# TopBP1 loads GINS onto the Metazoan Mcm2-7 Helicase during Replication Origin Firing

Inaugural Dissertation

for

the doctoral degree of

Dr. rer. nat.

from the Faculty of Biology

University of Duisburg-Essen

Submitted by

Milena Yasemin Jeanne Parlak

Born in Seeheim-Jugenheim

June 2023

**DOI:** 10.17185/duepublico/81308
**URN:** urn:nbn:de:hbz:465-20231214-062208-9

In the context of this doctoral work, the following article has been submitted:

Day M[*], Tetik B[*], Parlak M[*], Almeida-Hernandez Y, Sanchez-Garcia E, Kaschani F, Räschle M, Siegert H, Pearl L, Oliver A, Boos D
*Bipartite GINS binding mode of TopBP1 to activate the replicative helicase.*
bioRxiv (2023) doi: https://doi.org/10.1101/2023.03.31.535063
Under revision at Nature Communications

[*] equal contribution

The experiments underlying the present work were conducted at the Faculty of Biology at the Department of Molecular Genetics II at the University of Duisburg-Essen.

1. Examiner: Prof. Dr. Dominik Boos
2. Examiner: Prof. Dr. Hemmo Meyer
3. Examiner:

Chair of the Board of Examiners: Prof. Dr. Doris Hellerschmied-Jelinek

Date of the oral examination: 27.10.2023

# Table of Content

# Summary

During replication the genome must be duplicated accurately and exactly once per cell cycle to guarantee genetic homeostasis. The first step, called replication initiation, regulates when exactly and where in the genome DNA replication occurs. These regulatory steps and the essential proteins involved in this process are well-characterized in yeast. For the tight control of replication, initiation is divided into a two-step process that separates the loading of the replicative helicase in G1 phase from its activation in S phase of the cell cycle. During the first step, origin licensing, the Mcm2-7 helicase is loaded onto origin DNA and thereby forms the pre-replicative complex (pre-RC). Origin firing, the second step of replication initiation, mediates the conversion of the inactive pre-RC into the active Cdc45-Mcm2-7-GINS (CMG) helicase. The major regulatory step in this conversion is the assembly of an intermediate complex, the so-called pre-initiation complex (pre-IC). During pre-IC assembly the essential helicase components Cdc45 and GINS are recruited to the Mcm2-7 helicase. In yeast, Cdc45 recruitment to the helicase is mediated by the Sld3-Sld7 complex (metazoan Treslin-MTBP) and GINS is delivered as part of the so-called pre-loading complex (pre-LC). Although most of the core replication factors are conserved from yeast to humans, the molecular details of Cdc45 and GINS loading onto the pre-RC are not yet known in metazoans. Here we demonstrate that the essential firing factor TopBP1 interacts with GINS through two individual binding surfaces during S phase. TopBP1 is a multi-BRCT scaffold protein that fulfils multiple functions in metazoan cells. In particular, TopBP1 plays dual roles in replication origin firing and in the response to DNA damage. Our findings provide molecular details on how TopBP1 contributes to helicase activation during origin firing.

In this thesis we show that the previously described GINS interaction (GINI) domain of TopBP1 is highly conserved between metazoans. By protein sequence alignment between metazoan species, we narrow down the GINI domain to a conserved region between TopBP1 BRCT3 and BRCT4. Systematical mutations of this domain demonstrate that the GINI region is required for GINS binding *in vitro* but has only a moderate effect on replication in *Xenopus* egg extract. Structural analysis suggests that the GINI domain forms a short α helix. By integration of a mutation that breaks this

helix we show that the helix provides a binding surface for GINS. Furthermore, we verify the corresponding binding residues in the distal part of the A-domain of the GINS subunit Psf1. Interestingly, all GINI mutants are incapable to bind to GINS *in vitro* but still support replication in egg extract to relatively high amounts. This controversial observation suggests that a second GINS binding motif is present in TopBP1, compensating for the loss of the GINI domain function. Indeed, we identify a second GINS interaction surface in the BRCT4 domain of TopBP1. This TopBP1 BRCT4 domain binds to the base of the Psf1 A-domain and to the linker between the A- and the B-domain. Functional analysis demonstrates that simultaneous mutations in both GINS interaction sites in TopBP1 lead to severe defects of replication in *Xenopus* egg extract, suggesting a functional cooperation between the two GINS binding surfaces in TopBP1.

Within the active replisome the DNA polymerase ε (Polε) associates with the CMG helicase whereas TopBP1, Treslin and MTBP leave the complex before DNA synthesis starts. Our structural data suggests that simultaneous binding of GINS to TopBP1 BRCT4 and to the Polε subunit PolE2 is mutually exclusive. Therefore, we propose a model that once TopBP1 and GINS are associated with the helicase the B-domain of GINS Psf1 undergoes a conformational change that favors PolE2 binding over BRCT4 binding. We hypothesize that low-affinity GINS binding to the only remaining GINS interaction surface in TopBP1, induces TopBP1 ejection, and allows recycling of this limiting firing factor.

# Zusammenfassung

Um die genetische Homöostase zu gewährleisten, muss das Genom in jedem Zellzyklus komplett und exakt einmal repliziert werden. Der erste Schritt der Replikation ist die Replikations-Initiation. Diese reguliert, wann genau und wo im Genom die Replikation der DNA stattfindet. Die hier zugrunde liegenden Regulationsmechanismen sowie die Beteiligung essenzieller Proteine sind im Modellorganismus Hefe ausführlich charakterisiert. Für eine strikte Kontrolle der Replikation ist diese in einen zweistufigen Prozess unterteilt, welche das Laden der Helikase auf die DNA in der G1 Phase von ihrer Aktivierung in der S Phase trennt. Im ersten Schritt, der sogenannten Lizensierung, wird die Mcm2-7 Helikase auf die Replikationsursprünge geladen und bildet dadurch den Prä-Replikationskomplex (*pre-replicative complex*; pre-RC). Das Feuern der Replikationsursprünge im zweiten Schritt, vermittelt die Umwandlung des inaktiven pre-RC in die aktive Cdc45-Mcm2-7-GINS (CMG) Helikase. Der wichtigste regulatorische Schritt bei dieser Umwandlung ist der Aufbau eines Zwischenkomplexes, des sogenannten Prä-Initiationskomplexes (*pre-initiation complex*; pre-IC). Während der Assemblierung des pre-IC werden die essenziellen Helikase Komponenten Cdc45 und GINS zur Mcm2-7 Helikase rekrutiert. Aus der Hefe ist bekannt, dass die Rekrutierung von Cdc45 zu der Helikase durch den Sld3-Sld7 (Metazoa Treslin-MTBP) Proteinkomplex vermittelt wird. Des Weiteren wird GINS als Teil des Prä-Ladekomplexes (*pre-loading complex; pre-LC*) auf die Helikase geladen. Obwohl die meisten Kernreplikationsfaktoren konserviert in Metazoa-Zellen vorliegen, sind die molekularen Details der Cdc45- und GINS-Beladung auf den pre-RC noch nicht bekannt. Im Rahmen der vorliegenden Arbeit wird gezeigt, dass der essenzielle Feuerungsfaktor TopBP1 über zwei Bindungsoberflächen mit GINS interagiert. Das multi-BRCT-Gerüstprotein TopBP1 erfüllt in Metazoa-Zellen mehrere Funktionen. Insbesondere spielt TopBP1 duale Rollen beim Feuern der Replikationsursprünge und bei der Reaktion auf DNA-Schäden. Unsere Ergebnisse liefern molekulare Details darüber, wie TopBP1 zur Helikase-Aktivierung während des Ursprungsfeuers beiträgt.

Wir zeigen, dass die zuvor beschriebene GINS-Interaktions- (GINI) Domäne zwischen Metazoen hoch konserviert vorliegt. Durch Sequenzvergleich zwischen Metazoa Arten

konnten wir die GINI-Domäne auf eine konservierte Region zwischen TopBP1 BRCT3 und BRCT4 eingrenzen. Systematische Mutationen dieser Domäne zeigen, dass die GINI-Region für die GINS-Bindung *in vitro* erforderlich ist, aber nur einen moderaten Einfluss auf die Replikation im *Xenopus* Eiextrakt hat. Unsere Strukturanalysen legen nahe, dass die GINI-Domäne eine kurze α-Helix bildet. Durch die Integration einer Mutation, die zum Bruch dieser Helix führt, kann gezeigt werden, dass die Helix eine Bindungsoberfläche für GINS bildet. Darüber hinaus verifizieren wir die entsprechenden Bindungsreste im distalen Teil der A-Domäne der GINS-Untereinheit Psf1. Interessanterweise sind alle GINI-Mutanten nicht in der Lage, *in vitro* an GINS zu binden, unterstützen aber dennoch die Replikation in Eiextrakt in relativ hohem Maße. Diese kontroverse Beobachtung legt nahe, dass TopBP1 eine zweite Bindestelle für GINS enthält, die den funktionellen Verlust der GINI Domäne ausgleichen kann. Wir zeigen hier zum ersten Mal, dass die TopBP1 BRCT4 Domäne eine solche zweite GINS-Interaktionsoberfläche enthält. Die BRCT4 Domäne bindet hierbei an die Basis der Psf1 A-Domäne und an den Linker zwischen der A- und der B-Domäne. Unsere Funktionsanalysen zeigen, dass gleichzeitige Mutationen in beiden GINS-Interaktionsstellen in TopBP1 zu schwerwiegenden Replikationsdefekten im *Xenopus* Eiextrakt führen, was auf eine funktionelle Zusammenarbeit zwischen den beiden GINS-Bindungsoberflächen in TopBP1 schließen lässt.

Innerhalb des aktiven Replisoms assoziiert die DNA-Polymerase ε (Polε) mit der CMG-Helikase, wohingegen TopBP1, Treslin und MTBP den Komplex verlassen, bevor die DNA-Synthese beginnt. Unsere Strukturdaten legen nahe, dass sich die gleichzeitige Bindung von GINS an TopBP1 BRCT4 und an die Polε Untereinheit PolE2 gegenseitig ausschließt. Wir stellen daher ein Modell vor, bei dem die B-Domäne von GINS Psf1, sobald TopBP1 und GINS mit der Helikase assoziiert sind, eine Konformationsänderung erfährt, die die Bindung von PolE2 gegenüber der Bindung von BRCT4 bevorzugt. Wir nehmen an, dass die Bindung von GINS mit geringer Affinität an die einzige verbleibende GINS-Interaktionsoberfläche in TopBP1 den Auswurf von TopBP1 induziert. Dies könnte ein Mechanismus sein, der ein Recycling des limitierenden Feuerungsfaktors TopBP1 ermöglicht.

# Abbreviations

| | | |
|---|---|---|
| **A** | APC/C | anaphase-promoting complex/cyclosome |
| | APS | ammonium persulfate |
| | ATP | adenosine triphosphate |
| **B** | BRCA1 | breast cancer gene 1 |
| | BRCT | BRCA1 C-terminus |
| **C** | C-terminus | carboxyl-terminus |
| | CDK | cyclin-dependent kinase |
| | CMG | Cdc45-Mcm2-7-GINS |
| | CMGE | Cdc45-Mcm2-7-GINS-Polε |
| | ctrl | control |
| **D** | DAPI | 4', 6'-diamidino-2-phenylindole |
| | DDK | Dbf4-dependent kinase |
| | DMEM | Dulbecco's Modified Eagle Medium |
| | DMSO | dimethyl sulfoxide |
| | DNA | desoxyribonucleic acid |
| | dsDNA | double stranded DNA |
| | DTT | dithiothreitol |
| **F** | FCS | fetal calf serum |
| | Fkh | Forkhead |
| **G** | GINI | GINS interaction |
| | GINS | Go-Ichi-Ni-San |
| **H** | HEPS | 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid |
| **I** | IPTG | isopropyl-β-D-thiogalactopyranosid |
| **K** | kDa | kilodalton |
| | kb | kilobase |
| **L** | LB | lysogeny broth |
| **M** | Mcm | minichromosome maintenance protein complex |
| | ml | milliliter |

| | | |
|---|---|---|
| | mM | millimolar |
| **N** | N-terminus | amino-terminus |
| | nM | nanomolar |
| **P** | PBS | phosphate buffered saline |
| | PCNA | proliferating cell nuclear antigen |
| | PCR | polymerase chain reaction |
| | PMSF | phenylmethylsulfonyl fluoride |
| | Pol | polymerase |
| | pRb | retinoblastoma |
| | Pre-IC | pre-initiation complex |
| | Pre-LC | pre-loading complex |
| | Pre-RC | pre-replicative complex |
| **R** | rpm | rounds per minute |
| | RT | room temperature |
| **S** | SDS | sodium dodecyl sulfate |
| | SDS complex | Sld3/Sld7-Dpb11-Sld2 |
| | SDS-PAGE | sodium dodecyl sulfate polyacrylamide gel electrophoresis |
| | ssDNA | single stranded DNA |
| **T** | TBS | tris buffered saline |
| | TCA | trichloroacetic acid |
| | TEMED | tetramethylethylenediamine |
| | TMT complex | Treslin-MTBP-TopBP1 |
| | Tris | tris(hydroxymethyl)aminomethane |
| | TRP | temporal replication program |
| | TTR | timing transition region |
| **V** | v/v | volume percentage |
| | w/v | weight percentage |
| **W** | WT | wild type |
| **others** | µg | microgram |
| | µl | microliter |
| | µM | micromolar |
| | °C | degree Celsius |

# 1 Introduction

Complete and accurate genome duplication is essential in all living organisms to ensure genetic homeostasis. Genome duplication is divided into three steps: initiation, elongation, and termination. Replication initiation is a highly regulated process that guarantees precise duplication of the genome. This initial step leads to the establishment of two bidirectional replication forks at an origin of replication. During elongation, these replication forks bidirectionally synthesize new DNA. When two converging replication forks meet, the newly synthesized DNA is ligated into continuous strands, and the replisomes disassemble (termination). Replication initiation plays a crucial role for this molecular regulation and control.

Little is known of these molecular mechanisms in metazoan. The yeast model serves as a template for understanding replication initiation in eukaryotes. Replication initiation is a two-step process. The first step, origin licensing, involves loading the inactive replicative helicase onto the origin of replication. The activation of the helicase occurs during replication origin firing, the second step of replication initiation. Loading of Cdc45 and GINS, two helicase components, onto the Mcm2-7 helicase is key for this activation. The molecular details of helicase activation during origin firing in metazoan remain largely unknown. This research project provides a better molecular understanding of how the firing factor TopBP1 contributes to delivery of the helicase component GINS onto the replicative helicase.

## 1.1 Genome replication in higher eukaryotes

### 1.1.1 Basics of eukaryotic genome replication

Replication initiation is the initial step of genome replication. Each initiation event results in two bidirectional DNA replication forks in all domains of life, including viruses, prokaryotes, archaea, and eukaryotes. During the next step, replication elongation, the genomic DNA is copied by complementary base pairing to synthesize two semiconservative DNA double strands. Finally, replication termination takes place when two replisomes meet head-to-head. Consequently, replisomes disassemble, and the newly synthesized DNA is ligated into two continuous DNA double strands (dsDNA). Replication specifically occurs in S phase of the cell cycle. This coupling is

mediated by the rise of cyclin-dependent kinase (CDK) and Dbf4-dependent kinase (DDK) levels in S phase (details below).

Due to the large genome of eukaryotic cells, hundreds (yeast) or thousands (mammals) of origins are required to duplicate their genome during S phase. For this reason, it is crucial to ensure that no single inter-origin distance is too large to guarantee complete replication of the genome within one S phase (Hyrien et al., 2003). To prevent this so-called random gap problem, enough origins must fire, and they must be positioned close enough to guarantee duplication of the whole genome within the time of one S phase (Boos and Ferreira, 2019).

In addition to spatial regulation of replication through positioning of replication origins, higher eukaryotes have a particular complex temporal replication program (TRP) (Dimitrova and Gilbert, 1999; Leonard and Mechali, 2013). Since some firing factors are present in limiting amounts in somatic eukaryotic cells, not all origins can fire simultaneously (Collart et al., 2013; Mantiero et al., 2011; Tanaka et al., 2011a; Wong et al., 2011). Scattered replication throughout the genome would result in large inter-origin distances that cannot be rescued by an incoming opposing replication fork. The TRP therefore allows replication in so-called replication domains (Berezney et al., 2000; Nakamura et al., 1986). These replication domains consist of clusters of replicons. Each replicon, in turn, consists of approximately five origins that fire nearly synchronously. Replication domains become microscopically apparent as typical replication foci (Sporbert et al., 2002). Domain-wise replication could be a mechanism for locally efficient replication that minimizes isolated replication events. Isolated replication events would result in single replication forks that will not find opposing forks for termination. Single replication forks are at a high risk of stalling and collapsing, leaving behind single stranded DNA (ssDNA), prone to DNA damage and hence genome instability. To avoid such events, adjacent replication domains in the genome appear to replicate sequentially, indicating a domino-like model in which replication spreads from the first domain to the neighboring domains (Maya-Mendoza et al., 2010; Sadoni et al., 2004; Sporbert et al., 2002). When early and late replication domains are adjacent, they are separated by so-called timing transition regions (TTRs). These regions are indeed replicated unidirectionally, consistent with the two outermost forks travelling out of their domains (Berezney et al., 2000; Boos and Ferreira, 2019).

This simplified replication domino-like model suggests that mechanisms must exist to determine the timing of the first replication domain and mediate the spreading of

replication events that can be blocked at TTRs. Chromatin processes and parameters also appear to influence the timing of replication. An example of this is the replication of transcriptionally active open euchromatin early in S phase, whereas compact heterochromatin is replicates later (Boos and Ferreira, 2019; Marchal et al., 2019; Rhind, 2006; Yekezare et al., 2013). However, most of the molecular mechanisms and regulations that facilitate these processes are poorly understood.

### 1.1.2 Cell cycle regulation of DNA replication initiation

Eukaryotic cells have to ensure that the genomic information is copied accurately and precisely within the time of one cell cycle. Failure of accurate genome duplication by incomplete replication or over-replication would cause genetic alterations or cell death. Eukaryotic cells have therefore evolved a sophisticated molecular switch that ensures the activation of each replication origin exactly once per cell cycle. This switch strictly separates the loading and the activation of the replicative helicase on origins of replication during replication initiation.

Replication initiation is therefore divided into two separate steps; origin licensing and origin firing (Diffley, 2001). During origin licensing, the first step, the inactive replicative helicase is loaded onto origins and thereby forms the pre-replicative complex (pre-RC). Origin firing is the conversion of the formerly loaded pre-RC into active replisomes. Licensing and firing are temporally separated into distinct cell cycle phases in order to ensure that genome duplication occurs exactly once per cell cycle. The loading of pre-RCs during licensing takes place from late mitosis to G1 phase, whereas origin firing is tightly coupled to S phase (Diffley, 1996; Sclafani and Holzen, 2007). This temporal separation is essential to prevent the replication of genomic regions that have already been replicated. This regulation is guaranteed by oscillating activities of cyclin-dependent kinases (CDKs), the key regulators of the cell cycle. S phase CDK (S-CDK) levels are low in G1 phase due to the degradation of B-cyclins by the anaphase-promoting complex/cyclosome (APC/C). These low S-CDK levels allow origins to be licensed but not fired in G1 phase (Figure 1). During the G1/S phase transition, active G1 D-cyclin CDK4/6 mediates phosphorylation of the tumor suppressor retinoblastoma (pRb) (Malumbres and Barbacid, 2001). In turn, partial phosphorylation of pRb releases the transcription factor E2F (Johnson et al., 1993). Activated E2F promotes the transcription of cyclin-E and therefore increases S-CDK activity (Chen et al., 2009). At the end of G1 phase, high E-cyclin CDK levels further phosphorylate pRb resulting

in increased levels of E2F. This high level of E2F induces transcription of proteins involved in DNA replication (Classon and Harlow, 2002). The transition from G1 to S phase is further mediated by the inactivation of the APC/C, resulting in rising A-cyclin levels and a subsequent increase of S-CDK activity (Figure 1). High S-CDK levels in S phase inhibit origin licensing, whereas high S-CDK activity levels are required for origin firing (Siddiqui et al., 2013; Tanaka and Araki, 2010). This CDK dependency couples origin firing tightly to S phase (Diffley, 1996; Diffley, 2001; Diffley, 2004). The coupling of origin firing to S phase also depends on another cell cycle kinase, Dbf4-dependent kinase (DDK). DDK contains the catalytic subunit Cdc7 and a regulatory subunit Dbf4. DDK is inactive in G1 phase due to degradation of its regulatory subunit Dbf4 by APC/C and accumulates in S phase when APC/C becomes inactive (Figure 1).


**Figure 1: Two-step mechanism of replication initiation.**
Origin licensing takes place in G1 phase, when CDK and DDK activity is low and APC/C activity is high. At the beginning of S phase CDK and DDK activity rises and persists until the end of M phase. In M phase the APC/C promotes transition into anaphase. Adapted from (Galanti and Pfander, 2018).

This CDK- and DDK-dependent temporal separation of origin licensing and firing is conserved between yeast and humans (Sclafani and Holzen, 2007).

### 1.1.3 Molecular mechanisms of replication initiation in yeast

The molecular mechanisms and regulations of eukaryotic replication initiation are relatively well understood in budding yeast. This yeast model serves as a template to decipher the molecular details of replication initiation in vertebrates by investigating commonalities and distinctions with the yeast paradigm.

Origin licensing forms pre-RCs by loading the replicative Mcm2-7 helicase onto origin DNA. The helicase is inactive in pre-RCs to avoid premature replication in G1 phase. Replicative DNA helicases consist of ring-shaped hetero-hexamers. Two of these rings form a head-to-head formation around double stranded DNA by a tight interaction between their N-terminal domains (Evrin et al., 2009; Remus et al., 2009). Licensing is ATP-dependent and requires the essential replication factors origin recognition

complex (ORC), Cdt1 and Cdc6. ORC is a hetero-hexameric protein complex. Five of the six subunits form a crescent moon shape (Orc1-5), whereas Orc6 binds in the periphery. Orc1-5 contains an N-terminal AAA$^+$-ATPase and a C-terminal DNA-binding domain (Costa and Diffley, 2022). After sequence specific binding of ORC to origin DNA (Bell and Dutta, 2002) the AAA$^+$-ATPase Cdc6 binds to ORC. Binding of the licensing factor Cdt1 to the helicase stabilizes the formation of an open Mcm2-7 ring to facilitate the loading of the replicative helicase onto dsDNA (Frigola et al., 2017). The heptamer out of Mcm2-7 and Cdt1 is then recruited to the ORC/Cdc6 complex (Ticau et al., 2015). This loading requires ATP hydrolysis of the AAA$^+$-ATPase complexes ORC/Cdc6 (Coster and Diffley, 2017; Frigola et al., 2013; Kang et al., 2014; Siddiqui et al., 2013). In order to form a double hexamer, Mcm2-7 hexamers are loaded sequentially (Remus et al., 2009). After loading of the Mcm2-7 helicase and dissociation of Cdc6 and Cdt1, the double hexamer is called pre-RC (Figure 3 Licensing). To suppress licensing in S phase CDK inhibits the essential loading factors ORC, Cdc6, Cdt1 and Mcm2-7 (Drury et al., 1997; Drury et al., 2000; Elsasser et al., 1999; Labib et al., 1999; Nguyen et al., 2001).

During origin firing in the following S phase, the inactive helicase is converted into the active replisome. Helicase activation begins with the stepwise assembly of the so-called pre-initiation complex (pre-IC) onto pre-RCs. High levels of the two kinases S-CDK and DDK are required for this activation step (Zegerman and Diffley, 2007). Phosphorylation of the Mcm2-7 double hexamer by DDK mediates the recruitment of the firing factors Sld3 and Sld7 to the inactive pre-RC (Francis et al., 2009; Greiwe et al., 2022; Saleh et al., 2022). Sld7 is the consecutive binding partner of Sld3 and is present as a complex with Sld3 throughout the cell cycle, with Sld7 binding to the nonessential N-terminus of Sld3. The helicase component Cdc45 is delivered to the Mcm2-7 double hexamer by direct interaction with Sld3 (Itou et al., 2014). Phosphorylation of Sld3 and Sld2, the only two essential substrates of S-CDK, mediates their binding to two tandem BRCA1 C-terminus (BRCT) domains of Dpb11, thereby forming the Sld3/Sld7-Dpb11-Sld2 (SDS) complex (Figure 2) (Tanaka et al., 2007; Zegerman and Diffley, 2007). Tandem BRCT domains are sequence-specific phospho-protein binding domains. Dpb11 BRCT1/2 binds to Sld3 whereas Sld2 binds to BRCT4/5.

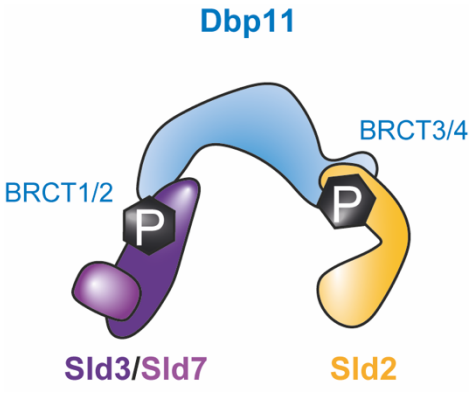



**Figure 2: Yeast SDS complex.**
CDK phosphorylation of Sld3 and Sld2 mediates the binding to two BRCT tandem domains in Dpb11, inducing the formation of the Sld3/Sld7-Dpb11-Sld2 (SDS) complex.

These binding activities facilitate two types of regulations. Firstly, binding of Sld2 to Dpb11 mediates the association of DNA polymerase ε (Polε) and the helicase component GINS to Dpb11 to form the pre-loading complex (pre-LC) (Muramatsu et al., 2010b). Secondly, the binding of phosphorylated Sld3 to Dpb11 recruits the pre-LC to the pre-RC. Upon successful delivery of the helicase components Cdc45 and GINS, the regulatory firing factors Sld3-Sld7, Sld2, and Dpb11 disassemble from the helicase. The remaining factors form the Cdc45-Mcm2-7-GINS (CMG) complex (Gambus et al., 2006; Kanemaki and Labib, 2006). Polε remains with the helicase and later fulfills essential functions in replisomes during DNA synthesis (Figure 3 Firing).

Before elongation can start, the origin DNA has to melt, and the two Mcm2-7 rings have to open to extrude one single DNA strand. The two rings encircle one single strand each, re-close, and dissociate into two single active helicases. For bidirectional replication, the two dissociated helicases migrate past each other in the N-terminal direction (Figure 3 Elongation) (Douglas et al., 2018; Fu et al., 2011). *In vitro* studies have shown that the kinases CDK and DDK and the essential initiation factors Sld3, Sld7, Sld2, Dpb11, Mcm10, Mcm2-7, Cdc45, GINS, and Polε are sufficient for replication origin firing (Yeeles et al., 2015). These factors form the core initiation factors in yeast.

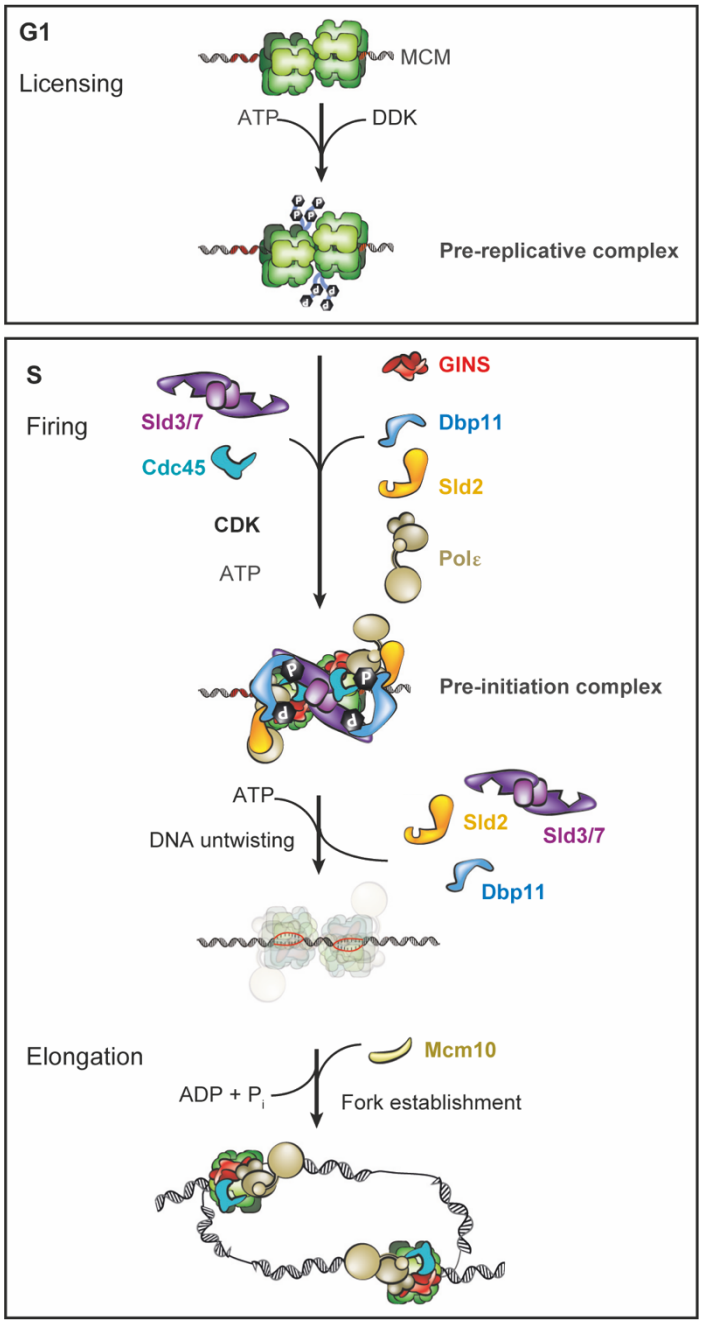



**Figure 3: DNA replication initiation in yeast.**
**Licensing:** loading of the inactive Mcm2-7 replicative helicase onto the origin of replication in G1 phase when CDK activity is low. The origin recognition complex (ORC) cooperates with the licensing factor Cdt1, bound to the helicase, and the licensing factor Cdc6. Recruitment of the helicase results in the formation of the pre-replicative complex (pre-RC). **Firing:** DDK phosphorylation of the Mcm2-7 helicase promotes recruitment of Sld3-Sld7 together with the helicase component Cdc45 to the pre-RC. Activation of CDK at G1-S transition mediates phospho-dependent binding of Sld3 and Sld2 to Dpb11. Phosphorylated Sld3 is required for the binding of the pre-loading complex (pre-LC), consisting of Sld2, Dpb11, GINS and Polε to the pre-RC by binding to Dpb11. Binding of these firing factors results in the formation of the pre-initiation complex (pre-IC). Addition of Mcm10 switches on ATPase-powered DNA unwinding by MCM, causing two Cdc45-Mcm2-7-GINS (CMG) complexes to cross their paths, which establishes bidirectional replication forks. **Elongation:** The SDS complex leaves the origin of replication whereas the helicase components Cdc45 and GINS together with Mcm10 activates the helicase. Adapted from (Costa and Diffley, 2022).

## 1.2 SDS complex as a focus of origin firing regulation

The SDS complex is a regulatory platform for origin firing in yeast. The complex mediates cell cycle and DNA damage checkpoint regulation and is subject to control mechanisms that ensure accurate and complete genome replication (Mantiero et al., 2011; Reusswig et al., 2016; Tanaka and Araki, 2011; Tanaka et al., 2011a)

Coupling of origin firing to S phase is CDK- and DDK-dependent. As described above, DDK phosphorylation of the Mcm2-7 helicase initiates Sld3-Sld7-Cdc45 recruitment, and CDK phosphorylated Sld3 and Sld2 bind to Dpb11, which in turn mediates pre-IC formation. In addition, the firing factor Sld2 is degraded during M phase, creating a temporal gap separation between licensing and origin firing. This degradation ensures that in the transition between low and high CDK and DDK levels, firing is still inhibited and cannot occur during G1 phase (Reusswig et al., 2016).

The SDS complex plays a key function in the inhibition of replication origin firing upon DNA damage or other sources of replication stress. DNA damage and replication stress activate the intra-S-checkpoint. The intra-S-checkpoint mediates phospho-dependent activation of the effector kinase Rad53 by the checkpoint kinase Mec1. This induces inhibitory phosphorylation of the SDS complex firing factor Sld3 and the regulatory Dbf4 subunit of DDK by Rad53 (Lopez-Mosqueda et al., 2010; Zegerman and Diffley, 2010). Inhibition of Dbf4 prevents phosphorylation of the Mcm2-7 helicase by DDK and therefore the binding of Sld3-Sld7, and the helicase component Cdc45 onto the helicase. Rad53-phosphorylated Sld3 cannot bind to Dpb11, resulting in the suppression of pre-IC formation and downstream CMG assembly (Boos et al., 2011; Guo et al., 2015; Kelly et al., 2022). Both, Dbf4 and Sld3, phosphorylation therefore inhibit origin firing under conditions that could cause genome instability.

Replication initiation is also influenced by the structure of the chromatin, suggesting that specific mechanisms must exist that regulate origin firing of certain genomic regions. Recently, it has been shown that Sld3 interacts with the acetyl transferase Esa1 at the silent HMLα locus. Acetylation of nucleosomes by Esa1 plays a crucial role in transcription and elongation by removing nucleosomes from DNA. This interaction controls both origin timing and transcriptional repression of the locus (Tanaka, 2021). As described above, compact heterochromatin is mainly replicated late in S phase. An exception is the pericentromeric heterochromatin that supports robust sister chromatid cohesion, which facilitates biorientation of sister kinetochores (Tanaka, 2002). It has been shown that the Ctf19 kinetochore complex facilitates the accumulation of DDK at

pericentromeres. This accumulation recruits Sld3-Sld7 to pericentromeres and initiates replication early in S phase (Natsume et al., 2013). Transcriptionally active and open euchromatin replicates early in S phase, indicating that transcription factors can stimulate origin activity. Indeed, binding of the transcription factors Forkhead 1 (Fkh1) and Fkh2 to replication origins induces replication initiation early in S phase (Knott et al., 2012).

Despite inhibition of replication initiation under situations that cause replication stress, complete genome duplication must be guaranteed to maintain genome integrity. Under replication stress conditions, stalling of a single replication fork can be rescued by a replication fork of a downstream origin. If, however, two neighboring replication forks collapse, the resulting unreplicated gap cannot be replicated by these forks. Consequently, new replisomes have to be established to complete replication in these unreplicated gaps. Since origin licensing is inhibited in S phase, it is not possible to load new Mcm2-7 helicases onto these unreplicated regions. To overcome this problem, a five- to twenty- fold excess of Mcm2-7 helicases is loaded in G1 phase compared to the number that actually fires in an unperturbed S phase (Woodward et al., 2006; Yekezare et al., 2013). These origins are called dormant origins because they remain inactive in a normal S phase. Dormant origins become activated in unreplicated regions caused by replication stress to rescue two stalled forks and are passively replicated if not needed (Blow et al., 2011; Ge and Blow, 2010; Ge et al., 2007; Letessier et al., 2011; Santocanale and Diffley, 1998; Yekezare et al., 2013).

### 1.3 Molecular mechanisms and regulations of initiation in vertebrates

Although the exact molecular mechanisms of vertebrate-specific replication regulation have not yet been elucidated, the orthologues of replication initiation core factors known from yeast have been identified, indicating conserved basic mechanisms. However, many of these orthologs contain additional vertebrate-specific regions probably required to facilitate the duplication of a much larger genome compared to yeast.

TopBP1, like yeast Dpb11, couples replication origin firing to S phase. In this process, TopBP1 binds Treslin the ortholog of Sld3 when phosphorylated by S-CDK. Like phosphorylated Sld3 that binds to Dpb11, phosphorylated Treslin binds to an equivalent phospho-specific binding domain in TopBP1 and is essential for DNA replication in vertebrates (Boos et al., 2011; Kumagai et al., 2011). Similarly to yeast,

the DNA damage checkpoint in vertebrates involves activation of the Rad53 vertebrate ortholog Chk1. Like in yeast, additional phosphorylation of vertebrate Sld3, Treslin, suppresses pre-IC formation by preventing its binding to TopBP1 (Boos et al., 2011; Guo et al., 2015; Kelly et al., 2022). MTBP, the counterpart of Sld7, was the last replication initiation factor identified in metazoans. Similarly to Sld7, MTBP binds to Treslin, the metazoan ortholog of Sld3. It has been shown that Treslin-MTBP-TopBP1 (TMT) form a stable protein complex in cell lysates (Figure 4) (Boos et al., 2013).

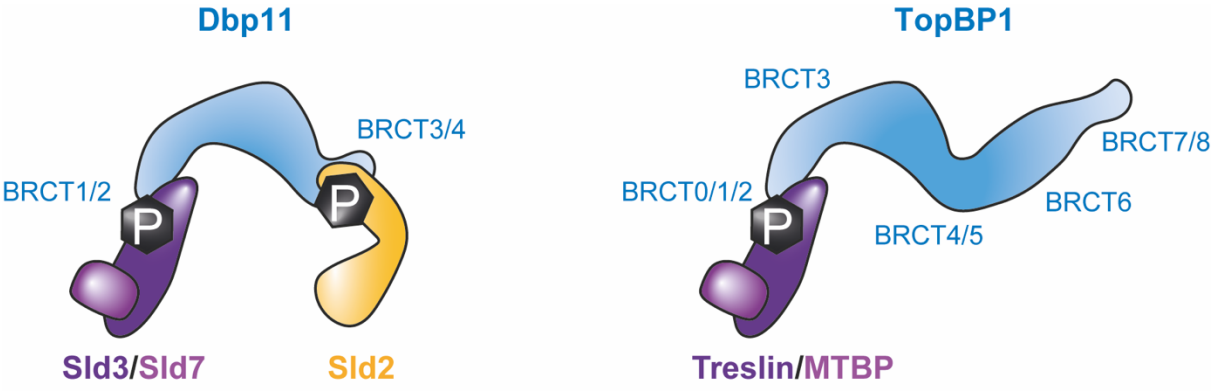



**Figure 4: Yeast SDS and metazoan TMT complexes.**
CDK phosphorylation of yeast Sld3 and Sld2 mediates the binding to Dpb11 BRCT1/2 and BRCT3/4, respectively. Treslin phosphorylation by CDK mediates phospho-dependent binding to TopBP1 BRCT0/1/2. Sld3 and Sld2 binding to Dpb11 and Treslin binding to TopBP1 result in the formation of the orthologous complexes Sld3/Sld7-Dpb11-Sld2 (SDS) and Treslin-MTBP-TopBP1 (TMT).

However, not all aspects of origin firing are conserved between metazoans and yeast. For example, the functional homolog of yeast Sld2, RecQL4, remains essential for replication initiation (Matsuno et al., 2006; Sangrithi et al., 2005), but the phospho-dependent binding to TopBP1 is probably not conserved. While Dpb11 BRCT3/4 is essential for replication due to its binding to Sld2, the equivalent TopBP1 BRCT4/5 domain is dispensable for replication in *Xenopus laevis* egg extract (Kumagai et al., 2010; Tak et al., 2006). All TMT complex proteins evolved additional protein domains that fulfill metazoan-specific functions. While MTBP is absolutely required for replication, yeast studies have shown that Sld7 is needed for normal progression of replication and for resistance against low doses of hydroxyurea but is not essential (Boos et al., 2013; Tanaka et al., 2011b). It has been suggested that MTBP is phosphorylated by at least three kinases: Cdk8/19-cyclin C kinase, M-CDK, and Chk1/2. While phosphorylation of MTBP at CDK consensus sites may promote origin firing, phosphorylation of checkpoint consensus sites inhibits replication (Ferreira et al., 2021). Treslin and Sld3 share three homologous domains, the middle domain, the

Sld3-Treslin domain and the adjacent TopBP1/Dpb11 interaction domain. Interestingly, Treslin mutants containing only these regions are not capable to rescue replication in siRNA rescue experiments. Furthermore, it has been shown that the vertebrate-specific N- and C-terminal domains are required for replication in human cells (Ferreira et al., 2022). Like Treslin and MTBP, TopBP1 is significantly larger than its yeast counterpart Dpb11. Vertebrate-specific functions of TopBP1 are the subject of this thesis and are therefore described in more detail in the following section.

In metazoans strict separation of origin licensing from origin firing is further regulated by restricting loading of the Mcm2-7 helicase to G1 phase. Like in yeast, active APC/C mediates the inactivation of the essential origin firing kinases S-CDK and DDK by degradation of S-cyclins and the regulatory Dbf4 subunit in G1 phase. In metazoans, APC/C additionally degrades the metazoan specific licensing inhibitor geminin in G1 phase. Geminin binding to Cdt1 inhibits pre-RC formation. In turn, degradation of geminin results in the release of Cdt1, which is required to load the replicative helicase onto origin DNA and is therefore indispensable for origin licensing. Geminin accumulates when APC/C activity is low in S phase, thereby preventing re-licensing (Tada et al., 2001; Wohlschlegel et al., 2000). In addition, the ubiquitin-mediated degradation of Cdt1 by CRL4 in S phase also helps to prevent re-licensing (Havens and Walter, 2011).

In conclusion, despite fundamental conservation between yeast and humans, metazoans have evolved specific regulatory mechanisms to guarantee the replication of their much larger and more complex genetic information.

### 1.3.1 TopBP1 a multi-BRCT domain scaffold protein

Like its yeast counterpart Dpb11, TopBP1 is involved in a variety of nuclear specific events, including DNA replication, DNA damage repair, cell cycle checkpoint activation, and regulation of transcription.

TopBP1 is a multi-BRCT domain scaffold protein containing nine BRCT domains (Garcia et al., 2005; Yamane et al., 1997). Although TopBP1 shares homologous sequences and functions with its yeast counterpart, Dpb11 is much smaller, containing only four BRCT domains (Figure 5 grey boxes). TopBP1 contains an additional BRCT0 domain at its very N-terminus, forming a triple N-terminal BRCT0/1/2 domain. TopBP1 BRCT3, BRCT5, and BRCT7/8 have no equivalents in Dpb11 (Figure 5 white boxes). As described above, TopBP1 couples origin firing to S phase by the binding of CDK

phosphorylated Treslin (Boos et al., 2011; Kumagai et al., 2010; Kumagai et al., 2011). However, the regulation of Sld2 and its metazoan counterpart RecQL4 has diverged. In yeast, the helicase component GINS is delivered by its binding to the GINS interaction (GINI) domain in Dpb11 as part of the pre-LC. Although the formation of such a pre-LC is not described in metazoans, a similar GINI domain has been found in TopBP1 (Figure 5) (Tanaka et al., 2013). The GINI domains in Dpb11 and TopBP1 share little sequence homology. In the context of this thesis, we provide a more detailed understanding of how TopBP1 recruits GINS to the helicase. Evolutionary modifications to regulate replication initiation are expected because metazoans have to deal with a much bigger genome compared to the yeast genome.

Despite their function in origin firing, Dpb11 and TopBP1 also share a functional role in DNA damage signaling and DNA repair pathways. The C-terminal ATR-activating domains of Dpb11 and TopBP1 mediate their binding to the checkpoint kinases Mec1 and ATR, respectively (Kumagai et al., 2006; Mordes et al., 2008b). This binding further activates the kinases and therefore crucially contributes to the DNA damage response. Dpb11 and TopBP1 also directly interact with DNA repair proteins. Colocalization of TopBP1 and Dpb11 with the DNA repair proteins 53BP1 and Rad9 (not to be confused with metazoan RAD9), respectively, is initiated by DNA double strand breaks (DSBs). Both repair proteins bind to BRCT1/2 of Dpb11/TopBP1, respectively, in a phospho-dependent manner, and 53BP1 additionally binds to TopBP1 BRCT5 (Figure 5) (Bigot et al., 2019; Pfander and Diffley, 2011).

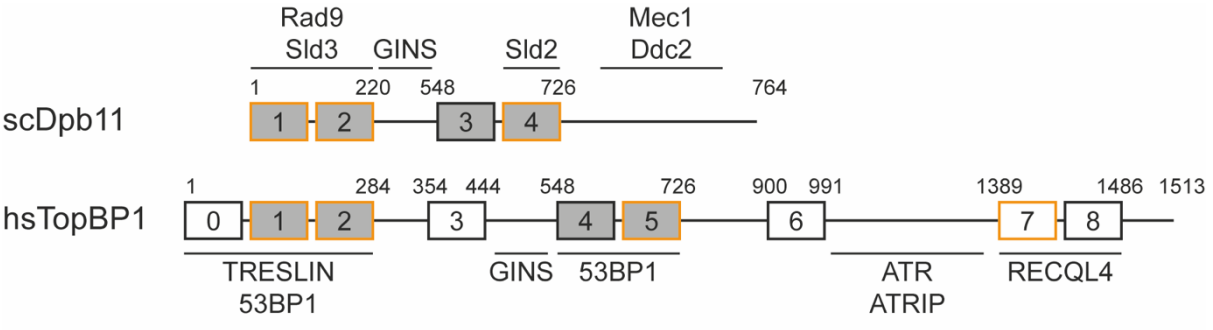



**Figure 5: Scheme of yeast Dpb11 and metazoan TopBP1 interactors.**
Domain model of *Saccharomyces cerevisiae* Dpb11 and *Homo sapiens* TopBP1. Dpb11 BRCT1/2 and BRCT3/4 are conserved in TopBP1 (grey boxes). Phosphorylation-dependent and -independent binding of homologous interacting proteins are shown in orange and black respectively.

The evolutionary modifications of TopBP1 allow interactions with other proteins, including additional DSB repair proteins (Table 1). It has been shown that TopBP1 associates with the DSB repair proteins, BRCA1, MDC1, BLM and CIP2A (Adam et

al., 2021; Greenberg et al., 2006; Sun et al., 2017; Wang et al., 2011). BRCA1 and TopBP1 are known to both interact with hyperphosphorylated FANCJ upon DSB. This binding contributes to DSB repair and S phase checkpoint activation (Sakasai et al., 2012). MDC1 association with TopBP1 facilitates TopBP1 recruitment to phosphorylated histone H2AX (γH2AX). γH2AX is known to be a biomarker for the presence of DSB in mammalian cells (Georgoulis et al., 2017). Binding of the ATP-dependent RecQ-like helicase BLM to TopBP1 is critical for the regulation of genome stability. BLM plays a crucial role in resolving homologous recombination intermediates to prevent genetic crossover events such as sister chromatid exchange (SCE). It has been reported that impaired binding of BLM to TopBP1 results in increased SCE with subsequent genome instability (Blackford et al., 2015; Sun et al., 2017). Association of TopBP1 with the phosphatase CIP2A also appears to be essential for maintaining genome stability, although the exact mechanisms are yet to be elucidated (Adam et al., 2021; Laine et al., 2021).

**Table 1: TopBP1 interacting proteins**

| BRCT | DNA replication | DNA repair | Cell cycle | Transcription |
|---|---|---|---|---|
| 0/1/2 | CDC45; Treslin-MTBP | BRCA1; MDC1; RAD9 (HUS1; RAD1); RHNO1; SMARCAD1 | 53BP1 | TCOF1 |
| inter 3/4 (GINI) | GINS | | | |
| 4/5 | | BLM; BRCA1 | 53BP1; MDC1 | TCOF1 |
| inter 4/5 | | CIP2A | | |
| 6 | CDC45 | PARP1 | FANCJ | E2F1; TOP2A; TOP2B |
| AAD | | ATR; ATRIP | ATR | |
| 7/8 | RECQL4 | BRCA1; PLK1 | PHF8 | TOP2A; TOP2B |

This illustrates the importance of TopBP1 in the response to DNA damage that has evolved from yeast to humans.

### 1.3.2 The helicase component GINS, a heterotetrameric complex

GINS along with Cdc45 and Mcm2-7 forms the CMG complex, which serves as the core of the replicative helicase (Botchan and Berger, 2010; Yeeles et al., 2015). The Mcm2-7 helicase forms a two-tiered ring structure around DNA. The N-tier of Mcm2-7

is supported by Cdc45 and GINS, forming a platform upon which the C-tier AAA+ ATPase motor can translocate along ssDNA in the 5' to 3' direction for bidirectional replication (Costa et al., 2011; Yuan et al., 2016).

GINS is a heterotetrameric complex consisting of four subunits: Sld5, Psf1, Psf2, and Psf3. The name GINS originates from the Japanese Go-Ichi-Ni-San (GINS), meaning five, one, two and three, representing the four related subunits. The GINS complex is highly conserved from archaea to eukaryotes (Goswami et al., 2018; Jones et al., 2021; Kanemaki et al., 2003; Kubota et al., 2003; Marinsek et al., 2006). It has been shown that each of the four GINS subunits consists of an α-helical A-domain and a smaller β-sheet rich B-domain (Carroni et al., 2017). The A-domains of Sld5 and Psf1 are located at their N-termini, and the B-domains at their C-termini. In contrast, the domain order is revered in Psf2 and Psf3. A-domains located at the same termini of Sld5 and Psf1, and Psf2 and Psf3, respectively, form heterodimers. This creates a vertical interface between the two heterodimers resulting in a 2-fold pseudo-symmetrical axis (Figure 6) (MacNeill, 2010).

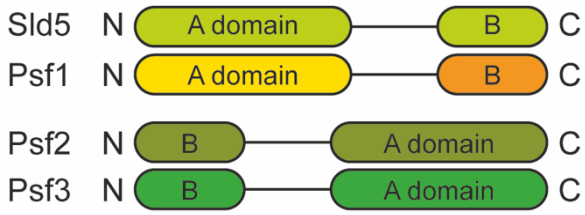



**Figure 6: Schematic domain structure of the eukaryotic GINS heterotetramer.**
The α-helical A-domain is located at the N-termini of Sld5 and Psf1 and the C-termini of Psf2 and Psf3. The β-sheet rich B-domain is located at the C-termini of Sld5 and Psf1 and at the N-termini at Psf2 and Psf3. Adapted from (MacNeill, 2010).

Two independent structural analyses of GINS have revealed that parts of Sld5, Psf1, and Psf3 are not visible in cryo-EM and X-ray crystallography, indicating that these domains are highly flexible (Chang et al., 2007; Kamada et al., 2007). Based on the visible structure of the GINS tetramer, it has been suggested that the flexible Sld5, Psf1, and Psf3 domains are adjacent. The colocalization of these flexible domains on the GINS surface indicates that they may form a binding surface for further replication proteins. *In vitro* assembly of the human replisome illustrated that the flexible B-domain of Psf1 adopts a discrete confirmation once integrated into the CMG complex. It appears to become ordered through its packing against Cdc45 (Chang et al., 2007; Day et al., 2023). The conformational change of the Psf1 B-domain further allows

interaction with the N-terminus of the Polε accessory subunit PolE2, resulting in the formation of the CMG-Polε complex (CMGE) (Jones et al., 2021).

## 1.4 *Xenopus laevis* egg extract to study DNA replication

In this thesis, *Xenopus laevis* egg extracts are used to investigate the role of TopBP1 in CMG formation. *Xenopus* egg extract is a powerful cell-free system that recapitulates the major nuclear transitions of the eukaryotic cell cycle *in vitro* under controls similar to those in the embryonic cell cycle. Since the pioneering invention of this model system in 1983 by Lohka and Masui, egg extract has facilitated many important discoveries about the vertebrate cell cycle, mitosis/meiosis, DNA replication, and other cellular processes (Lohka and Masui, 1983). Due to the lack of a plasma membrane, egg extract is amendable to biochemical manipulations, such as addition of proteins and chemical inhibitors or by immunodepletion of a protein of interest for loss of function studies. In this research project, the use of replicating *Xenopus* egg extract is essential to investigate the function of the TopBP1-GINS interaction and the formation of replisomes.

*Xenopus* egg extract has the ability to support cell cycle progression *in vitro*. As in other vertebrates, *Xenopus* eggs are arrested in metaphase of meiosis II. This arrest is mediated by the cytostatic factor (CSF). CSF facilitates this arrest by the inhibition of the APC/C. Consequently, securin and cyclin B cannot be degraded (Tunquist and Maller, 2003). Binding of securin to the protease separase prevents degradation of the cohesion ring that keeps the two siter chromatids together, hereby preventing the entry of the eggs into anaphase. In addition, high cyclin B-CDK1 (also called mitosis promoting factor; MPF) activity prevents meiotic exit (Gautier et al., 1990; Lohka et al., 1988). Fertilization induces an influx of $Ca^{2+}$ions into the zygote, which inactivates the CSF, promoting entry into anaphase and transition into interphase of the mitotic cell cycle (Jones, 1998). *Xenopus* eggs undergo twelve synchronous rounds of cell divisions, driven by the oscillation activity of the MPF (Murray and Kirschner, 1989). These twelve cell divisions are completed within seven hours after fertilization, without zygotic transcription. After this stage, the zygote reaches the so called Mid-Blastula transition and starts transcription (Newport and Kirschner, 1982). Therefore, *Xenopus* eggs contain large amounts of material to support cell progression without transcription before they reach the Mid-Blastula transition. Consequently, extracts prepared from

these eggs also contain abundant stockpiles of material that support nuclear assembly and DNA replication (Blow and Laskey, 1986).

