

UNIVERSITY OF BIRMINGHAM

DONSON, a Novel DNA Replication Initiation Factor

by

Georgia Kingsley-Lawrence

A thesis submitted to the University of Birmingham for the degree of DOCTOR OF PHILOSOPHY

Institute of Cancer and Genomic Sciences

College of Medical and Dental Sciences

University of Birmingham

March 2025

UNIVERSITY OF
BIRMINGHAM

University of Birmingham Research Archive

e-theses repository

This unpublished thesis/dissertation is copyright of the author and/or third parties. The intellectual property rights of the author or third parties in respect of this work are as defined by The Copyright Designs and Patents Act 1988 or as modified by any successor legislation.

Any use made of information contained in this thesis/dissertation must be in accordance with that legislation and must be properly acknowledged. Further distribution or reproduction in any format is prohibited without the permission of the copyright holder.

Abstract

DNA replication must be accurately regulated to prevent replication stress and ultimately genome instability. Regulating the loading and firing of origins of replication is therefore critical. Origin activation in yeast has been well characterised, however in higher eukaryotes this process is less well understood. In previous work, the Gambus lab identified the fork-protection factor DONSON to be interacting with factors of the post-termination replisome. DONSON has previously been shown to be essential for accurate progression of S-phase, both in human cells and Drosophila.

Using the cell-free *Xenopus laevis* egg extract system, through this PhD I have shown that DONSON is essential for the assembly of the active replicative helicase (CMG complex) at origins of replication, during the initiation phase. DONSON is required for the recruitment of GINS to the Mcm2-7 hexamer, and is required to stabilise Cdc45 here, to support helicase activation. To carry out this role, we have shown that DONSON interacts with the initiation factor TopBP1, which acts as a scaffold to recruit DONSON to the chromatin. Therefore, we propose that DONSON acts as the functional homologue of the yeast initiation factor Sld2. Once in contact with the DNA, we have shown that DONSON directly interacts with Sld5, of the GINS complex, and Mcm3 – bringing the two components together. This stabilises the entire CMG structure, and causes a misalignment of the two hexameric cores, in relation to one another, which aids in the initial melting of the origin DNA prior to duplex DNA unwinding during elongation. During this project, we have also successfully identified a DONSON peptide that can competitively inhibit DONSON's initiation function, which should therefore be further investigated for its potential therapeutic function.

As DONSON has previously been proposed to play an essential role in fork stabilisation following stress, and we have shown it is also essential during initiation, during this PhD I also looked to determine how it may be regulated during S phase. As we propose DONSON is the higher eukaryotic homologue of Sld2, and given that Sld2 is regulated through phosphorylation, we assessed the impact of CDK-dependent phosphorylation on DONSON. We have been able to show that phosphorylation of DONSON appears to be required for efficient DNA replication and CMG formation. Therefore, we propose that DONSON's functions are, at least in part, regulated by CDK- and ATR-dependent phosphorylation of the protein.

Acknowledgements

Firstly, to my fantastic supervisor, Prof Aga Gambus. Aga, you have believed in me, when even I didn't. You have pushed me to excel and have taught me invaluable lessons. Your patience with my sometimes scatty brain is very much appreciated, and the way you have taught me to think scientifically, to tenaciously pursue the research, and to develop my presentation and communication skills is invaluable. Thank you for being such an understanding and fun supervisor, I will always be grateful for my time in your lab and for all the wisdom that you have imparted to me!

To everyone in the Gambus lab, past and present: Cyntia, Angeliki, Divya, Becky, Alicja, Martina, Abi, Chongluan, Chaoyu, Paolo, Nev, Sarmi, Shaun… thank you! What a lab to be part of! Your kindness, fun and patience with me has made this experience so treasured. Knowing that we're in this together has really been a motivation for me and I am so grateful for all that you have taught me, all of the care you have shown me and all of the laughs we've had together. You have become some of my best friends. Particular thanks goes to Cyntia, I cannot express how valuable your teaching has been to me. You are one of the kindest, most gentle and most intelligent people I have ever met. You have taught me so much throughout my time in the Gambus lab and I know these skills will stand me in great stead for the rest of my career. Angeliki, you really have been my 'DONSON-partner' in this. I could not have done this without you. Thank you for being so patient and so utterly supportive of me. Thank you for sharing your expertise with me, and doing SO many purifications for the project! Martina, thank you for your care these past few months. Thank you for coming alongside me in this, for listening to me and for taking care of me! You are wonderful and have been a real source of fun and light in the final months of my PhD.

I of course want to thank all of my friends and family. Wow, you've had to put up with a lot! Thank you for being patient with me when all of my stories sound the same… blah, blah… commuting… blah, blah… tired! Thank you for always believing in me. Mum, thank you for being such an inspiration in so many ways. Thank you for pushing me to pursue my dreams, to prove people wrong and above all to show love. Thank you for all that you've done for me, for all that you've sacrificed and for **always** loving me. Dad, thank you for all of your care and support throughout my PhD. To my beautiful sister Freya and my

wonderful nieces Lilah and Selkie. I cannot express the joy that you bring to me. Thank you for cheering me on, for telling your teachers about what I do and for being a warm and loving space to get away and have fun! I am so grateful to all of you. Granny and Grandad, your care for me is so unrelenting. Your support of me from the very youngest age has been such a springboard for me. Thank you for all that you've given me over the years. Thank you for the safe space, thank you for the endless chocolate gifts to get me through and thank you for being proud of me no matter what I achieve. Lastly, to my darling husband Matt. All the things that I owe you my love – you know I could not have achieved this without you. What a start to our married life! Your constant and unfailing support and care of me is beyond measure. Your hugs, kind words and prayers for me are what have largely got me through this PhD, and that's why your name is on here too. We've been in this together, from start to finish and you deserve every recognition. You are wonderful!

Finally, to my God.

All to You I owe! Only You know everything it's taken for me to get this PhD. Only You were with me at every single step. For that, and so many other things, I am so grateful. Your care of me over the last three and a half years has kept me safe and has kept me sane! The privilege it has been to discover a part of Your beautiful design is beyond words – thank You for revealing to me everything at the right time. Thank You for taking me on this adventure and showing me ever more how wonderful You are and how much You love me. Thank you for putting me in this lab with these wonderful people and thank You for so clearly having Your hand on this project. You are my greatest inspiration and my greatest supporter. This PhD is completely dedicated to You. I cannot wait to see what You have for me next.

Ani oheveth otach Yahweh!



Table of Contents

List of Figures..... IX

List of Tables..... XI

Abbreviations..... XII

1 Introduction..... 1

1.1 Eukaryotic DNA replication overview..... 1

1.1.1 Initiation..... 2

1.1.2 Elongation..... 3

1.1.3 Termination..... 5

1.2 DNA Replication Initiation..... 7

1.2.1 Origin Licensing..... 7

1.2.2 Origin Firing..... 10

1.2.2.1 Initial DNA Melting..... 13

1.3 DNA Replication Stress..... 15

1.4 Regulation of DNA Replication by Phosphorylation..... 18

1.4.1 Replication Timing Control..... 20

1.5 DONSON..... 22

1.5.1 Orthologues..... 22

1.5.2 DONSON Gene and Protein Structure..... 23

1.5.3 DONSON's Association with the Replisome..... 24

1.5.3.1 Heterochromatin v Euchromatin..... 24

1.5.4 DONSON's Fork Protection Role..... 24

1.5.4.1 ICL Traversal..... 27

1.5.5 Primordial Dwarfism..... 28

1.5.6 Cancer..... 31

1.6 *Xenopus laevis* Egg Extract as a Model System..... 32

1.7 Aims..... 34

2 Materials and Methods..... 35

2.1 Tables..... 35

2.2 *Xenopus laevis* Egg Extract Work..... 42

2.2.1 Egg Extract Preparation..... 42

2.2.2 Sperm Preparation..... 43

2.2.3 *Xenopus laevis* Egg Extract Activation..... 44

2.2.4 Measuring Egg Extract Replication Efficiency..... 45

2.2.5 Immunodepletion of Protein from *X. laevis* Egg Extract..... 46

2.2.6 Chromatin Isolation to Analyse Chromatin-Associated Proteins..... 47

2.2.7 Dynabeads Preparation..... 48

2.2.8 Epoxy Bead Preparation..... 48

2.2.9 Small Scale Immunoprecipitation (IP) from Extract..... 48

2.2.10 Pull-down of Recombinant DONSON from Depleted Extract..... 49

2.2.11 Small Scale IP from Chromatin..... 49

2.2.12 IP from Nucleoplasm..... 50

2.2.13 Large Scale Chromatin IP of DONSON for Mass Spectrometry and CHROMASS..... 51

2.2.13.1 Mass Spectrometry Sample Preparation..... 51

2.2.14 Cdc45-FLAG IP from Chromatin..... 51

2.3 *In vitro* Assays..... 51

2.3.1 *In vitro* Kinase Assay..... 51

2.3.2 *In vivo* Phosphatase Assay..... 52

2.4 DNA Work..... 52

2.4.1 Site Directed Mutagenesis..... 52

2.4.2 PCR Product Visualisation..... 52

2.4.3 DNA Plasmid Processing..... 52

2.4.4 Plasmid Purification..... 53

2.5 Bacterial Work..... 53

2.5.1 Bacterial Strains and Growth Media..... 53
2.5.2 Bacterial Transformation..... 53
2.5.3 Expression of Recombinant *X. laevis* Proteins..... 54
2.5.3.1 Batch Protein Purification of *X. laevis* DONSON..... 54
2.5.4 Dialysis..... 55
2.5.5 Mass Photometry..... 55
2.5.5 *In vitro* Pull-Down of Recombinant Proteins..... 55
2.5.6 Antibody Purification..... 56
2.6 General Procedures..... 56
2.6.1 Protein Visualisation by SDS PAGE..... 56
2.6.2 Bradford Assay..... 57
2.6.3 Immunoblotting..... 57
2.6.4 Gel Filtration..... 58
2.6.4.1 Gel Filtration of Extract..... 58
2.6.4.2 Gel Filtration of Recombinant Proteins in Buffer..... 58
2.7 Computational Analysis..... 58
2.7.1 AlphaFold..... 58
2.7.2 ImageJ Analysis..... 58
2.7.3 Statistical Analysis..... 59
3 Results..... 60
3.1 Chapter 1: DONSON is a Novel Replication Initiation Factor Interacting with TopBP1..... 60
3.1.1 Introduction..... 60
3.1.2 Results..... 61
3.1.2.1 Characterisation of DONSON Chromatin Binding During Replication in *Xenopus* Egg Extract. 61
3.1.2.2 DONSON is Required to Support Replication..... 65
3.1.2.3 DONSON is Required for Formation of the Active Helicase..... 66
3.1.2.4 DONSON is Important for Late-Stage Initiation..... 70

3.1.2.5 DONSON Interacts with TopBP1..... 73

3.1.2.6 DONSON Forms a Homodimer and Interacts with Two TopBP1 Molecules..... 77

3.1.3 Discussion..... 81

3.2 Chapter 2: How does DONSON carry out its initiation function?..... 86

3.2.1 Introduction..... 86

3.2.2 Results..... 87

3.2.2.1 The Importance of DONSON's Flexible Regions..... 87

3.2.2.2 DONSON Directly Interacts with the CMG..... 88

3.2.2.3 DONSON Interacts with Mcm3..... 89

3.2.2.4 DONSON Interacts with Sld5..... 92

3.2.2.5 Utilising the N-terminal Peptide..... 95

3.2.2.6 DONSON-CMG Binding Stabilises the CMG and Causes Rotation to Cue DNA Melting..... 98

3.2.2.7 DONSON's Dimerisation is Required for DNA Synthesis..... 99

3.2.2.8 Do Known Patient Mutations Impact DONSON's Initiation Role?..... 101

3.2.3 Discussion..... 106

3.3 Chapter 3: Regulation of DONSON's Initiation Function through Phosphorylation..... 114

3.3.1 Introduction..... 114

3.3.2 Results..... 115

3.3.2.1 DONSON's Pre-Loading Complex Forms only Transiently..... 115

3.3.2.2 Does DONSON Impact TopBP1 in Egg Extract?..... 118

3.3.2.3 DONSON can be Phosphorylated..... 120

3.3.2.4 Does Phosphorylation Impact DONSON's Structure?..... 123

3.3.2.5 DONSON Phosphorylation is Important to Support DNA Replication..... 125

3.3.2.6 DONSON Phosphorylation is Important for its Interactions..... 129

3.3.3 Discussion..... 132

4 Discussion..... 138

4.1 DONSON's Role as the Higher Eukaryotic Sld2 Homologue..... 139

4.2 Relevance of this Research..... 148

4.3 Future Questions to Address..... 148

5 References..... 151

6 Appendix..... 166

List of Figures

Figure 1.1: Schematic of eukaryotic DNA replication…….. 2

Figure 1.2: Schematic of eukaryotic DNA origin licensing…… 9

Figure 1.3: Schematic of eukaryotic DNA origin firing…… 12

Figure 1.4: Schematic of DNA melting…….. 15

Figure 1.5: Schematic of ATR stress response activation…….. 17

Figure 1.6: Schematic of CDK activity during CMG formation…….. 21

Figure 1.7: Model of *Xenopus* DONSON protein…….. 23

Figure 1.8: Schematic of replication stress response with or without DONSON…….. 27

Figure 1.9: Map of human DONSON with known disease mutations mapped……. 31

Figure 1.10: Schematic of *Xenopus laevis* egg extract preparation…… 33

Figure 3.1.1: Purification of recombinant *Xenopus* laevis His- and MYC-tagged DONSON…… 61

Figure 3.1.2: Purification of *Xenopus laevis* α-DONSON antibodies from whole sera…….. 62

Figure 3.1.3: DONSON has the same chromatin binding kinetics during DNA replication reaction as other replisome components…… 63

Figure 3.1.4: DONSON has the same kinetics on chromatin as other replisome components when replication is perturbed…….. 64

Figure 3.1.5: DONSON interacts with replisome components on chromatin…….. 65

Figure 3.1.6: The immunodepletion of endogenous DONSON from egg extract impairs nascent DNA synthesis…….. 66

Figure 3.1.7: The immunodepletion of DONSON from egg extract impairs active helicase formation on chromatin…….. 67

Figure 3.1.8: Purified recombinant DONSON rescues nascent DNA synthesis and active helicase formation on chromatin in DONSON-depleted extract……. 68

Figure 3.1.9: Blocking the stress response cannot rescue nascent DNA synthesis and active helicase formation on chromatin…….. 69

Figure 3.1.10: DONSON chromatin-loading is reduced on CDK and DDK inhibition…….. 71

Figure 3.1.11: DONSON acts late during DNA replication initiation…….. 72

Figure 3.1.12: DONSON interacts with replication initiation factors…….. 73

Figure 3.1.13: DONSON interacts with TopBP1…… 75

Figure 3.1.14: The immunodepletion of endogenous TopBP1 decreases replication in egg extract. 77

Figure 3.1.15: rDONSON primarily exists in dimer form…… 78

Figure 3.1.16: DONSON interacts with TopBP1 via its BRCT3 domain…….. 80

Figure 3.2.1: DONSON's flexible regions are required to support nascent DNA synthesis…… 88

Figure 3.2.2: Isolating CMG monomers and dimers…….. 89

Figure 3.2.3: Dimeric DONSON bridges the double CMG…….. 89

Figure 3.2.4: DONSON directly interacts with Mcm3…….. 90

Figure 3.2.5: DONSON's interaction with Mcm3 is required for DNA synthesis…….. 92

Figure 3.2.6: DONSON directly interacts with Sld5…… 93

Figure 3.2.7: DONSON's interaction with Sld5 is required for DNA synthesis…… 94

Figure 3.2.8: DONSON's interaction with Sld5 is required for formation of the active helicase…… 95

Figure 3.2.9: DONSON's N-terminal helicase is sufficient to block DNA replication…….. 98

Figure 3.2.10 Schematic of initial DNA melting upon DONSON binding to the CMG…… 99

Figure 3.2.11: DONSON's dimerisation is required to support DNA synthesis…….. 101

Figure 3.2.12: A known DONSON patient mutation is positioned at the groove between DONSON monomers where TopBP1 is predicted to bind…… 102

Figure 3.2.13: Select DONSON patient mutations impact DNA replication initiation…… 103

Figure 3.2.14: Select DONSON patient mutations impact fork stability…… 105

Figure 3.3.1 Interaction between DONSON and other pre-landing complex components is transient…….. 117

Figure 3.3.2 Endogenous DONSON runs at a higher molecular weight than rDONSON…….. 118

Figure 3.3.3 DONSON depletion doesn't affect the molecular weight at which TopBP1 runs…… 119

Figure 3.3.4 DONSON can be in vivo and in vitro CDK phosphorylated…….. 122

Figure 3.3.5 DONSON's CDK phosphorylation impacts chromatin binding and structure…… 124

Figure 3.3.6 DONSON's CDK phosphorylation is required for efficient DNA synthesis…… 127

Figure 3.3.7 DONSON's CDK phosphorylation is required for efficient helicase formation…….. 128

Figure 3.3.8 DONSON's CDK phosphorylation is required for efficient TopBP1 interaction…….. 129

Figure 3.3.9 DONSON's CDK phosphorylation impacts 'pre-loading complex' structure…… 131

Figure 4.1.1 Schematic of how we propose DONSON supports CMG formation on chromatin in *Xenopus laevis* egg extract…… 143

Figure 4.1.2 Schematic of how we propose DONSON supports DNA melting in *Xenopus laevis* egg extract…… 145

Figure 4.1.3 Schematic of how DONSON may be released from liquid-liquid phase separated multimeric complexes in *Xenopus laevis* egg extract…… 147

Figure 4.3.1 DONSON degradation in human cells…… 150

List of Tables

Table 1.1 Known DONSON mutations implicated in developmental disorder…… 30

Table 2.1 Composition of Reagents Used…… 35

Table 2.2 Inhibitors Added to *X. laevis* Egg Extract…….. 37

Table 2.3 Antibodies…….. 37

Table 2.4 Plasmids…….. 38

Table 2.5 Primers…….. 39

Table 2.6 Recombinant Proteins…… 40

Table 2.7 Composition of PCR reaction mixture…….. 41

Table 2.8 Thermocycling Conditions…… 41

Table 3.1: DONSON interacts with terminated replisomes…….. 60

Table 3.2: DONSON Patient Mutations Sites are Conserved in Xenopus…… 102

Abbreviations

AAA+: ATPase Associated with diverse cellular Activities

ACS: ARS consensus sequence

ADP: Adenosine diphosphate

AIM: Auto-inducing media

ANIB: Acetate nuclear isolation buffer

APC/C: Anaphase Promoting Complex/Cyclosome

ARS: Autonomously replicating sequence

ATAD5: ATPase Family AAA Domain Containing 5

ATM: Ataxia-telangiectasia mutated

ATP: Adenosine triphosphate

ATR: Ataxia-Telangiectasia and Rad3-related

ATRi: ATR inhibitor

bp: base pair

BRCA1: Breast and ovarian cancer susceptibility protein 1

BRCT: BRCA1 C-terminal

BSA: Bovine Serum Albumins

Cdc: Cell division cycle protein

Cdh1: Cadherin 1

CDK: Cyclin dependent kinase

CDKi: CDK inhibitor

Cdt1: Chromatin licensing and DNA replication factor 1

CFS: Common fragile site

Chk1: Checkpoint kinase 1

CHX: Cycloheximide

Chromass: Chromatin mass spectrometry

Circ-RNA: Circular-RNA

CKI: CDK inhibitor

CMG: Cdc45/Mcm2-7/GINS complex

CRL4: Cullin-4-RING ubiquitin ligase

cryo-EM: Cryo-electron microscopy

Ctf4: Chromosome transmission fidelity 4

D: Aspartic acid

dCMGDo: DONSON-bound double CMG

DDK: Dbf4 dependent kinase

DDKi: DDK inhibitor

DDR: DNA damage response

DDT: DNA damage tolerance

DMF: Dimethylformamide

DNA: Deoxyribonucleic acid

DNMT: DNA (cytosine 5) methyl transferase

dNTP: Deoxyribonucleotide

DONSON: Downstream neighbour of Son

Dpb11: DNA polymerase B 11

DSB: Double strand break

dsDNA: Double stranded DNA

E: Glutamic acid

ER: Energy regenerator

ETAA1: Ewing tumour-associated antigen 1

F: Phenylalanine

FA: Fanconi Anaemia

FACT: Facilitates chromatin transcription complex

FANC: Fanconi anaemia complementation group protein

FEN1: Flap endonuclease 1

FPC: Fork protection complex

GEO: Gene Expression Omnibus

GINI: GINS-interacting

GINS: Go, Ichi, Nii, and San

H: Histone

Hbo1: Histone acetyltransferase bound to Orc1

Hd: Humpty Dumpty

HR: Homologous recombination

HSS: High speed supernatant

HU: Hydroxyurea

ICL: Interstrand crosslink

IgG: Immunoglobulin G

IP: Immunoprecipitation

iPOND: Isolation of proteins on nascent DNA

IPTG: Isopropyl-β-D-1-thiogalactopyranoside

K: Lysine

l: litre

LB: Luria-bertani broth

LFB: Licensing Factor Buffer

Lig1: DNA ligase 1

LLPS: Liquid-liquid phase separation

LRR1: Leucine rich repeat protein 1

LSS: Low speed supernatant

M: Methionine

Mass Spec: Mass spectrometry

MCM: Mini chromosome maintenance complex

MGORS: Meier Gorlin syndrome

MISSLA: Microcephaly, short stature and limb abnormalities

MMR: Marc's Modified Ringer's

MMS: Microcephaly Micromelia Syndrome

MPD: Microcephalic Primordial Dwarfism

MTBP: MDM2 binding protein

NPE: Nucleoplasmic extract

OCCM: ORC-Cdc6-Cdt1-MCM

OD: Optical density

OF: Okazaki fragment maturation

ORC: Origin recognition complex

P: Proline

PAGE: Polyacrylamide gel electrophoresis

PARP: Poly ADP ribose polymerase

PBS: Phosphate buffered saline

PCNA: Proliferating cell nuclear antigen

PCR: Polymerase chain reaction

PIP box: PCNA interacting protein box

PLA: Proximity-ligation assay

Plk1: Polo like kinase 1

PMSF: Phenylmethylsulfonyl fluoride

Pol: Polymerase

PP: Protein phosphatase

Pre-IC: Pre-initiation complex

Pre-LC: Pre-loading complex

Pre-RC: Pre-replication complex

PRMT5: Protein arginine N-methyltransferase 5

Psf: Partner of Sld five

PTM: Post translational modification

PVDF: Polyvinylidene

R: Arginine

RFC: Replication factor C

Rif1: Replication Timing Regulatory Factor 1

RING: Really interesting new gene

RNA: Ribonucleic acid

RNAP2: RNA polymerase II

RPA: Replication protein A

S: Serine

SCF: Skp1-Cdc53-F-box

SDS: Dodecyl sulfate sodium

SEM: Standard error of the mean

Sld: Synthetic Lethality with Dpb11

SMARCA: SWI/SNF-related matrix-associated actin-dependent regulator of chromatin subfamily A

SSB: Single strand break

ssDNA: Single stranded DNA

T: Threonine

TCA: Trichloroacetic acid

TCGA: The Cancer Genome Atlas

TopBP1: DNA topoisomerase 2 binding protein 1

Topo: Topoisomerase

TRAIP: TRAF interacting protein

TSS: Transcription termination sites

UEB: Un-activating Extraction Buffer

UFB: Ultrafine bridges

UV: Ultra violet

V: Valine

W: Tryptophan

wt: Wild type

WES: Whole exome sequencing

WHD: Winged helix domain

Y: Tyrosine

ZnF: Zinc Finger

1 Introduction

1.1 Eukaryotic DNA replication overview

Accurate DNA replication underpins precise cell division and is the foundation for the propagation of life. When the process goes awry and genome integrity is forfeited, mutation, cellular deterioration and disease can all be observed [1]. Genome replication is a semiconservative process by which two parental DNA strands act as templates for the synthesis of two daughter strands and results in two identical DNA molecules. Each contains one new, replicated strand and one old, template strand [2], [3]. The foundations of the replication process are evolutionarily conserved. However, the intricacy of both the machinery and regulation of the process increase with organism complexity and genome size [4]. DNA replication must be coordinated with other processes, including transcription and genome repair, and the complexity of this increases in eukaryotic replication [5], [6]. Chromatin structure also adds further complexity to the eukaryotic genome as it must be disassembled during replication and reconstituted following it. This involves histone eviction, regeneration and reestablishment, resulting in incorporation of parental and new histones to daughter DNA molecules [7].

Due to its importance, and complexity in higher eukaryotes, DNA replication must be well regulated. This is to ensure that all DNA is replicated during S phase and just once per cell cycle [8]. DNA replication is split into three distinct steps – initiation, elongation and termination (Figure 1.1).

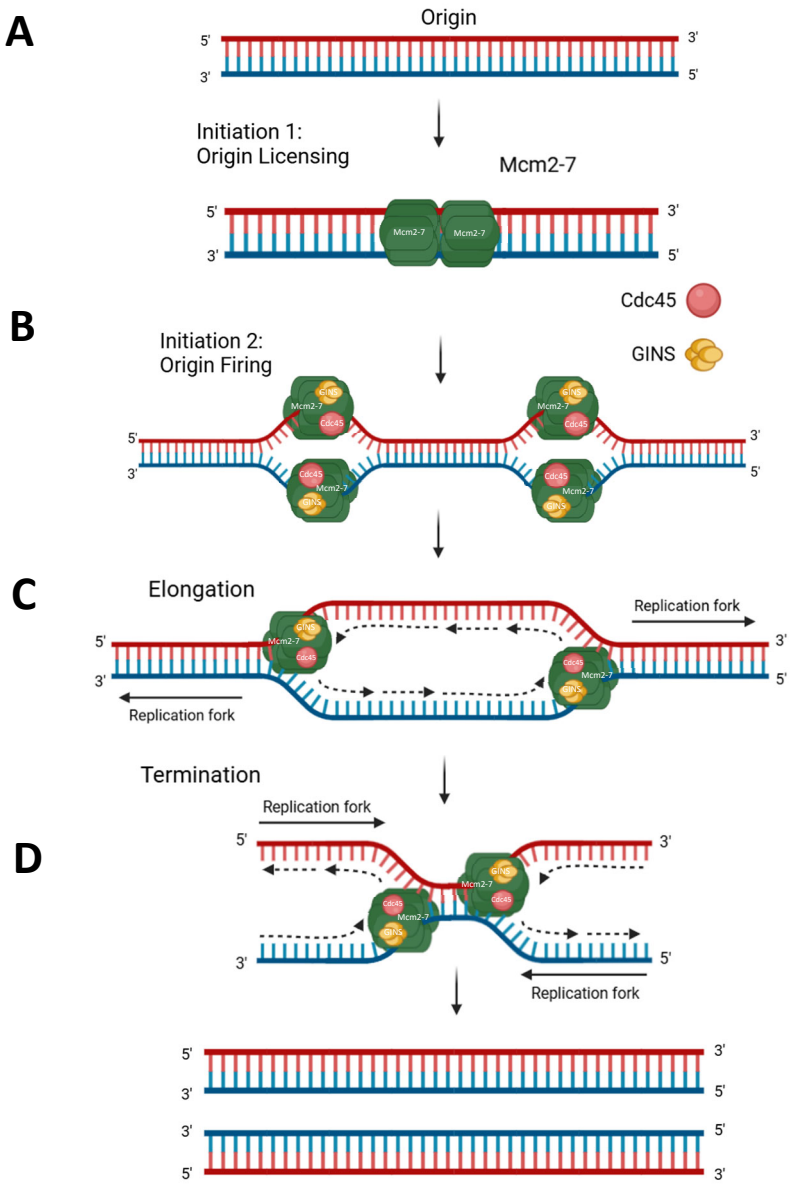


Figure 1.1 Schematic of eukaryotic DNA replication. Following origin recognition, Mcm2-7 hexamers are loaded onto DNA. The active helicase is formed via the recruitment of Cdc45 and GINS (CMG complex). The CMG, with the replication components surround it, unwinds the DNA to form two replication forks that progress along the DNA in opposite directions from each origin, replicating both strands of the DNA. Termination occurs upon the convergence of two forks from neighbouring origins. Adapted from [39].

1.1.1 Initiation

DNA replication initiates from loci on the DNA called replication origins. Here the replicative helicase is activated and proceeds to unwind the DNA. Whilst replication advances around bacterial circular chromosomes from just one replication origin, in eukaryotic chromosomes replication starts from numerous origins. The number of origins in a single eukaryotic linear chromosome spans from hundreds in yeast to many thousands in higher eukaryotes. With the exception of *S. cerevisiae* and *Saccharomycotina* species, eukaryotic origins do not have defined consensus sequences, likely instead

determined by DNA or chromatin features, such as transcription start sites, epigenetic marks on histones, DNA structures, CpG islands or AT-rich regions. In contrast, bacterial and budding yeast origins have defined DNA sequence elements [9], [10].

Initiation takes place in two distinct steps, origin licensing and origin firing. During licensing, the core of the replicative helicase is assembled at replication origins, in the form of two Minichromosome Maintenance 2-7 (Mcm2-7) protein complexes, which bind to origins in a head-to-head formation in relation to one another (Figure 1.1 A). This occurs during late M phase and through G1 phase. Following Mcm2-7 double hexamer loading, two active helicases are assembled by the recruitment of Cell division cycle protein 45 (Cdc45) and GINS complex (Psf1, Psf2, Psf3 and Sld5) (one each per Mcm2-7 hexamer) (Figure 1.1 B). Recruitment of Cdc45 and GINS to the Mcm2-7 hexamer forms the CMG complex (Cdc45, Mcm2-7 and GINS), which unwinds the DNA duplex, and occurs during S phase [11]. This requires activation by S phase kinases, Cyclin dependent kinase (CDK) enzymes and Drf1/Dbf4-dependent Cdc7 kinases (DDK). Two CMGs become activated for DNA replication via the addition of initiation factors that form a complex around each CMG, called the replisome. Each replisome and their resulting replication forks proceed away from each origin in opposite directions, forming a replication bubble, and supporting replication of the DNA.

1.1.2 Elongation

Following origin firing, each replication fork progresses along the DNA, facilitated by the active helicase unwinding the double-stranded DNA and supporting DNA replication as it proceeds (Figure 1.1 C) [12]. As the duplex DNA is unwound, single-stranded DNA is exposed and rapidly bound by replication protein A (RPA), to stabilise the DNA structure and prevent it being identified as DNA instability and protect it from nuclease attack [13]. Unwinding of the duplex DNA enables DNA polymerases at the fork to access and replicate each parental template. Proliferating cell nuclear antigen (PCNA) is a cofactor of DNA polymerases, loaded by the clamp loader Replication Factor C (RFC) [14]. PCNA is a ring-shaped protein and RFC uses its ATPase activity to open this ring and clamp it onto DNA. Once encircling the DNA, PCNA secures various polymerases to itself, including δ and Ɛ, thus increasing their processivity by maintaining

their chromatin binding [15]. Each CMG helicase translocates 3’ to 5’ along one template DNA strand: the ‘leading strand template’, so called because the daughter strand produced in reflection is created in the 5’ to 3’ direction [12]. On this strand, DNA polymerase Ɛ (Pol Ɛ) carries out continuous replication, producing a daughter strand in the 5’-3’ direction. DNA polymerases can only extend DNA, not start replication de novo, therefore a primase enzyme is required to prime the DNA for replication. On the leading strand, DNA primase – in association with Polymerase α (Pol α) – synthesises short RNA primers on the single stranded DNA, allowing Pol Ɛ to promote continuous DNA replication. However, there is also a suggestion that the new leading strand may start from the end of the first Okazaki fragment (OF) of the lagging strand of the fork progressing from the same origin in the opposite direction [16]. In contrast, the other parental strand – ‘the lagging strand template’ – is unwound in the 5’-3’ direction, meaning that it must be replicated discontinuously due to polymerases only replicating the DNA in this direction. This discontinuous replication relies on the priming activity of Pol α, which assembles short RNA-DNA primers along the length of the exposed parental strand. Ctf4 (yeast)/And-1 (human) is a homotrimeric replication factor that is required to recruit Pol α to the CMG, enabling it to create these short RNA primers. In human cells, this interaction also requires Cdc7 phosphorylation [17]. Following Pol α primer formation, DNA Polymerase δ (Pol δ) is then required to produce short sections of daughter DNA (Okazaki fragments) in a discontinuous manner [18]. Upon reaching the 5’ end of the previous OF, Pol δ replicates further, displacing the Pol α-produced RNA-DNA primer and forming a 5’ flap. Each resulting OF must then be processed to leave a continuous DNA molecule. The endonuclease FEN1 (flap endonuclease 1) cleaves the 5’ flap. A nick is left between each OF as a result, which is then ligated by ligase 1 (Lig1) [4].

The action of DNA unwinding by the helicase causes problems ahead of the fork where positive DNA supercoiling arises. This must therefore be resolved, to maintain DNA topology, relieve coiling tension and to prevent excessive supercoiling which can form alternative DNA structures, such as quadruplexes, and is transcriptionally deleterious [19], [20]. The majority of this is relieved by the enzyme Topoisomerase 1 (Topo1) an intrinsic component of the replisome [19]. Topo1 cleaves one strand of the parental duplex ahead of the replication fork, allowing the other strand to rotate freely. Less frequently, both parental strands are cleaved to relieve coiling tension. This is carried out by Topoisomerase 2 (Topo2) which,

following duplex cleavage, passes another double stranded DNA segment through the break to reverse the positive supercoil [21]. However, in the event that Topoisomerase enzymes are inhibited, the replication fork itself takes on the role by rotating the DNA. Although this enables the replisome to continue to translocate along the DNA, it results in intertwined sister chromatids behind the fork (pre-catenanes), which must be resolved by Topo2 before chromosomal segregation during mitosis [4].

As previously mentioned, the chromatin environment impacts the complexity of the DNA elongation process. As the replication fork progresses, nucleosomes ahead of the replication machinery must be disassembled to allow the replication machinery to progress and have access to the DNA. Nucleosome disassembly can be illustrated in three distinct steps. First, the interface between the $(H3\text{-}H4)_2$ tetramer and the (H2A-H2B) dimer is opened, after which the dimer is released from the tetramer and the DNA, and finally the tetramer is removed from the DNA [22]. This is coordinated by various histone chaperones, such as components of the human facilitates chromatin transcription (FACT) complex [23]. Immediately following replication fork proceeding and DNA synthesis, nucleosomes are reassembled on the DNA to compact the chromatin and prevent genomic instability. Parental histones, with their post-translational modifications (PTMs), from the parental strands are reestablished on the DNA, by cooperation between different histone chaperones FACT and Asf1, as well as replisome components. On the leading strand, Pol Ɛ supports this recycling, whilst on the lagging strand Mcm2-Ctf4-Pol α supports this process. Newly synthesised histones are also required. The $(H3\text{-}H4)_2$ tetramer is assembled first, followed by addition of (H2A-H2B) dimers. These newly synthesised histones are delivered to the DNA by chaperones, including CAF1, FACT or Rtt106. Following their deposit on DNA, new histones are modified to mirror parental PTMs – maintaining the epigenome [24].

1.1.3 Termination

As two forks from neighbouring origins converge, DNA replication is terminated (Figure 1.1 D). Upon termination, DNA unwinding and synthesis must be completed, pre-catenanes resolved and the replisome disassembled [25], [26]. Collectively, this enables chromosomes to faithfully segregate [27]. In bacteria, where there is a single origin of replication, replication forks and helicases terminate at a fixed region

opposite the origin. This region contains ten sites, which are called Ter sites in *E. coli* [28]. Collectively, these sites bind the terminator protein Tus, and form a replication fork trap. Here, forks enter but cannot leave, ensuring that roughly half of the chromosome is replicated in one direction and the other half in the opposite [29]. In organisms with multiple origins however, fork convergence occurs throughout the chromosome and termination occurs throughout S phase, whenever and wherever two forks converge. Although in eukaryotes termination sites are sequence non-specific, there appears to be bias towards regions adjacent to transcription termination sites (TSS) [30], [31]. Here, RNA polymerase II (RNAP2) pauses to terminate transcription, and this action can result in a collision with the replication machinery. This causes fork stalling and explains why replication termination may also be more likely to occur within these regions. The majority of termination occurs during early-mid S phase [32]. This can be explained by the fact that euchromatin is replicated first and is faster, meaning that there is a higher density of forks, including fork termination, in early-mid S phase.

During fork convergence, DNA is synthesised at the same rate throughout, indicating that converging forks do not collide, but instead pass each other. This action is aided by the PIF1 helicase [27]. PIF1 unwinds the DNA in the opposite direction to the CMG, allowing it to translocate past the opposite CMG. This enables each leading strand to extend past the other until they are just a few nucleotides away from the downstream OF [33]. The 3' end of the leading strand must then be ligated to the downstream OF and the final OF processed [26]. Whilst Pol δ displaces OFs during elongation, Pol Ɛ DNA synthesis is more prevalent towards the centre of replicons (where forks are more likely to converge) and therefore appears to be more likely to replicate the final DNA segments at termination sites [34].

Following fork convergence, replisomes are unloaded from the double stranded DNA. This begins with ubiquitylation of Mcm7 at the N-terminus, with K48-linked chains. Mcm7 is ubiquitylated by a ubiquitin ligase enzyme. In budding yeast Skp1-Cdc53-F-box (SCF) Cullin-RING complex with the substrate-targeting component Dia2, takes on this role, whilst in higher eukaryotes it is carried out by the Cullin2-based ubiquitin ligase CRL2 with the substrate receptor Leucine Rich Repeat 1 protein (LRR1) [26], [35]. As soon as the parental duplex between the converging CMGs is unwound, the lagging strand, that is excluded from the Mcm2-7 during elongation, loses its contact with the CMG. This allows the ubiquitin ligase to

interact with the CMG and ubiquitylate Mcm7 [35]–[38]. A 'ubiquitin threshold' of five or more proteins added in a chain to Mcm7 must be reached to promote CMG unloading [36]. Mcm7 K48 ubiquitylation promotes remodelling of the protein by the hexameric ATPase segregase p97, or Cdc48 in yeast, allowing its removal from the chromatin [39].

In the event that CMGs are not removed from the chromatin by the end of S phase, the ubiquitin ligase TRAIP removes them during mitosis [39]. This backup pathway only occurs in higher eukaryotes and was first discovered in *C. elegans* embryos [35]. TRAIP synthesises K6 and K63-linked ubiquitin chains on Mcm7. Following ubiquitylation by TRAIP, the CMG is disassembled by p97, along with its cofactor UBXN-3 [40]. Whilst TRAIP-dependent CMG disassembly can result in DNA breakage, if the complex is removed from unfinished forks and the DNA is exposed, this is a preferable outcome to the uncontrolled DNA damage that would occur were CMGs not removed by the end of mitosis [39].

1.2 DNA Replication Initiation

1.2.1 Origin Licensing

The first stage of DNA replication initiation, origin licensing, occurs during the final stages of mitosis and through G1 of the cell cycle, and involves origin licensing proteins that are strongly conserved between species. Licensing is characterised by the loading of two hexameric helicase cores onto the replication origin. In bacteria, the origin (oriC) is recognised by the initiator DnaA, which generates ssDNA adjacent to the oriC. Following this, DnaA recruits two hexameric helicases (DnaB) to each oriC. One hexamer encircles each of the two single strands of the DNA created by DnaA [41]. In eukaryotes, the Origin Recognition Complex (ORC) is the initiator (Figure 1.2 A) [41]. Although eukaryotic origins are traditionally thought of as sequence unspecific, in yeast ORC binds to T-rich regions labelled Autonomously Replicating Sequences (ARSs). These are comprised of a nonpalindromic ARS consensus sequence (ACS), which is essential for origin function [42]. In contrast, in higher eukaryotes the only DNA element identified at origins is a G-rich element with the propensity to form a non-canonical four-stranded helical structure (G quadruplex) [43].

The chromatin environment at origins also has an impact on origin licensing. Histone post-translational modifications and nucleosome positioning play a role in ORC localisation [44]. The local chromatin environment may require modification to enhance accessibility and enable origin licensing. A number of histone-modifying enzymes and chromatin remodellers associate with licensing factors for open chromatin formation [45], [46]. One example is histone acetyltransferase bound to Orc1 (Hbo1) , which primarily acetylates histone H4, producing more accessible DNA [47]. Likewise, histone H4 methylation supports ORC recruitment to nearby origins [48]. Ultimately, ORC must bind throughout the entire genome to stimulate replication at every part of DNA, however the DNA is difficult to access due to histones. Therefore, chromatin remodellers are required to support this.

ORC is comprised of six subunits (ORC1-6), with Orc1-5 essential for DNA binding. These subunits are positioned in a crescent-moon shape, whilst Orc6 binds on the periphery [49]. ORC is a member of the AAA+ family of ATPases and requires ATP binding to recognise and stably bind origin DNA via its C-terminus. Following ORC recruitment, Cdc6 – another AAA+ ATPase – is recruited to the origin. To recruit Cdc6, the winged helix domain (WHD) of Orc2 disengages from DNA, creating space for an Orc1-Cdc6 interaction, which results in the closure of the ORC ring, entrapping duplex DNA and creating a stable platform for Mcm2-7 hexamer recruitment [50]. Cdc10-dependent transcript 1 (Cdt1) is also required for Mcm2-7 loading, it binds to the complex and wraps around Mcm2, 6 and 4, keeping the Mcm2-5 gate open for DNA accessibility [51]. Importantly, it is the pre-assembled Mcm2-7 hexamer that is loaded, not the constituent subunits. The Cdt1-Mcm2-7 complex is then recruited to ORC-Cdc6 at origins, to form the OCCM (ORC-Cdc6-Cdt1-MCM), a prerequisite to the formation of the Mcm2-7 double hexamer on DNA (Figure 1.2 B) [52]. The ATPase activity of Cdc6 is required to bind ORC and recruit Cdt1, and subsequently to facilitate Mcm2-7 loading [53]. Cdc6 and Cdt1 are promptly released and a second round of Cdc6, Cdt1-MCM origin recruitment occurs, to form the double MCM hexamer on the DNA. This complex is stabilised by Orc6 [54]. The two Mcm2-7 hexamers associate in an inactive head-to-head formation (Figure 1.2 C) [55]. Both Cdc6 and Cdt1 have been visualised to shuttle between DNA-bound and soluble states, suggesting that each molecule is likely involved in numerous origin licensing events. Therefore, it is unsurprising that both proteins are highly regulated to ensure that inappropriate origin licensing is

prevented. Whilst the process of origin licensing is largely similar in yeast and higher eukaryotes, the mechanism is more complex in higher eukaryotes to allow for more flexible and increased control of the process [56].

To prevent re-replication of the DNA, loading of the Mcm2-7 double hexamers onto DNA is limited to late mitosis and G1 of the cell cycle, where Cdt1 and Cdc6 are plentiful and CDK activity is kept low. Once the cell cycle progresses into S phase, the factors required to load the double hexamers onto DNA are downregulated to ensure that origins fire only once during each cell cycle. In particular, Cdt1 is regulated via proteolysis and through sequestration of its activity. In higher eukaryotes, Cdt1 is regulated by the protein Geminin – not present in lower order eukaryotes. Geminin directly binds to Cdt1, preventing it from complexing with Mcm2-7 and so preventing chromatin loading of the hexamers and thus avoiding re-replication of the DNA [57]. Likewise, upon transition into S phase, the levels of the S phase kinases increase – required for formation of the active helicase and to promote origin firing. Phosphorylation of Cdc6 and Mcm2-7 hexamers not bound to the chromatin by S phase CDKs promotes degradation and/or nuclear export, likewise to prevent re-licensing of the DNA [58]. Therefore, during late-M and G1 Geminin and CDK activity is kept low, to enable licensing of the DNA. At the end of mitosis, Geminin is polyubiquitinated by the Anaphase Promoting Complex/Cyclosome (APC/C), which targets it for inactivation or degradation, whilst CDK activators, such as Cdc25 are also degraded [57]. Additionally, CDK inhibitors (CKIs) are accumulated during the end of mitosis and during G1 [41]. Collectively, this ensures that each step of initiation is regulated to prevent catastrophic re-replication of the DNA.

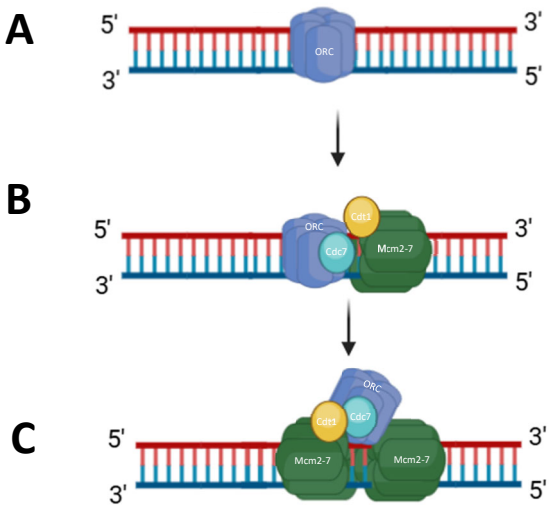


Figure 1.2 Schematic of eukaryotic DNA origin licensing. DNA origins are bound by the Origin Recognition Complex (ORC). Cdc6 promptly binds ORC and, along with Cdt1, recruits an Mcm2-7 hexamer to the DNA. Cdc6 and Cdt1 are recycled and recruit an additional Mcm2-7 hexamer to the origin, resulting in two Mcm2-7 hexamers bound in a head-to-head formation at each origin. Adapted from [41].

1.2.2 Origin Firing

Origin firing occurs during S phase and is characterised by the formation of two active helicases at origins, with corresponding bidirectional forks. Each active helicase is composed of a single Mcm2-7 hexameric core, activated by the presence of both Cdc45 and GINS complex, forming the CMG [11]. The recruitment of both Cdc45 and GINS requires the activity of S phase kinases and additional firing factors, which have predominantly been identified through genetic screens in budding yeast as Synthetic Lethality with Dpb11 (Sld)3, Sld7, Sld2, DNA polymerase B 11 (Dpb11), Pol Ɛ and α, and Mcm10 [59]. Despite this, until recently the mechanism of origin firing has not been well understood.

Once MCMs have been successfully loaded during origin licensing, at the G1-S transition kinase levels increase – both CDK and DDK. In *Saccharomyces cerevisiae,* the first step in origin firing is phosphorylation of the MCM hexamer by DDK (Figure 1.3 A) [60]. Following this, Sld3-Sld7 forms a complex with Cdc45 and they are collectively recruited to the hexamer (Figure 1.3 B) [61]. The crystallographic structure of Sld3-Sld7 oligomerisation reveals that Sld7 forms a homodimer, likely placing two copies of Sld3 in the appropriate orientation to read the DDK-dependent phosphorylation of Mcm4/6 within both rings of the double hexamer [62]. However, single molecule studies have suggested that recruitment of Cdc45 by Sld3 to each hexamer is sequential, rather than symmetrical [63]. In higher eukaryotes, the Sld3-Sld7 complex orthologue is Treslin-MTBP, which has been shown to be required for direct recruitment of Cdc45 onto the hexamer and a precursor for GINS recruitment (Figure 1.3 B) [64]. Importantly, under conditions of replication stress, DDK-phosphorylation of MCMs can be reversed by protein phosphatase 1 (PP1), which is targeted to the chromatin by Replication Timing Regulatory Factor 1 (Rif1) and promotes global downregulation of replication [65]. Sld2 is also phosphorylated, however by S phase CDKs rather than DDKs. This enables phosphorylated Sld2 to interact with Dpb11 through its carboxyl terminus (Figure 1.3 C). Dpb11 contains four BRCA1 C-terminal (BRCT) domains which bind to phosphorylated factors – in particular, BRCT1 and 2 bind phosphorylated Sld3, and BRCT3 and 4 bind phosphorylated Sld2 [66]. This enables the formation of the pre-loading complex (Dpb11, Sld2, GINS and DNA Pol ε) (pre-LC), which forms prior to DNA binding and is a prerequisite for CMG formation (Figure 1.3 C) [67]–[69]. Interestingly,

whilst Pol Ɛ is essential to pre-LC assembly and resulting CMG formation in yeast, in higher eukaryotes it appears to be dispensable for CMG assembly [70], [71].

Whether the orthologues of each pre-LC factor in higher eukaryotes are recruited to DNA as a complex is not yet understood. TopBP1 is homologous to Dpb11 (Cut5 in *Schizosaccharomyces pombe*, Mus101 in *C. elegans* and *Drosophila*), whilst the functional homologue of Sld2 has been under debate. The N-terminus of RecQ4 contains a region of homology to Sld2, however it functions during replication initiation only after CMG assembly [72], [73]. Markedly, TopBP1 contains more BRCT domains than its counterpart Dpb11. Domains 4-8 are dispensable for replication initiation and domains 1-2 have been shown to bind CDK-phosphorylated Treslin [74]. BRCT3 has also been shown to be essential for replication, however a binding partner for the domain hasn't previously been identified, suggesting that the Sld2 homologue may bind this region.

CDKs also phosphorylate MCM-bound Sld3. This enables Dpb11 to bind to Sld3 through its amino-terminus and dock the pre-LC onto Cdc45-bound Mcm2-7, depositing GINS for active helicase formation (Figure 1.3 D). Notably, the binding between Cdc45 and Mcm2-7 is only weak until GINS is recruited, at which point the stable CMG is formed [59], [74], [75]. This is likely due to the firing factors that recruit Cdc45 and GINS to Mcm2-7 serving to bridge the two helicases together on DNA [41]. Pol Ɛ has been observed fulfilling this role [76]. Apart from being required to replicate the leading strand, Pol Ɛ plays an essential non-catalytic role during initiation bridging GINS to the MCM hexamer, further supporting the interaction between them [76]. Specifically, the Dpb2 subunit of Pol Ɛ binds Psf1 and Mcm3, whilst the C-terminus of the Pol2 subunit binds Mcm2 and 5, further stabilising the interaction between the two complexes [76]. Stable interactions also occur directly between the elements of the CMG. Cdc45 and GINS have been observed to bind directly to Mcm2, 3 and 5 subunits of the hexamer, connecting the N-terminal and ATPase domains [77]. In yeast specifically, Psf2 binds to the A domain of Mcm5. Importantly, this causes the Mcm7 subunit to disengage from Mcm5 that it binds to prior to helicase activation. As a result, the dimerisation of the double hexamer is no longer supported and the two complexes separate slightly [78]. However, in Metazoa, Mcm7 is shorter and therefore may be unable to support stabilisation

of the double hexamer in the same way, indicating that the mechanism of helicase activation in higher eukaryotes may differ slightly compared to yeast. Collectively, the binding of GINS, Cdc45 and Pol Ɛ to the hexamer seals the Mcm2-5 DNA-entry gate, meaning that it cannot yet eject the lagging strand for DNA unwinding. To activate the complex for duplex DNA unwinding, additional factors must bind. In particular Mcm10 is required for the initial melting of the DNA [79].

Various factors required for origin firing have been observed at levels far below the number of origins actually used during S phase. These include Cdc45 and GINS. Upon overexpression of these factors, initiation from late-firing origins can be promoted [61], [80]. As these factors are rate limiting for initiation, this indicates that they contribute to ordered origin firing during S phase which is necessary for controlling replication timing [41].

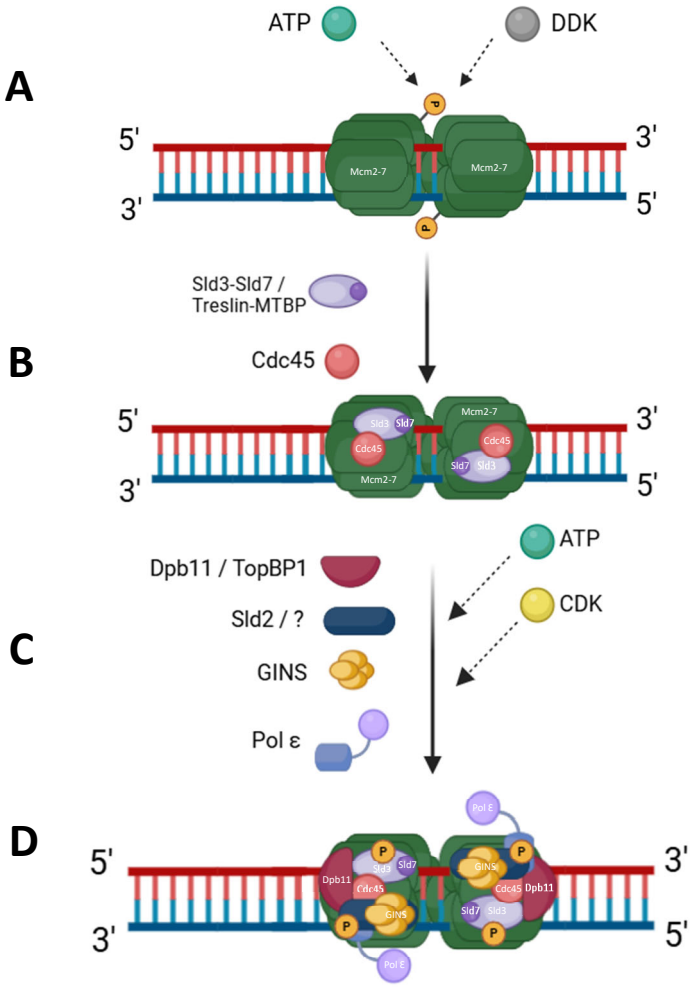


Figure 1.3 Schematic of eukaryotic DNA origin firing. Following Mcm2-7 double hexamer loading, DDKs phosphorylate MCM subunits. This allows recruitment of Cdc45 in its association with Sld3-Sld7 (yeast) / Treslin-MTBP (higher eukaryotes). Once bound, Sld3-Sld7 is phosphorylated by CDKs. CDKs also phosphorylate Sld2 which allows its interaction with Dpb11 (TopBP1 in higher eukaryotes). Dpb11 -Sld2-GINS-Pol Ɛ are then recruited to the DNA via an interaction between Dpb11 and the phosphorylated Sld3. Adapted from [41].

Critically, regulation to ensure only one round of replication per cell cycle is essential to maintain genomic and cellular stability. Once cells have entered into S phase and Geminin levels and CDK activity are increased, there is no chance to reload Mcm2-7 hexamers, to prevent re-licensing of the DNA. However, this means that to have enough licensed origins to replicate the entire length of the DNA, many more origins are licensed than support DNA replication. As observed in *Xenopus* egg extracts and HeLa cells, around 10-40 Mcm2-7 double hexamers are loaded onto each origin, while only about 10% are activated for replication. The unused origins are dubbed 'dormant', and this process is known as the MCM paradox [81], [82]. Dormant origins are most frequently utilised under conditions of replication stress in response to nearby stalled or damaged forks, ensuring the genome is not under-replicated [83]. However, insufficient origin licensing during G1 can occur and is a source of genome instability, by delaying S phase. In contrast, re-replication of chromosomes is prevented by origin licensing ending once S phase has begun. There is no known means to undo inappropriate origin licensing, therefore licensing controls are employed to prevent all origin licensing outside of G1 phase. For example, in human cells CDT1 has been observed to directly inhibit CMG progression in S phase until it is fully degraded by $CRL4^{Cdt2}$, to prevent relicensing at origins from which helicases have already progressed [84]. This ensures that DNA synthesis cannot take place until Mcm2-7 loading has been successfully completed and so deters re-replication.

1.2.2.1 Initial DNA Melting

Binding of Cdc45 and GINS to Mcm2-7 hexamers forms the pre-initiation complex (pre-IC), composed of the double hexamer with Sld2, Sld3/Treslin and Sld7/MTBP, Dpb11/TopBP1, Cdc45, Pol Ɛ and GINS. The pre-IC requires activation by additional factors before it can unwind duplex DNA and support replication. These factors include the fork protection complex (FPC) which acts to ensure unperturbed fork progression and in higher eukaryotes is composed of Timeless-Tipin and Claspin [85]. Additionally, Mcm10, RecQ4 and Polα-Primase are required to support DNA melting by the CMG.

As previously mentioned, in yeast the stable CMG and Pol Ɛ form a complex ((dCMG)-Pol Ɛ) that distorts the double helix and the linearisation of the central channel of the dimerised MCM hexamers [76]. As a result, the two MCM rings rotate, resulting in the N-terminal domains separating, leaving them tethered

to one another only via the Mcm6 zinc finger (ZnF) domain. This exposes the duplex DNA between the hexamers. Double hexamer loading during licensing requires hydrolysis of adenosine triphosphate (ATP), leaving adenosine diphosphate (ADP) stably bound to each complex. The initial untwisting of the duplex DNA is coupled to the release of this ADP from the hexamer. An additional ATP then binds each hexamer and is required for preliminary melting of the duplex DNA [86], [87]. In yeast, ATP binding has also been observed to result in the release of Dpb11, Sld3-7 and Sld2 from the DNA [86]. The double hexamer then separates into single hexamers, and each CMG complex is separated from its partner. Initial DNA untwisting then occurs, 6-7 bp is observed per CMG [86]. An additional factor, Mcm10, is required to promote hydrolysis of the ATP bound to each hexamer, which activates each CMG as a result. The CMG can assemble on the chromatin without the need for Mcm10, but this complex is inactive [59]. Each helicase ejects one strand of DNA from the central channel. Each of these strands then crosses over to the other helicase, becoming the translocation strand of this hexamer [86], [87]. Each helicase subsequently translocates in a 3'-5' direction on the leading strand template. The N-terminal domain of each MCM hexamer is positioned at the front of the helicase, indicating that each CMG passes the other by translocating from the origin in opposite directions. Origin firing only occurs if both helicases transition from binding dsDNA to binding ssDNA, further safeguarding the process. Following replisome activation, several factors are removed from the complex and do not travel with the fork throughout replication. In yeast these include Sld2, Sld3, Sld7 and Dpb11 [87]–[90].

The molecular mechanism by which DNA is melted is less well established in higher eukaryotes. However, interestingly a cryo-EM structure of a human Mcm2-7 double hexamer was determined, which showed the origin DNA partially melted [91]. This suggests that the initial melting of the DNA may occur much earlier in human cells – and perhaps additional higher eukaryotes – than in yeast. However, like in yeast, Mcm10 has also been seen to be required to support efficient metazoan DNA replication [92]. Mcm10 has been observed to bind to chromatin at a late stage of initiation in *Xenopus laevis* egg extract [93]. Here, Mcm10 is likely required to support initial DNA unwinding to promote DNA synthesis following formation of the CMG, as it is in yeast. However, upon knockdown of Mcm10, DNA replication can still proceed,

albeit at a much-reduced rate. Therefore, it appears that the DNA melting role of metazoan Mcm10 can be taken over by an alternative protein in its absence.

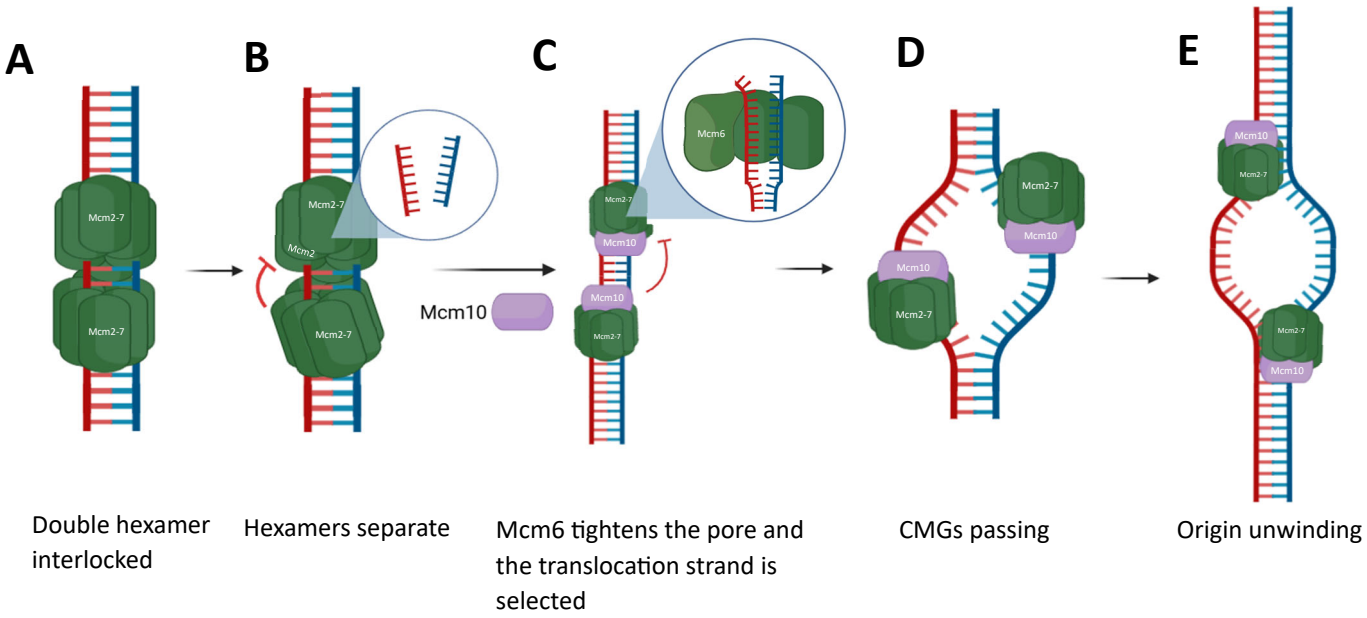


Figure 1.4 Schematic of DNA melting. Following CMG formation on the DNA (Mcm2-7 hexamers shown for simplicity in green) (A), the two hexamers separate slightly and Mcm2 causes the initial breakage between the base pairs (B), Mcm10 (purple) is recruited, the hexamers unpair and Mcm6 tightens the Mcm pore whilst selecting the translocation strand (C), the initially melted DNA is rewound and the two helicases pass one another on ssDNA (D), the helicases translocate on the DNA and replication is promoted (E). Adapted from [94].

1.3 DNA Replication Stress

The progression of the fork must remain highly regulated, to prevent or detect stress and genomic instability. To achieve this, factors are recruited that ensure the fork is stabilised, particularly those forming the fork protection complex (FPC). Timeless-Tipin (Tof1 and Csm3 in *S. cerevisiae*) and Mrc1 (yeast)/Claspin (higher eukaryotes) maintain fork stabilisation, regulation of fork speed and sister chromatid cohesion [85], [94]. Additionally, And-1 (Ctf4 in yeast), as previously mentioned, maintains the coupling of CMGs and polymerases [95]. However, when these maintenance pathways fail DNA replication can go awry and replication stress can occur.

DNA replication stress is characterised by the slowing or stalling of the replication fork along with DNA synthesis [96]. DNA replication stress is the dominant endogenous source of genome instability, which leads to ageing, neuropathology and cancer in humans [97], [98]. Replication forks frequently become stressed during S phase, causing them to stall and fail to terminate [5]. Replication stress can be generated from a wide range of physical obstacles and largely results in the formation of stretches of

ssDNA. Most frequently this occurs due to polymerase stalling and uncoupling from the CMG, which continues to unwind the DNA [96]. However, stress can also be induced by protein-DNA complexes or inter-strand DNA crosslinks, which can inhibit the helicase unwinding. Helicase-polymerase uncoupling results in stretches of single-stranded DNA (ssDNA) [99]. RPA will bind these ssDNA stretches and, if adjacent to newly replicated double-stranded DNA (dsDNA), this triggers the S phase checkpoint. This checkpoint is in place to determine if DNA has been accurately duplicated, and once triggered generates a signal to activate the replication stress response [40], [100]. As a result, the protein kinase ataxia-telangiectasia and Rad3-related (ATR) is recruited. Another protein kinase, Ataxia-telangiectasia mutated (ATM), is required for the DNA damage response and can also be activated by other forms of damage. ATR (Mec1 in yeast) is activated by a broader spectrum of damage than ATM and in particular is central to the DNA replication stress response. In human cells, ATR primarily exists in a complex with ATR-interacting protein (ATRIP) and following activation by RPA, ATR is further activated by TopBP1 and Ewing tumour-associated antigen 1 (ETAA1) [101], [102]. Upon its activation, ATR phosphorylates substrates that support the completion of DNA replication under stress (Figure 1.5). These include RPA and CHK1, which act to inhibit cell-cycle progression and suppress further origin firing, buying time for the replication stress to be resolved [96]. Activation of CHK1 leads to degradation of CDC25A phosphatase, resulting in the accumulation of phosphorylation on S phase Cyclin/CDKs, namely Cyclin A/CDK2. This phosphorylation, at sites such as Thr14 and Tyr15, is inhibitory to the origin firing function of the kinase [103]. ATR itself locally supports stabilisation and restart of the stalled forks once the stress has been removed. One method of replication downregulation is prevention of DDK-induced phosphorylation of MCM4, to inhibit the formation of the CMG during S phase. Upon ATR inhibition (ATRi), hyper-phosphorylation of MCM4 is observed in human cells, along with increased association of PSF3 with nascent DNA, both of which can be counteracted upon supplementation with a Cdc7 kinase inhibitor [17].

However, in the instance that the stalled fork cannot be restarted, new replication forks can be activated. This can occur through the firing of dormant origins, that have not been licensed, to allow for timely completion of replication [104]. Alternatively, the replisome can bypass the lesion, reprime the DNA and restart replication downstream, this leaves behind a ssDNA gap [105]. This is filled using the DNA damage

tolerance (DDT) pathway, involving specialised polymerases or the sister chromatid as a template, to complete replication and prevent the replication fork stalling for prolonged periods [106]. Fork reversal is an additional mechanism by which cells can tolerate replication stress following encounters with DNA lesions. Replication fork reversal requires recruitment of fork remodelling dsDNA translocases SMARCAL1, ZRANB3 and HLTF [107]. The remodeller recruited depends on the structure of the replication intermediate or type of replication challenge. During fork reversal, the two newly synthesised DNA strands are annealed to either remove or bypass the lesion to ensure replication can continue [108]. However, although fork reversal can stabilise stalled forks, they must be protected from nuclease cleavage to maintain genome stability [109], [110]. In addition to these mechanisms, the G2 checkpoint is a later safeguarding mechanism, triggered by any remaining DNA damage to ensure the cell does not enter mitosis with these issues [4].

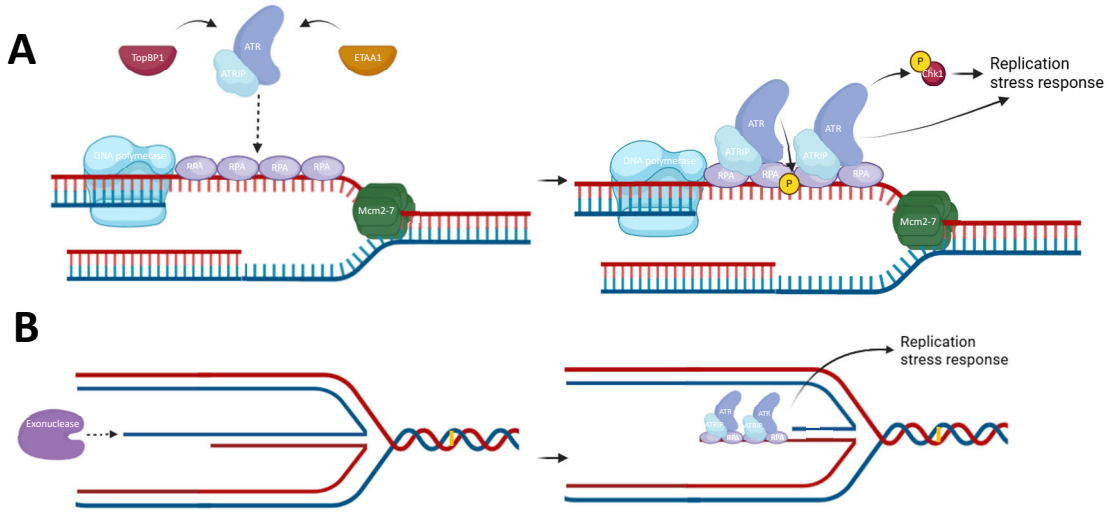


Figure 1.5 Schematic of ATR stress response activation. A) Helicase-polymerase uncoupling is presented, a common cause of replication stress and activation of the ATR stress response. Stretches of ssDNA are revealed as a result of uncoupling, which are coated with RPA. This triggers the recruited of ATR-ATRIP. ATR is further activated by either TopBP1 or ETAA1. This causes ATR to phosphorylate downstream effectors, such as RPA and CHK1 as presented here. As a result, replication is paused and DNA is repaired. B) The ATR stress response can also be triggered in response to DNA lesions, such as ICLs as shown here. In such cases, forks are stalled entirely and fork reversal is required. Specific exonucleases are then required to resect the DNA to produce the ssDNA required to activate the ATR stress response. Adapted from [110].

Although the cell initiates a robust stress response to stabilise and restart the replication fork, the fork may nonetheless fail and collapse. This is particularly likely if replication stress remains or factors involved in the stress response are lost. For example, in the absence of ATR, the replisome is likely to become disengaged [111]. This can result in complete fork collapse and under-replication of the DNA. If DNA is not

completely replicated during S phase this can lead to the formation of ultrafine bridges (UFB) in anaphase as well as the formation of secondary DNA structures, such as hairpins and quadruplexes, during the next cell cycle [112], [113]. Unreplicated regions of the genome are related with common fragile sites (CFS), loci at which the majority of the rearrangements occurring in human diseases are found [114]. The prolonged arrest of replication forks can also generate double strand breaks (DSBs) and likewise any nick or gap in the replicated DNA can be converted into a DSB [115]. Genome instability, of which DSBs are a form, is a hallmark of cancer cells. This therefore highlights the importance of fork protection and robust and efficient pathways that respond to stress.

1.4 Regulation of DNA Replication by Phosphorylation

Kinases are largely separated into families based on their amino acid substrate preference, for example CDKs have a specificity for serine-proline/threonine-proline motifs in a range of initiation factor substrates [116]. Whilst DDKs target serine and threonine residues in the Mcm2-7 complex [117]. For CDKs to act with enzymatic function, they require Cyclins as their regulatory subunits. The specific pairing between CDKs and Cyclins is dependent on the distinct cell cycle phase. Phosphorylation is utilised to enable interactions between replication components only under specific conditions and during the appropriate cell cycle phase. CDK levels are maintained at a low level during G1 to allow for Mcm2-7 hexamer loading. This is achieved via CDKs being degraded by the APC/C. Upon reduction of CDK-Cyclin levels, Cadherin 1 (Cdh1) is activated. Cdh1 competes with Cdc20 for binding to the APC. To promote origin licensing APC-Cdh1 supports accumulation of p27 and p21 (CKIs) and degradation of CDKs. However, not only does CDK and DDK activity need to be minimised, but any present phosphorylation of their substrates also needs to be removed. This is achieved through the activity of several phosphatases, which remove phosphorylation from target proteins, namely PP2A and PP1 (Cdc14 in yeast) [118]–[120].

In contrast, increased levels of S phase kinases are required to promote transition into S phase and support DNA replication. CKIs are degraded and result in the increase of Cyclin E-CDK2 and Cyclin A-CDK2. Cdh1 is in turn phosphorylated by these CDKs and Plk1, leading to its degradation and decrease in the activity of APC-Cdh1 [121], [122]. This marks the transition from G1 to S phase and stimulates formation

of the active helicase via recruitment of Cdc45 and GINS. DDK phosphorylation is essential for DNA replication in both yeast and metazoan, specifically for phosphorylation of chromatin-loaded Mcm4 and Mcm6 [123], [124]. This allows the recruitment of Sld3-Sld7/Treslin-MTBP with Cdc45 [125]. As previously mentioned, in yeast, Sld2 and Sld3 CDK-dependent phosphorylation is required to support CMG formation. Sld3 contains two sites that are CDK substrates, threonine 600 and serine 622, phosphorylation of which are required for its interaction with Dpb11, necessary to support replication [67]. These sites are also conserved in the Sld3 human orthologue Treslin/TICRR – threonine 669 and serine 1001 – indicating the importance of CDK phosphorylation for this process in higher eukaryotes. For recruitment of GINS to origins, CDK-dependent phosphorylation of Sld2 is also required, which likewise promotes its association with Dpb11. Strikingly, individual phosphorylation of either Sld2 or Sld3 has been shown to be insufficient to support either CMG formation or DNA replication in budding yeast, indicating the importance of phosphorylation for interactions at this stage [59]. Interestingly, whilst DDK activity is also essential for replication initiation, the order in which CDK and DDK activity occurs has been shown to be unimportant for efficient Cdc45 and GINS recruitment, and DNA replication [59]. Dephosphorylation of both Sld2 and Sld3 is also essential to support replication, in particular for the initial melting of the DNA. Protein phosphatase 2 (PP2A) specifically targets Sld3/Treslin. This supports dephosphorylation of Sld2 by protein phosphatase 4 (PP4) and ultimately removal of the factors from the CMG [126]. Dephosphorylation of both factors is essential for the MCM double hexamers to separate and for DNA melting to proceed.

Cyclin A-CDK2 forms at the onset of S phase, and allows DNA synthesis and progression to G2 [127]. In the presence of active CDKs, namely Cyclin A-CDK2, the ATPase activity of ORC and its interaction with Cdt1 in *S. cerevisiae*, and release from origins in higher eukaryotes, is prevented [128]–[130]. Phosphorylation of Cdc6 by CDKs results in its degradation, via proteolysis in *S. cerevisiae*, and nuclear export in higher eukaryotes [131]. Likewise, in *S. cerevisiae,* phosphorylation of Mcm2-7 hexamers and Cdt1 by late G1 Cyclins leads to their nuclear export [41]. Therefore, whilst CDK activity is a regulatory mechanism to prevent Mcm2-7 chromatin recruitment during S phase, to avoid re-replication of the DNA, low CDK activity is required for Mcm2-7 to be recruited to origins during G1.

1.4.1 Replication Timing Control

Apart from safeguarding the cell from DNA re-replication, phosphorylation plays an important role in the replication timing programme – dictating which regions of the genome are replicated during early or late S phase. Euchromatin is replicated first, whilst less accessible heterochromatin is replicated subsequently [132]. Likewise, a pattern in which a subset of origins (early) are fired prior to the remaining origins (late) is established [133]. Additionally, certain regions of the nuclei replicate with different timing. In particular, the nucleolus replicates late during S phase [134]. Collectively, the replication timing programme is largely promoted by S phase kinase-specific activity, which is required to promote DNA replication. For example, in regions with more compact chromatin structures (heterochromatin), this may delay DDK-dependent phosphorylation of Mcm2-7, and thus delay origin firing until later during S phase [135].

Progression from G1 to S phase is regulated by Cyclin D activating either CDK4 or CDK6, as well as Cyclin E-CDK2 [136]. Cyclin A-CDK2 regulates progression from G2 to M phase, while Cyclin B-CDK1 regulates mitosis and chromosome segregation [137]. Origins will fire later if the phosphorylation of the replication-specific substrates is prevented. For example, PP1 counteracts the DDK-dependent phosphorylation of Mcm2-7 at licensed origins. PP1 is recruited to the chromatin in a Rif1-dependent manner, and collectively their activity delays the firing of a subset of origins [138]. However, upon the DDK-activity increase during S phase, the interaction between Rif1 and PP1 can be disrupted, supporting further origin firing as S phase progresses [139], [140]. Likewise, increased origin firing can speed up S phase and increased S phase kinase activation promotes this. For example, the phosphatase Cdc25B/C dephosphorylates Cdk1 and allows it, with Cyclin B, to push cells into mitosis. This increases origin firing and therefore promotes a more rapid S phase [141].

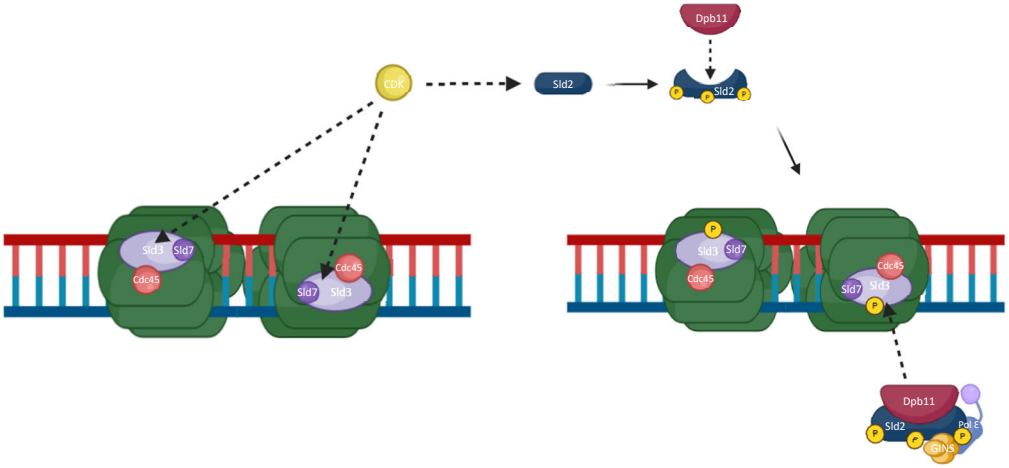


Figure 1.6 Schematic of CDK activity during CMG formation. Following DDK-dependent phosphorylation of the Mcm2-7 hexamer, Sld3-Sld7-Cdc45 is recruited (not shown). CDK phosphorylates Sld3 and Sld2. Sld2 phosphorylation causes a conformational change to the protein allowing Dpb11 to bind. Sld3 phosphorylation also allows Dpb11 to bind. Therefore, Dpb11, along with Sld2, GINS and Pol Ɛ are recruited to the hexamer through an interaction between Dpb11 and phosphorylated Sld3, forming the CMG. Adapted from [138].

As previously mentioned, the protein kinase Mec1/ATR functions in response to DNA damage during S phase, phosphorylating Rad53/Chk1/CHK1 and (indirectly) Chk2/CHK2 to downregulate new origin firing whilst damaged DNA is repaired [142]. This is achieved by CHK1 and CHK2 (only CHK1 in mammals) targeting the Cdc25 family of phosphatases. Under normal replication conditions Cdc25 acts to support activation of CDKs. Therefore, the CHK1 stress response targeting Cdc25A for degradation promotes inhibition of CDK and therefore downregulates DNA replication initiation [142]. The yeast homologue of CHK2, Rad53, also appears to inhibit DDK function by blocking the kinase from binding Mcm2-7, thus preventing formation of the active helicase [141]. CHK1 is also regulated at the G2-M boundary through the phosphorylation of Claspin by Polo like kinase 1 (Plk1). Claspin phosphorylation results in its degradation by the proteasome. This blocks CHK1-dependent inhibition of mitosis upon resolution of DNA damage [143]. By inhibiting origin firing, the checkpoint delays S phase progression to avoid the accumulation of topological stress and allow time for DNA repair [144]. To resume replication following repair, the checkpoint signalling cascade must be inactivated, this largely involves phosphatases, such as Pph3 and Ptc2/Ptc3 in yeast [145], [146].

1.5 DONSON

Downstream neighbour of Son (DONSON) is a protein whose gene was first identified when assessing the long arm of human chromosome 21 (21q22) [147]. When assessing mRNA levels by reverse transcription polymerase chain reaction, DONSON, here termed B17, was found expressed in a range of different cell types, including HeLa and foetal cells. The gene was predicted to contain seven exons and span 10 kb. The gene was first termed 'downstream neighbour of Son'/DONSON later that same year, when it was identified in a tail-to-tail orientation with its upstream neighbour Son [148]. Here, DONSON was found to be just 65 bp downstream from Son in mouse genomes, and 415 bp downstream in human genomes.

1.5.1 Orthologues

Despite the gene having been identified at the turn of the century, a function for DONSON was not proposed for many years. Through a mutagenesis screen in *Drosophila*, a protein homologous to DONSON, here named Humpty Dumpty (Hd), was identified as essential for DNA replication and genome stability in larval tissues [149]. Likewise, Hd was shown to be essential for cell proliferation during development, and upon mutation results in a 'thin eggshell' phenotype. Consistent with a DNA replication role, Hd protein expression levels were observed to peak during late G1-S phase, and in response to the E2F Transcription Factor 1 (E2F1). Therefore, Hd was labelled a gene important for DNA replication and subsequently orthologues were screened for. Interestingly, the gene DONSON was identified by sequence alignment as the Hd orthologue in mammals.

Using a computational approach followed by RNAi screening in HeLa cells, DONSON was found to share phenotype similarities with factors known to be involved in the DNA damage response (DDR) and therefore a potential role for DONSON in cell cycle progression was proposed [150]. DONSON has also been observed to be expressed in proliferating tissues particularly within the human embryonic brain, and has been seen co-expressed with DNA replication machinery components [151]. Likewise, in mice, Donson was observed as essential to the cell cycle, with knockout of the gene embryonic lethal. Collectively, this suggested an essential role for DONSON during DNA replication and the cell cycle, conserved across many organisms from plants to humans. Interestingly however, no DONSON orthologue was identified in yeast. Much of what is known about DONSON is due to its mutation being observed in a

number of developmental disorders, namely Microcephalic Primordial Dwarfism (MPD) and Meier Gorlin syndrome (MGORS) [151]–[155].

1.5.2 DONSON Gene and Protein Structure

Interestingly, the DONSON protein within human cells is observed at very low levels, and doesn't appear to have particularly recognisable structural features or enzymatic activity [150], [156]. Likewise, neither the gene sequence nor the protein structure are similar to any other known protein [157]. The DONSON gene has since been identified as having ten exons in the longest open reading frame (ORF) of the human gene, and eleven in the *Xenopus*. Using AlphaFold, the structure of the DONSON protein could be predicted. Based on this predicted structure, DONSON can be divided into six domains, with a long and predominantly unstructured N-terminus composed of 150 amino acids in humans (Figure 1.7), which wraps around a folded globular core and joined via a protruding flexible loop [158], [157]. The core of the protein is highly conserved throughout evolution, with a confident AlphaFold score, whereas the flexible N-terminus is not.

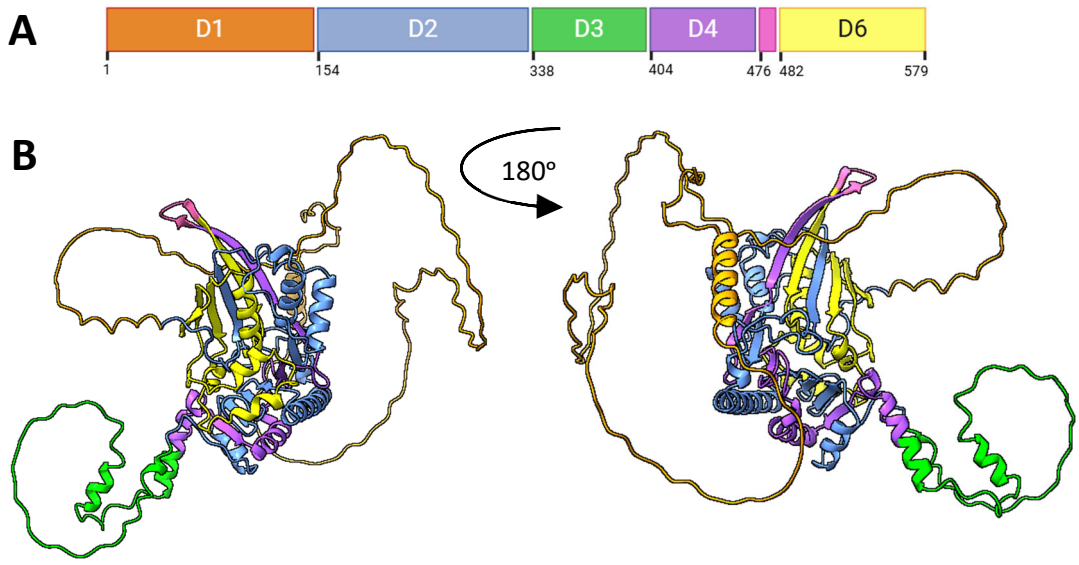


Figure 1.7 Model of *Xenopus* DONSON protein. A) DONSON is split into six domains of varying lengths. B) Each has been colour-coded and shown in the structural form as modelled via AlphaFold. As seen, the N-terminal domain (D1 – orange) is flexible and largely wraps around the globular portion of the protein.

1.5.3 DONSON's Association with the Replisome

Consistent with the identification that DONSON's peak expression occurs in S phase and that it was proposed to play a role in the DNA damage response, it has also been observed interacting with known replication machinery factors in human cells, both under normal conditions and upon DNA damage induction [35], [152]. Using mass spectrometry screening following an IP of Mcm3, DONSON was seen to interact specifically with Mcm2-7 subunits, as well as components of the GINS complex during replication and post-replication termination [35]. Likewise, through proximity-ligation assay (PLA), DONSON has been seen in the proximity of MCM2 and MCM7, as well as co-immunoprecipitating with Treslin, PCNA and RPA [152], [156].

1.5.3.1 Heterochromatin v Euchromatin

Genome structure is largely classified into two major domain types, the first being euchromatin characterised by accessibility, transcriptional activity and active histone modifications, such as H3K4me^3. The second type is heterochromatin, characterised by condensed DNA with inactive genes, repeated elements and inactive histone modifications, such as H3K9me^3 [156]. Sequences in euchromatin tend to be replicated during early S phase, whilst those in heterochromatin tend to be replicated during late. DONSON has been observed to not bind all replisomes, instead predominantly appearing bound to replisomes in euchromatin and during early S phase. In human cells, DONSON exhibits a >3-fold bias towards early S phase for its interaction with components of the GINS complex. Likewise, interaction between DONSON and H3K4me^3 can be observed throughout the nuclei and are threefold more frequent in early S phase than late, whilst very little interaction can be observed between DONSON and H3K9me^3 in either early or late S phase. The presence of DONSON-bound replisomes in active chromatin may suggest a potential cellular anticipation for replication stress in these open chromatin regions where replication-transcription collisions are more prone to happening and DNA damage is more likely, and therefore may suggest why DONSON preferentially binds replisomes in euchromatin [156].

1.5.4 DONSON's Fork Protection Role

As mentioned above, DNA replication stress is characterised by the slowing or stalling of the replication fork along with DNA synthesis and often involves the accumulation of ssDNA. The *Drosophila* orthologue,

Hd, has been shown to be required for genome stability, as upon its depletion a 'chromosome bridge' phenotype develops during the cleavage cycles of the embryos. This results from a sensitivity to replication defects, and results in incompletely replicated sister chromatids that cannot segregate [159]. Maternal Hd is required for maintenance of genome integrity during embryonic S phase and suggests that DONSON may perform a similar function for genome stability.

In human cells deficient for DONSON, ATR-dependent replication stress signalling is impaired, resulting in stalled and damaged forks not being repaired or restarted. Increased replication fork stalling prolongs S phase and increases origin firing. In line with this, DONSON downregulation results in a loss of suppression of new origin firing, suggesting a block to checkpoint activity in the absence of DONSON function [152]. This indicates a direct function for DONSON within the ATR-dependent stress response pathway. Reduced ATR and CHK1 phosphorylation are likewise observed upon downregulation of DONSON following hydroxyurea (HU) exposure. Additionally, RPA phosphorylation was strongly reduced upon DONSON depletion, suggesting that DONSON acts downstream of ATM/ATR, but upstream of RPA and CHK1.

DONSON appears to not only function in response to replication stress, but also once fork stalling has progressed to DNA breaks or DNA damage. DONSON has specifically been observed interacting with a subset of replisomes in human cells perturbed with hydroxyurea (HU) to stall forks, most of which are in early S phase, supporting a role for the protein in the DNA damage response (DDR) [152]. Phosphorylation of the H2AX histone variant at Serine 139 results in the γH2AX form and is an early cellular response to the induction of DNA damage. In contrast, 53BP1 foci form as a response to DNA DSBs. Upon depletion of DONSON, elevated levels of both γH2AX and 53BP1 foci could be observed in human cells in S phase [150], [152]. This indicates that the spontaneous DNA damage observed in the absence of DONSON is likely generated during DNA replication, a probable result of fork collapse. Interestingly, γH2AX phosphorylation exhibited in DONSON-depleted cells can be partially reduced upon the depletion of the nucleases responsible for the processing of stalled and damaged forks, MUS81 or XPF [152]. This indicates that the genomic instability observed upon DONSON knockdown results from nucleolytic processing, by these nucleases, of the resulting damaged forks. In addition to this, cells in

which DONSON is downregulated display various abnormalities, particularly irregular chromosomal morphology, such as misaligned chromosomes, chromatid gaps and breaks, spontaneous micronuclei formation and pulverisation of metaphase chromosomes [152]. DONSON has also been observed at the centrosome, required for accurate formation of bipolar spindles and this, as well as its requirement for fork stabilisation, may explain the abnormal chromosomal phenotypes observed upon its downregulation [150]. This suggests that DONSON is required to prevent transmission of DNA breaks into mitosis to prevent chromosomal damage and genome fragmentation.

In addition to a stress-response role, DONSON appears to regulate unperturbed forks, maintaining stability during replication [152]. Upon downregulation of DONSON in human cells, increased spontaneous fork stalling is observed in combination with fewer ongoing forks, despite an increase in BrdU-positive cells. Spontaneous DNA damage and fork instability are observed upon DONSON downregulation, characterised by fork asymmetry, suggesting that the increase in BrdU-positive cells likely reflects a prolonged S phase due to stalled forks, as opposed to a global reduction in synthesis. Collectively, this suggests that DONSON may stabilise ongoing forks. Interestingly, the amount of RPA recruited to chromatin is also increased upon DONSON's knockdown, therefore suggesting that ATR may not be efficiently recruited to these spontaneously formed DNA lesions without DONSON [152].

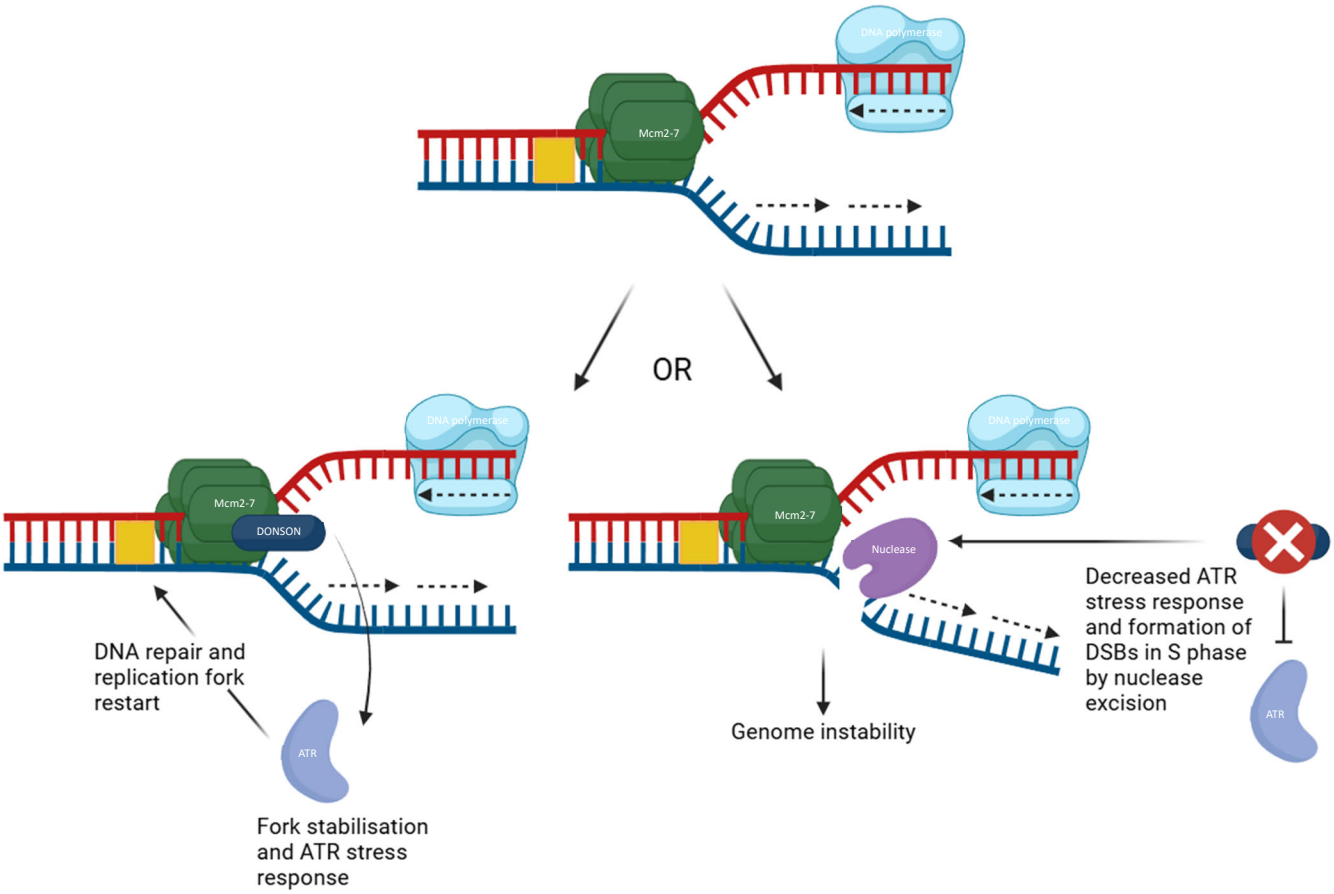


Figure 1.8 Schematic of replication stress response with or without DONSON. Upon replication stress (yellow square), possibly due to polymerase and helicase uncoupling the replication fork will stall (shown here with green Mcm2-7 hexamer only, for simplicity). With functional DONSON present, the fork is stabilised and the DNA can be repaired and replication restarted through the ATR stress response. However, without functional DONSON (red crossed), the ATR stress response is decreased and the fork not stabilised, leaving it open to nuclease excision. This leads to genome instability. Adapted from [152].

1.5.4.1 ICL Traversal

Interstrand crosslinks (ICLs) occur when a DNA damaging agent reacts with nucleotides on opposite DNA strands, promoting a covalent link between the two strands [160]. Unsurprisingly, these adducts cause problems for the replication machinery when it encounters them, as the duplex DNA cannot be unwound. Forks are often stalled by ICLs and therefore ICLs must be repaired. However, replication forks can also traverse past an ICL in an ATR-dependent manner, to prevent fork collapse and retain genome stability [160]. The Fanconi Anaemia Complementation Group M (FANCM) helicase is implicated in ICL traversal. Upon the downregulation of FANCM, ICL traversal is reduced. Interestingly, DONSON has also been observed to be important for ICL traversal, particularly for the survival of cells exposed to cisplatin, an ICL promoting molecule [156]. Upon DONSON downregulation, the frequency of ICL traversal in human cells

is markedly reduced. Strikingly, traversal frequency reduces further still upon knockdown of both DONSON and FANCM, indicating that the proteins are non-epistatic for this role.

As mentioned above, DONSON has largely been observed to interact with the CMG in euchromatin. In contrast, FANCM has been observed bound to the replisome in heterochromatin. This correlates with siRNA knockdown of DONSON in human cells which results in a decrease in the frequency of ICL traverse in early S phase cells only, and minimal change to traverse frequency in late S phase cells [156]. However, unlike FANCM, DONSON can be observed interacting with the replisome both with and without the GINS complex intact, whilst upon FANCM interaction, replisomes are remodelled and GINS is lost from the complex. GINS remaining intact in the DONSON-bound CMG, appears to correlate with unstressed replisomes not marked by ATR activation. This demonstrates that DONSON associates with replisomes in cells both with and without exposure to DNA damaging agents, further supporting dual roles in maintaining stability of unperturbed forks and acting in the ATR stress response pathway [156].

ATR plays a role in ICL traversal of both DONSON- and FANCM-replisomes. Upon ATR inhibition, the association of FANCM with the replisome is blocked [161]. Upon detection of ICL adjacent to FANCM-replisome, FANCM promotes the recruitment of the Fanconi Anaemia (FA) core complex to ICL sites [162]. Following recruitment, ATR phosphorylates Fanconi Anamia Complementation Group I (FANCI) as part of the complex, supporting recruitment of FA-associated nucleases to nick the DNA either side of the ICL and promote downstream repair [163]. In contrast, DONSON's association with ICLs increases upon inhibition of ATR. This is likely due to the replisome not being remodelled in the absence of ATR activity meaning that GINS is retained and the DONSON-bound replisome remains stalled at ICLs [156].

1.5.5 Primordial Dwarfism

Microcephalic primordial dwarfism (MPD) largely characterises a wide range of Mendelian syndromes, the principal features of which are intra-uterine and postnatal growth delay beside microcephaly, that occur in humans [164]. MPD predominantly results from mutation of genes encoding replisome components or proteins required to maintain genome stability [152]. Using whole exome sequencing (WES) on patients presenting with characteristics of MPD, mutation of DONSON has been implicated in

microcephalic dwarfism, as an autosomal recessive trait (Table 1.1). Within the study, all patients with DONSON mutations presented with similar clinical phenotypes, particularly microcephaly: reduced cerebral cortical size with decreased gyral folding. DONSON is required for correct embryonic development and its complete loss is lethal [151]. In mice, deletion of *Donson* in progenitors of specific neurons has been observed to result in widespread programmed cell death at early stages of cortical development. Likewise, they do not develop the hippocampus or dorsolateral neocortex [165]. This phenotype is conserved in humans where DONSON is expressed in the embryonic human brain in progenitor cells and is co-expressed with DNA replication machinery factors [151]. This may therefore explain the microcephaly observed in patients with MPD, as DONSON appears to be particularly important for the brain development. However, with MPD height is markedly less reduced than that observed in other microcephalic dwarfism-associated disorders [152]. Additionally, some patients present with minor skeletal abnormalities, but neither a recognisable facial phenotype nor organ abnormalities are observed.

In the most severe form of MPD, Microcephaly Micromelia Syndrome (MMS), the vast majority of DONSON transcripts are incorrectly spliced, whilst in less severe forms of MPD, the issues arise from hypomorphic mutations which result in a marked decrease in protein expression [151]–[153]. In some cases, this decreased protein expression can be explained by mRNA products that are out of frame and as a result likely undergo nonsense-mediated mRNA decay. Whilst in other cases, the mutation can directly compromise protein expression levels or subcellular localisation [152]. As mentioned above, downregulating DONSON in cells results in increased spontaneous fork stalling and decreased fork repair and restart resulting from impaired ATR-stress response activity. Therefore, phenotypes observed in patients with MPD DONSON mutations are likely due to a decrease in faithful and timely genome duplication which impacts the number of cells generated during intra-uterine development.

Table 1.1 Known DONSON mutations implicated in developmental disorders. Mutation type, site, resulting syndrome and references presented.

Mutation Type	Nucleotide Mutation	Phenotype	Reference
Substitution	c.832T>C	Microcephaly, short stature and limb abnormalities (MISSLA)	[152]
Substitution	c.876C>G	MISSLA	[152]
Substitution	c.1466A-C	MISSLA	[152]
Deletion	c.419-566	MISSLA	[152]
Deletion	c.428-566	MISSLA	[152]
Deletion	c.1251-1256	Microcephalic Dwarfism	[152]
Substitution	c.877C>T	Microcephalic Dwarfism	[152]
Duplication	c.1254dupT	Microcephalic Dwarfism	[152]
Duplication	c.1686dupT	Microcephalic Dwarfism	[152]
Substitution	c.1282C>T	Microcephalic Dwarfism	[152]
Deletion and substitution	c.1474-1475 deletion and 786-7T>C	Microcephalic Dwarfism	[152]
Double substitution	c.876C>G and 1047-9A>G	Microcephalic Dwarfism	[152]
Duplication and substitution	c.1628-1630dupAAA and 1032C>T	Microcephalic Dwarfism	[152]
Substitution	c.786-22A>G	Microcephalic Dwarfism	[152]
Substitution	c.1337T>C	Microcephalic Dwarfism	[152]
Substitution	c.1297C>T	Microcephalic Dwarfism	[152]
Duplication and substitution	c.1254dupT and 1510G>A	Microcephalic Dwarfism	[152]
Substitution	c.1047-9A>G	Microcephaly-Micromelia	[151]
Substitution	c.763C>T	Microcephaly-Micromelia	[152]
Deletion	c.346delG	MGORS	[166]
Substitution	c.1349A>G	MGORS	[166]
Substitution	c.683G>T	MGORS	[166]
Substitution	c.494T>C	MGORS	[154]
Substitution	c.607-36G>A	MGORS	[154]
Substitution	c.631C>T	MGORS	[154]
Substitution	c.670C>T	MGORS	[154]
Substitution	c.809A>G	MGORS	[154]
Substitution	c.1297C>T	MGORS	[152]
Substitution	c.1634C>T	MGORS	[154]

Meier Gorlin syndrome (MGORS) is a form of MPD that is characterised by variations in genes of known cell cycle factors particularly involved in DNA replication initiation, such as ORC1, ORC4, ORC6, CDT1, CDC6, CDC45, GMNN, MCM3, MCM5, MCM7, GINS2 and GINS3 [166]–[168]. MGORS clinical phenotypes include short stature, microtia and patella underdevelopment or absence [91]. Interestingly, DONSON mutation has also been observed in MGORS patients. Whole genome/exome sequencing of patients displaying characteristic MGORS phenotypes found missense mutations of DONSON in a number of cases [154], [155]. Many MGORS patient DONSON variants have been observed to result in reduced nuclear localisation of the protein or frameshifts resulting in a premature stop codon [154]. MGORS mutations are spread throughout the sequence of DONSON, not clustering at any part of the protein. It is interesting that DONSON mutation is implicated in MGORS, as the syndrome is largely characterised by mutation of DNA replication initiation-associated factors only.

