



HAL
open science

Evolution and Diversity of the TopoVI and TopoVI-like Subunits With Extensive Divergence of the TOPOVIBL subunit

Julia Brinkmeier, Susana M. Coelho, Henri-Marc Bourbon, Bernard de Massy

► **To cite this version:**

Julia Brinkmeier, Susana M. Coelho, Henri-Marc Bourbon, Bernard de Massy. Evolution and Diversity of the TopoVI and TopoVI-like Subunits With Extensive Divergence of the TOPOVIBL subunit. *Molecular Biology and Evolution*, 2022, 10.1093/molbev/msac227 . hal-04302026

HAL Id: hal-04302026

<https://hal.science/hal-04302026>

Submitted on 23 Nov 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.

Evolution and Diversity of the TopoVI and TopoVI-like Subunits With Extensive Divergence of the TOPOVIBL subunit

Julia Brinkmeier,¹ Susana Coelho ,² Bernard de Massy ,^{*} and Henri-Marc Bourbon³

¹Institut de Génétique Humaine (IGH), Centre National de la Recherche Scientifique, Univ Montpellier, Montpellier 34396, France

²Max Planck Institute for Developmental Biology, Tübingen 72076, Germany

³Centre de Biologie Intégrative, CNRS, Université de Toulouse, Toulouse 31400, France

***Corresponding author:** E-mail: bernard.de-massy@igh.cnrs.fr.

Associate editor: Harmit Malik

Abstract

Type II DNA topoisomerases regulate topology by double-stranded DNA cleavage and ligation. The TopoVI family of DNA topoisomerase, first identified and biochemically characterized in Archaea, represents, with TopoVIII and mini-A, the type IIB family. TopoVI has several intriguing features in terms of function and evolution. TopoVI has been identified in some eukaryotes, and a global view is lacking to understand its evolutionary pattern. In addition, in eukaryotes, the two TopoVI subunits (TopoVIA and TopoVIB) have been duplicated and have evolved to give rise to Spo11 and TopoVIBL, forming TopoVI-like (TopoVIL), a complex essential for generating DNA breaks that initiate homologous recombination during meiosis. TopoVIL is essential for sexual reproduction. How the TopoVI subunits have evolved to ensure this meiotic function is unclear. Here, we investigated the phylogenetic conservation of TopoVI and TopoVIL. We demonstrate that BIN4 and RHL1, potentially interacting with TopoVIB, have co-evolved with TopoVI. Based on model structures, this observation supports the hypothesis for a role of TopoVI in decatenation of replicated chromatids and predicts that in eukaryotes the TopoVI catalytic complex includes BIN4 and RHL1. For TopoVIL, the phylogenetic analysis of Spo11, which is highly conserved among Eukarya, highlighted a eukaryal-specific N-terminal domain that may be important for its regulation. Conversely, TopoVIBL was poorly conserved, giving rise to ATP hydrolysis-mutated or -truncated protein variants, or was undetected in some species. This remarkable plasticity of TopoVIBL provides important information for the activity and function of TopoVIL during meiosis.

Key words: sexual reproduction, meiosis, recombination, topoisomerase, SPO11, TOPOVIBL, TopoVI, BIN4, RHL1, MEI4, REC114.

Introduction

DNA topoisomerases play an important role in resolving topological constraints and were first discovered in bacteria almost 50 years ago (Wang 1971; Gellert et al. 1976). Ever since, DNA topoisomerases have been identified in many other organisms, such as archaea, eukaryotes, and viruses (Forterre and Gabelle 2009; Chen et al. 2013; Gabelle et al. 2014). They are ubiquitously expressed and play an essential role in various processes, such as replication, transcription, recombination, repair, and chromatin remodeling (Wang 2002; Vos et al. 2011; Chen et al. 2013). DNA topoisomerases have been extensively studied, and new family members have been progressively identified, adding insights into their structures and catalytic mechanisms.

Based on these features, DNA topoisomerases have been classified into two major types: type I and type II. Type I topoisomerases can transiently induce a break in

one DNA strand, followed by passage of a single strand and religation. Type II topoisomerases transiently break both DNA strands in a DNA double helix, followed by double-strand passage and break religation (Champoux 2001; Wang 2002; Chen et al. 2013). Type II topoisomerases can relax negatively and positively supercoiled DNA, and have decatenation activity. Type II topoisomerases are further divided into type IIA and type IIB topoisomerases, and they both work in an ATP- and magnesium-dependent manner to resolve topological constraints during transcription, replication, and recombination (Chen et al. 2013).

Topoisomerase VI (TopoVI) is a type IIB topoisomerase that was first identified in archaea (Bergerat et al. 1997) and forms a heterotetramer composed of two A and B subunits (A₂B₂) (Corbett et al. 2007; Graille et al. 2008). The A subunit (TopoVIA) carries the DNA cleavage activity and has two main domains: the 5Y-CAP (catabolite activator protein) DNA binding domain (also known as WHD, for winged-

© The Author(s) 2022. Published by Oxford University Press on behalf of Society for Molecular Biology and Evolution.

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial License (<https://creativecommons.org/licenses/by-nc/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact journals.permissions@oup.com

Open Access

helix domain), and the Toprim (Topoisomerase-primase) domain (fig. 1A). The B subunit (TopoVIB) contains the GHKL domain with the ATP binding/hydrolysis site, the helix-two-turn-helix (H2TH) domain, and the transducer domain that includes a long helical domain interacting with the N-terminal region of TopoVIA. The GHKL domain (ATPase) also known as the Bergerat fold (BF) contains three conserved motifs important for ATP binding that have been identified also in other proteins (i.e., DNA gyrase, heat-shock 90 family members, bacterial CheA histidine kinases, and MutL DNA mismatch protein family members) (Dutta and Inouye 2000). Biochemical analyses of archaeal TopoVI showed that ATP binding is involved in TopoVIB homodimerization and plays an important role in activating the DNA cleavage complex. Three conserved basic residues within the so-called KGRR loop and the C-ter Stalk/WKxY motif region of TopoVIB interact with DNA, with a preference for negative supercoiled DNA (Corbett et al. 2007; Wendorff and Berger 2018). The KGRR loop is part of the GHKL domain, whereas the C-ter Stalk/WKxY motif connects the transducer domain with TopoVIA and plays a role in DNA binding (Wendorff and Berger 2018). Studies on the H2TH domain of TopoVI showed that it plays a role in facilitating DNA strand passage (Wendorff and Berger 2018). The H2TH domain is followed by the transducer domain that contains a “switch lysine” required for ATP hydrolysis (Corbett et al. 2005).

In eukaryotes, TopoVI has been identified in vascular plants, green and red algae, and some protists (Hartung and Puchta 2001; Sugimoto-Shirasu et al. 2002; Yin et al. 2002; Malik, Ramesh et al. 2007). Differently from archaea, where TopoVI is the only Topo II activity (Forterre and Gadelle 2009), eukaryotes also have Type IIA topoisomerases, raising the question of the functional specificity of TopoVI relative to TopoIIA. In plants, TopoVI is needed for DNA endoreduplication, potentially requiring a decatenation activity (Hartung et al. 2002; Sugimoto-Shirasu et al. 2002; Yin et al. 2002; Kirik et al. 2007). *Arabidopsis thaliana* TopoVI has been proposed to play a role in the plant stress response by regulating gene expression (Simkova et al. 2012) and a recent study showed a role for *A. thaliana* TopoVI in heterochromatin organization (Méteignier et al. 2022). Potential TopoVI accessory factors (BIN4 and RHL1) have been identified in *A. thaliana* mutants that have similar dwarf phenotypes (Sugimoto-Shirasu et al. 2005; Breuer et al. 2007; Kirik et al. 2007). Furthermore, various independent Yeast-two-Hybrid (Y2H) experiments indicate that BIN4 and RHL1 may be part of the TopoVI complex in *A. thaliana* (Sugimoto-Shirasu et al. 2005; Breuer et al. 2007; Kirik et al. 2007).

The second specific feature of eukaryotes, compared with archaea, is the TopoVI-like (TopoVIL) complex that is composed of two proteins, SPO11 and TOPOVIBL, homologs to the TopoVIA and TopoVIB subunits (Bergerat et al. 1997; Keeney et al. 1997; Robert, Nore et al. 2016; Robert, Vrielynck et al. 2016; Vrielynck et al. 2016). The TopoVIL complex induces DSBs at meiosis onset, an essential step to initiate homologous

recombination and to create connections between homologous chromosomes for their proper segregation at meiosis I (Hunter 2015). Although no biochemical information about TopoVIL activity is currently available, *in vivo* data indicate that the TopoVIL complex can introduce DSBs, form a protein-DNA cleavage complex (like TopoVI), but differently from TopoVI, broken DNA ends are not predicted to be re-ligated (Bergerat et al. 1997; Keeney et al. 1997). Indeed, the phospho-tyrosine linkage of the cleavage complex is not reversed, and SPO11 is released from substrate as a SPO11-oligo intermediate by endonucleolytic cleavage (Neale et al. 2005; Lange et al. 2011). The broken ends are repaired by homologous recombination (Baudat et al. 2013). It results that SPO11 is not recycled and undergoes only one catalytic reaction. SPO11 is present in most eukaryotes examined and its phylogeny has been described (Ramesh et al. 2005; Malik, Pightling et al. 2007; Bloomfield 2016), consistent with the meiotic pathway conservation for sexual reproduction. Conversely, TOPOVIBL conservation and phylogeny remain to be determined because so far, TOPOVIBL has been identified only in few species, partly due to the high degree of divergence observed among orthologs (Robert, Nore et al. 2016).

Here, we extended the identification of TopoVI and TopoVIL families. For TopoVI, our analysis revealed new structural properties of TopoVIA and TopoVIB, and insights into the structural properties of BIN4 and RHL1 in eukaryotes. For TopoVIL, we revealed the widespread conservation of TOPOVIBL, with astonishing divergence, but a structurally conserved transducer subdomain that interacts with SPO11, particularly in metazoans. Lastly, we showed that REC114, a direct partner of TOPOVIBL, is broadly conserved in metazoans together with MEI4, a REC114 interacting protein.

Results and Discussion

Phylum- and Species-specific Conservation of TopoVIA, TopoVIB, BIN4, and RHL1 in Eukaryotes

The phylogeny of TopoVI, the first identified family of type IIB topoisomerases, was first limited to Archaea and Viridiplantae (Forterre and Gadelle 2009) and was then extended after the identification of other type IIB family members: TopoVIII, present in archaea, and conjugative plasmids of bacteria (Gadelle et al. 2014), and mini-A proteins often encoded by viruses (Takahashi et al. 2020).

The identified orthologs revealed the high conservation of both TopoVIA and TopoVIB subunits (Forterre et al. 2007). Within the DNA binding domain (5Y-CAP or WHD) of the TopoVIA subunit, specific residues are conserved, particularly the catalytic tyrosine residue (Y) (Bergerat et al. 1997). Within the Toprim domain, residues involved in Mg⁺⁺ binding are conserved (DxDxxG) (Nichols et al. 1999). TopoVIA also includes a conserved helix at its N terminus involved in the interaction with TopoVIB (Corbett et al. 2007; Graille et al. 2008).

