











# TopBP1 utilises a bipartite GINS binding mode to support genome replication

Received: 21 June 2023

Accepted: 7 February 2024

Published online: 27 February 2024

 Check for updates

Matthew Day<sup>1,2,10</sup>  , Bilal Tetik<sup>3,10</sup> , Milena Parlak<sup>3,10</sup>,  
Yasser Almeida-Hernández<sup>4,5</sup> , Markus Räschle<sup>6</sup> , Farnusch Kaschani<sup>7,8</sup>,  
Heike Siegert<sup>3</sup>, Anika Marko<sup>3</sup>, Elsa Sanchez-Garcia<sup>4,5</sup>, Markus Kaiser<sup>7,8</sup>,  
Isabel A. Barker<sup>2</sup>, Laurence H. Pearl<sup>2,9</sup>  , Antony W. Oliver<sup>2</sup>   &  
Dominik Boos<sup>3</sup>  

Activation of the replicative Mcm2-7 helicase by loading GINS and Cdc45 is crucial for replication origin firing, and as such for faithful genetic inheritance. Our biochemical and structural studies demonstrate that the helicase activator GINS interacts with TopBP1 through two separate binding surfaces, the first involving a stretch of highly conserved amino acids in the TopBP1-GINI region, the second a surface on TopBP1-BRCT4. The two surfaces bind to opposite ends of the A domain of the GINS subunit Psf1. Mutation analysis reveals that either surface is individually able to support TopBP1-GINS interaction, albeit with reduced affinity. Consistently, either surface is sufficient for replication origin firing in *Xenopus* egg extracts and becomes essential in the absence of the other. The TopBP1-GINS interaction appears sterically incompatible with simultaneous binding of DNA polymerase epsilon (Polε) to GINS when bound to Mcm2-7-Cdc45, although TopBP1-BRCT4 and the Polε subunit PolE2 show only partial competitiveness in binding to Psf1. Our TopBP1-GINS model improves the understanding of the recently characterised metazoan pre-loading complex. It further predicts the coordination of three molecular origin firing processes, DNA polymerase epsilon arrival, TopBP1 ejection and GINS integration into Mcm2-7-Cdc45.

Duplicating the genome accurately and exactly once each cell cycle lies at the heart of faithful genetic inheritance. The effective execution and regulation of DNA replication origin firing is pivotal to accurate genome duplication. Our understanding of how origin firing in eukaryotes

generates replisomes from their inactive precursors, pre-replicative complexes that comprise double hexamers of the Mcm2-7 helicase (MCM-DH), has increased. However, how the two helicase activators GINS and Cdc45 are loaded onto MCM-DHs remains less well understood.

<sup>1</sup>School of Biological and Behavioural Sciences, Blizard Institute, Queen Mary University of London, London E1 2AT, UK. <sup>2</sup>Cancer Research UK DNA Repair Enzymes Group, Genome Damage and Stability Centre, School of Life Sciences, University of Sussex, Falmer, Brighton BN1 9RQ, UK. <sup>3</sup>Molecular Genetics II, Center of Medical Biotechnology, University of Duisburg-Essen, Universitätsstraße 2-5, 45141 Essen, Germany. <sup>4</sup>Computational Bioengineering, Fakultät Bio- und Chemieingenieurwesen, Technical University Dortmund, Emil-Figge Str. 66, 44227 Dortmund, Germany. <sup>5</sup>Computational Biochemistry, Center of Medical Biotechnology, University of Duisburg-Essen, Universitätsstraße 2-5, 45141 Essen, Germany. <sup>6</sup>Molecular Genetics, Technical University Kaiserslautern, Paul-Ehrlich Straße 24, 67663 Kaiserslautern, Germany. <sup>7</sup>Analytics Core Facility Essen, Center of Medical Biotechnology, University of Duisburg-Essen, Universitätsstraße 2-5, 45141 Essen, Germany. <sup>8</sup>Chemical Biology, Center of Medical Biotechnology, University of Duisburg-Essen, Universitätsstraße 2-5, 45141 Essen, Germany. <sup>9</sup>Division of Structural Biology, Institute of Cancer Research, Chester Beatty Laboratories, 237 Fulham Road, London SW1E 6BT, UK. <sup>10</sup>These authors contributed equally: Matthew Day, Bilal Tetik, Milena Parlak. ✉ e-mail: [matthew.day@qmul.ac.uk](mailto:matthew.day@qmul.ac.uk); [Laurence.Pearl@sussex.ac.uk](mailto:Laurence.Pearl@sussex.ac.uk); [Antony.Oliver@sussex.ac.uk](mailto:Antony.Oliver@sussex.ac.uk); [dominik.boos@uni-due.de](mailto:dominik.boos@uni-due.de)

MCM-DHs form during origin licensing in the G1 phase of the cell cycle when the activity level of cyclin-dependent kinases (CDKs) are low<sup>1</sup>. A set of essential licensing factors sequentially loads two Mcm2-7 helicase hexamers onto origin DNA. Double-stranded DNA passes through the central channels of these helicase-inactive MCM-DHs<sup>2-4</sup>. In the following S phase, high CDK activity induces origin firing, converting each MCM-DH into two replisomes travelling in opposite directions. During this process, the MCM-DH separates, helicase activity is switched on, the dsDNA running inside the hexamers is unwound, and what will become the single lagging strand DNA template is excluded from each hexamer (Supplementary Fig. 1). The two active helicases pass each other, engaging with the DNA fork with their respective N-terminal faces first<sup>5,6</sup>.

Mcm2-7 helicase activation involves the loading of Cdc45 and GINS onto MCM-DHs to form the CMG (Cdc45-Mcm2-7-GINS) helicase<sup>7-10</sup>. CMG assembly occurs tightly coupled with partial Mcm2-7 double hexamer splitting and limited untwisting of the dsDNA inside the central CMG channels<sup>6,11,12</sup>. The subsequent assembly of mature replisomes involves several steps and many additional replisome components<sup>6,10</sup>.

In budding yeast, Cdc45 and GINS integration requires Dbf4-dependent kinase (DDK), S-CDK, the Sld3-Sld7 protein complex (Treslin-MTBP in higher eukaryotes), Dpb11 (TopBP1), Sld2 (potential human equivalents RecQL4 and DONSON) and DNA polymerase epsilon (DNA Pole)<sup>10,13,14</sup>. DDK binds MCM-DHs and phosphorylates the Mcm2-7 helicase, allowing Sld3-Sld7<sup>Treslin-MTBP</sup> association<sup>15-17</sup>. MCM-DH-bound Sld3-Sld7<sup>Treslin-MTBP</sup> recruits Cdc45 involving a direct Sld3<sup>Treslin</sup>-Cdc45 interaction<sup>14,18,19</sup>. Sld3-Sld7<sup>Treslin-MTBP</sup> is also involved in recruiting Dpb11<sup>TopBP1</sup>. For this, S-CDK phosphorylates Sld3<sup>Treslin</sup> on conserved CDK consensus sites, which can then bind to the N-terminal BRCT (breast cancer type 1 susceptibility) domains of Dpb11<sup>TopBP1</sup><sup>20-22</sup> (Fig. 1a(i)). The interaction between Sld3<sup>Treslin</sup> and Dpb11<sup>TopBP1</sup> serves to recruit the pre-loading complex (pre-LC) to MCM-DHs, comprising Dpb11<sup>TopBP1</sup>, GINS, Sld2<sup>RecQL4/DONSON</sup> and DNA Pole<sup>23</sup>. Pre-LC integrity requires a CDK phosphorylation-mediated interaction between phosphorylated Sld2<sup>RecQL4/DONSON</sup> and the C-terminal BRCT-pair of Dpb11<sup>TopBP1</sup> (Fig. 1a(ii)). Sld2<sup>RecQL4/DONSON</sup>, Dpb11<sup>TopBP1</sup> and Sld3-Sld7<sup>Treslin-MTBP</sup> then detach from the origin-bound Mcm2-7<sup>13,24</sup>, leaving GINS and Cdc45 behind to form the CMG helicase. DNA Pole associates with CMG in a dynamic fashion to form the CMGE complex (CMG-DNA Pole)<sup>25</sup>. The N-terminus of the Dpb2 subunit of DNA Pole and its vertebrate equivalent Pole2 bind to the Psf1 subunit of GINS<sup>26-28</sup>. It has been suggested that this interaction represents the essential activity of DNA Pole in replication origin firing<sup>25</sup>.

The precise molecular mechanism underpinning how the metazoan TopBP1<sup>Dpb11</sup>, Treslin<sup>Sld3</sup> and MTBP<sup>Sld7</sup> deliver GINS and Cdc45 to the Mcm2-7 helicase remains unclear. TopBP1<sup>Dpb11</sup> is a multi-BRCT domain scaffold protein containing nine BRCT domains<sup>29</sup>. The triple-BRCT module at the N-terminus of TopBP1<sup>Dpb11</sup> plus the BRCT4/5 pair share homology with budding yeast Dpb11 (Fig. 1a(i))<sup>30</sup>. As Dpb11, TopBP1 facilitates the coupling of origin firing to the cell cycle using its N-terminal BRCT modules<sup>21,22,31</sup>. TopBP1<sup>Dpb11</sup>-BRCT4/5 was, however, found to be non-essential to replication in *Xenopus* egg extracts<sup>31</sup>; perhaps surprisingly as the equivalent BRCT module of Dpb11<sup>TopBP1</sup> (BRCT3/4) is essential<sup>23</sup>. The minimal fragment of TopBP1<sup>Dpb11</sup> that supports replication includes BRCT0/1/2, BRCT3 (not conserved in Dpb11) and a short segment C-terminal to BRCT3 termed the GINS interaction or GINI region<sup>31,32</sup>. Yeast-two hybrid experiments showed that the GINI regions of TopBP1 and Dpb11, which have little sequence identity, both bind to GINS<sup>32</sup> (Fig. 1a(i)). As both TopBP1 and Dpb11 are limiting factors for origin firing<sup>33-35</sup>, it is likely that they are released and recycled to facilitate replication from multiple origins<sup>36</sup>. Both association and disassociation of Dpb11<sup>TopBP1</sup> are therefore important processes.

Recent evidence showed that TopBP1<sup>Dpb11</sup> and GINS are members of a metazoan pre-LC-like complex that also contains DNA Pole and a

new metazoan-specific origin firing factor, DONSON<sup>37-41</sup>. Pre-LC integrity appears to be important for metazoan origin firing because the interactions of DONSON with TopBP1<sup>Dpb11</sup> (via the essential TopBP1<sup>Dpb11</sup>-BRCT3 domain), GINS and Mcm3 are necessary for efficient origin firing. The TopBP1<sup>Dpb11</sup>-GINS interaction was not characterised in detail in these recent publications.

Here, we have used purified proteins to address how TopBP1<sup>Dpb11</sup> loads GINS onto MCM-DHs to activate the Mcm2-7 helicase during replication origin firing. Our structural, biochemical, and functional studies uncover the binding mode of GINS to TopBP1<sup>Dpb11</sup> and reveal a role for the TopBP1<sup>Dpb11</sup>-BRCT4 domain in replication, previously thought to be dispensable. Our data provide a model and framework for future investigation of CMG formation through GINS and Cdc45 loading by TopBP1<sup>Dpb11</sup> and the Treslin-MTBP<sup>Sld3-Sld7</sup> complex. Moreover, our study provides insight how TopBP1<sup>Dpb11</sup> acts at a molecular level. Such detailed understanding of TopBP1<sup>Dpb11</sup> is key to unravelling how this important scaffold protein can integrate its many different cellular roles in chromosome biology<sup>29</sup>.

## Results

### TopBP1 interacts with GINS in vivo and in vitro

As previous yeast two-hybrid experiments showed an interaction of the human and *Xenopus* TopBP1-GINI region with GINS (Psf1 and Psf3)<sup>32</sup>, we tested if a similar interaction could be detected in human cells. Three proximity biotinylation experiments using cells expressing APEX2-tagged TopBP1 resulted in a specific enrichment of Psf3 upon streptavidin pulldown and mass spectrometry, amongst a range of other known TopBP1 interactors (Supplementary Fig. 2a). Psf2 was also detected in one experiment. Immunoprecipitation from lysates of cells transfected with TopBP1-BRCT0-5 (aa, amino acids 1-766) showed a weak but specific signal for the GINS subunit Sld5 (Supplementary Fig. 2b, c). These experiments suggest that TopBP1 and GINS interact in human cells.

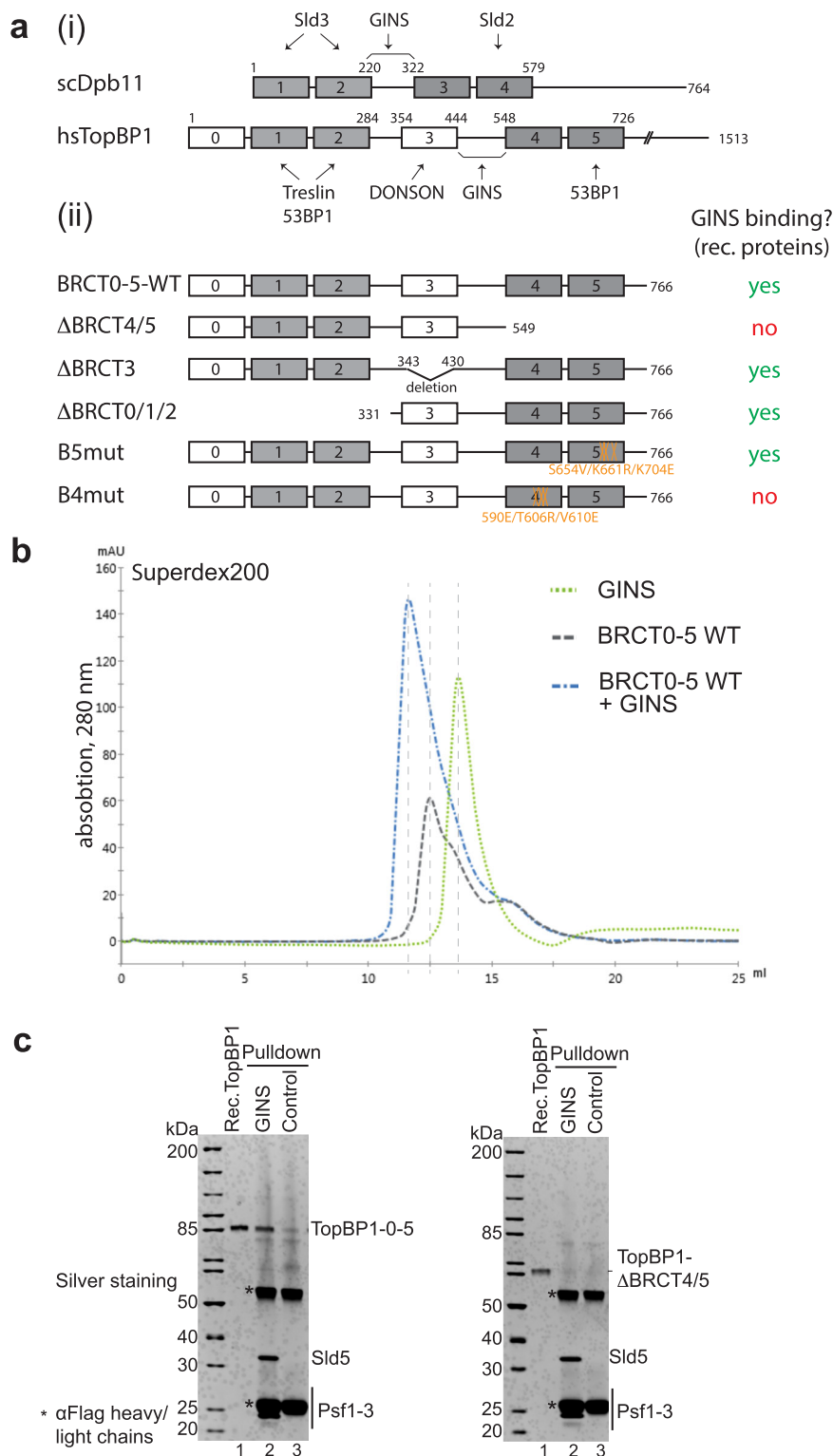
In addition, purified recombinant TopBP1-BRCT0-5 (TopBP1-0-5-WT) co-eluted with the GINS complex from a size exclusion column (Fig. 1b; Supplementary Fig. 3a). These results demonstrate that a TopBP1 fragment spanning the BRCT repeats 0-5 is sufficient to interact with GINS both in vivo and in vitro.

### Two binding motifs within TopBP1 contribute to GINS binding

We then used deletion and mutation analysis to narrow down the GINS binding region within the TopBP1-BRCT0-5 fragment. Removing the Treslin-binding BRCT0-2 module (TopBP1-ΔBRCT0/1/2) or BRCT3 (TopBP1-ΔBRCT3) from TopBP1 (Fig. 1a(ii)) did not affect binding (Supplementary Fig. 3b,c). In contrast, deleting BRCT4/5 downstream of the GINI region (TopBP1-ΔBRCT4/5) strongly impaired GINS interaction (Fig. 1c). We confirmed that the phospho-binding pocket of BRCT5 was not required for this interaction, using a triple point mutant that compromises its phospho-binding ability (B5mut)<sup>42</sup> (Fig. 1a(ii), Supplementary Fig. 3d,e). Our observations strongly suggested that a second GINS-binding motif, in addition to the GINI region, is present in the BRCT4/5 module. This was interesting, as no replication role for TopBP1-BRCT4/5 had been defined previously.