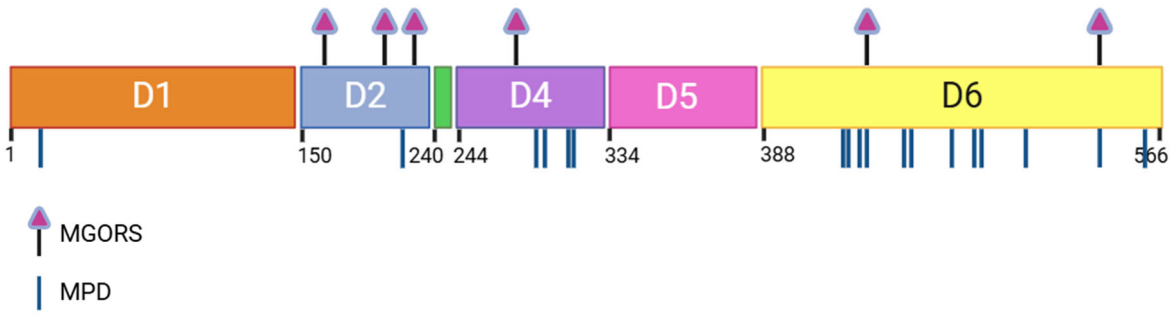


Figure 1.9 Map of human DONSON with known disease mutations mapped. The six domains of human DONSON are presented with the corresponding amino acid positions of each domain presented. Primordial dwarfism mutations observed in patients have been presented – specifically, Meier Gorlin (MGORS) mutation positions have been presented above the map with pink triangles, and remaining MPD mutations presented as blue circles below. Adapted from [152] and [157].

1.5.6 Cancer

Overexpression of DONSON has been implicated in a number of different cancer types. Analysis of The Cancer Genome Atlas (TCGA) revealed that in renal cell carcinoma, DONSON overexpression is related to poor prognosis for patients [169]. Likewise, according to TCGA and Gene Expression Omnibus (GEO), in prostate carcinoma DONSON is required for cell cycle progression and to maintain genomic stability in these cells [170]. Increased DONSON expression has also been observed in breast cancer tissue and cell lines, with overexpression correlating with lymph node metastasis and poor patient prognosis [171]. Elevated expression of DONSON circular-RNA (circ-RNA) has also been implicated in gastric cancer [172].

Circ-RNA is a form of noncoding RNA that forms a covalently closed continuous loop making it more resistant to nuclease degradation. Circ-DONSON is formed by back-splicing DONSON mRNA and appears to function as an oncogene in gastric cancer. Circ-RNAs are highly expressed in tumour cells where they can regulate numerous processes, including cell proliferation and survival. Circ-DONSON appears to recruit both chromatin remodellers and transcription factors shown to promote cancer development [173]–[176].

1.6 *Xenopus laevis* Egg Extract as a Model System

Xenopus laevis egg extract provides *in vitro* insight into biochemical and molecular mechanisms of replication and is a powerful tool for investigating complex activities that occur *in vivo.* Egg extracts are particularly useful as they are highly synchronous, cell-free and are able to recapitulate the principal nuclear transitions that occur in eukaryotes, therefore making this an ideal system for cell cycle research [177], [178], [179]. They are frequently used to study cellular events such as DNA replication and mitosis, activation of the S phase checkpoint, fork stalling and collapse, DNA repair, inter-strand crosslink repair, spindle microtubule dynamics, nuclear transport and apoptosis [177], [180].

The Gambus lab uses *Xenopus laevis* egg extract to study DNA replication, as it is rich in enzymes and precursors for DNA replication and chromatin assembly. In addition, transcription does not occur in the extract and it is the only system supporting study of vertebrate DNA replication outside the cell and which supports all phases of genome maintenance. Additionally, both replication and repair in *Xenopus* is highly conserved with mammalian systems, with functions uncovered in the egg extract system also conserved in human cells [179], [181].

Unfertilised *Xenopus* eggs are arrested at metaphase of meiosis II and form three different forms of egg extract based on centrifugation at different speeds – Low-Speed Supernatant (LSS), High-Speed Supernatant (HSS) and Nucleoplasmic Extract (NPE). Conventional formation of egg extract involves crushing the eggs via low speed centrifugation to collect the cytoplasm – this is LSS [179]. Throughout my PhD project, I have predominantly used LSS. Adding demembranated sperm DNA to egg extract promotes decondensation of sperm, as well as formation of a nuclear envelope around the sperm DNA and activation of one round of semi-conservative DNA replication. Egg extracts are supplemented with

cycloheximide (CHX) to ensure that only a single round of replication occurs and further cell cycle progression is prevented [178].

We have largely used LSS in our studies during my project and in the lab in general, to specifically investigate replication initiation mechanisms. As egg extracts transition into different stages of the cell cycle, they can be manipulated with the aid of specific drugs or proteins to either arrest or release into each phase in order to support the correct study conditions. The role of proteins of interest can be investigated by immunodepleting them from the extract and assessing the effect on DNA replication. However, it is important that supplementing the extract with the recombinant protein rescues the phenotypes observed.

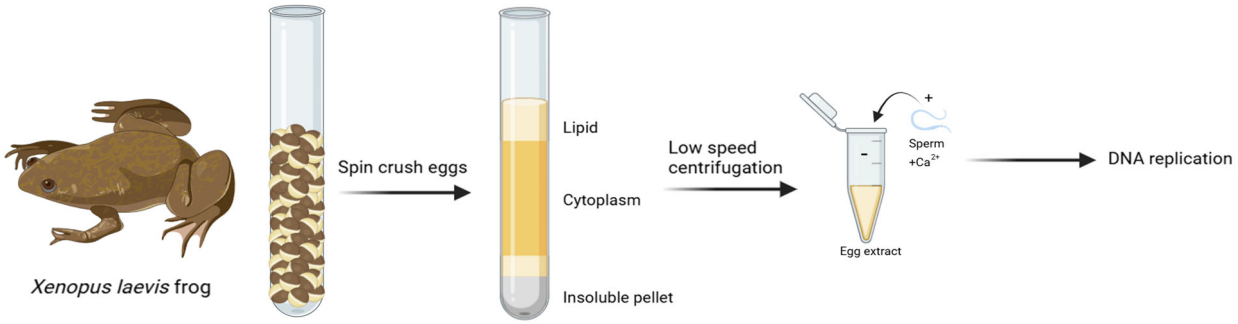


Figure 1.10 Schematic of *Xenopus laevis* egg extract preparation. Eggs retrieved from the frogs undergo multiple rounds of centrifugation to isolate and extract the cytoplasm, which can be activated and to which sperm DNA can be added to promote DNA replication. Adapted from [179].

1.7 Aims

Previous work in the Gambus lab identified DONSON interacting with the post-termination replisome, specifically co-immunoprecipitating with Mcm3 of the CMG. DONSON's interaction with the replisome has also been observed by others during elongation, as previously mentioned. Although DONSON's role as a fork protection factor is strongly reported, as well as its interaction with other known factors of the FPC, the molecular mechanism by which it carries out this function is currently unclear. Likewise, whether it has additional functions during different stages of DNA replication has not been reported. Of particular interest to us was the fact that DONSON interacts with the replisome during termination, suggesting that it may play a role in supporting replication termination and/or replisome disassembly.

The concept of additional roles for DONSON during replication has been speculated upon, additionally with regards to its mutation being implicated in Meier Gorlin syndrome, which is predominantly associated with mutation of DNA replication initiation factors. This was also of particular interest to us, as it suggests that either the syndrome is associated with mutation of factors involved later during replication, or that DONSON may play additional roles earlier in replication than previously proposed. In addition to this, the *Drosophila* orthologue, Hd, has been observed to localise to foci that frequently overlap with origins of replication [149].

Collectively, this led me to aim to determine the function of DONSON during DNA replication. To ascertain whether, in addition to its fork protection role, DONSON functions during replication initiation and/or termination. Secondly, I aim to determine whether DONSON has any further binding partners that support its function(s). Finally, I aim to ascertain whether post-translational modification of DONSON plays a role in these functions and potential complex formation with any partners identified. To achieve these aims I have used *Xenopus laevis* egg extract as a model system. Immunodepletion of DONSON from the extract, combined with addition of recombinant wild-type (WT) or mutated versions of DONSON enabled the study of the role of DONSON during DNA replication and to determine the mechanism of DONSON's action.

2 Materials and Methods

2.1 Tables

Table 2.1 Composition of Reagents Used

Buffer	*Composition*
Xenopus laevis Work	
10X Marc's Modified Ringer's (MMR)	1 M NaCl, 20 mM KCl, 10 mM $MgCl_2$, 20 mM $CaCl_2$, 1 mM EDTA, 50 mM HEPES, Adjust pH to 7.8 with NaOH Store at room temperature
Un-activating Extraction Buffer (UEB)	50 mM KCl, 50 mM HEPES, 5 mM $MgCl_2$, 5 mM EGTA, Adjust pH to 7.6 with KOH Store at 4°C
Cysteine solution	2.2% cysteine, 5 mM EGTA, adjust pH to 7.6 with KOH, Make up to 1 litre Freshly prepared
EB solution	50 mM KCl, 5 mM $MgCl_2$, 2 mM β-mercaptoethanol, 50 mM HEPES-KOH, pH 7.6 Store at 4°C
SuNaSp	0.25 M sucrose, 75 mM NaCl, 0.5 mM spermidine, 0.15 mM spermine, 15 mM HEPES-KOH pH 7.6 Store at 4°C
Acetate nuclear isolation buffer - 2X (ANIB) 100	100 mM HEPES pH 7.6, 200 mM KAc, 20 mM MgAc, 5 mM Mg-ATP (added from a stock of 250 mM ATP; 250 mM $MgCl_2$; pH 6.7 with NaOH), 1 mM spermidine, 0.6 mM spermine, 2 µg/ml of each: aprotinin, leupeptin, pepstatin, 50 mM β-glycerophosphate Store at 4°C
1x ANIB	Mix 1 volume of 2x ANIB100 with 1 volume of H_2O, 0.1% Triton X-100, 0.1 mM PMSF, 10 mM chloroacetamide Freshly prepared
Energy Regenerator (ER)	1 M phosphocreatine K salt, 600 µg/ml creatine phosphokinase, dissolve in 10 mM HEPES–KOH pH 7.6 Store at -20 °C
Cycloheximide	10 mg/ml in H_2O Store at -20 °C
$CaCl_2$	50 mM in H_2O, -20 °C
Cytochalasin D	10 mg/ml in DMSO Store at -20 °C
HEPES Solution	100 mM HEPES, pH 7.6 Store at -20 °C
Conjugation Buffer	20 mM sodium phosphate, 150 mM NaCl, pH7-9 Store at room temperature
Quenching Buffer	1 M Tris HCl, pH 7.5 Store at room temperature
100 mM BS3	25.8 mg BS3, dissolve in 450 µl conjugation buffer Freshly prepared
Licensing Factor Buffer (LFB1/50)	10% sucrose, 50 mM KCl, 40 mM HEPES pH 8, 20 mM K phosphate pH 8, 2 mM $MgCl_2$, 1 mM EGTA, 2 mM DTT, 1 µg/ml of each: aprotinin, leupeptin and pepstatin Store at 4°C
LFB1/100	10% sucrose, 100 mM KCl, 40 mM HEPES pH 8, 20 mM K phosphate pH 8, 2 mM $MgCl_2$, 1 mM EGTA, 2 mM DTT, 1 µg/ml each of each aprotinin, leupeptin and pepstatin Store at 4°C
Stop C	5 mM EGTA, 0.5% SDS, 20 mM Tris-HCl pH7.5 Store at 37°C
5% TCA solution	5% TCA, 0.5% Sodium pyrophosphate, make up to 5 L with H_2O Store at 4°C
10% TCA solution	10% TCA, 2% Sodium pyrophosphate, make up to 10 L with H_2O Store at 4°C

Protein Work	
1x PBS	137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 1.8 mM KH_2PO_4 Store at room temperature
1x PBST	137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 1.8 mM KH_2PO_4, 0.1% Tween-20 Store at room temperature
10x TBS	200 mM Tris Base, 1500 mM NaCl, pH 7.6 Store at room temperature
1x TBST	0.5 l of 10X TBS, 12.5 ml of 20% Tween 20, make up to 5 l with H_2O Store at room temperature
Tris Glycine running buffer	25 mM Tris-HCl pH 7.4, 250 mM Glycine, 0.1% SDS Store at room temperature
Transfer buffer	25 mM Tris-HCL pH 7.4, 190 mM Glycine, 20% methanol Store at room temperature
Coomassie Blue Stain	0.1% Brilliant Blue R250, 10% acetic acid, 50% methanol, 40% H_2O Mix overnight and filter using 3 mm Whatmann paper Store at room temperature
Destaining solution	10% acetic acid, 50% methanol, 40% H_2O Store at room temperature
Coupling buffer	0.1 M $NaHCO_3$, 0.5 M NaCl, dissolve in H_2O, pH 8.3 Store at 4°C
Blocking buffer	0.5 M ethanolamine, 0.5 M NaCl, pH 8.3 Store at room temperature
Washing buffer	0.5 M NaCl, 0.1% Triton X-100 in 1X PBS Store at 4°C
3% BSA+ 0.02% Azide solution	3 g BSA, 0.02% sodium azide, dissolve in in 100 ml 1X TBST Store at 4°C
PK buffer	0.5 M Tris-HCl, 0.1 M $MgCl_2$, 1 mM EDTA, 20 mM DTT, 0.1% Brij 35, pH to 7.5 Store at room temperature
10X ATP	Make a solution of 4 mM ATP in 1x PK buffer Store at -80°C
Aprotinin	10 mg/ml in H_2O Store at -20°C
Pepstatin	10 mg/ml in absolute ethanol Store at -20°C
Leupeptin	10 mg/ml in H_2O Store at -20°C
Bacterial Work	
Ampicillin	100 mg/ml in 50% ethanol Use at final concentration of 100 µg/ml Store at -20°C
Kanamycin	50 mg/ml in H_2O Use at final concentration of 50 µg/ml Store at -20°C
Isopropyl β-d-1-thiogalactopyranoside (IPTG)	100 mg/ml in H_2O Use at final concentration of 1 mM Store at -20°C
DONSON bacterial pellet resuspension buffer	500 mM NaCl, 50 mM NaH_2PO_4, 2 mM $MgCl_2$, 10% glycerol, pH 9 Freshly prepared

Table 2.2 Inhibitors Added to *X. laevis* Egg Extract

Name	Stock Concentration	Conditions	Final Concentration
MLN4924 (A01139, Active Biochem)	5 mM in DMSO	Added to extract 15 minutes after sperm chromatin addition	10 µM
NMS873 (17674, Cayman Chemical Company)	10 mM in DMSO	Added to extract 15 minutes after sperm chromatin addition	50 µM
Caffeine (C8960, Sigma)	100 mM in water	Added to extract alongside sperm chromatin	5 mM
Aphidicolin (A0781, Sigma)	8 mM in DMSO	Added to extract alongside sperm chromatin	40 µM
p27 Kip1 (Produced in-house)	5 µM in water	Added to extract alongside sperm chromatin	100 nM
Cdc7i PHA-767491 (942425-68-5, Fisher Scientific)	10 mM in water	Added to extract alongside sperm chromatin	100 µM
Geminin (purified in house)	0.5 mM in water	Added to extract alongside sperm chromatin	0.3 µM

Table 2.3 Antibodies

Antibody (Species)	*Source*	*Concentration*
Primary antibodies ***All primary antibodies were prepared at the appropriate concentration in 3% BSA+0.02% Azide solution and incubated overnight at 4°C.***		
DONSON – rabbit	This PhD project	1:1000
DONSON – sheep	This PhD project	1:1000
α-TopBP1 (Sheep)	Kindly provided by Prof J Blow [182]	1:750
α-Cut5 (Rabbit)	In house [183]	1:1000
α-Mcm7 (Sheep)	In house [183]	1:2000
α-Mcm10 (Sheep)	In house [93]	1:1000
α-Mcm3 (Sheep)	In house [184]	1:1000
α-Mcm4 (Rabbit)	In house [185]	1:1000
α-Cdc45 (Sheep)	In house [186]	1:500
α-PCNA (Mouse)	P8825-Sigma	1:2000
α-GINS (Sheep)	In house [187]	1:500
α-Psf2 (Sheep)	In house [186]	1:500
α-His (Mouse)	H1029, Sigma	1:1000
α-Rif1 (Mouse)	A300-568A, Bethyl	1:1000
α-PP1a (Mouse)	sc-271762, Santa Cruz	1:1000
α-PP2a (Mouse)	sc-56950, Santa Cruz	1:1000
α-RecQL4 (Sheep)	Kindly provided by Prof J Blow [182]	1:750
α-Treslin (Rabbit)	Kindly provided by Prof J Blow [182]	1:1000
α-MTBP (Rabbit)	Kindly provided by Prof J Blow [182]	1:2000
α-Pol Ɛ (Mouse)	PCRP-POLE3-4E1, DSHB	1:500
α-p97 ATPase (Mouse)	65278, Progen Biotechnik	1:1000
α-γH2Ax (Rabbit)	4418-APC-020, Trevigen	1:1000
Secondary antibodies ***All secondary antibodies were prepared at the appropriate concentration with 5% milk in 1X TBST and incubated for 1hr at room temperature.***		

α-Mouse IgG Peroxidase conjugated antibody produced in goat	A5278, Sigma	1:5000
α-Sheep IgG Peroxidase conjugated antibody produced in donkey	A3415, Sigma	1:10,000
α-Rabbit IgG Peroxidase conjugated antibody produced in goat	A0545, Sigma	1:25,000
α-Rabbit IgG Peroxidase, Rabbit TrueBlot produced in mouse	18-8816-31, Rockland	1:10,000
α-Sheep IgG Peroxidase, Sheep TrueBlot produced in mouse	18-8815-31, Rockland	1:10,000
α-Mouse IgG Peroxidase, Mouse TrueBlot produced in rat	18-8817-30, Rockland	1:10,000

Table 2.4 Plasmids

Gene of interest	*Parental vector*	*Antibiotic resistance*
His/MYC_DONSON	pET28a	Kanamycin
His/MYC_DONSON-Δ Loop (P334-L410)	pET28a	Kanamycin
His/MYC_DONSON-Δ N-terminal (A2-V155)	pET28a	Kanamycin
His/MYC_DONSON-D374A, E377A, W381A (DEW)	pET28a	Kanamycin
His/MYC_DONSON-S437A	pET28a	Kanamycin
His/MYC_DONSON- D374A, E377A, W381A, S437A (DEWS)	pET28a	Kanamycin
His/MYC_DONSON-Δ24	pET28a	Kanamycin
His/MYC_DONSON-R476A, Y481A, V473A (3A)	pET28a	Kanamycin
His/MYC_DONSON-W234L, M463T (WM)	pET28a	Kanamycin
His/MYC_DONSON-M463T	pET28a	Kanamycin
His/MYC_DONSON-P230S	pET28a	Kanamycin
His/MYC_DONSON-F171S	pET28a	Kanamycin
His/MYC_DONSON-R217C	pET28a	Kanamycin
His/MYC_DONSON-S26A, T33A, T140A, S142A (4A)	pET28a	Kanamycin
His/MYC_DONSON-S22A, S26A, T33A, S54A, T140A, S142A, S314A, S454A (8A)	pET28a	Kanamycin
His/MYC_DONSON-S22A, S26A, T33A, S36A, S39A, S54A, T140A, S142A, S314A, S454A (10A)	pET28a	Kanamycin
His/FLAG_TopBP1	pET28a	Kanamycin
His/FLAG_TopBP1-1-575 (N-ter)	pET28a	Kanamycin
His/FLAG_TopBP1-1163-1513 (C-ter)	pET28a	Kanamycin
His/FLAG_Cdc45	pRSFDuet-1	Ampicillin, Kanamycin

Table 2.5 Primers

Mutation	*Orientation*	*Primer sequence*
DONSON ΔLoop (P334-L410)	Forward	GACCATAAGCCTGAATCC
	Reverse	CTGGGAGAAAGTGATTCC
DONSON ΔN-terminal (A2-V155)	Forward	TCCGTGACTTTTCCTGCT
	Reverse	CATTTATTTCCAGGTGTACC
DONSON DEW (D374A, E377A, W381A)	Forward	CTCTGCGCTAGAAGAGATGGGAGTTGAAGATAAAATTAAAAAG
	Reverse	AAGCTCGCATCTTCCGCGCTTTCATCCGATGCCTC
DONSON S437A	Forward	AAACTGTAAGGCGATTGTCGCTGC
	Reverse	ATGAGGAAGTTCAGTAAAG
DONSON DEWS (D374A, E377A, W381A, S437A)	Forward	CTCTGCGCTAGAAGCGATGGGAGTTGAAGATAAAATTAAAAAG
	Reverse	AAGCTCGCATCTTCCGCGCTTTCATCCGATGCCTC
DONSON Δ24	Forward	AGGAGCAGGAGCGAGGCT
	Reverse	CAGATCCTCTTCTGAGATGAGTTTTTGTTCG
DONSON 3A (R476A, Y481A, V473A)	Forward	GAACTCTGGTGCGAAGGATCAGTTTAGTCTGG
	Reverse	ACCGCAGTCTTCGCATTCACACTCCTTGCCTTAAG
DONSON WM (W234L, M463T)	W234L Fwd	GTCGCTTCCTctgGTACAGCTTTTTC
	W234L Rv	GGATGAATCCAGTGCACAAG
	M463T Fwd	AGGAGCAACAACGCACGCGCTTA
	M463T Rv	CGGAATGCGACTGGAGAC
DONSON M463T	Forward	AGGAGCAACAACGCACGCGCTTA
	Reverse	CGGAATGCGACTGGAGAC
DONSON P230S	Forward	CTGGATTCATTCGTCGCTTCCTT
	Reverse	TGCACAAGACTTTGCTGAAAG
DONSON F171S	Forward	CCGCTTGCTGTCCACATCGTCTC
	Reverse	GTCTTTAAACTCCAGTCAGCAGG
DONSON R217C	Forward	CACAGACCTATGTTGTGCCTTTC
	Reverse	GATAACTTTGGCTCCTGAATAC
DONSON 4A (S26A, T33A, T140A, S142A)	S26A & T33A Fwd	GCTTCCAAGGCCGGCCTTTCTCCGTTTTCC
	S26A & T33A Rv	CTCGCTCCTGGCCCTCCGTCTACTTAATCG
	T140A & S142A Fwd	AAGCCCCCGAGAAGCCAGCTCCA
	T140A & S142A Rv	CGGCCTTTACAAGTGGATCGGCATCAG
DONSON 8A (S22A, S26A, T33A, S54A, T140A, S142A, S314A, S454A)	S22A & S26A Fwd	GAGGGCCAGGAGCGAGGCTTCC
	S22A & S26A Rv	CGTCTAGCTAATCGGAGAATCTCCGAGG
	T33A Fwd	GGCTTCCAAGGCCGGCCTTTCTC
	T33A Rv	TCGCTCCTGCTCCTCCGTC
	S54A Fwd	GCGGCCTTTCGCCCCTGGGCCGA
	S54A Rv	AGCCCGGGGACTCTCTTAATCACATCCC
	T140A, S142A Fwd	AAGCCCCCGAGAAGCCAGCTCCA
	T140A, S142A Rv	CGGCCTTTACAAGTGGATCGGCATCAG
	S314A Fwd	TGCGGTCATGGCGCCCACAACCA
	S314A Rv	GTAATGACATCACTTCCTGCAAGCCCGG
	S454A Fwd	CACTCTGCTGGCTCCAGTCGCATTC

	S454A Rv	GGAGGAAGCCCGGCTTGG
DONSON 10A (S22A, S26A, T33A, S36A, S39A, S54A, T140A, S142A, S314A, S454A)	S22A & S26A Fwd	GAGGGCCAGGAGCGAGGCTTCC
	S22A & S26A Rv	CGTCTAGCTAATCGGAGAATCTCCGAGG
	T33A, S36A, S39A Fwd	TCCGTTTGCCCCGGGGGATGTGATTAAG
	T33A, S36A, S39A Rv	GCAAGGCCGGCCTTGGAAGCCTCGCTCCT
	S54A Fwd	GCGGCCTTTCGCCCCTGGGCCGA
	S54A Rv	AGCCCGGGGACTCTCTTAATCACATCCC
	T140A & S142A Fwd	AAGCCCCCGAGAAGCCAGCTCCA
	T140A & S142A Rv	CGGCCTTTACAAGTGGATCGGCATCAG
	S314A Fwd	TGCGGTCATGGCGCCCACAACCA
	S314A Rv	GTAATGACATCACTTCCTGCAAGCCCGG
	S454A Fwd	CACTCTGCTGGCTCCAGTCGCATTC
	S454A Rv	GGAGGAAGCCCGGCTTGG

Table 2.6 Recombinant Proteins

Recombinant Protein	*Source*	*Addition to X. laevis egg extract*
His/MYC_DONSON_BL21	This PhD project	Added after sperm DNA from 0.05 mg/ml stock to 1/50th volume of the extract (1 µg/ml) Or Added to LFB1/50 at a concentration of 0.5 µM for *in vitro* IP
His/MYC_DONSON-Δ Loop (P334-L410)	This PhD project	Added after sperm DNA from 0.05 mg/ml stock to 1/50th volume of the extract (1 µg/ml)
His/MYC_DONSON-Δ N-terminal (A2-V155)	This PhD project	Added after sperm DNA from 0.05 mg/ml stock to 1/50th volume of the extract (1 µg/ml)
His/MYC_DONSON-D374A, E377A, W381A (DEW)	This PhD project	Added after sperm DNA from 0.05 mg/ml stock to 1/50th volume of the extract (1 µg/ml)
His/MYC_DONSON-S437A	This PhD project	Added after sperm DNA from 0.05 mg/ml stock to 1/50th volume of the extract (1 µg/ml)
His/MYC_DONSON- D374A, E377A, W381A, S437A (DEWS)	This PhD project	Added after sperm DNA from 0.05 mg/ml stock to 1/50th volume of the extract (1 µg/ml)
His/MYC_DONSON-Δ24	This PhD project	Added after sperm DNA from 0.05 mg/ml stock to 1/50th volume of the extract (1 µg/ml)
His/MYC_DONSON-R476A, Y481A, V473A (3A)	This PhD project	Added after sperm DNA from 0.05 mg/ml stock to 1/50th volume of the extract (1 µg/ml)

His/MYC_DONSON-W234L, M463T (WM)	This PhD project	Added after sperm DNA from 0.05 mg/ml stock to 1/50th volume of the extract (1 µg/ml)
His/MYC_DONSON-M463T	This PhD project	Added after sperm DNA from 0.05 mg/ml stock to 1/50th volume of the extract (1 µg/ml)
His/MYC_DONSON-P230S	This PhD project	Added after sperm DNA from 0.05 mg/ml stock to 1/50th volume of the extract (1 µg/ml)
His/MYC_DONSON-F171S	This PhD project	Added after sperm DNA from 0.05 mg/ml stock to 1/50th volume of the extract (1 µg/ml)
His/MYC_DONSON-R217C	This PhD project	Added after sperm DNA from 0.05 mg/ml stock to 1/50th volume of the extract (1 µg/ml)
His/MYC_DONSON-S26A, T33A, T140A, S142A (4A)	This PhD project	Added after sperm DNA from 0.05 mg/ml stock to 1/50th volume of the extract (1 µg/ml)
His/MYC_DONSON-S22A, S26A, T33A, S54A, T140A, S142A, S314A, S454A (8A)	This PhD project	Added after sperm DNA from 0.05 mg/ml stock to 1/50th volume of the extract (1 µg/ml)
His/MYC_DONSON-S22A, S26A, T33A, S36A, S39A, S54A, T140A, S142A, S314A, S454A (10A)	This PhD project	Added after sperm DNA from 0.05 mg/ml stock to 1/50th volume of the extract (1 µg/ml)
His/FLAG_TopBP1-1-575 (N-ter)	This PhD project	Added to LFB1/50 at a concentration of 0.5 µM for *in vitro* IP
His/FLAG_TopBP1-1163-1513 (C-ter)	This PhD project	Added to LFB1/50 at a concentration of 0.5 µM for *in vitro* IP
FLAG_GINS	Purified in-house [187]	Added to LFB1/50 at a concentration of 0.5 µM for *in vitro* IP
His/FLAG_Cdc45	This PhD project	70 nM added to egg extract at the same time as sperm for large scale IP from chromatin

Table 2.7 Composition of PCR reaction mixture

Components	Volume for 25 µl reaction (µl)	Final concentration
RepliQa HiFi Tough Mix (95200-100, Quanta Bio)	12.5 µl	1x
10 µM Forward Primer	0.75 µl	0.3 µM
10 µM Reverse Primer	0.75 µl	0.3 µM
Template DNA	variable	10 ng
Nuclease-Free Water	Make up to 25 µl	

Table 2.8 Thermocycling Conditions

Steps	Temperature	Time	Number of cycles
Initial denaturation	98°C	90 seconds	1
Denaturation	98°C	10 seconds	25
Annealing	variable	10-30 seconds	
Extension	72°C	5 sec/Kb	
Final extension	72°C	2 minutes	1
Hold	4°C	∝	

2.2 *Xenopus laevis* Egg Extract Work

2.2.1 Egg Extract Preparation

The extract was prepared arrested in metaphase of meiosis II, in order to study different aspects of the cell cycle. The eggs used for extract preparation need to be of quality, to produce extract of high quality for DNA replication. 3 days before laying, female frogs were primed with 150 units of Folligon, to increase the yield of eggs. The day before extract preparation, these same frogs were injected with 500 units of Chorulon (Merck, Sharp and Dohme), and they were placed into individual laying tanks containing 1xMMR buffer (Table 2.1). Cell cycle arrest of the eggs was maintained by the high salt concentration of the buffer. The temperature in the room in which the frogs are kept is maintained between 19-23°C. The morning after the final injection, good quality eggs with a clear circumferential difference between the larger dark animal pole and smaller and lighter coloured vegetal pole were collected for extract preparation.

These eggs were rinsed in 1xMMR buffer, to remove debris. Apoptosed eggs were removed with a Pasteur pipette. Eggs were then de-jellied with cysteine solution (Table 2.1), which was replenished 2-3 times for effective de-jellying. Eggs were rewashed with 1xMMR, followed by UEB (Table 2.1), which was supplemented with 2 mM DTT (A3668.0050, VWR). De-jellied eggs were then packed into 14ml round bottom polypropylene tubes (187261, Greiner), which each contained 1 ml UEB, supplemented with 10 µg/ml protease inhibitors: aprotinin, leupeptin and pepstatin, and 50 µg/ml Cytochalasin D (Table 2.1) – the excess buffer from the settling eggs was removed. The tubes were then centrifuged at room temperature (RT) in a Beckman JS 13.1 for 1 minute at 3000 rpm, to pack the eggs. White, apoptotic, swollen eggs float to the top and were removed with a Pasteur pipette. The remaining eggs were centrifuged at 10,000 rpm for 10 minutes at RT to separate the eggs into lipids (top layer), brown cytoplasmic section (middle layer) and an insoluble egg yolk pellet (bottom layer).

The side of the tube was then punctured with a 20 G needle, and the cytoplasmic layer collected with a 1 ml syringe and transferred to a 15 ml Falcon tube. From this point, the egg extract was kept on ice. The extract was then supplemented with 15% volume LFB1/50 (Table 2.1), 10 µg/ml protease inhibitors and 10 µg/ml Cytochalasin D, and mixed well with a Pasteur pipette. The extract was then transferred to SW50

ultracentrifuge tubes (344058, Beckmann) and spun at 30,000 rpm for 20 minutes at 4°C. This left a yellow lipid plug at the top, beneath which was a white membranous material, beneath was the golden cytoplasmic fraction and a black, insoluble pellet was at the bottom of the tube. Using a spatula cleaned with 70% ethanol, the lipid plug was removed. The cytoplasmic layer – including the membranes floating above – was carefully collected using P1000 and P200 pipettes, being sure not to disturb the yolk pellets at the bottom, which contain the mitochondria and would promote apoptosis upon freeze-thawing. The cytoplasmic layer was transferred into a 50 ml falcon tube. The collected extract was then supplemented with 2% glycerol solution and thoroughly mixed with a Pasteur pipette. The extract was then frozen as 20 µl beads by dropping into liquid nitrogen with a P200 pipette, and stored at -80°C. Extracts remain stable for more than 10 years without compromising quality.

Nucleoplasmic extract (NPE) was prepared by Dr N Gilhooly, as per [188].

When using the egg extract, the number of frozen beads corresponding to the volume required for the experiment were transferred to a microtube. Following thawing at room temperature, the extract was used immediately for the experiment.

2.2.2 Sperm Preparation

In order to carry out DNA replication, *Xenopus laevis* sperm DNA acts as the substrate when added to egg extract. Sperm nuclei were retrieved from male frogs in house. 5-9 days prior to testes removal, male frogs were primed with 150 units of Chorulon. On the day of testes removal, male frogs were placed in MS222 solution (A5040, Sigma Aldrich) in opaque chambers, until they lost reflexes and their heads dropped below the water surface, and unconsciousness was ensured. The abdomen was opened and the main arteries surrounding the heart were cut. Testes, within the lower abdominal region, were then carefully removed without damaging them and immediately transferred to EB solution (Table 2.1) on ice. Testes are pale in colour and are typically rounded in shape and 0.75-1.5 cm in length, with male frogs typically possessing two that can each vary in size.

Once in the lab, testes were washed in a petri dish with 20 ml EB solution. Using forceps cleaned with 70% ethanol, blood vessels and extraneous tissues were removed, without bursting them. Cleaned testes

were then transferred to a petri dish with 1 ml EB solution and were finely chopped with a razor blade. The resulting chopped material was then filtered through a 25 µl nylon filter. The resulting material appears cloudy and should be stored on ice. The material was then spun down at 2000 g in a swinging bucket rotor for 5 minutes at 4°C. The pellet fractions were then resuspended in 0.5 ml SuNaSp (Table 2.1), supplemented with 25 µl Lysolecithin (5 mg/ml) for each original testes, and were incubated at room temperature for 5 minutes. To confirm effective demembranation of sperm, 1 µl of material was mixed with 1 µl Hoechst 33258 (20 µg/ml) and viewed under a UV microscope. Those that are effectively demembranated will be stained bright blue with Hoechst, those that have not been effectively demembranated will not be stained blue. Demembranated sperm were centrifuged at 2000 g in swinging bucket rotor for 5 minutes at 4°C and the pellet was resuspended in 0.5 ml SuNaSp supplemented with 3% BSA – for each testis – to quench the lysolecithin. The samples were then centrifuged again and the pellet resuspended in 0.5 ml EB per testis. Centrifugation of the sample was repeated once more and the pellet resuspended in 100 µl EB supplemented with 30% glycerol per testis. A haemocytometer was used to count the number of sperm and large somatic-type nuclei. The counting was done twice and an average taken, for accuracy. The stock was then diluted to a final concentration of 400 ng DNA/µl with EB supplemented with 30% glycerol. The sperm DNA was stored as 50 µl aliquots at -80°C.

2.2.3 *Xenopus laevis* Egg Extract Activation

X. laevis egg extracts were prepared arrested at metaphase of meiosis II, with the addition of EGTA – a Ca^{2+} chelator – to enable maintenance of the arrest. To promote DNA replication, the arrested extract must be activated so that it can enter the mitotic interphase. Extracts were supplemented with 0.3 mM $CaCl_2$, this mimics the calcium waves generated during fertilisation and so enables the extract to enter the first mitotic interphase. Extract was also supplemented with $1/40^{th}$ volume of energy regenerator (ER) (Table 2.1), which provides sufficient energy to drive the reaction. $1/40^{th}$ volume of Cycloheximide (CHX) (Table 2.1) – 250 µg/ml final concentration – can also be added, to further synchronise the cycling of the extract or to stop Cyclin synthesis. $CaCl_2$ was left for 15 minutes in the extract at 23°C, which enables the release of the extract from metaphase arrest and activates the licensing system, promoting replication of

the sperm DNA. Once extracts were activated, they were supplemented with ~10 ng/µl of sperm DNA, to initiate the DNA replication reaction.

2.2.4 Measuring Egg Extract Replication Efficiency

Radiolabelled dNTPs can be incorporated into replicating DNA and so be used to measure replication efficiency in *X. laevis* egg extract, we use α[32P]-dATP (NEG512H250UC, Perkin Elmer). Egg extracts were activated as per 2.2.3, sperm DNA was then added and α[32P]-dATP (50 nCi/µl). Samples were then aliquoted to 3-5 µl and incubated at 23°C for the desired time course. At each timepoint, replication was terminated through the addition of 160 µl Stop C solution (Table 2.1), which contains SDS and 0.2 mg/ml proteinase K (740506, Macherey Nagel). Terminated reactions were incubated at 37°C for >30 mins, after which samples were added to 4 ml of ice cold 10% TCA solution and incubated at 4°C, to allow DNA precipitation. The precipitated DNA was then separated from the unincorporated α[32P]-dATP through filtration. 40 µl was taken from each sample, prior to filtration, and spot on a 25 mm Whatmann paper filter (1001-025, GE HEALTHCARE LIFE SCIENCES), to measure the total ^{32}P. The remaining sample was poured into a vacuum filter pump (Millipore), through glass fibre filters (11512083, Fisherbrand), and washed with 4 ml of fresh 10% TCA – to ensure all DNA was removed from the tubes. The glass filters were then washed with 5 ml of ice-cold 5% TCA, then <1ml 70% ethanol. Each glass fibre filter was then removed and placed alongside corresponding to the paper filters under an infrared lamp to dry. Each dry filter was inserted into a scintillation tube, and 0.5 ml of scintillant Ultima Gold F (6013179, Perkin Elmer) was then applied. The amount of DNA and the level of radioactivity was measured using a scintillation counter (Tri-Carb 2900TR).

The percentage of total ^{32}P incorporated into the nascent DNA was normalised by dividing the ^{32}P incorporated into the DNA (glass filter) by the total ^{32}P obtained from the paper filter. This normalised value for each sample was multiplied by 0.654 in order to calculate the total amount of DNA synthesised in ng/µl of extract. The value of 0.654 was estimated on the basis that egg extract contains an endogenous dNTP pool of ~50 µM, and the average molecular weight of dNMPs is 327 Da, with equal

quantities of all four dNMPs incorporated into nascent DNA. Therefore, the weight of dNMP incorporated in ng/µl = % total ^{32}P incorporated/100 × 50 × 10^{-6} × 4 × 327 × 10^{3} [189].

This assay can be used to measure DNA replication efficiency in both normal and protein-depleted egg extracts, with the addition of a number of difference inhibitors or recombinant proteins.

2.2.5 Immunodepletion of Protein (DONSON and TopBP1) from *X. laevis* Egg Extract

DONSON or TopBP1 were immunodepleted from egg extract to identify the functional role of the protein. The protein was depleted using DONSON- or TopBP1-specific antibodies crosslinked to magnetic beads. The following protocol describes coupling 540 µg of antibodies to 900 µl of magnetic beads. 900 µl of Protein A Dynabeads (10008D, Life Technologies) – used with antibodies raised in rabbit – were transferred to a 1.5 ml Eppendorf tube, washed with 1.5 ml HEPES solution (Table 2.1), and the supernatant removed. DONSON immunodepletions were performed using α-DONSON antibody raised in rabbit or α-Cut5 (TopBP1) antibodies raised in rabbit (Table 2.3) and affinity purified or nonspecific rabbit IgG (I5006, Sigma).

540 µg of antibodies were added to the washed beads and topped up to a volume of 900 µl with HEPES solution. The mixture was incubated rotating at 4°C overnight. The following day, the supernatant was removed and washed with PBST. The beads were then incubated with 25.8 mg BS3 crosslinker (S5799, Sigma), in BS3 conjugation buffer (Table 2.1) for 30 minutes at room temperature. Quenching buffer (Table 2.1) was then added and incubated rotating at room temperature for 15 minutes. The beads were then washed thrice with HEPES solution, once with HEPES solution + 1.5 µl PMSF, again with HEPES solution, and then twice with LFB1/50. The antibody coupled beads were then incubated with activated egg extracts. Effective immunodepletion requires 3 rounds of 40 minute incubation of egg extract with antibody-coupled beads at 1:2 bead:extract ratio. After incubation, the beads were then removed from the extract using a magnetic rack, the immunodepleted extracts were then collected and dropped as 10 µl beads into liquid nitrogen and stored at -80°C until use.

2.2.6 Chromatin Isolation to Analyse Chromatin-Associated Proteins

This assay was performed in S-phase of both normal and protein-depleted egg extracts, to study the chromatin binding dynamics of various replication-related proteins. Different recombinant proteins (Table 2.6) or inhibitors (Table 2.2) were added to analyse their effects on replication. Histones and a 'no DNA' – in which no sperm DNA was added – sample act as loading and protein precipitation controls respectively, and were assessed for each treatment tested.

Metaphase II-arrested egg extracts were activated (2.2.3) and 10 ng/µl of sperm DNA was added to the extract. Extracts were then supplemented with the appropriate treatments with corresponding controls. 20 µl of extract was typically used for each chromatin sample, and each was aliquoted into a 1.5 ml Treff tube (9607811903, TreffLab) and incubated at 23°C until the end of the time course. At the appropriate timepoint, the 20 µl sample was resuspended in 0.5 ml ice-cold 1x ANIB100 (Table 2.1), and underlayed with 100 µl of ANIB, containing 20% sucrose. Extract dilution with ANIB100 was required to reduce density and viscosity of the extract, allowing sedimentation of the chromatin through the sucrose cushion. Samples were spun in a swinging bucket rotor at 2100 g for 5 minutes at 4°C. Following centrifugation, the buffer above the sucrose cushion was removed and the tube washed twice with 100 µl 1X ANIB, to avoid contamination of the chromatin sample with diluted extract. The sucrose cushion was then carefully removed, leaving ~50 µl in the tube. The samples were then spun in a fixed angle rotor for 2 minutes at >10,000 g at 18°C, this compacts the pellet. The remaining sucrose buffer was then carefully removed – without disturbing the pellet – and the pellet was resuspended in 30 µl of 2X NuPAGE LDS loading buffer (NP0008, Invitrogen). The samples were then boiled and run on a 4-12% NuPAGE acrylamide gel (WG1403BOX, Invitrogen) and analysed by western blotting.

Following the gel run, the histone loading controls were washed 3 x 5 minutes in water, then submerged in Simply Blue Safe Stain (LC6060; Invitrogen) for >1 hour on a rocker. The gel was then destained in water until the bands were visible.

2.2.7 Dynabeads Preparation

100 µl of Dynabeads Protein A (10008D, Life Technologies) or Protein G (10004D, Life Technologies) were washed once with 1X PBST (Table 2.1) and 20 µg of affinity purified antibodies or IgG from serum were added, and the mix rotated ≥8 hours at 4°C. Following incubation, beads were washed once with 1 ml 1X PBST. 475 µl conjugation buffer (Table 2.1) and 25 µl stock BS3 (Table 2.1) were added, and the mix rotated at room temperature for 30 minutes. The crosslinking was quenched by adding 100 µl 1 M Tris HCl pH7.5, and the mix rotated at room temperature for 15 minutes. Beads were washed thrice with 1 ml 1X PBST. Beads were stored in 210 µl 1X PBST with 0.02% sodium azide at 4°C.

2.2.8 Epoxy Bead Preparation

450 µl Dynabeads M-270 epoxy (14302D, Life Technologies) were washed twice with 1 ml 0.1 M sodium phosphate pH 7.4 mixing at room temperature for 10 minutes, then removed. 300 µl of 3 M ammonium sulphate was added with 300 µg of antibody (DONSON, GINS or sheep IgG from serum), and incubated rotating at 4°C for 2 days. Following incubation, beads were washed 4 times with 1X PBS (Table 2.1), once with 1X PBS/0.5% triton rotating at room temperature for 10 minutes, then thrice with 1X PBS for 5 minutes. Beads were then resuspended in 900 µl 1X PBS containing 0.02% sodium azide and prepared beads were stored at 4°C until use.

2.2.9 Small Scale Immunoprecipitation (IP) from Extract

Metaphase II arrested egg extract was activated as above (2.2.3). Activated extract was then diluted in 4 volumes of LFB1/50 (Table 2.1) and centrifuged at maximum speed for 15 minutes at 4°C. A sample of the supernatant was taken and resuspended in 2x NuPAGE LDS loading buffer to create the input sample for western blot analysis. The remaining input was incubated with IgG rabbit or sheep serum (I5006 or I5131, Sigma) or purified antibodies against the protein of interest (anti-DONSON, Table 2.3), in a ratio of 1:100 µg of purified antibody: µl extract, on ice for 1 hour. During incubation, protein A or G Dynabeads (10002D or10004D, Invitrogen), for antibodies raised in rabbit or sheep respectively, were prepared by taking 15 µl of beads per sample and washing once with 1 ml of PBS and twice with 1 ml of LFB1/50 (Table 2.1). For each wash, tubes containing beads were resuspended in buffer, quickly spun and placed

on a magnetic rack (DynaMag™-2 Magnet, 12321D, Thermo Fisher Scientific) – all buffer was removed once the beads settled on the wall. Following incubation with antibodies, samples were incubated with the appropriate beads, by rotating for 1 hour at 4°C.

Proteins bound to beads were isolated by placing samples on the magnetic rack and removing the depleted material. Some depleted sample was taken from each and resuspended in 2x NuPAGE LDS loading buffer for western blotting analysis. Beads were then washed with 1 ml of LFB1/50 thrice, before being completely removed using the magnetic rack. Proteins immunoprecipitated from the input were eluted from the beads by boiling them in 30 µl of 2x NuPAGE LDS loading buffer for 5 minutes, before placing the tubes on a magnetic rack to separate from the beads. IP samples were then run on an SDS-PAGE gel, with input and depleted samples for visualisation by immunoblotting with the antibody against the protein of interest.

2.2.10 Pull-down of Recombinant DONSON from Depleted Extract

Metaphase II arrested extract, IgG- or DONSON-depleted, was activated as described above (2.2.3), and recombinant DONSON was added to the appropriate sample of extract to a final concentration of 16 nM, and incubated at room temperature for 15 minutes. Each sample was then diluted into 4 volumes of LFB1/50 (Table 2.1), centrifuged and the input sample was taken as above (2.2.9). Pierce α-c-Myc beads (88842, Thermo Scientific) or epoxy beads, prepared for the protein of interest (2.2.8) were prepared by taking 30 µl for each sample and washing them twice with 1X TBS-T (Table 2.1). Beads were then washed with 100 µl of LFB1/50 per sample, and placed on a magnetic rack to remove all buffer from the beads. Input prepared previously was then added to the appropriate prepared beads and incubated for 1 hour at 4°C. Following incubation, the depleted sample was taken, beads were washed and the immunoprecipitated protein samples were prepared as above (2.2.9).

2.2.11 Small Scale IP from Chromatin (DONSON, TopBP1, Treslin, GINS)

275 µl of extract per IP sample was activated as above (2.2.3) and optionally supplemented with replication-perturbing treatments (NMS, MLN, aphidicolin, caffeine) (2.2.6). Samples were then incubated at 23°C for the indicated time. Chromatin was then isolated in ANIB100 (Table 2.1), and chromatin pellets

were re-suspended in the same volume of original extract of ANIB100 containing 20% sucrose. Protein complexes were released from chromatin by digestion with 2 U/µl of Benzonase nuclease (E1014-25KU, Sigma) and sonicated using a Diagenode sonicator at medium setting for 5 minutes – 30 seconds on, 30 seconds off. The insoluble fraction was then centrifuged at 16k g for 10 minutes at 4°C.

30 µl of epoxy Dynabeads (14302D, Life Technologies) or 100 µl of protein A Dynabeads (10008D, Life Technologies), prepared as above (2.2.7 and 2.2.8), were incubated with 220 µl of digested chromatin rotating for 1-2 hours at 4°C. Following incubation, beads were washed twice with ANIB100, rotating at 4°C for 5 minutes, once with ANIB100 containing an additional 0.1% Triton X-100 and twice more with ANIB100. Proteins immunoprecipitated from the input were eluted from the beads by boiling them in 30 µl of 2x NuPAGE LDS loading buffer for 5 minutes, before placing the tubes on a magnetic rack to separate from the beads, and the IP samples run on an SDS-PAGE gel.

2.2.12 IP from Nucleoplasm

550 µl of egg extract per IP was activated (2.2.3) and optionally supplemented with replication-perturbing treatments. CDK inhibitor ($p27^{Kip1}$), CDC7 inhibitor (PHA-767491), aphidicolin and caffeine were added immediately after sperm DNA. The samples were incubated at 23°C for 1 hour. Nucleoplasm was isolated in ANIB100 (without Triton-X 100), underlayed with the same volume of original extract of ANIB100 (without Triton-X 100), containing 20% sucrose. Protein complexes were released from nucleoplasm through digestion with 2 U/µl of Benzonase nuclease (E1014-25KU, Sigma) incubated for 20 minutes at room temperature, and sonicated for 10 minutes at medium setting - 30 seconds on, 30 seconds off. The insoluble fraction was then centrifuged for 10 minutes at 16,000 g at 4°C. 50 µl of Dynabead epoxy beads prepared to either the protein of interest or IgG from sheep serum (15131, Sigma-Aldrich) (2.2.8) were added to each appropriate sample and incubated at 4°C for 2 hours with rotation. Following incubation, beads were washed twice with ANIB100 (containing 0.1% Triton-X 100) for 5 minutes at 4°C with rotation, once with LFB1/50 and twice more with ANIB100. Each IP sample was prepared by boiling washed beads with 50 µl of 2x NuPAGE LDS loading buffer for 5 min and run on an SDS-PAGE gel for immunoblotting analysis.

2.2.13 Large Scale Chromatin IP of DONSON for Mass Spectrometry and CHROMASS

3.9 ml *Xenopus laevis* egg extract was activated as above (2.2.3) and supplemented with 10 ng/µl of demembranated sperm DNA. After 15 minutes of incubation at 23ºC, 50 µM p97i NMS873 or 40 µM aphidicolin + 5 mM caffeine, was added and the samples incubated at 23ºC for a further 45 minutes. At 60 minutes, chromatin was isolated in ANIB100 buffer, underlayed with the same volume of original extract of ANIB100, containing 20% sucrose. The sample was centrifuged twice at 2200 g for 5 minutes at 4ºC and the sample washed 3 times with 1 ml ANIB100. Protein complexes were released from the chromatin through digestion with Benzonase nuclease (2 U/µl), sonicated and centrifuged, as described above (2.2.12). 450 µl of protein G Dynabeads prepared to either the protein of interest or IgG from sheep serum (2.2.7) were added to each appropriate sample and incubated at 4°C for 2 hours with rotation. Following incubation, beads were washed and IP samples prepared as above (2.2.12).

2.2.13.1 Mass Spectrometry Sample Preparation

Each sample was run on a 5-20% gradient gel (Invitrogen) for 1 cm and cut into 10 bands. Samples were submitted pre-plated for 10 fraction analysis.

Samples were analysed by mass spectrometry by Dr Richard Jones from MS Bioworks LLC.

2.2.14 Cdc45-FLAG IP from Chromatin

Recombinant FLAG-tagged Cdc45 was added to the egg extract and a FLAG-IP, silver staining, negative staining and cryo-EM of the sample were all performed by Dr P Passaretti, as per [190].

2.3 *In vitro* Assays

2.3.1 *In vitro* Kinase Assay

The capacity of recombinant DONSON, and mutant versions, to be phosphorylated was tested with the aid of CDK2/Cyclin A (PV3267; Thermo Scientific) and CDK2/Cyclin E1 (PV6295; Thermo Scientific). To test phosphorylation, an assay of PK buffer (Table 2.1), 10X ATP (Table 2.1), CDK/Cyclin and recombinant DONSON was set up. Samples were incubated for 30 minutes at 30ºC, after which the reaction was stopped either by the addition of 4X SDS loading buffer, or by putting samples on ice. The phosphorylated

samples are run on a 6% SuperSep Phos-tag gel (192-17401; Fujifilm) for 1 hr 15 min at 160 V and analysed by immunoblotting, or were used in further experiments.

This was performed by myself and Dr D Poovathumkadavil.

2.3.2 *In vivo* Phosphatase Assay

NPE was activated and incubated with Lambda Protein Phosphatase (P0753S, NEBiolabs) and run on a 6% SuperSep Phos-tag gel by Dr D Poovathumkadavil.

2.4 DNA Work

2.4.1 Site Directed Mutagenesis

Oligonucleotides required for creating mutant DONSON were designed using the online tool: NEBaseChanger. Table 2.5 shows the list of primers used for making these mutants, which were used to set up a PCR reaction using the composition listed in Table 2.7. A high fidelity Q5 DNA polymerase (Table 2.7) was used for amplification of DNA from plasmids containing the gene of interest. The annealing temperatures were calculated based on an online tool: NEB TmCalculator. The thermocycling conditions used are described in Table 2.8.

2.4.2 PCR Product Visualisation

DNA samples from PCR reactions were visualised on 0.8% agarose gels under a Safe view Mini box (Cleaver Scientific Ltd). The required amount of agarose was weighed out and melted in 1x TBE buffer (A4718, Merck). After the agarose had cooled to ~55°C, 1:000 SYBR Safe DNA gel stain (S33102, Thermo Scientific) was added, to allow for DNA visualisation under UV lamp, and the mixture was poured into a casting tray. DNA samples were loaded by mixing in a 1:5 ratio with 6x purple gel loading dye (B7025, New England Biolabs). For size estimation, 1 kb plus DNA ladder (10787018, Invitrogen) was used as a marker. Electrophoresis was carried out at 100 V for 1 h and the gel visualised with the safe view box.

2.4.3 DNA Plasmid Processing

The PCR product was subjected to enzymatic treatment with KLD (kinase, ligase, DpnI) enzyme mix (M0554S, NEB), for 30 minutes at room temperature. This allows for efficient phosphorylation, plasmid

DNA circularisation and removal of template DNA. The product was then transformed into bacteria (2.5.2).

2.4.4 Plasmid Purification

Recombinant plasmids containing the DONSON gene were transformed into competent *E. coli* derivatives, NEB 10-β cells (C3019H, NEW ENGLAND Biolabs) for DNA expression (2.5.2). The following day several colonies were picked and grown in 10 ml LB broth with kanamycin antibiotic (50 µg/ml). These cultures were left at 37°C overnight. The plasmids were extracted using QIAprep Spin Miniprep kit (27104, Qiagen) as per the manufacturer's instructions. The concentrations of the plasmids were confirmed using a nanodrop spectrophotometer. The incorporation of the mutation was confirmed through sequencing services by Plasmidsaurus. The successful mutants were then transformed into BL21 *E. coli* cells, and the proteins purified, dialysed into LFB1/50 and visualised by SDS-PAGE.

2.5 Bacterial Work

2.5.1 Bacterial Strains and Growth Media

E. coli derivatives BL21 (DE3) competent *E. coli* (C2527, NEB) and Rosetta (DE3) pLysS competent cells (70956, Merck) were the host strains that we used for protein production. All cultures were grown in autoclaved Luria Bertani (LB) broth/agar (L3022, Sigma/L2897, Sigma) or Auto induction media (AIM) (AIMLB0210, Formedium) supplemented with the appropriate amount of antibiotics (Table 2.1). The bacterial cultures were grown at 37°C in an incubator shaker. Cultures grown on LB agar plates were grown at 37°C overnight and then stored at 4°C.

2.5.2 Bacterial Transformation

20 µl of NEB 10-β cells (C3019H, NEW ENGLAND Biolabs) for plasmid purification, or 10 µl of BL21 or Rosetta cells (De3) for protein expression, were taken and 0.2 µl (~100 ng) of plasmid DNA was added, and cells were left on ice for 30 minutes, followed by 30 seconds heat shock at 42°C. The cells were then left on ice for 5 minutes, followed by addition of 250 µl of SOC medium (B9020S, NEB), after which cells were incubated at 37°C for 1 hour with constant shaking. 100 µl of bacterial culture was then plated on LB agar plates containing appropriate antibiotics and left overnight at 37°C in an incubator.

2.5.3 Expression of Recombinant *X. laevis* Proteins (DONSON, TopBP1 and Cdc45)

pet28a (kanamycin resistance) containing *X. laevis* His/MYC-DONSON was transformed into *E. coli* derivatives, BL21 (De3) cells, for protein expression (2.5.2). Colonies were then grown in 2x 10 ml LB broth at 37ºC overnight. 2 l of bacterial culture expressing each protein were then grown from this at 37ºC until OD value reached 0.4-0.6, and 1 mM IPTG (I6758, Sigma) was added for protein induction, and cultures were incubated at 37ºC for 2 hours. Cultures were then spun down at 6000 g for 15 minutes at 4ºC in a Beckman Avanti centrifuge. DONSON pellets were resuspended in resuspension buffer.

pet28a (kanamycin resistance) containing His/FLAG-TopBP1 (C-term/N-term) and pRSFDuet-1 (kanamycin resistance) containing *X. laevis* His-/FLAG-Cdc45 were transformed into *E. coli* derivatives, BL21 (De3) cells (TopBP1) and Rosetta (De3) cells (Cdc45) for protein expression by Dr A Skagia and Dr P Passaretti.

2.5.3.1 Batch Protein Purification of *X. laevis* DONSON

Resuspended pellets were thawed and 1 µl/ml Aprotinin, 1 µl/ml Pepstatin, 1 µl/ml Leupeptin, 0.1 mM PMSF, 1 mg/ml lysozyme and 25 U/ml Benzonase (E1014-25KU, Sigma) were added, and the samples rotated at 4ºC for 20-30 minutes, followed by sonication for 5 x 30 seconds. The homogenate was centrifuged at 16,000 g for 30 minutes at 4ºC. The resulting supernatant was mixed with Super Nickel NTA Agarose Affinity Resin (SUPER-NINTA100; Generon) for >3 h at 4°C. The mixture was then centrifuged at 200 g for 2 minutes and the supernatant was removed. The beads were then washed 3 times with resuspension buffer, containing 20 mM imidazole, 1 µl/ml Aprotinin, 1 µl/ml Pepstatin, 1 µl/ml Leupeptin, 0.1 mM PMSF. The protein was eluted from the beads using elution buffer (resuspension buffer containing 250 mM imidazole and 1 µl/ml Aprotinin, 1 µl/ml Pepstatin, 1 µl/ml Leupeptin, 0.1 mM PMSF). The fractions with the highest concentration of protein were dialysed into LFB1/50. The concentration of the proteins was determined by ocular density and they were stored at -80°C until use.

AKTA protein purification of TopBP1 (N- and C-term) and Cdc45 were performed by Dr P Passaretti and Dr A Skagia.

2.5.4 Dialysis

Elutions obtained from recombinant protein purifications were dialysed into LFB1/50 buffer – a more egg extract-friendly buffer – using Pur-A-Lyzer Maxi 25,000 dialysis columns (PURX25005, Merck). The columns were first incubated with distilled water inside for 5 minutes, to activate the cellulose membrane, enabling the exchange of buffer and free passage of molecules smaller than the size of the membrane. The protein sample was then incubated inside the columns at 4°C overnight, with constant stirring. Following dialysis, proteins were frozen into balls by dropping into liquid nitrogen and stored at -80°C until use.

2.5.4 Mass Photometry

Mass photometry of the recombinant DONSON proteins was performed by Dr P Passaretti, as per [190].

2.5.5 *In vitro* Pull-Down of Recombinant Proteins

A 50 µl mix per pull-down sample was prepared by combining 0.5 µM of each protein of interest (Table 2.6) and topping up to 50 µl with LFB1/50 (Table 2.1) and incubated at room temperature for 1 hour. Pre-phosphorylated proteins may be added (2.3.1). 5 µl of each sample was taken and 5 µl of 2x NuPAGE LDS loading buffer was added to prepare the input sample. Following incubation, 10 µl of the appropriate bead type was added (α-Myc (88842, Thermo Scientific), α-FLAG (A6001, Stratech Scientific Ltd), epoxy Dynabeads (2.2.8) or protein A/G Dynabeads (2.2.7)). The sample was incubated at room temperature for 1 hour. 5 µl of each sample was taken and 5 µl of 2x NuPAGE LDS loading buffer was added to prepare the depleted sample. Beads were washed 4 times with LFB1/50 buffer, adding 0.1% Triton-X to the second wash – if α-Myc beads were used, the first 2 washes were with 1X TBST (Table 2.1) and the second 2 with LFB1/50. Following washing, all buffer was removed and 30 µl 2x NuPAGE LDS loading buffer was added, samples were boiled at 95°C for 5 minutes and the IP sample was separated from the beads and run on an SDS-PAGE gel for immunoblotting analysis.

Performed by myself and Dr P Passaretti.

2.5.6 Antibody Purification

To purify antibodies against specific target proteins, the proteins were first purified (2.5.3). The proteins samples were sent to Eurogentec to raise antibodies in the selected host species, this was sent by the company in the form of the sera obtained. An antigen column, specific to the target protein, was then made using a HiTrap NHS-activated HP affinity column (17071701, Cytiva) – used for any primary amine-containing ligand. The column was activated with 1 mM ice cold HCl. The protein was dialysed (section 2.5.3) into coupling buffer (Table 2.1) and 1 ml of 1 mg/ml loaded into the column and incubated at room temperature for 30 minutes. The protein bound to the column due to a stable amide linkage forming from a covalent bond between the resin and the amine group of the ligand, which prepared the column for affinity purification. Unreacted NHS groups were blocked using blocking buffer (Table 2.1). The column was washed with 10 mM Tris-HCl, pH 8 then 0.1 M glycine, pH 2, twice. The column was then stored with 1X PBS at 4°C.

The antisera sent by Eurogentec was defrosted and diluted 1:4 with 1X PBS and filtered through 0.45 µm, then 0.2 µm syringe filters (SLHN033NB; Millipore, 15206869; Fisher Brand, respectively), to remove impurities. The sera was circulated through the prepared antigen column at 4°C overnight. The column was then washed with 50 ml washing buffer (Table 2.1), to remove excess sera. The antibodies were eluted to 1 ml with 0.1M glycine pH 2.6 into 1.5 ml Eppendorf tubes already containing 100ul 2M Tris-HCl; pH 8.5. The concentration of the purified antibodies was determined by Bradford assay and the elutions were stored at -20°C.

Antibodies were tested by small scale immunoprecipitation from extract, as above (2.2.9).

2.6 General Procedures

2.6.1 Protein Visualisation by SDS PAGE (Sodium Dodecyl Sulphate Poly Acrylamide Gel Electrophoresis)

All protein samples were analysed on a precast 4-12% NuPAGE Bis-Tris gel (WG1403BOX, Invitrogen). Protein samples were mixed with 2X NuPAGE LDS sample buffer (NP0008, Invitrogen) and were denatured by boiling at 95°C for 5 min, before loading onto the gel. 3 µl of PageRuler Plus Prestained Protein Ladder (26620, Invitrogen) was run alongside the samples, to determine protein sizes. The gel apparatus was

filled with 1X NuPAGE MOPS SDS running buffer (NP000102, Invitrogen) and electropheresed at 160 V for 1 hour 15 min.

The proteins run on the gels were visualised by staining with Coomassie stain (Table 2.1) for 1 hour on a rocker (SSM4; Stuart) at room temperature, to detect total protein content. Following incubation, the stain was discarded and the gel was destained with a destaining solution (Table 2.1), until the bands were visible.

2.6.2 Bradford Assay

To determine the concentration of proteins, a BSA standard assay of known concentration was serially diluted and compared to the protein of interest. A 10 mg/ml stock solution of BSA was prepared and diluted to 0.08, 0.06, 0.04, 0.03, 0.02, 0.01 mg/ml. 1 ml of Pierce Coomassie Plus (Bradford) Assay reagent (23238; Thermo Fisher Scientific) was added in a cuvette. These samples were analysed at 595 nm using a spectrometer (Biowave). A standard curve was created with calculated slope and intercept of the samples, based on known concentration and absorbance of each dilution. Protein samples were also run through the spectrometer at 595 nm. The resulting absorbance was translated into concentration of protein with the following formula: $(Absorbance\ at\ 595\ nm - graph\ intercept) \div graph\ slope = protein\ concentration\ [\frac{mg}{ml}]$.

To visually compare and establish protein concentrations. The stock solution of BSA was prepared at dilutions of 5, 2.5, 1.25 and 0.625 mg/ml, and purified proteins were also diluted with LFB1/50 buffer and 2X NuPAGE loading buffer. All samples were run on an acrylamide gel and visualised with Coomassie staining.

2.6.3 Immunoblotting

Protein samples were run on SDS-PAGE acrylamide gels and transferred onto Immobilon-P PVDF membranes (LC6060; Merck) via electroblotting at 110 V for 1 h 15 mins in transfer buffer (Table 2.1). The membranes with the proteins were then blocked in 5% non-fat dry milk in 1X TBST (Table 2.1), to prevent non-specific binding. The membranes were then rinsed in 1X TBST and incubated with the appropriate

primary antibody, prepared in 3% BSA + 0.02% sodium azide solution in 1X TBST, overnight at 4°C. The membranes were then washed in 1X TBST for 5 minutes, 3 times, to minimise the background and remove any unbound antibody. The membranes were then incubated for >1 hour with the peroxidase-labelled secondary antibodies (Table 2.3) specific to the species within which the primary antibody was raised. Following incubation, membranes were washed again with 1X TBST as before. The membrane was then incubated with chemi-luminescent peroxidase substrate: Western Bright ECL spray (K-12049-D50; Advansta). The membrane was then exposed to X-ray film (MOL7016, Scientific Laboratory Supplies).

2.6.4 Gel Filtration

2.6.4.1 Gel Filtration of Extract

50 µl of extract was activated as above (2.2.3), and resuspended in 4 times the volume with LFB1/50, and 100 µl of the supernatant was loaded into the gel filtration system (Superose 6 (29091598, Cytiva) or Superdex 200(28990946, Cytiva)). 50 µl fractions were eluted and run on an SDS-PAGE gel.
This was performed Dr P Passaretti.

2.6.4.2 Gel Filtration of Recombinant Proteins in Buffer

2 µM of recombinant protein was resuspended in 70 µl LFB1/50 and 50 µl was run through the gel filtration system as above (2.2.4.16.1). This was performed by Dr P Passaretti.

2.7 Computational Analysis

2.7.1 AlphaFold

Xenopus laevis protein sequences of interest were input into the AlphaFold3 server for a prediction of their interactions.

UCSF ChimeraX was used for visualising all models produced, as well as for making figures.

2.7.2 ImageJ Analysis

The density of pixels for Western blot bands were quantified using ImageJ software to produce a numeric value in arbitrary units, which was normalised to the appropriate loading control. IgG control samples

were always analysed with treatment samples and the fold difference between them calculated, along with the mean value and the standard error of the mean (SEM).

2.7.3 Statistical Analysis

All statistical analyses were calculated using the software Prism (GraphPad). All measurements were taken from individual samples, and no repeated measurements were taken.

3 Results

3.1 DONSON is a Novel Replication Initiation Factor Interacting with TopBP1

3.1.1 Introduction

To effectively carry out DNA replication, the core CMG helicase must be supported by a plethora of proteins at every stage – to regulate replication timing, maintain fork stability and to ensure accurate nascent DNA synthesis. DONSON has previously been seen to interact with and support the replisome during elongation, interacting with known members of the fork protection complex (FPC) namely Timeless-Tipin, to ensure efficient progression of the replisome and maintain genome integrity, as part of the ATR stress response [152]. In human cells, DONSON has also been seen to interact with core components of the helicase, where it has been observed in euchromatin more strongly than in heterochromatin [156]. This suggests a fluid, yet integral, association depending on conditions and context.

In the lab, we have largely been interested in the roles of different proteins involved in replication termination, and interestingly, we have also previously observed DONSON interacting with components of the CMG, during termination, Table 3.1 [35]. We were therefore interested in determining whether DONSON may play a role in the regulation of replisome disassembly, particularly due to its previously reported role in maintaining the replisomes at stalled forks.

Table 3.1: DONSON interacts with terminated replisomes. Egg extract was incubated with sperm DNA and supplemented with an inactive mutant of p97 and MLN4924 (a Cullin E3 ligase inhibitor) to block replisome disassembly in S-phase, chromatin was isolated and MCM3 was immunoprecipitated, and samples analysed by mass spectrometry. Total spectral count and percentage coverage is shown for a selection of replisome components. As seen in [35].

	Control (IgG)	IP (Mcm3)
Mcm3 (90 kDa)	17 (20%)	1111 (84%)
Mcm2 (100 kDa)	17 (17%)	975 (73%)
Mcm4 (97 kDa)	7 (7%)	947 (80%)
Mcm5 (82 kDa)	30 (18%)	927 (84%)
Mcm7 (82 kDa)	24 (10%)	905 (77%)
Mcm6 (93 kDa)	39 (18%)	888 (83%)
Cdc45 (66 kDa)	2 (3%)	277 (54%)
Psf1 (23 kDa)	0	62 (85%)
Psf2 (21 kDa)	0	33 (85%)
Psf3 (24 kDa)	0	55 (92%)
Sld5 (26 kDa)	0	44 (64%)
Ctf4 (125 kDa)	6 (17%)	554 (54%)
Spt16 (118 kDa)	3 (3%)	454 (55%)
SSRP (79 kDa)	0	229 (51%)
Timeless (149 kDa)	0	171 (32%)
Tipin (40 kDa)	0	54 (26%)
Top2a (179 kDa)	6 (3%)	484 (51%)
Claspin (146 kDa)	0	72 (21%)
PolA1 (165 kDa)	0	109 (33%)
PolA2 (67 kDa)	0	7 (11%)
PolE1 (261 kDa)	2 (1%)	543 (40%)
PolE2 (60kDa)	0	103 (55%)
DONSON (64 kDa)	0	68 (32%)

Therefore, the aim of this chapter of the thesis was to assess whether DONSON plays further roles in DNA replication than solely as a fork protection factor. I aimed to determine whether DONSON interacts with the replisome throughout DNA synthesis and if it supports replication in *Xenopus* egg extract, particularly during the termination stage of the replication reaction.

3.1.2 Results

3.1.2.1 Characterisation of DONSON Chromatin Binding During Replication in *Xenopus* Egg Extract

Given that DONSON has been observed associated with the replisome during elongation and termination, I wanted to visualise the kinetics of the protein on chromatin throughout replication. With this aim, antibodies specific to *Xenopus* DONSON were needed.

Sera against *Xenopus* DONSON had previously been raised in the lab and so I could purify the required antibodies from them using an antigen column. For this, I expressed full-length His- and MYC-tagged DONSON in *E. coli* and purified the protein using a Nickel column. Samples taken during purification were run in an SDS-PAGE gel and visualised through both Coomassie staining and by immunoblotting with sheep α-DONSON antibodies previously purified in the lab (Figure 3.1.1). Following successful protein purification, elution 2 was dialysed into coupling buffer and the concentration measured (data not shown).

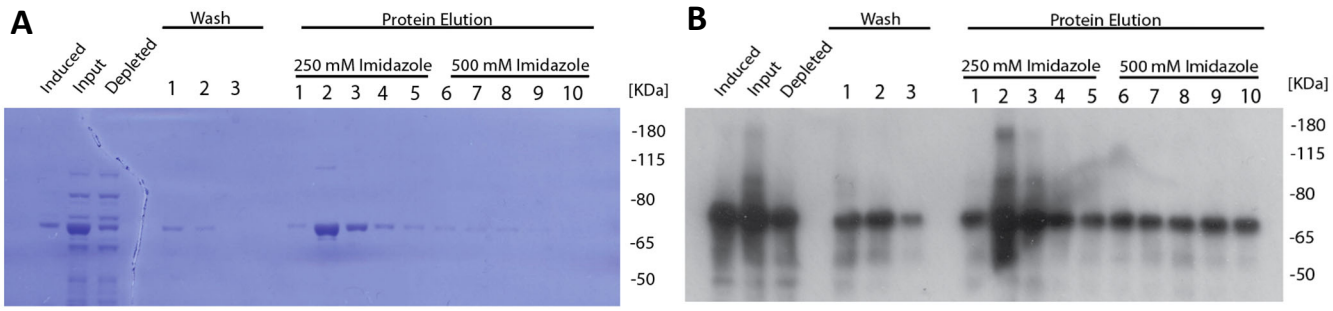


Figure 3.1.1: Purification of recombinant *Xenopus laevis* His- and MYC-tagged DONSON. Expression of His- and MYC-xDONSON was induced in *E. coli* (Induced) and purified as described in methods section using Nickel affinity purification. Supernatant from bacterial lysate containing soluble proteins (Input), unbound proteins (Depleted), samples from bead washes (Wash 1-3) and elutions using 2 concentrations of Imidazole (Protein elutions 1-10) are shown. Samples from the purification were run on two SDS-PAGE gels and visualised by Coomassie staining (A) and through immunoblotting with sheep α-full-length DONSON antibodies (B).

Next, the purified *Xenopus* DONSON (xDONSON) was used to create an antigen column, by incubating it with a HiTrap NHS-activated column, for purification of the antibodies from the whole sera. Sera against xDONSON was raised in both rabbit and sheep (Figure 3.1.2 A) and run through the DONSON antigen-column to elute the purified antibodies specific to DONSON. The most concentrated antibody fractions were kept and their purity tested by using them to immunoblot SDS-PAGE-resolved *Xenopus* egg extract (Figure 3.1.2 B). The ability to recognise natively folded DONSON by the sheep and rabbit antibodies was subsequently tested via immunoprecipitation (Figure 3.1.2 C). We observed that purified antibodies from both rabbit and sheep were able to IP DONSON from the extract with similar efficiency. However, we also observed that the DONSON sheep antibody was much cleaner for western blots and was therefore used for the remainder of the project for this technique.

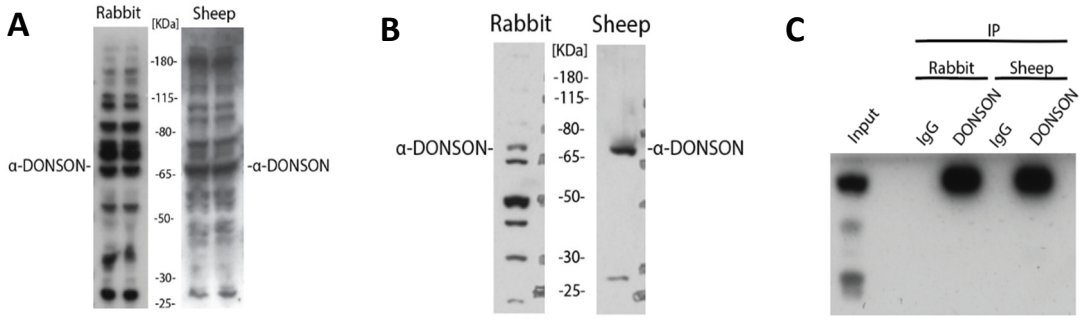


Figure 3.1.2: Purification of *Xenopus laevis* α-DONSON antibodies from whole sera. A and B) Testing rabbit and sheep α-DONSON sera and purified antibodies by immunoblotting. Sera against *Xenopus laevis* DONSON were raised in two rabbits (19 and 20) and one sheep. Egg extract samples were run on SDS-PAGE gels and visualised through immunoblotting with the specific sera (A) and purified antibodies (B). C) Testing affinity purified α-full-length DONSON antibodies by immunoprecipitation. Immunoprecipitation (IP) from whole egg extract was performed to test the ability to recognise native DONSON by the indicated purified antibodies, samples were blotted for using the α-full-length DONSON sheep antibody (B). Extract without antibodies (Input) and rabbit and sheep IgG antibodies were used as a negative control. α-full-length DONSON rabbit antibodies purified from both sera 19 and 20 were used interchangeably from this point.

With the antibodies purified, I could use them to detect DONSON binding on chromatin by immunoblotting. To assess this, *Xenopus laevis* egg extract was incubated with sperm DNA, and chromatin was isolated during the resulting replication reaction. The chromatin-bound proteins were subsequently analysed by immunoblotting with selected antibodies. As seen in Figure 3.1.3, our purified α-DONSON antibodies were detecting the protein at the peak of the replication reaction, with similar kinetics to those of known helicase components. As seen, Mcm7 binds very early during the reaction, different to the binding of all other factors blotted for. This represents origin licensing, where Mcm2-7 double-hexamers are loaded prior to the other factors of the CMG [55]. Psf2, one subunit of GINS, has peak binding at 60

minutes in this reaction – representative of the greatest number of active helicases bound to chromatin and the highest incidence of DNA replication [11]. Likewise, PCNA peak binding occurs at 60 minutes, for the same reason. DONSON peak binding also occurring at 60 minutes suggests that DONSON binds and is removed from chromatin at the same time as known replisome components. This indicates that DONSON may act as part of the replication machinery in egg extract.

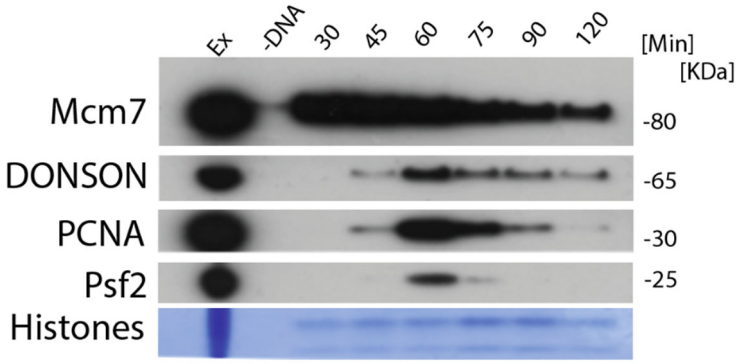


Figure 3.1.3: DONSON has the same chromatin binding kinetics during DNA replication reaction as other replisome components. Egg extract was incubated with sperm DNA to promote DNA replication. Chromatin was isolated at indicated timepoints throughout the replication reaction and chromatin-bound proteins were analysed by immunoblotting using indicated antibodies. Purified sheep α-full-length DONSON antibodies were used. An extract sample was used as a positive control for immunoblotting. A sample without DNA was also isolated and used as a chromatin specificity control. Histones were stained with colloidal Coomassie as a loading control. This is a representative experiment of more than ten repeats.

Having established that DONSON exhibits a similar chromatin binding pattern as known replisome components, I wanted to assess whether this was also the case under conditions where replication was perturbed. I had particular interest in both the post-termination replisome and the replisome during replication. With this in mind, DNA replication was set up in extract as previously, optionally supplemented with NMS973, to inhibit the activity of the p97 segregase, or MLN4924, to inhibit the activity of $Cul2^{Lrr1}$, both of which inhibit unloading of the terminated replisome [35], [191]. Alternatively, extract was optionally supplemented with aphidicolin, to inhibit DNA polymerase activity, and caffeine, to inhibit ATM/ATR kinase activity. This inhibits activation of the S phase checkpoint and so prevents inhibition of new origin firing, allowing us to sufficiently visualise perturbed replisomes on the chromatin during early stages of replication. As seen in Figure 3.1.4 A, upon inhibition of replisome disassembly from chromatin, all factors blotted for, including DONSON, are retained on chromatin for longer and accumulate at every timepoint – seen through higher levels of each protein compared to replication under normal conditions. Whilst we would expect a similar pattern upon addition of CULi (Figure 3.1.4 B),

the p97i is more efficient than the CULi in these experiments, explaining the differences in unloading dynamics [187]. Upon CULi addition, DONSON follows the same chromatin binding pattern as the replisome components blotted for, suggesting that DONSON may be disassembled from the chromatin with the replisome components. In contrast, as seen in Figure 3.1.4 C, upon DNA polymerase inhibition, binding of replisome components is reduced. Mcm2-7 subunit binding is increased, suggesting an accumulation of these molecules on chromatin due to replication being unable to proceed. The reduced chromatin binding of the remaining components, including DONSON, is representative of decreased formation of the replisome in the absence of DNA polymerase activity and checkpoint activation. When ATM/ATR activity is then inhibited, we see accumulation of all components blotted for, including DONSON. This represents uninhibited establishment of replisomes that are unable to progress without polymerase activity. Collectively, this indicates that DONSON chromatin binding exhibits the same pattern as those of known replisome components, further suggesting that DONSON may act as part of the replication machinery.

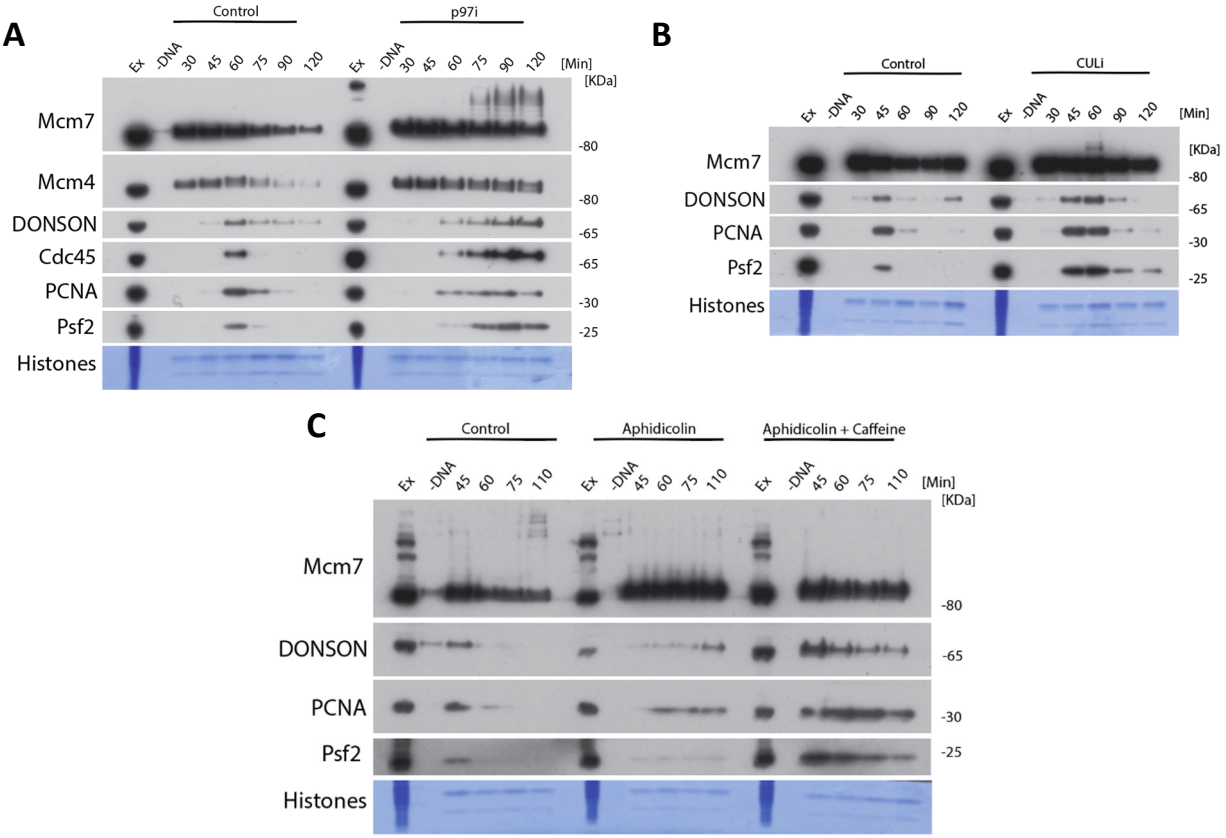


Figure 3.1.4: DONSON has the same kinetics on chromatin as other replisome components when replication is perturbed. Egg extract was incubated with sperm DNA to promote DNA replication and supplemented with a p97i (A), a Cullin inhibitor (B) or an inhibitor of DNA polymerases (aphidicolin) +/- an inhibitor of ATM/ATR (caffeine) (C). Chromatin was isolated as various timepoints and chromatin-bound proteins analysed by immunoblotting, positive and negative controls were performed as in previous experiment (Figure 3.1.3).

Having observed DONSON binding to and dissociating from chromatin with the same kinetics as replisome components, I wanted to determine whether DONSON actually directly binds any of these components during replication. To assess this, DNA replication was set up in egg extract as before and optionally supplemented with MLN4924. Chromatin was isolated at peak replication under normal unchallenged conditions (Figure 3.1.5 A), and later in replication when termination was perturbed (Figure 3.1.5 B). DONSON was immunoprecipitated, using the sheep α-DONSON antibodies purified above. As seen in Figure 3.1.5 A-B, as well as Mcm7, DONSON interacts with both Cdc45 and GINS – suggesting an interaction with the active helicase specifically. For this reason, I decided to immunoprecipitate GINS (Figure 3.1.5 C) to confirm whether DONSON would reciprocally co-immunoprecipitate. The egg extract was activated and supplemented with aphidicolin and caffeine, and chromatin was isolated. GINS was immunoprecipitated using sheep α-GINS antibodies and coimmunoprecipitation of other factors, including DONSON, was tested by immunoblotting with specific antibodies. As seen in Figure 3.1.5 C, many replisome components coimmunoprecipitate with GINS, as expected, including DONSON. This suggests that DONSON specifically interacts with the replisome on chromatin during replication in *Xenopus laevis* egg extract. This confirms previous observations in human cells [35], [152].

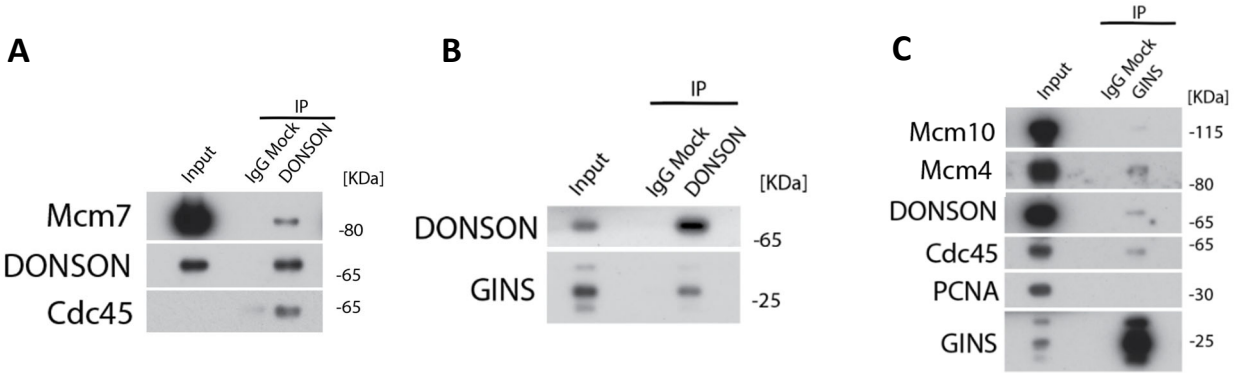


Figure 3.1.5: DONSON interacts with replisome components on chromatin. Egg extract was incubated with sperm DNA to promote DNA replication under normal unchallenged conditions (A) or supplemented with CULi (B), chromatin was isolated, proteins released from DNA through DNAse treatment and immunoprecipitation (IP) was performed with sheep α-DONSON antibodies. C) Activated egg extract was supplemented with aphidicolin and caffeine, chromatin was isolated, and IP was performed with sheep α-GINS antibodies. Co-immunoprecipitation of proteins was assessed by immunoblotting with indicated antibodies. Chromatin sample without antibodies (Input) and sheep IgG antibodies were used as a negative controls.

3.1.2.2 DONSON is Required to Support Replication

To determine the role that DONSON has during replication, we used the purified α-xDONSON antibodies to immunodeplete the endogenous protein from the extract. Extract was incubated with beads coupled

to either non-specific commercial IgG (control) or purified α-DONSON antibodies (protein A beads for rabbit antibodies, or G for sheep). Depletion efficiency was determined by assessing levels of the protein in whole extract by immunoblotting (Figure 3.1.6 A and C). With both antibodies, DONSON was successfully depleted to <5% of the endogenous levels. Following depletion, we tested the efficiency of DNA synthesis in the extracts by measuring α-[P32]dATP incorporation into nascent DNA (Figure 3.1.6 B and D). These results indicate that the absence of DONSON greatly affects the ability of the extract to synthesise DNA. In DONSON-depleted extract, the levels of replication were less than 25% of the IgG-depleted, when depleted with either rabbit or sheep antibody.

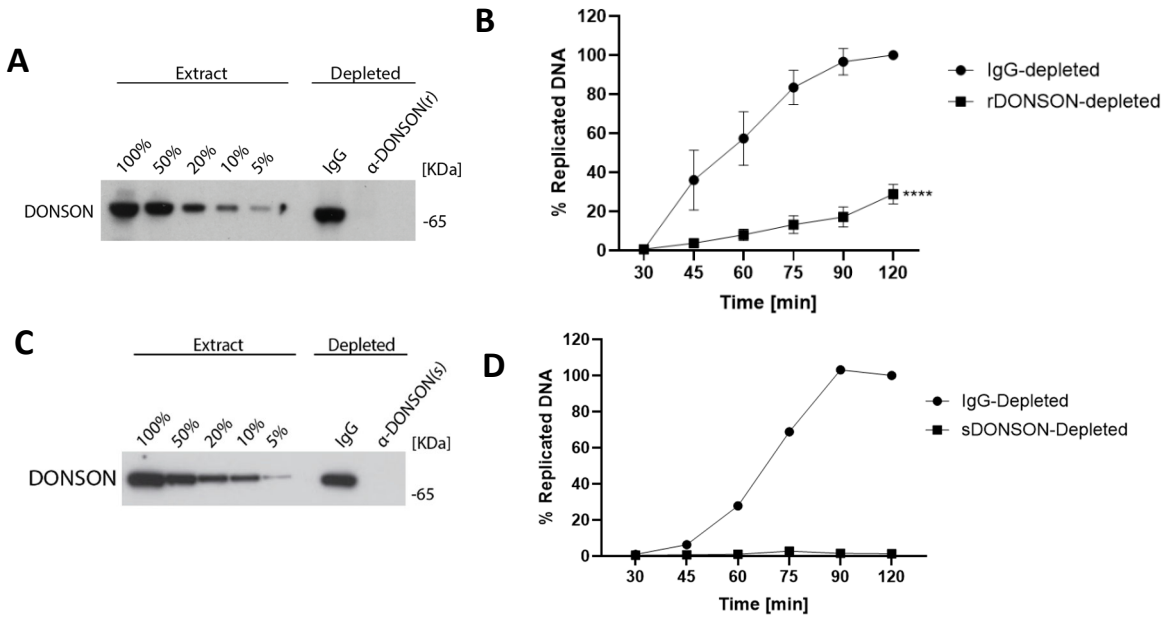


Figure 3.1.6: The immunodepletion of endogenous DONSON from egg extract impairs nascent DNA synthesis. *Xenopus* egg extract was incubated with DONSON rabbit (A and B) or sheep (C and D) α-DONSON antibodies coupled to beads to immunodeplete endogenous DONSON (sheep experiment performed by Dr A. Skagia). A and C) The levels of DONSON in extract were detected by western blotting to analyse depletion efficiency. Purified sheep anti-DONSON antibodies were used for immunoblotting. B and D) IgG- and DONSON-depleted extracts were incubated with sperm DNA and α-[P32]dATP. The DNA replication reaction was stopped at indicated timepoints and synthesised DNA was measured by α-[P32]dATP incorporation into nascent DNA. B) Two-way ANOVA comparing to IgG-depleted, ****$p < 0.0001$. (A and B are representative of more than ten repeats, C and D are representative of two repeats).

3.1.2.3 DONSON is Required for Formation of the Active Helicase

Having established a role for DONSON in supporting nascent DNA synthesis, I wanted to assess how it functions on the molecular level to achieve this. In particular, I was interested in whether there is an impact on the active helicase in the absence of DONSON, having observed DONSON interacting with

Cdc45 and GINS on chromatin. To determine this, rabbit IgG- and DONSON-depleted extracts were incubated with sperm DNA and chromatin was isolated at various timepoints throughout replication. As seen in Figure 3.1.7, Mcm4 is still loaded onto chromatin in the absence of DONSON, indicating that origin licensing is unaffected in its absence. However, it does appear to accumulate on chromatin, suggesting that replication does not progress after licensing in DONSON's absence. In contrast – whilst still being present in the extract – both Cdc45 and Psf2 levels on chromatin are greatly reduced in the absence of DONSON. Collectively, this suggests that DONSON is required to support the formation of the active helicase, specifically for Cdc45 and GINS chromatin loading.

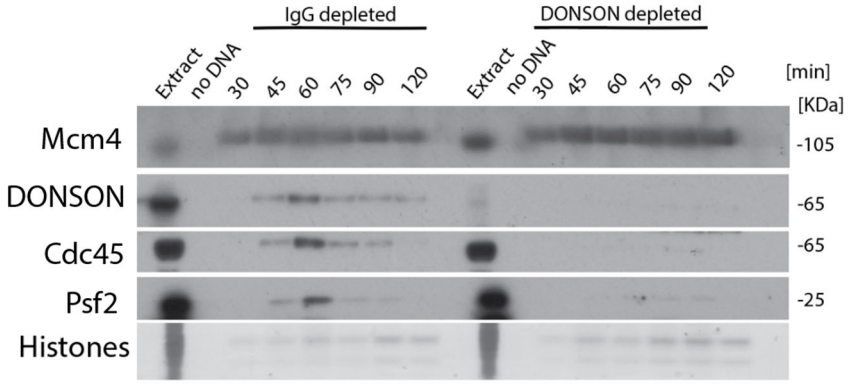


Figure 3.1.7: The immunodepletion of DONSON from egg extract impairs active helicase formation on chromatin. *Xenopus* egg extract immunodepleted of DONSON (using rabbit α-DONSON antibodies) was incubated with sperm DNA, chromatin was isolated and chromatin-bound proteins analysed by immunoblotting, positive and negative controls were performed as in previous experiments. This is a representative experiment of more than 10 repeats.

As this indicated a novel role for DONSON, I wanted to confirm that it was the absence of DONSON specifically that was causing this phenotype. To test this, I incubated IgG- and DONSON-depleted extracts as before, optionally supplemented both with purified recombinant DONSON protein and assessed the capacity of the extract to synthesise DNA. The concentration of DONSON I added to the extract was determined by testing a gradient of concentrations of recombinant DONSON (rDONSON). A concentration that visually matched the amount in the IgG-depleted extract (data not shown) was chosen (16 nM). As seen in Figure 3.1.8 A, the inability of DONSON-depleted extract to support nascent DNA synthesis is efficiently recovered when recombinant DONSON is introduced.

I similarly tested the capacity of the extract to support active helicase formation, both in the presence and absence of the recombinant protein, by assessing the loading of the factors on chromatin throughout replication, as before. As seen in Figure 3.1.8 B, introduction of rDONSON into the DONSON-depleted extract could successfully overcome the accumulation of Mcm7 on chromatin that occurs in its absence, thus suggesting that replication is no longer stalled after origin licensing. Likewise, upon rDONSON application, binding of both Cdc45 and GINS on chromatin is efficiently reestablished. Figure 3.1.8 C shows that Cdc45 and Psf2 chromatin binding rescue in the presence of rDONSON, is significant. This indicates that it is indeed DONSON that is causing the loss of active helicase formation in the depleted extract. Interestingly, binding of Mcm10 onto chromatin also appears to be reduced in the absence of DONSON. As Mcm10 has a role in DNA melting to promote elongation, this suggests that DONSON has a role between origin licensing and DNA replication elongation.

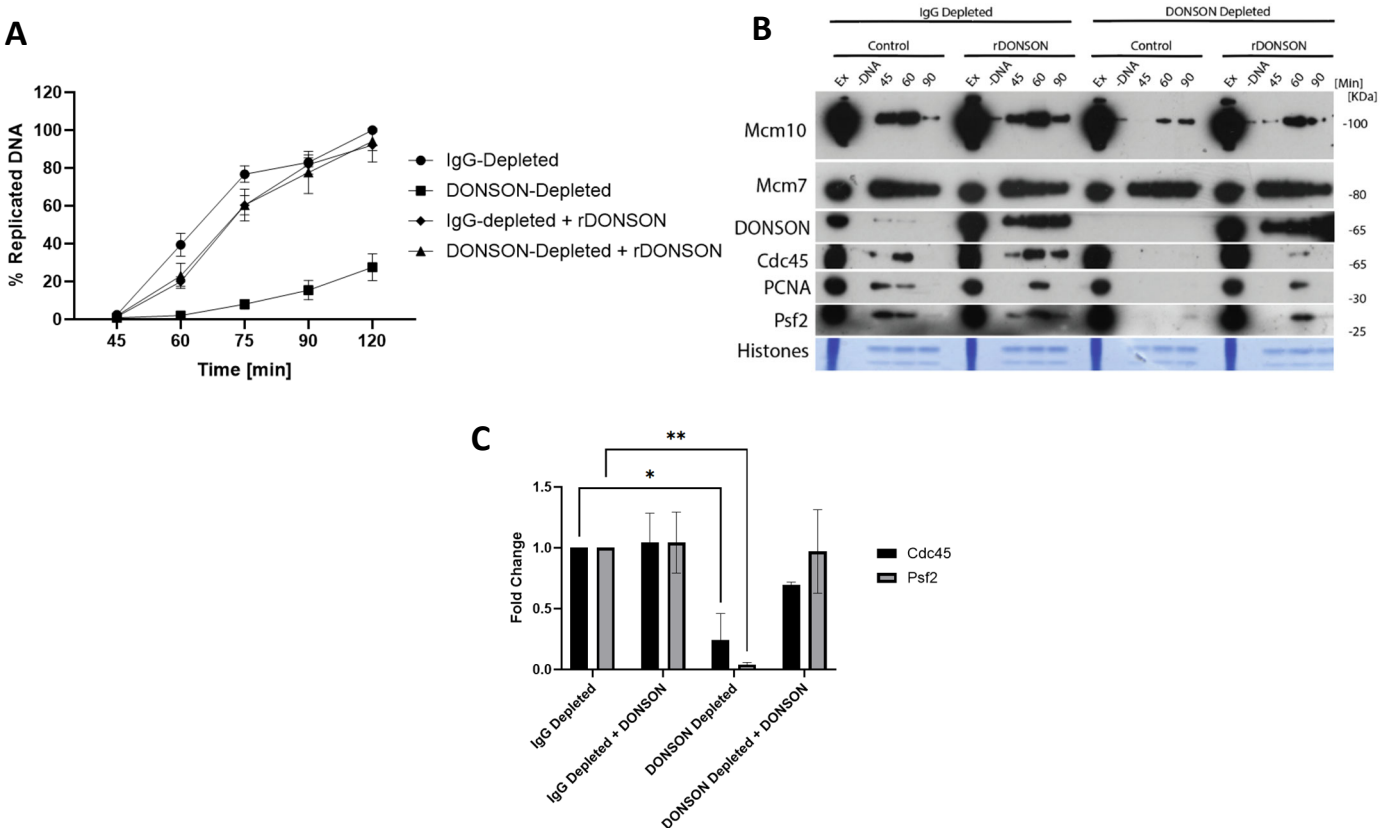


Figure 3.1.8: Purified recombinant DONSON rescues nascent DNA synthesis and active helicase formation on chromatin in DONSON-depleted extract. DNA replication was set up in IgG- and DONSON-depleted extract with the optional addition of recombinant xDONSON: A) and incubated with α-[P32]dATP. The DNA replication reaction was stopped at indicated timepoints and DNA synthesis was measured as in previous figures. Representative of more than ten repeats mean with SEM plotted. B) Chromatin was isolated at indicated timepoints and chromatin-bound proteins analysed by immunoblotting. Positive and negative controls were performed as in previous experiments. C) Quantification of fold change of chromatin-binding of specified proteins at peak replication in IgG- and DONSON-depleted extracts in n = 3 experiments (different extracts used). Fold-change levels normalised to the levels in IgG-depleted extract. Mean with SEM plotted. Two-way ANOVA comparing to IgG-depleted, *p = 0.015, **p = 0.002.

Collectively, this suggests that either DONSON has a role in initiation by supporting formation of the active helicase, or this phenotype is due to immediate activation of the S phase checkpoint following fork collapse in the absence of DONSON. As DONSON has previously been implicated in fork protection, I wanted to assess whether in its absence, forks were collapsing and further origin firing was being prevented. To determine this, I set up replication in depleted extracts as previously, and optionally supplemented with caffeine, to prevent activation of the S phase checkpoint. As seen in Figure 3.1.9 A, as expected, blocking ATM/ATR in non-depleted extracts resulted in a limited increase in the amount of nascent replicated DNA – representative of uncontrolled origin firing. However, no increase is observed in DONSON-depleted extract on the addition of caffeine. Likewise, as seen in Figure 3.1.9 B, caffeine does not rescue Cdc45 or GINS chromatin binding. This suggests that the DNA replication and helicase formation phenotypes observed in the absence of DONSON are due to a role for DONSON during initiation.