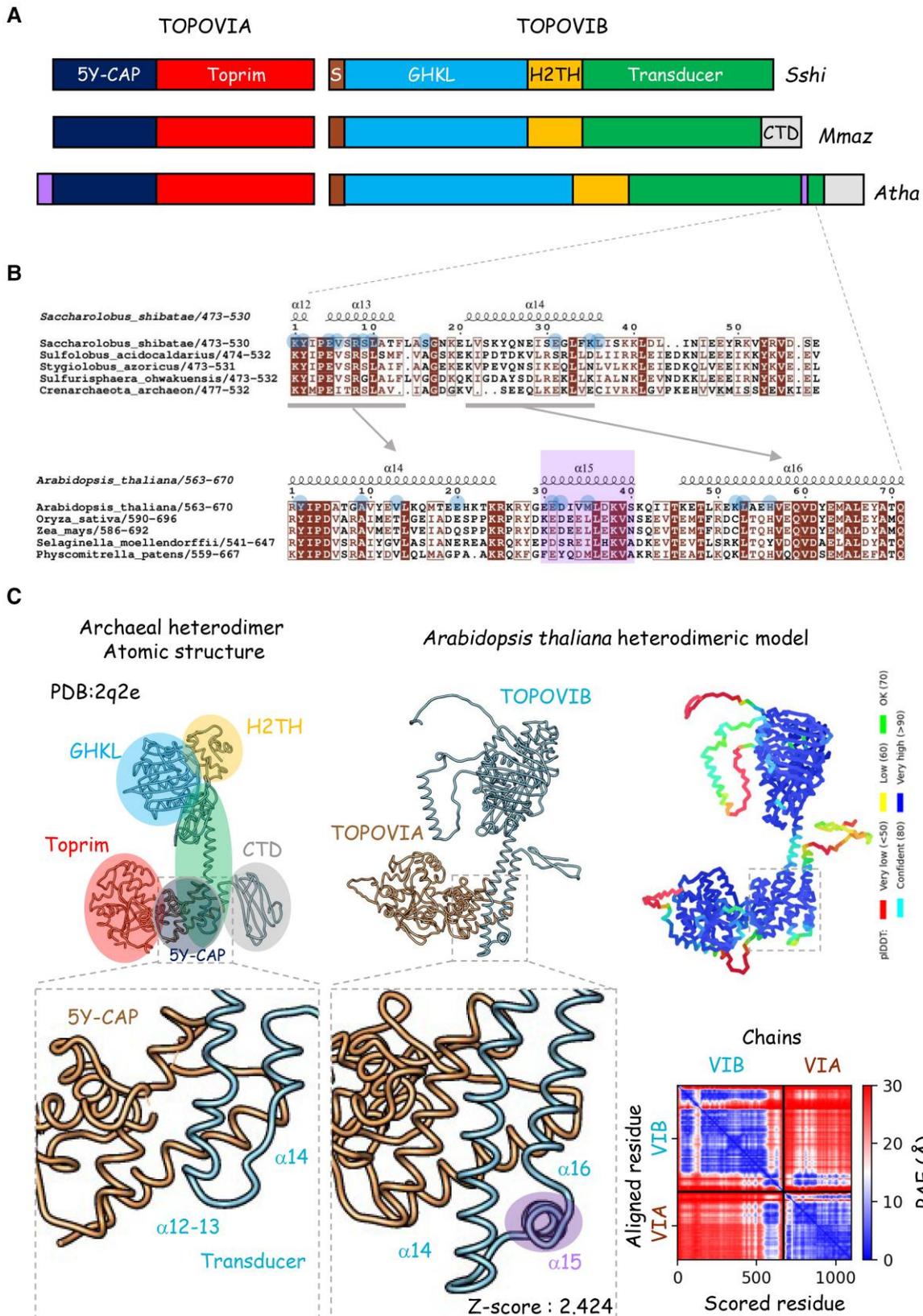


Fig. 1. Archaeal and *A. thaliana* TOPOVIB interact differently with TOPOVIA. (A) N- and C-terminal extensions in the eukaryotic TOPOVIA and TOPOVIB subunits. The domain organization of the *Saccharolobus shibatae* (*Sshi*), *Methanosarcina mazei* (*Mmaz*), and *Arabidopsis thaliana* (*Atha*) TopoVI subunits is shown. 5Y-CAP: Catabolite Activator Protein, Toprim: Topoisomerase-primase, S: N-terminal strap, GHKL: Gyrase B, Hsp90, Histidine Kinase, MutL, H2TH: Helix-2Turn-Helix domain, CTD: C-Terminal Domain. The CTD is present in TOPOVIB of some archaea (*M. mazei*) but not in others (*S. shibatae*). The additional α helices in *A. thaliana* TOPOVIA and TOPOVIB are highlighted in purple. (B) Alignment of archaea and plantae TOPOVIB transducer C-terminal subdomain showing the predicted additional helix $\alpha 15$. Residues interacting with TOPOVIA are highlighted with spheres. Predicted *A. thaliana* TOPOVIB structural elements are extracted from the AlphaFold database

Moreover, it was proposed that a short conserved motif in the C-terminal Toprim subdomain, called T2BI-box (for Type IIB topoisomerases Interaction), is involved in its activity based on phylogenetic studies of TopoVIA and miniTopoVIA variants (Takahashi et al. 2020). In TopoVIB, the conserved motifs include those in the GHKL domain for ATP binding and hydrolysis (BF motifs) (Bergerat et al. 1997), and residues in the transducer domain that includes the interaction interface with TopoVIA (Corbett et al. 2007; Graille et al. 2008). Several conserved motifs contribute to DNA binding, dimerization and ATP hydrolysis: the H2TH, the N-strap, a basic loop in the GHKL domain, the anchoring asparagine, the switch lysine, and the WKxY motif (Corbett and Berger 2003; Wendorff and Berger 2018).

Besides Viridiplantae, TopoVI homologs have been identified in several taxa among Eukarya (Malik, Pightling et al. 2007; Bloomfield 2016). We re-investigated their phylogeny and detected two novel structural hallmarks of eukaryal TopoVI (investigated genomes are provided as supplementary data, Supplementary Material online) (table 1): 1) an additional C-terminal helix (*A. thaliana* predicted helix $\alpha 15$) in TopoVIB, located between the two helices of the transducer domain involved in the interaction with TopoVIA (figs. 1A and 1B), and 2) a helix (helix $\alpha 1$) in TopoVIA, located N-terminally to the helix involved in the interaction with TopoVIB (fig. 1A and supplementary fig. S1A, Supplementary Material online). Based on an Alpha Fold (AF2) model structure of the *A. thaliana* TopoVIA-TopoVIB heterodimer, compared to *Methanosarcina mazei*, in *A. thaliana*, TopoVIB predicted helix $\alpha 15$ leads to a partially overlapping but distinct interaction between the A and B subunits (fig. 1C). The model gives confident or very high scores (pLDDT) excepted for residues located in loops or disordered regions such as the N- and C-terminal region of TopoVIB. The folding of the TopoVIB transducer domain and of TopoVIA N-terminal region of 5Y-CAP is validated by the low predicted alignment error (PAE) values (fig. 1C). The function of the additional TopoVIA N-terminal helix (helix $\alpha 1$ in supplementary fig. S1A, Supplementary Material online) is unknown. Based on model structures, it interacts with the Toprim domain (supplementary fig. S1B, Supplementary Material online). This interaction is also observed in the homodimer model (supplementary fig. S1C, Supplementary Material online) where the TopoVIA subunits are in a head to tail configuration as in *M. jannashii* TopoVIA (Nichols et al. 1999). Interestingly, these two modeled helices (helix $\alpha 1$ in TopoVIA and helix $\alpha 15$ in TopoVIB) we identified in Eukarya are also predicted to be present in Asgard archaea, a sister group of

eukaryotes (Spang et al. 2015; Seitz et al. 2016) (Zaremba-Niedzwiedzka et al. 2017) as shown by sequence alignments (supplementary figs. S2 and S4A, Supplementary Material online). Model structures highlighted the similarity between Asgard and *A. thaliana* for the folding of the predicted additional helical domain of TopoVIA (supplementary fig. S3A, Supplementary Material online), of the homotypic dimers (supplementary fig. S3B, Supplementary Material online) and of the predicted additional helix within the transducer of TopoVIB at the interface with TopoVIA (supplementary fig. S4B, Supplementary Material online). This illustrates the common ancestry between Asgard and Eukarya and establishes that these two helices are not eukaryal-specific. Analysis of eukaryotes in which TopoVI is present (Archaeplastida, some opisthokonts and stramenopiles, Alveolata and Rhizaria [SAR]) showed that TopoVIB contained the motifs required for dimerization and ATP binding/hydrolysis (N-strap; and N, G1, and G2 boxes within the GHKL domain) (supplementary figs. S5A, B, Supplementary Material online), the H2TH domain, and a conserved transducer domain, including the anchoring asparagine, the WKxY motif, and the switch lysine (supplementary fig. S6B, Supplementary Material online). Similarly, in TopoVIA, the residues important for the catalytic activity, DNA interaction (notably Mg^{++} binding), and the T2BI box were well conserved from Archaeplastida to SAR (supplementary fig. S6A, Supplementary Material online).

The eukaryal/Asgard-specific helices may represent interfaces for interaction with partners, thus potentially contributing to additional regulation of eukaryal TopoVI compared with archaeal TopoVI. Two putative TopoVI partners have been identified in *A. thaliana* based on genetic screens: BIN4 (also called MIDGET) and RHL1 (also called HYP7). Plants carrying mutations in either of these genes have the same phenotype as TopoVIA (SPO11-3) and TopoVIB mutants (Sugimoto-Shirasu et al. 2005; Breuer et al. 2007; Kirik et al. 2007). Moreover, in yeast two-hybrid assays BIN4 interacts with itself, with RHL1 and with TopoVIA, while RHL1 also interacts with TopoVIA (Sugimoto-Shirasu et al. 2005; Breuer et al. 2007; Kirik et al. 2007). In bimolecular fluorescence complementation assays (BiFC), RHL1 was found to interact with BIN4 and TopoVIA (Kirik et al. 2007). In co-immunoprecipitation experiments (CoIP) BIN4 was found to interact with TopoVIB (Météignier et al. 2022). These observations suggested that the activity of plant TopoVI may require RHL1 and BIN4 through interactions that remain to be determined. To gain insight into these observations, we analyzed RHL1 and BIN4 conservation in Eukarya and developed structural models.

(AF-Q9LZ03-F1-model_v3). (C) Interaction of the TOPOVIB transducer and TOPOVIA 5Y-CAP in *M. mazei* (PDB:2q2e), and as predicted in *A. thaliana* (AF2 model). The whole structures are shown in the upper panels, with magnified views of the interacting α helices in the lower panels. Archaeal TOPOVIA/B domains are circled and colored as in (A). The protein interaction Z-score assessment (PIZSA) is indicated. The protein complex is classified as a predicted stable association if the PIZSA Z-score is greater than 1.5. Model confidence (predicted "local Distance Difference Test"—pLDDT- score for each amino-acid residue is colored as indicated) is shown on the top right side with the corresponding predicted alignment error (PAE) plot on beneath.

Table 1. Summary of TopoVI and TopoVIL evolutionary Conservation and Diversity. Domains and Motifs Present (+) or Absent (–) or Present in Some Taxa and Absent in Others (+/–) are Shown for the two Subunits of TopoVI (TopoVIA and TopoVIB) and of TopoVIL (Spo11 and TopoVIBL). na, not applicable due to Lack of Domain; L, Linker Domain Substituting for the H2TH in TopoVIBL; 2H and 3H, 2 and 3 Helices, Respectively.