To prepare *Xenopus* egg extracts that efficiently support DNA replication, frogs are consecutively treated with the hormones pregnant mare serum gonadotropin and human chorionic gonadotropin to stimulate egg laying. These CSF-arrested eggs are crushed by low-speed centrifugation to separate the crude cytoplasm from the lipid layer and the yolk platelets (Figure 7 A). To remove further debris, particularly mitochondria that cause apoptosis, the crude cytoplasm is centrifuged again (Figure 7 B). This crude extract, consisting of the cytoplasm and membranes, is stably arrested in metaphase by the CSF until fertilization is mimicked by $Ca^{2+}$ addition. Addition of $Ca^{2+}$ induces repetitive MPF oscillations, equivalent to mitotic alternating mitoses and interphases. When *Xenopus* sperm DNA is added to this interphase extract, it is first reorganized into chromatin before interphase nuclei are formed (Figure 7 C). Low levels of MPF allow origin licensing. Nuclei form within about 30 minutes. The formation of nuclei allows the enrichment of replication factors in the nuclei, including CDK, which facilitates origin firing.

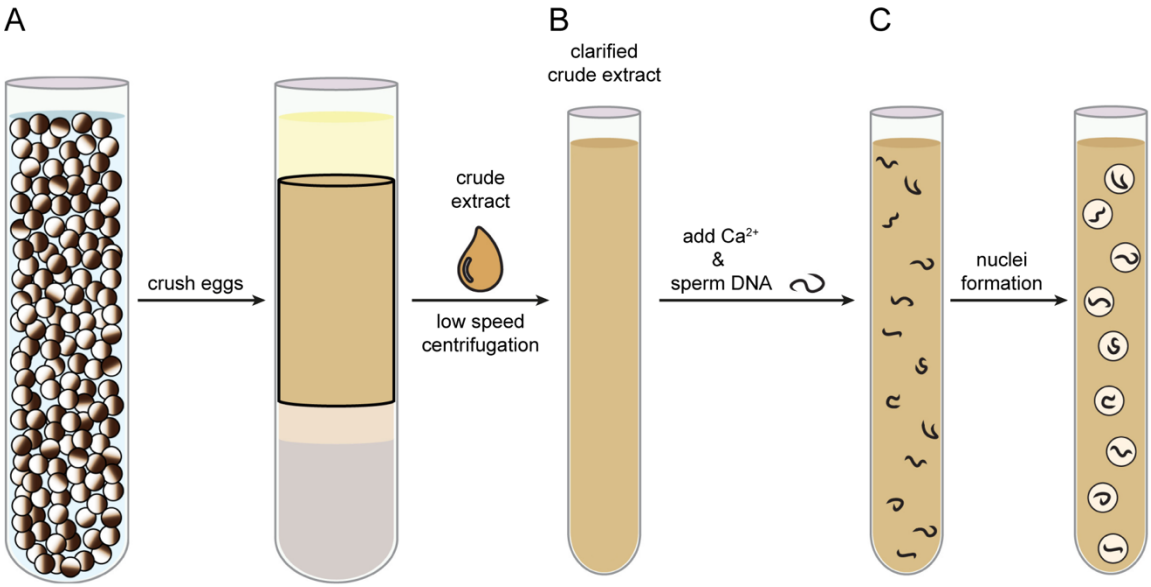



**Figure 7: Schematic representation of *Xenopus* egg extract preparation.**
A) Unfertilized *Xenopus laevis* eggs are crushed and the CSF-arrested crude extract is collected. B) Low speed centrifugation generates a clarified CSF-arrested crude extract, containing membranes. C) Clarified crude extract is released into interphase by $Ca^{2+}$ addition. Incubation of interphase extract with sperm DNA induces nuclei formation. Adapted from (Hoogenboom et al., 2017).

In this research project *Xenopus* egg extract is used to monitor replication origin firing on isolated chromatin, and DNA synthesis by incorporation of radioactive labeled dCTPs into nascent DNA.

# 2 Aim of the Study

To guarantee genetic homeostasis, the genome must be duplicated accurately and exactly once during each cell cycle. Due to the large eukaryotic genome this is a particularly complex process in eukaryotes. Replication origin firing serves as a critical step in this tightly regulated process of activating the replicative helicase during DNA replication. *In vitro* reconstitution experiments have defined the factors sufficient for replication origin firing in yeast (Yeeles et al., 2015; Zegerman and Diffley, 2007). Replication origin firing involves the assembly of the CMG complex, consisting of the Mcm2-7 helicase, Cdc45, and GINS, which facilitates DNA unwinding and supports the progression of DNA replication (Gambus et al., 2006; Kanemaki and Labib, 2006). In yeast the loading of Cdc45 is mediated by Sld3 and GINS is recruited as part of the pre-LC consisting of GINS, Dpb11, Sld2 and Polε (Itou et al., 2014; Muramatsu et al., 2010a). While in metazoan Cdc45 is presumably recruited by the metazoan counterpart of Sld3, Treslin, the metazoan pre-LC has not been described (Itou et al., 2014). Previous yeast two-hybrid experiments have revealed a GINS interaction domain loosely described between TopBP1 BRCT3 and BRCT4 that is required for efficient DNA replication in eukaryotes (Tanaka et al., 2013). However, the molecular details of the GINS-TopBP1 interaction that mediates the recruitment of GINS onto the Mcm2-7 helicase are still poorly understood.

In this project, we utilized a combination of various experimental approaches to investigate the interaction between TopBP1 and GINS, providing a more detailed understanding of their contribution to the activation of the CMG helicase. Structural predication and sequence alignment allowed the identification of potential binding surfaces between TopBP1 and GINS. Systematic mutations of these sites were used to determine their interaction *in vitro* and *in vivo*. Biochemical binding experiments and functional analyses revealed how TopBP1 associates with GINS and how this interaction effects replication in *Xenopus* egg extract. To assess the requirement of the TopBP1-GINS interaction in replication origin firing, we used mass spectrometry of chromatin-bound proteins.

In summary, our model provides a potential mechanism on how TopBP1 could be recycled after handing over GINS to the helicase.

# 3 Results

## 3.1 Measuring replication initiation and origin firing in interphase *Xenopus* egg extract

Cell-free extracts prepared from *Xenopus laevis* eggs were instrumental in investigating the eukaryotic cell cycle and a variety of other cell biological processes. In this thesis, egg extracts were used to study DNA replication initiation. For these analyses, the preparation and characterization of the egg extract was first established in the Boos laboratory.

Depending on the application, *Xenopus laevis* egg extract preparation methods differ. An egg extract method from the laboratory of Julian Blow was previously established in the Boos laboratory (Parlak, 2017). The Blow laboratory developed a method that was successfully used to investigate DNA replication *in vivo* in cell-free egg extract (Gillespie et al., 2012). For this, the conventional method of extract preparation has been optimized to support efficient DNA replication in these egg extracts. For the preparation of this extract, CSF-arrested eggs were collected and subsequently the cytosol was extracted. Compared to conventional extracts prepared from activated eggs, release from metaphase arrest was induced only after extract preparation. These extracts display a smaller variability in quality. To avoid apoptosis due to contamination with mitochondria, these egg extracts were further clarified by an additional centrifugation step. Optimal energy supply was guaranteed by the addition of a high concentration of an ATP-regenerating system. Furthermore, transcription was inhibited by addition of cycloheximide.

The following section recapitulates the establishment of CSF-arrested egg extract and complements it with additional characterizations.

### 3.1.1 Interphase *Xenopus laevis* egg extracts are capable to perform genome replication

The transition of sperm chromatin morphology was observed over time in *Xenopus laevis* egg extract. First, CSF-arrested egg extracts were released into interphase by addition of $CaCl_2$. Morphological changes of sperm chromatin added into the extract were followed over time in presence of DAPI.

Immediately after the addition of sperm DNA, chromatin showed a characteristic highly-condensed structure (Figure 8 A). After 20 minutes, sperm chromatin swelled, as indicated by the larger area of the DAPI-stained structure (Figure 8 B). This morphological change implied remodeling from sperm-specific, over-condensed chromatin into typical chromatin, which is found in somatic cells. Fully-formed nuclei could be observed after approximately 40 minutes (Figure 8 C).

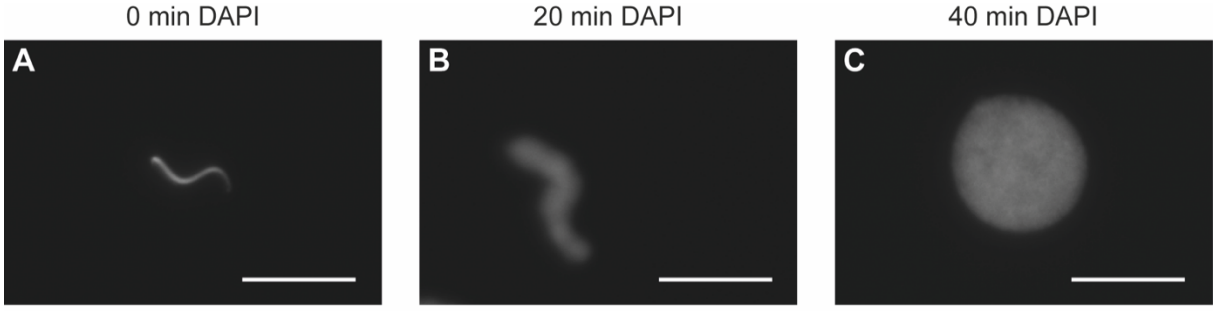



**Figure 8: Time course of nuclei assembly in $Ca^{2+}$ released *Xenopus laevis* egg extract.**
Sperm nuclei were incubated in $Ca^{2+}$ released egg extract by addition of 0.3 mM $CaCl_2$. Nuclei formation was monitored by DAPI staining followed by fluorescens microscopy. Scale bar: 20 µm. A) Condensed sperm DNA immediately after addition to egg extract. B) Swollen sperm chromatin after 20 minutes of incubation. C) Fully formed nuclei 40 minutes after sperm DNA addition to egg extract.

The observed chromatinization and subsequent formation of interphase nuclei showed that interphase extract was capable to form typical chromatin.

The cell-free *Xenopus laevis* egg extract system offers an exceptional opportunity to follow replication over time and to confirm genome replication in our system. For this, incorporation of radioactive labeled dCTPs into newly synthesized DNA was detected after defined time points by phosphor imaging. In addition, isolation of chromatin and chromatin-bound proteins was used to monitor the association and dissociation of replication factors, at different stages, from origin licensing to replisome formation.
Figure 9 A shows that within a delay of 30 minutes after sperm addition, which is required for nuclei formation, radioactive signals can be detected. In contrast, control extract without sperm DNA did not show radioactive signals. For quantification, signal intensities were measured using Fiji and the 0 minutes time points of samples with and without sperm DNA were subtracted as background (Figure 9 B). In later experiments, these replication assays were performed in triplicates for statistical analysis of quantified signal intensities.
True genome replication, as opposed to other replicative processes like DNA repair for example, is marked by licensing-dependent DNA replication, another essential

hallmark of replication initiation. Using the licensing inhibitor geminin is a classic method to achieve inhibition of origin licensing.

To test if replication in the egg extract was dependent on origin licensing, the licensing inhibitor geminin was added prior to sperm DNA and radioactive dCTPs. At concentrations above 250 nM geminin fully inhibited dCTP incorporation (Figure 9 C). Figure 9 D shows time-dependent chromatin isolation kinetics. Within the first 20 minutes (Figure 9 D lane 2) after sperm addition, isolated chromatin was associated with various licensing factors such as the Mcm2-7 helicase subunits Mcm2, Mcm6 and Mcm7, which were visualized by immunoblotting of chromatin fractions using antibodies against the indicated Mcm2-7 subunits. In contrast to these licensing factors, firing and replisome factors like polymerase ε (PolE2) and GINS (Sld5), associated with chromatin only after 30 to 45 minutes after sperm addition (Figure 9 D lane 3 and 4). To guarantee true DNA replication signals, control chromatin was isolated after 90 minutes from an egg extract containing the licensing inhibitor geminin (Figure 9 D lane 6). Neither signals for licensing markers nor signals for firing markers were observed as expected. The lower portion of the gel was stained with Coomassie as a loading control, showing equal loading of chromatin factions, indicated by similar intensity of histone bands.

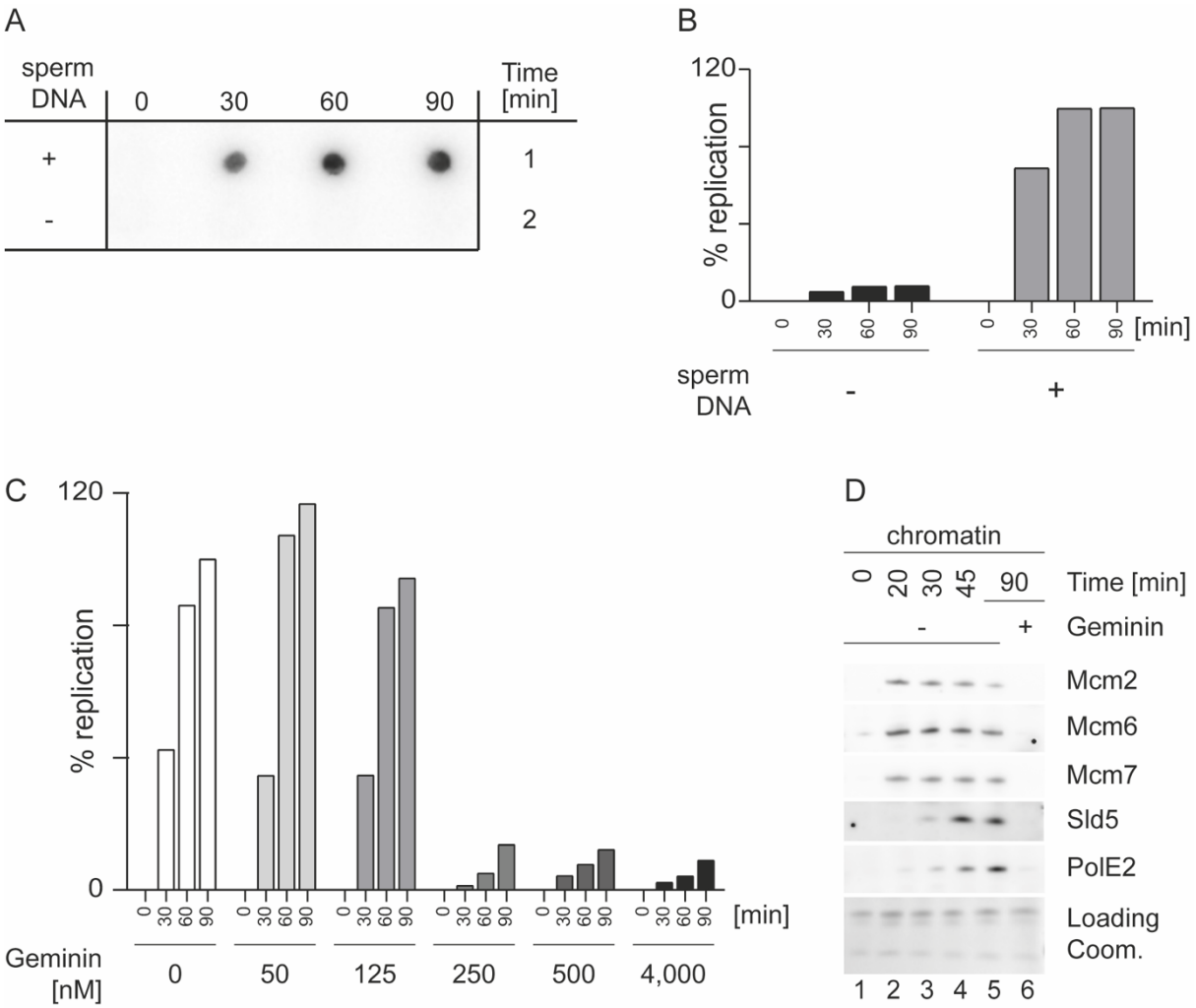



**Figure 9: DNA- and licensing-dependent genome replication in interphase *Xenopus* egg extract.** A) DNA replication in interphase egg extract. Egg extract was supplemented with $\alpha^{32}$P-dCTP. Start of the reaction was induced by the addition of 5 ng sperm DNA/µl egg extract (+ sperm). B) Quantification of experiment shown in A). C) Licensing dependency of radioactive signals. Titration of geminin into extract. D) Immunoblot analysis of chromatin-bound proteins upon isolation from interphase *Xenopus* egg extract. Indicated antibodies were used for detection of replication factors. Coom., Coomassie staining.

In conclusion, these data demonstrate template- and licensing-dependent genome replication, confirming true genome replication.

Replication of the genome is dependent on B-type polymerases like polymerases α, δ and ε. To test if replication in our egg extract system is dependent on these DNA polymerases, interphase egg extract was supplemented with different concentrations of the B-type polymerase inhibitor aphidicolin prior to sperm DNA and radioactive labeled dCTP. Incorporation of radioactive labeled dCTP into nascent DNA was inhibited by the addition of 10 µg aphidicolin per ml egg extract in a dose-dependent manner (Figure 10). To ensure maximal inhibition of replication, 50 µg aphidicolin per

ml egg extract were used in later experiments. These aphidicolin sensitive signals confirmed B-type polymerase dependent DNA replication in interphase egg extract.

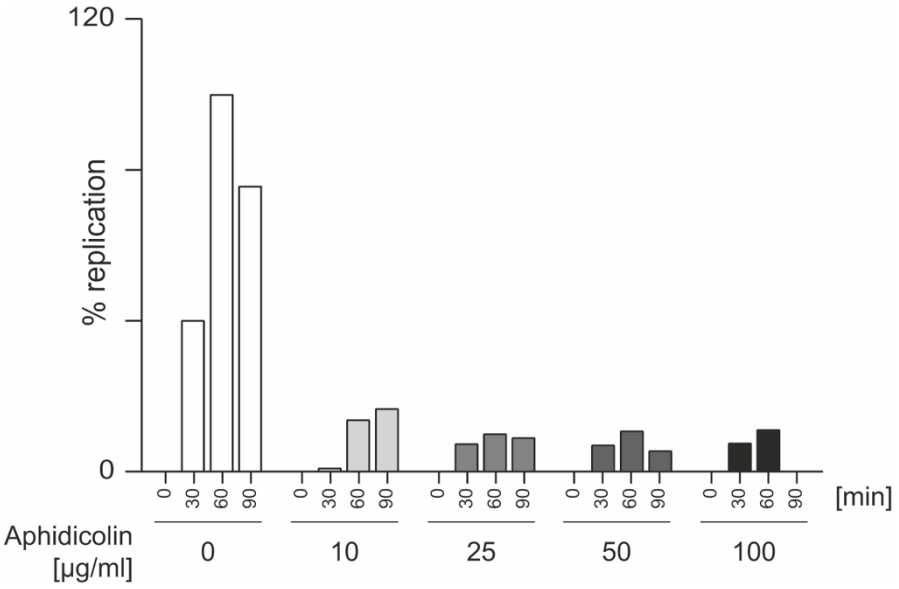



**Figure 10: Polymerase-dependent genome replication in interphase *Xenopus* egg extract.**
B-type polymerase-dependency of radioactive signals. Titration of aphidicolin into extract.

Taken together, these data confirmed three hallmarks of true genome replication in the experimental system: 1) DNA dependency, 2) licensing dependency, and 3) polymerase-dependent genome replication.

## 3.2 Proximity biotinylation identifies TopBP1-GINS interaction in human cells

To identify TopBP1 interactors, especially replication factors *in vivo*, proximity biotinylation was established in the Boos laboratory.

Biotin proximity labeling is a method to map interactors in the vicinity of a protein of interest. For this the APEX2 ascorbate peroxidase is fused to the protein of interest, and therefore is expressed together with the protein of interest in cells. To biotinylate proteins in proximity, the peroxidase catalyzes the oxidation of biotin phenol to reactive biotin phenol free radicals. These free radicals can interact with various functional groups of proteins in proximity and thereby label them with biotin. Subsequently, these biotinylated proteins are pulled down with streptavidin beads from cell lysates and can be identified by mass spectrometry. In contrast to classic immunoprecipitation experiments, proximity labeling combined with proteomics methods offers the

prospective to identify transient interactors and proteins in the vicinity of TopBP1 as well as direct TopBP1 binders.

### 3.2.1 Establishing APEX2-dependent proximity biotinylation

To systematically establish proximity biotinylation in human cells in the Boos laboratory, different transfection systems were tested first. Within these systems, the expression levels of proteins from integrated APEX2-tagged TopBP1 differ vastly. To accomplish a balance between signal sensitivity and signal specificity, two different cell lines stably expressing APEX2-tagged TopBP1, and transiently transfected cells were screened. Moreover, two TopBP1 expressing constructs containing either TopBP1 N- or C-terminally tagged with the peroxidase were used.

N- or C-terminal APEX2-tagged TopBP1 was either stably expressed in isogenic stable 293 Flip-In cells or transiently expressed in 293T cells. Furthermore, TopBP1 N-terminal tagged with APEX2 was stably expressed in HeLa Flip-In cells. Immunoblotting was used to assess the expression levels of APEX2-tagged TopBP1 in the different cell lines.

C-terminal tagged TopBP1 transiently transfected in 293T cells was expressed to similar levels as endogenous TopBP1 (Figure 11 A lane 2), whereas N-terminal tagged TopBP1 was expressed at significantly higher levels (Figure 11 A lane 1). Similar differences between N- and C-terminal APEX2 tagging was observed in 293 Flip-In cell lines stably expressing TopBP1 (Figure 11 B lane 3 and 4). Even lower expression levels were observed in HeLa Flip-In cell lines stably expressing N-terminal tagged TopBP1 (Figure 11 B lane 2). In order to test biotinylation by APEX2, biotinylation was induced by $H_2O_2$ addition in transiently-transfected 293T cells expressing either N- or C-terminal APEX2-tagged TopBP1. Subsequently, biotinylated proteins were purified using streptavidin immobilized on sepharose resin. Samples were immunoblotted using an HRP-coupled streptavidin antibody. Both constructs showed smears of biotinylated proteins (Figure 11 C lane 2 and 4) that were not present in controls omitting $H_2O_2$ (Figure 11 C lane 1 and 3) or in non-transfected control cells (Figure 11 C lane 5). N-terminal APEX2-fused TopBP1 showed higher signal intensities than C-terminal tagged TopBP1. Thus, higher expression levels of TopBP1 fused to the peroxidase caused also higher signal intensities of biotinylation levels. Therefore, transient transfected 239T cells and 293 Flip-In cells were used for further mass spectrometry experiments to identify TopBP1 interactors. The stably-expressing HeLa

cell lines were omitted due to low expression levels. Such low expression conditions may be tested in the future.

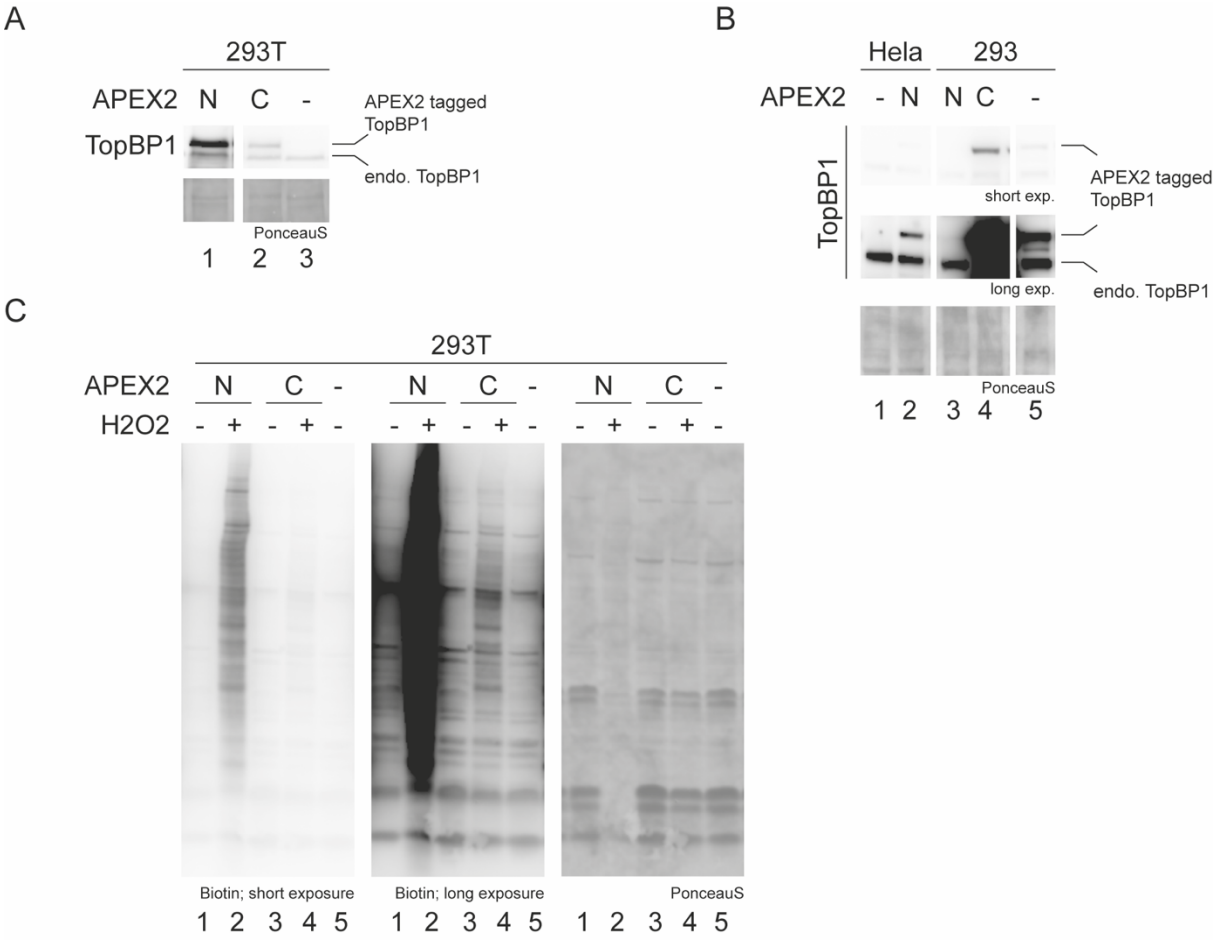



**Figure 11: APEX2-dependent proximity biotinylation.**
Immunoblot analysis of APEX2-tagged TopBP1 expression levels in A) transient transfected 293T cells and B) stably expressing HeLa Flip-In and 293 Flip-In cell lines. C) Immunoblot analysis of biotinylated proteins after streptavidin pull-down.

To identify TopBP1 interactors by proximity biotinylation followed by biotin pulldown, 239 Flip-In cell lines and transiently-transfected 239T cells expressing either N-terminal or C-terminal APEX2-fused TopBP1 were used. After induction of biotinylation, samples were sent for LC-MS/MS analysis, in collaboration with Dr. Farnusch Kaschani and Dr. Markus Kaiser. Label-free quantification (LFQ) intensities of quadruplicates of each condition were used for statistical analysis in Perseus. In brief, potential contaminants, reverse hits, and proteins identified only by site modifications were filtered out. To identify proteins with significantly changed relative abundance between conditions, LFQ intensities were log2 transformed and missing values of negative controls (-$H_2O_2$) were imputed with random values drawn from a normal distribution. Subsequently, LFQ intensities were filtered by valid values in a minimum of seven out of eight replicates. Next, two sample Student's t-tests using a

significant threshold of FDR<0.5 was used for statical analysis and visualized as volcano plots to show changes in protein enrichment between two conditions. In these scatter plots the x-axis shows changes of protein enrichment between two conditions. A value above zero describes an accumulation of a protein in one condition compared to another one. The y-axis reflects the p value determined for each protein by Student's t-tests. Due to the log2 transformation, small p values appear on the top section of the plot. Hence, comparing two conditions, significantly enriched proteins appear in the upper right corner, whereas the upper left corner represents significantly reduced proteins.

To assess the quality and reproducibility of the experimental system, four independent experiments using stable cell lines or transiently-transfected cells expressing N- or C-terminally APEX2-tagged TopBP1, were analyzed. Volcano plots show similar distributions of identified proteins. Enrichment of proteins became apparent in the upper-right corner of the scatter plot when the peroxidase was activated compared to the negative controls (Figure 12 red dots). A reduction of identified proteins shown in the upper left corner could also be observed (Figure 12 blue dots). Since a distinction was made between biotinylated and non-biotinylated proteins in these experiments, a significant reduction of biotinylated proteins is conceptionally not possible. It is more likely that these hits reflect background signals resulting from measurement errors in this method. However, the enrichment exceeds the reduction as would be expected for these analyses.

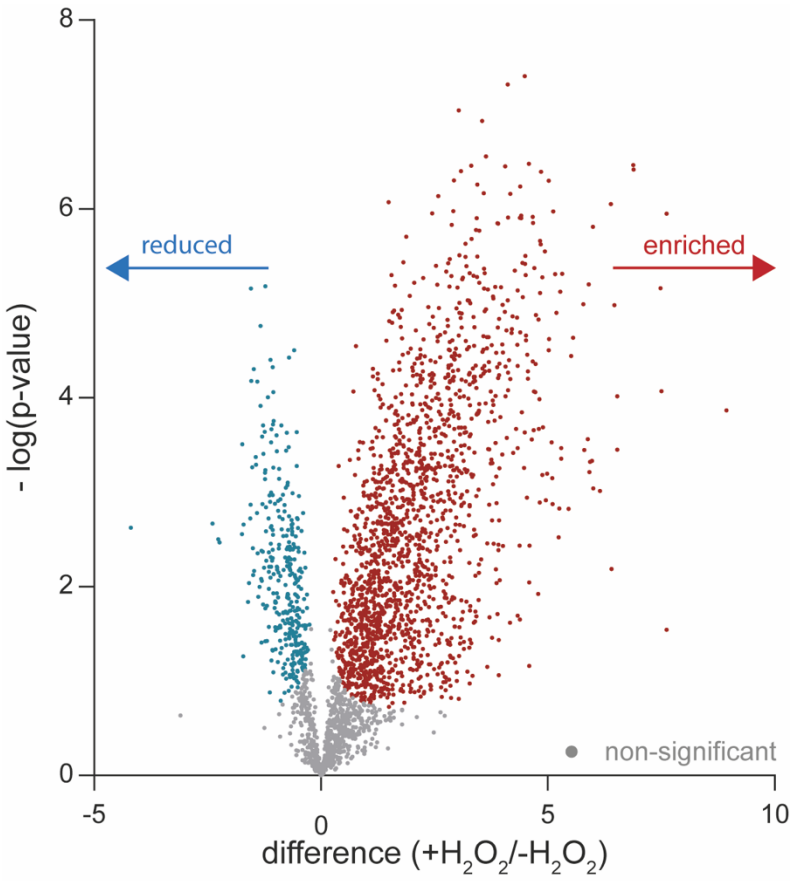



**Figure 12: Representative volcano plot of LC-MS/MS analysis after proximity biotinylation.** Red) Enriched proteins appeared in the upper-right corner of the scatter plot. Blue) Reduced proteins appeared in the upper-left corner. Grey) Non-significant hits.

### 3.2.2 APEX2-mediated proximity biotinylation allows determination of the TopBP1 proximitome

To determine the replication-specific proximitome of TopBP1, proximity biotinylation experiments followed my mass spectrometry analysis were established. The role of TopBP1 in replication initiation is not fully characterized. As described before TopBP1 contains vertebrate-specific regions, like the BRCT3 domain and the inter BRCT region between BRCT3 and BRCT4 which have unknown but essential functions. TopBP1 is an excellent candidate to establish the proximity biotinylation method because many TopBP1 interactors are already identified and can be used to verify these analyses.

To identify these interactors, several independent proximity biotinylation experiments were performed. For this, cell lines expressing APEX2-tagged TopBP1 at least up to endogenous levels were used. As shown in Figure 11 A and B, N-terminal tagged TopBP1 was overexpressed in both transiently-transfected 293T cell lines and stable 293 Flip-In cells, whereas C-terminal tagged TopBP1 was expressed to similar levels

as endogenous TopBP1 in transiently transfected cells. Induction of the peroxidase followed by streptavidin pulldown and mass spectrometry analysis were performed as described above. Two independent experiments using 293T cells that transiently expressed N-terminal tagged TopBP1, N #1 and N #2, showed significant enrichment of 2,066 and 2,396 proteins, respectively. Significant enrichment of 1,433 proteins was identified when C-terminal tagged TopBP1 was transiently expressed, whereas only 923 proteins were identified when APEX2-TopBP1 was stably expressed in 293 Flip-In cells (Figure 13 D). The observed difference in protein enrichment directly correlates with the expression levels of APEX2-tagged TopBP1 in the different expression systems (Figure 11 A and B).

To exclude unspecific hits and verify the reliability of the significantly enriched proteins, these lists of hits were compared to each other. For this analysis only proteins identified as significantly enriched by Student's t-test were used. Subsequently, numeric Venn diagrams were generated in Perseus. These numeric Venn diagrams identified the number of common and unique hits. Four parameters were compared to each other: 1) reliability of two independent experiments, 2) influence of the position of the APEX2 tag, 3) influence of expression levels on specificity and 4) identification of highly specific hits.

Comparing transiently expressed N-terminal APEX2 tagged TopBP1 revealed a similar number of significantly enriched proteins. Of these hits,1,812 were present in both experiments, corresponding to 88 % (N #1) and 76 % (N #2) of the total number respectively (Figure 13 A). This high level of common proteins suggests good reproducibility. Next, the difference between N- and C-terminal tagged TopBP1 was analyzed. Comparison of N #1 or N #2 with transiently expressed C-terminal tagged TopBP1, revealing 80 % and 95 % common hits respectively (Figure 13 B), indicating that the position of the APEX2 tag did not influence the reproducibility. To test the impact of expression levels on the specificity of the experimental system, N #1 or N #2 were compared to the low number of hits identified in stably-expressed APEX2-TopBP1. An overlap of 80 % and 87 % respectively suggests that endogenous expression levels did not influence biotinylation of specific proteins (Figure 13 C). It is more likely that overexpression of APEX2 leads to unspecify biotinylation. To eliminate these unspecific hits, a negative control expressing the peroxidase coupled to a nuclear localization signal (NLS) will be used in future experiments in the Boos laboratory.

Highly specific hits are likely to be present in all four experiments. A total of 462 common hits was identified in all four experiments (Appendix Table 1). This is equivalent to 50 % of the hits present in stably-expressed APEX2-TopBP1 that presumably reflects the highest specificity due to endogenous expression levels (Figure 13 D). In summary, these experiments provide evidence for reliable protein identification using the method.

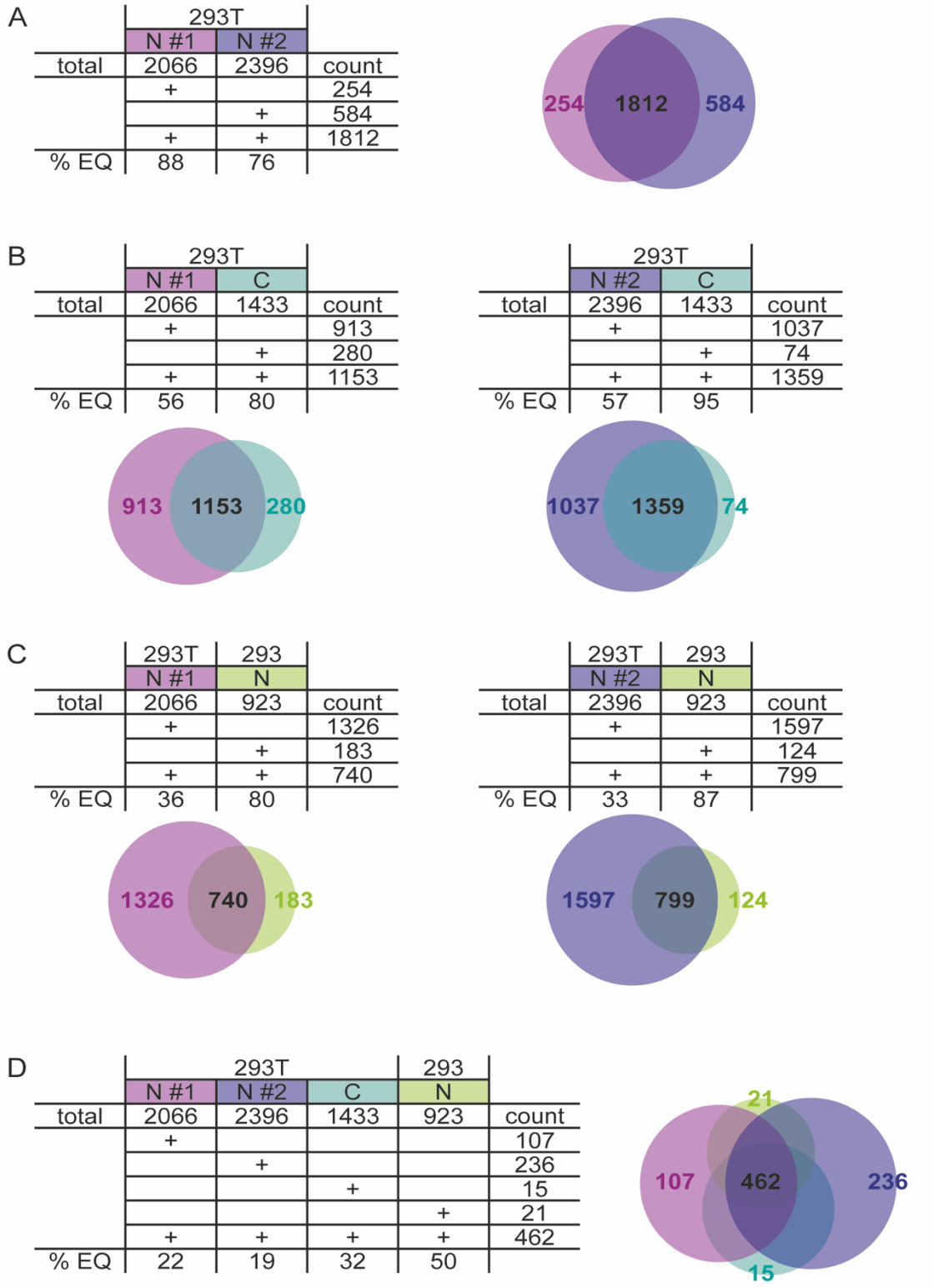



**Figure 13: Venn diagram of significantly enriched proteins in four individual proximity biotinylation experiments.**
A) Comparison of the number of significantly enriched proteins in two individual proximity biotinylation experiments using 293T cells transiently-expressing APEX2-TopBP1. B) Comparison of the number of significantly enriched proteins in proximity biotinylation experiments using 293T cells transiently-expressing either N- or C-terminal APEX2 tagged TopBP1. C) Comparison of the number of significantly enriched proteins in proximity biotinylation experiments using either 293 Flip-In cells stably or 293T cells transiently-expressing APEX2-TopBP1. D) Comparison of the number of significantly enriched proteins in four individual proximity biotinylation experiments.

Identification of known TopBP1 interactors were used to determine the specificity of the proximity biotinylation experimental system. For this aim, the four biotin labeling experiments were screened for 25 proteins that are known to physically interact with TopBP1, including the four GINS subunits (Figure 14 A). Out of these known interactors 68 %, 80 %, 56 % and 64 % were found in the individual experiments (Figure 14 B (i)). Within these hits, the GINS subunit Sld5 was found in one, Psf2 in three and Psf3 in all experiments. These high levels of identified TopBP1 interactors suggests good specificity.

In contrast to classical binding assays, proximity biotinylation offers the exceptional advantage to identify not only direct physical interactors of a protein of interest, but also allows the detection of proteins involved in the same pathway if they are in close vicinity. Here we refer to these proteins as indirect interactors. Identification of indirect TopBP1 interactors was used to further verify the specificity of the method. We define indirect TopBP1 interactors as proteins that are involved in the same pathway as a direct physical TopBP1 interactor. Indirect TopBP1 interactors were selected using the protein-protein interacting network STRING. For this, 50 proteins involved in a pathway of each direct physical TopBP1 interactor were determined by STRING. Next these lists of 50 proteins were compared to the 462 common proteins present in all experiments. As shown in Figure 14 B (ii) indirect TopBP1 interactors (green) were found for almost all direct physical TopBP1 interactor (black).

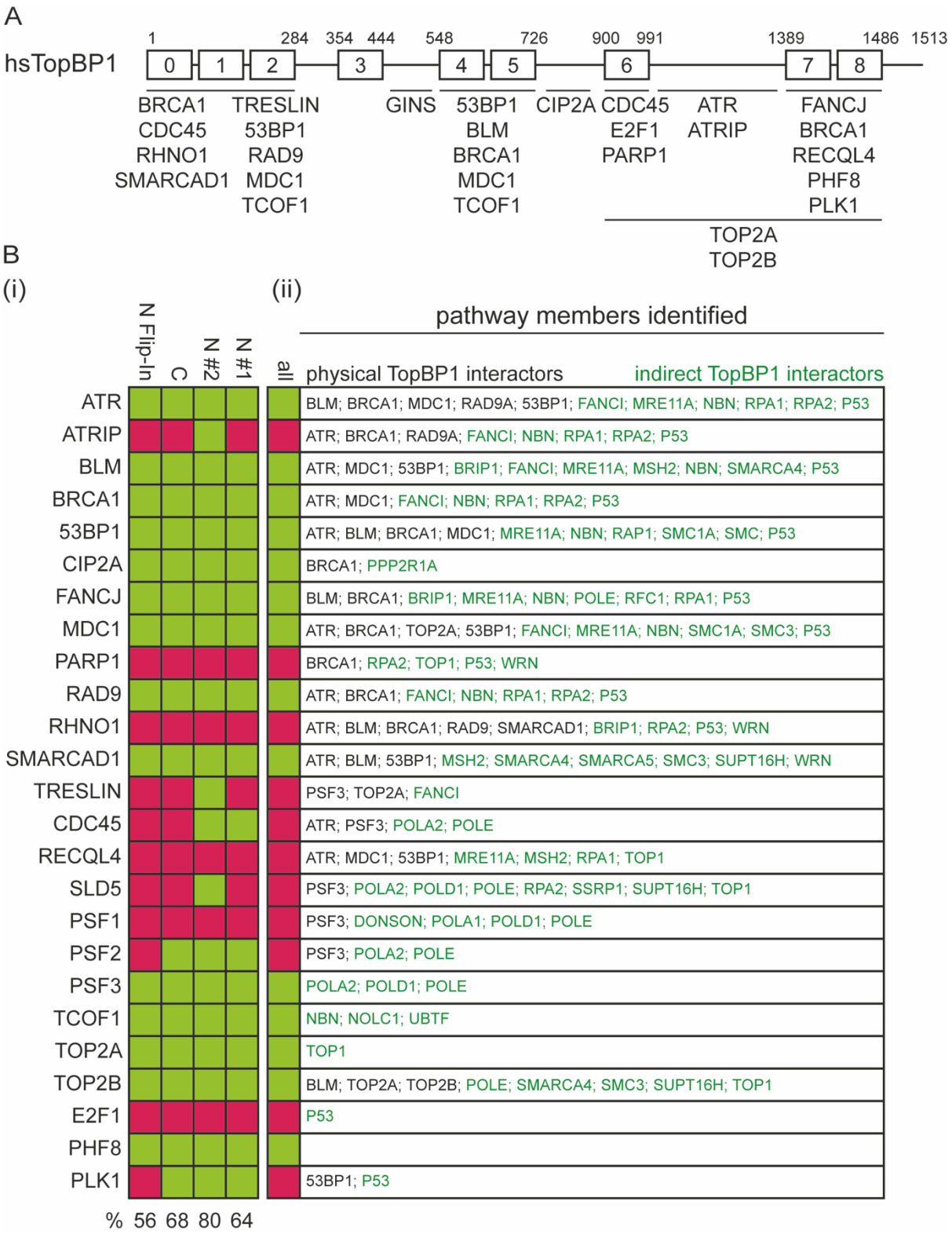



**Figure 14: Direct physical and indirect TopBP1 interactors identified by APEX2-mediated proximity biotinylation.**
A) Scheme of TopBP1 and physical TopBP1 interactors. B) Identification of TopBP1 interactors (i) physical TopBP1 interactors illustrated in A). (ii) Pathway members identified. Black = direct physical TopBP1 interactors and green = transient TopBP1 interactors.

The high level of direct physical and indirect TopBP1 interactors present in all experiments further supports the specificity of the experimental system.

### 3.2.3 Highly enriched proteins reveal TopBP1 pathways

The common 462 hits present in all four experiments represent a group of presumably highly specific and therefore especially interesting TopBP1 interactors. To investigate this set of hits further, pathway enrichment analysis was performed. Pathway enrichment using Perseus is based on the Gene Ontology (GO) database, a knowledgebase containing information on the function of genes. The 462 identified proteins showed enrichments in 58 of these GO-term pathways (Appendix Table 1). Some of these pathways are large groups of proteins containing more than 3,000 proteins and are therefore not suitable for verification of specific hits. Despite these broad pathways, a variety of TopBP1 specific pathways like DNA replication, DNA repair, DNA recombination, response to DNA damage stimulus, (regulation of) DNA metabolic processes and cell cycle were enriched in the common hits (Table 2 dark green). In addition, also pathways involving TopBP1 interactors are enriched (Table 2 light green). These enrichments suggest specific biotinylation of proteins involved in the same pathway as TopBP1. Interestingly, several unexpected pathways were also enriched, among them a variety of transcription-related pathways (Table 2 light and dark blue respectively). To exclude unspecific pathways resulting from background biotinylation, an APEX2-NLS will serve as a control in future experiments.

**Table 2: Pathway enrichment of common proteins identified.**

| GO term | P value | Enrichment | Total | In cluster | Cluster size | Ben. Ho. FDR |
|---|---|---|---|---|---|---|
| nucleoplasm | 2,95E-28 | 1,6306 | 1006 | 273 | 462 | 3,92E-26 |
| RNA processing | 6,02E-23 | 2,0749 | 417 | 144 | 462 | 1,82E-20 |
| nucleolus | 1,10E-20 | 2,0378 | 401 | 136 | 462 | 7,34E-19 |
| RNA metabolic process | 2,74E-20 | 1,5494 | 954 | 246 | 462 | 4,13E-18 |
| nitrogen compound metabolic process | 1,79E-19 | 1,3911 | 1339 | 310 | 462 | 1,80E-17 |
| cellular nitrogen compound metabolic process | 2,54E-19 | 1,3925 | 1329 | 308 | 462 | 1,92E-17 |
| nucleobase-containing compound metabolic process | 9,12E-19 | 1,469 | 1088 | 266 | 462 | 5,51E-17 |
| mRNA processing | 1,65E-18 | 2,2589 | 266 | 100 | 462 | 8,33E-17 |
| RNA splicing | 5,90E-18 | 2,3512 | 230 | 90 | 462 | 2,55E-16 |

| | | | | | | |
|---|---|---|---|---|---|---|
| macromolecule metabolic process | 3,19E-14 | 1,2617 | 1624 | 341 | 462 | 1,20E-12 |
| nucleus | 3,23E-13 | 1,3316 | 1268 | 281 | 462 | 1,43E-11 |
| intracellular non-membrane-bounded organelle | 7,70E-12 | 1,4447 | 836 | 201 | 462 | 2,56E-10 |
| mRNA metabolic process | 5,41E-11 | 1,7327 | 378 | 109 | 462 | 1,82E-09 |
| DNA metabolic process | 1,34E-09 | 1,7943 | 288 | 86 | 462 | 4,06E-08 |
| spliceosomal complex | 5,84E-08 | 2,1671 | 122 | 44 | 462 | 9,72E-07 |
| chromosome organization | 4,11E-08 | 1,668 | 317 | 88 | 462 | 1,13E-06 |
| DNA repair | 1,75E-07 | 1,9528 | 160 | 52 | 462 | 3,52E-06 |
| chromosome | 3,13E-07 | 2,2859 | 92 | 35 | 462 | 4,63E-06 |
| nucleobase-containing compound metabolic process | 6,96E-07 | 1,9357 | 149 | 48 | 462 | 1,24E-05 |
| transcription termination, DNA-dependent | 1,59E-06 | 2,546 | 59 | 25 | 462 | 2,66E-05 |
| primary metabolic process | 1,89E-06 | 1,1327 | 1846 | 348 | 462 | 3,01E-05 |
| biological regulation | 2,67E-06 | 1,8133 | 169 | 51 | 462 | 4,04E-05 |
| rRNA processing | 3,02E-06 | 2,1553 | 92 | 33 | 462 | 4,34E-05 |
| intracellular organelle | 1,47E-05 | 1,1011 | 2030 | 372 | 462 | 0,00016318 |
| preribosome | 1,61E-05 | 2,9133 | 33 | 16 | 462 | 0,00016435 |
| ribonucleoprotein complex | 3,39E-05 | 1,4269 | 379 | 90 | 462 | 0,00030037 |
| regulation of gene expression, epigenetic | 5,61E-05 | 2,027 | 83 | 28 | 462 | 0,00062703 |
| response to DNA damage stimulus | 8,94E-05 | 1,5142 | 250 | 63 | 462 | 0,00079424 |
| transcription from RNA polymerase I promoter | 8,28E-05 | 3,3048 | 20 | 11 | 462 | 0,00080673 |
| chromatin organization | 8,86E-05 | 1,5414 | 230 | 59 | 462 | 0,00081101 |
| nuclear transport | 7,61E-05 | 1,8287 | 115 | 35 | 462 | 0,00082122 |
| cellular metabolic process | 8,77E-05 | 1,1008 | 1894 | 347 | 462 | 0,00082751 |
| transcription from RNA polymerase II promoter | 0,00014343 | 1,7105 | 137 | 39 | 462 | 0,0011707 |

| | | | | | | |
|---|---|---|---|---|---|---|
| RNA polymerase complex | 0,00015691 | 3,6052 | 15 | 9 | 462 | 0,0012276 |
| regulation of DNA metabolic process | 0,00017986 | 1,9314 | 84 | 27 | 462 | 0,0013928 |
| transcription elongation, DNA-dependent | 0,00020093 | 2,2385 | 51 | 19 | 462 | 0,001517 |
| nucleocytoplasmic transport | 0,00026683 | 1,7547 | 113 | 33 | 462 | 0,0019654 |
| DNA replication | 0,00032826 | 1,8648 | 87 | 27 | 462 | 0,0023603 |
| chromatin modification | 0,0004073 | 1,5341 | 188 | 48 | 462 | 0,0028606 |
| RNA modification | 0,00048056 | 2,6705 | 27 | 12 | 462 | 0,0032984 |
| metabolic process | 0,0005742 | 1,0767 | 2009 | 360 | 462 | 0,0036896 |
| telomere organization | 0,00081807 | 2,2532 | 40 | 15 | 462 | 0,0048442 |
| nucleobase-containing compound transport | 0,00095063 | 1,6847 | 107 | 30 | 462 | 0,0054168 |
| DNA geometric change | 0,0019135 | 2,2318 | 35 | 13 | 462 | 0,010138 |
| meiotic cell cycle | 0,0021139 | 3,6052 | 10 | 6 | 462 | 0,010821 |
| DNA recombination | 0,0023538 | 1,7275 | 80 | 23 | 462 | 0,011848 |
| transcription initiation, DNA-dependent | 0,0030393 | 1,6273 | 96 | 26 | 462 | 0,015047 |
| protein modification by small protein conjugation or removal | 0,0033204 | 3,0043 | 14 | 7 | 462 | 0,015917 |
| transcription elongation factor complex | 0,0024075 | 2,5036 | 24 | 10 | 462 | 0,01601 |
| cytoplasm | 0,0028057 | 1,1184 | 1139 | 212 | 462 | 0,017769 |
| nuclear envelope | 0,0030164 | 1,8916 | 54 | 17 | 462 | 0,018236 |
| macromolecular complex | 0,0034245 | 1,1004 | 1305 | 239 | 462 | 0,018977 |
| cell cycle | 0,0046077 | 1,3122 | 261 | 57 | 462 | 0,021084 |
| nuclear chromosome | 0,0043308 | 2,1207 | 34 | 12 | 462 | 0,022154 |
| biosynthetic process | 0,0049219 | 1,1183 | 994 | 185 | 462 | 0,022185 |
| organelle | 0,0058059 | 1,0437 | 2228 | 387 | 462 | 0,028599 |
| cellular macromolecular complex assembly | 0,0084874 | 1,305 | 221 | 48 | 462 | 0,033288 |
| nucleus organization | 0,010018 | 2,072 | 29 | 10 | 462 | 0,03879 |

To narrow done the list of potential new TopBP1 interactors, the list of common hits was compared to the 10 % highest enriched proteins in each individual experiment. These 10 % best hits were determined by difference from the non-biotinylated control. These highly enriched proteins are unlikely to be the result of background biotinylation but are presumably true TopBP1 interactors. Comparing these 10 % best hits with the 462 proteins identified in all four experiments revealed that 62 %, 37 %, 52 % and 100 % were also present in the list of common hits (Table 3).