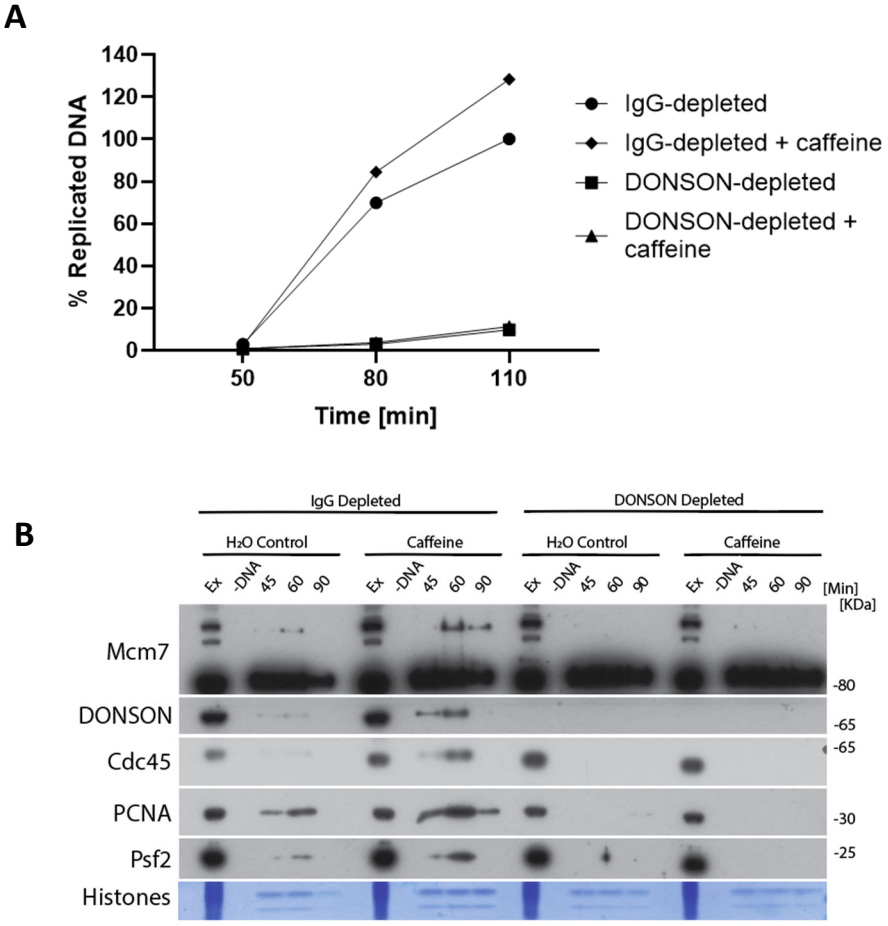


Figure 3.1.9: Blocking the stress response cannot rescue nascent DNA synthesis and active helicase formation on chromatin. DNA replication was set up in IgG- and DONSON-depleted extract with the optional addition of caffeine (an ATM/ATR inhibitor): A) and incubated with α-[P32]dATP. The DNA replication reaction was stopped at indicated timepoints and DNA synthesis was measured as in previous figures. N = 2. B) Chromatin was isolated at indicated timepoints and chromatin-bound proteins analysed by immunoblotting. Positive and negative controls were performed as in previous experiments. This is representative of two repeats.

3.1.2.4 DONSON is Important for Late-Stage Initiation

As our results suggest a role for DONSON in replication initiation, I wanted to determine the stage of initiation at which DONSON is functioning. To assess this, I incubated normal extract with sperm DNA and optionally supplemented with inhibitors of several stages of replication initiation and isolated chromatin, to determine whether DONSON's loading onto chromatin was perturbed by these inhibitors. As seen in Figure 3.1.10, under normal conditions successful chromatin-binding of replisome components and known initiation factors is observed and DONSON is also loaded. However, on the addition of geminin, an inhibitor of Mcm2-7 hexamer loading, most factors can no longer bind, including DONSON, further supporting the idea that DONSON acts during initiation after Mcm2-7 loading. Upon addition of PHA-767491, a DDK inhibitor (DDKi), fewer factors are unable to bind to chromatin and Mcm4 is visibly less phosphorylated. Notably, Cdc45 and GINS cannot bind, as these are loaded by initiation factors dependent on DDK phosphorylation [63]. DONSON chromatin-binding under this condition is partially reduced compared to normal S phase, suggesting some dependence on DDK activity for DONSON to bind chromatin. Interestingly, upon addition of $p27^{KIP1}$ inhibitor, which prevents CDK activity (CDKi), other factors have reduced capacity to load onto chromatin. These include RecQ4, TopBP1, PP1 and PP2 – factors known to play a role during initiation [68]. Cdc45 and GINS are also unable to bind – CDK activity is necessary for activation of factors required to load them [192]. DONSON's binding to chromatin is also inhibited in the absence of CDK activity. This suggests that DONSON loading onto chromatin requires CDK activity, and therefore DONSON appears to function late during initiation.

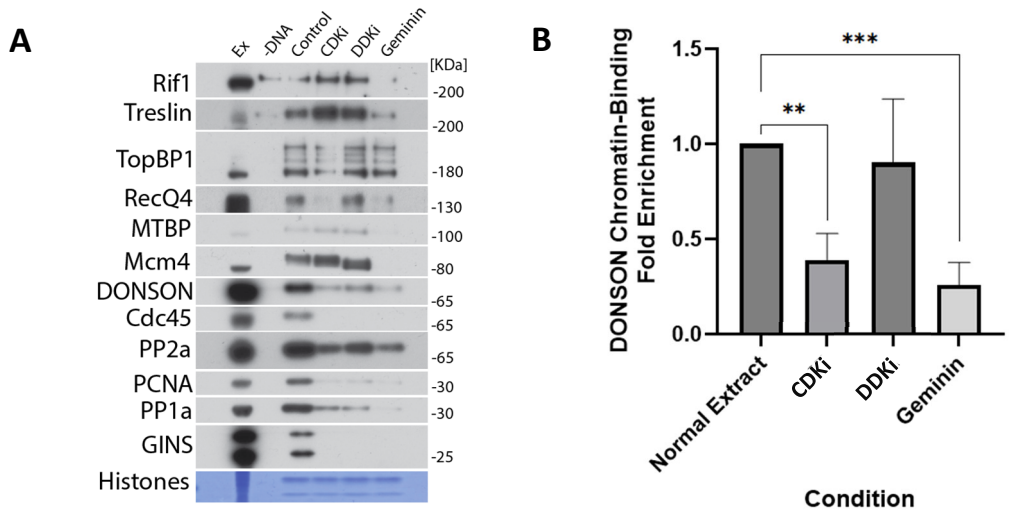


Figure 3.1.10: DONSON chromatin-loading is reduced on CDK and DDK inhibition. A) DNA replication was set up in extract and optionally supplemented with p27 (CDK inhibitor), PHA-767491 (DDK inhibitor) or geminin (origin licensing inhibitor). Chromatin was isolated at early replication and chromatin-bound proteins analysed by immunoblotting. Positive and negative controls were performed as in previous experiments. N = 4. B) Quantification of fold change of DONSON chromatin-binding under different conditions in n = 4 experiments (different extracts used). Fold-change levels normalised to the levels in normal extract sample. Mean with SEM plotted. Unpaired, two-tailed t-tests, ** p = 0.0046, *** p = 0.0008.

As I had observed DONSON binding chromatin in a CDK- and DDK-dependent manner, I wanted to investigate potential interactions between DONSON and known initiation factors on chromatin. Specifically, I wanted to assess if DONSON's depletion from the extract had any effect on the loading of these known initiation factors. To determine this, both IgG- and DONSON-depleted extracts were incubated with sperm DNA and chromatin was isolated throughout replication, as before. One sample, from peak replication in both depleted extracts, was analysed by mass spectrometry to determine the efficiency with which these factors were loaded onto chromatin. As seen in Figure 3.1.11 A, in the absence of DONSON, replisome formation on chromatin is inhibited, indicated by reduced peptide number for these components in the DONSON-depleted extract compared to the IgG-depleted. As seen before, peptide numbers for Mcm2-7 hexamer components are increased. However, we were unable to detect Mcm10 by mass spectrometry in either sample. Interestingly, the number of peptides of various initiation factors are also increased. This is also apparent in Figure 3.1.11 B, where isolated chromatin samples were tested by immunoblotting, showing increased binding of replication initiation factors on chromatin in DONSON depleted extract. This indicates that initiation is likely stalled following loading of

both hexamers and many initiation factors in the absence of DONSON. This further suggests that DONSON acts late during initiation, after loading of these other initiation factors.

A

	protein (Mw)	IgG depleted	DONSON depleted
Mcm2-7	Mcm2 (100 kDa)	109	176
	Mcm3 (90 kDa)	81	140
	Mcm4 (97 kDa)	73	98
	Mcm5 (82 kDa)	94	128
	Mcm6 (93 kDa)	94	107
	Mcm7 (82 kDa)	70	102
replisome	Polε (261 kDa)	22	0
	Polα (165 kDa)	15	0
	Polδ (125 kDa)	18	0
	Cdc45 (66 kDa)	8	0
	Psf3 (24 kDa)	2	0
	Psf4 (26 kDa)	2	0
	SPT16 (118 kDa)	42	3
	AND-1 (125 kDa)	9	0
	Timeless (149 kDa)	4	0
	PCNA (29 kDa)	45	9
initiation factors	RecQ4 (169 kDa)	4	4
	Treslin (221 kDa)	12	26
	MTBP (96 kDa)	13	24
	Rif1 (257 kDa)	76	98
	TopBP1 (169 kDa)	7	17

B

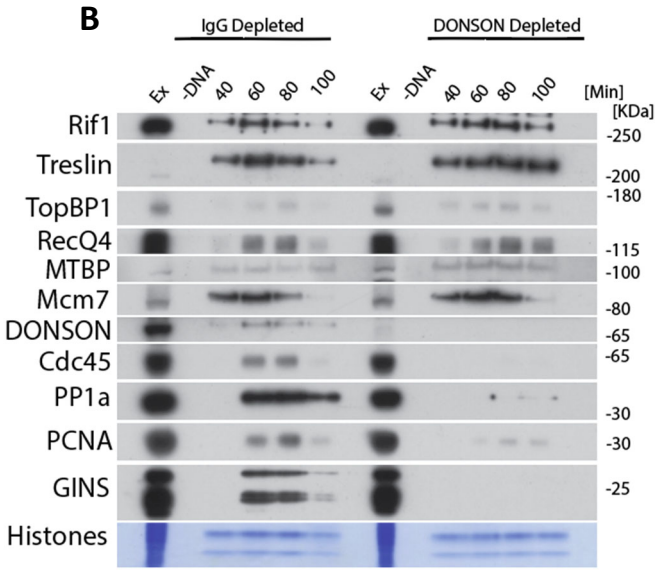
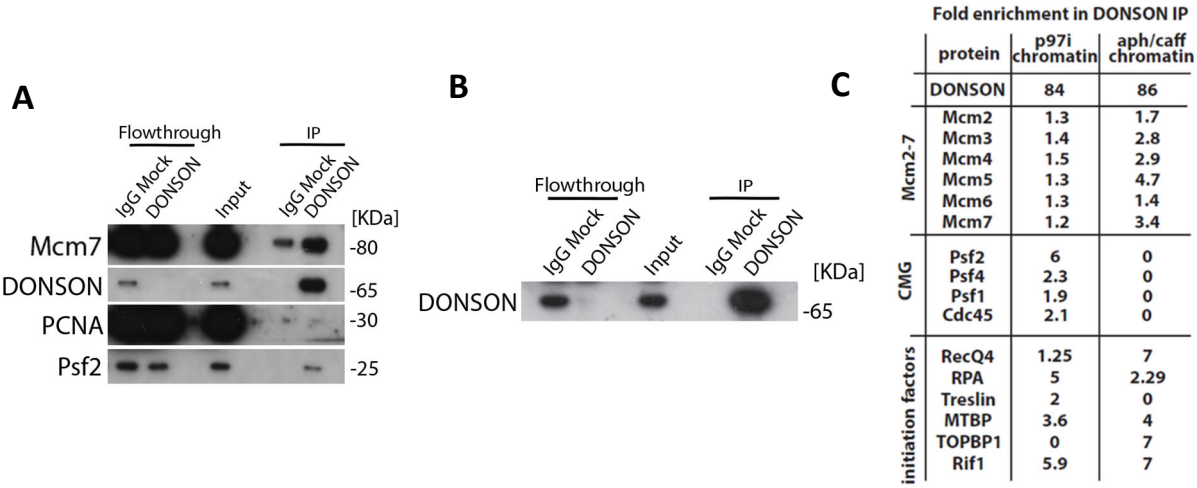


Figure 3.1.11: DONSON acts late during DNA replication initiation. DNA replication was set up in IgG- and DONSON-depleted extract as before and: A) Isolated at peak replication, run on an SDS-PAGE gel and samples analysed by mass spectrometry. Peptide numbers of proteins found on chromatin in each chromatin sample are shown. B) Chromatin was isolated at each timepoint indicated and chromatin-bound proteins analysed by immunoblotting. Positive and negative controls were performed as in previous experiments.

With the suggestion that DONSON carries out its function after other known initiation factors, I looked to ascertain whether it directly interacts with any of them on chromatin. With this in mind, I set up a DNA replication reaction as before, and optionally supplemented with either p97i, to assess terminated replisomes, or aphidicolin and caffeine, to assess replisomes accumulated during early S phase elongation. Chromatin was isolated, DONSON immunoprecipitated (Figures 3.1.12 A and B) and samples sent for mass spectrometry analysis, to identify which, if any, of the initiation factors DONSON interacts with under these conditions. As seen in Figures 3.1.12 A and B, DONSON has been efficiently removed from the flowthrough samples in both conditions, suggesting that all the DONSON present in the input has been immunoprecipitated effectively. The fold enrichment of peptide number in the DONSON IP compared to the IgG mock IP is shown. As seen in Figure 3.1.12 C, whilst DONSON appears to specifically co-immunoprecipitate most indicated factors under both conditions, with aphidicolin and caffeine no interactions are observed with Cdc45 and GINS. The 0 value represents no coimmunoprecipitation in either the IgG- or DONSON-IP under these conditions. This may be due to the small size of both Cdc45

and GINS making them difficult to detect via mass spec. Under both conditions, it is clear that DONSON is interacting with initiation factors, indicating that DONSON is binding to the replisome when these factors are present. Of all the initiation factors, the most striking difference in interaction at termination and initiation stages is with TopBP1, which appears to only interact with DONSON at early-stage replication. This could suggest an interaction that may be required for DONSON loading onto chromatin to carry out its initiation role.



	protein	p97i chromatin	aph/caff chromatin
	DONSON	84	86
Mcm2-7	Mcm2	1.3	1.7
	Mcm3	1.4	2.8
	Mcm4	1.5	2.9
	Mcm5	1.3	4.7
	Mcm6	1.3	1.4
	Mcm7	1.2	3.4
CMG	Psf2	6	0
	Psf4	2.3	0
	Psf1	1.9	0
	Cdc45	2.1	0
initiation factors	RecQ4	1.25	7
	RPA	5	2.29
	Treslin	2	0
	MTBP	3.6	4
	TOPBP1	0	7
	Rif1	5.9	7

Figure 3.1.12: DONSON interacts with replication initiation factors. Egg extract was incubated with sperm DNA to promote DNA replication as before and supplemented with A) p97i or B) aphidicolin and caffeine (DNA polymerase and ATM/ATR inhibitors), chromatin was isolated and immunoprecipitation (IP) was performed with sheep IgG and α-DONSON antibodies. Co-immunoprecipitation of proteins was assessed by immunoblotting with indicated antibodies (A). Chromatin input sample without antibodies (Input), flowthrough from which the IP beads were removed (flowthrough) and nonspecific sheep IgG antibodies were used as a negative control. C) Samples from each IP were run on an SDS-PAGE gel and analysed by mass spectrometry. Peptide count fold enrichment of each protein indicated are shown, normalised to peptide count in IgG mock sample.

3.1.2.5 DONSON Interacts with TopBP1

With the possibility of an initiation-specific interaction between TopBP1 and DONSON on chromatin, we wanted to follow this up, first through confirmation by western blotting. With this in mind, we set up replication in egg extract, supplementing with aphidicolin and caffeine. We isolated chromatin and immunoprecipitated DONSON or TopBP1 – assessing interactions by immunoblotting with the specific antibodies indicated. As seen in Figure 3.1.13 A and B, TopBP1 and DONSON reciprocally immunoprecipitate on chromatin under these conditions, along with other known factors of the replisome, indicating an interaction in the context of the replisome. Having previously seen a dependence on DONSON for the loading of Cdc45 onto chromatin, we also wanted to assess the interaction between DONSON and Treslin, one initiation factor known to recruit Cdc45 to the chromatin. To evaluate this, we

immunoprecipitated Treslin as we had done DONSON and TopBP1 in Figures 3.1.13 A and B. However, Figure 3.1.13 C indicates that DONSON may co-immunoprecipitate less efficiently with Treslin than with TopBP1, although this requires further investigation.

Therefore, I wanted to further assess the interaction between DONSON and TopBP1 in particular. Specifically, I wanted to confirm whether a direct interaction can occur outside of the context of the replisome formed on chromatin. To address this, I activated DNA replication by incubating extract with sperm DNA and optionally supplemented with $p27^{KIP1}$ or PHA-767491 to perturb initiation as before. I isolated the nuclei, containing chromatin and nucleoplasm, and immunoprecipitated DONSON from these samples, immunoblotting for other factors to determine their interaction. As seen in Figure 3.1.13 D and E, of all the potential interaction partners, only TopBP1 interacts with DONSON in this context, and this interaction is significantly reduced when CDK activity is prevented. This suggests that DONSON and TopBP1 interact in the nuclei prior to DONSON's loading onto chromatin, and may indicate that DONSON is recruited to chromatin during initiation in combination with TopBP1.

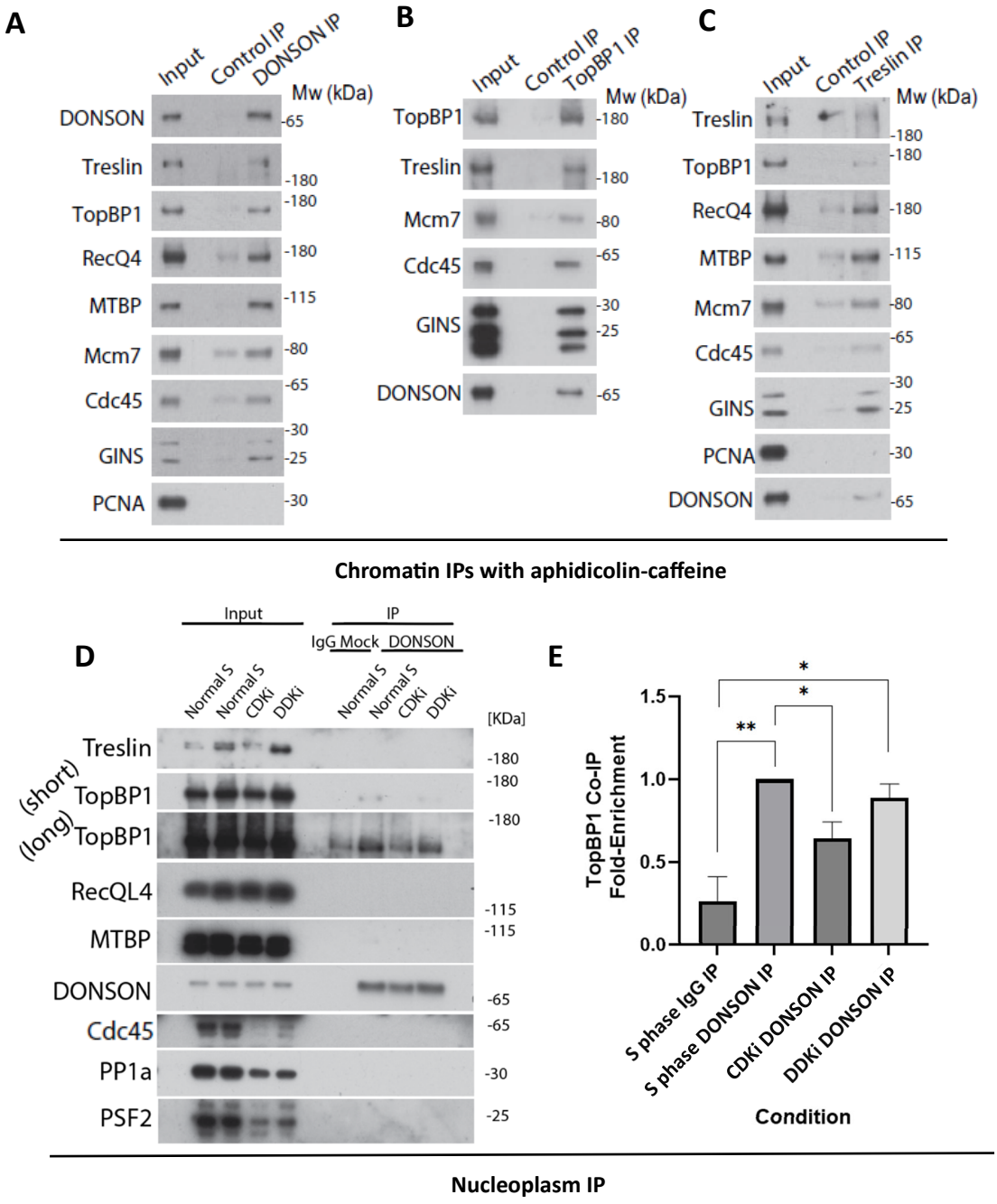


Figure 3.1.13: DONSON interacts with TopBP1. A-C) Egg extract was incubated with sperm DNA to promote DNA replication, chromatin was isolated and immunoprecipitation (IP) was performed with α-DONSON, α-TopBP1 or α-Treslin antibodies, as indicated. Co-immunoprecipitation of proteins was assessed by immunoblotting with indicated antibodies. Chromatin sample without antibodies (Input) and rabbit or sheep IgG antibodies were used as a negative control. Experiments A-C were performed by Dr C. Fernandez-Cuesta. D) DNA replication was set up in extract with the optional addition of replication initiation inhibitors ($p27^{KIP1}$ or PHA-767491). Replicating nuclei were isolated in the middle of S-phase with a buffer without detergent to allow retention of part of the nucleoplasm. DONSON was immunoprecipitated from this (Input) and samples were analysed by immunoblotting with indicated antibodies. Nonspecific sheep IgG antibodies were used as a negative control. E) Quantification of fold change of coimmunoprecipitation of TopBP1 under different conditions in n = 3 experiments (different extracts used). Fold-change levels normalised to the levels in IgG IP sample. Mean with SEM plotted. Unpaired, two-tailed t-tests, * $p < 0.022$, ** $p = 0.0076$.

Having observed an interaction with TopBP1 that is not only specific to chromatin, we set out to determine whether DONSON's binding to chromatin is actually dependent on TopBP1. To do this, Dr Skagia prepared an extract partially depleted of TopBP1. As seen in Figure 3.1.14 A, we were able to deplete the extract to around 25% TopBP1 remaining. We then tested the replication efficiency of the extract and observed that with partial depletion of the protein, replication efficiency dropped to around 50% of the IgG-depleted extract level (Figure 3.1.14 B) [74]. Using this extract, I was then able to assess the impact of partial depletion of TopBP1 on DONSON's chromatin binding. To do this, I initiated a replication reaction in IgG- and TopBP1-depleted extracts as previously, and isolated chromatin at various timepoints during replication. I immunoblotted for different factors using antibodies as shown (Figure 3.1.14 C) and observed that chromatin binding of most factors tested decreased. DONSON chromatin binding was also greatly reduced, suggesting that upon the reduction of TopBP1, DONSON is not efficiently recruited to the chromatin. As we have also observed them interacting in the nucleus, this may suggest that TopBP1 could be required for DONSON to be loaded onto chromatin. As previously reported, we also observe a decrease in GINS chromatin binding upon TopBP1 partial depletion [193]. As we have also previously observed DONSON-dependence for GINS loading, this may suggest that GINS is recruited to chromatin during initiation in a manner that is dependent on both DONSON and TopBP1.

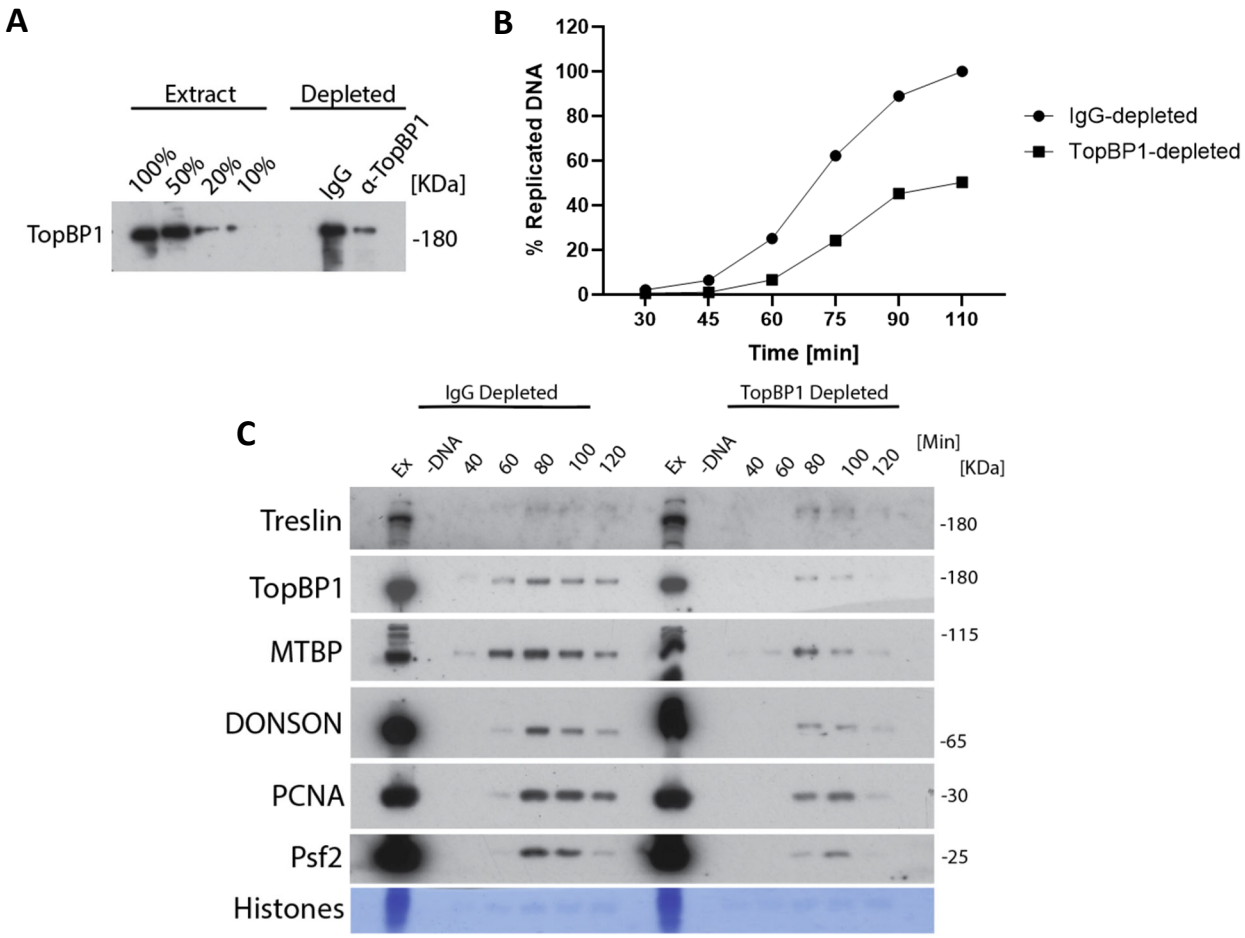


Figure 3.1.14: The immunodepletion of endogenous TopBP1 decreases replication in egg extract. A) *Xenopus* egg extract was immunodepleted of endogenous TopBP1. The levels of TopBP1 in extract were detected by western blotting to analyse depletion efficiency. Purified rabbit α-TopBP1 antibodies were used for immunoblotting. B) IgG- and TopBP1-depleted extracts were incubated with sperm DNA and α-[P32]dATP. The DNA replication reaction was stopped at indicated timepoints and synthesised DNA was measured by α-[P32]dATP incorporation. (A and B, n =2, Performed by Dr. A Skagia.). C) DNA replication was set up in IgG- and TopBP1-depleted extracts, chromatin was isolated at each timepoint indicated and chromatin-bound proteins analysed by immunoblotting. Positive and negative controls were performed as in previous experiments.

3.1.2.6 DONSON Forms a Homodimer and Interacts with Two TopBP1 Molecules

Having observed that TopBP1 is likely required for chromatin-loading of DONSON, we wanted to explore whether the two proteins directly interact. To model TopBP1's interaction with DONSON, we needed first to determine whether DONSON primarily exists as a monomer or whether it forms higher order complexes. We first assessed our WT rDONSON by gel filtration to observe the apparent molecular mass at which the protein was running through the column. Our recombinant DONSON protein was diluted in LFB1/50 and the sample run through a Superose 6 gel filtration column. Eluted fractions were then run on an SDS-PAGE gel and DONSON was immunoblotted using α-DONSON antibodies from indicated species. Interestingly, as seen in Figures 3.1.15 A and B, the size at which our WT rDONSON runs, blotted for with

either antibody, is greater than the known molecular weight of the protein (~66 kDa). This suggests that rDONSON may preferentially form multimeric complexes. However, this could be representative of the protein forming less globular and more flexible structures. Therefore, for a more accurate insight, we assessed our recombinant WT DONSON by mass photometry. As seen in Figure 3.1.15 C, recombinant WT DONSON primarily adopts a dimeric form, measuring at 128.3 ±15.4 kDa, in close agreement to the expected molecular mass of a DONSON dimer at 132 kDa. This is visualised structurally in Figure 3.1.15 E, which is a representation of two molecules with six different domains as indicated in 1D in Figure 3.1.15 D.

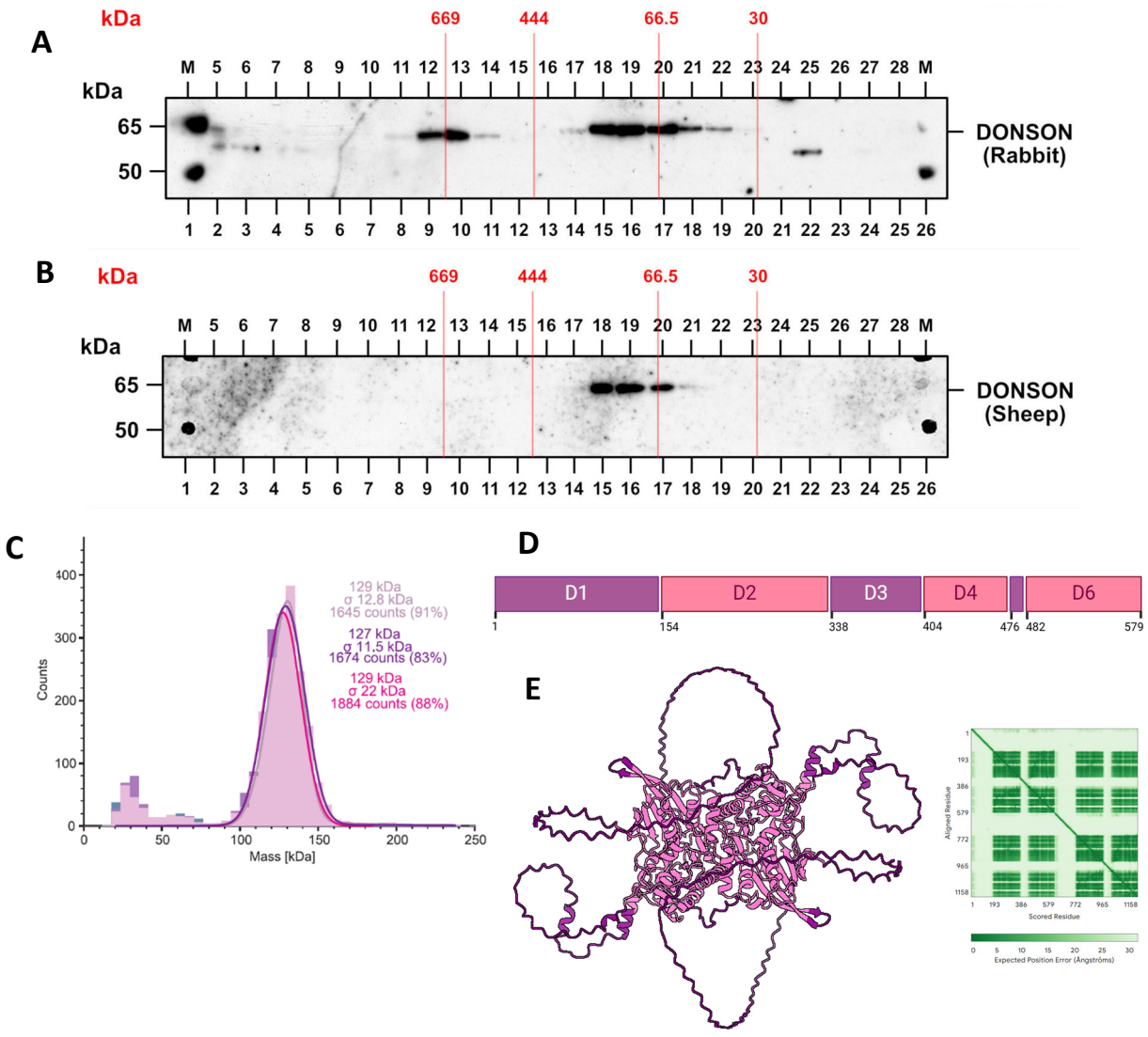


Figure 3.1.15: rDONSON primarily exists in dimer form. A-B) WT rDONSON was diluted in LFB1/50 and analysed by gel filtration using a Superose 6 column, 50 µl fractions were collected, run on an SDS-PAGE gel and α-DONSON antibodies (A: rabbit, B: sheep) were used or immunoblotting. Molecular weight markers are shown above the gel in red. Performed by Dr P Passaretti. C) Purified WT xDONSON was assessed by mass photometry. The mass range of the compound is shown, n=3. D) 1D map of *Xenopus* DONSON structure, including 6 domains indicated. E) Model of DONSON dimer with each domain coloured to match D, created using AlphaFold3. ipTM = 0.63 pTM = 0.64. Corresponding expected position error plot shown.

Having identified that DONSON preferentially exists in dimer form, we could better predict how it may interact with TopBP1. To determine this, we modelled the interaction between the proteins using AlphaFold3. Figure 3.1.16 A depicts the 1D structure of TopBP1, with the BRCT domains numbered. BRCT 1, 2 and 3 have previously been reported to be essential to support DNA replication [74]. As we are interested in the initiation function of DONSON, we have included these domains specifically in Figure 3.1.16 C. The AlphaFold predicted structure depicts each DONSON monomer (light and dark grey) flipped 180° compared to one another, in the dimer form (Figure 3.1.16 B and C). Interestingly, the BRCT3 domain is predicted to fit in the groove formed between the two DONSON monomers of the dimer. As the dimer is symmetrical, we can model two BRCT3 domains fitting at the top and bottom between the monomers (Figure 3.1.16 B). Moreover, the three BRCT domains required to support replication initiation are predicted to fit within the interface between the two DONSON molecules in the dimer. With this prediction, we wanted to determine whether the two proteins do interact directly through these regions. To assess this, we purified truncated forms of the TopBP1 protein – the N-terminus (containing the first 575 amino acids of the protein, including BRCT 1, 2, 3) and the C-terminus (containing the last 350 amino acids of the protein, including BRCT 7, 8) – and WT DONSON and GINS (Figure 3.1.16 D). We were then able to combine the proteins and peptides *in vitro*, and immunoprecipitate DONSON. We ran the IP samples in an SDS-PAGE gel and assessed coimmunoprecipitation of the TopBP1 peptides by blotting with indicated antibodies (Figure 3.1.16 E). We observed that, whilst the N-terminal TopBP1 peptide is able to coimmunoprecipitate with DONSON (1st IP lane), the C-terminal peptide cannot (3rd IP lane). This suggests that, as predicted by AlphaFold, only the first three BRCT domains of TopBP1 do interact with DONSON. As we had observed DONSON interacting with GINS, we wanted to assess whether this interaction was dependent on TopBP1, as it is in yeast [67]. Therefore, we also added recombinant GINS in our reaction containing FLAG-tagged Psf2. We observed that GINS weakly interacts with recombinant DONSON *in vitro*, particularly when in complex with the N-terminal fragment of TopBP1, but this reaction is not robust and can only be detected using strong FLAG antibody. This indicates that the principal interaction occurs between DONSON and TopBP1, likely forming a platform for GINS recruitment to chromatin.

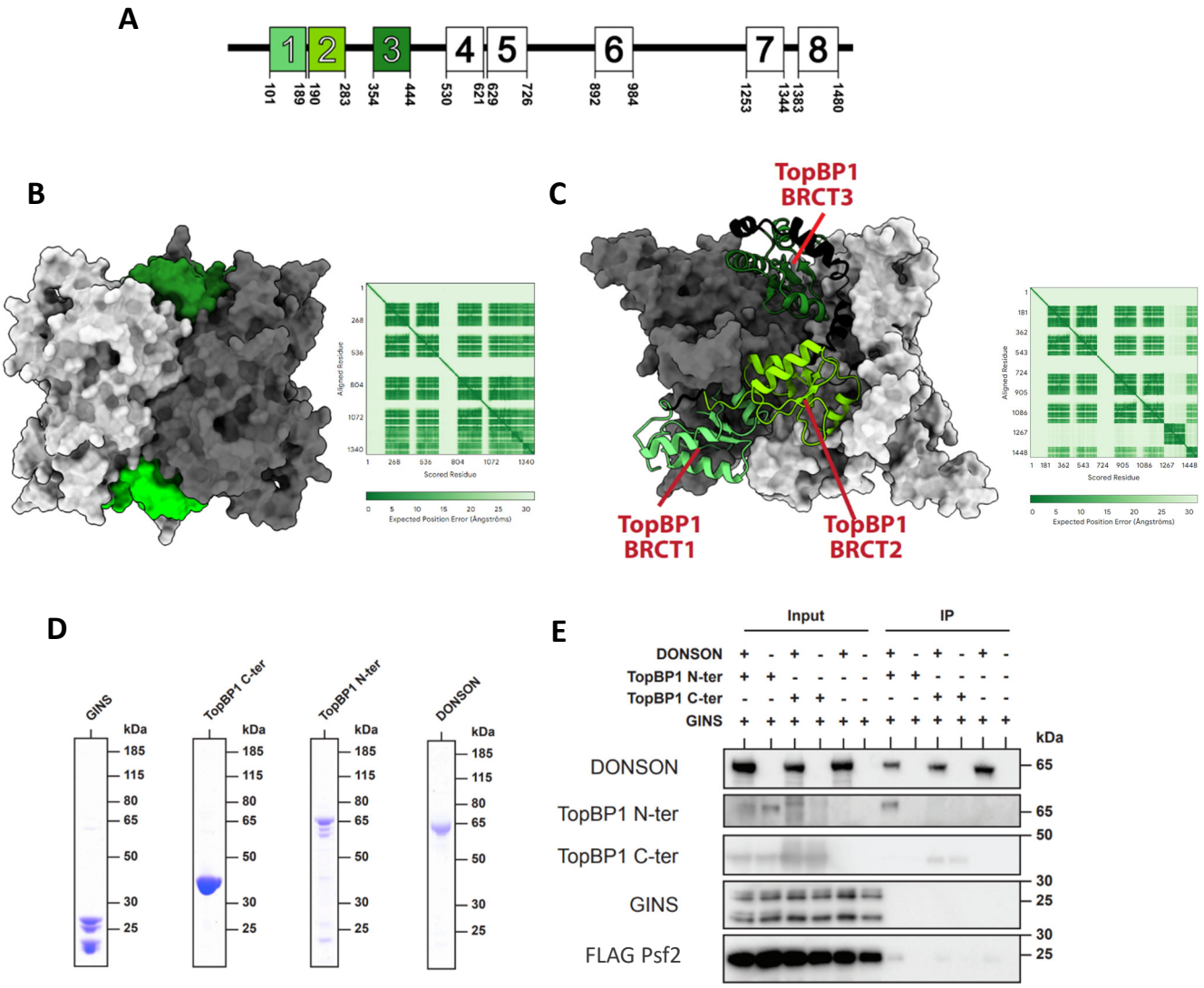


Figure 3.1.16: DONSON interacts with TopBP1 via its BRCT3 domain. A) Model of TopBP1 structure, with BRCT domains indicated. B) AlphaFold prediction of the DONSON dimer (grey) interacting with TopBP1 BRCT3 (green). ipTM = 0.67 pTM = 0.66. Corresponding expected position error plot shown. C) AlphaFold prediction of the interaction of TopBP1 BRCT1-3 (indicated) – domains essential for replication – with dimeric DONSON (grey). ipTM = 0.52pTM = 0.57. Corresponding expected position error plot shown. D) Purified proteins run on SDS-PAGE gel and visualised with Coomassie staining. E) In vitro IP with WT rDONSON, truncated N-terminal rTopBP1, truncated C-terminal rTopBP1 and WT rGINS. Proteins were diluted and mixed in LFB1/50 in combinations indicated and samples were run in an SDS-PAGE gel and blotted with specified antibodies. Performed by Dr P Passaretti.

3.1.3 Discussion

This chapter of my thesis focuses on the role of DONSON in the establishment of the active helicase and support of DNA replication. Using *Xenopus laevis* egg extract as a model system, preliminary mass spec data from the lab identified DONSON as an interactor of the terminated helicase [35]. The interaction with the replicative helicase has also been observed in human cells throughout S phase [156]. I have also observed DONSON interacting with replisome components on chromatin. In addition to this, I have identified that DONSON binds to chromatin with the same kinetics as known components of the active helicase throughout replication in normal S phase and when replication is perturbed under different conditions. Additionally, I have demonstrated that DONSON is essential for DNA replication initiation, specifically to support formation of the active helicase on DNA. We suggest that DONSON is likely required to directly recruit GINS. Finally, I have shown that DONSON acts as a dimer to support active helicase formation. DONSON may perform this function in a TopBP1-dependent manner.

The results in this chapter, demonstrate that DONSON likely acts as the functional homologue of the yeast protein Sld2, which is required for recruitment of GINS to chromatin during DNA replication initiation in a Dpb11 (TopBP1)-dependent manner [67]–[69]. This is indicated particularly by the depletion of DONSON, where both DNA replication and GINS recruitment to chromatin during initiation is inhibited. The same results have been observed by the Walter and Tanaka groups who both used *Xenopus laevis* egg extract in similar experiments [194], [195]. In addition to this, the Labib group proposed the same role for DNSN-1 (DONSON) in *C. elegans* [196]. The *Drosophila* orthologue of DONSON, Hd, has likewise previously been observed to be essential for DNA replication [159]. Interestingly, the Tanaka group additionally propose that DONSON recruits Cdc45 to the chromatin, suggesting that they interact in the nucleus prior to interaction on chromatin. In line with this, in Figure 3.1.8 B we observe that upon DONSON depletion, Cdc45 chromatin binding, not just GINS, is lost. However, we have never observed DONSON interacting with Cdc45 outside of the whole replicate on chromatin. Therefore, we propose, along with the Walter and Labib groups, that this is due to GINS recruitment by DONSON being required to stabilise Cdc45 on the chromatin in the CMG form, as opposed to directly recruiting Cdc45 to the chromatin. Similarly, work from the Chistol group suggests that GINS recruitment is required to stabilise Cdc45 on the chromatin and

whilst two copies of Cdc45 are recruited to the chromatin via a dimerisation scaffold, that this scaffold is not DONSON [197]. This will be discussed further in the next chapter. In yeast, Sld3-Sld7 directly interacts with Cdc45 to recruit it to the Mcm2-7 hexamer [61]. Sld2 is not required for recruitment of Cdc45 to the chromatin, but its recruitment of GINS is indeed required for CMG formation and to stabilise the Cdc45-Mcm2-7 binding [59], [75], [101]. In yeast, the firing factors required for CMG formation are required to bridge the two helicases together to form a stable construct, therefore supporting Cdc45 on chromatin [41]. We likewise propose that DONSON adopts this role in higher eukaryotes and this is likely the reason we see reduced Cdc45 binding on chromatin in DONSON's absence.

In yeast, recruitment of GINS by Sld2 is dependent on its interaction with Dpb11 [67]. Our finding that DONSON may act as the Sld2 homologue is somewhat surprising, as the helicase RecQL4 has traditionally been associated with this role in higher eukaryotes. RecQL4 is structurally the closest vertebrate homologue of Sld2, with sequence homology between *S. cerevisiae* Sld2 and the N-terminus of *Xenopus* RecQ4 [72], [73]. RecQ4 has been implicated in DNA replication in both *Xenopus* and *Drosophila* where, upon its downregulation, normal replication becomes severely defective [198]. Specifically, these defects have been observed during initiation, where Sld2 also functions, as opposed to during elongation [199], [200]. In addition to this, RecQ4 has been observed to interact with TopBP1, at its BRCT7-8 domains [193]. However, BRCT4-8 of TopBP1 have been shown to be dispensable for DNA replication initiation, indicating that the RecQ4-TopBP1 interaction is not required for initiation in the same way as Sld2-Dpb11 is [74]. Likewise, RecQ4 has been shown to play a role in replication initiation only after the CMG has been successfully assembled in *Xenopus*, specifically after TopBP1/Cut5, Sld5 and Cdc45 have been loaded onto chromatin [72]. Additionally, RecQ4 has been shown to be essential for chromatin loading of Pol α, following CMG formation [73]. Collectively, this work has concluded that RecQ4 plays a later role in DNA replication initiation than its previously proposed homologue Sld2 [201]. Interestingly, in *C. elegans* Sld-2, RecQ4 and DNSN-1 are all present. It has therefore been previously suggested that over evolution the helicase domain (RecQ4) may have separated from the non-catalytic domain (Sld2) [202]. Likewise, as both Sld2 and DONSON are required for GINS recruitment to chromatin, it is suggested that DONSON and RecQ4 developed early during eukaryotic evolution, after which Sld2 was lost [196]. This is a similar story

for the initiation factor Treslin, whose homologue in yeast is Sld3. In *C. elegans* both SLD-3 and the Treslin N-domain-containing protein TRES-1 are present, indicating that Treslin may have developed from Sld3 in higher eukaryotes, and that this may be a pattern of protein development in *C. elegans* [196], [202].

Sld3-Sld7 and the higher eukaryotic homologues Treslin-MTBP, have been observed to recruit Cdc45 to the chromatin, in an interaction that requires TopBP1 [64], [203]–[205]. Dpb11 interacts with phosphorylated Sld3 already bound to the chromatin [206]. Phosphorylated Treslin has been identified as the binding partner for TopBP1's BRCT1-2 domains, previously shown to be essential for DNA replication [74]. However, the binding partner for TopBP1's third domain (BRCT3) which is also essential for replication, has not previously been identified. In Figure 3.1.16 we show that not only does DONSON interact with TopBP1, but specifically that this interaction involves the first three BRCT domains of TopBP1, and may be specific to BRCT3. Together, this supports our proposal that DONSON is the homologue of yeast Sld2, as it appears to directly bind TopBP1 at a domain that is essential for replication.

Although the results from this chapter indicate that DONSON acts during replication initiation, the protein has previously been proposed to stabilise replication forks upon replication stress [152]. Importantly, we confirmed that the loss of DNA replication and fork stability we observe upon DONSON's downregulation in our experiments is not due to spontaneous fork stalling and DNA damage, as the phenotype could not be rescued in DONSON-depleted extract upon addition of caffeine, an ATM/ATR inhibitor. Likewise in *Drosophila*, DONSON's functional orthologue, Hd, is required to maintain genome stability, as well as to support DNA replication [159]. Embryos containing mutated Hd inherited from their mothers, exhibit incompletely replicated chromosomes and arrested nuclei, suggesting a defect in DNA replication and subsequent cell cycle progression. Similarly, DONSON has been implicated in forms of Microcephalic Primordial Dwarfism (MPD), including a subtype called Meier Gorlin Syndrome (MGORS) [152], [154], [155]. MGORS is classically associated with proteins involved during replication initiation. Mutation of factors associated with later stages of replication are not classically implicated in development of MGORS [207]. Therefore, this may support our findings that DONSON also plays a role during replication initiation.

Addressing the question of whether DONSON travels with the fork throughout replication to perform both roles is important. Figures 3.1.3 and 3.1.4 indicate that DONSON binds chromatin with the same kinetics as other replisome components during normal S phase and when replication is perturbed at different stages. Additionally, our IPs from chromatin during elongation and post-termination (Table 3.1 and Figure 3.1.5) all indicate that DONSON interacts with replisome components at the stages tested for, suggesting that DONSON may travel with the replication fork throughout replication. In line with this, through isolation of proteins on nascent DNA (iPOND) in human cells, DONSON has been observed enriched on newly replicated DNA along with Mcm2-7 proteins. This suggests that the DONSON interacts with the replisome once DNA replication has already proceeded [152]. Additionally, in untreated HeLa cells, two forms of replisome can be observed throughout S phase: those containing DONSON, as well as those without [156]. This indicates that, under certain conditions, DONSON can be observed bound to the replisome at various points during S phase, which may suggest that DONSON travels with the replisome. In my experiments, using chromatin isolation, we can identify DONSON binding chromatin when peak replication occurs during S phase. However, neither of these techniques can determine whether DONSON remains bound to chromatin throughout elongation, or whether the observed signal comes from new firing origins. In contrast, using single molecule analysis in egg extract, the Chistol group have been able to visualise the dynamics with which DONSON binds to and is subsequently released from the chromatin. They identified that TopBP1, RecQL4, and DONSON all transiently bind the CMG to support initiation, but none of them travel with the replication fork [197]. Interestingly, once the DONSON dimer has recruited two GINS complexes to each double hexamer, under normal conditions it can be visualised rapidly dissociating from the chromatin. DNA synthesis does not begin before DONSON's dissociation from the chromatin, indicating that it must be removed for the helicase to begin unwinding DNA. This indicates that although DONSON can be observed associating with replisome components and on the chromatin throughout S phase, that this is representative of incidences of initiation, or re-recruitment upon replication stress, and that DONSON does not travel with the fork after helicase activation. In addition, the Tanaka group observed that following replication stress, DONSON is not required to promote replication fork restart [195]. As DONSON has previously been shown to be required to maintain fork stability as part

of the ATR stress response in human cells, this indicates that there may be differences between model systems used [152]. However, this needs further investigation.

Altogether, this chapter has shown that DONSON is required to support initiation of DNA replication, in particular to support formation of the CMG on chromatin. DONSON's role in supporting DNA replication appears to be CDK- and TopBP1-dependent and we propose that DONSON and TopBP1 together support recruitment of GINS to chromatin. This will be discussed further in the next chapter.

3.2 How does DONSON carry out its initiation function?

3.2.1 Introduction

Since starting work to understand DONSON's function during DNA replication, DONSON's role as a novel replication initiation factor became well established. Work in *Xenopus laevis* egg extract, *human* HCT116 cells, *mouse* embryonic stem cells and *Caenorhabditis elegans*, all successfully identified DONSON as being essential for formation of the active helicase during DNA replication initiation [194]–[196], [208]. However, the mechanism by which DONSON supports GINS binding to the Mcm2-7 hexamer to form the active helicase was yet to be determined.

Discovering the structural interactions of the components within the replisome complex had the potential to shed light on how DONSON acts during replication initiation. The structure of the CMG and active replisomes on DNA has been identified in many organisms, through replisome reconstitution [59], [76], [86], [209] and cryo-electron microscopy (Cryo-EM) in *Saccharomyces cerevisiae* [210], [211], *Drosophila melanogaster* [212] and human cells [213]–[215]. Using the yeast replisome reconstitution system, protein factors involved in the process have been successfully identified, including Sld2 – the protein for which we propose DONSON is the higher eukaryotic homologue – and Dpb11 – the TopBP1 homologue [59], [86]. Additionally, the structure of human TopBP1 has been established in small fragments [216]–[220], along with its interaction with GINS via the GINI domain [193]. However, the structure of the *Xenopus* CMG and DONSON, as well as DONSON's molecular interaction with the replisome in any organism, was yet to be determined.

Therefore, the aim of this thesis chapter was to determine the way in which DONSON directly interacts with the CMG, identifying which parts of the protein are particularly important for its initiation function. With particular interest in how DONSON interacts with GINS to bring it to the chromatin and so determine how it carries out this initiation function.

3.2.2 Results

3.2.2.1 The Importance of DONSON's Flexible Regions

Having observed that helicase formation is dependent on DONSON, we wanted to determine the part of DONSON required to support this interaction. Although no *Xenopus* structure had been identified when I began my project, using the AlphaFold predicted structure we could observe two flexible regions – one at the N-terminus and one extending from the central globular domain. We wanted to test a potential role for these flexible domains for active helicase formation and for stabilisation of the complex.

To assess this, we expressed truncated versions of the His- and MYC-tagged DONSON – ΔP334-L410 (ΔLoop) and ΔA2-V155 (ΔN) – in *E. coli* and purified the proteins as with the WT, using a Nickel column. Samples taken during purification were run in an SDS-PAGE gel and visualised by Coomassie staining (Figure 3.2.1 A, B). Following successful protein purification, I set up DNA replication in IgG- and DONSON-depleted extracts, and optionally supplemented with the WT, ΔLoop or ΔN rDONSON, and assessed the ability of the extracts to support DNA synthesis. As seen in Figure 3.2.1 D, whilst DNA replication is rescued by WT rDONSON in DONSON-depleted extract to similar levels as the IgG-depleted, with either of the two truncated mutants DNA synthesis remains low, ~15-35% of the IgG-. This suggests an importance for both these flexible regions to support efficient DNA replication and that they may be required for DONSON's initiation function in particular.

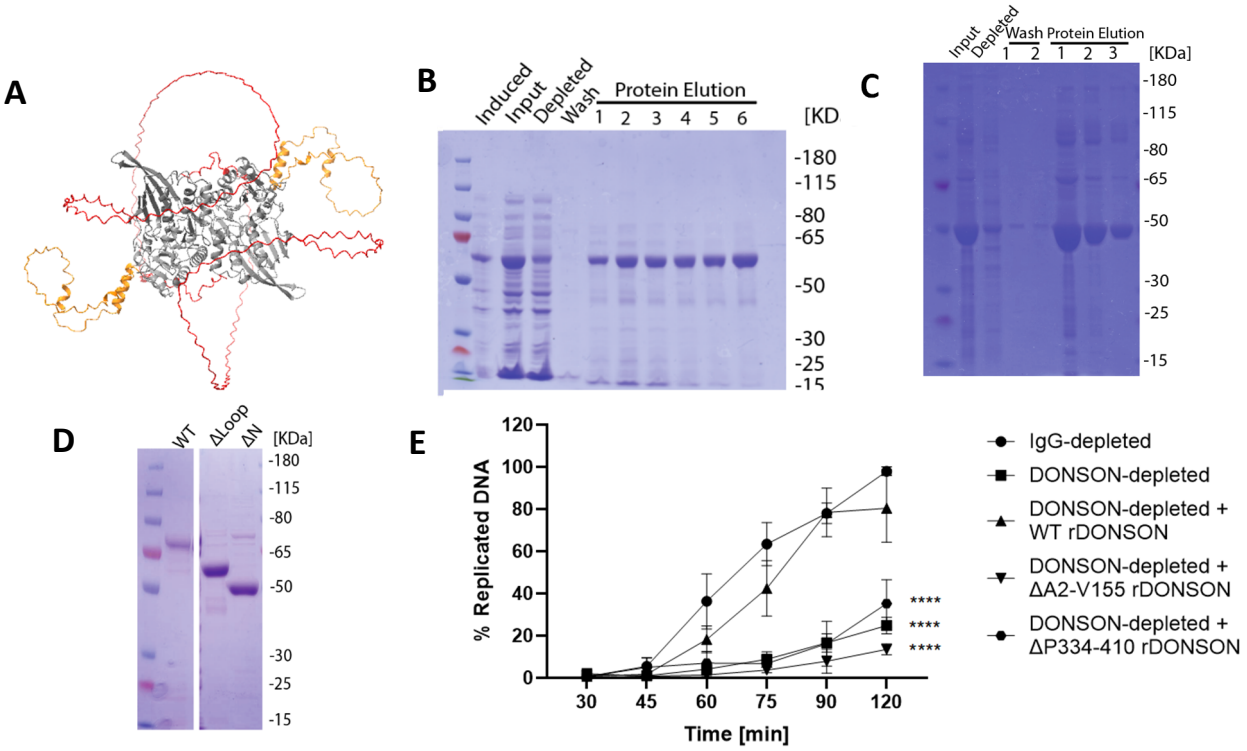


Figure 3.2.1: DONSON's flexible regions are required to support nascent DNA synthesis. A) AlphaFold3 model of dimeric DONSON with the flexible N-terminus and the protruding loop highlighted red and orange respectively. ipTM = 0.63 pTM = 0.64. Corresponding expected position error plot shown in Figure 3.1.15. B-C) Purification of His- and MYC-tagged DONSON ΔP334-L410 (B) and ΔA2-V155 mutants (C). *E. coli* carrying the mutated DONSON plasmids were grown (uninduced) and expression was induced in *E. coli* (Induced) and purified as described in methods section using Nickel affinity purification. Supernatant from bacterial lysate containing soluble proteins (Input), unbound proteins (Depleted), washes (wash) and elutions using 250 mM of Imidazole (Protein elutions 1-6) are shown. Samples from the purification were run on an SDS-PAGE gel and visualised by Coomassie staining. D) Samples of each protein at the same concentration were run in an SDS-PAGE gel and visualised by Coomassie staining. E) DNA replication was set up in IgG- and DONSON-depleted extract with the optional addition of WT rDONSON or ΔP334-L410 and ΔA2-V155 mutants, and incubated with α- [P32] dATP. The DNA replication reaction was stopped at indicated timepoints and DNA synthesis was measured as in previous figures. Mean and SEM are plotted. Two-way ANOVA comparing to IgG-depleted, ****p < 0.0001.

3.2.2.2 DONSON Directly Interacts with the CMG

Having observed a requirement for the flexible loop and N-terminal tail in supporting DNA replication, we wanted to identify exactly how DONSON interacts with replisome components, with the view to shed light on how it promotes DNA synthesis. To assess the structures of *Xenopus* CMG and replisome, Dr P. Passaretti purified a FLAG-tagged Cdc45 protein and incubated it with the egg extract. This enabled isolation of the CMG on chromatin, via an α-FLAG IP. The complexes formed could then be visualised by negative staining (Figure 3.2.2). Strikingly, we observed CMG monomers, but also higher order assemblies that after further investigations turned out to be a DONSON-bound double CMG (dCMGDo).

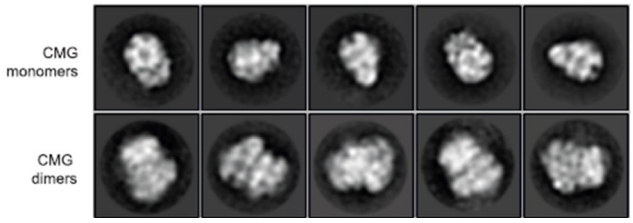


Figure 3.2.2: Isolating CMG monomers and dimers. Performed by Dr P Passaretti. DNA replication was set up in extract as before, with the optional addition of Cdc45-TEV-His_{10}-$FLAG_5$, aphidicolin and caffeine. Chromatin was isolated at 60 minutes, digested with benzonase and the tagged-Cdc45 immunoprecipitated. The sample was assessed with negative stain, using uranyl acetate. 2D averages of single (row 1) and double (row 2) complexes are presented.

The CMG-dimer assemblies were analysed by Dr Milos Cvetkovic and Dr Alessandro Costa by cryo-EM imaging. From the data produced, complexes of the double CMG, with dimeric DONSON bridging across their N-termini could be observed (Figure 3.2.3). The globular regions of DONSON position at the interface between the two helicases, in particular bridging between Mcm3 and GINS.

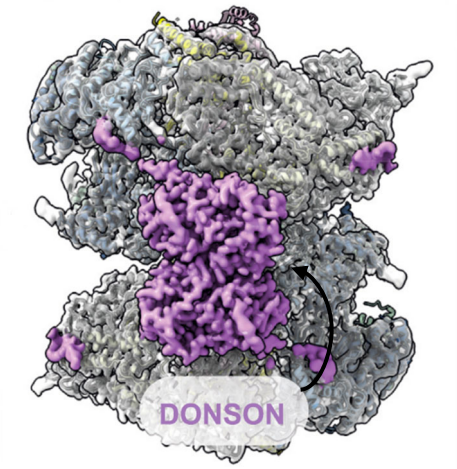


Figure 3.2.3: Dimeric DONSON bridges the double CMG. Performed by Dr P Passaretti, Dr M Cvetkovic and Dr A Costa. DONSON-bound double CMG atomic model from the unassigned Cryo-EM data. Dimeric DONSON is depicted in purple – indicated by the arrow.

3.2.2.3 DONSON Interacts with Mcm3

Having observed this interesting assembly, Dr Milos Cvetkovic, Dr Agata Butryn and Dr Alessandro Costa built an atomic model of this structure to identify specific interactions between DONSON and individual factors of the CMG. Figure 3.2.4 A shows the model of a single DONSON-bound CMG from the cryo-EM data. As seen, within the complex DONSON interacts directly with Mcm3 at two sites – the blue shaded area showing Mcm3. DONSON's structure is represented in purple. The first interaction between DONSON and Mcm3 (Figure 3.2.4 B), occurs between the ATPase domain of Mcm3 and an alpha helix motif within

the flexible loop that I truncated in Figure 3.2.1. This is shown in the cryo-EM model (dark purple) which also closely resembles the AlphaFold prediction of these interactions (light purple). The cryo-EM data indicates a particularly well-resolved Trp381 in the DONSON alpha-helix that engages Lys285 of Mcm3. The AlphaFold model predicts two further interactions between Mcm3 and DONSON within this region – Asp374 and Glu377, as seen in light purple. This suggests that the importance of this flexible loop in supporting DNA synthesis, as seen in Figure 3.2.1 D, may be due to its direct interaction with Mcm3, without which the active helicase may not be established. The second interaction observed in the cryo-EM data involves the N-terminal A-domain of Mcm3 (Figure 3.2.4 C), between DONSON Ser437 and Mcm3 Gln27. Interestingly, this Mcm3 site is proximal to one engaged by Psf3 in the CMG. This suggests that DONSON may bind to this site of Mcm3 to specifically enable Psf3 to be deposited adjacently and establish the CMG.

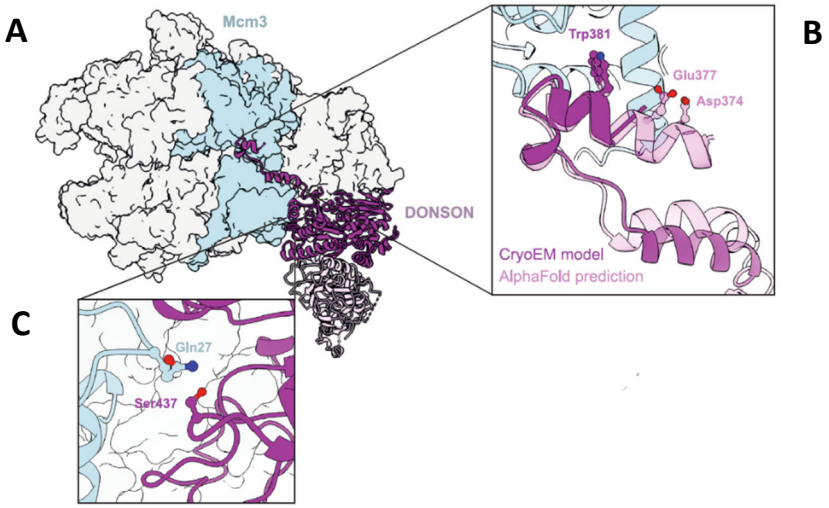


Figure 3.2.4: DONSON directly interacts with Mcm3. Performed by Dr Passaretti and researchers from Dr Alessandro Costa's lab at the Francis Crick Institute. A) One isolated DONSON-bound CMG is presented. B) The flexible loop of DONSON is presented engaging with the ATPase domain of Mcm3. The darker purple represents the atomic model from the cryo-EM density. The lighter purple represents the AlphaFold prediction. Both models identified DONSON Trp381 as an Mcm3 ATPase interacting site. AlphaFold also predicts Asp374 and Glu377 to engage with the Mcm3 ATPase domain. C) The atomic model from the cryo-EM data is presented, identifying the DONSON Ser437 interacts with the Mcm3 Gln27.

Having observed these interactions between DONSON and Mcm3, I wanted to assess whether they are required to support efficient DNA replication. To determine this, I expressed in *E. coli* three mutated versions of DONSON, with the identified sites mutated to alanine. The first at the 'loop' sites - D374A, E377A, W381A (DEW), the second at S437A, and the third being a combination of the two D374A, E377A, W381A, S437A (DEWS). The DEWS mutant was designed specifically to determine whether the

combination of sites together is required to support replication, even if separately they aren't essential. We purified the proteins as previously. Samples taken during purification were run in an SDS-PAGE gel and visualised by Coomassie staining (Figure 3.2.5 A, B, C). I then set up replication in IgG- and DONSON-depleted extracts as before and optionally incubated with the WT and mutated DONSON proteins to assess their ability to support nascent DNA synthesis. As seen in Figure 3.2.5 E, each mutant had reduced capacity to support efficient replication. Whilst the WT protein could successfully rescue DNA replication as observed before, the S437A mutant could rescue to ~70%. By comparison, the DEW mutant could rescue to <60% of the IgG-depleted extract level, and the DEWS mutant, could not rescue much higher than in the DONSON-depleted extract itself, ~40%. This suggests that all our identified DONSON-Mcm3 interactions are important to support efficient replication. However, the DONSON alpha-helix within the flexible loop region appears to be of greater importance for efficient replication, and the combination of the two sites allows for an increase in efficiency and may represent increased stabilisation of the complex.

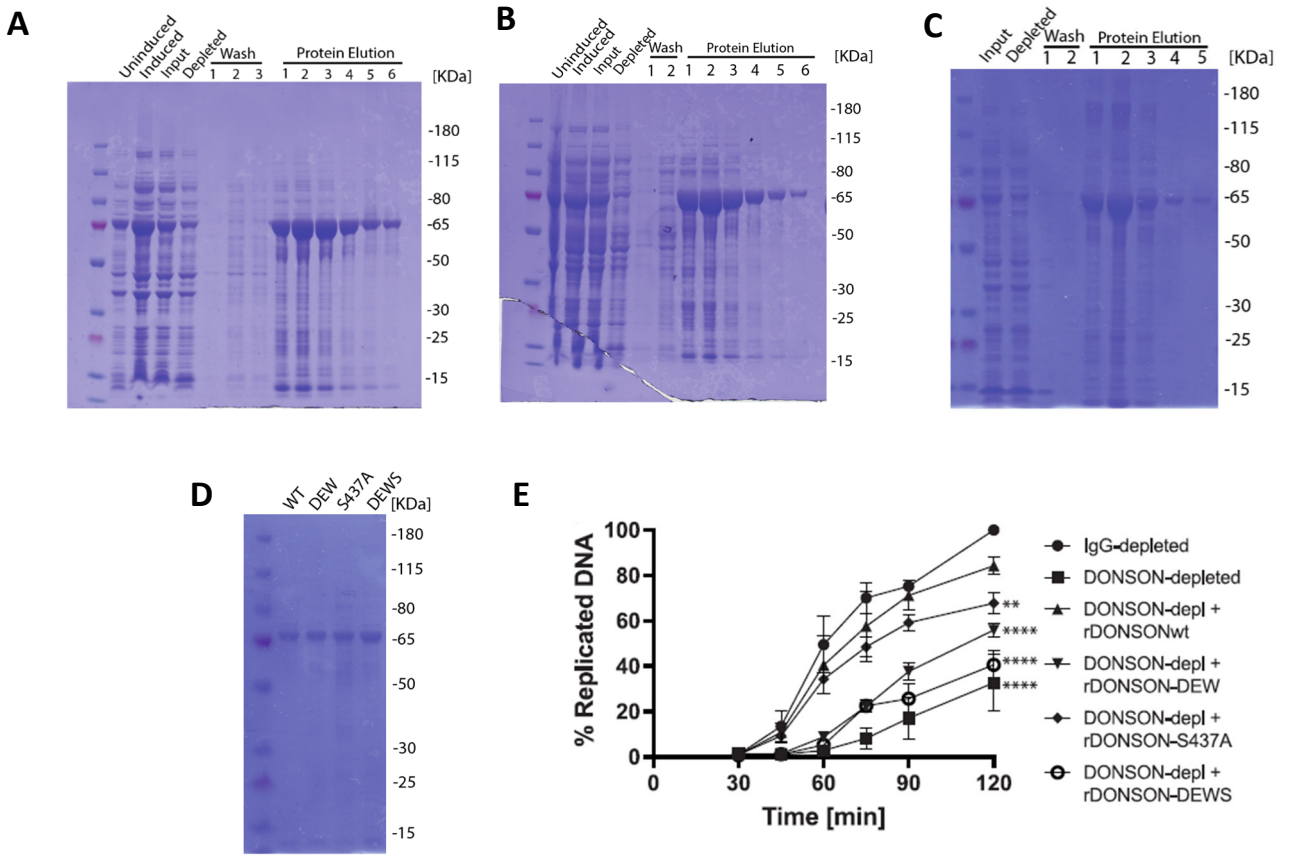


Figure 3.2.5: DONSON's interaction with Mcm3 is required for DNA synthesis. A-C) Purification of His- and MYC-tagged DONSON D374A, E377A, W381A (DEW) (A), S437A (B) and D374A, E377A, W381A, S437A (DEWS) (C) mutants. *E. coli* carrying the mutated DONSON plasmids were grown (uninduced) and expression was induced in *E. coli* (Induced) and purified as described in methods section using Nickel affinity purification. Supernatant from bacterial lysate containing soluble proteins (Input), unbound proteins (Depleted) and elutions using 250 mM of Imidazole (Protein elutions 1-6) are shown. Samples from the purification were run on an SDS-PAGE gel and visualised by Coomassie staining. D) Samples of each protein at the concentration added into extract were run on an SDS-PAGE gel and visualised by Coomassie staining, for comparison of accurate concentration calculation. E) DNA replication was set up in IgG- and DONSON-depleted extract with the optional addition of WT, DEW, S437A or DEWS rDONSON and incubated with α-[P32]dATP. The DNA replication reaction was stopped at indicated timepoints and DNA synthesis was measured as in previous figures. Two-way ANOVA comparing to IgG-depleted, **** $p < 0.0001$. ** $p = 0.0046$.

3.2.2.4 DONSON Interacts with Sld5

The second important point of DONSON's interaction is Sld5 (Figure 3.2.6 A), at two sites. The first interaction occurs between the A-domain of Sld5 and residues 6-22 of DONSON (Figure 3.2.6 B) represented by the dark purple atomic model, which strongly matches the AlphaFold predicted structure. Interestingly, these residues form part of the region I truncated in the ΔN mutant (Figure 3.2.1) and may explain why the mutant was unable to support replication. The second interaction occurs between a hairpin loop (amino acids 473-482) within the core of DONSON, and the B-domain of Sld5 (Figure 3.2.6 C). Despite this strong indication of this interaction, it was not possible to identify single interacting residues

from either protein. AlphaFold predicted this interface with high confidence and suggested three sites in particular that interact with Sld5 in this region: Arg476, Tyr481, Val473.

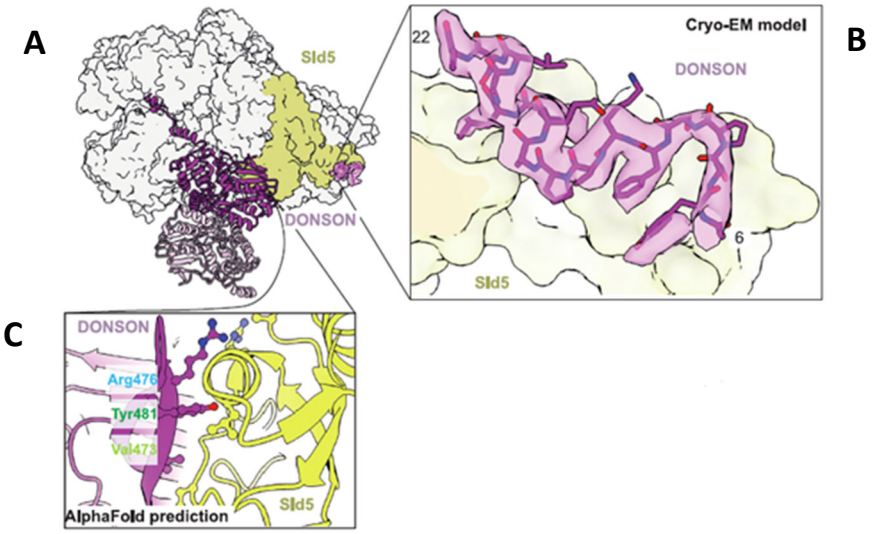


Figure 3.2.6: DONSON directly interacts with Sld5. A) One isolated DONSON-bound CMG is presented. B) The atomic model from the cryo-EM density of the N-terminal tail of DONSON is presented engaging with Sld5 A-domain. C) The AlphaFold predicted interaction between 3 sites within the DONSON globular domain – Arg476, Tyr481, Val473 – and the Sld5 B-domain.

As with the Mcm3 interaction sites, I wanted to ascertain whether these interactions observed between DONSON and Sld5 were required to support replication. To achieve this, I mutated DONSON at these sites, truncating the first 24 amino acids (Δ24), to cover the range of the helix at the N-terminus, or changing the three sites of the hairpin loop to alanine – R476A, Y481A, V473A (3A) – and expressed them in *E. coli*. We purified the proteins as before and visualised samples by Coomassie staining (Figure 3.2.7 A, B). I then tested replication in IgG- and DONSON-depleted extracts as before, optionally incubating with the WT protein or these mutants, to assess their capacity to support nascent DNA synthesis. As seen in Figure 3.2.7 D, whilst the 3A mutant can successfully rescue replication to similar levels to the IgG-depleted extract, the Δ24 mutant has no capacity to support DNA replication, with no increase from the DONSON-depleted extract. This suggests either that the mutation of the three predicted sites within the hairpin loop is not enough to disrupt this interaction, or that this interaction is not essential to support replication initiation. In contrast, the N-terminal helix does appear to be essential to support replication initiation. This suggests this is the mechanism by which DONSON binds GINS to recruit it to the chromatin for CMG assembly.

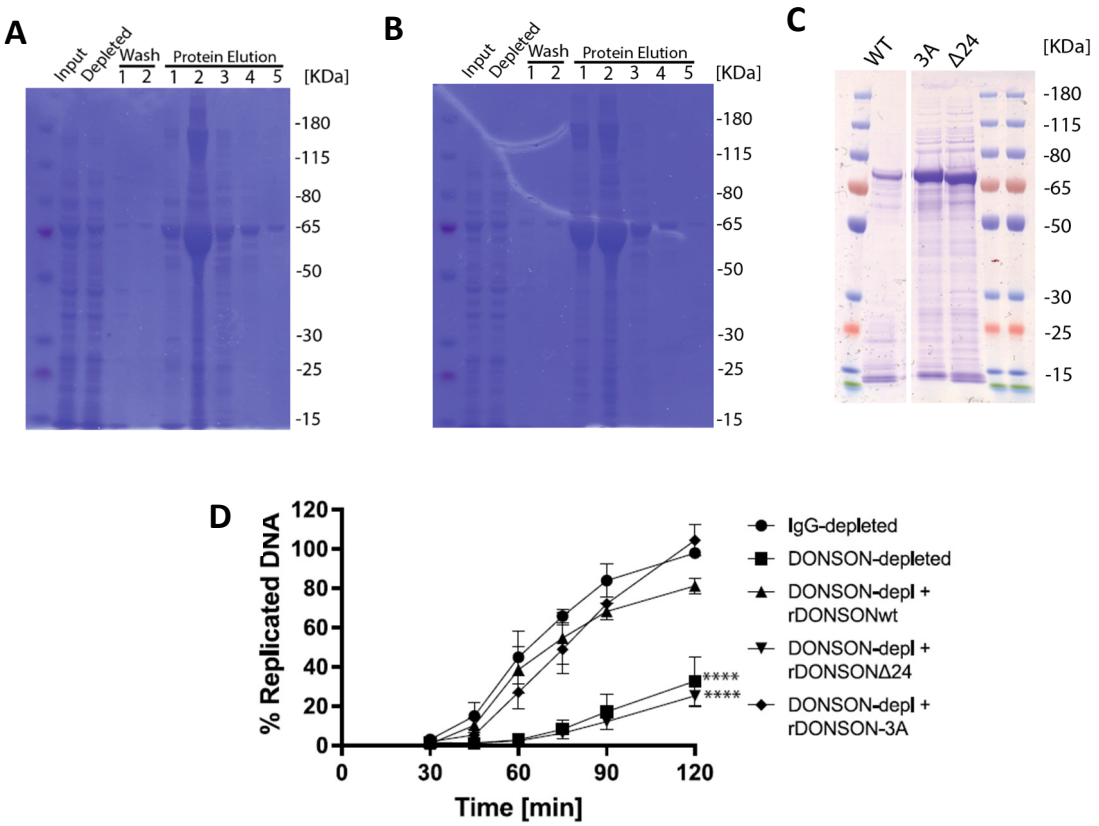


Figure 3.2.7: DONSON's interaction with Sld5 is required for DNA synthesis. A-B) Purification of His- and MYC-tagged DONSON R476A, Y481A, V473A (3A) (A) and Δ24(B) mutants. Expression of xDONSON mutants were induced in *E. coli* (Induced) and purified as described in methods section using Nickel affinity purification. Supernatant from bacterial lysate containing soluble proteins (Input), unbound proteins (Depleted) and elutions using 250 mM of Imidazole (Protein elutions 1-6) are shown. Samples from the purification were run on an SDS-PAGE gel and visualised by Coomassie staining. C) Samples of each protein at the concentration added into extract were run on an SDS-PAGE gel and visualised by Coomassie staining, for comparison of accurate concentration calculation. D) DNA replication was set up in IgG- and DONSON-depleted extract with the optional addition of WT, 3A or Δ24 rDONSON and incubated with α-[P32]dATP. The DNA replication reaction was stopped at indicated timepoints and DNA synthesis was measured as in previous figures. Two-way ANOVA comparing to IgG-depleted, **** $p < 0.0001$.

Having identified the importance of the first 24 amino acids of DONSON in supporting DNA replication, I sought to determine whether this is specifically for formation of the active helicase on the chromatin, as we suspected. To test this, I incubated IgG- and DONSON-depleted extracts with sperm DNA as before and optionally supplemented with either the WT protein or the Δ24 mutant. I isolated chromatin from all extract samples at different timepoints throughout replication, ran the samples in an SDS-PAGE gel and assessed the loading of the factors on chromatin throughout replication by immunoblotting with antibodies specific to those factors. As seen in Figure 3.2.8 A, and as we had previously observed in Figure 3.1.8, depletion of DONSON from the egg extract results in reduced chromatin binding of both Cdc45 and GINS proteins, whilst addition of the WT protein could rescue this binding. In contrast, on addition of the Δ24 mutant, chromatin binding of both Cdc45 and GINS remain at similar levels as in the DONSON-

depleted extract. This indicates the importance of the N-terminal helix for formation of the active helicase on chromatin during replication initiation. Interestingly, we did not observe an interaction between DONSON and Cdc45 in our cryo-EM data, suggesting that although DONSON doesn't recruit Cdc45 to chromatin, it's recruitment of GINS is necessary to stabilise it there. Chromatin binding of the other factors blotted for appears to be marginally increased upon addition of the Δ24 mutant. Likewise, DONSON itself retains the capacity to bind chromatin even without these N-terminal 24 amino acids. This suggests that the helix is required only for GINS recruitment and all factors that bind prior to DONSON – as established in Figure 3.1.11 – are still able to do so in the absence of this helix.

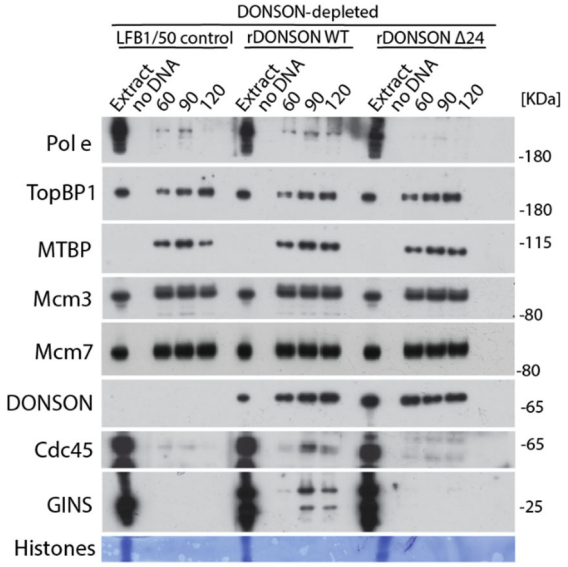


Figure 3.2.8: DONSON's interaction with Sld5 is required for formation of the active helicase. DNA replication was set up in IgG- and DONSON-depleted extract with the optional addition of recombinant WT or Δ24 xDONSON: A) Chromatin was isolated at indicated timepoints and chromatin-bound proteins analysed by immunoblotting. Positive and negative controls were performed as in previous experiments. N=3.

3.2.2.5 Utilising the N-terminal Peptide

Having established that the DONSON N-terminal helix is required for formation of the active helicase, we wanted to assess whether this peptide alone is sufficient to compete with endogenous DONSON and block its initiation function. To determine this, we collaborated with Professor Andrew Wilson's lab, who synthesised two forms of the peptide (amino acids 7-22) one in the linear form (Figure 3.2.9 A) and one supported in the helical form found within the endogenous protein (through lactam bonding) (Figure 3.2.9 B). Using these peptides, I could assess whether their addition into egg extract is sufficient to block replication and active helicase formation. To test DNA replication, I set replication up in the egg extract

and supplemented with the peptides at different final concentration – 0.5, 1, 2 or 5 mM – and measured DNA replication efficiency at the peak of replication (45 minutes) and at the end of replication reaction (90 minutes). As seen in Figure 3.2.9 C and D, addition of either form of the peptide into the extract reduces the efficiency of DNA replication. The peptide in the lactam form (black) has a greater capacity to reduce DNA replication efficiency at all concentrations than the peptide in the linear form (grey). Suggesting that the helical structure of the peptide is important for efficient function of the protein. At early replication (Figure 3.2.9 C), just 2 mM of the linear-form peptide is required to permit only very low levels of replication, whilst by 5 mM addition replication appears to be completely blocked. In contrast, the lactam-form peptide at this timepoint efficiently blocks the majority of replication even at the lowest concentration tested (0.5 mM) – permitting around 15% of the control replication levels (0 mM at 90 min). At the later timepoint, the differences between linear and lactam are more pronounced. While with the linear form ~30% remains upon 5 mM addition, with the lactam peptide 1 mM is sufficient to block ~50% of the control replication levels. Upon 2 mM addition, ~5% of the control replication level remains and with 5 mM, no replication occurs. This suggests that the peptide, particularly in the helical form, is very efficient at blocking DNA replication in *Xenopus* egg extract.

To further assess this novel finding, I tested whether the peptides could block active helicase formation on chromatin. I set up replication in extract as before and optionally supplemented with 2 mM of either peptide. This concentration was chosen based on the efficiency of both peptides to block replication (Figure 3.2.9 C and D). I isolated chromatin at early, mid and late replication and assessed the binding of numerous factors by immunoblotting with their specific antibodies. As seen in Figure 3.2.9 E, both forms of the peptide have an impact on binding of several factors tested for. Mcm7 binding appears to be largely unaltered by the linear form of the peptide. However, when the lactam form is applied, Mcm7 binding and retention on chromatin is increased. This is likely due to replication being unable to proceed following hexamer loading, causing Mcm2-7 stalling on chromatin. This is also the binding pattern with MTBP, a loading factor of Cdc45, suggesting that the peptide, particularly in the lactam form, also blocks replication after MTBP has bound to chromatin. In contrast, the loading of TopBP1, PCNA, Cdc45 and GINS proteins is reduced when the helical peptide is applied. For TopBP1 and PCNA, binding is mildly reduced

and suggests that both proteins have bound to chromatin irrespective of DONSON initiation activity, but may require CMG formation to be stabilised there. Whereas for Cdc45 and GINS, binding appears totally absent when the lactam peptide is applied. This suggests that 2 mM of the helical peptide is sufficient to prevent the formation of the active helicase entirely. However, whilst the binding for most factors appears the same when the linear peptide is applied as in the normal extract, peak binding of Cdc45 and GINS proteins occurs later in replication. This suggests that this form of the peptide is less efficient at blocking helicase formation but can nonetheless inhibit and slow down the initiation reaction. Interestingly, the binding of endogenous DONSON to chromatin in the presence of either form of this peptide is unaffected. Suggesting that whether DONSON binds Sld5 or not, it will be recruited to the chromatin. Collectively, the addition of the peptide appears to mimic what we observe upon DONSON depletion from the extract. We hypothesise that this is due to the peptide inhibiting the interaction between the endogenous DONSON and the GINS complex. However, we have not yet been able to visualise this and it must be directly tested. As this peptide consists of just the very N-terminus of the DONSON protein, then it is unsurprising that both TopBP1 and DONSON can still bind to the chromatin in its presence, as they interact via the DONSON-dimer interface. Likewise, Pol Ɛ loading appears to be unaffected by either form of the peptide. DONSON may be involved in recruitment of Pol Ɛ, utilising a distinct part of the protein that is unaffected by competition with the peptide. Whilst both forms of the peptide show interesting results, the lactam form shows particular promise to prevent formation of the active helicase and block DNA replication initiation. This suggests the peptide could be utilised to prevent replication and cell cycling and therefore requires further investigation.

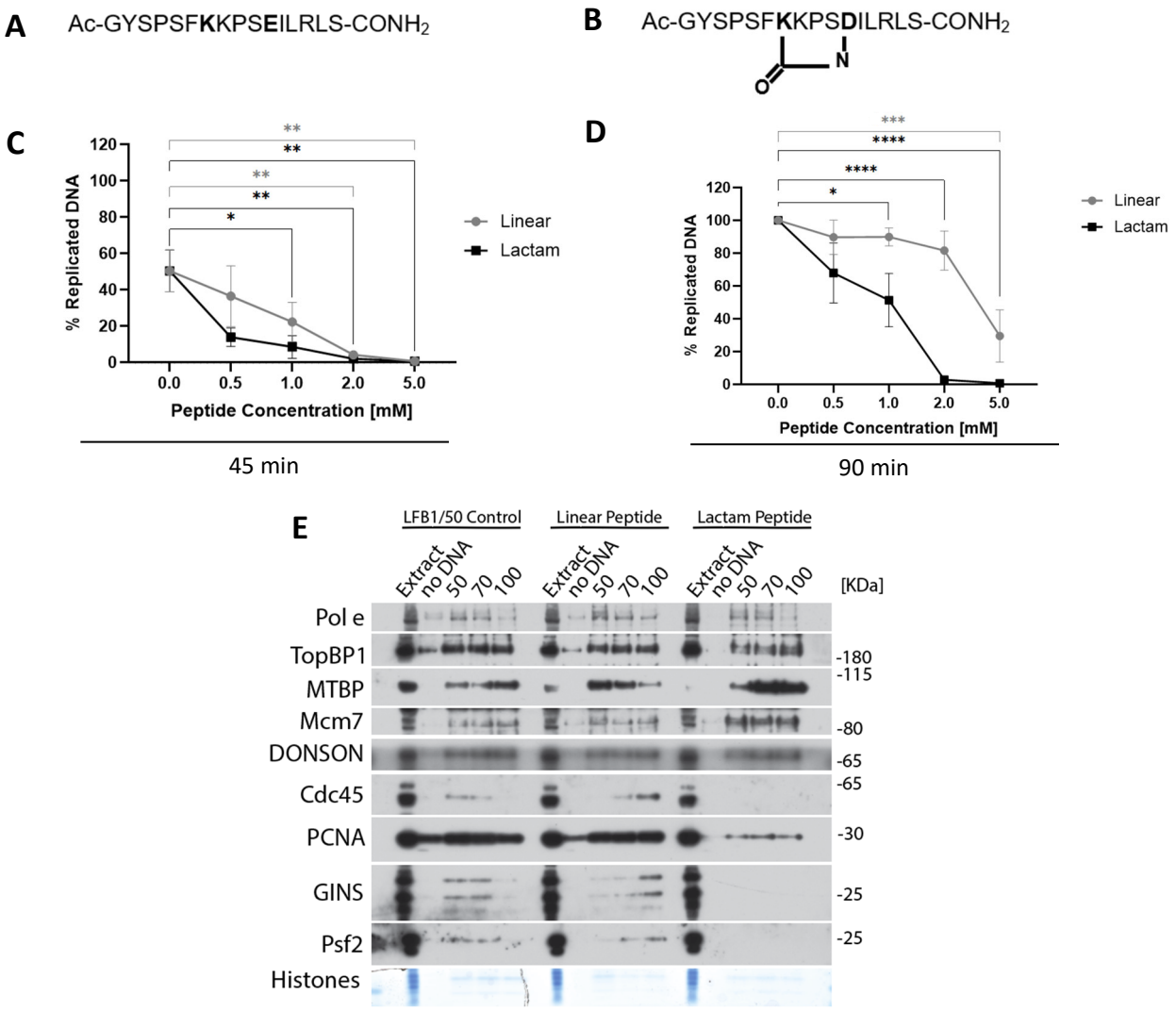


Figure 3.2.9: DONSON's N-terminal helicase is sufficient to block DNA replication. A-B) Dr D. Gimenez-Ibanez synthesised and purified the helix at the N-terminus of the xDONSON protein, in either linear (A) or lactam (B) – supported in the helical structure – form. C-E) DNA replication was set up in extract with the optional addition of either peptide C,D) at 0.5, 1, 2 or 5 mM, and incubated with α- [P32] dATP. The DNA replication reaction was stopped at 45 min (C) or 90 min (D) and DNA synthesis was measured as in previous figures. N = 3, mean with SEM presented; Two-way ANOVA comparing to 0 mM at 90 min timepoint, *p < 0.028, **p < 0.0097, ***p = 0.0011, ****p < 0.0001. E) 2 mM was added and chromatin was isolated at indicated timepoints and chromatin-bound proteins analysed by immunoblotting. Positive and negative controls were performed as in previous experiments. N=4.

3.2.2.6 DONSON-CMG Binding Stabilises the CMG and Causes Rotation to Cue DNA Melting

Having observed key interaction sites between DONSON and the CMG, we also further investigated the overall structure of the DONSON-bound double CMG. Upon inspection of the cryo-EM model structure, we observed that the double hexamer is stabilised by dimeric DONSON. Within the construct, each of the two globular cores of the dimer contacts a single CMG and provides stability to the entire DONSON-bound double-CMG. This suggests a key role for dimerisation of DONSON in stabilisation of the double-CMG once it has recruited GINS to the chromatin. Dr Cvetkovic and Dr Costa compared the Mcm2-7 double

hexamer interface when DONSON is bound to Mcm2-7 double hexameric complexes previously observed [60], [78], [87], [221]–[223]. The most striking observation was that the Mcm2-7 hexamer inter-ring interactions are loosened in the presence of DONSON – the two rings appear to hover over one another, rather than directly interacting. When assessing the C-terminal ATPase side, a ~60° clockwise rotation of one Mcm2-7 ring was also observed, causing the Mcm7 side of one hexamer to swing out, resulting in the two rings becoming misaligned and the two central channels separating. Additionally, with DONSON bound, it appears that Mcm2 overlaps with the position that, in the double hexamer form, duplex DNA occupies. Mcm2 ZnF of one hexamer ring in the DONSON-bound double CMG partially occludes the N-terminal pore of the other ring. Collectively, this suggests that with DONSON bound to the CMG, a rotation of the two Mcm2-7 rings is promoted, misaligning their central pores and likely promoting duplex DNA untwisting/melting. Further details can be found in [190].

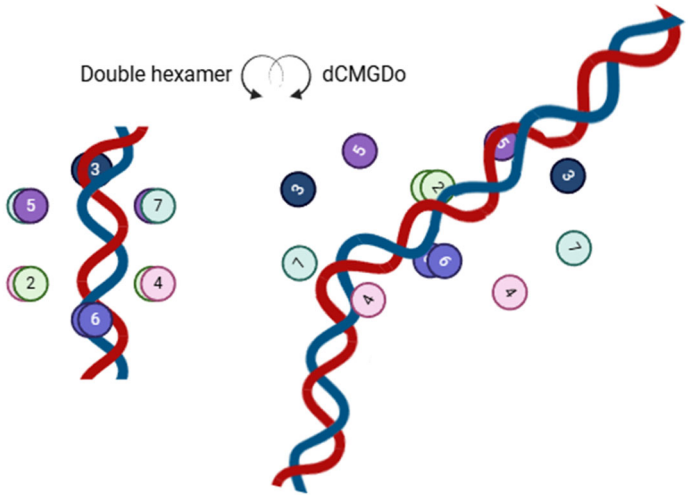


Figure 3.2.10 Schematic of initial DNA melting upon DONSON binding to the CMG. Each MCM subunit is presented overlapped in the double hexamer formation (left) and misaligned (right). As the two hexamers misalign, once DONSON is bound, the DNA becomes kinked due to Mcm2 (green in the centre) occludes the central pore of the two rings. Adapted from [195].

3.2.2.7 DONSON's Dimerisation is Required for DNA Synthesis

As we had observed significant stabilisation of the double CMG structure provided by dimeric DONSON, we wanted to assess the importance of this dimer form. To do this, we needed to create a mutant that does not readily form the dimer. To begin, we identified two MPD patient mutations that map at the dimerisation interface, suggesting that they may lead to disease phenotypes through destabilisation of the dimer. These are W234L and M463T. W234 engages with W234 of the opposite DONSON monomer

to enable dimerisation, and corresponds to W228 in humans which, when mutated to Leu, is dominant negative and has been identified in a patient diagnosed with Femoral-Facial syndrome. M463 maps at the centre of DONSON's hydrophobic pocket and makes contact with the dimerisation surface in *cis*. It corresponds to M446 in humans which, when mutated to Thr, is recessive and causes Primordial Dwarfism in patients. We expressed this mutant (WM) in *E. coli* and purified it using a Nickel column. Samples taken during purification were run in an SDS-PAGE gel and visualised by Coomassie staining (Figure 3.2.11 A). As shown in Figure 3.1.15, using mass photometry we observed that our WT rDONSON preferentially exists in the dimer form. Therefore, using mass photometry, we looked to determine whether our 'dimer' mutant exists primarily in monomer form instead. We assessed the molecular mass through mass photometry of a range of concentrations of the protein – 5, 10 and 20 nM. As seen in Figure 3.2.11 B, a greater concentration of this mutant is required to form the dimer. At 5 nM, the mutant primarily exists in the monomer form; it exists equally between monomer and dimer at 10 nM; whilst at 20 nM it primarily forms the dimer. This suggests that these mutations are sufficient to abrogate efficient dimer formation of DONSON at low DONSON concentrations. Therefore, we could use this mutant to assess whether dimerisation of DONSON is required to support DNA replication. We set up replication in DONSON-depleted extract and optionally incubated with either the WT protein or the WM mutant at different final concentrations – 0, 1.5, 2, 4, 8, 16 nM. We stopped replication in early S phase and assessed DNA replication by measuring incorporation of α-[P32]dATP, as in previous figures. As seen in Figure 3.2.11 D, whilst little replication is observed in the 1.5-4 nM range for either protein, at 8 nM WT DONSON is noticeably more efficient at supporting replication than WM, and this difference is even more striking at 16 nM. This suggests that dimerisation is indeed key for supporting efficient DNA replication. The inability of the W228L and M446T mutants to dimerise may be the reason for the phenotypes observed in patients expressing them.

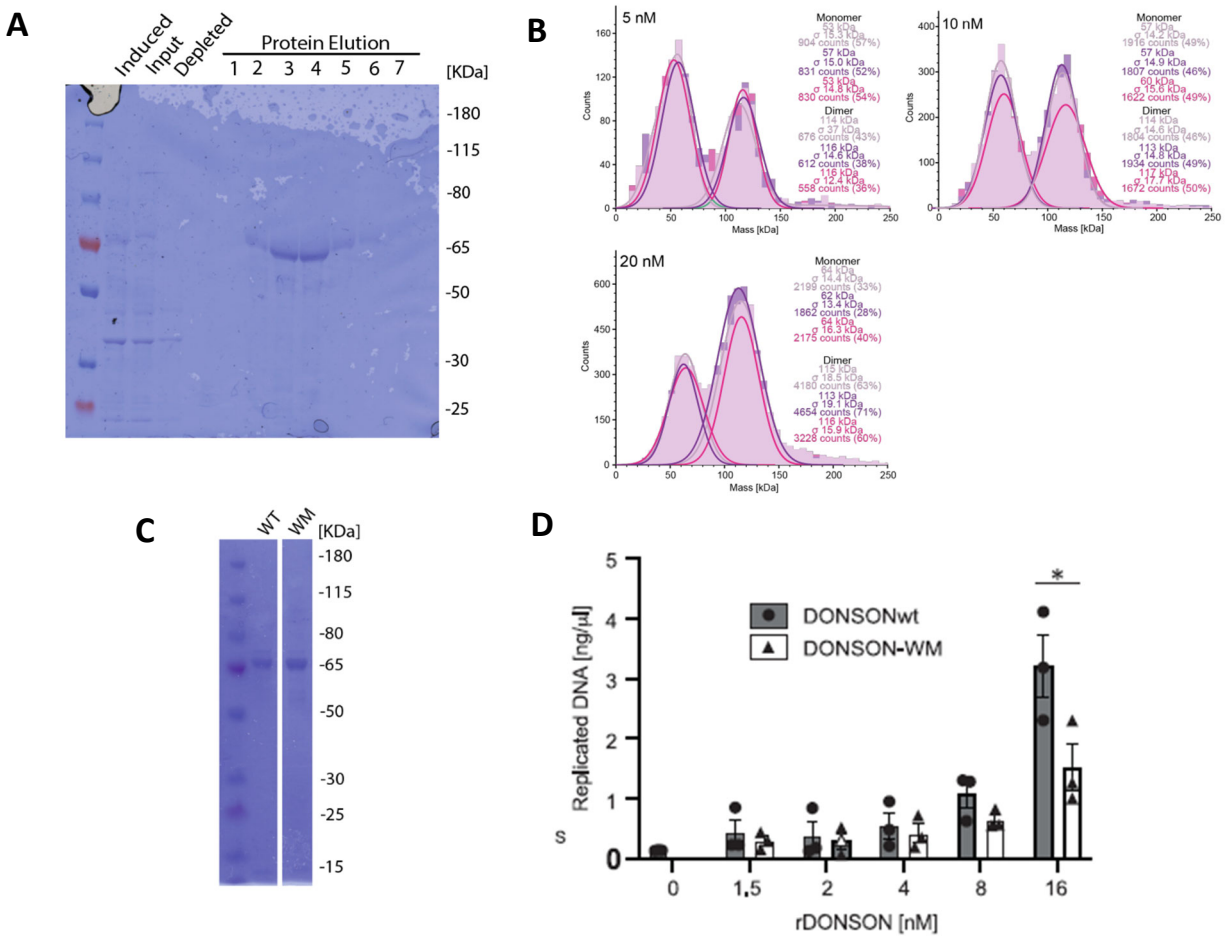


Figure 3.2.11: DONSON's dimerisation is required to support DNA synthesis. A) Purification of His- and MYC-tagged DONSON W234L, M463T (WM). Expression of xDONSON mutants were induced in *E. coli* (Induced) and purified as described in methods section using Nickel affinity purification. Supernatant from bacterial lysate containing soluble proteins (Input), unbound proteins (Depleted) and elutions using 250 mM of Imidazole (Protein elutions 1-7) are shown. Samples from the purification were run on an SDS-PAGE gel and visualised by Coomassie staining. B) Performed by Dr Passaretti. Purified xDONSON WM at dilutions 5, 10 and 20 nM were assessed by mass photometry. The mass range of the compounds analysed is shown at each concentration tested. C) Samples of each protein at the concentration added into extract were run on an SDS-PAGE gel and visualised by Coomassie staining, for comparison of accurate concentration calculation. D) DNA replication was set up in DONSON-depleted extract with the optional addition of recombinant WT or WM xDONSON at dilutions from 1.5 nM to 16 nM, and incubated with α- [P32] dATP. The DNA replication reaction was stopped in early S phase, and DNA

3.2.2.8 Do Known Patient Mutations Impact DONSON's Initiation Role?

Having observed two patient mutations impacting dimerisation and DNA replication efficiency, we set out to determine whether other known patient mutations result in similar phenotypes at the molecular level. To do this, we assessed the AlphaFold predicted structure and observed another known patient mutation, R217C – implicated in Meier Gorlin syndrome – at this interface too, at the groove between the dimer around where TopBP1 is predicted to bind (Figure 3.2.12).

