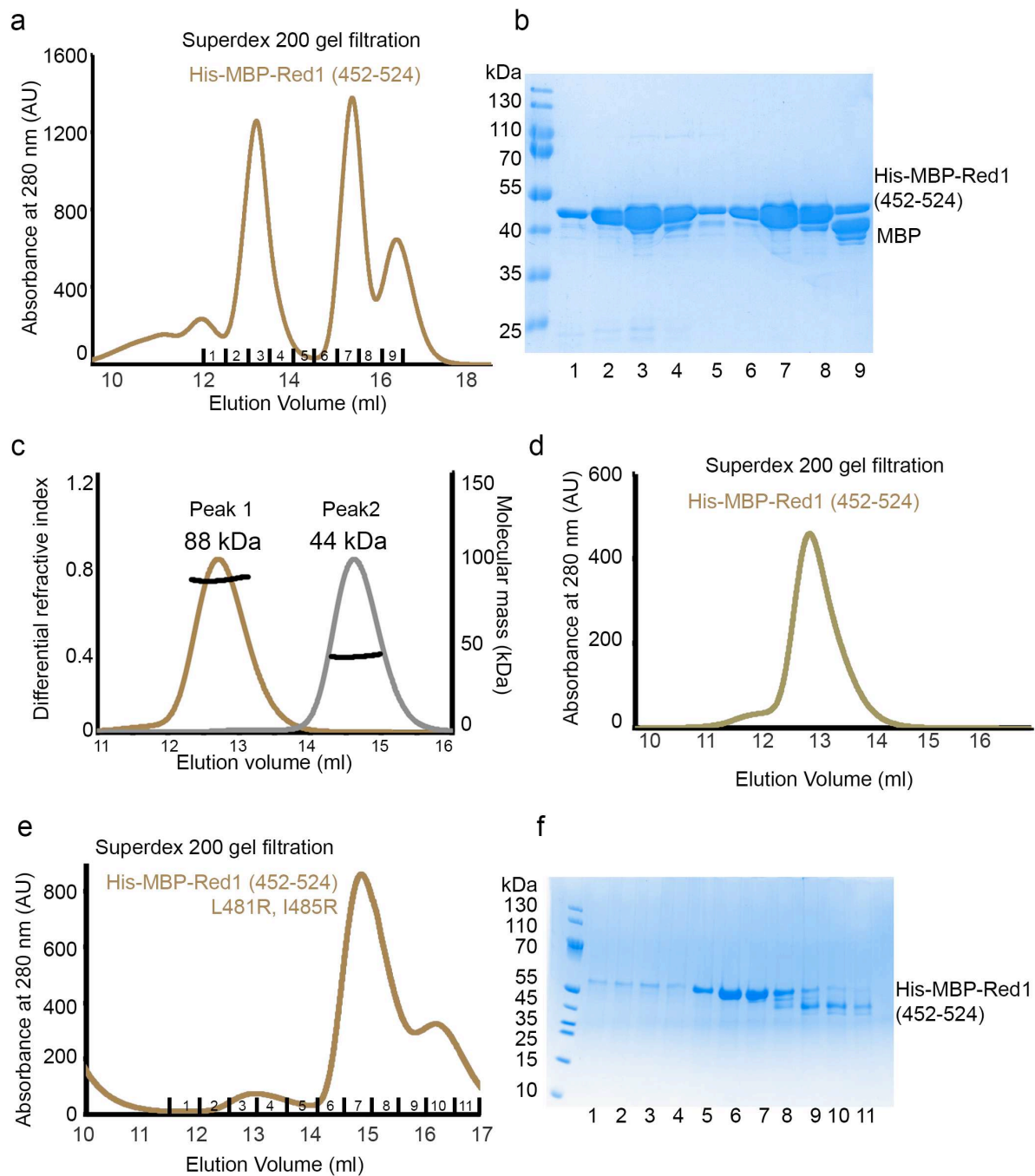
### Supplementary Figure 10

**a** Comparison of the conserved Red1-binding surface among *S. pombe* (this study), human (PDB code-6F7P - <http://doi.org/10.2210/pdb6F7P/pdb>) and *A. thaliana* (PDB code-3AX1-<http://doi.org/10.2210/pdb3AX1/pdb>) structures and the AlphaFold2 model of *D. melanogaster* ARS2 (AF-Q9V9K7). **b** MBP-pull down experiment with human His-GST-ARS2 and His-MBP-ZFC3H1. ZFC3H1<sup>20-33</sup> is sufficient for ARS2<sup>147-871</sup> binding (lane 2). E23R mutation abolished the interaction (lane 3) while I27L retains a WT level of binding (lane 4). **c** MBP-pull down experiment, showing that human His-GST-ARS2<sup>147-871</sup> can interact with both His-MBP-ZFC3H1<sup>12-33</sup> and His-MBP-Red1<sup>20-40</sup>. Source data are provided as a Source Data file.