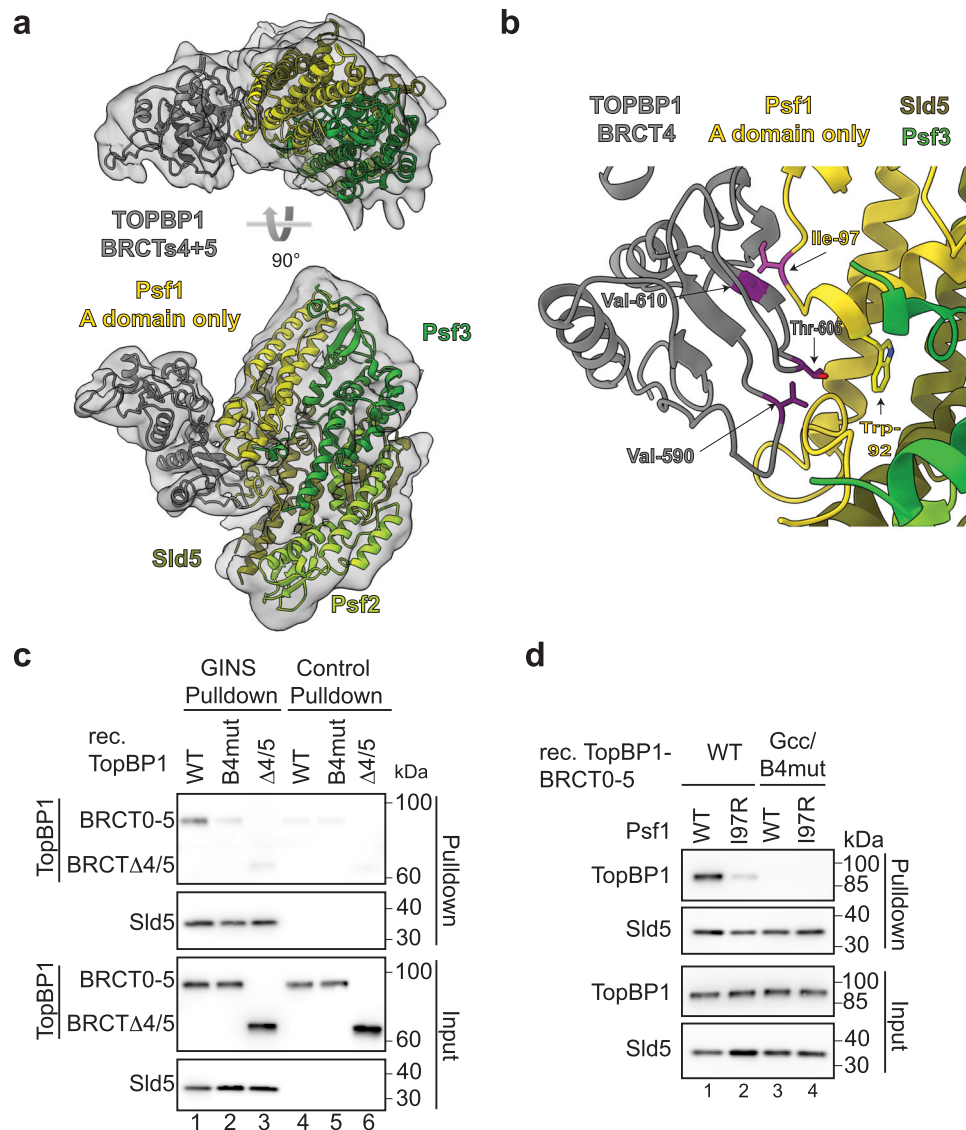
### Cryo-EM structure of TopBP1-BRCT4/5 in complex with GINS

To reveal how TopBP1 interacts with the GINS complex, peak fractions of TopBP1-BRCTΔ0/1/2 in complex with GINS eluting from the size exclusion column (Supplementary Fig. 3b) were used for cryo-EM. Movies were collected on an FEI Titan Krios microscope equipped with a Gatan K3 camera. 18,945 movies were processed using a combination of CryoSPARC<sup>43</sup> and RELION<sup>44</sup> (Supplementary Fig. 4, Supplementary Table 1) to produce a final 3-dimensional reconstruction, derived from 208,115 particles at 4.6 Å based on Fourier shell correlation (FSC) values. Density corresponding to the GINS tetramer could be readily identified in the resultant map, allowing facile placement of an X-ray



**Fig. 1 | GINS and TopBP1 interact.** **a (i)** Domain model of human (hs) TopBP1, its conservation with BRCT1/2 and 3/4 domains (grey boxes) of budding yeast Dpb11 (scDpb11). Arrows point to binding domains for interacting proteins. **(ii)** Overview of TopBP1 wild type (WT) and BRCT deletion ( $\Delta$ ) and point mutants. GINS binding capabilities are labelled in red and green. Numbers, amino acid position; capital letters indicate amino acid substitutions. Amino acid substitutions of point mutants in BRCT4 (B4mut) and 5 (B5mut) are coloured orange. **b** Elution profiles of size

exclusion chromatography (Superdex200) of individual recombinant TopBP1-BRCT0-5 protein (grey) and GINS tetramers (green), or of both upon incubation (blue). **c** Recombinant GINS immobilised on Flag-beads and control beads (Flag peptide-coupled beads) were used to pull down recombinant TopBP1-0-5-WT or TopBP1- $\Delta$ BRCT4/5 (compare a(ii)). SDS gels were silver stained. kDa specifies molecular weight marker bands. The experiment was done more than three times with similar results. Source data are provided as a Source data file.



**Fig. 2 | Protein structure of the GINS-TopBP1-BRCT4/5 complex.** **a** Structural model of the GINS-TopBP1-BRCT4/5 complex. Crystal structures of the GINS subunits (PDB:2E9X) (shades of green, yellow), and the central BRCT4/5 domain of TopBP1 (PDB:3UEN) (grey) were docked into the cryo-EM volume shown as transparent volume. **b** Zoom-in of the BRCT4-Psf1 interface. The residues in stick representation appear crucial for the interface. Mutations to break the interaction are coloured pink (Psf1-I97R) and purple (TopBP1-B4mut; compare Fig. 1a(ii)). **c** Pull-down of the indicated recombinant TopBP1-BRCT0-5-strep versions (see

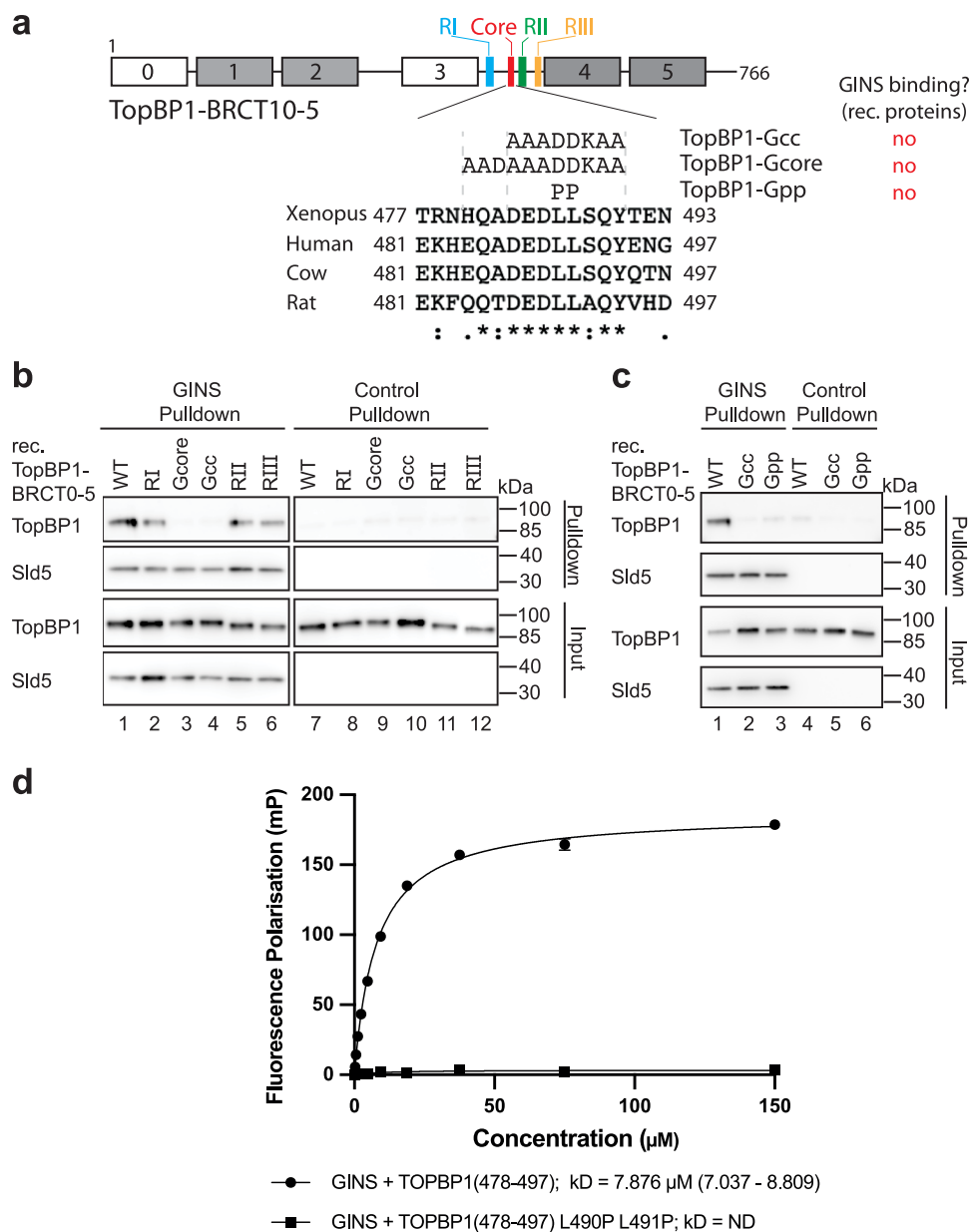
Fig. 1a(ii)) by Flag-beads-immobilised recombinant GINS or Flag peptide-coupled control beads. Analysis was done by immunoblotting. The experiment was done more than three times with similar results. **d** Pull-down of the indicated recombinant TopBP1-BRCT0-5-WT-strep using immobilised GINS-WT or GINS carrying a Ile97 to arginine mutation in Psf1. TopBP1-Gcc/B4mut-strep (Figs. 1a(ii) and 3a) was used as a non-GINS binding control. The experiment was done twice with similar results. Source data are provided as a Source data file.

crystal structure by rigid-body docking (PDB:2E9X) (Fig. 2a). As features of the individual GINS subunits were clearly identifiable, this also allowed unambiguous orientation of the pseudo-symmetric complex. The crystal structure appeared to account for all the density seen with the exception of a protrusion away from the main body of the GINS complex. This remaining density was too large to represent just the B domain of Psf1 (absent from the docked X-ray crystal structure), or the single BRCT3 of TopBP1, but was readily accounted for by docking of the X-ray crystal structure for TopBP1-BRCT4/5 (PDB:3UEN). The two BRCT domains were arranged such that BRCT4 interacted directly with the Psf1 subunit of GINS.

The GINS interacting region within the BRCT4 repeat of TopBP1 faces towards the A domain of Psf1 and the linker region connecting domains A and B (Fig. 2a, b), close to where the B domain docks to the CMG helicase (B domain not resolved). Note that some

ambiguity exists as to the precise molecular details due to the moderate resolution. In BRCT4, the interaction surface involves a region around beta strands  $\beta 3$  and  $\beta 4$ . This region diverts from the canonical BRCT fold in that it lacks an alpha helix ( $\alpha 2$ )<sup>44</sup> as if to create space for GINS binding (Supplementary Fig. 5a, b). The region shows high sequence conservation in BRCT4 equivalents across a wide range of metazoans (Supplementary Fig. 5c). The GINS interaction surface in BRCT4 consists of residues found at the edge of the central four-stranded  $\beta$ -sheet, with additional interactions made by residues in the loops connecting the strands, including amino acids in the variant region between strands  $\beta 3$  and  $\beta 4$ . Together, TopBP1-BRCT4 binds to the A domain of Psf1 and involves the linker region located towards the B domain. To our knowledge, this represents a novel mode for a BRCT domain-mediated interaction.





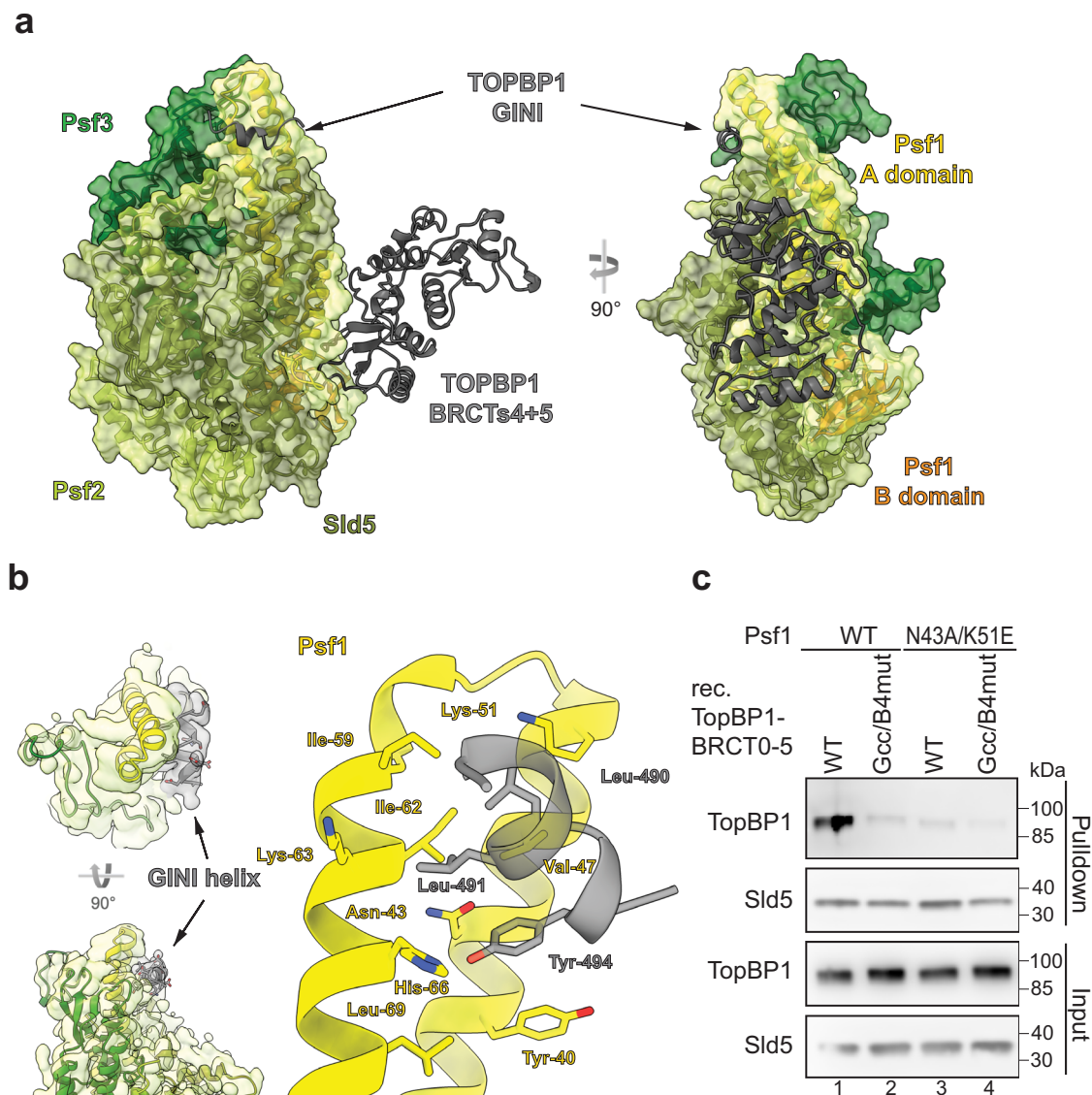
**Fig. 3 | The conserved core of the TopBP1-GINI domain is required for a stable interaction with GINS.** **a** Domain model and T-COFFEE alignment (one-letter aa code) of TopBP1 proteins from humans (Q92547) *Xenopus laevis* (Q7ZZY3), *Bos taurus* (A0A3Q1LWE4) and *Rattus norvegicus* (A0A816GFZ6). Coloured boxes indicate GINI region mutants in sub-regions I (blue), II (green) and III (orange), and in the GINI core helix (red) (details in Supplementary Fig. 7). Amino acid substitutions, names of GINI core region mutations and corresponding GINS binding capabilities are shown. ., \* and \* indicate low conservation, high conservation and identical

amino acids, respectively. **b/c** Pull-down of recombinant TopBP1-BRCT0-5-strep-WT or the indicated GINI region mutants (**a**) by bead-immobilised GINS or Flag peptide-coupled control beads. Analysis was done by immunoblotting. The experiments were done more than three times with similar results. Source data are provided as a Source data file. **d** Fluorescence polarisation measurements including dissociation constants (kD) calculations using recombinant GINS tetramer and fluorescently labelled TopBP1-GINI-core peptides (WT or Gpp (**a**)). For each peptide  $n = 3$  independent experiments and data are presented as mean values  $\pm$  SEM.

TopBP1 residues Val590, Thr606 and Val610 appeared to be involved in the interface, providing a series of direct hydrophobic interactions with the side chains of Trp92 and Ile97 in Psf1 (Fig. 2b). We mutated the three TopBP1 residues to glutamate, arginine and glutamate, respectively, and tested if the resulting protein (TopBP1-B4mut; Fig. 1a(ii)) co-purified with GINS. Here, TopBP1-B4mut was strongly compromised in its ability to bind GINS (Fig. 2c) but could still interact with phospho-53BP1 (Supplementary Fig. 6a). We also altered the opposing Psf1 interface by mutating Ile97 to an arginine (Psf1-I97R). Recombinant Psf1-I97R bound TopBP1 weaker than the wild type protein (Fig. 2d), but could still form GINS tetramers (Supplementary Fig. 6b).