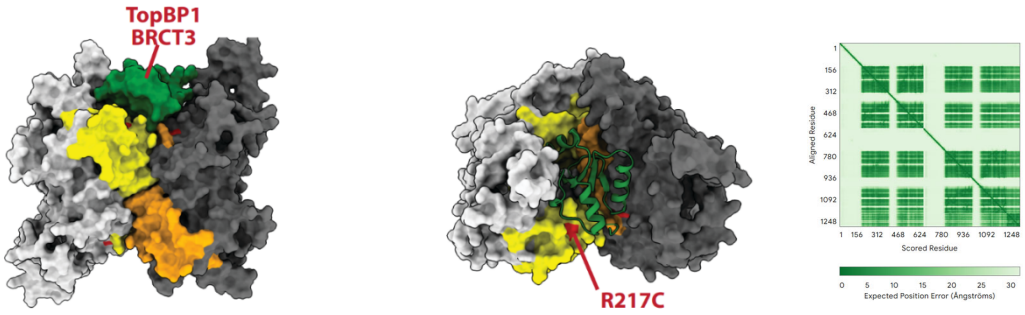


Figure 3.2.12: A known DONSON patient mutation is positioned at the groove between DONSON monomers where TopBP1 is predicted to bind. AlphaFold shows two DONSON molecules forming the dimer. ipTM = 0.65 pTM = 0.65. Corresponding expected position error plot shown. The TopBP1 BRCT3 domain is labelled. DONSON D2 domains are highlighted in yellow for one molecule and orange for the other – these contain known Meier-Gorlin mutations. One mutation site in particular is positioned at the groove between the two DONSON molecules: R217C – highlighted with a red arrow.

Having identified a known Meier-Gorlin mutation and a known Primordial Dwarfism mutation at the dimer interface, we wanted to assess whether these mutants were able to support DNA replication. I additionally wanted to assess other Meier Gorlin mutations, due to initiation factors in particular being mutated in this condition. We chose five different known mutations with sites conserved in *Xenopus* to test in egg extract, two MPD and three MGORS (Table 3.2). However, we could not successfully purify one MPD mutated protein, not included in the table (data not shown).

Table 3.2: DONSON Patient Mutations Sites are Conserved in *Xenopus*. Human sites mutated in the indicated syndrome are presented, with the conserved *Xenopus* site which we mutated to test in egg extract.

Disease	**Human Mutation Site**	***Xenopus* Equivalent**
MPD	M446T	M463T
MGORS	R211C	R217C
MGORS	F165S	F171S
MGORS	P224S	P230S

I mutated DONSON at these sites – R217C, F171S and P230S (MGORS) and M463T (MPD), previously tested in combination with W234L – and expressed them in *E. coli*. We purified the proteins as before and visualised samples by Coomassie staining (Figure 3.2.13 A-D). I then tested replication in IgG- and

DONSON-depleted extracts as before, optionally incubating with the WT protein or these mutants, to assess their ability to support nascent DNA synthesis. As seen in Figure 3.2.13 F, both F171S and P224S support efficient DNA replication, with nascent DNA synthesis levels similar to both the IgG-depleted extract and the WT rescued DONSON-depleted extract. Interestingly, the M463T mutant, which was unable to support efficient replication when in combination with the W234T mutant, supports efficient replication when mutated alone. This suggests that whilst this site is important for DONSON dimerisation, it may work in combination with W234 for DONSON's initiation role. In contrast, the R217C mutant is unable to support efficient DNA replication, with levels peaking at around 50% of the IgG-mock depleted extract. This suggests that the R217 site is specifically required to support efficient DNA replication. Although we cannot conclude that this is due to R217 being required for the DONSON-TopBP1 interaction, its position at the groove between the DONSON molecules suggests that it is likely required to support the interaction.

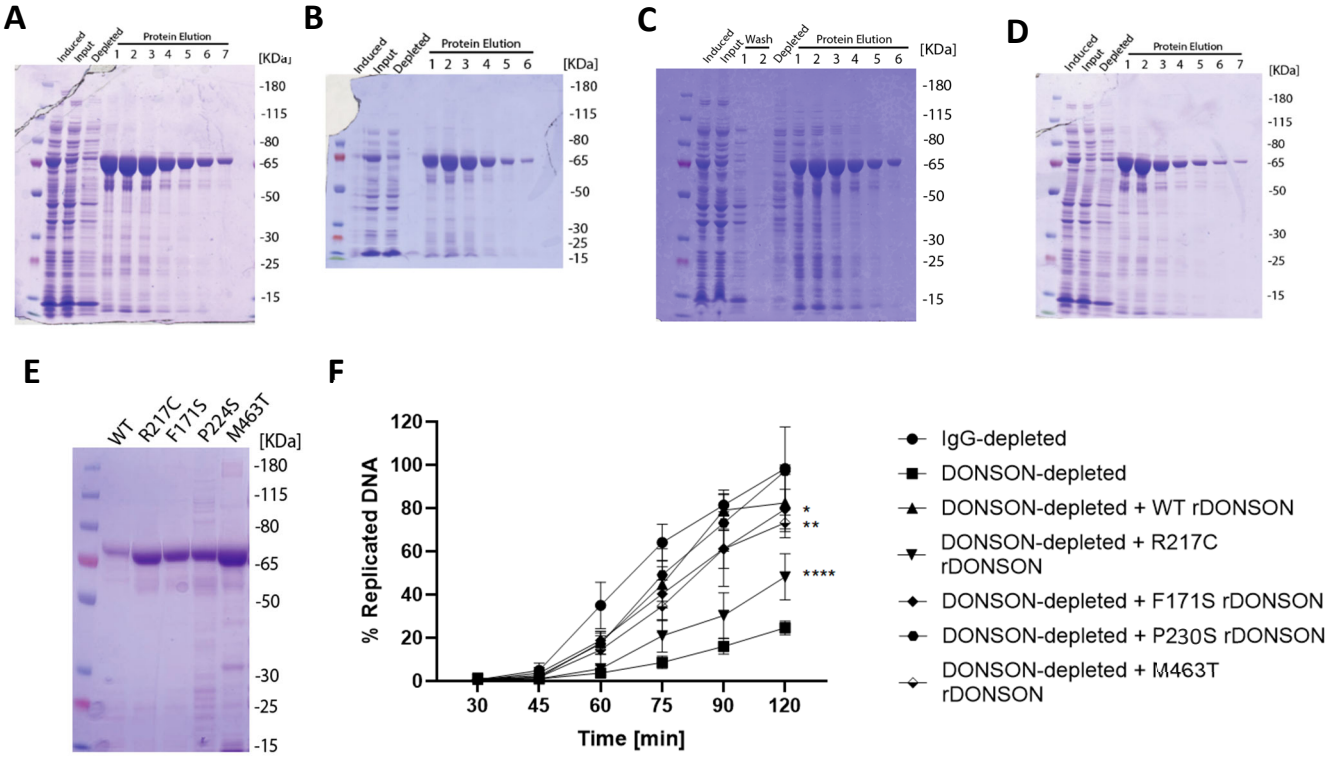


Figure 3.2.13: Select DONSON patient mutations impact DNA replication initiation. A-D) Purification of His- and MYC-tagged DONSON R217C (A), F171S (B), M463T (C), P230S (D) mutants. Expression of xDONSON mutants were induced in *E. coli* (Induced) and purified as described in methods section using Nickel affinity purification. Supernatant from bacterial lysate containing soluble proteins (Input), unbound proteins (Depleted) and elutions using 250 mM of Imidazole (Protein elutions 1-6) are shown. Samples from the purification were run on an SDS-PAGE gel and visualised by Coomassie staining. E) Samples of each protein at the concentration added into extract were run on an SDS-PAGE gel and visualised by Coomassie staining, for comparison of accurate concentration calculation. F) DNA replication was set up in IgG- and DONSON-depleted extract with the optional addition of WT rDONSON or indicated mutants and incubated with α- [P32] dATP. The DNA replication reaction was stopped at indicated timepoints and DNA synthesis was measured as in previous figures. n = 3. Mean and SEM plotted. Two-way ANOVA comparing to IgG-depleted, **** p< 0.0001, ** p = 0.0014 (F171S), * p = 0.0146 (M463T).

Having tested the ability of these mutants to support DNA synthesis, and seeing an impact only upon mutation of R217, I wanted to test whether these mutants may impact other post-initiation functions of DONSON. To determine this, I looked to assess whether there are suggestions of replication stress (γ-H2AX signal on chromatin) or whether chromatin binding of replication-supporting factors is perturbed in the presence of these mutants. I set up replication in IgG- and DONSON-depleted extracts and optionally incubated with WT or the mutant DONSON proteins, as previously. I isolated chromatin at several timepoints and assessed the binding of numerous factors by immunoblotting with specific antibodies, with particular interest in factors associated with DNA damage and fork instability – p97 and γ-H2AX. As seen in Figure 3.2.14, none of the mutants tested appear to have an impact upon formation of the active helicase except R217C, which agrees with the reduced level of synthesised DNA (poor α-GINS antibody blotting should be noted in the M463T sample). We observed a delay in chromatin binding of Cdc45 and GINS when R217C is applied, compared to both the IgG-depleted and the WT rescued DONSON-depleted extract. This suggests that the mutant is unable to efficiently support formation of the active helicase and, given its position at the groove between the molecules, this may be due to insufficient binding to TopBP1. Likewise, Mcm7 binding appears unaffected by any mutants tested other than R217C, where it is increased, similarly suggesting that due to the active helicase not forming, Mcm2-7 hexamers are retained on chromatin. TopBP1 and p97 chromatin binding levels also appear to be unaffected by any of the Meier-Gorlin mutants tested, and p97 binding is very similar for M463T Primordial Dwarfism mutant as for the WT. Perhaps most intriguing, the levels of γ-H2AX – indicative of DNA damage and genome instability – do not differ greatly between any of the mutants tested and the WT. Importantly, low level DNA damage signal can consistently be observed during replication in egg extract [224]. Although the level of γ-H2AX in the IgG-depleted sample is less than in the DONSON-depleted sample in Figure 3.2.14 A, this may be the result of inconsistent protein transfer or immunoblotting. Upon DONSON depletion, the level of DNA damage does not markedly increase, likely due to replication not occurring in the absence of DONSON.

Using this as our base level of damage, none of the mutants appear to alter the levels of damage signalling. This requires further investigation.

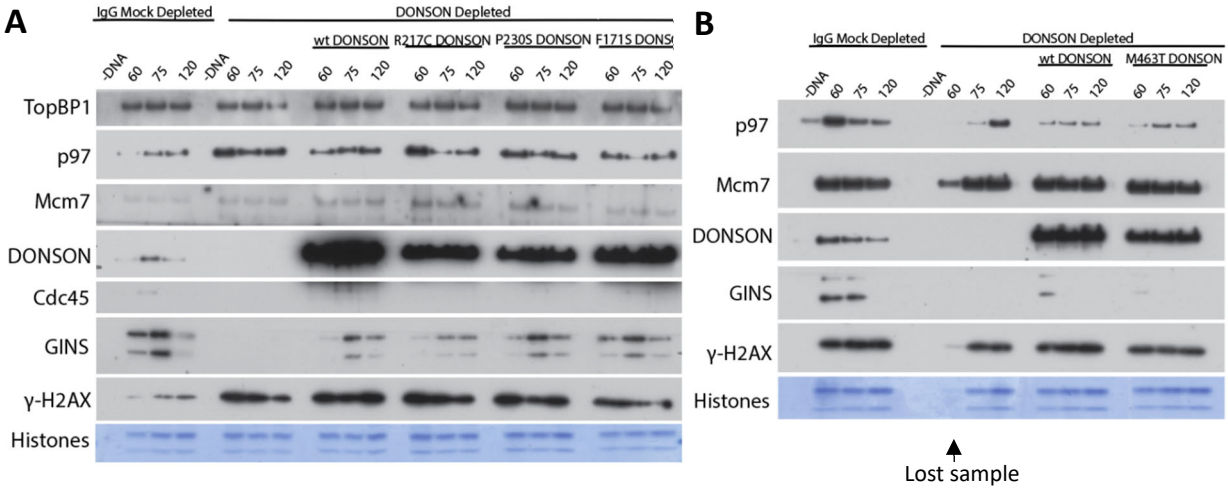


Figure 3.2.14: Select DONSON patient mutations impact fork stability. DNA replication was set up in IgG- and DONSON-depleted extract with the optional addition of recombinant WT or R217C, P230S, F171S Meier Gorlin patient mutants (A) or M463T primordial dwarfism patient mutant (B). Chromatin was isolated at indicated timepoints and chromatin-bound proteins analysed by immunoblotting. Positive and negative controls were performed as in previous experiments.

3.2.3 Discussion

This chapter of my thesis has focussed on the interactions of DONSON in the context of the replisome on chromatin, specifically identifying the role DONSON has in establishing the CMG at origins. The work from the previous chapter showed that DONSON is responsible for recruitment of GINS to the Mcm2-7 hexamers and stabilisation of Cdc45 here. In this chapter we discovered the positions to which DONSON binds within CMG factors on chromatin, identifying interactions of most importance to its role, and utilising a short peptide to block DNA replication and CMG formation at origins. In addition, we identified the importance of dimerisation for DONSON's initiation role, and tested a number of different disease mutants with molecular phenotypes that cannot support replication initiation.

Prior to an identification of DONSON's molecular structure or direct interactions in any organism, a structure was predicted through the AlphaFold programme. Through this, a folded globular core encircled by an unstructured N-terminus and joined via a protruding flexible loop could be identified, which we predicted were likely important for stabilising DONSON's interactions with other proteins [157], [225]. Having established that the N-terminal and protruding flexible loop regions of the protein are important to support DNA replication (Figure 3.2.1 D), we set out to determine how DONSON directly interacts with components of the CMG and establish how it supports formation of the active helicase. Using AlphaFold-Multimer as a predictive tool for DONSON interaction partners, the Walter lab predicted that DONSON interacts with SLD5, TopBP1, POLE2 (the catalytic subunit of Pol Ɛ) and MCM3, and that these interactions are conserved across various species [194]. Structural analysis, Cryo-EM in particular, shed light on the specific interaction sites between DONSON and the CMG, likewise identifying interactions with Mcm3 and Sld5 [190]. By Cryo-EM, the Labib lab also identified interactions between *C. elegans* DNSN-1 and the GINS subunits SLD-5, PSF-2 and PSF-3, as well as with MCM-3 [196]. Collectively, this made our structural investigation all the more interesting.

Interestingly, the overarching observation from our cryo-EM data is that dimeric DONSON interacts with both CMGs at each active origin, specifically interacting with Mcm3 and Sld5. The two regions of Mcm3 we identified DONSON interacting with within the CMG are the ATPase domain, within the core of the

protein, and the A-domain at the N-terminus (Figure 3.2.4). Interestingly, Mcm3 has been identified as critical for recruitment of the Mcm2-7 hexamer to ORC [226]. Likewise, the ATPase domain of the protein, whilst being required for translocation of the complex, is also involved in the interaction with other MCMs within the hexamer, with Psf3 also interacting with the hexamer at the Mcm3-5 interface strengthening this junction [227]. Additionally, the A-domain of MCM3 has been shown to interact with the N-terminus of MCM7 in human cells [78]. Collectively, this suggests that DONSON's association with Mcm3 enables it to efficiently access other components of the core of the helicase, providing an ideal platform to deposit GINS. Like our data, using mass photometry the Walter lab also identified DONSON as readily homodimerising, and as a result predict that dimerised DONSON brings two copies of each replisome component to the chromatin. However, they did not visualise DONSON's interactions with the replisome components in dimer form. Rather, they used AlphaFold to predict interactions between monomeric DONSON and a single CMG. In this way, they were able to predict an interaction between DONSON's flexible loop and the AAA+ domain of MCM3. This fits with our cryo-EM data which indicates a particularly well-resolved interaction between DONSON's Trp381 in the α-helix of its flexible loop and Lys285 in the ATPase domain of Mcm3. We also used AlphaFold to identify two further sites within this domain (Figure 3.2.4 B). In addition to interactions at the AAA+ domain, our cryo-EM data indicated further interactions between DONSON's globular core and the A-domain of Mcm3, which were not identified in the AlphaFold predicted structure by the Walter lab. They subsequently mutated their predicted interaction sites in the α-helix of DONSON's flexible loop and performed co-IPs with the predicted interactors, observing that all interactions were still viable, except with MCM3, which resulted in the other factors being unable to bind to the chromatin. Likewise, upon mutation of our identified Mcm3-interacting sites in DONSON, we observed a decrease in the efficiency of replication, suggesting that DONSON's interaction with Mcm3 is essential for GINS recruitment to the chromatin and resultant formation of the CMG.

Our cryo-EM data also identified an interaction between the N-terminus of DONSON, residues 6-22, and the A-domain (α-helical) at the N-terminus of Sld5 (Figure 3.2.6). Whilst our AlphaFold modelling predicted an interaction between DONSON's globular domain and the B-domain (β-strand) at the C-

terminus of Sld5. The α-domain of Sld5 is primarily involved in helix–helix interactions with adjacent GINS subunits via their α-domains. Whereas, the B-domain of Sld5 is required to stabilise the entire GINS complex, essential for replication activity [228], [229]. Using AlphaFold, the Walter lab also predicted that the N-terminus of DONSON binds to SLD5. When they modelled the interactions with all components of the GINS complex, they also observed that this interaction with SLD5 extends to DONSON's globular domain. Upon mutation of DONSON's predicted interaction sites, they not only found that GINS no longer co-immunoprecipitated DONSON, but that replication was no longer supported either. We likewise observed the same (Figure 3.2.7 D). Upon truncation of the N-terminal 24 amino acids, which includes the sixteen we had observed binding to Sld5, we observed that DNA replication was no longer supported in the extract. However, mutation of the sites predicted to interact with the B-domain of Sld5 did not abrogate the ability of the extract to support replication. Collectively, this indicates that the DONSON dimer is responsible for recruiting GINS to the hexamer via Mcm3, and that these interactions are essential for CMG formation and resulting DNA replication. What they also suggest, is that *in silico* modelling is a very powerful and largely accurate tool for predicting interactions between complex factors.

As mentioned above, the Labib lab also used cryo-EM to determine the interactions of DONSON, here called DNSN-1, within the *C. elegans* replisome [196]. They also observed DNSN-1 acting in dimer form, however interestingly they observed the dimeric DONSON interacting with just one CMG. This is likely due to the use of a short DNA substrate. Within this dimer, they identified that the promoter of just one monomer interacts with the CMG. This led them to suggest that, as we observed in *Xenopus*, that the other monomer may interact with the other Mcm2-7 at the origin, simultaneously forming two CMGs. As previously mentioned, whilst they observed some of the same interactions – DONSON binding to SLD-5 and MCM-3 – they also observed additional interactions between DONSON and PSF-2 and -3. The DNSN-1 dimer forms a large interface with the N-terminus of the MCM2-7 hexamer, with one α-helix of DONSON interacting with SLD-5 and PSF-2, and another with the AAA+ domain of MCM-3. This is particularly interesting as it suggests that DONSON directly interacts with more components of the GINS complex in *C. elegans* than we have observed in *Xenopus*. This may indicate that DONSON interacts with multiple GINS

subunits prior to chromatin binding, for stable recruitment to the hexamer. As we had done, the Labib group truncated DNSN-1 at its N-terminus, to prevent interaction with GINS, and had observed no DNA replication initiation. Mutation of DNSN-1's MCM-3 interacting sites likewise abrogated formation of the CMG on chromatin. Supporting the idea that the requirement for DONSON's interactions with Sld5 and Mcm3 is conserved throughout species.

Related to DONSON's role in supporting replication initiation, in this chapter I have also presented our findings that DONSON is required for the initial melting of DNA at the end of the initiation process. By comparing the cryo-EM structure obtained in our work to the previously determined structures of the double hexamer, it was observed that upon dimeric DONSON binding to the double hexamer, the interaction between the rings is loosened, causing the duplex DNA within this channel to become exposed and untwisted [190]. Interestingly, the Walter lab also predict a rotation of the MCM double hexamers in relation to one another upon binding of dimeric DONSON. However, by assessing their predicted structure, they propose that clockwise rotation of the two MCMs is actually required to enable dimeric DONSON to bind both MCM3 monomers in the double helix.

Whilst our cryo-EM data implicates DONSON in the initial melting of the duplex DNA within the Mcm2-7 central channel, it is Mcm10 that is responsible for separating the double CMGs to eject the lagging strand from the central channel and promote DNA unwinding [230]. Interestingly, Mcm10 binding in yeast has been observed to occur at the N-terminus of the hexamer, at the homo-dimerisation interface of the double hexamers. Specifically, Mcm10 appears to bind to Mcm2 [230]. Intriguingly, in our cryo-EM data, upon DONSON binding the position of Mcm2 is altered so that it partially occludes the N-terminal channel of the partner MCM ring, where the duplex DNA previously occupied. Given that movement of Mcm2 over the central channel of the hexamer can only occur following duplex DNA removal, this may suggest that DONSON can perform initial DNA melting, or that its binding aids the subsequent action of Mcm10 in origin melting. Additionally, Sld3/Treslin has been observed to bind to Mcm2 too [231]. This suggests that initiation factors required for the loading of Cdc45 and GINS may be removed from the CMG prior to, or as the result of, Mcm10 binding. In yeast, ATP hydrolysis at this stage is also coupled with the release of

several initiation factors, including Dpb11, Sld3-7 and Sld2 from the DNA [86]. Our data indicating that DONSON is required to stabilise the double CMG suggests that, like its yeast orthologue, it may need to be removed prior to Mcm10 binding to allow for double CMG separation. Mcm10 stimulates the hydrolysis of the ATP bound to each hexamer following initial DNA untwisting, to enable to ATP hydrolysis DNA translocation function of the CMG. Intriguingly, the Labib group observed that RNAi depletion of DNSN-1 is synthetic lethal with knockdown of MCM-10 [196]. This suggests that in the absence of MCM-10, DNSN-1 may contribute to the helicase activation step in a redundant pathway to MCM-10. Likewise, in *Xenopus*, Mcm10 is not essential for replication initiation, perhaps because DONSON may be able to support DNA melting in its absence [187].

The removal of Cdc45 and GINS-loading factors from chromatin to enable DNA unwinding has also been proposed by the Labib lab. In particular, they observed that binding of DONSON with GINS and MCM-3 appears to be incompatible with binding of Pol α to the complex [196]. This indicates that the two factors cannot bind at the same time and therefore suggests that DONSON is removed from the CMG prior to DNA unwinding and replication. Interestingly, Mcm10 is required for binding of Pol α to the DNA and suggests that DONSON removal and Mcm10-Pol α recruitment is sequential. Interestingly, the Chistol group used single molecule analysis to assess the timing of several factors binding to the DNA. Here, they observed that RecQL4 is required for DONSON's dissociation from the DNA [197]. Interestingly, RecQL4 has been observed to be required for binding of Mcm10 to the CMG [232]. The Chistol group also observed that DONSON disengages the DNA promptly after GINS has been recruited. Upon RecQL4 depletion from *Xenopus* egg extract, they observed a 2.5-fold increase in the delay between GINS recruitment and origin firing, suggesting that RecQL4 is required for efficient DNA unwinding and replication. Likewise, depletion of RecQL4 led to a 3-fold increase in the time DONSON was bound to chromatin. Collectively, this suggests that DONSON is removed prior to Mcm10-CMG binding and this is likely regulated by RecQL4, which enables the binding of both Mcm10 and Pol α to activate the replisome for DNA unwinding and replication. Of note, as mentioned above, the binding of DONSON to the hexamers causes an adjustment to their position in relation to one another, misaligning their channels. As the two CMGs must pass one another as they begin to unwind the duplex DNA, it may be that this

misaligned position following DONSON binding allows them to immediately pass one another and unwind the DNA upon DONSON's dissociation.

As well as the structural significance of DONSON binding to the CMG, in this chapter I have proposed the potential for DONSON-related factors to be utilised to downregulate DNA replication. This is perhaps the most exciting finding presented within this chapter, laying the foundation for potential therapeutic uses. The peptide was designed based on the importance of the N-terminal region for interacting with Sld5 (Figure 3.2.6 B) to support replication (Figure 3.2.7 D) and CMG formation of chromatin (Figure 3.2.8 A). Using the purified peptide, I have shown in this chapter that the lactam version in particular has the capacity to abrogate CMG formation on the chromatin (Figure 3.2.9 E) and as a result perturb efficient DNA replication (Figure 3.2.9 C and D). The principle of peptides blocking factors required to support replication is extensively adopted in numerous chemotherapies [233]–[235], particularly as peptides tend to have a lower toxicity than chemical compounds. One key way in which therapeutic peptides work is through competing with the endogenous substrate of a protein factor involved in cancer cell survival [232]. This is how we believe our N-terminal peptide is functioning: competing with endogenous DONSON to bind Sld5. This could therefore be translated into downregulating DNA replication and therefore cell cycle progression in cancer cells. However, as DONSON is essential for DNA replication, it is important to determine how the peptide could differentiate between cancer cells and healthy cells, as well as the correct concentration of peptide to use, to minimise the impact on healthy cells. Likewise, it is key to determine whether DONSON's interaction with Sld5 is required for its role in the ATR stress-response as, as observed in human cells, downregulation of DONSON is implicated in DNA damage and abnormal chromatin structures that could themselves be problematic [152].

Within this chapter, I also tried to determine whether DONSON appears to have dual function during replication, during initiation and as a fork protection factor. One key way to determine this, was to identify a separation of function mutant, that successfully supports replication initiation, but cannot maintain fork stability. Figure 3.2.13 F, shows the impact that a number of known MPD patient mutants have at the molecular level on DNA replication efficiency. Testing these mutants in this way allowed us to assess

whether characteristics of the syndrome observed were in fact the result of inefficient replication. We observed that just one mutant tested, R217C (MGORS), was defective in supporting efficient replication. To assess whether the remaining mutants may be unable to stabilise the replication fork later in replication we assessed markers of DNA damage. As seen in Figure 3.2.14, the levels of γ-H2AX on chromatin in the presence of these mutants is comparable to the levels in both the IgG-depleted extract and in the presence of the WT protein. This indicates that no increase in DNA damage is observed as the result of these DONSON mutations. However, a number of patients express these mutant forms of DONSON, and therefore we would expect to see a defect at the molecular level – either due to low expression or due to faulty function of the protein [152], [156], [157], [236]. The way DONSON was first implicated in MPD was using siRNA to downregulate its expression in human cells. siRNA can take hours to successfully downregulate the target protein. Therefore, with sufficient DONSON levels prior to complete downregulation, DNA replication initiation could be supported, however later stages of replication and the cell cycle would be impacted. Therefore, it may be that we have not assessed sufficient cell cycles to observe the impact of these mutants at a later stage of replication or cell cycle regulation. Likewise, it would be interesting to assess the levels of 53BP1 on chromatin, to determine whether more advanced DNA damage can be observed with these mutants. However, to further study DONSON's role in fork stability we could also adopt the previously used siRNA mechanism or use a specific DONSON degron cell line to accurately degrade DONSON at a specific desired time within the cell cycle.

With the indication that DONSON supports replication initiation and is promptly removed from the chromatin, we questioned the cause of the phenotypes observed in MPD patients. The question of whether the small brain and body size observed in MPD patients results from a reduction in DNA replication rather than reduced fork stabilisation and ATR activation, was of particular interest. IP during S phase in human cells has shown that DONSON indeed interacts with Timeless and Tipin [152]. Likewise, in mouse cells a dual role for DONSON, during initiation as well as in fork progression and stability and checkpoint signalling, has been observed [194]. In a mouse embryonic model homozygous for a known patient MPD mutation, the mutation expressed caused a decrease in DONSON protein levels. At the

molecular level, fork asymmetry was increased and ATR signalling was reduced. Interestingly, in the same cells the chromatin loading of both Cdc45 and GINS is impaired, indicating that although ATR signalling is decreased in the absence of DONSON, uncontrolled origin firing does not occur, but instead normal replication initiation is also decreased. As we have shown that TopBP1 is required for the recruitment of DONSON to chromatin (Figure 3.1.14 C), and that upon the depletion of DONSON from the extract, TopBP1 stalls on the chromatin (Figure 3.1.11), it may be that in cases where DONSON mutations cause decreased protein levels, that TopBP1 is recruited to the chromatin, but stalls there as replication cannot proceed efficiently. TopBP1 is required to further active ATR following RPA activation upon replication stress [101]. Therefore if, in the absence of DONSON, TopBP1 is stalled on the chromatin it would not be able to efficiently activate ATR to promote fork stabilisation and DNA repair, leading to the aberrant chromatin structures, fork stalling and instability phenotypes observed upon DONOSN downregulation. Future work to determine whether DONSON indeed has a dual role during DNA replication is required to determine the extent to which DONSON is required to support the ATR stress response, and whether MPD phenotypes resulting from DONSON mutation are instead the result of inefficient DNA replication.

Altogether, this chapter has shown that DONSON directly interacts with Sld5 of the GINS complex and Mcm3 of the Mcm2-7 hexamer, indicating that DONSON does indeed recruit GINS to the hexamer to support the formation of the CMG. In particular, the N-terminal helix of DONSON is essential to support CMG formation and replication and can be utilised as a peptide to prevent DNA replication in *Xenopus* egg extract. Whilst we have been unable to characterise a separation of function mutant, to assess whether DONSON is required for both replication initiation and the ATR stress response, we have identified that two MPD patient mutations likely cause the syndrome due to them being unable to support dimerisation of the protein, required for its initiation function. Having established DONSON's key role in initiation, the regulation of its function will be discussed in the next chapter.

3.3 Regulation of DONSON's Initiation Function through Phosphorylation

3.3.1 Introduction

As presented in previous chapters, DONSON recruits GINS to chromatin to form and stabilise the CMG in a TopBP1-dependent manner. Other groups propose that DONSON, like its orthologue Sld2, forms a 'pre-loading complex' (pre-LC) with GINS, TopBP1 and Pol Ɛ in *Xenopus* [194], [195], and that phosphorylation is not essential for this complex formation. In this chapter we investigate the existence and regulation of formation of such a complex.

In budding yeast, formation of the pre-LC begins with phosphorylation of Sld2. Upon the G1-S phase transition, Cyclin-dependent kinase (CDK) activity increases to enable origin firing only after G1 phase. CDK phosphorylation is required for helicase activation, with Sld2 a key CDK substrate necessary for this to occur. Sld2 has eleven clustered CDK phosphorylation motifs [237]. Phosphorylation of ten of these sites results in a conformational change of the protein to reveal Threonine 84 (Thr84), phosphorylation of which then enables interaction with Dpb11 [68]. Whilst it has been proposed that CDK phosphorylation of DONSON is not required for its initiation role [194] we were interested in assessing whether it is in fact this PTM that determines whether the pre-LC will stably form and for DONSON to carry out its initiation role. Therefore, the aim of this chapter was to assess whether CDK phosphorylation of DONSON is required for its role during replication initiation. Of particular interest, is the extent to which phosphorylation enables interaction with TopBP1 as the potential foundation of a 'vertebrate pre-LC-like' complex formation.

3.3.2 Results

3.3.2.1 DONSON's Pre-Loading Complex Forms only Transiently

Having identified interactions between DONSON and Sld5 and Mcm3, confirming its Sld2-like role in *Xenopus*, we sought to determine whether DONSON also forms a pre-LC as a pre-requisite for active helicase formation. Our previous data suggested that although DONSON can be observed interacting with TopBP1 in the nucleoplasm as well as on chromatin, interactions between DONSON and GINS only occur stably on chromatin (Figure 3.1.5 B, Figure 3.1.13 A, B, D) and therefore may form only transiently outside of this context under particular conditions. In addition to this, immunodepletion of DONSON from egg extract does not even slightly co-immunodeplete either GINS or TopBP1 (data not shown). Collectively, this is contrary to what is proposed by other labs working with *Xenopus* egg extract [194], [195]. We therefore looked to confirm these preliminary findings. I began by assessing interactions between DONSON, TopBP1 and GINS in cytoplasmic egg extract. With this aim, I performed an IP of DONSON from the egg extract, using both rabbit and sheep DONSON-specific antibodies (Figure 3.3.1 A). Although both IPs produced a strong band for DONSON, we observed no clear co-immunoprecipitation with either TopBP1 or GINS. Whilst there does appear to be a possible increased band for TopBP1 co-IP with the DONSON sheep antibody compared to IgG, there isn't with the rabbit. As recombinant DONSON had been supplemented into the egg extract and immunoprecipitated by groups reporting DONSON's involvement in pre-LC formation [194], [195], we assessed whether we could observe these interactions with the recombinant protein instead of the endogenous. I began by using α-MYC beads – our recombinant protein is MYC-tagged. However, I observed extensive non-specific immunoprecipitation of many proteins, even in the absence of DONSON (data not shown). Therefore, using the same sheep antibody as in Figure 3.3.1 A, I performed DONSON IPs in DONSON-depleted extract with the optional addition of WT or Δ24 rDONSON (Figure 3.3.1 B). I used our Δ24 mutant as a control for the specificity of GINS in this reaction, as it should disrupt the interaction with GINS. Unsurprisingly, we observed no co-immunoprecipitation with Mcm7 from the cytoplasm, which acts as a negative control. GINS did not coimmunoprecipitate with either the Δ24 mutant or WT DONSON. With the WT IP, a faint band indicating TopBP1 co-IP can be observed, which is absent upon IP of Δ24. This suggests that DONSON and TopBP1 may be able to

transiently interact in the cytoplasm – however this requires further investigation. This supports our previous finding that of all replication factors we tested, DONSON interacts only with TopBP1 outside of the chromatin (Figure 3.1.13 D). This may also suggest a difference in binding between the endogenous and the recombinant forms of DONSON. Where DONSON has been reported to stably form the pre-LC, recombinant DONSON has been used and either studied in the nuclear fraction of *Xenopus* egg extract [194], or at a much higher concentration (ranging from 60 nM – 8.8 µM) than we have used and observed DONSON at in the endogenous form (~40 nM) [195]. This may therefore explain the differences that we see.

To investigate further whether these proteins likely form a pre-loading complex in extract we assessed the extract by gel filtration. Interphase cytoplasmic (Figure 3.3.1 C) and nucleoplasmic (Figure 3.3.1 D) extract were resolved through a Superose 6 gel filtration column, to visualise potential protein complexes. In both cytoplasmic and nucleoplasmic extract GINS runs at a much smaller molecular weight than either TopBP1 or DONSON. The GINS complex has a molecular weight of 98 kDa and is globular in shape. As we observe Psf2 running around this size and not spread across many fractions, it appears that in both cytoplasmic and nucleoplasmic extract it runs as a tetramer of just the four GINS subunits [238]. As there is no overlap with the larger sizes of DONSON, this suggests that GINS does not complex with DONSON in either the cytoplasm or the nucleus.

As seen in Figure 3.3.1 D, in nucleoplasm TopBP1 largely runs in the same fractions as DONSON, with both proteins appearing strongly in fractions 13-18. As seen in Figure 3.3.1 C, in cytoplasmic extract, the bands for TopBP1 are far less clear and appear to be more spread, likely due to TopBP1 being less concentrated and more difficult to detect in the cytoplasmic fraction. However, both proteins also appear in the same spread of fractions as one another, namely 9-18. Of particular interest was the molecular weight at which both proteins were running – far higher than the individual weights (DONSON at 65 kDa and TopBP1 at 179 kDa) or combined. In addition to this, the spread of both proteins is wide. More flexible structures will be eluted from the column in a wider range of fractions than globular. Therefore, this indicates that TopBP1 and DONSON may form a large and flexible complex, perhaps with other replication factors not

blotted for, in the extract. In contrast, DONSON doesn't appear to complex with GINS in the extract and any stable interaction appears to be confined to the chromatin.

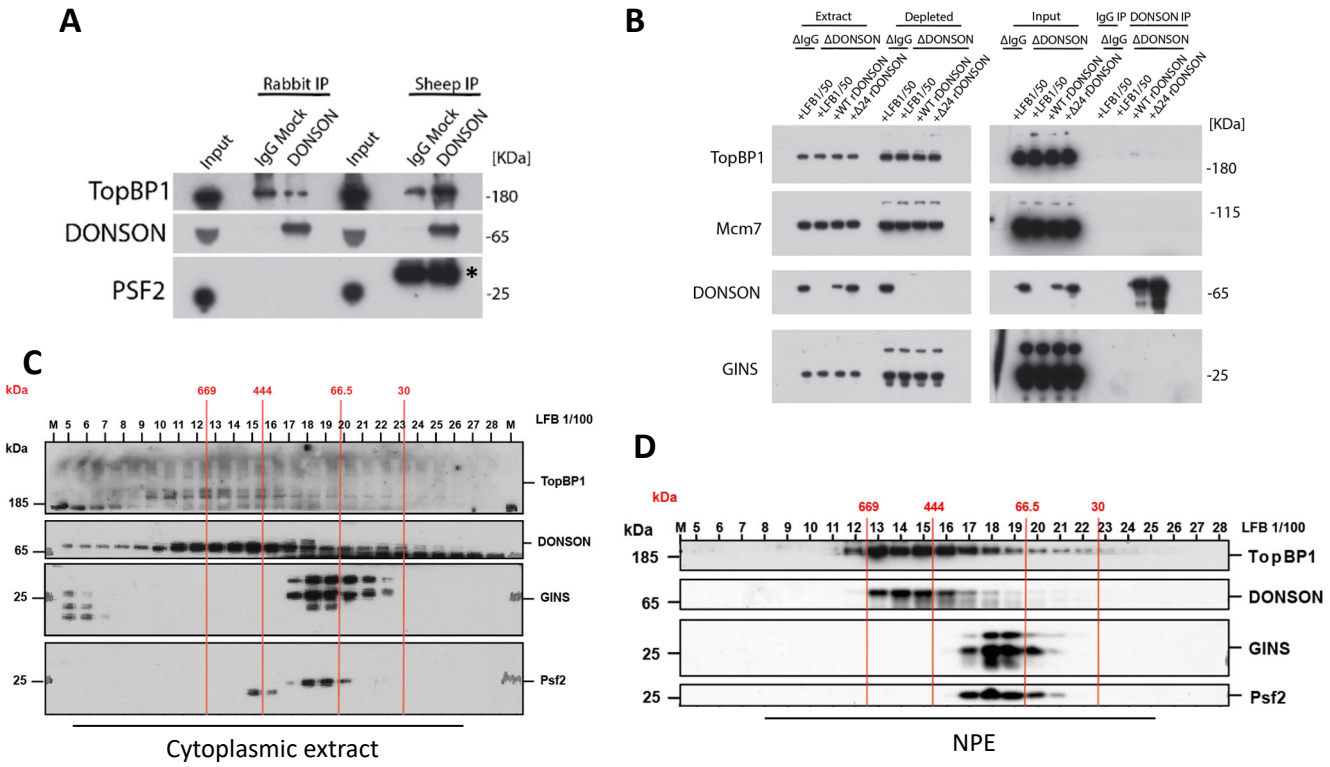


Figure 3.3.1 Interaction between DONSON and other pre-landing complex components is transient. A) IP of endogenous DONSON from whole egg extract was performed and TopBP1 and GINS were blotted for using specific antibodies. Extract without antibodies (Input) and rabbit and sheep IgG antibodies were used as controls. * shows the IgG non-specific light band of the IgG and α-DONSON sheep antibodies, which have been identified by the PSF2 antibody. B) WT or Δ24 rDONSON were optionally added to DONSON-depleted extract and IP of DONSON was performed with the same sheep antibodies as in A. Other factors were blotted for using specific antibodies. Extract from each condition without antibodies (Input) and IgG-depleted extract were used as controls. C-D) Performed by Dr P. Passaretti. Red numbers at the top of each blot indicate molecular weight markers for the column. Black numbers at the top indicate the gel filtration fractions run, marker = M. C) Interphase egg extract was run through a Superose 6 column to separate proteins by size and shape. Fractions were run on an SDS-PAGE gel and blotted for indicated antibodies. D) NPE was run through a Superose 6 column to separate proteins by size and shape. Fractions were run and indicated proteins blotted for.

Strikingly, comparing the separation through the gel filtration of endogenous DONSON in extract and rDONSON in buffer (Figure 3.3.2 A), we observed a large difference between the size at which each protein runs. Whilst the endogenous DONSON primarily runs at >400kDa, as observed previously, the recombinant protein appears in fractions of a much lower molecular weight, ~250-60 kDa, that more closely correspond to the dimeric form of DONSON. This indicates that the endogenous protein does exist within a much larger, more flexible complex. Additionally, it may have post-translational modifications not present on the recombinant. Interestingly, when applied to extract, the size at which the recombinant protein runs increases to the molecular weight at which we observe the endogenous protein in the

cytoplasmic extract (Figure 3.3.2 B). This suggests that upon addition to the extract, the recombinant protein is incorporated into a complex the same as the endogenous exists in.

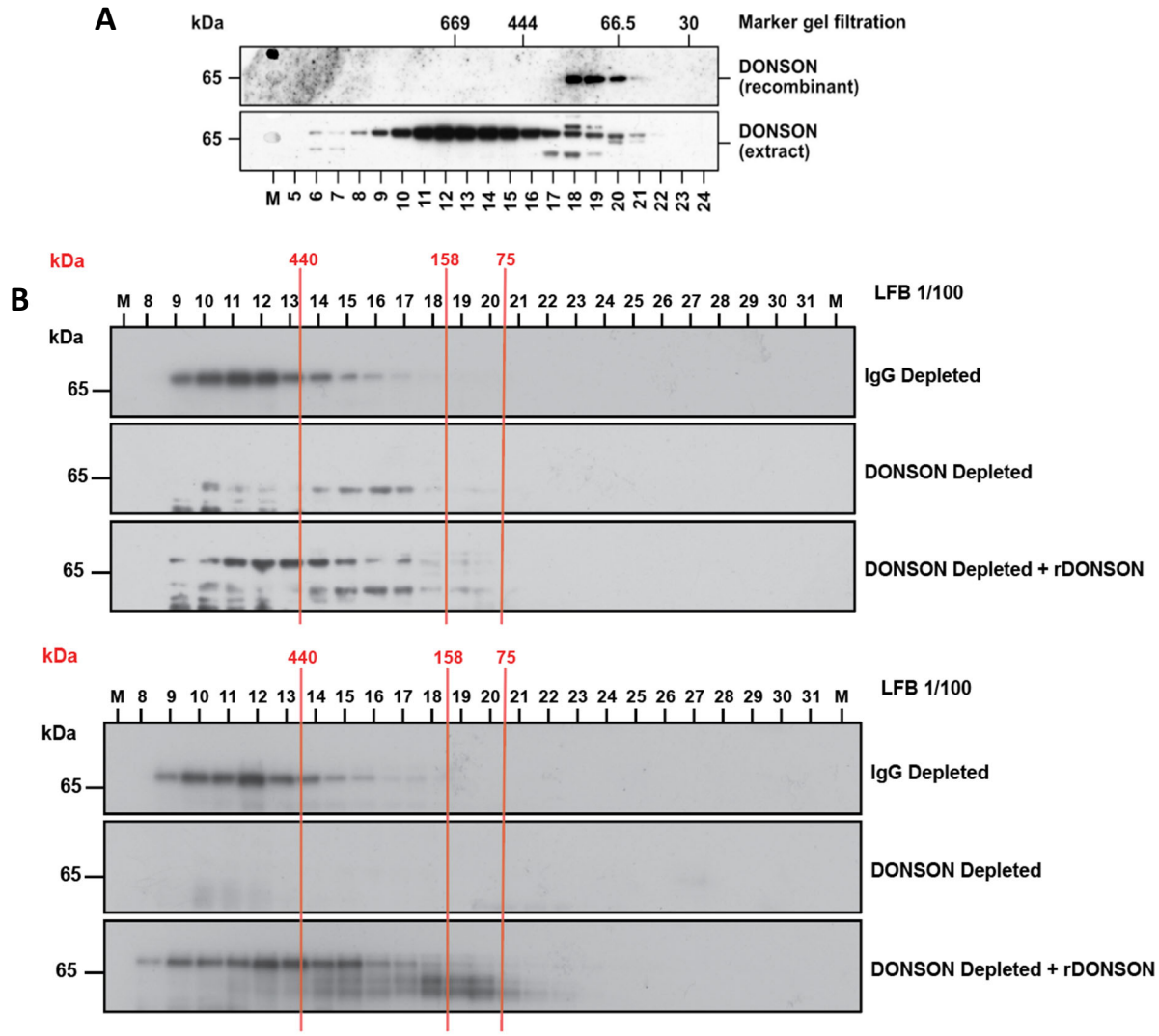


Figure 3.3.2 Endogenous DONSON runs at a higher molecular weight than rDONSON. The recombinant DONSON protein in LFB1/50, and activated egg extract were run through a superose 6 column, as in previous figures. Fractions were run in an SDS-PAGE gel and DONSON was blotted for with specific antibodies. B) IgG- and DONSON-depleted extracts, with the WT rDONSON optionally added into the extract were run through a Superdex 200 gel filtration column and elutions run in an SDS-PAGE gel and blotted for with α-DONSON sheep antibody. N =2, both repeats shown.

3.3.2.2 Does DONSON Impact TopBP1 in Egg Extract?

With some evidence for an interaction between TopBP1 and DONSON (Figures 3.1.13, 3.1.16, 3.3.1 B-D, 3.3.2 B), we looked to determine how this interaction may impact TopBP1's form and interactions within the extract, having observed that TopBP1 can load onto chromatin before/independently of DONSON (Figure 3.1.11 B). With this in mind, we took IgG- and DONSON-depleted extracts and optionally supplemented the DONSON-depleted with rDONSON. Extract samples were then diluted with LFB1/100

and run through a gel filtration column as before. The samples were then run on an SDS-PAGE gel and the molecular weight of certain factors were assessed by immunoblotting with the antibodies specified. As seen in Figure 3.3.3, the size at which TopBP1 runs in the extract is unaltered in the absence of DONSON, compared to the IgG-depleted extract. This suggests that DONSON does not alter the complexes that TopBP1 forms within the extract and may indicate that TopBP1's interactions in complex in the cytoplasm are why we do not strongly co-IP TopBP1 with DONSON here.

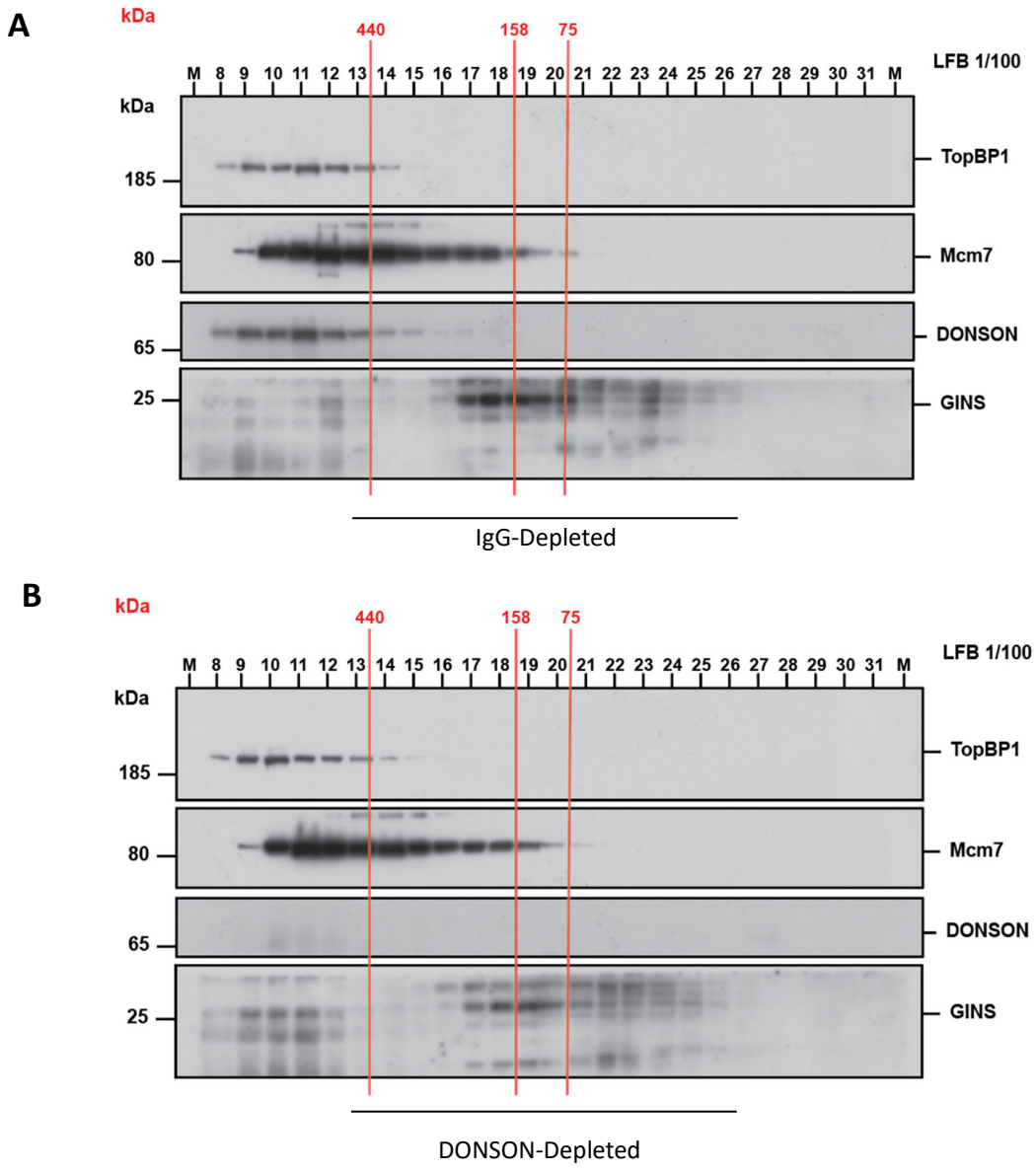


Figure 3.3.3 DONSON depletion doesn't affect the molecular weight at which TopBP1 runs. IgG- (A) and DONSON-depleted (B) extracts were run through a Superdex 200 gel filtration column and elutions run in an SDS-PAGE gel and blotted for with α-DONSON sheep antibody. Representative of two repeats.

3.3.2.3 DONSON can be Phosphorylated

Sld2, the yeast orthologue of DONSON, requires phosphorylation to function during initiation, and this is one way in which initiation is regulated. Likewise, the DONSON protein sequence contains numerous classic CDK sites (Figure 3.3.4 B). It would therefore make sense for DONSON to also be regulated through phosphorylation to regulate and safeguard replication in higher eukaryotes. Therefore, we first looked to test if DONSON can be phosphorylated. With this in mind, we isolated DONSON from chromatin and the nucleoplasm and assessed the protein by mass spectrometry to identify the specific sites at which DONSON could be phosphorylated. As seen in Figure 3.3.4 A, we identified six sites that were phosphorylated, a number of which are classic CDK substrate motifs, serine-proline/threonine-proline. As seen in Figure 3.3.4 A, the identification of the sites *in vivo* (with the endogenous protein) came from four experiments, all under different conditions. Despite the consistently low sequence coverage of DONSON – it is not an abundant protein in either the chromatin or the nuclear fraction – we could still map DONSON phosphorylation at six sites. As seen, all these *in vivo*-mapped sites occur within the flexible N-terminal region of DONSON (Figure 3.3.4 C) and therefore may indicate that phosphorylation at these sites is required for interaction with other factors or for complex stabilisation. Of all the sites identified, T140 and S142 are most frequently identified as phosphorylated, which may suggest that these sites are most accessible to CDKs in extract. Of particular interest was that these sites occur within a region of the protein close to the area predicted to interact with the TopBP1 BRCT3 domain (Figure 3.1.16 B-C). Suggesting that phosphorylation may play a role in stabilisation of this interaction.

Given this, we then looked to phosphorylate our recombinant protein with commercial CDK (Figure 3.3.4 A). To assess this, we optionally incubated our rDONSON with CDK2/Cyclin A or CDK2/Cyclin E – both of which are active during S phase – to determine the extent of phosphorylation. As seen in Figure 3.3.4 B, once samples were run on a phostag gel, a clear band shift of DONSON when incubated with either CDK/Cyclin combination can be observed. This suggests that DONSON is efficiently phosphorylated *in vitro* by S phase CDKs. To ascertain whether we can visualise *in vivo* DONSON phosphorylation on phos-tag gels, NPE was optionally incubated with phosphatase enzymes and samples run on a phostag gel, with

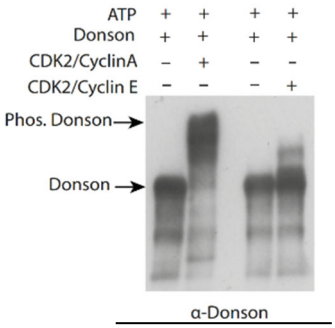
phosphorylated DONSON for size comparison. NPE is made by activating egg extract and extracting the nuclei during late S phase [188]. Therefore, we would expect CDK activity to be present. As seen in Figure 3.3.4 C, whilst in cytoplasmic extract DONSON phosphorylation is not observed, in NPE without phosphatase enzymes applied, there is a band above DONSON. When comparing to the *in vitro* phosphorylated rDONSON sample – although the gel has not run straight – the phosphorylation of the endogenous sample in NPE resembles the phosphorylation of the rDONSON. However, this requires confirmation with a better run gel. We are confident that this is phosphorylation of the protein as this band is absent upon treatment with the phosphatase enzyme. This suggests that a proportion of DONSON is indeed phosphorylated in the nucleus.

As we could observe DONSON phosphorylation by commercial Cyclin-CDKs, we sent the *in vitro* sample for mass spec and observed thirteen further sites phosphorylated on the protein (Figure 3.3.4 D). Interestingly, the majority of the sites are classic CDK-sites and four of these sites are conserved and found phosphorylated in human DONSON. This indicates that, like Sld2, CDK phosphorylation of DONSON may be important for its function.

A

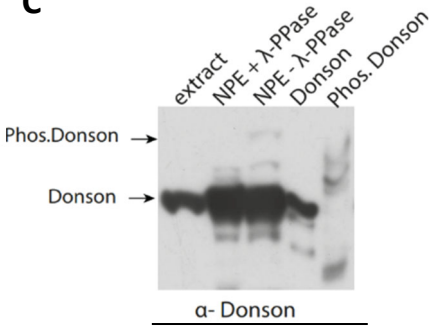
		DONSON IP (Chr.) S-ph Aphidicolin and Caffeine	DONSON IP (Nuc.) S-ph Aphidicolin and Caffeine	DONSON IP (Chr.) S-ph p97i	Mcm3 IP (Chr.) S-ph Aphidicolin, Caffeine and p97i
All DONSON peptides		26 unique peptides 40 unique spectra 86 total spectra 46% coverage	40 unique peptides 89 unique spectra 283 total spectra 47% coverage	26 unique peptides 49 unique spectra 84 total spectra 54% coverage	20 unique peptides 34 unique spectra 74 total spectra 36% coverage
KPSEILRLSR		0	0	0	0
KPSEILRL**S**R	**S22**	0	0	0	1
SRSEASKTGLSPFSPGDVIK		0	1	0	0
SRSEASKTGLSPFSPGDVIK	**S26**	0	1	0	0
SRSEASKTGLSPFSPGDVIK		0	1	0	0
SRSEASK**T**GLSPFSPGDVIK	**T33**	0	1	0	0
RVPGLRPFSPGPNK		3	7	1	4
RVPGLRPF**S**PGPNK	**S54**	0	1	0	0
TESPEKPAPACEIPK		6	12	2	2
TESPEKPAPACEIPK	**T140**	2	4	1	2
TESPEKPAPACEIPK		4	10	1	2
TE**S**PEKPAPACEIPK	**S142**	4	6	2	2

B



In vitro

C



In vivo

D

Xenopus Classic CDK Sites	*In vivo* Mapped	*In vitro* mapped	Human equivalent
S9			S9
	S22	S22	
	S26		
	T33	T33	
S36		S36	
S39		S39	S34
S54	S54	S54	
		T75	
S78		S78	
		T86	
		S94	
S112		101, 106, 112, 120	
	T140	T140	T135
S142	S142	S142	
T252			S253
S314		S314	S308
		S396	
S454		S454	S437
S528			

E

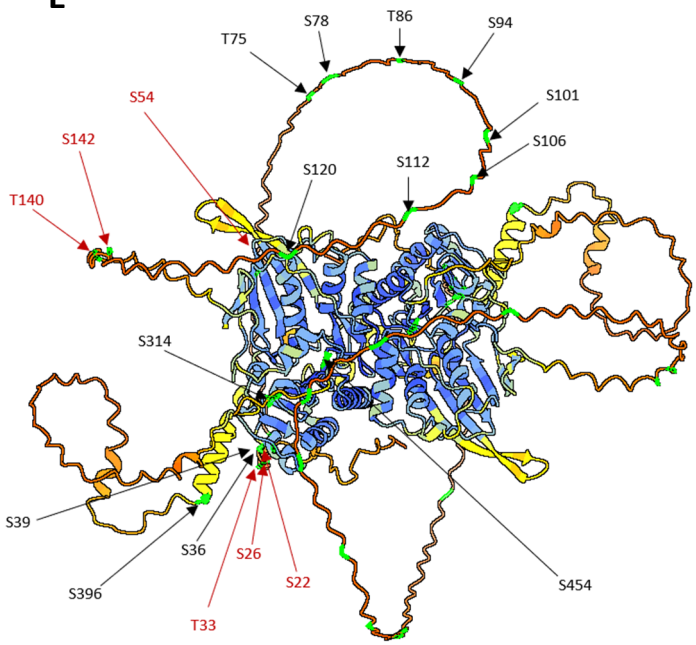


Figure 3.3.4 DONSON can be *in vivo* and *in vitro* CDK phosphorylated. A) Samples of DONSON-IPs from chromatin (chr.) and nucleoplasm (nuc.) in S-phase were analysed by mass spectrometry. Additionally, an IP of Mcm3 from S-phase chromatin was also analysed. Number of DONSON peptides containing identified phosphorylation sites either in phosphorylated (red) or non-phosphorylated (black) form are reported. The number of DONSON unique peptides, unique spectra (includes peptides with differential modifications), total spectra and protein coverage is also presented for comparison between experiments. Only peptides with >80% confidence of phosphorylation identification are reported. B-C) Performed by Dr D Poovathumkadavil. B) WT rDONSON was optionally *in vitro* phosphorylated with either CDK2/CyclinA or CDK2/CyclinE, separated on phos-tag gel and visualised by immunoblotting with DONSON-specific antibodies. C) Egg extract was activated and optionally supplemented with CDK2/CyclinA to phosphorylate DONSON (lane 5), nucleoplasmic extract (NPE) was optionally supplemented with λ-phosphatase (λ-PPase) (lanes 2 and 3), separated on a phos-tag gel and DONSON was visualised with DONSON-specific sheep antibodies. Activated extract without CDK2/CyclinA or λ-PPase, and WT rDONSON were used as controls. D) Presentation of classic CDK sites within DONSON, all endogenous DONSON sites found through IP and mass spec CDK phosphorylated, and all rDONSON sites found through *in vitro* phosphorylation and mass spec. E) AlphaFold3 model of the DONSON dimer with all CDK phosphorylation sites mapped marked in green. Arrows point to these sites on one monomer of the dimer. Red arrows point to sites identified from *in vivo* experiments, black point to those only identified *in vitro*. ipTM = 0.63 pTM = 0.64. Corresponding expected position error plot shown in Figure 3.1.15.

3.3.2.4 Does Phosphorylation Impact DONSON's Structure?

Having observed CDK phosphorylation of DONSON both *in vitro* and *in vivo*, we wanted to determine whether it is important for its initiation function. We had previously observed a decrease in DONSON's chromatin binding upon inhibition of CDK activity by addition of $p27^{KIP1}$ CDK inhibitor (Figure 3.1.10 A). To analyse this in more detail, I set up replication in egg extract and optionally supplemented with $p27^{KIP1}$ to prevent CDK activity. I isolated chromatin throughout replication and ran samples on an SDS-PAGE gel, blotting for DONSON and other indicated factors. As seen in Figure 3.3.5 A, Mcm7 chromatin binding is largely unaltered by CDK inactivity – a slight increase in binding is observed, explained by replication being unable to proceed and hexamers remaining stalled on chromatin. We observe no chromatin binding of either Cdc45 or Psf2 upon application of $p27^{KIP1}$. This is unsurprising, as CDK activity is required for formation of the active helicase on chromatin. Interestingly, DONSON's chromatin binding throughout replication is reduced when CDK activity is inhibited, particularly at the beginning of the reaction, when significant replisome binding can be observed in the control reaction initiation. This suggests that CDK phosphorylation is likely important for DONSON's recruitment to chromatin. However, it is not clear whether this is the result of blocking DONSON phosphorylation directly, or due to phosphorylation of another factor required for its loading. Therefore, we looked to determine the extent to which phosphorylation of DONSON itself changes its appearance. To ascertain this, we assessed our *in vitro* phosphorylated rDONSON in LFB1/50 by gel filtration. The fractions running at a very small molecular weight, likely show a precipitated portion of the recombinant protein. As previously (Figure 3.3.2 A), our rDONSON runs at around the expected molecular weight of the dimer. Interestingly, *in vitro* phosphorylation of rDONSON doesn't appear to alter the molecular weight at which the protein runs (Figure 3.3.5 B). This was particularly interesting to us as we had reproducibly observed that CDK activity is required for DONSON's recruitment to the DNA, as well as its interaction with TopBP1. This led us to question whether, rather than supporting increased weight DONSON complexes, phosphorylation instead releases the protein from higher order complexes (Figure 3.3.1) enabling it to interact with components required for its initiation role. To determine this, I began by assessing whether DONSON phosphorylation alters its structure. Using AlphaFold, I modelled phosphorylated dimeric DONSON, with ten sites of those

we had identified by mass spec in the phosphorylated form. The ten chosen were composed of the six *in vivo* sites we had mapped, plus four that we had mapped *in vitro* and were classic CDK sites and largely conserved in human DONSON, as we predicted that these were the most likely to be readily phosphorylated. We observed that the majority of these phosphorylated sites cluster together at the top and bottom of the dimer interface (Figure 3.3.5 C) at the sites that we had previously indicated were likely occupied by TopBP1's BRCT3 domain (Figure 3.1.16 B-C). As seen in Figure 3.3.5 D, the phosphorylated dimer largely overlaps with the structure of the unphosphorylated, apart from these two regions at the top and the bottom of the interface where the flexible regions have become more condensed into the globular domain of the protein. Therefore, phosphorylation may be required to alter the interactions of DONSON's flexible regions, perhaps preventing higher order interactions and supporting the protein in the dimeric form.

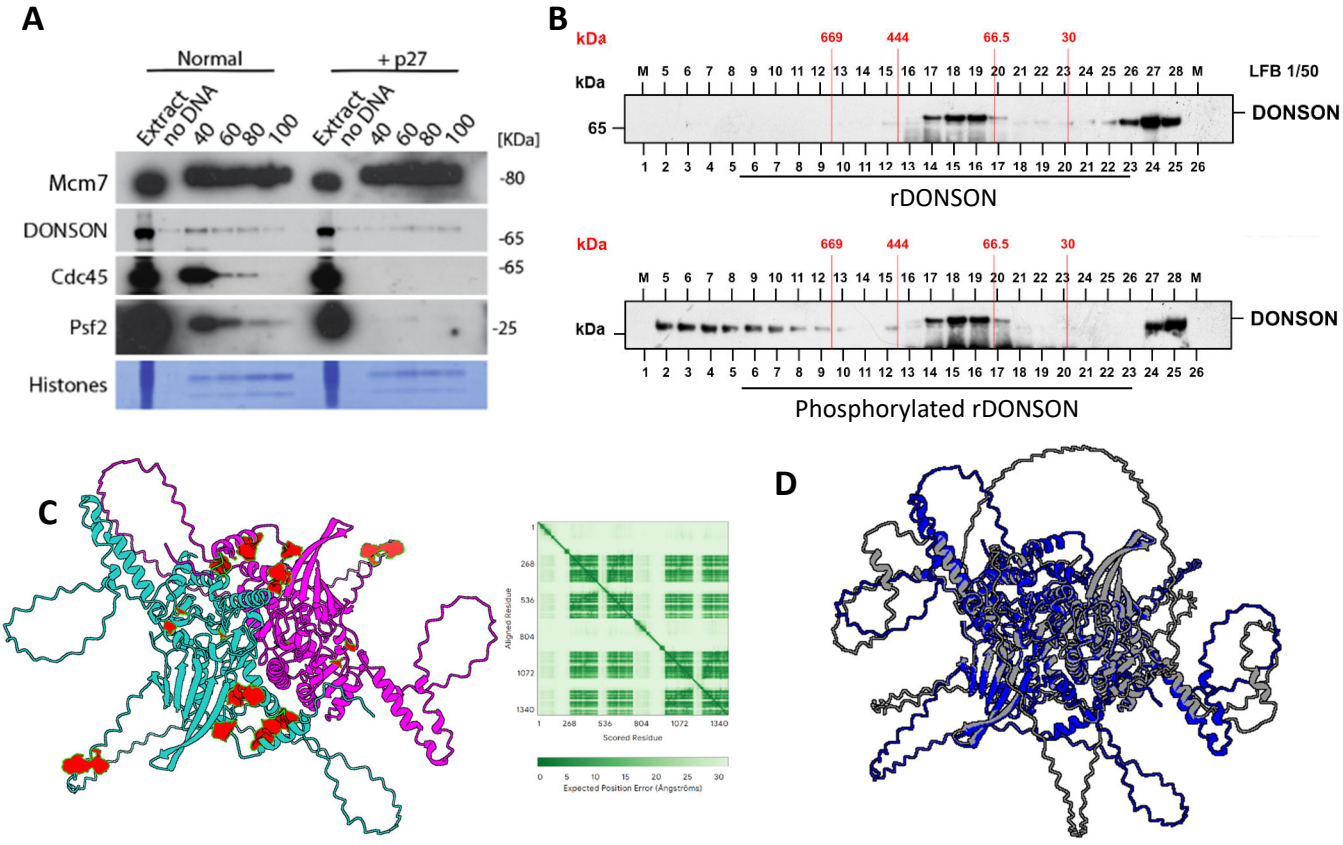


Figure 3.3.5 DONSON's CDK phosphorylation impacts chromatin binding and structure. A) DNA replication was set up in egg extract and optionally supplemented with CDKi (p27). Chromatin was isolated at indicated timepoints and chromatin-bound proteins analysed by immunoblotting. Positive and negative controls were performed as in previous experiments. B) WT rDONSON was optionally *in vitro* CDK phosphorylated, diluted in LFB1/50 and run through a Superose 6 column. Fractions were run in an SDS-PAGE gel and DONSON was blotted for with specific antibodies. C) AlphaFold3 model of DONSON dimer (each monomer shown in cyan or magenta) with 10 sites phosphorylated (shown in red) – 6 sites identified *in vivo* (S22, S26, T33, S54, T140, S142) and 4 of the sites identified *in vitro* (S36, S39, S314, S454). ipTM = 0.53 pTM = 0.55. Corresponding expected position error plot shown. D) AlphaFold3 model of the DONSON dimer with the 10 sites phosphorylated (blue), overlaid onto the unphosphorylated DONSON dimer (grey). RMSD = 0.429 Å.

3.3.2.5 DONSON Phosphorylation is Important to Support DNA Replication

As our results indicated that phosphorylation may impact DONSON's ability to be made available for initiation, as well as its chromatin binding, we looked to determine whether it impacts DNA replication efficiency to test its requirement to support initiation. To assess this, I began mutating DONSON's phosphorylation sites we had identified *in vivo*. The first candidates were T140 and S142, which had repeatedly come up as phosphorylated in our mass spec data, as well as S26 and T33 which are positioned at the very N-terminus of the protein. I expressed in *E. coli* the mutated version of DONSON (4A), with the chosen sites mutated to alanine, to prevent their phosphorylation. We purified the protein as previously and visualised the samples by Coomassie (Figure 3.3.6 A). I then set up replication in IgG- and DONSON-depleted extracts as before and optionally incubated with the WT or 4A DONSON to assess its ability to support nascent DNA synthesis. As seen in Figure 3.3.6 B, this mutant is able to support efficient nascent DNA synthesis. We therefore proceeded to mutate further sites that we had identified, with the idea that DONSON may need to be phosphorylated at numerous sites to achieve the necessary conformational change to support replication, like Sld2. To this effect, I mutated a further four sites identified at the N-terminus – S22, S54, S314 and S454 – all to alanine to prevent phosphorylation (8A) – and purified the protein from *E. coli* (Figure 3.3.6 C). I tested the ability of this protein to support efficient nascent DNA synthesis as before (Figure 3.3.6 D), and once again observed levels with a similar efficiency to the WT protein. I decided to mutate two further sites, both that were in the classic CDK form and one of which is conserved in human DONSON – S36 and S39 – (10A) and purified the protein from *E. coli* (Figure 3.3.6 E). I tested the ability of the mutant to support replication. Promisingly, I observed a significant drop in efficiency of the protein to support nascent DNA synthesis, when testing the mutant from a number of independent purifications and in a number of DONSON depleted extracts (Figure 3.3.6 F). Notably, the WT rDONSON is not rescuing replication in the depleted extract as we had previously observed, this could be due to a less successful purification of the protein. However, the 10A mutant, which was produced from an independent purification is even less able to support efficient DNA replication, reproducibly. This suggests that CDK phosphorylation of DONSON may be required to regulate DONSON replication activity, possibly resulting in a conformational change that supports interactions

necessary for recruitment to the chromatin. However, we wanted to check whether this may in fact be due to an unwanted conformational change from the WT rDONSON resulting from mutation of multiple sites to alanine. Therefore, I modelled the 10A mutant (pink) and compared it to the structure of the WT DONSON dimer (grey). As seen in Figure 3.3.6 G, the overlapping similarity of the two dimers is strong. This indicates that the drop in replication efficiency may indeed be due to the lack of phosphorylation of the protein instead of a byproduct of mutation. However, this requires structural investigation.

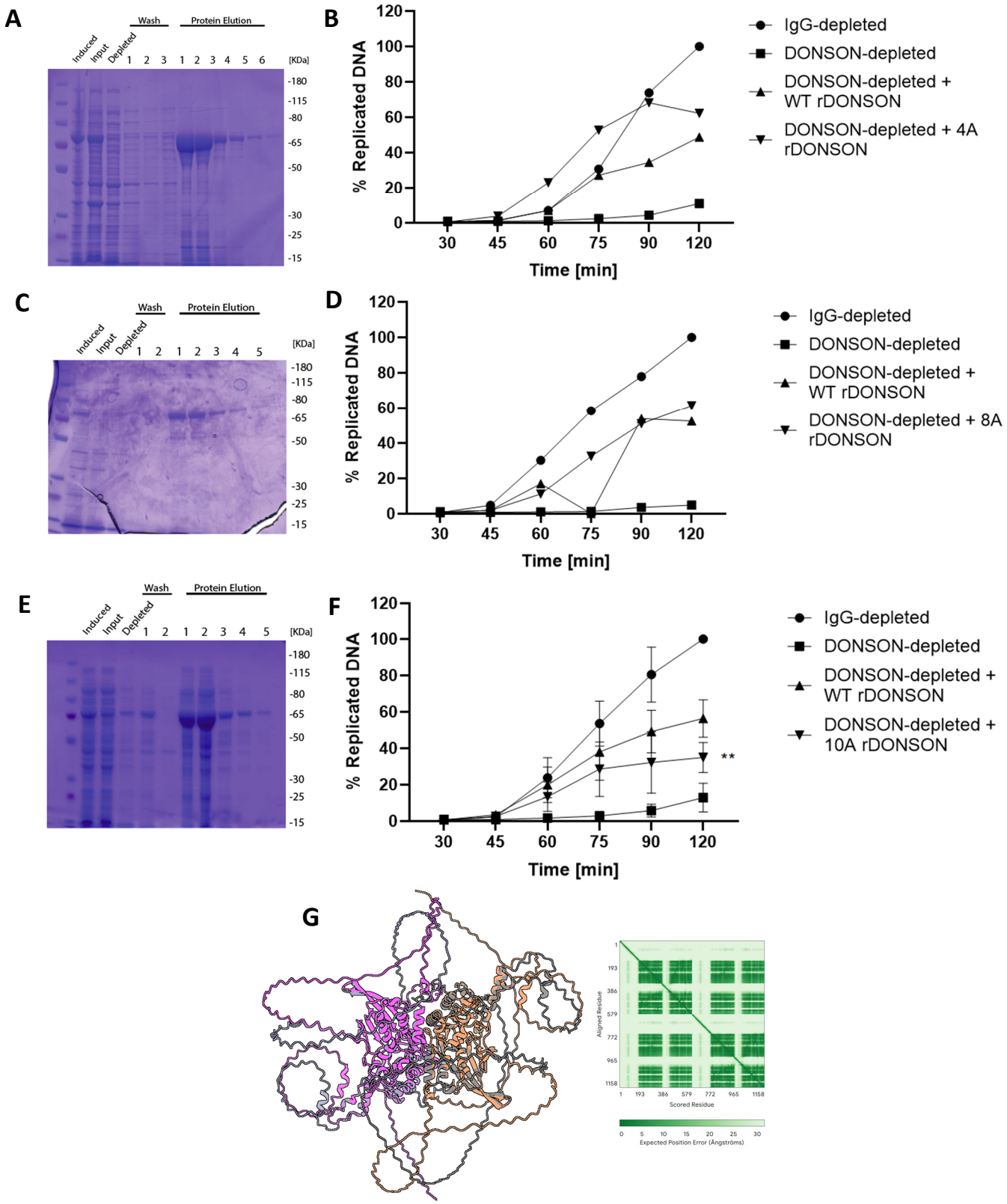


Figure 3.3.6 DONSON's CDK phosphorylation is required for efficient DNA synthesis. A, C, E) Purification of His- and MYC-tagged DONSON S26A, T33A, T140A, S142A (4A) (A), S22A, S26A, T33A, S54A, T140A, S142A, S314A, S454A (8A) (C), S22A, S26A, T33A, S36A, S39A, S54A, T140A, S142A, S314A, S454A (10A) (E) mutants. Expression of xDONSON mutants were induced in *E. coli* (Induced) and purified as described in methods section using Nickel affinity purification. Supernatant from bacterial lysate containing soluble proteins (Input), unbound proteins (Depleted), washed beads (Wash) and elutions using 250 mM of Imidazole (Protein elutions 1-5/6) are shown. Samples from the purification were run on an SDS-PAGE gel and visualised by Coomassie staining. B, D, F) DNA replication was set up in IgG- and DONSON-depleted extract with the optional addition of WT rDONSON or indicated mutants and incubated with α-[P32]dATP. The DNA replication reaction was stopped at indicated timepoints and DNA synthesis was measured as in previous figures. B and D) n = 2. F) n = 3. Mean and SEM plotted. Two-way ANOVA comparing to IgG-depleted, **p = 0.0044. G) AlphaFold3 model of the DONSON dimer with the 10 sites mutated (pink and coral – each colour for one monomer). ipTM = 0.63 pTM = 0.64. Corresponding expected position error plot shown. Mutated dimer overlaid onto the WT DONSON dimer (grey). RMSD = 0.392 Å.

Having observed a decrease in replication with the 10A phosphorylation mutant, we wanted to ascertain whether this is due to DONSON phosphorylation being required for formation of the active helicase. To assess this, interphase IgG- and DONSON-depleted extract were optionally supplemented with either WT or 10A rDONSON. I isolated chromatin throughout replication and observed chromatin binding of indicated factors. The binding of Mcm7, MTBP and PCNA remain consistent in each condition, unaffected by the mutation of DONSON. Similarly, Pol Ɛ binds chromatin with the same pattern in each condition, also suggesting that phosphorylation at these ten sites does not impact its recruitment to the chromatin. Promisingly however, both Cdc45 and GINS chromatin recruitment is reduced when 10A DONSON is added, indicating that DONSON phosphorylation may indeed be necessary for formation of the active helicase. Interestingly, where we had previously observed a decrease in DONSON chromatin binding when CDK activity is blocked, when these sites cannot be phosphorylated, DONSON is efficiently recruited to the chromatin. This suggests that general CDK phosphorylation of the protein, rather than specific phosphorylation of DONSON at these sites may be required for its chromatin recruitment, and therefore that DONSON phosphorylation is instead required for formation of the CMG. However, TopBP1 chromatin binding appears to be somewhat increased, indicating that TopBP1 is stalled on the chromatin – although it is important to confirm that this is not the result of a loading artefact. This suggests that TopBP1 can still load 10A DONSON onto the chromatin, yet appears unable to support DONSON's recruitment of GINS. However, this requires further investigation.

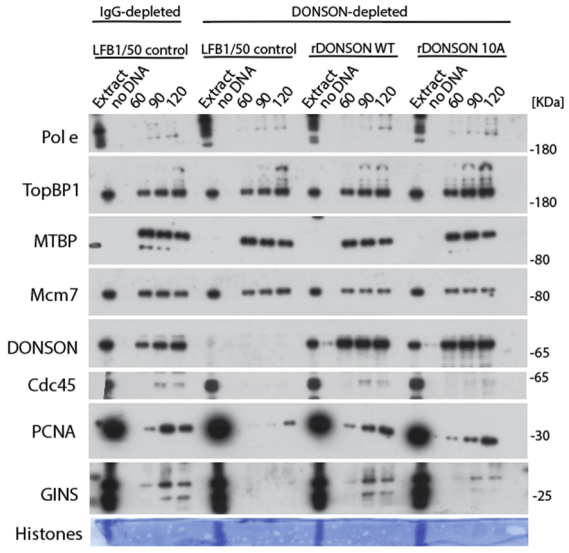


Figure 3.3.7 DONSON's CDK phosphorylation is required for efficient helicase formation. A) DNA replication was set up in IgG- and DONSON-depleted extract with the optional addition of recombinant WT or 10A mutant. Chromatin was isolated at indicated timepoints and chromatin-bound proteins analysed by immunoblotting. Positive and negative controls were performed as in previous experiments.

3.3.2.6 DONSON Phosphorylation is Important for its Interactions

Having observed perturbation of DNA replication initiation in the absence of DONSON phosphorylation, I wanted to determine whether this is due to the requirement of phosphorylation to support DONSON's interaction in a pre-LC-like complex, ahead of chromatin recruitment. To assess this, I performed DONSON IPs in DONSON-depleted extract with the optional addition of WT or 10A rDONSON (Figure 3.3.8 A). As seen previously (Figure 3.3.1 B), we observed no coimmunoprecipitation of GINS with WT DONSON, and similarly no interaction with 10A. Likewise, as previously observed, a faint band of TopBP1 can be seen co-immunoprecipitating with the WT DONSON. In contrast, this band is lost when 10A is applied to the extract, suggesting that phosphorylation may be required for the TopBP1-DONSON interaction. However, as the TopBP1 band with the WT IP is so weak, this requires further investigation.

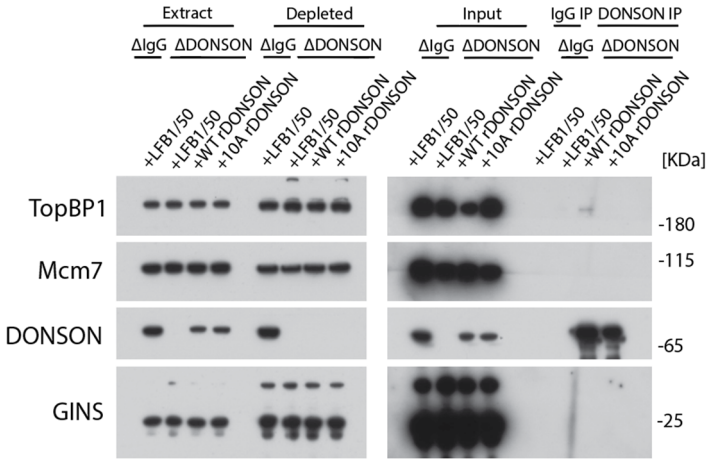


Figure 3.3.8 DONSON's CDK phosphorylation is required for efficient TopBP1 interaction. WT or 10A rDONSON were optionally added to DONSON-depleted extract and IP of DONSON was performed. Other factors were blotted for using specific antibodies. Extract from each condition without antibodies (Input) and IgG-depleted extract were used as controls.

As this indicates that phosphorylation may play a role in enabling or stabilising the interaction between DONSON and TopBP1, I once again used AlphaFold3, to determine the differences between the WT protein, the phosphorylated WT and the 10A mutant in the possible structure of a vertebrate pre-LC-like complex. I modelled DONSON in the dimer form with two Sld5 subunits, as we had previously observed that DONSON interacts with Sld5 specifically to recruit GINS to the chromatin. Likewise, I modelled two TopBP1 molecules, up to the end of the BRCT4 domain, as BRCT1-3 have been shown to be required to support replication [74], while the GINS-interacting domain in human TopBP1 is positioned between domains 3 and 4 [193]. I also modelled Pol Ɛ as it is the final proposed component of the pre-loading complex. In particular, I modelled the catalytic domain of Pol Ɛ (POLE2) due to its size and importance in

DNA replication. As seen in Figure 3.3.9 A, and proposed previously, the DONSON dimer forms the central binding platform for TopBP1 and Sld5 binding, and seen here Pol Ɛ. Of note, the complex is sparse and appears disordered. In contrast, when phosphorylated at the ten sites chosen above, the entire formation of the complex is changed. Most strikingly, the complex becomes denser and more ordered, with all factors condensed around the DONSON core dimer. Interestingly, the position of the two Sld5 monomers twists vertically, and when viewed from the side, can be seen to allow Pol Ɛ to condense into DONSON and position itself more vertically also. Intriguingly, the majority of the phosphorylated sites appear to directly bind to the N-terminal region of TopBP1, whilst the BRCT3 domain remains where it is positioned in the unphosphorylated form, suggesting that phosphorylation may be key for inducing additional interactions between DONSON and TopBP1, creating a more stable scaffold for GINS and Pol Ɛ to bind and indeed form a pre-LC-like complex. The complex formed in the presence of the 10A mutant dimer bears a striking resemblance to that formed around the unphosphorylated WT dimer. Indicating that the results we have observed with this mutant may be due to decreased stable interactions with TopBP1, GINS and Pol Ɛ resulting from lack of phosphorylation in this protein. Whilst interesting, these preliminary findings require biochemical investigation and cryo-EM structure identification.

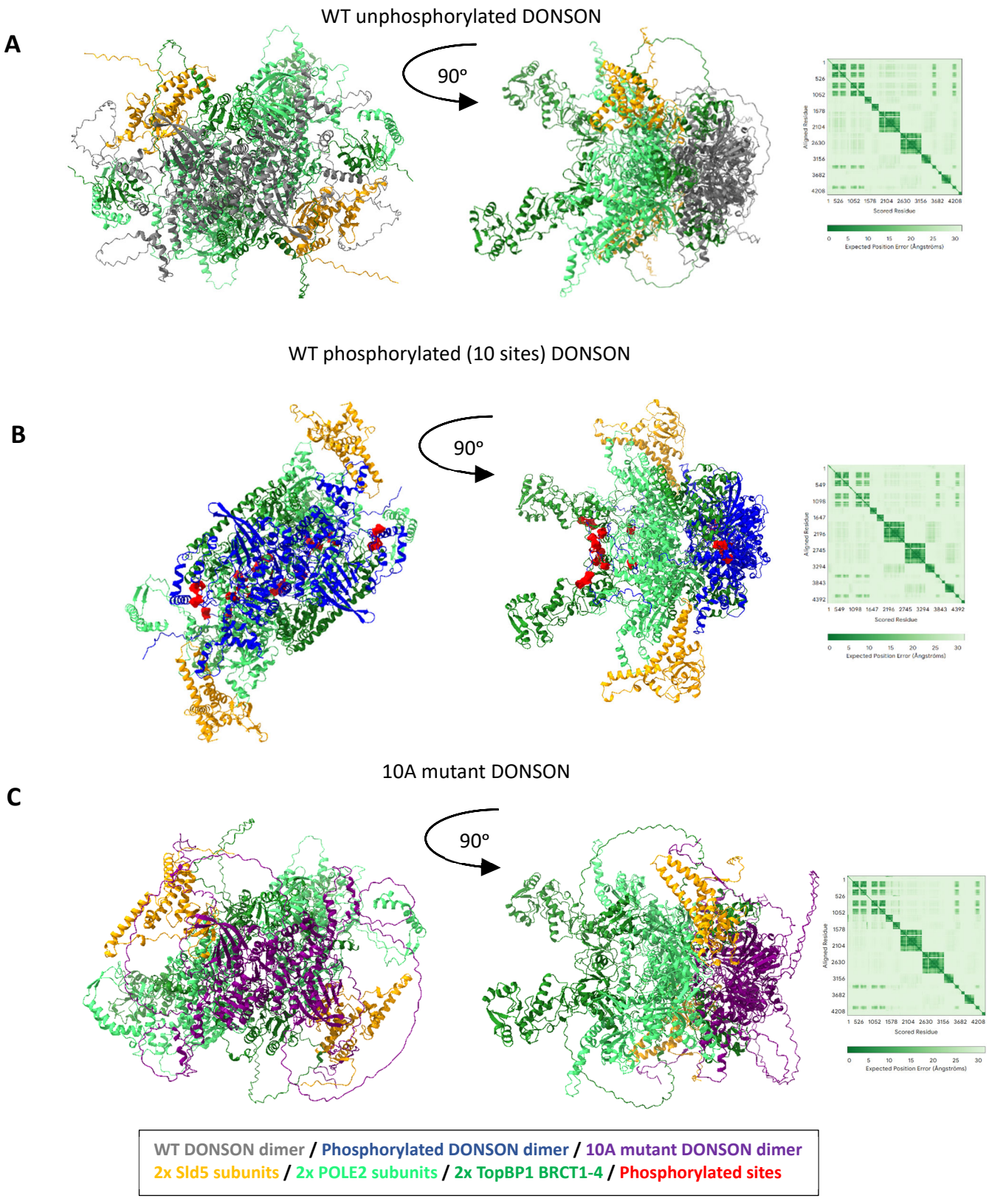


Figure 3.3.9 DONSON's CDK phosphorylation impacts 'pre-loading complex' structure. A) AlphaFold3 modelling of WT DONSON dimer (grey) with 2x Sld5 (yellow), 2x TopBP1 (1-630) (dark green) and Pol Ɛ catalytic domain (light green). ipTM = 0.37 pTM = 0.41. Corresponding expected position error plot shown. B) AlphaFold3 modelling of DONSON dimer phosphorylated at 10 sites as indicated in Figure 3.3.5 D (blue, phosphorylated amino acids marked in red) with 2x Sld5 (yellow), 2x TopBP1 (1-630) (dark green) and Pol Ɛ catalytic domain (light green). ipTM = 0.35 pTM = 0.39. Corresponding expected position error plot shown. C) AlphaFold3 modelling of 10A mutant DONSON dimer (purple) with 2x Sld5 (yellow), 2x TopBP1 (1-630) (dark green) and Pol Ɛ catalytic domain (light green). ipTM = 0.36 pTM = 0.4. Corresponding expected position error plot shown. A-C) Left, complex viewed at the dimer interface. Right, rotated 90°.

3.3.3 Discussion

In this chapter, I have largely been addressing how DONSON is regulated throughout replication and the cell cycle – primarily focussing on the impact of phosphorylation of the protein. One principal reason for addressing this was to determine the potential reasons for why we were not observing a stable pre-loading, or pre-LC-like, complex when others were [195], [196]. In line with this, I have demonstrated that DONSON can be phosphorylated *in vivo* and *in vitro*, and that phosphorylation at these sites is required to support efficient nascent DNA synthesis and for CMG formation on the chromatin. Modelling suggests that this may be due to phosphorylation being required for interactions with components of a vertebrate pre-LC-like complex and accurate positioning of the factors for interaction with the chromatin.

A pre-LC consisting of the functional analogues of TopBP1, DONSON, GINS and Pol Ɛ has long been proposed and accepted in budding yeast, with each factor associating with origins in a mutually dependent manner [69]. Here, the pre-LC forms independently of DNA but in a CDK-dependent manner. Phosphorylation of Sld2 is required for activation of DNA replication, particularly for its interaction with Dpb11, which forms the basis of the complex. In higher eukaryotes, an analogous pre-LC was yet to be identified. Through their *in silico* investigation of DONSON, the Walter lab proposed that, like its yeast homologue Sld2, DONSON also interacts with TopBP1 to provide the scaffold for a pre-LC prior to DNA binding [194]. In *Xenopus* cytoplasmic egg extract, they also observed reciprocal interactions between a FLAG-tagged recombinant DONSON and all other components of the proposed pre-LC, whilst upon perturbation of these interactions, DNA replication could not be supported. Likewise, using *Xenopus laevis* egg extract, the Tanaka group identified a 'pre-LC-like' complex which begins with formation of a subcomplex containing DONSON, GINS, Pol Ɛ and Cdc45, before being recruited to chromatin by an essential interaction with TopBP1 [195]. However, through chromatin isolation in egg extract, they observed that TopBP1 appears to bind chromatin independent to DONSON, GINS and Pol Ɛ. Therefore, suggesting that, in contrast to the yeast pre-LC, the vertebrate pre-LC-like complex does not contain TopBP1 until it associates with the DNA. Using single-molecular imaging, the Chistol group, like the Tanaka group, observed that TopBP1 is recruited to the Mcm2-7 hexamer in egg extract prior to GINS, specifically one minute before and independent of DONSON [197]. Whilst we have observed interactions with GINS

on chromatin (Figure 3.1.5 B and C, and 13 A), and TopBP1 on chromatin (Figure 3.1.13 A and B) and in the nucleoplasm (Figure 3.1.13 D), as well as both *in vitro* (Figure 3.1.16 E), we have only observed an interaction, and a very weak one, with TopBP1 in the cytoplasm (Figure 3.3.1 B and 8 A). In contrast to the Walter and Tanaka groups, this suggests to us that if there is a vertebrate pre-LC, that it is fragile and forms only transiently prior to binding to the DNA.

To address why our findings did not match with those proposing that DONSON forms a vertebrate pre-LC, we began by assessing the ability of the endogenous, compared to the recombinant protein, to interact with other pre-LC factors. Particularly as both groups used recombinant DONSON either only in NPE, or with higher concentrations than we have observed the endogenous protein to be expressed at [194], [195]. As seen in Figure 3.3.1 A, when assessing these interactions in extract isolating the endogenous DONSON, we could observe no robust interaction with either GINS or TopBP1. In contrast, when applying the recombinant protein, we observe coimmunoprecipitation of TopBP1 in the cytoplasmic extract (Figure 3.3.1 B). This suggests that there is a difference between the endogenous and recombinant DONSON proteins. Assessing the endogenous (in extract) and recombinant (in buffer) proteins by gel filtration (Figure 3.3.2) indicated that the recombinant protein runs at a lower molecular weight than the endogenous. However once added to extract, the recombinant runs at the same molecular weight as the endogenous. Although this suggests that the recombinant protein acts like the endogenous in extract, we haven't tested our rDONSON in NPE. The features of NPE differ from the cytoplasmic extract, notably it contains higher CDK activity. It may be that in replicating or active extract, formation of multimeric complexes by rDONSON is prevented and the protein is left available to dimerise and interact with pre-LC factors. This may explain the differences between our results and those of the Walter lab. In contrast, the differences with the Tanaka group results may instead be explained by a much higher concentration of rDONSON being added to the extract than we observe the endogenous at. Whilst we estimate an average endogenous concentration in egg extract of 40 nM, they use a range from 60 nM-8.8 μM [195]. This may result in an artificially large amount being available for the 'pre-LC-like' interactions in extract.

Interestingly, the interaction that we observe between DONSON and TopBP1 in nucleoplasm is significantly reduced upon the addition of a CDKi (Figure 3.1.13 D-E). Likewise, as previously mentioned, the chromatin binding of DONSON also appears to require CDK activity (Figure 3.1.10 and 3.3.5 A). This indicates that the recruitment of DONSON to chromatin during initiation is both TopBP1- and CDK-dependent. Interestingly, Sld2 requires phosphorylation at multiple sites to enable interaction with Dpb11 and resultant recruitment to chromatin [67]. In yeast this is a necessary step for replication regulation, limiting initiation to only appropriate conditions and timing [68]. Having also observed phosphorylation of DONSON at many sites, we propose that CDK-dependent phosphorylation of DONSON is likewise required to regulate initiation in higher eukaryotes, to regulate the rate of CMG formation and to prevent re-replication of the DNA. In contrast, both the Walter and Tanaka groups suggest that phosphorylation of DONSON is not required for an interaction with TopBP1 [194], [195]. However, in line with our data (Figures 3.1.10, 3.3.5 A) the Tanaka group suggests that recruitment of DONSON to chromatin is. In a similar experiment in egg extract to Figure 3.1.10, the Walter lab also identified that DONSON chromatin-binding is abrogated upon the addition of CDKi, DDKi and geminin [194]. They concluded that DONSON functions late during replication, however did not suggest that this role is dependent on its phosphorylation by CDKs, but instead that general CDK activity is required prior to loading of the protein.

To further assess the potential impact of phosphorylation on accessibility of the protein and having identified several CDK-phosphorylated sites on DONSON, I mutated the protein to prevent phosphorylation at many of these sites. With this mutant, we observed that DNA replication efficiency decreased, as well as CMG formation on chromatin and finally interaction with TopBP1 in non-replicating cytoplasmic extract (Figures 3.3.6 D, 7, 8 A). Were DONSON's interaction with TopBP1 and its recruitment to chromatin only dependent on phosphorylation of other factors, mutating these phosphorylation sites would have made no difference to replication, CMG formation or interaction with TopBP1. Therefore, this indicates that phosphorylation of DONSON itself is likely required for its initiation function. However, as seen in Figure 3.3.6, phosphorylation at many sites is required for its function, suggesting that phosphorylation of specific sites is not necessarily important, but rather robust overall phosphorylation to alter the structure of the protein, similar to Sld2, is [68].

Having identified the importance of phosphorylation for DONSON's initiation function, we assessed whether this might specifically be the reason why we haven't observed sufficient interactions between DONSON and the other factors of the pre-LC. To address this, I turned to AlphaFold. I modelled the interactions between the components of the pre-LC. Specifically, I wanted to visualise the difference that phosphorylation of the protein makes to these interactions and the structure of the complex. As seen in Figure 3.3.9 B, phosphorylation of numerous sites that we had mapped by mass spec appears to condense the entire complex – supporting increased interactions and a more stable form of the complex. This is particularly interesting as BRCT3 of TopBP1, that we propose DONSON binds to, has not been predicted to bind phosphor-peptides [239]. Based on the modelling, in the context of the other pre-LC components, these phosphorylated sites appear to be important for binding POLE2 and the Treslin-MTBP-interacting domains of TopBP1 specifically. Interestingly, in yeast Cdk-dependent phosphorylation is not all that is required for Sld2-Dpb11 interaction, but does enhance efficient interaction [237]. This therefore suggests to us that a vertebrate pre-LC is likely a fragile and transient complex that requires CDK phosphorylation of DONSON as one of its conditions for formation.

We have observed DONSON separating through the gel filtration column at a high molecular size, possibly suggesting a multimeric complex. However, we have also observed that DONSON's role in supporting CMG formation and stabilisation requires only a dimeric form of the protein. Additionally, phosphorylation appears to be important for DONSON's initiation interactions. Therefore, we suggest that phosphorylation may play a role in releasing DONSON from a large complex in extract and enabling its interaction with TopBP1. Interestingly, DONSON's N-terminal disordered region is the part of the protein in which we have observed the greatest likelihood of phosphorylation (Figure 3.3.3). In recent years, many proteins containing disordered regions have been identified as undergoing assimilation into phase-separated condensates [240], [241]. These structures are characterised by their lack of membranes and formation through liquid-liquid phase separation (LLPS). LLPS occurs due to multivalent interactions between proteins and/or nucleic acids, often via the intrinsically disordered regions of the proteins, forming dense liquid-like spots within the cell. These condensates have been observed to support various biological processes, including transcription, chromatin organisation and the DNA damage response [242].

Whilst DONSON has not previously been proposed to form phase-separated condensates, it has been observed to function as part of the ATR stress response [152]. TopBP1 has also been shown to be a principal activator of ATR in response to replication stress and has in fact been shown to extensively self-assemble into phase-separated condensates to amplify activation of ATR [101], [243]. Although we have not yet investigated whether DONSON forms condensates, or whether it may interact with TopBP1 and perhaps other factors involved in the ATR stress response, this is a possibility worth exploring. This is of particular interest, as preliminary data from my project has indicated that upon mutation of an ATR-phosphorylation site within DONSON, the DNA replication efficiency of the extract increases (data not shown). Likewise, we have identified interactions between DONSON and numerous chromatin remodelling factors, as well as ETAA1, which is known to directly activate ATR following preliminary activation by RPA [102]. Collectively, this indicates that DONSON is involved in the ATR stress response, perhaps functioning to support chromatin remodelling to downregulate DNA replication and transcription during DNA repair. To support this function, DONSON may likewise be recruited to a large complex, perhaps supported by other implicated factors that form a scaffold for DONSON, such as TopBP1.

Related to DONSON's propensity to be phosphorylated, the PTM has been observed to significantly impact LLPS [244]. Apart from altering the charge of the protein, phosphorylation also impacts the structure, distribution and interaction potential. At disordered regions of proteins specifically, phosphorylation alters their charge distribution, resulting in changes to their protein-protein binding affinities [245]. Through these mechanisms, phosphorylation can either support or abrogate the formation and maintenance of liquid-like condensates, specific to the amino acid and protein. Therefore, it may be that the endogenous DONSON primarily exists in these large condensates and requires CDK-dependent phosphorylation at the beginning of S phase to promote release of the monomeric or dimeric protein for its function during replication initiation.

Collectively, this chapter has shown that phosphorylation of DONSON is important for its role in supporting DNA replication in *Xenopus* egg extract, as well as in its interaction with TopBP1. We also propose that there may be a possibility that DONSON primarily exists in large complexes in the extract,

requiring phosphorylation of the protein to release monomeric or dimeric DONSON and interact with TopBP1 to recruit GINS to the chromatin and support CMG formation. However, this is inconclusive and requires further investigation.

4 Discussion

This thesis has brought new insights into the process of DNA replication initiation, in particular the formation of the active helicase. The aim of my project was to determine whether DONSON has further roles during replication than solely as a fork protection factor. DONSON is an essential protein to maintain fork stability; however, it is not yet clear the way in which DONSON supports this and neither has it been previously proposed to support replication initiation.

Although in recent years the mechanism by which the CMG is formed on chromatin has been investigated, not all the factors required in higher eukaryotes have been determined. In this thesis I have discovered that DONSON is essential to recruit GINS to the chromatin to enable formation of the CMG and support DNA replication in *Xenopus laevis* egg extract. Critically, I have shown that the roles that DONSON plays during replication are far more extensive than previously proposed, and that phosphorylation is one mechanism by which the initiation function of the protein is likely regulated. Furthermore, utilising the novel finding that the N-terminus of DONSON is required to interact with the GINS complex and recruit it to the chromatin, I have been able to utilise this peptide to abrogate CMG formation and DNA replication, shedding light onto a potential new inhibitor of DNA replication and the cell cycle. Collectively, my findings suggest that DONSON is an integral component of the initiation stage of replication, without which replication cannot successfully commence, indicating a potential molecular mechanism for the phenotypes observed in patients with MPD- or MGORS-DONSON mutations.

Initially, we thought that DONSON may function during DNA replication termination, having previously observed it interacting with Mcm3 in post-replication conditions [35]. During my project, I have confirmed that DONSON directly interacts with known components of the replisome on chromatin. DONSON has previously been observed to interact with MCM2-7 and GINS subunits as well as with CDC45 in human cells [152], [156]. I have observed DONSON interacting with each of these subunits on chromatin in *Xenopus* egg extract, but not outside of this context in the nuclei or cytoplasm. Likewise, I have shown that DONSON is a component of the replisome, binding to and dissociating from the chromatin with similar dynamics to those of other known replisome components, particularly factors of the CMG.

However, DONSON's binding during early replication, as well as the perturbation of replication on its depletion from the extract, led us to consider whether DONSON may be playing a critical role during early steps of replication. During my PhD, I have successfully demonstrated DONSON's role in the recruitment of GINS to chromatin during initiation, which it appears to carry out in a TopBP1-dependent manner – suggesting it may be the higher eukaryotic homologue of yeast Sld2 [69].