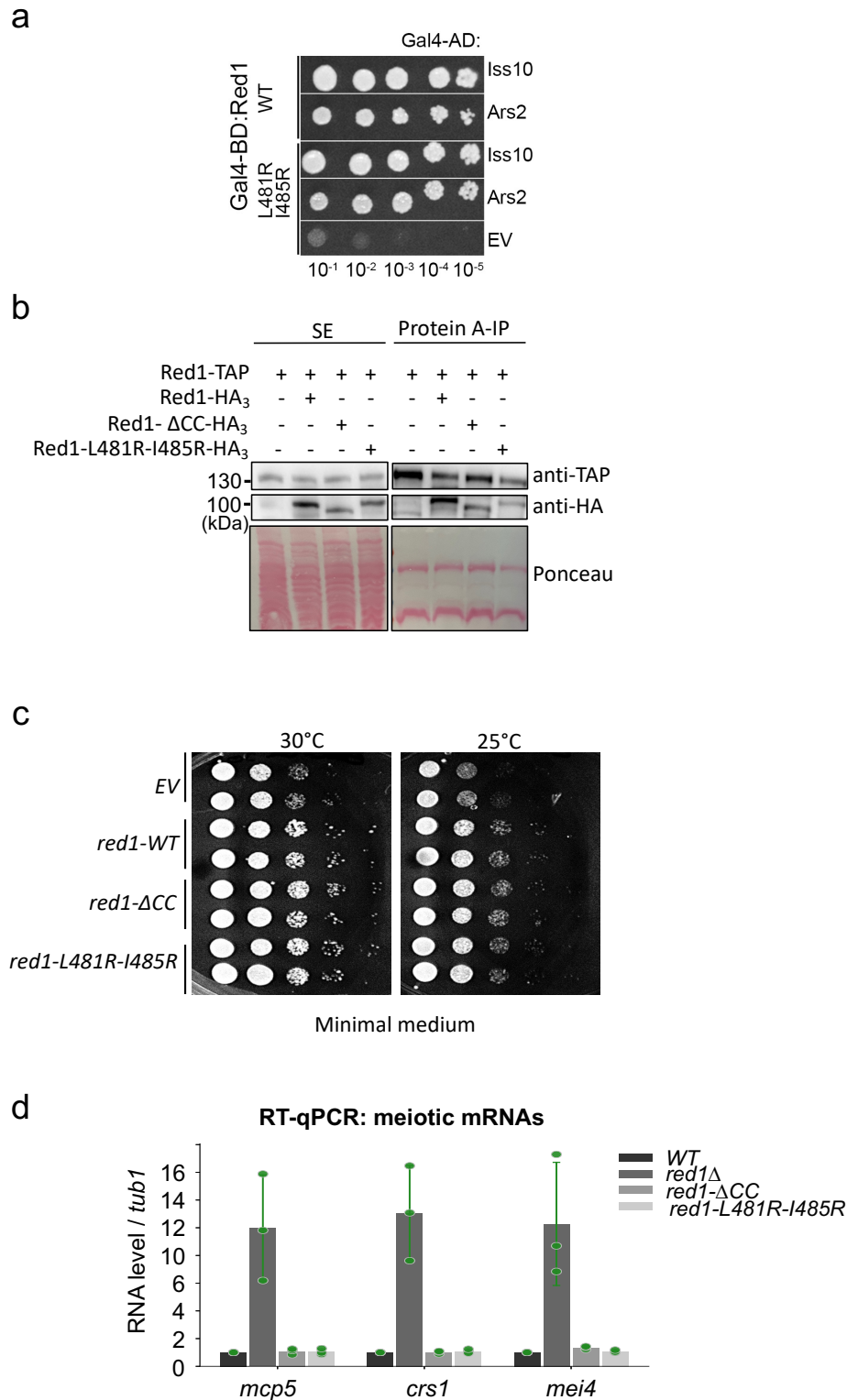
### Supplementary Figure 11

AlphaFold2 models of the *S. pombe*, *S. octosporus* and *S. cryophilus* Red1 C terminus. Only the dimeric coiled-coil regions, corresponding to the alignment in Fig. 7c are shown. For *S. octosporus*, longer chains, predicted with high confidence are shown. The structures are colored according to the AlphaFold per-residue confidence score (pLDDT). Predicted aligned error plot is shown for each model.