**Table 3: 10 % hits in intersecting proteins identified.**

| | 293T | | | 293 Flip-In |
|---|---|---|---|---|
| | N #1 | N #2 | C | N |
| 10 % hits in total | 173 | 348 | 146 | 58 |
| 10 % hits in 462 | 107 | 129 | 76 | 58 |
| EQ | 62 % | 37 % | 52 % | 100 % |

These analyses illustrate various possibilities to study potential new TopBP1 interactors and their pathways. Moreover, the common hit and the 10 % hit lists are a good resource for future investigations.

### 3.3 TopBP1-GINS interaction is required to support replication in *Xenopus* egg extract

To activate the replicative helicase during origin firing, the helicase components GINS and Cdc45 have to be delivered to the inactive pre-RC. The molecular mechanism how TopBP1 is recruiting GINS to the Mcm2-7 helicase is largely unknown. As described above, four independent biotin labeling experiments using cells expressing APEX2-tagged TopBP1 resulted in specific enrichment of the GINS subunits Sld5, Pfs2 and Psf3. To examine the physiological importance of the TopBP1-GINS interaction, the earlier established *Xenopus* egg extract system was used to test the interaction betweenTopBP1 and GINS *in vivo*.

For this, a TopBP1 depletion/rescue system was established in the Boos laboratory. First, depletion efficiency was tested by depleting interphase egg extract with anti-TopBP1 antibodies or control antibodies bound to magnetic protein G beads. Different amounts of non-depleted egg extract were loaded side-by-side with depleted extract. Control-depleted extract showed similar TopBP1 signal intensities compared to 100 % untreated extract as expected (Figure 15 A lane 1 and 5). In contrast, TopBP1 depleted

extract did not show TopBP1 signals (Figure 15 A lane 6 and 7). This showed that more than 90 % of endogenous TopBP1 was removed from the extract (Figure 15 A lane 4, 6 and 7). To test whether TopBP1 is required for efficient replication in the cell-free system, egg extracts immunodepleted of TopBP1 were supplemented with sperm DNA and radioactive labeled $\alpha^{32}$P-dCTP and then the accumulation of newly synthesized DNA was tracked over time. As expected, depletion of TopBP1 resulted in a strong reduction of radioactive labeled dCTP signals compared to extracts depleted with control-IgG antibodies (mock; Figure 15 B). Moreover, efficient DNA replication in TopBP1-depleted extracts could be rescued by addition of recombinant TopBP1 BRCT0-5-WT, a truncated fragment of TopBP1 that is known to support replication in *Xenopus* egg extract (Kumagai et al., 2010) (Figure 15 B).

These immunodepletion experiments indicate that endogenous TopBP1 can successfully be removed from *Xenopus* egg extract resulting in a strong suppression of replication. Furthermore, replication similar to physiological levels can be observed when TopBP1 BRCT0-5-WT is added back to the depleted extract. In collaboration with Bilal Tetik the binding of TopBP1 to GINS was test *in vitro* using pulldown assays. Biochemical binding assays using GINS immobilized on magnetic anti-FLAG beads have shown that a TopBP1 fragment containing all functional domains that also exists in yeast Dpb11 bound to GINS (Figure 15 C). Removing the vertebrate specific BRCT3 domain in TopBP1 did not affect this binding (Figure 15 D). Similar, deletion of the Treslin interaction domain BRCT0/1/2 did not compromise the interaction with GINS, indicated by a clear shift upon complex formation using size-exclusion chromatography (Appendix Figure 1; in collaboration with Matthew Day).

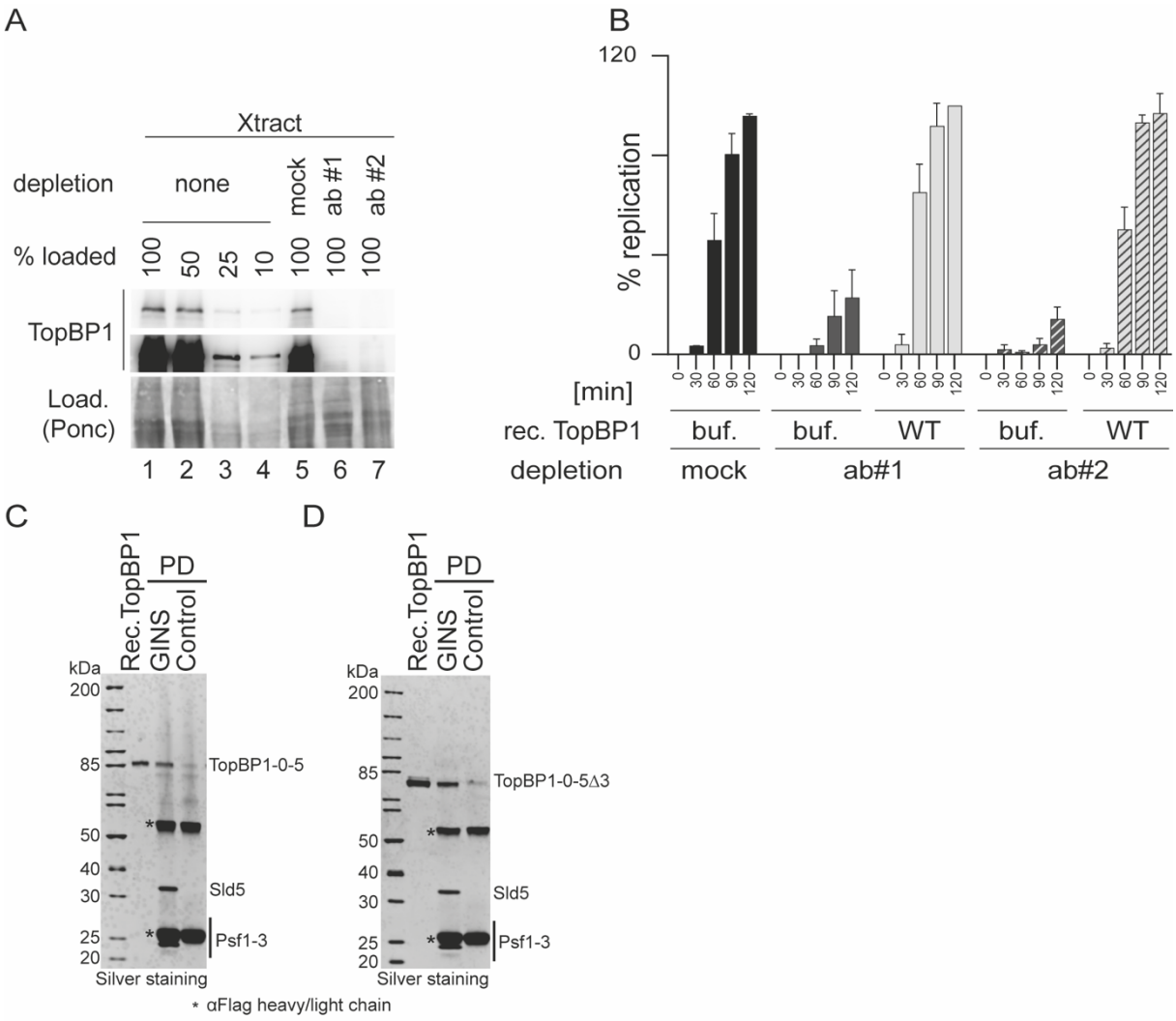



**Figure 15: TopBP1 BRCT0-5 interacts with GINS *in vivo* and *in vitro.***
A) The indicated relative amounts (100 % - 10 %) of immunodepleted and non-depleted interphase egg extract (Xtract) were immunoblotted using TopBP1 antibody #2. TopBP1 antibodies #1 and #2 and IgG (mock) were used for immunodepletion. Ponc, PonceauS staining. B) Replication analysis of interphase egg extract immunodepleted with indicated antibodies upon adding back buffer (buf.), or recombinant TopBP1 BRCT0-5 wild type (WT). C) and D) Immobilized recombinant GINS was used to pull down TopBP1 BRCT0-5 (C) or TopBP1 BRCTΔ3 (D). SDS gels were silver stained.

These data confirmed the interaction of GINS with a TopBP1 fragment containing BRCT0-5 *in vivo* and showed that BRCT0/1/2 and BRCT3 are dispensable for GINS binding *in vitro*.

### 3.3.1 TopBP1 GINI domain is required for GINS binding

Previous yeast two-hybrid experiments revealed a binding between the GINS subunit Psf1 and the GINS interaction (GINI) region in TopBP1. This GINI region had been loosely mapped between TopBP1 BRCT3 and BRCT4, a region with little sequence similarity to the know GINI region in yeast Dpb11 (Tanaka et al., 2013). Therefore, a multiple amino acid sequence alignment across this region was performed using

T-COFFEE. This alignment implied a high degree of conserved amino acids between vertebrate species. Almost identical amino acids were identified within the section from Glu484 to Tyr494. Such highly conserved regions are thought to have functional values. To narrow down the GINI domain sequence, this highly conserved region and various other regions between BRCT3 and BRCT4 were mutated. The following TopBP1 mutants were generated: TopBP1-Gcc (GINI center core), Gcore (GINI core), and RI-III (region I-III) (Figure 16).

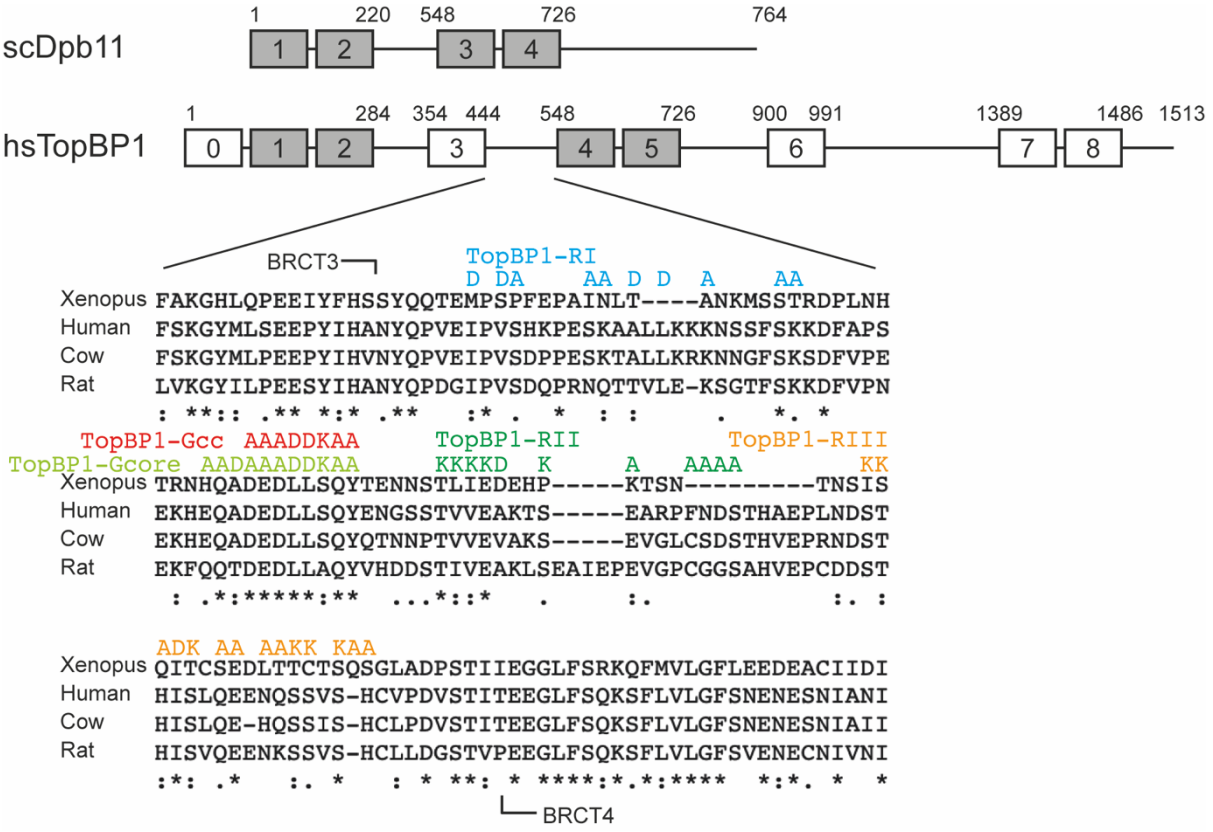



**Figure 16: Alignment of TopBP1 inter-BRCT region between BRCT3 and BRCT4**
Top: Schematic comparison between *Saccharomyces cerevisiae* Dpb11 (scDpb11) and *Homo sapiens* TopBP1 (hsTopBP1). Bottom: Alignment of TopBP1 from different vertebrate species using T-COFFEE and relevant position to the TopBP1 BRCT domains. Xenopus, *Xenopus laevis* (Q7ZZY3), Cow, *Bos taurus* (A0A3Q1LWE4), Rat, *Rattus norvegicus* (A0A8I6GFZ6). Colored labeling shows the amino acids mutated between TopBP1 BRCT3 and BRCT4.

All mutants purified from Sf9 insect cells were generated in the background of the TopBP1 BRCT0-5 fragment shown to be sufficient to bind to GINS *in vitro* and support replication in egg extract (Figure 15 B and C).

The capacity of TopBP1 mutants to support replication was tested in immunodepleted egg extracts (Figure 17 A and B). While TopBP1 BRCT0-5 WT and RI-III supported replication to similar levels as the mock control, Gcore and Gcc mutants showed moderate reduction in replication (Figure 17 C). The moderate effect was somewhat

surprising given that recruitment of GINS to the helicase is essential for replication. In collaboration with Bilal Tetik the capacity of these TopBP1 mutants to bind to GINS was tested by pulldown assays. Pulldown of TopBP1 WT, Gcore, Gcc and RI-III mutants revealed that RI-III interacted with GINS like TopBP1 WT (Figure 17 D lane 1, 2, 5 and 6) whereas Gcc and Gcore did not coelute with GINS (Figure 17 D lane 3 and 4). The control pulldown with empty beads did not show any signals using the same TopBP1 fragments (Figure 17 D lane 7-12), confirming the specificity of the binding assay.

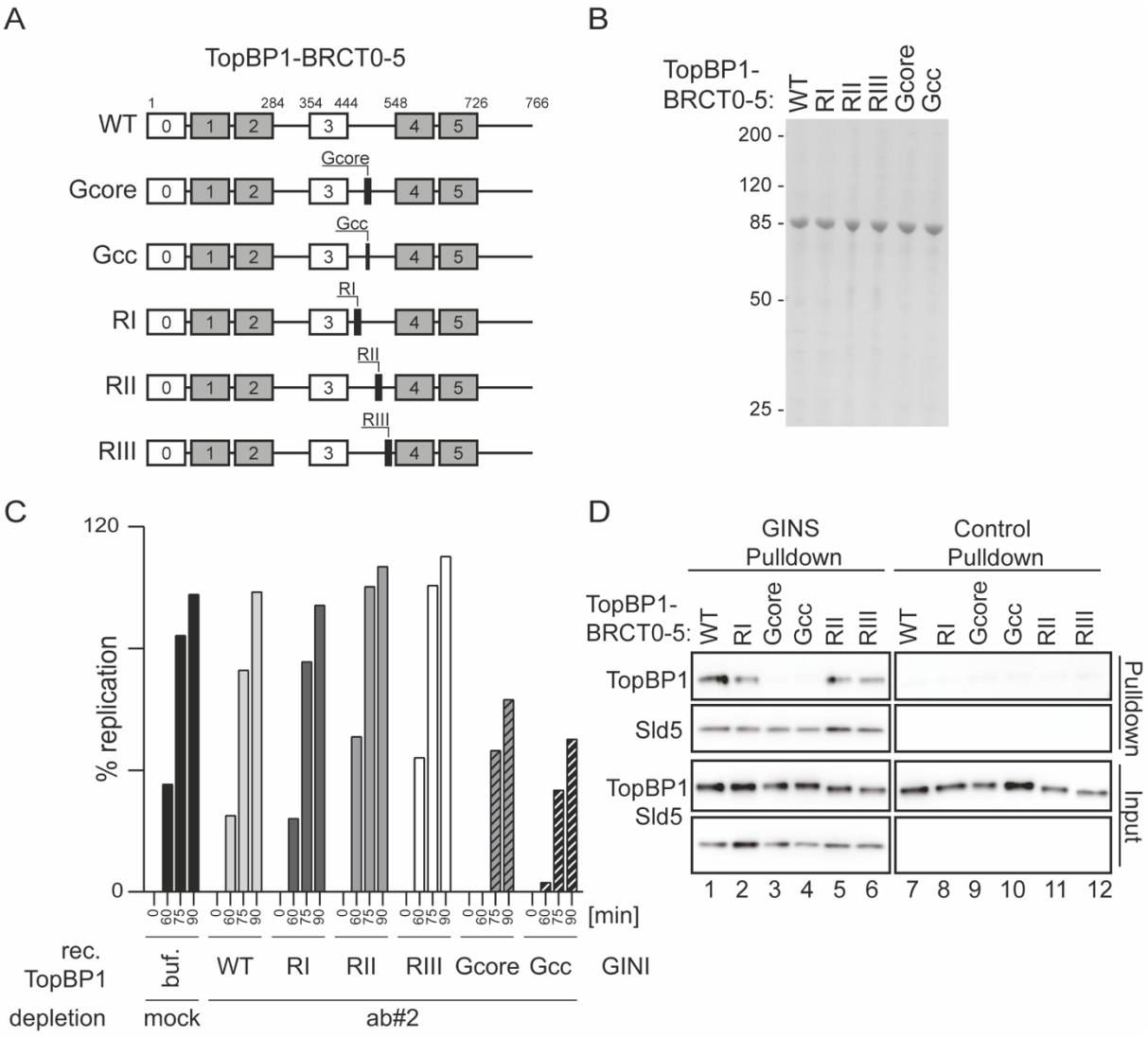



**Figure 17: The conserved GINI core of TopBP1 is required for normal replication in *Xenopus* egg extract and for *in vitro* binding to GINS.**
A) Scheme of TopBP1 WT and indicated mutants. B) Coomassie-stained SDS-PAGE of the indicated recombinant TopBP1 BRCT0-5 fragments. C) Replication analysis of interphase egg extract immunodepleted with antibody #2 upon adding back buffer (buf.), or recombinant TopBP1 BRCT0-5 wild type (WT) or TopBP1 BRCT0-5 mutants shown in B. D) Pulldown of TopBP1 BRCT0-5 wild type (WT) or indicated TopBP1 BRCT0-5 mutants using recombinant GINS immobilized on beads. Immunoblotting analysis using the indicated antibodies.

These experiments demonstrate that the TopBP1 GINI core region is required to bind to GINS *in vitro* and for normal replication in *Xenopus* egg extract.

### 3.3.2 The GINI-core helix binds a surface comprising the Psf1 and Psf3 regions

To characterize the interaction between the TopBP1 GINI core and GINS in more detail, structural predication by AlphaFold2 was used to identify the potential binding surface. The predicted AlphaFold2 structure revealed a short α helix within the GINI core region. To investigate if the helix core is required for GINS binding, the central leucine-leucine motif was mutated to two prolines to break the helix (Figure 18 A). Next, the capability to support replication in interphase egg extract of this TopBP1 Gpp (GINI LL491/492PP) mutant was tested. Compared to the WT, the Gpp helix break mutant showed a moderate reduction in replication, similar to the signals observed using the Gcc mutant (Figure 18 B and Appendix Figure 2 A). Biochemical binding assays using GINS immobilized on magnetic beads demonstrated that TopBP1 Gpp, like TopBP1 Gcc, compared to the WT was deficient to bind to GINS (Figure 18 C lane 1-3).

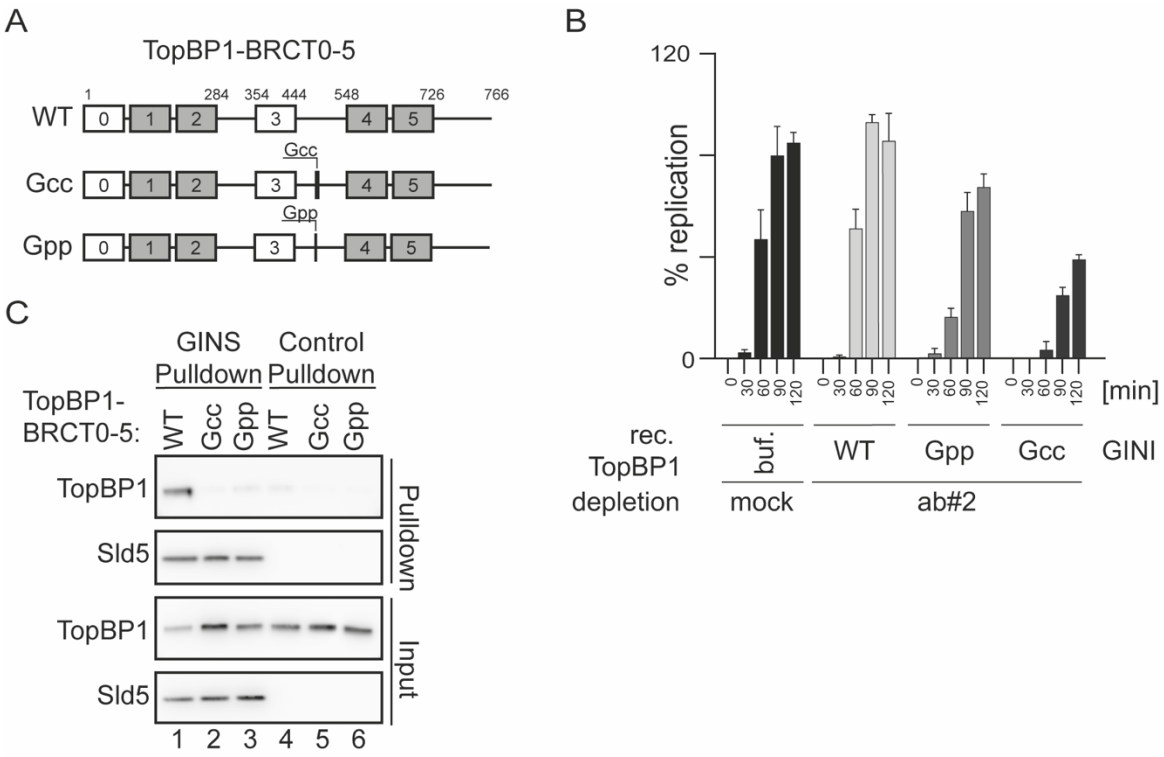



**Figure 18: The TopBP1 GINI core helix is required for GINS binding and normal replication in *Xenopus* egg extract.**
A) Scheme of TopBP1 BRCT0-5 wild type (WT), conserved core (Gcc) and helix break (Gpp) mutants. B) Replication analysis of interphase egg extract immunodepleted with antibody #2 upon adding back buffer (buf.), or recombinant TopBP1 BRCT0-5 wild type (WT) or TopBP1 BRCT0-5 mutants. C) Pulldown of TopBP1 BRCT0-5 wild type (WT) or indicated TopBP1 BRCT0-5 mutants using recombinant GINS immobilized on beads. Immunoblotting analysis using the indicated antibodies.

This data suggests that the GINI core helix predicted by AlphaFold2 is required to support normal replication in *Xenopus* egg extract and for binding to GINS *in vitro*.

To test if the GINI core helix contributes to the GINS interaction surface, fluorescence polarization (FP) was used to measure the interaction between a fluorescently labeled GINI core peptide and the GINS tetramer. In collaboration with Matthew Day this FP assays determined that the GINI helix, but not the Gpp helix break mutant, was able to bind to GINS (Appendix Figure 3). This data suggested that the TopBP1 GINI core helix constitutes a binding site for GINS. To identify and confirm the sites in GINS that binds to the GINI core helix, potential binding sites were determined using AlphaFold2 and computational molecular docking simulations (in collaboration with Dr. Yasser Almeida Hernández). Both analyses suggested an interaction of the GINI core helix with the distal part of the GINS Psf1 A-domain. BS3-crosslinking mass spectrometry experiments confirmed the proximity of the GINI core helix and the GINS Psf1 A-domain (Appendix Figure 4; in collaboration with Heike Siegert). Moreover, cryo-EM mapping of the TopBP1-GINS interaction allowed a short helical element, consistent

with the structural prediction (in collaboration with Matthew Day). Docking of the cryo-EM map and the predicted structure suggested that the hydrophobic side of the GINI-core helix interacts across two helices of the Psf1 A-domain. In this model the side chains of Leu490 and Leu491 of the GINI core helix made contact to Val47, Lys51, Ile59, Ile62, Lxs63 and His66 of Psf1. Additionally, the side chain of Tyr494 of the GINI helix is inserted into a small pocket formed by Tyr40, Asn43, Leu69 and His66 of Psf1 (Figure 19).

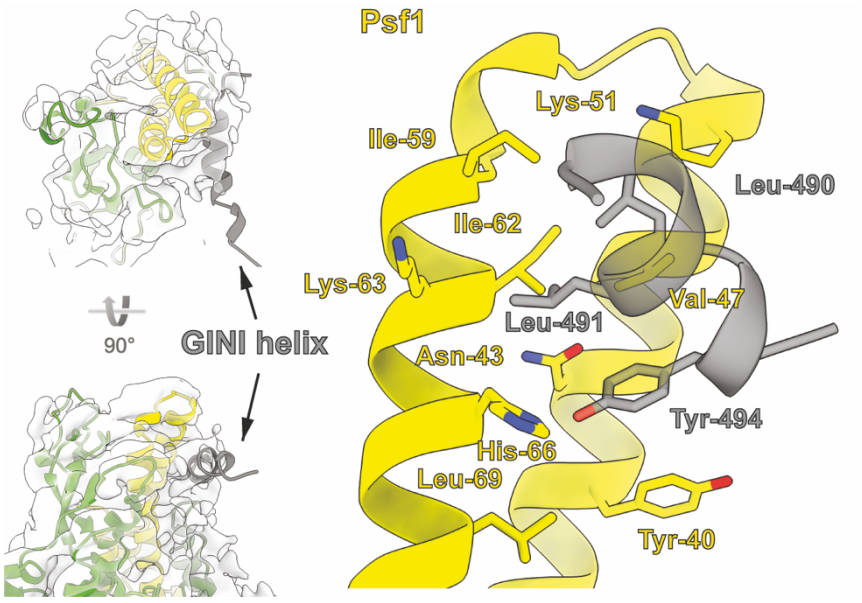



**Figure 19: Cryo-EM volume and AplphaFold2 predication of the interaction between the GINI core helix and GINS**
Cryo-EM volume showing an extra-helical density ascribed to the GINI helix (left) alongside a structural mode (right) based on this cryo-EM and an AlphaFold2 predicted structure.

This data suggested two pairs of residues in Psf1, to be critical for the interaction between the GINI helix and Psf1: Asn43/Lys51 and Lys63/His66. As part of the pocket, Asn43 and His66 make contacts to Tyr494 in the GINI domain. Additionally, the lysine residues in both pairs could form part of the hydrophobic grove that accommodates the leucine residues of the GINI core helix. To validate the interaction between these two amino acid pairs of Psf1 and the GINI core helix, both pairs were mutated to alanine and glutamate: Psf1-NK43/51/AE and Psf1-KH63/66AE. Only Psf1-NK43/51AE was still able to form the GINS tetramer, but its ability to bind to TopBP1 BRCT0-5 was severely compromised (Appendix Figure 5 A; in collaboration with Bilal Tetik). This data provides further evidence of the interaction between the TopBP1 GINI core helix and the distal end of the Psf1 A-domain and the adjacent region in Psf3.

In summary, these experiments demonstrate that the GINI core helix is essential for the binding to GINS *in vitro*. Interestingly, mutations of the GINI core helix have only moderate effects on replication in *Xenopus* egg extract, suggesting that TopBP1 might contains an additional GINS binding motif.

### 3.3.3 TopBP1 BRCT4 is a novel GINS interaction interface

Evidence for a second GINS interaction surface in TopBP1 came from interaction studies. As described above, a TopBP1 BRCT0-5 fragment was sufficient to bind to GINS *in vitro* (Figure 20 C). As shown before, deletion of either the triple BRCT0/1/2 domain or BRCT3 did not compromise the binding to GINS (Figure 15 D). In contrast, deletion of BRCT4/5 abrogates the binding of TopBP1 to immobilized GINS (Figure 20 B). This observation was surprising because it is known that a TopBP1 fragment comprising BRCT0-3 is sufficient to support replication in *Xenopus* egg extract. To confirm this, the ΔBRCT4/5 deletion mutant was added to TopBP1-depleted interphase extract. TopBP1 ΔBRCT4/5 indeed rescued replication to the same level as TopBP1 BRCT0-5 (Figure 20 C and Appendix Figure 2 A).

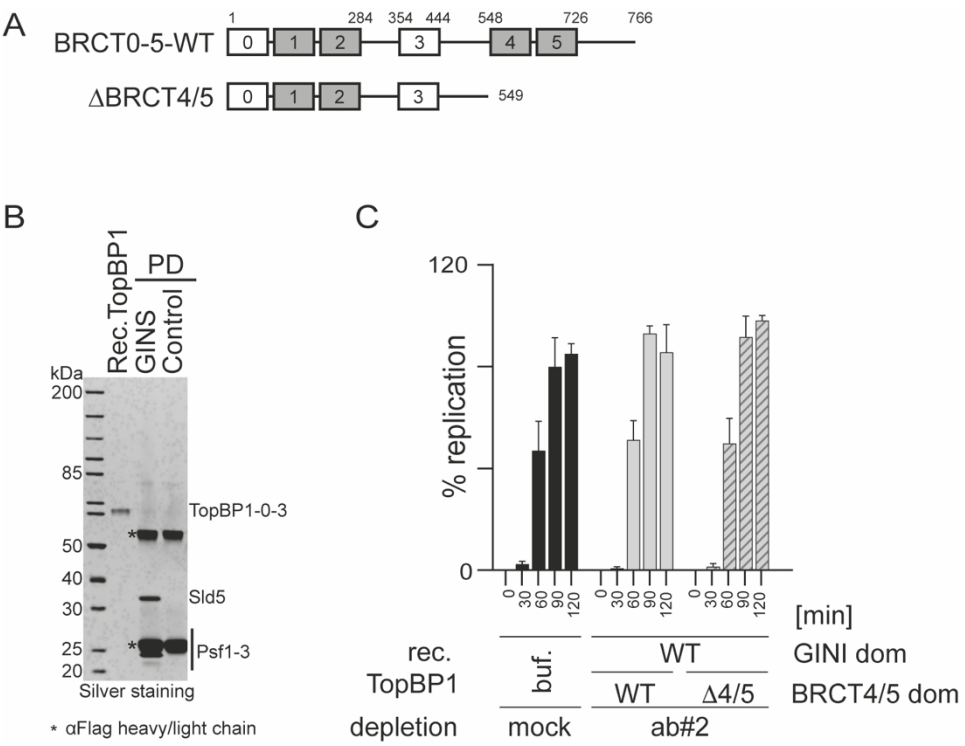



**Figure 20: TopBP1 BRCT4/5 is required for GINS binding *in vitro* but not for replication in *Xenopus* egg extract.**
A) Scheme of TopBP1 BRCT0-5 wild type and TopBP1 ΔBRCT4/5. B) Immobilized recombinant GINS was used to pull down TopBP1 BRCTΔ4/5. SDS gels were silver-stained. C) Replication analysis of interphase egg extract immunodepleted with antibody #2 upon adding back buffer (buf.), or recombinant TopBP1 BRCT0-5 wild type (WT) or TopBP1 BRCTΔ4/5.

These contrasting results suggest a low-affinity binding of GINS to TopBP1 that is not stable in pulldown analyses but is sufficient for replication in *Xenopus* egg extract.

Due to the conflicting results, the binding of BRCT4/5 to GINS appeared interesting and was therefore characterized in more detail. To narrow down the residues in TopBP1 BRCT4/5 that interact with GINS, the crystal structures of the GINS complex and the BRCT4/5 domain of TopBP1 was docked into the volume determined by cryo-EM. This data indicated that BRCT4 directly interacts with the Psf1 subunit of GINS and faces towards the A-domain and in the linker region between the A- and the B-domain (Figure 21 A). The model suggested that the interacting residues of BRCT4 are located at the edge of the central four-stranded β-sheet and in the loop connecting the strands. More precisely, the model predicted amino acid residues Val590, Thr606, and Val610 to be involved in a hydrophobic interface that directly interacts with the Psf1 side chains Trp92 and Ile97 (Figure 21 B). For verification, the indicated TopBP1 residues were mutated to glutamate, arginine, and glutamate, respectively. Pulldown experiments using immobilized GINS, demonstrated that neither this BRCT4 mutant (B4mut) nor ΔBRCT4/5 bound to GINS (Figure 21 C lane 1-3). Like the BRCT4/5 deletion mutant, the point mutant did not show a defect in supporting replication in interphase *Xenopus* egg extract (Figure 21 D). Furthermore, mutation of the opposing GINS Psf1 residue Ile97 to an arginine showed a weaker binding to TopBP1 compared to TopBP1 BRCT0-5 wild type (Appendix Figure 5 B; in collaboration with Bilal Tetik).

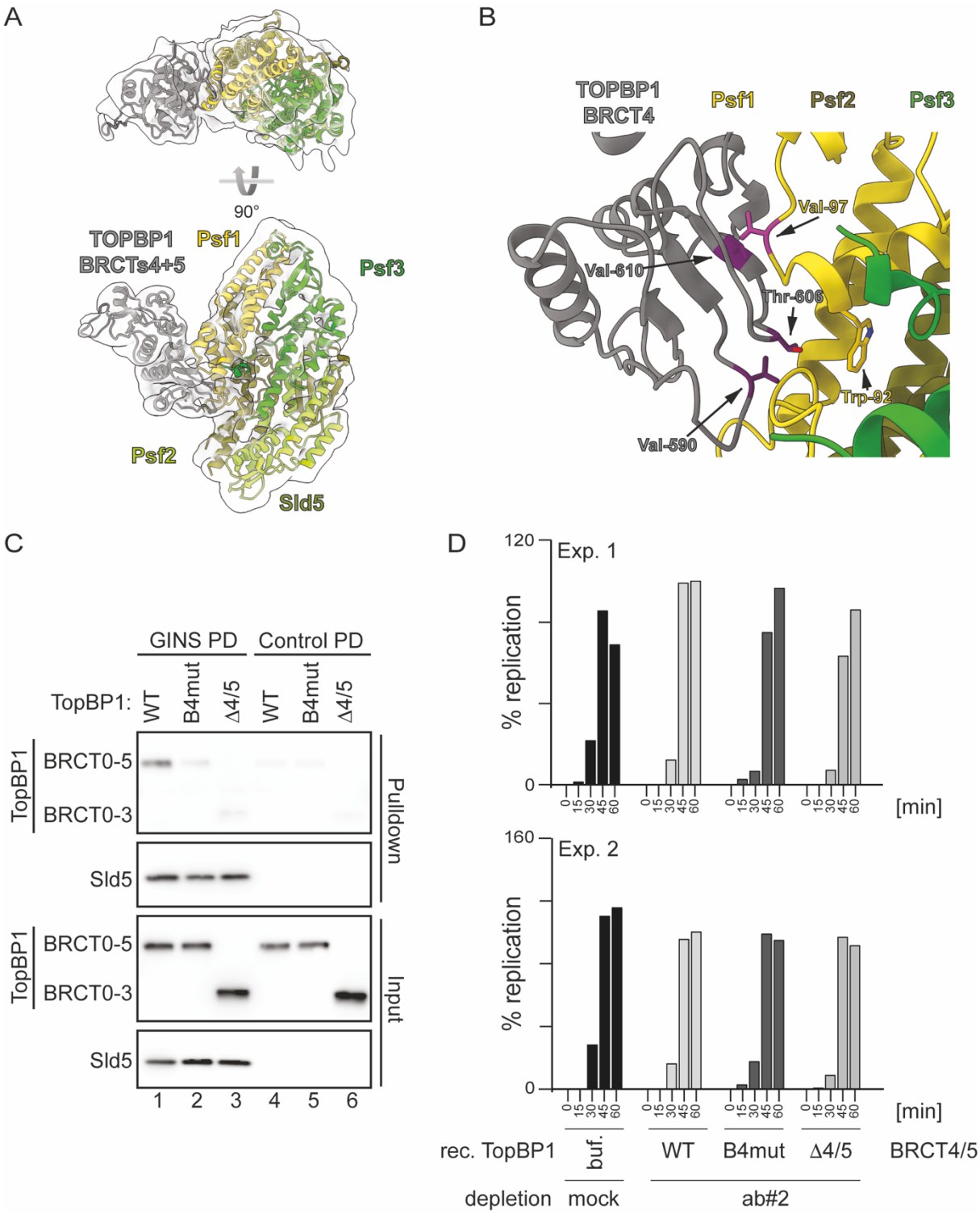



**Figure 21: TopBP1 BRCT4 binds to GINS Psf1 A-domain.**
A) Model of the GINS TopBP1 BRCT4/5 complex generated by docking the crystal structures of the GINS complex and the central BRCT4/5 domain of TopBP1 into the volume determined by cryo-EM. B) Pulldown of the indicated TopBP1 mutants using immobilized recombinant GINS. Immunoblotting analysis using the indicated antibodies. C) Replication analysis of interphase egg extract immunodepleted with antibody #2 upon adding back buffer (buf.), recombinant TopBP1 BRCT0-5 wild type (WT) or indicated TopBP1 mutants.

In conclusion, we describe for the first time a specific molecular function of TopBP1 BRCT4. That BRCT4 is not required for replication in egg extract was surprising. We

hypothesize that the residual binding of GINI is sufficient but that in TopBP1 WT, both interaction surfaces collaborate to bind GINS to support replication.

### 3.3.4 Cooperation of TopBP1 GINI and BRCT4 to bind GINS

To verify if both the GINI and the BRCT4 binding activities cooperate, the capability of single and double mutants to support replication in *Xenopus* egg extract was tested. If the two GINS binding motifs cooperate, TopBP1 double mutants would be deficient in supporting replication in interphase egg extract. As observed before, deleting or mutating the BRCT4 domain resulted in $\alpha^{32}$P-dCTP incorporation levels similar to TopBP1 BRCT0-5 WT (Figure 21 B and Figure 22 B) whereas mutations in the GINI domain caused a moderate reduction in replication (Figure 18 B and Figure 22 B). In contrast, combining the BRCT4/5 deletion with either the Gcc or Gpp mutation, reduced replication to nearly background levels (Figure 22 B (i) and Appendix Figure 2 A). Double mutants harboring the BRCT4 point mutation in combination with the GINI center core mutation were affected to the same levels as the BRCT4/5 deletion double mutant (Figure 22 B (ii)). Consistent with this data, estimated binding energies indicated that the affinity of each individual binding surface is lower than the energy required for simultaneous binding (in collaboration with Dr. Yasser Almeida Hernández; Appendix Figure 2 B). In collaboration with Bilal Tetik, the binding of these mutants was tested in pulldown assays. These biochemical binding assays confirmed that neither the single nor the double mutants bound to GINS compared to TopBP1 BRCT0-5 WT (Figure 22 C lane 1-6).

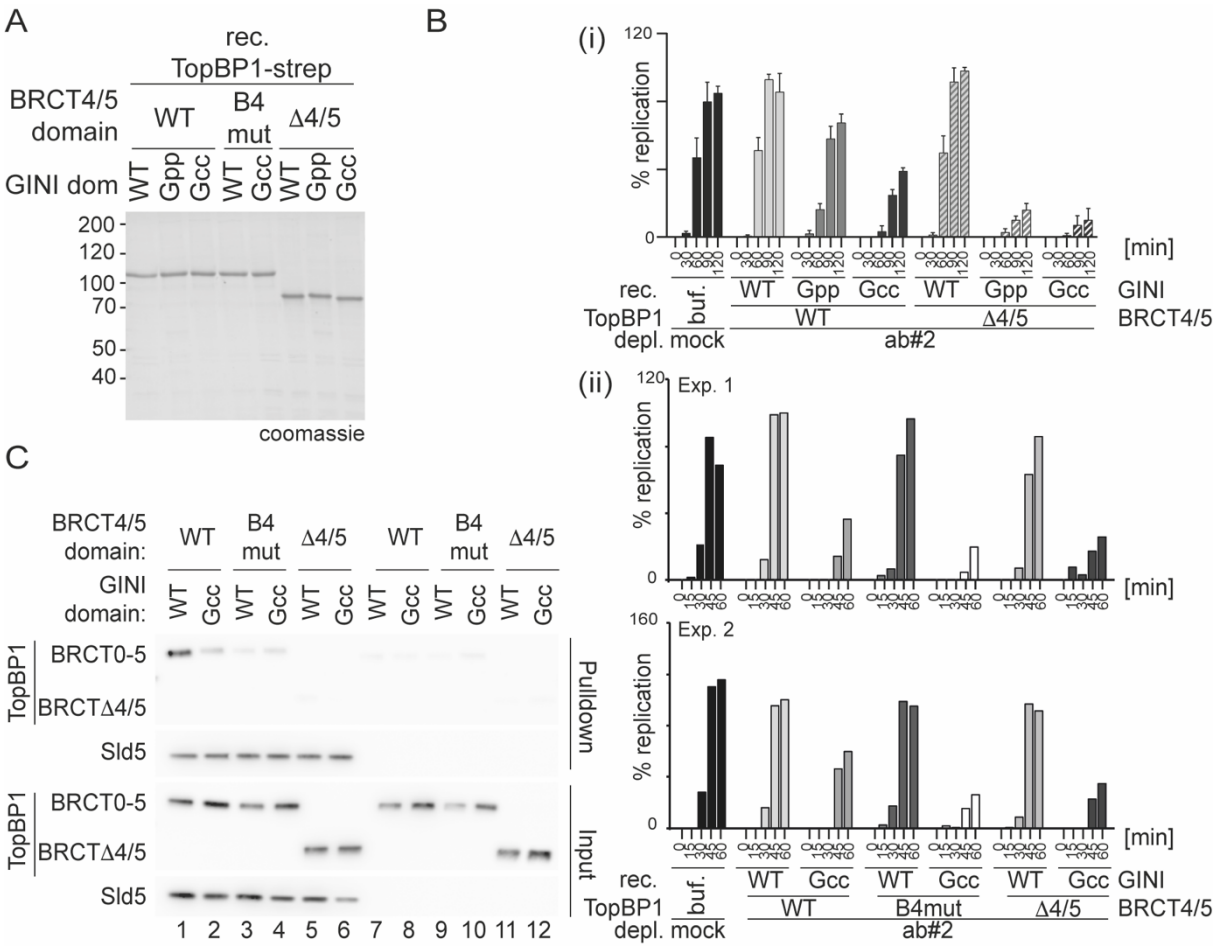



**Figure 22: Cooperation of both GINS binding surfaces in TopBP1 is required to support DNA replication in *Xenopus* egg extract.**
A) Coomassie stained SDS-PAGE gel of the indicated recombinant TopBP1 BRCT0-5 proteins. B) Replication analysis of interphase egg extract immunodepleted with antibody #2 upon adding back buffer (buf.), recombinant TopBP1 BRCT0-5 wild type (WT) or indicated TopBP1 mutants. C) Pulldown of the indicated TopBP1 mutants using immobilized recombinant GINS. Immunoblotting analysis using the indicated antibodies.

These results indicate that either a functional GINI or BRCT4 domain is sufficient to support replication whereas double mutants are compromised in supporting replication in *Xenopus* egg extract. Taken together, these observations strengthen the model that both GINS binding surfaces in TopBP1 cooperate to form a composite interaction surface.

### 3.3.5 TopBP1 interaction with GINS is required for origin firing

Chromatin isolation was used to investigate if the TopBP1-GINS interaction is required for origin firing. Isolated chromatin and chromatin-bound proteins were analyzed by immunoblotting and mass spectrometry. For Western blotting, either TopBP1 BRCT0-5 WT, Gcc, ΔBRCT4/5 or Gcc-ΔBRCT4/5 was added to TopBP1-immunodepleted extract prior to sperm addition. All samples were supplemented with

the B-type polymerase inhibitor aphidicolin to prevent termination. Consequently, replisomes accumulate, and signals of chromatin-bound proteins increase. Comparing TopBP1 BRCT0-5 WT samples with and without geminin demonstrated licensing-dependent chromatin binding of the helicase Mcm2-7 (Figure 23 A lane 2 and 3). Chromatin binding of the helicase subunits Mcm4 and Mcm6 was not affected by a buffer control compared to a sample containing TopBP1 BRCT0-5 WT, indicating that licensing occurred independent of TopBP1 (Figure 23 A lane 1 and 2). The replisome markers Cdc45, GINS (Sld5), Polε (PolE2), and PCNA were detected in immunodepleted extract containing TopBP1 BRCT0-5 WT (Figure 23 A lane 2). Similar to their effects in rescue experiments, chromatin samples using the ΔBRCT4/5 or Gcc single mutant showed moderate and strong reduction in replisome marker signals, respectively (Figure 23 A lane 4 and 5). In the chromatin fraction isolated from extracts supplemented with buffer or the Gcc-ΔBRCT4/5 double mutant, signal intensities for replisome markers were significantly reduced (Figure 23 A lane 1 and 6).

For mass spectrometry using the CHROMASS protocol, TopBP1-depleted extracts were supplemented with TopBP1 BRCT0-5 WT or the Gcc-ΔBRCT4/5 double mutant. Similar to immunoblotting analyses, all samples were supplemented with aphidicolin to allow accumulation of replisomes on chromatin. To control licensing-dependent replication, one sample was supplemented with geminin in addition to TopBP1 BRCT0-5 WT. CHROMASS analysis also showed geminin-sensitive chromatin binding of Mcm2-7. Irrespective of the addition of buffer or TopBP1 BRCT0-5 WT, all six subunits of the Mcm2-7 helicase were identified, confirming that licensing occurred independent of TopBP1 (Figure 23 B and D). On the other hand, replisome markers such as GINS and Cdc45 were identified on the chromatin from extract containing the wild type and were significantly reduced when buffer or the Gcc-ΔBRCT4/5 mutant was added (Figure 23 B and E). In contrast to immunoblotting, the CHROMASS protocol can detect more pathway proteins, including TopBP1-related pathways that can be linked with DNA repair at the replication fork (Figure 23 B and C). These pathway proteins are a good source for future investigations.

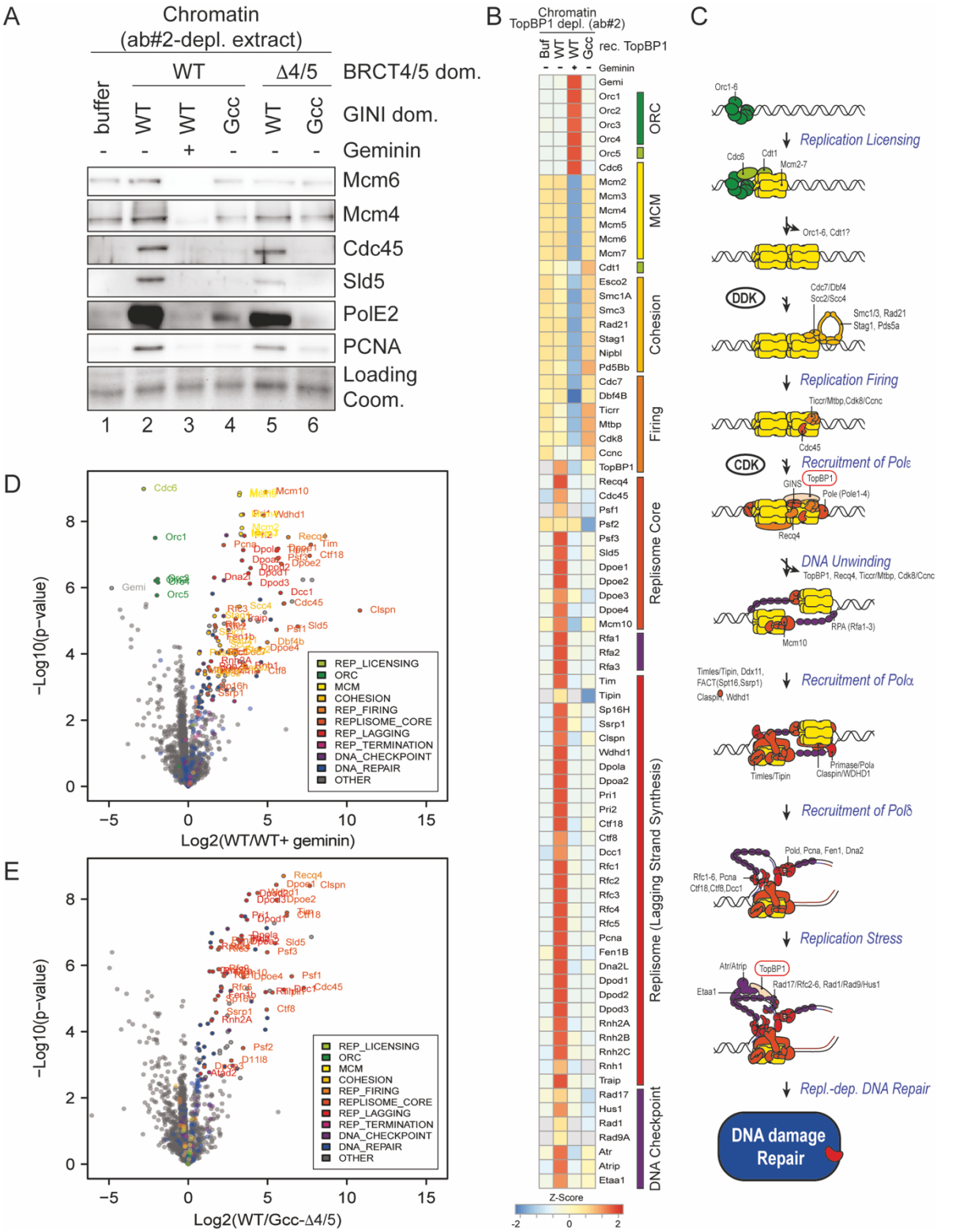



**Figure 23: GINS binding to TopBP1 is required for origin firing in *Xenopus* egg extract.**
A) Immunoblot analysis of chromatin-bound proteins isolated from TopBP1-immunodepleted egg extract upon addition of buffer (buf.) or indicated TopBP1 BRCT0-5 proteins. B) - E) Label-free quantification of chromatin-bound proteins by mass spectrometry upon chromatin isolation from interphase egg extract after replacing endogenous TopBP1 with indicated TopBP1 mutants. B) Semi-quantitative heat map, C) schematic model of DNA replication, D) and E) volcano plots highlighting a selection of differentially detected proteins comparing D) wild type (WT) and geminin-treatment (WT + geminin) or E) wild type (WT) with double mutant (Gcc-ΔBRCT4/5).

In summary, these data demonstrate that the interaction between TopBP1 and GINS is required for replisome formation upstream of CMG formation and therefore presumably also for origin firing. This in turn leads to the conclusion that the two GINS binding motifs in TopBP1 cooperate to assemble the CMG helicase.

## 3.4 *In vitro* pre-IC reconstitution as an outlook to investigate origin firing

Pre-IC assembly is the major regulatory step of CMG formation. Our data demonstrated how TopBP1 recruits GINS to the replicative helicase and therefore provides further detailed understanding of CMG formation. To investigate the molecular details of pre-IC formation in vertebrates, the long-term aim would be to reconstitute origin firing in vertebrates. Within this research project, pre-RC isolated from *Xenopus* egg extract were characterized and could serve in the future as a substrate for pre-IC reconstitution.