		Nter helix	5Y-CAP		TOPRIM			Cter
			TopoVIB or BL-interacting domain	Catalytic Y	Mg++ binding	T2BI box		
TopoVIA	Archaea	-	+	+	+	+	-	
	Asgard	+	+	+	+	+	-	
	Eukarya	+	+	+	+	+	-	
Spo11	Eukarya	+/-	+	+	+	+	*+ (<i>C. elegans</i>)	
		GHKL		H2TH	Transducer			CTD
		N strap	Bergerat Fold (N, G1, G2 boxes)	Anchoring N	WKxY	Switch K	Spo11-interacting domain	
TopoVIB	Archaea	+	+	+	+	+	+ (2H)	+/-
	Asgard	+	+	+	+	+	+ (3H)	+
	Eukarya	+	+	+	+	+	+ (3H)	+
TopoVIBL	Eukarya	+/-	+/-	- (L)	+/-	+	+ (3H)	+ (Rec114 interacting domain)
	GHKL	+	+	- (L)	+	+	+ (3H)	+ (Rec114 interacting domain)
	GHKL-like	-	-	- (L)	-	+	+ (3H)	+ (Rec114 interacting domain)
	No GHKL	na	na	na	na	na	+ (3H)	+ (Rec114 interacting domain)

This analysis led to three main conclusions: 1) BIN4 and RHL1 are evolutionarily conserved in plants and also in some opisthokonts (including choanoflagellates, a sister group of metazoans) and in SAR. We did not find any BIN4 or RHL1 homolog in metazoans, fungi, amoebozoans, foraminiferans and excavates (fig. 2A); 2) their presence is strictly correlated with that of TopoVI (A and B) with striking concomitant losses in some alveolates and stramenopiles (fig. 2A); 3) analysis of BIN4 and RHL1 orthologs revealed for each protein a conserved central domain predicted to be mostly composed of β -strands (fig. 2B and supplementary fig. S7A, Supplementary Material online). BIN4 and RHL1 models showed a striking structural overlap when monomers were compared (supplementary fig. 7B, Supplementary Material online, right panel). In the model of the BIN4/RHL1 heterodimer, the interaction between monomers involves residues of each conserved core (fig. 2C and supplementary fig. S7A, Supplementary Material online). We tested the conservation of this interaction by assaying the properties of the BIN4 and RHL1 orthologs from the model brown alga *Ectocarpus siliculosus* (Coelho et al. 2020). We found that *E. siliculosus* BIN4 and RHL1 interacted and that the conserved domain of each protein was necessary and sufficient for the interaction (supplementary fig. S8A, B and C, Supplementary Material online): In BIN4, among the three domains tested, only the central domain (BIN4 trunc2) corresponding to the conserved core domain, showed interaction with the RHL1 conserved core domain. In RHL1, among the two domains tested, only the conserved core domain (RHL1 trunc1) showed interaction with BIN4 (supplementary

fig. S8B and C, Supplementary Material online). We also uncovered a similarity between our BIN4/RHL1 model and the interacting β -barrel domains of the Ctf8/Dcc1 heterodimer (fig. 2C). This heterodimer binds to the C-terminal end of Ctf18, a subunit of the RFC-CTF18 complex (Wade et al. 2017). Dcc1 also possesses three winged-helix domains (WHD), one of which binds ssDNA and dsDNA. The Ctf8/Dcc1 complex thus creates a bridge between the DNA and RFC (Wade et al. 2017). The RFC-CTCF18 complex plays a role in loading PCNA at replication forks (Bermudez et al. 2003) and in sister chromatid cohesion (Mayer et al. 2001; Kawasumi et al. 2021). The similarity with BIN4 and RHL1 is thus particularly interesting since BIN4 and RHL1 also bind to DNA in vitro (Sugimoto-Shirasu et al. 2005; Breuer et al. 2007). Furthermore, we detected by modelling an interaction between the C-terminal domain of TopoVIB and the interface of β -strands from the BIN4/RHL1 heterodimer from *A. thaliana* (fig. 3A). In TopoVIB, this interaction involves residues in the transducer domain, and in the C-terminal domain for the interaction with RHL1 (zoom in fig. 3A and supplementary fig. S6B, Supplementary Material online) but neither the WKxY motif nor the H2TH domain. Similar model structures were obtained with *E. siliculosus* proteins (fig. 3B). However, Y2H assays reported interactions of *A. thaliana* BIN4 and RHL1 with TopoVIA, not TopoVIB (Sugimoto-Shirasu et al. 2005; Breuer et al. 2007), and a CoIP assay identified an interaction between BIN4 and TopoVIB (Métégnier et al. 2022). In our Y2H conditions, we could not detect any interaction between *A. thaliana* BIN4 or RHL1 and TopoVIB (supplementary

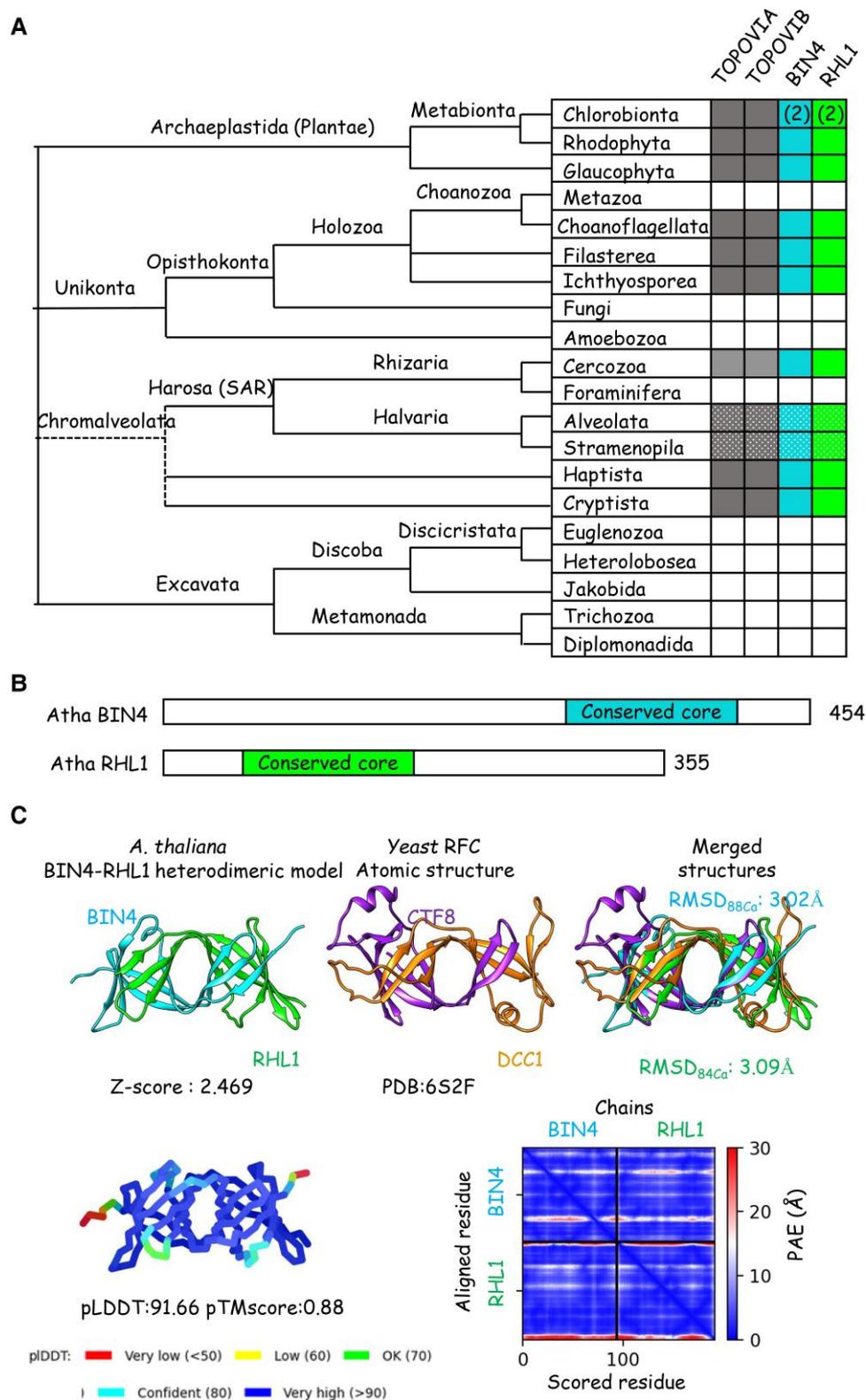


FIG. 2. TopoVI holoenzyme subunit conservation among eukaryotic taxa. (A) TOPOVIA, TOPOVIB, BIN4, and RHL1 conservation in Eukarya. The presence of TOPOVIA, TOPOVIB, BIN4, and RHL1 orthologs is shown by filled boxes. Partial occurrences in Alveolata (e.g., present in Dinophyta but absent in Perkinsozoa and Apicomplexa) and in Oomycota (e.g., all four genes are present in *Aphanomyces astaci* while undergoing pseudogenization concomitantly in the closely related *A. invadens*) are indicated by dotted boxes. The presence of BIN4 and RHL1 duplicates in Chlorobionta (i.e., *Physcomitrium patens*) is also indicated. The tree on the left side does not represent the genetic distances, branch lengths being arbitrary and only used to illustrate the consensus view of the phylogenetic relationships among the major eukaryotic groups. Dashed lines indicate that Chromalveolata is currently not considered to be monophyletic. (B) Schematic representations of *A. thaliana* (Atha) BIN4 and RHL1. The position of the conserved cores are shown. (C) *Arabidopsis thaliana* BIN4 and RHL1 are predicted to interact through evolutionarily conserved β -barrel domains that have structural similarities with the RFC CTF8/DCC1 dimerization module. Upper panels: Heterodimeric