### Supplementary Figure 12

**a** Superdex 200 gel filtration elution profile of His-MBP-Red1<sup>452-524</sup> revealing two elution peaks (third peak corresponds to MBP). **b** SDS-PAGE analysis of fractions 1-10 of the Superdex 200 gel filtration elution profile of His-MBP-Red1<sup>452-524</sup> shown in **a**. **c** Molecular mass determination of the His-MBP-Red1<sup>452-524</sup> by MALLS. The measured molecular masses of 88 kDa and 44 kDa correspond to a dimer and monomer of Red1 (peak 1 and 2 of gel filtration shown in **a**). Calculated molecular mass of a monomer is 53 kDa. The two samples were injected at 9 and 5 mg/ml, respectively. **d** 2<sup>nd</sup> round Superdex 200 gel filtration of His-MBP-Red1<sup>452-524</sup> originating from the first elution peak shown in **a**. **e** Superdex 200 gel filtration profile of His-MBP-Red1<sup>452-524</sup> (L481R, I485R) mutant. The protein mostly elutes in a peak corresponding to the monomer. A large portion of the mutated protein also elutes in the void volume, indicating that monomeric His-MBP-Red1<sup>452-524</sup> becomes unstable. **f** SDS-PAGE-Red1<sup>452-524</sup> (L481R, I485R) shown in **e**. Source data are provided as a Source Data file.



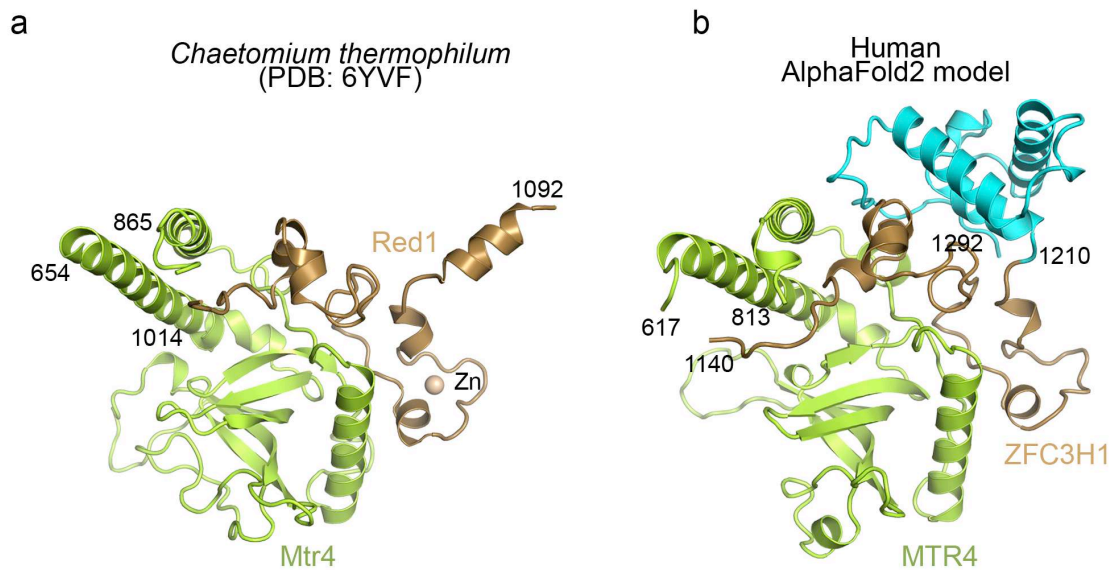
### Supplementary Figure 13

**a** Y2H assay showing that the L481R, I485R double mutant in Red1 does not impact the interactions of FL Red1 with Iss10 and Ars2. Growth was assayed on medium lacking histidine and with 30mM 3AT. **b** Co-immunoprecipitation experiments of Red1-TAP and Red1-HA<sub>3</sub>, Red1-ΔCC-HA<sub>3</sub> or Red1-L481R-I485R-HA<sub>3</sub>, analysed by western blot. Soluble cell (SE) extracts were prepared from exponentially growing cells in minimal liquid medium. **c** Growth assays comparing the growth of *red1*Δ cells expressing from a plasmid wild-type Red1, Red1-ΔCC or Red1-L481R-I485R. Cells were

grown on minimal medium, at temperatures of 30 or 25°C. **d** Quantitative RT-PCR monitoring the level of *mcp5*, *crs1* and *mei4* meiotic mRNAs in *red1* $\Delta$  cells expressing from a plasmid wild-type Red1, Red1- $\Delta$ CC or Red1-L481R-I485R mutants. Total RNAs were purified from exponentially growing cells in liquid minimal medium. Error bars represent standard deviation from three independent biological replicates. Source data are provided as a Source Data file.