4.1 DONSON's Role as the Higher Eukaryotic Sld2 Homologue

The order in which the factors required for replication initiation are loaded onto the chromatin is generally conserved throughout species, beginning with recruitment of the Mcm2-7 hexamer by ORC, Cdc6 and Cdt1 in G1 phase [221], [222], [246], [247]. To form the active helicase, Cdc45 and GINS must be recruited to each hexamer at the beginning of S phase, which requires a plethora of initiation factors. The activity of these factors depends on high levels of S phase kinases, and we propose that DONSON is no exception.

To begin, the Dbf4-dependent kinase (DDK) Cdc7, with Dbf4, selectively hyperphosphorylates Mcm4 of the loaded double hexamers [60]. In yeast, this is required for the recruitment of Sld3-Sld7 with Cdc45, whilst in *Xenopus* egg extract, this is required for the recruitment of Treslin-MTBP with Cdc45 [125], [182]. In *Xenopus* egg extract, Treslin-MTBP has previously been proposed to bind chromatin prior to the start of S phase, with their levels peaking during S phase, the same as the *Drosophila* homologue of DONSON, Hd [159], [182]. Upon DDK inhibition in yeast, Sld3-Sld7 and Cdc45 cannot load onto chromatin. Interestingly however, in *Xenopus* egg extract, DDK does not appear to be essential for the recruitment of Treslin-MTBP to Mcm2-7 hexamers, but rather stabilises the complex here [182]. Likewise, we can still observe Treslin-MTBP chromatin binding in the presence of DDKi. However, Treslin has also been proposed to be a direct target of DDK activity in *Xenopus* egg extract, with chromatin-associated Treslin migrating faster upon DDKi [182]. In line with this, we have also observed a possible faster migration of MTBP upon application of PHA-767491. However, having observed Treslin-MTBP nonetheless binding chromatin in the presence of DDKi, if Treslin-MTBP is a direct target of DDK, it does not appear to be essential for recruitment. Rather perhaps, it may be required to stabilise the complex to the Mcm2-7 hexamer. This suggests that

DDK-dependent phosphorylation of Treslin may be more important for the stabilisation of this interaction than recruitment.

In yeast, upon DDK inhibition, GINS (Psf1) is not loaded onto chromatin either [125]. This suggests that the chromatin recruitment of the GINS-specific loading factors Sld2 and Dpb11, is dependent on DDK activity. Therefore, if DONSON is indeed the functional homologue of Sld2, its initiation function may be (indirectly) regulated in an alternative manner, specifically with greater importance placed on the activity of DDK. Using *Xenopus laevis* egg extract we have also observed that DONSON recruitment to chromatin is slightly downregulated upon DDK inhibition with PHA-767491. Whilst this suggests that DDK-dependent phosphorylation is required for the chromatin-loading activity of Sld2 and Dpb11, and likely DONSON and TopBP1, DDK can not phosphorylate Sld2 [248]. This suggests that Sld2 and DONSON recruitment to chromatin occurs late during initiation, after Sld3-Sld7-Cdc45/Treslin-MTBP-Cdc45 have loaded and is therefore indirectly dependent upon DDK activity. It is also likely that all these factors stabilise each other at the origins.

In yeast, following DDK-dependent recruitment of Sld3-Sld7-Cdc45 to chromatin, CDK targets Sld2 and Sld3 for phosphorylation to promote their interaction with Dpb11 [67], [68]. Upon inhibition of CDK activity Sld3-Sld7, along with Cdc45, can still be loaded onto chromatin bound Mcm2-7 hexamers [59]. However, chromatin loading of other firing factors is prevented: Dpb11, Sld2, Pol Ɛ, GINS, Mcm10. In contrast, in *Xenopus* egg extract, upon addition of $p27^{KIP1}$ to prevent CDK activity, Treslin-MTBP chromatin-loading increased, whereas Cdc45 did not [182], [249]. In this project, I also observe this binding pattern. This suggests that replication is blocked following Treslin-MTBP loading and either an additional, CDK-dependent, factor is required for Cdc45 chromatin loading in *Xenopus* compared to yeast, or CDK activity is required to specifically stabilise Cdc45 once loaded. Serine-1001 of Sld3/Treslin is well conserved and is a site of CDK phosphorylation required for interaction with TopBP1 and its recruitment to the Mcm2-7 hexamer [204], [249]. Therefore, it is not surprising that upon CDK inhibition, TopBP1 recruitment to chromatin was abrogated in egg extract. In contrast, following CDK inhibition I observed that TopBP1 binding to chromatin was reduced rather than ablated. Additionally, we observe a substantial reduction in chromatin binding of both DONSON and GINS. This therefore suggests that whilst CDK

phosphorylation is required only to increase the efficiency of TopBP1 chromatin-loading, it is essential for any chromatin-loading of either DONSON or GINS. Of note, addition of CDKi in the nucleoplasm abrogated the direct interaction observed between DONSON and TopBP1, indicating that the reason we see a loss of DONSON on chromatin under such conditions is likely due to a loss of interaction with TopBP1. Additionally, the interaction between our CDK-phosphorylation mutant (10A) and TopBP1 is decreased in comparison to WT DONSON. This may suggest that not only is general S phase CDK activity required for DONSON's chromatin loading, but direct phosphorylation of DONSON may be required for its interaction with TopBP1 and subsequent recruitment to chromatin. However, it should be noted that our 10A mutant could still load onto chromatin. Therefore, it may be that CDK-dependent phosphorylation of DONSON is required only for an efficient interaction of DONSON with TopBP1. Alternatively, the additional CDK sites on DONSON that we have not mutated may still support a sufficient interaction with TopBP1, GINS or the hexamer to allow DONSON to bind chromatin but not to fulfil its functions. Finally, CDK phosphorylation of an additional factor (likely Treslin) may be required for this recruitment to occur and therefore with Treslin still being phosphorylated, when DONSON is mutated, DONSON can still be recruited. Nevertheless, phosphorylation of DONSON is required to efficiently recruit GINS to the chromatin for CMG formation and therefore further investigation is needed to understand the mechanism of this regulation.

Not only have I established that interaction with TopBP1 and effective chromatin recruitment of DONSON and GINS is CDK-dependent, but I have also observed that the efficiency of DNA replication and active helicase formation appears to depend on direct CDK phosphorylation of DONSON. Therefore, assessing whether DONSON's phosphorylation is required for the stabilisation, not just formation of, the CMG on chromatin was of interest. As previously mentioned, both our results and previous studies in *Xenopus* egg extract, showed that upon addition of a CDKi, Cdc45 could no longer be observed binding to chromatin. In contrast, CDKi addition led to increased binding of Treslin-MTBP, likely indicating that the two proteins are stalled and replication cannot commence without CDK activity. Whilst this may suggest that Treslin's CDK-dependent phosphorylation may be necessary for its interaction with Cdc45, Cdc45 is not observed binding chromatin in the absence of DONSON when CDK activity is permitted. Likewise, upon addition of

our 10A phosphorylation mutant DONSON, limited Cdc45 chromatin loading can be seen. Therefore, this instead suggests that through its recruitment of GINS, phosphorylated DONSON in particular, may be required to stabilise CMG proteins on the chromatin. However, through *in vitro* study of yeast replication initiation factors, Sld2 was observed to interact directly with the Mcm2-7 hexamer and strikingly Sld2 phosphorylation had no impact on this interaction [250]. This indicates that, outside of the context of the chromatin in yeast, phosphorylation of DONSON's proposed homologue is not required for its interaction with the hexamer. Although we haven't assessed direct interactions between our phosphorylation mutant DONSON (10A) and any components of the Mcm2-7 hexamer on chromatin, we have noted that the mutant can still bind chromatin. In contrast, upon addition of the 10A mutant, neither GINS nor Cdc45 can be loaded onto chromatin as efficiently. This indicates that DONSON's phosphorylation is likely more important for retention of GINS and Cdc45 in the CMG than for DONSON's recruitment to the chromatin. With Mcm7, whilst we still observe some inhibition of replication it is not as severe as in the absence of DONSON, indicating that there is still some fork progression with this mutant. Therefore, this suggests that in the presence of our phosphorylation mutant, a limited number of origins are able to fire. As a result, the forks likely replicate further along the DNA than usual, as indicated by the inefficient loading of Cdc45 and GINS. Collectively, we are still observing dormant origins being dissociated from the chromatin, as well as DNA synthesis with this mutant, but this is likely occurring with many fewer forks than normal. Therefore, this suggests that phosphorylation of DONSON plays a critical role in the efficiency of replication. Of note, DONSON's phosphorylation may therefore impact genome stability, as with fewer forks supporting replication, there will be a higher impact on the ability of replication to proceed if the forks stall.

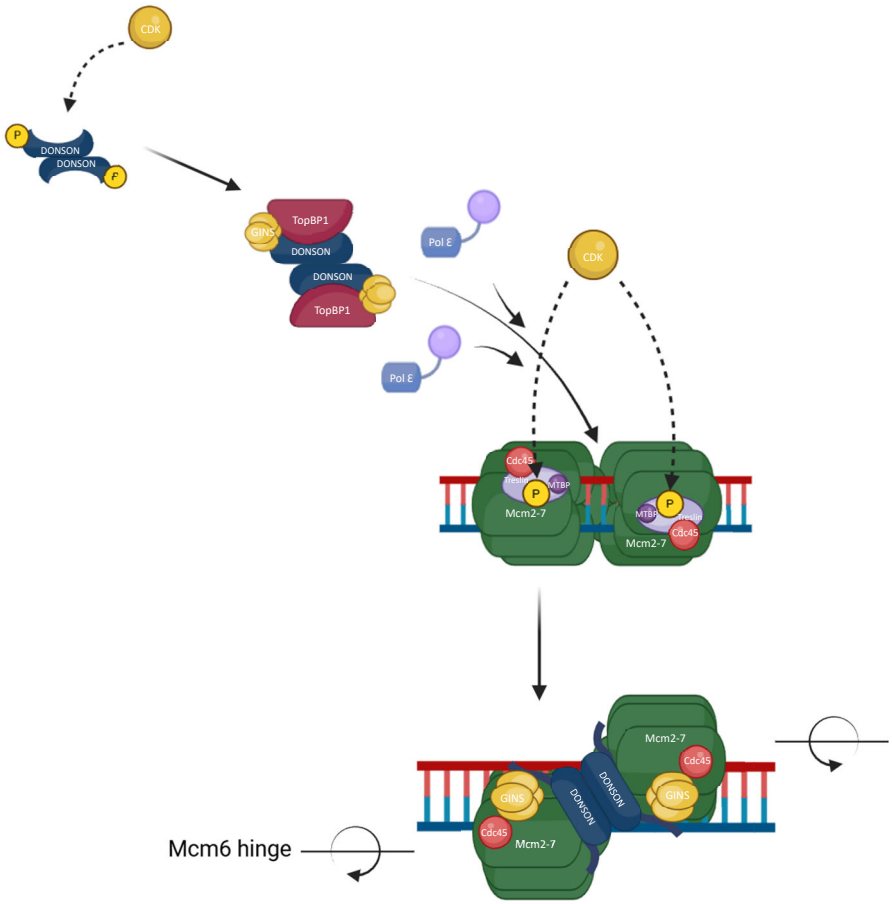


Figure 4.1.1 Schematic of how we propose DONSON supports CMG formation on chromatin in *Xenopus laevis* egg extract. Following Mcm2-7 double hexamer loading, DDKs phosphorylate MCM subunits. This promotes recruitment of two Treslin-MTBP-Cdc45 complexes (not shown). CDKs phosphorylate Treslin, and also phosphorylate DONSON which allows its interaction with TopBP1. Collectively, a DONSON dimer supporting by two TopBP1 molecules bring GINS, and potentially Pol Ɛ to the double hexamer for CMG formation. Once here, DONSON stabilises the double helicase and promotes misalignment of the two hexamers to promote origin DNA melting. Adapted from [190].

During this project, we have also identified that upon dimeric DONSON's binding of the Mcm2-7 hexamer and formation of the active helicase, there is a shift in the alignment of the two Mcm2-7 central channels in *Xenopus* egg extract [190]. This realignment of the hexameric channels puts the CMG in a position beneficial for duplex DNA melting. Likewise, in yeast, a rotation of the hexameric core upon CMG formation is observed which, as a result, enables duplex DNA melting and separation of the double hexamer [87]. In yeast and *Xenopus* egg extract, Mcm10 is recruited to support the complete separation of the two CMGs at each origin, allowing for lagging strand ejection and for the helicases to cross paths [93], [230]. The binding of ATP to the hexamer that occurs prior to Mcm10 function is coupled to the release of several initiation factors, including Dpb11, Sld3-7 and Sld2 from the DNA [86]. In addition, Sld3 has been observed to directly bind to Mcm2 in yeast, the site at which Mcm10 is observed to bind the

hexamer, and given that the binding site sequence is conserved in Treslin, it is likely that this is how Treslin contacts the hexamer in higher eukaryotes also [230], [231]. Furthermore, DONSON has also been observed to dissociate from the chromatin rapidly following recruitment of GINS [197]. This removal appears to be somewhat dependent on RecQL4, which has interestingly been observed to be required for the binding of MCM10 to the CMG in human cells [197], [232]. Collectively therefore, this suggests that the initiation factors required for Cdc45 and GINS recruitment to the chromatin must be removed prior to, or upon, the binding of Mcm10 in yeast and higher eukaryotes. However, Mcm10 has been shown not to actually be essential for DNA replication in higher eukaryotes and dual depletion of MCM-10 and DNSN-1 in worms is synthetic lethal [187], [196]. This therefore indicates that DONSON and Mcm10 may function in alternative pathways to both support duplex DNA unwinding. DONSON's ability to misalign the hexamers may play an integral role in this and is perhaps sufficient to support further DNA unwinding in the absence of Mcm10. Interestingly, upon DONSON depletion we actually observed a reduction of Mcm10 binding to chromatin; notably however, we have not assessed Mcm10 chromatin binding in the presence of our 10A phosphorylation mutant. This indicates that both DONSON's binding to, and removal from, the chromatin must be tightly regulated as it impacts downstream initiation events. Interestingly, dephosphorylation of both Sld3 and Sld2 is necessary for the two proteins to dissociate from chromatin and to support efficient DNA replication [126]. Therefore, it may be that DONSON's phosphorylation is required to maintain its interaction at the CMG and upon its dephosphorylation it is removed from the chromatin. Whilst we haven't specifically observed a change in the efficiency of removal of our 10A mutant compared to the WT from the chromatin, we have not tested timepoints at which we have observed removal of either our WT or 10A. As we have observed the requirement for DONSON's phosphorylation to stabilise the CMG on the chromatin, it may be that phosphorylation is actually required to stabilise the structure and allow the misalignment of the hexameric channels. Therefore, it may also be required to support Mcm10 recruitment. Collectively, this requires further investigation.

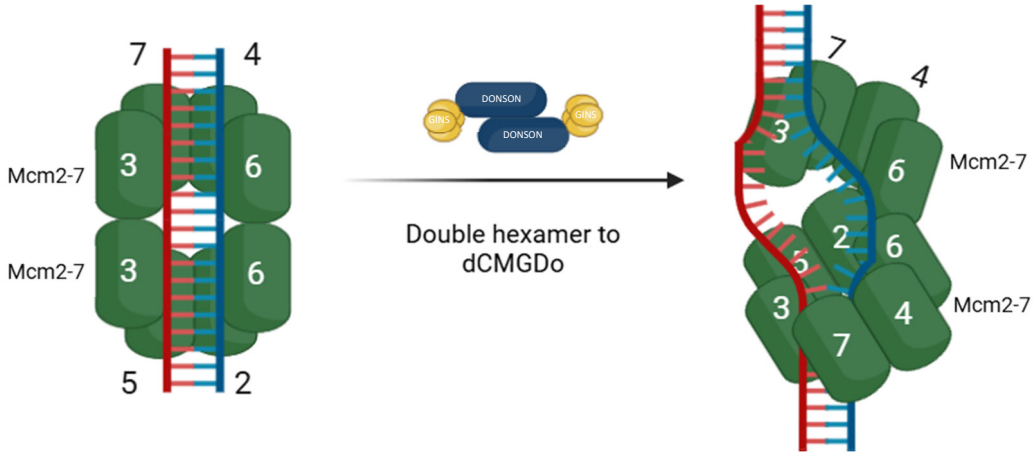


Figure 4.1.2 Schematic of how we propose DONSON supports DNA melting in *Xenopus laevis* egg extract. The double hexamer (here shown without the 'front' two subunits for simplicity) is in a head-to-head formation bound to duplex DNA. Upon dimeric DONSON (dark blue with yellow GINS complex) binding to the double hexamer, and supporting CMG formation, each hexamer rotates around the DNA via Mcm6 positioning the Mcm2 of each hexamer to occlude the central channel of the opposite hexamer and promote origin melting. Adapted from [190].

Collectively, the results from my thesis and the current published data, lead me to speculate as to the scope of DONSON's functions during DNA replication in higher eukaryotes. If DONSON functions to recruit GINS to Mcm2-7 hexamers, support initial DNA melting and is also required to stabilise forks and function within the ATR stress response pathway, what determines the role in which DONSON functions at any one time? Having observed extensive phosphorylation across the protein, as well as it being required to support initiation, we propose that DONSON is largely regulated via phosphorylation. However, the question as to which of these proposed roles for DONSON we think phosphorylation is responsible for regulating is of note. TopBP1 is another protein with roles in both DNA replication initiation and in the regulation of the ATR stress response. It is of particular interest to me that TopBP1 and DONSON are both involved in replication initiation and in activation of the ATR stress response and possibly suggests that TopBP1 may act as a scaffold for DONSON in both situations. Although not reported in my thesis, I have observed that DONSON is phosphorylated by ATR and that this site appears to be important for regulation of origin firing. Likewise, by mass spec, I have observed DONSON interacting with the ATR activator ETAA1. TopBP1 contains many phosphorylation-specific interacting domains and its function is largely regulated via the phosphorylation of its interaction partners in a cell cycle-dependent manner [251]. Likewise, it is thought that TopBP1 is regulated through subcellular localisation or protein concentration – of both TopBP1 itself, and its interaction partners [252]. As DONSON does not possess any enzymatic activity or

known PTM-specific interaction domains, it is likely that it is regulated through subcellular localisation and via partners that do. Therefore, phosphorylation of DONSON by CDK and ATR is likely one way in which DONSON's roles throughout the cell cycle are regulated.

Through gel filtration of cytoplasmic extract and NPE, we have observed DONSON primarily existing within complexes of very large molecular weight. Interestingly, we also observe TopBP1 primarily running within the same eluted fractions, although no strong direct interaction with DONSON was observed in extract. Notably, TopBP1 has been observed to assemble into phase-separated condensates of high molecular weight to amplify ATR activation [101], [243]. Therefore, it may be that DONSON also exists in such complexes due to its role in the stress response, with TopBP1 acting as its scaffold [253]. Upon DONSON depletion from the extract, we observed that TopBP1 is retained at this high molecular weight. However, we also observed that TopBP1 can be recruited to chromatin upon DONSON depletion. This may therefore suggest that, just as DONSON is not required for TopBP1's chromatin recruitment, TopBP1's existence in these multimeric complexes is not dependent on DONSON, but rather the other way around. Given that we primarily observe DONSON in these high molecular weight complexes, I hypothesise that DONSON is principally retained here and requires CDK-dependent phosphorylation to be released to achieve its replication initiation functions. Furthermore, that DONSON's CDK phosphorylation is required to release TopBP1, and that Treslin's phosphorylation is further required for TopBP1 to support active helicase formation on the chromatin.

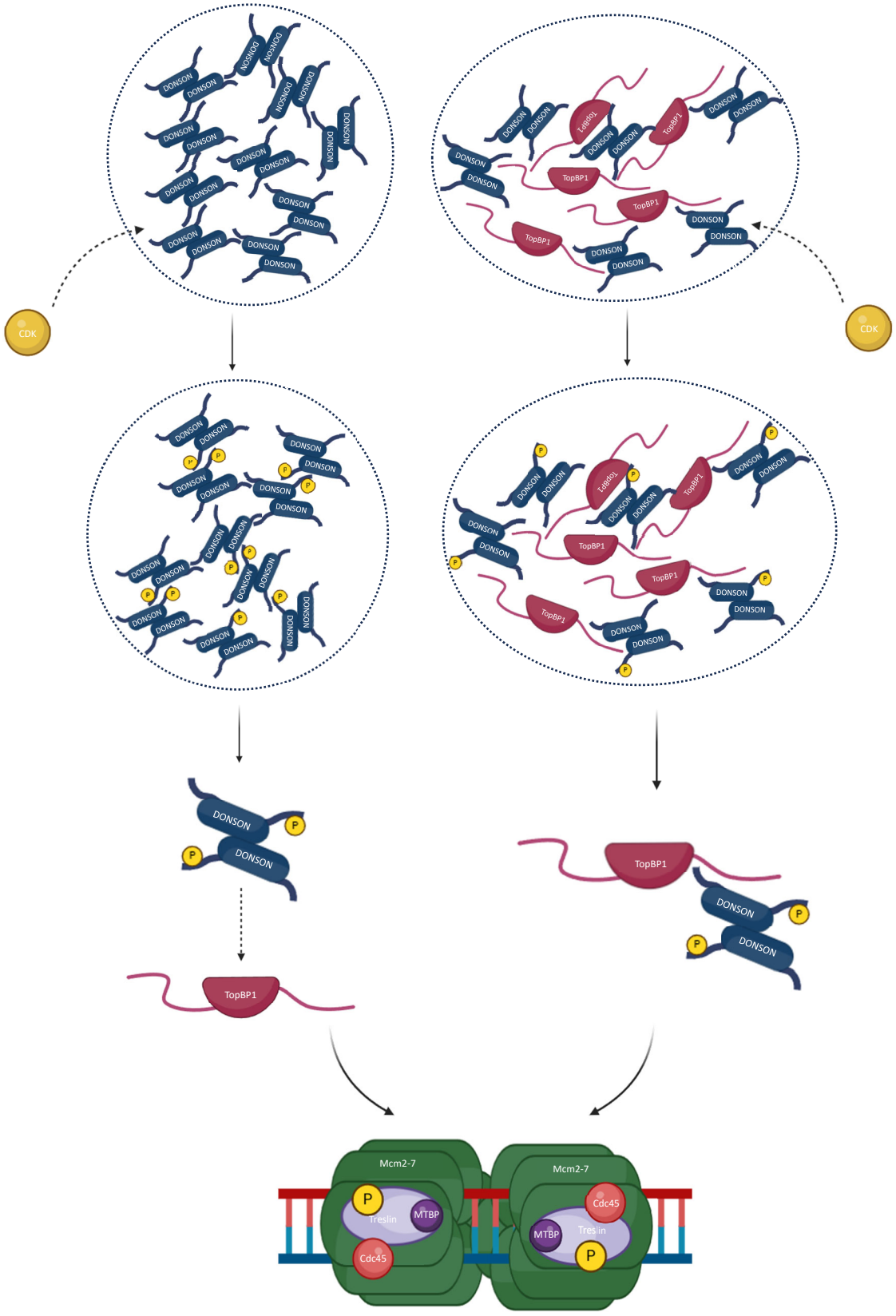


Figure 4.1.3 Schematic of how DONSON may be released from liquid-liquid phase separated multimeric complexes in *Xenopus laevis* egg extract. DONSON (shown in dark blue) may exist in large multimeric complexes either alone or in combination with another, likely scaffolding, protein (here TopBP1 is presented, in dark pink). CDK-dependent phosphorylation of DONSON will alter the charge or formation of the protein causing the multimeric bonds to be broken and the protein released, either alone or in combination with TopBP1. Dimeric DONSON would then be available to be recruited to DNA during initiation for CMG formation. Adapted from [253].

4.2 Relevance of this Research

Identifying DONSON's essential role in DNA replication initiation provides insight into the molecular mechanism of the process in higher eukaryotes. This therefore enables insight into how the process can be regulated and manipulated for the potential treatment of diseases in which replication is either under or over stimulated and the cell cycle is perturbed. I have been able to preliminarily investigate the potential of blocking DONSON using a peptide of the N-terminus of the protein and have seen some promising results blocking both CMG assembly on the chromatin and DNA replication levels. Based on this data, future work could provide insight into the potential of the N-terminal lactam peptide in downregulating replication and cell cycle progression. However, it would also be interesting to identify how the remaining patient mutations impact DONSON at the molecular level. In particular, identifying a separation of function mutant would be beneficial for the assessment of DONSON's potential multifunctional roles within replication. Additionally, investigating the way in which DONSON can be regulated has the potential to shed new light on further ways in which it can be manipulated to either promote or limit efficient DNA replication in patients.

4.3 Future Questions to Address

Having established that DONSON is required for functions in addition to its stress response role, the way in which it is regulated to achieve each of its purposes needs to be established. With data suggesting that both CDK- and ATR-dependent phosphorylation of DONSON affects its function, we suggest that phosphorylation may at least in part regulate DONSON. Therefore, future research should focus on which roles of DONSON its phosphorylation impacts. The particular roles, we suggest it may affect, and therefore that should be investigated are as follows:

1. DONSON's retention in a large multimeric complex, either self-interacting or with TopBP1.

It is of interest to determine whether CDK phosphorylation is sufficient to release DONSON from these large complexes. In particular, whether *in vitro* phosphorylated DONSON, when added into the extract, can still form these large complexes, when assessed by gel filtration. Likewise, assessing whether TopBP1

depletion abrogates DONSON's ability to form these large complexes, could shed light on whether TopBP1 acts as a scaffold for DONSON within these high molecular weight structures.

2. DONSON's release from such a multimeric complex specifically to enable its interaction with initiation factors and support DNA replication initiation.

If we do observe a shift to a lower molecular weight of DONSON in extract upon phosphorylation with commercial CDKs, testing whether our 10A mutant would be resistant to these conditions would further support this finding. Likewise, *in vitro* phosphorylating DONSON and assessing its interactions with TopBP1 and GINS, via IP, on chromatin could provide insight into the extent to which phosphorylation is required for the interactions likely required for its recruitment to chromatin.

3. Stabilisation of a vertebrate pre-LC, and alteration of the structure of the factors for efficient positioning on the chromatin, as well as stabilisation of duplex CMGs at origins.

It is of particular interest to assess our 10A mutant by cryo-EM, to determine whether the interactions we have previously observed with the endogenous protein are altered. In particular, assessing whether the interactions proposed by AlphaFold are reliable, as the suggestion that phosphorylation of DONSON may be important for stabilisation of the pre-LC and positioning the factors for efficient recruitment to the chromatin is significant. Likewise, assessing the interactions between 10A DONSON and CMG components – particularly Sld5 and Mcm3 – on chromatin, may shed light on whether phosphorylation is required to stabilise the CMG on chromatin.

4. DONSON's removal from the chromatin following effective CMG stabilisation and hexamer misalignment, to allow for duplex DNA unwinding.

Assessing the impact of the 10A mutant on the interactions of RecQL4, TopBP1 and Mcm10 with the pre-initiation complex, could indicate whether DONSON's phosphorylation is important for origin unwinding. In particular, whether it is required to support the removal of initiation factors from the chromatin, including its own, as well as to support the recruitment of DNA unwinding factors, particularly Mcm10. Additionally, as mentioned above, performing a chromatin isolation experiment with both our WT and

10A mutant at later timepoints, may indicate whether our phosphorylation mutant is removed from chromatin with altered efficiency, compared to our WT.

As well as continuing this work and addressing these questions in *Xenopus laevis* egg extract, it is important to translate our data into human cells, particularly if the view is to determine whether we can extend these findings into a therapeutic approach. Although our dimeric recombinant protein is able to efficiently support replication and CMG formation, a more reliable system for testing the function of DONSON may be to use human cells with ectopic expression, that can be postranslationally modified in an analogous way to the endogenous protein. Therefore, specifically to assess the impact of CDK-dependent phosphorylation of DONSON, human cells may be a better model. Over the past few months, we have been attempting to reconstitute previous data from other labs, showing that DONSON is required to support DNA replication in human cells [152]. To assess this, we have been using DONSON-specific degron-tagged HCT116 human cells (Figure 4.3.1). We have observed that degradation of DONSON promisingly results in DNA content remaining at 2N (data not shown). However, we have not yet reproducibly observed rescue of DNA replication upon transfection with WT DONSON. In the future, we propose assessing the impact of CDK-dependent phosphorylation of DONSON on DNA replication and interactions of the proposed pre-LC components in human cells. Additionally, it would be particularly interesting to assess whether DONSON can be observed in liquid-like condensates within human cells, and whether these contain any other factors, particularly TopBP1 and those involved in the ATR stress response.

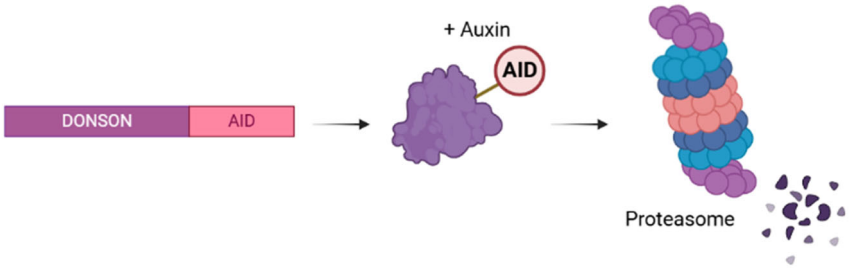


Figure 4.3.1 DONSON degradation in human cells. Schematic depicting HCT116 human colon cancer cells carrying the auxin-inducible degron (AID) tagged DONSON. Cells are cultured and cell cycling is established. Auxin is applied and proteasomal degradation of the endogenous DONSON occurs, resulting in cells deplete of DONSON ready for mutant testing.

5 References

[1] W. P. Vermeij, J. H. J. Hoeijmakers, and J. Pothof, “Genome Integrity in Aging: Human Syndromes, Mouse Models, and Therapeutic Options,” *Annu. Rev. Pharmacol. Toxicol.*, vol. 56, pp. 427–445, 2016, doi: 10.1146/annurev-pharmtox-010814-124316.

[2] F. Crick and J. Watson, “Molecular Structure of Nucleic Acids: A Structure for Deoxyribose Nucleic Acid,” *Nature*, vol. 171, no. 4356, pp. 737–738, 1953.

[3] M. Meselson and F. W. Stahl, “ The replication of DNA in Escherichia coli ,” *Proc. Natl. Acad. Sci.*, vol. 44, no. 7, pp. 671–682, 1958, doi: 10.1073/pnas.44.7.671.

[4] S. Priego Moreno, R. M. Jones, and A. Gambus, “DNA Replication and Cell Cycle Control,” in *DNA Repair and Replication*, 2018, p. 16.

[5] D. Cortez, “Replication-Coupled DNA Repair,” *Mol. Cell*, vol. 74, no. 5, pp. 866–876, 2019, doi: 10.1016/j.molcel.2019.04.027.

[6] I. Tsirkas *et al.*, “Transcription-replication coordination revealed in single live cells,” *Nucleic Acids Res.*, vol. 50, no. 4, pp. 2143–2156, 2022.

[7] C. Alabert and A. Groth, “Chromatin replication and epigenome maintenance,” *Nat. Rev.*, vol. 13, no. 2, pp. 153–167, 2012.

[8] Y. J. Machida, J. L. Hamlin, and A. Dutta, “Right place, right time, and only once: Replication initiation in metazoans,” *Cell*, vol. 123, no. 1, pp. 13–24, 2005, doi: 10.1016/j.cell.2005.09.019.

[9] B. Ekundayo and F. Bleichert, “Origins of DNA replication,” *PLoS Genet.*, vol. 15, no. 9, pp. 1–21, 2019.

[10] A. C. Leonard and M. Mechali, “DNA Replication Origins,” *Cold Spring Harb. Perspect. Biol.*, vol. 5, no. 10, pp. 1–17, 2013.

[11] S. E. Moyer, P. W. Lewis, and M. R. Botchan, “Isolation of the Cdc45/Mcm2–7/GINS (CMG) complex, a candidate for the eukaryotic DNA replication fork helicase,” *Proc. Natl. Acad. Sci.*, vol. 103, no. 27, pp. 10236–10241, Jul. 2006, doi: 10.1073/pnas.0602400103.

[12] M. L. Bochman and A. Schwacha, “The Mcm Complex: Unwinding the Mechanism of a Replicative Helicase,” *Microbiol. Mol. Biol. Rev.*, vol. 73, no. 4, pp. 652–683, 2009.

[13] H. P. Nasheuer, A. M. Meaney, T. Hulshoff, I. Thiele, and N. O. Onwubiko, “Replication Protein A, the Main Eukaryotic Single-Stranded DNA Binding Protein, a Focal Point in Cellular DNA Metabolism,” *Int. J. Mol. Sci.*, vol. 25, no. 1, 2024, doi: 10.3390/ijms25010588.

[14] Z. Yuan, R. Georgescu, N. Y. Yao, O. Yurieva, M. E. O’Donnell, and H. Li, “Mechanism of PCNA loading by Ctf18-RFC for leading-strand DNA synthesis,” *Science (80-.).*, vol. 385, no. 6708, Aug. 2024, doi: 10.1126/science.adk5901.

[15] G.-L. Moldovan, B. Pfander, and S. Jentsch, “PCNA, the Maestro of the Replication Fork,” *Cell*, vol. 129, no. 4, pp. 665–679, 2007, [Online]. Available: 10.1016/j.cell.2007.05.003

[16] V. Aria and J. T. P. Yeeles, “Mechanism of Bidirectional Leading-Strand Synthesis Establishment at Eukaryotic DNA Replication Origins,” *Mol. Cell*, vol. 73, no. 2, pp. 199-211.e10, Jan. 2019, doi: 10.1016/j.molcel.2018.10.019.

[17] T. Moiseeva, B. Hood, S. Schamus, M. J. O’Connor, T. P. Conrads, and C. J. Bakkenist, “ATR kinase inhibition induces unscheduled origin firing through a Cdc7-dependent association between GINS and And-1,” *Nat. Commun.*, vol. 8, no. 1, 2017, doi: 10.1038/s41467-017-01401-x.

[18] L. Pellegrini, “The Pol α-Primase Complex,” in *The Eukaryotic Replisome: a Guide to Protein Structure and Function*, 62nd ed.Springer, Dordrecht, 2012, pp. 157–169. doi: 10.1007/978-94-007-4572-8_9.

[19] J. J. Champoux, “DNA TOPOISOMERASES : Structure, Function and Mechanism,” *Annu. Rev. Biochem.*, vol. 70, pp. 369–413, 2001.

[20] H. P. Patel *et al.*, "DNA supercoiling restricts the transcriptional bursting of neighboring eukaryotic genes," *Mol. Cell*, vol. 83, no. 10, pp. 1573-1587.e8, 2023, doi: 10.1016/j.molcel.2023.04.015.

[21] J. M. Berger, S. J. Gamblin, S. C. Harrison, and J. C. Wang, "Structure and mechanism of DNA topoisomerase II," *Nature*, vol. 379, pp. 225–232, 1996.

[22] H. Zhao *et al.*, "Nucleosome Assembly and Disassembly in vitro Are Governed by Chemical Kinetic Principles," *Front. Cell Dev. Biol.*, vol. 9, Oct. 2021, doi: 10.3389/fcell.2021.762571.

[23] M. J. McCauley *et al.*, "Human FACT subunits coordinate to catalyze both disassembly and reassembly of nucleosomes," *Cell Rep.*, vol. 41, no. 13, p. 111858, Dec. 2022, doi: 10.1016/j.celrep.2022.111858.

[24] W. Zhang, J. Feng, and Q. Li, "The replisome guides nucleosome assembly during DNA replication," *Cell Biosci.*, vol. 10, no. 1, p. 37, Dec. 2020, doi: 10.1186/s13578-020-00398-z.

[25] M. Maric, T. Maculins, G. De Piccoli, and K. Labib, "Cdc48 and a ubiquitin ligase drive disassembly of the CMG helicase at the end of DNA replication," *Science (80-.).*, vol. 346, no. 6208, pp. 1253596-1–12, 2014.

[26] J. M. Dewar and J. Walter, "Mechanisms of DNA replication termination," *Nat. Rev. Mol. Cell Biol.*, vol. 18, no. 8, pp. 507–516, 2017.

[27] J. M. Dewar, M. Budzowska, and J. Walter, "The mechanism of DNA replication termination in vertebrates," *Nature*, vol. 525, no. 7569, pp. 345–350, 2015.

[28] C. Neylon, A. V Kralicek, T. M. Hill, and N. E. Dixon, "Replication Termination in Escherichia coli: Structure and Antihelicase Activity of the Tus-Ter Complex," *Microbiol. Mol. Biol. Rev.*, vol. 69, no. 3, pp. 501–526, 2005.

[29] J. U. Dimude, S. L. Midgley-Smith, M. Stein, and C. J. Rudolph, "Replication Termination: Containing Fork Fusion-Mediated Pathologies in Escherichia coli," *Genes (Basel).*, vol. 7, no. 40, pp. 1–22, 2016.

[30] R. Bailey, S. Priego Moreno, and A. Gambus, "Termination of DNA replication forks: 'Breaking up is hard to do,'" *Nucleus*, vol. 6, no. 3, pp. 187–196, 2015.

[31] Y.-H. Chen *et al.*, "Transcription shapes DNA replication initiation and termination in human cells," *Nat. Struct. Mol. Biol.*, vol. 26, no. 1, pp. 67–77, 2019.

[32] A. Gambus, "Termination of Eukaryotic Replication Forks," in *DNA Replication: From Old Principles to New Discoveries*, 2018, pp. 163–187.

[33] T. D. Deegan, J. Baxter, M. A. Ortiz Bazan, J. T. P. Yeeles, and K. P. M. Labib, "Pif1-Family Helicases Support Fork Convergence during DNA Replication Termination in Eukaryotes," *Mol. Cell*, vol. 74, pp. 231–244, 2019.

[34] Y. Daigaku *et al.*, "A global profile of replicative polymerase usage," *Nat. Struct. Mol. Biol.*, vol. 22, no. 3, pp. 192–198, 2015.

[35] R. Sonneville *et al.*, "CUL-2LRR1 and UBXN-3 drive replisome disassembly during DNA replication termination and mitosis," *Nat. Cell Biol.*, vol. 19, no. 5, pp. 468–479, 2017.

[36] T. D. Deegan, P. P. Mukherjee, R. Fujisawa, C. Polo Rivera, and K. Labib, "CMG helicase disassembly is controlled by replication fork DNA, replisome components and a ubiquitin threshold," *Elife*, vol. 9, no. e60371, pp. 1–33, 2020.

[37] E. Low, G. Chistol, M. S. Zaher, O. V Kochenova, and J. C. Walter, "The DNA replication fork suppresses CMG unloading from chromatin before termination," *GENES Dev.*, vol. 34, pp. 1534–1545, 2020.

[38] M. Jenkyn-Bedford *et al.*, "A Conserved Mechanism for Regulating Replisome Disassembly in Eukaryotes," *Nature*, 2021, [Online]. Available: https://doi.org/10.1038/s41586-021-04145-3

[39] S. Priego Moreno and A. Gambus, "Mechanisms of eukaryotic replisome disassembly," *Biochem. Soc. Trans.*, vol. 48, pp. 823–836, 2020.

[40] A. Franz *et al.*, “Chromatin-associated degradation is defined by UBXN-3/FAF1 to safeguard DNA replication fork progression,” *Nat. Commun.*, vol. 7, no. 11593, pp. 1–2, 2016.

[41] A. Costa and J. F. X. Diffley, “The Initiation of Eukaryotic DNA Replication,” *Annu. Rev. Biochem.*, vol. 91, pp. 107–131, 2022, doi: 10.1146/annurev-biochem-072321-110228.

[42] F. Chang, J. F. Theis, J. Miller, C. A. Nieduszynski, C. S. Newlon, and M. Weinreich, “ Analysis of Chromosome III Replicators Reveals an Unusual Structure for the ARS318 Silencer Origin and a Conserved WTW Sequence within the Origin Recognition Complex Binding Site ,” *Mol. Cell. Biol.*, vol. 28, no. 16, pp. 5071–5081, 2008, doi: 10.1128/mcb.00206-08.

[43] S. Hoshina *et al.*, “Human origin recognition complex binds preferentially to G-quadruplex-preferable RNA and single-stranded DNA,” *J. Biol. Chem.*, vol. 288, no. 42, pp. 30161–30171, 2013, doi: 10.1074/jbc.M113.492504.

[44] M. Prioleau and D. M. Macalpine, “DNA replication origins — where do we begin ?,” pp. 1683–1697, 2016, doi: 10.1101/gad.285114.116.ical.

[45] N. Sugimoto *et al.*, “Cdt1-binding protein GRWD1 is a novel histone-binding protein that facilitates MCM loading through its influence on chromatin architecture,” *Nucleic Acids Res.*, vol. 43, no. 12, pp. 5898–5911, 2015, doi: 10.1093/nar/gkv509.

[46] N. Sugimoto, T. Yugawa, M. Iizuka, T. Kiyono, and M. Fujita, “Chromatin remodeler sucrose nonfermenting 2 homolog (SNF2H) is recruited onto DNA replication origins through interaction with Cdc10 protein-dependent transcript 1 (Cdt1) and promotes pre-replication complex formation,” *J. Biol. Chem.*, vol. 286, no. 45, pp. 39200–39210, 2011, doi: 10.1074/jbc.M111.256123.

[47] B. Miotto and K. Struhl, “HBO1 histone acetylase is a coactivator of the replication licensing factor Cdt1,” *Genes Dev.*, vol. 22, no. 19, pp. 2633–2638, 2008, doi: 10.1101/gad.1674108.

[48] H. Talasz, H. H. Lindner, B. Sarg, and W. Helliger, “Histone H4-Lysine 20 Monomethylation Is Increased in Promoter and Coding Regions of Active Genes and Correlates with Hyperacetylation,” *J. Biol. Chem.*, vol. 280, no. 46, pp. 38814–38822, 2005.

[49] J. W. Semple *et al.*, “An essential role for Orc6 in DNA replication through maintenance of pre-replicative complexes,” *EMBO J.*, vol. 25, no. 21, pp. 5150–5158, 2006, doi: 10.1038/sj.emboj.7601391.

[50] X. Feng, Y. Noguchi, M. Barbon, B. Stillman, C. Speck, and H. Li, “The structure of ORC–Cdc6 on an origin DNA reveals the mechanism of ORC activation by the replication initiator Cdc6,” *Nat. Commun.*, vol. 12, no. 1, pp. 1–12, 2021, doi: 10.1038/s41467-021-24199-1.

[51] S. A. Samel *et al.*, “A unique DNA entry gate serves for regulated loading of the eukaryotic replicative helicase MCM2-7 onto DNA,” *Genes Dev.*, vol. 28, no. 15, pp. 1653–1666, 2014, doi: 10.1101/gad.242404.114.

[52] Z. Yuan *et al.*, “Structural basis of Mcm2-7 replicative helicase loading by ORC-Cdc6 and Cdt1,” *Nat. Struct. Mol. Biol.*, vol. 24, no. 3, pp. 316–324, 2017, doi: 10.1038/nsmb.3372.

[53] A. Fernández-Cid *et al.*, “An ORC/Cdc6/MCM2-7 Complex Is Formed in a Multistep Reaction to Serve as a Platform for MCM Double-Hexamer Assembly,” *Mol. Cell*, vol. 50, no. 4, pp. 577–588, May 2013, doi: 10.1016/j.molcel.2013.03.026.

[54] S. Ticau, L. J. Friedman, N. A. Ivica, J. Gelles, and S. P. Bell, “Single-molecule studies of origin licensing reveal mechanisms ensuring bidirectional helicase loading,” *Cell*, vol. 161, no. 3, pp. 513–525, 2015, doi: 10.1016/j.cell.2015.03.012.

[55] I.-E. Symeonidou, S. Taraviras, and Z. Lygerou, “Control over DNA replication in time and space,” *FEBS Lett.*, vol. 586, no. 18, pp. 2803–2812, 2012.

[56] A. Errico and V. Costanzo, “Differences in the DNA replication of unicellular eukaryotes and metazoans: known unknowns,” *EMBO Rep.*, vol. 11, no. 4, pp. 270–278, Apr. 2010, doi: 10.1038/embor.2010.27.

[57] A. Li and J. J. Blow, “Cdt1 downregulation by proteolysis and geminin inhibition prevents DNA re-

replication in Xenopus," *EMBO J.*, vol. 24, no. 2, pp. 395–404, Jan. 2005, doi: 10.1038/sj.emboj.7600520.

[58] V. Q. Nguyen, C. Co, and J. J. Li, "Cyclin-dependent kinases prevent DNA re-replication through multiple mechanisms," *Nature*, vol. 411, no. 6841, pp. 1068–1073, Jun. 2001, doi: 10.1038/35082600.

[59] J. T. P. Yeeles, T. D. Deegan, A. Janska, A. Early, and J. F. X. Diffley, "Regulated eukaryotic DNA replication origin firing with purified proteins," *Nature*, vol. 519, no. 7544, pp. 431–435, 2015, doi: 10.1038/nature14285.

[60] J. F. Greiwe *et al.*, "Structural mechanism for the selective phosphorylation of DNA-loaded MCM double hexamers by the Dbf4-dependent kinase," *Nat. Struct. Mol. Biol.*, vol. 29, no. 1, pp. 10–20, Jan. 2022, doi: 10.1038/s41594-021-00698-z.

[61] S. Tanaka, R. Nakato, Y. Katou, K. Shirahige, and H. Araki, "Origin association of Sld3, Sld7, and Cdc45 proteins is a key step for determination of origin-firing timing," *Curr. Biol.*, vol. 21, no. 24, pp. 2055–2063, 2011, doi: 10.1016/j.cub.2011.11.038.

[62] H. Itou, Y. Shirakihara, and H. Araki, "The quaternary structure of the eukaryotic DNA replication proteins Sld7 and Sld3," *Acta Crystallogr. Sect. D Biol. Crystallogr.*, vol. 71, no. 8, pp. 1649–1656, Aug. 2015, doi: 10.1107/S1399004715010457.

[63] L. De Jesús-Kim, L. J. Friedman, M. Lõoke, C. K. Ramsoomair, J. Gelles, and S. P. Bell, "DDK regulates replication initiation by controlling the multiplicity of Cdc45-GINS binding to Mcm2-7," *Elife*, vol. 10, Feb. 2021, doi: 10.7554/eLife.65471.

[64] D. Boos *et al.*, "Regulation of DNA replication through Sld3-Dpb11 interaction is conserved from yeast to humans," *Curr. Biol.*, vol. 21, no. 13, pp. 1152–1157, 2011, doi: 10.1016/j.cub.2011.05.057.

[65] R. C. Alver, G. S. Chadha, P. J. Gillespie, and J. J. Blow, "Reversal of DDK-Mediated MCM Phosphorylation by Rif1-PP1 Regulates Replication Initiation and Replisome Stability Independently of ATR/Chk1," *Cell Rep.*, vol. 18, no. 10, pp. 2508–2520, 2017, doi: 10.1016/j.celrep.2017.02.042.

[66] S. Tanaka, Y. Komeda, T. Umemori, Y. Kubota, H. Takisawa, and H. Araki, "Efficient Initiation of DNA Replication in Eukaryotes Requires Dpb11/TopBP1-GINS Interaction," *Mol. Cell. Biol.*, vol. 33, no. 13, pp. 2614–2622, Jul. 2013, doi: 10.1128/MCB.00431-13.

[67] P. Zegerman and J. F. X. Diffley, "Phosphorylation of Sld2 and Sld3 by cyclin-dependent kinases promotes DNA replication in budding yeast," *Nature*, vol. 445, pp. 281–285, 2007.

[68] S. Tanaka, T. Umemori, K. Hirai, S. Muramatsu, Y. Kamimura, and H. Araki, "CDK-dependent phosphorylation of Sld2 and Sld3 initiates DNA replication in budding yeast," *Nature*, vol. 445, no. 7125, pp. 328–332, Jan. 2007, doi: 10.1038/nature05465.

[69] S. Muramatsu, K. Hirai, Y.-S. S. Tak, Y. Kamimura, and H. Araki, "CDK-dependent complex formation between replication proteins Dpb11, Sld2, Pol ε, and GINS in budding yeast," *Genes Dev.*, vol. 24, no. 6, pp. 602–612, Mar. 2010, doi: 10.1101/gad.1883410.

[70] S. Waga, T. Masuda, H. Takisawa, and A. Sugino, "DNA polymerase ε is required for coordinated and efficient chromosomal DNA replication in Xenopus egg extracts," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 98, no. 9, pp. 4978–4983, 2001, doi: 10.1073/pnas.081088798.

[71] S. Vipat, D. Gupta, S. Jonchhe, H. Anderspuk, E. Rothenberg, and T. N. Moiseeva, "The non-catalytic role of DNA polymerase epsilon in replication initiation in human cells," *Nat. Commun.*, vol. 13, no. 1, pp. 1–15, 2022, doi: 10.1038/s41467-022-34911-4.

[72] M. N. Sangrithi *et al.*, "Initiation of DNA replication requires the RECQL4 protein mutated in Rothmund-Thomson syndrome," *Cell*, vol. 121, no. 6, pp. 887–898, Jun. 2005, doi: 10.1016/j.cell.2005.05.015.

[73] K. Matsuno, M. Kumano, Y. Kubota, Y. Hashimoto, and H. Takisawa, "The N-Terminal Noncatalytic Region of Xenopus RecQ4 Is Required for Chromatin Binding of DNA Polymerase α in the Initiation of DNA Replication," *Mol. Cell. Biol.*, vol. 26, no. 13, pp. 4843–4852, Jul. 2006, doi: 10.1128/MCB.02267-

05.

[74] A. Kumagai, A. Shevchenko, A. Shevchenko, and W. G. Dunphy, “Treslin collaborates with TopBP1 in triggering the initiation of DNA replication,” *Cell*, vol. 140, pp. 349–359, 2010.

[75] I. Ilves, T. Petojevic, J. J. Pesavento, and M. R. Botchan, “Activation of the MCM2-7 Helicase by Association with Cdc45 and GINS Proteins,” *Mol. Cell*, vol. 37, no. 2, pp. 247–258, 2010, doi: 10.1016/j.molcel.2009.12.030.

[76] P. Goswami *et al.*, “Structure of DNA-CMG-Pol epsilon elucidates the roles of the non-catalytic polymerase modules in the eukaryotic replisome,” *Nat. Commun.*, vol. 9, no. 1, p. 5061, Nov. 2018, doi: 10.1038/s41467-018-07417-1.

[77] A. Costa *et al.*, “The structural basis for MCM2–7 helicase activation by GINS and Cdc45,” *Nat. Struct. Mol. Biol.*, vol. 18, no. 4, pp. 471–477, Apr. 2011, doi: 10.1038/nsmb.2004.

[78] N. Li *et al.*, “Structure of the eukaryotic MCM complex at 3.8 Å,” *Nature*, vol. 524, no. 7564, pp. 186–191, Aug. 2015, doi: 10.1038/nature14685.

[79] F. Van Deursen, S. Sengupta, G. De Piccoli, A. Sanchez-Diaz, and K. Labib, “Mcm 10 associates with the loaded DNA helicase at replication origins and defines a novel step in its activation,” *EMBO J.*, vol. 31, no. 9, pp. 2195–2206, 2012, doi: 10.1038/emboj.2012.69.

[80] D. Mantiero, A. Mackenzie, A. Donaldson, and P. Zegerman, “Limiting replication initiation factors execute the temporal programme of origin firing in budding yeast,” *EMBO J.*, vol. 30, no. 23, pp. 4805–4814, Nov. 2011, doi: 10.1038/emboj.2011.404.

[81] M. Das, S. Singh, S. Pradhan, and G. Narayan, “MCM Paradox: Abundance of Eukaryotic Replicative Helicases and Genomic Integrity,” *Mol. Biol. Int.*, vol. 2014, pp. 1–11, Oct. 2014, doi: 10.1155/2014/574850.

[82] B. Al Ahmad Nachar and F. Rosselli, “Safeguarding DNA Replication: A Golden Touch of MiDAS and Other Mechanisms,” *Int. J. Mol. Sci.*, vol. 23, no. 19, p. 11331, Sep. 2022, doi: 10.3390/ijms231911331.

[83] D. McIntosh and J. J. Blow, “Dormant origins, the licensing checkpoint, and the response to replicative stresses,” *Cold Spring Harb. Perspect. Biol.*, vol. 4, no. 10, 2012, doi: 10.1101/cshperspect.a012955.

[84] N. Ratnayeke, Y. Baris, M. Chung, J. T. P. Yeeles, and T. Meyer, “CDT1 inhibits CMG helicase in early S phase to separate origin licensing from DNA synthesis,” *Mol. Cell*, vol. 83, no. 1, pp. 26-42.e13, Jan. 2023, doi: 10.1016/j.molcel.2022.12.004.

[85] A. R. Leman and E. Noguchi, “Local and global functions of Timeless and Tipin in replication fork protection,” *Cell Cycle*, vol. 11, no. 21, pp. 3945–3955, Nov. 2012, doi: 10.4161/cc.21989.

[86] M. E. Douglas *et al.*, “The mechanism of eukaryotic CMG helicase activation,” *Nature*, vol. 555, no. 7695, pp. 265–268, 2018, doi: 10.1038/nature25787.

[87] J. S. Lewis *et al.*, “Mechanism of replication origin melting nucleated by CMG helicase assembly,” *Nature*, vol. 606, no. 7916, pp. 1007–1014, Jun. 2022, doi: 10.1038/s41586-022-04829-4.

[88] M. Kanemaki and K. Labib, “Distinct roles for Sld3 and GINS during establishment and progression of eukaryotic DNA replication forks,” *EMBO J.*, vol. 25, no. 8, pp. 1753–1763, Apr. 2006, doi: 10.1038/sj.emboj.7601063.

[89] Y. Kamimura, H. Masumoto, A. Sugino, and H. Araki, “Sld2, Which Interacts with Dpb11 in Saccharomyces cerevisiae , Is Required for Chromosomal DNA Replication,” *Mol. Cell. Biol.*, vol. 18, no. 10, pp. 6102–6109, Oct. 1998, doi: 10.1128/MCB.18.10.6102.

[90] S. Tanaka and H. Araki, “Regulation of the initiation step of DNA replication by cyclin-dependent kinases,” *Chromosoma*, vol. 119, no. 6, pp. 565–574, Dec. 2010, doi: 10.1007/s00412-010-0291-8.

[91] S. A. de Munnik *et al.*, “Meier-Gorlin syndrome,” *Orphanet J. Rare Dis.*, vol. 10, no. 1, p. 114, Dec. 2015, doi: 10.1186/s13023-015-0322-x.

[92] R. Mayle, L. Langston, K. R. Molloy, D. Zhang, B. T. Chait, and M. E. O'Donnell, "Mcm10 has potent strand-annealing activity and limits translocase-mediated fork regression," *Proc. Natl. Acad. Sci.*, vol. 116, no. 3, pp. 798–803, Jan. 2019, doi: 10.1073/pnas.1819107116.

[93] G. S. Chadha, A. Gambus, P. J. Gillespie, and J. J. Blow, "Xenopus Mcm10 is a CDK-substrate required for replication fork stability," *Cell Cycle*, vol. 15, no. 16, pp. 2183–2195, Aug. 2016, doi: 10.1080/15384101.2016.1199305.

[94] A. R. Leman, C. Noguchi, C. Y. Lee, and E. Noguchi, "Human Timeless and Tipin stabilize replication forks and facilitate sister-chromatid cohesion," *J. Cell Sci.*, vol. 123, pp. 660–670, 2009.

[95] M. L. Kilkenny, A. C. Simon, J. Mainwaring, D. Wirthensohn, S. Holzer, and L. Pellegrini, "The human CTF4-orthologue AND-1 interacts with DNA polymerase a/primase via its unique C-terminal HMG box," *Open Biol.*, vol. 7, no. 170217, pp. 1–15, 2017.

[96] M. K. Zeman and K. A. Cimprich, "Causes and consequences of replication stress," *Nat. Cell Biol.*, vol. 16, no. 1, pp. 1–8, 2014.

[97] H. Gaillard, E. Herrera-Moyano, and A. Aguilera, "Transcription-Associated Genome Instability," *Chem. Rev.*, vol. 113, no. 11, 2013, [Online]. Available: https://pubs.acs.org/doi/10.1021/cr400017yhttps://pubs.acs.org/doi/10.1021/cr400017y

[98] H. Gaillard, T. García-Muse, and A. Aguilera, "Replication stress and cancer," *Nat. Rev. Cancer*, vol. 15, no. 5, pp. 276–280, 2015, doi: 10.1038/nrc3916.

[99] M. Pacek and J. C. Walter, "A requirement for MCM7 and Cdc45 in chromosome unwinding during eukaryotic DNA replication," *EMBO J.*, vol. 23, no. 18, pp. 3667–3676, 2004, doi: 10.1038/sj.emboj.7600369.

[100] T. S. Byun, M. Pacek, M. C. Yee, J. C. Walter, and K. A. Cimprich, "Functional uncoupling of MCM helicase and DNA polymerase activities activates the ATR-dependent checkpoint," *Genes Dev.*, vol. 19, no. 9, pp. 1040–1052, 2005, doi: 10.1101/gad.1301205.

[101] A. Kumagai, J. Lee, H. Y. Yoo, and W. G. Dunphy, "TopBP1 activates the ATR-ATRIP complex," *Cell*, vol. 124, no. 5, pp. 943–955, 2006, doi: 10.1016/j.cell.2005.12.041.

[102] P. Haahr *et al.*, "Activation of the ATR kinase by the RPA-binding protein ETAA1," *Nat. Cell Biol.*, vol. 18, no. 11, pp. 1196–1207, Nov. 2016, doi: 10.1038/ncb3422.

[103] K. Uto, D. Inoue, K. Shimuta, N. Nakajo, and N. Sagata, "Chk1, but not Chk2, inhibits Cdc25 phosphatases by a novel common mechanism," *EMBO J.*, vol. 23, no. 16, pp. 3386–3396, Aug. 2004, doi: 10.1038/sj.emboj.7600328.

[104] A. M. Woodward *et al.*, "Excess Mcm2–7 license dormant origins of replication that can be used under conditions of replicative stress," *J. Cell Biol.*, vol. 173, no. 5, pp. 673–683, Jun. 2006, doi: 10.1083/jcb.200602108.

[105] I. Elvers, F. Johansson, P. Groth, K. Erixon, and T. Helleday, "UV stalled replication forks restart by re-priming in human fibroblasts," *Nucleic Acids Res.*, vol. 39, no. 16, pp. 7049–7057, 2011, doi: 10.1093/nar/gkr420.

[106] N. Mailand, I. Gibbs-Seymour, and S. Bekker-Jensen, "Regulation of PCNA-protein interactions for genome stability," *Nat. Rev. Mol. Cell Biol.*, vol. 14, no. 5, pp. 269–282, 2013, doi: 10.1038/nrm3562.

[107] L. A. Poole and D. Cortez, "Functions of SMARCAL1, ZRANB3, and HLTF in maintaining genome stability," *Crit. Rev. Biochem. Mol. Biol.*, vol. 52, no. 6, pp. 696–714, Nov. 2017, doi: 10.1080/10409238.2017.1380597.

[108] A. Quinet, D. Lemaçon, and A. Vindigni, "Replication Fork Reversal: Players and Guardians," *Mol. Cell*, vol. 68, no. 5, pp. 830–833, Dec. 2017, doi: 10.1016/j.molcel.2017.11.022.

[109] M. B. Adolph and D. Cortez, "Mechanisms and regulation of replication fork reversal," *DNA Repair (Amst).*, vol. 141, p. 103731, Sep. 2024, doi: 10.1016/j.dnarep.2024.103731.

[110] J. C. Saldivar, D. Cortez, and K. A. Cimprich, “The essential kinase ATR: ensuring faithful duplication of a challenging genome,” *Nat. Rev. Mol. Cell Biol.*, vol. 18, no. 10, pp. 622–636, Oct. 2017, doi: 10.1038/nrm.2017.67.

[111] R. L. Ragland *et al.*, “RNF4 and PLK1 are required for replication fork collapse in ATR-deficient cells,” *Genes Dev.*, vol. 27, no. 20, pp. 2259–2273, 2013, doi: 10.1101/gad.223180.113.

[112] Y. Liu, C. F. Nielsen, Q. Yao, and I. D. Hickson, “The origins and processing of ultra fine anaphase DNA bridges,” *Curr. Opin. Genet. Dev.*, vol. 26, pp. 1–5, 2014, doi: 10.1016/j.gde.2014.03.003.

[113] A. Moreno *et al.*, “Unreplicated DNA remaining from unperturbed S phases passes through mitosis for resolution in daughter cells,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 113, no. 39, pp. E5757–E5764, 2016, doi: 10.1073/pnas.1603252113.

[114] I. D. Hickson and R. Bhowmick, “The ‘enemies within’: Regions of the genome that are inherently difficult to replicate,” *F1000Research*, vol. 6, no. May, pp. 1–7, 2017, doi: 10.12688/f1000research.11024.1.

[115] Y. Saintigny *et al.*, “Characterization of homologous recombination induced by replication inhibition in mammalian cells,” *EMBO J.*, vol. 20, no. 14, pp. 3861–3870, 2001, doi: 10.1093/emboj/20.14.3861.

[116] M. Mann, S.-E. Ong, M. Grønborg, H. Steen, O. N. Jensen, and A. Pandey, “Analysis of protein phosphorylation using mass spectrometry: deciphering the phosphoproteome,” *Trends Biotechnol.*, vol. 20, no. 6, pp. 261–268, Jun. 2002, doi: 10.1016/S0167-7799(02)01944-3.

[117] Larasati and B. P. Duncker, “Mechanisms governing DDK regulation of the initiation of DNA replication,” *Genes (Basel).*, vol. 8, no. 1, pp. 1–12, 2017, doi: 10.3390/genes8010003.

[118] C. Wurzenberger and D. W. Gerlich, “Phosphatases: providing safe passage through mitotic exit,” *Nat. Rev. Mol. Cell Biol.*, vol. 12, no. 8, pp. 469–482, Aug. 2011, doi: 10.1038/nrm3149.

[119] S. A. Touati, L. Hofbauer, A. W. Jones, A. P. Snijders, G. Kelly, and F. Uhlmann, “Cdc14 and PP2A Phosphatases Cooperate to Shape Phosphoproteome Dynamics during Mitotic Exit,” *Cell Rep.*, vol. 29, no. 7, pp. 2105-2119.e4, Nov. 2019, doi: 10.1016/j.celrep.2019.10.041.

[120] B. L. Powers and M. C. Hall, “Re-examining the role of Cdc14 phosphatase in reversal of Cdk phosphorylation during mitotic exit,” *J. Cell Sci.*, vol. 130, no. 16, pp. 2673–2681, Aug. 2017, doi: 10.1242/jcs.201012.

[121] A. W. Lau *et al.*, “Regulation of APCCdh1 E3 ligase activity by the Fbw7/cyclin E signaling axis contributes to the tumor suppressor function of Fbw7,” *Cell Res.*, vol. 23, no. 7, pp. 947–961, Jul. 2013, doi: 10.1038/cr.2013.67.

[122] H. Fukushima *et al.*, “SCF-Mediated Cdh1 Degradation Defines a Negative Feedback System that Coordinates Cell-Cycle Progression,” *Cell Rep.*, vol. 4, no. 4, pp. 803–816, Aug. 2013, doi: 10.1016/j.celrep.2013.07.031.

[123] L. I. Francis, J. C. W. Randell, T. J. Takara, L. Uchima, and S. P. Bell, “Incorporation into the prereplicative complex activates the Mcm2–7 helicase for Cdc7–Dbf4 phosphorylation,” *Genes Dev.*, vol. 23, no. 5, pp. 643–654, Mar. 2009, doi: 10.1101/gad.1759609.

[124] Y.-J. Sheu and B. Stillman, “The Dbf4–Cdc7 kinase promotes S phase by alleviating an inhibitory activity in Mcm4,” *Nature*, vol. 463, no. 7277, pp. 113–117, Jan. 2010, doi: 10.1038/nature08647.

[125] T. D. Deegan, J. T. Yeeles, and J. F. Diffley, “Phosphopeptide binding by Sld3 links Dbf4-dependent kinase to MCM replicative helicase activation,” *EMBO J.*, vol. 35, no. 9, pp. 961–973, May 2016, doi: 10.15252/embj.201593552.

[126] F. Jenkinson, K. W. Tan, B. Schöpf, M. M. Santos, and P. Zegerman, “Dephosphorylation of the pre-initiation complex is critical for origin firing,” *Mol. Cell*, vol. 83, no. 1, pp. 12-25.e10, Jan. 2023, doi: 10.1016/j.molcel.2022.12.001.

[127] D. Coverley, C. Pelizon, S. Trewick, and R. A. Laskey, “Chromatin-bound Cdc6 persists in S and G2

phases in human cells, while soluble Cdc6 is destroyed in a cyclin A-cdk2 dependent process," *J. Cell Sci.*, vol. 113, no. 11, pp. 1929–1938, Jun. 2000, doi: 10.1242/jcs.113.11.1929.

[128] S. Chen and S. P. Bell, "CDK prevents Mcm2–7 helicase loading by inhibiting Cdt1 interaction with Orc6," *Genes Dev.*, vol. 25, no. 4, pp. 363–372, Feb. 2011, doi: 10.1101/gad.2011511.

[129] E. Liu, X. Li, F. Yan, Q. Zhao, and X. Wu, "Cyclin-dependent Kinases Phosphorylate Human Cdt1 and Induce Its Degradation," *J. Biol. Chem.*, vol. 279, no. 17, pp. 17283–17288, Apr. 2004, doi: 10.1074/jbc.C300549200.

[130] M. Makise, M. Takehara, A. Kuniyasu, N. Matsui, H. Nakayama, and T. Mizushima, "Linkage between Phosphorylation of the Origin Recognition Complex and Its ATP Binding Activity in Saccharomyces cerevisiae," *J. Biol. Chem.*, vol. 284, no. 6, pp. 3396–3407, Feb. 2009, doi: 10.1074/jbc.M804293200.

[131] P. Saha *et al.*, "Human CDC6/Cdc18 Associates with Orc1 and Cyclin-cdk and Is Selectively Eliminated from the Nucleus at the Onset of S Phase," *Mol. Cell. Biol.*, vol. 18, no. 5, pp. 2758–2767, May 1998, doi: 10.1128/MCB.18.5.2758.

[132] A. E. Vouzas and D. M. Gilbert, "Replication timing and transcriptional control: beyond cause and effect — part IV," *Curr. Opin. Genet. Dev.*, vol. 79, p. 102031, Apr. 2023, doi: 10.1016/j.gde.2023.102031.

[133] T. Nakatani *et al.*, "Emergence of replication timing during early mammalian development," *Nature*, vol. 625, no. 7994, pp. 401–409, Jan. 2024, doi: 10.1038/s41586-023-06872-1.

[134] E. X. Kwan *et al.*, "Ribosomal DNA replication time coordinates completion of genome replication and anaphase in yeast," *Cell Rep.*, vol. 42, no. 3, p. 112161, Mar. 2023, doi: 10.1016/j.celrep.2023.112161.

[135] H. Fu, A. Baris, and M. I. Aladjem, "Replication timing and nuclear structure," *Curr. Opin. Cell Biol.*, vol. 52, pp. 43–50, Jun. 2018, doi: 10.1016/j.ceb.2018.01.004.

[136] U. Galderisi, F. P. Jori, and A. Giordano, "Cell cycle regulation and neural differentiation," *Oncogene*, vol. 22, no. 33, pp. 5208–5219, Aug. 2003, doi: 10.1038/sj.onc.1206558.

[137] A. Satyanarayana and P. Kaldis, "Mammalian cell-cycle regulation: several Cdks, numerous cyclins and diverse compensatory mechanisms," *Oncogene*, vol. 28, no. 33, pp. 2925–2939, Aug. 2009, doi: 10.1038/onc.2009.170.

[138] D. Boos and P. Ferreira, "Origin Firing Regulations to Control Genome Replication Timing," *Genes (Basel).*, vol. 10, no. 3, p. 199, Mar. 2019, doi: 10.3390/genes10030199.

[139] A. Davé, C. Cooley, M. Garg, and A. Bianchi, "Protein Phosphatase 1 Recruitment by Rif1 Regulates DNA Replication Origin Firing by Counteracting DDK Activity," *Cell Rep.*, vol. 7, no. 1, pp. 53–61, Apr. 2014, doi: 10.1016/j.celrep.2014.02.019.

[140] S. Hiraga *et al.*, "Rif1 controls DNA replication by directing Protein Phosphatase 1 to reverse Cdc7-mediated phosphorylation of the MCM complex," *Genes Dev.*, vol. 28, no. 4, pp. 372–383, Feb. 2014, doi: 10.1101/gad.231258.113.

[141] H. Beck *et al.*, "Cyclin-Dependent Kinase Suppression by WEE1 Kinase Protects the Genome through Control of Replication Initiation and Nucleotide Consumption," *Mol. Cell. Biol.*, vol. 32, no. 20, pp. 4226–4236, Oct. 2012, doi: 10.1128/MCB.00412-12.

[142] T. H. Stracker, T. Usui, and J. H. J. Petrini, "Taking the time to make important decisions: The checkpoint effector kinases Chk1 and Chk2 and the DNA damage response," *DNA Repair (Amst).*, vol. 8, no. 9, pp. 1047–1054, Sep. 2009, doi: 10.1016/j.dnarep.2009.04.012.

[143] F. Bassermann, D. Frescas, D. Guardavaccaro, L. Busino, A. Peschiaroli, and M. Pagano, "The Cdc14B-Cdh1-Plk1 Axis Controls the G2 DNA-Damage-Response Checkpoint," *Cell*, vol. 134, no. 2, pp. 256–267, Jul. 2008, doi: 10.1016/j.cell.2008.05.043.

[144] F. Jenkinson and P. Zegerman, "Roles of phosphatases in eukaryotic DNA replication initiation control," *DNA Repair (Amst).*, vol. 118, p. 103384, Oct. 2022, doi: 10.1016/j.dnarep.2022.103384.

[145] S. J. Szyjka *et al.*, “Rad53 regulates replication fork restart after DNA damage in Saccharomyces cerevisiae,” *Genes Dev.*, vol. 22, no. 14, pp. 1906–1920, Jul. 2008, doi: 10.1101/gad.1660408.

[146] A. Travesa, A. Duch, and D. G. Quintana, “Distinct Phosphatases Mediate the Deactivation of the DNA Damage Checkpoint Kinase Rad53,” *J. Biol. Chem.*, vol. 283, no. 25, pp. 17123–17130, Jun. 2008, doi: 10.1074/jbc.M801402200.

[147] D. Slavov *et al.*, “Criteria for gene identification and features of genome organization: Analysis of 6.5 Mb of DNA sequence from human chromosome 21,” *Gene*, vol. 247, no. 1–2, pp. 215–232, 2000, doi: 10.1016/S0378-1119(00)00089-5.

[148] S. L. Wynn *et al.*, “Organization and conservation of the Gart/Son/Donson locus in mouse and human genomes,” *Genomics*, vol. 68, no. 1, pp. 57–62, 2000, doi: 10.1006/geno.2000.6254.

[149] J. L. Bandura, E. L. Beall, M. Bell, H. R. Silver, M. R. Botchan, and B. R. Calvi, “humpty dumpty Is Required for Developmental DNA Amplification and Cell Proliferation in Drosophila,” *Curr. Biol.*, vol. 15, pp. 755–759, 2005.

[150] F. Fuchs *et al.*, “Clustering phenotype populations by genome-wide RNAi and multiparametric imaging,” *Mol. Syst. Biol.*, vol. 6, no. 370, pp. 1–13, 2010.

[151] G. D. Evrony *et al.*, “Integrated genome and transcriptome sequencing identifies a noncoding mutation in the genome replication factor DONSON as the cause of microcephaly-micromelia syndrome,” *Genome Res.*, vol. 27, pp. 1323–1335, 2017.

[152] J. J. Reynolds *et al.*, “Mutations in DONSON disrupt replication fork stability and cause microcephalic dwarfism,” *Nat. Genet.*, vol. 49, no. 4, pp. 537–549, 2017.

[153] H. A. Abdelrahman, A. John, B. R. Ali, and L. Al-Gazali, “Further Delineation of the MicrocephalyMicromelia Syndrome Associated with Loss-of-Function Variants in DONSON,” *Mol. Syndromol.*, vol. 10, pp. 171–176, 2019.

[154] K. M. Knapp *et al.*, “Linked-read genome sequencing identifies biallelic pathogenic variants in DONSON as a novel cause of Meier-Gorlin syndrome,” *J. Med. Genet.*, vol. 57, pp. 195–202, 2020.

[155] G. Nerakh, V. S. Vineeth, K. Tallapaka, L. Nair, A. Dalal, and S. Aggarwal, “Microcephalic primordial dwarfism with predominant Meier-Gorlin phenotype, ichthyosis, and multiple joint deformities-Further expansion of DONSON Cell Cycle-opathy phenotypic spectrum,” *Am. J. Med. Genet.*, vol. 188, pp. 2139–2146, 2022.

[156] J. Zhang *et al.*, “DONSON and FANCM associate with different replisomes distinguished by replication timing and chromatin domain,” *Nat. Commun.*, vol. 11, no. 3951, pp. 1–15, 2020.

[157] E. Nielsen-Dandoroff, M. S. G. Ruegg, and L. S. Bicknell, “The expanding genetic and clinical landscape associated with Meier-Gorlin syndrome,” *Eur. J. Hum. Genet.*, vol. 31, no. 8, pp. 859–868, Aug. 2023, doi: 10.1038/s41431-023-01359-z.

[158] G. Kingsley *et al.*, “DONSON facilitates Cdc45 and GINS chromatin association and is essential for DNA replication initiation,” *Nucleic Acids Res.*, vol. 51, no. 18, pp. 9748–9763, 2023, doi: 10.1093/nar/gkad694.

[159] S. Lesly, J. L. Bandura, and B. R. Calvi, “Rapid DNA synthesis during early drosophila embryogenesis is sensitive to maternal humpty dumpty protein function,” *Genetics*, vol. 207, no. 3, pp. 935–947, 2017, doi: 10.1534/genetics.117.300318.

[160] K. Mutreja *et al.*, “ATR-Mediated Global Fork Slowing and Reversal Assist Fork Traverse and Prevent Chromosomal Breakage at DNA Interstrand Cross-Links,” *Cell Rep.*, vol. 24, no. 10, pp. 2629-2642.e5, 2018, doi: 10.1016/j.celrep.2018.08.019.

[161] J. Huang *et al.*, “Remodeling of Interstrand Crosslink Proximal Replisomes Is Dependent on ATR, FANCM, and FANCD2,” *Cell Rep.*, vol. 27, no. 6, pp. 1794-1808.e5, 2019, doi: 10.1016/j.celrep.2019.04.032.

[162] M. Castella *et al.*, “FANCI Regulates Recruitment of the FA Core Complex at Sites of DNA Damage Independently of FANCD2,” *PLoS Genet.*, vol. 11, no. 10, pp. 1–27, 2015, doi: 10.1371/journal.pgen.1005563.

[163] R. S. Cheung, M. Castella, A. Abeyta, P. R. Gafken, N. Tucker, and T. Taniguchi, “Ubiquitination-Linked Phosphorylation of the FANCI S/TQ Cluster Contributes to Activation of the Fanconi Anemia I/D2 Complex,” *Cell Rep.*, vol. 19, no. 12, pp. 2432–2440, 2017, doi: 10.1016/j.celrep.2017.05.081.

[164] A. Klingseisen and A. P. Jackson, “Mechanisms and pathways of growth failure in primordial dwarfism,” *Genes Dev.*, vol. 25, no. 19, pp. 2011–2024, 2011, doi: 10.1101/gad.169037.

[165] S. Venkataramanappa *et al.*, “The microcephaly gene Donson is essential for progenitors of cortical glutamatergic and GABAergic neurons,” *PLOS Genet.*, vol. 17, no. 3, p. e1009441, Mar. 2021, doi: 10.1371/journal.pgen.1009441.

[166] L. C. Burrage *et al.*, “De Novo GMNN Mutations Cause Autosomal-Dominant Primordial Dwarfism Associated with Meier-Gorlin Syndrome,” *Am. J. Hum. Genet.*, vol. 97, no. 6, pp. 904–913, Dec. 2015, doi: 10.1016/j.ajhg.2015.11.006.

[167] A. L. Fenwick *et al.*, “Mutations in CDC45 , Encoding an Essential Component of the Pre-initiation Complex, Cause Meier-Gorlin Syndrome and Craniosynostosis,” *Am. J. Hum. Genet.*, vol. 99, no. 1, pp. 125–138, Jul. 2016, doi: 10.1016/j.ajhg.2016.05.019.

[168] A. Vetro *et al.*, “MCM5: a new actor in the link between DNA replication and Meier-Gorlin syndrome,” *Eur. J. Hum. Genet.*, vol. 25, no. 5, pp. 646–650, May 2017, doi: 10.1038/ejhg.2017.5.

[169] N. Klumper *et al.*, “Downstream neighbor of SON (DONSON) is associated with unfavorable survival across diverse cancers with oncogenic properties in clear cell renal cell carcinoma,” *Transl. Oncol.*, vol. 13, no. 11, pp. 1–8, 2020.

[170] N. Klümper *et al.*, “Downstream neighbor of son (DONSON) expression is enhanced in phenotypically aggressive prostate cancers,” *Cancers (Basel).*, vol. 12, no. 11, pp. 1–11, 2020, doi: 10.3390/cancers12113439.

[171] Y. Qi *et al.*, “Downstream Neighbor of Son Overexpression is Associated With Breast Cancer Progression and a Poor Prognosis,” *J. Breast Cancer*, vol. 25, no. 4, pp. 327–343, 2022, doi: 10.4048/jbc.2022.25.e26.

[172] L. Ding *et al.*, “Circular RNA circ-DONSON facilitates gastric cancer growth and invasion via NURF complex dependent activation of transcription factor SOX4,” *Mol. Cancer*, vol. 18, no. 45, pp. 1–11, 2019.

[173] E. Liu, X. Sun, J. Li, and C. Zhang, “miR-30a-5p inhibits the proliferation, migration and invasion of melanoma cells by targeting SOX4,” *Mol. Med. Rep.*, Jun. 2018, doi: 10.3892/mmr.2018.9166.

[174] J. Zou and Y. Xu, “MicroRNA-140 Inhibits Cell Proliferation in Gastric Cancer Cell Line HGC-27 by Suppressing SOX4,” *Med. Sci. Monit.*, vol. 22, pp. 2243–2252, Jun. 2016, doi: 10.12659/MSM.896633.

[175] Q. Liu *et al.*, “UCA1 promotes cell proliferation and invasion and inhibits apoptosis through regulation of the miR129–SOX4 pathway in renal cell carcinoma,” *Onco. Targets. Ther.*, vol. Volume 11, pp. 2475–2487, May 2018, doi: 10.2147/OTT.S160192.

[176] N. Wang *et al.*, “CXCL1 derived from tumor-associated macrophages promotes breast cancer metastasis via activating NF-κB/SOX4 signaling,” *Cell Death Dis.*, vol. 9, no. 9, p. 880, Aug. 2018, doi: 10.1038/s41419-018-0876-3.

[177] G. Almouzni and M. Méchali, “Xenopus egg extracts: A model system for chromatin replication,” *Biochim. Biophys. Acta - Gene Struct. Expr.*, vol. 951, no. 2–3, pp. 443–450, Dec. 1988, doi: 10.1016/0167-4781(88)90118-2.

[178] P. J. Gillespie, A. Gambus, and J. J. Blow, “Preparation and use of Xenopus egg extracts to study DNA replication and chromatin associated proteins,” *Methods*, vol. 57, no. 2, pp. 203–213, Jun. 2012, doi:

10.1016/j.ymeth.2012.03.029.

[179] W. S. Hoogenboom, D. Klein Douwel, and P. Knipscheer, “Xenopus egg extract: A powerful tool to study genome maintenance mechanisms,” *Dev. Biol.*, vol. 428, no. 2, pp. 300–309, Aug. 2017, doi: 10.1016/j.ydbio.2017.03.033.

[180] M. Blum and T. Ott, “Xenopus: An Undervalued Model Organism to Study and Model Human Genetic Disease,” *Cells Tissues Organs*, vol. 205, no. 5–6, pp. 303–313, 2018, doi: 10.1159/000490898.

[181] J. Sparks and J. C. Walter, “Extracts for Analysis of DNA Replication in a Nucleus-Free System,” *Cold Spring Harb. Protoc.*, vol. 2019, no. 3, p. pdb.prot097154, Mar. 2019, doi: 10.1101/pdb.prot097154.

[182] I. Volpi, P. J. Gillespie, G. S. Chadha, and J. J. Blow, “The role of DDK and Treslin–MTBP in coordinating replication licensing and pre-initiation complex formation,” *Open Biol.*, vol. 11, no. 10, Oct. 2021, doi: 10.1098/rsob.210121.

[183] S. Priego Moreno, R. Bailey, N. Campion, S. Herron, and A. Gambus, “Polyubiquitylation drives replisome disassembly at the termination of DNA replication,” *Science (80-.).*, vol. 346, no. 6208, pp. 477–481, 2014.

[184] G. A. Khoudoli, P. J. Gillespie, G. Stewart, J. S. Andersen, J. R. Swedlow, and J. J. Blow, “Temporal Profiling of the Chromatin Proteome Reveals System-wide Responses to Replication Inhibition,” *Curr. Biol.*, vol. 18, no. 11, pp. 838–843, Jun. 2008, doi: 10.1016/j.cub.2008.04.075.

[185] T. A. Prokhorova and J. J. Blow, “Sequential MCM/P1 Subcomplex Assembly Is Required to Form a Heterohexamer with Replication Licensing Activity,” *J. Biol. Chem.*, vol. 275, no. 4, pp. 2491–2498, Jan. 2000, doi: 10.1074/jbc.275.4.2491.

[186] A. Gambus, G. A. Khoudoli, R. C. Jones, and J. J. Blow, “MCM2-7 Form Double Hexamers at Licensed Origins in Xenopus Egg Extract,” *J. Biol. Chem.*, vol. 286, no. 13, pp. 11855–11864, 2011.

[187] Z. Tarcan, D. Poovathumkadavil, A. Skagia, and A. Gambus, “The p97 segregase cofactor Ubxn7 facilitates replisome disassembly during S-phase,” *J. Biol. Chem.*, vol. 298, no. 8, p. 102234, Aug. 2022, doi: 10.1016/j.jbc.2022.102234.

[188] R. Lebofsky, T. Takahashi, and J. C. Walter, “DNA Replication in Nucleus-Free Xenopus Egg Extracts,” in *DNA Replication Methods and Protocols*, S. Vengrova and J. Dalgaard, Eds., Vol 521.Humana Press, 2009, pp. 229–252. doi: 10.1007/978-1-60327-815-7_13.

[189] J. J. Blow and R. A. Laskey, “Initiation of DNA replication in nuclei and purified DNA by a cell-free extract of Xenopus eggs,” *Cell*, vol. 47, no. 4, pp. 577–587, Nov. 1986, doi: 10.1016/0092-8674(86)90622-7.

[190] M. A. Cvetkovic *et al.*, “The structural mechanism of dimeric DONSON in replicative helicase activation,” *Mol. Cell*, vol. 83, no. 22, pp. 4017-4031.e9, 2023, doi: 10.1016/j.molcel.2023.09.029.

[191] G. Fullbright, H. B. Rycenga, J. D. Gruber, and D. T. Long, “p97 Promotes a Conserved Mechanism of Helicase Unloading during DNA Cross-Link Repair,” *Mol. Cell. Biol.*, vol. 36, no. 23, pp. 2983–2994, Dec. 2016, doi: 10.1128/MCB.00434-16.

[192] S. V. Srinivasan, D. Dominguez-Sola, L. C. Wang, O. Hyrien, and J. Gautier, “Cdc45 Is a Critical Effector of Myc-Dependent DNA Replication Stress,” *Cell Rep.*, vol. 3, no. 5, pp. 1629–1639, May 2013, doi: 10.1016/j.celrep.2013.04.002.

[193] M. Day *et al.*, “TopBP1 utilises a bipartite GINS binding mode to support genome replication,” *Nat. Commun.*, vol. 15, no. 1, p. 1797, Feb. 2024, doi: 10.1038/s41467-024-45946-0.

[194] Y. Lim *et al.*, “In silico protein interaction screening uncovers DONSON’s role in replication initiation,” *Science (80-.).*, vol. 381, no. 6664, 2023, doi: 10.1126/science.adi3448.

[195] Y. Hashimoto, K. Sadano, N. Miyata, H. Ito, and H. Tanaka, “Novel role of DONSON in CMG helicase assembly during vertebrate DNA replication initiation,” *EMBO J.*, vol. 42, no. 17, 2023, doi: 10.15252/embj.2023114131.

[196] Y. Xia *et al.*, “DNSN-1 recruits GINS for CMG helicase assembly during DNA replication initiation in Caenorhabditis elegans,” *Science (80-.).*, vol. 381, no. 6664, 2023, doi: 10.1126/science.adi4932.

[197] R. Terui, S. E. Berger, L. A. Sambel, D. Song, and G. Chistol, “Single-molecule imaging reveals the mechanism of bidirectional replication initiation in metazoa,” *Cell*, vol. 187, no. 15, pp. 3992-4009.e25, 2024, doi: 10.1016/j.cell.2024.05.024.

[198] C. Capp, J. Wu, and T. Hsieh, “RecQ4: the second replicative helicase?,” *Crit. Rev. Biochem. Mol. Biol.*, vol. 45, no. 3, pp. 233–242, Jun. 2010, doi: 10.3109/10409231003786086.

[199] J. Wu, C. Capp, L. Feng, and T. Hsieh, “Drosophila homologue of the Rothmund–Thomson syndrome gene: Essential function in DNA replication during development,” *Dev. Biol.*, vol. 323, no. 1, pp. 130–142, Nov. 2008, doi: 10.1016/j.ydbio.2008.08.006.

[200] Y. Xu *et al.*, “dRecQ4 Is Required for DNA Synthesis and Essential for Cell Proliferation in Drosophila,” *PLoS One*, vol. 4, no. 7, p. e6107, Jul. 2009, doi: 10.1371/journal.pone.0006107.

[201] L. Padayachy, S. G. Ntallis, and T. D. Halazonetis, “RECQL4 is not critical for firing of human DNA replication origins,” *Sci. Rep.*, vol. 14, no. 1, p. 7708, Apr. 2024, doi: 10.1038/s41598-024-58404-0.

[202] V. Gaggioli, E. Zeiser, D. Rivers, C. R. Bradshaw, J. Ahringer, and P. Zegerman, “CDK phosphorylation of SLD-2 is required for replication initiation and germline development in C. elegans,” *J. Cell Biol.*, vol. 204, no. 4, pp. 507–522, Feb. 2014, doi: 10.1083/jcb.201310083.

[203] G. Zonderland *et al.*, “The TRESLIN-MTBP complex couples completion of DNA replication with S/G2 transition,” *Mol. Cell*, vol. 82, no. 18, pp. 3350-3365.e7, Sep. 2022, doi: 10.1016/j.molcel.2022.08.006.

[204] D. Boos, M. Yekezare, and J. F. X. Diffley, “Identification of a Heteromeric Complex That Promotes DNA Replication Origin Firing in Human Cells,” *Science (80-.).*, vol. 340, no. 6135, pp. 981–984, May 2013, doi: 10.1126/science.1237448.

[205] A. Kumagai and W. G. Dunphy, “MTBP, the partner of Treslin, contains a novel DNA-binding domain that is essential for proper initiation of DNA replication,” *Mol. Biol. Cell*, vol. 28, pp. 2998–3012, 2017.

[206] H. Araki, “Initiation of chromosomal DNA replication in eukaryotic cells; Contribution of Yeast genetics to the elucidation,” *Genes Genet. Syst.*, vol. 86, no. 3, pp. 141–149, 2011, doi: 10.1266/ggs.86.141.

[207] L. S. Bicknell *et al.*, “Mutations in the pre-replication complex cause Meier-Gorlin syndrome,” *Nat. Genet.*, vol. 43, no. 4, pp. 356–360, 2011, doi: 10.1038/ng.775.

[208] C. Evrin, V. Alvarez, J. Ainsworth, R. Fujisawa, C. Alabert, and K. P. Labib, “DONSON is required for CMG helicase assembly in the mammalian cell cycle,” *EMBO Rep.*, vol. 24, no. 11, Nov. 2023, doi: 10.15252/embr.202357677.

[209] J. T. P. Yeeles, A. Janska, A. Early, and J. F. X. Diffley, “How the Eukaryotic Replisome Achieves Rapid and Efficient DNA Replication,” *Mol. Cell*, vol. 65, no. 1, pp. 105–116, Jan. 2017, doi: 10.1016/j.molcel.2016.11.017.

[210] R. Georgescu *et al.*, “Structure of eukaryotic CMG helicase at a replication fork and implications to replisome architecture and origin initiation,” *Proc. Natl. Acad. Sci.*, vol. 114, no. 5, Jan. 2017, doi: 10.1073/pnas.1620500114.

[211] D. Baretić, M. Jenkyn-Bedford, V. Aria, G. Cannone, M. Skehel, and J. T. P. Yeeles, “Cryo-EM Structure of the Fork Protection Complex Bound to CMG at a Replication Fork,” *Mol. Cell*, vol. 78, no. 5, pp. 926-940.e13, Jun. 2020, doi: 10.1016/j.molcel.2020.04.012.

[212] P. Eickhoff *et al.*, “Molecular Basis for ATP-Hydrolysis-Driven DNA Translocation by the CMG Helicase of the Eukaryotic Replisome,” *Cell Rep.*, vol. 28, no. 10, pp. 2673-2688.e8, Sep. 2019, doi: 10.1016/j.celrep.2019.07.104.

[213] N. J. Rzechorzek, S. W. Hardwick, V. A. Jatikusumo, D. Y. Chirgadze, and L. Pellegrini, “CryoEM structures of human CMG–ATPγS–DNA and CMG–AND-1 complexes,” *Nucleic Acids Res.*, vol. 48, no. 12, pp. 6980–6995, Jul. 2020, doi: 10.1093/nar/gkaa429.

[214] M. L. Jones, Y. Baris, M. R. Taylor, and J. T. Yeeles, “Structure of a human replisome shows the organisation and interactions of a DNA replication machine,” *EMBO J.*, vol. 42, no. 22, Nov. 2023, doi: 10.15252/embj.2023115685.

[215] Y. Baris, M. R. G. Taylor, V. Aria, and J. T. P. Yeeles, “Fast and efficient DNA replication with purified human proteins,” *Nature*, vol. 606, no. 7912, pp. 204–210, Jun. 2022, doi: 10.1038/s41586-022-04759-1.

[216] R. S. G. I. (RSGI) Nagashima, T., Hayashi, F., Yokoyama, S., “The third BRCA1 C-terminus (BRCT) domain of Similar to S.pombe rad4+/cut5+ product.” Nov. 26, 2004. doi: 10.2210/pdb1wf6/pdb.

[217] M. Rappas, A. W. Oliver, and L. H. Pearl, “Structure and function of the Rad9-binding region of the DNA-damage checkpoint adaptor TopBP1,” *Nucleic Acids Res.*, vol. 41, no. 8, pp. 4741–4741, Apr. 2013, doi: 10.1093/nar/gkt202.

[218] L. Sun *et al.*, “Structural Insight into BLM Recognition by TopBP1,” *Structure*, vol. 25, no. 10, pp. 1582-1588.e3, Oct. 2017, doi: 10.1016/j.str.2017.08.005.

[219] C. C. Y. Leung, E. Kellogg, A. Kuhnert, F. Hänel, D. Baker, and J. N. M. Glover, “Insights from the crystal structure of the sixth BRCT domain of topoisomerase IIβ binding protein 1,” *Protein Sci.*, vol. 19, no. 1, pp. 162–167, Jan. 2010, doi: 10.1002/pro.290.

[220] Y. Huo, L. Bai, M. Xu, and T. Jiang, “Crystal structure of the N-terminal region of human Topoisomerase IIβ binding protein 1,” *Biochem. Biophys. Res. Commun.*, vol. 401, no. 3, pp. 401–405, Oct. 2010, doi: 10.1016/j.bbrc.2010.09.066.

[221] Y. Noguchi *et al.*, “Cryo-EM structure of Mcm2-7 double hexamer on DNA suggests a lagging-strand DNA extrusion model,” *Proc. Natl. Acad. Sci.*, vol. 114, no. 45, Nov. 2017, doi: 10.1073/pnas.1712537114.

[222] F. Abid Ali *et al.*, “Cryo-EM structure of a licensed DNA replication origin,” *Nat. Commun.*, vol. 8, no. 1, p. 2241, Dec. 2017, doi: 10.1038/s41467-017-02389-0.

[223] J. Li *et al.*, “The human pre-replication complex is an open complex,” *Cell*, vol. 186, no. 1, pp. 98-111.e21, Jan. 2023, doi: 10.1016/j.cell.2022.12.008.

[224] P. C. Burcham, “Internal hazards: baseline DNA damage by endogenous products of normal metabolism,” *Mutat. Res. Toxicol. Environ. Mutagen.*, vol. 443, no. 1–2, pp. 11–36, Jul. 1999, doi: 10.1016/S1383-5742(99)00008-3.

[225] R. Huber, “Flexibility and rigidity, requirements for the function of proteins and protein pigment complexes,” *Biochem. Soc. Trans.*, vol. 15, no. 6, pp. 1009–1020, Dec. 1987, doi: 10.1042/bst0151009.

[226] J. Frigola, D. Remus, A. Mehanna, and J. F. X. Diffley, “ATPase-dependent quality control of DNA replication origin licensing,” *Nature*, vol. 495, no. 7441, pp. 339–343, Mar. 2013, doi: 10.1038/nature11920.

[227] Z. Yuan *et al.*, “Structure of the eukaryotic replicative CMG helicase suggests a pumpjack motion for translocation,” *Nat. Struct. Mol. Biol.*, vol. 23, no. 3, pp. 217–224, Mar. 2016, doi: 10.1038/nsmb.3170.

[228] Y. P. Chang, G. Wang, V. Bermudez, J. Hurwitz, and X. S. Chen, “Crystal structure of the GINS complex and functional insights into its role in DNA replication,” *Proc. Natl. Acad. Sci.*, vol. 104, no. 31, pp. 12685–12690, Jul. 2007, doi: 10.1073/pnas.0705558104.

[229] K. Kamada, Y. Kubota, T. Arata, Y. Shindo, and F. Hanaoka, “Structure of the human GINS complex and its assembly and functional interface in replication initiation,” *Nat. Struct. Mol. Biol.*, vol. 14, no. 5, pp. 388–396, May 2007, doi: 10.1038/nsmb1231.

[230] S. S. Henrikus *et al.*, “Unwinding of a eukaryotic origin of replication visualized by cryo-EM,” *Nat. Struct. Mol. Biol.*, vol. 31, no. 8, pp. 1265–1276, Aug. 2024, doi: 10.1038/s41594-024-01280-z.

[231] D. Fang, Q. Cao, and H. Lou, “Sld3-MCM Interaction Facilitated by Dbf4-Dependent Kinase Defines an Essential Step in Eukaryotic DNA Replication Initiation,” *Front. Microbiol.*, vol. 7, Jun. 2016, doi:

10.3389/fmicb.2016.00885.

[232] J. S. Im, S. Y. Park, W. H. Cho, S. H. Bae, J. Hurwitz, and J. K. Lee, "RecQL4 is required for the association of Mcm10 and Ctf4 with replication origins in human cells," *Cell Cycle*, vol. 14, no. 7. pp. 1001–1009, 2015. doi: 10.1080/15384101.2015.1007001.

[233] X. Chen, J. L. Zaro, and W.-C. Shen, "Fusion protein linkers: Property, design and functionality," *Adv. Drug Deliv. Rev.*, vol. 65, no. 10, pp. 1357–1369, Oct. 2013, doi: 10.1016/j.addr.2012.09.039.

[234] L. Li, W. Sun, Z. Zhang, and Y. Huang, "Time-staggered delivery of docetaxel and H1-S6A,F8A peptide for sequential dual-strike chemotherapy through tumor priming and nuclear targeting," *J. Control. Release*, vol. 232, pp. 62–74, Jun. 2016, doi: 10.1016/j.jconrel.2016.04.021.

[235] G. L. Bidwell, E. Perkins, J. Hughes, M. Khan, J. R. James, and D. Raucher, "Thermally Targeted Delivery of a c-Myc Inhibitory Polypeptide Inhibits Tumor Progression and Extends Survival in a Rat Glioma Model," *PLoS One*, vol. 8, no. 1, p. e55104, Jan. 2013, doi: 10.1371/journal.pone.0055104.

[236] E. Karaca *et al.*, "Biallelic and De Novo Variants in DONSON Reveal a Clinical Spectrum of Cell Cycle-opathies with Microcephaly, Dwarfism and Skeletal Abnormalities," *Am. J. Med. Genet. Part A*, vol. 179, no. 10, pp. 2056–2066, 2019, doi: 10.1002/ajmg.a.61315.

[237] H. Masumoto, S. Muramatsu, Y. Kamimura, and H. Araki, "S-Cdk-dependent phosphorylation of Sld2 essential for chromosomal DNA replication in budding yeast," *Nature*, vol. 415, no. 6872, pp. 651–655, 2002, doi: 10.1038/nature713.

[238] M. Carroni *et al.*, "New insights into the GINS complex explain the controversy between existing structural models," *Sci. Rep.*, vol. 7, no. 1, p. 40188, Jan. 2017, doi: 10.1038/srep40188.

[239] M. Day, A. W. Oliver, and L. H. Pearl, "Phosphorylation-dependent assembly of DNA damage response systems and the central roles of TOPBP1," *DNA Repair (Amst).*, vol. 108, p. 103232, Dec. 2021, doi: 10.1016/j.dnarep.2021.103232.

[240] S. Chong and M. Mir, "Towards Decoding the Sequence-Based Grammar Governing the Functions of Intrinsically Disordered Protein Regions," *J. Mol. Biol.*, vol. 433, no. 12, p. 166724, Jun. 2021, doi: 10.1016/j.jmb.2020.11.023.

[241] Y. Shin and C. P. Brangwynne, "Liquid phase condensation in cell physiology and disease," *Science (80-.).*, vol. 357, no. 6357, Sep. 2017, doi: 10.1126/science.aaf4382.

[242] P.-H. Peng, K.-W. Hsu, and K.-J. Wu, "Liquid-liquid phase separation (LLPS) in cellular physiology and tumor biology.," *Am. J. Cancer Res.*, vol. 11, no. 8, pp. 3766–3776, 2021.

[243] C. Frattini *et al.*, "TopBP1 assembles nuclear condensates to switch on ATR signaling," *Mol. Cell*, vol. 81, no. 6, pp. 1231-1245.e8, Mar. 2021, doi: 10.1016/j.molcel.2020.12.049.

[244] Z. Lao *et al.*, "Insights into the Atomistic Mechanisms of Phosphorylation in Disrupting Liquid–Liquid Phase Separation and Aggregation of the FUS Low-Complexity Domain," *J. Chem. Inf. Model.*, vol. 62, no. 13, pp. 3227–3238, Jul. 2022, doi: 10.1021/acs.jcim.2c00414.

[245] M. Modic, M. Adamek, and J. Ule, "The impact of IDR phosphorylation on the RNA binding profiles of proteins," *Trends Genet.*, vol. 40, no. 7, pp. 580–586, Jul. 2024, doi: 10.1016/j.tig.2024.04.004.

[246] D. Remus, F. Beuron, G. Tolun, J. D. Griffith, E. P. Morris, and J. F. X. Diffley, "Concerted Loading of Mcm2–7 Double Hexamers around DNA during DNA Replication Origin Licensing," *Cell*, vol. 139, no. 4, pp. 719–730, Nov. 2009, doi: 10.1016/j.cell.2009.10.015.

[247] C. Evrin *et al.*, "A double-hexameric MCM2-7 complex is loaded onto origin DNA during licensing of eukaryotic DNA replication," *Proc. Natl. Acad. Sci.*, vol. 106, no. 48, pp. 20240–20245, Dec. 2009, doi: 10.1073/pnas.0911500106.

[248] K.-U. Reusswig, F. Zimmermann, L. Galanti, and B. Pfander, "Robust Replication Control Is Generated by Temporal Gaps between Licensing and Firing Phases and Depends on Degradation of Firing Factor Sld2," *Cell Rep.*, vol. 17, no. 2, pp. 556–569, Oct. 2016, doi: 10.1016/j.celrep.2016.09.013.

[249] A. Kumagai, A. Shevchenko, A. Shevchenko, and W. G. Dunphy, “Direct regulation of Treslin by cyclin-dependent kinase is essential for the onset of DNA replication,” *J. Cell Biol.*, vol. 193, no. 6, pp. 995–1007, Jun. 2011, doi: 10.1083/jcb.201102003.

[250] I. Bruck, D. M. Kanter, and D. L. Kaplan, “Enabling Association of the GINS Protein Tetramer with the Mini Chromosome Maintenance (Mcm)2–7 Protein Complex by Phosphorylated Sld2 Protein and Single-stranded Origin DNA,” *J. Biol. Chem.*, vol. 286, no. 42, pp. 36414–36426, Oct. 2011, doi: 10.1074/jbc.M111.282822.