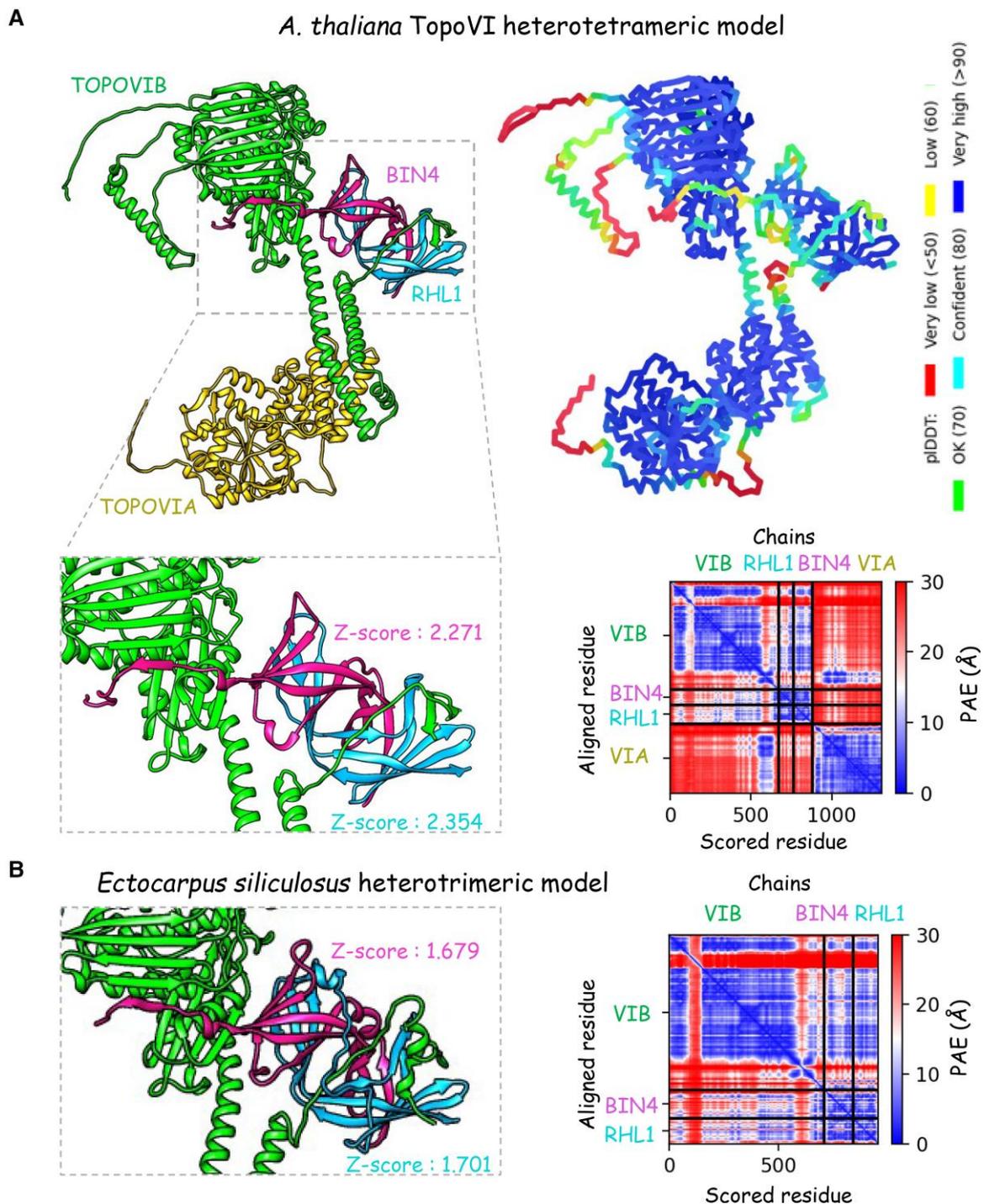


Fig. 3. RHL1 and BIN4 are predicted to interact with the TOPOVIB CTD and GHKL domains through their respective evolutionarily conserved (CORE) regions. (A) Modeled interactions between *A. thaliana* TopoVI holoenzyme subunits (as determined by AF2). The predicted interface between TOPOVIB (green), RHL1^{CORE} (blue), and BIN4^{CORE} (pink) is magnified in the lower left panel. SPO11-3/TOPOVIA (yellow) has been included to show distinct predicted interfaces (see fig. 1). Interaction Z-scores between the TOPOVIB CTD/GHKL domains and RHL1^{CORE} or BIN4^{CORE}, respectively, are indicated. Model confidence (pLDDT) (upper left) is shown in the upper right panel with the corresponding PAE plot on beneath. (B) Conservation of the predicted interface between TopoVI holoenzyme subunits in *E. siliculosus*. A magnification of the evolutionarily conserved interacting domains is shown on the left panel, with chain coloring as in (A). Z-scores are indicated. The corresponding PAE plot is shown in the right panel.

model of the BIN4/RHL1 dimerization module with the interaction Z-score (left). Atomic structure of the CTF8/DCC1 interface (PDB:6S2F) (middle). Merging of the CTF8/DCC1 and BIN4/RHL1 dimerization modules (right). The docking root-mean-square-deviation (RMSD) values are indicated. Lower panels: Model confidence as shown with predicted LDDT (left) and alignment error (right). Global pLDDT and pTM (predicted “Template Modeling”) scores are indicated.

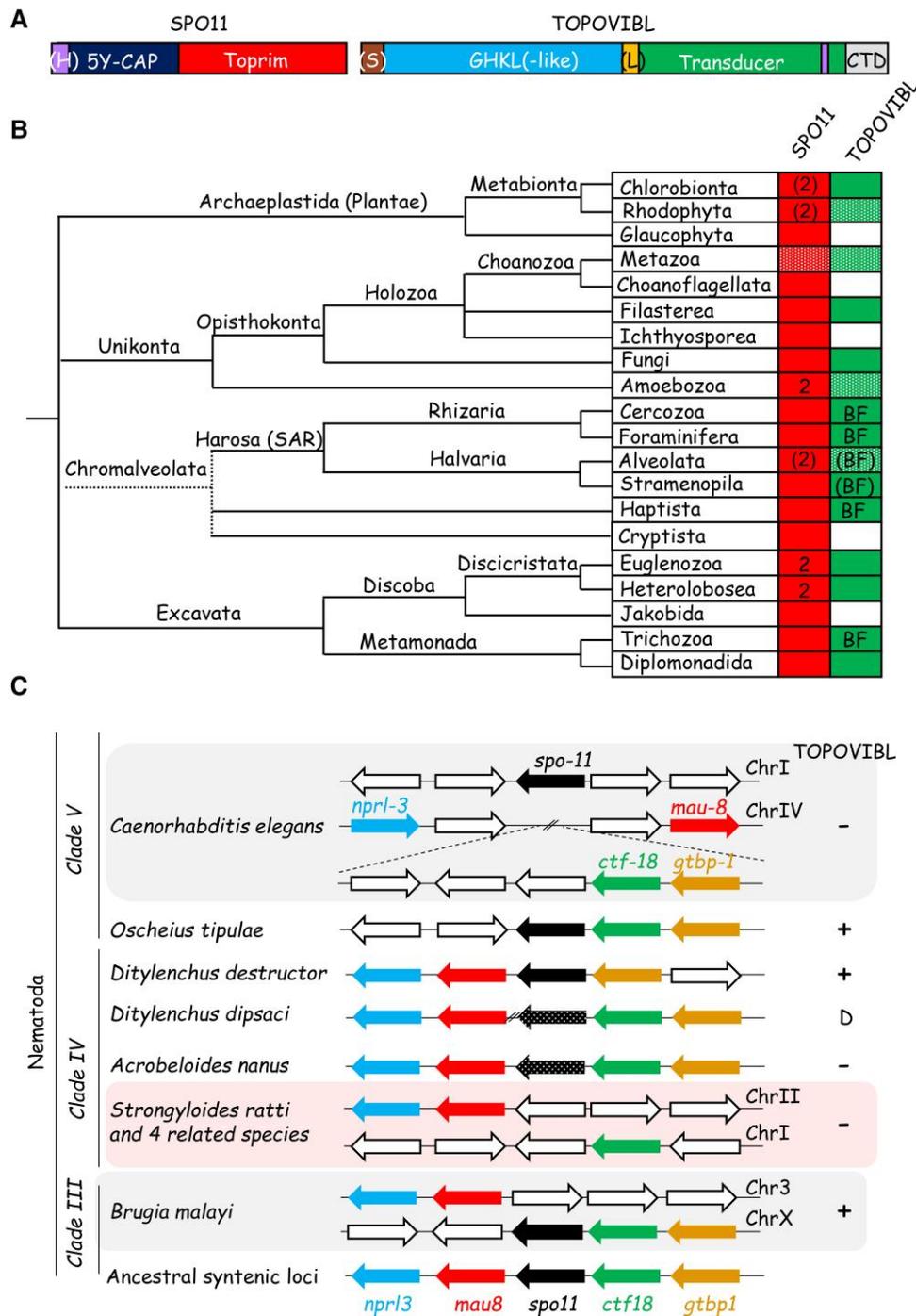


FIG. 4. Taxa-specific conservation of the TOPOVIBL BF motifs. (A) Domains of SPO11 and TOPOVIBL. Some SPO11 orthologs have an additional N-terminal α helix (H, purple box). The structure of TOPOVIBL orthologs is variable: with or without the N-terminal Strap (S) motif, GHKL domain with or without the conserved BF motifs (termed GHKL-like), and with a variable C-terminal domain (CTD). The region between the GHKL(-like) and the transducer domain is not conserved, and is called "Linker" domain (L). (B) Conservation of TopoVIL subunits in Eukarya. The presence (filled box) or partial occurrence (dotted box) of SPO11 and TOPOVIBL is indicated on the right. Some taxa have two SPO11 paralogs. In some groups, TOPOVIBL contains conserved BF motifs. (BF) indicates BF motif heterogeneity among species within a phylum. A white box indicates that no TOPOVIBL ortholog could be identified. The phylogenetic relationships are described as in [figure 2](#). (C) *spo11* and *top6bl* loci have been lost or are undergoing degeneration (D) processes (pseudogenization) in some nematodes (Clade IV and V species). Gene synteny are shown, centered on chromosomal regions homologous to the presumptive ancestral *spo11* linkage group (present in *D. destructor*), with transcription units indicated by arrows. Colored arrows only show units homologous to *C. elegans* genes. Pseudogenes are indicated with dotted arrows.

[fig. S8D and E, Supplementary Material](#) online, diploids a and c). However, we reasoned that since BIN4/RHL1 are predicted to form a complex, the expression of both BIN4 and RHL1 might be necessary for the interaction with TOPOVIB. Indeed, we could show by Y2H that an interaction was detected when all three proteins were co-expressed in yeast ([supplementary fig. S8D and E, Supplementary Material](#) online, diploids b and d), supporting the proposed model.

The implication of these two putative TopoVIB partners could be important for the eukaryal TopoVI activity. As A.

thaliana TopoVIA/B mutants are defective in endoreduplication ([Hartung et al. 2002; Sugimoto-Shirasu et al. 2002](#)), it was proposed that the TopoVI decatenation activity is involved in resolving catenated chromatids generated by endoreduplication. In fact, although TopoVI possesses both relaxation and decatenation activities, it was shown that *M. mazei* TopoVI is preferentially a decatenase in vitro ([McKie et al. 2022](#)). Interestingly, the functional link with DNA replication is also supported by the interaction detected by immunoprecipitation between *Pyrococcus abyssi* TopoVI and the replication fork components PCNA, NusC

and RFC (Ren et al. 2009). In Eukarya, some factors may also recruit TopoVI to replication forks, and BIN4 and RHL1 may stabilize such interactions. BIN4 was also shown to interact with the methionine synthase MAT3 with a potential role on H3K9 methylation (Méteignier et al. 2022).