**Supplementary Figure 15**

**a** Crystal structure of the *Chaetomium thermophilum* Mtr4-Red1 complex (PDB code- 6YVF-<http://doi.org/10.2210/pdb6YVF/pdb>). **b** AlphaFold2 model of human MTR4-ZFC3H1 complex. Only residues with pLDDT score above 70 are shown. The MTR4-binding region of ZFC3H1 packs against an additional helical domain shown in cyan.



**Supplementary Table 1. NMR and refinement statistics for the minimal heterodimer Iss10/Red1.**

	Iss10/Red1
<b>Number of structures in ensemble</b>	20
<b>Distance Constraints</b>	
Total	1993
Protein Intra-residue	692
Protein Inter-residue	
Sequential ( $ i-j =1$ )	321
Short Range ( $1< i-j <4$ )	142
Medium Range ( $3< i-j <6$ )	79
Long Range ( $ i-j >5$ )	74
Intermolecular	87
Ambiguous	528
Hydrogen bond	70
<b>Total dihedral angle constraints</b>	
Backbone ( $\phi, \psi$ )	154 (77,77)
Sidechain ( $\chi_1, \chi_2$ )	58 (46,12)
<b>Structure Statistics</b>	
<b>Violations (mean and SD)</b>	
Distance Constraints ( $\text{\AA}$ ) <sup>a</sup>	$0.023 \pm 0.002$
Dihedral Angle Constraints ( $^\circ$ ) <sup>b</sup>	$2.28 \pm 0.09$
RDC, Q <sup>c</sup>	$0.24 \pm 0.01$
<b>Deviations from idealized geometry</b>	
Bond Lengths ( $\text{\AA}$ )	$0.003 \pm 0.000$
Bond Angles ( $^\circ$ )	$0.45 \pm 0.01$
Impropers ( $^\circ$ )	$1.2 \pm 0.1$
<b>Ramachandran plot (%)<sup>d</sup></b>	
Most favored	92.9 (88.2)
Additionally favored	6.6 (10.7)
Generally allowed	0.4 (0.7)
Disallowed	0.1 (0.4)
<b>Average pairwise rmsd (<math>\text{\AA}</math>)</b>	
Protein backbone 2 $^\circ$ structure	$0.5 \pm 0.2$
Protein heavy 2 $^\circ$ structure	$1.0 \pm 0.2$
Protein backbone all	$1.2 \pm 0.3$
Protein heavy all	$1.3 \pm 0.2$

<sup>a</sup> No distance violation greater than 0.5  $\text{\AA}$ .

<sup>a</sup> No dihedral restraint violation greater than 10 $^\circ$ .

<sup>c</sup> Calculated for each structure in the ensemble using the method described by <sup>3</sup>.

<sup>d</sup> Determined by using PROCHECK-NMR<sup>4</sup>. Values do not include the expression tag. Values that include the tag are in parentheses.

**Supplementary Table 2. Crystallographic data collection and refinement statistics.**

	<b>Red1-Ars2</b>
<b>Data collection</b>	
Space group	$P2_1$
Cell dimensions	
$a, b, c$ (Å)	64.7, 128.3, 89.3
$\alpha \beta \gamma$ (°)	90, 108, 90
Resolution (Å)	85-2.8 (3-2.8) <sup>a</sup>
$R_{\text{merge}}$ (%)	29.6 (148)
$I / \sigma I$	5.3 (1.5)
$CC_{1/2}$ (%)	99 (55)
Completeness (%)	76 (18)
Redundancy	5.6 (6.2)
<b>Refinement</b>	
Resolution (Å)	84.8-2.8
No. reflections	25473
$R_{\text{work}}/R_{\text{free}}$	0.245/0.302
$B$ -factors	53
R.m.s. deviations	
Bond lengths (Å)	0.002
Bond angles (°)	1.17

<sup>a</sup> Values in parentheses are for highest resolution shell.