### 3.4.1 Isolation of chromatin bound pre-RCs from *Xenopus* egg extract

To guarantee isolation of pre-RCs, formation of downstream intermediates or replisomes had to be inhibited. Two methods to isolate pre-RCs from egg extract were established in the Boos laboratory. To block initiation after licensing, CDK levels were decreased either with the CDK inhibitor roscovitine or Triton X-100. As a detergent, Triton X-100 prevents nuclei formation and therefore the rise of CDK levels required for origin firing. Immunoblotting of chromatin bound pre-RCs isolated from either roscovitine- or Triton X-100-treated extracts showed geminin-sensitive chromatin binding of Mcm2-7 subunits (Figure 24 A and B).

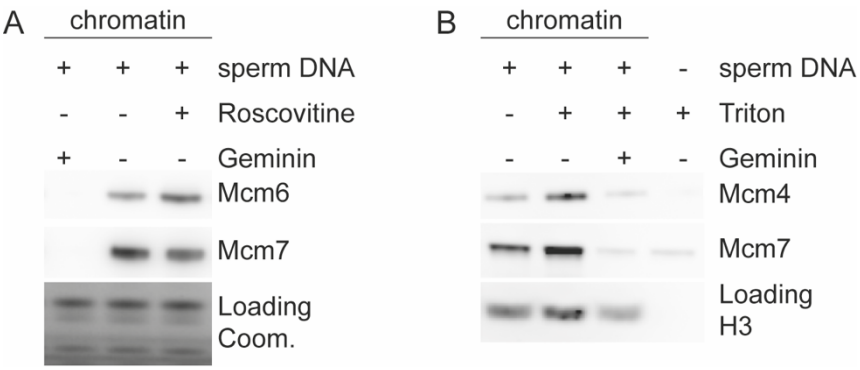



**Figure 24: Isolation of pre-RC containing chromatin from *Xenopus* egg extract.**
A) and B) Immunoblot analysis of chromatin-bound pre-RCs isolated from interphase egg extract treated with A) roscovitine (Mcm6 and Mcm7) or B) Triton X-100 (Mcm4 and Mcm7). Loading was controlled by A) Coom., Coomassie straining or B) histone H3 immunoblotting.

In summary, two methods to isolate pre-RCs from *Xenopus laevis* egg extract were successfully established in the Boos laboratory.

### 3.4.2 Vertebrate pre-RCs form salt-sensitive double hexamers

A hallmark of yeast pre-RCs is salt-stability of the double hexamer. Tight interaction of the N-termini of two Mcm2-7 hexamers is crucial for bidirectional replisome formation in yeast.

To test if pre-RCs isolated from interphase egg extract form salt-stable double hexamers, pre-RCs were solubilized, separated by size-exclusion chromatography, and visualized by cryo-EM to investigate double hexamer formation. To solubilize pre-RCs from chromatin, isolated chromatin was treated with benzonase to digest the DNA. Subsequently, soluble fractions and pellets were separated by centrifugation and visualized by immunoblotted. Similar levels of Mcm2-7 signals were detected in a non-separated input sample and in the insoluble pellet fraction of an untreated sample. In contrast Mcm2-7 signals appeared in the soluble fraction upon treatment with benzonase (Figure 25 A lane 3, 5 and 6). Control samples using either the licensing inhibitor geminin or omitting sperm DNA did not show signals for Mcm2-7 subunits (Figure 25 A lane 1 and 2). Next, these solubilized pre-RCs were separated by size exclusion chromatography. Eluted fractions were collected and analyzed by SDS-PAGE followed by silver staining or immunoblotting. Silver staining as well as immunoblotting showed a peak fraction of eluted proteins of approximately 670 kilodaltons (Figure 25 B). In contrast to these results, a previous publication by the Blow laboratory showed that pre-RCs isolated from *Xenopus* egg extract have a molecular mass of approximately 945 kilodalton (Gambus et al., 2011). Since the calculated molecular mass of a single hexamer is approximately 530 kilodalton and the results from the Blow laboratory suggested a mass of 945 kilodalton for a double hexamer, it was not clear if single or double hexamers were isolated. In collaboration with the Costa Laboratory, solubilized pre-RCs were analyzed by cryo-EM. This analysis showed the characteristic double hexamer shape known from yeast (Figure 25 C). However, only 0.5 % of the particles observed were double hexamers whereas the rest appeared to be single hexamers. This was surprising because yeast studies demonstrated that the pre-RC double hexamer is a stable complex resistant to high-salt concentrations. Therefore, salt stability of pre-RCs isolated from egg extract was tested by supplementing the isolation buffer and sucrose cushion with increasing salt

concentrations. Immunoblotting of these chromatin fraction showed decreased Mcm2-7 signals at 350 mM KCl or more (Figure 25 D lane 7 and 9) whereas lower salt concentrations did not affect the Mcm2-7 signal intensities (Figure 25 D lane 1, 3 and 5). This salt sensitivity differs from yeast pre-RCs that are stable at least up to 500 mM salt concentration (Remus et al., 2009). Pre-RCs submitted to size-exclusion chromatography were isolated using 100 mM KCl and are therefore supposed to be stable. However, if vertebrate pre-RCs are in general less stable, the processing of the samples using size-exclusion chromatography and cryo-EM could lead to dissociation of the double hexamer.

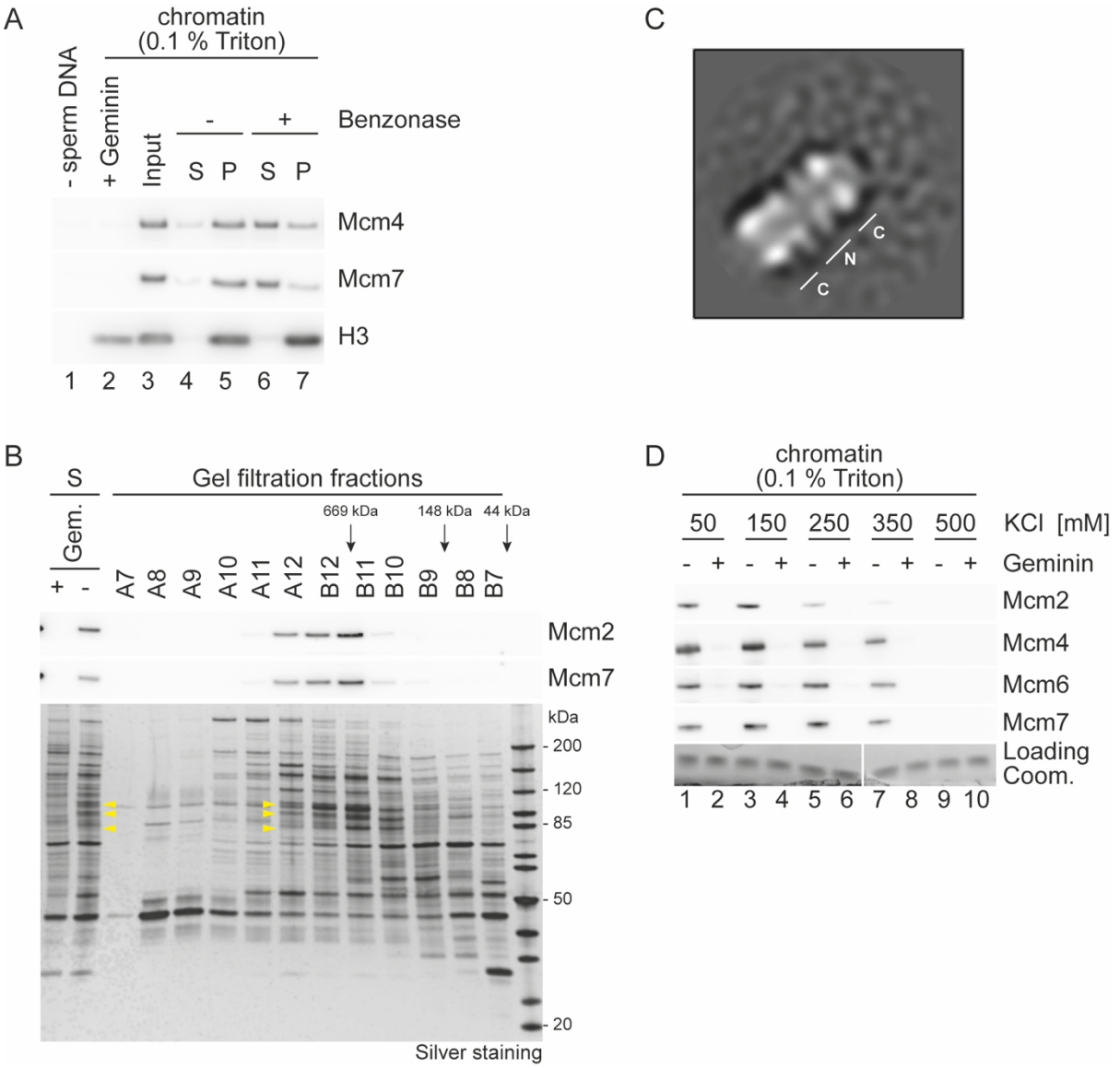



**Figure 25: Vertebrate pre-RCs are salt-sensitive.**
A) Immunoblot analysis of solubilized pre-RCs isolated from Triton X-100 treated interphase egg extract. Loading was controlled by histone H3 immunoblotting. B) SDS-PAGE analysis of elution fractions after size-exclusion chromatography of solubilized pre-RCs isolated from Triton X-100-treated interphase egg extract. Yellow arrow heads indicated Mcm2-7 subunits. C) Cryo-EM imaging of a single particle of a pre-RC double hexamer. D) Immunoblot analysis of pre-RCs isolated from Triton X-100-treated interphase egg extract using increasing salt concentrations. Loading was controlled by Coomassie staining.

Taken together these results suggested that pre-RCs isolated from *Xenopus* egg extract are less stable than yeast pre-RCs. These observations were surprising because bidirectional replication mediated by a stable double hexamer is fundamental for correct DNA replication. An explanation for less stable vertebrates could be that the formation of nucleosomes is keeping the two hexamers in close proximity, but this is speculation and would need further investigations.

### 3.4.3 Functionality of isolated pre-RCs from *Xenopus* egg extract

To obtain more detailed information about the functionality of isolated pre-RCs, their capability to support replication in egg extract was tested. For this, incorporation of radioactive labeled dCTP into pre-RC containing chromatin or control sperm DNA was measured in untreated or licensing-deficient interphase egg extract. Both DNA templates started replication within a delay of 30 minutes in the untreated extract. While pre-RC containing chromatin also supported replication in licensing-deficient extract, incorporation of radioactive dCTPs was inhibited upon sperm DNA addition to geminin-treated extract (Figure 26).

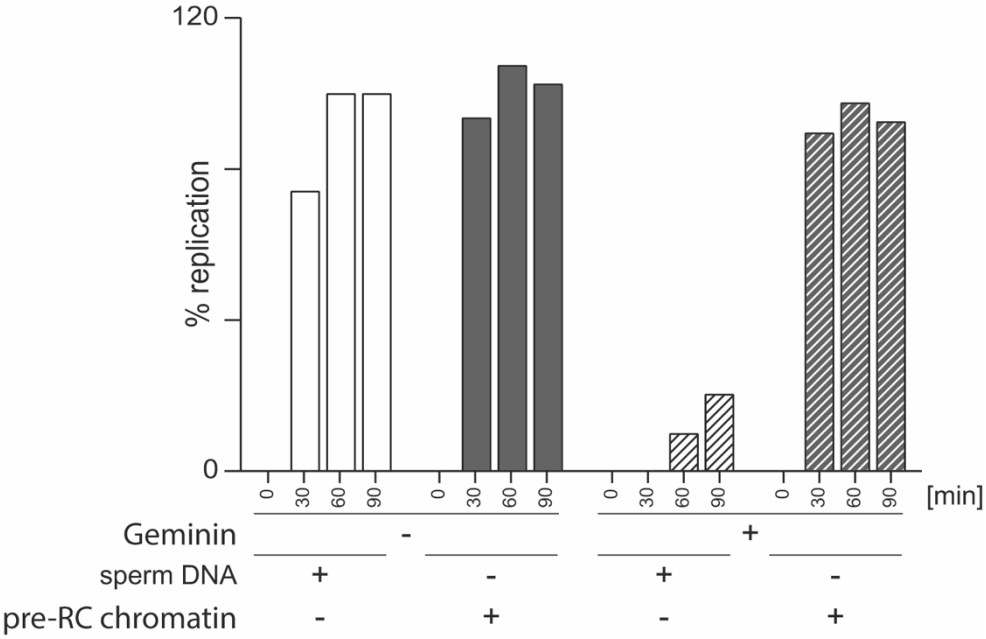



**Figure 26: Pre-RC-containing chromatin supports replication in licensing-deficient *Xenopus* egg extract.**
Replication analysis of interphase egg extract upon addition of pre-RC-containing chromatin isolated from roscovitine-treated egg extract or control sperm DNA.

These results demonstrated that functional pre-RCs could be isolated from *Xenopus* egg extract. Taken together this data provides a good basis for future pre-IC reconstitution experiments to further unravel the detailed molecular mechanisms of vertebrate origin firing.

# 4 Discussion

In this thesis, the interaction between TopBP1 and GINS was characterized at a molecular level, providing valuable insights into how their association supports vertebrate replication origin firing. Recent studies on yeast origin firing have advanced our understanding of how the Mcm2-7 helicase acts as a molecular machine that contributes to build replication forks (Costa and Diffley, 2022; De Jesus-Kim et al., 2021; Deegan et al., 2016; Douglas et al., 2018; Douglas and Diffley, 2016; Greiwe et al., 2022; Heller et al., 2011; Lewis et al., 2022; Looke et al., 2017; Saleh et al., 2022; Yeeles et al., 2015). However, the activation of the replicative helicase by loading GINS and Cdc45 to form the active CMG helicase has received limited attention in these studies. Addressing this gap, our investigation offers novel molecular insights into the loading of GINS onto pre-RCs mediated by the regulatory protein TopBP1.

Origin firing, the crucial regulatory step in replication initiation, plays a pivotal role in the formation of the CMG complex. Acting as a regulatory platform, TopBP1, along with MTBP and Treslin, orchestrates the coupling of replication origin firing to the S phase and is subject to modulation by diverse regulatory pathways, including the DNA damage checkpoint. As part of this research project, we have generated a dataset comprising potential new interactors of TopBP1, which can be used in future studies to deepen our understanding of the functions of TopBP1.

Although the origin firing step is fundamental in every cell cycle the molecular details of metazoan origin firing are poorly understood (Boos and Ferreira, 2019). In this project we isolated pre-RCs from *Xenopus* egg extract that can be used as substrate for *in vitro* pre-IC assembly together with other firing factors that were purified in the Boos laboratory. In particular, the Boos Laboratory is interested in the regulation platform consisting of TopBP1, Treslin, and MTBP that mediates the activation of the replicative helicase.

## 4.1 Molecular mechanisms of CMG formation

The formation of the CMG complex during origin firing is a crucial step in DNA replication initiation and plays a fundamental role in ensuring accurate and efficient genome duplication. Activation of the helicase begins with the stepwise assembly of

the pre-IC onto its precursor pre-RC. This critical step is dependent on high levels of CDK and DDK activity (Zegerman and Diffley, 2007). The recruitment of the yeast and metazoan firing factors Sld3-Sld7 and Treslin-MTBP to the inactive pre-RC is facilitated by the phosphorylation of the Mcm2-7 double hexamer by DDK (Francis et al., 2009; Greiwe et al., 2022; Saleh et al., 2022). The delivery of the helicase component Cdc45 to the Mcm2-7 double hexamer is thought to occur through a direct interaction with Sld3 and Treslin, respectively (Ferreira et al., 2022; Itou et al., 2014). Like in yeast, the coupling of origin firing to S phase is mediated by phospho-dependent binding of Sld3/Treslin to Dpb11/TopBP1. Phosphorylation at two conserved CDK phosphorylation sites in Sld3 and Treslin mediates the binding to the N-terminal tandem BRCT1/2 of Dbp11 and the homologue triple BRCT0/1/2 module of TopBP1 (Boos et al., 2011; Kumagai et al., 2011). In yeast this binding facilitates the recruitment of the pre-LC, including the helicase component GINS, to the pre-RC (Muramatsu et al., 2010a). The formation of such a pre-LC is not described in metazoan. After the helicase components Cdc45 and GINS are delivered to the pre-RC, the regulatory firing factors Treslin/Sld3, MTBP/Sld7 and TopBP1/Dpb11 dissociate, resulting in the formation of the active CMG helicase (Gambus et al., 2006; Kanemaki and Labib, 2006).

Electron microscopy studies of the CMG complex in *Drosophila* and yeast have provided insights into its structure. These studies have shown that Cdc45 and GINS, interact with the MCM2-7 ring. Cdc45 binds to the N-terminal tier of the MCM2-7 ring, contributing to the stabilization of the Mcm2:5 interface. On the other hand, GINS binds to the side of the MCM2-7 ring, specifically stabilizing the Mcm5:3 interface. These interactions between Cdc45, GINS, and the MCM2-7 ring play a crucial role in the assembly and activation of the CMG complex during DNA replication initiation (Abid Ali et al., 2016; Costa et al., 2011; Yuan et al., 2016). Although the archaea Mcm ring alone is sufficient to unwind DNA in biochemical assays, crystal structure analyses determined a complex similar to the yeast CMG in *Sulfolobus solfataricus* (Chong et al., 2000; Meagher et al., 2019; Xu et al., 2016). Recent cryo-EM structures of human CMGs purified from HEK293 cells and, from *in vitro* reconstituted replisomes have shown that the binding of Cdc45 to the Mcm2:5 interface, and the association of GINS with the Mcm3:5 is conserved (Jones et al., 2021; Rzechorzek et al., 2020). These studies reveal that the CMG helicase shows a striking resemblance across species, emphasizing its vital role in replication origin firing. However, the precise molecular

mechanisms by which the two helicase activators, GINS and Cdc45, are loaded onto pre-RCs are not yet fully elucidated.

### 4.1.1 The TopBP1 GINI core helix interacts with GINS Psf1 and Psf3

TopBP1 is a key regulator that couples activation of the replicative helicase to the S phase of the cell cycle. As described above this coupling is mediated by the binding of phosphorylated Treslin/Sld3 to the N-terminal BRCT mode in TopBP1/Dpb11 (Boos et al., 2011; Kumagai et al., 2011). Yeast Dpb11 BRCT3/4 and metazoan TopBP1 BRCT4/5 also share homology (Makiniemi et al., 2001). Dpb11 BRCT3/4 mediates the essential phospho-dependent binding of Sld2 whereas the equivalent TopBP1 BRCT4/5 domain was found to be dispensable for replication in *Xenopus* egg extract. However, the TopBP1 fragment that supports replication in egg extract contains the metazoan-specific BRCT3 domain and the inter-BRCT domain between BRCT3 and BRCT4 (Kumagai et al., 2010). This inter-BRCT region contains the GINS interaction (GINI) domain in TopBP1 (Tanaka et al., 2013). Interaction of this GINI domain with the subunit Psf1 of the helicase component GINS is required for the integration of GINS onto pre-RC, facilitating CMG formation and activation of the helicase (Takayama et al., 2003). Tanaka et al. have further characterized this interaction using yeast two-hybrid experiments and biochemical binding assays. Yeast two-hybrid experiments have revealed the presence of the GINI domain between Dpb11 BRCT2 and BRCT3, as well as in the homologous domain between BRCT3 and BRCT4 in TopBP1. Immunoprecipitation of a truncated Dpb11 fragment from yeast cell lysates demonstrated that the inter-BRCT domain of Dpb11 is sufficient to interact with Psf1. The detections of all four GINS subunits in these experiments indicated that the GINI domain interacts with the GINS complex *in vivo* (Tanaka et al., 2013). Biochemical binding assays were used to confirm this interaction *in vitro*. They could further demonstrate that the GINI domain is important for cell viability.

In this project we use sequence alignment and systematic mutations within the inter-BRCT domain to narrow down the GINI domain to amino acid 487-495 in TopBP1. Furthermore, AlphaFold2 predication and cryo-EM reveal a short α-helix motif within the GINI domain. Mutations in the GINI domain that break the helix lead to disruption of GINS binding *in vitro* but has only a moderate effect on replication in *Xenopus* egg extract. Here we identify two amino acids, asparagine and lysine, at position 43 and 51 in the GINS subunit Psf1 to interact with the GINI helix in TopBP1. These residues

are positioned in the distal end of the GINS Psf1 A domain close to the Psf3 subunit. Our data therefore describe the interaction between GINS Psf1 and Psf3 and the GINI helix in TopBP1 on a molecular level. This observation is consistent with the previous yeast two-hybrid experiments (Tanaka et al., 2013).

### 4.1.2 Cooperation between GINI and the novel interaction site BRCT4

Our study reveals a novel cooperative interaction between the GINI core helix and a previously unknown GINS binding surface in the BRCT4 domain of TopBP1. Cooperation of these two regions form a composite high-affinity interface that is required for stable interaction between TopBP1 and GINS *in vitro*. Our *in vitro* experiments demonstrate that each individual binding site alone is unable to form a stable complex with GINS, yet they can still support substantial levels of replication in *Xenopus* egg extracts. This suggests that even a relatively weak TopBP1-GINS interaction is sufficient for origin firing, potentially compensated by contacts between other pre-LC factors *in vivo*. While the GINI region and BRCT4 are not individually required for origin firing, their combined action is essential, indicating that their molecular roles are not distinct beyond GINS binding. It is plausible that their shared essential role is to facilitate the chaperoning of GINS onto pre-RCs, and the redundancy between the two binding sites provides an explanation for the previously overlooked significance of the TopBP1 BRCT4/5 domain in replication processes. Our model demonstrates that both GINS binding surfaces in TopBP1 are accessible to GINS when integrated into the CMG complex. This observation is consistent with the binding and delivery of GINS to the pre-RC using both binding surfaces in TopBP1. This model also allows simultaneous binding of the TopBP1 BRCT0/1/2 domain to phosphorylated Treslin, that in turn recruits Cdc45 to the pre-RC (Figure 27 A).

Budding yeast and *Xenopus* studies have shown that TopBP1, Treslin and MTBP are limiting replication initiation factors, indicating that these factors have to be recycled for late origin firing (Collart et al., 2013; Mantiero et al., 2011). In order to recycle these firing factors, they have to be released from the pre-RC after delivering the helicase components Cdc45 and GINS (Jenkinson and Zegerman, 2022). However, such a release mechanism is not yet known. Our data shows that the two individual GINS interaction sites in TopBP1 have only low binding affinities. This observation allows a model where the release of one site could lead to destabilization of the complex and thus to the dissociation of TopBP1. Since the two binding sites in TopBP1 are

accessible for GINS when integrate into the CMG complex, no steric clash would induce TopBP1 ejection. However, further structural investigations in this research project suggests that integration of Polε into the CMG (CMGE) is not compatible with TopBP1 BRCT4 GINS interaction and could force TopBP1 to leave the complex (in collaboration with Matthew Day; Appendix Figure 6). How the replacement of TopBP1 by Polε occurs mechanistically, and whether Polε displaces TopBP1 or TopBP1 has to leave the complex to allow Polε binding, remains unclear. Our cryo-EM data and previous crystallization studies of GINS show that the B-domain of Psf1, that is needed for CMG formation but not for the formation of the GINS tetramer, is flexible (Chang et al., 2007; Kamada et al., 2007). We hypothesize here a model where the reconfiguration of this B-domain mediates TopBP1 dissociation from the CMG. Once GINS is integrated into the CMG the former flexible Psf1 B domain adopts a discrete conformation packing around Cdc45 (Figure 27 B). This reconfiguration excludes binding of the TopBP1 BRCT4 domain to Psf1 and may favor the binding to Polε. As a result, the low affinity binding of GINS to only one interaction site in TopBP1 would lead to TopBP1 dissociation (Figure 27 C). This model suggests a three-step process: first, delivery of GINS by TopBP1 to the helicase, second, association of Polε and, third, ejection of TopBP1.

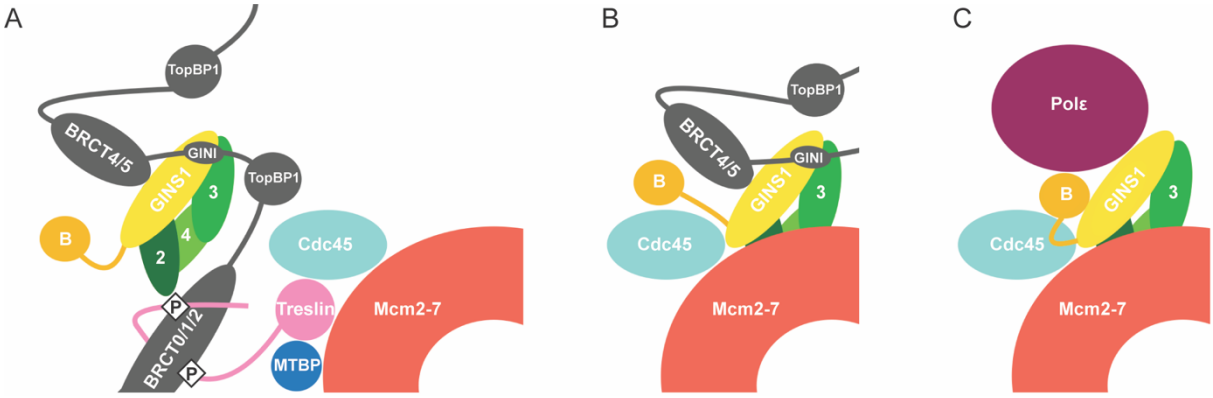



**Figure 27: Hypothetical model of the TopBP1-dependent loading of GINS onto pre-RCs.**
A) TopBP1 is recruited through a phosphorylation dependent interaction between the N-terminal BRCT0/1/2 module and Treslin. This allows delivery of the GINS tetramer to the Mcm2-7-Cdc45 complex. B) Psf1 B-domain orders, packing against Ccd45, simultaneously locking the Mcm2-7 ring and weakening the interface between TopBP1 BRCT4 and the GINS complex. C) TopBP1 dissociates, allowing Polε to bind to form the CMGE complex. Adapted from (Day et al., 2023)

This mechanistical change illustrates that the coupling of origin firing to S phase remains fundamental, yet flexible. Similar changes were observed in the model organisms *C. elegans* and fission yeast for the replication factors Treslin and RecQL4,

indicating that partial redundancy of S phase kinases and their substrates support these mechanistical changes (Fukuura et al., 2011; Gaggioli et al., 2014).

## 4.2 Control of genome duplication by replication origin firing

In eukaryotes, genome duplication is a particularly complex challenge, as multiple origins of replication are required to duplicate their large genomes. The presence of multiple origins of replication introduces the possibility of encountering significant inter-origin distances that exceed the replication capacity of two replication forks during one S phase. In most eukaryotes including metazoan the positioning of origins is not strictly defined by specific DNA sequences, resulting in smaller and larger inter-origin distances. To prevent this random gap problem and ensure accurate duplication of the genome within the duration of one S phase, origins must fire reliable, and they must be placed close enough to each other. (Figure 28 A 1.). Origins that do not fire or are misplaced increase the gap size between two replication forks and result in unreplicated DNA sections (Figure 28 A 2. and 3.) (Boos and Ferreira, 2019).

Due to limiting amounts of some firing factors including the regulation platform composed of Treslin, MTBP and TopBP1 not all origins can fire simultaneously (Collart et al., 2013; Mantiero et al., 2011; Tanaka et al., 2011a; Wong et al., 2011). To avoid scattered replication events that would cause large inter-origin distances, replication is organized in replication domains (Berezney et al., 2000; Nakamura et al., 1986). It become apparent that neighboring domains fire sequentially, suggesting a domino-like model where replication progresses from the initial domain to the adjacent domains (Figure 28 B )(Maya-Mendoza et al., 2010; Sadoni et al., 2004; Sporbert et al., 2002). This could be a mechanism that prevents isolated replication events and therefore also the generation of large inter-origin distances. However, timing transition regions, the region between two adjacent domains that start replication with a significant time delay are an exception. In this regions replication must be blocked and firing of late replicating domains has to be initiated to a later time point (Figure 28 C) (Boos and Ferreira, 2019).

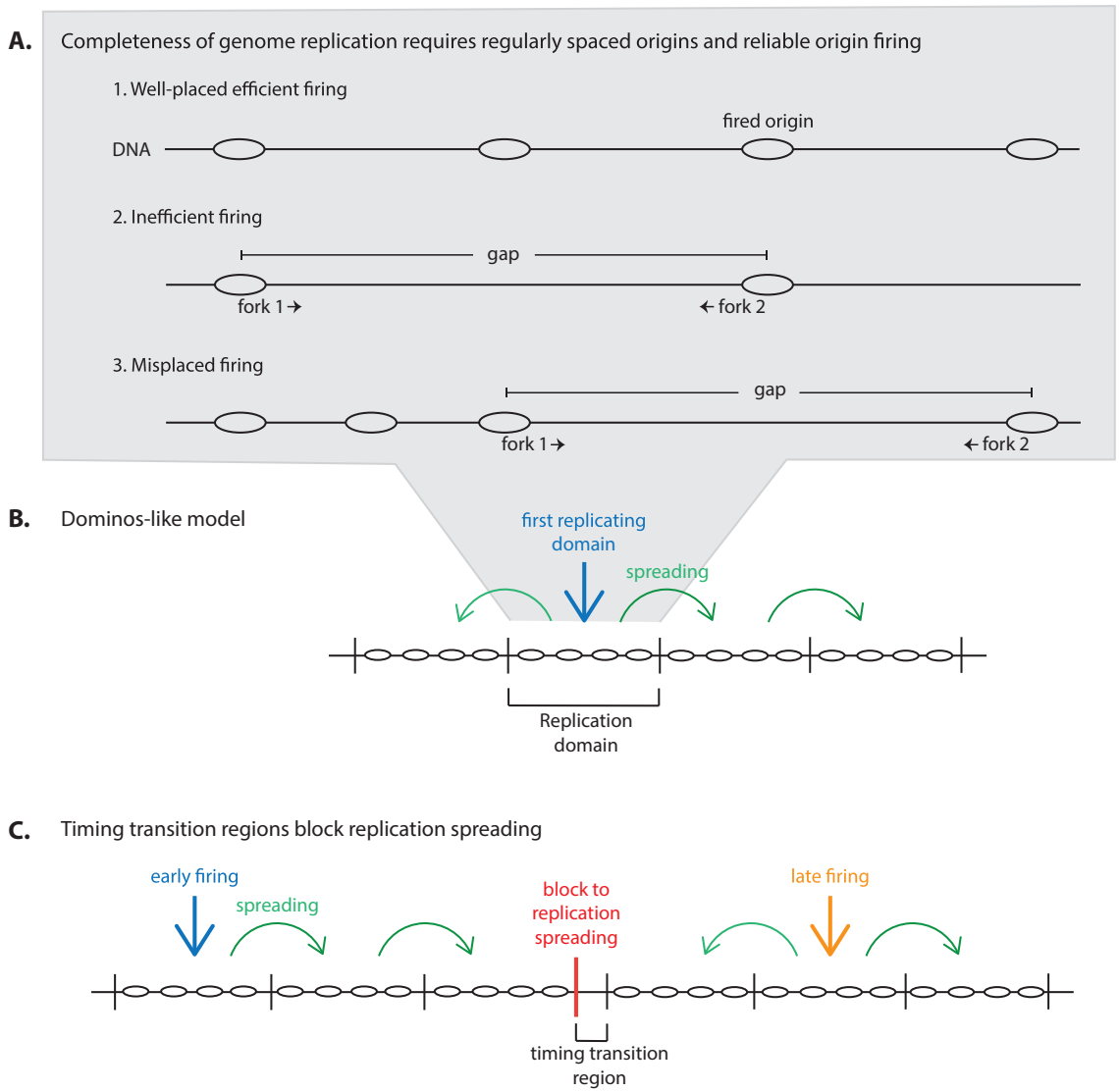



**Figure 28: Proper control of origin firing is required for complete genome duplication.**
A) Well-placed origins and efficient firing are required for complete DNA replication (1.) Replication problems can arise if not enough origins fire (2.) or if origins are misplaced (3.) Such misregulations can result in gaps in the replicating genome that are too big to be replicated by two forks during an S phase. B) and C) Schematic representation of the domino-like model in which replication spreads from the first replicating domain to its neighbors. C) This replication spreading is blocked at timing transition regions. Timing of the later domain at timing transition regions requires a new timing signal (Boos and Ferreira, 2019).

The domino-like model of genome replication is highly simplified. It remains unclear what determines the start of the first replication event and how spreading and blockage is achieved. However, limiting amounts of firing factors, including Treslin, MTBP and TopBP1 seem to be involved in replication timing. In addition, Treslin and MTBP interact with transcriptionally active chromatin known to be replicated early in S phase, suggesting that they contribute to the timing of replication (details below).

### 4.2.1 The TMT complex as a regulatory platform for replication origin firing

Like the yeast SDS complex the metazoan Treslin-MTBP-TopBP1 (TMT) complex mediates the activation of the CMG helicase specifically in S phase. For this the Treslin-MTBP complex promotes the recruitment of the helicase component Cdc45 to DDK phosphorylated Mcm2-7 (Francis et al., 2009; Greiwe et al., 2022; Saleh et al., 2022; Volpi et al., 2021). Rise of CDK levels in S phase mediates the loading of the helicase component GINS by phospho-dependent binding of Treslin to TopBP1 (Boos et al., 2011; Kumagai et al., 2010; Tanaka and Araki, 2013; Zegerman and Diffley, 2007). Under DNA damage conditions, which can result in genome alterations and subsequently genome instability, the interaction between TopBP1 and Treslin is inhibited. Additional phosphorylation of Treslin by the checkpoint kinase Chk1 prevents binding to TopBP1 and thereby also the activation of the replicative helicase (Boos et al., 2011; Guo et al., 2015; Kelly et al., 2022).

Although, these regulatory firing factors are essential for replication origin firing, none of these proteins are part of the active replisome. In fact, reconstitution of the human replisome using proteins purified from insect cells has demonstrated that these regulators are not required for *in vitro* assembly of the active helicase, although they are essential *in vivo* (Boos et al., 2011; Ferreira et al., 2021; Jones et al., 2021; Kumagai et al., 2010).

In contrast to genomic DNA replication in archaea, eukaryotic cells depend on the strict coupling of replication to the S phase of the cell cycle, that ensures accurate and complete genome duplication. As described above this coupling and the formation of the active CMG complex depends on the regulation of the TMT (metazoan) or SDS (yeast) complexes by DDK and CDK. Whilst a homologue of the CMG helicase also exists in archaea, they do not possess a regulatory TMT/SDS complex or their regulatory cell cycle kinases DDK and CDK (Barry and Bell, 2006). A simple but hypothetical explanation for this is, that these regulatory replication factors evolved with the eukaryotic cell cycle. In addition to this strict coupling of replication to S phase the eukaryote SDS/TMT complex is targeted by the effector kinases Rad53/Chk1 in order to inhibit replication upon DNA damage and replication stress. Such checkpoints that are important for genome stability in eukaryotes are not known in archaea (Marshall and Santangelo, 2020).

Apart from the conserved function of TopBP1, Treslin and MTBP, they evolved addition metazoan specific regulations in DNA replication and damage signaling pathways. Besides the phospho-dependent binding of Treslin to TopBP1, CDK phosphorylation of Treslin controls S phase length by altering fork number and the rate of progression through the replication timing program (Sansam et al., 2015). In addition to Chk1 phosphorylation of Treslin which inhibits replication, Treslin is known to be directly involved in the ATR kinase-mediated checkpoint response (Hassan et al., 2013). Furthermore, the vertebrate-specific N- and C-terminal domains of Treslin appear to be required for normal replication in human cells. Although deletion of either of these vertebrate-specific domains has only mild effects on replication, a core Treslin protein lacking both termini is insufficient to support replication and origin firing in U2OS cells. This suggests that the two vertebrate-specific domains cooperate to facilitate replication origin firing in parallel pathways (Ferreira et al., 2022). Another vertebrate specific function of Treslin is its contribution to a mechanism for epigenetic control of DNA replication (Sansam et al., 2018). Acetylated histones mediate binding of transcriptional regulatory proteins such as BRD2 and BRD4. The recruitment of multiprotein complexes to acetylated chromatin by these BRD proteins enhances transcriptional activation (Fujisawa and Filippakopoulos, 2017; Rahman et al., 2011). As described above transcriptionally active chromatin is replicated early in S phase (Boos and Ferreira, 2019; Marchal et al., 2019; Rhind, 2006; Yekezare et al., 2013). Treslin binding to BRD2 and BRD4 results in the recruitment of Treslin to acetylated chromatin and is required for normal spatiotemporal replication in metazoan (Sansam et al., 2018).

In contrast to yeast Sld7, the consecutive binding partner of Treslin, MTBP, is essential for replication. Mutants of MTBP that are unable to bind to Treslin result in a decrease in the overall levels of Treslin-MTBP complex in human cells. This suggests that the binding between MTBP and Treslin is required for the stability or proper assembly of the complex (Boos et al., 2013; Kumagai and Dunphy, 2017). Furthermore, dimerization of MTBP by its C-terminal domain facilitates normal replication in human cells (Kohler et al., 2019). This suggests that dimerization of MTBP serves as a control to ensure simultaneous activation of both Mcm2-7 helicases and loading of Cdc45 and GINS. This mechanism would prevent the activation of single replisomes and would ensure bidirectional replication. In contrast to yeast Sld7, MTBP is phosphorylated by at least three kinases, which modulate origin firing and therefore may balance

replication activity, indicating that MTBP evolved to an origin firing regulation platform (Ferreira et al., 2021). Like Treslin, MTBP forms spatially regulated associations with multiple genomic signals to facilitate the initiation of DNA replication. In particular, the C-terminus of MTBP is known to bind to various transcriptionally active DNA structures, including G-quadruplex DNA, AP-1 motifs, and histone H3K4me2 chromatin (Kumagai and Dunphy, 2020). These observations are consistent with a coordinated association of MTBP with multiple genomic signals to promote initiation of replication.

In addition to TopBP1's essential role in origin firing, TopBP1 is known to play a critical role in regulating the response to DNA damage. TopBP1 binding to the checkpoint kinase ATR, further stimulates kinase activity required for regulation of the checkpoint response (Kumagai et al., 2006). As a multi-BRCT scaffold protein, TopBP1 contains numerous binding domains for additional interactors to fulfill its dual role in replication initiation and DNA damage signaling. Since the vertebrate-specific functions of TopBP1 are the subject of this thesis, its regulation is discussed in more detail in the following section.

Taken together these mechanistical changes in the metazoan TMT complex enable more complex regulations of replication initiation and DNA damage signaling. Given the increased complexity of metazoan genomes, these superior regulatory mechanisms seem essential for maintaining genetic homeostasis.

### 4.2.2 Molecular functions of TopBP1

Our research uncovers a novel molecular aspect of TopBP1, which is particularly intriguing given its multifunctional nature and the requirement for precise regulation. TopBP1 as many metazoan orthologs is much larger than its yeast counterpart Dpb11. In addition to the conserved N-terminal BRCT0/1/2 domain and the BRCT4/5 pair, TopBP1 contains five metazoan-specific BRCT domains and additional inter-BRCT domains (Garcia et al., 2005; Yamane et al., 1997). TopBP1 plays dual roles in replication initiation and in the regulation of DNA damage signaling. In metazoans the two checkpoint kinases ATM and ATR are essential upstream regulators in the DNA damage signaling pathways. While ATM is primarily recruited to DSB, ATR activation is mainly induced by replication stress and single strand breaks.

The above-described activation of ATR by TopBP1 is mediated by the interaction of the ATR-activation domain (AAD) located between BRCT6 and BRCT7 in TopBP1

(Kumagai et al., 2006). Interaction of the AAD either directly with ATR or the ATR-interacting protein (ATRIP) activates the checkpoint kinase (Mordes et al., 2008a). How exactly TopBP1 is recruited to ATR at sites of damaged DNA is not completely understood but it become apparent that ATR activation by TopBP1 requires interaction with the RAD9-HUS1-RAD (9-1-1) checkpoint complex. Therefore, RAD9 binds phospho-dependently to the N-terminal BRCT0/1/2 motif in TopBP1 (Delacroix et al., 2007; Lee et al., 2007). This RAD9-TopBP1 interaction is required for the activation of ATR but not for the recruitment of TopBP1 to stalled replication forks (Lee and Dunphy, 2010; Yan and Michael, 2009). The Bloom syndrome helicase BLM that binds phospho-dependently to TopBP1 BRCT5 is considered to play a role in TopBP1 recruitment to ATR (Blackford et al., 2015).

Our mass spectrometry data confirm several known TopBP1 interacting proteins, including ATR, RAD9 and BLM. This data indicates identification of specific TopBP1 interacting proteins and can therefore serve as a resource to investigate these interactions in detail or even identify new TopBP1 interacting proteins.

Apart from ATR activation, TopBP1 is also known to interact with ATM at DSB. Phosphorylation of TopBP1 serine-1131 by ATM strongly enhances ATR activity at DSB (Yoo et al., 2007). In addition to this ATM phosphorylation site in TopBP1, TopBP1 contains numerous motifs that are also known to be phosphorylated by ATM. ATM preferentially phosphorylates substrates in response to DNA damage on SQ or TQ motifs (Traven and Heierhorst, 2005).

Interestingly, one of these SQ motifs is located adjacent to the GINI core helix (Sokka et al., 2010). This SQ side is therefore a potential candidate to regulate GINS-TopBP1 interaction upon activation of ATM. In collaboration with Matthew Day, we tested if a fluorescently labeled TopBP1 peptide containing phosphorylated serine at the SQ site impacts the binding of the GINS tetramer (data not shown). Indeed, phosphorylation of TopBP1 Ser492 reduces interaction with the GINS tetramer significantly in FP assays. This may indicate a mechanism that regulates GINS binding to TopBP1 upon DNA damage. To investigate this hypothesis further the here established proximity biotinylation method was used for initial experiments (data not shown). Preliminary data indicates that transient expression of APEX2-TopBP1 wild type in irradiated cells leads to the enrichment of proteins involved in the DNA damage signaling pathway. In particular, MDC1, ATR and CIP2A were increased in comparison to untreated cells. ATR, MDC1 and CIP2A are known TopBP1 interactors involved in DNA damage

signaling. MDC1 association with TopBP1 facilitates TopBP1 recruitment to phosphorylated histone H2AX (γH2AX). γH2AX is known as a biomarker for the presence of DSB in mammalian cells (Georgoulis et al., 2017). Association of TopBP1 with the phosphatase CIP2A also appears to be essential to maintain genome stability, although the exact mechanisms are not yet elucidated (Adam et al., 2021; Laine et al., 2021). This data provides a first hint to verify the damage/repair proximitome of TopBP1. To test this hypothesis further, we generated a TopBP1 mutant that cannot be phosphorylated at this SQ site (AQ). In theory this AQ mutant cannot be targeted by ATM at this particular site and therefore should not show enrichment of DNA damage proteins after irradiation. Initial irradiation experiments using cells transiently expressing the TopBP1-AQ mutant cannot confirm this theory. However, this experiment is only performed once and has to be replicated. Besides, the DNA damage response does not relay on only one ATM site in TopBP1. Therefore, it is most likely that the effect is simply not strong enough to be detected by this method.

Despite these observations we verify the replication proximitome of TopBP1 that confirms TopBP1-GINS interaction in human cells. The establishment of the proximity biotinylation method was successful and will therefore be used in other projects in the Boos laboratory.

## 4.3 Isolated *Xenopus* pre-RCs as a tool to investigate replication origin firing

Replication origin firing is a highly regulated process that leads to the stepwise conversion of the inactive pre-RC to the active CMG. Although origin firing is fundamental in each cell cycle, the molecular mechanisms of this step are poorly understood in metazoans (Boos and Ferreira, 2019). Yeast *in vitro* reconstitution of origin firing has revealed that high levels of DDK and CDK are required for the formation of the pre-IC intermediate (Zegerman and Diffley, 2007). It has been shown that the core firing factors Sld3, Sld7, Sld2, Dpb11, Polε together with the helicase Mcm2-7, Cdc45 and GINS are sufficient for origin firing (Yeeles et al., 2015; Yeeles et al., 2017).

Recently, the structure of the human pre-RC isolated from G1 phase cells was resolved (Li et al., 2023). In comparison to the yeast pre-RC, the human pre-RC appears to play a crucial role in initial melting of dsDNA. (Abid Ali et al., 2017; Greiwe et al., 2022; Li et al., 2023). This initial melting is required to maintain stable coupling

between the two opposing human Mcm2-7 hexamers on DNA (Li et al., 2023). In contrast to yeast pre-RCs, which are capable to slide on DNA, the tight binding of the human pre-RC to DNA suggests that the assembly of the double hexamer is immobile (Gros et al., 2015; Remus et al., 2009). Interestingly, melting of the dsDNA in yeast is induced later in replication when the CMGE double hexamer is formed (Lewis et al., 2022). In addition, Li et al. showed that the ATP/ADP occupancy between the adjacent Mcm2-7 subunits differ from the arrangement observed in yeast pre-RCs (Greiwe et al., 2022; Noguchi et al., 2017). These mechanistical differences indicate regulatory changes in double hexamer formation between yeast and humans.

In this scientific project we isolate pre-RCs from *Xenopus* egg extract. In contrast to yeast and human pre-RC the double hexamers seem to be salt-sensitive (Li et al., 2023; Remus et al., 2009). However, our data show that pre-RCs isolated from egg extract are capable to support replication in licensing-deficient egg extract. These findings provide a good basis to study origin firing in vertebrates.

The long-term aim is to use these isolated pre-RCs to assembly pre-ICs and probably CMGs *in vitro*. In parallel projects in the Boos Laboratory Treslin-MTBP and the helicase component Cdc45 are purified from insect cells. Together with recombinant GINS and TopBP1, and the kinases CDK and DDK, biochemical binding assays could be used to stepwise assemble pre-ICs on the isolated pre-RCs. Further investigation of the TMT regulator complex could provide valuable insights into its role in CMG formation and the mechanisms of recycling these limiting factors after Cdc45 and GINS recruitment to the helicase. In particular, this approach can be used to investigate the here proposed TopBP1 ejection model in more detail by competition experiments between TopBP1 and Polε binding to GINS.

# 5 Material & Methods

## 5.1 *Xenopus laevis* egg extract methods

*Xenopus laevis* egg extract methods were initially adapted from the Blow lab (Gillespie et al., 2012) and established in the Boos lab during my master thesis. For isolation of chromatin from interphase egg extract and subsequent mass spectrometry, methods were further adapted from the laboratory of Dr. Markus Räschle (Raschle et al., 2015).

### 5.1.1 Generation of replicating *Xenopus laevis* egg extract

*Xenopus laevis* egg extracts were prepared from metaphase II arrested eggs. Therefore, frogs were primed with 50 units of pregnant mare serum gonadotrophin (Sigma-Aldrich) 3 days before and with 500 units of human chorionic gonadotrophin (Sigma-Aldrich) ~16 hours before the eggs were required. After the second injection frogs were placed in individual laying tanks in 2 l of 1x MMR buffer (100 mM NaCl, 2 mM KCl, 1 mM MgCl2, 2 mM $CaCl_2$, 0.1 mM EDTA, 5 mM HEPES-NaOH, pH 7.8). Eggs from individual frogs were collected in separate glass beakers. Subsequently, apoptotic and activated eggs were removed to keep only metaphase-arrested, high-quality eggs. To remove non-egg debris eggs were rinsed several times in 1x MMR. After rinsing, MMR was replaced by dejellying solution (2 % cysteine (w/v) in H2O, pH 7.8 with NaOH). The solution was swirled gently to remove the jelly coat surrounding the eggs. Dejellyed eggs were rinsed twice with XBE2 (100 mM KCl, 2 mM $MgCl_2$, 0.1 mM $CaCl_2$, 1.71 % (w/v) sucrose, 5 mM K-EGTA, 10 mM HEPES-KOH, pH 7.7) and subsequently transferred to 4 ml centrifuge tubes containing 1 ml XBE2, 10 µg/ml protease inhibitors (aprotinin, leupeptin and pepstatin) and 100 µg/ml cytochalasin D. Excess buffer was removed from the settling eggs before packing the eggs by centrifugation at 1,400 g in a swinging bucket rotor for 1 minute at 16 °C. Excess buffer was removed again before spin-crushing the eggs by centrifugation at 16,000 g in a swinging bucket rotor for 10 minutes at 16 °C. The dirty brown cytoplasmic layer was then collected by using a 20G needle and transferred to a 2.2 ml centrifuge tube. From this point, the extract was kept on ice. Next, the extract was supplemented with 10 µg/ml cytochalasin D, 10 µg/ml protease inhibitors, 1:80 dilution of energy regenerator (1 M 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 $MgCl_2$, 1 mM EGTA, 2 mM DTT, 20 mM $K_2HPO_4$/ $KH_2PO_4$ pH 8.0, 40 mM HEPES-KOH, pH 8.0) to 15 % (v/v). To clarify the extract, it was centrifuged at 84,000 g in a pre-cooled swinging bucket rotor for 20 minutes at 4 °C. After centrifugation the lipid plug from the top of the tube was removed with an ethanol-cleaned and dried small-size spatula. Then the golden cytoplasmic layer was collected, including the wispy membranes floating immediately below the lipid plug. Subsequently, the extract was supplemented with 2 % glycerol (v/v) before freezing. The extract was dropped with a small pipette tip into a plastic dish containing liquid nitrogen for snap-freezing. The extract beats were stored at -80 °C. Adapted from (Parlak, 2017).

### 5.1.2 Chromatin isolation from *Xenopus laevis* egg extract

To isolate chromatin from *Xenopus laevis* egg extract, extracts were supplemented with 1/40 energy generator (1 M phosphocreatine K salt, 600 µg/ml creatine phosphokinase in 10 mM HEPES-KOH pH 7.6), 250 µg/ml cycloheximide and released into interphase by adding 0.3 mM $CaCl_2$ for 15 minutes at 23 °C. To start the replication reaction, a maximum of 20 ng sperm DNA/µl extract was added. For each chromatin sample 10-20 µl extract was used. All time points were aliquoted into different tubes and incubated at 23°C for the required length of time. Reactions were stopped by adding 500 µl of ice-cold NIB (50 mM HEPES-KOH pH 7.6, 50 mM KCl, 5 mM $MgCl_2$, 0.5 mM spermidine, 0.15 mM spermine, 0.1 % Triton X-100, 2 mM DTT, 1 µg/ml leupeptin, 1 µg/ml pepstatin, 1 µg/ml chymostatin, 0.1 mM PMSF) to the extract and mixed thoroughly but gently. After the reactions were stopped, diluted extracts were underlaid slowly with 100 µl of NIB + 30 % sucrose to create a cushion. The reaction was centrifuged for 10 minutes at 2,500 g at 4 °C in a swinging-bucket rotor. After centrifugation, the buffer was removed from above the cushion and rinsed three times with 200 µl NIB by adding and removing the buffer on top of the cushion before removing most of the cushion. About 15 µl were left in the tubes and centrifuged at 10,000 g for 2 minutes to focus the pellet. The remaining liquid was removed from the tube by using a fine pipette tip. For subsequent separation by SDS PAGE, 1 x SDS sample buffer (6.5 % glycerol, 715 mM β-mercaptoethanol, 3 % SDS, 62.5 mM Tris-HCl pH 7.9, 0.005 % Bromphenol blue) was added to each tube and boiled 10 minutes at 95°C. For solubilization of pre-RCs from isolated chromatin, chromatin

fractions were resuspended in one-fourth original extract volume NIB supplemented with 30 % sucrose and 2 U Benzonase per µl buffer. Adapted from (Parlak, 2017).

### 5.1.3 Immunodepletion of *Xenopus laevis* egg extract

For XTopBP1 depletion from *Xenopus laevis* egg extract antibodies, were coupled for 1 hour at RT to G dynabeads. Detailed information about antibody usage is listed in Table 4. Before depletion, 1/40 energy generator (1 M phosphocreatine K salt, 600 µg/ml creatine phosphokinase in 10 mM HEPES-KOH pH 7.6), 250 µg/ml cycloheximide were added, and egg extracts were released into interphase by adding 0.3 mM $CaCl_2$ for 15 minutes at 23 °C. Depletion reactions were incubated on ice for 45 min. After removing the beads, depleted egg extracts were either used fresh or aliquoted, snap-frozen and stored at -80 °C for replication assays and chromatin isolation.