Early Origin of the TopoVIL complex and TOPOVIBL divergence

High Conservation of SPO11 Orthologs

SPO11 is conserved in Eukarya, in agreement with the observation that most eukaryotes undergo sexual reproduction (Malik, Ramesh et al. 2007). Dictyostelids, which lack a *Spo11* gene, are the only known exception. This observation raises a yet unsolved puzzle for meiotic recombination which is dependent on *Spo11* in most species but not in dictyostelids (Goodenough and Heitman 2014; Bloomfield 2018). It is thought that early after the emergence of the last eukaryotic common ancestor (LECA), its descendants possessed two SPO11 copies (SPO11-1 and SPO11-2) that are detected based on the conservation of the two major domains (5Y-CAP and Toprim) (Sprink and Hartung 2014). Loss of one copy occurred several times independently during evolution, and only some species have retained both paralogs (SPO11-1 and SPO11-2 in Plantae, and also in some amoebozoans, alveolates and excavates). Conversely, many eukaryotes have only one SPO11 (SPO11-1 in opisthokonts, some amoebozoans, rhizarians, haptophytes, cryptophytes, and excavates; and only SPO11-2 in stramenopiles, red and green algae) (Malik, Ramesh et al. 2007). SPO11 splice variants are conserved in plants (Sprink and Hartung 2014; Ku et al. 2020), and in mammals (Romanienko and Camerini-Otero 1999). Interestingly these conserved splice variants correspond to two isoforms, one with and one without the N-terminal helix required for SPO11 interaction with TOPOVIBL. The variant without this helix should be catalytically inactive. In mice, expression kinetics data showed that the variant with the N-terminal helix (called SPO11 β) is predominant at prophase onset, when meiotic DSBs form, and the variant without this helix (SPO11 α) is predominant later in prophase (Bellani et al. 2010). Therefore, splicing regulation might contribute to modulate meiotic DSB activity.

In our investigation of SPO11 orthologs, we identified SPO11 in all eukaryal supergroups (fig. 4B), with the exception of dictyostelids and we identified another taxon where SPO11 is apparently absent: Tylenchomorpha (Nematoda Clade IV) (fig. 4C). Syntenic analyses showed the specific loss of SPO11 in *Strongyloides ratti* and four related species or undergoing degeneration in *Acrobeloides nanus* as well as *Ditylenchus dipsaci*, but being apparently still functional in the related *D. destructor* (fig. 4C). We also observed some species with two SPO11 paralogs, and with a high degree of conservation including the catalytic site (catalytic tyrosine, Mg⁺⁺ binding site, and T2BI box) (figs. 4A and B, table 1 and supplementary fig. S9, Supplementary Material online). Moreover, we found that a conserved predicted additional

N-terminal helix, absent in the characterized archaeal TopoVIA, and adjacent to the helix interacting with TOPOVIBL, was present in taxa-specific paralogs, reminiscent of the previously described predicted additional N-terminal helix in Asgard archaeal and eukaryal TopoVIA (supplementary fig. S10A, Supplementary Material online). *Mus musculus* SPO11 model shows the interaction of the predicted additional N-terminal helix with the Toprim domain (supplementary fig. S10B, Supplementary Material online), thus similar to *A. thaliana* and Asgard TopoVIA models (supplementary figs. S1B and S3A, Supplementary Material online). In *S. cerevisiae*, cross-linking and mass spectrometry on the Spo11 core complex (Spo11/Rec102/Rec104/Ski8) identified interactions between the N-terminal helix of Spo11 and both Rec102 and Rec104 (Claeys Bouuaert, Tischfield et al. 2021) but no or weak interactions between the N-terminal helix and the Toprim domain of Spo11. One potential interpretation for these differences is that the N-terminal helix of Spo11 could fold in distinct configurations, that is with the Toprim domain as suggested by models or with a partner as shown in *S. cerevisiae*. These alternative configurations may be involved in the regulation of Spo11 activity. The conserved T2BI motif is of unknown function. It has been proposed for *Methanococcus jannaschii* TopoVI to regulate its activity (Takahashi et al. 2020), based on the proximity of some T2BI motif residues and the catalytic tyrosine of TopoVIA (Nichols et al. 1999). It is intriguing to note that in *S. cerevisiae*, Ski8 interacts with Spo11 in a region overlapping with the T2BI motif (Arora et al. 2004; Claeys Bouuaert, Tischfield et al. 2021). Based on analysis of the *A. thaliana* Ski8 ortholog, it was however concluded that the meiotic role of Ski8 observed in *S. cerevisiae* is not conserved (Jolivet et al. 2006). The putative role of the T2BI motif may thus be executed through molecular interactions that differ in distinct species.

We observed an exceptional feature of *C. elegans* SPO-11: a sub-genus specific predicted additional helix at the C-terminal end that might be implicated in partner interaction (supplementary fig. S10, Supplementary Material online). The presence of this extension in several *Caenorhabditis* species (clade V) as well as in some Clade IV nematodes (Tylenchoidea) correlates with the apparent loss of TOPOVIBL as determined by syntenic analyses (supplementary fig. S11B, Supplementary Material online).

Evolution and High Diversity of TOPOVIBL homologs

TOPOVIBL phylogeny is more complex due to its high level of divergence, as already observed in a preliminary phylogenetic analysis within plantae and metazoans (Robert, Nore et al. 2016). In fact, we observed a large heterogeneity between species, from the presence of a conserved to a highly divergent B subunit to species where TOPOVIBL is undetectable. The divergence pattern suggests independent events, leading to the loss of similarity, and likely functional alterations relative to its TopoVIB ancestor subunit (figs. 4A and B, and table 1). Importantly, it has been

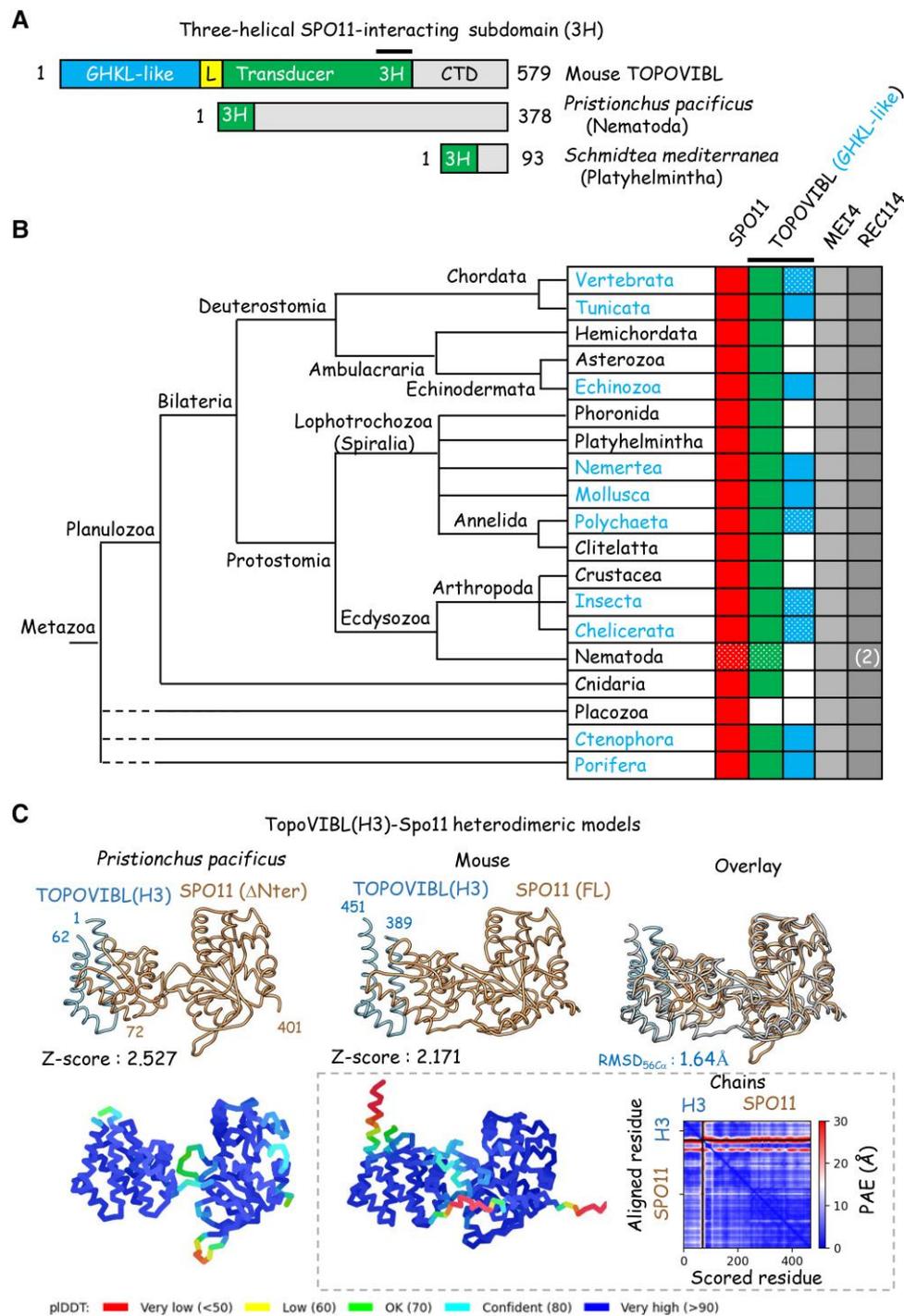


Fig. 5. Conservation of meiotic DSB proteins among metazoans: phyla-specific loss of the TOPOVIBL GHKL-like domain and drastic reduction of the SPO11-interacting transducer domain. (A) Schematic representation of metazoan TOPOVIBL proteins with a GHKL-like domain, a transducer including the predicted three α helices (3H) that interact with SPO11 and a C-terminal domain (CTD) (*M. musculus*), or with only 3H and CTD (*P. pacificus*), or with 3H and a minimal CTD including one α -helix that might interact with REC114 (*S. mediterranea*) (see alignments in [supplementary figs. S17–S19, Supplementary Material](#) online). (B) Schematic phylogenetic tree showing the presence (filled boxes) or not (empty boxes) of SPO11, TOPOVIBL, REC114 and MEI4 homologs in the indicated metazoan phyla. Dotted boxes indicate heterogeneity (in the same taxon, some species contain SPO11 and/or TOPOVIBL and others do not). Taxa in which TOPOVIBL comprises a GHKL-like domain are indicated in blue letters and with blue filled boxes. Note that in some species among Vertebrata, Polychaeta, Insecta, and Chelicerata, TOPOVIBL orthologs lack the GHKL-like domain and have a transducer limited to the 3H domain (see [supplementary fig. S18, Supplementary Material](#) online). The phylum (Nematoda) with species (e.g., *C. elegans*) having two REC114 paralogs is indicated. The tree on the left side does not represent the genetic distances, branch lengths being arbitrary and only used to illustrate the consensus view of the phylogenetic relationships among the major metazoan taxa. Dashed lines indicate taxa for which the relationships are still debated. (C) Modeled structures of TOPOVIBL–SPO11 interaction in *P. pacificus* (left panel) and *M. musculus* (middle panel). AF2 models only include TOPOVIBL 3H domains (in blue) and predict tight interaction with the modeled helices α 2–3 of SPO11 (in brown). Overlaid structures are shown in the right panel. The interaction Z-scores and RMSD values (for the merged heterotypic structures) are indicated. Model confidence (pLDDT) is shown beneath of each predicted structure. The corresponding PAE plot for the modeled mouse heterodimer is shown in the lower right panel.

predicted that TopoVIL activity differs from that of TopoVI at least for the religation step, which is absent or repressed in the TopoVIL reaction (Robert, Vrielynck et al. 2016). Moreover, as some Type IIB topoisomerases can cleave DNA without ATP (Gadelle et al. 2014), the ATP binding and hydrolysis activity may not be an absolute prerequisite for TopoVIL function and/or this activity or may be substituted by alternative molecular interactions.