[251] J. Bagge, V. H. Oestergaard, and M. Lisby, “Functions of TopBP1 in preserving genome integrity during mitosis,” *Semin. Cell Dev. Biol.*, vol. 113, pp. 57–64, May 2021, doi: 10.1016/j.semcdb.2020.08.009.

[252] M. Sokka, S. Parkkinen, H. Pospiech, and J. E. Syväoja, “Function of TopBP1 in Genome Stability,” in *Genome Stability and Human Diseases*, 2010, pp. 119–141. doi: 10.1007/978-90-481-3471-7_7.

[253] S. Mehta and J. Zhang, “Liquid–liquid phase separation drives cellular function and dysfunction in cancer,” *Nat. Rev. Cancer*, vol. 22, no. 4, pp. 239–252, Apr. 2022, doi: 10.1038/s41568-022-00444-7.

6 Appendix

Mass Spectrometry Results

1. Mass spec results of depleted extract chromatin samples (as in Figure 3.1.11 A)
2. Mass spec results of chromatin IP samples (as in Figure 3.1.12 C)

Publications

1. G. Kingsley *et al.*, “DONSON facilitates Cdc45 and GINS chromatin association and is essential for DNA replication initiation,” *Nucleic Acids Res.*, vol. 51, no. 18, pp. 9748–9763, 2023, doi: 10.1093/nar/gkad694.

2. M. A. Cvetkovic *et al.*, “The structural mechanism of dimeric DONSON in replicative helicase activation,” *Mol. Cell*, vol. 83, no. 22, pp. 4017-4031.e9, 2023, doi: 10.1016/j.molcel.2023.09.029.

1. Mass spec results of depleted extract chromatin samples (as in Figure 3.1.11 A)

Protein	*IgG Depleted Peptides*	*DONSON Depleted Peptides*	*Fold Enrichment*
Smarca5	44	N.C.	-44
Nup205	28	N.C.	-28
Fanci	26	N.C.	-26
Tpr	22	N.C.	-22
Pole	21.5	N.C.	-21.5
Pola1	19	N.C.	-19
Dnmt1	18	N.C.	-18
Cytosine-specific methyltransferase	17	N.C.	-17
SALL4	17	N.C.	-17
Pol3	16	N.C.	-16
Xsal-3	15	N.C.	-15
Ubtf	12	N.C.	-12
Hsp90aa1.1	10	N.C.	-10
Rfc4	10	N.C.	-10
Mi-2 histone deacetylase complex protein 66	9	N.C.	-9
Rbbp4-a	9	N.C.	-9
Lgals3	9	N.C.	-9
Wdhd1	9	N.C.	-9
Nup88	8	N.C.	-8
P60	8	N.C.	-8
Chromatin assembly factor 1 subunit A-A	8	N.C.	-8
CDC45	8	N.C.	-8
Claspin	7	N.C.	-7
Rfc1	7	N.C.	-7
Tmpo	6	N.C.	-6
Hic2	6	N.C.	-6
Zhx3	6	N.C.	-6
Uhrf1	6	N.C.	-6
Prim1	6	N.C.	-6
Orc4	6	N.C.	-6
Rpl28	6	N.C.	-6
Pole2	6	N.C.	-6
Rfc5.s	11	2	-5.5
Tim1	5	N.C.	-5
Psma2	5	N.C.	-5
Psmb7	5	N.C.	-5
Dscc1	5	N.C.	-5
Adnp	5	N.C.	-5
Pold3	5	N.C.	-5
Smc3	5	N.C.	-5
PCNA	49	11	-4.5
NPAT	4	N.C.	-4

Elongation factor 1-alpha	4	N.C.	-4
ppp1cc-a	4	N.C.	-4
Rmi1	4	N.C.	-4
Vapb	4	N.C.	-4
Chtf18	4	N.C.	-4
Parp1	4	N.C.	-4
UBC9	4	N.C.	-4
Pola2	4	N.C.	-4
Orc3	4	N.C.	-4
Gnb2l1	7	2	-3.5
Coilin	16	5	-3.2
Ppp1cb	3	N.C.	-3
Rad23b	6	2	-3
Srsf2	6	2	-3
L13	3	N.C.	-3
Atic	3	N.C.	-3
Arf-1	3	N.C.	-3
Paics	3	N.C.	-3
Usp7	3	N.C.	-3
Smarca4	3	N.C.	-3
Mdh1	3	N.C.	-3
Psmb1	3	N.C.	-3
Psmb2	3	N.C.	-3
Ehmt2	3	N.C.	-3
Rplp0	3	N.C.	-3
Nfyc	3	N.C.	-3
Ddi2	3	N.C.	-3
Rpl18a	3	N.C.	-3
Dnaja1	3	N.C.	-3
Lig1	3	N.C.	-3
Mprip	3	N.C.	-3
Helb	3	N.C.	-3
Gins3	3	N.C.	-3
Gins4	3	N.C.	-3
Chtf8	3	N.C.	-3
Rfc2	23	8	-2.9
Rad50	11	4	-2.8
Psma-7	8	3	-2.7
Arpc4	8	3	-2.7
Proteasome subunit alpha type	8	3	-2.7
Rfc3	23	9	-2.6
Tubulin alpha	21	8	-2.6
Atrx	5	2	-2.5
Flnb	17	7	-2.4
Etaa	9	4	-2.3
Hdac1-a	9	4	-2.3
Psmb5	7	3	-2.3
Chd8	7	3	-2.3

Cabin1	11	5	-2.2
Chd4	30	14	-2.1
Hcfc1	8	4	-2
MBD3	6	3	-2
Cdca7	6	3	-2
Rpl14	4	2	-2
Psmd7	4	2	-2
Uchl5	2	N.C.	-2
Psmd3	2	N.C.	-2
Baz2a	2	N.C.	-2
Idh1	2	N.C.	-2
Smarcc2	2	N.C.	-2
Rps16	2	N.C.	-2
Dnajb6-b	2	N.C.	-2
Nusap1-a	2	N.C.	-2
Gys1	2	N.C.	-2
Rps23	2	N.C.	-2
Adnp2	2	N.C.	-2
Rpl26	2	N.C.	-2
Rcc2	2	N.C.	-2
Rpl22	2	N.C.	-2
Rps13	2	N.C.	-2
Rpl11	2	N.C.	-2
Rpl6	2	N.C.	-2
Srsf5	2	N.C.	-2
Upk1b	2	N.C.	-2
Dynll1	2	N.C.	-2
Chmp1a	2	N.C.	-2
Ssh2	2	N.C.	-2
Rpl35a	2	N.C.	-2
Pot1	2	N.C.	-2
Asf1ab	2	N.C.	-2
Cbr3	2	N.C.	-2
Habp4.L	2	N.C.	-2
Smc1a	2	N.C.	-2
Kpna3	2	N.C.	-2
Acad11	2	N.C.	-2
Fancf	2	N.C.	-2
Plk1	2	N.C.	-2
Anln	2	N.C.	-2
Prkra-a	2	N.C.	-2
Rps19	2	N.C.	-2
Proteasome subunit beta	2	N.C.	-2
Orc2	2	N.C.	-2
Rab5a	2	N.C.	-2
Zbtb16	2	N.C.	-2
Gatad2a.S	17	9	-1.9
ISWI	136	74	-1.8
Arpc2-b	16	9	-1.8
Hic2	11	6	-1.8

Numa1	27	16	-1.7
Svil	22	13	-1.7
Efhd1	10	6	-1.7
Rps4	5	3	-1.7
Dbd1	5	3	-1.7
Proteasome subunit alpha	5	3	-1.7
Psma4	5	3	-1.7
Rts	5	3	-1.7
Blm	13	8	-1.6
Rbbp7	13	8	-1.6
Dppa2	13	8	-1.6
Xrrc5	8	5	-1.6
Mel28	21	14	-1.5
Gsn	3	2	-1.5
Topo2	72	49	-1.5
TPR	38	25	-1.5
SSRP1	18	12	-1.5
Mre11	6	4	-1.5
Pole3	6	4	-1.5
Arpc5l	6	4	-1.5
Ruvbl2	6	4	-1.5
H1f0-a	3	2	-1.5
Rps11	3	2	-1.5
Hells	3	2	-1.5
Rpl23a	3	2	-1.5
Cdc6b	3	2	-1.5
Proteasome subunit beta	3	2	-1.5
Tcf20	3	2	-1.5
Rab10	3	2	-1.5
RPA (70kDa)	156	140	-1.1
MCM4	78	93	+1.2
MCM6	90	111	+1.2
MCM5	97	125	+1.3
MCM7	73	94	+1.3
MCM2	110	159	+1.4
Nup153	51	75	+1.5
Cingulin like 1	51	75	+1.5
Nup53	4	6	+1.5
Zmym3	4	6	+1.5
Mta1	4	6	+1.5
Fancc	4	6	+1.5
Ube3a	4	6	+1.5
Ywhaz	2	3	+1.5
birc5.1-a	2	3	+1.5
Ddost	2	3	+1.5
Vdac2	2	3	+1.5
Arpc5	2	3	+1.5
Ppib	2	3	+1.5

Zona pellucida C glycoprotein	2	3	+1.5
Rps14	2	3	+1.5
Histone H2A	33	49	+1.5
Nup43	17	25	+1.5
Tubb2b	15	22	+1.5
MCM3	86	141	+1.6
Nup98	8	13	+1.6
Abcf1	8	13	+1.6
nup50.L	5	8	+1.6
Nupl1	9	14	+1.6
Dbnl-1	9	14	+1.6
Nup62	9	14	+1.6
B4	100	167	+1.7
Tjp1	3	5	+1.7
Ruvbl1	3	5	+1.7
faap100.L	6	10	+1.7
Treslin	15	25	+1.7
H1foo	100	167	+1.7
Enah	7	12	+1.7
Lamin-L	8	14	+1.8
Topbp1-A	9	16	+1.8
Neurabin	5	9	+1.8
Nup50	6	11	+1.8
Cltc	7	13	+1.9
Prdx1	N.C.	2	+2
Ccdc176	N.C.	2	+2
Lamin-A	N.C.	2	+2
slc37a2	N.C.	2	+2
Cat.2	N.C.	2	+2
KIF22-B	N.C.	2	+2
sqstm1	N.C.	2	+2
Hmbox1	N.C.	2	+2
Habp4	N.C.	2	+2
T-complex protein 1 subunit delta	N.C.	2	+2
Cenpc	N.C.	2	+2
Rpl24	N.C.	2	+2
Rpl27	N.C.	2	+2
Anxa7	N.C.	2	+2
Rpl31	N.C.	2	+2
Ssb-b	N.C.	2	+2
Pdia5	N.C.	2	+2
Pa2g4	N.C.	2	+2
guanine nucleotide binding protein	N.C.	2	+2
Rab18	N.C.	2	+2
Rpl4-b	N.C.	2	+2
Uchl1	N.C.	2	+2
XMAN1	N.C.	2	+2

Prkdc	N.C.	2	+2
Rps5	N.C.	2	+2
Gpi	N.C.	2	+2
Mov10	N.C.	2	+2
Perilipin	N.C.	2	+2
Wdr18	N.C.	2	+2
Cap1	N.C.	2	+2
Annexin	N.C.	2	+2
Lap3	N.C.	2	+2
Dynll2	N.C.	2	+2
Cndp2	N.C.	2	+2
HBA-L4	N.C.	2	+2
Rnf123	N.C.	2	+2
Plastin 3	N.C.	2	+2
Cyp	N.C.	2	+2
Anxa2	N.C.	2	+2
Top3a	9	18	+2
Lamin B1	7	14	+2
Cell division cycle associated 2	5	10	+2
Ldhc	3	6	+2
Gmcl1	2	4	+2
Arg1	2	4	+2
Ybx2-a	2	4	+2
Cct6a	2	4	+2
Programmed cell death 6-interacting protein	2	4	+2
Cct8	2	4	+2
SUV39H1	5	11	+2.2
2-phosphoglycerate dehydratase	6	14	+2.3
Fructose-bisphosphate aldolase	3	7	+2.3
Dsp	9	26	+2.9
Cenpt	N.C.	3	+3
Myo7a	N.C.	3	+3
Psmc4	N.C.	3	+3
Rps8	N.C.	3	+3
Orc5	N.C.	3	+3
Tmem209	N.C.	3	+3
Myp	N.C.	3	+3
Psmd4	N.C.	3	+3
Cct3	N.C.	3	+3
serine/threonine-protein phosphatase 2A	N.C.	3	+3

Creatine kinase isozyme IV	N.C.	3	+3
Atp2a2	N.C.	3	+3
Zc3hc1	N.C.	3	+3
Eif3l	N.C.	3	+3
Bcap31	N.C.	3	+3
Vwa5a.2	N.C.	3	+3
Sun1	N.C.	3	+3
Mtbp	9	27	+3
Vasp	4	12	+3
Rpn2	2	6	+3
tfap2d	N.C.	4	+4
Eif3a	N.C.	4	+4
Erlin-2-A	N.C.	4	+4
Evpl	N.C.	4	+4
Cse1l	N.C.	4	+4
Eif4enif1	N.C.	4	+4
ATP synthase subunit alpha	4	19	+4.8
Srrm1.L	N.C.	5	+5
Wwc1	N.C.	5	+5
Psmd14	N.C.	5	+5
Proteasome subunit alpha	N.C.	5	+5
Tpx2-a	N.C.	5	+5
Nucleolin	N.C.	5	+5
Top1.2	N.C.	5	+5
Trim28	N.C.	5	+5
Vitellogenin-A2	N.C.	5	+5
Coronin	N.C.	6	+6
Vasp	N.C.	6	+6
hsp90b1.S	N.C.	6	+6
slc37a2	N.C.	6	+6
Ipo5	N.C.	6	+6
Polq	N.C.	6	+6
Pdia3	N.C.	6	+6
Blm	N.C.	7	+7
Hal.2	N.C.	8	+8
Histone H2A	N.C.	8	+8
Pom121	N.C.	8	+8
Arpc2	N.C.	9	+9
Rae1/Gle2	N.C.	13	+13
Hspa5	N.C.	20	+20
Smarca5	N.C.	43	+43
Histone H4	N.C.	88	+88

2. Mass spec results of chromatin IP samples (as in Figure 3.1.12 C)

	p97i Chromatin			*Aph/caff chromatin*		
Protein	Control peptides	DONSON peptides	Fold enrichment	Control peptides	DONSON peptides	Fold enrichment
DONSON	N.C.	84	+84	N.C.	86	+86
Pik3c2a	N.C.	51	+51	-	-	-
Ubtf	N.C.	40	+40	N.C.	24	+24
Epabp-b	N.C.	40	+40	-	-	-
NPAT	N.C.	40	+40	N.C.	32	+32
Mel28	N.C.	37	+37	-	-	-
Tacc3	N.C.	33	+33	-	-	-
Cytosine-specific methyltransferase	5	121	+24.2	N.C.	23	+23
Etaa	N.C.	24	+24	N.C.	37	+37
Stim1	N.C.	23	+23	N.C.	11	+11
Cspp1	N.C	21	+21	N.C.	17	+17
Cenpt	N.C.	20	+20	-	-	-
Sart1	N.C.	19	+19	N.C.	6	+6
Ppp1cb	N.C.	19	+19	-	-	-
Ppp1ca	N.C.	18	+18	-	-	-
Gsn	N.C.	18	+18	3	20	+6.7
Nup153	N.C.	18	+18	-	-	-
Sltm	N.C.	17	+17	-	-	-
Chd7	N.C.	17	+17	-	-	-
Mycbp2	N.C.	17	+17	-	-	-
Perpherin	N.C.	16	+16	-	-	-
Wdr33	N.C.	16	+16	-	-	-
Cingulin like 1	N.C.	16	+16	-	-	-
Ncka5l	N.C.	15	+15	-	-	-
TopoI	N.C.	13	+13	-	-	-
Srsf6	N.C.	13	+13	-	-	-
Sec16a	N.C.	13	+13	-	-	-
coREST	N.C.	13	+13	-	-	-
Pola1	N.C.	12	+12	N.C.	10	+10
SALL4	N.C.	12	+12	N.C.	3	+3
KHC	N.C.	11	+11	-	-	-
Cltb-prov	N.C.	11	+11	-	-	-
Smarca4/BRG1	N.C.	11	+11	-	-	-
Srrm2	N.C.	10	+10	-	-	-
Tsc2/tuberin	N.C.	10	+10	-	-	-
Usp10-a	N.C.	9	+9	-	-	-
Pelp1	N.C.	9	+9	-	-	-
Rad50	N.C.	9	+9	5	76	+15.2
Np-a	N.C.	9	+9	-	-	-
Camsap1	N.C.	9	+9	-	-	-
Spats2	N.C.	9	+9	N.C.	5	+5
Nek9	N.C.	9	+9	-	-	-
Ppp1r9a	N.C.	8	+8	-	-	-
Ybx2-b	N.C.	8	+8	-	-	-
Fancm	N.C.	8	+8	-	-	-
Mdn1	N.C.	8	+8	-	-	-
Msh6	N.C.	8	+8	N.C.	12	+12

Atxn2l	N.C.	8	+8	-	-	-
Mta1	N.C.	8	+8	-	-	-
Cltb	N.C.	8	+8	-	-	-
Hcfc1	N.C.	8	+8	-	-	-
Zbtb17	N.C.	8	+8	-	-	-
Scaper	3	23	+7.3	-	-	-
Fanca	2	14	+7	-	-	-
Zmym3	N.C.	7	+7	-	-	-
Jmjd1c	N.C.	7	+7	-	-	-
Srsf4	N.C.	7	+7	-	-	-
Prpf40a	N.C.	7	+7	-	-	-
Srsf5	N.C.	7	+7	-	-	-
Shank2	N.C.	7	+7	-	-	-
Amot	N.C.	7	+7	-	-	-
Mapkbp1	N.C.	7	+7	-	-	-
Clathrin	98	615	+6.3	-	-	-
Gins2	N.C.	6	+6	-	-	-
Plk1	N.C.	6	+6	N.C.	3	+3
Abcf1	N.C.	6	+6	-	-	-
Protein phosphatase slingshot homolog	N.C.	6	+6	-	-	-
Phf14	N.C.	6	+6	-	-	-
Cdc5l	N.C.	6	+6	-	-	-
Fbl	N.C.	6	+6	-	-	-
Dock7	N.C.	6	+6	-	-	-
Cpsf1	N.C.	6	+6	-	-	-
Tjp1	50	294	+5.9	-	-	-
Rif1	9	53	+5.9	N.C.	7	+7
B52-prov	3	17	+5.7	-	-	-
Nucleoprotein TPR	4	22	+5.5	-	-	-
Fmn2	3	16	+5.3	-	-	-
Rpa1	N.C.	5	+5	100	229	+2.29
Fanca complementation group	N.C.	5	+5	-	-	-
Eftud2	N.C.	5	+5	-	-	-
Parp1	N.C.	5	+5	-	-	-
XNop56	N.C.	5	+5	-	-	-
Rfc4	N.C.	5	+5	N.C.	4	+4
Dynll2	N.C.	5	+5	-	-	-
Sumo1a	N.C.	5	+5	-	-	-
Tubgcp2	N.C.	5	+5	-	-	-
Tubgcp3	N.C.	5	+5	-	-	-
Shank2	N.C.	5	+5	-	-	-
Parn	N.C.	5	+5	-	-	-
Zc3h13	N.C.	5	+5	-	-	-
Kaz	N.C.	5	+5	-	-	-
Fip1l1	N.C.	5	+5	-	-	-

Kiaa1429	N.C.	5	+5	-	-	-
Thrap3	4	26	+4.5	-	-	-
Wdr76	5	22	+4.4	-	-	-
Rpl7a	20	85	+4.3	-	-	-
Nup153	31	128	+4.1	-	-	-
Eif4enif1	32	131	+4	-	-	-
Ybx2-a	3	12	+4	-	-	-
Kiaa1797	N.C.	4	+4	-	-	-
Hdlbp	N.C.	4	+4	-	-	-
H2afy2	N.C.	4	+4	-	-	-
Cdk11b	N.C.	4	+4	-	-	-
Adenine nucleotide translocase	N.C.	4	+4	-	-	-
Mdn1	N.C.	4	+4	-	-	-
Csnk2a2	N.C.	4	+4	-	-	-
Znf605	N.C.	4	+4	-	-	-
Srsf2	N.C.	4	+4	-	-	-
Clta	N.C.	4	+4	-	-	-
Nkap	N.C.	4	+4	-	-	-
Myl6	N.C.	4	+4	-	-	-
Shroom2	N.C.	4	+4	-	-	-
Mllt4	N.C.	4	+4	-	-	-
Dgcr8	N.C.	4	+4	-	-	-
Lsm14a-a	N.C.	4	+4	-	-	-
Mki67	37	145	+3.9	-	-	-
Xtr	28	100	+3.6	-	-	-
Mtbp	5	18	+3.6	N.C.	4	+4
Fxr1	15	53	+3.5	-	-	-
Filip1	2	7	+3.5	-	-	-
Chtf18	7	23	+3.3	-	-	-
Cabin1	29	93	+3.2	N.C.	16	+16
Npm2	12	38	+3.2	-	-	-
Atrx	5	16	+3.2	-	-	-
Rpl4-a	31	97	+3.1	-	-	-
Sep1	2	6	+3	-	-	-
Hdac1	N.C.	3	+3	-	-	-
Eif4g1	N.C.	3	+3	-	-	-
Psmc-3	N.C.	3	+3	-	-	-
Erlin2-a	N.C.	3	+3	N.C.	6	+6
Camk2	N.C.	3	+3	-	-	-
Ywhaq	N.C.	3	+3	-	-	-
Importin (alpha)	N.C.	3	+3	9	25	+2.8
Ubr5	N.C.	3	+3	-	-	-
Prpf19	N.C.	3	+3	-	-	-
Cpsf2	N.C.	3	+3	-	-	-
Sgk223	N.C.	3	+3	-	-	-
Asf1ab	N.C.	3	+3	-	-	-
Sun1	N.C.	3	+3	-	-	-
Nusap1	N.C.	3	+3	-	-	-

Oocyte zinc finger protein	N.C.	3	+3	-	-	-
Eif6	N.C.	3	+3	-	-	-
Srpk1	N.C.	3	+3	-	-	-
Fn1	N.C.	3	+3	-	-	-
Transformer 2	N.C.	3	+3	-	-	-
Hells	N.C.	3	+3	N.C.	21	+21
Xcdc6	N.C.	3	+3	-	-	-
RNA helicase II	N.C.	3	+3	-	-	-
Otud4	N.C.	3	+3	-	-	-
Pacs1	N.C.	3	+3	-	-	-
Sfpq	N.C.	3	+3	-	-	-
Dapk1	N.C.	3	+3	-	-	-
Cep85	N.C.	3	+3	-	-	-
Aim1	N.C.	3	+3	-	-	-
Irr1.L	N.C.	3	+3	-	-	-
Rpl4-b	39	112	+2.9	6	N.C.	-6
RNPC7	9	26	+2.9	N.C.	13	+13
Smarca5	94	263	+2.8	-	-	-
Prrc2c	79	217	+2.7	-	-	-
Eif4e1b	3	8	+2.7	N.C.	3	+3
Wdhd1	51	135	+2.6	-	-	-
Sugp2	15	39	+2.6	-	-	-
FFA-1	77	190	+2.5	-	-	-
Hira	41	103	+2.5	-	-	-
Timeless	10	25	+2.5	-	-	-
Dppa2	2	5	+2.5	N.C.	6	+6
Dynll1	2	5	+2.5	-	-	-
Flii	2	5	+2.5	-	-	-
Orc4	2	5	+2.5	-	-	-
Nup160	25	62	+2.5	2	22	+11
Pol epsilon	68	165	+2.4	N.C.	2	+2
Histone H2A	28	68	+2.4	3	6	+2
Usp7	7	17	+2.4	N.C.	5	+5
Plekhg3	5	12	+2.4	-	-	-
Ranbp2	47	106	+2.3	47	209	+4.4
Trip12	18	42	+2.3	-	-	-
Ehmt	7	16	+2.3	-	-	-
Gins4	6	14	+2.3	-	-	-
Plekhg1	3	7	+2.3	-	-	-
Bat2	105	226	+2.2	-	-	-
Rpl5a	59	127	+2.2	-	-	-
Chd4	16	104	+2.2	-	-	-
SSRP1	47	102	+2.2	N.C.	5	+5
Habp4	40	88	+2.2	-	-	-
Rps2	41	19	+2.2	-	-	-
Tcf20	10	22	+2.2	-	-	-
Smarcc1	6	13	+2.2	-	-	-
Snrnp200	5	11	+2.2	-	-	-
PRDX2	5	11	+2.2	N.C.	3	+3
Ahctf1	115	247	+2.1	N.C.	10	+10

Baz1b	109	217	+2.1	N.C.	6	+6
Rps4	59	126	+2.1	-	-	-
Histone H3.2	27	58	+2.1	-	-	-
POM121	33	68	+2.1	-	-	-
Aurora kinase B-A	10	21	+2.1	N.C.	7	+7
Smtn	9	19	+2.1	N.C.	9	+9
Seh1l-b	7	15	+2.1	-	-	-
Lgals3	27	13	+2.1	33	15	-2.2
Treslin	23	45	+2	-	-	-
Williams syndrome transcription factor	46	93	+2	-	-	-
Rpl6	9	18	+2	-	-	-
Tpx2-a	8	16	+2	-	-	-
Hic2	6	12	+2	-	-	-
Tfam	6	12	+2	-	-	-
Rbbp7	4	8	+2	-	-	-
Ankrd43	2	4	+2	-	-	-
Ldhb	2	4	+2	-	-	-
C3orf15	2	4	+2	-	-	-
Dctn1	N.C.	2	+2	-	-	-
Cofilin-1-A	N.C.	2	+2	-	-	-
Kif16b	N.C.	2	+2	-	-	-
Ubap2	N.C.	2	+2	-	-	-
Whsc1	N.C.	2	+2	-	-	-
Methionine aminopeptidase	N.C.	2	+2	-	-	-
Hnrnpc	N.C.	2	+2	6	N.C.	-6
Srp72	N.C.	2	+2	-	-	-
Znf638	N.C.	2	+2	-	-	-
Rnps1-b	N.C.	2	+2	-	-	-
Twinfilin-1	N.C.	2	+2	-	-	-
Hmg20a	N.C.	2	+2	-	-	-
Mlf1ip	N.C.	2	+2	-	-	-
Baz1a	109	271	+1.9	N.C.	5	+5
Gins1	7	13	+1.9	-	-	-
Rpl17	23	41	+1.8	-	-	-
TopoII	390	680	+1.7	-	-	-
Myo18a	115	186	+1.6	N.C.	10	+10
ISWI	195	320	+1.6	N.C.	46	+46
Histone H2B	57	89	+1.6	13	9	-1.4
SPT16	135	209	+1.5	4	42	+10.5
Rps3a-a	74	111	+1.5	4	N.C.	-4
Srgap2	36	53	+1.5	N.C.	9	+9
Vasp	57	84	+1.5	17	42	+2.5
Rpl7	46	30	+1.5	-	-	-
Eef2	161	218	+1.4	-	-	-
Supervillin	227	307	+1.4	10	82	+8.2
MCM3	137	186	+1.4	41	115	+2.8
Rpsa	78	108	+1.4	N.C.	3	+3

Rplp0	63	86	+1.4	N.C.	3	+3
Enah	167	212	+1.3	15	91	+6.1
MCM5	119	159	+1.3	21	98	+4.7
Capza1	86	113	+1.3	19	50	+2.6
Rps18	26	35	+1.3	-	-	-
Kif20a	13	17	+1.3	-	-	-
Nup43	7	9	+1.3	N.C.	3	+3
MCM7	139	162	+1.2	39	131	+3.4
Nup205	88	107	+1.2	22	144	+6.5
Cullin2	13	15	+1.2	-	-	-
Psmd3	8	6	-1.3	-	-	-
PSMD2	46	31	-1.5	N.C.	11	+11
Ythdf1	5	3	-1.7	-	-	-
Lzpb-a	47	25	-1.9	-	-	-
Gp37	46	23	-2	N.C.	7	+7
Hspa8	6	3	-2	-	-	-
Traip	4	2	-2	-	-	-
Ipo5	4	2	-2	-	-	-
Pl10-prov	2	N.C.	-2	-	-	-
GStm1	2	N.C.	-2	-	-	-
Peptidyl-prolyl cis-trans isomerase	2	N.C.	-2	N.C.	2	+2
Psmb3	2	N.C.	-2	-	-	-
Anp32b	2	N.C.	-2	-	-	-
Sptbn1	5	2	-2.5	-	-	-
Zp4	33	11	-3	33	14	-2.4
Profilin	6	2	-3	4	2	-2
Trim28	6	2	-3	-	-	-
Mdh1	3	N.C.	-3	-	-	-
Zp3.2	3	N.C.	-3	-	-	-
Prpf39.1	3	N.C.	-3	-	-	-
Psmd4	3	N.C.	-3	-	-	-
Glutathione S transferase	3	N.C.	-3	-	-	-
Ankrd17	13	4	-3.3	-	-	-
ZPY1	39	11	-3.5	-	-	-
Cse1l	29	8	-3.6	-	-	-
Psmd13	8	2	-4	-	-	-
Cand1	4	N.C.	-4	-	-	-
Gcn1l1	4	N.C.	-4	-	-	-
Pum1	4	N.C.	-4	-	-	-
Eif3e	4	N.C.	-4	-	-	-
Macf1	4	N.C.	-4	-	-	-
Prpf3	4	N.C.	-4	-	-	-
Aaas	4	N.C.	-4	-	-	-
Zp2	27	6	-4.5	N.C.	2	+2
Translational endoplasmic reticulum ATPase	5	N.C.	-5	-	-	-
Psmd12	5	N.C.	-5	-	-	-

Numa1	11	2	-5.5	-	-	-
Ube3a	6	N.C.	-6	N.C.	3	+3
St13	6	N.C.	-6	7	N.C.	-7
Zpy2	32	5	-6.4	-	-	-
Psmc5	7	N.C.	-7	-	-	-
Picalm.2	7	N.C.	-7	-	-	-
Eif3b	8	N.C.	-8	-	-	-
P44s10	8	N.C.	-8	-	-	-
Psmc1	10	N.C.	-10	-	-	-
Psmd11	11	N.C.	-11	N.C.	2	+2
Arhgap21-b	12	N.C.	-12	-	-	-
Programmed cell death protein	18	N.C.	-18	-	-	-
Plectin	-	-	-	N.C.	96	+96
ATR	-	-	-	N.C.	55	+55
Dbnl-a	-	-	-	N.C.	43	+43
Lamin B3 S	-	-	-	N.C.	40	+40
Pol3	-	-	-	N.C.	36	+36
Importin (beta)	-	-	-	N.C.	30	+30
Nup88B	-	-	-	N.C.	27	+27
Drebrin1	-	-	-	N.C.	26	+26
Incenp-a	-	-	-	N.C.	25	+25
Nup214	-	-	-	3	70	+23.3
Capza2	-	-	-	N.C.	21	+21
Incenp-b	-	-	-	N.C.	20	+20
Nup188	-	-	-	N.C.	17	+17
Smarcal1	-	-	-	N.C.	16	+16
Flnb	-	-	-	N.C.	16	+16
Rfc3	-	-	-	N.C.	16	+16
Neurabin	-	-	-	3	45	+15
Dna2	-	-	-	N.C.	14	+14
Ttc26	-	-	-	N.C.	14	+14
Inf2	-	-	-	N.C.	14	+14
C3	-	-	-	N.C.	13	+13
Nup133	-	-	-	N.C.	13	+13
Erlin2-b	-	-	-	N.C.	12	+12
Spice1	-	-	-	N.C.	11	+11
Orc1	-	-	-	N.C.	10	+10
Sytl2	-	-	-	N.C.	10	+10
Pcm1	-	-	-	N.C.	10	+10
Arpc5	-	-	-	N.C.	10	+10
Orc1	-	-	-	N.C.	10	+10
Seh1l-a	-	-	-	N.C.	9	+9
SIPA1L1	-	-	-	N.C.	9	+9
ZPAX	-	-	-	N.C.	8	+8
Tdrd3	-	-	-	N.C.	8	+8
Mprip	-	-	-	N.C.	8	+8
Cdca9 (borealin 2)	-	-	-	N.C.	8	+8
Flna	-	-	-	24	188	+7.8
B4	-	-	-	3	23	+7.7
Rfc2	-	-	-	N.C.	7	+7

Kif4	-	-	-	N.C.	7	+7
Arghap5	-	-	-	N.C.	7	+7
Ribophorin 2	-	-	-	N.C.	7	+7
Lamin B2	-	-	-	N.C.	7	+7
Rts	-	-	-	N.C.	7	+7
Rfc2	-	-	-	N.C.	7	+7
Topbp1	-	-	-	N.C.	7	+7
Itpr1	-	-	-	N.C.	7	+7
Nup107	-	-	-	8	55	+6.9
Rangap1	-	-	-	7	48	+6.9
Gphn	-	-	-	6	39	+6.5
Topoisomerase 3a	-	-	-	N.C.	6	+6
Fanci	-	-	-	N.C.	6	+6
Diaph1	-	-	-	N.C.	6	+6
Mre11	-	-	-	N.C.	6	+6
Luzp1	-	-	-	N.C.	6	+6
Ipo7	-	-	-	N.C.	6	+6
Cgn	-	-	-	17	98	+5.8
Nup155	-	-	-	16	90	+5.6
Nup85				5	27	+5.4
Topoisomerase 3b	-	-	-	N.C.	5	+5
Zbtb33	-	-	-	N.C.	5	+5
Vapb	-	-	-	N.C.	5	+5
Rpa2	-	-	-	9	43	+4.8
Lamin III	-	-	-	33	142	+4.3
Nup93	-	-	-	18	75	+4.2
Tpx2-b	-	-	-	N.C.	4	+4
Fancd2	-	-	-	N.C.	4	+4
Orc2	-	-	-	N.C.	4	+4
XHp1-gamma	-	-	-	N.C.	4	+4
Fam83h	-	-	-	N.C.	4	+4
Smpd4	-	-	-	N.C.	4	+4
Nup53	-	-	-	N.C.	4	+4
Nckipsd	-	-	-	N.C.	4	+4
Dhx35	-	-	-	N.C.	4	+4
Rfc2	-	-	-	N.C.	4	+4
Mi-2 histone deacetylase complex protein 66	-	-	-	N.C.	4	+4
Galectin	-	-	-	N.C.	4	+4
Cap1	-	-	-	N.C.	4	+4
Tpx2-b	-	-	-	N.C.	4	+4
GATAD2A	-	-	-	N.C.	4	+4
Rpa3	-	-	-	2	6	+3
Rps3a-b	-	-	-	N.C.	3	+3
Ppp1cc-a	-	-	-	N.C.	3	+3
Rad26	-	-	-	N.C.	3	+3
Msh2	-	-	-	N.C.	3	+3
Senp1	-	-	-	N.C.	3	+3
RPS3a-b	-	-	-	N.C.	3	+3

Upk3a	-	-	-	N.C.	3	+3
Smc1a	-	-	-	N.C.	3	+3
Ddost	-	-	-	N.C.	3	+3
Vcp	-	-	-	N.C.	3	+3
Eprs	-	-	-	N.C.	3	+3
PolD3	-	-	-	N.C.	3	+3
Gnb2l1	-	-	-	N.C.	3	+3
Uhrf1	-	-	-	N.C.	3	+3
Msh2	-	-	-	N.C.	3	+3
Rad26	-	-	-	N.C.	3	+3
SALL4	-	-	-	N.C.	3	+3
Rps19	-	-	-	N.C.	3	+3
Nup50.L	-	-	-	N.C.	3	+3
Rcc1	-	-	-	N.C.	3	+3
MCM4	-	-	-	37	107	+2.9
Lima1	-	-	-	31	90	+2.9
Dbnl-b	-	-	-	25	67	+2.7
Nupl1	-	-	-	2	5	+2.5
Kpnb1	-	-	-	35	85	+2.4
Capping protein beta	-	-	-	20	48	+2.4
Cttn	-	-	-	92	191	+2.1
Ubiquitin A-52	-	-	-	21	45	+2.1
Arpc1a	-	-	-	13	27	+2.1
Coronin	-	-	-	52	105	+2
Ran	-	-	-	3	6	+2
Prrc2b	-	-	-	N.C.	2	+2
Ppp1r12c	-	-	-	N.C.	2	+2
Zp2	-	-	-	N.C.	2	+2
Cyp	-	-	-	N.C.	2	+2
Akap17a	-	-	-	N.C.	2	+2
Ddx41	-	-	-	N.C.	2	+2
Vitellogenin	-	-	-	N.C.	2	+2
Nup37	-	-	-	N.C.	2	+2
MCM2	-	-	-	77	131	+1.7
xlZPC	-	-	-	7	11	+1.6
MCM6	-	-	-	80	115	+1.4
Pold2	-	-	-	6	4	-1.5
Prim1	-	-	-	6	4	-1.5
PCNA	-	-	-	5	3	-1.7 (???)
Rae1/Gle2	-	-	-	8	4	-2
Rpl9	-	-	-	2	N.C.	-2
Rpl11	-	-	-	2	N.C.	-2
Ywhaz	-	-	-	2	N.C.	-2
Hemoglobin subunit epsilon	-	-	-	2	N.C.	-2
Rpl9	-	-	-	2	N.C.	-2
Dld	-	-	-	2	N.C.	-2
Rpl11	-	-	-	2	N.C.	-2
Plch2	-	-	-	2	N.C.	-2
Trap1	-	-	-	2	N.C.	-2

Ndst1	-	-	-	2	N.C.	-2
Plekhm3	-	-	-	2	N.C.	-2
Sub1	-	-	-	2	N.C.	-2
Tubb4	-	-	-	18	8	-2.3
Programmed cell death 6-interacting protein	-	-	-	17	7	-2.4
Hsc70	-	-	-	134	44	-3
Pde4dip	-	-	-	3	N.C.	-3
Sri	-	-	-	3	N.C.	-3
serine/threonine-protein phosphatase 2A	-	-	-	3	N.C.	-3
Nup54	-	-	-	41	12	-3.4
Acly	-	-	-	4	N.C.	-4
Hnrnpa1	-	-	-	4	N.C.	-4
Cct7	-	-	-	4	N.C.	-4
Slc25a4	-	-	-	4	N.C.	-4
Gca	-	-	-	5	N.C.	-5
Anxa7	-	-	-	10	2	-5
Heterogeneous nuclear ribonucleoprotein A3 homolog 1	-	-	-	5	N.C.	-5
Jup	-	-	-	5	N.C.	-5
Hnrnph1	-	-	-	6	N.C.	-6
Heterogeneous nuclear ribonucleoprotein A2 homolog 2	-	-	-	6	N.C.	-6
Rps23	-	-	-	6	N.C.	-6
Histone H4	-	-	-	47	7	-6.7
ATP5a1	-	-	-	16	2	-8
Hsp90beta	-	-	-	81	10	-8.1
Ube2l3	-	-	-	10	N.C.	-10
Tuba1a	-	-	-	10	N.C.	-10
Mdh2	-	-	-	13	N.C.	-13
Hspa1a	-	-	-	11	N.C.	-11
Hsp70	-	-	-	16	N.C.	-16
L-lactate dehydrogenase	-	-	-	22	N.C.	-22
Hspa9	-	-	-	23	N.C.	-23
Eno3	-	-	-	23	N.C.	-23
Hsp90aa1.1	-	-	-	25	N.C.	-25
Tubb	-	-	-	176	7	-25.1
Fructose-bisphosphate aldolase	-	-	-	27	N.C.	-27
Tuba1b	-	-	-	126	4	-31.5
Eno1	-	-	-	35	N.C.	-35

Eef2.1	-	-	-	36	N.C.	-36
Hspd1	-	-	-	37	N.C.	-37
Hspa1b	-	-	-	38	N.C.	-38
ATP5b	-	-	-	41	N.C.	-41
Tuba4a	-	-	-	136	N.C.	-136

Nucleic Acids Research, 2023 **1**
https://doi.org/10.1093/nar/gkad694

DONSON facilitates Cdc45 and GINS chromatin association and is essential for DNA replication initiation

Georgia Kingsley, Aggeliki Skagia, Paolo Passaretti, Cyntia Fernandez-Cuesta, Alicja Reynolds-Winczura, Kinga Koscielniak and Agnieszka Gambus*

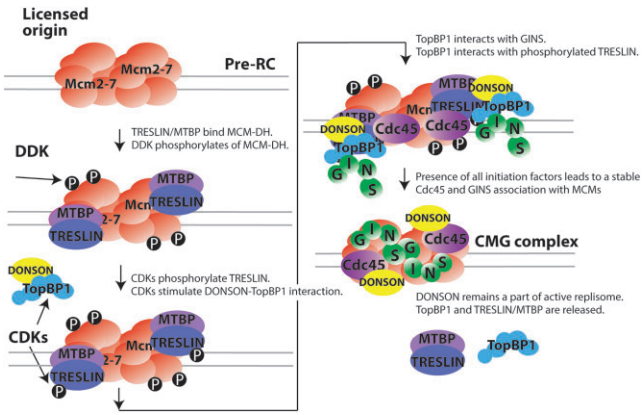
Institute of Cancer and Genomic Sciences, Birmingham Centre for Genome Biology, University of Birmingham, UK

Received May 25, 2023; Revised August 02, 2023; Editorial Decision August 03, 2023; Accepted August 17, 2023

ABSTRACT

Faithful cell division is the basis for the propagation of life and DNA replication must be precisely regulated. DNA replication stress is a prominent endogenous source of genome instability that not only leads to ageing, but also neuropathology and cancer development in humans. Specifically, the issues of how vertebrate cells select and activate origins of replication are of importance as, for example, insufficient origin firing leads to genomic instability and mutations in replication initiation factors lead to the rare human disease Meier-Gorlin syndrome. The mechanism of origin activation has been well characterised and reconstituted in yeast, however, an equal understanding of this process in higher eukaryotes is lacking. The firing of replication origins is driven by S-phase kinases (CDKs and DDK) and results in the activation of the replicative helicase and generation of two bi-directional replication forks. Our data, generated from cell-free *Xenopus laevis* egg extracts, show that DONSON is required for assembly of the active replicative helicase (CMG complex) at origins during replication initiation. DONSON has previously been shown to be essential during DNA replication, both in human cells and in *Drosophila*, but the mechanism of DONSON's action was unknown. Here we show that DONSON's presence is essential for replication initiation as it is required for Cdc45 and GINS association with Mcm2–7 complexes and helicase activation. To fulfil this role, DONSON interacts with the initiation factor, TopBP1, in a CDK-dependent manner. Following its initiation role, DONSON also forms a part of the replisome during the elongation stage of DNA replication. Mutations in *DONSON* have recently been shown to lead to the Meier-Gorlin syndrome; this novel replication initiation role of DONSON therefore provides the explanation for the phenotypes caused by *DONSON* mutations in patients.

GRAPHICAL ABSTRACT



INTRODUCTION

Faithful cell division is the basis for the propagation of life. The replication, repair and epigenetic regulation of the human genome are three of nine interrelated pathways involved in ageing (1,2) and defects in the maintenance of genome integrity are the foremost drivers of cellular deterioration (2). The mechanisms of origin firing during DNA replication initiation and replication fork progression have been well characterised and reconstituted in yeast, however, equal understanding of these processes in higher eukaryotes is lacking. Specifically, the issue of how vertebrates, including humans, select and activate origins of replication are not resolved conclusively and are of importance as, for example, areas with a paucity of replication origins are more prone to genomic instability (3), while mutations in factors involved in origin firing lead to the rare human disease Meier-Gorlin syndrome (4).

DNA replication occurs in three stages: initiation, elongation and termination. DNA replication initiates from

*To whom correspondence should be addressed. Tel: +44 1214158688; Email: a.gambus@bham.ac.uk

© The Author(s) 2023. Published by Oxford University Press on behalf of Nucleic Acids Research.
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

Downloaded from https://academic.oup.com/nar/advance-article/doi/10.1093/nar/gkad694/7252675 by guest on 29 August 2023

thousands of replication origins. These are the positions within the genome where replicative helicases become activated and start unwinding DNA while moving in opposite directions, away from each other, creating two DNA replication forks. The replicative helicase is composed of Cdc45, the Mcm2–7 hexamer and the GINS complex (CMG complex) (5); it is positioned at the tip of replication forks and forms a platform for replisome assembly (6). Activation of the CMG helicase (initiation stage) is divided into two cell cycle stages (Figure 1A): loading of the inactive core of Mcm2–7 in the form of double hexamers in the G1 stage of the cell cycle (origin licensing) and subsequent activation of a proportion of the double hexamers by Cdc45 and GINS during S-phase (origin firing). Once established, the replication forks replicate chromatin (elongation stage) until they encounter forks, moving in an opposing direction, which initiated from neighboring origins (Figure 1A). At this point, the termination of replication forks takes place.

During replication initiation, double hexamers of Mcm2–7 are first phosphorylated by Drf1/Dbf4-dependent Cdc7 kinase (DDK) (7,8) (Figure 1A). In *Saccharomyces cerevisiae*, DDK phosphorylation of Mcm2–7 promotes the recruitment of Sld3/Sld7 and Cdc45 to origins. This is followed by S-phase cyclin-dependent kinase (S-CDK) phosphorylation of Sld3 and Sld2, which supports the formation and recruitment to chromatin of the pre-landing complex (formed of Dpb11, Sld2, GINS and DNA Polε), allowing for CMG helicase assembly (9,10) (Figure 1A). DDK-dependent MCM phosphorylation is reversed by protein phosphatase 1 (PP1), which is targeted to chromatin by Rif1 (11–13). In higher eukaryotes, Treslin/MTBP complex is homologous to Sld3/Sld7 and TopBP1 protein to Dpb11. Both Treslin/MTBP and TopBP1 were shown to be essential for the loading of Cdc45 and GINS onto the Mcm2–7 hexamer in higher eukaryotes (14–18). In contrast, the functional homologue of Sld2 is still under debate. The N-terminus of RecQ4/RecQL4 helicase (*X.l.* RTS) contains a region of homology to Sld2 and RecQ4 is essential for initiation of DNA replication (19,20). However, in the *X. leavis* egg extract system, which provided much of this pioneering research, immunodepletion of RecQ4 does not affect the binding of Cdc45 and GINS to Mcm2–7. Rather, RecQ4 was found to promote a later stage of conversion of the pre-initiation complex (20–24).

Mutations in *DONSON* lead to microcephalic primordial dwarfism (MPD), which is a collective term for a group of human disorders characterised by intra-uterine and post-natal growth delay alongside marked microcephaly (4,25). DONSON was shown to interact with the replisome and is proposed to protect stalled DNA replication forks from the action of nucleases (25). Cells with siRNA downregulated DONSON (siDONSON) exhibit accumulation of stalled and asymmetric DNA replication forks and generation of double-stranded DNA breaks, which can be partially rescued by co-depletion of nucleases that can act on stalled forks (25). Moreover, DONSON was shown to be important for proper activation of the S-phase checkpoint response (25). Intriguingly, DONSON was shown to interact with only a subset of replisomes—specifically ones that were activated in early S-phase rather than late S-phase (26). During early S-phase, DONSON was shown to help replication forks traverse inter-strand crosslinks (ICL) (26). Interestingly, the *Drosophila* homologue of DONSON, *humpty dumpty* (*hd*), was also found to be essential for DNA replication, especially in early cleavage divisions of fly embryos, as *hd* knockout embryos arrested during the first few nuclear cleavage cycles with incomplete chromosome segregation (27,28). This suggests the evolutionary conservation of DONSON importance for cell proliferation in the developing embryo.

Altogether, we know several processes during DNA replication that DONSON is involved in and the consequences of DONSON's downregulation, however, the mechanisms by which DONSON delivers its functions and how they are regulated are unknown. We have therefore set out to determine the molecular mechanism of DONSON's action during DNA replication in the *Xenopus laevis* egg extract system, which is the only higher eukaryotic cell-free system able to undertake a full round of cell-cycle regulated DNA replication. We have found that *X.l.* DONSON indeed interacts with replisomes during DNA replication but surprisingly, we found that DONSON is required for initiation of DNA replication – more specifically for the activation of replicative helicase by loading of GINS and Cdc45.

MATERIALS AND METHODS

Inhibitors

MLN4924 (A01139, Active Biochem) was dissolved in DMSO at 5 mM and added to the extract 15 min after addition of sperm chromatin at 10 μM. NMS873 (17674, Cayman Chemical Company) was dissolved in DMSO at 10 mM and added to the extract 15 min after addition of sperm nuclei at 50 μM. Caffeine (C8960, Sigma) was dissolved in water at 100 mM and added to the extract along with demembranated sperm chromatin at 5 mM. Aphidicolin was dissolved in DMSO at 8 mM and added to the extract along with demembranated sperm chromatin at 40 μM. EcoR1 (R6011, Promega) was purchased at stock 12 U/μl and added to the extract at 0.05 U/μl. Mitomycin C (Calbiochem) was dissolved in water at 5 mM in water and used at 500 μM. Etoposide (Calbiochem) was dissolved in DMSO at 10 mM and used at 200 μM. Camptothecin was dissolved in water at 10 mM and used at 500 μM. Cdc7i PHA-767491 was dissolved in water and added at 100 μM.

Recombinant proteins

pET28a-DONSON, pET28a-DONSON-ΔLoop(P334-L410), pET28a-DONSON-ΔN(A2-G155), pET28a-DONSON-M463T, pET28a-DONSON-E521K, pET28a-DONSON-F171S, pET28a-DONSON-P230S and pET28a-DONSON-R217C vectors were used for protein expression in 2 l of BL21 (DE3) (1 mM IPTG was added at $OD_{600} = 0.6$, followed by incubation overnight at 18°C). Frozen bacterial pellets were lysed in resuspension buffer (500 mM NaCl, 50 mM NaH_2PO_4, 2 mM $MgCl_2$, 10% glycerol, pH 9, 1 μg/ml of each: aprotinin, leupeptin and pepstatin) and supplemented with 1 mg/ml lysozyme, 0.05% Tween-20 and BitNuclease. After sonication, the lysate was clarified by centrifugation at 14 000 g for 30 min at 4°C, and supernatants containing soluble proteins were

Downloaded from https://academic.oup.com/nar/advance-article/doi/10.1093/nar/gkad694/7252675 by guest on 29 August 2023

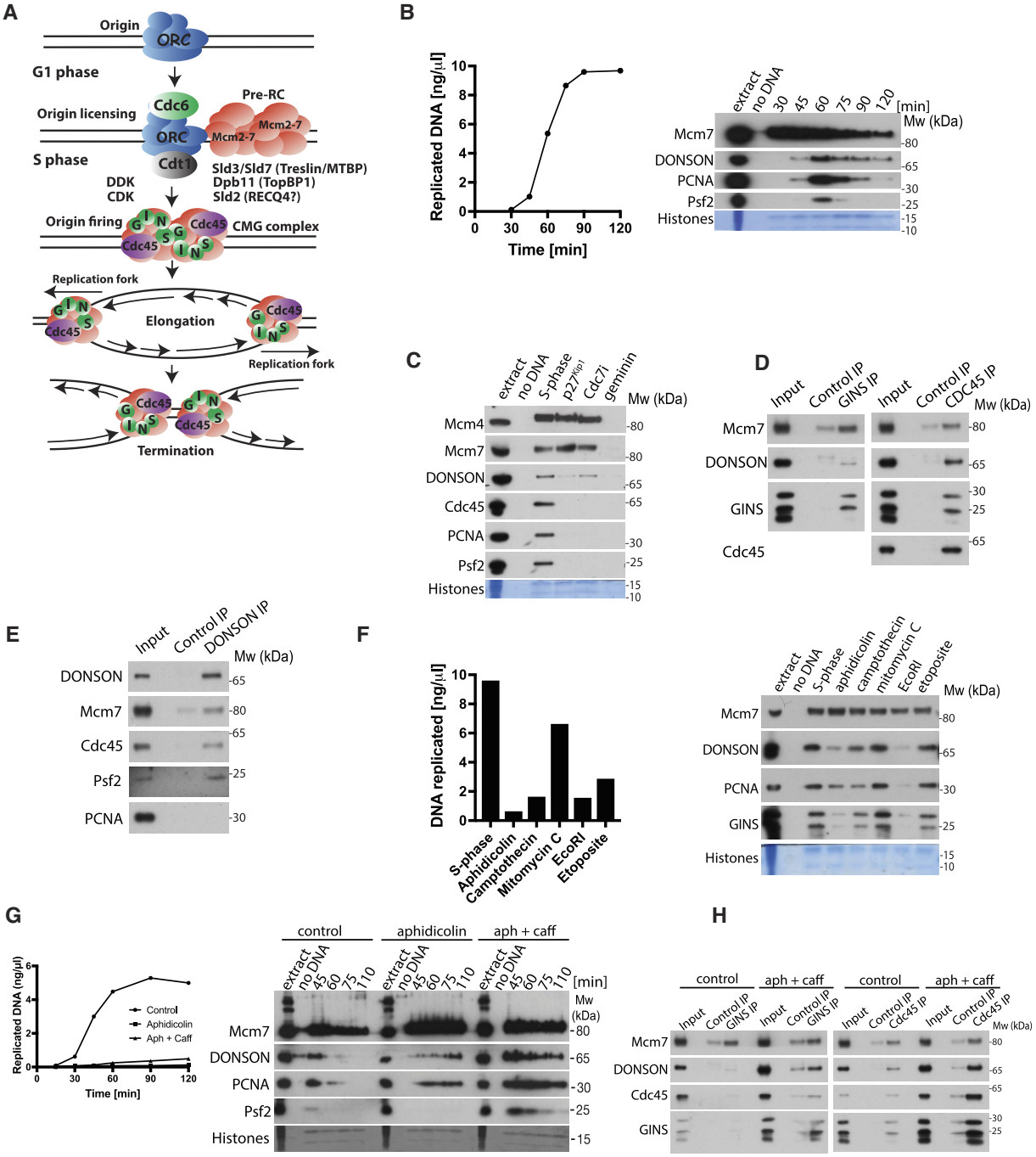


Figure 1. DONSON interacts with replisome on replicating chromatin in *Xenopus laevis* egg extract. (**A**) Model of eukaryotic DNA replication. *S. cerevisiae* initiation factors are listed with human homologues names in brackets. (**B**) DNA replication was set up in *Xenopus* egg extract and synthesis of nascent DNA followed by incorporation of α^{32}P-dATP into newly synthesised DNA (left). At the same indicated times chromatin was isolated and chromatin bound factors resolved on SDS-PAGE and immunoblotted with indicated antibodies (right). 'No DNA' control served as chromatin specificity control, while Coomassie stained histones from the bottom of the gel serve as a loading and purity control. (**C**) Replication reaction was set up in egg extract in optional presence of indicated replication initiation inhibitors. Chromatin fractions were isolated in the middle of S-phase and analysed by immunoblotting as in (B). (**D**) Chromatin replicating in egg extract was isolated in the middle of S-phase, chromatin bound protein complexes were released from chromatin by digestion of DNA with benzonase and replisomes immunoprecipitated using α-Cdc45 or α-GINS specific antibodies. Non-specific commercial IgG antibodies served as non-specific control. Input and immunoprecipitated samples were resolved on SDS-PAGE and immunoblotted with indicated antibodies. (**E**) Chromatin was prepared as in (D), but immunoprecipitation was carried out with α-DONSON antibody. (**F**) Replication reaction was set up in egg extract and treated optionally with indicated replication stress and DNA damaging agents. Chromatin was isolated after 60 min of reaction progression (middle/late stage of replication) and immunoblotted with indicated antibodies as in (B) (right), while the ability of the extract to synthesise nascent DNA was assessed after 90 min of replication reaction progression (left). (**G**) Replication reaction was set up in egg extract and optionally supplemented with aphidicolin and caffeine, as indicated. Progression of DNA synthesis during replication reaction was analysed (left), while chromatin bound factors were analysed by immunoblotting as in (B) (right). (**H**) Chromatin replicating in egg extract in optional presence of aphidicolin and caffeine was isolated, chromatin protein complexes released by DNA digestion and Cdc45 or GINS immunoprecipitated as in (D).

Downloaded from https://academic.oup.com/nar/advance-article/doi/10.1093/nar/gkad694/7252675 by guest on 29 August 2023

then incubated with 2 ml of pre-washed Super Ni-NTA Affinity Resin (SUPER-NINTA100, Generon) for 2 h, rotating at 4°C. The beads were then washed 3× with 30 ml of Lysis buffer, both supplemented with 15 mM imidazole. The beads were transferred to 10 ml columns (Poly-Prep Chromatography Column, Bio-Rad) and eluted with Lysis Buffer supplemented with 250 mM imidazole. Most concentrated elution fractions were dialysed against LFB1/50 buffer (10% sucrose, 50 mM KCl, 40 mM HEPES pH 8.0, 20 mM K phosphate pH 8, 2 mM $MgCl_2$, 1 mM EGTA, 2 mM DTT, 1 μg/ml of each: aprotinin, leupeptin and pepstatin) and snap frozen in liquid nitrogen. Recombinant p27(KIP1) was described previously (8) and used at 100 nM, while geminin^{DEL} was described in (29).

Antibodies

α-PCNA (P8825) and α-HIS (H1029) were purchased from Sigma; α-RIF1 (A300-568A) was purchased from Bethyl; α-PP1α (sc-271762) and α-PP2α (sc-56950) were purchased from Santa Cruz Biotechnology.

Affinity purified α-Cdc45, α-Psf2 (30), α-Mcm4 antibodies (31), α-Mcm7 (32), α-RecQ4 and α-TopBP1 antibodies were previously described (33). Similarly, α-GINS (34), α-Treslin and α-MTBP were described in (18).

Xenopus full-length DONSON protein was purified as described above and antibodies raised against such prepared antigens in sheep or rabbit. The resulting antibody sera was purified in-house against the purified antigen. The specificity of each new antibody is presented in Supplementary Figure S1.

Xenopus laevis egg extract preparation

All of the work with *Xenopus laevis* was approved by Animal Welfare and Ethical Review Body (AWERB) at University of Birmingham and approved by UK Home Office in form of Project License issued for Dr Agnieszka Gambus. *Xenopus laevis* egg extract was prepared as previously described (35).

Immunodepletion

DONSON immunodepletions with sheep antibodies were performed using Dynabeads Protein G (10004D, Life Technologies) coupled to antibodies against DONSON or nonspecific sheep IgGs (I5131, Sigma), with two rounds of 45 min incubation at 4°C. The DONSON antibodies were coupled at 600 μg per 1 ml of beads. Effective immunodepletion required two rounds of 45 min incubation of egg extract with antibody coupled beads at 50% beads ratio.

DONSON immunodepletions with rabbit antibodies were performed using Dynabeads Protein A (10008D, Life Technologies) coupled to *Xenopus* DONSON antibodies raised in rabbit and affinity purified or nonspecific rabbit IgG (I5006, Sigma). The DONSON antibodies were coupled at 600 μg per 1 ml of beads. Effective immunodepletion required two rounds of 45 min incubation of egg extract with antibody coupled beads at 50% beads ratio.

TopBP1 immunodepletions were performed using Dynabeads Protein G (10004D, Life Technologies) coupled to antibodies against TopBP1 or nonspecific sheep IgGs (I5131, Sigma). The TopBP1 antibodies were coupled at 600 μg per 1 ml of beads. Effective immunodepletion required two rounds of 45 min incubation of egg extract with antibody coupled beads at 50% beads ratio.

DNA synthesis assay

Interphase *X. laevis* egg extract was supplemented with 10 ng/μl of demembranated sperm chromatin and incubated at 23°C for indicated time. Synthesis of nascent DNA was then measured by quantification of α^{32}P-dATP (PerkinElmer) incorporation into newly synthesised DNA, as described before (35). The extract contains endogenous dNTP pools of ∼50 μM (36). The total amount of DNA synthesized, expressed as ng DNA / μl extract, can then be calculated by multiplying percent total ^{32}P incorporated by a factor of 0.654 (36). This calculation assumes an average molecular weight of 327 Da for dNMPs and equal quantities of all four dNTPs incorporated into DNA (weight of dNMP incorporated in ng/μl = percent total ^{32}P incorporated/$100 \times 50 \times 10^{-6} \times 4 \times 327 \times 10^3$) (36).

For quantification of replication efficiency between different experimental repeats in different extracts, the quantity of DNA replicated at the end of the reaction in IgG-depleted extract was set as 100% and the remaining values normalised to this.

Chromatin isolation time-course

Interphase *X. laevis* egg extract was supplemented with 10–15 ng/μl of demembranated sperm DNA and subjected to indicated treatments. The reaction was incubated at 23°C for indicated length of time and chromatin was isolated in ANIB100 buffer (50 mM HEPES pH 7.6, 100 mM KOAc, 10 mM MgOAc, 2.5 mM Mg-ATP, 0.5 mM spermidine, 0.3 mM spermine, 1 μg/ml of each aprotinin, leupeptin and pepstatin, 25 mM β-glycerophosphate and 10 mM 2-chloroacetamide (Merck)) as described previously (35).

During the chromatin isolation procedure, a sample without addition of sperm DNA (no DNA) is processed in an analogous way, usually at the end of the time course, to serve as a chromatin specificity control. The bottom of the PAGE gel on which the chromatin samples were resolved is cut off and stained with Colloidal Coomassie (SimplyBlue, Life Technologies) to stain histones which provide loading controls and indications of sample contamination with egg extract (cytoplasm).

Western blot quantification

The quantification of western blots is provided to indicate reproducibility of trends in experiments rather than to provide absolute values of increases or decreases in a signal. The quantified experiments were performed in different preparations of extracts and independently immunodepleted extracts to confirm that observed phenotypes are not specific for one extract prep.

The density of pixels of each band of the western blot and scanned Coomassie stained histones within the gel were quantified using Image J software. The numeric value in arbitrary units for each western blot band was normalised to

Downloaded from https://academic.oup.com/nar/advance-article/doi/10.1093/nar/gkad694/7252675 by guest on 29 August 2023

appropriate loading control (bands of Coomassie stained histones from the same sample). The analysis of IgG-depleted (control) and treatment samples was always done together and the fold difference between them calculated. Fold change from three repeated experiments is plotted on the graph with mean value and standard error of the mean (SEM) as calculated by GraphPad PRISM.

Immunoprecipitation from chromatin

275 μl of egg extract per IP was induced into interphase and mixed with 10–15 ng/μl demembranated sperm nuclei and optionally supplemented with the indicated treatments. The reaction was incubated at 23°C for the indicated time. Chromatin was isolated in ANIB100 (50 mM HEPES pH 7.6, 100 mM KOAc, 10 mM MgOAc, 2.5 mM Mg-ATP, 0.5 mM spermidine, 0.3 mM spermine, 1 μg/ml of each aprotinin, leupeptin and pepstatin, 25 mM β-glycerophosphate, 0.1 mM Na_3VO_4, 0.1% triton and 10 mM 2-chloroacetamide), and the chromatin pellets re-suspended in the same volume of original extract of ANIB100 containing 20% sucrose. Protein complexes were released from chromatin by digestion with 2 U/μl of Benzonase nuclease (E1014-25KU, Sigma) and sonicated for 5 min using a Diagenode sonicator with settings: 30 s on, 30 s off, medium setting. The insoluble fraction was then spun in a microfuge at 4°C, 10 min, 16k g.

Prepared beads:

- 30 μl of Dynabeads M-270 epoxy (14302D, Life Technologies) coupled covalently to 20 μg of affinity purified DONSON antibodies or IgG from sheep serum (I5131, Sigma);
- 100 μl of Dynabeads Protein A (10008D, Life Technologies) covalently coupled to 20 μg of affinity purified DONSON, affinity purified Treslin, or IgG from rabbit serum (I5006, Sigma) using BS3 crosslinker (S5799, Sigma);

were incubated with 220 μl digested chromatin at 4°C for 1–2 hours with rotation. Following the incubation time, beads were washed for 5 min rotating at 4°C twice with ANIB100, once with ANIB100 containing an additional 0.1% Triton X-100 and finally twice with ANIB100 buffer. Each sample was prepared by boiling in 50 μl of 2x NuPAGE LDS loading buffer (Life Technologies) for 5 min.

Immunoprecipitation from nucleoplasm

550 μl of egg extract per IP was induced into interphase and mixed with 10–15 ng/μl demembranated sperm chromatin and optionally supplemented with the indicated treatments. CDK inhibitor ($p27^{Kip1}$) and CDC7 inhibitor (PHA-767491) were added immediately after sperm DNA. The reaction was incubated at 23°C for 45 min. Nucleoplasm was isolated in ANIB100 (50 mM HEPES pH 7.6, 100 mM KOAc, 10 mM MgOAc, 2.5 mM Mg-ATP, 0.5 mM spermidine, 0.3 mM spermine, 1 μg/ml of each aprotinin, leupeptin and pepstatin, 25 mM β-glycerophosphate and 10 mM 2-chloroacetamide), and the nuclear pellets re-suspended in the same volume of original extract of ANIB100 containing 20% sucrose and 0.1% triton. Protein complexes were released from nucleoplasm by digestion with 2 U/μl of Benzonase nuclease (E1014-25KU, Sigma) and sonicated for 5 min using a Diagenode sonicator with settings: 30s on, 30s off, medium setting. The insoluble fraction was then spun in a microfuge at 4°C, 10 min, 16 000 rcf.

100 μl of Dynabeads Protein A (10008D, Life Technologies) covalently coupled to 20 μg of affinity purified DONSON or IgG from rabbit serum (I5006, Sigma) using BS3 crosslinker (S5799, Sigma); were incubated with 400 μl digested nucleoplasm at 4°C for 1–2 h with rotation. Following the incubation time, beads were washed for 5 min rotating at 4°C twice with ANIB100, once with LFB1/50 and finally twice with ANIB100 buffer. Each sample was prepared by boiling in 50 μl of 2× NuPAGE LDS loading buffer (Life Technologies) for 5 min.

Large scale immunoprecipitation of DONSON for mass spectrometry and CHROMASS

3.9 ml of *X. laevis* egg extract was activated and supplemented with 10 ng/μl of demembranated sperm DNA, 50 μM p97 inhibitor NMS873 or 40 μM aphidicolin + 5 mM caffeine and incubated at 23°C for 60 min. Chromatin was isolated in ANIB/100 buffer. Immunoprecipitation of DONSON was performed as described previously for p97 (37) and the immunoprecipitated material was analysed by mass spectrometry with Dr Richard Jones from MS Bioworks LLC.

Similarly, chromatin isolated from 10 μl of egg extract was processed in analogous way for mass spectrometry.

Sample preparation

Each sample was run on a 5–20% gradient gel (Invitrogen) for 1 cm and cut into 10 bands. Samples were submitted pre-plated for ten fraction analysis. Gel pieces were processed using a robot (ProGest, DigiLab) with the following protocol:

- Washed with 25 mM ammonium bicarbonate followed by acetonitrile.
- Reduced with 10 mM dithiothreitol at 60°C followed by alkylation with 50 mM iodoacetamide at RT.
- Digested with trypsin (Promega) at 37°C for 4 h.
- Quenched with formic acid and the supernatant was analysed directly without further processing.

Mass spectrometry

The gel digests were analysed by nano LC/MS/MS with a Waters M-class HPLC system interfaced to a ThermoFisher Oribitrap Fusion Lumos. Peptides were loaded on a trapping column and eluted over a 75 μm analytical column at 350 nl/min; both columns were packed with Luna C18 resin (Phenomenex). A 30 min gradient was employed. The mass spectrometer was operated in data-dependent mode, with MS and MS/MS performed in the Orbitrap at 60 000 resolution and 15 000 resolution, respectively. Advanced Peak Detection was turned on. The instrument was run with a 3 s cycle for MS and MS/MS. Proteome Discoverer v1.4 was used for peak generation.

Downloaded from https://academic.oup.com/nar/advance-article/doi/10.1093/nar/gkad694/7252675 by guest on 29 August 2023

Data processing

Data were searched using Mascot (Matrix Science, London, UK; version 2.8.0.1) with the following parameters:

Enzyme: Trypsin Fully Specific
Database: Uniprot Xenopus (forward and reverse appended with common contaminants) released on 15 April 2014. 79 274 (including reverse and CON) entries in the database were searched.
Fixed modification: Carbamidomethyl (C)
Variable modifications: Oxidation (M), Acetyl (Protein N-term), Deamidation (NQ), GlyGly (K), Phospho (STY)
Mass values: Monoisotopic
Peptide Mass Tolerance: 10 ppm
Fragment Mass Tolerance: 0.02 Da
Max Missed Cleavages: 2

Mascot DAT files were parsed into the Scaffold (version Scaffold_5.1.0, Proteome Software Inc., Portland, OR) software for validation, filtering and to create a non-redundant list per sample. Data were filtered with 1% protein and peptide false discovery rate (FDR) and requiring at least two unique peptides per protein.

Peptide identifications were accepted if they could be established at greater than 34.0% probability to achieve an FDR <1.0% by the Percolator posterior error probability calculation (38). Protein identifications were accepted if they could be established at greater than 99.0% probability to achieve an FDR less than 1.0% and contained at least two identified peptides. Protein probabilities were assigned by the Protein Prophet algorithm (39). Proteins that contained similar peptides and could not be differentiated based on MS/MS analysis alone were grouped to satisfy the principles of parsimony.

For calculation of fold enrichment for proteins with 0 peptides detected in control immunoprecipitation, that number was changed to 1 to allow for fold enrichment calculation.

AlphaFold analyses

AlphaFold (AF) modelling was carried out using the AF-multimer software (40,41) installed on BlueBEAR, the high-performance computing (HPC) cluster at the University of Birmingham. *Xenopus laevis* DONSON protein structures were predicted using the standalone version of AlphaFold (2.1.1-foss-2021a-CUDA-11.3.1). The AlphaFold models used were selected after producing five conformations ranked by their per-residue confidence scores (pLDDT) for each conformation (42). Differently, Human DONSON AF structure was downloaded from UniProt. UniProt protein sequences used are Q5U4U4 and Q9NYP3 for *Xenopus* and human respectively. Molecular graphics and analyses performed with UCSF ChimeraX, (43,44).

RESULTS

DONSON interacts with replisomes on chromatin in egg extract

DONSON was previously reported to interact with replisomes in human cells, especially those activated in early S-phase (25,26) and we have found DONSON interacting with terminated replisomes on chromatin in *Xenopus* egg extract (Supplementary Figure S1A). In this experiment, (described previously in (37)), a DNA replication reaction was set up in the egg extract, supplemented with inhibitors to block unloading of ubiquitylated terminated replisomes: cullin E3 ubiquitin ligases inhibitor (MLN4924) to block $Cul2^{Lrr1}$ from ubiquitylating Mcm7 within the terminated CMG helicases, and an ATPase-dead mutant of p97 segregase (p97mut). Such treatments led to the accumulation of large quantities of post-termination replisomes on chromatin. They were immunoprecipitated with Mcm3 antibodies and co-immunoprecipitated proteins were analysed by mass spectrometry (37). Many replisome components were identified this way and, interestingly, we also found multiple peptides of DONSON. As DONSON was reported to protect stalled replication forks in human cells (25), we hypothesised that it may also play a role in regulation of replisome disassembly during DNA replication. Therefore, we raised two antibodies, in sheep (α-DONSONs) and in rabbit (α-DONSONr) against recombinant *X.l.* DONSON purified from bacteria (Supplementary Figure S1B). We used these antibodies to analyse the pattern of DONSON chromatin binding during DNA replication in *Xenopus* egg extract.

Firstly, we analysed an unperturbed DNA replication reaction (Figure 1B). This reaction was initiated in egg extract by addition of the substrate DNA (demembranated sperm chromatin), and the progress of the replication reaction followed by incorporation of α^{32}P-dATP into the synthesised nascent DNA was measured in parallel to isolation of the chromatin bound proteins for western blot analysis (Figure 1B). As expected, we could observe Mcm7 binding to chromatin from the beginning of the reaction as origin licensing takes place within the first 5 min of the replication reaction in our system. The replisome components were present on chromatin at 60 min, which coincides with the fast rate of DNA synthesis in this experiment and thus a large number of replication forks replicating DNA. Finally, we could observe DONSON binding to chromatin at a similar time as the replisome factors but found that it was retained at a lower level for a longer time (Figure 1B). We decided therefore to test whether the observed DONSON interaction with chromatin depends on active replication (Figure 1C). For this, we inhibited DNA replication either through inhibition of origin licensing with geminin (Mcm4 and Mcm7 not present on chromatin), or through inhibition of CDK activity with $p27^{Kip1}$ (Mcm4 loaded and phosphorylated by DDK; no Cdc45, PCNA and Psf2 loading), or through inhibition of DDK activity with Cdc7i PHA-767491 (Mcm4 loaded but not phosphorylated; replication fork factors not present). In all situations where the DNA replication reaction was inhibited, we observed reduced binding of DONSON to chromatin (Figure 1C), suggesting that DONSON's interaction with chromatin is mostly dependent on active replication. The reduction of binding was reproducibly less profound upon Cdc7 inhibition suggesting that DONSON's chromatin interaction depends more on CDK activity. We verified also that once on chromatin during S-phase, DONSON indeed forms part of the replisome, as it can co-immunoprecipitate with

Downloaded from https://academic.oup.com/nar/advance-article/doi/10.1093/nar/gkad694/7252675 by guest on 29 August 2023

GINS (Psf2) and Cdc45 from a replicating chromatin fraction (Figure 1D and E).

As discussed previously, we observed DONSON interacting with immunoprecipitated terminated replisomes (Supplementary Figure S1A). To follow up this observation, we analysed DONSON chromatin binding when replisome unloading was inhibited either by inhibition of $Cul2^{Lrr1}$ activity using MLN4924 neddylation inhibitor (Supplementary Figure S1D) or by inhibiting the activity of p97 segregase with the NMS973 inhibitor (Supplementary Figure S1C). In both cases, we observed that DONSON accumulates on chromatin with similar timings as terminated replisomes. Altogether, these data suggest that DONSON continues interacting with replisomes at later stages of the replication.

As DONSON was reported in human cells to be important for cell survival when challenged with replication stress, and to fully activate S-phase checkpoint (25,26), we next analysed whether DONSON's chromatin binding activity is stimulated by different forms of replication stress and DNA damage (Figure 1F). In this case, the DNA replication reactions were set up in the presence of different stressors: aphidicolin—inhibitor of replicative polymerases; camptothecin – inhibitor of Topoisomerase 1, causing ssDNA breaks and DNA–protein crosslinks (DPCs); mitomycin C (MMC)—causing interstrand crosslinks (ICLs); EcoRI – restriction enzyme inducing double-stranded DNA breaks (DSBs); etoposide—inhibitor of Topoisomerase 2, causing DSBs and DPCs. All these treatments reduced the treated extract's ability to synthesise nascent DNA, albeit to different levels, and showed differing levels of reduction in replisome components binding to replicating chromatin. In all treatments, DONSON's chromatin binding pattern was most similar to that of other replisome components (PCNA and GINS), and we did not observe stimulation of DONSON chromatin association upon any of the treatments (Figure 1F).

To learn more about DONSON's interaction with chromatin upon replication stress and S-phase checkpoint activation, we also set up replication reactions in the presence of aphidicolin alone or with aphidicolin and caffeine together. Caffeine is a potent ATM/ATR inhibitor and in the presence of aphidicolin and caffeine, replication forks fire and stall due to inhibition of polymerases; however, as the checkpoint is not active, origins fire uncontrollably leading to an accumulation of high numbers of stalled replication forks on chromatin, while very little nascent DNA is being synthesised (Figure 1G). Importantly, under these conditions we can observe that DONSON accumulates on chromatin together with the stalled replication forks, suggesting that S-phase checkpoint activity is not needed for DONSON chromatin binding.

Finally, we wanted to determine whether DONSON still interacts with replisomes upon S-phase checkpoint inhibition. To achieve this, we isolated replicating chromatin challenged with aphidicolin/caffeine as above, digested DNA to release protein complexes from chromatin and immunoprecipitated GINS complex or Cdc45 (Figure 1H). Reassuringly, we observed that DONSON co-immunoprecipitated with GINS and Cdc45 together with other replisome factors, irrespectively of checkpoint inhibition. Comparing DONSON interactions with the replisome components on normal S-phase chromatin and chromatin from aphidicolin/caffeine treated extract, we can detect the same interactions, although there are more replisomes and more DONSON accumulated upon aphidicolin/caffeine treatment and thus we detect their interactions more robustly.

DONSON is essential for replication in *Xenopus* egg extract

Having established that DONSON is a component of normal replisomes in the *Xenopus* system, we aimed to determine the importance of DONSON for replication in the egg extract. To this end, we immunodepleted DONSON from egg extract using each of the raised antibodies. Both antibodies were able to immunodeplete DONSON to a level lower than 5% of DONSON remaining in the extract (Figure 2A, B). We then analysed the ability of IgG-depleted and DONSON-depleted extracts to synthesise nascent DNA. In both cases, the replication capacity of DONSON-depleted extract was very strongly diminished (Figure 2A, B). We also analysed chromatin binding of DNA replication factors in IgG- and DONSON-depleted extracts and could observe no DONSON, but also no Cdc45 and GINS binding to chromatin in the DONSON-depleted extract, which went hand in hand with the inhibition of DNA synthesis (Figure 2A, B). Importantly, we could rescue both the DNA synthesis and Cdc45 and GINS chromatin biding in DONSON-depleted extract through addition of recombinant *X.l.* DONSON purified from bacteria (Figure 2C and Supplementary Figure S1E). This rescue confirms that the only essential replication factor depleted in DONSON-depleted extract was DONSON itself. Addition of recombinant DONSON to DONSON-depleted egg extract readily rescues the levels of DNA synthesis in the extract, but fewer replication forks seem to be established to support this DNA synthesis. It is likely that DONSON purified from bacteria lacks important post-translational modifications and is not as active as the endogenous DONSON.

DONSON is essential for replication initiation

As DONSON was proposed to stabilise stalled replication forks in human cells (25), we next tested whether the inhibition of DNA replication in DONSON-depleted extract is due to the collapse of early replication forks, which could lead to S-phase checkpoint activation and inhibition of any further origin firing. To do so, we supplemented IgG- or DONSON-depleted extract with caffeine to block potential checkpoint activation (Figure 3A). Such treatment of DONSON-depleted extract did not rescue its ability to synthesise DNA, and importantly, did not rescue Cdc45 and GINS chromatin binding (Figure 3A) suggesting that it is not checkpoint activation that inhibits Cdc45 and GINS chromatin binding in DONSON-depleted extract. We also verified that DONSON depletion is not blocking the formation of the nuclear envelope around chromatin in the egg extract, which is essential for replication in our system (45) (data not shown).

During DNA replication initiation, origins are first licensed by loading of Mcm2–7 double hexamers onto chromatin. Mcm2–7 are then phosphorylated by DDKs and the

Downloaded from https://academic.oup.com/nar/advance-article/doi/10.1093/nar/gkad694/7252675 by guest on 29 August 2023

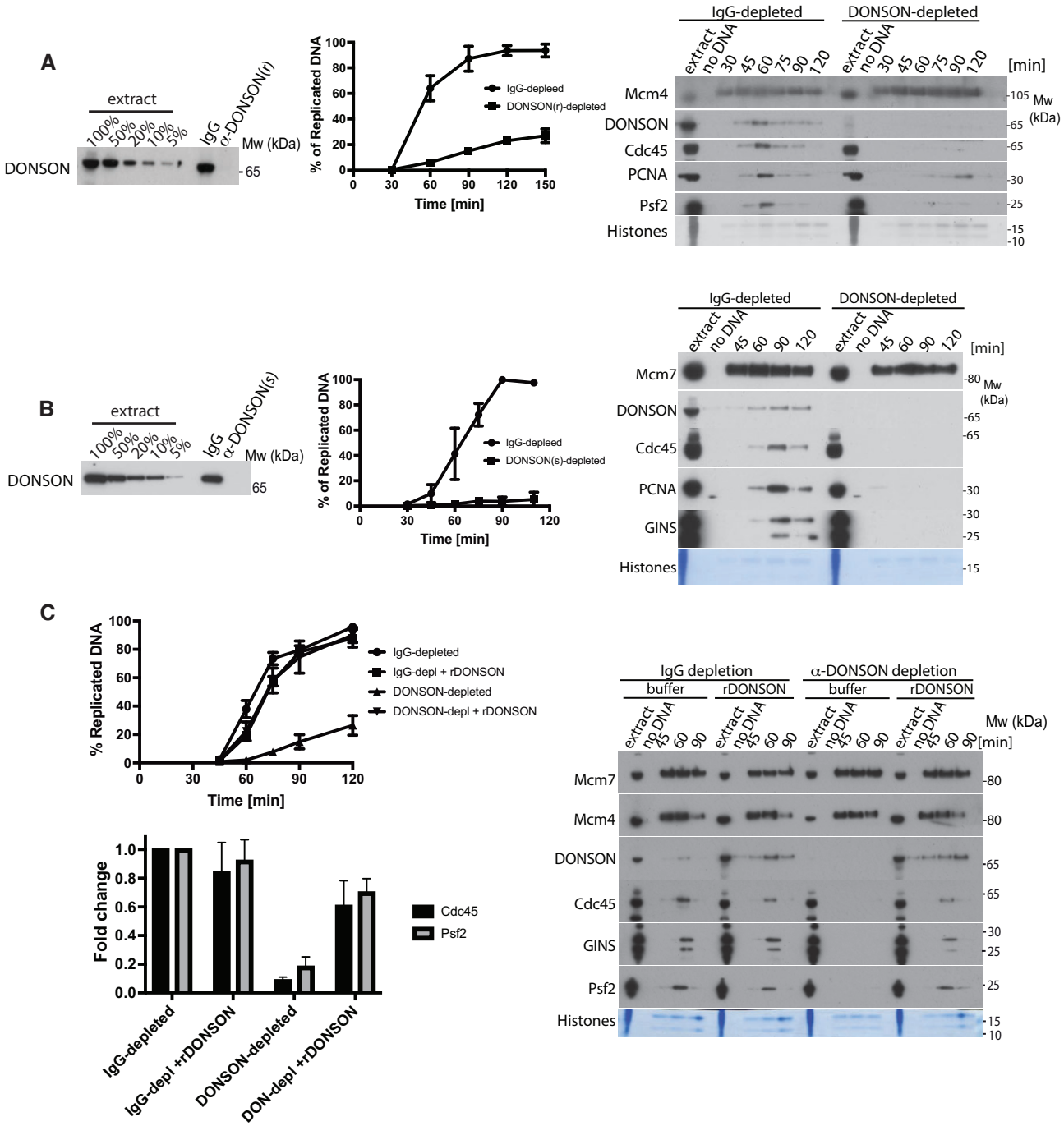


Figure 2. DONSON is essential for DNA replication in *X.laevis* egg extract. (**A**, **B**) *Xenopus* egg extract was immunodepleted of DONSON using either rabbit α-DONSON antibodies (**A**) or sheep α-DONSON antibodies (**B**). Nonspecific rabbit or sheep commercial IgGs were used for control IgG– depletions. The efficiency of depletion was established by immunoblotting in comparison to titration of original extract (left). The ability of immunodepleted extracts to synthesise nascent DNA was analysed by incorporation of α^{32}P-dATP into newly synthesised DNA (middle). Mean of $n = 5$ for rabbit DONSON depletion, $n = 3$ for sheep DONSON depletion is presented with SEM. The binding of replication fork factors in the absence of DONSON was analysed by immunoblotting of isolated chromatin fractions as in Figure 1B (right). (**C**) Replication reaction was set up in IgG- or DONSON-depleted extracts with optional addition of recombinant wt DONSON. The level of nascent DNA synthesis was followed as above (left top, $n = 3$, Mean with SEM is presented) and chromatin was isolated at indicated times and analysed by immunoblotting with indicated antibodies as in (A) (right). Quantification of fold change of chromatin bound Cdc45 and Psf2 at the peak of replication in IgG- and DONSON-depleted extract with optional addition of rDONSON over $n = 3$. The levels were normalised to level in IgG-depleted extract. Mean with SEM (left, bottom).

Downloaded from https://academic.oup.com/nar/advance-article/doi/10.1093/nar/gkad694/7252675 by guest on 29 August 2023

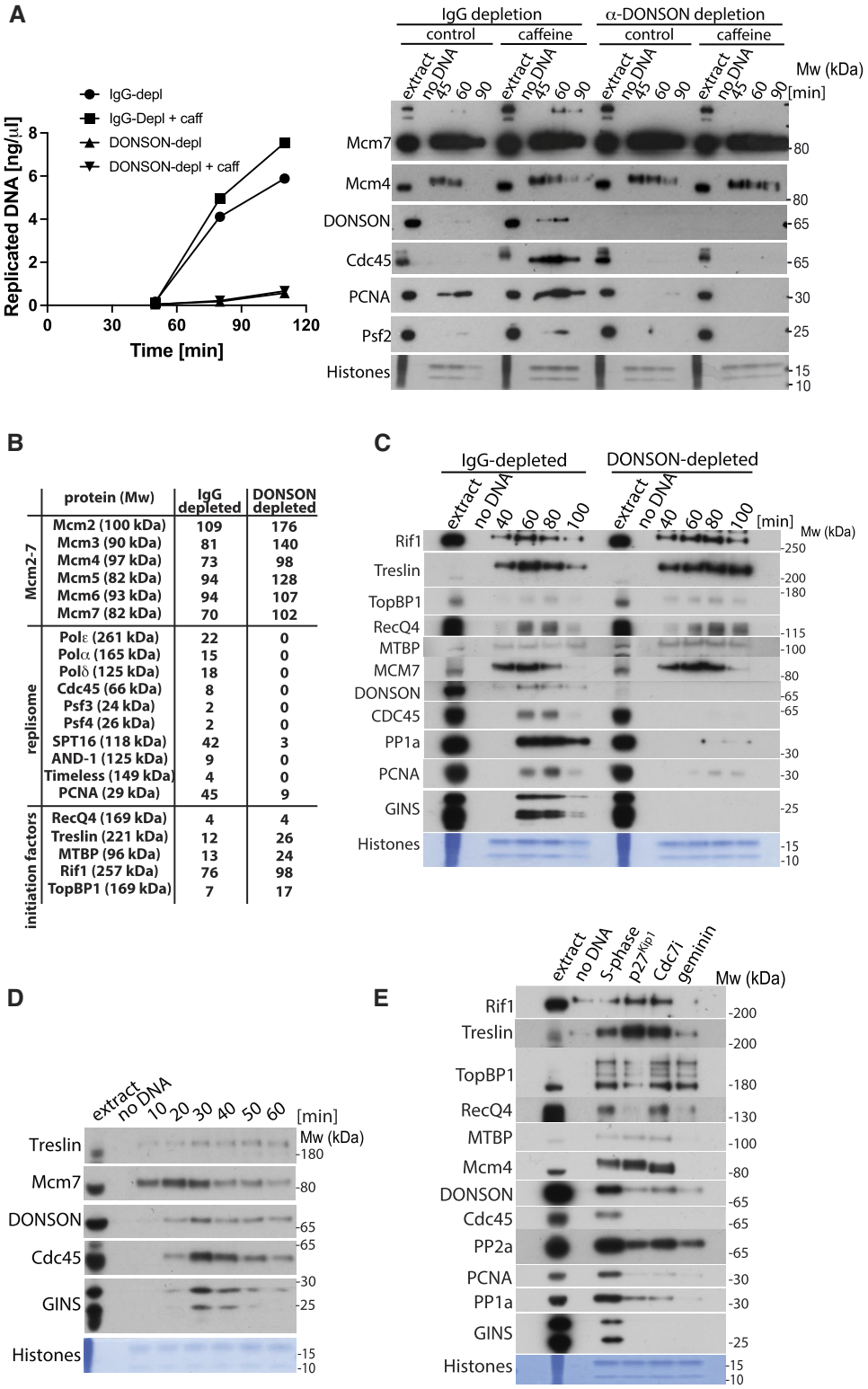


Figure 3. DONSON is essential for replication initiation in egg extract. (**A**) Replication reaction was set up in IgG- or DONSON-depleted egg extract optionally supplemented with caffeine. DNA synthesis was analysed throughout the reaction by incorporation of α^{32}P-dATP into newly synthesised DNA (left) and chromatin bound factors analysed at indicated times with indicated antibodies (right). (**B**) Egg extract was immunodepleted of DONSON with rabbit antibodies as in Figure 2A, chromatin was isolated in the middle of S-phase (45 min) and whole chromatin proteome of IgG- and DONSON-depleted reaction analysed by mass spectrometry. The total spectral count of selected replication factors is presented. (**C**) Replication reaction was set up in IgG- or DONSON-depleted egg extract and chromatin isolated at indicated times. Chromatin bound factors were analysed by immunoblotting with indicated antibodies as in Figure 1B. (**D**) Replication reaction was set up in *Xenopus* egg extract, chromatin isolated at indicated times during early stages of replication reaction and analysed as above. (**E**) Replication reaction was set up in egg extract in optional presence of indicated replication initiation inhibitors. Chromatin fractions were isolated in the middle of S-phase and analysed by immunoblotting as in (A).

Downloaded from https://academic.oup.com/nar/advance-article/doi/10.1093/nar/gkad694/7252675 by guest on 29 August 2023

CDKs to stimulate Cdc45 and GINS association through modification of several initiation factors. Our analyses of chromatin in DONSON-depleted extracts show that the origin licensing stage happens normally as we observe normal levels of chromatin-bound components of Mcm2–7 complexes (Figure 2). Moreover, analysis of chromatin-bound Mcm4 in DONSON-depleted extract revealed the same levels of phosphorylation as in IgG-depleted extract, suggesting that the DDK phosphorylation step is executed with similar efficiency in the absence of DONSON (Figure 2A and C, Mcm4). To understand, therefore, how DONSON functions to support DNA replication in the egg extract we first analysed by mass spectrometry the proteome of replicating chromatin in IgG- and DONSON-depleted extracts (CHROMASS) (Figure 3B). While we observed a complete absence of replisome components (Cdc45, GINS, polymerases, AND-1 etc.) in the DONSON-depleted extract, we also observed increased levels of Mcm2–7 components, presumably as they are not unloaded from chromatin due to inhibition of replication. Interestingly, we also observed higher levels of some of the initiation factors Treslin/MTBP complex and TopBP1 (Figure 3B), suggesting that the replication reaction is inhibited in DONSON-depleted extract at the stage of initiation. We confirmed these findings by western blotting in IgG- and DONSON-depleted extracts (Figure 3C), showing that without DONSON initiation factors like: Rif1, Treslin/MTBP, TopBP1 and RecQ4 all remain bound to chromatin throughout the time course of the reaction, while the replisome components are missing (Figure 3C).

DONSON's cooperation with other replication initiation factors

As DONSON seems to be involved in the initiation stage of DNA replication, we next examined in more detail the timings of chromatin association for DONSON and other initiation factors (Figure 3D). We set up an unchallenged DNA replication reaction in the egg extract, but this time analysed a pattern of protein binding to chromatin early on during the reaction, before onset of replication forks firing. As shown previously (Figure 1B), we can see binding of Mcm7 from the earliest time point (10 min) due to rapid origin licensing. Treslin starts binding to chromatin early (10–20 min), which is consistent with a previous report showing that the Treslin/MTBP complex binds chromatin before replication initiation (18). In this replication reaction, Cdc45 and GINS start binding chromatin at 20 min with a peak at 30–40 min. DONSON's pattern of chromatin binding mostly resembles that of Cdc45 and GINS (Figure 3D).

We also compared chromatin binding of DONSON and other initiation factors upon inhibition of replication initiation at different stages of the reaction: inhibition of licensing (geminin treatment), inhibition of DDK activity (PHA-767491 Cdc7i) and inhibition of CDK activity ($p27^{Kip1}$ inhibitor) (Figure 3E). Treatment with geminin was shown previously to significantly reduce chromatin binding of Treslin/MTBP and RecQ4 (18,20,46), while inhibition of CDK activity with $p27^{Kip1}$ led to Treslin/MTBP hyper-loading onto chromatin (18,46) and partially reduced loading of RecQ4 (18,20,46). Inhibition of DDK activity was shown to have a less profound of an effect on Treslin/MTBP chromatin binding (18). Our observations agree with these previous findings. Previous studies concerning TopBP1 however, showed that chromatin binding was either unaffected by all such treatments (47) or reduced to different levels by all of them (18,46). In our hands, we observe that TopBP1 chromatin binding is partially reduced after CDK inhibition. Some of these differences are likely to be the effect of different antibodies used in different studies. Interestingly, DONSON chromatin binding is reduced upon inhibition of S-phase kinases DDK and CDK, especially with the latter, and its chromatin binding pattern resembles much more the replisome components rather than other initiation factors (Figure 3E), suggesting that it is acting relatively late in the initiation cascade.

In an effort to understand the mechanism by which DONSON exerts its initiation function, we first compared interactors of DONSON on chromatin at early or late replication stages. To do so, we accumulated DONSON and other replisome components on chromatin either at the termination stage, by inhibiting replisome unloading with p97i (akin to the experiment in Supplementary Figure S1C), or at an earlier stage of replication by supplementing extract with aphidicolin and caffeine (akin to Figure 1F); here, the replisomes are stalled, but new origins keep firing due to checkpoint inhibition. We immunoprecipitated DONSON and analysed co-immunoprecipitated factors by mass spectrometry. We then calculated the enrichment of replisome and replication initiation factors in both IPs (Supplementary Figure S2A). Interestingly, we observed enrichment of peptides of TopBP1 and RecQ4 in the DONSON IP from aphidicolin/caffeine treated sample. We then analysed interactions between DONSON and initiation factors within the aphidicolin/caffeine chromatin fraction by immunoblotting (Figure 4A). Here we observed that DONSON interacts with the replisome components, as we have seen in Figure 1, but interestingly also with TopBP1, RecQ4 and Treslin/MTBP (Figure 4A). Importantly, reciprocal immunoprecipitation of TopBP1 and Treslin from aphidicolin/caffeine treated chromatin fraction shows analogous interactions with DONSON (Figure 4B). Altogether these data suggest again that DONSON is not only a part of the mature replisome, but also plays a role at the initiation stage.

DONSON interacts with TopBP1

We have shown that DONSON forms a part of the replisome during the initiation stage of DNA replication as it can interact with replication initiation factors on chromatin (Figure 4A and B). We next wanted to understand whether any of these interactions can occur directly with DONSON, outside the context of a replisome. We therefore isolated nuclei containing nucleoplasm (chromatin and nucleoplasm), in a buffer without addition of detergent, and immunoprecipitated DONSON. Interestingly, out of all potential interacting proteins tested, DONSON could interact with TopBP1 and this interaction was strongly reduced by $p27^{Kip1}$ CDK inhibitor (Figure 4C). We therefore partially depleted TopBP1 from egg extract to 25% of its original concentration (Supplementary Figure S2B) and

Downloaded from https://academic.oup.com/nar/advance-article/doi/10.1093/nar/gkad694/7252675 by guest on 29 August 2023

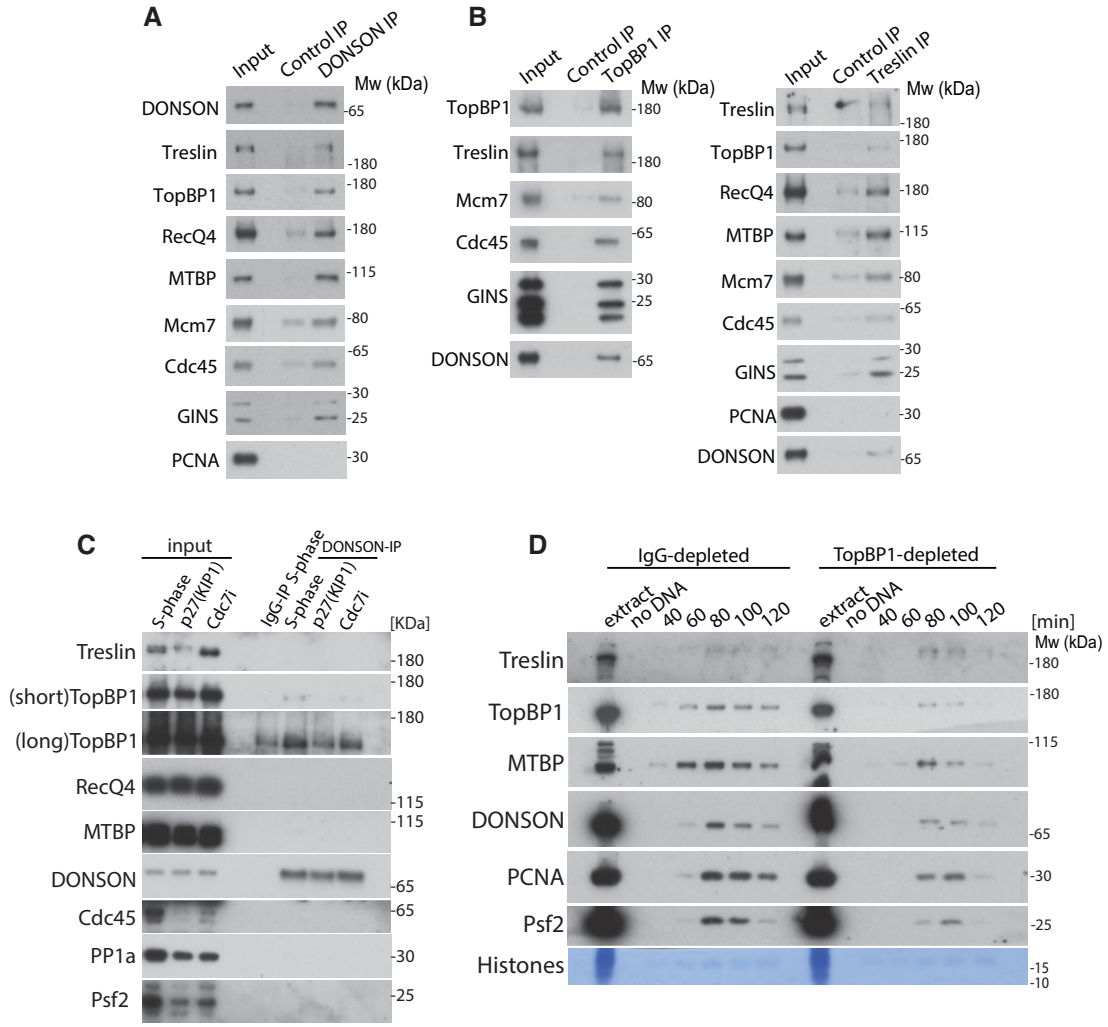


Figure 4. DONSON interacts with replication initiation factors. (**A**) Chromatin replication reaction was carried on in presence of aphidicolin and caffeine. Chromatin was isolated after 60 min of replication reaction, protein complexes were released from chromatin by digestion of DNA with benzonase and non-specific control antibodies or α-DONSON(s) antibodies used for immunoprecipitation. The input and immunoprecipitated samples were analysed by immunoblotting with indicated antibodies. (**B**) As (A) but α-TopBP1 or α-Treslin antibodies were used for immunoprecipitation. (**C**) DNA replication reaction was set up in egg extract with optional addition of replication initiation inhibitors ($p27^{KIP1}$ or Cdc7i). In the middle of S-phase, replicating nuclei were isolated in a buffer without detergent, which allows to retain part of the nucleoplasm. From this input material, DONSON was immunoprecipitated and immunoprecipitation samples analysed by immunoblotting with indicated antibodies. As a control, non-specific antibody immunoprecipitation from unchallenged S-phase input was performed in parallel. (**D**) TopBP1 was immunodepleted from egg extract to 25% of the original level. Replication reaction was set up in IgG- and TopBP1-depleted extract and chromatin isolated at indicated times of reaction and chromatin fractions analysed as above.

analysed whether chromatin binding of DONSON depends on TopBP1 (Figure 4D). With this, we observed that the level of chromatin-bound DONSON was strongly reduced. Altogether, all these data suggest that DONSON interacts with chromatin in a TopBP1- and CDK-dependent manner and acts as a replication initiation factor, essential for Cdc45 and GINS recruitment to the CMG complex.