**Table 4: Antibodies used for *Xenopus laevis* egg extract depletion**

| Antibody | Antibody concentration/µl beads | Antibody concentration/µl extract |
|---|---|---|
| XTopBP1 #1 | 0.5 µg/µl | 0.675 µg/µl |
| XTopBP1 #2 | 0.5 µg/µl | 0.675 µg/µl |

### 5.1.4 Chromatin isolation from XTopBP1 depleted *Xenopus* egg extract

XTopBP1 depleted interphase egg extract was supplemented with 50 µg/ml aphidicolin to avoid termination and therefore accumulate of replisomes onto chromatin. The reaction was aliquoted into 15 µl samples before adding buffer or 6 ng/µl recombinant TopBP1. Prior to sperm DNA addition, reactions were pre-incubated with XBE2 buffer (100 mM KCl, 2 mM $MgCl_2$, 0.1 mM $CaCl_2$, 1.71 % (w/v) sucrose, 5 mM K-EGTA, 10 mM HEPES-KOH, pH 7.7) or 2.25 µM geminin 10 minutes on ice. The replication reaction was started by adding 9 ng sperm DNA/µl. Reactions were incubated for 75 minutes at 23 °C after sperm addition. By diluting the extracts with 300 µl ice cold EBL salt (10 mM HEPES-KOH pH 7.7, 50 mM KCl, 2.5 mM $MgCl_2$) + 250 mM sucrose + 0.6 % Triton X-100 the reaction was stopped. Subsequently, the reaction was loaded onto 150 µl ELB salt + 25 % sucrose and centrifuged for 10 minutes at 2,500 g at 4 °C in a swingout rotor. After removing the liquid above the sucrose cushion, the surface of the cushion was rinsed twice with 200 µl EBL salt containing 250 mM sucrose before discarding most of the cushion. About 15 µl were left in the tubes and centrifuged for

2 minutes at 10,000 g at 4 °C in a fixed angle. The remaining liquid was removed from the tube using a gel-loading pipette tip. For subsequent separation by SDS-PAGE, chromatin pellets were resuspended in 24 µl 1x SDS sample buffer (6.5 % glycerol, 715 mM β-mercaptoethanol, 3 % SDS, 62.5 mM Tris-HCl pH 7.9, 0.005 % Bromophenol blue) and boiled for 10 minutes at 95 °C. For mass spectrometry analysis of isolated chromatin pellets, samples were further processed as described in section 5.1.5 CHROMASS.

### 5.1.5 CHROMASS

For mass spectrometry analysis chromatin was isolated as described in section 5.1.4. Processing of the chromatin pellets and analyzation by CHROMASS (Raschle et al., 2015) was performed by the lab of Markus Räschle. In brief, chromatin pellets were resuspended in 50 µl denaturing buffer (8 M Urea; 100 mM Tris pH 8). Proteins were reduced by incubating the samples with 5 mM dithiothreitol (final concentration) at 22 °C for 30 min. Next, peptides were alkylated by adding 20 mM iodoacetamide (final concentration) for 30 minutes at 37 °C. Subsequently samples were supplemented with 25 mM DTT (final concentration) and incubated for 5 minutes at 22 °C. 500 ng Lys-C was added for 3 hours at 37 °C to digest the proteins. Urea concentration was diluted to 1 M by the addition of dilution buffer (100 mM ammonium bicarbonate). 500 ng Trypsin per sample was added and samples were incubated at 37 °C overnight. Digested peptides were acidified with 1 % trifluoroacetic acid (final concentration) and desalted on Empore C18 material, according to Rappsilber et al. (Rappsilber et al., 2007). Eluted peptides were dried in a vacuum concentrator and reconstituted with 9 µl of buffer A (0.1 % formic acid) containing 0.2 % acetonitrile and 0.01 % trifluoro acetic acid. For MS analysis, 4 µl 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 µm inner diameter packed in-house with C18 resin ReproSilPur 120, 1.9 µm diameter Dr. Maisch) using a 3-hour non-linear gradient at a flow rate of 250 nl/minute 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.

MS data files from single-shot experiments were processed with MaxQuant (version 2.0.1.0) using a non-redundant *Xenopus laevis* data base (Raschle et al., 2015). Raw data were normalized using the label-free quantitation (LFQ) algorithm implemented in MaxQuant. MS Data with Perseus (version 1.6.15.0) (Tyanova et al., 2016). Protein groups were filtered to eliminate contaminants, reverse hits, and proteins identified by site only. For the heat map, LFQ intensities were log2 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 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 log2 transformed, and missing values were imputed with random values drawn from a normal distribution centered around the detection limit of the MS instrument (Perseus imputation; width=0.3, down shift = 1.8). Two sample Student's t-tests with permutation-based FDR control were carried out in Perseus. For these tests, three valid values in at least one quadruplicate of either of the tested conditions were required. FDR was adjusted for multiple testing by the Benjamini-Hochberg procedure using a significance threshold of FDR<0.05. Data visualization was carried out in R (Day et al., 2023).

### 5.1.6 Replication assay

To assess the efficiency of recombinant TopBP1, and determine the concentrations of inhibitors and detergents, replication assays were conducted, and replication signals were quantified. For TopBP1 rescue experiments, TopBP1 depleted extract was supplemented with 50 nCi/µl $\alpha^{32}$P-dCTP and aliquoted into different tubes. Subsequently, 6 ng recombinant TopBP1 protein per µl extract was added to TopBP1 depleted extract (section 5.1.3). To test inhibitor and detergent concentrations untreated extract was first supplemented with a 1/40 energy generator (1 M phosphocreatine K salt, 600 µg/ml creatine phosphokinase in 10 mM HEPES-KOH pH 7.6), 250 µg/ml cycloheximide and released into interphase by adding 0.3 mM $CaCl_2$ for 15 minutes at 23 °C. Extract was supplemented with 50 nCi/µl $\alpha^{32}$P-dCTP and aliquoted into different tubes. Subsequently, buffer, inhibitor or detergent was added to extract samples. To start replication reactions, 5 ng sperm DNA (kindly provided by the lab of O Stemmann) per µl extract was added. 1-3 µl extract was

spotted for each time point onto a glass fiber membrane. Unbound $\alpha^{32}$P-dCTP was removed by rinsing the membrane three times 15 minutes with ice-cold 5 % TCA and once with ice-cold ethanol. Newly synthesized DNA was detected by phospho-imaging.

## 5.2 Biochemical methods

### 5.2.1 SDS-PAGE

To separate proteins by size, discontinuous SDS-PAGE was used to analyze protein samples. 5-14 % and 10 % self-made polyacrylamide or 4-12 % Criterion™ XT Bis-Tris Precast and 3-8 % Criterion™ XT Tris-Acetate Precast gels were used to separate protein samples. Stacking gels for self-made gels were composed of 125 mM Tris (pH 6.9), 0.1 % SDS and 5.12 % acrylamide. Protein samples were denatured by 10 minutes boiling at 95 °C in SDS loading buffer (6.5 % glycerol, 715 mM β-mercaptoethanol, 3 % SDS, 62.5 mM Tris-HCl pH 7.9, 0.005 % Bromphenol blue). Samples were loaded onto gels for separation at constant electrical voltage of 200 V for 45-75 min. Laemmli running buffer (25 mM Tris base, 192 mM glycine, 0.5 g/l SDS) was used for self-made gels, MOPS or MES buffer for Bis-Tris and Tricine buffer for Tris-Acetate gels.

### 5.2.2 Western blot

To transfer proteins onto a nitrocellulose membrane, proteins were first separated by SDS-PAGE as described in section 5.2.1. Proteins were transferred with a constant voltage of 90 V for 90 minutes in an electrophoresis chamber filled with Towbin transfer buffer (25 mM Tris base, 192 mM glycine, 20 % methanol). The chamber was cooled during the transfer using external ice and a sealed ice block. After transferring the proteins, the membrane was stained with Ponceau S in order to confirm the successful transfer of proteins and to mark the protein leader for later discrimination of protein sizes. Subsequently, membranes were rinsed with distilled water and blocked for 10-30 minutes with 1.5 % BSA in TBST (135 mM NaCl, 2.5 mM KCl, 14 mM Tris HCl, 11 mM Tris base and 0.01 % Tween-20) at RT to avoid unspecific binding of antibodies. Proteins were detected by incubating the membrane overnight at 4 °C with respective antibodies listed in Table 5. After incubation of the membrane with the primary antibody, membranes were washed three times for 10 minutes with TBST and incubated with an HRP-conjugated secondary antibody for 2 hours at RT. Finally, the

membranes were washed three times with TBST for 10 min. For detection of the secondary antibody by chemiluminescence, the membranes were covered in a 1:1 dilution of the Clarity™ Western ECL Blotting Substrates, a peroxidase solution, and a luminol/enhancer solution. Signals were detected at different time points by Amersham Imager 600.

**Table 5: Antibodies used for protein detection**

| Antibody/Species | Company/Order # | Preparation | Incubation |
|---|---|---|---|
| anti-H3<br>goat | Abcam<br>Ab12079 | 1:1,000 in TBST with<br>1.5 % BSA | overnight at 4 °C |
| anti-Mcm2<br>rabbit | Santa Cruz Biotech<br>sc-9839 | 1:2,000 in TBST with<br>1.5 % BSA | overnight at 4 °C |
| anti-Mcm4<br>rabbit | Abcam<br>ab4459 | 1:1,000 in TBST with<br>1.5 % BSA | overnight at 4 °C |
| anti-Mcm6<br>rabbit | Bethyl Laboratories<br>A300-194A | 1:2,000 in TBST with<br>1.5 % BSA | overnight at 4 °C |
| anti-Mcm7<br>mouse | Abcam<br>ab2360 | 1:250 in TBST with<br>1.5 % BSA | overnight at 4 °C |
| anti-PCNA<br>mouse (PC10) | Tissue Culture Facility<br>Francis Crick Institute | 1:1,000 in TBST with<br>1.5 % BSA | overnight at 4 °C |
| anti-XCdc45<br>rabbit | self-made | 1:1,000 in TBST with<br>1.5 % BSA | overnight at 4 °C |
| anti-XPolE2<br>rabbit | self-made | 1:1,000 in TBST with<br>1.5 % BSA | overnight at 4 °C |
| anti-XSld5<br>rabbit | self-made | 1:1,000 in TBST with<br>1.5 % BSA | overnight at 4 °C |
| anti-XTopBP1 #2<br>rabbit | self-made | 1:3,000 in TBST with<br>1.5 % BSA | overnight at 4 °C |
| anti-mouse HRP<br>goat | Jackson Laboratories<br>115-035-174 | 1:5,000 in TBST with<br>1.5 % BSA | 2 hours at RT |
| anti-rabbit HRP<br>donkey | Jackson Laboratories<br>711-035-152 | 1:10,000 in TBST with<br>1.5 % BSA | 2 hours at RT |
| anti-goat HRP<br>donkey | Jackson Laboratories<br>705-035-003 | 1:10,000 in TBST with<br>1.5 % BSA | 2 hours at RT |

### 5.2.3 Coomassie staining of proteins separated by SDS-PAGE

The proteins separated by SDS-PAGE were detected by incubating the gels in Colloidal Coomassie (900 mM $(NH_4)_2SO_4$, 12 % Phosphoric Acid (v/v), 1 mg/ml Coomassie Brilliant G-250) supplemented with 20 % methanol. Subsequently gels

were destained with distilled water. After destaining, gels were imaged with an Amersham Imager 600.

### 5.2.4 Silver staining of proteins separated by SDS-PAGE

Silver staining of SDS PAGE was performed according to the Blum method. The separated proteins were detected by incubating gels for 45 minutes in fixation solution (50 % methanol, 12 % acidic acid and 0.02 % formaldehyde). Gels were washed twice for 10 minutes, first in 50 % ethanol and afterward in 30 % ethanol. Next, gels were sensitized for 1 minute in 0.014 % sodium thiosulfate reagent and washed three times in $H_2O$ for 20 seconds. Afterwards gels were incubated for 20 minutes in staining solution (0.1 % $AgNO_3$, 0.03 % formaldehyde) and washed twice for 20 seconds in $H_2O$. Subsequently gels were incubated with developer solution (60 g/l $Na_2CO_3$, 0.02% formaldehyde, 2 g/l $Na_2S_2O_3$ x $5H_2O$). The reaction was stopped by adding 5 % acidic acid when protein staining became visible.

### 5.2.5 Analytical size exclusion chromatography

A Superose 6 Increase 3.2/300 column installed to an ETTAN FPLC system was used for analytic size exclusion chromatography. The column was equilibrated with buffer (50 mM HEPES pH 7.4, 100 mM KCl, 5 mM $MgCl_2$, 0.1 % Triton X-100, 2.5 mM MgATP, 0.1 mM PMSF and 1 x protease inhibitor cocktail). 50 µl protein samples were centrifuged at 4 °C for 10 minutes at 10,000 g to remove precipitated proteins. Soluble protein sample was loaded onto the column and size exclusion chromatography was performed with a flow rate of 40 µl/minute at 4 °C. 100 µl elution fraction were collected and analyzed by SDS PAGE.

### 5.2.6 Pulldown of recombinant TopBP1 fragments by immobilized GINS

For *in vitro* pulldown experiments using recombinant GINS and TopBP1, 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 2,400 ng (Coomassie-scale) GINS via 3 x FLAG-Sld5. 100 nM final concentration of TopBP1-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, 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. The reaction was

incubated for 45 minutes at 4 °C rotating. Then, beads were washed three times for 5 minutes with interaction buffer and boiled in 50 µl Laemmli buffer (10 % SDS, 200 mM Tris-HCL pH 6.0, 20 % glycerol, 2.16 M β-mercaptoethanol, 0.01 % bromophenol blue) before SDS PAGE. Pulldown of recombinant TopBP1 fragments was performed by Bilal Tetik (Day et al., 2023).

## 5.3 Purification of recombinant proteins

### 5.3.1 Purification of TopBP1-Strep from Sf9 insect cells

Sf9 insect cells were grown as a suspension culture 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 x $10^6$ cells/ml) were infected with a 1:100 dilution of recombinant baculovirus carrying an expression construct for TopBP1-1-766-Strep WT or mutants and were incubated at 27 °C for 72 hours. 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 44800 g for 45 minutes, and the supernatant was loaded onto the StrepTrapHP-1ml column (Cytiva, 28907546) using an Äkta FPLC system. The bound protein was eluted with 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). Recombinant TopBP1 proteins were purified by Bilal Tetik (Day et al., 2023).

### 5.3.2 Purification of *Xenopus laevis* geminin from *E. coli*

GST-tagged *Xenopus* geminin (Xgeminin) in pGEX was purified from 6 l of Rosetta *E. coli* culture. Expression of GST-Xgeminin was induced with 1 mM IPTG at OD600 = 0.8 at 25 °C overnight. Cells were harvested by centrifugation at 6000 rpm (JLA-8.1, Avanti JXN-26 Beckman Culture) for 15 min. Subsequently, cells were resuspended in lysis buffer (5 mg/ml lysozyme, 20 mM HEPES pH 7.7, 200 mM NaCl, 5 mM β-mercaptoethanol, 5 % glycerol) and lysed by sonication. Cell suspension was centrifuged at 20000 rpm (SS-34, Sorvall RC 6 PLUS) for 20 minutes at 4 °C. Supernatant was incubated with glutathione-sepharose (GE Healthcare 17513201) for 3 hours before elution with lysis buffer + 50 mM glutathione (Applichem A2084,0025). Peak fractions were pooled and dialyzed into XBE2 buffer (100 mM KCl, 2 mM $MgCl_2$, 0.1 mM $CaCl_2$, 1.71 % (w/v) sucrose, 5 mM K-EGTA, 10 mM HEPES-KOH, pH 7.7)

and concentrated to 360 µM Xgeminin. Aliquots were frozen at - 80 °C (Day et al., 2023).

## 5.4 Molecular biology methods

### 5.4.1 Polymerase Chain Reaction

Polymerase chain reaction (PCR) was used to amplify DNA fragments using specific oligonucleotide primers. PCRs were performed using Phusion High-Fidelity DNA Polymerase (NEB) according to the manufacturer's introduction in a thermal cycler with the following cycling condition:

**Table 6: PCR cycling conditions**

| Time | Temperature | Cycles |
|---|---|---|
| 5 min | 98 °C | Initial denaturation |
| 20 sec | 98 °C | 10 x |
| 20 sec | 59 °C | |
| 1 min | 72 °C | |
| 20 sec | 98 °C | 20x |
| 20 sec | 72 °C | |
| 1 min | 72 °C | |
| 5 min | 72 °C | |

### 5.4.2 Agarose gel electrophoresis

DNA was separated by size using agarose gel electrophoresis. Gels were prepared by boiling 0.75-2 % agarose in TAE buffer (8 mM Tris base, 0.02 mM EDTA, 0.4 mM acetic acid). After cooling down the solution, DNA stain HDGreen (3 µl/50 ml TAE) was added and gels were casted into an agarose gel electrophoresis chamber. After the gels were polymerized, the chamber was filled with TAE buffer. DNA samples were supplemented with 1x DNA loading buffer and loaded onto the gel. For verification of the size of DNA fragments the DNA size standard O'GeneRuler 1 kb DNA Ladder (Thermo Scientific) was loaded as a reference. DNA fragments were separated according to size by applying a constant voltage of 80-120 V for 30-60 minutes. A Gel Doc system (Bio-Rad) was used to visualize DNA bound to HDGreen by excitation with UV light. DNA was cut out and purified for cloning using the NucleoSpin PCR and Gel-Clean Up Kits (Machery-Nagle) according to the manufacturer's instructions.

### 5.4.3 Cloning

For the integration DNA fragments into plasmid vectors, DNA fragments were first amplified by PCR using primes with introduced restriction sites for the restriction enzymes FseI (forward primer) or AscI (reverse primer). DNA fragments and plasmid vector were FseI and AscI digested. Subsequently, DNA inserts were ligated into the vector using the T4 DNA ligase (NEB). The reaction was incubated overnight at 14 °C. Finally, vector plasmids with integrated DNA fragments were purified using NucleoSpin PCR and Gel-Clean Up Kits (Machery-Nagle) according to the manufacturer's instructions.

### 5.4.4 Transformation of chemical competent *E. coli*

NEB5α and Rosetta *E. coli* cells were thawed on ice to amplify DNA plasmids. 10-20 µl chemically competent cells were supplemented with an appropriate amount of transforming DNA and incubated for 30 minutes on ice. Cells were heat shocked at 42 °C for 45 seconds and placed back on ice for 2 minutes. *E. coli* cells were supplemented with 90-180 µl LB-medium and incubated at 37 °C for one hour. Finally, cells were plated onto LB plates containing ampicillin for selection, and incubated overnight at 37 °C.

### 5.4.5 Isolation of vector Plasmids from *E. coli*

For isolation of plasmid DNA from *E. coli* single colonies were picked from LB plates and incubated in 5 ml LB-medium supplemented with ampicillin. After overnight incubation at 37 °C, plasmid DNA was isolated using either NucleoSpin Plasmid Kits or NucleoSpin Midi Kit (Machery-Nagle) according to the manufacturer's instructions.

## 5.5 Cell culture

Cell lines were kept in incubators at 37 °C and 5 % $CO_2$. Media and solutions used for maintaining cell lines in culture were sterile-filtered. Biosafety cabinets were sterilized with 70 % ethanol before and after cells were treated or split. Media and solutions were pre-warmed to 37 °C in a water bath before use. DMEM (Dulbecco's Modified Eagle Medium) containing 10 % fetal calf serum and 1 % penicillin/streptomycin for culturing cell lines and split twice weekly. Therefore, the spent cell culture media was discarded, and cells were washed with DPBS (Dulbecco's Phosphate Buffered Saline). Cells were dissociated from the by incubating with TrypLE Express for 2-5 minutes at 37 °.

Subsequently, cells were dissolved from the dish by gently dispersing media over the cell layer surface several times. Cells were resuspended, and the appropriate volume was pipetted into new cell culture dishes to achieve the desired cell concentration. Finally, the cells were returned to the incubator.

### 5.5.1 Generation of stabile HeLa Flip-In cell lines

To generate stable HeLa Flip-In T-Rex cell lines (kind gift from Dr. Gerhard Müller-Newen (Fahrenkamp et al., 2015)) stably expressing APEX-TopBP1, cells were seeded into 6 cm cell culture dishes. 3 x $10^6$ cells were seeded per dish, containing 2 ml Dulbecco's Modified Eagle Medium (DMEM) supplemented with 10 % fetal calf serum (FCS). The next day, transfection reactions were prepared: Solution A contained 4 µg pOG44 Flp-Recombinase Expression Vector (Thermo Fisher), 4 µg pcDNA5 vector harboring the gen of interest and 500 µl Minimum Essential Media (Opti-MEM). Solution B contained 500 µl Opti-MEM supplemented with 20 µl Lipofectamine 2000 transfection reagent. Both reactions were first incubated separately for 5 minutes at room temperature. Subsequently, both reactions were mixed and incubated for another 15-20 minutes at room temperature. The transfection mix was added dropwise to the cells. After incubation for 24 hours at 37 °C with 5 % $CO_2$ cells were split into two 15 cm dishes, and the medium was exchanged to a selection medium (DMEM, 10 % FCS, 200 µg/ml hygromycin B). During the following two weeks, cells were cultivated in selection medium.

### 5.5.2 Generation of stabile 293 Flip-In cell lines

To generate isogenic stable 293 Flip-In cell lines (kindly provided by the Meyer Laboratory) stably expressing APEX tagged TopBP1, cells were seeded into a 6-well plate. 1.6 x $10^5$ cells were seeded per well, containing 2 ml Dulbecco's Modified Eagle Medium (DMEM) supplemented with 10 % fetal calf serum (FCS). The next day, 1.8 µg pOG44 Flp-Recombinase Expression Vector (Thermo Fisher), 0.2 µg pcDNA5 vector harboring the gen of interest, 200 µl Jet Prime buffer and 3.6 µl Jet Prime transfection reagent (Polyplus) was added. The reaction was mixed and incubated for 15-20 minutes at room temperature. The transfection mix was added dropwise to the cells. After incubation for 4 hours at 37 °C with 5 % $CO_2$ the medium was exchanged for DMEM containing 10 % FCS. The next day cells were split into two 10 cm dishes.

During the following two weeks, cells were selected in DMEM containing 10 %FCS, 1 % penicillin/streptomycin, 100 µg/ml hygromycin B and 15 µg/µl blasticidin.

### 5.5.3 Transient transfection of 239 T cells

0.8 x $10^6$ 239T cells were seeded onto a 6 cm dish 24 hours before transfection. For calcium phosphate transfection, 4.4 µg plasmid DNA listed in Table 7 were incubated for 5 minutes with 2 M $CaCl_2$, 200 µl sterile filtered water and 2x HEPES buffered saline. While incubating the transfection mix, cells were supplemented with 25 µl chloroquine. Subsequently, cells were transfected with the transfection mix. Media was replaced by fresh DMEM containing 10 % FCS and 1 % penicillin/streptomycin after 12-16 hours. For proximity biotinylation, cells were processed as described in section 5.5.4 12 hours after cells were supplemented with fresh medium.

**Table 7: Plasmids for transfection**

| Plasmid name | Vector backbone | Recombinant protein |
|---|---|---|
| BT764 | pcDNA5 | APEX2-TopBP1 |
| MP784 | pcDNA5 | TopBP1-APEX2 |

### 5.5.4 APEX2 proximity biotinylation and denaturing enrichment

Proximity labelling was performed in isogenic stable 293 Flip-In (Thermo Fisher Scientific) 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 x $10^6$ cells were seeded onto a 6 cm plate. For each condition, four dishes were used. The day after transfection 90 % of confluent cells were incubated for 30 minutes with 500 µM biotin phenol at 37 °C. Subsequently, cells were incubated for 1 minute at room temperature with or without 1 mM $H_2O_2$. The biotinylation reaction was stopped by discarding the medium and washing cells tree times with quenching buffer (10 mM sodium azide, 10 mM sodium ascorbate and 5 mM Trolox in Dulbecco's PBS (DPBS; Thermo Fisher Scientific). Cells were removed from the dishes with the dishes with DPBS, harvest into reaction tubes, and washed again with DPBS. Next, cells were lysed at 95 °C for 10 minutes 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). Samples were incubated for 15 minutes at 4 °C and centrifuged at

21,000 g for 2 minutes at 4 °C. Supernatant was incubated overnight with 15 µl streptavidin sepharose beads (GE Healthcare) per sample at 4 °C. After incubation beads were washed once with TNN buffer + 0.1 % SDS and twice with 25 mM ammonium bicarbonate and then processed for mass spectrometry (Day et al., 2023).

### 5.5.5 Mass spectrometry of streptavidin-purified proteins

Biotinylated proteins from biotin pulldown experiments were enriched on 15 µl streptavidin sepharose beads. Captured proteins were washed twice with $H_2O$ prior to on-bead digestion to remove MS-incompatible buffer components. The beads were taken up in 100 µl 0.8 M Urea, 50 mM ammonium bicarbonate buffer (ABC) and supplemented with 5 mM Dithiothreitol (DTT). After incubation at 37 °C for 30 minutes, 10 mM Iodoacetamide (IAM) was added and incubated for another 30 minutes at room temperature in the dark while shaking at 1,500 rpm (Thermomixer C, Eppendorf). In order to quench unreacted IAM, the DTT concentration was raised to a final of 11 mM. Trypsin digestion was initiated by adding 300 ng Trypsin to each sample and placing the samples at 37 °C while shaking at 1,150 rpm overnight. The tryptic digestion was stopped the next day by adding a final of 1 % (v/v) formic acid (FA) to the samples. The beads were collected by centrifugation (650 g, 5 minutes), and 100 µl of the supernatant was transferred to an Eppendorf Lo-bind 1.5 ml tube. The remaining beads were incubated with 50 µl 1 % formic acid for 5 minutes at room temperature and 1,150 rpm. Beads were collected by centrifugation at 650 g for 5 minutes. The supernatant was combined with the first one. The sepharose beads were discarded. In order to remove residual sepharose beads, the combined peptide-containing solutions were passed over pre-equilibrated (50 µl 0.5 % formic acid) home-made 2-disc Glass microfiber StageTip (disc material: GE Healthcare; pore size: 1.2 µM; thickness: 0.26 mm; 50 g; 2 minutes). The cleared tryptic digests were then desalted on home-made C18 StageTips as described in Rappsilber et al. (Rappsilber et al., 2007). 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 µl) and directly used for LC-MS/MS experiments (Day et al., 2023).

Sample preparation for LC-MS/MS and mass spectrometry was performed by Dr. Farnusch Kaschani, Svenja Heimann and Jenny Bormann.

## 5.6 Bioinformatical methods

### 5.6.1 Quantification of signal intensities from replication assays

Signals from replication assays were quantified using the free software Fiji/ImageJ. Radioactive signals on glass fiber membranes were phosphor imaged first. Only samples from the same glass fiber membrane were compared to each other. The row-integrated density was measured for each signal by placing a circle around the signal. Zero-minute time points of each condition was subtracted as background as well as background signals of the buffer control. Processed signal intensities were normalized to TopBP1 BRCT0-5 wild type.

### 5.6.2 Statistical analysis in GraphPad

Graphs were generated with the software GraphPad Prism. For mass spectrometry analysis, processed data generated in Perseus were plotted as volcano scatter plots. For replication analysis in *Xenopus* egg extract, either single experiments or three biological replicates were shown. The standard error of the mean is plotted for statistical analysis of biological replicates.

### 5.6.3 Statistical analysis in Perseus

Statistical analysis of proximity biotinylation experiments were generated in Perseus (version 1.6.15.0). Protein groups were filtered to eliminate contaminants, reverse hits, and proteins identified by site only. For volcano plots, LFQ intensities were log2 transformed. To identify proteins with significant abundance changes between the two conditions, missing values of the control samples without $H_2O_2$ were imputed with random values drawn from a normal distribution centered around the detection limit of the MS instrument (Perseus imputation; width=0.3, down shift = 1.8). Two sample Student's t-tests with permutation-based FDR control were carried out in Perseus. For these tests, seven valid values in total were required. FDR was adjusted for multiple testing by the Benjamini-Hochberg procedure using a significance threshold of FDR<0.05. Data visualization was carried out in GraphPad Prism.

## 5.7 Material

**Table 8: Buffer and Solutions**

| Buffer/ Solution | Composition |
|---|---|
| 1x MMR | 100mM NaCl<br>2 mM KCl<br>1 mM $MgCl_2$<br>2 mM $CaCl_2$<br>0.1 mM EDTA<br>5 mM HEPES-NaOH, pH 7.8 |
| 1x SDS sample buffer | 6.5 % glycerol<br>715 mM β-mercaptoethanol<br>3 % SDS<br>62.5 mM Tris-HCl pH 7.9<br>0.005 % Bromphenol blue |
| 6x DNA loading buffer | 48 % glycerol<br>6 mM EDTA<br>Orange G as desired |
| Colloidal Coomassie | 900 mM $(NH_4)_2SO_4$<br>12 % Phosphoric Acid (v/v)<br>1 mg/ml Coomassie Brilliant G-250 |
| Dejelly buffer | 2 % cysteine (w/v) in $H_2O$, pH 7.8 with NaOH |
| Denaturing buffer | 8 M Urea<br>100 mM Tris pH 8 |
| Developer solution | 60 g/l $Na_2CO_3$<br>0.02% formaldehyde<br>2 g/l $Na_2S_2O_3$ x $5H_2O$ |
| Dilution buffer | 100 mM ammonium bicarbonate |
| EBL salt | 10 mM HEPES-KOH, pH 7.7<br>50 mM KCl<br>2.5 mM $MgCl_2$ |
| Energy regenerator | 1 M phosphocreatine K salt<br>600 µg/ml creatine phosphokinase in 10 mM HEPES-KOH pH 7.6 |
| Fixation solution | 50 % methanol<br>12 % acidic acid<br>0.02 % formaldehyde |
| Laemmli Running Buffer | 25 mM Tris base<br>192 mM glycine<br>0.5 g/l SDS |
| LFB1/50 | 10 % (w/v) sucrose<br>50 mM KCl<br>2 mM $MgCl_2$<br>1 mM EGTA<br>2 mM DTT<br>20 mM $K_2HPO_4$/ $KH_2PO_4$ pH 8.0 |

| | |
|---|---|
| | 40 mM HEPES-KOH, pH 8.0 |
| Nuclear Isolation Buffer (NIB) | 50 mM HEPES-KOH pH 7.6<br>50 mM KCl<br>5 mM $MgCl_2$<br>0.5 mM spermidine<br>0.15 mM spermine<br>0.1 % Triton X-100<br>2 mM DTT<br>1 µg/ml leupeptin<br>1 µg/ml pepstatin<br>1 µg/ml chymostatin<br>0.1 mM PMSF |
| Quenching buffer | 10 mM sodium azide<br>10 mM sodium ascorbate<br>5 mM Trolox<br>in DPBS |
| Sensitizer solution | 0.014 % sodium thiosulfate |
| Staining solution | 0.1 % $AgNO_3$<br>0.03 % formaldehyde |
| TAE | 8 mM Tris base<br>0.02 mM EDTA<br>0.4 mM acetic acid |
| TBST | 135 mM NaCl<br>2.5 mM KCl<br>14 mM Tris HCl<br>11 mM Tris base<br>0.01 % Tween-20 |
| TNN | 20 mM Tris pH 7.9<br>200 mM NaCl<br>0.5 % NP-40 alternative |
| Towbin | 25 mM Tris base<br>192 mM glycine<br>20 % methanol (v/v) |
| Tris-buffered Saline (TBS) | 135 mM NaCl<br>2.5 mM KCl<br>14 mM Tris HCl<br>11 mM Tris base |
| TSD | 50 mM Tris pH 7.5<br>5 mM DTT<br>1 % SDS |
| XBE2 | 100 mM KCl<br>2 mM $MgCl_2$<br>0.1 mM $CaCl_2$<br>1.71 % (w/v) sucrose<br>5 mM K-EGTA<br>10 mM HEPES-KOH, pH 7.7 |

**Table 9: Chemical and Components**

| Chemical/ Component | Company |
|---|---|
| (+)-Sodium L-ascorbate | Sigma-Aldrich |
| $(NH_4)_2SO_4$ | Carl Roth |
| 2 % Bis-Acrylamide | Bio-Rad Laboratories GmbH |
| 40 % Acrylamide Solution | Bio-Rad Laboratories GmbH |
| Acetic Acid | VWR Chemicals |
| Agarose (for gel electrophoresis) | Carl Roth |
| $AgNO_3$ | Sigma-Aldrich |
| Ammonium Bicarbonate | Sigma-Aldrich |
| Aphidicolin | Sigma-Aldrich |
| Aprotinin | Sigma-Aldrich |
| APS | Merck |
| AscI (10,000 units/ml) | New England BioLabs |
| Benzonase | Sigma-Aldrich |
| Blasticidin | Thermos Fisher Scientific |
| Bromphenol blue | AppliChem |
| BSA | Carl Roth |
| $CaCl_2$ | Carl Roth |
| ClarityTM Western ECL Substrate | Bio-Rad Laboratories GmbH |
| cOmplete, EDTA-free Protease Inhibitor Cocktail | La Roche |
| Coomassie brilliant Blue | AppliChem |
| Creatine Phosphokinase | Sigma-Aldrich |
| Cysteine | Sigma-Aldrich |
| Cytochalasin D | Sigma-Aldrich |
| DAPI | Life Technologies |
| DMEM (Dulbecco's Modified Eagle Medium) | Thermos Fisher Scientific |
| DMSO (Dimethyl Sulfoxide) | AppliChem |
| DPBS (Dulbecco’s Phosphate Buffered Saline) | Thermos Fisher Scientific |
| DTT Dithiothreitol | Sigma-Aldrich |
| EDTA | AppliChem |
| EGTA | Sigma-Aldrich |
| Fetal Calf Serum | Thermos Fisher Scientific |
| Formaldehyde | Carl Roth |
| FseI (2,000 units/ml) | New England BioLabs |
| Gibco Penicillin-Streptomycin (10,000 U/ml) | Thermos Fisher Scientific |

| | |
|---|---|
| Glycerol | AppliChem |
| HDGreen Plus | Intas Science Imaging |
| HEPES | Carl Roth |
| Human Chorionic Gonadotrophin | Sigma-Aldrich |
| Hydrogen Peroxidase Solution 3 % | Sigma-Aldrich |
| Hygromycin B | InvovGen |
| IPTG | Thermo Fisher Scientific |
| Isopropanol | Bernd Kraft |
| Jet Prime Buffer | Polyplus |
| Jet Prime Transfection Reagent | Polyplus |
| $K_2HPO_4$ | Merck |
| KCl | Carl Roth |
| $KH_2PO_4$ | Sigma-Aldrich |
| LB-Agar | Carl Roth |
| Leupeptin | Merck |
| Lipofectamine 2000 | Invitrogen |
| Methanol | Fisher Chemicals |
| $MgCl_2$ | Sigma-Aldrich |
| $Na_2CO_3$ | Merck |
| $Na_2S_2O_3$ x $5H_2O$ | Sigma-Aldrich |
| NaCl | Bernd Kraft |
| NaOH | Bernd Kraft |
| NEB 5-alpha | New England BioLabs |
| NP-40 (alternative) | Calbiochem |
| O'GeneRuler 1 kb | Thermo Fisher Scientific |
| Opti-MEM | Thermo Fisher Scientific |
| PageRulerTM Prestained Protein Ladder | Thermo Fisher Scientific |
| PageRulerTM Unstained Protein Ladder | Thermo Fisher Scientific |
| Pepstatin | Carl Roth |
| Phosphocreatine K salt | Sigma-Aldrich |
| Phosphoric Acid | Applichem |
| Phusion High-Fidelity DNA Polymerase | New England BioLabs |
| PMSF | Sigma-Aldrich |
| Ponceau S | Sigma-Aldrich |
| Pregnant Mare Serum Gonadotrophin | Sigma-Aldrich |
| Roscovitine | Merck |

| | |
|---|---|
| SDS | Carl Roth |
| Sodium Azide pure | AppliChem |
| Sodium Thiosulfate | Sigma-Aldrich |
| Spermidine | Sigma-Aldrich |
| Spermine | Sigma-Aldrich |
| Streptavidin Sepharose Beads | GE Healthcare |
| Sucrose | AppliChem |
| TEMED | Merck |
| Trichloroacetic Acid | AppliChem |
| Tris Hydrochloride | Carl Roth |
| Tris Base | Carl Roth |
| Triton X-100 | AppliChem |
| TrypLE Express | Thermo Fisher Scientific |
| Trypsin | Sigma-Aldrich |
| Tween-20 | Sigma-Aldrich |
| Urea | Carl Roth |
| XT MES Running Buffer | Bio-Rad Laboratories GmbH |
| XT MOPS Running Buffer | Bio-Rad Laboratories GmbH |
| XT Tricine Running Buffer | Bio-Rad Laboratories GmbH |
| $\alpha^{32}$P-dCTP | Perkin Elmer |
| β-mercaptoethanol | Carl Roth |

**Table 10: List of Kits**

| Kit | Company |
|---|---|
| NucleoBond Xtra Midi/Maxi | Machery-Nagel |
| NucleoSpin Gel- and PCR-Clean up | Machery-Nagel |
| NucleoSpin Plasmid | Machery-Nagel |

**Table 11: List of Plasmids**

| Plasmid name | Vector backbone | Recombinant protein |
|---|---|---|
| BT702 | pLib | TopBP1 1-766 LL489/490PP-Strep |
| BT703 | pLib | TopBP1 1-766 SQ492AQ |
| BT704 | pLib | TopBP1 1-766 SQ492EQ |
| BT707 | pLib | TopBP1 1-766-Strep |
| BT708 | pLib | TopBP1 1-766 RegionI-Strep |

| | | |
|---|---|---|
| BT709 | pLib | TopBP1 1-766 RegionII-Strep |
| BT710 | pLib | TopBP1 1-766 RegionIII-Strep |
| BT711 | pLib | TopBP1 1-766 Gcore-Strep |
| BT712 | pLib | TopBP1 1-766 Gcc-Strep |
| BT759 | pLib | TopBP1 1-549-Strep |
| BT760 | pLib | TopBP1 1-549 Gcc-Strep |
| BT761 | pLib | TopBP1 1-549 LL489/490PP-Strep |
| BT764 | pcDNA5 | APEX2-TopBP1 |
| BT770 | pLib | TopBP1 1-766 B4mut-Strep |
| BT772 | pLib | TopBP1 1-766 Gcc/B4mut-Strep |
| MP783 | pcDNA5 | APEX2 |
| MP784 | pcDNA5 | TopBP1-APEX2 |

**Table 12: List of Primers**

| Name | Sequence | Purpose |
|---|---|---|
| TopBP1 5'F | AATGGCCGGCCAatgtccagaaatgacaaagaacc | cloning |
| TopBP1 3'A_open | tatGGCGCGCCtGTGTACTCTAGGTCGTTTGAT | cloning |

**Table 13: Equipment and Devices**

| Equipment/Devices | Company |
|---|---|
| AL54 Analytical Balance | Mettler Toledo GmbH |
| Amersham Imager 600 | GE Healthcare |
| Avanti JXN-26 Centrifuge | Beckman Coulter |
| Axio vert. A1 microscope | Zeiss |
| Centrifuge 5427 R | Eppendorf |
| Centrifuge 5810 R | Eppendorf |
| Countess II | Thermo Fisher Scientific |
| Criterion Blotter | Bio-Rad Laboratories GmbH |
| Easypet | Eppendorf AG |
| Eppendorf Research Plus | Eppendorf AG |
| Eraser raytest | GE Healthcare |
| ETTAN | GE Healthcare |
| gel electrophoresis camber | Bio-Rad Laboratories GmbH |
| New BrunswickTM Galaxy 170R | Eppendorf AG |
| Optima MAX-XP Ultracentrifuge | Beckman Coulter |

| | |
|---|---|
| Phosphor Imager Fujifilm FLA-3000 | GE Healthcare |
| Pioneer Plus Precision | Ohaus Europe GmbH |
| Pipetboy acu 2 | INTEGRA Biosciences GmbH |
| Power Pac Basic Power Supply | Bio-Rad Laboratories GmbH |
| Power Pac HC Power Supply | Bio-Rad Laboratories GmbH |
| Serological Pipettes | Sarstedt |
| Thermomixer Comfort | Eppendorf AG |
| Ultra-Clear Tubes | Beckman Coulter |
| Vortex Genie 2 | Scientific Industries, Inc. |
| Water bath WNB 14 | Memmert GmbH + Co. KG |

**Table 14: Software**

| Name | Provider |
|---|---|
| Adobe Illustrator 2023 | Adobe |
| Fiji | Open source |
| GraphPad Prism | GraphPad Software |
| Microsoft Office | Microsoft |
| Perseus 1.6.15.0 | MaxQuant |
| SecMan Pro | DNA Star |
| SeqBuilder | DNA Star |

# 6 References

Abid Ali, F., M.E. Douglas, J. Locke, V.E. Pye, A. Nans, J.F.X. Diffley, and A. Costa. 2017. Cryo-EM structure of a licensed DNA replication origin. *Nat Commun*. 8:2241.

Abid Ali, F., L. Renault, J. Gannon, H.L. Gahlon, A. Kotecha, J.C. Zhou, D. Rueda, and A. Costa. 2016. Cryo-EM structures of the eukaryotic replicative helicase bound to a translocation substrate. *Nat Commun*. 7:10708.

Adam, S., S.E. Rossi, N. Moatti, M. De Marco Zompit, Y. Xue, T.F. Ng, A. Alvarez-Quilon, J. Desjardins, V. Bhaskaran, G. Martino, D. Setiaputra, S.M. Noordermeer, T.K. Ohsumi, N. Hustedt, R.K. Szilard, N. Chaudhary, M. Munro, A. Veloso, H. Melo, S.Y. Yin, R. Papp, J.T.F. Young, M. Zinda, M. Stucki, and D. Durocher. 2021. The CIP2A-TOPBP1 axis safeguards chromosome stability and is a synthetic lethal target for BRCA-mutated cancer. *Nat Cancer*. 2:1357-1371.

Barry, E.R., and S.D. Bell. 2006. DNA replication in the archaea. *Microbiol Mol Biol Rev*. 70:876-887.

Bell, S.P., and A. Dutta. 2002. DNA replication in eukaryotic cells. *Annu Rev Biochem*. 71:333-374.

Berezney, R., D.D. Dubey, and J.A. Huberman. 2000. Heterogeneity of eukaryotic replicons, replicon clusters, and replication foci. *Chromosoma*. 108:471-484.

Bigot, N., M. Day, R.A. Baldock, F.Z. Watts, A.W. Oliver, and L.H. Pearl. 2019. Phosphorylation-mediated interactions with TOPBP1 couple 53BP1 and 9-1-1 to control the G1 DNA damage checkpoint. *Elife*. 8.

Blackford, A.N., J. Nieminuszczy, R.A. Schwab, Y. Galanty, S.P. Jackson, and W. Niedzwiedz. 2015. TopBP1 interacts with BLM to maintain genome stability but is dispensable for preventing BLM degradation. *Mol Cell*. 57:1133-1141.

Blow, J.J., X.Q. Ge, and D.A. Jackson. 2011. How dormant origins promote complete genome replication. *Trends Biochem Sci*. 36:405-414.

Blow, J.J., and R.A. Laskey. 1986. Initiation of DNA replication in nuclei and purified DNA by a cell-free extract of Xenopus eggs. *Cell*. 47:577-587.

Boos, D., and P. Ferreira. 2019. Origin Firing Regulations to Control Genome Replication Timing. *Genes (Basel)*. 10.

Boos, D., L. Sanchez-Pulido, M. Rappas, L.H. Pearl, A.W. Oliver, C.P. Ponting, and J.F. Diffley. 2011. Regulation of DNA replication through Sld3-Dpb11 interaction is conserved from yeast to humans. *Curr Biol*. 21:1152-1157.

Boos, D., M. Yekezare, and J.F. Diffley. 2013. Identification of a heteromeric complex that promotes DNA replication origin firing in human cells. *Science*. 340:981-984.

Botchan, M., and J. Berger. 2010. DNA replication: making two forks from one prereplication complex. *Mol Cell*. 40:860-861.

Carroni, M., M. De March, B. Medagli, I. Krastanova, I.A. Taylor, H. Amenitsch, H. Araki, F.M. Pisani, A. Patwardhan, and S. Onesti. 2017. New insights into the GINS complex explain the controversy between existing structural models. *Sci Rep*. 7:40188.

Chang, Y.P., G. Wang, V. Bermudez, J. Hurwitz, and X.S. Chen. 2007. Crystal structure of the GINS complex and functional insights into its role in DNA replication. *Proc Natl Acad Sci U S A*. 104:12685-12690.

Chen, H.Z., S.Y. Tsai, and G. Leone. 2009. Emerging roles of E2Fs in cancer: an exit from cell cycle control. *Nat Rev Cancer*. 9:785-797.

Chong, J.P., M.K. Hayashi, M.N. Simon, R.M. Xu, and B. Stillman. 2000. A double-hexamer archaeal minichromosome maintenance protein is an ATP-dependent DNA helicase. *Proc Natl Acad Sci U S A*. 97:1530-1535.

Classon, M., and E. Harlow. 2002. The retinoblastoma tumour suppressor in development and cancer. *Nat Rev Cancer*. 2:910-917.

Collart, C., G.E. Allen, C.R. Bradshaw, J.C. Smith, and P. Zegerman. 2013. Titration of four replication factors is essential for the Xenopus laevis midblastula transition. *Science*. 341:893-896.

Costa, A., and J.F.X. Diffley. 2022. The Initiation of Eukaryotic DNA Replication. *Annu Rev Biochem*. 91:107-131.

Costa, A., I. Ilves, N. Tamberg, T. Petojevic, E. Nogales, M.R. Botchan, and J.M. Berger. 2011. The structural basis for MCM2-7 helicase activation by GINS and Cdc45. *Nat Struct Mol Biol*. 18:471-477.

Coster, G., and J.F.X. Diffley. 2017. Bidirectional eukaryotic DNA replication is established by quasi-symmetrical helicase loading. *Science*. 357:314-318.

Day, M., B. Tetik, M. Parlak, Y. Almeida-Hernández, M. Räschle, F. Kaschani, H. Siegert, A. Marko, E. Sanchez-Garcia, M. Kaiser, I.A. Barker, L.H. Pearl, A.W. Oliver, and D. Boos. 2023. TopBP1 utilises a bipartite GINS binding mode to activate the replicative helicase. *bioRxiv*:2023.2003.2031.535063.

De Jesus-Kim, L., L.J. Friedman, M. Looke, C.K. Ramsoomair, J. Gelles, and S.P. Bell. 2021. DDK regulates replication initiation by controlling the multiplicity of Cdc45-GINS binding to Mcm2-7. *Elife*. 10.

Deegan, T.D., J.T. Yeeles, and J.F. Diffley. 2016. Phosphopeptide binding by Sld3 links Dbf4-dependent kinase to MCM replicative helicase activation. *EMBO J*. 35:961-973.

Delacroix, S., J.M. Wagner, M. Kobayashi, K. Yamamoto, and L.M. Karnitz. 2007. The Rad9-Hus1-Rad1 (9-1-1) clamp activates checkpoint signaling via TopBP1. *Genes Dev*. 21:1472-1477.

Diffley, J.F. 1996. Once and only once upon a time: specifying and regulating origins of DNA replication in eukaryotic cells. *Genes Dev*. 10:2819-2830.

Diffley, J.F. 2001. DNA replication: building the perfect switch. *Curr Biol*. 11:R367-370.

Diffley, J.F. 2004. Regulation of early events in chromosome replication. *Curr Biol*. 14:R778-786.

Dimitrova, D.S., and D.M. Gilbert. 1999. The spatial position and replication timing of chromosomal domains are both established in early G1 phase. *Mol Cell*. 4:983-993.

Douglas, M.E., F.A. Ali, A. Costa, and J.F.X. Diffley. 2018. The mechanism of eukaryotic CMG helicase activation. *Nature*. 555:265-268.

Douglas, M.E., and J.F.X. Diffley. 2016. Recruitment of Mcm10 to Sites of Replication Initiation Requires Direct Binding to the Minichromosome Maintenance (MCM) Complex. *J Biol Chem*. 291:5879-5888.

Drury, L.S., G. Perkins, and J.F. Diffley. 1997. The Cdc4/34/53 pathway targets Cdc6p for proteolysis in budding yeast. *EMBO J*. 16:5966-5976.

Drury, L.S., G. Perkins, and J.F. Diffley. 2000. The cyclin-dependent kinase Cdc28p regulates distinct modes of Cdc6p proteolysis during the budding yeast cell cycle. *Curr Biol*. 10:231-240.

Elsasser, S., Y. Chi, P. Yang, and J.L. Campbell. 1999. Phosphorylation controls timing of Cdc6p destruction: A biochemical analysis. *Mol Biol Cell*. 10:3263-3277.

Evrin, C., P. Clarke, J. Zech, R. Lurz, J. Sun, S. Uhle, H. Li, B. Stillman, and C. Speck. 2009. A double-hexameric MCM2-7 complex is loaded onto origin DNA during licensing of eukaryotic DNA replication. *Proc Natl Acad Sci U S A*. 106:20240-20245.

Fahrenkamp, D., H.S. de Leur, A. Kuster, N. Chatain, and G. Muller-Newen. 2015. Src family kinases interfere with dimerization of STAT5A through a phosphotyrosine-SH2 domain interaction. *Cell Commun Signal*. 13:10.

Ferreira, P., V. Hofer, N. Kronshage, A. Marko, K.U. Reusswig, B. Tetik, C. Diessel, K. Kohler, N. Tschernoster, J. Altmuller, N. Schulze, B. Pfander, and D. Boos. 2021. MTBP phosphorylation controls DNA replication origin firing. *Sci Rep*. 11:4242.

Ferreira, P., L. Sanchez-Pulido, A. Marko, C.P. Ponting, and D. Boos. 2022. Refining the domain architecture model of the replication origin firing factor Treslin/TICRR. *Life Sci Alliance*. 5.

Francis, L.I., J.C. Randell, T.J. Takara, L. Uchima, and S.P. Bell. 2009. Incorporation into the prereplicative complex activates the Mcm2-7 helicase for Cdc7-Dbf4 phosphorylation. *Genes Dev*. 23:643-654.

Frigola, J., J. He, K. Kinkelin, V.E. Pye, L. Renault, M.E. Douglas, D. Remus, P. Cherepanov, A. Costa, and J.F.X. Diffley. 2017. Cdt1 stabilizes an open MCM ring for helicase loading. *Nat Commun*. 8:15720.

Frigola, J., D. Remus, A. Mehanna, and J.F. Diffley. 2013. ATPase-dependent quality control of DNA replication origin licensing. *Nature*. 495:339-343.

Fu, Y.V., H. Yardimci, D.T. Long, T.V. Ho, A. Guainazzi, V.P. Bermudez, J. Hurwitz, A. van Oijen, O.D. Scharer, and J.C. Walter. 2011. Selective bypass of a lagging strand roadblock by the eukaryotic replicative DNA helicase. *Cell*. 146:931-941.

Fujisawa, T., and P. Filippakopoulos. 2017. Functions of bromodomain-containing proteins and their roles in homeostasis and cancer. *Nat Rev Mol Cell Biol*. 18:246-262.

Fukuura, M., K. Nagao, C. Obuse, T.S. Takahashi, T. Nakagawa, and H. Masukata. 2011. CDK promotes interactions of Sld3 and Drc1 with Cut5 for initiation of DNA replication in fission yeast. *Mol Biol Cell*. 22:2620-2633.

Gaggioli, V., E. Zeiser, D. Rivers, C.R. Bradshaw, J. Ahringer, and P. Zegerman. 2014. CDK phosphorylation of SLD-2 is required for replication initiation and germline development in C. elegans. *J Cell Biol*. 204:507-522.

Galanti, L., and B. Pfander. 2018. Right time, right place-DNA damage and DNA replication checkpoints collectively safeguard S phase. *EMBO J*. 37.

Gambus, A., R.C. Jones, A. Sanchez-Diaz, M. Kanemaki, F. van Deursen, R.D. Edmondson, and K. Labib. 2006. GINS maintains association of Cdc45 with MCM in replisome progression complexes at eukaryotic DNA replication forks. *Nat Cell Biol*. 8:358-366.

Gambus, A., G.A. Khoudoli, R.C. Jones, and J.J. Blow. 2011. MCM2-7 form double hexamers at licensed origins in Xenopus egg extract. *J Biol Chem*. 286:11855-11864.