In several eukaryotic taxa, TOPOVIBL is readily identifiable due to its high conservation with TopoVIB, specifically in the GHKL ATP binding/hydrolysis domain. The comparison of archaeal TopoVIB and TOPOVIBL from several chromalveolates highlighted the conservation of the N-strap required for dimerization, and of the N, G1, and G2 boxes required for ATP binding and hydrolysis (supplementary figs. S4 and S12, Supplementary Material online). In these species, the overall TOPOVIBL structure was highly similar to that of archaeal TopoVIB, with the notable exception of the absence of the H2TH domain and the presence of a predicted additional helix in the C-terminal transducer subdomain (supplementary fig. S12, table 1, Supplementary Material online). However, this additional helix was present also in TopoVIB from Asgard archaea and eukaryotes (fig. 1B and supplementary fig. S4A, Supplementary Material online). These highly conserved TOPOVIBL proteins can be differentiated from TopoVIB by the absence of the H2TH domain, which in TopoVI has DNA binding activity and is important for strand passage (Wendorff and Berger 2018). Conversely, in other eukaryotic taxa, TOPOVIBL was highly divergent from TopoVIB. Specifically, they lost the N-strap, the three conserved boxes of the GHKL, the BF motifs (thus referred to as GHKL-like), and the H2TH domain. Therefore, the three predicted helices within the C-terminal transducer subdomain that interacts with SPO11 were the major remaining conserved feature between these TOPOVIBL and TopoVIB (supplementary fig. S13, Supplementary Material online). The WKxY motif was only partially conserved, with mainly a remaining tryptophane residue (supplementary fig. S13, Supplementary Material online).

We observed the divergence of TOPOVIBL with the specific loss of the conserved BF motifs within some taxa, such as Oomycota (Stramenopila). Although some species, such as *Aphanomyces astaci* (Saprolegniales) had a highly conserved GHKL, in Albuginales (e.g., *Albugo candida* and *Albugo laibachii*), mutations disrupted both the N-strap and the conserved boxes (N, G1 and G2) required for ATP binding and hydrolysis (supplementary fig. S14A, Supplementary Material online). This indicates that TOPOVIBL function may have evolved by simultaneously losing a dimerization interface (the N-strap) and the ATPase activity (the BF motifs). However, the structural organization of the transducer region was conserved (supplementary fig. S14, Supplementary Material online). In *A. astaci*, in addition to the conservation of the GHKL BF motifs, models of the GHKL domain could be reconstituted with a high score when compared to *Saccharolobus shibatae* TopoVIB. This applied to both *A. astaci* TopoVIB and TOPOVIBL (supplementary fig. S15, Supplementary

Material online). In *A. astaci* TopoVIB, a large insertion is mainly disordered and gives low pLDDT scores and high PAE values. The models also showed specifically the conservation of the N-strap helix (zoom in supplementary fig. S15, Supplementary Material online panel A). The ATP binding site also should be conserved because its structural models were highly similar between *A. astaci* TopoVIB and TOPOVIBL (supplementary fig. S16, Supplementary Material online).

The TOPOVIBL organization plasticity was further revealed by one additional evolutionary outcome: the separation of the GHKL-like domain and the transducer in some species, for instance in *S. cerevisiae*. In budding yeast, REC102 was identified as homologous to the TOPOVIBL transducer domain (Robert, Nore et al. 2016). It was then proposed that REC104, which does not contain any identifiable feature of a GHKL domain but which interacts with REC102, could substitute for the GHKL domain (Robert, Vrielynck et al. 2016). This was shown by the biochemical characterization of the REC102/REC104 complex (Claeys Bouuaert, Tischfield et al. 2021). Although we could detect proteins with homology to the transducer domain of TOPOVIB, it is currently very challenging to determine in an unbiased manner a GHKL-like domain given the poor conservation observed when it is identified (see supplementary fig. S13, Supplementary Material online). REC104 could be identified as a partner, based on the genetic data that allowed showing its essential role for meiotic DSB formation in *S. cerevisiae* (Galbraith and Malone 1992).

The high TOPOVIBL diversity was recapitulated in metazoans: few phyla have a TOPOVIBL with a GHKL domain lacking *bona fide* BF motifs, thus named GHKL-like (but with a conserved 3D fold, as determined by modeling) while others have only retained structural features in the transducer domain, specifically the three conserved helices (3H) predicted to be involved in the interaction with SPO11 (fig. 5 and supplementary fig. S17, Supplementary Material online). All metazoan TOPOVIBL proteins also contain a C-terminal domain (CTD) that includes an alpha helix interacting with REC114 in *M. musculus* TOPOVIBL (Nore et al. 2021) (supplementary fig. S17, Supplementary Material online). REC114 is an evolutionary conserved protein essential for meiotic DSB formation (Kumar et al. 2010). Its direct interaction with the TopoVIL complex in *S. cerevisiae* and *M. musculus* suggests that it plays a direct role in promoting DSB activity (Claeys Bouuaert, Pu et al. 2021; Nore et al. 2021). An even more drastic evolutionary step is the complete loss of the GHKL-like domain, previously observed in *S. cerevisiae*, *Schizosaccharomyces pombe* and *Drosophila melanogaster* (Robert, Nore et al. 2016). We detected TOPOVIBL lacking any GHKL-like domain in several distant clades, such as Hemichordata, Platyhelmintha, Insecta, and Cnidaria (fig. 5B), and also within groups, such as in Amphibia (Chordata), Polychaeta (Annelida) and Chelicerata (Arthropoda) (supplementary fig. S18B, Supplementary Material online). This indicates repeated independent partial or complete losses of the GHKL-like domain. Therefore, in these species, TOPOVIBL is only

composed of a truncated transducer domain, limited to the 3H SPO11-interacting domain, and a C-terminal domain (fig. 5A, supplementary fig. S19A, Supplementary Material online). We identified a remarkably short TOPOVIBL form in *Schmidtea mediterranea*, where it is essentially composed of the 3H domain and a predicted C-terminal helix (fig. 5A and supplementary fig. S18A, Supplementary Material online). Even in the absence of a GHKL-like domain, the interaction between the 3H domain and SPO11 could be modeled in *Pristionchus pacificus* and *C. monodelphis* (fig. 5C and supplementary fig. S19B, Supplementary Material online). As indicated in the Metazoa tree (fig. 5B), we could not detect any TOPOVIBL protein, even when searching for transducer-only proteins, in Placozoa (*Trichoplax adhaerens* and *Hoilungia hongkongensis*) and the nematode *C. elegans*. TOPOVIBL remained undetected only in very close relatives to *C. elegans*, but not in the more diverging *C. monodelphis*, *C. auriculariae*, and *C. parvicauda* (supplementary fig. S11B, Supplementary Material online). Syntheny analyses show simultaneous loss or degeneration of *spo11* and *top6bl* in Strongyloididae, *A. nanus* and *D. dipsaci* (supplementary fig. S11B, Supplementary Material online and fig. 4). Strikingly, both *spo11* and *top6bl* are undergoing pseudogenization in *D. dipsaci*, as expected for a close functional partnership. In *C. elegans*, *C. auriculariae* and *C. parvicauda*, several meiotic genes, such as *dmc-1*, *mnd-1*, and *hop-2*, have been lost (Rillo-Bohn et al. 2021) (our own observations); however, the loss of the gene encoding TOPOVIBL does not appear to be correlated. Understanding the potential implication of TOPOVIBL loss for SPO11 activity would require first to determine the precise role of the interactions between TOPOVIBL/SPO11 and TOPOVIBL/REC114. As biochemical information is available only in archaea (Corbett and Berger 2005; Graille et al. 2008), *S. cerevisiae* (Claeys Bouuaert, Pu et al. 2021; Claeys Bouuaert, Tischfield et al. 2021) and *M. musculus* (Nore et al. 2021), one could hypothesize that in eukaryotes these interactions may act to modulate SPO11 dimerization and/or binding to DNA, and/or the catalytic site position. In Placozoa and *C. elegans*, alternative strategies may have been selected, for instance by involving other partners for binding to SPO11. Interestingly, while the placozoan *T. adhaerens* and *H. hongkongensis* have a single REC114 homolog, *C. elegans* has two REC114 paralogs (DSB-1 and DSB-2) of which one (DSB-1) interacts with SPO11-1 in Y2H assays (Hinman et al. 2021). This suggests that REC114 function is maintained but acts by interacting with SPO11 rather than with TOPOVIBL.

As mouse REC114 is a direct partner of TOPOVIBL (Nore et al. 2021) and of MEI4 (Kumar et al. 2018), we also explored its evolutionary conservation and identified both REC114 and MEI4 orthologs in all metazoans examined (fig. 5, supplementary figs. S20 and S21, Supplementary Material online).

Concluding Remarks

The phylogenetic reconstruction of TopoVI highlights a new and poorly explored property of TopoVI: the contribution

of interacting partners. The concerted evolution of TopoVIA/B with BIN4 and RHL1 suggests that these two proteins are part of a large holoenzyme complex and are directly involved in regulating TopoVI catalytic activity. This information may help to understand why and how TopoVI has been retained in some eukaryotes. This could indicate a specific need linked to the resolution of catenated DNAs in DNA replication, in cells undergoing endoreduplication. One TopoVI-specific, but not exclusive feature is its decatenase activity. As suggested by recent biochemical studies, this may be linked to specific substrate recognition. The proposition of TopoVI binding to DNA crossing (Wendorff and Berger 2018) is interesting with respect to the topological status of sister chromatids after DNA replication (McKie et al. 2022) and some factors may tether TopoVI to replication forks or intermediates. Our observation of a similarity between BIN4/RHL1 and the Ctf8/Dcc1 heterodimer which acts at replication forks motivates to explore this hypothesis.

The molecular evolution from TopoVI of a DNA cleavage activity to initiate recombination during meiosis remains an enigma: why other type II topoisomerases have not evolved to achieve this function? Is there a specific property of TopoVI that could account for this observation? In meiosis, TopoVIL activity takes place at meiotic prophase I onset, when the DNA has been replicated. A temporal link with DNA replication has been shown in *S. cerevisiae* (Borde et al. 2000) opening the possibility that the TopoVIL substrate could be catenated sister chromatids. However, a requirement for the presence of sister chromatids is not supported by the observation that in *S. cerevisiae* and *S. pombe*, DNA replication is dispensable for meiotic DSB formation (Murakami and Nurse 2001; Blitzblau et al. 2012). TopoVIL activity and therefore the TopoVIL substrate remain to be characterized at the biochemical level. Given the extreme diversity of TopoVIBL in Eukarya (table 1), one can anticipate a range of biochemical activities for TopoVIL, from enzymes with the ability to religate the broken ends in species with a highly conserved GHKL to a DNA cleavage-only activity for others. For those species where TopoVIL may have maintained a DNA cleavage-ligation activity, some additional meiotic factors should be involved to prevent the religation of broken ends and thus to allow the DSB ends to engage into homologous recombination during meiosis. For those species with TopoVIBL containing a poorly conserved GHKL domain (i.e., GHKL-like), the question as to how the SPO11/TOPOVIBL complex dimerizes would be important to answer. Indeed, the common expected features for all TopoVIL are to act as a dimer to introduce two concerted nicks and to form a protein-DNA covalent complex, as observed for TopoVI (Corbett and Berger 2005; Graille et al. 2008). The stability of the dimer, via Spo11 and/or TopoVIBL, is expected to be important and could be a regulatory step for the activity: dissociation of the dimer interface(s) may compromise religation. Minimal versions of TopoVIBL seem to include a three-helical (H3) domain for interaction with Spo11, and a

putative domain of interaction with Rec114 in Metazoa. Partners such as Rec114 and Mei4, may regulate TopoVIL activity by inducing changes in conformation and/or stability of the complex. It has been proposed that a REC114/MEI4 complex with a 2/1 stoichiometry could stabilize the whole holoenzyme complex as a dimer (Nore et al. 2021). Because, Spo11 forms a covalent complex with the DNA (Keeney et al. 1997), in the absence of religation, this complex is *a priori* a dead-end product and the enzyme is not recycled. On the other end, the absence of religation means a high risk for genome instability since the broken ends should be repaired by homologous recombination. These questions also apply to non-meiotic cells, since the expression of some meiotic genes such as *Spo11* has been detected in cancer cells and proposed to participate to tumorigenesis (Jay et al. 2021).