DONSON's protein structure and mutations

We next wanted to explore whether DONSON's function may be conserved through evolution. The AlphaFold prediction of *X.l.* DONSON's structure suggests that DONSON is formed of a long, mostly unstructured N-terminus (D1), wrapping around a folded globular core with one protruding flexible loop (D3) (Figure 5A–C). Interestingly, the folded globular core of DONSON is highly conserved throughout evolution, while the unstructured N-terminus is not (Supplementary Figure S3A and B) and the flexible loop is partially conserved. We therefore wanted to examine whether this flexible parts of DONSON are important for its function. We purified DONSON without the N-terminus (DONSONΔN) or without the flexible loop (D3) (DONSONΔloop) (Supplementary Figure S3C) and tested whether they can rescue replication in the egg extract immunodepleted of DONSON. First, we titrated wt DONSON into DONSON-depleted extract to establish the minimal concentration needed for a DONSON rescue (Supplementary Figure S3D) and then tested the mutants (Figure 5E). Although both deletions generated proteins that were well expressed, soluble and could be easily purified (Supplementary Figure S3C), they could not rescue the ability of

Downloaded from https://academic.oup.com/nar/advance-article/doi/10.1093/nar/gkad694/7252675 by guest on 29 August 2023

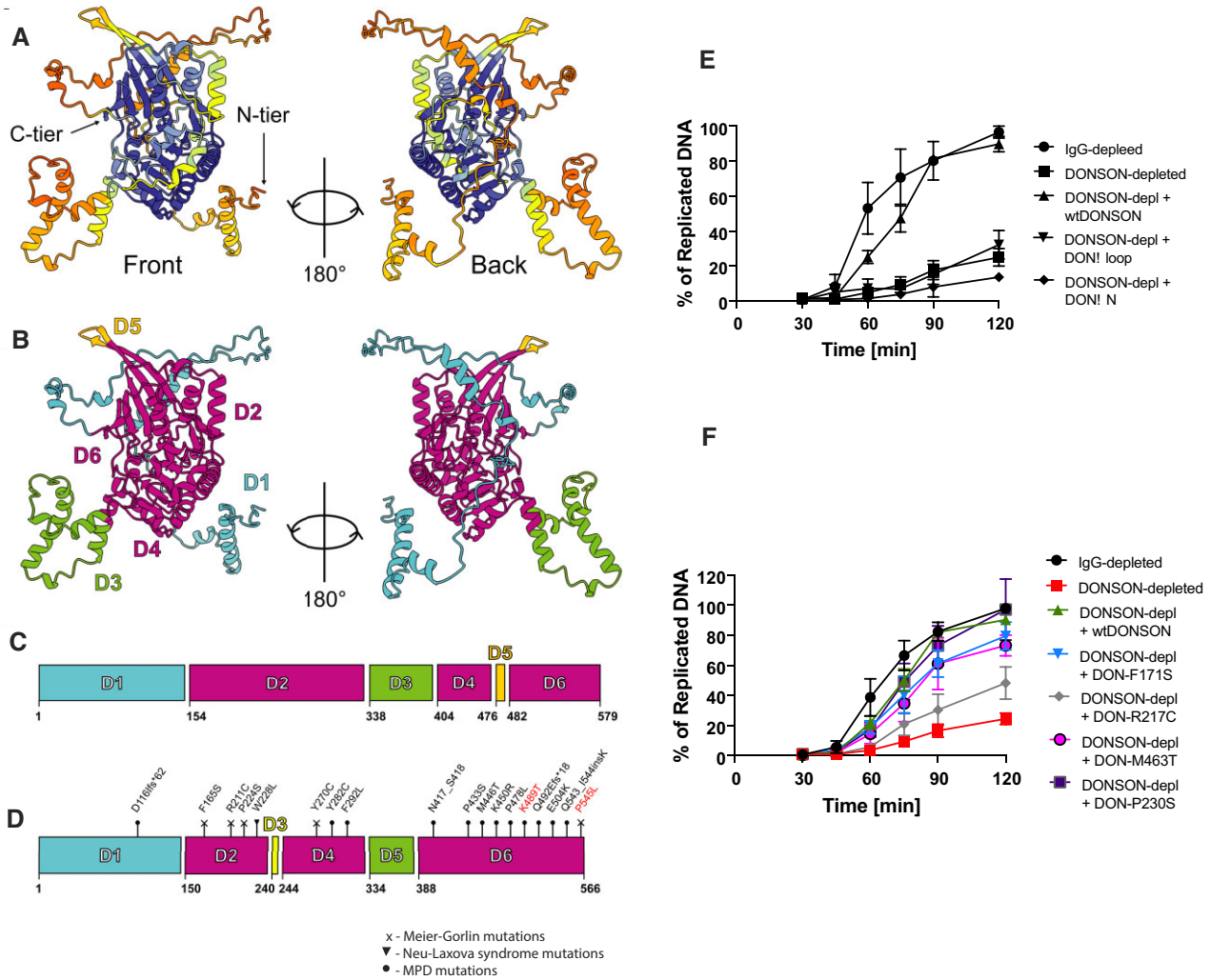


Figure 5. DONSON 3D structure organisation. (**A**) AlphaFold prediction of DONSON 3D structure. High confidence prediction in blue, low confidence prediction in yellow/orange. (**B, C**) Based on the predicted structure DONSON can be divided into six domains—N-terminal D1 domain which is unstructured, structured D2, flexible loop D3, structured D4 and D6 with a short turn motif D5. The 1D model of DONSON with colour coded domains (C) and 3D representation of these domains separation (B). (**D**) Mapping of the disease mutations described in (48) on human DONSON structure. Meier-Gorlin, Microcephalic Primordial Dwarfism MPD syndrome and Neu-Laxova syndrome mutations are mapped. The mutations marked in red do not have *Xenopus* homologous sequence. (**E**) DONSON mutant lacking unstructured N-terminus (DONSONΔN) or the flexible loop D3 from (B), (DONSONΔloop), was purified (Supplementary Figure 3C) and tested for rescue of DONSON-depleted egg extract in its ability to synthesise nascent DNA. (**F**) Indicated point mutants of DONSON (*Xenopus* DONSON with mutations homologous to patient derived human mutations) were purified (Supp Figure 3C) and tested as in (E).

the DONSON-depleted extract to synthesise nascent DNA (Figure 5E).

We also mapped known patient mutations causing microcephalic primordial dwarfism (MPD), specifically Meier-Gorlin syndrome in humans (48), onto the model structure of human DONSON (Figure 5D). Interestingly, most of the Meier-Gorlin mutations cluster within the human DONSON D2 domain (corresponding to the first half of the *X.l.*DONSON D2 domain). To test the effect of DONSON disease-causing mutations on replication initiation, we cloned and purified a selection of DONSON point mutations that were reported to cause MPD or Meier-Gorlin syndrome in humans (Supplementary Figure S3C). We chose two mutations leading to MPD: M446T (*Xenopus* M463T), which is a homozygous mutation found in 3 patients, E504K (*Xenopus* E521K), which was found in 2 patients as a compound heterozygous mutation combined with a stop codon introducing mutation. Interestingly, we could not efficiently express *X.l.* DONSON-E521K in bacteria, most likely due to protein destabilisation (Supplementary Figure S3C), while *X.l.* DONSON-M446T purified well and could fully restore the ability of the DONSON-depleted extract to synthesise nascent DNA (Supplementary Figure S3C and Figure 5F). We also chose three mutations leading to Meier-Gorlin syndrome: R211C (*Xenopus* R217C), which is a homozygous mutation found in two patients, and F165S (*Xenopus* F171S), which was found in one patient as a compound heterozygous mutation combined with an intronic splice site mutation, and another compound heterozygous mutation P224S (*Xenopus* P230S).

Downloaded from https://academic.oup.com/nar/advance-article/doi/10.1093/nar/gkad694/7252675 by guest on 29 August 2023

All three mutants were readily purified (Supplementary Figure S3C) and while *X.l.* DONSON-F171S and DONSON-P230S could rescue the ability of DONSON-depleted extract to synthesise nascent DNA, *X.l.* DONSON-R217C was partially defective (Figure 5F), suggesting that this mutation may disrupt the functionality of DONSON during replication initiation.

Altogether, these investigations revealed that flexible and poorly conserved fragments of DONSON are nevertheless essential for DONSON activity and that some of the patient mutations leading to the Meier-Gorlin syndrome phenotype can indeed disrupt the essential function of DONSON during DNA replication initiation.

DISCUSSION

Based on the data presented here we would like to propose that DONSON is a novel DNA replication initiation factor, as it is clearly essential for loading of Cdc45 and GINS onto origins during initiation. In the absence of DONSON, the replicative helicase, the CMG complex, cannot be formed and without the helicase activity, replication forks cannot be assembled. Mechanistically, our data suggest that DONSON acts through direct interaction with TopBP1, and this interaction is dependent on CDK activity. TopBP1, and its yeast homologue Dpb11/Cut5, is a master regulator of replication initiation. TopBP1 it is thought to bridge together Treslin/MTBP in complex with Cdc45, and RecQ4 in complex with GINS (Figure 1A) (14–20). However, immunodepletion of RecQ4 from *Xenopus* egg extract does not inhibit Cdc45 and GINS chromatin binding (20), suggesting that RecQ4, despite the partial homology to Sld2, does not play an equivalent role in higher eukaryotes.

TopBP1 acts as a scaffold protein, able to make numerous protein-protein interactions, facilitated in most cases through the multiple BRCT domains. BRCT domains most often interact with phosphorylated partners and often act in pairs. Yeast Dpb11 (*S. cerevisiae*) and Cut5 (*S. pombe*) contain 4 BRCT domains each, with BRCT1 + 2 interacting with phosphorylated Sld3 (9) and BRCT3 + 4 interacting with phosphorylated Sld2 (49,50). However, the human and *Xenopus* TopBP1 homologues are much larger proteins, containing eight BRCT domains (1-8), with an additional N-terminal BRCT domain identified in human TopBP1, named BRCT0 (51). This evolution of multiple additional BRCT domains suggests that higher eukaryotic TopBP1 is likely to have additional interactors, functions and regulations than the yeast counterparts. Like the yeast protein however, TopBP1 binds with phosphorylated Treslin (at S1001 in human and S976 in *Xenopus*) through its BRCT1 + 2 domains (14). TopBP1 also contains a GINS interacting motif (GINI), which lies between BRCT3 and BRCT4 domains (52), and was proposed recently to interact directly with GINS using BRCT4 + 5 domains (53). Interestingly, although BRCT4 + 5 are homologous to fungal BRCT3 + 4, it is BRCT7 + 8 that binds RecQ4, but this interaction is not dependent on CDK phosphorylation in either species (21,54). Moreover, TopBP1 BRCT7 + 8 are not essential for replication initiation at all (46), strengthening the argument that RecQ4 cannot be the functional homologue of yeast Sld2.

AlphaFold prediction of the DONSON-TopBP1 interaction (not shown) suggests that DONSON interacts with BRCT3 of TopBP1. Interestingly, it was shown that although the N-terminal fragment of TopBP1 spanning BRCT1 + 2 is sufficient to interact with Treslin, it is not sufficient to support DNA replication in the *Xenopus* egg extract. Instead, the TopBP1 fragment including BRCT1-3 is essential and sufficient to support replication initiation (46). This suggests an essential initiation role for BRCT3 in TopBP1, which thus far was not assigned to interact with any specific initiation factor. DONSON could therefore be such a factor. Interestingly, DONSON's interaction with TopBP1 is dependent on CDK activity, as would be expected for the functional homologue of Sld2. It remains to be determined however, whether this interaction is driven by direct phosphorylation of DONSON by CDKs, or perhaps phosphorylation of another factor.

We propose here that DONSON likely fills the role of a functional homologue of Sld2 in higher eukaryotes. In yeast, Dpb11-Sld2 forms a 'pre-landing complex' together with GINS and DNA Pol ε (55). In our analyses of DONSON interactions, we could detect DONSON interacting with GINS and DNA Pol ε, but only from chromatin fractions, where many interactions can be stimulated by formation of replisomes around the CMG complex. The only robust interaction that we found for DONSON within the nucleoplasmic fraction was that with TopBP1. It is not yet clear, therefore, whether DONSON and TopBP1 indeed form an analogous pre-landing complex. How else could DONSON act during origin firing to stimulate Cdc45 and GINS interaction with Mcm2–7? The molecular mechanism underpinning how TopBP1 and Treslin/MTBP deliver GINS and Cdc45 to Mcm2–7 remains unknown. It is clear from the work of many groups that Treslin/MTBP bind Mcm2–7 early in the initiation reaction. Their interaction is strengthened by DDK activity but does not require CDK activity; on the contrary, they are hyperloaded upon inhibition of CDKs (18,46). Phosphorylation of Treslin/MTBP by CDKs stimulates their interaction with N-terminal BRCT repeats of TopBP1 and is essential for helicase activation (14–15,46). TopBP1 can interact with GINS and is clearly essential for Cdc45 and GINS chromatin binding (47,53). We show here that DONSON interaction with TopBP1 depends on CDK activity. It is likely therefore that upon CDK activation, both TopBP1 and DONSON are recruited to phosphorylated Treslin/MTBP at Mcm2–7 double hexamers. The presence of all these initiation factors together allows for stable recruitment of Cdc45 and GINS into a CMG complex and helicase activation. TopBP1 is a limiting factor for origin firing (56) and is likely to be released following helicase activation, to facilitate firing of further origins, while DONSON is likely to be retained as part of the replisome in *Xenopus* egg extract.

Our data clearly indicate that DONSON plays an essential role in initiation of DNA replication, but DONSON has been shown previously in human cells to regulate the progression of replication forks, especially in response to replication stress (25,26). However, the molecular function of DONSON has largely been investigated in highly unstable HeLa cells (25,26), which often overexpress replication factors to support fast proliferation. Moreover, DONSON

Downloaded from https://academic.oup.com/nar/advance-article/doi/10.1093/nar/gkad694/7252675 by guest on 29 August 2023

was always downregulated using siRNA (which takes several days to efficiently downregulate the protein level) or through analysis of patient-derived mutants, which destabilise the protein and lead to continuous lower levels of DONSON (25,26). Finally, the effects of DONSON downregulation on DNA replication were analysed by DNA fibre analyses or creation of DNA damage, which would miss cells that do not initiate DNA replication. Therefore, it is possible that the initiation phenotype of DONSON downregulation had been missed thus far in human cells and awaits careful dissection and determination. Interestingly, DONSON was also proposed to specifically form part of early firing replisomes, and not the ones in late S-phase (26). This observation raises the possibility that DONSON specifically activates early firing origins, in contrast to Rif1, which is proposed to regulate firing of mainly late replication origins (57). Such a specified role for DONSON would also explain why it is so essential during embryogenesis in both *Xenopus* egg extract and in *Drosophila* embryos (27,28), as both of these exemplify embryonic systems, which have very fast DNA replication and lack replication timing programmes, so all origins can be regarded as early firing.

Downregulation of DONSON in human cells has been shown to diminish the ability of the fired replication forks to progress past replication obstacles and causes activation of S-phase checkpoint (25,26). We observed here that DONSON interacts with replisomes not only during replication initiation but also at later stages of the DNA replication reaction. DONSON accumulates with the rest of the replisome on chromatin upon inhibition of unloading of post-termination replisomes and interacts with post-termination replisomes. It is very likely therefore that after facilitating the initiation of replication machinery, DONSON travels with the replisomes and regulates fork progression. To study such a function of DONSON, we need to develop separation-of-function mutants, that can deliver the initiation function but are defective in the fork progression function. It is possible that some patient mutations may deliver such tools for the future.

Mutations in DONSON have been known for a while to be the underlying genetic problems, which lead to microcephalic primordial dwarfism (MPD). Interestingly, mutations in DONSON were also recently shown to specifically cause Meier-Gorlin syndrome (48,58), the subtype of MPD caused by dysregulation of DNA origin firing and replication initiation. The discovery of a replication initiation role for DONSON therefore explains the phenotypes caused by these mutations. It is interesting to speculate that if DONSON has multiple roles during DNA replication, different mutations within the protein can specifically disrupt these different functions. Most DONSON patient mutations have been shown to destabilise the protein and lead to much lower levels of DONSON protein in the cells (25). Nevertheless, it is possible that some point mutations (e.g. R211C) disrupt specifically its interaction with TopBP1, while others may destabilise interactions with other replication factors important for fork progression and checkpoint activation. More research is needed to answer all these questions.

Finally, in the last two years, researchers have reached a better understanding of the importance of DONSON in cancer development. DONSON was found to be a biomarker for risk stratification in clear cell renal carcinoma (ccRCC) (59), and in prostate carcinoma (PCa) (60). It has also been linked with a malignant phenotype in ccRCC cell culture models (61). Furthermore, it was shown to be a driver of breast cancer progression and a potential target for subsequent therapies (62). Understanding the different molecular functions of DONSON during DNA replication provides us with a solid foundation for translation of DONSON's potential as a biomarker and as a therapy target in the future.

DATA AVAILABILITY

The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE repository with the dataset identifier:

Project Name: Chromatin proteome of S-phase in IgG- and DONSON-depleted egg extract

Project accession: PXD042435

Project DOI: 10.6019/PXD042435

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

This work was supported by the MRC funded IMPACT PhD studentship for G.K. and BBSRC MIBTP studentship for C.F.C. Wellcome Trust Investigator Award (215510/Z/19/Z) funded P.P. and A.G, while BBSRC BB/T001860/1 funded A.S. and A.R.-W. Authors would like to thank Dr Rebecca Jones and Dr Neville Gilhooly for critical reading and discussions of the manuscript.

FUNDING

Medical Research Council [Impact PhD Studentship]; Biotechnology and Biological Sciences Research Council [BB/T001860/1, MIBTP PhD Studentship]; Wellcome Trust [215510/Z/19/Z]. Funding for open access charge: Wellcome Trust.

Conflict of interest statement. None declared.

REFERENCES

1. Lopez-Otin,C., Blasco,M.A., Partridge,L., Serrano,M. and Kroemer,G. (2013) The hallmarks of aging. *Cell*, **153**, 1194–1217.
2. Vermeij,W.P., Hoeijmakers,J.H. and Pothof,J. (2016) Genome integrity in aging: human syndromes, mouse models, and therapeutic options. *Annu. Rev. Pharmacol. Toxicol.*, **56**, 427–445.
3. Zeman,M.K. and Cimprich,K.A. (2014) Causes and consequences of replication stress. *Nat. Cell Biol.*, **16**, 2–9.
4. Klingseisen,A. and Jackson,A.P. (2011) Mechanisms and pathways of growth failure in primordial dwarfism. *Genes Dev.*, **25**, 2011–2024.
5. Moyer,S.E., Lewis,P.W. and Botchan,M.R. (2006) Isolation of the Cdc45/Mcm2–7/GINS (CMG) complex, a candidate for the eukaryotic DNA replication fork helicase. *Proc. Nat. Acad. Sci. U.S.A.*, **103**, 10236–10241.
6. Gambus,A., Jones,R.C., Sanchez-Diaz,A., Kanemaki,M., van Deursen,F., Edmondson,R.D. and Labib,K. (2006) GINS maintains association of Cdc45 with MCM in replisome progression complexes at eukaryotic DNA replication forks. *Nat. Cell Biol.*, **8**, 358–366.

Downloaded from https://academic.oup.com/nar/advance-article/doi/10.1093/nar/gkad694/7252675 by guest on 29 August 2023

7. Jares,P., Luciani,M.G. and Blow,J.J. (2004) A Xenopus Dbf4 homolog is required for Cdc7 chromatin binding and DNA replication. *BMC Mol. Biol.*, **5**, 5.
8. Walter,J.C. (2000) Evidence for sequential action of cdc7 and cdk2 protein kinases during initiation of DNA replication in Xenopus egg extracts. *J. Biol. Chem.*, **275**, 39773–39778.
9. Zegerman,P. and Diffley,J.F. (2007) Phosphorylation of Sld2 and Sld3 by cyclin-dependent kinases promotes DNA replication in budding yeast. *Nature*, **445**, 281–285.
10. Tanaka,S., Umemori,T., Hirai,K., Muramatsu,S., Kamimura,Y. and Araki,H. (2007) CDK-dependent phosphorylation of Sld2 and Sld3 initiates DNA replication in budding yeast. *Nature*, **445**, 328–332.
11. Poh,W.T., Chadha,G.S., Gillespie,P.J., Kaldis,P. and Blow,J.J. (2014) Xenopus Cdc7 executes its essential function early in S phase and is counteracted by checkpoint-regulated protein phosphatase 1. *Open Biol.*, **4**, 130138.
12. Alver,R.C., Chadha,G.S., Gillespie,P.J. and Blow,J.J. (2017) Reversal of DDK-Mediated MCM Phosphorylation by Rif1-PP1 Regulates Replication Initiation and Replisome Stability Independently of ATR/Chk1. *Cell Rep.*, **18**, 2508–2520.
13. Hiraga,S., Alvino,G.M., Chang,F., Lian,H.Y., Sridhar,A., Kubota,T., Brewer,B.J., Weinreich,M., Raghuraman,M.K. and Donaldson,A.D. (2014) Rif1 controls DNA replication by directing Protein Phosphatase 1 to reverse Cdc7-mediated phosphorylation of the MCM complex. *Genes Dev.*, **28**, 372–383.
14. Kumagai,A., Shevchenko,A., Shevchenko,A. and Dunphy,W.G. (2011) Direct regulation of Treslin by cyclin-dependent kinase is essential for the onset of DNA replication. *J. Cell Biol.*, **193**, 995–1007.
15. Boos,D., Sanchez-Pulido,L., Rappas,M., Pearl,L.H., Oliver,A.W., Ponting,C.P. and Diffley,J.F. (2011) Regulation of DNA replication through Sld3-Dpb11 interaction is conserved from yeast to humans. *Curr. Biol.*, **21**, 1152–1157.
16. Boos,D., Yekezare,M. and Diffley,J.F. (2013) Identification of a heteromeric complex that promotes DNA replication origin firing in human cells. *Science*, **340**, 981–984.
17. Kumagai,A. and Dunphy,W.G. (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.
18. Volpi,I., Gillespie,P.J., Chadha,G.S. and Blow,J.J. (2021) The role of DDK and Treslin-MTBP in coordinating replication licensing and pre-initiation complex formation. *Open Biol.*, **11**, 210121.
19. Thangavel,S., Mendoza-Maldonado,R., Tissino,E., Sidorova,J.M., Yin,J., Wang,W., Monnat,R.J. Jr, Falaschi,A. and Vindigni,A. (2010) Human RECQ1 and RECQ4 helicases play distinct roles in DNA replication initiation. *Mol. Cell. Biol.*, **30**, 1382–1396.
20. Sangrithi,M.N., Bernal,J.A., Madine,M., Philpott,A., Lee,J., Dunphy,W.G. and Venkitaraman,A.R. (2005) Initiation of DNA replication requires the RECQL4 protein mutated in Rothmund-Thomson syndrome. *Cell*, **121**, 887–898.
21. Matsuno,K., Kumano,M., Kubota,Y., Hashimoto,Y. and Takisawa,H. (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.
22. Sanuki,Y., Kubota,Y., Kanemaki,M.T., Takahashi,T.S., Mimura,S. and Takisawa,H. (2015) RecQ4 promotes the conversion of the pre-initiation complex at a site-specific origin for DNA unwinding in Xenopus egg extracts. *Cell Cycle*, **14**, 1010–1023.
23. Kliszczak,M., Sedlackova,H., Pitchai,G.P., Streicher,W.W., Krejci,L. and Hickson,I.D. (2015) Interaction of RECQ4 and MCM10 is important for efficient DNA replication origin firing in human cells. *Oncotarget*, **6**, 40464–40479.
24. Im,J.S., Park,S.Y., Cho,W.H., Bae,S.H., Hurwitz,J. and Lee,J.K. (2015) RecQL4 is required for the association of Mcm10 and Ctf4 with replication origins in human cells. *Cell Cycle*, **14**, 1001–1009.
25. Reynolds,J.J., Bicknell,L.S., Carroll,P., Higgs,M.R., Shaheen,R., Murray,J.E., Papadopoulos,D.K., Leitch,A., Murina,O., Tarnauskaite,Z. *et al.* (2017) Mutations in DONSON disrupt replication fork stability and cause microcephalic dwarfism. *Nat. Genet.*, **49**, 537–549.
26. Zhang,J., Bellani,M.A., James,R.C., Pokharel,D., Zhang,Y., Reynolds,J.J., McNee,G.S., Jackson,A.P., Stewart,G.S. and Seidman,M.M. (2020) DONSON and FANCM associate with different replisomes distinguished by replication timing and chromatin domain. *Nat. Commun.*, **11**, 3951.
27. Lesly,S., Bandura,J.L. and Calvi,B.R. (2017) Rapid DNA synthesis during early Drosophila embryogenesis is sensitive to maternal humpty dumpty protein function. *Genetics*, **207**, 935–947.
28. Bandura,J.L., Beall,E.L., Bell,M., Silver,H.R., Botchan,M.R. and Calvi,B.R. (2005) humpty dumpty is required for developmental DNA amplification and cell proliferation in Drosophila. *Curr. Biol.*, **15**, 755–759.
29. Hodgson,B., Li,A., Tada,S. and Blow,J.J. (2002) Geminin becomes activated as an inhibitor of Cdt1/RLF-B following nuclear import. *Current Biology : CB*, **12**, 678–683.
30. Gambus,A., Khoudoli,G.A., Jones,R.C. and Blow,J.J. (2011) MCM2–7 form double hexamers at licensed origins in Xenopus egg extract. *J. Biol. Chem.*, **286**, 11855–11864.
31. Prokhorova,T.A. and Blow,J.J. (2000) Sequential MCM/P1 subcomplex assembly is required to form a heterohexamer with replication licensing activity. *J. Biol. Chem.*, **275**, 2491–2498.
32. Priego Moreno,S., Jones,R.M., Poovathumkadavil,D., Scaramuzza,S. and Gambus,A. (2019) Mitotic replisome disassembly depends on TRAIP ubiquitin ligase activity. *Life Sci Alliance*, **2**, e201900390.
33. Moreno,S.P., Bailey,R., Campion,N., Herron,S. and Gambus,A. (2014) Polyubiquitylation drives replisome disassembly at the termination of DNA replication. *Science*, **346**, 477–481.
34. Tarcan,Z., Poovathumkadavil,D., Skagia,A. and Gambus,A. (2022) The p97 segregase cofactor Ubxn7 facilitates replisome disassembly during S-phase. *J. Biol. Chem.*, **298**, 102234.
35. Gillespie,P.J., Gambus,A. and Blow,J.J. (2012) Preparation and use of Xenopus egg extracts to study DNA replication and chromatin associated proteins. *Methods*, **57**, 203–213.
36. Blow,J.J. and Laskey,R.A. (1986) Initiation of DNA replication in nuclei and purified DNA by a cell-free extract of Xenopus eggs. *Cell*, **47**, 577–587.
37. Sonneville,R., Moreno,S.P., Knebel,A., Johnson,C., Hastie,C.J., Gartner,A., Gambus,A. and Labib,K. (2017) CUL-2LRR-1 and UBXN-3 drive replisome disassembly during DNA replication termination and mitosis. *Nat. Cell Biol.*, **19**, 468–479.
38. Kall,L., Storey,J.D. and Noble,W.S. (2008) Non-parametric estimation of posterior error probabilities associated with peptides identified by tandem mass spectrometry. *Bioinformatics*, **24**, i42–i48.
39. Nesvizhskii,A.I., Keller,A., Kolker,E. and Aebersold,R. (2003) A statistical model for identifying proteins by tandem mass spectrometry. *Anal. Chem.*, **75**, 4646–4658.
40. Jumper,J., Evans,R., Pritzel,A., Green,T., Figurnov,M., Ronneberger,O., Tunyasuvunakool,K., Bates,R., Zidek,A., Potapenko,A. *et al.* (2021) Highly accurate protein structure prediction with AlphaFold. *Nature*, **596**, 583–589.
41. Evans,R.M., Brooke,E.J., Wehlin,S.A., Nomerotskaia,E., Sargent,F., Carr,S.B., Phillips,S.E. and Armstrong,F.A. (2016) Mechanism of hydrogen activation by [NiFe] hydrogenases. *Nat. Chem. Biol.*, **12**, 46–50.
42. Jumper,J. and Hassabis,D. (2022) Protein structure predictions to atomic accuracy with AlphaFold. *Nat. Methods*, **19**, 11–12.
43. Pettersen,E.F., Goddard,T.D., Huang,C.C., Meng,E.C., Couch,G.S., Croll,T.I., Morris,J.H. and Ferrin,T.E. (2021) UCSF ChimeraX: structure visualization for researchers, educators, and developers. *Protein Sci.*, **30**, 70–82.
44. Goddard,T.D., Huang,C.C., Meng,E.C., Pettersen,E.F., Couch,G.S., Morris,J.H. and Ferrin,T.E. (2018) UCSF ChimeraX: meeting modern challenges in visualization and analysis. *Protein Sci.*, **27**, 14–25.
45. Gillespie,P.J., Khoudoli,G.A., Stewart,G., Swedlow,J.R. and Blow,J.J. (2007) ELYS/MEL-28 chromatin association coordinates nuclear pore complex assembly and replication licensing. *Curr. Biol.*, **17**, 1657–1662.
46. Kumagai,A., Shevchenko,A., Shevchenko,A. and Dunphy,W.G. (2010) Treslin collaborates with TopBP1 in triggering the initiation of DNA replication. *Cell*, **140**, 349–359.
47. Van Hatten,R.A., Tutter,A.V., Holway,A.H., Khederian,A.M., Walter,J.C. and Michael,W.M. (2002) The Xenopus Xmus101 protein is required for the recruitment of Cdc45 to origins of DNA replication. *J. Cell Biol.*, **159**, 541–547.
48. Nerakh,G., Vineeth,V.S., Tallapaka,K., Nair,L., Dalal,A. and Aggarwal,S. (2022) Microcephalic primordial dwarfism with

Downloaded from https://academic.oup.com/nar/advance-article/doi/10.1093/nar/gkad694/7252675 by guest on 29 August 2023

predominant Meier-Gorlin phenotype, ichthyosis, and multiple joint deformities-Further expansion of DONSON Cell Cycle-opathy phenotypic spectrum. *Am. J. Med. Genet. A*, **188**, 2139–2146.
49. Tak,Y.S., Tanaka,Y., Endo,S., Kamimura,Y. and Araki,H. (2006) A CDK-catalysed regulatory phosphorylation for formation of the DNA replication complex Sld2-Dpb11. *EMBO J.*, **25**, 1987–1996.
50. Masumoto,H., Muramatsu,S., Kamimura,Y. and Araki,H. (2002) S-Cdk-dependent phosphorylation of Sld2 essential for chromosomal DNA replication in budding yeast. *Nature*, **415**, 651–655.
51. Garcia,V., Furuya,K. and Carr,A.M. (2005) Identification and functional analysis of TopBP1 and its homologs. *DNA Repair (Amst.)*, **4**, 1227–1239.
52. Tanaka,S., Komeda,Y., Umemori,T., Kubota,Y., Takisawa,H. and Araki,H. (2013) Efficient initiation of DNA replication in eukaryotes requires Dpb11/TopBP1-GINS interaction. *Mol. Cell. Biol.*, **33**, 2614–2622.
53. Day,M., Tetik,B., Parlak,M., Almeida-Hernández,Y., Räschle,M., Kaschani,F., Siegert,H., Marko,A., Sanchez-Garcia,E., Kaiser,M. *et al.* (2023) TopBP1 utilises a bipartite GINS binding mode to activate the replicative helicase. bioRxiv doi: https://doi.org/10.1101/2023.03.31.535063, 03 April 2023, preprint: not peer reviewed.
54. Ohlenschlager,O., Kuhnert,A., Schneider,A., Haumann,S., Bellstedt,P., Keller,H., Saluz,H.P., Hortschansky,P., Hanel,F., Grosse,F. *et al.* (2012) The N-terminus of the human RecQL4 helicase is a homeodomain-like DNA interaction motif. *Nucleic Acids Res.*, **40**, 8309–8324.
55. Tanaka,S. and Araki,H. (2013) Helicase activation and establishment of replication forks at chromosomal origins of replication. *Cold Spring Harb. Perspect. Biol.*, **5**, a010371.
56. Collart,C., Allen,G.E., Bradshaw,C.R., Smith,J.C. and Zegerman,P. (2013) Titration of four replication factors is essential for the Xenopus laevis midblastula transition. *Science*, **341**, 893–896.
57. Hayano,M., Kanoh,Y., Matsumoto,S., Renard-Guillet,C., Shirahige,K. and Masai,H. (2012) Rif1 is a global regulator of timing of replication origin firing in fission yeast. *Genes Dev.*, **26**, 137–150.
58. Knapp,K.M., Sullivan,R., Murray,J., Gimenez,G., Arn,P., D'Souza,P., Gezdirici,A., Wilson,W.G., Jackson,A.P., Ferreira,C. *et al.* (2020) Linked-read genome sequencing identifies biallelic pathogenic variants in DONSON as a novel cause of Meier-Gorlin syndrome. *J. Med. Genet.*, **57**, 195–202.
59. Klumper,N., Blajan,I., Schmidt,D., Kristiansen,G., Toma,M., Holzel,M., Ritter,M. and Ellinger,J. (2020) Downstream neighbor of SON (DONSON) is associated with unfavorable survival across diverse cancers with oncogenic properties in clear cell renal cell carcinoma. *Transl. Oncol.*, **13**, 100844.
60. Klumper,N., von Danwitz,M., Stein,J., Schmidt,D., Schmidt,A., Kristiansen,G., Muders,M., Holzel,M., Ritter,M., Alajati,A. *et al.* (2020) Downstream Neighbor of SON (DONSON) expression is enhanced in phenotypically aggressive prostate cancers. *Cancers (Basel)*, **12**, 3439.
61. Yamada,Y., Nohata,N., Uchida,A., Kato,M., Arai,T., Moriya,S., Mizuno,K., Kojima,S., Yamazaki,K., Naya,Y. *et al.* (2020) Replisome genes regulation by antitumor miR-101-5p in clear cell renal cell carcinoma. *Cancer Sci.*, **111**, 1392–1406.
62. Qi,Y., Wu,H., Liu,C., Zheng,D., Yan,C., Hu,W., Zhang,X. and Dai,X. (2022) Downstream Neighbor of Son Overexpression is Associated With Breast Cancer Progression and a Poor Prognosis. *J. Breast Cancer*, **25**, 327–343.

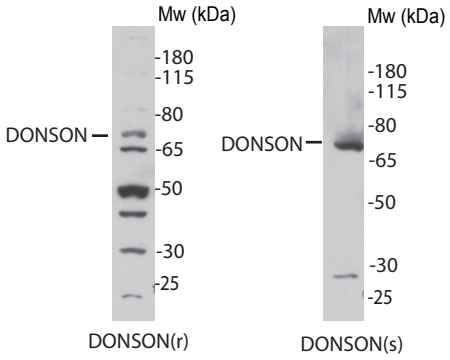
Downloaded from https://academic.oup.com/nar/advance-article/doi/10.1093/nar/gkad694/7252675 by guest on 29 August 2023

© The Author(s) 2023. Published by Oxford University Press on behalf of Nucleic Acids Research.
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

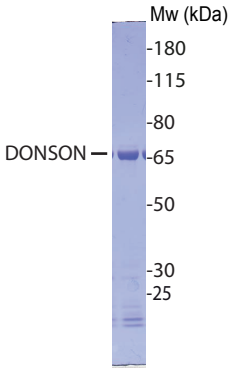
A

	Control (IgG)	IP (Mcm3)
Mcm3 (90 kDa)	17 (20%)	1111 (84%)
Mcm2 (100 kDa)	17 (17%)	975 (73%)
Mcm4 (97 kDa)	7 (7%)	947 (80%)
Mcm5 (82 kDa)	30 (18%)	927 (84%)
Mcm7 (82 kDa)	24 (10%)	905 (77%)
Mcm6 (93 kDa)	39 (18%)	888 (83%)
Cdc45 (66 kDa)	2 (3%)	277 (54%)
Psf1 (23 kDa)	0	62 (85%)
Psf2 (21 kDa)	0	33 (85%)
Psf3 (24 kDa)	0	55 (92%)
Sld5 (26 kDa)	0	44 (64%)
Ctf4 (125 kDa)	6 (17%)	554 (54%)
Spt16 (118 kDa)	3 (3%)	454 (55%)
SSRP (79 kDa)	0	229 (51%)
Timeless (149 kDa)	0	171 (32%)
Tipin (40 kDa)	0	54 (26%)
Top2a (179 kDa)	6 (3%)	484 (51%)
Claspin (146 kDa)	0	72 (21%)
PolA1 (165 kDa)	0	109 (33%)
PolA2 (67 kDa)	0	7 (11%)
PolE1 (261 kDa)	2 (1%)	543 (40%)
PolE2 (60kDa)	0	103 (55%)
DONSON (64 kDa)	0	68 (32%)

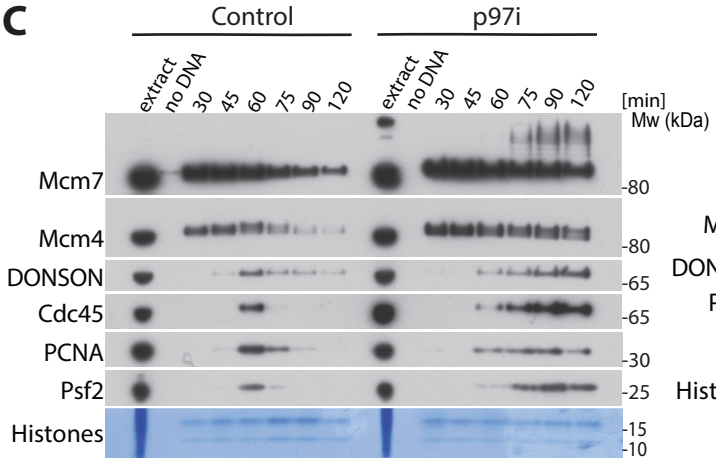
B



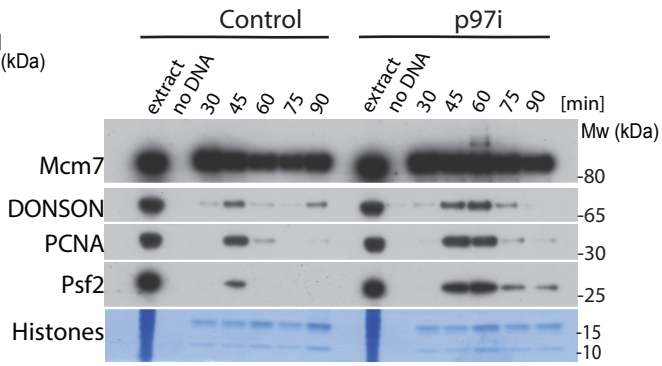
E



C



D



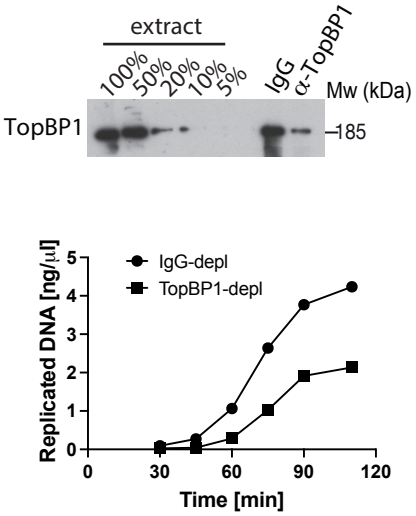
Supp Figure 1. DONSON interacts with terminated replisomes. **(A)** DNA replication reaction was established in Xenopus laevis egg extract supplemented with inactive mutant of p97 segregase and MLN4924 cullin E3 ligase inhibitor to block disassembly of replisomes in S-phase (as described in Sonneville et al 2017). Chromatin containing retained terminated replisomes was isolated, fragmented with benzonase and Mcm3 immunoprecipitaed. Proteins interacting with Mcm3 were analysed by mass spectrometry. Total spectral count and percentage coverage is presented for selected replisome components. **(B)** Polyclonal rabbit and shhep antibodies were raised against recombinant X.l.DONSON purified from bacteria. Protein band corresponding to DONSON within egg extract is ndicated. **(C)** Replicaton reaction was set up in egg extract supplemented with p97i NMS932 and chromatin isolated at indicated times. Chromatin samples were analysed by western blotting with indicated antibodies as in Figure 1. **(D)** As in (C) but the replication reaction was supplemented with cullin E3 ligase inhibitor MLN4924. **(E)** X.laevis 6HIS-DONSON expressed in bacteria and affinity purified.

A

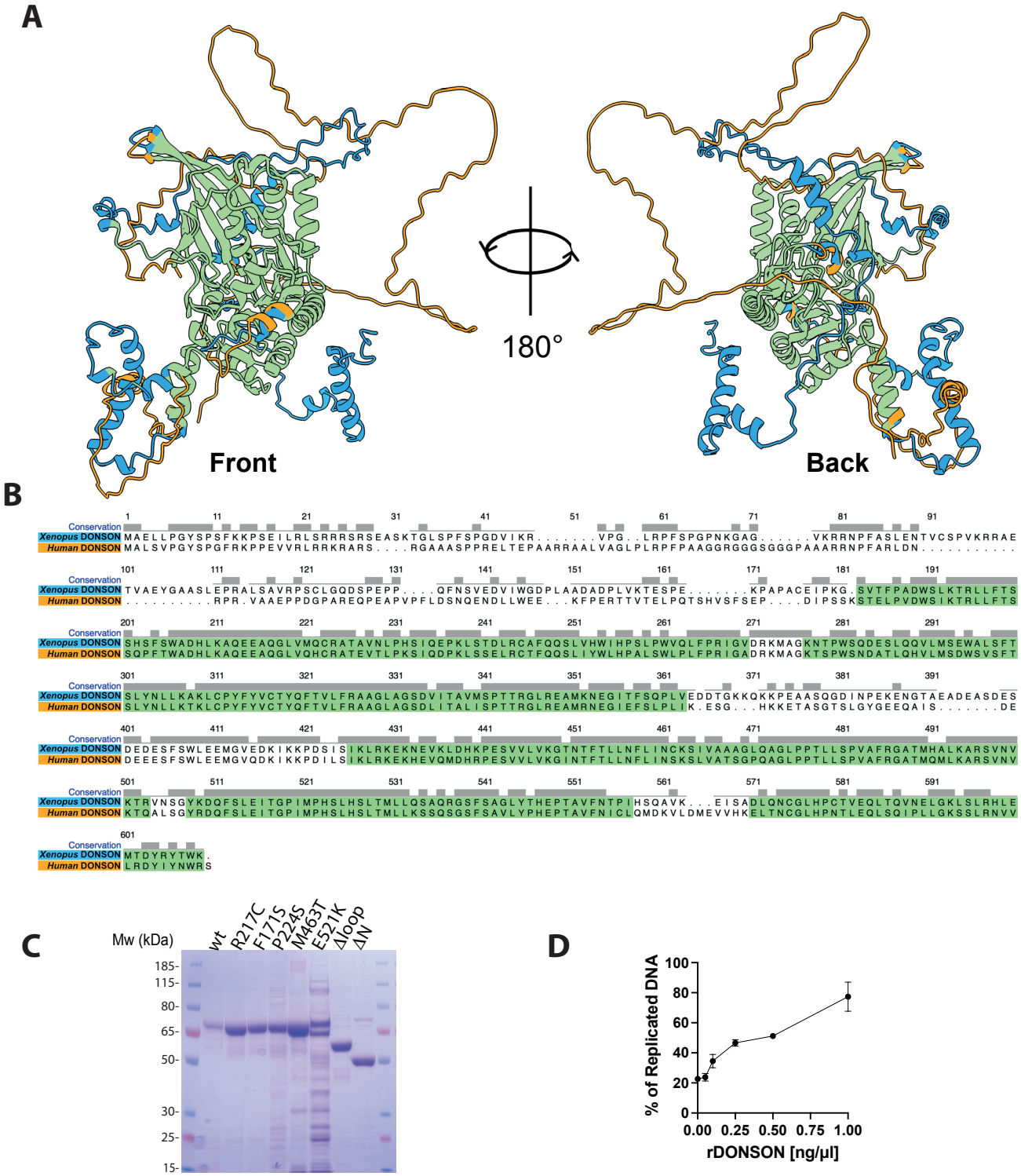
Fold enrichment in DONSON IP

	protein	p97i chromatin	aph/caff chromatin
	DONSON	84	86
Mcm2-7	Mcm2	1.3	1.7
	Mcm3	1.4	2.8
	Mcm4	1.5	2.9
	Mcm5	1.3	4.7
	Mcm6	1.3	1.4
	Mcm7	1.2	3.4
CMG	Psf2	6	0
	Psf4	2.3	0
	Psf1	1.9	0
	Cdc45	2.1	0
initiation factors	RecQ4	1.25	7
	RPA	5	2.29
	Treslin	2	0
	MTBP	3.6	4
	TOPBP1	0	7
	Rif1	5.9	7

B



Supp Figure 2. (A) DNA replication was set up in egg extract either supplemented with p97i to accumulate post-termination replisome or with aohidicolin and caffeine to accumulate stalled and continously estabishing replisomes. Chromatin sample was isolated either in late (for p97 inhibitor, 90 min) or middle (for aphidicolin/caffeine, 60 min) S-phase, protein complexes released from chromatn by DNA digestion and non-specific control and DONSON(s) antibodies used for immunoprecipiatation. The fold enrichment of presented replication fators in DONSON IP over nonspecific conrol IP was coalculated and is preseneted. **(B)** TopBP1 was immunodepleted from egg extract. The efficiency of immunodepletion was analysed by western blotting against titration of original egg extract (top) and the ability of the immunodepleted extract to synthesise nascent DNA analysed by incorporation of α^{32}P-dATP into synthesised DNA (bottom).



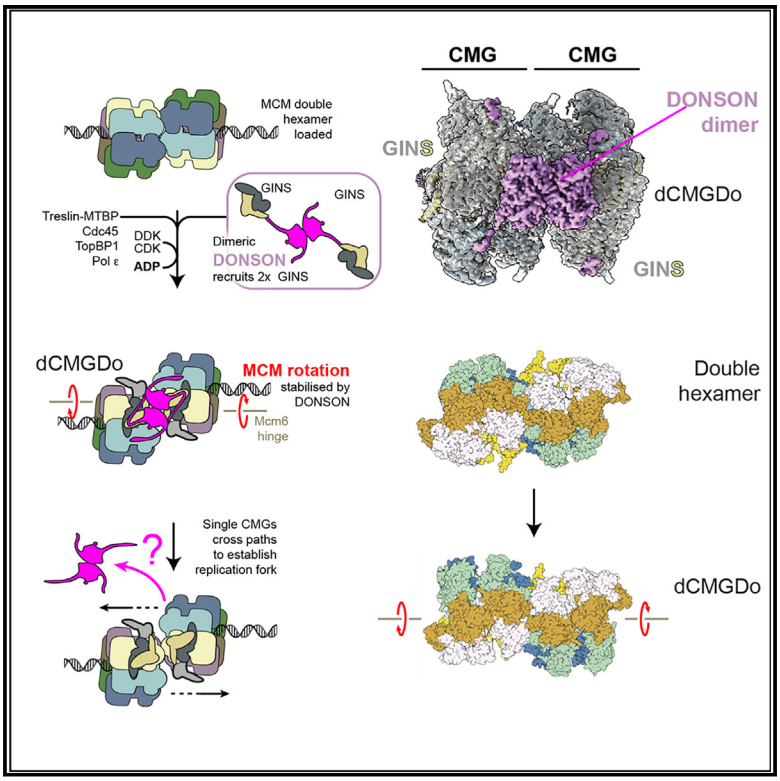
Supp Figure 3. Conservation of DONSON structure **(A)** AlphaFold overlay of 3D structure of Xenopus DONSON (blue) and human DONSON (orange). Overlapping structure in green. **(B)** Comparison of XenopusDONSON (blue) and human DONSON (orange) primary sequences. Conserved sequence in green. **(C)** Purification of DONSON truncation and patient mutations indicted. **(D)** Titration of recombinant wt DONSON into DONSON-depleted egg extract. The % of control replication level of IgG-depleted egg extract is presented as mean of n=2 with SEM.

Molecular Cell

Article

The structural mechanism of dimeric DONSON in replicative helicase activation

Graphical abstract



Highlights

- Purified complex from replicating chromatin frog egg extracts is cryo-EM tractable
- Structure of a double CMG bound by a DONSON dimer
- MCM double hexamer misaligned and ready to replicate bidirectionally
- Patients' mutations impairing DONSON dimerization affect DNA replication

Authors

Milos A. Cvetkovic, Paolo Passaretti, Agata Butryn, ..., Grant S. Stewart, Agnieszka Gambus, Alessandro Costa

Correspondence

a.gambus@bham.ac.uk (A.G.), alessandro.costa@crick.ac.uk (A.C.)

In brief

Cvetkovic, Passaretti, et al. determine the cryo-EM structure of a double CMG bound by dimeric DONSON, explaining how the CMG helicase is assembled symmetrically in metazoans. MCM rings become misaligned and poised to cross paths supporting replication origin activation. Disrupting DONSON dimer affects DNA replication, explaining etiology of primordial dwarfism.

Cvetkovic et al., 2023, Molecular Cell *83*, 4017–4031
November 16, 2023 © 2023 The Authors. Published by Elsevier Inc.
https://doi.org/10.1016/j.molcel.2023.09.029

CellPress

Molecular Cell

CellPress
OPEN ACCESS

Article

The structural mechanism of dimeric DONSON in replicative helicase activation

Milos A. Cvetkovic,[1,4] Paolo Passaretti,[2,4] Agata Butryn,[1] Alicja Reynolds-Winczura,[2] Georgia Kingsley,[2] Aggeliki Skagia,[2] Cyntia Fernandez-Cuesta,[2] Divyasree Poovathumkadavil,[2] Roger George,[3] Anoop S. Chauhan,[2] Satpal S. Jhujh,[2] Grant S. Stewart,[2] Agnieszka Gambus,[2,*] and Alessandro Costa[1,5,*]

[1]Macromolecular Machines Laboratory, The Francis Crick Institute, London NW1 1AT, UK
[2]Institute of Cancer and Genomic Sciences, Birmingham Centre for Genome Biology, University of Birmingham, Birmingham B15 2TT, UK
[3]Structural Biology Science Technology Platform, The Francis Crick Institute, London NW1 1AT, UK
[4]These authors contributed equally
[5]Lead contact
*Correspondence: a.gambus@bham.ac.uk (A.G.), alessandro.costa@crick.ac.uk (A.C.)
https://doi.org/10.1016/j.molcel.2023.09.029

SUMMARY

The MCM motor of the replicative helicase is loaded onto origin DNA as an inactive double hexamer before replication initiation. Recruitment of activators GINS and Cdc45 upon S-phase transition promotes the assembly of two active CMG helicases. Although work with yeast established the mechanism for origin activation, how CMG is formed in higher eukaryotes is poorly understood. Metazoan Downstream neighbor of Son (DONSON) has recently been shown to deliver GINS to MCM during CMG assembly. What impact this has on the MCM double hexamer is unknown. Here, we used cryoelectron microscopy (cryo-EM) on proteins isolated from replicating *Xenopus* egg extracts to identify a double CMG complex bridged by a DONSON dimer. We find that tethering elements mediating complex formation are essential for replication. DONSON reconfigures the MCM motors in the double CMG, and primordial dwarfism patients' mutations disrupting DONSON dimerization affect GINS and MCM engagement in human cells and DNA synthesis in *Xenopus* egg extracts.

INTRODUCTION

DNA replication is essential for cell proliferation. It occurs only once per cell cycle, and errors in the control mechanisms promote chromosomal instability, causing genetic disease and cancer. To ensure that DNA is copied only once, eukaryotic cells temporally separate DNA loading of the minichromosome maintenance (MCM) replicative helicase from its activation.[1,2] MCM is a ring-shaped ATPase that becomes loaded as a double hexamer onto origin DNA before replication initiation.[3–7] The DNA unwinding function in the double hexamer is inactive. Activation requires the recruitment of a set of firing factors, including GINS and Cdc45 that bind MCM to assemble the functional form of the helicase, named CMG (Cdc45-MCM-GINS).[8–11] Genetics and biochemical reconstitution with purified proteins provided a clear picture of the pathway and molecular mechanism for CMG formation and activation in yeast.[2,12] This work established that CMG assembly can be staged in three steps, namely double-hexamer phosphorylation, Cdc45 recruitment, and GINS recruitment. In the first step, double hexamers are phosphorylated by the Dbf4-dependent kinase (DDK).[13–19] The second step toward CMG assembly is Cdc45 recruitment, mediated by Sld3 (bound to Sld7), which specifically recognizes double hexamer phosphorylation written by DDK.[20,21] GINS recruitment instead depends on the activity of cyclin-dependent kinase Clb5-Cdc28 (CDK) that phosphorylates Sld2, promoting its association with Dpb11, Pol epsilon, and GINS. CDK also phosphorylates Sld3, and this modification is also read by Dpb11. Concomitant recognition of phospho-Sld2 and phospho-Sld3 by Dpb11 recruits GINS to the double hexamer.[22–24] Cdc45 and GINS engagement promotes ATP binding by MCM, which in turn triggers release of Dpb11, Sld3/7, and Sld2.[25]

Structural work established that GINS and Cdc45 are recruited to form two CMG complexes in a step known to require the Pol epsilon replicative polymerase.[25,26] This results in a double CMG (dCMG)-Pol epsilon assembly that distorts the DNA double helix and changes in the register of dimerizing MCM rings. As the two rings rotate with respect to one another, the two N-terminal domains of MCM become almost fully detached, remaining tethered via the Mcm6 zinc finger (ZnF) domain and exposing B-form duplex DNA between the two hexamers. Within each helicase ring, ATP binding triggers a change in the DNA grip, leading to untwisting and melting of the double helix.[27] In a subsequent step, Mcm10 activates the ATP hydrolysis function of MCM, causing the two helicases to eject one DNA strand from their central channel and cross paths, so that the strand ejected

This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

from one hexamer becomes the translocation strand of the other.[25,28,29] This mechanism requires that both helicases are activated at the same time for bidirectional replication to occur. Structural evidence supports this symmetric helicase activation mechanism. In fact, cryoelectron microscopy (cryo-EM) reconstructions obtained from biochemical reconstitution efforts, or isolation of chromatin-bound complexes from cells, show that dimeric MCM assemblies can either be found as symmetric inactive double hexamers or dCMG-Pol epsilon complexes that nucleate DNA melting symmetrically.[6,7,27,30,31] Never was an asymmetric intermediate observed, which contains an inactive MCM bound to an activated CMG-Pol epsilon. Although key for the establishment of bidirectional replication, the structural mechanism for the conversion from double hexamer to dCMG remains only partially understood. Furthermore, only for yeast but not for higher eukaryotes has the complete list of factors required for CMG assembly and activation been determined. In fact, work with higher eukaryotes led to identifying almost all orthologs of the factors involved in yeast replication initiation, including the Dpb11 homolog, TopBP1, and the Sld3-7 homolog, Treslin-MTBP.[32–34] The exception is a homolog for yeast Sld2, whose identity remains unclear. Previous studies noted that the RecQL4 factor, which contains an Sld2 homology domain, is essential for replication initiation. However, RecQL4 is not required for Cdc45 and GINS interaction with MCMs, rather it acts downstream of CMG formation.[35–40] Recent work with frog egg extracts led to the suggestion that DONSON (Downstream neighbor of Son) is the elusive functional homolog of yeast Sld2 in higher eukaryotes. Like Sld2, DONSON directly contacts GINS, and this interaction supports CMG assembly and triggers DNA replication initiation. Moreover, like Sld2, chromatin association by DONSON requires double hexamer formation, the kinase activity of DDK and CDK as well as the presence of TopBP1/Dpb11.[41,42] Mutations in DONSON are the cause of microcephalic primordial dwarfism, a group of disorders characterized by growth delays, often related to defects in the initiation step of DNA replication.[43,44] A role for DONSON in origin activation is remarkable as it indicates a function separate from its established role during elongation. In fact, DONSON was previously implicated in processes required after origin activation, including proper activation of S-phase checkpoint response, replication fork protection from the action of nucleases, and traversal of inter-strand cross-links.[43,45] To understand the role of DONSON in GINS recruitment and the symmetric activation of the MCM replicative helicase, we used the *Xenopus laevis* egg extract system to stall replication at an early stage and imaged replication protein assemblies isolated from chromatin with cryo-EM.

RESULTS

CMG assemblies isolated from replicating frog egg extracts

To understand the mechanism for helicase activation that supports replication origin firing in vertebrates, we devised a protocol to isolate the activated replicative helicase from replication reactions established using the *Xenopus laevis* egg extract system. First, we used bacterial overexpression to produce *Xenla*Cdc45 fused to C-terminal, tobacco etch virus (TEV) protease-cleavable, tandem His_{10}-$FLAG_5$ tags (Figure S1). Demembranated sperm DNA was added to an activated egg extract supplemented with recombinant Cdc45. Cdc45-TEV-His_{10}-$FLAG_5$ becomes incorporated into isolated chromatin (Figure 1A) and rescues DNA replication in extracts depleted for the endogenous protein (Figure 1B). We repeated the replication experiments in the presence of caffeine-and-aphidicolin to capture an early step in replication. Caffeine shuts down the DNA damage checkpoint by inhibiting the two master control kinases in DNA damage response, ataxia telangiectasia mutated (ATM) and ataxia telangiectasia and Rad3-Related (ATR).[46] Aphidicolin is a potent inhibitor of family B DNA polymerases (including the Pol alpha, delta, and epsilon replicative polymerases).[47] The combined use of these two inhibitors leads to uncontrolled origin firing but stops DNA synthesis from happening. In these conditions, we expected accumulation of stable intermediates of replication initiation and stalled replication forks. In fact, western blot analysis performed on isolated chromatin showed a marked increase in initiation factors and components of the established replisome when caffeine or caffeine/aphidicolin were used (Figure 1C). After supplementing these egg extracts with recombinant Cdc45, we isolated chromatin, digested DNA using benzonase, and immunoprecipitated Cdc45, using anti-FLAG M2 beads. Silver-stain SDS-PAGE revealed multiple proteins co-purified with recombinant Cdc45 (Figure 1D), and mass spectrometry analysis identified several replication factors (Table S1). These include all six subunits of the MCM complex (within the first ten hits), the four GINS subunits, RPA, and components of the replicative polymerase assemblies, Pol alpha/primase, Pol epsilon, and Pol delta. Hence, recombinant Cdc45 can be incorporated in a CMG complex competent for replication fork establishment. Mass spectrometry also identified DONSON as the member of a Cdc45-containing assembly. To establish whether the Cdc45-containing complexes represented structural intermediates of origin activation or rather the established replication fork, we performed negative stain electron microscopy combined with single-particle two-dimensional (2D) analysis. The resulting 2D averages revealed recognizable views of a single CMG, similar to the negative stain averages observed after origin firing reconstituted *in vitro* with purified yeast proteins.[25] We also observed two CMGs with N-terminal MCM sides engaged in dimerization, which were highly reminiscent of the initiation intermediate, dCMG-Pol epsilon, caught in the act of nucleating DNA melting before helicases cross paths[27] (Figure 1E). From our negative stain analysis, it was unclear whether either of the two CMG assemblies was found associated with DONSON or any components of the established replication fork.

A dimeric CMG-DONSON complex revealed by cryo-EM

Seeking to characterize the dimeric CMG assemblies, we repeated the same electron microscopy experiment in cryogenic conditions. 2D classification combined with *ab initio* structure determination and three-dimensional (3D) classification confirmed the presence of single CMG and dCMG assemblies (Figure S2).

The structure of the single CMG was refined to 3.0 Å average resolution, with 2.6 Å local resolution for the GINS, Cdc45, and part of N-terminal MCM, whereas the ATPase domain was

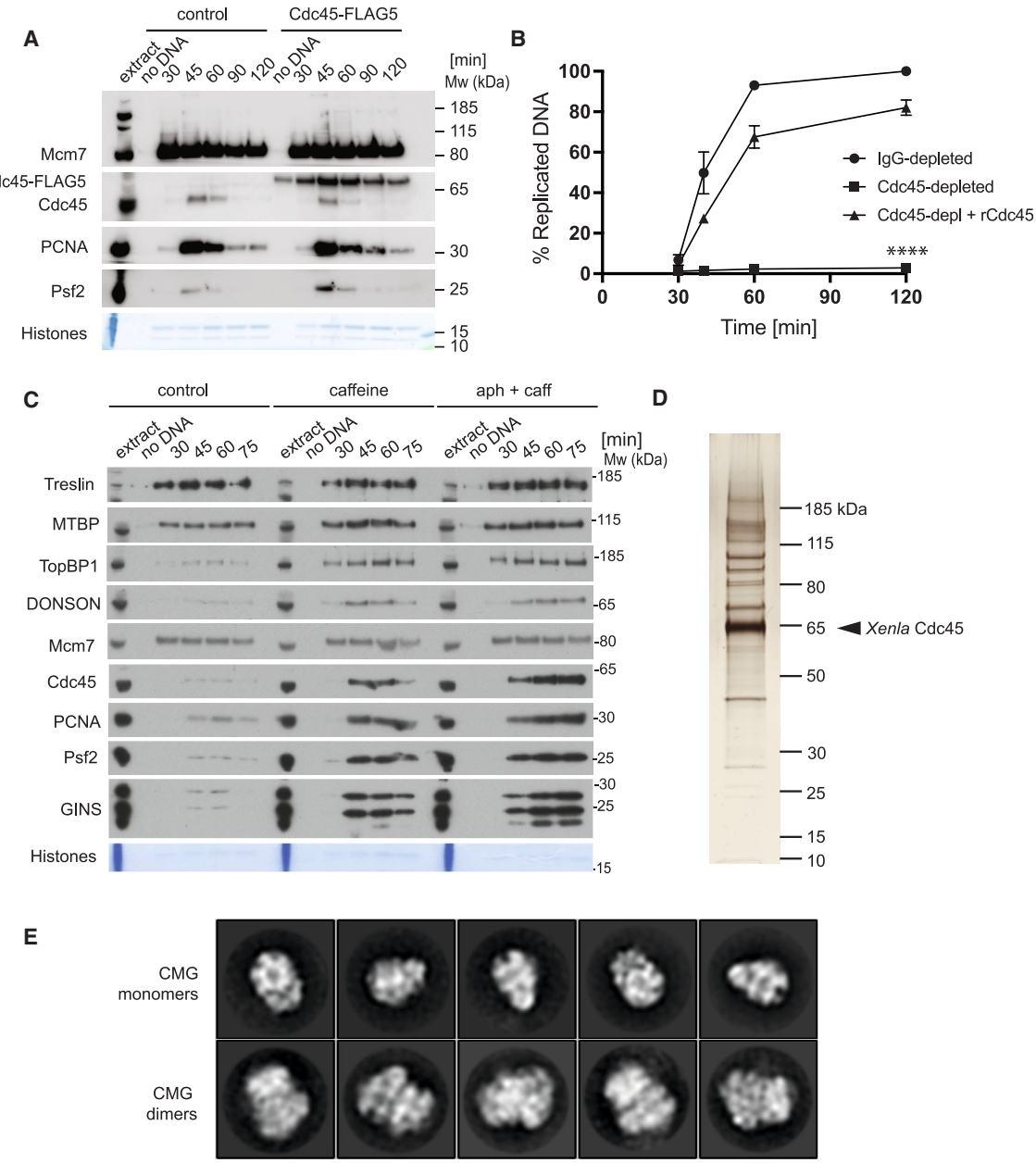


Figure 1. Isolation of replication complexes from chromatin replicated in *Xenopus laevis* egg extract

(A) Cdc45-TEV-His_{10}-$FLAG_5$ becomes chromatin incorporated. DNA replication reaction was set up in *Xenopus* egg extract with optional addition of 70 nM Cdc45-TEV-His_{10}-$FLAG_5$, and chromatin was isolated at indicated times. Chromatin-bound factors were resolved on SDS-PAGE and immunoblotted with indicated antibodies. "No DNA" control served as chromatin specificity control, whereas Coomassie stained histones serve as a loading and purity control.

(B) DNA replication reaction was set up in IgG- or Cdc45-depleted extract optionally supplemented with recombinant Cdc45-TEV-His_{10}-$FLAG_5$ as in (A). The synthesis of nascent DNA was followed by incorporation of α^{32}P-dATP into newly synthesized DNA. n = 3, mean with SEM presented. Two-way ANOVA comparing to IgG-depleted, **** $p < 0.0001$.

(C) DNA replication reaction was set up in *Xenopus* egg extract with optional addition of 5 mM caffeine and 40 μM aphidicolin. Chromatin was isolated at indicated times and analyzed as in (A).

(D) DNA replication reaction was set up in *Xenopus* egg extract with optional addition of 70 nM Cdc45-TEV-His_{10}-$FLAG_5$, 5 mM caffeine and 40 μM aphidicolin. After 60 min of the reaction, chromatin was isolated, digested with benzonase, and Cdc45-TEV-His_{10}-$FLAG_5$ immunoprecipitated with M2 FLAG beads. The immunoprecipitated sample was resolved on SDS-PAGE and silver stained.

(E) Negative stain 2D averages of single (first row) and double CMG (second row).

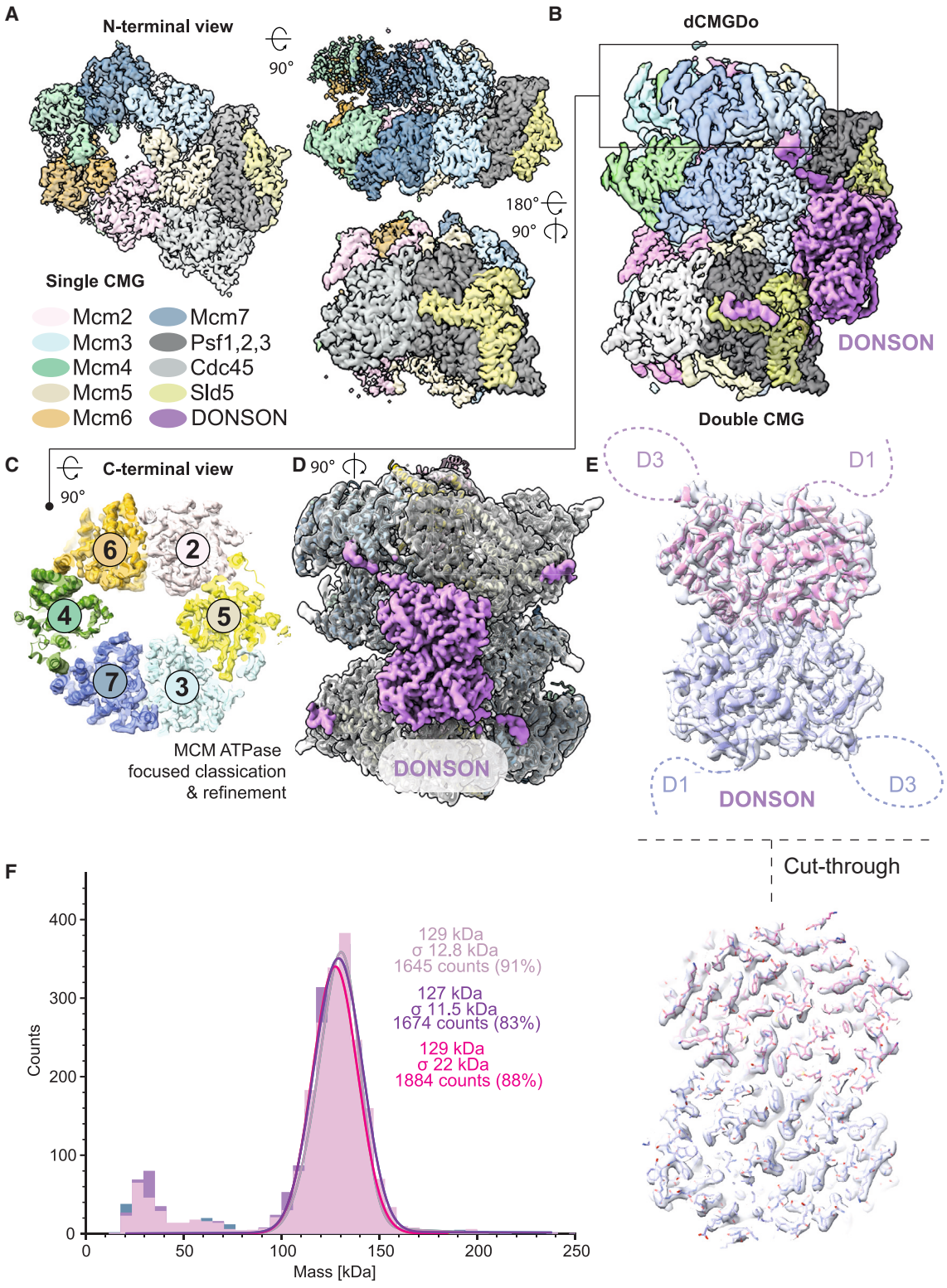


Figure 2. Cryo-EM analysis of Cdc45 containing complexes leads to identification of a DONSON-bound double CMG assembly

(A) The N-terminal end-on view, the side, and the rotated side view of single CMG.

(B) The side view of a double CMG in a configuration distinct from what was previously observed with the yeast CMG assembled at an origin of replication.

(legend continued on next page)

more flexible (Figures 2A and S3). No DNA density was observed bound to CMG, as expected given that Cdc45-containing complexes were purified from chromatin after DNA digestion using benzonase. Neither DONSON nor any of the other replication fork components were observed, possibly because their association with CMG involves flexible elements or because these factors had become dissociated in the conditions used during purification and cryo-EM sample preparation.

The structure of dCMG contained 2-fold symmetry across the dimerizing N-terminal domain and was refined to 3.1 Å average resolution (Figure 2B). As for the single CMG, the highest local resolution (2.7–3.7 Å) was detected for the N-terminal MCM region, GINS and Cdc45, whereas the ATPase domain was more flexible (Figure S3). Resolution of the ATPase tier was improved to 3.5 Å after performing symmetry expansion, signal subtraction, and focused classification on one ATPase domain, and refinement of the most populated conformers (Figures 2C, S2, and S3). Although no DNA was found in the complex, an extensive globular density feature, resolved to 2.7 Å, could be recognized straddling across two CMG assemblies (Figure 2D). Within each CMG assembly, this density bridges between Mcm3 and GINS. Because of the GINS interaction, we speculated that the bridging element could correspond to the DONSON factor that promotes incorporation of GINS into the CMG complex. To test our hypothesis, we docked the AlphaFold2 model for a DONSON monomer into the cryo-EM density left unoccupied in the dCMG complex.[48] Automated docking revealed two non-overlapping solutions, with identical cross-correlation scores, showing that the unoccupied density in the dCMG complex corresponds to a 2-fold symmetric homodimer of DONSON (Figure 2E). We named this 24-mer complex dCMGDo (for double CMG-DONSON) and subjected the complete atomic structure to real space refinement (Table 1).

A constitutive DONSON homodimer engages two GINS sites

Dimeric DONSON could deposit two GINS assemblies onto a phosphorylated double hexamer on the path to CMG formation. Alternatively, monomeric DONSON could deliver two GINS assemblies in independent events and homo-dimerize only when incorporated in the dCMGDo. To discriminate between these two possibilities, we overexpressed and purified recombinant *Xenopus laevis* DONSON from *E. coli* cells. Mass photometry analysis showed that recombinant DONSON is primarily dimeric, having a measured molecular mass of 128.3 ±15.4 kDa, in striking agreement with what is expected for a DONSON homodimer (132 kDa, Figure 2F). A truncation variant that lacks residues 1–155, forming the unstructured D1 loop, still assembles as a homodimer (Figure S4A). Thus, isolated recombinant DONSON homodimerizes via the globular domain, as seen in the dCMGDo assembly.

To understand the interaction between GINS and DONSON, we analyzed the cryo-EM structure in detail. We found that DONSON contacts two separate sites on GINS. A first interface involves the A-domain of Sld5, which we found decorated by unoccupied density. AlphaFold2 multimer predicted a high-confidence interaction between this Sld5 domain and residues 7–24 of DONSON.[48] Based on the cryo-EM density, we could build an accurate atomic model for residues 6–22, which strikingly matches the predicted structure (Figure 3A). The second interface is provided by the B-domain of Sld5, engaged by a hairpin loop (amino acids 473–482), emanating from the DONSON central core (Figure 3B). Single residues involved in this interface cannot be clearly identified based on the cryo-EM map. In fact, the contact between the two surfaces could only be observed in the non-sharpened density map. Nonetheless, AlphaFold2 multimer predicted this interface with high confidence, providing a model that accurately recapitulates the relative orientation between GINS and DONSON, observed by cryo-EM.[48] To assess the importance of these interactions, we generated two DONSON variants, with changes in elements interacting with Sld5 A-domain and B-domain, respectively. We named these variants DONSONΔ24 (containing a truncation of residues 1–24) and DONSON 3A (containing changes Val473Ala, Arg476Ala, and Tyr481Ala) (Figure S1B). To ensure that these changes do not alter DONSON's core structure, we analyzed both variants using mass photometry. We found that both DONSONΔ24 as well as DONSON 3A form homodimers, like the wild-type (WT) protein (Figures S4B and S4C). Thus, any defect in DONSON's function should be ascribed to defective GINS interaction and not protein misfolding. To test the functionality of our mutants during DNA replication, we immunodepleted DONSON from the egg extract using an antibody known to deplete endogenous DONSON to levels below 5%, which strongly diminishes the extract's ability to synthesize nascent DNA.[42] We found that although WT recombinant DONSON purified from *E. coli* rescued replication to the levels of control IgG-depleted extract, DONSONΔ24 was unable to support DNA replication in the DONSON-depleted extract. However, DONSON 3A variant was proficient in supporting nascent DNA synthesis (Figure 3C), in agreement with the cryo-EM observation that the hairpin loop of DONSON is flexible and only loosely engages the Sld5 B-domain. We confirmed that WT DONSON and DONSON 3A supported loading of Cdc45 and GINS onto chromatin, but DONSONΔ24 was defective (Figure S5A). To test the prediction that a DONSON variant lacking the N-terminal Sld5 interaction element fails to engage GINS, we immunoprecipitated recombinant FLAG-tagged GINS after incubation with different DONSON variants and probed the interaction with an anti-DONSON antibody. We found that WT DONSON and DONSON 3A could associate with recombinant GINS, but DONSONΔ24 failed to interact (Figure S5B).

Dimeric DONSON engages two Mcm3 sites

We then analyzed the interaction between MCM and DONSON in the dCMGDo cryo-EM structure. A first interaction involves the

(C) A view of the MCM ATPase improved after focused classification and refinement.
(D) Unassigned density after building the double CMG atomic model.
(E) Cartoon representation of a DONSON homodimer built into the cryo-EM density represented with transparent surface rendering and a cut-through view.
(F) Mass photometry analysis demonstrates that purified recombinant DONSON homodimerizes.

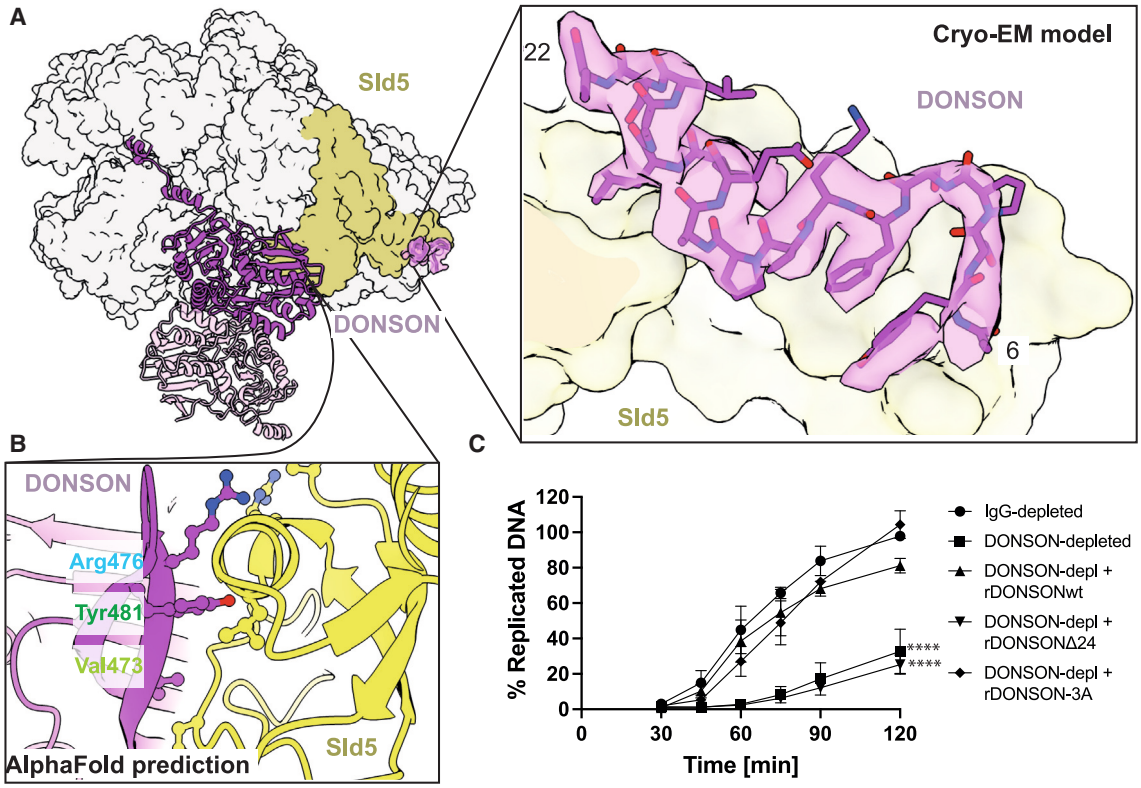


Figure 3. DONSON binding to Sld5

(A) One isolated CMG from the dCMGDo structure, engaged by dimeric DONSON. Only one DONSON protomer binds one CMG. The zoomed in view shows the N-terminal tail of DONSON, tethered to the DONSON core via an unstructured (invisible) linker domain, and bound to Sld5.

(B) AlphaFold prediction of the engagement interface between the DONSON globular dimerization domain and the Sld5 B-domain.

(C) DNA replication reaction was set up in IgG- or DONSON-depleted extract optionally supplemented with 16 nM recombinant wild-type or indicated DONSON mutants. The synthesis of nascent DNA was followed by incorporation of α^{32}P-dATP into newly synthesized DNA at indicated times. n = 3, mean with SEM presented. Two-way ANOVA comparing with IgG-depleted, **** $p < 0.0001$.

outer perimeter of the ATPase domain of Mcm3 and an alpha helix element inserted within a flexible loop (D3 loop, residues 345–395) emanating from the DONSON core. This DONSON alpha helix features a well-resolved Trp381 engaging Mcm3 Lys285. Our model strikingly resembles the AlphaFold2 multimer model,[48] which also predicted additional Mcm3 interactions with DONSON residues Asp374 and Glu377 (not well resolved in the experimental map, Figure 4A). A second interaction involves the N-terminal A-domain of Mcm3, at a site proximal to that engaged by GINS subunit Psf3 in the CMG (Figure 4B). To test the importance of the interaction sites with Mcm3, we generated three DONSON variants. The DONSON DEW variant, containing Asp374, Glu377, and Trp381 changed to alanine, targeted the Mcm3 ATPase domain interaction (Figure 4A). Ser437Ala targeted a polar interaction of DONSON with Mcm3 Gln27, embedded in a pocket that appears otherwise hydrophobic (Figure 4B). A third variant, DONSON DEWS, combined DONSON DEW and Ser437Ala. We confirmed that these three variants can homodimerize like the WT proteins (Figures S4D–S4F) and then asked whether changes in the DONSON-Mcm3 interaction had any effect on DNA replication. Recombinant DONSON containing the Ser437Ala mutation showed a mild defect in rescuing DNA replication after immuno-depletion of the endogenous protein. Instead, the DONSON DEW variant was severely, although not totally, deficient at rescuing replication. The DONSON DEWS variant, which targeted both N-terminal and ATPase interaction sites on Mcm3, was completely unable to rescue replication (Figure 4C). We confirmed that DONSON Ser437Ala supported loading of Cdc45 and GINS onto chromatin, but DONSON DEW and DONSON DEWS were defective (Figure S5C).

MCM dimerization in dCMGDo

We compared the MCM dimerization interface in dCMGDo with dimeric complexes observed before[6,7,17,27,30,31] and found that the hexamer interface is completely reconfigured upon DONSON binding. First, the inter-ring interactions are significantly loosened. In fact, by measuring the distance between the N-terminal ZnF domains, we conclude that the two MCM rings hover over one another rather than engaging in direct interactions. With the shortest inter-ring distances ranging from 5 to 7 Å, any contacts, if they exist, are likely water mediated. We then inspected the complex from the C-terminal ATPase side to observe the structural changes upon double hexamer to

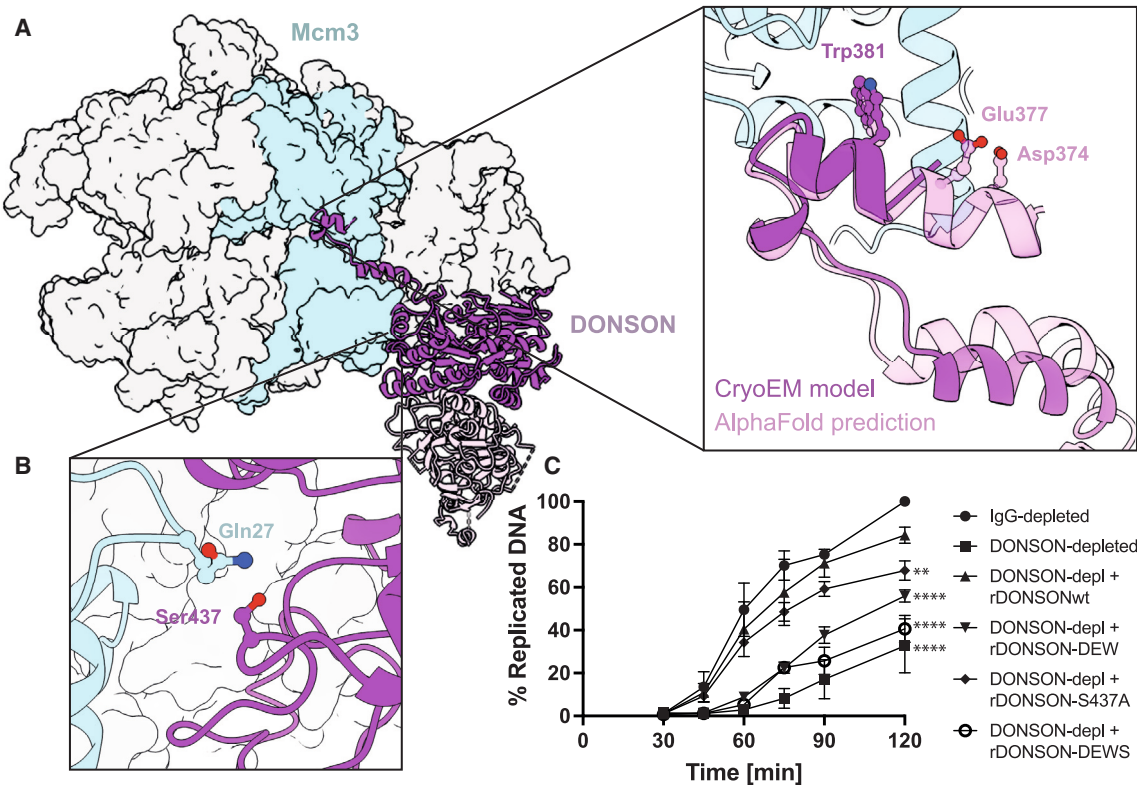


Figure 4. DONSON binding to Mcm3

(A) One isolated CMG from the dCMGDo structure, engaged by dimeric DONSON. The zoomed in view shows the D3 loop of DONSON engaging the Mcm3 ATPase domain. The solid cartoon represents the atomic model built into the cryo-EM density. The transparent cartoon represents the AlphaFold prediction. The two models identify Trp381 as an Mcm3 ATPase interacting element. AlphaFold also identifies Asp374 and Glu 377 as ATPase-engaged.

(B) Ser437 engages Mcm3 A-domain residue, Gln27, according to the atomic model built into the cryo-EM density.

(C) DNA replication reaction was set up in IgG- or DONSON-depleted extract optionally supplemented with 16 nM recombinant wild-type or indicated DONSON mutants. The synthesis of nascent DNA was followed by incorporation of α^{32}P-dATP into newly synthesized DNA at indicated times. n = 3, mean with SEM presented. Two-way ANOVA comparing to IgG-depleted, **** $p < 0.0001$. ** $p = 0.0046$.

dCMGDo conversion. We detected an ~60° clockwise rotation of one MCM ring pivoting around two aligned N-terminal ZnF domains of Mcm2 and Mcm6 across the double hexamer (Figure 5A; Video S1). As a result, the Mcm7 side of one hexamer swings out causing the two MCM rings to become misaligned, to the point that the two central channels are no longer co-axial and rather identify topologically separate entities (Figure 5B). This reconfiguration moves the two Mcm5 ZnF domains past one other, whereas they still remain in close proximity (Figure 5C). The ZnF domains of two Mcm2 subunits move from a distal to a proximal position at the inter-ring interface. In this configuration, the Mcm2 ZnF of one MCM ring appears poised to partially occlude the N-terminal pore of the opposed MCM ring (Figure 5D). A comparison with the published structures shows that the Mcm2 in dCMGDo overlaps with the position occupied by duplex DNA in the double hexamer. Here, the same side of Mcm2 that touches the DNA backbone in the double hexamer,[6,7,30,31] instead faces the opposed Mcm2 ZnF across dCMGDo (Figure S6A). In summary, double hexamer to dCMGDo transition triggers a rotation of the two MCM rings so that their central pores become misaligned. This causes each Mcm2 ZnF to plug the opposed MCM ring pore so that a DNA interacting element in Mcm2 becomes occluded. These changes upon dCMGDo formation would make it impossible for DNA to thread through the juxtaposed N-terminal domains of MCM while maintaining a double helical character. The clockwise rotation of two MCM rings of the double hexamer observed in the dCMGDo, and the prediction that duplex DNA is untwisted at the dimerization interface, matches a hypothetical fork-establishment mechanism by Eric Enemark.[49] This mechanism was shown using an ingenious physical model for double hexamer separation, built with wood sticks and strings. In this model, the two DNA strands were pulled on to mimic 3′ to 5′ single-stranded DNA spooling established for CMG and the two MCM rings were allowed to pivot around one MCM hinge dimer across the double hexamer. As a result, the opposed side of MCM (for us, Mcm7) swings out, and the DNA at the N-terminal dimerization interface is unwound and sheared apart (Figure 5A). This strikingly matches our conformational transitions modeled after the dCMGDo structure.

It should be noted that the dCMGDo complex was released from chromatin via DNA digestion, and we considered the possibility that retaining DNA could limit the ability of two MCMs to become offset as observed in our structure. We note however

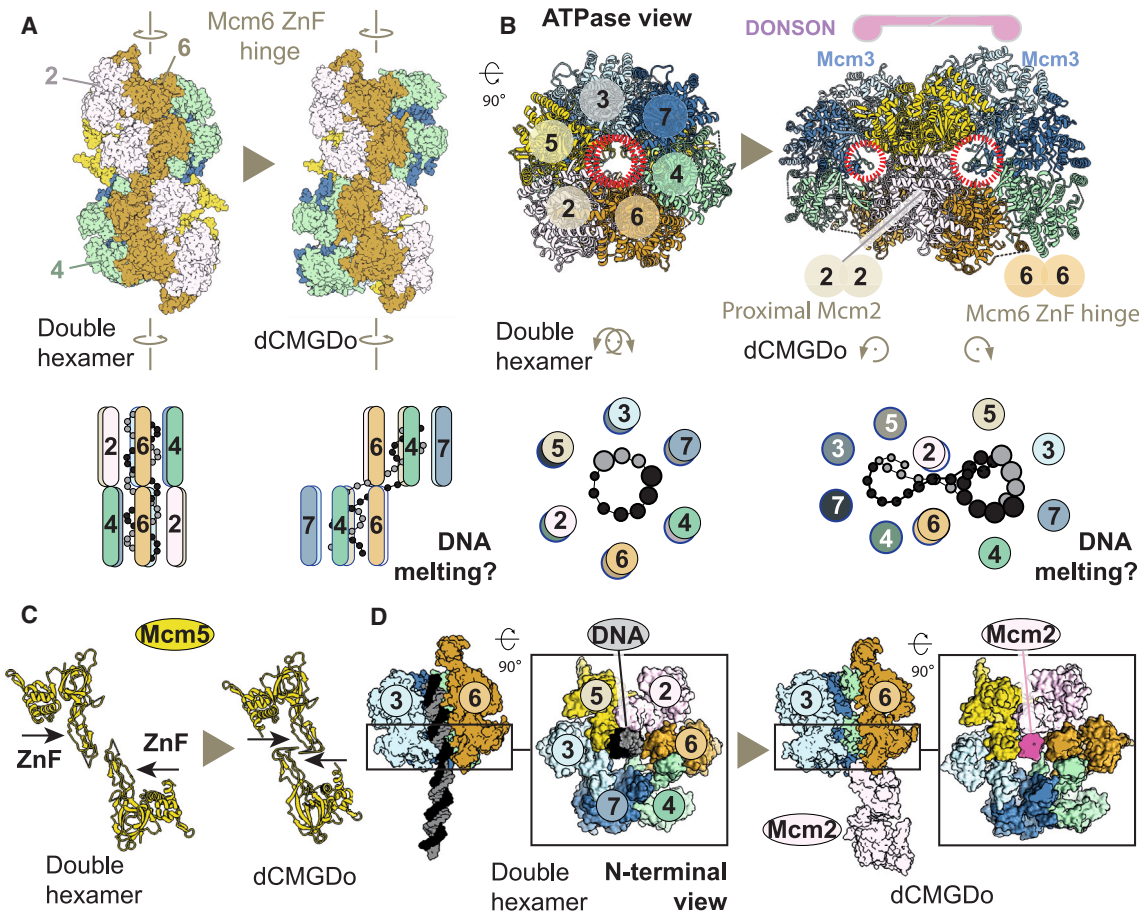


Figure 5. Conformational changes within the MCM occurring upon conversion from double hexamer to dCMGDo

(A) Two MCM rings, viewed from the side, rotate clockwise with respect to one another pivoting around Mcm6. This movement causes the Mcm7 side of the ring to swing out. This movement agrees with a theoretical model proposed previously.[49]

(B) The double hexamer to dCMGDo conversion viewed from an end-on view of the ATPase ring. The two MCM central channels are no longer co-axial in dCMGDo.

(C) Two N-terminal ZnF domains cross paths upon conversion from double hexamer to dCMGDo.

(D) Reconfiguration of the N-terminal MCM ring pore upon double hexamer to dCMGDo transition. Mcm2 from one ring in dCMGDo occludes the MCM central channel occupied by duplex DNA in the double hexamer.

that Mcm3 and Sld5 interactions with dimeric DONSON, essential for replication, can only be established if one MCM swings out. Instead, due to steric clashes, the DONSON dimer could not be engaged if two CMGs maintained a double hexamer inter-ring register (Figure S6B). Likewise, DONSON dimer engagement would not be possible if two CMGs interacted as observed in the yeast dCMG-Pol epsilon structure, as DONSON binding to one CMG complex would be too far to reach the Mcm3 A-domain and Sld5 B-domain of the second CMG (Figure S6C).

The role of dimeric DONSON

Inspection of the dCMGDo structure established that a unique reconfiguration of the double hexamer appears to be captured by (or imparted upon) the engagement of dimeric DONSON. At least for the globular core of the protein, each protomer in the DONSON dimer contacts one CMG complex alone, rather than bridging across the two helicases (Figure 6A). Hence, stability of the whole dCMGDo structure, featuring two offset MCM rings, appears dependent on DONSON homodimerization. We looked for amino acid changes that would disrupt dimer formation in our structure and came to realize that two residues, whose human counterparts are mutated in patients, map at the dimer interface in our structure (Figure 6B). One is Met463, which maps at the center of a hydrophobic pocket and makes direct contact with dimerization surface residues in cis. It corresponds to Met446 in humans, whose mutation to Thr is recessive and causes primordial dwarfism in patients.[43,50] The second is Trp234, engaged in a hydrophobic interaction with equivalent Trp234 from the opposed protomer. It corresponds to Trp228 in humans, whose mutation to Leu is dominant negative, and was identified in a patient clinically diagnosed with Femoral-Facial syndrome.[51] We created a human cell line containing GFP-tagged and FLAG-tagged DONSON either WT or containing the Met446Thr or

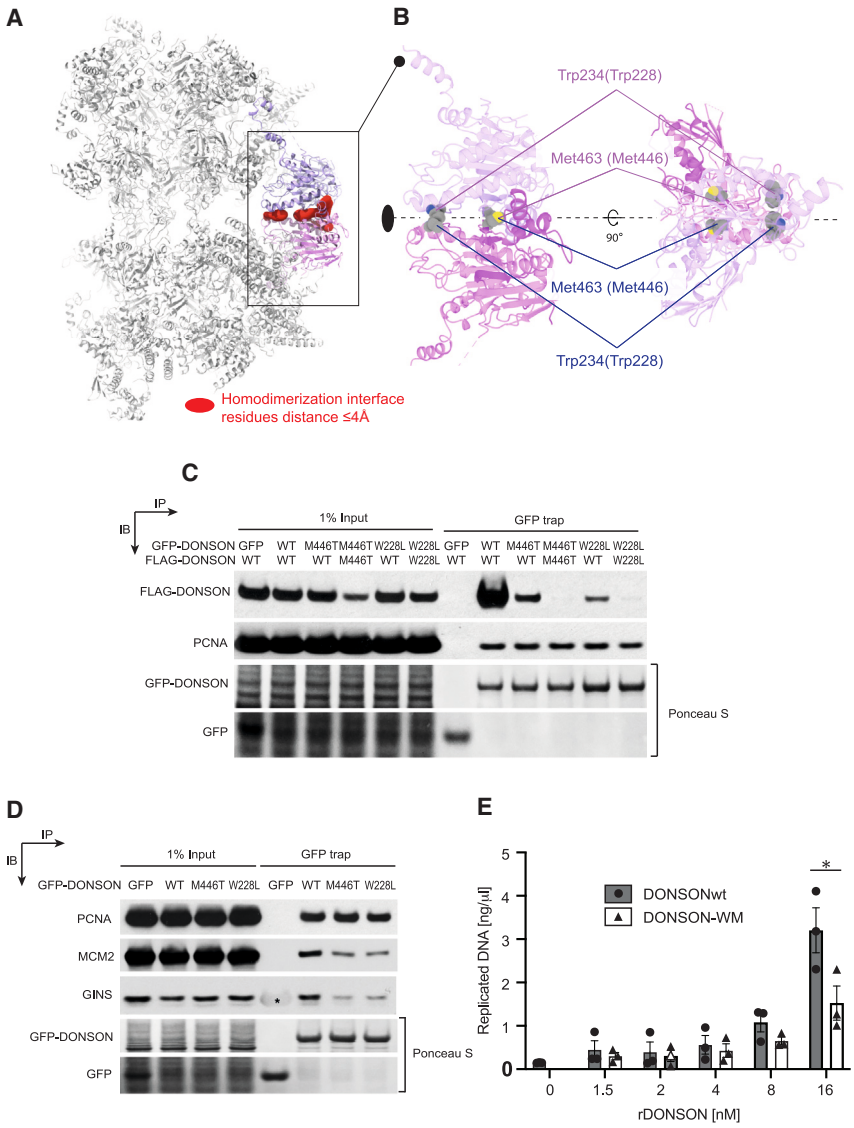


Figure 6. The role of DONSON dimerization in replication initiation

(A) CMG dimerization through the DONSON dimer. Red surface identified distances of 4 Å or less between surface residues of one monomeric CMG-DONSON subcomplex with the symmetry-related subcomplex. The only direct contacts keeping the dCMGDo complex together occur through dimerizing DONSON subunits.

(B) Two views of DONSON reveal that residues Trp234 (human Trp228 mutated in patients) and Met463 (human Met446 mutated in patients) map at the homodimerization interface.

(C) The patient-associated DONSON Met446Thr and Trp228Leu mutations disrupt dimerization. Cells were transfected with different combinations of WT and mutant DONSON tagged with either GFP or FLAG (as indicated). GFP-DONSON was purified from cell extracts using GFP-Trap and co-purified proteins were subjected to SDS-PAGE and western blotting with the antibodies indicated. Ponceau S stain was used to visualize the affinity purified GFP-tagged DONSON on the nitrocellulose filter.

(D) The DONSON Met446Thr and Trp228Leu mutations disrupt binding to the GINS and MCM complex but not PCNA. Cells were transfected with WT and mutant DONSON tagged with GFP (as indicated). GFP-DONSON was purified from cell extracts using GFP-Trap and co-purified proteins were subjected to SDS-PAGE and western blotting with the antibodies indicated. Ponceau S stain was used to visualize the affinity purified GFP-tagged DONSON on the nitrocellulose filter. (*) The antibody raised to the entire GINS complex cross-reacts with purified GFP protein.

(E) DNA replication assay exploring the ability of a WM variant to rescue DNA replication. The replication reaction was set up in DONSON-depleted extract optionally supplemented with indicated concentrations of recombinant wild-type DONSON or DONSON WM mutant. Synthesis of nascent DNA was followed by incorporation of α^{32}P-dATP into newly synthesized DNA in early S-phase. n = 3, mean with SEM presented; paired, nonparametric, two-tailed t test DONSON WT vs. DONSON WM p = 0.0312.

Trp228Leu amino acid changes. The cell lines prepared were either expressing both tagged proteins as mutants (homozygous) or one WT and one mutant (heterozygous). We used western blotting with anti-FLAG M2 antibodies to test for protein expression levels and found a minor reduction in homozygous DONSON Trp228Leu compared with homozygous WT DONSON and a more significant decrease in Met446Thr expression, although the signal remained robust. We then immunoprecipitated DONSON using GFP trap and probed for dimer formation with the M2 antibody that would recognize FLAG-tagged DONSON. We found that heterozygous cells containing WT and Met446Thr alleles or WT and Trp228Leu alleles were significantly defective for DONSON dimer formation. This effect was exacerbated when using homozygous Met446Thr or Trp228Leu DONSON cell lines, with dimer formation completely impaired (Figure 6C). We then asked whether dimer disruption observed with patients' mutations had any effect on DONSON's ability to interact with the established replisome (proliferating cell nuclear antigen [PCNA]), GINS, or with MCM. Although PCNA interaction was unperturbed with all mutants tested, we found that both DONSON M446T and W228L variants have a defect in both GINS and MCM binding (Figure 6D). Finally, we wanted to test whether changes in the *Xenopus* proteins that recapitulate homologous patients' mutations have any effect on DNA replication. To this end, we used bacterial expression to produce a recombinant DONSON variant (DONSON WM) containing both Met463 and Trp234 mutations. We used mass photometry to show that this variant presented defects in homodimerization (Figure S7). Indeed, although the WT protein formed homodimers in a range of protein concentrations, the DONSON WM's ability to homodimerize was concentration dependent. We observed that although DONSON WM measured at 20 nM mainly formed homodimers, it was primarily monomeric at 5 nM (the concentration used to measure

all other DONSON mutants). Given this observation, we decided to titrate WT and DONSON WM mutant in DNA replication experiments after immunodepletion of endogenous DONSON and measure nascent DNA synthesis in early S-phase. Although little DNA synthesis can be observed in the 1.5–4 nM concentration range, we observed that WT DONSON is more efficient at rescuing replication at 8 and 16 nM concentration (Figure 6E). Collectively, our data provide a link between primordial dwarfism, DONSON's homodimerization and the ability to incorporate GINS into the CMG helicase, in a key step on the path to replication initiation.

DISCUSSION

We used the *Xenopus* egg extract system to isolate and image structural intermediates of replication initiation with cryo-EM. We employed conditions to activate origins but halt DNA synthesis. By doing this, we captured an initiating replisome assembly that is specific to higher eukaryotes. It contains DONSON, a replication factor that is absent in yeast and has recently been established as the assembly factor that incorporates GINS into the CMG complex.[41,42] We used structural and biochemical methods to show that DONSON assembles as a homodimer, and we identified essential elements that support its engagement to both MCM and GINS. These structural elements map within flexible loops that emanate from the dimerization core of DONSON. We also identified interaction interfaces that involve the DONSON dimer core and structured domains in both GINS and MCM. These interactions however play a less important role in GINS recruitment and origin activation.

Our observations raise the question of whether DONSON homodimerization is at all important for GINS incorporation into the CMG. To address this question, we turned to cell lines of patients affected by primordial dwarfism. We identified two patients' mutations resulting in amino acid changes at the DONSON dimer interface, and we discovered that dimerization of these protein variants is defective. The same dimerization defect was observed using recombinant expression of a *Xenopus* DONSON variant that contains the homologous amino acid changes. As primordial dwarfism is often associated with mutations in proteins involved in replication initiation,[52] our findings support a role for DONSON dimerization in origin activation. To corroborate this notion, we asked whether, in human cells, the DONSON variants defective for dimerization are still able to interact with GINS and MCM, as observed in our structure. We found that interaction with both GINS and MCM is partially impaired, indicating a direct role for DONSON homodimerization in efficient CMG formation. We then used *Xenopus* egg extracts to show that interfering with DONSON homodimerization makes DNA replication rescue inefficient. Our data invite a model whereby dimeric DONSON recruits two GINS molecules to MCM concurrently, which matches our structural observation of a symmetric dCMGDo initiation intermediate. This mechanism mirrors previous findings with yeast proteins, suggesting that conversion from an inactive MCM double hexamer to a DNA melting CMG-Pol epsilon homodimer occurs symmetrically. Future structural work will establish whether homodimerization is also used by yeast Sld2 as a mechanism to recruit two GINS factors to the double hexamer on the path to origin DNA melting.