### Supplementary Table 3. List of *S. pombe* strains used in this study.

Name	Genotype	Source
SPV8	<i>h+ leu1-32 ori1 ade6 M216 ura4-D18 imr1R::ura4+</i>	Lab Stock
SPV2180	<i>h+ leu1-32 ori1 ade6 M216 ura4-D18 imr1R::ura4+ red1::TAP KanR</i>	Lab Stock
SPV3106	<i>h+ leu1-32 ori1 ade6 M216 ura4-D18 imr1R::ura4+ red1Δ::hphR</i>	Lab Stock
SPV5699	<i>h+ leu1-32 ori1 ade6 M216 ura4-D18 red1::TAP KanR</i>	This study
SPV5714	<i>h+ leu1-32 ori1 ade6 M216 ura4-D18 red1Δ30-207::TAP KanR</i>	This study
SPV5785	<i>h+ leu1-32 ori1 ade6 M216 ura4-D18 red1E32R::TAP KanR Clone 29</i>	This study
SPV5786	<i>h+ leu1-32 ori1 ade6 M216 ura4-D18 red1E32R::TAP KanR Clone 30</i>	This study
SPV5787	<i>h+ leu1-32 ori1 ade6 M216 ura4-D18 red1L205R::TAP KanR Clone 4</i>	This study
SPV5788	<i>h+ leu1-32 ori1 ade6 M216 ura4-D18 red1L205R::TAP KanR Clone 14</i>	This study
SPV5994	<i>h- leu1-32 ori1 ade6 M216 ura4-D18 imr1R::ura4+ iss10Δ::hphR</i>	This study
SPV5995	<i>h+ leu1-32 ori1 ade6 M216 ura4-D18 imr1R::ura4+ iss10Δ::hphR</i>	This study
SPV6013	<i>h- leu1-32 ori1 ade6 M216 ura4-D18 imr1R::ura4+ red1L205R::TAP KanR iss10Δ::hphR Clone 3</i>	This study
SPV6015	<i>h+ leu1-32 ori1 ade6 M216 ura4-D18 imr1R::ura4+ red1L205R::TAP KanR iss10Δ::hphR Clone 7</i>	This study
SPV6017	<i>h+ leu1-32 ori1 ade6 M216 ura4-D18 red1::TAP KanR iss10::GFP NatR Clone 5</i>	This study
SPV6018	<i>h+ leu1-32 ori1 ade6 M216 ura4-D18 red1::TAP KanR iss10::GFP NatR Clone 6</i>	This study
SPV6020	<i>h+ leu1-32 ori1 ade6 M216 ura4-D18 red1L205R::TAP KanR iss10::GFP NatR Clone 5</i>	This study
SPV6021	<i>h+ leu1-32 ori1 ade6 M216 ura4-D18 red1L205R::TAP KanR iss10::GFP NatR Clone 6</i>	This study
SPV6023	<i>h- leu1-32 ori1 ade6 M216 ura4-D18 imr1R::ura4+ ars2-GFP Clone 1</i>	This study
SPV6024	<i>h- leu1-32 ori1 ade6 M216 ura4-D18 imr1R::ura4+ ars2-GFP Clone 2</i>	This study
SPV6026	<i>h+ leu1-32 ori1 ade6 M216 ura4-D18 red1E32R::TAP KanR ars2-GFP Clone 1</i>	This study
SPV6027	<i>h+ leu1-32 ori1 ade6 M216 ura4-D18 red1E32R::TAP KanR ars2-GFP Clone 2</i>	This study
SPV6029	<i>h+ leu1-32 ori1 ade6 M216 ura4-D18 imr1R::ura4+ red1::TAP KanR ars2-GFP Clone 1</i>	This study
SPV6030	<i>h+ leu1-32 ori1 ade6 M216 ura4-D18 imr1R::ura4+ red1::TAP KanR ars2-GFP Clone 2</i>	This study
SPV6065	<i>h+ leu1-32 ori1 ade6 M216 ura4-D18 imr1R::ura4+ red1Δ::hphR + plasmide pRJ 81 XL Clone 1</i>	This study
SPV6066	<i>h+ leu1-32 ori1 ade6 M216 ura4-D18 imr1R::ura4+ red1Δ::hphR + plasmide pRJ 81 XL Clone 2</i>	This study
SPV6067	<i>h+ leu1-32 ori1 ade6 M216 ura4-D18 imr1R::ura4+ red1Δ::hphR + plasmide pRJ 81 XL Clone 3</i>	This study
SPV6068	<i>h+ leu1-32 ori1 ade6 M216 ura4-D18 imr1R::ura4+ red1Δ::hphR + plasmide pRJ 81 XL red1-3HA Clone 1</i>	This study
SPV6069	<i>h+ leu1-32 ori1 ade6 M216 ura4-D18 imr1R::ura4+ red1Δ::hphR + plasmide pRJ 81 XL red1-3HA Clone 2</i>	This study
SPV6070	<i>h+ leu1-32 ori1 ade6 M216 ura4-D18 imr1R::ura4+ red1Δ::hphR + plasmide pRJ 81 XL red1-3HA Clone 3</i>	This study
SPV6071	<i>h+ leu1-32 ori1 ade6 M216 ura4-D18 imr1R::ura4+ red1Δ::hphR + plasmide pRJ 81 XL red1-452-524Δ-3HA Clone 1</i>	This study