Materials and Methods

Homolog Identification, Alignments, Models, and Syntheny

Most TopoVIA, TopoVIB, BIN4, RHL1, SPO11, TOPOVIBL, REC114, and MEI4 homologs were identified from series of PSI-BLAST and HHMER analyses at the Max Planck Institute (MPI) using the MPI Toolkit (<https://toolkit.tuebingen.mpg.de>). Primary sequences were corrected through genomic translation and alternative exon prediction (<https://web.expasy.org/translate/>). Multiple sequence alignments (MSAs) generated by MAFFT 7.0 (<http://mafft.cbrc.jp/alignment/server>) with the auto mode and default parameters were used as inputs. MSA inputs included only previously validated homologs (as described hereafter), and alternative query sequences were systematically used as the MSA header sequence to improve detection of remote homologs. To discriminate among GHKL-containing homologs, MSA inputs included either full-length proteins or only their transducer domain, particularly the three-helical (H3) SPO11-interacting subdomain. Candidate proteins were validated or not by secondary structure prediction with PSIPRED (<http://bioinf.cs.ucl.ac.uk/psipred>). Several DSB proteins were identified using PSI-BLAST or tblastn at NCBI (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>), notably using transcriptome shotgun assembly (TSA) sequence databases, and dedicated BLAST servers, particularly from the *Caenorhabditis* Genome Project (CGP) (<http://blast.caenorhabditis.org/>). Remote homologs were validated by AlphaFold2 (AF2) modeling (see below), particularly on the basis of their capacity to stably interact with the SPO11 5Y-CAP domain through their H3 domain (Z-score > 1.5). Protein–protein interaction Z-scores were calculated with PIZSA (<http://cospi.iiserpune.ac.in/pizsa/>) with a distance threshold of 4.0 Angstroms.

MSAs were colored with JALVIEW 2.11.1.7 (<https://www.jalview.org/>) using the ClustalX similarity scheme, or with ESPript 3.0 (<http://espript.ibcp.fr>) using the % equivalent similarity score scheme, with the header sequence predicted secondary structure and the

Black&White coloring scheme. Structural elements in [figure 1B](#), [supplementary figs. S1A, S2, S4–6, S7A, S9–10, S11A, S12–14](#) and [S17, S18A, S19A, and S20–21](#), [Supplementary Material](#) online were extracted from the AlphaFold database (<https://alphafold.ebi.ac.uk/>) or from AF2 modeling. Protein modeling was done with an advanced version of AF2 from ColabFold (<https://github.com/sokrypton/ColabFold>). The per-residue confidence score (pLDDT; values greater than 90 indicating high confidence) and distance error for every pair of residues (PAE) were provided by the ColabFold website. The relative position of two domains is confidently predicted then the PAE values are less than 5 Angstroms. Models are available upon request to HMB (via e-mailing to B.d.M.). Models were superimposed with MatchMaker from CHIMERA 1.15 (<https://www.cgl.ucsf.edu/chimera/>). ATP binding predictions were done with I-TASSER (<https://zhanggroup.org/I-TASSER/>). Protein structures were compared with DALI (<http://ekhidna2.biocenter.helsinki.fi/dali/>) and PDBFold (<https://www.ebi.ac.uk/msd-srv/ssm/>).

Syntheny among nematodes was determined through series of BLASTp analyses using SPO11 or TOPOVIBL homologs as query sequences, then by inspection of the flanking loci using the corresponding Genome browsers, as provided by the WormBase Parasite website (<https://parasite.wormbase.org/index.html>). Homologies among recovered loci were identified through comparison to the *C. elegans* genome (https://wormbase.org/tools/blast_blat).

cDNA Cloning

The potential *E. siliculosus* TOPOVI, BIN4 and RHL1 (protein and gene) sequences were identified based on a combination of sequence alignments (H.-M.B., unpublished) and the RNA sequencing data available in the Genome viewer of *OrcAE* (*E. siliculosus* V2 data, <https://bioinformatics.psb.ugent.be/orcae/>). For the validation of the potential coding sequences of *E. siliculosus* genes, the cDNAs were synthesized by reverse transcription, as previously described in [Grey et al. \(2016\)](#), using RNA samples from a diploid sporophyte (Ec702) and a haploid sporophyte (Ec32, male) undergoing apomeiosis. This was followed by PCR amplification of the cDNA samples using gene-specific primers ([supplementary table S1](#), [Supplementary Material](#) online), as previously described in ([Grey et al. 2016](#)). The Gateway Gene Cloning system (Invitrogen) and synthesized cDNA (by GeneArt) optimized for expression in *E. coli* (compatible with codon usage of *S. cerevisiae*) were used for the cloning of full-length and truncated versions of *E. siliculosus* TOPOVI, BIN4, and RHL1 genes into pGADH-GW or pAS2dd.

Yeast two Hybrid Assays

Yeast two hybrid assays were performed as previously described in [Imai et al. \(2017\)](#). A diploid clone that expresses *M. musculus* Gal-4 AD-TOPOVIBL and Gal4 BD-SPO11 β was used as a control for positive interactions (positive control) ([Robert, Nore et al. 2016](#)). For testing interactions upon expression of three proteins, the third partner (RHL1

or BIN4) was expressed without fusion to Gal4 from the pAG422 vector carrying the *ADE2* gene (gift from E. Bertrand). Interactions were assayed based on the expression of the *HIS3* gene.

Acknowledgement

B.d.M. was funded by Agence Nationale de la Recherche (ANR- 18-CE11-0024-01), Prize Coups d'Élan for French Research from the Fondation Bettencourt-Schueller, European Research Council (ERC) Executive Agency under the European Community's Seventh Framework Programme (FP7/2007-2013 Grant Agreement no. [322788]), and CNRS. H.M.B. is currently retired and was granted as a CNRS researcher. We thank Thomas Robert for discussions and critical reading of the manuscript.

Data Availability

All data are incorporated into the article and its online [Supplementary Material](#).

Supplementary Material

[Supplementary data](#) are available at *Molecular Biology and Evolution* online.

References

- 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.
- Baudat F, Imai Y, de Massy B. 2013. Meiotic recombination in mammals: localization and regulation. *Nat Rev Genet*. **14**:794–806.
- Bellani MA, Boateng KA, McLeod D, Camerini-Otero RD. 2010. The expression profile of the major mouse SPO11 isoforms indicates that SPO11beta introduces double strand breaks and suggests that SPO11alpha has an additional role in prophase in both spermatocytes and oocytes. *Mol Cell Biol*. **30**:4391–4403.
- Bergerat A, de Massy B, Gabelle D, Varoutas PC, Nicolas A, Forterre P. 1997. An atypical topoisomerase II from Archaea with implications for meiotic recombination [see comments]. *Nature*. **386**:414–417.
- Bermudez VP, Maniwa Y, Tappin I, Ozato K, Yokomori K, Hurwitz J. 2003. The alternative Ctf18-Dcc1-Ctf8-replication factor C complex required for sister chromatid cohesion loads proliferating cell nuclear antigen onto DNA. *Proc Natl Acad Sci U S A*. **100**:10237–10242.
- Blitzblau HG, Chan CS, Hochwagen A, Bell SP. 2012. Separation of DNA replication from the assembly of break-competent meiotic chromosomes. *PLoS Genet*. **8**:e1002643.
- Bloomfield G. 2016. Atypical ploidy cycles, Spo11, and the evolution of meiosis. *Semin Cell Dev Biol*. **54**:158–164.
- Bloomfield G. 2018. Spo11-independent meiosis in social amoebae. *Annu Rev Microbiol*. **72**:293–307.
- Borde V, Goldman AS, Lichten M. 2000. Direct coupling between meiotic DNA replication and recombination initiation. *Science*. **290**:806–809.
- Breuer C, Stacey NJ, West CE, Zhao Y, Chory J, Tsukaya H, Azumi Y, Maxwell A, Roberts K, Sugimoto-Shirasu K. 2007. BIN4, A novel component of the plant DNA topoisomerase VI complex, is required for endoreduplication in Arabidopsis. *Plant Cell*. **19**:3655–3668.
- Champoux JJ. 2001. DNA Topoisomerases: structure, function, and mechanism. *Annu Rev Biochem*. **70**:369–413.
- Chen SH, Chan NL, Hsieh TS. 2013. New mechanistic and functional insights into DNA topoisomerases. *Annu Rev Biochem*. **82**:139–170.
- Claeys Bouuaert C, Pu S, Wang J, Oger C, Daccache D, Xie W, Patel DJ, Keeney S. 2021. DNA-driven condensation assembles the meiotic DNA break machinery. *Nature*. **592**:144–149.
- Claeys Bouuaert C, Tischfield SE, Pu S, Mimitou EP, Arias-Palomo E, Berger JM, Keeney S. 2021. Structural and functional characterization of the Spo11 core complex. *Nat Struct Mol Biol*. **28**:92–102.
- Coelho SM, Peters AF, Muller D, Cock JM. 2020. Ectocarpus: an evo-devo model for the brown algae. *EvoDevo*. **11**:19.
- Corbett KD, Benedetti P, Berger JM. 2007. Holoenzyme assembly and ATP-mediated conformational dynamics of topoisomerase VI. *Nat Struct Mol Biol*. **14**:611–619.
- Corbett KD, Berger JM. 2003. Structure of the topoisomerase VI-B subunit: implications for type II topoisomerase mechanism and evolution. *EMBO J*. **22**:151–163.
- Corbett KD, Berger JM. 2005. Structural dissection of ATP turnover in the prototypical GHL ATPase TopoVI. *Structure*. **13**:873–882.
- Corbett KD, Schoeffler AJ, Thomsen ND, Berger JM. 2005. The structural basis for substrate specificity in DNA topoisomerase IV. *J Mol Biol*. **351**:545–561.
- Dutta R, Inouye M. 2000. GHKL, an emergent ATPase/kinase superfamily. *Trends Biochem Sci*. **25**:24–28.
- Forterre P, Gabelle D. 2009. Phylogenomics of DNA topoisomerases: their origin and putative roles in the emergence of modern organisms. *Nucleic Acids Res*. **37**:679–692.
- Forterre P, Gribaldo S, Gabelle D, Serre MC. 2007. Origin and evolution of DNA topoisomerases. *Biochimie*. **89**:427–446.
- Gabelle D, Krupovic M, Raymann K, Mayer C, Forterre P. 2014. DNA Topoisomerase VIII: a novel subfamily of type IIB topoisomerases encoded by free or integrated plasmids in Archaea and Bacteria. *Nucleic Acids Res*. **42**:8578–8591.
- Galbraith AM, Malone RE. 1992. Characterization of REC104, a gene required for early meiotic recombination in the yeast *Saccharomyces cerevisiae*. *Dev Genet*. **13**:392–402.
- Gellert M, Mizuuchi K, O'Dea MH, Nash HA. 1976. DNA Gyrase: an enzyme that introduces superhelical turns into DNA. *Proc Natl Acad Sci U S A*. **73**:3872–3876.
- Goodenough U, Heitman J. 2014. Origins of eukaryotic sexual reproduction. *Cold Spring Harb Perspect Biol*. **6**:a016154.
- Graille M, Cladiere L, Durand D, Lecoite F, Gabelle 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.
- Grey C, Espeut J, Ametsitsi R, Kumar R, Luksza M, Brun C, Verlhac MH, Suja JA, de Massy B. 2016. SKAP, an outer kinetochore protein, is required for mouse germ cell development. *Reproduction*. **151**:239–251.
- Hartung F, Angelis KJ, Meister A, Schubert I, Melzer M, Puchta H. 2002. An archaeobacterial topoisomerase homolog not present in other eukaryotes is indispensable for cell proliferation of plants. *Curr Biol*. **12**:1787–1791.
- Hartung F, Puchta H, (0 co-authors) 2001. Molecular characterization of homologues of both subunits A (SPO11) and B of the archaeobacterial topoisomerase 6 in plants. *Gene*. **271**:81–86.
- Hinman AW, Yeh HY, Roelens B, Yamaya K, Woglar A, Bourbon HG, Chi P, Villeneuve AM. 2021. *Caenorhabditis elegans* DSB-3 reveals conservation and divergence among protein complexes promoting meiotic double-strand breaks. *Proc Natl Acad Sci U S A*. **118**(33):e2109306118.
- Hunter N. 2015. Meiotic recombination: the essence of heredity. *Cold Spring Harb Perspect Biol*. **7**(12):a016618.
- Imai Y, Baudat F, Taillepiere 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.
- Jay A, Reitz D, Namekawa SH, Heyer WD. 2021. Cancer testis antigens and genomic instability: more than immunology. *DNA Repair (Amst)*. **108**:103214.