### TopBP1 requires the conserved core of the GINI region for GINS binding

As the predicted interface between the TopBP1-GINI region and GINS<sup>32</sup> was not evident in our cryo-EM map we used additional interaction experiments to validate a role for this motif in GINS binding. The GINI region had been loosely defined as sitting between TopBP1-BRCT3 and BRCT4 (Fig. 1a(i)). Multiple sequence alignment across this region revealed a high degree of conservation between vertebrate species with the section from Glu484 to Tyr494 consisting of almost completely identical amino acids (Fig. 3a, Supplementary Fig. 7). We mutated this core, and other regions of conservation, to generate the mutants TopBP1-Gcc (GINI centre core), Gcore (GINI core) and RI-III



**Fig. 4 | The TopBP1-GINI domain interacts with Psf1-A domain.** **a** Structural model of TopBP1-BRCT3-5 complexed to the GINS tetramer predicted using AlphaFold2-Multimer. All chains are represented as cartoon. The transparent volume represents the surface of the GINS complex (not cryo-EM density). Of TopBP1, only the GINI helix (residues 485–496) and BRCT4/5 (residues 551–766), are shown. **b** Left: Cryo-EM volume showing extra helical density ascribed to the GINI

helix (grey). Right: a structural model based on this cryo-EM and an AlphaFold2-Multimer predicted structure. **c** Pull-down of recombinant TopBP1-BRCT0-5-WT-strep using immobilised GINS-WT or GINS carrying the indicated Psf1 mutations. TopBP1-BRCT0-5-Gcc/B4mut-strep served as a non-GINS binding control. Analysis was done by immunoblotting. The experiment was done twice with similar results. Source data are provided as a Source data file.

(regions I-III) (Fig. 3a, Supplementary Fig. 7). Pull-down experiments revealed that TopBP1-Gcc and Gcore, but not RI-III, were compromised in GINS interaction (Fig. 3b). Secondary structure prediction and AlphaFold2 models of TopBP1 suggested the presence of a short helix within this core region for some vertebrate species. By mutation of the central leucine-leucine motif to two consecutive prolines to break the helix (TopBP1-Gpp; Fig. 3a) we could show that the interaction with GINS was strongly compromised, with no effect on binding to 53BP1 (Fig. 3c, Supplementary Fig. 6a). To test if this helix constituted a GINS-interaction surface, we used fluorescence polarisation (FP). A fluorescently labelled GINI-core peptide (Ala478-Gly497) bound the GINS tetramer with a dissociation constant of  $K_d$  of  $-8 \mu\text{M}$  (Fig. 3d). A peptide harbouring the PP mutation failed to bind (Fig. 3d). We therefore concluded that the GINI region is a GINS interaction site in its own right, but with moderate binding affinity.

### The GINI-core helix binds Psf1 and 3

We next sought to identify the site in GINS that interacts with the TopBP1-GINI region. Structural prediction using AlphaFold2-Multimer suggested that the GINI-core helix interacted with the distal part of the Psf1 A domain (Fig. 4a, Supplementary Fig. 8). Moreover, blind molecular docking simulations with the GINI core region (487-DEDLLSQY-494) resulted in the top-10 solutions clustering in the region of the AlphaFold-predicted GINI binding site (Supplementary Fig. 9).

A BS<sup>3</sup>-cross-linking mass spectrometry experiment with purified recombinant TopBP1-GINS also confirmed the predicted proximity of the TopBP1-GINI core helix region (residues 487–494) and the Psf1-A domain. Several GINI-core-proximal residues cross-linked with Psf1-Lys63 and Thr61 (Supplementary Fig. 10a, Supplementary Data 1); both cross-linked Psf1 residues are situated at the distal end of the Psf1 A domain (Supplementary Fig. 10b, c). The GINI-core-proximal residues

also cross-linked with lysines 74 and 80 of Psf3; each Psf3 residue also being in close proximity to the predicted site of interaction. Given that the GINI-core helix itself contains no lysine residues, our data are consistent with the hypothesis that this helix binds to both the distal region of the Psf1 A domain and the adjacent region of Psf3 (Supplementary Fig. 10 provides a detailed description).

With all this information in hand, we could improve the resolution of our cryo-EM data, using an alternative GINS-expression construct, which omits the Psf1-B domain, and served to improve detail around the predicted interaction region (Supplementary Fig. 11). The resultant map showed a volume of extra density that allowed a short helical element to be incorporated into our model, consistent with the AlphaFold2-Multimer predictions (Fig. 4b). The hydrophobic side of the amphipathic GINI-core helix can be seen to pack across two helices of the Psf1-A domain, making interactions with the side chains of Leu490 and Leu491 of the GINI region and Val47, Lys51, Ile59, Ile62, Lys63 and His66 of Psf1. In addition, the side chain of Tyr494 in the GINI region is inserted into a small pocket on the surface of Psf1 formed by Tyr40, Asn43 and Leu69 with the hydroxyl group picking up a hydrogen bond with the side chain of His66.

For validation, we mutated Asn43 and Lys51, to Ala and Glu, respectively. Asn43 forms part of the pocket accepting Tyr494 from the GINI region, whilst Lys51 is positioned such that it could form a salt bridge with either Asp487 or Glu495 and part of the hydrophobic groove that accommodates the lucines of the GINI-core helix. Psf1-N43A/K51E readily formed the GINS tetramer (Supplementary Fig. 6b) but was severely compromised in its ability to interact with TopBP1-BRCTO-5-WT (Fig. 4c), again consistent with binding of the GINI region to the distal end of the Psf1-A domain and the adjacent regions in Psf3.

### Simultaneous inactivation of the GINI core region and BRCT4 is required to suppress replication origin firing

Our pulldown interaction studies with GINS and TopBP1-BRCTO-5 carried out at 150 mM sodium chloride suggested that the GINS interacting regions GINI and BRCT4 of TopBP1 are individually required for GINS interaction (Figs. 2c and 3c). However, pulldowns in sodium acetate revealed specific interactions of the individual interaction regions, albeit at reduced levels compared to TopBP1-WT (Fig. 5a, Supplementary Fig. 12a). Here, the GINI region bound to GINS more efficiently than BRCT4. Estimation of binding free energies suggested that the affinities are lower when the surfaces interact individually rather than simultaneously (Fig. 5b). Together, this suggests a model in which both interactions have relatively moderate binding affinities, but cooperate to form a composite interaction surface with high avidity.

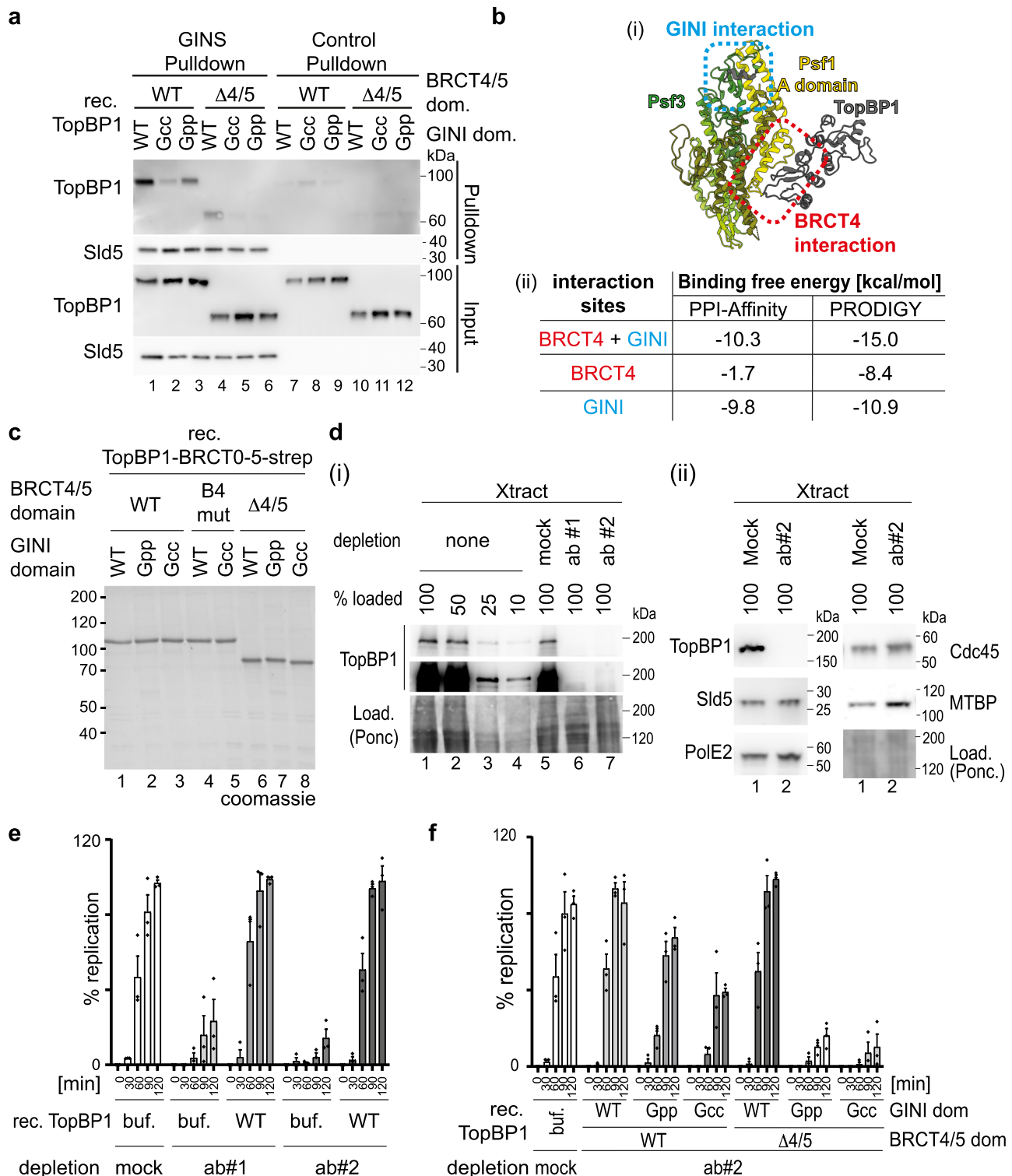
We next asked how the TopBP1-GINI and BRCT4 regions contribute to genome replication, by testing if mutational inactivation of GINI and BRCT4 (Fig. 5c) could suppress replication in *Xenopus* egg extracts. Immunodepleting TopBP1 with two independent antibodies (anti-TopBP1-#1 and #2) effectively removed the endogenous TopBP1 (Fig. 5d(i)), but neither GINS, MTBP, DNA Pole nor Cdc45 (Fig. 5d(ii)), from the extracts. TopBP1 depletion strongly reduced incorporation of radioactive  $\alpha$ -<sup>32</sup>P-dCTP when compared to a mock IgG depletion control (Fig. 5e). Adding recombinant TopBP1-BRCTO-5-WT effectively rescued replication in depleted extracts (Fig. 5e), whereas addition of either TopBP1-BRCTO-5-Gcc or TopBP1-BRCTO-5-Gpp (BRCT4 intact) led to moderately reduced nucleotide incorporation by 21% and 55% (120 min time point), respectively (Fig. 5f, Supplementary Fig. 13a), consistent with the observed reduced affinity of BRCT4 for GINS in the absence of a functional GINI region (Fig. 5a, lanes 1–3). TopBP1- $\Delta$ BRCT4/5 and TopBP1-B4mut (GINI domain intact) were not detectably compromised in their abilities to support replication (Fig. 5f; Supplementary Fig. 13b), as observed before<sup>31,32</sup>, consistent with the GINI region binding GINS stronger than BRCT4 (Fig. 5a, lanes 2, 4; Supplementary Fig. 12a). Combination of BRCT4/5 deletion or

mutation (B4mut) with either TopBP1-BRCTO-5-Gcc or -Gpp decreased replication to nearly background levels (Fig. 5f, Supplementary Fig. 13), consistent with these double mutants eliminating GINS binding (Fig. 5a, lanes 5, 6; Supplementary Fig. 12a). These results indicate that both TopBP1 binding regions, GINI and BRCT4, are individually capable of supporting replication, whilst simultaneous inactivation of the GINS binding regions strongly suppresses replication.

We next tested if the TopBP1-GINS interaction is required for replication origin firing. We isolated chromatin from *Xenopus* egg extracts upon TopBP1 immunodepletion and add-back of TopBP1-BRCTO-5-WT or mutants, 75 min after starting replication (a time point when genome replication has progressed significantly). To increase signals, we added aphidicolin, which prevents termination, allowing replisomes to accumulate on chromatin. Mass spectrometry using the CHROMASS protocol<sup>45</sup> and immunoblotting showed chromatin binding of Mcm2-7 subunits in TopBP1-depleted egg extract (Fig. 6a, c, d, Supplementary Fig. 14, Supplementary Data 2). Addition of geminin, an inhibitor of origin licensing<sup>46</sup>, prevented Mcm2-7 binding (Fig. 6a, c, d). Together, this confirmed that origin licensing is independent of TopBP1. Several replisome markers, including GINS and Cdc45, were detected on chromatin in TopBP1-depleted extracts containing TopBP1-WT, but were significantly reduced when the TopBP1-Gcc- $\Delta$ BRCT4/5 mutant or buffer were added (Fig. 6a, b, d). Thus, the GINS binding sites in TopBP1 are required for origin firing (replisome formation), more precisely for a step between MCM-DH formation and CMG formation. Chromatin samples upon complementation with the single mutant TopBP1-Gcc increased the levels of individual replisome markers weakly but detectably over geminin controls (Pole2 in Fig. 6d and Cdc45 in Supplementary Fig. 12b). The TopBP1- $\Delta$ BRCT4/5 mutant supported strong replisome formation with only mild effect on Sld5, PCNA, Pole2 and Cdc45 compared to WT (Fig. 6d, lanes 2, 5; Supplementary Fig. 12b). Together, this analysis showed that both binding sites are individually capable of supporting origin firing. They do so to different degrees, consistent with the GINI domain mediating stronger GINS interaction than BRCT4. Together, we conclude that both GINS binding sites in TopBP1 act in concert to assemble the CMG helicase. We noticed a surprisingly strong origin firing defect with the TopBP1-Gcc mutant, given that it supports replication to a significant level. It is feasible that compensatory fork acceleration partially compensates for decreased replisome numbers in the Gcc mutant. Our observation that deleting BRCT4 has only minor effects on replisome formation is also consistent with the scenario that the BRCT4-GINS interaction has a role downstream of origin firing that aggravates the DNA synthesis defects (Fig. 5f, Fig. S13) of GINI site mutants.

### Mutually exclusive binding of TopBP1 and DNA Pole to GINS in the context of CMG

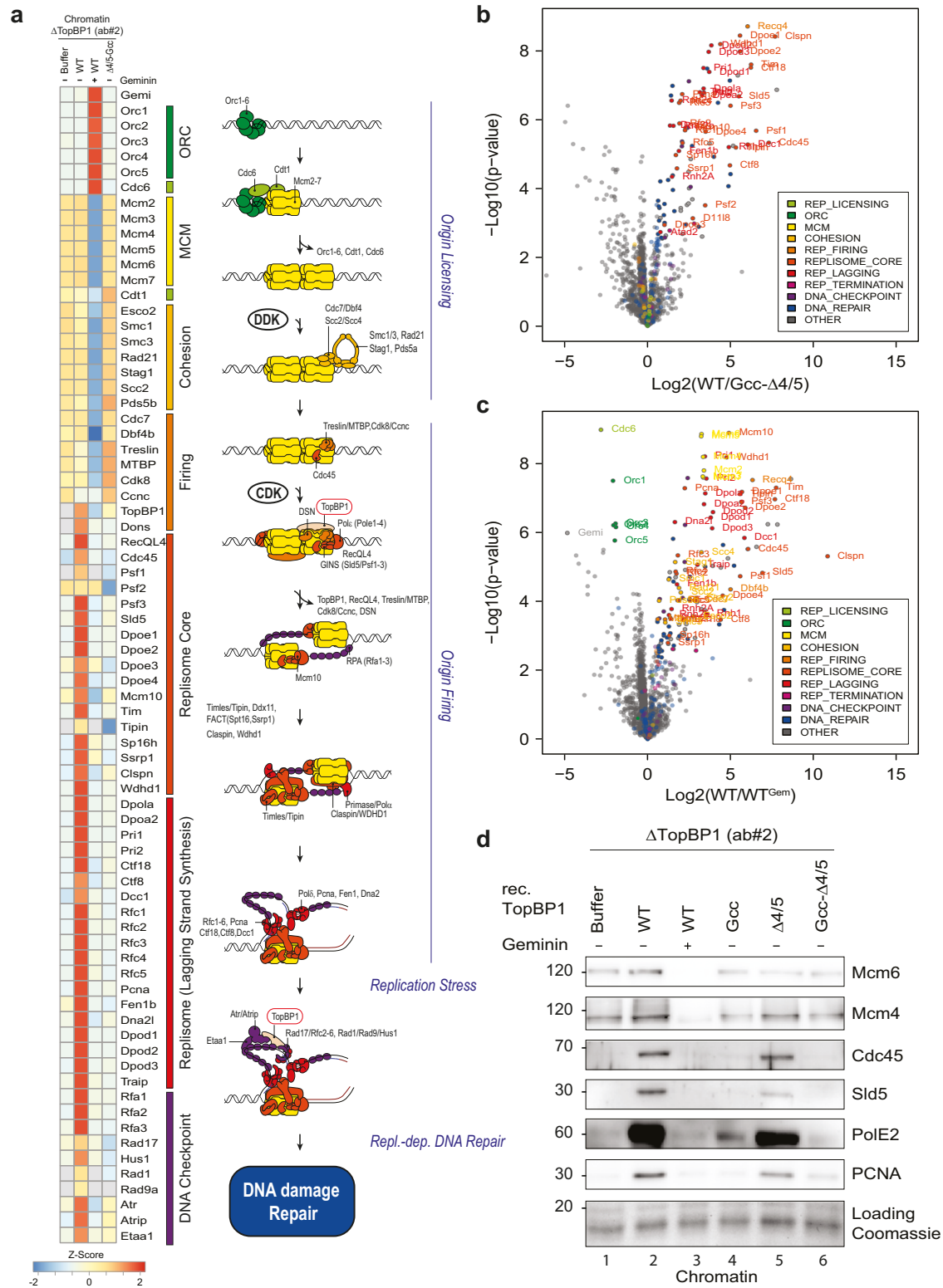
We next explored the relationship between the interactions of TopBP1 and DNA Pole with GINS. These proteins are part of the recently characterised metazoan pre-LC<sup>37–40</sup>. Analysis of pre-LC suggests that GINS may simultaneously bind DNA Pole and TopBP1. When GINS is complexed with DONSON in the pre-LC context, binding sites in Psf1 for the TopBP1-GINI and BRCT4 regions are accessible, as suggested by published AlphaFold multimer models<sup>37</sup>. These pre-LC models are less informative on how DNA Pole affects the TopBP1-GINS interaction. The N-terminal region of Pole2 (Pole2-N; aa 1–75) binds GINS mainly at the Psf1-B domain, with minor interactions with residues in the A-domain that partly overlap with the TopBP1-BRCT4 interaction surface (Fig. 7a–c). To test if the Psf1-B domain is required for TopBP1 interaction, we did pulldowns with GINS lacking the Psf1-B domain (Psf1- $\Delta$ B). These indicated that the Psf1-B domain is dispensable for interaction with TopBP1-WT (Supplementary Fig. 15a(i)). This could be due to a residual affinity of BRCT4 for Psf1- $\Delta$ B or due to the GINI-Psf1 interaction. We therefore analysed specifically the interaction of GINS with BRCT4 using the TopBP1-Gcc mutant. TopBP1-Gcc bound to GINS