SPV6072	<i>h+ leu1-32 ori1 ade6 M216 ura4-D18 imr1R::ura4+ red1Δ::hphR + plasmide pRJ 81 XL red1-452-524Δ-3HA Clone 2</i>	This study
SPV6073	<i>h+ leu1-32 ori1 ade6 M216 ura4-D18 imr1R::ura4+ red1Δ::hphR + plasmide pRJ 81 XL red1-452-524Δ-3HA Clone 3</i>	This study
SPV6074	<i>h+ leu1-32 ori1 ade6 M216 ura4-D18 imr1R::ura4+ red1Δ::hphR + plasmide pRJ 81 XL red1-L481R-I485R-3HA Clone 1</i>	This study
SPV6075	<i>h+ leu1-32 ori1 ade6 M216 ura4-D18 imr1R::ura4+ red1Δ::hphR + plasmide pRJ 81 XL red1-L481R-I485R-3HA Clone 2</i>	This study
SPV6076	<i>h+ leu1-32 ori1 ade6 M216 ura4-D18 imr1R::ura4+ red1Δ::hphR + plasmide pRJ 81 XL red1-L481R-I485R-3HA Clone 3</i>	This study
SPV6095	<i>h+ leu1-32 ori1 ade6 M216 ura4-D18 imr1R::ura4+ red1::TAP-KanR + plasmide pRJ 81 XL Clone 1</i>	This study
SPV6096	<i>h+ leu1-32 ori1 ade6 M216 ura4-D18 imr1R::ura4+ red1::TAP-KanR + plasmide pRJ 81 XL Clone 2</i>	This study
SPV6098	<i>h+ leu1-32 ori1 ade6 M216 ura4-D18 imr1R::ura4+ red1::TAP-KanR + plasmide pRJ 81 XL red1-3HA Clone 4</i>	This study
SPV6099	<i>h+ leu1-32 ori1 ade6 M216 ura4-D18 imr1R::ura4+ red1::TAP-KanR + plasmide pRJ 81 XL red1-3HA Clone 5</i>	This study
SPV6101	<i>h+ leu1-32 ori1 ade6 M216 ura4-D18 imr1R::ura4+ red1::TAP-KanR + plasmide pRJ 81 XL red1-452-524Δ-3HA Clone 7</i>	This study
SPV6102	<i>h+ leu1-32 ori1 ade6 M216 ura4-D18 imr1R::ura4+ red1::TAP-KanR + plasmide pRJ 81 XL red1-452-524Δ-3HA Clone 8</i>	This study
SPV6104	<i>h+ leu1-32 ori1 ade6 M216 ura4-D18 imr1R::ura4+ red1::TAP-KanR + plasmide pRJ 81 XL red1-L481R-I485R-3HA Clone 10</i>	This study
SPV6105	<i>h+ leu1-32 ori1 ade6 M216 ura4-D18 imr1R::ura4+ red1::TAP-KanR + plasmide pRJ 81 XL red1-L481R-I485R-3HA Clone 11</i>	This study
SPV6107	<i>h<sup>+</sup> leu1-32 ura4-D18 ade6-210 otr1R(Sph1)::ade6<sup>+</sup> 3xFlag-clr4 red1::13Myc-hphR + plasmide pJR 81 XL Clone 1</i>	This study
SPV6108	<i>h<sup>+</sup> leu1-32 ura4-D18 ade6-210 otr1R(Sph1)::ade6<sup>+</sup> 3xFlag-clr4 red1::13Myc-hphR + plasmide pJR 81 XL Clone 2</i>	This study
SPV6110	<i>h<sup>+</sup> leu1-32 ura4-D18 ade6-210 otr1R(Sph1)::ade6<sup>+</sup> 3xFlag-clr4 red1::13Myc-hphR + plasmide pJR 81 XL red1-3HA Clone 4</i>	This study
SPV6111	<i>h<sup>+</sup> leu1-32 ura4-D18 ade6-210 otr1R(Sph1)::ade6<sup>+</sup> 3xFlag-clr4 red1::13Myc-hphR + plasmide pJR 81 XL red1-3HA Clone 5</i>	This study
SPV6113	<i>h<sup>+</sup> leu1-32 ura4-D18 ade6-210 otr1R(Sph1)::ade6<sup>+</sup> 3xFlag-clr4 red1::13Myc-hphR + plasmide pJR 81 XL red1-452-524Δ-3HA Clone 7</i>	This study
SPV6114	<i>h<sup>+</sup> leu1-32 ura4-D18 ade6-210 otr1R(Sph1)::ade6<sup>+</sup> 3xFlag-clr4 red1::13Myc-hphR + plasmide pJR 81 XL red1-452-524Δ-3HA Clone 8</i>	This study
SPV6116	<i>h<sup>+</sup> leu1-32 ura4-D18 ade6-210 otr1R(Sph1)::ade6<sup>+</sup> 3xFlag-clr4 red1::13Myc-hphR + plasmide pJR 81 XL red1-L481R-I485R-3HA Clone 10</i>	This study
SPV6117	<i>h<sup>+</sup> leu1-32 ura4-D18 ade6-210 otr1R(Sph1)::ade6<sup>+</sup> 3xFlag-clr4 red1::13Myc-hphR + plasmide pJR 81 XL red1-L481R-I485R-3HA Clone 11</i>	This study