- Jolivet S, Vezon D, Froger N, Mercier R. 2006. Non conservation of the meiotic function of the Ski8/Rec103 homolog in *Arabidopsis*. *Genes Cells*. **11**:615–622.
- Kawasumi R, Abe T, Psakhye I, Miyata K, Hirota K, Branzei D. 2021. Vertebrate CTF18 and DDX11 essential function in cohesion is bypassed by preventing WAPL-mediated cohesin release. *Genes Dev*. **35**:1368–1382.
- Keeney S, Giroux CN, 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.
- Kirik V, Schrader A, Uhrig JF, Hulskamp M. 2007. MIDGET Unravels functions of the *Arabidopsis* topoisomerase VI complex in DNA endoreduplication, chromatin condensation, and transcriptional silencing. *Plant Cell*. **19**:3100–3110.
- Ku JC, Ronceret A, Golubovskaya I, Lee DH, Wang C, Timofejeva L, Kao YH, Gomez Angoa AK, Kremling K, Williams-Carrier R, et al. 2020. Dynamic localization of SPO11-1 and conformational changes of meiotic axial elements during recombination initiation of maize meiosis. *PLoS Genet*. **16**:e1007881.
- Kumar R, Bourbon HM, de Massy B. 2010. Functional conservation of Mei4 for meiotic DNA double-strand break formation from yeasts to mice. *Genes Dev*. **24**:1266–1280.
- Kumar R, Oliver C, Brun C, Juarez-Martinez AB, 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.
- Lange J, Pan J, Cole F, Thelen MP, Jasin M, Keeney S. 2011. ATM Controls meiotic double-strand-break formation. *Nature*. **479**:237–240.
- Malik SB, Pightling AW, Stefaniak LM, Schurko AM, Logsdon JM, Jr. 2007. An expanded inventory of conserved meiotic genes provides evidence for sex in *Trichomonas vaginalis*. *PLoS One* **3**:e2879.
- Malik SB, Ramesh MA, Hulstrand AM, Logsdon JM, Jr. 2007. Protist homologs of the meiotic Spo11 gene and topoisomerase VI reveal an evolutionary history of gene duplication and lineage-specific loss. *Mol Biol Evol* **24**:2827–2841.
- Mayer ML, Gygi SP, Abersold R, Hieter P. 2001. Identification of RFC(Ctf18p, Ctf8p, Dcc1p): an alternative RFC complex required for sister chromatid cohesion in *S. cerevisiae*. *Mol Cell*. **7**:959–970.
- McKie SJ, Desai PR, Seol Y, Allen AM, Maxwell A, Neuman KC. 2022. Topoisomerase VI is a chirally-selective, preferential DNA decatenase. *elife*. **11**:e67021.
- Méteignier LV, Lecampion C, Velay F, Vriet C, Dimnet L, Rougée M, Breuer C, Soubigou-Taconnat L, Sugimoto K, Barneche F, et al. 2022. Topoisomerase VI participates in an insulator-like function that prevents H3K9me2 spreading. *Proc Natl Acad Sci U S A*. **119**:e2001290119.
- Murakami H, Nurse P. (21324755 co-authors). 2001. Regulation of premeiotic S phase and recombination-related double-strand DNA breaks during meiosis in fission yeast. *Nat Genet* **28**:290–293.
- Neale MJ, Pan J, Keeney S. 2005. Endonucleolytic processing of covalent protein-linked DNA double-strand breaks. *Nature*. **436**:1053–1057.
- Nichols MD, DeAngelis K, Keck JL, Berger JM. 1999. Structure and function of an archaeal topoisomerase VI subunit with homology to the meiotic recombination factor Spo11. *EMBO J*. **18**:6177–6188.
- Nore A, Juarez-Martinez AB, Clément J, Brun C, Diagouraga B, Grey C, Bourbon HM, Kadlec J, Robert T, de Massy B. 2021. TOPOVIBL-REC114 interaction regulates meiotic DNA double-strand breaks. *Biorxiv*.
- Ramesh MA, Malik SB, Logsdon JM, Jr. 2005. A phylogenomic inventory of meiotic genes; evidence for sex in giardia and an early eukaryotic origin of meiosis. *Curr Biol* **15**:185–191.
- Ren B, Kuhn J, Meslet-Cladiere L, Briffotiaux J, Norais C, Lavigne R, Flament D, Ladenstein R, Myllykallio H. 2009. Structure and function of a novel endonuclease acting on branched DNA substrates. *EMBO J*. **28**:2479–2489.
- Rillo-Bohn R, Adilardi R, Mitros T, Avsaroglu B, Stevens L, Kohler S, Bayes J, Wang C, Lin S, Baskevitch KA, et al. 2021. Analysis of meiosis in *Pristionchus pacificus* reveals plasticity in homolog pairing and synapsis in the nematode lineage. *elife*. **10**:e70990.
- Robert T, Nore A, Brun C, Maffre C, Crimi B, Guichard V, Bourbon HM, de Massy B. 2016. The TopoVIB-like protein family is required for meiotic DNA double-strand break formation. *Science*. **351**:943–949.
- Robert T, Vrielynck N, Mezard C, de Massy B, Grelon M. 2016. A new light on the meiotic DSB catalytic complex. *Semin Cell Dev Biol*. **54**:165–176.
- Romanienko PJ, Camerini-Otero RD. 1999. Cloning, characterization, and localization of mouse and human SPO11. *Genomics*. **61**:156–169.
- Seitz KW, Lazar CS, Hinrichs KU, Teske AP, Baker BJ. 2016. Genomic reconstruction of a novel, deeply branched sediment archaeal phylum with pathways for acetogenesis and sulfur reduction. *ISME J*. **10**:1696–1705.
- Simkova K, Moreau F, Pawlak P, Vriet C, Baruah A, Alexandre C, Hennig L, Apel K, Laloi C. 2012. Integration of stress-related and reactive oxygen species-mediated signals by Topoisomerase VI in *Arabidopsis thaliana*. *Proc Natl Acad Sci U S A*. **109**:16360–16365.
- Spang A, Saw JH, Jorgensen SL, Zaremba-Niedzwiedzka K, Martijn J, Lind AE, van Eijk R, Schleper C, Guy L, Ettema TJG. 2015. Complex archaea that bridge the gap between prokaryotes and eukaryotes. *Nature*. **521**:173–179.
- Sprink T, Hartung F. 2014. The splicing fate of plant SPO11 genes. *Front Plant Sci*. **5**:214.
- Sugimoto-Shirasu K, Roberts GR, Stacey NJ, McCann MC, Maxwell A, Roberts K. 2005. RHL1 is an essential component of the plant DNA topoisomerase VI complex and is required for ploidy-dependent cell growth. *Proc Natl Acad Sci U S A*. **102**:18736–18741.
- Sugimoto-Shirasu K, Stacey NJ, Corsar J, Roberts K, McCann MC. 2002. DNA Topoisomerase VI is essential for endoreduplication in *Arabidopsis*. *Curr Biol*. **12**:1782–1786.
- Takahashi TS, Da Cunha V, Krupovic M, Mayer C, Forterre P, Gabelle D. 2020. Expanding the type IIB DNA topoisomerase family: identification of new topoisomerase and topoisomerase-like proteins in mobile genetic elements. *NAR Genom Bioinform* **2**:lqz021.
- Vos SM, Tretter EM, Schmidt BH, Berger JM. 2011. All tangled up: how cells direct, manage and exploit topoisomerase function. *Nat Rev Mol Cell Biol*. **12**:827–841.
- Vrielynck N, Chambon A, Vezon D, Pereira L, Chelysheva L, De Muyt A, Mezard C, Mayer C, Grelon M. 2016. A DNA topoisomerase VI-like complex initiates meiotic recombination. *Science*. **351**:939–943.
- Wade BO, Liu HW, Samora CP, Uhlmann F, Singleton MR. 2017. Structural studies of RFC(C)(tf18) reveal a novel chromatin recruitment role for Dcc1. *EMBO Rep*. **18**:558–568.
- Wang JC. 1971. Interaction between DNA and an *Escherichia coli* protein omega. *J Mol Biol*. **55**:523–533.
- Wang JC. 2002. Cellular roles of DNA topoisomerases: a molecular perspective. *Nat Rev Mol Cell Biol*. **3**:430–440.
- Wendorff TJ, Berger JM. 2018. Topoisomerase VI senses and exploits both DNA crossings and bends to facilitate strand passage. *elife*. **7**:e31724.
- Yin Y, Cheong H, Friedrichsen D, Zhao Y, Hu J, Mora-Garcia S, Chory J. (22133800 co-authors). 2002. A crucial role for the putative *Arabidopsis* topoisomerase VI in plant growth and development. *Proc Natl Acad Sci U S A* **99**:10191–10196.
- Zaremba-Niedzwiedzka K, Caceres EF, Saw JH, Backstrom D, Juzokaite L, Vancaester E, Seitz KW, Anantharaman K, Starnawski P, Kjeldsen KU, et al. 2017. Asgard archaea illuminate the origin of eukaryotic cellular complexity. *Nature*. **541**:353–358.