**Fig. 5 | Both GINS binding surfaces in TopBP1 cooperate in supporting genome replication.** **a** Immobilized recombinant Flag-GINS or Flag peptide-coupled beads were used to pull down the indicated recombinant TopBP1-BRCT0-5-strep versions (Figs. 1a and 3a) using 100 mM NaAc in the binding buffer. The pulldown was analysed by immunoblotting. The experiment was done twice with similar results. **b** (i) shows the AlphaFold2-Multimer model of the GINS-TopBP1 complex that was used to estimate the binding free energy using PPI-Affinity and PRODIGY shown in (ii). Estimations were done for the binding interfaces involving TopBP1-GINI (blue) and TopBP1-BRCT4 (red), both together and alone. **c** Coomassie-stained SDS PAGE gel of the indicated recombinant TopBP1-BRCT0-5 proteins (Figs. 1a(ii) and 3a). **d** The indicated relative amounts (100%–10%) of immunodepleted and non-

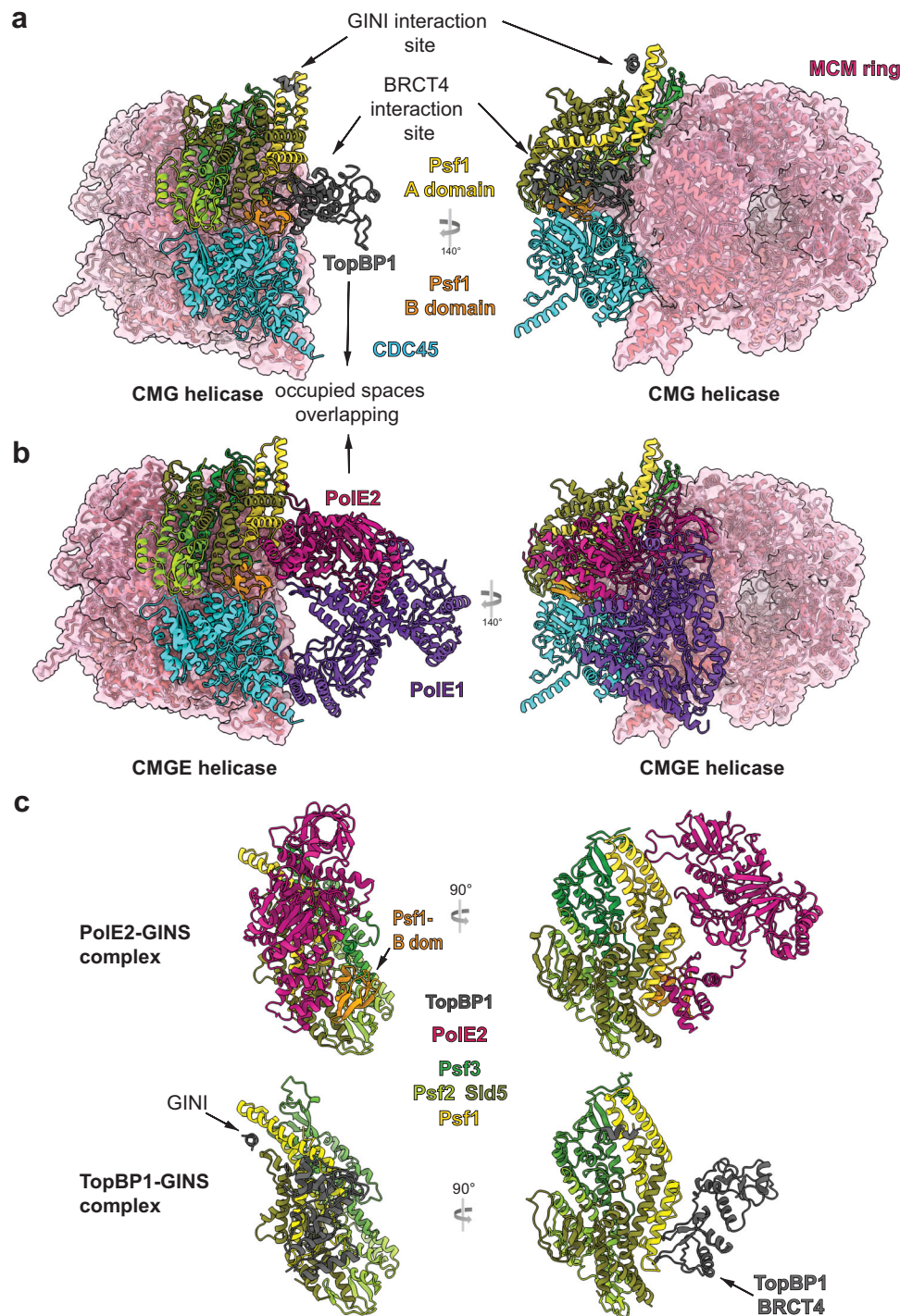
depleted *Xenopus* egg extracts (Xtract) were analysed by immunoblotting using anti-TopBP1 #2 and other indicated antibodies. TopBP1 antibodies #1 and #2, or unspecific IgG (mock) were used for immunodepletions. Ponc, ponceau staining. Error bars, SEM,  $n = 3$ . The experiments were done more than three times (i) or twice (ii) with similar results. **e, f** Replication analyses (radioactive nucleotide incorporation) using *Xenopus* egg extracts immunodepleted with antibodies anti-TopBP1#1, #2 or unspecific IgG (mock). Buffer (buf), recombinant wild-type (WT) or mutant TopBP1-BRCT0-5-strep (c) were added. Diamonds, Individual data points of  $n = 3$  independent experiments, Error bars, SEM. Source data are provided as a Source data file.





**Fig. 6 | GINS binding to TopBP1 is required for the origin firing step of replication.** **a–c** Heat map showing the intensity of chromatin-bound proteins recovered from TopBP1-depleted interphase *Xenopus* egg extracts supplemented wild type or mutant TopBP1 determined by quantitative mass spectrometry. Chromatin was isolated 75 min after sperm addition. All extracts contained 50 mg/ml aphidicolin and, where indicated, 2.25 mM geminin. The heat map shows  $\text{log}_2$  transformed label-free quantification (LFQ) intensities z-score normalized across rows and mean averaged over four independent replicates ( $n = 4$ ). For the intensities of all proteins see Supplementary Data 2. **b, c** Volcano plots highlighting a selection of

differentially detected proteins in samples described in (a).  $\text{Log}_2$  transformed LFQ intensities from four replicates each were analysed by a two-sided Student's t-test. Significantly changed proteins with an Benjamini–Hochberg adjusted  $p$ -value of equal or less than 0.05 are reported in Supplementary Data 2. For clarity, only proteins involved in DNA replication are labelled (all significantly enriched proteins are labelled in Supplementary Fig. 14). **d** Experiment as in (a) was analysed by chromatin isolation and immunoblotting with the indicated antibodies. Coomassie staining of the gel part containing histones served as a loading control.



**Fig. 7 | Exclusive binding of TopBP1 and DNA polymerase epsilon to GINS in the CMG context.** **a**, **b** Structure of human CMG (PDB 7PFO), in panel (a) with our GINS-TopBP1 superposed, in panel (b) with DNA Pol $\epsilon$  (CMGE; PDB 7PFO) both shown from the same angle, demonstrating accessibility of the TopBP1 binding sites in

CMG, and strong overlap between BRCT4 and PoIE2 occupied spaces. The pink volume constitutes surface representation. **c** In-detail views of GINS-TopBP1 interaction sites from the cryo-EM model in this manuscript (8OK2) alongside the GINS-PoIE2 interaction sites seen in the larger CMGE structure (7PFO).

partially dependently on the Psfl-B domain, whereas the B domain had no influence on TopBP1-WT and B4mut binding (Supplementary Fig. 15a(ii)). The moderate effect of the B domain deletion on BRCT4 binding is consistent with the partial overlap of the BRCT4 and PoIE2 binding sites in Psfl. Consistently, TopBP1-Gcc, but neither TopBP1-WT nor B4mut, was partially competed off of immobilised GINS by ten times molar excess of PoIE2-N (Supplementary Fig. 15b(ii)). TopBP1-WT was largely resistant to PoIE2 competition also in stringent pulldown

conditions (Supplementary Fig. 15b(i)). Together, in pre-LC, TopBP1 likely binds GINS through the GINI region, potentially also involving a residual affinity of BRCT4 for GINS despite the presence of DNA Pol $\epsilon$ . TopBP1 and GINS are likely further stabilised in pre-LCs via their interactions with DONSON<sup>37,38,41</sup>.

Next, we analysed the CMGE helicase. Both TopBP1 binding sites in GINS are accessible in the CMGE helicase, showing that TopBP1 could bind in the CMG context (Fig. 7a, b)<sup>27</sup>. However, the binding of DNA

Pole and TopBP1 seem mutually exclusive. Superposition of CMGE and CMG-TopBP1 revealed potential major steric clashes between TopBP1 and Pole2 around the BRCT4 interaction region (Fig. 7a–c). Based on this steric exclusion and the observed direct competition between Pole2 and BRCT4, we propose that TopBP1-BRCT4 disengages before DNA Pole associates to form CMGE. Because both TopBP1-GINS binding surfaces must cooperate for a strong GINS interaction BRCT4 disengagement should significantly weaken the TopBP1-GINS complex. Subsequent full dissociation of TopBP1 from CMGE may additionally require mechanisms that dislodge DONSON and Treslin-MTBP from CMGE.

The transition of GINS from TopBP1-bound to a stable CMG requires reconfiguration of the Psf1-B domain<sup>47</sup>. In CMG, the Psf1-B domain adopts a discrete conformation tightly packing against Cdc45 (Fig. 7a)<sup>27</sup>. In contrast, the B domain appears flexible in the GINS tetramer and in the TopBP1-GINS complex, because our cryo-EM data did not detect the B domain at the same map contour level as the rest of GINS (Supplementary Fig. 16a, b). A previous crystallization study of the GINS tetramer also suggested Psf1-B domain flexibility<sup>48</sup>. Thus, integration of GINS in the CMG appears to order the B domain. In this conformation, the B domain binds to the N-terminus of Pole2 once DNA Pole associates to form CMGE<sup>27</sup>.

Together, we envision that the Psf1-B domain reconfigures into a well-ordered state upon CMG assembly. TopBP1-BRCT4 must then dissociate to exchange with DNA Pole for CMGE formation, which likely helps TopBP1 dissociate. Once CMG has formed, DNA Pole seems to dynamically associate with CMG<sup>25</sup>. These observations suggest that the three processes (1) handover of GINS from TopBP1 to MCM-DHs, (2) DNA Pole association and (3) TopBP1 dissociation may be mechanistically intertwined.

## Discussion

Our data provide an understanding of the interaction between TopBP1<sup>Dpb11</sup> and GINS at the molecular level, and how this supports replication origin firing. In the past decade, insight into origin firing has largely come from studies focussed on understanding how the yeast Mcm2-7 helicase operates as a molecular machine to build replication forks<sup>6,10–12,14,16,17,19,49–52</sup>. Most recently, the aspect of how GINS and Cdc45 are loaded to activate the helicase through formation of the CMG complex has been addressed in more detail by describing the metazoan pre-LC<sup>37–41</sup>. Our studies here reveal an important aspect of pre-LC, the essential role of the TopBP1<sup>Dpb11</sup>-GINS interaction (Supplementary Fig. 17).

Formation of the CMG complex serves as a key regulatory step of origin firing. In metazoa and yeast, CMG formation involves the factors Treslin/Sld3, MTBP/Sld7, RecQL4/DONSON/Sld2, TopBP1/Dpb11 and the Mcm2-7 helicase itself. Apart from coupling origin firing to the S phase via S-CDK and DDK, these signal integrators are targeted by the DNA damage checkpoint and other regulatory pathways to mediate DNA damage-dependent origin firing control, execution of the temporal replication programme and fine-tuning of origin firing<sup>53–67</sup>. Such regulation helps avoid replication stress that is key to structural and numerical chromosomal instability in cancer cells<sup>68</sup>.

TopBP1<sup>Dpb11</sup> is a scaffold protein. To mediate S phase coupling of origin firing, the N-terminal triple BRCT0-2 module of TopBP1<sup>Dpb11</sup> binds Treslin<sup>Sld3</sup>, in a strictly phospho-dependant manner, via two CDK sites that are also conserved in budding yeast Sld3<sup>Treslin21,22</sup>. Despite this level of conservation, some aspects of origin firing in vertebrates are likely to have changed mechanistically when compared to the yeast paradigm: (1) DDK strengthens the interaction between TopBP1<sup>Dpb11</sup> and the Treslin-MTBP<sup>Sld3-Sld7</sup> complex, albeit via an unclarified mechanism<sup>69</sup>, whereas a similar role for DDK in yeast has not been reported. (2) an N-terminal fragment that excludes the BRCT4/5 module of TopBP1<sup>Dpb11</sup> effectively supports replication in *Xenopus* egg extracts<sup>31,32</sup>, whereas the equivalent BRCT3/4 module of Dpb11<sup>TopBP1</sup> is

essential, through CDK-mediated binding to phospho-Sld2<sup>RecQL4/DONSON</sup><sup>23</sup>. The metazoan-specific DONSON protein may replace some roles of Sld2<sup>37–40</sup>. These mechanistic differences exemplify evolutionary flexibility in S-phase coupling, supported by findings in the *C. elegans* and fission yeast model organisms where the contributions of Treslin<sup>Sld3</sup> and RecQL4<sup>Sld2</sup> equivalents have changed during evolution<sup>70,71</sup>. Such flexibility may be supported by the partially redundant roles of S phase kinases (S-CDK and DDK) and their substrates (Treslin<sup>Sld3</sup>, RecQL4<sup>Sld2</sup>, Mcm2-7).

TopBP1<sup>Dpb11</sup> supports CMG formation via an interaction with the GINS complex. In budding yeast, Dpb11<sup>TopBP1</sup> and GINS are recruited to MCM-DHs as part of a pre-LC comprising Dpb11<sup>TopBP1</sup>, GINS, DNA Pole and Sld2<sup>RecQL4/DONSON</sup><sup>72</sup>. Dpb11<sup>TopBP1</sup> appears to dock the pre-LC onto Sld3<sup>Treslin-Sld7</sup><sup>MTBP</sup>-Cdc45 via the CDK-mediated Dpb11<sup>TopBP1</sup>-Sld3<sup>Treslin</sup> interaction; DDK is required for initial binding of Sld3<sup>Treslin-Sld7</sup><sup>MTBP</sup>-Cdc45 onto MCM-DHs<sup>14,19</sup>. In metazoans, RecQL4<sup>Sld2</sup> is not part of the stable protein complex defined as pre-LC, which instead contains DONSON. DONSON is a scaffold protein with a series of interaction sites for TopBP1<sup>Dpb11</sup>, GINS, DNA Pole and Mcm3 (Supplementary Fig. 17a)<sup>37–40</sup>. Our described TopBP1<sup>Dpb11</sup>-GINS interaction probably occurs in the context of pre-LCs.

The GINI region of TopBP1<sup>Dpb11</sup> that had been broadly defined as the region between BRCT3 and BRCT4, bound to Psf1 and 3 in Y2H experiments<sup>32</sup>. We show here that a short alpha-helix found at the core of the conserved GINI region (GINI-core helix) forms an essential part of the binding surface that is necessary and sufficient for detectable interaction with GINS. This helix cooperates with a second binding site situated in BRCT4 to form a composite high-affinity GINS-interaction surface; involving a hereinto unknown mode of BRCT-interaction that relies on a specific type of BRCT fold found in TopBP1-BRCT4, which lacks helix 2. Notably, our in vitro experiments show that each individual GINS interaction site is capable of binding GINS, albeit with reduced efficiency compared to the complete composite interaction surface. This was consistent with the two GINS binding sites being able to support significant levels of replication in *Xenopus* egg extracts, which is in line with an earlier study focussing on the GINI region<sup>32</sup>. These results indicate that even a low affinity TopBP1<sup>Dpb11</sup>-GINS interaction is sufficient to allow origin firing. In vivo, contacts between other pre-LC components may compensate for a weak TopBP1<sup>Dpb11</sup>-GINS interaction. A high degree of inter-dependence between association of the components was also observed for the yeast pre-LC complex<sup>72</sup>. Simultaneous mutation of both GINS binding domains in TopBP1<sup>Dpb11</sup> suppressed origin firing. This suggests that neither interaction site has a unique molecular role beyond GINS binding. The simplest scenario is that their common essential molecular role is to bind GINS chaperoning it onto MCM-DHs.

An alternative interpretation of our TopBP1<sup>Dpb11</sup> mutant analysis seems feasible: TopBP1-Δ4/5 and B4mut only resulted in no DNA synthesis defect and a weak origin firing defect. This scenario is consistent with the possibility that the BRCT4-GINS interaction plays an unidentified role in replication, for example in elongation, which could contribute to the observed effect that BRCT4 inactivating mutants aggravate the defects of GINI site mutants in DNA replication assays (nucleotide incorporation).