Garcia, V., K. Furuya, and A.M. Carr. 2005. Identification and functional analysis of TopBP1 and its homologs. *DNA Repair (Amst)*. 4:1227-1239.

Gautier, J., J. Minshull, M. Lohka, M. Glotzer, T. Hunt, and J.L. Maller. 1990. Cyclin is a component of maturation-promoting factor from Xenopus. *Cell*. 60:487-494.

Ge, X.Q., and J.J. Blow. 2010. Chk1 inhibits replication factory activation but allows dormant origin firing in existing factories. *J Cell Biol*. 191:1285-1297.

Ge, X.Q., D.A. Jackson, and J.J. Blow. 2007. Dormant origins licensed by excess Mcm2-7 are required for human cells to survive replicative stress. *Genes Dev*. 21:3331-3341.

Georgoulis, A., C.E. Vorgias, G.P. Chrousos, and E.P. Rogakou. 2017. Genome Instability and gammaH2AX. *Int J Mol Sci*. 18.

Gillespie, P.J., A. Gambus, and J.J. Blow. 2012. Preparation and use of Xenopus egg extracts to study DNA replication and chromatin associated proteins. *Methods*. 57:203-213.

Goswami, P., F. Abid Ali, M.E. Douglas, J. Locke, A. Purkiss, A. Janska, P. Eickhoff, A. Early, A. Nans, A.M.C. Cheung, J.F.X. Diffley, and A. Costa. 2018. Structure of DNA-CMG-Pol epsilon elucidates the roles of the non-catalytic polymerase modules in the eukaryotic replisome. *Nat Commun*. 9:5061.

Greenberg, R.A., B. Sobhian, S. Pathania, S.B. Cantor, Y. Nakatani, and D.M. Livingston. 2006. Multifactorial contributions to an acute DNA damage response by BRCA1/BARD1-containing complexes. *Genes Dev*. 20:34-46.

Greiwe, J.F., T.C.R. Miller, J. Locke, F. Martino, S. Howell, A. Schreiber, A. Nans, J.F.X. Diffley, and A. Costa. 2022. Structural mechanism for the selective phosphorylation of DNA-loaded MCM double hexamers by the Dbf4-dependent kinase. *Nat Struct Mol Biol*. 29:10-20.

Gros, J., C. Kumar, G. Lynch, T. Yadav, I. Whitehouse, and D. Remus. 2015. Post-licensing Specification of Eukaryotic Replication Origins by Facilitated Mcm2-7 Sliding along DNA. *Mol Cell*. 60:797-807.

Guo, C., A. Kumagai, K. Schlacher, A. Shevchenko, A. Shevchenko, and W.G. Dunphy. 2015. Interaction of Chk1 with Treslin negatively regulates the initiation of chromosomal DNA replication. *Mol Cell*. 57:492-505.

Hassan, B.H., L.A. Lindsey-Boltz, M.G. Kemp, and A. Sancar. 2013. Direct role for the replication protein treslin (Ticrr) in the ATR kinase-mediated checkpoint response. *J Biol Chem*. 288:18903-18910.

Havens, C.G., and J.C. Walter. 2011. Mechanism of CRL4(Cdt2), a PCNA-dependent E3 ubiquitin ligase. *Genes Dev*. 25:1568-1582.

Heller, R.C., S. Kang, W.M. Lam, S. Chen, C.S. Chan, and S.P. Bell. 2011. Eukaryotic origin-dependent DNA replication in vitro reveals sequential action of DDK and S-CDK kinases. *Cell*. 146:80-91.

Hoogenboom, W.S., D. Klein Douwel, and P. Knipscheer. 2017. Xenopus egg extract: A powerful tool to study genome maintenance mechanisms. *Dev Biol*. 428:300-309.

Hyrien, O., K. Marheineke, and A. Goldar. 2003. Paradoxes of eukaryotic DNA replication: MCM proteins and the random completion problem. *Bioessays*. 25:116-125.

Itou, H., S. Muramatsu, Y. Shirakihara, and H. Araki. 2014. Crystal structure of the homology domain of the eukaryotic DNA replication proteins Sld3/Treslin. *Structure*. 22:1341-1347.

Jenkinson, F., and P. Zegerman. 2022. Roles of phosphatases in eukaryotic DNA replication initiation control. *DNA Repair (Amst)*. 118:103384.

Johnson, D.G., J.K. Schwarz, W.D. Cress, and J.R. Nevins. 1993. Expression of transcription factor E2F1 induces quiescent cells to enter S phase. *Nature*. 365:349-352.

Jones, K.T. 1998. Ca2+ oscillations in the activation of the egg and development of the embryo in mammals. *Int J Dev Biol*. 42:1-10.

Jones, M.L., Y. Baris, M.R.G. Taylor, and J.T.P. Yeeles. 2021. Structure of a human replisome shows the organisation and interactions of a DNA replication machine. *EMBO J*. 40:e108819.

Kamada, K., Y. Kubota, T. Arata, Y. Shindo, and F. Hanaoka. 2007. Structure of the human GINS complex and its assembly and functional interface in replication initiation. *Nat Struct Mol Biol*. 14:388-396.

Kanemaki, M., and K. Labib. 2006. Distinct roles for Sld3 and GINS during establishment and progression of eukaryotic DNA replication forks. *EMBO J*. 25:1753-1763.

Kanemaki, M., A. Sanchez-Diaz, A. Gambus, and K. Labib. 2003. Functional proteomic identification of DNA replication proteins by induced proteolysis in vivo. *Nature*. 423:720-724.

Kang, S., M.D. Warner, and S.P. Bell. 2014. Multiple functions for Mcm2-7 ATPase motifs during replication initiation. *Mol Cell*. 55:655-665.

Kelly, R.L., A.M. Huehls, A. Venkatachalam, C.J. Huntoon, Y.J. Machida, and L.M. Karnitz. 2022. Intra-S phase checkpoint kinase Chk1 dissociates replication proteins Treslin and TopBP1 through multiple mechanisms during replication stress. *J Biol Chem*. 298:101777.

Knott, S.R., J.M. Peace, A.Z. Ostrow, Y. Gan, A.E. Rex, C.J. Viggiani, S. Tavare, and O.M. Aparicio. 2012. Forkhead transcription factors establish origin timing and long-range clustering in S. cerevisiae. *Cell*. 148:99-111.

Kohler, K., L. Sanchez-Pulido, V. Hofer, A. Marko, C.P. Ponting, A.P. Snijders, R. Feederle, A. Schepers, and D. Boos. 2019. The Cdk8/19-cyclin C transcription regulator functions in genome replication through metazoan Sld7. *PLoS Biol*. 17:e2006767.

Kubota, Y., Y. Takase, Y. Komori, Y. Hashimoto, T. Arata, Y. Kamimura, H. Araki, and H. Takisawa. 2003. A novel ring-like complex of Xenopus proteins essential for the initiation of DNA replication. *Genes Dev*. 17:1141-1152.

Kumagai, A., and W.G. Dunphy. 2017. MTBP, the partner of Treslin, contains a novel DNA-binding domain that is essential for proper initiation of DNA replication. *Mol Biol Cell*. 28:2998-3012.

Kumagai, A., and W.G. Dunphy. 2020. Binding of the Treslin-MTBP Complex to Specific Regions of the Human Genome Promotes the Initiation of DNA Replication. *Cell Rep*. 32:108178.

Kumagai, A., J. Lee, H.Y. Yoo, and W.G. Dunphy. 2006. TopBP1 activates the ATR-ATRIP complex. *Cell*. 124:943-955.

Kumagai, A., A. Shevchenko, A. Shevchenko, and W.G. Dunphy. 2010. Treslin collaborates with TopBP1 in triggering the initiation of DNA replication. *Cell*. 140:349-359.

Kumagai, A., A. Shevchenko, A. Shevchenko, and W.G. Dunphy. 2011. Direct regulation of Treslin by cyclin-dependent kinase is essential for the onset of DNA replication. *J Cell Biol*. 193:995-1007.

Labib, K., J.F. Diffley, and S.E. Kearsey. 1999. G1-phase and B-type cyclins exclude the DNA-replication factor Mcm4 from the nucleus. *Nat Cell Biol*. 1:415-422.

Laine, A., S.G. Nagelli, C. Farrington, U. Butt, A.N. Cvrljevic, J.P. Vainonen, F.M. Feringa, T.J. Gronroos, P. Gautam, S. Khan, H. Sihto, X. Qiao, K. Pavic, D.C. Connolly, P. Kronqvist, L.L. Elo, J. Maurer, K. Wennerberg, R.H. Medema, H. Joensuu, E. Peuhu, K. de Visser, G. Narla, and J. Westermarck. 2021. CIP2A Interacts with TopBP1 and Drives Basal-Like Breast Cancer Tumorigenesis. *Cancer Res*. 81:4319-4331.

Lee, J., and W.G. Dunphy. 2010. Rad17 plays a central role in establishment of the interaction between TopBP1 and the Rad9-Hus1-Rad1 complex at stalled replication forks. *Mol Biol Cell*. 21:926-935.

Lee, J., A. Kumagai, and W.G. Dunphy. 2007. The Rad9-Hus1-Rad1 checkpoint clamp regulates interaction of TopBP1 with ATR. *J Biol Chem*. 282:28036-28044.

Leonard, A.C., and M. Mechali. 2013. DNA replication origins. *Cold Spring Harb Perspect Biol*. 5:a010116.

Letessier, A., G.A. Millot, S. Koundrioukoff, A.M. Lachages, N. Vogt, R.S. Hansen, B. Malfoy, O. Brison, and M. Debatisse. 2011. Cell-type-specific replication initiation programs set fragility of the FRA3B fragile site. *Nature*. 470:120-123.

Lewis, J.S., M.H. Gross, J. Sousa, S.S. Henrikus, J.F. Greiwe, A. Nans, J.F.X. Diffley, and A. Costa. 2022. Mechanism of replication origin melting nucleated by CMG helicase assembly. *Nature*. 606:1007-1014.

Li, J., J. Dong, W. Wang, D. Yu, X. Fan, Y.C. Hui, C.S.K. Lee, W.H. Lam, N. Alary, Y. Yang, Y. Zhang, Q. Zhao, C.L. Chen, B.K. Tye, S. Dang, and Y. Zhai. 2023. The human pre-replication complex is an open complex. *Cell*. 186:98-111 e121.

Lohka, M.J., M.K. Hayes, and J.L. Maller. 1988. Purification of maturation-promoting factor, an intracellular regulator of early mitotic events. *Proc Natl Acad Sci U S A*. 85:3009-3013.

Lohka, M.J., and Y. Masui. 1983. Formation in vitro of sperm pronuclei and mitotic chromosomes induced by amphibian ooplasmic components. *Science*. 220:719-721.

Looke, M., M.F. Maloney, and S.P. Bell. 2017. Mcm10 regulates DNA replication elongation by stimulating the CMG replicative helicase. *Genes Dev*. 31:291-305.

Lopez-Mosqueda, J., N.L. Maas, Z.O. Jonsson, L.G. Defazio-Eli, J. Wohlschlegel, and D.P. Toczyski. 2010. Damage-induced phosphorylation of Sld3 is important to block late origin firing. *Nature*. 467:479-483.

MacNeill, S.A. 2010. Structure and function of the GINS complex, a key component of the eukaryotic replisome. *Biochem J*. 425:489-500.

Makiniemi, M., T. Hillukkala, J. Tuusa, K. Reini, M. Vaara, D. Huang, H. Pospiech, I. Majuri, T. Westerling, T.P. Makela, and J.E. Syvaoja. 2001. BRCT domain-containing protein TopBP1 functions in DNA replication and damage response. *J Biol Chem*. 276:30399-30406.

Malumbres, M., and M. Barbacid. 2001. To cycle or not to cycle: a critical decision in cancer. *Nat Rev Cancer*. 1:222-231.

Mantiero, D., A. Mackenzie, A. Donaldson, and P. Zegerman. 2011. Limiting replication initiation factors execute the temporal programme of origin firing in budding yeast. *EMBO J*. 30:4805-4814.

Marchal, C., J. Sima, and D.M. Gilbert. 2019. Control of DNA replication timing in the 3D genome. *Nat Rev Mol Cell Biol*. 20:721-737.

Marinsek, N., E.R. Barry, K.S. Makarova, I. Dionne, E.V. Koonin, and S.D. Bell. 2006. GINS, a central nexus in the archaeal DNA replication fork. *EMBO Rep*. 7:539-545.

Marshall, C.J., and T.J. Santangelo. 2020. Archaeal DNA Repair Mechanisms. *Biomolecules*. 10.

Matsuno, K., M. Kumano, Y. Kubota, Y. Hashimoto, and H. Takisawa. 2006. The N-terminal noncatalytic region of Xenopus RecQ4 is required for chromatin binding of DNA polymerase alpha in the initiation of DNA replication. *Mol Cell Biol*. 26:4843-4852.

Maya-Mendoza, A., P. Olivares-Chauvet, A. Shaw, and D.A. Jackson. 2010. S phase progression in human cells is dictated by the genetic continuity of DNA foci. *PLoS Genet*. 6:e1000900.

Meagher, M., L.B. Epling, and E.J. Enemark. 2019. DNA translocation mechanism of the MCM complex and implications for replication initiation. *Nat Commun*. 10:3117.

Mordes, D.A., G.G. Glick, R. Zhao, and D. Cortez. 2008a. TopBP1 activates ATR through ATRIP and a PIKK regulatory domain. *Genes Dev*. 22:1478-1489.

Mordes, D.A., E.A. Nam, and D. Cortez. 2008b. Dpb11 activates the Mec1-Ddc2 complex. *Proc Natl Acad Sci U S A*. 105:18730-18734.

Muramatsu, S., K. Hirai, Y.S. Tak, Y. Kamimura, and H. Araki. 2010a. CDK-dependent complex formation between replication proteins Dpb11, Sld2, Pol (epsilon, and GINS in budding yeast. *Genes Dev*. 24:602-612.

Muramatsu, S., K. Hirai, Y.S. Tak, Y. Kamimura, and H. Araki. 2010b. CDK-dependent complex formation between replication proteins Dpb11, Sld2, Pol (epsilon}, and GINS in budding yeast. *Genes Dev*. 24:602-612.

Murray, A.W., and M.W. Kirschner. 1989. Cyclin synthesis drives the early embryonic cell cycle. *Nature*. 339:275-280.

Nakamura, H., T. Morita, and C. Sato. 1986. Structural organizations of replicon domains during DNA synthetic phase in the mammalian nucleus. *Exp Cell Res*. 165:291-297.

Natsume, T., C.A. Muller, Y. Katou, R. Retkute, M. Gierlinski, H. Araki, J.J. Blow, K. Shirahige, C.A. Nieduszynski, and T.U. Tanaka. 2013. Kinetochores coordinate pericentromeric cohesion and early DNA replication by Cdc7-Dbf4 kinase recruitment. *Mol Cell*. 50:661-674.

Newport, J., and M. Kirschner. 1982. A major developmental transition in early Xenopus embryos: II. Control of the onset of transcription. *Cell*. 30:687-696.

Nguyen, V.Q., C. Co, and J.J. Li. 2001. Cyclin-dependent kinases prevent DNA re-replication through multiple mechanisms. *Nature*. 411:1068-1073.

Noguchi, Y., Z. Yuan, L. Bai, S. Schneider, G. Zhao, B. Stillman, C. Speck, and H. Li. 2017. Cryo-EM structure of Mcm2-7 double hexamer on DNA suggests a lagging-strand DNA extrusion model. *Proc Natl Acad Sci U S A*. 114:E9529-E9538.

Parlak, M. 2017. A first step towards *in vitro* reconstitution of origin firing in vertebrates (Master Thesis, University Duisburg-Essen).

Pfander, B., and J.F. Diffley. 2011. Dpb11 coordinates Mec1 kinase activation with cell cycle-regulated Rad9 recruitment. *EMBO J*. 30:4897-4907.

Rahman, S., M.E. Sowa, M. Ottinger, J.A. Smith, Y. Shi, J.W. Harper, and P.M. Howley. 2011. The Brd4 extraterminal domain confers transcription activation independent of pTEFb by recruiting multiple proteins, including NSD3. *Mol Cell Biol*. 31:2641-2652.

Rappsilber, J., M. Mann, and Y. Ishihama. 2007. Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. *Nat Protoc*. 2:1896-1906.

Raschle, M., G. Smeenk, R.K. Hansen, T. Temu, Y. Oka, M.Y. Hein, N. Nagaraj, D.T. Long, J.C. Walter, K. Hofmann, Z. Storchova, J. Cox, S. Bekker-Jensen, N. Mailand, and M. Mann. 2015. Proteomics reveals dynamic assembly of repair complexes during bypass of DNA cross-links. *Science*. 348:1253671.

Remus, D., F. Beuron, G. Tolun, J.D. Griffith, E.P. Morris, and J.F. Diffley. 2009. Concerted loading of Mcm2-7 double hexamers around DNA during DNA replication origin licensing. *Cell*. 139:719-730.

Reusswig, K.U., F. Zimmermann, L. Galanti, and B. Pfander. 2016. 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.

Rhind, N. 2006. DNA replication timing: random thoughts about origin firing. *Nat Cell Biol*. 8:1313-1316.

Rzechorzek, N.J., S.W. Hardwick, V.A. Jatikusumo, D.Y. Chirgadze, and L. Pellegrini. 2020. CryoEM structures of human CMG-ATPgammaS-DNA and CMG-AND-1 complexes. *Nucleic Acids Res*. 48:6980-6995.

Sadoni, N., M.C. Cardoso, E.H. Stelzer, H. Leonhardt, and D. Zink. 2004. Stable chromosomal units determine the spatial and temporal organization of DNA replication. *J Cell Sci*. 117:5353-5365.

Sakasai, R., A. Sakai, M. Iimori, S. Kiyonari, K. Matsuoka, Y. Kakeji, H. Kitao, and Y. Maehara. 2012. CtIP- and ATR-dependent FANCJ phosphorylation in response to DNA strand breaks mediated by DNA replication. *Genes Cells*. 17:962-970.

Saleh, A., Y. Noguchi, R. Aramayo, M.E. Ivanova, K.M. Stevens, A. Montoya, S. Sunidhi, N.L. Carranza, M.J. Skwark, and C. Speck. 2022. The structural basis of Cdc7-Dbf4 kinase dependent targeting and phosphorylation of the MCM2-7 double hexamer. *Nat Commun*. 13:2915.

Sangrithi, M.N., J.A. Bernal, M. Madine, A. Philpott, J. Lee, W.G. Dunphy, and A.R. Venkitaraman. 2005. Initiation of DNA replication requires the RECQL4 protein mutated in Rothmund-Thomson syndrome. *Cell*. 121:887-898.

Sansam, C.G., D. Goins, J.C. Siefert, E.A. Clowdus, and C.L. Sansam. 2015. Cyclin-dependent kinase regulates the length of S phase through TICRR/TRESLIN phosphorylation. *Genes Dev*. 29:555-566.

Sansam, C.G., K. Pietrzak, B. Majchrzycka, M.A. Kerlin, J. Chen, S. Rankin, and C.L. Sansam. 2018. A mechanism for epigenetic control of DNA replication. *Genes Dev*. 32:224-229.

Santocanale, C., and J.F. Diffley. 1998. A Mec1- and Rad53-dependent checkpoint controls late-firing origins of DNA replication. *Nature*. 395:615-618.

Sclafani, R.A., and T.M. Holzen. 2007. Cell cycle regulation of DNA replication. *Annu Rev Genet*. 41:237-280.

Siddiqui, K., K.F. On, and J.F. Diffley. 2013. Regulating DNA replication in eukarya. *Cold Spring Harb Perspect Biol*. 5.

Sokka, M., S. Parkkinen, H. Pospiech, and J.E. Syvaoja. 2010. Function of TopBP1 in genome stability. *Subcell Biochem*. 50:119-141.

Sporbert, A., A. Gahl, R. Ankerhold, H. Leonhardt, and M.C. Cardoso. 2002. DNA polymerase clamp shows little turnover at established replication sites but sequential de novo assembly at adjacent origin clusters. *Mol Cell*. 10:1355-1365.

Sun, L., Y. Huang, R.A. Edwards, S. Yang, A.N. Blackford, W. Niedzwiedz, and J.N.M. Glover. 2017. Structural Insight into BLM Recognition by TopBP1. *Structure*. 25:1582-1588 e1583.

Tada, S., A. Li, D. Maiorano, M. Mechali, and J.J. Blow. 2001. Repression of origin assembly in metaphase depends on inhibition of RLF-B/Cdt1 by geminin. *Nat Cell Biol*. 3:107-113.

Tak, Y.S., Y. Tanaka, S. Endo, Y. Kamimura, and H. Araki. 2006. A CDK-catalysed regulatory phosphorylation for formation of the DNA replication complex Sld2-Dpb11. *EMBO J*. 25:1987-1996.

Takayama, Y., Y. Kamimura, M. Okawa, S. Muramatsu, A. Sugino, and H. Araki. 2003. GINS, a novel multiprotein complex required for chromosomal DNA replication in budding yeast. *Genes Dev*. 17:1153-1165.

Tanaka, S. 2021. Interaction of replication factor Sld3 and histone acetyl transferase Esa1 alleviates gene silencing and promotes the activation of late and dormant replication origins. *Genetics*. 217:1-11.

Tanaka, S., and H. Araki. 2010. Regulation of the initiation step of DNA replication by cyclin-dependent kinases. *Chromosoma*. 119:565-574.

Tanaka, S., and H. Araki. 2011. Multiple regulatory mechanisms to inhibit untimely initiation of DNA replication are important for stable genome maintenance. *PLoS Genet*. 7:e1002136.

Tanaka, S., and H. Araki. 2013. Helicase activation and establishment of replication forks at chromosomal origins of replication. *Cold Spring Harb Perspect Biol*. 5:a010371.

Tanaka, S., Y. Komeda, T. Umemori, Y. Kubota, H. Takisawa, and H. Araki. 2013. Efficient initiation of DNA replication in eukaryotes requires Dpb11/TopBP1-GINS interaction. *Mol Cell Biol*. 33:2614-2622.

Tanaka, S., R. Nakato, Y. Katou, K. Shirahige, and H. Araki. 2011a. Origin association of Sld3, Sld7, and Cdc45 proteins is a key step for determination of origin-firing timing. *Curr Biol*. 21:2055-2063.

Tanaka, S., T. Umemori, K. Hirai, S. Muramatsu, Y. Kamimura, and H. Araki. 2007. CDK-dependent phosphorylation of Sld2 and Sld3 initiates DNA replication in budding yeast. *Nature*. 445:328-332.

Tanaka, T., T. Umemori, S. Endo, S. Muramatsu, M. Kanemaki, Y. Kamimura, C. Obuse, and H. Araki. 2011b. Sld7, an Sld3-associated protein required for efficient chromosomal DNA replication in budding yeast. *EMBO J*. 30:2019-2030.

Tanaka, T.U. 2002. Bi-orienting chromosomes on the mitotic spindle. *Curr Opin Cell Biol*. 14:365-371.

Ticau, S., L.J. Friedman, N.A. Ivica, J. Gelles, and S.P. Bell. 2015. Single-molecule studies of origin licensing reveal mechanisms ensuring bidirectional helicase loading. *Cell*. 161:513-525.

Traven, A., and J. Heierhorst. 2005. SQ/TQ cluster domains: concentrated ATM/ATR kinase phosphorylation site regions in DNA-damage-response proteins. *Bioessays*. 27:397-407.

Tunquist, B.J., and J.L. Maller. 2003. Under arrest: cytostatic factor (CSF)-mediated metaphase arrest in vertebrate eggs. *Genes Dev*. 17:683-710.

Tyanova, S., T. Temu, P. Sinitcyn, A. Carlson, M.Y. Hein, T. Geiger, M. Mann, and J. Cox. 2016. The Perseus computational platform for comprehensive analysis of (prote)omics data. *Nat Methods*. 13:731-740.

Vizcaino, J.A., A. Csordas, N. del-Toro, J.A. Dianes, J. Griss, I. Lavidas, G. Mayer, Y. Perez-Riverol, F. Reisinger, T. Ternent, Q.W. Xu, R. Wang, and H. Hermjakob. 2016. 2016 update of the PRIDE database and its related tools. *Nucleic Acids Res*. 44:D447-456.

Volpi, I., P.J. Gillespie, G.S. Chadha, and J.J. Blow. 2021. The role of DDK and Treslin-MTBP in coordinating replication licensing and pre-initiation complex formation. *Open Biol*. 11:210121.

Wang, J., Z. Gong, and J. Chen. 2011. MDC1 collaborates with TopBP1 in DNA replication checkpoint control. *J Cell Biol*. 193:267-273.

Wohlschlegel, J.A., B.T. Dwyer, S.K. Dhar, C. Cvetic, J.C. Walter, and A. Dutta. 2000. Inhibition of eukaryotic DNA replication by geminin binding to Cdt1. *Science*. 290:2309-2312.

Wong, P.G., S.L. Winter, E. Zaika, T.V. Cao, U. Oguz, J.M. Koomen, J.L. Hamlin, and M.G. Alexandrow. 2011. Cdc45 limits replicon usage from a low density of preRCs in mammalian cells. *PLoS One*. 6:e17533.

Woodward, A.M., T. Gohler, M.G. Luciani, M. Oehlmann, X. Ge, A. Gartner, D.A. Jackson, and J.J. Blow. 2006. Excess Mcm2-7 license dormant origins of replication that can be used under conditions of replicative stress. *J Cell Biol*. 173:673-683.

Xu, Y., T. Gristwood, B. Hodgson, J.C. Trinidad, S.V. Albers, and S.D. Bell. 2016. Archaeal orthologs of Cdc45 and GINS form a stable complex that stimulates the helicase activity of MCM. *Proc Natl Acad Sci U S A*. 113:13390-13395.

Yamane, K., M. Kawabata, and T. Tsuruo. 1997. A DNA-topoisomerase-II-binding protein with eight repeating regions similar to DNA-repair enzymes and to a cell-cycle regulator. *Eur J Biochem*. 250:794-799.

Yan, S., and W.M. Michael. 2009. TopBP1 and DNA polymerase-alpha directly recruit the 9-1-1 complex to stalled DNA replication forks. *J Cell Biol*. 184:793-804.

Yeeles, J.T., T.D. Deegan, A. Janska, A. Early, and J.F. Diffley. 2015. Regulated eukaryotic DNA replication origin firing with purified proteins. *Nature*. 519:431-435.

Yeeles, J.T., A. Janska, A. Early, and J.F. Diffley. 2017. How the Eukaryotic Replisome Achieves Rapid and Efficient DNA Replication. *Mol Cell*. 65:105-116.

Yekezare, M., B. Gomez-Gonzalez, and J.F. Diffley. 2013. Controlling DNA replication origins in response to DNA damage - inhibit globally, activate locally. *J Cell Sci*. 126:1297-1306.

Yoo, H.Y., A. Kumagai, A. Shevchenko, A. Shevchenko, and W.G. Dunphy. 2007. Ataxia-telangiectasia mutated (ATM)-dependent activation of ATR occurs through phosphorylation of TopBP1 by ATM. *J Biol Chem*. 282:17501-17506.

Yuan, Z., L. Bai, J. Sun, R. Georgescu, J. Liu, M.E. O'Donnell, and H. Li. 2016. Structure of the eukaryotic replicative CMG helicase suggests a pumpjack motion for translocation. *Nat Struct Mol Biol*. 23:217-224.

Zegerman, P., and J.F. Diffley. 2007. Phosphorylation of Sld2 and Sld3 by cyclin-dependent kinases promotes DNA replication in budding yeast. *Nature*. 445:281-285.

Zegerman, P., and J.F. Diffley. 2010. Checkpoint-dependent inhibition of DNA replication initiation by Sld3 and Dbf4 phosphorylation. *Nature*. 467:474-478.

# List of Figures

**Appendix Figures**

# List of Tables



**Appendix Table**

# Appendix

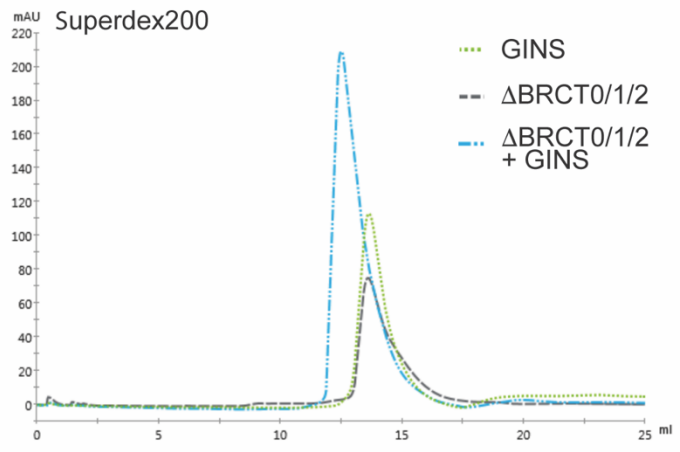
## 6.1 Appendix Figures

Figures shown in the appendix are part of the article published in bioRxiv:

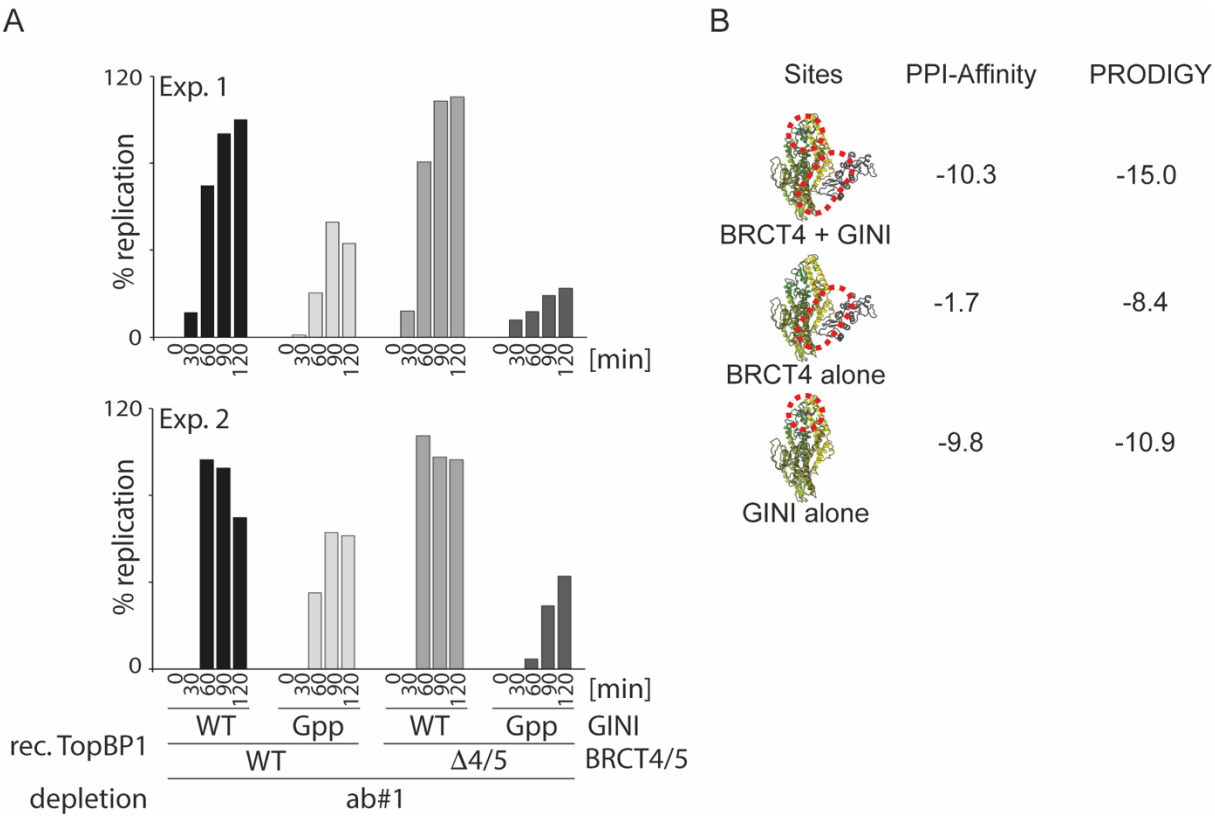
Day M, Tetik B, Parlak M, Almeida-Hernandez Y, Sanchez-Garcia E, Kaschani F, Räschle M, Siegert H, Pearl L, Oliver A, Boos D

*Bipartite GINS binding mode of TopBP1 to activate the replicative helicase.*

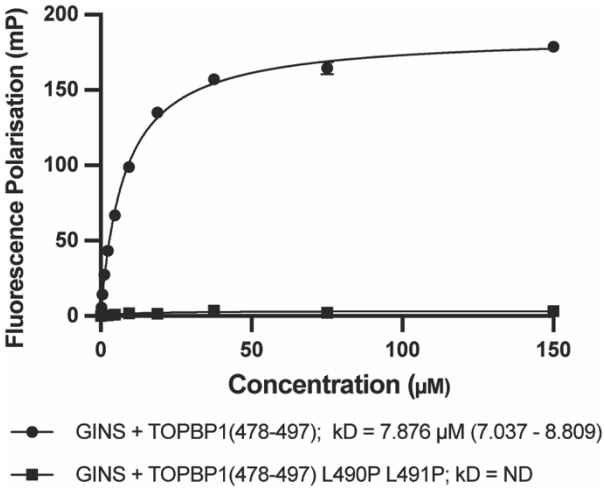
bioRxiv (2023) doi: https://doi.org/10.1101/2023.03.31.535063



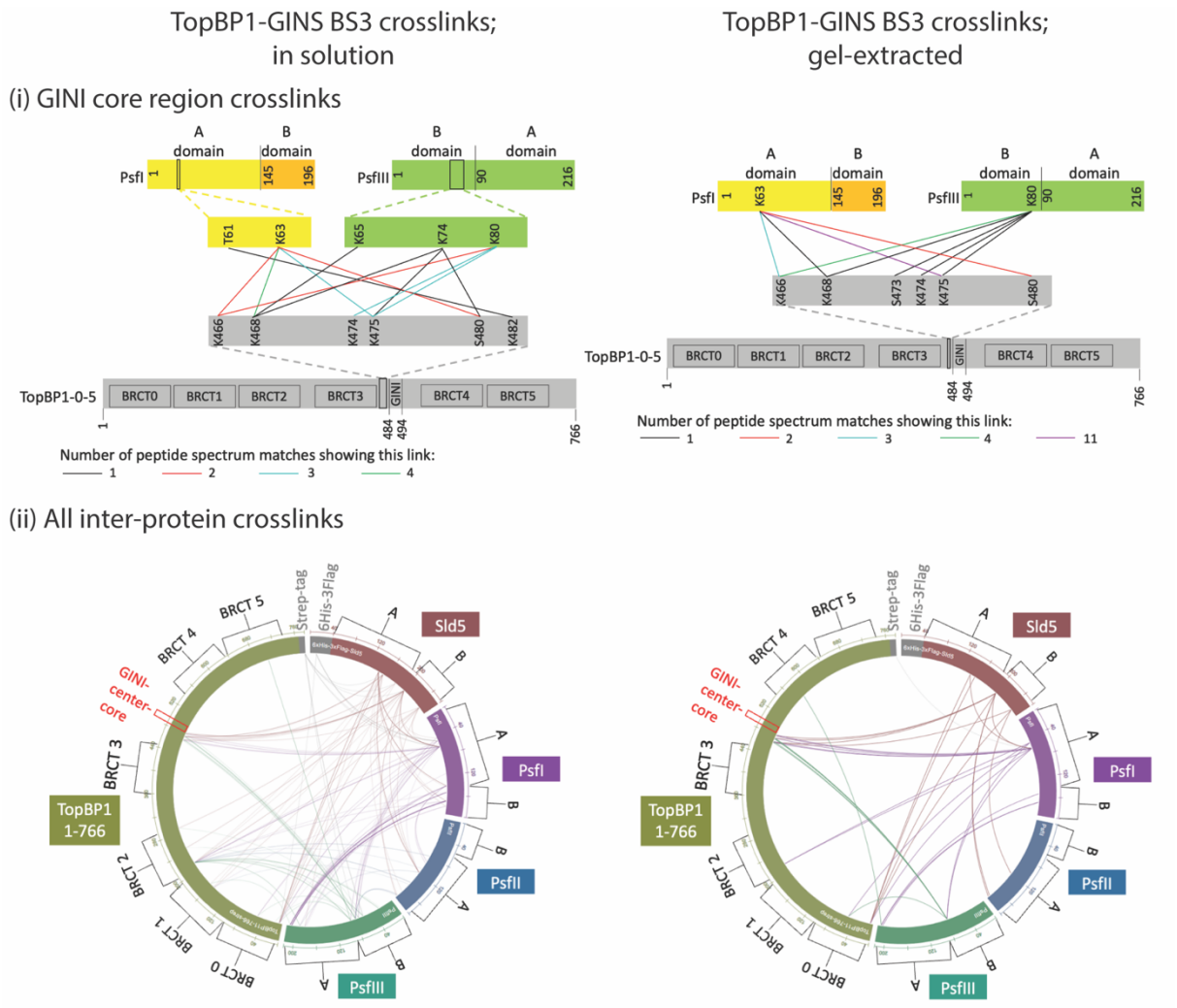

**Appendix Figure 1: Characterization of the GINS-TopBP1 interaction.**
Recombinant TopBP1-ΔBRCT0-2 was incubated with recombinant GINS before size-exclusion chromatography using a Superdex200 increase 10 300 column. The individual recombinant proteins were run separately for comparison. Elution profiles are shown with GINS colored green, TopBP1 grey and the complex is colored cyan. The clear shift upon complex formation suggests the N-terminal BRCT0-2 module of TopBP1 is dispensable for the interaction.



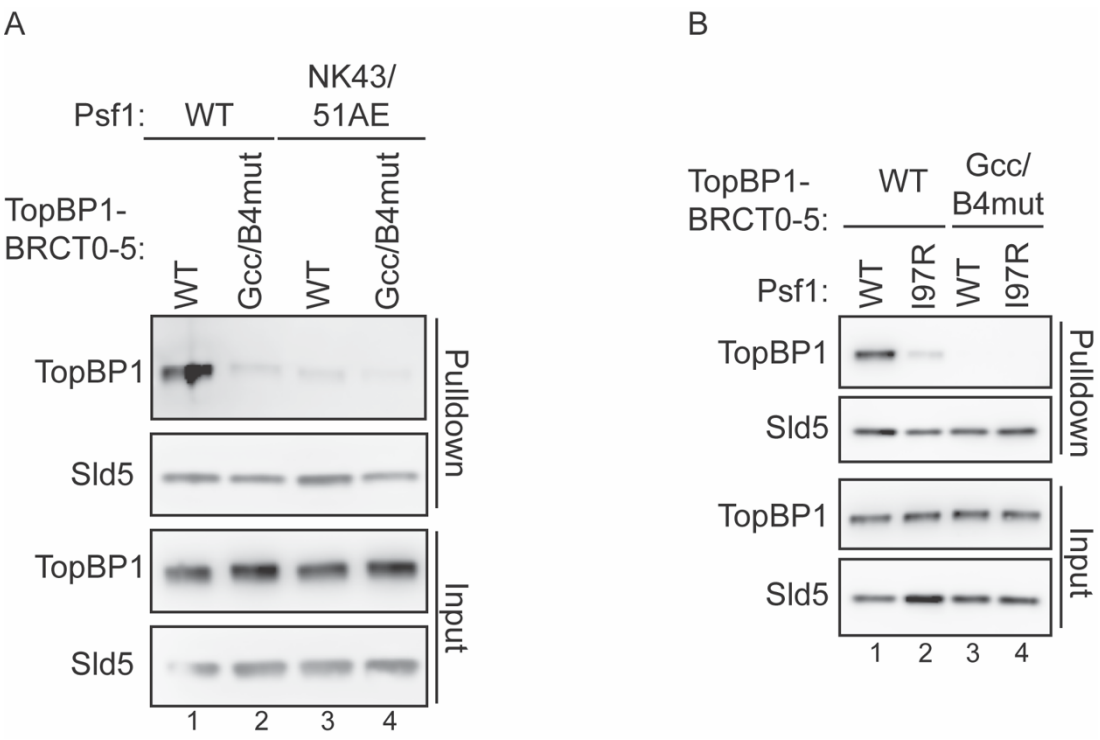

**Appendix Figure 2: Both GINS binding surfaces in TopBP1 cooperate in supporting DNA replication.**
A) Replication analysis of interphase egg extract immunodepleted with antibody #1 upon adding back recombinant TopBP1 BRCT0-5 wild type (WT), or indicated TopBP1 mutants. B) PPI-Affinity and PRODIGY were used to estimate the free energy (in kcal/mol) of the interactions involving TopBP1 GINI and BRCT4.



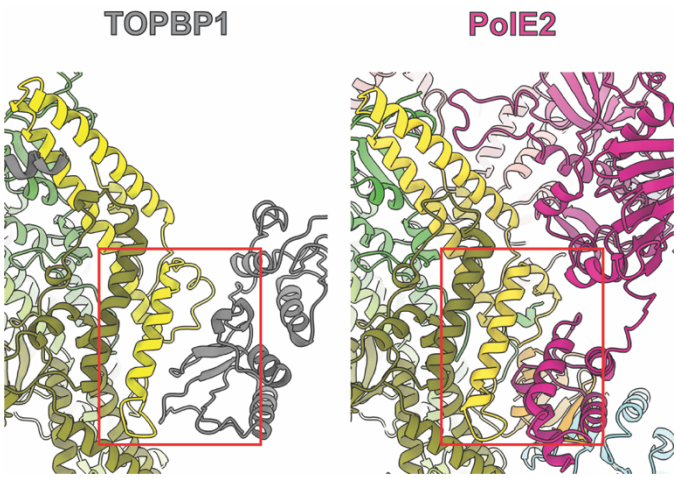

**Appendix Figure 3: Fluorescence polarization experiments using recombinant GINS tetramer and fluorescently labelled TopBP1-GINI-core.**
Dots show wild type TopBP1-GINI peptide interaction with GINS. Squares show helix break mutant TopBP1 peptide interaction with GINS.




**Appendix Figure 4: Crosslinking mass spectrometry to analyze the GINS-TopBP1 complex.** Schematic representation of BS3-mediated cross-links between TopBP1-BRCT0-5-Strep and GINS (detailed description below). (i) shows crosslinks involving the GINI domain of TopBP1, (ii) shows circle plots of all inter-protein crosslinks detected. Recombinant TopBP1-BRCT0-5-Strep (amino acids 1-766) and GINS were crosslinked with 600µM or 2500µM of BS3 crosslinker and analyzed by mass spectrometry. Crosslinked samples were either analyzed directly (in solution, left) or after resolving cross links by SDS-PAGE and in-gel digestion (right). ProXL (Protein Cross-Linking Database) was used to analyze and visualize the data. Cross-links in (i) are colored according to the number of peptide spectrum matches showing this link.




**Appendix Figure 5: Pulldown of TopBP1 using immobilized GINS.**
A) Pulldown of the indicated recombinant TopBP1-BRCT0-5-strep versions using immobilized GINS-WT or GINS carrying the indicated Psf1 mutations. Analysis was done by immunoblotting. B) Pulldown of the indicated recombinant TopBP1-BRCT0-5-strep versions using immobilized GINS-WT or GINS carrying a Ile97 to arginine mutation in Psf1.




**Appendix Figure 6: Exclusive binding of TopBP1 and DNA polymerase ε to GINS in the CMG context.**
Closeup of interaction sites on GINS for TopBP1 BRCT4 and PolE2.

## 6.2 Appendix Tables

**Appendix Table 1: Pathway enrichment of common proteins**

| GO-term | Ben. Ho. FDR |
|---|---|
| nucleoplasm | 3,92E-26 |
| AAAS; ACIN1; ADAR; AFF4; AKAP17A; AKAP8; AKAP8L; ANLN; ANXA11; AP3D1; APRT; ARID1A; ASF1B; ATAD2; ATR; BCLAF1; BLM; BPTF; BRCA1; C9orf78; CASC5; CBX3; CDC27; CDC5L; CDC73; CDCA2; CDK11B; CDK11A; CDK12; CENPF; CHD4; COASY; COIL; CPSF1; CPSF6; CSTF2; DDB1; DDX17; DDX21; DDX42; DDX5; DHX15; DHX16; DHX38; DHX8; DHX9; DKC1; DNMT1; EFTUD2; EIF4A3; ELAC2; EPB41L2; ERCC3; ESF1; EWSR1; FANCI; FOXC1; FOXK1; FUBP1; FUBP1; FUS; GATAD2B; GEMIN5; GINS3; GNL3; GTF2F2; GTF2H1; GTF2I; GTF3C1; GTF3C3; H2AFX; HIST1H2AA; HIST2H2AB; HAT1; HCFC1; HEXIM1; HIST1H2BN; HIST1H2BL; HIST1H2BM; HIST1H2BH; HIST2H2BF; HIST1H2BC; HIST1H2BD; H2BFS; HIST1H2BK; HIST1H2BA; HIST1H4A; HLTF; HNRNPA1; HNRNPA2B1; HNRPA2B1; HNRNPA3; HNRNPAB; HNRNPC; HNRNPD; HNRNPDL; HNRNPF; HNRNPH3; HNRNPK; HNRNPL; HNRNPR; HSP90AA1; HSPA8; HTATSF1; HUWE1; IK; ILF2; ILF3; INTS9; IPO5; IWS1; KDM3A; KDM3B; KHDRBS1; KHSRP; KIF22; KIF4A; KMT2D; KPNA2; KPNB1; LAS1L; LEO1; LIG3; LUC7L3; MASTL; MATR3; MCMBP; MDC1; MED1; MED12; TNRC11; MIF; MORC2; MORC3; MRE11A; MSH2; MSH6; MTA2; NACC1; NASP; NBN; NCAPD3; NCL; NELFA; NIFK; NIPBL; NOC2L; NOLC1; NOP56; NUMA1; NUP205; OGT; PABPN1; PAPOLA; PARP1; PDCD4; PDS5A; PELP1; PHB; PHF6; PHF8; POLA2; POLD1; POLE; POLR1A; POLR1B; POLR1E; POLR2A; POLR2B; POLR2E; POLR2L; POLR3A; PPM1G; PPP1R10; PQBP1; PRC1; PRCC; PRKDC; PRPF40A; PRPF4B; PRPF8; PSIP1; PSMC5; PSME3; PTBP1; RACGAP1; RAD50; RAD9A; RAI1; RANBP3; RBBP6; RBM10; RBM12; RBM22; RBM25; RBM4; RBMX; RFC1; RNPS1; RPA1; RPA2; RPS27A; UBC; UBB; UBA52; RRN3; SAFB; SART3; SCAF11; SCAF4; SET; SETD1A; SETD2; SETX; SF1; SF3A1; SF3B1; SF3B2; SF3B3; SF3B4; SIN3A; SKIV2L2; SLFN11; SLU7; SMARCA4; SMARCA5; SMARCAD1; SMARCC1; SMC1A; SMC3; SMNDC1; SNRNP200; SNRNP70; SNRPE; SNW1; SPEN; SREK1; SRI; SRRM2; SRRT; SRSF1; SRSF10; SRSF2; SRSF6; SRSF9; SSRP1; STAG2; SUGP2; SUMO1; SUMO2; SUPT16H; SUPT5H; SYMPK; TAF15; TARDBP; TDP43; TBCB; TCEB3; TCF20; THOC2; THRAP3; TLE1; TOP1; TOP2A; TOP2B; TP53; TP53BP1; TPR; TPX2; TRIM27; TRIM28; TWISTNB; TXN; U2AF1; U2AF2; U2SURP; UBR5; UBTF; USP48; UTP20; WAPAL; WBP11; WDR75; WRN; XPO1; XPO5; XRCC1; YLPM1; ZFR; ZNF318; ZNF638 | |
| RNA processing | 1,82E-20 |
| ACIN1; ADAR; AKAP17A; ALKBH5; CDC5L; CDC73; CDK12; CIRH1A; CPSF1; CPSF6; CSTF2; DDX17; DDX21; DDX47; DDX5; DDX52; DDX54; DHX15; DHX16; DHX37; DHX38; DHX8; DHX9; DKC1; EFTUD2; EIF4A3; ELAC2; ERCC3; FTSJ3; FUS; GEMIN5; GTF2F2; GTF2H1; HEATR1; HNRNPA1; HNRNPA2B1; HNRPA2B1; HNRNPA3; HNRNPC; HNRNPD; HNRNPDL; HNRNPF; HNRNPH3; HNRNPK; HNRNPL; HNRNPR; HSPA8; INTS9; IWS1; KHDRBS1; KHSRP; KIAA1429; LAS1L; LEO1; LUC7L3; MPHOSPH10; NAF1; NOL11; NOL8; NOL9; NOLC1; NOP2; NOP56; NOP58; NSUN2; PABPN1; PAF1; PAPD5; PAPOLA; PARN; PDCD11; POLR2A; POLR2B; POLR2E; POLR2L; POP1; PPP2R1A; PQBP1; PRPF38B; PRPF40A; PRPF4B; PRPF8; PTBP1; PUS7; PWP2; RBM10; RBM22; RBM25; RBM3; RBM4; RBMX; RBMXL1; RNPS1; RPL7; RRP1; RRP1B; RTCB; SART3; SCAF11; SETX; SF1; SF3A1; SF3B1; SF3B2; SF3B3; SF3B4; SKIV2L2; SLU7; SMC1A; SMNDC1; SNRNP200; SNRNP70; SNRPE; SNW1; SREK1; SRRM2; SRRT; SRSF1; SRSF10; SRSF2; SRSF6; SRSF9; SUGP2; SUPT5H; SUPT6H; SYMPK; TARDBP; TDP43; TBL3; THOC2; THRAP3; TRA2A; TRA2B; TRMT1L; U2AF1; U2AF2; U2SURP; UTP14A; UTP15; UTP20; UTP3; WBP11; WDR3; WDR36; XRN2; ZNF638 | |
| nucleolus | 7,34E-19 |
| ACIN1; ADAR; AFF4; BAZ2A; BCLAF1; BLM; BMS1; CDC5L; CDK12; CIRH1A; COIL; DDX17; DDX18; DDX21; DDX24; DDX27; DDX42; DDX47; DDX5; DDX50; DDX52; DDX54; DHX15; DHX37; | |