**Supplementary Table 4. List of primers used for the *S. pombe* analysis.**

gene	Sequence 5' > 3'
<b>strain construction</b>	
<i>Ars2-GFP</i> tag for	TTCTATAGAACTTATCAAGATTTGGATGCGCCTAACCAGGAGTTCCAGAATTAGATTA Ccggatccccgggtaattaa
<i>Ars2-GFP</i> tag rev	TAACGTCTATACAGGAAATGATATTAACCAAAAATCTTTCTGACTAAGAACATAAAATTT Agaattcgagctcgtaattaa
<i>Iss10-GFP</i> tag for	AACGACTATGAATTAGCAGAGGAAATGAGAGGCGTCGTCATCATAGCTTTTATGATT ATcggatccccgggtaattaa
<i>Iss10-GFP</i> tag rev	TCAAATATTCCGCCAAGCAATTTTTAAATTTGTTGCGCGCACTTACGGAAGAGTATAGC TAgaattcgagctcgtaattaa
<i>red1Δ30-207</i> for	GGTCAATAAATCTCGATGAGTTGCGAAAAAGGCTTTAGAATCAAAAAAGAAGATGA AGAAGATGAGAGTAATGACTCTCGCCAGGGTTTTCCAGTCACGAC
<i>red1Δ30-207</i> rev	AGTCCGAGCTTTAAAAATAAGTATGAACAACCGAAGGCTCGATGCCCTCAGCAATA AAATCGTTAAAGCGCACTCCTACAGCGGATAACAATTTACACAGGA
<i>red1-3HA</i> tag for	acgtcccggactatgcaggatcctatcctatgacgtccagattacgctTAGGGATCCCCGGGTAAA
<i>red1-3HA</i> tag rev	catacggatagcccgcatagtcaggaaacatcgatgggaaccggcggaATGGAACCTTTGCAACAGG
<i>red1Δ452-524</i> for	GTTGCATGTTTACCTAAAG
<i>red1Δ452-524</i> rev	TAATATCTGCTCAGTTGTATC
<i>red1 L481R-I485R</i> for	cttaagaGAAAAAGAGGAAGGAGAAC
<i>red1 L481R-I485R</i> for	tatttctAATGTCAATCTCGGATTTTTTAATC
<b>genotyping</b>	
<i>red1Δ30-207</i> for	TTGCTACATTCAGTATACCTGTTCA
<i>red1Δ30-207</i> rev	TTCGCTAATTTCCGCATCTT
<i>Ars2-GFP</i> tag for	GGATCCCAAAGTGGTAAGCA
<i>Ars2-GFP</i> tag rev	TTCTATCTGTGCAGCGTTTCG
<i>Iss10-GFP</i> tag for	GCCTCCTATTGAAACGCACT
<i>Iss10-GFP</i> tag rev	TGCGAAACCAATTTATGCAA
<i>red1-3HA</i> tag for	TTGAAATATCGGTGCTTACCT
<i>red1-3HA</i> tag rev	CAGTACTGGCAAGGGAGACATTCC
<i>red1Δ452-524</i> for	gtagctgagggtgacaacaa
<i>red1 L481R-I485R</i> for	GAG ATT GAC ATT AGA AAT AAC TTA AGA
<i>red1 crib</i> rev	atctgttccgtgtccat
<b>qPCR</b>	
<i>tub1</i> for	GTA CTG G C C C A T A C C G T G A T
<i>tub1</i> rev	C G A A T G G A A G A C G A G A A A G C
<i>ssm4</i> for	T G C A A G A G G A A A C T C A A A G G
<i>ssm4</i> rev	T T C C T C C C A C T T G T T T T G A
<i>spo5</i> for	T T A T G G C G T G C T C G T G A A T A
<i>spo5</i> rev	T T C G A C T C C A A T G G C A C A T A
<i>mug8</i> for	G G C T G A C C A A A T G A A T G C T T
<i>mug8</i> rev	G A T C T G C C C A A C C A A C A A T C
<i>mug9</i> for	C T C T G T T G G C T C C T T C A A G C
<i>mug9</i> rev	T A A T G T T T G A G C A C G C A T G G