Consistent with published Y2H experiments<sup>32</sup>, our structural data indicate that the GINI core helix region interacts with both Psf1 and Psf3, at a position close to the distal end of the Psf1 A domain, whereas BRCT4 interacts with both the base of the Psf1-A domain and the linker that connects it to the B domain. Notably, the two receiving interfaces in GINS are fully accessible both in the free GINS tetramer and in the CMG context, consistent with a model where TopBP1<sup>Dpb11</sup> recruits GINS using both motifs (Supplementary Fig. 17b). TopBP1<sup>Dpb11</sup> may simultaneously bind phosphorylated Treslin<sup>Sld3</sup>-MTBP<sup>Sld7</sup> via its N-terminal BRCT0-2 module<sup>21,22,73</sup>, with Treslin<sup>Sld3</sup> bringing Cdc45 along through its Sld3-Treslin homology domain<sup>18,74</sup>. One open question in the field



had been the biological role of the BRCT3 domain in TopBP1, which has no equivalent in Dpb11. We now know that an essential role of BRCT3 is the binding to DONSON<sup>37,38</sup>.

After delivery of GINS and Cdc45 to form CMG, the subsequent release of pre-LC components including and of Treslin<sup>Sld3</sup>-MTBP<sup>Sld7</sup> is likely to be important. Whether DNA Pole is transferred from pre-LC to CMG to form CMGE during pre-LC dissociation is an interesting aspect. Release may help recycle these limiting factors to allow their use to fire late or dormant origins<sup>35</sup>. Moreover, their release from CMG may be essential for converting CMGs into replisomes. If release is essential, these firing factors constitute a firing-inhibitory intermediate as long as they are CMG-bound, and their controlled dissociation could be part of regulating origin firing. Dissociation could be triggered by the Mcm10 and RecQL4 firing factors that have roles downstream of CMG formation, or by CMG binding replisome factors like DNA polymerase alpha.

Dissociation and disassembly of Treslin<sup>Sld3</sup>-MTBP<sup>Sld7</sup>-TopBP1<sup>Dpb11</sup> may involve the action of phosphatases to counteract cell cycle kinase phosphorylation<sup>36,69,75</sup>. Moreover, the low affinities of the two individual GINS-interacting sites in TopBP1<sup>Dpb11</sup> offer the idea that release of one site would significantly destabilise the TopBP1-GINS complex to trigger TopBP1<sup>Dpb11</sup> dissociation. Competition of DNA Pole with TopBP1 binding may also be involved in TopBP1<sup>Dpb11</sup> dissociation, because simultaneous binding of the two proteins to CMG is sterically impossible without significant conformational reconfigurations. Whether TopBP1<sup>Dpb11</sup> dissociation is a prerequisite for DNA Pole association, or whether the polymerase acts to displace TopBP1<sup>Dpb11</sup> remains unclear. The reconfiguration of the Psf1-B domain during CMG formation suggests another potential mechanism for TopBP1<sup>Dpb11</sup> release. The B domain seems flexible in the TopBP1<sup>Dpb11</sup>-GINS complex and establishes an extensive set of interactions with Cdc45 when it reconfigures to form CMG, which might favour Pole binding over TopBP1<sup>Dpb11</sup>-BRCT4.

Recent insights have suggested that formation of the CMG complex is tightly associated with (i) splitting of Mcm2-7 double hexamers into single hexamers; (ii) the exchange of ADP for ATP in the Mcm2-7 nucleotide-binding sites; and (iii) limited untwisting of dsDNA inside the central channels<sup>6,11,12</sup>. How TopBP1<sup>Dpb11</sup>, Treslin<sup>Sld3</sup> and MTBP<sup>Sld7</sup> relates to these steps will be key to understanding whether their mechanistic roles go beyond just a simple chaperoning function, guiding Cdc45 and GINS to MCM-DHs. For example, the reported homo-dimerization of TopBP1<sup>Dpb11</sup> and Treslin<sup>Sld3</sup>-MTBP<sup>Sld7</sup> and DONSON might play a role in the cooperative conversion of Mcm2-7 hexamers into replisomes, thus guaranteeing bi-directional replication<sup>41,69,76</sup>.

Taken together, our data provide insight into, and a testable hypothesis for, the series of molecular events leading to CMG formation, and its regulation; potentially leading to future therapeutic routes for the treatment of cancer, due the high reliance of tumours on high levels of replication initiation<sup>77</sup>. Targeted interference with origin firing could be one potential avenue, consistent with the testing of DDK inhibitors as candidates in clinical trials<sup>78</sup>.

## Methods

Uncropped images of SDS PAGE gels are available as a Source data file. Information on antibodies and plasmids used are available in Supplementary Data 3.

### Protein purification

**Purification of recombinant TopBP1.** Purification of TopBP1-331-766 (BRCT3-5) for structural studies was done using bacterial expression. BL21(DE3) cells carrying a pET17 plasmid encoding HIS-(3C)-TopBP1-331-766-Strep was grown in TurboBroth at 37 °C to an OD<sub>600</sub> of 2.0 and induced by with 0.5 mM IPTG prior to overnight incubation at 16 °C. Lysis by sonication was done in lysis buffer (25 mM HEPES pH 7.5, 200 mM NaCl, 0.5 mM TCEP, 10 U DNASE Turbo and complete, EDTA-

free Protease Inhibitor Cocktail (Merck). The resulting lysate was clarified by centrifugation at 40,000 × g for 60 min at 4 °C. The supernatant was applied to a 5 ml HiTrap TALON crude column (GE Healthcare, Little Chalfont, UK), washed with lysis buffer, followed by a wash with lysis buffer containing 10 mM imidazole. Elution followed by increasing imidazole to 250 mM. The eluted protein was separated on a 5 ml HiTrap STREP column (GE Healthcare, Little Chalfont, UK), and eluted with 2 mM desthiobiotin. The eluate was concentrated before separation on a Superdex200increase 10 300 column (GE Healthcare, Little Chalfont, UK) equilibrated in 10 mM HEPES pH 7.5, 150 mM NaCl, 0.5 mM TCEP.

Recombinant TopBP1-1-766-Strep-WT (BRCT0-5) and mutants for interaction studies and experiments in *Xenopus* egg extract were purified using Sf9 insect cells. The cells were grown in suspension in insect cell media (Pan biotech, P04-850 500) at 27 °C. Baculoviruses were generated using pLIB-based constructs (Addgene, 80610). 500 ml Sf9 cells (1 × 10<sup>6</sup>/ml) were infected with a 1:100 dilution of recombinant baculovirus carrying TopBP1-1-766-Strep-WT or mutants, and were incubated at 27 °C for 72 h. Cell pellets were lysed by douncing in 30 ml of lysis buffer (20 mM HEPES pH 8.0, 150 mM NaCl, 0.1% (v/v) Tween-20, 0.5 mM TCEP, protease inhibitor cocktail (Roche Complete protease inhibitor cocktail, 05056489001)). The lysate was centrifuged at 44,800 × g for 45 min and the supernatant was loaded onto the StrepTrapHP-Iml column (Cytiva, 28907546). Elution was done using 2.5 mM desthiobiotin in elution buffer (20 mM HEPES pH 8.0, 150 mM NaCl, 0.01% Tween-20 (v/v), 0.5 mM TCEP, 2% (v/v) glycerol, 2.5 mM desthiobiotin).

**Purification of GINS.** GINS for structural and fluorescence polarisation studies was purified in Sf9 insect cells Psf2, Psf3 and HIS-(3C)-Sld5 were expressed together with Psf1 or Psf1 (1–151; B domain deleted) from a single baculovirus, respectively, produced using pBig1a. Cell pellets were resuspended in lysis buffer containing 25 mM HEPES pH 7.5, 200 mM NaCl, 0.5 mM TCEP, 10 U DNASE Turbo and complete, EDTA-free Protease Inhibitor Cocktail (Merck), then disrupted by sonication. The resulting lysate was clarified by centrifugation at 40,000 × g for 60 min at 4 °C. The supernatant was separated on a 5 ml HiTrap TALON crude column (GE Healthcare, Little Chalfont, UK), using washing in lysis buffer supplemented with 10 mM imidazole. Elution was done using a gradient up to 250 mM imidazole. Peak fractions were pooled and concentrated. For subsequent size exclusion chromatography (analytical and for EM) a Superdex200increase 10 300 column (GE Healthcare, Little Chalfont, UK) equilibrated in 10 mM HEPES pH 7.5, 150 mM NaCl, 0.5 mM TCEP was used while for FP experiments a Superdex200 16 600 column (GE Healthcare, Little Chalfont, UK) equilibrated in 25 mM HEPES pH 7.5, 150 mM NaCl, 1 mM EDTA, 0.25 mM TCEP, 0.02% (v/v) Tween 20 was used.

GINS for pulldown experiments was purified using 6xHis-3xFlag-Sld5. Sf9 insect cells in suspension cultures were co-infected with 1:50 dilution of each four recombinant baculoviruses for expressing 6xHis-3xFlag-Sld5, Psf1, Psf2 and Psf3, respectively. The viruses were made using pLib-based plasmids. 72 h post infection, pellets from one-litre cells were lysed by douncing in 60 ml of lysis buffer (20 mM HEPES pH 8.0, 300 mM NaCl, 0.1% (v/v) Tween-20, 25 mM imidazole, 0.5 mM TCEP, protease inhibitor cocktail (Roche Complete protease inhibitor cocktail, 05056489001)). The lysate was centrifuged at 44,800 × g for 45 min and the supernatant was incubated with 2.5 ml Ni-NTA agarose beads (Qiagen, 30210) for 1 h at 4 °C. The bound protein was eluted with six bead volumes elution buffer (20 mM HEPES pH 8.0, 300 mM NaCl, 0.01% (v/v) Tween-20, 250 mM imidazole, 0.5 mM TCEP, 2% (v/v) glycerol). The eluate was dialysed into 20 mM HEPES pH 8.0, 150 mM NaCl, 0.01% (v/v) Tween-20, 0.5 mM TCEP, 2% (v/v) glycerol, and then applied onto a HiTrap HP Q column (Cytiva, 17115301). To elute the bound protein, a linear gradient of NaCl (from 150 mM to 1 M) in elution buffer (20 mM HEPES pH 8.0, 0.01% (v/v) Tween-20, 0.5 mM



TCEP, 2% (v/v) glycerol) was applied. Peak fractions were pooled and separated by size exclusion chromatography (equilibration buffer: 20 mM HEPES pH 8.0, 300 mM NaCl, 0.01% (v/v) Tween-20, 0.5 mM TCEP, 2% (v/v) glycerol).

**Purification of PolE2-N and geminin from *E. coli*.** PolE2-N-1-75 (amino acids 1–75) with N-terminal MBP-TEV2 tag in pMal, and GST tagged *Xenopus* geminin (Xgeminin) in pGEX were expressed in Rosetta *E. coli* culture. Expressions were induced with 1 mM IPTG at OD<sub>600</sub> = 0.6 at 20 °C (PolE2-N) or 25 °C (Xgeminin) overnight. Cells were harvested by centrifugation. For PolE2-N, cell pellets were resuspended in lysis buffer (20 mM Hepes pH 8.0, 150 mM NaCl, 0.5 mM TCEP, 2 (v/v) glycerol and protease inhibitor cocktail (Roche Complete protease inhibitor cocktail, 05056489001)) and lysed by sonication. The cell lysate was clarified by centrifugation. The supernatant was incubated with amylose resin (New England Biolabs, E8021S) for 1 h before elution with lysis buffer + 10 mM maltose. The peak fraction was dialysed into TopBP1 buffer (20 mM HEPES pH 8.0, 150 mM NaCl, 0.5 mM TCEP, 0.01% (v/v) Tween-20 and 2% (v/v) glycerol). For Xgeminin, lysis by sonication was carried out in 5 mg/ml lysozyme, 20 mM Hepes pH 7.7, 200 mM NaCl, 5 mM β-mercaptoethanol, 5% glycerol. After a clarifying centrifugation step, the supernatant was incubated with glutathione-sepharose (0.5 ml per liter culture; GE Healthcare 17513201) for 3 h before elution with lysis buffer + 50 mM glutathione (Applchem A2084,0025). Peak fractions were pooled and dialysed into XBE2 buffer (100 mM KCl, 2 mM MgCl<sub>2</sub>, 0.1 mM CaCl<sub>2</sub>, 1.71% w:v sucrose, 5 mM K-EGTA, 10 mM HEPES-KOH, pH 7.7) and concentrated to 360 μM Xgeminin. Aliquots were shock frozen and stored at –80 °C.

#### Size exclusion chromatography for isolating the TopBP1-GINS complex

Recombinant TopBP1 and GINS proteins were mixed in an equimolar ratio to give a final concentration of 20 μM of each component and incubated on ice for at least 30 min prior to application to a Superdex200increase 10 300 column (GE Healthcare, Little Chalfont, UK) equilibrated in 10 mM HEPES pH 7.5, 150 mM NaCl, 0.5 mM TCEP.

#### Fluorescence polarisation experiments

Fluorescein-labelled peptides (WT: Flu-GYGAPSEKHEQAEDLLSQ YENG or LLPP: Flu-GYGAPSEKHEQAEDPPSQYENG) (Peptide Protein Research Ltd, Bishops Waltham, UK) at a concentration of 100 nM, were incubated at room temperature with increasing concentrations of GINS in 25 mM HEPES pH 7.5, 150 mM NaCl, 1 mM EDTA, 0.25 mM TCEP, 0.02% (v/v) Tween 20 in a black 96-well polypropylene plate (VWR, Lutterworth, UK). Fluorescence polarisation was measured in a POLARstar Omega multimode microplate reader (BMG Labtech GmbH, Offenburg, Germany). Binding curves represent the mean of three independent experiments, with error bars representing SEM.

#### Immunoprecipitations from transiently transfected 293T cells

Transient transfections of 6xMyc-TeV2-TopBP1 into 293T cells (ATCC CRL-11268) were carried out using PEI (polyethyleneimine) according to a protocol kindly shared by David Cortez' lab. 4 μg plasmid DNA in 100 μl DMEM (ThermoFisher, 41965039) without penicillin-streptomycin were combined with 2.4 μl Polyethylenimin (Sigma, 408727; 10 mg/ml) and incubated for 20 min before addition to 4 × 10<sup>6</sup> cells on a 10-cm dish. Transfected cells were used 72 h post transfection. 72 h post transfection, 293T cells were lysed by douncing in ten times cell pellet volume of lysis buffer (20 mM HEPES pH 8.0, 150 mM NaCl, 0.1% (v/v) Tween-20, 0.5 mM TCEP, 2% (v/v) glycerol, protease inhibitor cocktail (Roche, Complete protease inhibitor cocktail, 05056489001)), and centrifuged at 20,000 × g for 15 min at 4 °C. The soluble supernatant was added to magnetic anti-Myc beads (5 μl slurry per sample; ThermoFischer, 88842) and incubated for 2 h at 4 °C rotating. Cells equivalent to 50% of a 15 cm dish were used per sample.

Beads were washed three times in lysis buffer with 5 min incubation each, and finally boiled in 50 μl Laemmli loading buffer (6.5% glycerol, 715 mM β-mercaptoethanol, 3% SDS, 62.5 mM Tris-HCl pH 7.9, 0.005% bromophenol blue).

#### Pulldown experiments of recombinant PolE2-N or TopBP1 using immobilised GINS

Magnetic anti-Flag beads (1 μl slurry for western blot/silver staining-scale experiments, 4 μl for Coomassie-scale experiments) were coupled with 600 ng (western/silver-scale) or 2400 ng (coomassie-scale) GINS via 3xFlag-Sld5. 100 nM final concentration of TopBP1-BRCTO-5-Strep (WT or mutants) in 20 μl (western/silver-scale) or 80 μl (coomassie-scale) of interaction buffer (20 mM HEPES pH 8.0, 150 mM NaCl (or 100 mM NaCH<sub>3</sub>COO when indicated), 0.01% (v/v) Tween-20, 0.5 mM TCEP, 2% (v/v) glycerol) were added to GINS-coupled beads in 5 μl (western/silver-scale) or 20 μl (coomassie-scale) interaction buffer. For PolE2-N, 1 μM PolE2-N protein in 20 μl were used for GINS-coupled beads in 5 μl reaction buffer. After incubation for 45 min at 4 °C rotating, beads were washed three times for 5 min with interaction buffer and boiled in 50 μl Laemmli buffer before SDS PAGE.

#### Pulldown from cell lysates with recombinant TopBP1-1-766-strep (BRCTO-5)

For pulldown assays from soluble lysates of non-transfected 293T cells, 5 μl streptactin Sepharose HP (Cytiva, 28935599) beads were coupled with 10 μg recombinant TopBP1-1-766-strep-WT or mutants in coupling buffer (20 mM HEPES pH 8.0, 150 mM NaCl, 0.1% (v/v) Tween-20, 0.5 mM TCEP, 2% (v/v) glycerol) for 45 min at 4 °C before washing two times with coupling buffer and then once in cell lysis buffer (20 mM HEPES pH 8.0, 150 mM NaCl, 0.1% (v/v) Tween-20, 0.5 mM TCEP, 2% (v/v) glycerol, protease inhibitor cocktail (Roche Complete protease inhibitor cocktail, 05056489001)). Control pulldowns were done with Flag-peptide-coupled beads or, when indicated, using GINS-coupled beads and the non-GINS binding TopBP1-Gcc-B4mut. 293T cells were lysed in ten times pellet volume of cell lysis buffer, and centrifuged at 20,000 × g for 15 min at 4 °C. Soluble lysate from cells equivalent to 50% of a 15 cm tissue culture plate was added to the TopBP1-coupled beads and incubated for 2 h at 4 °C rotating. Beads were washed three times with 5 min incubation in lysis buffer and boiled in 50 μl Laemmli buffer before SDS PAGE.

#### Generation of replicating *Xenopus* egg extracts

Our work with *Xenopus laevis* frogs complied with ethical regulations. The protocols used in this study, namely to handle the frogs, collect their eggs for cytosolic extract generation, and to prepare sperm, were approved by the Landesamt für Natur-, Umwelt- und Verbraucherschutz, Nordrhein-Westfalen (81-02.05.40.20.050).