DHX9; DKC1; EIF4A3; ESF1; EWSR1; EZR; FOXK1; FTSJ3; FXR1; GNL3; GTF3C1; GTPBP4; H1FX; HEATR1; HLTF; HSPA8; IK; ILF2; ILF3; IPO5; LAS1L; LEO1; LYAR; MDN1; MED1; MED12; TNRC11; MKI67; MPHOSPH10; MYBBP1A; NBN; NCL; NIFK; NOC2L; NOL11; NOL8; NOL9; NOLC1; NOP2; NOP56; NOP58; NSUN2; NVL; PAPD5; PARN; PARP1; PDCD11; PELP1; PHF6; PHF8; POLR1A; POLR1B; POLR1E; PPAN-P2RY11; PPAN; PPP1CC; PRKDC; PTBP1; PWP1; PWP2; RBBP6; RBM4; RFC1; RPL19; RPL4; RPL7; RPS27A; UBC; UBB; UBA52; RRN3; RRP1; RRP12; RRP1B; RRS1; RSL1D1; RTF1; SCAF11; SENP3; SENP3-EIF4A1; SETX; SIN3A; SKIV2L2; SMARCA4; SMARCA5; SMC1A; SRSF9; SSRP1; STAG2; SUMO1; TBL3; TCERG1; TCOF1; TCOF1; TEX10; TLE1; TOP1; TOP2A; TP53; TRA2A; TSR1; TWISTNB; UBTF; URB1; URB2; USP36; UTP14A; UTP15; UTP20; UTP3; WDR3; WDR36; WDR43; WDR75; WRN; XPO1; XRN2; ZNF207

| RNA metabolic process | 4,13E-18 |
|---|---|

AAAS; ACIN1; ADAR; AFF4; AKAP17A; ALKBH5; ARID1A; ASF1B; ATAD2; BAZ1A; BAZ1B; BAZ2A; BCLAF1; BPTF; BRD2; CBX3; CCNK; CDC5L; CDC73; CDK12; CHD1; CHD4; CIRH1A; CPSF1; CPSF6; CSTF2; CTR9; DDX17; DDX18; DDX21; DDX24; DDX27; DDX42; DDX47; DDX5; DDX50; DDX52; DDX54; DHX15; DHX16; DHX37; DHX38; DHX8; DHX9; DIDO1; DKC1; DNMT1; EFTUD2; EIF4A3; ELAC2; ERCC3; ESF1; ETF1; EWSR1; FARSA; FLNA; FTSJ3; FUBP1; FUS; GATAD2B; GEMIN5; GTF2F2; GTF2H1; GTF2I; GTF3C1; GTF3C3; HEATR1; HELLS; HEXIM1; HLTF; HNRNPA1; HNRNPA2B1; HNRPA2B1; HNRNPA3; HNRNPC; HNRNPD; HNRNPDL; HNRNPF; HNRNPH3; HNRNPK; HNRNPL; HNRNPR; HSPA8; HTATSF1; ICE1; ILF2; ILF3; INTS9; IWS1; KDM3A; KDM3B; KHDRBS1; KHSRP; KIAA1429; KMT2D; LAS1L; LEO1; LUC7L3; MCTS1; MED1; MED12; TNRC11; MEPCE; MPHOSPH10; MYBBP1A; NAF1; NELFA; NIFK; NOC2L; NOL11; NOL8; NOL9; NOLC1; NOP2; NOP56; NOP58; NSUN2; NUP133; NUP155; NUP160; NUP205; NUP210; PABPN1; PAF1; PAPD5; PAPOLA; PARN; PARP1; PAXBP1; PDCD11; PELP1; PHF6; PHF8; POLR1A; POLR1B; POLR1E; POLR2A; POLR2B; POLR2E; POLR2L; POLR3A; POP1; PPP1R10; PPP2R1A; PQBP1; PRPF38B; PRPF40A; PRPF4B; PRPF8; PSIP1; PSMC5; PTBP1; PUS7; PWP1; PWP2; RBM10; RBM22; RBM25; RBM3; RBM4; RBMX; RBMXL1; RFC1; RNPS1; RPL10; RPL19; RPL28; RPL32; RPL4; RPL7; RPS27A; UBC; UBB; UBA52; RRN3; RRP1; RRP1B; RTCB; RTF1; SAFB; SART3; SCAF11; SETD1A; SETD2; SETX; SF1; SF3A1; SF3B1; SF3B2; SF3B3; SF3B4; SKIV2L2; SLU7; SMARCA5; SMARCC1; SMC1A; SMNDC1; SNRNP200; SNRNP70; SNRPE; SNW1; SREK1; SRRM2; SRRT; SRSF1; SRSF10; SRSF2; SRSF6; SRSF9; SSRP1; SUGP2; SUPT16H; SUPT5H; SUPT6H; SYMPK; TARDBP; TDP43; TBL3; TCEB3; TCERG1; TCF20; TCOF1; THOC2; THRAP3; TP53; TP53BP1; TPR; TRA2A; TRA2B; TRIM27; TRIM28; TRMT1L; TWISTNB; TXN; U2AF1; U2AF2; U2SURP; UBTF; UHRF1; UTP14A; UTP15; UTP20; UTP3; VARS; WBP11; WDR3; WDR36; XRN2; YWHAB; ZMYM2; ZNF318; ZNF638

| nitrogen compound metabolic process | 1,80E-17 |
|---|---|

AAAS; ACIN1; ADAR; AFF4; AKAP17A; AKAP8; ALKBH5; APRT; ARID1A; ASF1B; ATAD2; ATP5J2; ATP5J2-PTCD1; ATP5L; ATR; BAZ1A; BAZ1B; BAZ2A; BCLAF1; BLM; BPTF; BRCA1; BRD2; BRIP1; CBX3; CCNK; CDC5L; CDC73; CDK12; CENPF; CHAF1A; CHD1; CHD4; CIRH1A; COASY; CPSF1; CPSF6; CSTF2; CTR9; DDB1; DDX17; DDX18; DDX21; DDX24; DDX27; DDX42; DDX47; DDX5; DDX50; DDX52; DDX54; DHX15; DHX16; DHX37; DHX38; DHX8; DHX9; DIDO1; DKC1; DNMT1; EFTUD2; EIF4A3; EIF5A; EIF5AL1; ELAC2; ERCC3; ESF1; ETF1; EWSR1; FANCI; FARSA; FLNA; FOXC1; FTSJ3; FUBP1; FUS; GATAD2B; GEMIN5; GINS3; GNAS; GTF2F2; GTF2H1; GTF2I; GTF3C1; GTF3C3; H2AFX; HIST1H2AA; HIST2H2AB; HAT1; HEATR1; HELLS; HEXIM1; HIST1H4A; HLTF; HNRNPA1; HNRNPA2B1; HNRPA2B1; HNRNPA3; HNRNPC; HNRNPD; HNRNPDL; HNRNPF; HNRNPH3; HNRNPK; HNRNPL; HNRNPR; HSP90AA1; HSPA8; HTATSF1; HUWE1; ICE1; ILF2; ILF3; INTS9; IWS1; KDM3A; KDM3B; KHDRBS1; KHSRP; KIAA1429; KIF22; KMT2D; KPNA2; KPNB1; LAS1L; LEO1; LIG3; LUC7L3; MCMBP; MCTS1; MDC1; MED1; MED12; TNRC11; MEPCE; MPHOSPH10; MRE11A; MSH2; MSH6; MTA2; MYBBP1A; NAF1; NASP; NBN; NCAPD3; NELFA; NIFK; NOC2L; NOL11; NOL8; NOL9; NOLC1; NOP2; NOP56; NOP58; NSUN2; NUP133; NUP155; NUP160; NUP205; NUP210; PABPN1; PAF1; PAPD5;

PAPOLA; PARN; PARP1; PAXBP1; PDCD11; PELP1; PHB; PHF6; PHF8; POLA2; POLD1; POLE; POLR1A; POLR1B; POLR1E; POLR2A; POLR2B; POLR2E; POLR2L; POLR3A; POP1; PPP1R10; PPP2R1A; PQBP1; PRKDC; PRPF38B; PRPF40A; PRPF4B; PRPF8; PRPS1; PSIP1; PSMC5; PSME3; PTBP1; PUS7; PWP1; PWP2; RAD50; RAD9A; RBM10; RBM22; RBM25; RBM3; RBM4; RBMX; RBMXL1; RFC1; RNPS1; RPA1; RPA2; RPL10; RPL19; RPL28; RPL32; RPL4; RPL7; RPS27A; UBC; UBB; UBA52; RRN3; RRP1; RRP1B; RTCB; RTF1; SAFB; SART3; SCAF11; SET; SETD1A; SETD2; SETX; SF1; SF3A1; SF3B1; SF3B2; SF3B3; SF3B4; SHMT2; SIN3A; SKIV2L2; SLU7; SMARCA4; SMARCA5; SMARCAD1; SMARCC1; SMC1A; SMC3; SMNDC1; SNRNP200; SNRNP70; SNRPE; SNW1; SREK1; SRRM2; SRRT; SRSF1; SRSF10; SRSF2; SRSF6; SRSF9; SSRP1; STOML2; SUGP2; SUMO1; SUPT16H; SUPT5H; SUPT6H; SYMPK; TARDBP; TDP43; TBL3; TCEB3; TCERG1; TCF20; TCOF1; THOC2; THRAP3; TK1; TOP1; TOP2A; TOP2B; TOPBP1; TP53; TP53BP1; TPR; TRA2A; TRA2B; TRIM27; TRIM28; TRMT1L; TWISTNB; TXN; U2AF1; U2AF2; U2SURP; UBR5; UBTF; UHRF1; USP7; UTP14A; UTP15; UTP20; UTP3; VARS; WBP11; WDR3; WDR36; WRN; WRNIP1; XRCC1; XRN2; YWHAB; ZMYM2; ZNF318; ZNF638

| cellular nitrogen compound metabolic process | 1,92E-17 |
|---|---|

AAAS; ACIN1; ADAR; AFF4; AKAP17A; AKAP8; ALKBH5; APRT; ARID1A; ASF1B; ATAD2; ATP5J2; ATP5J2-PTCD1; ATP5L; ATR; BAZ1A; BAZ1B; BAZ2A; BCLAF1; BLM; BPTF; BRCA1; BRD2; BRIP1; CBX3; CCNK; CDC5L; CDC73; CDK12; CENPF; CHAF1A; CHD1; CHD4; CIRH1A; COASY; CPSF1; CPSF6; CSTF2; CTR9; DDB1; DDX17; DDX18; DDX21; DDX24; DDX27; DDX42; DDX47; DDX5; DDX50; DDX52; DDX54; DHX15; DHX16; DHX37; DHX38; DHX8; DHX9; DIDO1; DKC1; DNMT1; EFTUD2; EIF4A3; EIF5A; EIF5AL1; ELAC2; ERCC3; ESF1; ETF1; EWSR1; FANCI; FARSA; FLNA; FTSJ3; FUBP1; FUS; GATAD2B; GEMIN5; GINS3; GNAS; GTF2F2; GTF2H1; GTF2I; GTF3C1; GTF3C3; H2AFX; HIST1H2AA; HIST2H2AB; HAT1; HEATR1; HELLS; HEXIM1; HIST1H4A; HLTF; HNRNPA1; HNRNPA2B1; HNRPA2B1; HNRNPA3; HNRNPC; HNRNPD; HNRNPDL; HNRNPF; HNRNPH3; HNRNPK; HNRNPL; HNRNPR; HSPA8; HTATSF1; HUWE1; ICE1; ILF2; ILF3; INTS9; IWS1; KDM3A; KDM3B; KHDRBS1; KHSRP; KIAA1429; KIF22; KMT2D; KPNA2; KPNB1; LAS1L; LEO1; LIG3; LUC7L3; MCMBP; MCTS1; MDC1; MED1; MED12; TNRC11; MEPCE; MPHOSPH10; MRE11A; MSH2; MSH6; MTA2; MYBBP1A; NAF1; NASP; NBN; NCAPD3; NELFA; NIFK; NOC2L; NOL11; NOL8; NOL9; NOLC1; NOP2; NOP56; NOP58; NSUN2; NUP133; NUP155; NUP160; NUP205; NUP210; PABPN1; PAF1; PAPD5; PAPOLA; PARN; PARP1; PAXBP1; PDCD11; PELP1; PHB; PHF6; PHF8; POLA2; POLD1; POLE; POLR1A; POLR1B; POLR1E; POLR2A; POLR2B; POLR2E; POLR2L; POLR3A; POP1; PPP1R10; PPP2R1A; PQBP1; PRKDC; PRPF38B; PRPF40A; PRPF4B; PRPF8; PRPS1; PSIP1; PSMC5; PSME3; PTBP1; PUS7; PWP1; PWP2; RAD50; RAD9A; RBM10; RBM22; RBM25; RBM3; RBM4; RBMX; RBMXL1; RFC1; RNPS1; RPA1; RPA2; RPL10; RPL19; RPL28; RPL32; RPL4; RPL7; RPS27A; UBC; UBB; UBA52; RRN3; RRP1; RRP1B; RTCB; RTF1; SAFB; SART3; SCAF11; SET; SETD1A; SETD2; SETX; SF1; SF3A1; SF3B1; SF3B2; SF3B3; SF3B4; SHMT2; SIN3A; SKIV2L2; SLU7; SMARCA4; SMARCA5; SMARCAD1; SMARCC1; SMC1A; SMC3; SMNDC1; SNRNP200; SNRNP70; SNRPE; SNW1; SREK1; SRRM2; SRRT; SRSF1; SRSF10; SRSF2; SRSF6; SRSF9; SSRP1; STOML2; SUGP2; SUMO1; SUPT16H; SUPT5H; SUPT6H; SYMPK; TARDBP; TDP43; TBL3; TCEB3; TCERG1; TCF20; TCOF1; THOC2; THRAP3; TK1; TOP1; TOP2A; TOP2B; TOPBP1; TP53; TP53BP1; TPR; TRA2A; TRA2B; TRIM27; TRIM28; TRMT1L; TWISTNB; TXN; U2AF1; U2AF2; U2SURP; UBR5; UBTF; UHRF1; USP7; UTP14A; UTP15; UTP20; UTP3; VARS; WBP11; WDR3; WDR36; WRN; WRNIP1; XRCC1; XRN2; YWHAB; ZMYM2; ZNF318; ZNF638

| nucleobase-containing compound metabolic process | 5,51E-17 |
|---|---|

AAAS; ACIN1; ADAR; AFF4; AKAP17A; AKAP8; ALKBH5; APRT; ARID1A; ASF1B; ATAD2; ATP5J2; ATP5J2-PTCD1; ATP5L; ATR; BAZ1A; BAZ1B; BAZ2A; BCLAF1; BLM; BPTF; BRCA1; BRD2; BRIP1; CBX3; CCNK; CDC5L; CDC73; CDK12; CENPF; CHAF1A; CHD1; CHD4; CIRH1A; COASY; CPSF1; CPSF6; CSTF2; CTR9; DDB1; DDX17; DDX18; DDX21; DDX24; DDX27; DDX42; DDX47; DDX5; DDX50; DDX52; DDX54; DHX15; DHX16; DHX37; DHX38; DHX8; DHX9; DIDO1; DKC1; DNMT1; EFTUD2; EIF4A3; ELAC2; ERCC3; ESF1; ETF1; EWSR1; FANCI; FARSA; FLNA; FTSJ3;

FUBP1; FUS; GATAD2B; GEMIN5; GINS3; GNAS; GTF2F2; GTF2H1; GTF2I; GTF3C1; GTF3C3; H2AFX; HIST1H2AA; HIST2H2AB; HAT1; HEATR1; HELLS; HEXIM1; HIST1H4A; HLTF; HNRNPA1; HNRNPA2B1; HNRPA2B1; HNRNPA3; HNRNPC; HNRNPD; HNRNPDL; HNRNPF; HNRNPH3; HNRNPK; HNRNPL; HNRNPR; HSPA8; HTATSF1; HUWE1; ICE1; ILF2; ILF3; INTS9; IWS1; KDM3A; KDM3B; KHDRBS1; KHSRP; KIAA1429; KIF22; KMT2D; KPNA2; KPNB1; LAS1L; LEO1; LIG3; LUC7L3; MCMBP; MCTS1; MDC1; MED1; MED12; TNRC11; MEPCE; MPHOSPH10; MRE11A; MSH2; MSH6; MTA2; MYBBP1A; NAF1; NASP; NBN; NCAPD3; NELFA; NIFK; NOC2L; NOL11; NOL8; NOL9; NOLC1; NOP2; NOP56; NOP58; NSUN2; NUP133; NUP155; NUP160; NUP205; NUP210; PABPN1; PAF1; PAPD5; PAPOLA; PARN; PARP1; PAXBP1; PDCD11; PELP1; PHB; PHF6; PHF8; POLA2; POLD1; POLE; POLR1A; POLR1B; POLR1E; POLR2A; POLR2B; POLR2E; POLR2L; POLR3A; POP1; PPP1R10; PPP2R1A; PQBP1; PRKDC; PRPF38B; PRPF40A; PRPF4B; PRPF8; PRPS1; PSIP1; PSMC5; PTBP1; PUS7; PWP1; PWP2; RAD50; RAD9A; RBM10; RBM22; RBM25; RBM3; RBM4; RBMX; RBMXL1; RFC1; RNPS1; RPA1; RPA2; RPL10; RPL19; RPL28; RPL32; RPL4; RPL7; RPS27A; UBC; UBB; UBA52; RRN3; RRP1; RRP1B; RTCB; RTF1; SAFB; SART3; SCAF11; SET; SETD1A; SETD2; SETX; SF1; SF3A1; SF3B1; SF3B2; SF3B3; SF3B4; SIN3A; SKIV2L2; SLU7; SMARCA4; SMARCA5; SMARCAD1; SMARCC1; SMC1A; SMC3; SMNDC1; SNRNP200; SNRNP70; SNRPE; SNW1; SREK1; SRRM2; SRRT; SRSF1; SRSF10; SRSF2; SRSF6; SRSF9; SSRP1; STOML2; SUGP2; SUMO1; SUPT16H; SUPT5H; SUPT6H; SYMPK; TARDBP; TDP43; TBL3; TCEB3; TCERG1; TCF20; TCOF1; THOC2; THRAP3; TK1; TOP1; TOP2A; TOP2B; TOPBP1; TP53; TP53BP1; TPR; TRA2A; TRA2B; TRIM27; TRIM28; TRMT1L; TWISTNB; TXN; U2AF1; U2AF2; U2SURP; UBR5; UBTF; UHRF1; USP7; UTP14A; UTP15; UTP20; UTP3; VARS; WBP11; WDR3; WDR36; WRN; WRNIP1; XRCC1; XRN2; YWHAB; ZMYM2; ZNF318; ZNF638

| | |
|---|---|
| mRNA processing | 8,33E-17 |

ACIN1; ADAR; AKAP17A; ALKBH5; CDC5L; CDC73; CDK12; CPSF1; CPSF6; CSTF2; DDX17; DDX47; DDX5; DHX15; DHX16; DHX38; DHX8; DHX9; EFTUD2; EIF4A3; ERCC3; FUS; GEMIN5; GTF2F2; GTF2H1; HNRNPA1; HNRNPA2B1; HNRPA2B1; HNRNPA3; HNRNPC; HNRNPD; HNRNPF; HNRNPH3; HNRNPK; HNRNPL; HNRNPR; HSPA8; IWS1; KHDRBS1; KHSRP; KIAA1429; LEO1; PABPN1; PAF1; PAPD5; PAPOLA; PARN; PDCD11; POLR2A; POLR2B; POLR2E; POLR2L; PQBP1; PRPF38B; PRPF40A; PRPF4B; PRPF8; PTBP1; RBM10; RBM22; RBM25; RBM4; RBMX; RBMXL1; RNPS1; SART3; SCAF11; SF1; SF3A1; SF3B1; SF3B2; SF3B3; SF3B4; SKIV2L2; SLU7; SMC1A; SMNDC1; SNRNP200; SNRNP70; SNRPE; SNW1; SREK1; SRRM2; SRSF1; SRSF10; SRSF2; SRSF6; SRSF9; SUGP2; SUPT5H; SUPT6H; SYMPK; TARDBP; TDP43; THOC2; THRAP3; TRA2A; TRA2B; U2AF1; U2AF2; WBP11; XRN2

| | |
|---|---|
| RNA splicing | 2,55E-16 |

ACIN1; AKAP17A; CDC5L; CDK12; CPSF1; CSTF2; DDX17; DDX47; DDX5; DHX15; DHX16; DHX38; DHX8; DHX9; EFTUD2; EIF4A3; FUS; GEMIN5; GTF2F2; HNRNPA1; HNRNPA2B1; HNRPA2B1; HNRNPA3; HNRNPC; HNRNPD; HNRNPF; HNRNPH3; HNRNPK; HNRNPL; HNRNPR; HSPA8; IWS1; KHSRP; KIAA1429; LUC7L3; MPHOSPH10; PABPN1; PAPOLA; POLR2A; POLR2B; POLR2E; POLR2L; PPP2R1A; PQBP1; PRPF38B; PRPF40A; PRPF4B; PRPF8; PTBP1; RBM10; RBM22; RBM25; RBM4; RBMX; RBMXL1; RNPS1; RTCB; SART3; SCAF11; SF1; SF3A1; SF3B1; SF3B2; SF3B3; SF3B4; SKIV2L2; SLU7; SMC1A; SMNDC1; SNRNP200; SNRNP70; SNRPE; SNW1; SREK1; SRRM2; SRSF1; SRSF10; SRSF2; SRSF6; SRSF9; SUGP2; SUPT6H; TARDBP; TDP43; THOC2; THRAP3; TRA2A; TRA2B; U2AF1; U2AF2; WBP11; ZNF638

| | |
|---|---|
| macromolecule metabolic process | 1,20E-12 |

AAAS; ACIN1; ADAR; AFF4; AGTPBP1; AKAP17A; AKAP8; ALKBH5; ARID1A; ASF1B; ATAD2; ATR; BAZ1A; BAZ1B; BAZ2A; BCLAF1; BLM; BPTF; BRCA1; BRD2; BRIP1; CBX3; CCNK; CCT2; CCT3; CDC27; CDC5L; CDC73; CDK11B; CDK11A; CDK12; CENPF; CHAF1A; CHD1; CHD4; CIRH1A; CPSF1; CPSF6; CRKL; CSNK2B; CSNK2B-LY6G5B-1181; CSNK2B-LY6G5B--991; CSTF2; CTR9; DDB1; DDX17; DDX18; DDX21; DDX24; DDX27; DDX42; DDX47; DDX5; DDX50;

DDX52; DDX54; DHX15; DHX16; DHX37; DHX38; DHX8; DHX9; DIDO1; DKC1; DNMT1; EEF1B2; EFTUD2; EIF4A3; EIF5A; EIF5AL1; ELAC2; ERCC3; ERLIN2; ESF1; ETF1; EWSR1; EZR; FANCI; FARSA; FLNA; FOXC1; FTSJ3; FUBP1; FUS; GANAB; GATAD2B; GEMIN5; GINS3; GNAS; GTF2F2; GTF2H1; GTF2I; GTF3C1; GTF3C3; H2AFX; HIST1H2AA; HIST2H2AB; HAT1; HEATR1; HELLS; HEXIM1; HIST1H2BN; HIST1H2BL; HIST1H2BM; HIST1H2BH; HIST2H2BF; HIST1H2BC; HIST1H2BD; H2BFS; HIST1H2BK; HIST1H2BA; HIST1H4A; HLTF; HNRNPA1; HNRNPA2B1; HNRPA2B1; HNRNPA3; HNRNPC; HNRNPD; HNRNPDL; HNRNPF; HNRNPH3; HNRNPK; HNRNPL; HNRNPR; HSP90AA1; HSPA8; HSPE1; HSPE1-MOB4; HTATSF1; HUWE1; ICE1; ILF2; ILF3; INTS9; IWS1; KDM3A; KDM3B; KHDRBS1; KHSRP; KIAA1429; KIF22; KMT2D; KPNA2; KPNB1; LAS1L; LEO1; LIG3; LUC7L3; MCMBP; MCTS1; MDC1; MED1; MED12; TNRC11; MEPCE; MORC3; MPHOSPH10; MRE11A; MSH2; MSH6; MTA2; MYBBP1A; NAF1; NASP; NBN; NCAPD3; NELFA; NIFK; NOC2L; NOL11; NOL8; NOL9; NOLC1; NOP2; NOP56; NOP58; NSUN2; NUP133; NUP155; NUP160; NUP205; NUP210; OGT; PABPN1; PAF1; PAPD5; PAPOLA; PARN; PARP1; PAXBP1; PDCD11; PDIA3; PDIA4; PEF1; PELP1; PHB; PHF6; PHF8; POLA2; POLD1; POLE; POLR1A; POLR1B; POLR1E; POLR2A; POLR2B; POLR2E; POLR2L; POLR3A; POP1; PPIL4; PPM1G; PPP1CC; PPP1R10; PPP2R1A; PPP6C; PQBP1; PRKDC; PRPF38B; PRPF40A; PRPF4B; PRPF8; PSIP1; PSMC5; PSME3; PTBP1; PUS7; PWP1; PWP2; RAD50; RAD9A; RBBP6; RBM10; RBM22; RBM25; RBM3; RBM4; RBMX; RBMXL1; RFC1; RNPS1; RPA1; RPA2; RPL10; RPL19; RPL28; RPL32; RPL4; RPL7; RPN1; RPS27A; UBC; UBB; UBA52; RRN3; RRP1; RRP1B; RTCB; RTF1; SAFB; SART3; SCAF11; SET; SETD1A; SETD2; SETX; SF1; SF3A1; SF3B1; SF3B2; SF3B3; SF3B4; SIN3A; SKIV2L2; SLC25A6; SLU7; SMARCA4; SMARCA5; SMARCAD1; SMARCC1; SMC1A; SMC3; SMNDC1; SNRNP200; SNRNP70; SNRPE; SNW1; SREK1; SRI; SRRM2; SRRT; SRSF1; SRSF10; SRSF2; SRSF6; SRSF9; SSRP1; STAG2; STOML2; SUGP2; SUMO1; SUMO2; SUPT16H; SUPT5H; SUPT6H; SYMPK; TARDBP; TDP43; TBCB; TBL3; TCEB3; TCERG1; TCF20; TCOF1; TCP1; THOC2; THRAP3; TK1; TOP1; TOP2A; TOP2B; TOPBP1; TP53; TP53BP1; TPR; TRA2A; TRA2B; TRIM27; TRIM28; TRMT1L; TTK; TWISTNB; TXN; U2AF1; U2AF2; U2SURP; UBA3; UBE2M; UBR5; UBTF; UHRF1; USP36; USP48; USP7; UTP14A; UTP15; UTP20; UTP3; VARS; WBP11; WDR3; WDR36; WRN; WRNIP1; XPO1; XPO5; XRCC1; XRN2; YWHAB; ZMYM2; ZNF318; ZNF638

| | |
|---|---|
| nucleus | 1,43E-11 |

AAAS; ACIN1; ADAR; AFF4; AGTPBP1; AKAP17A; AKAP8; AKAP8L; ALKBH5; ANP32E; ANXA7; ARID1A; ASF1B; ATAD5; ATP5J2; ATP5J2-PTCD1; BAZ1B; BAZ2A; BCLAF1; BLM; BMS1; BPTF; BRCA1; BRD2; BRIP1; BTAF1; CASC5; CBX3; CCNK; CDC27; CDC5L; CDC73; CDK11B; CDK11A; CDK12; CENPF; CHD1; CHD4; COIL; CPSF6; CSNK2B; CSNK2B-LY6G5B-1181; CSNK2B-LY6G5B--991; CSTF2; DDB1; DDX17; DDX21; DDX42; DDX5; DDX52; DDX54; DHX15; DHX16; DHX38; DHX8; DHX9; DIDO1; DKC1; DNMT1; DONSON; EIF5A; EIF5AL1; ELAC2; EPB41L2; ERCC3; EWSR1; FAM98B; FANCI; FLNA; FOXC1; FOXK1; FTSJ3; FUBP1; FUS; GEMIN5; GNL3; GTF2F2; GTF2I; GTF3C1; GTPBP4; H1FX; H2AFX; HIST1H2AA; HIST2H2AB; HAT1; HEATR1; HELLS; HEXIM1; HIST1H2BN; HIST1H2BL; HIST1H2BM; HIST1H2BH; HIST2H2BF; HIST1H2BC; HIST1H2BD; H2BFS; HIST1H2BK; HIST1H2BA; HIST1H4A; HLTF; HNRNPA2B1; HNRPA2B1; HNRNPA3; HNRNPAB; HNRNPC; HNRNPD; HNRNPDL; HNRNPF; HNRNPH3; HNRNPK; HNRNPL; HNRNPUL2; HNRNPUL2-BSCL2; HSP90AA1; HSPA8; HTATSF1; HUWE1; IK; ILF2; ILF3; IPO5; IWS1; KDM3A; KHDRBS1; KIF22; KMT2D; LIG3; LUC7L3; LYAR; MASTL; MCMBP; MDC1; MDN1; MED1; MED12; TNRC11; MKI67; MORC2; MRE11A; MYBBP1A; NACC1; NAF1; NASP; NBN; NCAPD3; NCL; NIFK; NIPBL; NOC2L; NOL8; NOP58; NSUN2; NUMA1; NVL; OGT; PABPN1; PAPOLA; PARN; PARP1; PAXBP1; PDCD11; PDCD4; PDIA3; PEBP1; PELP1; PHB; PHF6; PHF8; PHIP; POLD1; POLE; POLR1A; POLR1E; POLR2A; POLR2B; POLR2E; POLR2L; PPAN-P2RY11; PPAN; PPIL4; PPM1G; PPP1CC; PPP1R10; PPP2R1A; PQBP1; PRC1; PRCC; PRPF8; PSIP1; PSMC5; PSME3; PTBP1; PWP1; RACGAP1; RAD50; RAD54L2; RAD9A; RAI1; RANBP3; RAVER1; RBM10; RBM22; RBM3; RBM4; RBMX; RBMXL1; RFC1; RIF1; RNPS1; RPA1; RPA2; RPL4; RPL7; RRN3; RRP1; RRP1B; RRS1; RTCB; SAFB; SART3; SCAF11; SENP3; SENP3-EIF4A1; SET; SETD1A; SETD2; SETX; SF1; SF3B1; SIN3A; SKIV2L2; SLC25A6; SLFN11; SLU7;

SMARCA4; SMARCA5; SMC1A; SMC3; SMNDC1; SNRNP200; SNRNP70; SNRPE; SNW1; SRRT; SRSF1; SRSF10; SRSF2; SSRP1; STAG2; SUMO1; SUMO2; SUPT16H; SUPT5H; SUPT6H; TAF15; TARDBP; TDP43; TBL3; TCEB3; TCERG1; TCF20; TCOF1; TCOF1; THRAP3; TLE1; TMPO; TMPO; TOP1; TOP2A; TOP2B; TOPBP1; TP53; TP53BP1; TPR; TPX2; TRA2A; TRA2B; TRIM27; TRIM28; TRMT1; TWISTNB; TXN; U2AF2; U2SURP; UBA3; UBR5; UHRF1; URB2; USP36; USP7; UTP15; UTP3; WAPAL; WBP11; WDR3; WRNIP1; XPO1; XPO5; XRCC1; XRN2; YLPM1; YWHAB; ZNF207

| intracellular non-membrane-bounded organelle | 2,56E-10 |
|---|---|

AAAS; ACIN1; ACTA1; ACTC1; ACTG2; ACTA2; ACTR1A; ADAR; AFF4; AKAP8; ANLN; ANXA11; ATAD3A; ATR; BAZ1A; BAZ1B; BAZ2A; BCLAF1; BLM; BMS1; BRCA1; CASC5; CBX3; CCT3; CDC27; CDC5L; CDCA2; CDK12; CENPF; CHD4; CIRH1A; CNN3; COIL; DDX17; DDX18; DDX21; DDX24; DDX27; DDX42; DDX47; DDX5; DDX50; DDX52; DDX54; DHX15; DHX37; DHX9; DIDO1; DKC1; EIF4A3; ESF1; EWSR1; EZR; FLNA; FOXK1; FTSJ3; FXR1; GNL3; GTF2F2; GTF3C1; GTPBP4; H1FX; H2AFX; HIST1H2AA; HIST2H2AB; HEATR1; HIST1H4A; HLTF; HNRNPA1; HNRNPK; HSPA8; IK; ILF2; ILF3; IPO5; KHSRP; KIF22; KIF4A; LAS1L; LEO1; LYAR; MASTL; MDC1; MDN1; MED1; MED12; TNRC11; MKI67; MORC3; MPHOSPH10; MSH2; MSH6; MYBBP1A; NBN; NCL; NIFK; NOC2L; NOL11; NOL8; NOL9; NOLC1; NOP2; NOP56; NOP58; NSUN2; NUMA1; NUP133; NVL; PAPD5; PARN; PARP1; PDCD11; PDS5A; PELP1; PHF6; PHF8; POLR1A; POLR1B; POLR1E; PPAN-P2RY11; PPAN; PPP1CC; PPP2R1A; PQBP1; PRC1; PRKDC; PRPF4B; PTBP1; PWP1; PWP2; RACGAP1; RBBP6; RBM4; RFC1; RIF1; RPL10; RPL19; RPL28; RPL32; RPL4; RPL7; RPS27A; UBC; UBB; UBA52; RRN3; RRP1; RRP12; RRP1B; RRS1; RSL1D1; RTF1; SCAF11; SENP3; SENP3-EIF4A1; SETD1A; SETD2; SETX; SF1; SHMT2; SIN3A; SKIV2L2; SMARCA4; SMARCA5; SMARCC1; SMC1A; SMC3; SMNDC1; SRSF9; SSRP1; STAG2; STOML2; SUMO1; SUPT16H; SYMPK; TBCB; TBL3; TCERG1; TCOF1; TCOF1; TCP1; TEX10; TLE1; TOP1; TOP2A; TOPBP1; TP53; TP53BP1; TPR; TPX2; TRA2A; TSR1; TTK; TWISTNB; UBTF; URB1; URB2; USP36; UTP14A; UTP15; UTP20; UTP3; WAPAL; WDR3; WDR36; WDR43; WDR75; WRN; XPO1; XRN2; ZFR; ZNF207

| mRNA metabolic process | 1,82E-09 |
|---|---|

ACIN1; ADAR; AKAP17A; ALKBH5; CDC5L; CDC73; CDK12; CPSF1; CPSF6; CSTF2; DDX17; DDX47; DDX5; DHX15; DHX16; DHX38; DHX8; DHX9; EFTUD2; EIF4A3; ERCC3; ETF1; FUS; GEMIN5; GTF2F2; GTF2H1; HNRNPA1; HNRNPA2B1; HNRPA2B1; HNRNPA3; HNRNPC; HNRNPD; HNRNPF; HNRNPH3; HNRNPK; HNRNPL; HNRNPR; HSPA8; IWS1; KHDRBS1; KHSRP; KIAA1429; LEO1; PABPN1; PAF1; PAPD5; PAPOLA; PARN; PDCD11; POLR2A; POLR2B; POLR2E; POLR2L; PPP2R1A; PQBP1; PRPF38B; PRPF40A; PRPF4B; PRPF8; PTBP1; RBM10; RBM22; RBM25; RBM4; RBMX; RBMXL1; RNPS1; RPL10; RPL19; RPL28; RPL32; RPL4; RPL7; RPS27A; UBC; UBB; UBA52; SART3; SCAF11; SF1; SF3A1; SF3B1; SF3B2; SF3B3; SF3B4; SKIV2L2; SLU7; SMC1A; SMNDC1; SNRNP200; SNRNP70; SNRPE; SNW1; SREK1; SRRM2; SRSF1; SRSF10; SRSF2; SRSF6; SRSF9; SUGP2; SUPT5H; SUPT6H; SYMPK; TARDBP; TDP43; THOC2; THRAP3; TRA2A; TRA2B; U2AF1; U2AF2; WBP11; XRN2

| DNA metabolic process | 4,06E-08 |
|---|---|

ACIN1; AKAP8; ATR; BAZ1A; BAZ1B; BAZ2A; BLM; BRCA1; BRIP1; CDC5L; CENPF; CHAF1A; CHD1; CHD4; DDB1; DHX9; DKC1; DNMT1; ERCC3; FANCI; GINS3; GNAS; GTF2H1; GTPBP4; H2AFX; HIST1H2AA; HIST2H2AB; HAT1; HELLS; HIST1H4A; HLTF; HUWE1; KIF22; KPNA2; KPNB1; LIG3; MCMBP; MDC1; MRE11A; MSH2; MSH6; MTA2; NASP; NBN; NCAPD3; NOL8; PARP1; PDS5A; PHB; POLA2; POLD1; POLE; POLR2A; POLR2B; POLR2E; POLR2L; PPP2R1A; PRKDC; RAD50; RAD9A; RFC1; RPA1; RPA2; RPS27A; UBC; UBB; UBA52; SET; SETD2; SETX; SIN3A; SMARCA4; SMARCA5; SMARCAD1; SMC1A; SMC3; SSRP1; STAG2; STOML2; SUMO1; SUPT16H; SUPT6H; TK1; TOP1; TOP2A; TOP2B; TOPBP1; TP53; TP53BP1; TRIM28; UBR5; UHRF1; USP7; WAPAL; WRN; WRNIP1; XRCC1; XRN2; YLPM1

| spliceosomal complex | 9,72E-07 |
|---|---|

AKAP17A; CDC5L; DDX5; DHX15; DHX16; DHX38; DHX8; EFTUD2; EIF4A3; HNRNPA1; HNRNPA2B1; HNRPA2B1; HNRNPA3; HNRNPC; HNRNPDL; HNRNPF; HNRNPH3; HNRNPK; HNRNPR; HSPA8; PRPF38B; PRPF4B; PRPF8; RBM22; RBM25; RBMX; SF1; SF3A1; SF3B1; SF3B2; SF3B3; SF3B4; SKIV2L2; SLU7; SMNDC1; SNRNP200; SNRNP70; SNRPE; SNW1; SREK1; SRRM2; SRSF1; SRSF2; U2AF1; U2AF2

| chromosome organization | 1,13E-06 |
|---|---|
| ACIN1; AKAP8; ANP32E; ARID1A; ASF1B; BAZ1A; BAZ1B; BAZ2A; BLM; BPTF; BRD2; CASC5; CBX3; CDC73; CENPF; CHAF1A; CHD1; CHD4; CTR9; DDB1; DKC1; DNMT1; GATAD2B; H1FX; H2AFX; HIST1H2AA; HIST2H2AB; HAT1; HELLS; HIST1H2BN; HIST1H2BL; HIST1H2BM; HIST1H2BH; HIST2H2BF; HIST1H2BC; HIST1H2BD; H2BFS; HIST1H2BK; HIST1H2BA; HIST1H4A; HLTF; HNRNPC; HUWE1; KDM3A; KDM3B; KIF22; KMT2D; LEO1; MCMBP; MRE11A; MSH2; MTA2; NASP; NBN; NCAPD3; NIPBL; NOC2L; OGT; PAF1; PARP1; PDS5A; PHB; PHF8; POLA2; POLD1; POLE; PRKDC; RAD50; RFC1; RPA1; RPA2; RTF1; SAFB; SART3; SET; SETD1A; SETD2; SIN3A; SMARCA4; SMARCA5; SMARCAD1; SMARCC1; SMC1A; SMC3; SSRP1; STAG2; SUPT16H; SUPT5H; SUPT6H; TOP1; TOP2A; TOP2B; TP53; UBTF; UHRF1; USP7; UTP3; WRN; ZNF207 | |
| DNA repair | 3,52E-06 |
| ATR; BAZ1B; BLM; BRCA1; BRIP1; CDC5L; CHAF1A; DDB1; ERCC3; FANCI; GTF2H1; H2AFX; HIST1H2AA; HIST2H2AB; HLTF; HUWE1; KIF22; LIG3; MDC1; MRE11A; MSH2; MSH6; NBN; PARP1; POLD1; POLE; POLR2A; POLR2B; POLR2E; POLR2L; PRKDC; RAD50; RAD9A; RFC1; RPA1; RPA2; RPS27A; UBC; UBB; UBA52; SETD2; SETX; SMARCA5; SMC1A; SMC3; SSRP1; SUMO1; SUPT16H; TOPBP1; TP53; TP53BP1; TRIM28; UBR5; UHRF1; USP7; WRN; XRCC1 | |
| chromosome | 4,63E-06 |
| AKAP8; ATR; BAZ1A; BAZ1B; BLM; BRCA1; CDCA2; H2AFX; HIST1H2AA; HIST2H2AB; HIST1H4A; KIF4A; MDC1; MKI67; MPHOSPH10; MSH2; MSH6; NIFK; NUMA1; PDS5A; PRPF4B; RBBP6; RRS1; SETD1A; SETD2; SETX; SMARCA5; SMARCC1; SMC1A; SMC3; SSRP1; STAG2; SUPT16H; TOP2A; TOPBP1; WAPAL; ZFR | |
| nucleobase-containing compound metabolic process | 1,24E-05 |
| AAAS; ACIN1; ADAR; AFF4; AKAP17A; AKAP8; ALKBH5; APRT; ARID1A; ASF1B; ATAD2; ATP5J2; ATP5J2-PTCD1; ATP5L; ATR; BAZ1A; BAZ1B; BAZ2A; BCLAF1; BLM; BPTF; BRCA1; BRD2; BRIP1; CBX3; CCNK; CDC5L; CDC73; CDK12; CENPF; CHAF1A; CHD1; CHD4; CIRH1A; COASY; CPSF1; CPSF6; CSTF2; CTR9; DDB1; DDX17; DDX18; DDX21; DDX24; DDX27; DDX42; DDX47; DDX5; DDX50; DDX52; DDX54; DHX15; DHX16; DHX37; DHX38; DHX8; DHX9; DIDO1; DKC1; DNMT1; EFTUD2; EIF4A3; ELAC2; ERCC3; ESF1; ETF1; EWSR1; FANCI; FARSA; FLNA; FTSJ3; FUBP1; FUS; GATAD2B; GEMIN5; GINS3; GNAS; GTF2F2; GTF2H1; GTF2I; GTF3C1; GTF3C3; H2AFX; HIST1H2AA; HIST2H2AB; HAT1; HEATR1; HELLS; HEXIM1; HIST1H4A; HLTF; HNRNPA1; HNRNPA2B1; HNRPA2B1; HNRNPA3; HNRNPC; HNRNPD; HNRNPDL; HNRNPF; HNRNPH3; HNRNPK; HNRNPL; HNRNPR; HSPA8; HTATSF1; HUWE1; ICE1; ILF2; ILF3; INTS9; IWS1; KDM3A; KDM3B; KHDRBS1; KHSRP; KIAA1429; KIF22; KMT2D; KPNA2; KPNB1; LAS1L; LEO1; LIG3; LUC7L3; MCMBP; MCTS1; MDC1; MED1; MED12; TNRC11; MEPCE; MPHOSPH10; MRE11A; MSH2; MSH6; MTA2; MYBBP1A; NAF1; NASP; NBN; NCAPD3; NELFA; NIFK; NOC2L; NOL11; NOL8; NOL9; NOLC1; NOP2; NOP56; NOP58; NSUN2; NUP133; NUP155; NUP160; NUP205; NUP210; PABPN1; PAF1; PAPD5; PAPOLA; PARN; PARP1; PAXBP1; PDCD11; PELP1; PHB; PHF6; PHF8; POLA2; POLD1; POLE; POLR1A; POLR1B; POLR1E; POLR2A; POLR2B; POLR2E; POLR2L; POLR3A; POP1; PPP1R10; PPP2R1A; PQBP1; PRKDC; PRPF38B; PRPF40A; PRPF4B; PRPF8; PRPS1; PSIP1; PSMC5; PTBP1; PUS7; PWP1; PWP2; RAD50; RAD9A; RBM10; RBM22; RBM25; RBM3; RBM4; RBMX; RBMXL1; RFC1; RNPS1; RPA1; RPA2; RPL10; RPL19; RPL28; RPL32; RPL4; RPL7; RPS27A; UBC; UBB; UBA52; RRN3; RRP1; RRP1B; RTCB; RTF1; SAFB; SART3; SCAF11; SET; SETD1A; SETD2; SETX; SF1; SF3A1; SF3B1; SF3B2; SF3B3; | |

SF3B4; SIN3A; SKIV2L2; SLU7; SMARCA4; SMARCA5; SMARCAD1; SMARCC1; SMC1A; SMC3; SMNDC1; SNRNP200; SNRNP70; SNRPE; SNW1; SREK1; SRRM2; SRRT; SRSF1; SRSF10; SRSF2; SRSF6; SRSF9; SSRP1; STOML2; SUGP2; SUMO1; SUPT16H; SUPT5H; SUPT6H; SYMPK; TARDBP; TDP43; TBL3; TCEB3; TCERG1; TCF20; TCOF1; THOC2; THRAP3; TK1; TOP1; TOP2A; TOP2B; TOPBP1; TP53; TP53BP1; TPR; TRA2A; TRA2B; TRIM27; TRIM28; TRMT1L; TWISTNB; TXN; U2AF1; U2AF2; U2SURP; UBR5; UBTF; UHRF1; USP7; UTP14A; UTP15; UTP20; UTP3; VARS; WBP11; WDR3; WDR36; WRN; WRNIP1; XRCC1; XRN2; YWHAB; ZMYM2; ZNF318; ZNF638

| transcription termination, DNA-dependent | 2,66E-05 |
|---|---|

CPSF1; CSTF2; DHX38; ERCC3; GTF2H1; PABPN1; PAPOLA; POLR1A; POLR1B; POLR1E; POLR2E; POLR2L; POLR3A; RNPS1; SETX; SNRPE; SRSF1; SRSF2; SRSF6; SRSF9; TWISTNB; U2AF1; U2AF2; UBTF; XRN2

| primary metabolic process | 3,01E-05 |
|---|---|

AAAS; ABCD3; ACIN1; ADAR; AFF4; AGTPBP1; AKAP17A; AKAP8; ALDH2; ALKBH5; APRT; ARID1A; ASF1B; ATAD2; ATP5J2; ATP5J2-PTCD1; ATP5L; ATR; BAZ1A; BAZ1B; BAZ2A; BCLAF1; BLM; BPTF; BRCA1; BRD2; BRIP1; CBX3; CCNK; CCT2; CCT3; CDC27; CDC5L; CDC73; CDK11B; CDK11A; CDK12; CENPF; CHAF1A; CHD1; CHD4; CIRH1A; COASY; CPSF1; CPSF6; CRKL; CSNK2B; CSNK2B-LY6G5B-1181; CSNK2B-LY6G5B--991; CSTF2; CTR9; DDB1; DDX17; DDX18; DDX21; DDX24; DDX27; DDX42; DDX47; DDX5; DDX50; DDX52; DDX54; DHX15; DHX16; DHX37; DHX38; DHX8; DHX9; DIDO1; DKC1; DNMT1; EEF1B2; EFTUD2; EIF4A3; EIF5A; EIF5AL1; ELAC2; ERCC3; ERLIN2; ESF1; ETF1; EWSR1; FANCI; FARSA; FLNA; FOXC1; FTSJ3; FUBP1; FUS; GANAB; GATAD2B; GEMIN5; GINS3; GNAS; GOLT1B; GTF2F2; GTF2H1; GTF2I; GTF3C1; GTF3C3; H2AFX; HIST1H2AA; HIST2H2AB; HAT1; HEATR1; HELLS; HEXIM1; HIST1H2BN; HIST1H2BL; HIST1H2BM; HIST1H2BH; HIST2H2BF; HIST1H2BC; HIST1H2BD; H2BFS; HIST1H2BK; HIST1H2BA; HIST1H4A; HLTF; HNRNPA1; HNRNPA2B1; HNRPA2B1; HNRNPA3; HNRNPC; HNRNPD; HNRNPDL; HNRNPF; HNRNPH3; HNRNPK; HNRNPL; HNRNPR; HSP90AA1; HSPA8; HSPE1; HSPE1-MOB4; HTATSF1; HUWE1; ICE1; ILF2; ILF3; INTS9; IWS1; KDM3A; KDM3B; KHDRBS1; KHSRP; KIAA1429; KIF22; KMT2D; KPNA2; KPNB1; LAS1L; LEO1; LIG3; LUC7L3; MCMBP; MCTS1; MDC1; MED1; MED12; TNRC11; MEPCE; MIF; MORC2; MORC3; MPHOSPH10; MRE11A; MSH2; MSH6; MTA2; MYBBP1A; NAF1; NASP; NBN; NCAPD3; NELFA; NIFK; NOC2L; NOL11; NOL8; NOL9; NOLC1; NOP2; NOP56; NOP58; NSUN2; NUP133; NUP155; NUP160; NUP205; NUP210; OGT; PABPN1; PAF1; PAPD5; PAPOLA; PARN; PARP1; PAXBP1; PDCD11; PDIA3; PDIA4; PEF1; PELP1; PHB; PHF6; PHF8; POLA2; POLD1; POLE; POLR1A; POLR1B; POLR1E; POLR2A; POLR2B; POLR2E; POLR2L; POLR3A; POP1; PPIL4; PPM1G; PPP1CC; PPP1R10; PPP2R1A; PPP6C; PQBP1; PRKDC; PRPF38B; PRPF40A; PRPF4B; PRPF8; PRPS1; PSIP1; PSMC5; PSME3; PTBP1; PUS7; PWP1; PWP2; RAD50; RAD9A; RBBP6; RBM10; RBM22; RBM25; RBM3; RBM4; RBMX; RBMXL1; RFC1; RNPS1; RPA1; RPA2; RPL10; RPL19; RPL28; RPL32; RPL4; RPL7; RPN1; RPS27A; UBC; UBB; UBA52; RRN3; RRP1; RRP1B; RTCB; RTF1; SAFB; SART3; SCAF11; SET; SETD1A; SETD2; SETX; SF1; SF3A1; SF3B1; SF3B2; SF3B3; SF3B4; SHMT2; SIN3A; SKIV2L2; SLC25A6; SLU7; SMARCA4; SMARCA5; SMARCAD1; SMARCC1; SMC1A; SMC3; SMNDC1; SNRNP200; SNRNP70; SNRPE; SNW1; SREK1; SRI; SRRM2; SRRT; SRSF1; SRSF10; SRSF2; SRSF6; SRSF9; SSRP1; STAG2; STOML2; SUGP2; SUMO1; SUMO2; SUPT16H; SUPT5H; SUPT6H; SYMPK; TARDBP; TDP43; TBCB; TBL3; TCEB3; TCERG1; TCF20; TCOF1; TCP1; THOC2; THRAP3; TK1; TOP1; TOP2A; TOP2B; TOPBP1; TP53; TP53BP1; TPR; TRA2A; TRA2B; TRIM27; TRIM28; TRMT1L; TTK; TWISTNB; TXN; U2AF1; U2AF2; U2SURP; UBA3; UBE2M; UBR5; UBTF; UHRF1; USP36; USP48; USP7; UTP14A; UTP15; UTP20; UTP3; VARS; WBP11; WDR3; WDR36; WRN; WRNIP1; XRCC1; XRN2; YWHAB; ZMYM2; ZNF318; ZNF638

| biological regulation | 4,04E-05 |
|---|---|

AAAS; ACIN1; ACTA1; ACTC1; ACTG2; ACTA2; ADAR; AGTPBP1; AHSA1; AKAP17A; AKAP8; ALDH2; ALKBH5; ANLN; ANXA7; AP3B1; AP3D1; APRT; ARID1A; ASF1B; ATAD2; ATAD3A; ATR; BAZ1A; BAZ1B; BAZ2A; BCAP31; BCLAF1; BLM; BPTF; BRCA1; BRD2; BRIP1; BTAF1; CASC5; CBX3; CCNK; CDC27; CDC5L; CDC73; CDCA2; CDK11B; CDK11A; CDK12; CENPF; CHD1; CHD4; CIRH1A; CRKL; CSNK2B; CSNK2B-LY6G5B-1181; CSNK2B-LY6G5B--991; CTR9; DDB1; DDX17; DDX47; DDX5; DDX54; DHX9; DKC1; DNMT1; EIF4A3; EIF5A; EIF5AL1; ERCC3; ESF1; ETF1; EWSR1; EZR; FANCI; FLNA; FOXC1; FOXK1; FUBP1; FXR1; GNAS; GNL3; GOLT1B; GTF2F2; GTF2H1; GTF2I; GTPBP4; H2AFX; HIST1H2AA; HIST2H2AB; HAT1; HEATR1; HELLS; HEXIM1; HIST1H2BN; HIST1H2BL; HIST1H2BM; HIST1H2BH; HIST2H2BF; HIST1H2BC; HIST1H2BD; H2BFS; HIST1H2BK; HIST1H2BA; HIST1H4A; HLTF; HNRNPA2B1; HNRPA2B1; HNRNPAB; HNRNPC; HNRNPD; HNRNPDL; HNRNPF; HNRNPK; HSP90AA1; HSPA8; HSPE1; HSPE1-MOB4; HTATSF1; ICE1; ILF2; ILF3; IPO5; IWS1; KDM3A; KDM3B; KHDRBS1; KHSRP; KIAA1524; KIF22; KIF4A; KMT2D; KPNA2; KPNB1; LAMB1; LEO1; LIG3; MASTL; MATR3; MCTS1; MDC1; MED1; MED12; TNRC11; MEPCE; MIF; MORC3; MPHOSPH10; MRE11A; MSH2; MSH6; MTA2; MYBBP1A; NACC1; NBN; NELFA; NIFK; NIPBL; NOC2L; NOL8; NOP2; NUP133; NUP155; NUP160; NUP205; NUP210; OGT; PABPN1; PAF1; PARP1; PAXBP1; PDCD4; PDIA3; PDIA4; PDS5A; PELP1; PHB; PHF6; PHF8; PHIP; POLA2; POLD1; POLE; POLR1A; POLR1B; POLR1E; POLR2A; POLR2B; POLR2E; POLR2L; POLR3A; PPM1G; PPP1CC; PPP2R1A; PQBP1; PRC1; PRCC; PRKDC; PRPF40A; PSIP1; PSMC5; PSME3; PTBP1; RACGAP1; RAD50; RAD9A; RAI1; RBM22; RBM25; RBM3; RBM4; RBMX; RFC1; RIF1; RNPS1; RPA1; RPA2; RPS27A; UBC; UBB; UBA52; RRN3; RRP1B; RSL1D1; RTF1; SAFB; SART3; SET; SETD1A; SETD2; SETX; SF1; SF3A1; SF3B3; SF3B4; SHMT2; SIN3A; SLC25A6; SLFN11; SMARCA4; SMARCA5; SMARCAD1; SMARCC1; SMC1A; SMC3; SNRNP70; SNW1; SPEN; SRI; SRRT; SRSF1; SRSF10; SRSF6; SRSF9; SSRP1; STAG2; STOML2; SUMO1; SUMO2; SUPT16H; SUPT5H; SUPT6H; SYMPK; TAF15; TARDBP; TDP43; TBL3; TCEB3; TCERG1; TCF20; TCP1; THRAP3; TLE1; TMPO; TOP1; TOP2A; TP53; TP53BP1; TPR; TPX2; TRA2B; TRIM27; TRIM28; TTK; TWISTNB; TXN; U2AF2; UBA3; UBE2M; UBR5; UBTF; UHRF1; USP36; USP48; USP7; UTP15; UTP20; WAPAL; WDR36; WRN; WRNIP1; XPO1; XRN2; YLPM1; YWHAB; ZMYM2; ZNF207; ZNF318; ZNF638