<i>mcp5</i> for	T T T G G A T T A A C C C G C A T T G T
<i>mcp5</i> rev	T G G C G T T C T T T A A G G C A T C T
<i>rec8</i> for	G T T G A A G T T G G A C G G G A T G T
<i>rec8</i> rev	T T C T A C C C T A C T C G G C A T C G
<i>crs1</i> for	A G G C C G A C G G A A A T T T A T G
<i>crs1</i> rev	A C C A T T T G T G G G C A A A G A A C
<i>mei4</i> for	A A A A G C G A C C T T C A A G C A A A
<i>mei4</i> rev	T T G C A T C G T T T G A G A C T T C G
<i>cti6 PROMPT</i> for	C A T C G G A G A A C C T G G T T T G T
<i>cti6 PROMPT</i> rev	T G C A T T C C C A A A T C A A G T C
<i>rlp402 PROMPT</i> for	G A A A G A G G A A C A A T G T T T A G A A G T C
<i>rlp402 PROMPT</i> rev	A A G C C C T A C G A T T G C T G T G
<i>pho1</i> for	G A C T T C C C G T A G C C C C T A T C
<i>pho1</i> rev	T C C A A C A G C A T C G A A A G C A G
<i>pho1ncRNA</i> for	T G A G A T T T A C G G G A A G T G C C
<i>pho1ncRNA</i> rev	A G C G A C T A T T C C C C T G T T G T
<i>byr2</i> for	T C A T C T T C A C C G G A A C G C G
<i>byr2</i> rev	A G C G G A C A C T T C G T T T C T G
<i>nam1-L</i> for	C C A G T A G T A T C T G A A C G G T G
<i>nam1-L</i> rev	C A T G G A C G T G G A A A C T C T C G
<i>nam1</i> for	T C G G C A A T T C A A C C A T A A C A
<i>nam1</i> rev	T G G C T T C A T T C C T T C C T T T G
<b>RT</b>	
<i>RT pho1</i>	A A T A A C T C G A C A C G G C C A C T
<i>RT pho1ncRNA</i>	C A A A C G A A C C A A C A A C G C C
<i>RT byr2</i>	A G C G G A C A C T T C G T T T C T G
<i>RT nam1</i>	C A A T C A C G A C G G A T C G T A C A

## Supplementary Reference

1. Soltysova, M. *et al.* Structural insight into DNA recognition by bacterial transcriptional regulators of the SorC/DeoR family. *Acta Crystallogr. Sect. D Struct. Biol.* **77**, 1411-1424 (2021).
2. Terwilliger, T. C. Maximum-likelihood density modification. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **56**, 965-972 (2000).
3. Cornilescu, G., Marquardt, J. L., Ottiger, M. & Bax, A. Validation of protein structure from anisotropic carbonyl chemical shifts in a dilute liquid crystalline phase. *J. Am. Chem. Soc.* **120**, 6836-6837 (1998).
4. Laskowski, R. A., Rullmann, J. A. C., MacArthur, M. W., Kaptein, R. & Thornton, J. M. AQUA and PROCHECK-NMR: Programs for checking the quality of protein structures solved by NMR. *J. Biomol. NMR* **8**, 477-486 (1996).