*Xenopus laevis* egg extracts were prepared from metaphase II arrested eggs as described in ref. 79. After washing with MMR (100 mM NaCl, 2 mM KCl, 1 mM MgCl<sub>2</sub>, 2 mM CaCl<sub>2</sub>, 0.1 mM EDTA, 5 mM HEPES-NaOH, pH 7.8) eggs were dejellied (2% cysteine w:v in H<sub>2</sub>O, pH 7.8 with NaOH). Eggs were rinsed in XBE2 (100 mM KCl, 2 mM MgCl<sub>2</sub>, 0.1 mM CaCl<sub>2</sub>, 1.71% w:v sucrose, 5 mM K-EGTA, 10 mM HEPES-KOH, pH 7.7). Then, the eggs were transferred into centrifuge tubes containing 1 ml XBE2 + 10 μg/ml protease inhibitor (aprotinin Sigma A6279, leupeptin Merck 108975, pepstatin Carl Roth 2936.1) + 100 μg/ml cytochalasin D (Sigma C8273). For packing, eggs were centrifuged for 1 min at 1400 × g at 16 °C in a swingout rotor. Excess buffer, activated and apoptotic eggs were removed before crushing the eggs at 16,000 × g for 10 min at 16 °C. The extract was collected with a 20 G needle and supplemented with 10 μg/ml cytochalasin D, 10 μg/ml protease inhibitors, 1:80 dilution of energy regenerator (1M phosphocreatine K salt, 600 μg/ml creatine phosphokinase in 10 mM HEPES-KOH pH 7.6) and LFB1/50 (10% w-v sucrose, 50 mM KCl, 2 mM MgCl<sub>2</sub>, 1 mM EGTA, 2 mM DTT, 20 mM K<sub>2</sub>HPO<sub>4</sub>/KH<sub>2</sub>PO<sub>4</sub> pH 8.0, 40 mM

HEPES–KOH, pH 8.0) to 15% (v:v). The extract was cleared at  $84,000 \times g$  at  $4^\circ\text{C}$  for 20 min in a swingout rotor. The cytoplasmic layer was collected and supplemented with 2% glycerol v:v. The extract was frozen by dropping 20  $\mu\text{l}$  aliquots into liquid nitrogen and stored at  $-80^\circ\text{C}$ .

### Immunodepletion of TopBP1 from *Xenopus* egg extract

For XTopBP1 depletion from *Xenopus* egg extract, 0.5  $\mu\text{g}$  antibody #1, #2 or IgG (rabbit; self-made) were coupled per  $\mu\text{l}$  magnetic Protein G dynabeads (Life Technologies 10004D) for 1 h at RT. Freshly thawed extracts were supplemented with 1/40 energy regenerator and 250  $\mu\text{g}/\text{ml}$  cycloheximide and released into interphase by adding 0.3 mM  $\text{CaCl}_2$  for 15 min at  $23^\circ\text{C}$ . Subsequently, extracts were depleted for 45 min on ice with 0.675  $\mu\text{g}$  antibody per  $\mu\text{l}$  extract. After depletion the extract was aliquoted, snap frozen and stored at  $-80^\circ\text{C}$  for replication assays and chromatin isolations.

### *Xenopus* egg extract replication assays

For TopBP1 addback experiments, 6 ng/ $\mu\text{l}$  recombinant TopBP1 protein or buffer were added to TopBP1 immunodepleted interphase *Xenopus* egg extract containing 50 nCi/ $\mu\text{l}$   $\alpha^{32}\text{P}$ -dCTP (Perkin Elmer NEG513A250UC). To start the replication reaction the egg extracts were supplemented with 5 ng sperm DNA (kindly provided by the lab of O. Stemmann) per  $\mu\text{l}$  extract and incubated at  $23^\circ\text{C}$ . 1  $\mu\text{l}$  extract was spotted for each time point onto a glass fibre membrane. Unbound  $\alpha^{32}\text{P}$ -dCTP was removed by rinsing the membrane three times 15 min with ice cold 5% TCA in water and once with ice cold EtOH. Newly replicated DNA was detected by phospho-imaging. Standard error of the mean was calculated from three independent experiments.

### Chromatin isolation from *Xenopus* egg extracts and analysis by mass spectrometry (CHROMASS) and western blotting

For chromatin isolation, XTopBP1-depleted interphase egg extract was supplemented with 50  $\mu\text{g}/\text{ml}$  aphidicolin and aliquoted into 15  $\mu\text{l}$  samples before addition of buffer or recombinant TopBP1 to a final concentration 6 ng/ $\mu\text{l}$ . 2.25  $\mu\text{M}$  Xgeminin or buffer were added before incubation for 10 min on ice. To start the replication, 9 ng sperm DNA/ $\mu\text{l}$  extract was added and incubated for 75 min at  $23^\circ\text{C}$ . Reactions were stopped with 300  $\mu\text{l}$  ice cold ELB salt (10 mM HEPES–KOH pH 7.7, 50 mM KCl, 2.5 mM  $\text{MgCl}_2$ ) + 250 mM sucrose + 0.6% triton X-100. The diluted extract was loaded onto a sucrose cushion (150  $\mu\text{l}$  ELB salt + 25% sucrose) and centrifuged for 10 min at  $2500 \times g$  at  $4^\circ\text{C}$  in a swingout rotor. The supernatant above the sucrose cushion was removed and the cushion surface was washed twice with 200  $\mu\text{l}$  ELB salt containing 250 mM sucrose. The cushion was removed leaving about 15  $\mu\text{l}$  behind, followed by centrifugation for 2 min at  $10,000 \times g$  at  $4^\circ\text{C}$  in a fixed angle rotor. For western blotting, the chromatin pellet was resuspended in 24  $\mu\text{l}$  1x Laemmli SDS sample buffer (6.5% glycerol, 715 mM  $\beta$ -mercaptoethanol, 3% SDS, 62.5 mM Tris-HCl pH 7.9, 0.005% bromophenol blue) and 5  $\mu\text{l}$  were separated by denaturing SDS PAGE. For MS analysis, four biological replicates of chromatin pellets were analysed by CHROMASS<sup>45</sup>. In brief, chromatin pellets were resuspended in 50  $\mu\text{l}$  denaturation buffer (8 M Urea; 100 mM Tris pH 8). Dithiothreitol (DTT) was added to a final concentration of 5 mM and samples were incubated at  $22^\circ\text{C}$  for 30 min. To alkylate peptides, iodoacetamide (20 mM) was added and samples were incubated at  $37^\circ\text{C}$  for 30 min. DTT (25 mM) was added, and samples were incubated at  $22^\circ\text{C}$  for 5 min. 500 ng Lys-C (add Supplier) were added, and samples were incubated at  $37^\circ\text{C}$  for 3 h. Samples were diluted with 100 mM ammonium bicarbonate to adjust the urea concentration to 1 M. Trypsin (Sigma, 500 ng per sample) was added and samples were incubated over night at  $37^\circ\text{C}$ . Digested peptides were acidified with 1% trifluoroacetic acid (TFA) and desalted on Empore C18 material according to Rappsilber et al. <https://www.nature.com/articles/nprot.2007.261>. Eluted peptides were dried in a vacuum concentrator and reconstituted with 9  $\mu\text{l}$

of buffer A (0.1% formic acid) containing 0.2% acetonitrile and 0.01% trifluoroacetic acid. For MS analysis, 4  $\mu\text{l}$  of the reconstituted peptides were separated on an EASY-nLC™ 1200 chromatography system (Thermo Scientific) coupled to a Q Exactive HF-X Orbitrap LC-MS/MS System (Thermo Fisher Scientific) via a Nanoflex ion source. Peptide separation was carried out in analytical columns (50 cm, 75  $\mu\text{m}$  inner diameter packed in-house with C18 resin ReproSilPur 120, 1.9  $\mu\text{m}$  diameter Dr. Maisch) using a 3-h nonlinear gradient at a flow rate of 250 nl/min using buffer A (aqueous 0.1% formic acid) and buffer B (80% acetonitrile, 0.1% formic acid). MS data was acquired in data-dependent fashion using a Top15 method. The exact parameters for chromatography and MS instrument settings can be retrieved from the raw files available at ProteomeXchange (PXD040024).

MS data files from single-shot experiments were processed with MaxQuant (version 2.0.1.0) using a non-redundant *Xenopus laevis* data base available at ProteomeXchange (PXD040024)<sup>45</sup>. Raw data were normalized using the label-free quantitation (LFQ) algorithm implemented in MaxQuant. MS Data with Perseus (version 1.6.15.0)<sup>80</sup>. Protein groups were filtered to eliminate contaminants, reverse hits, and proteins identified by site only. For the heat map (Fig. 6a) LFQ intensities were log<sub>2</sub> transformed and for each protein z-scores were calculated across all replicates ( $N = 4$ ) of all four conditions. Subsequently, the average of the z-scores was calculated for each condition and selected proteins were plotted (see supplementary data 2 for the z-score of all proteins). Proteins were manually annotated and sorted according to their function in DNA replication and DNA repair. To identify proteins with significant abundance changes between the four conditions, LFQ intensities were log<sub>2</sub> transformed and missing values were imputed with random values drawn from a normal distribution centred around the detection limit of the MS instrument (Perseus imputation; width = 0.3, down shift = 1.8). Two sample Student's t-tests were carried out in Perseus. Student's t-tests were carried out in Perseus. For these tests three valid values in at least one quadruplicate of either of the tested conditions was required. FDR was adjusted for multiple testing by the Benjamini–Hochberg procedure using a significance threshold of  $\text{FDR} < 0.05$  (see supplementary data 2). Data visualisation was carried out in R. All scripts are available upon request.

### Determination of TopBP1-proximal proteins by APEX2 biotinylation

**Sample preparation.** Proximity labelling was performed in isogenic stable 293 Flip-In (Thermo Fisher; R75007) cells stably expressing TopBP1 N-terminal tagged with APEX2 or in 293T cells transiently expressing either N-terminal or C-terminally APEX2-tagged TopBP1. 24 h before transfection,  $0.8 \times 10^6$  cells were seeded onto a 6-cm plate. For each condition, four dishes were separately transfected (biological replicates). One day after transfection, cells were incubated for 30 min with 500  $\mu\text{M}$  biotin phenol at  $37^\circ\text{C}$ , before incubation for 1 min at RT with or without (controls) 1 mM  $\text{H}_2\text{O}_2$ . The medium was discarded, and cells were washed three times with quenching buffer (10 mM sodium azide, 10 mM sodium ascorbate and 5 mM Trolox in Dulbecco's PBS (Life Technologies 14190169)). Cells were rinsed off the dishes with DPBS, transferred into reaction tubes before washing again with DPBS. Cells were lysed at  $95^\circ\text{C}$  for 10 min in TSD buffer (50 mM Tris pH 7.5, 5 mM DTT and 1% SDS). Samples were diluted ten times with TNN buffer (20 mM Tris pH 7.9, 200 mM NaCl and 0.5% NP-40 alternative) and benzonase (75 U per sample) (Sigma E1014). Samples were incubated for 15 min at  $4^\circ\text{C}$  and centrifuged at  $21,000 \times g$  for 2 min at  $4^\circ\text{C}$ . Supernatants were incubated over night with 15  $\mu\text{l}$  streptavidin sepharose beads (Sigma GE17-5113-01) at  $4^\circ\text{C}$ . Beads were washed once with TNN buffer + 0.1% SDS and twice with 25 mM ammonium bicarbonate before processing for mass spectrometry. For this, captured proteins were washed two times with  $\text{H}_2\text{O}$  prior to on-bead digestion to remove MS incompatible buffer components. The beads were taken

up in 100  $\mu$ l 0.8 M urea, 50 mM ammonium bicarbonate buffer (ABC) and supplemented with 5 mM DTT. After incubation at 37 °C for 30 min, 10 mM iodoacetamide (IAM) was added and incubated for 30 min at room temperature in the dark while shaking at 1500 rpm (Thermomixer C, Eppendorf). The IAM was quenched with 11 mM DTT. Trypsin was added to a total of 300 ng Trypsin per sample. The samples were incubated over night at 37 °C, shaking at 1150 rpm before stopping by 1% (vol/vol) formic acid (FA). After bead collection by centrifugation (650  $\times$  g, 5 min) 100  $\mu$ l supernatant were transferred to a Eppendorf Lo-bind 1.5 ml tube. The remaining beads were incubated with 50  $\mu$ l 1% formic acid for (1150 rpm, 5 min rpm, RT) and after collecting the beads once more by centrifugation (650  $\times$  g, 5 min) the supernatant was combined with the first one. The sepharose beads were discarded. To remove residual sepharose beads, the combined peptide containing solutions were passed over pre-equilibrated (50  $\mu$ l 0.5% formic acid) home-made 2-disc Glass microfiber StageTip (disc material: GE Healthcare; pore size: 1.2  $\mu$ m; thickness: 0.26 mm; 50  $\times$  g, 2 min). The cleared tryptic digests were then desalted on home-made C18 StageTips as described<sup>81</sup>. Briefly, peptides were immobilized and washed on a 2 disc C18 (Empore) StageTip. After elution from the StageTips, samples were dried using a vacuum concentrator (Eppendorf) and the peptides were taken up in 0.1% formic acid solution (10–15  $\mu$ l) and directly used for LC-MS/MS experiments.

**LC-MS/MS settings.** MS Experiments were performed on an Orbitrap Fusion LUMOS instrument (Thermo) coupled to an EASY-nLC 1200 ultra-performance liquid chromatography (UPLC) system (Thermo). The UPLC was operated in the one-column mode. The analytical column was a fused silica capillary (75  $\mu$ m  $\times$  46 cm) with an integrated fritted emitter (CoAnn Technologies) packed in-house with Kinetex 1.7  $\mu$ m C18-XB core shell beads (Phenomenex). The analytical column was encased by a column oven (Sonation PRSO-V2) and attached to a nanospray flex ion source (Thermo). The column oven temperature was set to 50 °C during sample loading and data acquisition. The LC was equipped with two mobile phases: solvent A (0.2% FA, 2% Acetonitrile, ACN, 97.8% H<sub>2</sub>O) and solvent B (0.2% FA, 80% ACN, 19.8% H<sub>2</sub>O). All solvents were of UPLC grade (Honeywell). Peptides were directly loaded onto the analytical column with a maximum flow rate that would not exceed the set pressure limit of 980 bar (usually around 0.4–0.6  $\mu$ l/min). Peptides were subsequently separated on the analytical column by running a 105 min gradient of solvent A and solvent B (start with 3% B; gradient 3% to 9% B for 6:30 min; gradient 9% to 30% B for 62:30 min; gradient 30% to 50% B for 24:00 min; 50% to 100% B for 2:30 min; 100% for 9:30 min) at a flow rate of 300 nl/min. The mass spectrometer was controlled by the Orbitrap Fusion Lumos Tune Application (version 3.3.2782.28) and operated using the Xcalibur software (version 4.3.73.11). The mass spectrometer was set in the positive ion mode. The ionization potential (spray voltage) was set to 2.5 kV. Source fragmentation was turned off. Precursor ion scanning was performed in the Orbitrap analyser (FT; Fourier transform mass spectrometer) in the scan range of  $m/z$  370–1500 and at a resolution of 240,000 with the internal lock mass option turned on (lock mass was 445.120025  $m/z$ , polysiloxane)<sup>82</sup>. AGC (automatic gain control) was set to standard and acquisition time to auto. Product ion spectra were recorded in a data-dependent fashion in the IT (IT; ion trap mass spectrometer) in a variable scan range (auto setting) and at rapid scan rate. Peptides were analysed using the setting top speed (repeating cycle of full precursor ion scan (AGC target set to 300%; acquisition time set to auto) followed by dependent MS2 scans for 3 s (minimum intensity threshold  $4 \times 10^3$ )). The MS2 precursor ions were isolated using the quadrupole (isolation window 1.6  $m/z$ ) and fragmentation was achieved by Higher-energy C-trap dissociation (HCD) (normalized collision mode (stepped setting) and normalized collision energy set to 27, 32, 40%). During MS2 data acquisition dynamic ion exclusion was

set to 20 s. Only charge states between 2–7 were considered for fragmentation.

**Data processing.** RAW spectra were submitted to an Andromeda<sup>83</sup> search in MaxQuant (v2.0.3.0) using the default settings (Cox and Mann, 2008)<sup>84</sup>. Label-free quantification and match-between-runs was activated<sup>85</sup>. The MS/MS spectra data were searched against the Uniprot *H. sapiens* reference database (one protein per gene; UP000005640\_9606\_OPPG.fasta, 20585 entries, downloaded 1/10/2022) and a dedicated database containing the APEX2-myc sequence (ACE\_0741\_SOI\_v01.fasta; 1 entry). All searches included a contaminants database search (as implemented in MaxQuant, 246 entries). The contaminants database contains known MS contaminants and was included to estimate the level of contamination. Andromeda searches allowed oxidation of methionine residues (16 Da) and acetylation of the protein N-terminus (42 Da) as dynamic modifications and the static modification of cysteine (57 Da, alkylation with iodoacetamide). Enzyme specificity was set to Trypsin/P with two missed cleavages allowed. The instrument type in Andromeda searches was set to Orbitrap and the precursor mass tolerance was set to  $\pm$  20 ppm (first search) and  $\pm$  4.5 ppm (main search). The MS/MS match tolerance was set to  $\pm$ 0.5 Da. The peptide spectrum match FDR and the protein FDR were set to 0.01 (based on target-decoy approach). Minimum peptide length was 7 aa. For protein quantification unique and razor peptides were allowed. Modified peptides were allowed for quantification. The minimum score for modified peptides was 40. Label-free protein quantification was switched on, and unique and razor peptides were considered for quantification with a minimum ratio count of 2. Retention times were recalibrated based on the built-in nonlinear time-rescaling algorithm. MS/MS identifications were transferred between LC-MS/MS runs with the match-between-runs option in which the maximal match time window was set to 0.7 min and the alignment time window set to 20 min. The quantification is based on the value at maximum of the extracted ion current. At least two quantification events were required for a quantifiable protein. Further analysis and filtering of the results was done in Perseus v1.6.10.0.<sup>80</sup> For quantification, we combined related biological replicates to categorical groups and investigated only those proteins that were found in at least one categorical group in a minimum of 3 out of 4 biological replicates. Comparison of protein group quantities (relative quantification) between different MS runs is based solely on the LFQ's as calculated by MaxQuant, MaxLFQ algorithm<sup>85</sup>. Statistical evaluation was done by two-sided Student's t-testing (FDR = 0.05, SO = 0.1)

#### Analysis of the TopBP1-GINS by cross-linking mass spectrometry

**Sample preparation.** 50  $\mu$ g (1.15  $\mu$ g/ $\mu$ l) of each GINS and TopBP1-BRCT0-5-strep (amino acids 1–766) were incubated for 45 min on ice in cross-linking buffer (20 mM HEPES pH 8.0, 150 mM NaCl, 0.01% Tween-20, 2% Glycerol, 0.5 mM TCEP). 600  $\mu$ M or 2500  $\mu$ M of BS<sup>3</sup> cross-linker were added and incubated for 30 min at 35 °C shaking. The reaction was stopped by adding ammonium bicarbonate to a final concentration of 100 mM. 90% of the cross-linked sample was used for mass spectrometry analysis of cross-links. 10% of the reaction were separated by SDS-PAGE. The coomassie-stained gel was cut in slices above the molecular weight of monomeric TopBP1-BRCT0-5-strep for subsequent mass spectrometry.