| rRNA processing | 4,34E-05 |
|---|---|

CIRH1A; DDX21; DDX47; DDX52; DKC1; EIF4A3; FTSJ3; HEATR1; LAS1L; MPHOSPH10; NAF1; NOL11; NOL8; NOL9; NOLC1; NOP2; NOP56; NOP58; PAPD5; PDCD11; PWP2; RPL7; RRP1; RRP1B; SKIV2L2; TBL3; UTP14A; UTP15; UTP20; UTP3; WBP11; WDR3; WDR36

| intracellular organelle | 0,00016318 |
|---|---|

AAAS; ABCD3; ACIN1; ACTA1; ACTC1; ACTG2; ACTA2; ACTR1A; ADAR; AFF4; AGTPBP1; AHSA1; AKAP17A; AKAP8; AKAP8L; ALKBH5; ANLN; ANP32E; ANXA11; ANXA7; AP3B1; AP3D1; ARID1A; ASF1B; ATAD3A; ATAD5; ATP5J2; ATP5J2-PTCD1; ATP5L; ATR; BAZ1A; BAZ1B; BAZ2A; BCAP31; BCLAF1; BLM; BMS1; BPTF; BRCA1; BRD2; BRIP1; BTAF1; CASC5; CBX3; CCNK; CCT3; CDC27; CDC5L; CDC73; CDCA2; CDK11B; CDK11A; CDK12; CENPF; CHD1; CHD4; CIRH1A; CNN3; COIL; CPSF6; CRKL; CSNK2B; CSNK2B-LY6G5B-1181; CSNK2B-LY6G5B--991; CSTF2; DDB1; DDX17; DDX18; DDX21; DDX24; DDX27; DDX42; DDX47; DDX5; DDX50; DDX52; DDX54; DHX15; DHX16; DHX37; DHX38; DHX8; DHX9; DIDO1; DKC1; DNMT1; DONSON; EIF4A3; EIF5A; EIF5AL1; ELAC2; EPB41L2; ERCC3; ERLIN2; ESF1; EWSR1; EZR; FAM98B; FANCI; FLNA; FOXC1; FOXK1; FTSJ3; FUBP1; FUS; FXR1; GANAB; GEMIN5; GNL3; GOLT1B; GTF2F2; GTF2I; GTF3C1; GTPBP4; H1FX; H2AFX; HIST1H2AA; HIST2H2AB; HAT1; HCFC1; HEATR1; HELLS; HEXIM1; HIST1H2BN; HIST1H2BL; HIST1H2BM; HIST1H2BH; HIST2H2BF; HIST1H2BC; HIST1H2BD; H2BFS; HIST1H2BK; HIST1H2BA; HIST1H4A; HLTF; HNRNPA1; HNRNPA2B1; HNRPA2B1; HNRNPA3; HNRNPAB; HNRNPC; HNRNPD; HNRNPDL; HNRNPF; HNRNPH3; HNRNPK; HNRNPL; HNRNPUL2; HNRNPUL2-BSCL2; HSP90AA1; HSPA8; HSPE1; HSPE1-MOB4; HTATSF1; HUWE1; IK; ILF2; ILF3; IPO5; ISOC1; IWS1; KDM3A; KHDRBS1; KHSRP; KIF22; KIF4A; KMT2D; LAS1L; LEO1; LIG3; LUC7L3; LYAR; MASTL; MCMBP; MDC1; MDN1; MED1; MED12; TNRC11; MKI67; MORC2; MORC3; MPHOSPH10; MRE11A; MSH2; MSH6; MYBBP1A;

| | |
|---|---|
| NACC1; NAF1; NASP; NBN; NCAPD3; NCL; NIFK; NIPBL; NOC2L; NOL11; NOL8; NOL9; NOLC1; NOP2; NOP56; NOP58; NSUN2; NUMA1; NUP133; NVL; OGT; PABPN1; PAPD5; PAPOLA; PARN; PARP1; PAXBP1; PDCD11; PDCD4; PDIA3; PDIA4; PDS5A; PEBP1; PELP1; PHB; PHF6; PHF8; PHIP; POLD1; POLE; POLR1A; POLR1B; POLR1E; POLR2A; POLR2B; POLR2E; POLR2L; PPAN-P2RY11; PPAN; PPIL4; PPM1G; PPP1CC; PPP1R10; PPP2R1A; PQBP1; PRC1; PRCC; PRKDC; PRPF4B; PRPF8; PSIP1; PSMC5; PSME3; PTBP1; PWP1; PWP2; RACGAP1; RAD50; RAD54L2; RAD9A; RAI1; RANBP3; RAVER1; RBBP6; RBM10; RBM22; RBM3; RBM4; RBMX; RBMXL1; RFC1; RIF1; RNPS1; RPA1; RPA2; RPL10; RPL19; RPL28; RPL32; RPL4; RPL7; RPN1; RPS27A; UBC; UBB; UBA52; RRN3; RRP1; RRP12; RRP1B; RRS1; RSL1D1; RTCB; RTF1; SAFB; SART3; SCAF11; SENP3; SENP3-EIF4A1; SET; SETD1A; SETD2; SETX; SF1; SF3B1; SHMT2; SIN3A; SKIV2L2; SLC25A6; SLFN11; SLU7; SMARCA4; SMARCA5; SMARCC1; SMC1A; SMC3; SMNDC1; SNRNP200; SNRNP70; SNRPE; SNW1; SRI; SRRT; SRSF1; SRSF10; SRSF2; SRSF9; SSRP1; STAG2; STOML2; SUMO1; SUMO2; SUPT16H; SUPT5H; SUPT6H; SYMPK; TAF15; TARDBP; TDP43; TBCB; TBL3; TCEB3; TCERG1; TCF20; TCOF1; TCOF1; TCP1; TEX10; THRAP3; TLE1; TMEM33; TMEM43; TMPO; TMPO; TOP1; TOP2A; TOP2B; TOPBP1; TP53; TP53BP1; TPR; TPX2; TRA2A; TRA2B; TRIM27; TRIM28; TRMT1; TSR1; TTK; TWISTNB; TXN; U2AF2; U2SURP; UBA3; UBR5; UBTF; UHRF1; URB1; URB2; USP36; USP48; USP7; UTP14A; UTP15; UTP20; UTP3; VARS; WAPAL; WBP11; WDR3; WDR36; WDR43; WDR75; WRN; WRNIP1; XPO1; XPO5; XRCC1; XRN2; YLPM1; YWHAB; ZFR; ZNF207; ZNF638 | |
| preribosome | 0,00016435 |
| FTSJ3; HEATR1; MPHOSPH10; NOP56; NOP58; PDCD11; PWP2; RRP1; RRP1B; TBL3; UTP14A; UTP15; UTP20; UTP3; WDR3; WDR36 | |
| ribonucleoprotein complex | 0,00030037 |
| AKAP17A; BRCA1; CDC5L; CPSF6; DDX5; DHX15; DHX16; DHX38; DHX8; DHX9; DKC1; EFTUD2; EIF4A3; FTSJ3; FXR1; GTF3C1; HEATR1; HNRNPA1; HNRNPA2B1; HNRPA2B1; HNRNPA3; HNRNPAB; HNRNPC; HNRNPC; HNRNPD; HNRNPDL; HNRNPF; HNRNPH3; HNRNPK; HNRNPL; HNRNPR; HSPA8; ILF2; ILF3; KHSRP; LUC7L3; MPHOSPH10; NAF1; NCL; NOP56; NOP58; PABPN1; PDCD11; POP1; PQBP1; PRPF38B; PRPF4B; PRPF8; PWP2; RBM22; RBM25; RBM4; RBMX; RBMXL1; RPL10; RPL19; RPL28; RPL32; RPL4; RPL7; RPS27A; UBC; UBB; UBA52; RRP1; RRP1B; SF1; SF3A1; SF3B1; SF3B2; SF3B3; SF3B4; SKIV2L2; SLU7; SMNDC1; SNRNP200; SNRNP70; SNRPE; SNW1; SREK1; SRRM2; SRSF1; SRSF2; TBL3; TOP1; U2AF1; U2AF2; UTP14A; UTP15; UTP20; UTP3; WDR3; WDR36; XPO1 | |
| regulation of gene expression, epigenetic | 0,00062703 |
| ARID1A; ATAD2; BAZ1B; BAZ2A; BRCA1; DNMT1; ERCC3; GNAS; GTF2H1; HAT1; HELLS; HIST1H4A; KMT2D; MTA2; PHF8; POLR1A; POLR1B; POLR1E; POLR2E; POLR2L; SIN3A; SMARCA4; SMARCA5; TRIM27; TRIM28; TWISTNB; UBTF; UHRF1 | |
| response to DNA damage stimulus | 0,00079424 |
| ATAD5; ATR; BAZ1B; BLM; BRCA1; BRIP1; CCNK; CDC5L; CHAF1A; DDB1; ERCC3; FANCI; GTF2H1; H2AFX; HIST1H2AA; HIST2H2AB; HLTF; HUWE1; KIF22; LIG3; MASTL; MCTS1; MDC1; MIF; MRE11A; MSH2; MSH6; NBN; NIPBL; PARP1; POLD1; POLE; POLR2A; POLR2B; POLR2E; POLR2L; PRKDC; PSMC5; PSME3; RAD50; RAD9A; RFC1; RIF1; RPA1; RPA2; RPS27A; UBC; UBB; UBA52; SETD2; SETX; SMARCA5; SMC1A; SMC3; SNW1; SSRP1; SUMO1; SUPT16H; TOP2A; TOPBP1; TP53; TP53BP1; TRIM28; UBR5; UHRF1; USP7; WRN; XRCC1 | |
| transcription from RNA polymerase I promoter | 0,00080673 |
| ERCC3; GTF2H1; POLR1A; POLR1B; POLR1E; POLR2E; POLR2L; RRN3; TCOF1; TWISTNB; UBTF | |
| chromatin organization | 0,00081101 |

ANP32E; ARID1A; ASF1B; BAZ1A; BAZ1B; BAZ2A; BPTF; BRD2; CASC5; CBX3; CDC73; CHAF1A; CHD1; CHD4; CTR9; DDB1; DNMT1; GATAD2B; H1FX; H2AFX; HIST1H2AA; HIST2H2AB; HAT1; HELLS; HIST1H2BN; HIST1H2BL; HIST1H2BM; HIST1H2BH; HIST2H2BF; HIST1H2BC; HIST1H2BD; H2BFS; HIST1H2BK; HIST1H2BA; HIST1H4A; HLTF; HNRNPC; HUWE1; KDM3A; KDM3B; KMT2D; LEO1; MTA2; NASP; NOC2L; OGT; PAF1; PHB; PHF8; RTF1; SAFB; SART3; SET; SETD1A; SETD2; SIN3A; SMARCA4; SMARCA5; SMARCAD1; SMARCC1; SSRP1; SUPT16H; SUPT5H; SUPT6H; TOP1; TP53; UBTF; UHRF1; USP7; UTP3

| nuclear transport | 0,00082122 |
|---|---|

AAAS; ADAR; ALKBH5; CPSF1; DHX38; EIF5A; EIF5AL1; HNRNPA1; IPO5; KPNA2; KPNB1; MYBBP1A; NOC2L; NOP58; NUP133; NUP155; NUP160; NUP205; PABPN1; PDIA3; PHIP; PPP1R10; RNPS1; SET; SRSF1; SRSF10; SRSF2; SRSF6; SRSF9; THOC2; TPR; U2AF1; U2AF2; WRN; XPO1; XPO5

| cellular metabolic process | 0,00082751 |
|---|---|

AAAS; ABCD3; ACIN1; ADAR; AFF4; AGTPBP1; AKAP17A; AKAP8; ALDH2; ALKBH5; ANXA7; APRT; ARID1A; ASF1B; ATAD2; ATP5J2; ATP5J2-PTCD1; ATP5L; ATR; BAZ1A; BAZ1B; BAZ2A; BCLAF1; BLM; BPTF; BRCA1; BRD2; BRIP1; CBX3; CCNK; CCT2; CCT3; CDC27; CDC5L; CDC73; CDK11B; CDK11A; CDK12; CENPF; CHAF1A; CHD1; CHD4; CIRH1A; COASY; CPSF1; CPSF6; CRKL; CSNK2B; CSNK2B-LY6G5B-1181; CSNK2B-LY6G5B--991; CSTF2; CTR9; DDB1; DDX17; DDX18; DDX21; DDX24; DDX27; DDX42; DDX47; DDX5; DDX50; DDX52; DDX54; DHX15; DHX16; DHX37; DHX38; DHX8; DHX9; DIDO1; DKC1; DNMT1; EEF1B2; EFTUD2; EIF4A3; EIF5A; EIF5AL1; ELAC2; ERCC3; ERLIN2; ESF1; ETF1; EWSR1; EZR; FANCI; FARSA; FLNA; FTSJ3; FUBP1; FUS; GANAB; GATAD2B; GEMIN5; GINS3; GNAS; GTF2F2; GTF2H1; GTF2I; GTF3C1; GTF3C3; H2AFX; HIST1H2AA; HIST2H2AB; HAT1; HEATR1; HELLS; HEXIM1; HIST1H4A; HLTF; HNRNPA1; HNRNPA2B1; HNRPA2B1; HNRNPA3; HNRNPC; HNRNPD; HNRNPDL; HNRNPF; HNRNPH3; HNRNPK; HNRNPL; HNRNPR; HSP90AA1; HSPA8; HSPE1; HSPE1-MOB4; HTATSF1; HUWE1; ICE1; ILF2; ILF3; INTS9; IWS1; KDM3A; KDM3B; KHDRBS1; KHSRP; KIAA1429; KIF22; KMT2D; KPNA2; KPNB1; LAS1L; LEO1; LIG3; LUC7L3; MCMBP; MCTS1; MDC1; MED1; MED12; TNRC11; MEPCE; MIF; MORC2; MORC3; MPHOSPH10; MRE11A; MSH2; MSH6; MTA2; MYBBP1A; NAF1; NASP; NBN; NCAPD3; NELFA; NIFK; NOC2L; NOL11; NOL8; NOL9; NOLC1; NOP2; NOP56; NOP58; NSUN2; NT5DC1; NUP133; NUP155; NUP160; NUP205; NUP210; OGT; PABPN1; PAF1; PAPD5; PAPOLA; PARN; PARP1; PAXBP1; PDCD11; PDIA3; PDIA4; PELP1; PHB; PHF6; PHF8; POLA2; POLD1; POLE; POLR1A; POLR1B; POLR1E; POLR2A; POLR2B; POLR2E; POLR2L; POLR3A; POP1; PPIL4; PPM1G; PPP1CC; PPP1R10; PPP2R1A; PPP6C; PQBP1; PRKDC; PRPF38B; PRPF40A; PRPF4B; PRPF8; PRPS1; PSIP1; PSMC5; PSME3; PTBP1; PUS7; PWP1; PWP2; RAD50; RAD9A; RBBP6; RBM10; RBM22; RBM25; RBM3; RBM4; RBMX; RBMXL1; RFC1; RNPS1; RPA1; RPA2; RPL10; RPL19; RPL28; RPL32; RPL4; RPL7; RPN1; RPS27A; UBC; UBB; UBA52; RRN3; RRP1; RRP1B; RTCB; RTF1; SAFB; SART3; SCAF11; SET; SETD1A; SETD2; SETX; SF1; SF3A1; SF3B1; SF3B2; SF3B3; SF3B4; SHMT2; SIN3A; SKIV2L2; SLC25A6; SLU7; SMARCA4; SMARCA5; SMARCAD1; SMARCC1; SMC1A; SMC3; SMNDC1; SNRNP200; SNRNP70; SNRPE; SNW1; SREK1; SRRM2; SRRT; SRSF1; SRSF10; SRSF2; SRSF6; SRSF9; SSRP1; STAG2; STOML2; SUGP2; SUMO1; SUMO2; SUPT16H; SUPT5H; SUPT6H; SYMPK; TARDBP; TDP43; TBCB; TBL3; TCEB3; TCERG1; TCF20; TCOF1; TCP1; THOC2; THRAP3; TK1; TOP1; TOP2A; TOP2B; TOPBP1; TP53; TP53BP1; TPR; TRA2A; TRA2B; TRIM27; TRIM28; TRMT1L; TTK; TWISTNB; TXN; U2AF1; U2AF2; U2SURP; UBA3; UBE2M; UBR5; UBTF; UHRF1; USP36; USP48; USP7; UTP14A; UTP15; UTP20; UTP3; VARS; WBP11; WDR3; WDR36; WRN; WRNIP1; XRCC1; XRN2; YWHAB; ZMYM2; ZNF318; ZNF638

| transcription from RNA polymerase II promoter | 0,0011707 |
|---|---|

AFF4; CCNK; CPSF1; CSTF2; DDX21; DHX38; ERCC3; FLNA; FUBP1; GTF2F2; GTF2H1; GTF2I; ICE1; MED1; NELFA; PABPN1; PAPOLA; PARP1; POLR2A; POLR2B; POLR2E; POLR2L; PSMC5;

RBMX; RNPS1; SNRPE; SRSF1; SRSF2; SRSF6; SRSF9; SSRP1; SUPT16H; SUPT5H; TARDBP; TDP43; TCEB3; TCERG1; TP53BP1; U2AF1; U2AF2

| RNA polymerase complex | 0,0012276 |
|---|---|
| POLR1A; POLR1B; POLR1E; POLR2A; POLR2B; POLR2E; POLR2L; POLR3A; TWISTNB | |
| regulation of DNA metabolic process | 0,0013928 |
| ATR; BLM; BRCA1; GTPBP4; H2AFX; HIST1H2AA; HIST2H2AB; KPNA2; LIG3; MRE11A; MSH2; MSH6; NBN; PDS5A; PPP2R1A; RAD50; RPA2; SMARCAD1; SMC1A; SMC3; STAG2; STOML2; SUPT6H; TP53; TRIM28; UBR5; WAPAL; WRNIP1; YLPM1 | |
| transcription elongation, DNA-dependent | 0,001517 |
| ERCC3; GTF2F2; GTF2H1; NELFA; POLR1A; POLR1B; POLR1E; POLR2A; POLR2B; POLR2E; POLR2L; POLR3A; SETD2; SSRP1; SUPT16H; SUPT5H; TCEB3; TWISTNB; UBTF | |
| nucleocytoplasmic transport | 0,0019654 |
| AAAS; ADAR; ALKBH5; CPSF1; DHX38; EIF5A; EIF5AL1; HNRNPA1; IPO5; KPNA2; KPNB1; MYBBP1A; NOP58; NUP133; NUP155; NUP160; NUP205; PABPN1; PDIA3; PHIP; PPP1R10; RNPS1; SET; SRSF1; SRSF10; SRSF2; SRSF6; SRSF9; THOC2; TPR; U2AF1; U2AF2; XPO1; XPO5 | |
| DNA replication | 0,0023603 |
| ATR; BAZ1A; BLM; BRCA1; CHAF1A; DKC1; GINS3; LIG3; MCMBP; MRE11A; NASP; NOL8; POLA2; POLD1; POLE; RAD50; RAD9A; RFC1; RPA1; RPA2; SET; SIN3A; SSRP1; SUPT16H; TOP1; WRN; WRNIP1 | |
| chromatin modification | 0,0028606 |
| ANP32E; ARID1A; ASF1B; BAZ1A; BAZ1B; BAZ2A; BPTF; BRD2; CASC5; CBX3; CDC73; CHD1; CHD4; CTR9; DDB1; DNMT1; GATAD2B; HAT1; HELLS; HIST1H2BN; HIST1H2BL; HIST1H2BM; HIST1H2BH; HIST2H2BF; HIST1H2BC; HIST1H2BD; H2BFS; HIST1H2BK; HIST1H2BA; HIST1H4A; HLTF; HNRNPC; HUWE1; KDM3A; KDM3B; KMT2D; LEO1; MTA2; NASP; OGT; PAF1; PHB; PHF8; RTF1; SIN3A; SMARCA4; SMARCA5; SMARCAD1; SMARCC1; SSRP1; SUPT5H; SUPT6H; TOP1; UBTF; UHRF1; USP7; UTP3 | |
| RNA modification | 0,0032984 |
| ADAR; ALKBH5; DKC1; FTSJ3; KIAA1429; MEPCE; NAF1; NOP56; NOP58; NSUN2; PARN; PUS7 | |
| metabolic process | 0,0036896 |
| AAAS; ABCD3; ACIN1; ACTA1; ACTC1; ACTG2; ACTA2; ADAR; AFF4; AGTPBP1; AKAP17A; AKAP8; ALDH2; ALKBH5; ANXA7; AP3D1; APRT; ARID1A; ASF1B; ATAD2; ATP5J2; ATP5J2-PTCD1; ATP5L; ATR; BAZ1A; BAZ1B; BAZ2A; BCLAF1; BLM; BPTF; BRCA1; BRD2; BRIP1; CBX3; CCNK; CCT2; CCT3; CDC27; CDC5L; CDC73; CDK11B; CDK11A; CDK12; CENPF; CHAF1A; CHD1; CHD4; CIRH1A; COASY; CPSF1; CPSF6; CRKL; CSNK2B; CSNK2B-LY6G5B-1181; CSNK2B-LY6G5B--991; CSTF2; CTR9; DDB1; DDX17; DDX18; DDX21; DDX24; DDX27; DDX42; DDX47; DDX5; DDX50; DDX52; DDX54; DHX15; DHX16; DHX37; DHX38; DHX8; DHX9; DIDO1; DKC1; DNMT1; EEF1B2; EFTUD2; EIF4A3; EIF5A; EIF5AL1; ELAC2; ERCC3; ERLIN2; ESF1; ETF1; EWSR1; EZR; FANCI; FARSA; FLNA; FOXC1; FTSJ3; FUBP1; FUS; GANAB; GATAD2B; GEMIN5; GINS3; GNAS; GNL3; GOLT1B; GTF2F2; GTF2H1; GTF2I; GTF3C1; GTF3C3; GTPBP4; H2AFX; HIST1H2AA; HIST2H2AB; HAT1; HEATR1; HELLS; HEXIM1; HIST1H2BN; HIST1H2BL; HIST1H2BM; HIST1H2BH; HIST2H2BF; HIST1H2BC; HIST1H2BD; H2BFS; HIST1H2BK; HIST1H2BA; HIST1H4A; HLTF; HNRNPA1; HNRNPA2B1; HNRPA2B1; HNRNPA3; HNRNPC; HNRNPD; HNRNPDL; HNRNPF; HNRNPH3; HNRNPK; HNRNPL; HNRNPR; HSP90AA1; HSPA8; HSPE1; HSPE1-MOB4; HTATSF1; HUWE1; ICE1; ILF2; ILF3; INTS9; IWS1; KDM3A; KDM3B; KHDRBS1; KHSRP; KIAA1429; KIF22; KIF4A; KMT2D; KPNA2; KPNB1; LAS1L; LEO1; LIG3; | |

LUC7L3; MCMBP; MCTS1; MDC1; MDN1; MED1; MED12; TNRC11; MEPCE; MIF; MORC2; MORC3; MPHOSPH10; MRE11A; MSH2; MSH6; MTA2; MYBBP1A; NAF1; NASP; NBN; NCAPD3; NELFA; NIFK; NOC2L; NOL11; NOL8; NOL9; NOLC1; NOP2; NOP56; NOP58; NSUN2; NT5DC1; NUP133; NUP155; NUP160; NUP205; NUP210; OGT; PABPN1; PAF1; PAPD5; PAPOLA; PARN; PARP1; PAXBP1; PDCD11; PDIA3; PDIA4; PDS5A; PEF1; PELP1; PHB; PHF6; PHF8; POLA2; POLD1; POLE; POLR1A; POLR1B; POLR1E; POLR2A; POLR2B; POLR2E; POLR2L; POLR3A; POP1; PPIL4; PPM1G; PPP1CC; PPP1R10; PPP2R1A; PPP6C; PQBP1; PRKDC; PRPF38B; PRPF40A; PRPF4B; PRPF8; PRPS1; PSIP1; PSMC5; PSME3; PTBP1; PUS7; PWP1; PWP2; RAD50; RAD9A; RBBP6; RBM10; RBM22; RBM25; RBM3; RBM4; RBMX; RBMXL1; RFC1; RNPS1; RPA1; RPA2; RPL10; RPL19; RPL28; RPL32; RPL4; RPL7; RPN1; RPS27A; UBC; UBB; UBA52; RRN3; RRP1; RRP1B; RTCB; RTF1; SAFB; SART3; SCAF11; SET; SETD1A; SETD2; SETX; SF1; SF3A1; SF3B1; SF3B2; SF3B3; SF3B4; SHMT2; SIN3A; SKIV2L2; SLC25A6; SLU7; SMARCA4; SMARCA5; SMARCAD1; SMARCC1; SMC1A; SMC3; SMNDC1; SNRNP200; SNRNP70; SNRPE; SNW1; SREK1; SRI; SRRM2; SRRT; SRSF1; SRSF10; SRSF2; SRSF6; SRSF9; SSRP1; STAG2; STOML2; SUGP2; SUMO1; SUMO2; SUPT16H; SUPT5H; SUPT6H; SYMPK; TARDBP; TDP43; TBCB; TBL3; TCEB3; TCERG1; TCF20; TCOF1; TCP1; THOC2; THRAP3; TK1; TOP1; TOP2A; TOP2B; TOPBP1; TP53; TP53BP1; TPR; TRA2A; TRA2B; TRIM27; TRIM28; TRMT1L; TTK; TWISTNB; TXN; U2AF1; U2AF2; U2SURP; UBA3; UBE2M; UBR5; UBTF; UHRF1; USP36; USP48; USP7; UTP14A; UTP15; UTP20; UTP3; VARS; WAPAL; WBP11; WDR3; WDR36; WRN; WRNIP1; XPO1; XPO5; XRCC1; XRN2; YLPM1; YWHAB; ZMYM2; ZNF318; ZNF638

| telomere organization | 0,0048442 |
|---|---|
| BLM; DKC1; HIST1H4A; MRE11A; NBN; PARP1; POLA2; POLD1; POLE; PRKDC; RAD50; RFC1; RPA1; RPA2; WRN | |
| nucleobase-containing compound transport | 0,0054168 |
| AAAS; ALKBH5; CPSF1; DHX38; EIF4A3; EIF5A; EIF5AL1; HNRNPA1; HNRNPA2B1; HNRPA2B1; IWS1; KHSRP; NUP133; NUP155; NUP160; NUP205; NUP210; PABPN1; RNPS1; SLC25A6; SRSF1; SRSF10; SRSF2; SRSF6; SRSF9; SUPT6H; THOC2; TPR; U2AF1; U2AF2; XPO1; XPO5 | |
| DNA geometric change | 0,010138 |
| BLM; BRIP1; CHD1; CHD4; DHX9; ERCC3; MRE11A; NBN; RAD50; SETX; TOP2A; TOP2B; WRN | |
| meiotic cell cycle | 0,010821 |
| H2AFX; HIST1H2AA; HIST2H2AB; MRE11A; NBN; NUMA1; STAG2; ZNF318 | |
| DNA recombination | 0,011848 |
| ATR; BLM; BRCA1; H2AFX; HIST1H2AA; HIST2H2AB; LIG3; MDC1; MRE11A; MSH2; MSH6; NBN; POLA2; POLD1; POLE; PRKDC; RAD50; RFC1; RPA1; RPA2; SETX; TOP2A; TOP2B; TP53BP1; WRN | |
| transcription initiation, DNA-dependent | 0,015047 |
| ERCC3; GTF2F2; GTF2H1; GTF2I; MED1; MED12; TNRC11; PARP1; POLR1A; POLR1B; POLR1E; POLR2A; POLR2B; POLR2E; POLR2L; RPS27A; UBC; UBB; UBA52; RRN3; RTF1; SMARCA5; SNW1; THRAP3; TP53; TRIM28; TWISTNB; TXN; UBTF; YWHAB | |
| protein modification by small protein conjugation or removal | 0,015917 |
| AAAS; BLM; BRCA1; CDC27; CDC73; CTR9; DDB1; HLTF; HUWE1; LEO1; MDC1; NUP133; NUP155; NUP160; NUP205; NUP210; PAF1; PARP1; PSMC5; PSME3; RBBP6; RPA1; RPS27A; UBC; UBB; UBA52; SMC1A; SMC3; STAG2; SUMO1; SUMO2; TCEB3; TPR; TRIM27; TRIM28; UBA3; UBE2M; UBR5; UHRF1; USP36; USP48; USP7; WRN | |
| transcription elongation factor complex | 0,01601 |
| AFF4; CDC73; CTR9; ICE1; LEO1; NELFA; PAF1; RTF1; SUPT5H; TCEB3 | |

| cytoplasm | 0,017769 |
|---|---|

AAAS; ACIN1; ACTA1; ACTC1; ACTG2; ACTA2; ACTR1A; ADAR; AGTPBP1; AHSA1; AKAP17A; AKAP8L; ANP32E; ANXA11; APRT; BCAP31; BCLAF1; BLM; BPTF; BRCA1; BRD2; BRIP1; C15orf39; C9orf78; CASC5; CCT2; CCT3; CDC27; CDC5L; CDCA2; CDK11B; CDK11A; CENPF; CHD1; CHD4; CNN3; COASY; COIL; CSNK2B; CSNK2B-LY6G5B-1181; CSNK2B-LY6G5B--991; DDB1; DDX17; DDX18; DDX21; DDX24; DDX27; DDX42; DDX47; DDX5; DDX50; DDX52; DDX54; DHX15; DHX16; DHX37; DHX38; DHX8; DHX9; DIDO1; DKC1; EEF1B2; EFTUD2; EIF4A3; EIF5A; EIF5AL1; ERLIN2; ETF1; EWSR1; EZR; FAM98B; FANCI; FARSA; FLNA; FOXC1; FUS; FXR1; GANAB; GEMIN5; GNAS; GTF2I; GTPBP4; HAT1; HEXIM1; HLTF; HNRNPA1; HNRNPA2B1; HNRPA2B1; HNRNPAB; HNRNPDL; HNRNPK; HNRNPL; HNRNPR; HSP90AA1; HSPA8; HSPE1; HSPE1-MOB4; HUWE1; ILF2; ILF3; INTS9; IPO5; IWS1; KDM3A; KHDRBS1; KIAA1524; KIF22; KIF4A; KPNA2; KPNB1; LAS1L; MASTL; MCMBP; MCTS1; MDN1; MIF; MKI67; MORC2; MSH6; MYBBP1A; NACC1; NAF1; NASP; NELFA; NIFK; NOL8; NOLC1; NOP56; NOP58; NSUN2; NUMA1; NUP205; NVL; OGT; PABPN1; PAF1; PAPD5; PAPOLA; PARN; PDCD4; PDIA3; PDIA4; PEBP1; PEF1; PELP1; PHB; POLA2; POLD1; POLE; POLR1A; POLR1B; PPM1G; PPP1CC; PQBP1; PRC1; PRPF40A; PSMC5; PSME3; PWP2; RACGAP1; RAD9A; RANBP3; RAVER1; RBBP6; RBM22; RBM25; RBM3; RBM4; RIF1; RNPS1; RPL19; RPL28; RPL32; RPL4; RPL7; RPN1; RPS10; RPS10-NUDT3; RPS27A; UBC; UBB; UBA52; RTCB; SART3; SENP3; SENP3-EIF4A1; SET; SETX; SIN3A; SLFN11; SLU7; SMC1A; SMC3; SMNDC1; SRI; SRRT; SRSF1; SRSF10; SRSF2; SSRP1; SUMO1; SYMPK; TAF15; TBCB; TCEB3; TCOF1; TCOF1; TCP1; TEX10; TLE1; TMEM33; TMPO; TOP2A; TOP2B; TOPBP1; TP53; TP53BP1; TPR; TPX2; TRIM27; TXN; UBE2M; UBR5; USP48; UTP15; UTP20; WAPAL; WBP11; XPO1; XPO5; YLPM1; YWHAB; ZFR; ZNF318; ZNF638

| nuclear envelope | 0,018236 |
|---|---|

AAAS; ANXA11; ANXA7; CBX3; CENPF; KPNB1; NUP133; NUP155; NUP160; NUP205; NUP210; PARP1; RTCB; TMPO; TMPO; TPR; XPO1

| macromolecular complex | 0,018977 |
|---|---|

AAAS; ACTA1; ACTC1; ACTG2; ACTA2; ACTR1A; AFF4; AKAP17A; ANP32E; AP3B1; AP3D1; ARID1A; ASF1B; ATP5J2; ATP5J2-PTCD1; ATP5L; BAZ1A; BAZ2A; BPTF; BRCA1; CASC5; CCT2; CCT3; CDC27; CDC5L; CDC73; CDK12; CENPF; CHAF1A; CHD4; CIRH1A; CPSF1; CPSF6; CSNK2B; CSNK2B-LY6G5B-1181; CSNK2B-LY6G5B--991; CSTF2; CTR9; DDB1; DDX5; DHX15; DHX16; DHX38; DHX8; DHX9; DKC1; EEF1B2; EFTUD2; EIF4A3; EIF5A; EIF5AL1; EMC1; ERCC3; ERLIN2; ETF1; EZR; FAM98B; FARSA; FLNA; FTSJ3; FXR1; GANAB; GATAD2B; GEMIN5; GNAS; GTF2F2; GTF2H1; GTF3C1; GTF3C3; H1FX; H2AFX; HIST1H2AA; HIST2H2AB; HAT1; HEATR1; HIST1H2BN; HIST1H2BL; HIST1H2BM; HIST1H2BH; HIST2H2BF; HIST1H2BC; HIST1H2BD; H2BFS; HIST1H2BK; HIST1H2BA; HIST1H4A; HNRNPA1; HNRNPA2B1; HNRPA2B1; HNRNPA3; HNRNPAB; HNRNPC; HNRNPC; HNRNPD; HNRNPDL; HNRNPF; HNRNPH3; HNRNPK; HNRNPL; HNRNPR; HSP90AA1; HSPA8; ICE1; ILF2; ILF3; INTS9; IPO5; KHDRBS1; KHSRP; KIF22; KIF4A; KMT2D; KPNB1; LAMB1; LAS1L; LEO1; LIG3; LUC7L3; MED1; MED12; TNRC11; MPHOSPH10; MRE11A; MSH2; MSH6; MTA2; MYBBP1A; NAF1; NASP; NBN; NCAPD3; NCL; NELFA; NIPBL; NOL11; NOP56; NOP58; NUMA1; NUP133; NUP155; NUP160; NUP205; NUP210; OGT; PABPN1; PAF1; PARP1; PDCD11; PELP1; PHF6; POLA2; POLD1; POLE; POLR1A; POLR1B; POLR1E; POLR2A; POLR2B; POLR2E; POLR2L; POLR3A; POP1; PPP1CC; PPP1R10; PPP2R1A; PQBP1; PRC1; PRKDC; PRPF38B; PRPF4B; PRPF8; PRPS1; PSMC5; PSME3; PWP2; RACGAP1; RAD50; RAD9A; RBM22; RBM25; RBM4; RBMX; RBMXL1; RFC1; RPA1; RPA2; RPL10; RPL19; RPL28; RPL32; RPL4; RPL7; RPN1; RPRD2; RPS27A; UBC; UBB; UBA52; RRP1; RRP1B; RTCB; RTF1; SENP3; SENP3-EIF4A1; SET; SETD1A; SF1; SF3A1; SF3B1; SF3B2; SF3B3; SF3B4; SIN3A; SKIV2L2; SLC25A6; SLU7; SMARCA4; SMARCA5; SMARCC1; SMC1A; SMC3; SMNDC1; SNRNP200; SNRNP70; SNRPE; SNW1; SPEN; SREK1; SRRM2; SRSF1; SRSF2; SUMO1; SUPT5H; TBCB; TBL3; TCEB3; TCP1; TEX10; THOC2; THRAP3; TLE1; TMEM33; TOP1; TOP2A; TOP2B; TP53; TP53BP1; TPR; TPX2; TWISTNB; U2AF1; U2AF2; UTP14A; UTP15; UTP20; UTP3; WAPAL; WDR3; WDR36; XPO1; XPO5; YWHAB; ZNF207

| cell cycle | 0,021084 |
|---|---|
| AAAS; ACTR1A; ANLN; ANXA11; ATR; BLM; BRCA1; BRIP1; CASC5; CCNK; CDC27; CDC73; CENPF; CHAF1A; CSNK2B; CSNK2B-LY6G5B-1181; CSNK2B-LY6G5B--991; FANCI; FOXC1; GTF2H1; GTPBP4; H2AFX; HIST1H2AA; HIST2H2AB; HEXIM1; HIST1H4A; HSP90AA1; HSPA8; KHDRBS1; LIG3; MASTL; MCTS1; MDC1; MED1; MIF; MRE11A; MSH2; MYBBP1A; NASP; NBN; NCAPD3; NIPBL; NOLC1; NUMA1; NUP133; NUP155; NUP160; NUP205; NUP210; PDCD4; PDS5A; PHF8; PHIP; POLA2; POLD1; POLE; PPM1G; PPP1CC; PPP2R1A; PRCC; PRPF40A; PSMC5; PSME3; RAD9A; RFC1; RIF1; RPA1; RPA2; RPS27A; UBC; UBB; UBA52; SET; SIN3A; SLFN11; SMARCA4; SMC1A; SMC3; STAG2; SUPT5H; TARDBP; TDP43; TOP2A; TP53; TPR; TPX2; TTK; UBA3; UHRF1; WAPAL; XPO1; ZNF207; ZNF318 | |
| nuclear chromosome | 0,022154 |
| BAZ1A; BLM; BRCA1; H2AFX; HIST1H2AA; HIST2H2AB; HIST1H4A; MSH2; MSH6; NIFK; RRS1; SETX; SMC1A; TOPBP1 | |
| biosynthetic process | 0,022185 |
| AAAS; AFF4; ALDH2; AP3D1; APRT; ARID1A; ASF1B; ATAD2; ATP5J2; ATP5J2-PTCD1; ATP5L; ATR; BAZ1A; BAZ1B; BAZ2A; BCLAF1; BLM; BPTF; BRCA1; BRD2; CBX3; CCNK; CDC5L; CDC73; CENPF; CHAF1A; CHD1; CHD4; CIRH1A; COASY; CPSF1; CSTF2; CTR9; DDB1; DDX17; DDX21; DDX5; DDX54; DHX38; DIDO1; DKC1; DNMT1; EEF1B2; EIF5A; EIF5AL1; ERCC3; ESF1; ETF1; EWSR1; FLNA; FUBP1; GATAD2B; GINS3; GNAS; GTF2F2; GTF2H1; GTF2I; GTF3C1; GTF3C3; HELLS; HEXIM1; HLTF; HNRNPD; HNRNPDL; HSPA8; HTATSF1; ICE1; ILF2; ILF3; IWS1; KDM3A; KDM3B; KHDRBS1; KHSRP; KMT2D; LEO1; LIG3; MCMBP; MCTS1; MED1; MED12; TNRC11; MIF; MRE11A; MYBBP1A; NASP; NELFA; NIFK; NOC2L; NOL11; NOL8; NUP133; NUP155; NUP160; NUP205; NUP210; PABPN1; PAF1; PAPOLA; PARP1; PAXBP1; PELP1; PHB; PHF6; PHF8; POLA2; POLD1; POLE; POLR1A; POLR1B; POLR1E; POLR2A; POLR2B; POLR2E; POLR2L; POLR3A; PPP1R10; PQBP1; PRPS1; PSIP1; PSMC5; PWP1; RAD50; RAD9A; RBM3; RBMX; RFC1; RNPS1; RPA1; RPA2; RPL10; RPL19; RPL28; RPL32; RPL4; RPL7; RPN1; RPS27A; UBC; UBB; UBA52; RRN3; RTF1; SAFB; SET; SETD1A; SETD2; SETX; SF1; SHMT2; SIN3A; SMARCA5; SMARCC1; SNRPE; SNW1; SRRT; SRSF1; SRSF2; SRSF6; SRSF9; SSRP1; STOML2; SUPT16H; SUPT5H; SUPT6H; TARDBP; TDP43; TCEB3; TCERG1; TCF20; TCOF1; THRAP3; TK1; TOP1; TP53; TP53BP1; TPR; TRIM27; TRIM28; TWISTNB; TXN; U2AF1; U2AF2; UBTF; UHRF1; WRN; WRNIP1; XRN2; YWHAB; ZMYM2; ZNF318; ZNF638 | |
| organelle | 0,028599 |
| AAAS; ABCD3; ACIN1; ACTA1; ACTC1; ACTG2; ACTA2; ACTR1A; ADAR; AFF4; AGTPBP1; AHSA1; AKAP17A; AKAP8; AKAP8L; ALDH2; ALKBH5; ANLN; ANP32E; ANXA11; ANXA7; AP3B1; AP3D1; APRT; ARID1A; ASF1B; ATAD2; ATAD3A; ATAD5; ATP5J2; ATP5J2-PTCD1; ATP5L; ATR; BAZ1A; BAZ1B; BAZ2A; BCAP31; BCLAF1; BLM; BMS1; BOLA2; BOLA2B; BPTF; BRCA1; BRD2; BRIP1; BTAF1; CASC5; CBX3; CCNK; CCT2; CCT3; CDC27; CDC5L; CDC73; CDCA2; CDK11B; CDK11A; CDK12; CENPF; CHD1; CHD4; CIRH1A; CNN3; COASY; COIL; CPSF6; CRKL; CSNK2B; CSNK2B-LY6G5B-1181; CSNK2B-LY6G5B--991; CSTF2; CUTA; DDB1; DDX17; DDX18; DDX21; DDX24; DDX27; DDX42; DDX47; DDX5; DDX50; DDX52; DDX54; DHX15; DHX16; DHX37; DHX38; DHX8; DHX9; DIDO1; DKC1; DNMT1; DONSON; EIF4A3; EIF5A; EIF5AL1; ELAC2; EPB41L2; ERCC3; ERLIN2; ESF1; EWSR1; EZR; FAM98B; FANCI; FLNA; FOXC1; FOXK1; FTSJ3; FUBP1; FUS; FXR1; GANAB; GEMIN5; GNAS; GNL3; GOLT1B; GTF2F2; GTF2I; GTF3C1; GTPBP4; H1FX; H2AFX; HIST1H2AA; HIST2H2AB; HAT1; HCFC1; HEATR1; HELLS; HEXIM1; HIST1H2BN; HIST1H2BL; HIST1H2BM; HIST1H2BH; HIST2H2BF; HIST1H2BC; HIST1H2BD; H2BFS; HIST1H2BK; HIST1H2BA; HIST1H4A; HLTF; HNRNPA1; HNRNPA2B1; HNRPA2B1; HNRNPA3; HNRNPAB; HNRNPC; HNRNPD; HNRNPDL; HNRNPF; HNRNPH3; HNRNPK; HNRNPL; HNRNPUL2; HNRNPUL2-BSCL2; HSP90AA1; HSPA8; HSPE1; HSPE1-MOB4; HTATSF1; HUWE1; IK; ILF2; ILF3; IPO5; ISOC1; IWS1; KDM3A; KHDRBS1; KHSRP; KIF22; KIF4A; KMT2D; KPNB1; LAMB1; LAS1L; LEO1; LIG3; LUC7L3; LYAR; MASTL; MCMBP; MDC1; MDN1; MED1; MED12; | |

TNRC11; MIF; MKI67; MORC2; MORC3; MPHOSPH10; MRE11A; MSH2; MSH6; MYBBP1A; NACC1; NAF1; NASP; NBN; NCAPD3; NCL; NIFK; NIPBL; NOC2L; NOL11; NOL8; NOL9; NOLC1; NOP2; NOP56; NOP58; NSUN2; NUMA1; NUP133; NVL; OGT; PABPN1; PAPD5; PAPOLA; PARN; PARP1; PAXBP1; PDCD11; PDCD4; PDIA3; PDIA4; PDS5A; PEBP1; PEF1; PELP1; PHB; PHF6; PHF8; PHIP; POLD1; POLE; POLR1A; POLR1B; POLR1E; POLR2A; POLR2B; POLR2E; POLR2L; PPAN-P2RY11; PPAN; PPIL4; PPM1G; PPP1CC; PPP1R10; PPP2R1A; PQBP1; PRC1; PRCC; PRKDC; PRPF4B; PRPF8; PSIP1; PSMC5; PSME3; PTBP1; PWP1; PWP2; RACGAP1; RAD50; RAD54L2; RAD9A; RAI1; RANBP3; RAVER1; RBBP6; RBM10; RBM22; RBM3; RBM4; RBMX; RBMXL1; RFC1; RIF1; RNPS1; RPA1; RPA2; RPL10; RPL19; RPL28; RPL32; RPL4; RPL7; RPN1; RPS27A; UBC; UBB; UBA52; RRN3; RRP1; RRP12; RRP1B; RRS1; RSL1D1; RTCB; RTF1; SAFB; SART3; SCAF11; SENP3; SENP3-EIF4A1; SET; SETD1A; SETD2; SETX; SF1; SF3B1; SHMT2; SIN3A; SKIV2L2; SLC25A6; SLFN11; SLU7; SMARCA4; SMARCA5; SMARCC1; SMC1A; SMC3; SMNDC1; SNRNP200; SNRNP70; SNRPE; SNW1; SPEN; SRI; SRRT; SRSF1; SRSF10; SRSF2; SRSF9; SSRP1; STAG2; STOML2; SUMO1; SUMO2; SUPT16H; SUPT5H; SUPT6H; SYMPK; TAF15; TAGLN2; TARDBP; TDP43; TBCB; TBL3; TCEB3; TCERG1; TCF20; TCOF1; TCOF1; TCP1; TEX10; THRAP3; TLE1; TMEM33; TMEM43; TMPO; TMPO; TOP1; TOP2A; TOP2B; TOPBP1; TP53; TP53BP1; TPR; TPX2; TRA2A; TRA2B; TRIM27; TRIM28; TRMT1; TSR1; TTK; TWISTNB; TXN; U2AF2; U2SURP; UBA3; UBE2M; UBR5; UBTF; UHRF1; URB1; URB2; USP36; USP48; USP7; UTP14A; UTP15; UTP20; UTP3; VARS; WAPAL; WBP11; WDR3; WDR36; WDR43; WDR75; WRN; WRNIP1; XPO1; XPO5; XRCC1; XRN2; YLPM1; YWHAB; ZFR; ZNF207; ZNF638

| | |
|---|---|
| cellular macromolecular complex assembly | 0,033288 |

ADAR; ANLN; ASF1B; BMS1; BRD2; CASC5; CCT2; CENPF; CHAF1A; CSNK2B; CSNK2B-LY6G5B-1181; CSNK2B-LY6G5B--991; GEMIN5; H1FX; H2AFX; HIST1H2AA; HIST2H2AB; HAT1; HIST1H2BN; HIST1H2BL; HIST1H2BM; HIST1H2BH; HIST2H2BF; HIST1H2BC; HIST1H2BD; H2BFS; HIST1H2BK; HIST1H2BA; HIST1H4A; HSP90AA1; KIF4A; LUC7L3; MCTS1; NASP; NUP205; POLR1E; PRPF8; PSIP1; PWP2; RACGAP1; RBMX; RPA1; RPA2; SART3; SCAF11; SET; SETX; SF1; SF3A1; SLU7; SMARCA5; SMNDC1; SNRNP200; SNRPE; SRSF1; SRSF10; SRSF6; SRSF9; TCP1; TMEM33; TSR1

| | |
|---|---|
| nucleus organization | 0,03879 |

KDM3A; NOLC1; NUMA1; NUP133; NUP205; POLR1B; RRN3; SUMO1; TMEM33; TPR

# Data Availability

The mass spectrometry proteomics data for the proximity dependent biotinylation experiments, and the chromatin mass spectrometry (CHROMASS) experiments have been deposited to the ProteomeXchange Consortium via the PRIDE (Vizcaino et al., 2016) partner repository (https://www.ebi.ac.uk/pride/archive/) with the dataset identifiers PXD040000 (PDB experiments) and PXD040024 (CHROMASS experiments).

# Acknowledgments

First, I want to thank my supervisor Prof. Dr. Dominik Boos for the opportunity to work on this interesting topic of replication initiation. I am especially grateful for his support, constant feedback, and the possibility to ask him questions. Whenever it was needed, his door was open.

I sincerely thank Prof. Dr. Hemmo Meyer for reviewing my thesis.

A special thanks goes to Bilal Tetik and Dr. Matthew Day who provided the biochemical binding studies and the structural data presented in this thesis. Furthermore, I want to thank Dr. Markus Räschle for the MS analysis of isolated chromatin and his patience while explaining Perseus to an absolute rookie.

I also want to thank Dr. Johannes van den Boom, Helen Müschenborn and Dr. Bojana Kravic from the Meyer lab for their help with the frogs, proximity biotinylation and generation of stable cell lines.

A warm thank you to the lab girls Anika Marko, Verena Höfer, Eman Zaffar and Katerina Pravi not just for their support in the lab. Without you and my favorite question of the day “lunch?” it would not have been the same. I would especially like to thank Anika for her help with the huge proximity biotinylation experiments. Without you I would have never finished it. And thank you Eman for taking over the ungrateful task of reading my methods section. Verena, I will never forget our time in New York. Who would think that finding dinner in New York could be that difficult.

I thank Prof. Dr. Stefan Westermann for scientific support. I also want to thank all the members of the Westermann group for our regular scientific discussion even though Dr. Nicolay Kornakov does not believe in CDK inhibitors. In particular, I want to thank Christian Cozma, Jennifer Harris and Alexander Dudziak for smoothing the language in my thesis. And of course, I always enjoyed catching fresh air with you, Christian.

I would also like to thank Stefanie Raabe for great administration and the nice chats in our office when you came back from the third floor.

I would like to thank my thesis advisory committee, Dr. Peter Bieling, Dr. Alex Bird and Matías Hernández for their advice and scientific discussions on my project.

Many thanks to the amazing IMPRS-LM coordinators Dr. Lucia Sorani and Christa Hornemann. Chista, thank you for reviewing all my applications for a coordinator

position. I am looking forward to meeting both of you. I am sure you will have more professional advice for me.

A special thanks goes to my family, who always believed in me. Especially my mother, who is convinced that her children can become anything they want. I also want to thank my stepfather; I persevered as you told me. Thank you to my siblings for supportive calls, voicemails, and funny pictures of my nephew. There are so many other members of my family who deserve my gratitude, thank you for your support and for always believing in me.
I also want to thank the absolute best friends one can have. You are like a second family to me.

Finally, I want to thank my boyfriend Max for his endless support and patience during the whole process of finishing my thesis. Especially at the very end I would have been lost without you. I am incredibly thankful to have you in my life.

# Declarations

In accordance with § 7 (para. 2, clause d and f) of the Regulations Governing the Doctoral Proceedings of the Faculty of Biology for awarding the doctoral degree Dr. rer. nat., I hereby declare that I have written the herewith submitted dissertation independently using only the materials listed, and have cited all sources taken over verbatim or in content as such.

Essen, _____ _____

Signature of the doctoral candidate

In accordance with § 7 (para. 2, clause e and g) of the Regulations Governing the Doctoral Proceedings of the Faculty of Biology for awarding the doctoral degree Dr. rer. nat., I hereby declare that I have undertaken no previous attempts to attain a doctoral degree, that the current work has not been rejected by any other faculty, and that I am submitting the dissertation only in this procedure.

Essen, _____ _____

Signature of the doctoral candidate

In accordance with § 6 (para. 2, clause g) of the Regulations Governing the Doctoral Proceedings of the Faculty of Biology for awarding the doctoral degree Dr. rer. nat., I hereby declare that I represent the field to which the topic *"TopBP1 loads GINS onto the Metazoan Mcm2-7 Helicase during Replication Origin Firing"* is assigned in research and teaching and that I support the application of Milena Yasemin Jeanne Parlak

Prof. Dr. Dominik Boos

Name of the scientific supervisor/
Member of the University of Duisburg-Essen

Essen, _____ _____

Signature of the scientific supervisor/
Member of the University of Duisburg-Essen