**Sample processing for mass spectrometry.** Sample preparation of cross-linked samples for LC/MS/MS is based on the SP3 protocol<sup>86</sup>. 30  $\mu$ g of protein from each cross-linking sample was taken up in 100  $\mu$ l 1 $\times$  SP3 lysis buffer (final concentrations: 5% (wt/vol) SDS; 10 mM TCEP; 200  $\mu$ l 40 mM chloracetamide; 200 mM HEPES pH 8) and heated for 5 min at 90 °C. After cooling the samples to room temperature (on ice) a mix of 150  $\mu$ g hydrophobic (#65152105050250) and 150  $\mu$ g



hydrophilic (#45152105050250) SeraMag Speed Beads (Cytiva) was added (bead to protein ratio 10 to 1) and gently mixed. Then 100  $\mu\text{L}$  100% vol/vol Ethanol (EtOH) was added before incubation for 20 min at 24 °C shaking vigorously. The beads were collected on a magnet and the supernatant aspirated. The beads were then washed 4 times with 180  $\mu\text{L}$  80% EtOH (collection time on the magnet minimum of 4 min). The beads were then finally taken up in 100  $\mu\text{L}$  25 mM ammoniumbicarbonate (ABC) containing 1  $\mu\text{g}$  Trypsin (Protein:Trypsin ratio 30:1). To help bead dissociation, samples were incubated for 5 min in a sonification bath (preheated to 37 °C). Samples were incubated over night shaking at vigorously. Samples were acidified with formic acid (FA, final 1% vol/vol) before collection on a magnet. The supernatants were transferred to a fresh Eppendorf tube, before removing trace beads using a magnet for 5 min. The tryptic digests were then desalted on home-made C18 StageTips as described<sup>81</sup>. Briefly, peptides were immobilized and washed on a 2 disc C18 StageTip. Samples were then dried using a vacuum concentrator (Eppendorf) and the peptides were taken up in 0.1% formic acid solution (10  $\mu\text{L}$ ) and directly used for LC-MS/MS experiments.

**LC-MS/MS settings cross-linking mass spectrometry.** MS Experiments were performed on an Orbitrap Fusion LUMOS instrument (Thermo) coupled to an EASY-nLC 1200 ultra-performance liquid chromatography (UPLC) system (Thermo). The UPLC was operated in the one-column mode. The analytical column was a fused silica capillary (75  $\mu\text{m}$   $\times$  41 cm) with an integrated fritted emitter (CoAnn Technologies) packed in-house with Kinetex 1.7  $\mu\text{m}$  C18-XB core shell beads (Phenomenex). The analytical column was encased by a column oven (Sonation PRSO-V2) and attached to a nanospray flex ion source (Thermo). The column oven temperature was set to 50 °C during sample loading and data acquisition. The LC was equipped with two mobile phases: solvent A (0.2% FA, 2% Acetonitrile, ACN, 97.8% H<sub>2</sub>O) and solvent B (0.2% FA, 80% ACN, 19.8% H<sub>2</sub>O). All solvents were of UPLC grade (Honeywell). Peptides were directly loaded onto the analytical column with a maximum flow rate that would not exceed the set pressure limit of 980 bar (usually around 0.4–0.6  $\mu\text{L}/\text{min}$ ). Peptides were subsequently separated on the analytical column by running a 70 min gradient of solvent A and solvent B (start with 2% B; gradient 2% to 6% B for 5:00 min; gradient 6% to 25% B for 42:00 min; gradient 25% to 40% B for 15:00 min; 40% to 98% B for 1:00 min; 98% for 7:00 min) at a flow rate of 350 nL/min. The mass spectrometer was controlled by the Orbitrap Fusion Lumos Tune Application (version 3.3.2782.28) and operated using the Xcalibur software (version 4.3.73.11). The mass spectrometer was set in the positive ion mode. The ionization potential (spray voltage) was set to 2.5 kV. Source fragmentation was turned off. Precursor ion scanning was performed in the Orbitrap analyser (FT; Fourier transform mass spectrometer) in the scan range of  $m/z$  370–1600 and at a resolution of 120,000 with the internal lock mass option turned on (lock mass was 445.120025  $m/z$ , polysiloxane)<sup>82</sup>. AGC (automatic gain control) was set to standard and acquisition time to auto. Product ion spectra were recorded in a data-dependent fashion in the FT in a variable scan range (auto setting) and at 15,000 resolution. Peptides were analysed using a top speed regime (repeating cycle of full precursor ion scan (AGC target set to standard; acquisition time 200 ms) followed by dependent MS2 scans for 5 s (minimum intensity threshold  $2 \times 10^4$ ). The MS2 precursor ions were isolated using the quadrupole (isolation window 2.0  $m/z$ ) and fragmentation was achieved by Higher-energy C-trap dissociation (HCD) (normalized collision mode set to stepped and normalized collision energy set to 27, 30, 33%). During MS2 data acquisition dynamic ion exclusion was set to 60 s. Only charge states between 3–7 were considered for fragmentation.

**Data processing protocol.** The searches for cross-linked peptides were performed with two different search engines: MetaMorpheus

(MM) version 0.0.320<sup>87</sup> and pLink2 version v2.3.9<sup>88</sup>. Searches were performed on the Thermo RAW files using the database ACE\_0704\_SOI\_plus\_con\_v01.fasta (244 entries). The database contains the sequences for TopBP1 (1–766), Psfl, PsfII, PsfIII, Sld5 and 239 known contaminating proteins found in MS samples. The peptide spectrum match FDR for MM was 0.01 and for pLink2 0.05 (based on target-decoy approach, decoys are generated by the software). The settings for MetaMorpheus were: cross-linker name = DSS\_KSTY (note: DSS and BS<sup>3</sup> have identical cross-linker size); cross-linker type = non-cleavable; cross-linker mass = 138.06808; cross-linker modification site 1 = K; cross-linker modification site 2 = KSTY; protease = trypsin; maximum missed cleavages = 3; minimum peptide length = 5; maximum peptide length = 60; initiator methionine behaviour = Variable; max modification isoforms = 1024; fixed modifications = Carbamidomethyl on C, variable modifications = Oxidation on M; parent mass tolerance(s) =  $\pm 10$  ppm; product mass tolerance =  $\pm 20$  ppm. The settings for pLink2 were: cross-linker name = DSS-BS<sup>3</sup>-KSTY; cross-linker type = non-cleavable; cross-linker mass = 138.06808; cross-linker modification site 1 = K; cross-linker modification site 2 = KSTY; protease = trypsin; maximum missed cleavages = 3; minimum peptide length = 6; maximum peptide length = 60; fixed modifications = Carbamidomethyl on C, variable modifications = Oxidation on M; parent mass tolerance(s) =  $\pm 20$  ppm; product mass tolerance =  $\pm 20$  ppm. The results from both searches were converted to the ProXL XML format using the respective converters (metaMorph2ProxlXML.jar, plink2toProxlXML.jar, follow links on the ProXL website: <https://proxl-ms.org/>) and uploaded to the ProXL server<sup>89</sup>. Analysis and evaluation of cross-links was performed on our local ProXL Server. The results from both searches was analysed together.

### Computational modelling

Locally implemented AlphaFold2-Multimer version 2.2.0<sup>90</sup> was used to predict the binding interfaces of GINS and TopBP1 using the full-length sequences of each GINS subunit and residues 301–766 of human TopBP1. 5 models were obtained, using 5 seeds per model, by default and ranked with the combined score of ipTM+pTM implemented in AlphaFold2-Multimer<sup>90,91</sup>. The top ranked model used for figure production and statistical plots were produced using Alphapickle 1.5.4 for the top seed from each model.

Blind molecular docking for insight into the GINI binding region used the GINI sequence DEDLLSQY of TopBP1 (residues 487–494) in the servers Cluspro<sup>92</sup> (accessed on 13/06/2022), HPEPDOCK<sup>93</sup> (accessed on 13/06/2022), and MDockPeP<sup>94–96</sup> (accessed on 13/06/2022) web servers with 200 solutions per server. We further extended, refined, and re-scored the sampling, by generating 10 additional conformations per web server solution (for a total of 6000 solutions) with FlexPepDock<sup>97</sup>. The top-10 best-scored solutions were selected for analysis and visualization.

The binding free energies of GINS with BRCT4 and GINI were estimated based on the best scored AlphaFold2 model using the web servers PRODIGY<sup>98</sup> and PPI-Affinity<sup>99</sup>. The region of GINI employed for the estimation comprised residues 481 to 496. Both PRODIGY and PPI-Affinity analyse the residues in the binding interface for free energy estimation. PRODIGY uses a linear amino acid chain model, whilst PPI-Affinity uses a structure-based machine learning algorithm.

### Cryo-electron microscopy

**Grid preparation.** Quantifoil 1.2/1.3, 300 mesh copper grids (Quantifoil) were glow discharged using a Tergeo Plasma Cleaner (Pie Scientific) with an indirect plasma treatment for 30 s. Grids were loaded into a Leica EM GP2 (Leica microsystems) and 3  $\mu\text{L}$  of peak fractions from size exclusion chromatography were diluted to 0.1 mg/ml and applied to the front of the grid, with an additional 0.5  $\mu\text{L}$  buffer applied to the grids rear, before back blotting for 3 s and plunging into an ethane propane mix.



**Electron microscopy and data processing.** For dataset 1, grids were stored in liquid nitrogen prior to imaging at 300 kV on a FEI Titan Krios (Thermo Fisher Scientific) equipped with K3 detector (Gatan). 18,945 movies were collected, using data acquisition software EPU (Thermo Fisher Scientific), at a magnification of 105,000 and a pixel size of 0.85 Å using a total dose of 50 e<sup>-</sup>/Å<sup>2</sup>. These were motion corrected in 5 × 5 patches using MOTIONCOR2<sup>100</sup> implemented in RELION4.0<sup>44</sup> before importing into cryoSPARC<sup>43</sup>. Micrographs were CTF corrected prior to blob picking and extraction of an initial set of particles. Subsequent filtering by 2D classification removed most of these leaving a cleaned set of particles that were used to produce 4 initial models. The 2D classes and initial models appeared to have at least two distinct subpopulations, one containing just the GINS complex alone and one that also had some additional density. In order to distinguish between these two sets of particles heterorefinement was performed to split the two groups. The first population of particles containing only the GINS complex were then imported into RELION4.0 before several cycles of 3D classification, to yield a final set of 111,455 particles that gave a volume at 5.9 Å resolution used to produce the GINS alone volume used in (Supplementary Fig. 16a). The second group of TOPBP1-GINS particles were used to train a TOPAZ model before implementing further picking using TOPAZ<sup>101</sup>. The resulting particles were again filtered by rounds of 2D classification and heterorefinement before exporting the particles, that appeared to consist of the full TOPBP1-GINS complex, into RELION4.0. Several cycles of 3D classification were then performed with classes appearing to lack the additional density eliminated, to yield a final set of 208,115 particles. 3D refinement and post-processing of these gave a volume at 4.71 Å resolution into which previously solved crystal structures of the GINS complex (29EX) and the central BRCT4/5 module of TopBP1 (3UEN) could be docked.

For dataset 2, grids were stored in liquid nitrogen prior to imaging at 300 kV on a FEI Titan Krios (Thermo Fisher Scientific) equipped with Fakcon 4i detector (Thermo Fisher Scientific). 8808 movies were collected, using data acquisition software EPU (Thermo Fisher Scientific), at a magnification of 120,000 and a pixel size of 0.74 Å using a total dose of 39.69 e<sup>-</sup>/Å<sup>2</sup>. The same processing schedule was followed as for dataset one to yield a final set of 154,278 particles. 3D refinement and post-processing of these gave a volume at 4.1 Å resolution into which the previously solved model could be docked alongside the predicted GINI helix from the AlphaFold2-Multimer model. Minor adjustments to the model were made in Coot before refinement of the final model using PHENIX<sup>102</sup>.

### Data presentation

Figure panels for all structural data were produced using ChimeraX-1.3.

### Reporting summary

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

### Data availability

The raw mass spectrometry proteomics data for the proximity-dependent biotinylation experiments, the cross-linking experiments and the chromatin mass spectrometry (CHROMASS) experiments generated in this study, have been deposited to the ProteomeXchange Consortium via the PRIDE<sup>103</sup> partner repository (<https://www.ebi.ac.uk/pride/archive/>) with the dataset identifiers PXD040000 (PDB experiments), PXD040156 (CL-MS experiments) and PXD040024 (CHROMASS experiments). Cryo-EM maps and refined coordinates generated in this study have been deposited in the EMDB and Protein Databank with the identifiers EMD-16916 and PDB ID 8OK2, respectively. Published Protein Database entries used in this study: 2E9X (GINS crystal structure); 7PFO (core human replisome); 3UEN (TopBP1-BRCT4/5 crystal structure); 6XTX (human DNA-bound CMG). Source data are provided with this paper.

## References

- Siddiqui, K., On, K. F. & Diffley, J. F. Regulating DNA replication in eukarya. *Cold Spring Harb. Perspect. Biol.* **5**, a010108 (2013).
- Remus, D. et al. Concerted loading of Mcm2-7 double hexamers around DNA during DNA replication origin licensing. *Cell* **139**, 719–730 (2009).
- Evryn, C. et al. A double-hexameric MCM2-7 complex is loaded onto origin DNA during licensing of eukaryotic DNA replication. *Proc. Natl Acad. Sci. USA* **106**, 20240–20245 (2009).
- Li, N. et al. Structure of the eukaryotic MCM complex at 3.8 Å. *Nature* **524**, 186–191 (2015).
- Georgescu, R. et al. Structure of eukaryotic CMG helicase at a replication fork and implications to replisome architecture and origin initiation. *Proc. Natl Acad. Sci. USA* **114**, E697–E706 (2017).
- Douglas, M. E., Ali, F. A., Costa, A. & Diffley, J. F. X. The mechanism of eukaryotic CMG helicase activation. *Nature* **555**, 265–268 (2018).
- Moyer, S. E., Lewis, P. W. & Botchan, M. R. Isolation of the Cdc45/Mcm2-7/GINS (CMG) complex, a candidate for the eukaryotic DNA replication fork helicase. *Proc. Natl Acad. Sci. USA* **103**, 10236–10241 (2006).
- Ilves, I., Petojevic, T., Pesavento, J. J. & Botchan, M. R. Activation of the MCM2-7 helicase by association with Cdc45 and GINS proteins. *Mol. Cell* **37**, 247–258 (2010).
- Costa, A. et al. DNA binding polarity, dimerization, and ATPase ring remodeling in the CMG helicase of the eukaryotic replisome. *Elife* **3**, e03273 (2014).
- Yeeles, J. T., Deegan, T. D., Janska, A., Early, A. & Diffley, J. F. Regulated eukaryotic DNA replication origin firing with purified proteins. *Nature* **519**, 431–435 (2015).
- Lewis, J. S. et al. Mechanism of replication origin melting nucleated by CMG helicase assembly. *Nature* **606**, 1007–1014 (2022).
- Meagher, M., Epling, L. B. & Enemark, E. J. DNA translocation mechanism of the MCM complex and implications for replication initiation. *Nat. Commun.* **10**, 3117 (2019).
- Kanemaki, M. & Labib, K. Distinct roles for Sld3 and GINS during establishment and progression of eukaryotic DNA replication forks. *EMBO J.* **25**, 1753–1763 (2006).
- Heller, R. C. et al. Eukaryotic origin-dependent DNA replication in vitro reveals sequential action of DDK and S-CDK kinases. *Cell* **146**, 80–91 (2011).
- Francis, L. I., Randell, J. C., Takara, T. J., Uchima, L. & Bell, S. P. Incorporation into the prereplicative complex activates the Mcm2-7 helicase for Cdc7-Dbf4 phosphorylation. *Genes Dev.* **23**, 643–654 (2009).
- Greiwe, J. F. et al. Structural mechanism for the selective phosphorylation of DNA-loaded MCM double hexamers by the Dbf4-dependent kinase. *Nat. Struct. Mol. Biol.* **29**, 10–20 (2022).
- Saleh, A. et al. The structural basis of Cdc7-Dbf4 kinase dependent targeting and phosphorylation of the MCM2-7 double hexamer. *Nat. Commun.* **13**, 2915 (2022).
- Itou, H., Muramatsu, S., Shirakihara, Y. & Araki, H. Crystal structure of the homology domain of the eukaryotic DNA replication proteins Sld3/Treslin. *Structure* **22**, 1341–1347 (2014).
- Deegan, T. D., Yeeles, J. T. & Diffley, J. F. Phosphopeptide binding by Sld3 links Dbf4-dependent kinase to MCM replicative helicase activation. *EMBO J.* **35**, 961–973 (2016).
- Zegerman, P. & Diffley, J. F. X. Phosphorylation of Sld2 and Sld3 by cyclin-dependent kinases promotes DNA replication in budding yeast. *Nature* **445**, 281–285 (2007).
- Kumagai, A., Shevchenko, A. & Dunphy, W. G. Direct regulation of Treslin by cyclin-dependent kinase is essential for the onset of DNA replication. *J. Cell Biol.* **193**, 995–1007 (2011).

22. Boos, D. et al. Regulation of DNA replication through Sld3-Dpb11 interaction is conserved from yeast to humans. *Curr. Biol.* **21**, 1152–1157 (2011).
23. Tak, Y. S., Tanaka, Y., Endo, S., Kamimura, Y. & Araki, H. A CDK-catalysed regulatory phosphorylation for formation of the DNA replication complex Sld2-Dpb11. *EMBO J.* **25**, 1987–1996 (2006).
24. Gambus, A. et al. GINS maintains association of Cdc45 with MCM in replisome progression complexes at eukaryotic DNA replication forks. *Nat. Cell Biol.* **8**, 358–366 (2006).
25. Sengupta, S., van Deursen, F., de Piccoli, G. & Labib, K. Dpb2 integrates the leading-strand DNA polymerase into the eukaryotic replisome. *Curr. Biol.* **23**, 543–552 (2013).
26. Goswami, P. et al. Structure of DNA-CMG-Pol epsilon elucidates the roles of the non-catalytic polymerase modules in the eukaryotic replisome. *Nat. Commun.* **9**, 5061 (2018).
27. Jones, M. L., Baris, Y., Taylor, M. R. G. & Yeeles, J. T. P. Structure of a human replisome shows the organisation and interactions of a DNA replication machine. *EMBO J.* **40**, e108819 (2021).
28. Xu, Z. et al. Synergism between CMG helicase and leading strand DNA polymerase at replication fork. *Nat. Commun.* **14**, 5849 (2023).
29. Day, M., Oliver, A. W. & Pearl, L. H. Phosphorylation-dependent assembly of DNA damage response systems and the central roles of TOPBP1. *DNA Repair* **108**, 103232 (2021).
30. Makiniemi, M. et al. BRCT domain-containing protein TopBP1 functions in DNA replication and damage response. *J. Biol. Chem.* **276**, 30399–30406 (2001).
31. Kumagai, A., Shevchenko, A. & Dunphy, W. G. Treslin collaborates with TopBP1 in triggering the initiation of DNA replication. *Cell* **140**, 349–359 (2010).
32. Tanaka, S. et al. Efficient initiation of DNA replication in eukaryotes requires Dpb11/TopBP1-GINS interaction. *Mol. Cell Biol.* **33**, 2614–2622 (2013).
33. Mantiero, D., Mackenzie, A., Donaldson, A. & Zegerman, P. Limiting replication initiation factors execute the temporal programme of origin firing in budding yeast. *EMBO J.* **30**, 4805–4814 (2011).
34. Tanaka, S., Nakato, R., Katou, Y., Shirahige, K. & Araki, H. Origin association of Sld3, Sld7, and Cdc45 proteins is a key step for determination of origin-firing timing. *Curr. Biol.* **21**, 2055–2063 (2011).
35. Collart, C., Allen, G. E., Bradshaw, C. R., Smith, J. C. & Zegerman, P. Titration of four replication factors is essential for the *Xenopus laevis* midblastula transition. *Science* **341**, 893–896 (2013).
36. Jenkinson, F., Tan, K. W., Schopf, B., Santos, M. M. & Zegerman, P. Dephosphorylation of the pre-initiation complex is critical for origin firing. *Mol. Cell* **83**, 12–25 e10 (2023).
37. Lim, Y. et al. In silico protein interaction screening uncovers DONSON's role in replication initiation. *Science* **381**, eadi3448 (2023).
38. Xia, Y. et al. DNSN-1 recruits GINS for CMG helicase assembly during DNA replication initiation in *Caenorhabditis elegans*. *Science* **381**, eadi4932 (2023).
39. Kingsley, G. et al. DONSON facilitates Cdc45 and GINS chromatin association and is essential for DNA replication initiation. *Nucleic Acids Res.* **51**, 9748–9763 (2023).
40. Hashimoto, Y., Sadano, K., Miyata, N., Ito, H. & Tanaka, H. Novel role of DONSON in CMG helicase assembly during vertebrate DNA replication initiation. *EMBO J.* **42**, e114131 (2023).
41. Cvetkovic, M. A. et al. The structural mechanism of dimeric DONSON in replicative helicase activation. *Mol. Cell* **83**, 4017–4031.e9 (2023).
42. Bigot, N. et al. Phosphorylation-mediated interactions with TOPBP1 couple 53BP1 and 9-1-1 to control the G1 DNA damage checkpoint. *Elife* **8**, e44353 (2019).
43. Punjani, A., Rubinstein, J. L., Fleet, D. J. & Brubaker, M. A. cryo-SPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nat. Methods* **14**, 290–296 (2017).
44. Leung, C. C. et al. Structural insights into recognition of MDC1 by TopBP1 in DNA replication checkpoint control. *Structure* **21**, 1450–1459 (2013).
45. Raschle, M. et al. DNA repair. Proteomics reveals dynamic assembly of repair complexes during bypass of DNA cross-links. *Science* **348**, 1253671 (2015).
46. McGarry, T. J. & Kirschner, M. W. Geminin an inhibitor of DNA replication, is degraded during mitosis. *Cell* **93**, 1043–1053 (1998).
47. Kamada, K., Kubota, Y., Arata, T., Shindo, Y. & Hanaoka, F. Structure of the human GINS complex and its assembly and functional interface in replication initiation. *Nat. Struct. Mol. Biol.* **14**, 388–396 (2007).
48. Chang, Y. P., Wang, G., Bermudez, V., Hurwitz, J. & Chen, X. S. Crystal structure of the GINS complex and functional insights into its role in DNA replication. *Proc. Natl Acad. Sci. USA* **104**, 12685–12690 (2007).
49. De Jesus-Kim, L. et al. DDK regulates replication initiation by controlling the multiplicity of Cdc45-GINS binding to Mcm2-7. *Elife* **10**, e65471 (2021).
50. Looke, M., Maloney, M. F. & Bell, S. P. Mcm10 regulates DNA replication elongation by stimulating the CMG replicative helicase. *Genes Dev.* **31**, 291–305 (2017).
51. Douglas, M. E. & Diffley, J. F. X. Recruitment of Mcm10 to sites of replication initiation requires direct binding to the minichromosome maintenance (MCM) complex. *J. Biol. Chem.* **291**, 5879–5888 (2016).
52. Costa, A. & Diffley, J. F. X. The Initiation of eukaryotic DNA replication. *Annu. Rev. Biochem.* **91**, 107–131 (2022).
53. Zegerman, P. & Diffley, J. F. X. Checkpoint-dependent inhibition of DNA replication initiation by Sld3 and Dbf4 phosphorylation. *Nature* **467**, 474–478 (2010).
54. Lopez-Mosqueda, J. et al. Damage-induced phosphorylation of Sld3 is important to block late origin firing. *Nature* **467**, 479–483 (2010).
55. Natsume, T. et al. Kinetochores coordinate pericentromeric cohesion and early DNA replication by Cdc7-Dbf4 kinase recruitment. *Mol. Cell* **50**, 661–674 (2013).
56. Reuswig, K. U., Zimmermann, F., Galanti, L. & Pfander, B. Robust replication control is generated by temporal gaps between licensing and firing phases and depends on degradation of firing factor Sld2. *Cell Rep.* **17**, 556–569 (2016).
57. Tanaka, S. Interaction of replication factor Sld3 and histone acetyltransferase Esa1 alleviates gene silencing and promotes the activation of late and dormant replication origins. *Genetics* **217**, 1–11 (2021).
58. Sansam, C. G. et al. A mechanism for epigenetic control of DNA replication. *Genes Dev.* **32**, 224–229 (2018).
59. Wittig, K. A., Sansam, C. G., Noble, T. D., Goins, D. & Sansam, C. L. The CRL4DTL E3 ligase induces degradation of the DNA replication initiation factor TICRR/TRESLIN specifically during S phase. *Nucleic Acids Res.* **49**, 10507–10523 (2021).
60. Johnson, M. C., Can, G., Santos, M. M., Alexander, D. & Zegerman, P. Checkpoint inhibition of origin firing prevents inappropriate replication outside of S-phase. *Elife* **10**, e63589 (2021).
61. Ferreira, P. et al. MTBP phosphorylation controls DNA replication origin firing. *Sci. Rep.* **11**, 4242 (2021).
62. Ciardo, D. et al. Polo-like kinase 1 (Plk1) regulates DNA replication origin firing and interacts with Rif1 in *Xenopus*. *Nucleic Acids Res.* **49**, 9851–9869 (2021).
63. Bohly, N. et al. Increased replication origin firing links replication stress to whole chromosomal instability in human cancer. *Cell Rep.* **41**, 111836 (2022).

64. Toledo, L. I. et al. ATR prohibits replication catastrophe by preventing global exhaustion of RPA. *Cell* **155**, 1088–1103 (2013).
65. Helmrich, A., Ballarino, M., Nudler, E. & Tora, L. Transcription-replication encounters, consequences and genomic instability. *Nat. Struct. Mol. Biol.* **20**, 412–418 (2013).
66. Rhind, N. & Gilbert, D. M. DNA replication timing. *Cold Spring Harb. Perspect. Biol.* **5**, a010132 (2013).
67. Boos, D. & Ferreira, P. Origin firing regulations to control genome replication timing. *Genes* **10**, 199 (2019).
68. Burrell, R. A. et al. Replication stress links structural and numerical cancer chromosomal instability. *Nature* **494**, 492–496 (2013).
69. Volpi, I., Gillespie, P. J., Chadha, G. S. & Blow, J. J. The role of DDK and Treslin-MTBP in coordinating replication licensing and pre-initiation complex formation. *Open Biol.* **11**, 210121 (2021).
70. Gaggioli, V. et al. CDK phosphorylation of SLD-2 is required for replication initiation and germline development in *C. elegans*. *J. Cell Biol.* **204**, 507–522 (2014).
71. Fukuura, M. et al. CDK promotes interactions of Sld3 and Drc1 with Cut5 for initiation of DNA replication in fission yeast. *Mol. Biol. Cell* **22**, 2620–2633 (2011).
72. Muramatsu, S., Hirai, K., Tak, Y. S., Kamimura, Y. & Araki, H. CDK-dependent complex formation between replication proteins Dpb11, Sld2, Pol  $\epsilon$ , and GINS in budding yeast. *Genes Dev.* **24**, 602–612 (2010).
73. Rappas, M., Oliver A. W. & Pearl L. H. Structure and function of the Rad9-binding region of the DNA-damage checkpoint adaptor TopBP1. *Nucleic Acids Res.* **39**, 313–324 (2011).
74. Ferreira, P., Sanchez-Pulido, L., Marko, A., Ponting, C. P. & Boos, D. Refining the domain architecture model of the replication origin firing factor Treslin/TICRR. *Life Sci. Alliance* **5**, e202101088 (2022).
75. Chou, D. M., Petersen, P., Walter, J. C. & Walter, G. Protein phosphatase 2A regulates binding of Cdc45 to the prereplication complex. *J. Biol. Chem.* **277**, 40520–40527 (2002).
76. Kohler, K. et al. The Cdk8/19-cyclin C transcription regulator functions in genome replication through metazoan Sld7. *PLoS Biol.* **17**, e2006767 (2019).
77. Zimmerman, K. M., Jones, R. M., Petermann, E. & Jeggo, P. A. Diminished origin-licensing capacity specifically sensitizes tumor cells to replication stress. *Mol. Cancer Res.* **11**, 370–380 (2013).
78. Rainey, M. D., Quachthithu, H., Gaboriau, D. & Santocanale, C. DNA replication dynamics and cellular responses to ATP competitive CDC7 kinase inhibitors. *ACS Chem. Biol.* **12**, 1893–1902 (2017).
79. Gillespie, P. J., Gambus, A. & Blow, J. J. Preparation and use of *Xenopus* egg extracts to study DNA replication and chromatin associated proteins. *Methods* **57**, 203–213 (2012).
80. Tyanova, S. et al. The Perseus computational platform for comprehensive analysis of (prote)omics data. *Nat. Methods* **13**, 731–740 (2016).
81. Rappsilber, J., Mann, M. & Ishihama, Y. Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. *Nat. Protoc.* **2**, 1896–1906 (2007).
82. Olsen, J. V. et al. Parts per million mass accuracy on an Orbitrap mass spectrometer via lock mass injection into a C-trap. *Mol. Cell Proteom.* **4**, 2010–2021 (2005).
83. Cox, J. et al. Andromeda: a peptide search engine integrated into the MaxQuant environment. *J. Proteome Res.* **10**, 1794–1805 (2011).
84. Cox, J. & Mann, M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat. Biotechnol.* **26**, 1367–1372 (2008).
85. Cox, J. et al. Accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction, termed MaxLFQ. *Mol. Cell Proteom.* **13**, 2513–2526 (2014).
86. Hughes, C. S. et al. Single-pot, solid-phase-enhanced sample preparation for proteomics experiments. *Nat. Protoc.* **14**, 68–85 (2019).
87. Lu, L. et al. Identification of MS-cleavable and noncleavable chemically cross-linked peptides with metaMorpheus. *J. Proteome Res.* **17**, 2370–2376 (2018).
88. Chen, Z. L. et al. A high-speed search engine pLink 2 with systematic evaluation for proteome-scale identification of cross-linked peptides. *Nat. Commun.* **10**, 3404 (2019).
89. Riffle, M., Jaschob, D., Zelter, A. & Davis, T. N. ProXL (protein cross-linking database): a platform for analysis, visualization, and sharing of protein cross-linking mass spectrometry data. *J. Proteome Res.* **15**, 2863–2870 (2016).
90. Evans, R. et al. Protein complex prediction with AlphaFold-Multimer. Preprint at *bioRxiv* <https://doi.org/10.1101/2021.10.04.463034> (2022).
91. Jumper, J. et al. Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589 (2021).
92. Kozakov, D. et al. The ClusPro web server for protein-protein docking. *Nat. Protoc.* **12**, 255–278 (2017).
93. Zhou, P., Jin, B., Li, H. & Huang, S. Y. HPEPDOCK: a web server for blind peptide-protein docking based on a hierarchical algorithm. *Nucleic Acids Res.* **46**, W443–W450 (2018).
94. Yu, C. et al. A mechanism for preventing asymmetric histone segregation onto replicating DNA strands. *Science* **361**, 1386–1389 (2018).
95. Xu, X., Yan, C. & Zou, X. MDockPeP: an ab-initio protein-peptide docking server. *J. Comput. Chem.* **39**, 2409–2413 (2018).
96. Yan, C., Xu, X. & Zou, X. Fully blind docking at the atomic level for protein-peptide complex structure prediction. *Structure* **24**, 1842–1853 (2016).
97. Raveh, B., London, N. & Schueler-Furman, O. Sub-angstrom modeling of complexes between flexible peptides and globular proteins. *Proteins* **78**, 2029–2040 (2010).
98. Xue, L. C., Rodrigues, J. P., Kastiris, P. L., Bonvin, A. M. & Vangone, A. PRODIGY: a web server for predicting the binding affinity of protein-protein complexes. *Bioinformatics* **32**, 3676–3678 (2016).
99. Romero-Molina, S. et al. PPI-affinity: a web tool for the prediction and optimization of protein-peptide and protein-protein binding affinity. *J. Proteome Res.* **21**, 1829–1841 (2022).
100. Zheng, S. Q. et al. MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat. Methods* **14**, 331–332 (2017).
101. Bepler, T. et al. Positive-unlabeled convolutional neural networks for particle picking in cryo-electron micrographs. *Nat. Methods* **16**, 1153–1160 (2019).
102. Liebschner, D. et al. Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. *Acta Crystallogr D. Struct. Biol.* **75**, 861–877 (2019).
103. Vizcaino, J. A. et al. 2016 update of the PRIDE database and its related tools. *Nucleic Acids Res.* **44**, D447–456 (2016).

## Acknowledgements

We thank Nora Cronin, Rebecca Thompson and Daniel Maskell, Yehuda Halfon, Pascale Schelleberger and Fabienne Beuron for assistance with cryo-EM. We also thank Svenja Heimann and Jenny Bormann for technical assistance with cross-linking and APEX2 mass spectrometry and Ken L. Dreger for our local ProXL server. We further thank Stefan Westermann, Hemmo Meyer and Olaf Stemmann for sharing reagents and expertise on biochemistry and *Xenopus* egg extracts. The lab of D.B. was supported by the SFB1430 (Project-ID 424228829, FOR2800) and RTG1739 of the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) and the NRW Rückkehrerförderprogramm (state of North-Rhine-Westphalia), the labs of A.W.O and L.H.P. by Cancer Research UK Programme Grants C302/A14532 and C302/A24386, and the lab of M.R. by the FOR2800 research group (DFG). E.S.G. acknowledges funding from the DFG under Germany's Excellence Strategy (EXC 2033 – 390677874 – RESOLV), the instrumentation programme of the

DFG (436586093) and CRC1430 (Project-ID 424228829). M.K. and F.K. are grateful for funding by the DFG (DFG, INST 20876/322-1 FUGG and CRC1430).

### Author contributions

M.D. performed cryo-EM and biochemical experiments with assistance by I.A.B. B.T. performed biochemical binding studies by pull-down, most mutation analyses, some *Xenopus* egg extract experiments, and Colab AlphaFold predictions that identified the GINI binding site in GINS. M.P. performed experiments using *Xenopus* egg extracts and proximity biotinylation experiments. A.M. generated many point mutant cDNAs and assisted in experiments by B.T. and M.P. Y.H.A. and E.S.G. performed computational modelling and AlphaFold2-multimer predictions and analyses. H.S. performed and analysed cross-linking-mass spectrometry under supervision of B.T. and including mass spectrometry by F.K. and M.K. M.R. performed CHROMASS experiments together with M.P. F.K. and M.K. performed mass spectrometry for proximity biotinylation and cross-linking experiments. A.O.W., L.H.P., M.D. and D.B. conceptualised and supervised the project. D.B. and M.D. wrote the manuscript. D.B. coordinated the work and performed individual revision experiments.

### Funding

Open Access funding enabled and organized by Projekt DEAL.

### 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/s41467-024-45946-0>.

**Correspondence** and requests for materials should be addressed to Matthew Day, Laurence H. Pearl, Antony W. Oliver or Dominik Boos.

**Peer review information** *Nature Communications* thanks the anonymous reviewers for their contribution to the peer review of this work. A peer review file is 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 licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence 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 licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024