What could be the structural mechanism whereby DONSON brings two GINS to assemble the dCMGDo? And how could this molecular intermediate mature to establish two diverging replication forks? Direct docking of DONSON to the MCM double hexamer, as seen in our structure, is impossible due to steric clashes. Thus, based on our functional and structural evidence, we propose that in the first holo-helicase assembly step, unstructured elements of dimeric DONSON, engaging Sld5 and the Mcm3 ATPase, support the incorporation of two copies of GINS into the two CMGs. We envisage that CMG assembly alone would lead to the rotation of two MCMs rings and nucleation of DNA unwinding inside of the two ATPase rings, similar to what we previously observed for the double CMG complex with yeast proteins.[27] After this stage in origin activation, the two helicases must cross paths, so that the N-terminal dimerization interface of the double hexamer becomes poised at the front of the advancing replication fork.[25,28,29] With our structure of a *Xenopus* dCMGDo activation intermediate, we established that the globular DONSON dimer captures the two helicases in a fixed relative orientation where the two MCM hexamers are no longer co-axial. The only way for this structure to exist on DNA involves the unwinding of the Watson and Crick strands at the dimerization interface[49] (Figure 7). What happens after dCMGDo formation is more speculative. For example, lagging strand ejection could occur after dCMGDo formation with the two MCM ring pores misaligned. This step in helicase activation might also involve the poorly defined function of RecQL4 at initiation, which is known to act after CMG formation.[35–40] Exposure of the lagging strand template would in turn trigger the recruitment of single-stranded DNA interacting factors Mcm10, RPA, and Pol alpha. Recruitment of these elongation factors could prompt the release of DONSON, which would in turn let go of the two CMG assemblies. The two helicases at this point would be free to continue along their trajectory toward replication fork establishment. In support of this model, the recent structure of Pol alpha bound to CMG shows that the primase subunit Pri2 engages the A-domain of Mcm3,[53] on a surface that overlaps with the DONSON docking site (Figure S6D). Thus, Pol alpha recruitment to the CMG could compete out dimeric DONSON, and its release from origins could finalize dCMG separation. A DONSON dimer could either remain tethered via a flexible element to one of the two diverging CMGs or be recruited to the advancing replisome at a later stage, when its fork protection functions are required. In this context, it is possible that robust DONSON engagement to a single CMG, similar to what was observed in our structure, will selectively occur at a stalled replication fork, when Okazaki fragments are not being synthesized.

Our previous work established that single-particle cryo-EM can be used to visualize intermediates in replication origin activation reactions, reconstituted *in vitro* with purified yeast protein.[3,17,27] With that approach, we used biochemical tools to stage helicase activation, for example by omitting the essential yeast Mcm10 factor that activates DNA unwinding on the path to origin activation. Our current work extends our structural analysis in three ways. First, using the *Xenopus* egg extract, we analyze the full complexity of the cellular replication

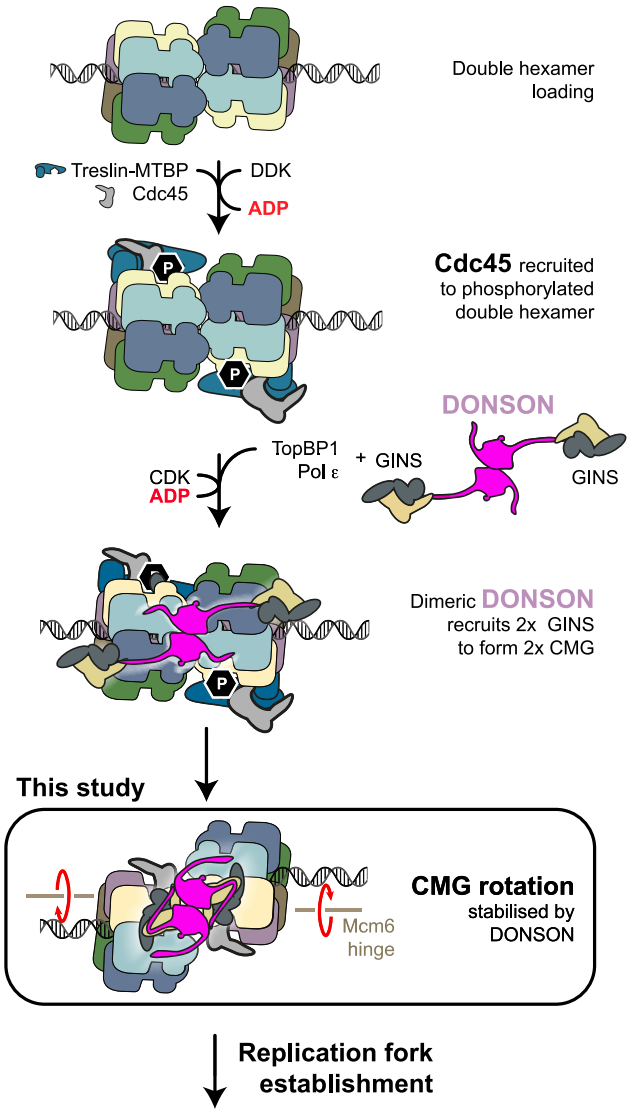


Figure 7. A model for the incorporation of two GINS factors into two CMG assemblies, mediated by dimeric DONSON

Loaded MCM double hexamers are phosphorylated by DDK.

The phosphorylation is read by Treslin-MTBP, which in turn promote Cdc45 recruitment onto the MCM double hexamer. A DONSON dimer engages two copies of the GINS firing factors. It recruits them to the MCM double hexamer by engaging the Mcm3 ATPase via the D3 loop. The DONSON dimer then locks onto the double hexamer, delivering GINS to form the CMG. This leads to a reconfiguration in the double hexamer structure, with two MCM rings that rotate clockwise with respect to one another and become offset ready for replication fork establishment.

process. Second, we introduced purification approaches to target our structural analysis to the CMG replicative helicase assembly and ignore licensed origins that have not started the initiation process. This reduces sample heterogeneity to levels manageable with single-particle cryo-EM. Third, by using an ATM/ATR inhibitor and a replicative-polymerase inhibitor, we showed that we could control intersecting pathways to ask detailed questions regarding structural mechanism. In future work, we aim to image replication reactions in the context of chromatin.

Limitations of the study

Three items limit our study. First, the dCMGDo complex was purified after digesting chromatin with benzonase, and hence, it does not contain DNA. Spatial constraints imposed by symmetric binding of functionally important DONSON elements imply that the relative orientation of the two CMGs is unlikely to change when DNA is retained inside the protein complex. That said, the presence of DNA will likely impart changes to the ATPase tier, including changes in nucleotide occupancy. Imaging the dCMGDo complex on DNA is our future aim. Second, we used aphidicolin to block replicative polymerase activity. Although this strategy captures an early replication intermediate, it does not allow for performing meaningful kinetic studies on the dCMGDo, following complex formation and separation as DNA replication progresses. Third, it remains important to address whether competitive MCM binding by Pol alpha and DONSON has a role in initiating DNA replication.

While our manuscript was under review, two articles were published, predicting[54] or describing[55] the structure of monomeric CMG bound to DONSON. These studies agree with our structure-function characterization and with articles cited in our paper.[41,42] A prediction that two assembled helicases rotate to achieve full DONSON engagement[54] matches our structural observations. Possible competitive CMG binding for DONSON and Pol alpha in regulating fork establishment and progression is also noted.[55]

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS
  - *Xenopus laevis* egg extract model system
  - Cell lines
  - *Escherichia coli* strains
  - Experimental source material for *in vitro* work
- METHOD DETAILS
  - *Xenopus laevis* egg extract
  - Demembranated *Xenopus* sperm
  - DNA synthesis assay
  - Egg extract immunodepletion
  - Chromatin isolation time course
  - Recombinant protein purification
  - Large-scale immunoprecipitation from chromatin
  - Mass spectrometry
  - Mass-spectrometry data processing
  - GFP-Trap affinity purification from human cells
  - Mass photometry

Table 1. *Xenla* CMG cryo-EM data collection and processing

	Xenla sCMG consensus (EMDB: EMD-18195)	*Xenla* dCMGDo consensus (EMDB: EMD-18191) *Xenla* dCMGDo without ATPase (PDB: 8Q6O)	*Xenla* dCMGDo MCM ATPase (EMDB: EMD-18192) *Xenla* dCMGDo MCM ATPase (PDB: 8Q6P)
Data collection and processing			
Nominal magnification	130,000	130,000	130,000
Voltage (keV)	300	300	300
Electron exposure (e^{-}/Å^2)	33	33	33
Defocus range (μm)	−1.3 to −3.4	−1.3 to −3.4	−1.3 to −3.4
Pixel size (Å)	0.95	0.95	0.95
Symmetry imposed	C1	C2	C1
Initial particle number	4,140,303	4,140,303	4,140,303
Final particle numbers	137,591	100,667	144,989
Global map resolution at 0.143 FSC threshold (Å)	3.03	3.14	3.53
Map resolution range at 0.5 FSC (Å)[a]	2.5–9.5	2.7–8.3	2.5–8.5
Map sharpening *B*-factors (Å^2)	−91	−93	−113
Model refinement[b]			
Model-map CC (CC_{mask}/CC_{box}/CC_{peaks}/CC_{volume})	–	0.81/0.66/0.56/0.81	0.78/0.55/0.35/0.77
Model-map FSC, 0.143 threshold (Å) (masked/unmasked)	–	3.1/3.1	3.5/3.7
Model-map FSC, 0.5 threshold (Å) (masked/unmasked)	–	3.4/3.6	3.9/6.4
Model composition			
Non-hydrogen atoms	–	51,983	16,795
Protein residues	–	6,454	2,104
Ligands	–	10 × Zn^{2+}	5 × ATP
RMS deviations			
Bond lengths (Å)	–	0.002	0.002
Bond angles (°)	–	0.492	0.484
Mean *B*-factors (Å^2)			
Protein	–	122.07	108.87
Ligands	–	260.17	83.98
Validation[c]			
MolProbity score	–	1.35	1.35
Clashscore	–	4.9	5.07
Poor rotamers (%)	–	0.05	0.05
CaBLAM outliers (%)	–	1.20	1.48
Ramachandran plot[c]			
Favored (%)	–	97.6	97.6
Allowed (%)	–	2.4	2.4
Outliers (%)	–	0.0	0.0

[a]Estimated with cryoSPARC.[64]
[b]Calculated with comprehensive validation (cryo-EM) tool in *PHENIX*.[69]
[c]MolProbity[77] validation in *PHENIX*.

- ○ Sample preparation and data collection for negative stain electron microscopy
- ○ Negative-stain EM image processing
- ○ Sample preparation and data collection for cryo-EM
- ○ Cryo-EM image processing
- ○ Model building and refinement
- ○ Visualization of models and maps

● QUANTIFICATION AND STATISTICAL ANALYSIS

- DNA synthesis analysis
- EM data analysis

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.molcel.2023.09.029.

ACKNOWLEDGMENTS

We acknowledge the Crick Structural Biology STP for support with computation (A. Nans, A. Purkiss, and P. Walker) and cryo-EM support (A. Nans and D. Benton). This work was funded jointly by the Wellcome Trust, MRC, and CRUK at the Francis Crick Institute (FC001065). A.C. receives funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement no. 820102). M.A.C. was supported by a Marie Skłodowska-Curie Postdoctoral Fellowship (agreement no. 845939). A Wellcome Trust Investigator Award (215510/Z/19/Z) to A.G. funded P.P., BBSRC (BB/T001860/1) was awarded to A.G. and funded A.S. and A.R.-W., and MRC funded IMPACT PhD studentship for G.K. and BBSRC MIBTP studentship for C.F.-C. Work by G.S.S.'s group is funded by a CRUK programme grant (C17183/A23303), which also funded S.S.J. An MRC project grant MR/W001152/1 awarded to G.S.S. funded A.S.C.

AUTHOR CONTRIBUTIONS

Investigation—performed all electron and cryoelectron microscopy experiments, image processing, molecular modeling, mass photometry, data interpretation, and writing, M.A.C.; investigation—performed replicative complexes isolation, *in vitro* biochemistry, data interpretation, and writing, P.P.; investigation—performed atomic model building, and helped with biochemistry, data interpretation, and writing, A.B.; investigation—performed mutant generation and biochemistry in *X.l.* egg extract, A.R.-W.; investigation—performed mutant purification and biochemistry in *X.l.* egg extract, G.K.; investigation—performed mutants generation and purification, A.S.; investigation—performed biochemistry in *X.l.* egg extract, C.F.-C.; investigation—performed biochemistry in *X.l.* egg extract, D.P.; investigation—performed mass photometry analysis R.G.; investigation—performed work with human cell lines, A.S.C.; investigation— performed work with human cell lines, S.S.J.; supervision, funding acquisition, writing, and investigation—generated cell lines and performed GFP-Trap affinity purification, G.S.S.; conceptualization, supervision, funding acquisition, and writing, A.G.; conceptualization, supervision, funding acquisition, and investigation—cryoelectron microscopy screening, writing, A.C.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: August 15, 2023
Revised: September 15, 2023
Accepted: September 25, 2023
Published: October 10, 2023

REFERENCES

1. Attali, I., Botchan, M.R., and Berger, J.M. (2021). Structural mechanisms for replicating DNA in eukaryotes. Annu. Rev. Biochem. *90*, 77–106. https://doi.org/10.1146/annurev-biochem-090120-125407.
2. Costa, A., and Diffley, J.F.X. (2022). The initiation of eukaryotic DNA replication. Annu. Rev. Biochem. *91*, 107–131. https://doi.org/10.1146/annurev-biochem-072321-110228.
3. Miller, T.C.R., Locke, J., Greiwe, J.F., Diffley, J.F.X., and Costa, A. (2019). Mechanism of head-to-head MCM double-hexamer formation revealed by cryo-EM. Nature *575*, 704–710. https://doi.org/10.1038/s41586-019-1768-0.
4. Remus, D., Beuron, F., Tolun, G., Griffith, J.D., Morris, E.P., and Diffley, J.F. (2009). Concerted loading of Mcm2-7 double hexamers around DNA during DNA replication origin licensing. Cell *139*, 719–730. https://doi.org/10.1016/j.cell.2009.10.015.
5. Evrin, C., Clarke, P., Zech, J., Lurz, R., Sun, J., Uhle, S., Li, H., Stillman, B., and Speck, C. (2009). A double-hexameric MCM2-7 complex is loaded onto origin DNA during licensing of eukaryotic DNA replication. Proc. Natl. Acad. Sci. USA *106*, 20240–20245. https://doi.org/10.1073/pnas.0911500106.
6. Noguchi, Y., Yuan, Z., Bai, L., Schneider, S., Zhao, G., Stillman, B., Speck, C., and Li, H. (2017). Cryo-EM structure of Mcm2-7 double hexamer on DNA suggests a lagging-strand DNA extrusion model. Proc. Natl. Acad. Sci. USA *114*, E9529–E9538. https://doi.org/10.1073/pnas.1712537114.
7. Abid Ali, F., Douglas, M.E., Locke, J., Pye, V.E., Nans, A., Diffley, J.F.X., and Costa, A. (2017). Cryo-EM structure of a licensed DNA replication origin. Nat. Commun. *8*, 2241. https://doi.org/10.1038/s41467-017-02389-0.
8. Moyer, S.E., Lewis, P.W., and Botchan, M.R. (2006). Isolation of the Cdc45/Mcm2-7/GINS (CMG) complex, a candidate for the eukaryotic DNA replication fork helicase. Proc. Natl. Acad. Sci. USA *103*, 10236–10241.
9. Ilves, I., Petojevic, T., Pesavento, J.J., and Botchan, M.R. (2010). Activation of the MCM2-7 helicase by association with Cdc45 and GINS proteins. Mol. Cell *37*, 247–258.
10. Pacek, M., Tutter, A.V., Kubota, Y., Takisawa, H., and Walter, J.C. (2006). Localization of MCM2-7, Cdc45, and GINS to the site of DNA unwinding during eukaryotic DNA replication. Mol. Cell *21*, 581–587.
11. Gambus, A., Jones, R.C., Sanchez-Diaz, A., Kanemaki, M., van Deursen, F., Edmondson, R.D., and Labib, K. (2006). GINS maintains association of Cdc45 with MCM in replisome progression complexes at eukaryotic DNA replication forks. Nat. Cell Biol. *8*, 358–366. https://doi.org/10.1038/ncb1382.
12. Bell, S.P., and Labib, K. (2016). Chromosome duplication in Saccharomyces cerevisiae. Genetics *203*, 1027–1067. https://doi.org/10.1534/genetics.115.186452.
13. Francis, L.I., Randell, J.C., Takara, T.J., Uchima, L., and Bell, S.P. (2009). Incorporation into the prereplicative complex activates the Mcm2-7 helicase for Cdc7-Dbf4 phosphorylation. Genes Dev. *23*, 643–654. https://doi.org/10.1101/gad.1759609.
14. Sheu, Y.J., and Stillman, B. (2006). Cdc7-Dbf4 phosphorylates MCM proteins via a docking site-mediated mechanism to promote S phase progression. Mol. Cell *24*, 101–113. https://doi.org/10.1016/j.molcel.2006.07.033.
15. Sheu, Y.J., and Stillman, B. (2010). The Dbf4-Cdc7 kinase promotes S phase by alleviating an inhibitory activity in Mcm4. Nature *463*, 113–117. https://doi.org/10.1038/nature08647.
16. Labib, K. (2010). How do Cdc7 and cyclin-dependent kinases trigger the initiation of chromosome replication in eukaryotic cells? Genes Dev. *24*, 1208–1219. https://doi.org/10.1101/gad.1933010.
17. Greiwe, J.F., Miller, T.C.R., Locke, J., Martino, F., Howell, S., Schreiber, A., Nans, A., Diffley, J.F.X., and Costa, A. (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. https://doi.org/10.1038/s41594-021-00698-z.
18. Saleh, A., Noguchi, Y., Aramayo, R., Ivanova, M.E., Stevens, K.M., Montoya, A., Sunidhi, S., Carranza, N.L., Skwark, M.J., and Speck, C. (2022). The structural basis of Cdc7-Dbf4 kinase dependent targeting and phosphorylation of the MCM2-7 double hexamer. Nat. Commun. *13*, 2915. https://doi.org/10.1038/s41467-022-30576-1.
19. Cheng, J., Li, N., Huo, Y., Dang, S., Tye, B.K., Gao, N., and Zhai, Y. (2022). Structural Insight into the MCM double hexamer activation by Dbf4-Cdc7 kinase. Nat. Commun. *13*, 1396. https://doi.org/10.1038/s41467-022-29070-5.

20. Deegan, T.D., Yeeles, J.T., and Diffley, J.F. (2016). Phosphopeptide binding by Sld3 links Dbf4-dependent kinase to MCM replicative helicase activation. EMBO J. *35*, 961–973. https://doi.org/10.15252/embj.201593552.

21. Tanaka, S., Nakato, R., Katou, Y., Shirahige, K., and Araki, H. (2011). Origin association of Sld3, Sld7, and Cdc45 proteins is a key step for determination of origin-firing timing. Curr. Biol. *21*, 2055–2063. https://doi.org/10.1016/j.cub.2011.11.038.

22. Muramatsu, S., Hirai, K., Tak, Y.S., Kamimura, Y., and Araki, H. (2010). CDK-dependent complex formation between replication proteins Dpb11, Sld2, Pol (epsilon}, and GINS in budding yeast. Genes Dev. *24*, 602–612. https://doi.org/10.1101/gad.1883410.

23. Zegerman, P., and Diffley, J.F. (2007). Phosphorylation of Sld2 and Sld3 by cyclin-dependent kinases promotes DNA replication in budding yeast. Nature *445*, 281–285. https://doi.org/10.1038/nature05432.

24. Tanaka, S., Umemori, T., Hirai, K., Muramatsu, S., Kamimura, Y., and Araki, H. (2007). CDK-dependent phosphorylation of Sld2 and Sld3 initiates DNA replication in budding yeast. Nature *445*, 328–332. https://doi.org/10.1038/nature05465.

25. Douglas, M.E., Ali, F.A., Costa, A., and Diffley, J.F.X. (2018). The mechanism of eukaryotic CMG helicase activation. Nature *555*, 265–268. https://doi.org/10.1038/nature25787.

26. Goswami, P., Abid Ali, F., Douglas, M.E., Locke, J., Purkiss, A., Janska, A., Eickhoff, P., Early, A., Nans, A., Cheung, A.M.C., et al. (2018). Structure of DNA-CMG-Pol epsilon elucidates the roles of the non-catalytic polymerase modules in the eukaryotic replisome. Nat. Commun. *9*, 5061. https://doi.org/10.1038/s41467-018-07417-1.

27. Lewis, J.S., Gross, M.H., Sousa, J., Henrikus, S.S., Greiwe, J.F., Nans, A., Diffley, J.F.X., and Costa, A. (2022). Mechanism of replication origin melting nucleated by CMG helicase assembly. Nature *606*, 1007–1014. https://doi.org/10.1038/s41586-022-04829-4.

28. Eickhoff, P., Kose, H.B., Martino, F., Petojevic, T., Abid Ali, F., Locke, J., Tamberg, N., Nans, A., Berger, J.M., Botchan, M.R., et al. (2019). Molecular basis for ATP-hydrolysis-driven DNA translocation by the CMG helicase of the eukaryotic replisome. Cell Rep. *28*, 2673–2688.e8. https://doi.org/10.1016/j.celrep.2019.07.104.

29. Georgescu, R., Yuan, Z., Bai, L., de Luna Almeida Santos, R., Sun, J., Zhang, D., Yurieva, O., Li, H., and O'Donnell, M.E. (2017). Structure of eukaryotic CMG helicase at a replication fork and implications to replisome architecture and origin initiation. Proc. Natl. Acad. Sci. USA *114*, E697–E706. https://doi.org/10.1073/pnas.1620500114.

30. Li, N., Zhai, Y., Zhang, Y., Li, W., Yang, M., Lei, J., Tye, B.K., and Gao, N. (2015). Structure of the eukaryotic MCM complex at 3.8 A. Nature *524*, 186–191. https://doi.org/10.1038/nature14685.

31. Li, J., Dong, J., Wang, W., Yu, D., Fan, X., Hui, Y.C., Lee, C.S.K., Lam, W.H., Alary, N., Yang, Y., et al. (2023). The human pre-replication complex is an open complex. Cell *186*, 98–111.e21. https://doi.org/10.1016/j.cell.2022.12.008.

32. Boos, D., Yekezare, M., and Diffley, J.F. (2013). Identification of a heteromeric complex that promotes DNA replication origin firing in human cells. Science *340*, 981–984. https://doi.org/10.1126/science.1237448.

33. Boos, D., Sanchez-Pulido, L., Rappas, M., Pearl, L.H., Oliver, A.W., Ponting, C.P., and Diffley, J.F. (2011). Regulation of DNA replication through Sld3-Dpb11 interaction is conserved from yeast to humans. Curr. Biol. *21*, 1152–1157. https://doi.org/10.1016/j.cub.2011.05.057.

34. Kumagai, A., Shevchenko, A., Shevchenko, A., and Dunphy, W.G. (2011). Direct regulation of Treslin by cyclin-dependent kinase is essential for the onset of DNA replication. J. Cell Biol. *193*, 995–1007. https://doi.org/10.1083/jcb.201102003.

35. Thangavel, S., Mendoza-Maldonado, R., Tissino, E., Sidorova, J.M., Yin, J., Wang, W., Monnat, R.J., Jr., Falaschi, A., and Vindigni, A. (2010). Human RECQ1 and RECQ4 helicases play distinct roles in DNA replication initiation. Mol. Cell. Biol. *30*, 1382–1396. https://doi.org/10.1128/MCB.01290-09.

36. Sangrithi, M.N., Bernal, J.A., Madine, M., Philpott, A., Lee, J., Dunphy, W.G., and Venkitaraman, A.R. (2005). Initiation of DNA replication requires the RECQL4 protein mutated in Rothmund-Thomson syndrome. Cell *121*, 887–898. https://doi.org/10.1016/j.cell.2005.05.015.

37. Matsuno, K., Kumano, M., Kubota, Y., Hashimoto, Y., and Takisawa, H. (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. https://doi.org/10.1128/MCB.02267-05.

38. Sanuki, Y., Kubota, Y., Kanemaki, M.T., Takahashi, T.S., Mimura, S., and Takisawa, H. (2015). RecQ4 promotes the conversion of the pre-initiation complex at a site-specific origin for DNA unwinding in Xenopus egg extracts. Cell Cycle *14*, 1010–1023. https://doi.org/10.1080/15384101.2015.1007003.

39. Kliszczak, M., Sedlackova, H., Pitchai, G.P., Streicher, W.W., Krejci, L., and Hickson, I.D. (2015). Interaction of RECQ4 and MCM10 is important for efficient DNA replication origin firing in human cells. Oncotarget *6*, 40464–40479. https://doi.org/10.18632/oncotarget.6342.

40. Im, J.S., Ki, S.H., Farina, A., Jung, D.S., Hurwitz, J., and Lee, J.K. (2009). Assembly of the Cdc45-Mcm2-7-GINS complex in human cells requires the Ctf4/And-1, RecQL4, and Mcm10 proteins. Proc. Natl. Acad. Sci. USA *106*, 15628–15632. https://doi.org/10.1073/pnas.0908039106.

41. Hashimoto, Y., Sadano, K., Miyata, N., Ito, H., and Tanaka, H. (2023). Novel role of DONSON in CMG helicase assembly during vertebrate DNA replication initiation. EMBO J. *42*, e114131. https://doi.org/10.15252/embj.2023114131.

42. Kingsley, G., Skagia, A., Passaretti, P., Fernandez-Cuesta, C., Reynolds-Winczura, A., Koscielniak, K., and Gambus, A. (2023). DONSON facilitates Cdc45 and GINS chromatin association and is essential for DNA replication initiation. Nucleic Acids Res. https://doi.org/10.1093/nar/gkad694.

43. Reynolds, J.J., Bicknell, L.S., Carroll, P., Higgs, M.R., Shaheen, R., Murray, J.E., Papadopoulos, D.K., Leitch, A., Murina, O., Tarnauskaitė, Ž., et al. (2017). Mutations in DONSON disrupt replication fork stability and cause microcephalic dwarfism. Nat. Genet. *49*, 537–549. https://doi.org/10.1038/ng.3790.

44. Nerakh, G., Vineeth, V.S., Tallapaka, K., Nair, L., Dalal, A., and Aggarwal, S. (2022). Microcephalic primordial dwarfism with predominant Meier-Gorlin phenotype, ichthyosis, and multiple joint deformities-Further expansion of DONSON Cell Cycle-opathy phenotypic spectrum. Am. J. Med. Genet. A *188*, 2139–2146. https://doi.org/10.1002/ajmg.a.62725.

45. Zhang, J., Bellani, M.A., James, R.C., Pokharel, D., Zhang, Y., Reynolds, J.J., McNee, G.S., Jackson, A.P., Stewart, G.S., and Seidman, M.M. (2020). DONSON and FANCM associate with different replisomes distinguished by replication timing and chromatin domain. Nat. Commun. *11*, 3951. https://doi.org/10.1038/s41467-020-17449-1.

46. Sarkaria, J.N., Busby, E.C., Tibbetts, R.S., Roos, P., Taya, Y., Karnitz, L.M., and Abraham, R.T. (1999). Inhibition of ATM and ATR kinase activities by the radiosensitizing agent, caffeine. Cancer Res. *59*, 4375–4382.

47. Krokan, H., Wist, E., and Krokan, R.H. (1981). Aphidicolin inhibits DNA synthesis by DNA polymerase alpha and isolated nuclei by a similar mechanism. Nucleic Acids Res. *9*, 4709–4719. https://doi.org/10.1093/nar/9.18.4709.

48. Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A., et al. (2021). Highly accurate protein structure prediction with AlphaFold. Nature *596*, 583–589. https://doi.org/10.1038/s41586-021-03819-2.

49. Meagher, M., Epling, L.B., and Enemark, E.J. (2019). DNA translocation mechanism of the MCM complex and implications for replication initiation. Nat. Commun. *10*, 3117. https://doi.org/10.1038/s41467-019-11074-3.

50. Milner, R.D., Khallouf, K.A., Gibson, R., Hajianpour, A., and Mathew, C.G. (1993). A new autosomal recessive anomaly mimicking Fanconi's anaemia phenotype. Arch. Dis. Child. *68*, 101–103. https://doi.org/10.1136/adc.68.1.101.

51. Karaca, E., Posey, J.E., Bostwick, B., Liu, P., Gezdirici, A., Yesil, G., Coban Akdemir, Z., Bayram, Y., Harms, F.L., Meinecke, P., et al. (2019). Biallelic and de novo variants in DONSON reveal a clinical spectrum of cell cycle-opathies with microcephaly, dwarfism and skeletal abnormalities. Am. J. Med. Genet. A *179*, 2056–2066. https://doi.org/10.1002/ajmg.a.61315.

52. Bellelli, R., and Boulton, S.J. (2021). Spotlight on the replisome: aetiology of DNA replication-associated genetic diseases. Trends Genet. *37*, 317–336. https://doi.org/10.1016/j.tig.2020.09.008.

53. Jones, M.L., Aria, V., Baris, Y., and Yeeles, J.T.P. (2023). How Pol alpha-primase is targeted to replisomes to prime eukaryotic DNA replication. Mol. Cell *83*, 2911–2924.e16. https://doi.org/10.1016/j.molcel.2023.06.035.

54. Lim, Y., Tamayo-Orrego, L., Schmid, E., Tarnauskaite, Z., Kochenova, O.V., Gruar, R., Muramatsu, S., Lynch, L., Schlie, A.V., Carroll, P.L., et al. (2023). In silico protein interaction screening uncovers DONSON's role in replication initiation. Science *381*, eadi3448. https://doi.org/10.1126/science.adi3448.

55. Xia, Y., Sonneville, R., Jenkyn-Bedford, M., Ji, L., Alabert, C., Hong, Y., Yeeles, J.T.P., and Labib, K.P.M. (2023). DNSN-1 recruits GINS for CMG helicase assembly during DNA replication initiation in Caenorhabditis elegans. Science *381*, eadi4932. https://doi.org/10.1126/science.adi4932.

56. Gambus, A., Khoudoli, G.A., Jones, R.C., and Blow, J.J. (2011). MCM2-7 form double hexamers at licensed origins in Xenopus egg extract. J. Biol. Chem. *286*, 11855–11864. https://doi.org/10.1074/jbc.M110.199521.

57. Tarcan, Z., Poovathumkadavil, D., Skagia, A., and Gambus, A. (2022). The p97 segregase cofactor Ubxn7 facilitates replisome disassembly during S-phase. J. Biol. Chem. *298*, 102234. https://doi.org/10.1016/j.jbc.2022.102234.

58. Priego Moreno, S., Jones, R.M., Poovathumkadavil, D., Scaramuzza, S., and Gambus, A. (2019). Mitotic replisome disassembly depends on TRAIP ubiquitin ligase activity. Life Sci. Alliance *2*, 2. https://doi.org/10.26508/lsa.201900390.

59. Moreno, S.P., Bailey, R., Campion, N., Herron, S., and Gambus, A. (2014). Polyubiquitylation drives replisome disassembly at the termination of DNA replication. Science *346*, 477–481. https://doi.org/10.1126/science.1253585.

60. Volpi, I., Gillespie, P.J., Chadha, G.S., and Blow, J.J. (2021). The role of DDK and Treslin-MTBP in coordinating replication licensing and pre-initiation complex formation. Open Biol. *11*, 210121. https://doi.org/10.1098/rsob.210121.

61. Wagner, T., Merino, F., Stabrin, M., Moriya, T., Antoni, C., Apelbaum, A., Hagel, P., Sitsel, O., Raisch, T., Prumbaum, D., et al. (2019). SPHIRE-crYOLO is a fast and accurate fully automated particle picker for cryo-EM. Commun. Biol. *2*, 218. https://doi.org/10.1038/s42003-019-0437-z.

62. Kimanius, D., Dong, L., Sharov, G., Nakane, T., and Scheres, S.H.W. (2021). New tools for automated cryo-EM single-particle analysis in RELION-4.0. Biochem. J. *478*, 4169–4185. https://doi.org/10.1042/BCJ20210708.

63. Zivanov, J., Nakane, T., Forsberg, B.O., Kimanius, D., Hagen, W.J., Lindahl, E., and Scheres, S.H. (2018). New tools for automated high-resolution cryo-EM structure determination in RELION-3. eLife *7*. https://doi.org/10.7554/eLife.42166.

64. Punjani, A., Rubinstein, J.L., Fleet, D.J., and Brubaker, M.A. (2017). cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. Nat. Methods *14*, 290–296. https://doi.org/10.1038/nmeth.4169.

65. Zhang, K. (2016). Gctf: real-time CTF determination and correction. J. Struct. Biol. *193*, 1–12. https://doi.org/10.1016/j.jsb.2015.11.003.

66. Rohou, A., and Grigorieff, N. (2015). CTFFIND4: fast and accurate defocus estimation from electron micrographs. J. Struct. Biol. *192*, 216–221. https://doi.org/10.1016/j.jsb.2015.08.008.

67. Bepler, T., Morin, A., Rapp, M., Brasch, J., Shapiro, L., Noble, A.J., and Berger, B. (2019). Positive-unlabeled convolutional neural networks for particle picking in cryo-electron micrographs. Nat. Methods *16*, 1153–1160. https://doi.org/10.1038/s41592-019-0575-8.

68. Emsley, P., Lohkamp, B., Scott, W.G., and Cowtan, K. (2010). Features and development of coot. Acta Crystallogr. D Biol. Crystallogr. *66*, 486–501. https://doi.org/10.1107/S0907444910007493.

69. Liebschner, D., Afonine, P.V., Baker, M.L., Bunkóczi, G., Chen, V.B., Croll, T.I., Hintze, B., Hung, L.W., Jain, S., McCoy, A.J., et al. (2019). Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. Acta Crystallogr. D Struct. Biol. *75*, 861–877. https://doi.org/10.1107/S2059798319011471.

70. Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., and Ferrin, T.E. (2004). UCSF Chimera–a visualization system for exploratory research and analysis. J. Comput. Chem. *25*, 1605–1612.

71. Goddard, T.D., Huang, C.C., Meng, E.C., Pettersen, E.F., Couch, G.S., Morris, J.H., and Ferrin, T.E. (2018). UCSF ChimeraX: meeting modern challenges in visualization and analysis. Protein Sci. *27*, 14–25. https://doi.org/10.1002/pro.3235.

72. Gillespie, P.J., Gambus, A., and Blow, J.J. (2012). Preparation and use of Xenopus egg extracts to study DNA replication and chromatin associated proteins. Methods *57*, 203–213. https://doi.org/10.1016/j.ymeth.2012.03.029.

73. Blow, J.J., and Laskey, R.A. (1986). Initiation of DNA replication in nuclei and purified DNA by a cell-free extract of Xenopus eggs. Cell *47*, 577–587. https://doi.org/10.1016/0092-8674(86)90622-7.

74. Simon, A.C., Sannino, V., Costanzo, V., and Pellegrini, L. (2016). Structure of human Cdc45 and implications for CMG helicase function. Nat. Commun. *7*, 11638. https://doi.org/10.1038/ncomms11638.

75. Käll, L., Storey, J.D., and Noble, W.S. (2008). Non-parametric estimation of posterior error probabilities associated with peptides identified by tandem mass spectrometry. Bioinformatics *24*, i42–i48. https://doi.org/10.1093/bioinformatics/btn294.

76. Nesvizhskii, A.I., Keller, A., Kolker, E., and Aebersold, R. (2003). A statistical model for identifying proteins by tandem mass spectrometry. Anal. Chem. *75*, 4646–4658. https://doi.org/10.1021/ac0341261.

77. Chen, V.B., Arendall, W.B., 3rd, Headd, J.J., Keedy, D.A., Immormino, R.M., Kapral, G.J., Murray, L.W., Richardson, J.S., and Richardson, D.C. (2010). MolProbity: all-atom structure validation for macromolecular crystallography. Acta Crystallogr. D Biol. Crystallogr. *66*, 12–21. https://doi.org/10.1107/S0907444909042073.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Sheep polyclonal anti-Psf2	Gambus lab	Gambus et al.[56]
Sheep polyclonal anti-Cdc45	Gambus lab	Gambus et al.[56]
Sheep polyclonal anti-GINS	Gambus lab	Tarcan et al.[57]
Rabbit polyclonal anti-DONSON	Gambus lab	Kingsley et al.[42]
Sheep polyclonal anti-DONSON	Gambus lab	Kingsley et al.[42]
Sheep polyclonal anti-Mcm7	Gambus lab	Priego Moreno et al.[58]
Sheep polyclonal anti-RecQ4	Gambus lab	Moreno et al.[59]
Sheep polyclonal anti-TopBP1	Gambus lab	Moreno et al.[59]
Sheep polyclonal anti-Treslin	Blow lab	Volpi et al.[60]
Sheep polyclonal anti-MTBP	Blow lab	Volpi et al.[60]
IgG from rabbit serum	Sigma-Aldrich	RRID: AB_1163659
Polyclonal rabbit anti-FLAG	Sigma-Aldrich	RRID: AB_439687
Monoclonal mouse anti-Mcm2	BD Transduction labs	RRID: AB_2141952
Monoclonal mouse anti-PCNA	Santa Cruz Biotechnology	RRID: AB_628110
Monoclonal mouse anti-PCNA	Sigma Aldrich	RRID: AB_477413
Bacterial and virus strains		
Rosetta™(DE3) pLysS Competent Cells	Novagen (Merck)	70956
BL21(DE3) Competent Cells	Novagen (Merck)	69451
Chemicals, peptides, and recombinant proteins		
Aphidicolin	Bio-Techne Limited	Cat# 5736/1
Caffeine	Sigma-Aldrich	CAS: 58-08-2
ANTI-FLAG® M2 Magnetic Beads	Sigma-Aldrich	Cat# M8823
3X FLAG Peptide	Stratech Scientific Ltd	A6001
HisTrap HP	Cytiva	Cat# 17524801
HiTrap Heparin HP	Cytiva	Cat# 17040701
Cdc45-TEV-His_{10}-$FLAG_5$	This study	Addgene agreement 83268
DONSON WT	Gambus lab	Kingsley et al.[42]
DONSON DEW (D374A; E377A; W381A)	This study	Addgene agreement 83268
DONSON S437A	This study	Addgene agreement 83268
DONSON WM (W234L; M463T)	This study	Addgene agreement 83268
DONSON 3A (V473A; R476A; Y481A)	This study	Addgene agreement 83268
DONSON Δ24 (Δ1-24)	This study	Addgene agreement 83268
DONSON DEWS (D374A; E377A; W381A; S437A)	This study	Addgene agreement 83268
DONSON ΔD1 (Δ1-155)	Gambus lab	Kingsley et al.[42]
GINS: (Psf1, Psf2, Psf3 and Sld5)	Gambus lab	Tarcan et al.[57]
Graphene oxide	Sigma-Aldrich	763705-25ML
Deposited data		
Xenla dCMGDo consensus cryo-EM map	This study	EMDB: EMD-18191
Xenla dCMGDo ATPase cryo-EM map	This study	EMDB: EMD-18192
Xenla sCMG consensus cryo-EM map	This study	EMDB: EMD-18195
Xenla dCMGDo – without MCM ATPase PDB coordinates	This study	PDB: 8Q6O
Xenla dCMGDo – MCM ATPase PDB coordinates	This study	PDB: 8Q6P

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Xenopus laevis Cdc45-FLAG interactions on chromatin during DNA replication – Mass spectrometry data project	This study	PRIDE: PXD044422
Experimental models: Cell lines		
293FT	Thermo Fisher Scientific	R70007
Experimental models: Organisms/strains		
Xenopus laevis	Gambus lab	N/A
Oligonucleotides		
DONSON DEW (D374A; E377A; W381A) forward: CTCTGCGCTAGAAGAGATGGGAGTTGAAGA TAAAATTAAAAAG;	Sigma-Aldrich	N/A
DONSON DEW (D374A; E377A; W381A) reverse: AAGCTCGCATCTTCCGCGCTTTCATC CGATGCCTC	Sigma-Aldrich	N/A
DONSON S437A forward: AAACTGTAAGGC GATTGTCGCTGC	Sigma-Aldrich	N/A
DONSON S437A reverse: TGAGGAAGTT CAGTAAAG	Sigma-Aldrich	N/A
DONSON W234L forward: GTCGCTTCCTCT GGTACAGCTTTTTC	Sigma-Aldrich	N/A
DONSON W234L reverse: GGATGAATCCAG TGCACAAG	Sigma-Aldrich	N/A
DONSON M463T forward: AGGAGCAACAA CGCACGCGCTTA	Sigma-Aldrich	N/A
DONSON M463T reverse: CGGAATGCGA CTGGAGAC	Sigma-Aldrich	N/A
DONSON 3A (V473A; R476A; Y481A) forward: GAACTCTGGTGCGAAGGATCAGTTTAGTCTGG;	Sigma-Aldrich	N/A
DONSON 3A (V473A; R476A; Y481A) reverse: ACCGCAGTCTTCGCATTCACACTCCTTG CCTTAAG	Sigma-Aldrich	N/A
DONSON Δ24 (Δ1-24) forward: AGGAGCAGG AGCGAGGCT	Sigma-Aldrich	N/A
DONSON Δ24 (Δ1-24) reverse: CAGATCCTCT TCTGAGATGAGTTTTTGTTCG	Sigma-Aldrich	N/A
DONSON ΔD1 (Δ1-155) forward: TCCGTGAC TTTTCCTGCT	Sigma-Aldrich	N/A
DONSON ΔD1 (Δ1-155) reverse: CATTTATTT CCAGGTGTACC	Sigma-Aldrich	N/A
Recombinant DNA		
pET28a-DONSON	Gambus lab	Addgene: 208797
pET28a-DONSON ΔD1	Gambus lab	Addgene: 208804
pET28a-DONSON-WM (W234L, M463T)	This study	Addgene: 208798
pET28a-DONSON-DEWS (D374A, E377A, W381A, S437A)	This study	Addgene: 208799
pET28a-DONSON-3A (V473A, R476A, Y481A)	This study	Addgene: 208800
pET28a-DONSON Δ24 (1-24 removed)	This study	Addgene: 208801
pET28a-DONSON-S437A	This study	Addgene: 208803
pET28a-DONSON-DEW (D374A, E377A, W381A)	This study	Addgene: 208802
Software and algorithms		
crYOLO 1.4.0	Wagner et al.[61]	https://cryolo.readthedocs.io/en/stable/

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Relion 3.0.4 and 4.0.0	Kimanius et al.[62]; Zivanov et al[63]	https://www3.mrc-lmb.cam.ac.uk/relion//index.php?title=Main_Page
cryoSPARC v4.2.1	Punjani et al.[64]	https://cryosparc.com/
Gctf	Zhang et al.[65]	https://sbgrid.org/software/titles/gctf
CtfFind-4.1	Rohou et al.[66]	https://grigoriefflab.umassmed.edu/ctffind4
Topaz-0.2.5	Bepler et al.[67]	https://cb.csail.mit.edu/cb/topaz/
Coot v0.9.8.1	Emsley et al.[68]	https://www2.mrc-lmb.cam.ac.uk/personal/pemsley/coot/
PHENIX v1.20.1	Liebschner et al.[69]	http://www.phenix-online.org/
UCSF Chimera V1.16	Pettersen et al.[70]	https://www.cgl.ucsf.edu/chimera/
UCSF ChimeraX	Goddard et al.[71]	https://www.rbvi.ucsf.edu/chimerax/
Mascot v 2.8.0.1	Matrix Science, London, UK	https://www.matrixscience.com/
Scaffold v 5.1.0	Proteome Software Inc., Portland	https://proteomesoftware.com/
AcquireMP v 2.5	Refeyn Ltd., Oxford, UK	https://www.refeyn.com/
DiscoverMP v 2.5	Refeyn Ltd., Oxford, UK	https://www.refeyn.com/
GraphPad Prism v 10.0.2	GraphPad Software LLC	https://www.graphpad.com/

RESOURCE AVAILABILITY

Lead contact

Further Information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Alessandro Costa (alessandro.costa@crick.ac.uk).

Materials availability

Plasmids generated in this study are deposited to Addgene with accession numbers indicated in the key resources table.

Data and code availability

- Datasets are available as indicated below:
- Mass spectrometry analysis of immunoprecipitated Cdc45-TEV-His_{10}-$FLAG_5$ were deposited in PRIDE database and publicly available as of the date of publication. Accession numbers are listed here and in the key resources table. **Project Name**: *Xenopus laevis* Cdc45-FLAG interactions on chromatin during DNA replication. **Project accession**: PRIDE: PXD044422 **Project** https://doi.org/10.6019/PXD044422x
- Cryo-EM density maps used in model building have been deposited in the Electron Microscopy Data Bank (EMDB) https://www.ebi.ac.uk/pdbe/emdb, under the following accession numbers: EMDB: EMD-18195 for *Xenla* sCMG consensus, EMDB: EMD-18191 for *Xenla* dCMGDo consensus, EMDB: EMD-18192 for *Xenla* dCMGDo MCM ATPase.
- Atomic coordinates have been deposited in the Protein Data Bank (PDB), http://www.pdb.org, with the following accession numbers: PDB: 8Q6O for *Xenla* dCMGDo without ATPase, PDB: 8Q6P for *Xenla* dCMGDo MCM ATPase.
- All original western blots are available at: Mendeley Data: https://doi:10.17632/3d28nxpy9s.1 The full mass spectrometry data for Figure 1 is deposited in PRIDE: PXD044422.

- This paper does not report original code.
- Any additional information required to reanalyse the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

Xenopus laevis egg extract model system

African clawed toads (*Xenopus laevis*) were maintained in Birmingham animal house. The animals are housed five per tank at 20°C and fed five times a week. For preparation of egg extracts only fully grown females were used as they are the only sex able to lay eggs. For preparation of demembranated sperm, only mature males were used.

All the work with *Xenopus laevis* was approved by Animal Welfare and Ethical Review Body (AWERB) at University of Birmingham and approved by UK Home Office in the form of the Project License issued for Dr Agnieszka Gambus.

CellPress
OPEN ACCESS

All experiments performed conformed to the UK Home Office standards.

Cell lines

293FT (Thermo Fisher Scientific) cells were routinely grown in DMEM supplemented with 10% FCS, 5% L-glutamine and 5% penicillin-streptomycin. The cell line was routinely tested for mycoplasma by PCR.

Escherichia coli strains

For recombinant expression of DONSON variants, BL21 (DE3) strain was used. Cells were grown in 2 L of LB media. After cells reached OD_{600}=0.6, 1 mM IPTG was added, followed by overnight incubation at 18°C.

Cdc45-TEV-His_{10}-$FLAG_5$ was expressed in Rosetta (DE3) pLysS competent cells (Novagen, Merck Millipore). Bacteria were grown in auto inducing media (Formedium) at 37°C until OD_{600}=0.6, then moved to 18°C overnight.

Experimental source material for *in vitro* work

The source organism of recombinant proteins expressed in bacteria and analysed was *Xenopus laevis*.

METHOD DETAILS

Xenopus laevis egg extract

Xenopus laevis egg extracts were prepared as previously described.[72] Metaphase II arrested *X. laevis* egg extract is made from unfertilised female frog oocytes. In order to increase the amount of stage six (mature) oocytes, ten frogs were primed with 150 units follicle stimulating hormone Foligon (Intervet) two to seven days before eggs were required. The day before the eggs collection, frogs were injected with 400-600 units of serum gonadotropin Chorulon (Intervet), and subsequently transferred to laying tanks containing 2.5 L of 1x MMR (0.1 M NaCl, 2 mM KCl, 1 mM $MgCl_2$, 2 mM $CaCl_2$, 0.1 mM EDTA, 5 mM HEPES, pH to 7.8 with NaOH). Frogs were kept in the tanks laying eggs overnight at $\leq$ 23°C. Eggs from different frogs were collected the next morning in a 1 L glass beaker. Apoptotic and immature eggs were not collected. Eggs were rinsed with 1x MMR, and most of the buffer was subsequently removed. Eggs were then de-jellied by rinsing them in cysteine solution (2.2% cysteine, 5mM EGTA, pH to 7.6 with KOH). De-jellied eggs were rinsed again with 1x MMR, and then washed in UEB buffer (50 mM KCl, 50 mM HEPES, 5 mM $MgCl_2$, 5 mM EGTA, 2 mM DTT, pH to 7.6 with KOH) and white/swollen apoptotic eggs floating on the top were removed with a Pasteur pipette. The de-jellied eggs were packed into 14 ml round bottom polypropylene tubes (187261; Greiner) with 1 ml UEB containing 10 μg/ml protease inhibitors: aprotinin, leupeptin and pepstatin and 50 μg/ml Cytochalasin D (C8273-5MG, Sigma) and the excess buffer from the settling eggs was removed. The tubes were then spun to pack the eggs in a Beckman JS13 rotor at 800 g for 1 min, room temperature (RT). The white apoptotic swollen eggs that float to the top were removed using a Pasteur pipette followed by another centrifugation at high speed at 10,000 g for 10 min at RT. This results in separation of the eggs into a lipid layer at the top, brown cytoplasmic fraction in the middle and an insoluble egg yolk pellet at the bottom. The cytoplasmic layer was collected using a 20 G needle and a 1 ml syringe via side puncture and from this point the extract was kept on ice. Extract was supplemented with 10 μg/ml protease inhibitors, 10 μg/ml Cytochalasin D and 15% of LFB1/50 (10% sucrose, 50 mM KCl, 40 mM HEPES pH 8, 20 mM K phosphate pH 8, 2 mM $MgCl_2$, 1 mM EGTA, 2 mM DTT, 1 μg/ml of each: aprotinin, leupeptin and pepstatin). The extract was gently and well mixed using a Pasteur pipette and transferred to SW55 ultracentrifuge tubes (344058, Beckmann) which were then subjected to a final clarifying spin at 30,000 g for 20 min at 4°C. After the spinning, the yellow lipid plug from the top was removed with a clean spatula, and the pale yellow cytoplasmic fraction collected, taking extra care to not disturb the layer bellow it containing mitochondria (mitochondria will lyse after freezing-thawing the extract and will promote apoptosis rendering the extract useless). The collected extract was then supplemented with 1% v:v of glycerol, and frozen in liquid nitrogen in small beads by pipetting it dropwise and stored at -80°C.

Demembranated *Xenopus* sperm

Xenopus laevis demembranated sperm was prepared as previously described.[72] 15 male frogs were euthanised by placing them into individual chambers containing MS222 anaesthetic solution. Testes were then localised in the lower abdominal region and carefully cut out and placed in EB solution (50 mM KCl, 5 mM $MgCl_2$, 2 mM β-mercaptoethanol, 50 mM HEPES-KOH, pH 7.6) on ice. Testes were then washed with approximately 20 ml of EB solution in a 9 cm Petri dish, and cleaned by carefully removing any blood vessels and extraneous tissue using dissection forceps. Cleaned testes were then transferred to a clean 9 cm Petri dish containing 2 ml EB solution, and chopped as finely as possible with a razor blade. Chopped material was kept on ice, and once all testes were chopped, the resultant mixture was filtered through a 25 mm mesh nylon filter mounted on a small funnel. The filtered material was then transferred to a 15 ml Falcon tube and centrifuged at 2,000 g in a swinging bucket rotor for 5 minutes at 4°C. Pelleted sperm nuclei were resuspended in 0.5 ml of SuNaSp per testis (0.25 M sucrose, 75 mM NaCl, 0.5 mM spermidine, 0.15 mM spermine, 15 mM HEPES-KOH pH 7.6) and subsequently supplemented with 25 μl of lysolecithin (5 mg/ml) per testis and incubated for 5 min at RT. Demembranation of the sperm was determined by mixing 1 μl of the sperm sample with 1 μl of Hoechst 33258 (20 μg/ml) and imaged by UV microscopy. Demembranated sperm stained bright blue with Hoechst, while non demembranated sperm did not. If less than 95% of the sperm was demembranated, the mixture was re-spun, pellet resuspended in fresh SuNaSp and lysolecithin treatment repeated as before. Once demembranated, the homogenate was centrifuged at 2,000 g in a swinging bucket rotor for 5 minutes at 4°C, and

lysolecithin was quenched by resuspending the pellet in 0.5 ml per testis of SuNaSp containing 3% BSA. Demembranated sperm nuclei were washed twice with EB solution. After the second wash, the pellet was resuspended in 100 μl per testis of EB containing 30% glycerol. The sperm DNA concentration in the final mixture was estimated by counting the number of sperm using a hemocytometer (counting repeated four times). Somatic-type nuclei were counted as two sperm nuclei as they are diploid. Given that the *Xenopus* haploid genome corresponds to 3 pg of DNA, concentration of DNA in the homogenate could be estimated after counting the number of haploid nuclei. Finally, demembranated sperm nuclei were aliquoted into 80 μl fractions and stored at -80°C

DNA synthesis assay

Interphase *Xenopus laevis* egg extract was supplemented with 10 ng/μl of demembranated sperm chromatin and incubated at 23°C for indicated time. Synthesis of nascent DNA was then measured by quantification of α^{32}P-dATP (PerkinElmer) incorporation into newly synthesised DNA, as described before.[72] The extract contains endogenous dNTP pools of ~50 μM.[73] The total amount of DNA synthesized, expressed as ng DNA/μl extract, can then be calculated by multiplying percent total ^{32}P incorporated by a factor of 0.654.[73] This calculation assumes an average molecular weight of 327 Da for dNMPs and equal quantities of all four dNTPs incorporated into DNA (weight of dNMP incorporated in ng/μl = percent total ^{32}P incorporated/100 × 50 × 10^{-6} × 4 × 327 × 10^{3}).[73]

Egg extract immunodepletion

DONSON immunodepletions with rabbit antibodies were performed using Dynabeads Protein A (10008D, Life Technologies) coupled to *Xenopus* DONSON antibodies raised in rabbit and affinity purified, or nonspecific rabbit IgG (I5006, Sigma-Aldrich). The DONSON antibodies were coupled at 600 μg per 1 ml of beads. Effective immunodepletion required two rounds of 45 min incubation of egg extract with antibody coupled beads at 50% beads ratio.

Cdc45 immunodepletions were performed using Dynabeads Protein A (10008D, Life Technologies) coupled to Cdc45 antibodies raised in rabbit and affinity purified, or nonspecific rabbit IgG (I5006, Sigma). The Cdc45 antibodies were coupled at 600 μg per 1 ml of beads and incubated with 660 μl of extract. Complete immunodepletion required three rounds of 20 min incubation at room temperature with the extract.

Chromatin isolation time course

Interphase *Xenopus laevis* egg extract was supplemented with 10 ng/μl of demembranated sperm DNA and subjected to indicated treatments. The reaction was incubated at 23°C for indicated length of time and chromatin was isolated in ANIB100 buffer (50 mM HEPES pH 7.6, 100 mM KOAc, 10 mM $Mg(OAc)_2$, 2.5 mM Mg-ATP, 0.5 mM spermidine, 0.3 mM spermine, 1 μg/ml of each aprotinin, leupeptin and pepstatin, 25 mM β-glycerophosphate and 10 mM 2-chloroacetamide (Merck)) as described previously.[72] During the chromatin isolation procedure, a sample without addition of sperm DNA (no DNA) is processed in an analogous way, usually at the end of the time course, to serve as a chromatin specificity control. The bottom of the PAGE gel on which the chromatin samples were resolved is cut off and stained with Colloidal Coomassie (SimplyBlue, Life Technologies) to stain histones which provide loading controls and indications of sample contamination with egg extract (cytoplasm). Antibodies used: anti-PCNA (Sigma Aldrich, P8825), anti-MCM2 (BD Transduction labs, 610700), anti-FLAG (Sigma Aldrich, F7425), anti-Psf2 and anti-Cdc45,[56] anti-GINS,[57] sheep anti-DONSON and rabbit anti-DONSON,[42] anti-Mcm7,[58] anti-RecQ4 and anti-TopBP1,[59] anti-Treslin and anti-TopBP1.[60]

Recombinant protein purification

Xenopus laevis Cdc45 gene was cloned in a bacterial pRSFDuet-1 T7 expression plasmid (Novagen) fused to a C-terminal, TEV-cleavable-His_{10}-FLAG tag and codon optimized for expression in *E. coli* by Pellegrini's lab.[74] The vector was modified by adding further four FLAG tags via PCR to have *Xenopus laevis* Cdc45 fused with a TEV-cleavable-His_{10}-$FLAG_5$ tag.

Cdc45-TEV-His_{10}-$FLAG_5$ was expressed in Rosetta (DE3) pLysS competent cells (Novagen, Merck Millipore). Bacteria were grown in auto inducing media (Formedium) at 37°C until OD=0.6, then moved to 18°C overnight. Bacteria were pelleted at 6000 g for 10 min at 4°C, resuspended in 50 mM Tris, 100 mM NaCl, 10% sucrose and frozen. Lysis was performed by bringing the suspension to 500 mM NaCl, 15 mM imidazole, 1 mM PMSF, 0.1% Triton X-100, 1 mg/ml lysozyme and adding one protease inhibitor cocktail tablet (Roche cOmplete), then freezing and thawing in liquid nitrogen three times followed by incubation for 1h at 4°C. The lysate was spun at 20,000 g for 30 min at 4°C and the soluble fraction incubated with 30 U/ml benzonase nuclease (E1014-25KU, Sigma) for 1 h at 4°C. Finally, the lysate was spun at 20,000 g for 30 min at 4°C and the soluble part collected.

Purifications were performed using an ÄKTAprime plus system. The lysate was loaded onto a nickel-affinity column (Cytiva HisTrap HP) equilibrated in His binding buffer (50 mM Tris pH 7.5, 500 mM NaCl and 0.1 mM PMSF) and Cdc45-TEV-His_{10}-$FLAG_5$ was eluted with an imidazole gradient of 3-80% over 100 ml at 1 ml/min (His elution buffer: 50 mM Tris pH 7.5, 500 mM NaCl, 500 mM imidazole and 0.1 mM PMSF). The eluted fractions corresponding to the elution peak measured at 280 nm were collected and dialysed into Heparin binding buffer (50 mM Tris pH 7.5, 1 mM EDTA, 1 mM DTT and 0.1 mM PMSF) overnight. The protein was then loaded onto a heparin column (Cytiva HiTrap Heparin HP) equilibrated in Heparin binding buffer and eluted with a salt gradient of 5-40% over 100 ml at 1 ml/min (Hep elution buffer: 50 mM Tris pH 7.5, 2 M NaCl, 500 mM imidazole and 0.1 mM PMSF). The protein was then dialysed in LFB1/50 buffer (10% sucrose, 50 mM KCl, 40 mM HEPES pH 8.0, 20 mM K phosphate pH 8.0, 2 mM $MgCl_2$, 1 mM EGTA, 2 mM DTT, 1 μg/ml of each: aprotinin, leupeptin, and pepstatin) and stored at -80°C at final concentration of 42 mM.

pET28a-DONSON, pET28a-DONSON-WM (W234L, M463T), pET28a-DONSON-DEWS (D374A, E377A, W381A, S437A), pET28a-DONSON-3A (V473A, R476A, Y481A), pET28a-DONSONΔ24 (1-24 removed), pET28a-DONSON-S437A, pET28a-DONSON-DEW (D374A, E377A, W381A) vectors were used for protein expression in 2L of BL21 (DE3) (1 mM IPTG was added at OD_{600} = 0.6, followed by incubation overnight at 18°C). Frozen bacterial pellets were lysed in resuspension buffer (500mM NaCl, 50mM NaH2PO4, 2mM MgCl2, 10% glycerol, pH 9, 1 μg/ml of each: aprotinin, leupeptin and pepstatin) and supplemented with 1 mg/ml lysozyme, 0.05% Tween-20, and BitNuclease. After sonication, the lysate was clarified by centrifugation at 14,000 g for 30 min at 4°C, and supernatant fractions containing soluble proteins were then incubated with 2 ml of pre-washed Super Ni-NTA Affinity Resin (SUPER-NINTA100, Generon) for 2 hours rotating at 4°C. The beads were then washed three times with 30 ml of Lysis buffer, both supplemented with 15 mM imidazole. The beads were transferred to 10 ml columns (Poly-Prep Chromatography Column, Bio-Rad) and eluted with Lysis Buffer supplemented with 250 mM imidazole. Most concentrated elution fractions were dialysed against LFB1/50 buffer (10% sucrose, 50 mM KCl, 40 mM HEPES pH 8.0, 20 mM K phosphate pH 8, 2 mM $MgCl_2$, 1 mM EGTA, 2 mM DTT, 1 μg/ml of each: aprotinin, leupeptin and pepstatin) and snap frozen in liquid nitrogen.

$FLAG_5$, His_6-tagged GINS was purified as described in Tarcan et al.[57]

Large-scale immunoprecipitation from chromatin

4 ml of *Xenopus laevis* egg extract were activated into interphase by adding energy regenerator (50 mM HEPES pH 7.6, 306 mg/ml creatine phosphate, 3.6 mg/ml creatine phosphokinase), 0.25 mg/ml cycloheximide and 0.3 mM $CaCl_2$ and supplemented with 10 ng/μl of demembranated sperm DNA, 70 nM recombinant Cdc45, 40 μM aphidicolin, 5 mM caffeine and incubated at 23°C for 60 min. Chromatin was isolated in ANIB100 buffer (50 mM HEPES pH 7.6, 100 mM KOAc, 10 mM $Mg(OAc)_2$, 2.5 mM Mg-ATP, 0.5 mM spermidine, 0.3 mM spermine, 1 μg/ml of each aprotinin, leupeptin and pepstatin, 25 mM β-glycerophosphate, 0.1 mM Na_3VO_4 and 10 mM 2-chloroacetamide) as described previously.[72]

Chromatin pellets were resuspended in ANIB100 containing 20% sucrose. Protein complexes were released from chromatin by digestion with 0.4 U/μl benzonase nuclease (E1014-25KU, Sigma) and sonicated for 5 min using a sonicator with settings: 30 s on, 30 s off, low power (Bioruptor Diagenode UCD-200). Immunoprecipitation was performed using 100 μl of anti-FLAG M2 magnetic beads (Sigma-Aldrich). Before elution the sample was washed four times with 1 ml of ANIB100 20% sucrose, ANIB100 20% sucrose 0.1% Triton X-100, ANIB100 and elution buffer (25 mM HEPES pH 7.5, 100 mM KOAc, 5 mM $Mg(OAc)_2$, 1 mM ATP and 0.02% NP-40), respectively. The sample was eluted adding 250 μM 3xFLAG peptide (Stratech) to 200 μl of elution buffer.

The immunoprecipitated sample was separated via SDS-PAGE (NuPAGE™ 4 to 12%, Bis-Tris, Invitrogen) and silver stained (SilverQuest™, Invitrogen).

Mass spectrometry

Gel digests were analysed by nano LC/MS/MS with a Waters M-class HPLC system interfaced to a ThermoFisher Oribitrap Fusion Lumos. Peptides were loaded on a trapping column and eluted over a 75 μm analytical column at 350 nl/min; both columns were packed with Luna C18 resin (Phenomenex). A 30 min gradient was employed. The mass spectrometer was operated in data-dependent mode, with MS and MS/MS performed in the Orbitrap at 60,000 resolution and 15,000 resolution, respectively. Advanced Peak Detection was turned on. The instrument was run with a 3 s cycle for MS and MS/MS. Proteome Discoverer v1.4 was used for peak generation.

Mass-spectrometry data processing

Data were searched using Mascot (Matrix Science, London, UK; version 2.8.0.1) with the following parameters:

Enzyme: Trypsin Fully Specific

Database: Uniprot Xenopus (forward and reverse appended with common contaminants) released on 04/15/2014. 79,274 (including reverse and CON) entries in the database were searched.

Fixed modification: Carbamidomethyl (C)

Variable modifications: Oxidation (M), Acetyl (Protein N-term), Deamidation (NQ), GlyGly (K), Phospho (STY)

Mass values: Monoisotopic

Peptide Mass Tolerance: 10 ppm

Fragment Mass Tolerance: 0.02 Da

Max Missed Cleavages: 2

Mascot DAT files were parsed into the Scaffold (version Scaffold_5.1.0, Proteome Software Inc., Portland, OR) software for validation, filtering and to create a non-redundant list per sample. Data were filtered with 1% protein and peptide false discovery rate (FDR) and requiring at least two unique peptides per protein.

Peptide identifications were accepted if they could be established at greater than 34.0% probability to achieve an FDR less than 1.0% by the Percolator posterior error probability calculation.[75] Protein identifications were accepted if they could be established at greater than 99.0% probability to achieve an FDR less than 1.0% and contained at least two identified peptides. Protein probabilities were assigned by the Protein Prophet algorithm.[76] Proteins that contained similar peptides and could not be differentiated based on MS/MS analysis alone were grouped to satisfy the principles of parsimony. For calculation of fold enrichment for proteins with no peptides detected in control immunoprecipitation, that number was changed to one to allow for fold enrichment calculation.

GFP-Trap affinity purification from human cells

For GFP-Trap pulldown experiments with 293FT cells, confluent cells were transfected with 12 μg of plasmid DNA using Lipofectamine 2000 and harvested 48 h post-transfection. Cells were incubated in lysis buffer (150 mM NaCl, 50 mM Tris HCl pH 7.5, 2 mM $MgCl_2$, 0.5-1% NP40, 250 U/ml Benzonase (Merck) and EDTA-free protease inhibitor cocktail [Roche]) for 30 min with rotation at 4°C. Cell lysates were then pre-cleared at 65,000 g at 4°C for 30 min. For GFP-Trap, 5-7 mg of lysate was incubated with GFP-Trap agarose beads (ChromoTek) for 5 h at 4°C. The resulting GFP-Trap complexes were washed with wash buffer (100-300 mM NaCl, 50 mM Tris HCl pH 7.5, 0.5% NP40, and complete protease inhibitor cocktail [Roche]) and analysed by SDS–PAGE.

Immunoblotting was performed with antibodies Flag (Sigma-Aldrich, F7425; 1:5000), GINS (Gambus lab; 1:500), MCM2 (BD Transduction labs, 610700; 1:2000), CDC45 (Cell Signaling Technology, 11881; 1:500), PCNA (Santa Cruz Biotechnology, sc-56; 1:1000), GFP (Abcam, ab6556; 1:2000) and TOPBP1 (Fortis Life Science, A300-111A; 1:500).

Mass photometry

Solution-phase mass determination of native DONSON and mutants thereof were acquired using the TwoMP (Refeyn) mass photometer. Experimental data were obtained in the form of mass photometry videos recorded for one minute using the AcquireMP v2.5 software (Refeyn) on precleaned, high sensitivity microscope slides. Prior to analysis of target proteins, a mass calibration was determined using urease, BSA and aldolase as standards spanning a mass range of 66 to 544 kDa. Each standard was used at a final concentration of 2 to 5 nM by diluting ten-fold into 18 μl of Dulbecco PBS buffer. A linear mass calibration was obtained after fitting of Gaussian functions to the ratiometric contrast values of each standard using the DiscoverMP v2.5 software (Refeyn). In a similar fashion, DONSON proteins, were diluted ten-fold into Dulbecco PBS buffer to achieve a final concentration of 3 to 5 nM. Mass photometry videos were recorded for one minute and analysed in the DiscoverMP v2.5 software.

Sample preparation and data collection for negative stain electron microscopy

Negative stain EM (NS-EM) samples were prepared on 300-mesh copper grids with carbon film (EM Resolution C300Cu). After glow discharging the grids for 60 s at 25 mA using GloQube Plus unit (Quorum), 4 μl of the elution from the large-scale immunoprecipitation of chromatin were applied to the grids and incubated for 1 min. Excess sample was blotted away and the procedure of applying 4 μl sample followed by blotting was repeated two more times. After third application of the sample and the final blotting, grids were stained by successive stirring in three 35 μl drops of 2% uranyl acetate for 20 s each. Excess stain was removed by blotting and the grids were left to dry for a few minutes before being stored for imaging.

All NS-EM data were acquired on a FEI Tecnai LaB6 G2 Spirit transmission electron microscope, operating at 120 kV. Micrographs were collected with a Gatan Rio 4k x 4k CMOS camera at nominal magnification of 29,000 (physical pixel size 3.1 Å/pix) and within a -0.5 to -2.0 μm defocus range.

Negative-stain EM image processing

A subset of particles was picked using a crYOLO 1.4.0[61] model trained for previous CMG negative stain dataset and the picks were manually curated in crYOLO. This subset was then used for training a new crYOLO model for particle picking. Subsequent processing steps were carried out using Relion-3.0.4[63] and cryoSPARC v4.2.1.[64] In Relion-3.0.4 the CTF for each micrograph was estimated using Gctf[65] particles were extracted from micrographs and 2D classification was performed. *Ab initio* reconstructions and refinements were performed in cryoSPARC v4.2.1.[64]

Sample preparation and data collection for cryo-EM

Graphene oxide dispersion was prepared by diluting 10 μl of the stock (2 mg/ml, Sigma Aldrich 763705) in 80 μl of water, resuspending the mixture by pipetting and spinning down remaining aggregates at 650 g for 1.5 min in a bench centrifuge. Clear top layer was then carefully transferred to a new tube. UltrAuFoil R1.2/1.3 grids (Quantifoil Micro Tools) were glow discharged for 5 min at 45 mA using GloQube Plus unit (Quorum), 4 μl of freshly prepared graphene oxide dispersion was applied to each grid and incubated for 2-3 min. After blotting graphene oxide, three droplets of water, 20 μl each, were sequentially picked up with the grid and blotted away, the first two with the front side of the grid (where graphene oxide had been applied) and the last one with the back side of the grid. Grids were then dried by incubating at room temperature before the sample was applied.

Grids with freshly applied graphene oxide were prepared for cryo-EM in a Vitrobot Mark IV (FEI Thermo Scientific) by triple application of the elution sample from the large-scale immunoprecipitation of chromatin and subsequent freezing in liquid ethane. 4 μl of the sample was applied on grids, incubated for 1 min at 22°C at 100% humidity, quickly blotted away (0.5 s blotting time), followed by two more rounds of sample application and blotting. After the third sample application and incubation, excess sample was blotted away for 4.5 s and grids were plunged into liquid ethane.

Cryo-EM data collection was conducted on an in-house Thermo Scientific Titan Krios transmission electron microscope operating at 300 kV, equipped with a Thermo Scientific Falcon 4i direct electron detector camera and Selectris energy filter. EPU software (Thermo Scientific) was used for automatic acquisition of movies in counting mode with physical pixel size of 0.95 Å and a total electron dose of 33 e/$Å^2$ during 5.44 s exposure time. 1674 EER frames were initially collected per movie and they were subsequently processed in 27 fractions, with 1.22 e/ $Å^2$ per fraction. A total of 82,135 movies were collected in a single session using an energy

CellPress
OPEN ACCESS

filter slit width of 10 eV and a defocus range set at -1.3 to -3.4 μm. First 7,002 movies were collected with a defocus range set at -1.3 to -2.8 μm.

Cryo-EM image processing

RELION-4.0.0[62] and cryoSPARC v4.2.1 were used for image processing (Figure S2). The movies were corrected for drift, dose-weighted and the contrast transfer function (CTF) parameters were estimated using CtfFind-4.1[66] in RELION. Particles were first manually picked in cryoSPARC on a small subset of 28 micrographs across the defocus range and these 594 particles were used for creating an initial Topaz-0.2.5[67] model for picking, through cryoSPARC. The model was then improved through several iterative cycles consisting of particle extraction (binned by two, with box size 256 px at 1.9 Å/px), cleaning up with 2D classifications in cryoSPARC or RELION and training of a new Topaz model with cleaned-up particles used as an input. To obtain better 2D-average-based separation of good particles from noise, contamination, particle aggregations or particles close to the edge of graphene oxide layers, the subset of micrographs used for iterative cycles of particle picking, extraction and 2D classifications was increased to 1,060, at first instance, and then to 28,744 micrographs. Topaz models were trained on a smaller subset of 60 or 108 micrographs with 1,000–3,000 input particles, and input particles were manually curated before training to remove obviously wrong picks. The final Topaz model (Topaz model #1) was trained on a subset of 108 micrographs across the defocus range, with 2,488 particles. Another Topaz model (Topaz model #2) was trained for picking particles in low-defocus micrographs and the training was conducted on a subset of 85 micrographs with 1822 particles and a defocus range between -1.3 and -1.9 μm. The full set of collected micrographs was curated in RELION based on CTF parameters (rlnCtfMaxResolution and rlnCtfFigureOfMerit) to remove poor-quality micrographs and micrographs without a graphene oxide coverage. This yielded a total of 67,761 micrographs that were subsequently divided in two subsets by defocus values: a subset of 56,876 micrographs with defocus $\leq$ -1.9 μm (high-defocus subset), and a subset of 10,885 micrographs with defocus > -1.9 μm (low-defocus subset). Using the Topaz model #1, 3,560,024 particles were picked from the high-defocus subset of micrographs. Similarly, using the Topaz model #2, 580,279 particles were picked from the low-defocus subset, making the total number of particles initially picked from the whole dataset 4,140,303. 2D-classification-based clean-up of the particles from the two subsets of micrographs was performed separately.

Using RELION, the high-defocus subset of micrographs was divided into seven groups and picked particles were extracted with a binning factor of three and a box size of 180 px (2.85 Å/px). The respective seven groups of particles were then subjected to successive rounds of 2D classifications in parallel to remove contaminants, aggregated and smaller particles as well as particles close to the edge of graphene oxide layers. The remaining particles were joined and re-extracted with a binning factor of two and a box size of 294 pix (1.9 Å/px). The low-defocus subset of micrographs was divided into two groups and the proceeding steps of particle extraction, several rounds of 2D classification and final joining and re-extraction of remaining particles were conducted in the same way as described for the high-defocus subset of micrographs. In the end, re-extracted particles coming from the high- and the low-defocus subset were joined, totalling 954,780 particles. These particles were submitted for 3D multi-reference classification with image alignment in RELION, using regularisation parameter T value of four, C1 symmetry and four low-pass filtered reference maps: (1) a double CMG-DONSON (dCMGDo) map generated by *ab initio* reconstruction and homogeneous refinement in cryoSPARC of 34,344 particles binned by two, coming from the highest-resolution dCMGDo 2D classes obtained during preliminary processing of the first 33,441 micrographs of this dataset; (2) a single CMG map generated by *ab initio* reconstruction, followed by heterogeneous and homogeneous refinements (C1 symmetry) in cryoSPARC, from a small, trial cryo-EM dataset (2301 micrographs) of the same sample collected on our in-house Thermo Scientific Talos Arctica transmission electron microscope; the remaining two reference maps, (3) and (4), which were poorly defined and did not look like CMG, were among those generated by *ab initio* reconstruction and refinements in cryoSPARC of particles binned by two obtained during preliminary processing of the first 33,441 micrographs. In this way, particles of the dCMGDo complex (219,456 particles; 23%) and particles of the single CMG (361,156 particles; 37%) were separated from the particles with poor features that did not contain complete CMG.

Particles from the dCMGDo 3D class (219,456 particles) were unbinned and re-extracted in RELION with a box size of 540 pix (0.95 Å/px), followed by *ab initio* reconstruction, heterogeneous and non-uniform refinement with applied C2 symmetry) in cryoSPARC. The refined map was low-pass filtered and used as a reference map for a 3D classification in RELION of the unbinned dCMGDo set containing 219,456 particles, performed with image alignment, C1 symmetry, regularisation parameter T value of four and with two classes. A 3D class with 128,581 particles (56%) was selected, auto-refined using C2 symmetry and post-processed in RELION to 4.2 Å resolution. These particles were then subjected to several rounds of CTF refinement and auto-refinement (C2 symmetry), as well as to two rounds of Bayesian polishing. CTF-refined and polished dCMGDo particles were then further processed in cryoSPARC, by performing *ab initio* reconstruction and heterogeneous refinement with two classes and C1 symmetry, which yielded a new set of 100,667 particles. These particles were then subjected to non-uniform refinement, with applied C2 symmetry alignment, per-particle defocus optimisation and per-group CTF parameters optimisation, leading to a 3.2 Å resolution map. Further local refinement with a mask encompassing the whole dCMGDo complex, with applied C2 symmetry and pose/shift gaussian prior during alignment, resulted in a consensus map with improved resolution of 3.1 Å, sharpened with *B* factor of -93 Å^2.

The set of 128,581 CTF-refined and polished dCMGDo particles was auto-refined (C2 symmetry) and post-processed in RELION yielding a 4.07 Å resolution map. To better resolve MCM ATPase domains that featured conformational heterogeneity, a C2 symmetry expansion procedure was applied, followed by particle subtraction in RELION, where a mask encompassing ATPase domains of one MCM hexamer was used to subtract all the remaining signal. The mask was created from the auto-refined dCMGDo map, after

segmentation and subtraction of all segmented regions except one MCM-ATPase in UCSF Chimera.[70] Subtracted particles were then subjected to focused 3D classification without image alignment, using the same mask applied in particle subtraction, with ten classes, C1 symmetry and regularisation parameter T value of 16. Two 3D classes with 144,989 particles were selected and the particles were submitter to local refinement in cryoSPARC with applied C1 symmetry, pose/shift gaussian prior during alignment and dynamic masking. This yielded the MCM-ATPases map at 3.5 Å resolution, sharpened with *B* factor of -113 Å^2.

Particles from the single CMG subset (361,156) were re-extracted with a box size of 240 pix (1.9 Å/px) and submitted for two rounds of 3D classification in RELION, with four and two classes respectively, regularisation parameter T value of four, C1 symmetry, applied image alignment and with the low-pass filtered single CMG map from the screening Talos Arctica dataset mentioned earlier used as a reference. In between two rounds of 3D classification, the particles were unbinned (box size of 480 px, with 0.95 Å/px). 3D classes where single CMG looked isotropic were selected (137,591 particles) and the particles were submitted to cryoSPARC for *ab initio* reconstruction and non-uniform refinement with per-particle defocus optimisation, per-group CTF parameters optimisation and with C1 symmetry. Finally, local refinement with a mask encompassing the whole complex, applied pose/shift gaussian prior during alignment and with C1 symmetry was performed. This resulted in the 3.0 Å resolution map of a single CMG, sharpened with B-factor of -91 Å^2.

Model building and refinement

We used AlphaFold-predicted[48] models to initially place all *Xenopus laevis* CMG subunits into the sharpened consensus map of the dimeric CMG complex: Mcm2 (AF-P55861-F1), Mcm3 (AF-P49739-F1), Mcm4 (AF-P30664-F1), Mcm5 (AF-P55862-F1), Mcm6 (AF-Q5FWY4-F1), Mcm7 (AF-Q7ZXB1-F1), Pfs1 (AF-Q7ZT47-F1), Psf2 (AF-Q7ZT46-F1), Pfs3 (AF-Q7ZT01-F1), Cdc45 (AF-Q9YHZ6-F1) and Sld5 (AF-Q7ZT48-F1). AlphaFold Multimer was used to predict the interaction between DONSON and Mcm3 or Sld5. We first rigid-body docked protein chains comprising one CMGDo monomer into the corresponding segments of the EM density using UCSF Chimera.[70] Since inter-domain movements within MCM subunits were apparent, they were fitted as separate rigid bodies (A domain, B/C domain, ATPase domain and winged-helix domain (WHD)). After rigid-body docking, all subunits were fitted into the density using the jiggle-fit tool and flexible fitting in real space (Geman-McClure Alpha set to 0.1) with all-chain self-restraints applied (4.3 Å). This was performed in *Coot*.[68] For the N-terminal domains of MCM subunits, Pfs1, Psf2, Pfs3, Cdc45 and Sld5 sharpened consensus map was used. MCM ATPase and WHD domains were fitted to the sharpened ATPase map. Any parts of the predicted model that comprised unstructured regions or corresponded to poorly resolved regions in the map were trimmed off at this point. In particular, N-terminal hairpins (NtHp) of Mcm5, Mcm6 and Mcm7 as well as helix-2 insert (h2i) and pre-sensor 1 β-hairpin (PS1BH) elements in Mcm4 and Mcm7 could not be modeled. After flexible fitting, we used *Coot* to optimize the fit of well-resolved chain fragments to their corresponding maps (map weight set to 10), applying torsion and Ramachandran restraints to maintain backbone and sidechain geometry. Due to the poor density of the Mcm4 ATPase and Mcm6 ZnF regions, only flexible fitting and no manual adjustments were performed for these regions. Residual density for Mcm5 WHD was visible, but it was not included in the final model due to poor map quality in that region. The resolution of the ATPase map was sufficient to include nucleotides in the ATPase sites, although the exact nucleotide state could not have been established. We therefore assigned ATP to all occupied sites (all except for unoccupied 7/4 site), but we cannot rule out the possibility that they represent ADP or mixed ATP/ADP states. Prior to structure refinement, we fitted a second CMGDo monomer (without MCM ATPase) into the symmetry-equivalent part of the consensus reconstruction to create the complete symmetry-expanded dCMGDo complex. We then refined the structures in real space with *PHENIX*[69] using default settings and rotamer fit set to "outliers". The quality of the resulting atomic models was evaluated with MolProbity[77] and Comprehensive validation (cryo-EM) tools, both integrated into *PHENIX* (summarized in Table 1).

Visualization of models and maps

UCSF Chimera[70] and UCSF ChimeraX[71] were used for visualizing all maps and models in this study. ChimeraX was also used for making figures.

QUANTIFICATION AND STATISTICAL ANALYSIS

DNA synthesis analysis

For quantification of replication efficiency between different experimental repeats in different extracts, the quantity of DNA replicated at the end of the reaction in IgG-depleted extract was set as 100% and the remaining values normalised to this. The mean with SEM was plotted using GraphPad Prism 9.2 or 10. Statistical significance of differences between time-courses replication curves were estimated using 2way ANOVA comparing each time to IgG-depleted replication curve. For titration of recombinant DONSON, paired, nonparametric, two-tailed t-test was performed using GraphPad Prism 10.

EM data analysis

Quantification, statistical analysis and validation pertaining to processing of negative stain and cryo-EM images are implemented in the software described in the image processing section of the methods details. Global resolution estimation of refined cryo-EM maps are based on the 0.143 cutoffs of the Fourier Shell Correlation between two half maps refined independently.

Molecular Cell, Volume *83*

Supplemental information

The structural mechanism of dimeric DONSON in replicative helicase activation

Milos A. Cvetkovic, Paolo Passaretti, Agata Butryn, Alicja Reynolds-Winczura, Georgia Kingsley, Aggeliki Skagia, Cyntia Fernandez-Cuesta, Divyasree Poovathumkadavil, Roger George, Anoop S. Chauhan, Satpal S. Jhujh, Grant S. Stewart, Agnieszka Gambus, and Alessandro Costa

Supplementary information

The structural mechanism of dimeric DONSON in replicative helicase activation

Milos A. Cvetkovic, Paolo Passaretti, Agata Butryn, Alicja Reynolds-Winczura, Georgia Kingsley, Aggeliki Skagia, Cyntia Fernandez-Cuesta, Divyasree Poovathumkadavil, Roger George, Anoop S. Chauhan, Satpal S. Jhujh, Grant S. Stewart, Agnieszka Gambus, Alessandro Costa

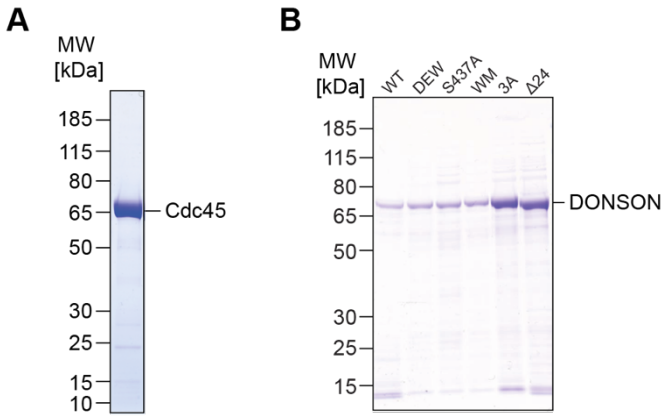
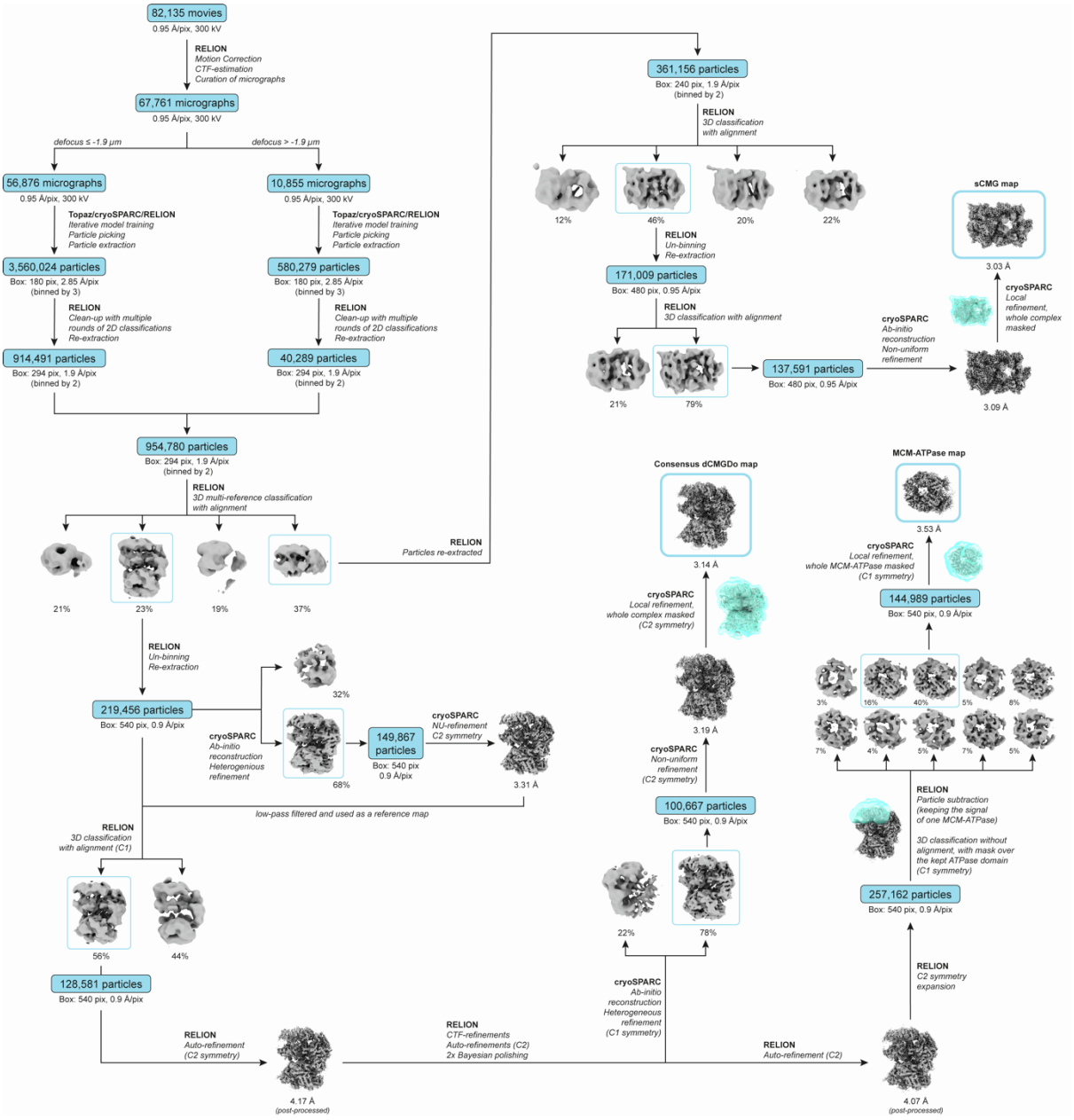
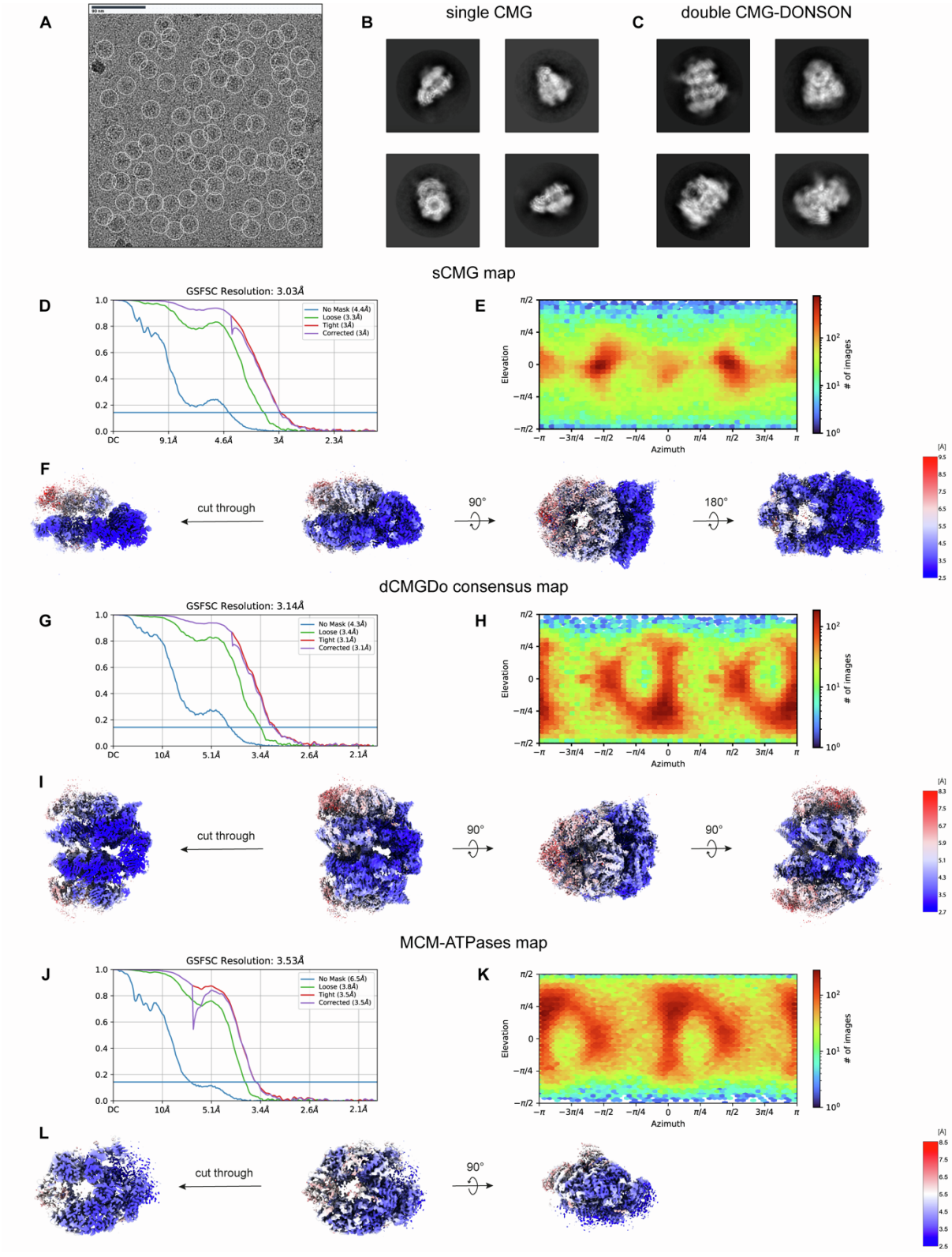


Figure S1. Recombinant proteins created in this study, related to Figure 1. Purified proteins resolved on SDS-PAGE gel and stained with Coomassie: **(A)** *Xenla* Cdc45-TEV-His_{10}-$FLAG_5$. **(B)** *Xenla* DONSON variants: wild-type (WT); Asp374Ala, Glu377Ala and Trp381Ala (DEW); Ser437Ala; Trp234Leu, Met463Thr (WM); Val473Ala, Arg476Ala, Tyr481Ala (3A); deletion of residues 1 to 24 (Δ24).

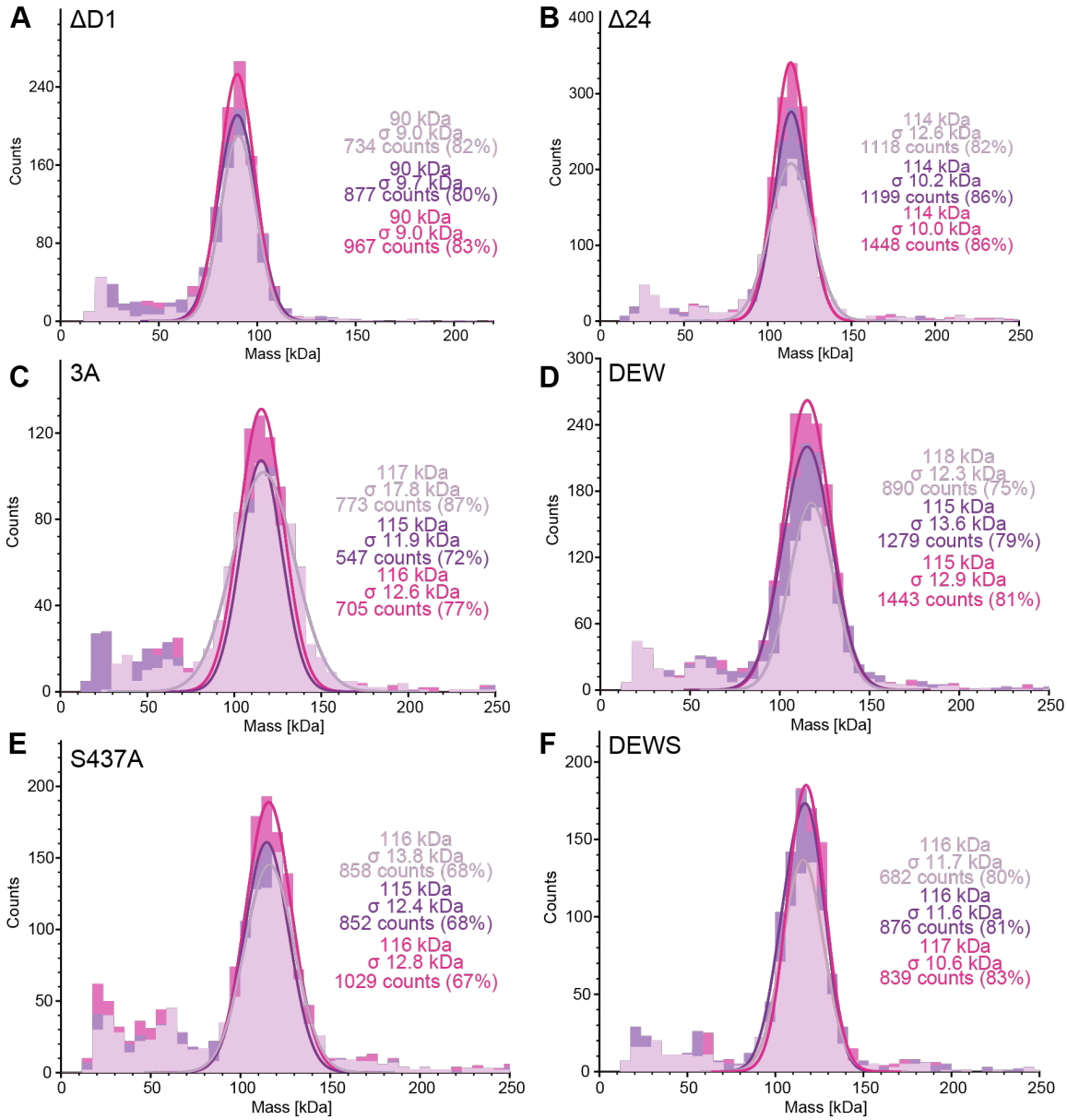


Supplementary Figure 2. Cryo-EM image processing pipeline, related to Figure 2.



Supplementary Figure 3. Cryo-EM analysis of single CMG and dCMGDo, related to Figure 2. **(A)** A representative micrograph. **(B)** 2D averages of single CMG. **(C)** 2D averages of dCMGDo. **(D)** Fourier shell correlation plot for single CMG. **(E)** Angular distribution for single CMG. **(f)** Cryo-EM map coloured according to local resolution for single CMG. **(G)** Fourier shell correlation plot for dCMGDo consensus map. **(H)** Angular distribution for dCMGDo consensus map. **(I)** Cryo-EM map coloured according to local resolution for

dCMGDo consensus map. **(J)** Fourier shell correlation plot for MCM ATPase map isolated from the dCMGDo. **(K)** Angular distribution for MCM ATPase map. **(L)** Cryo-EM map coloured according to local resolution for MCM ATPase map.



Supplementary Figure 4. Mass photometry analysis of various *Xenopus laevis* DONSON mutants, related to Figures 3 and 4. All variants shown in this figure maintain a dimeric character despite the amino acid changes. **(A)** Truncation of residues 1 to 155 (D1). **(B)** Truncation of residues 1 to 24 (Δ24). **(C)** Val473Ala, Arg476Ala, Tyr481Ala (3A). **(D)** Asp374Ala, Glu377Ala, Trp381Ala (DEW). **(E)** Ser437Ala. **(F)** Asp374Ala, Glu377Ala, Trp381Ala, Ser437Ala (DEWS).

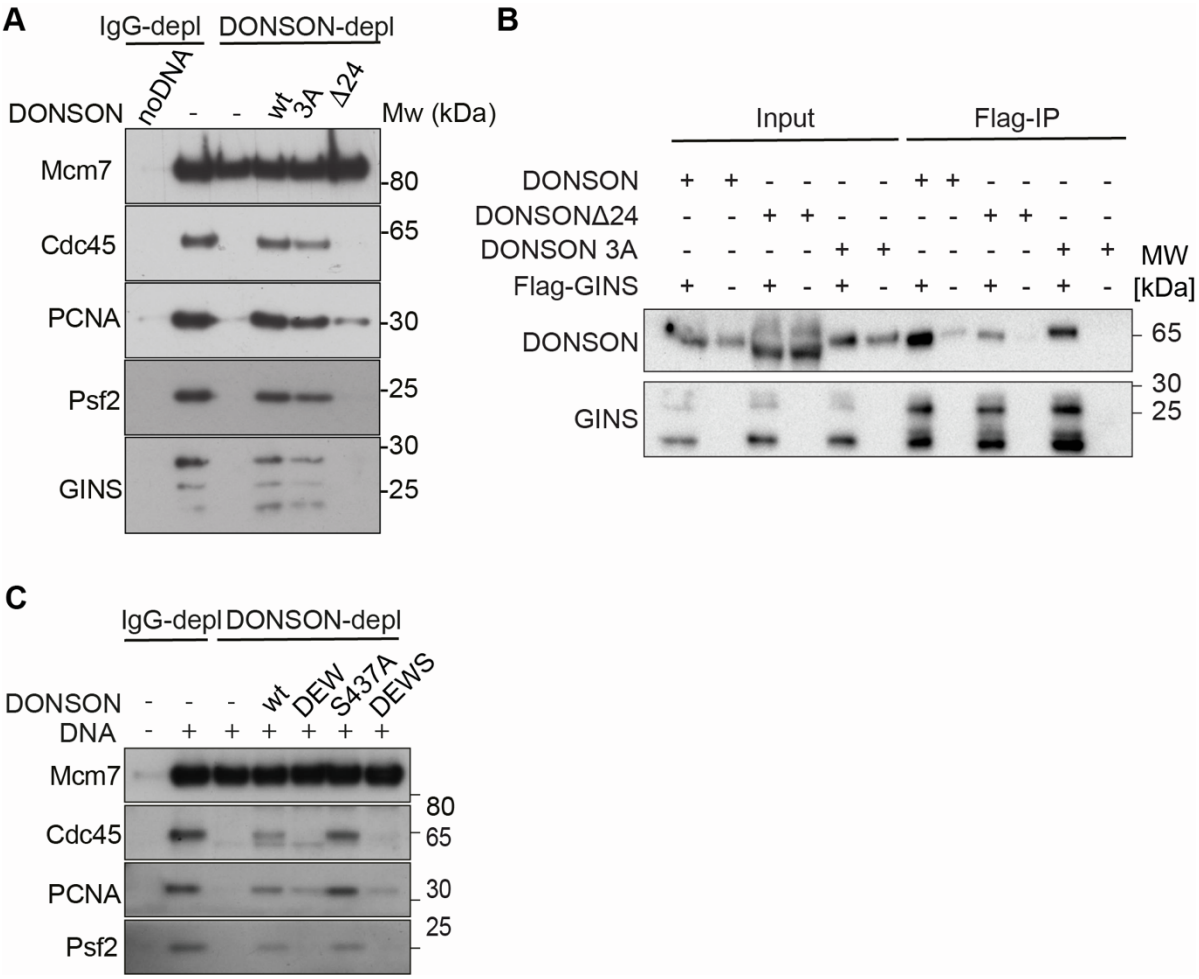
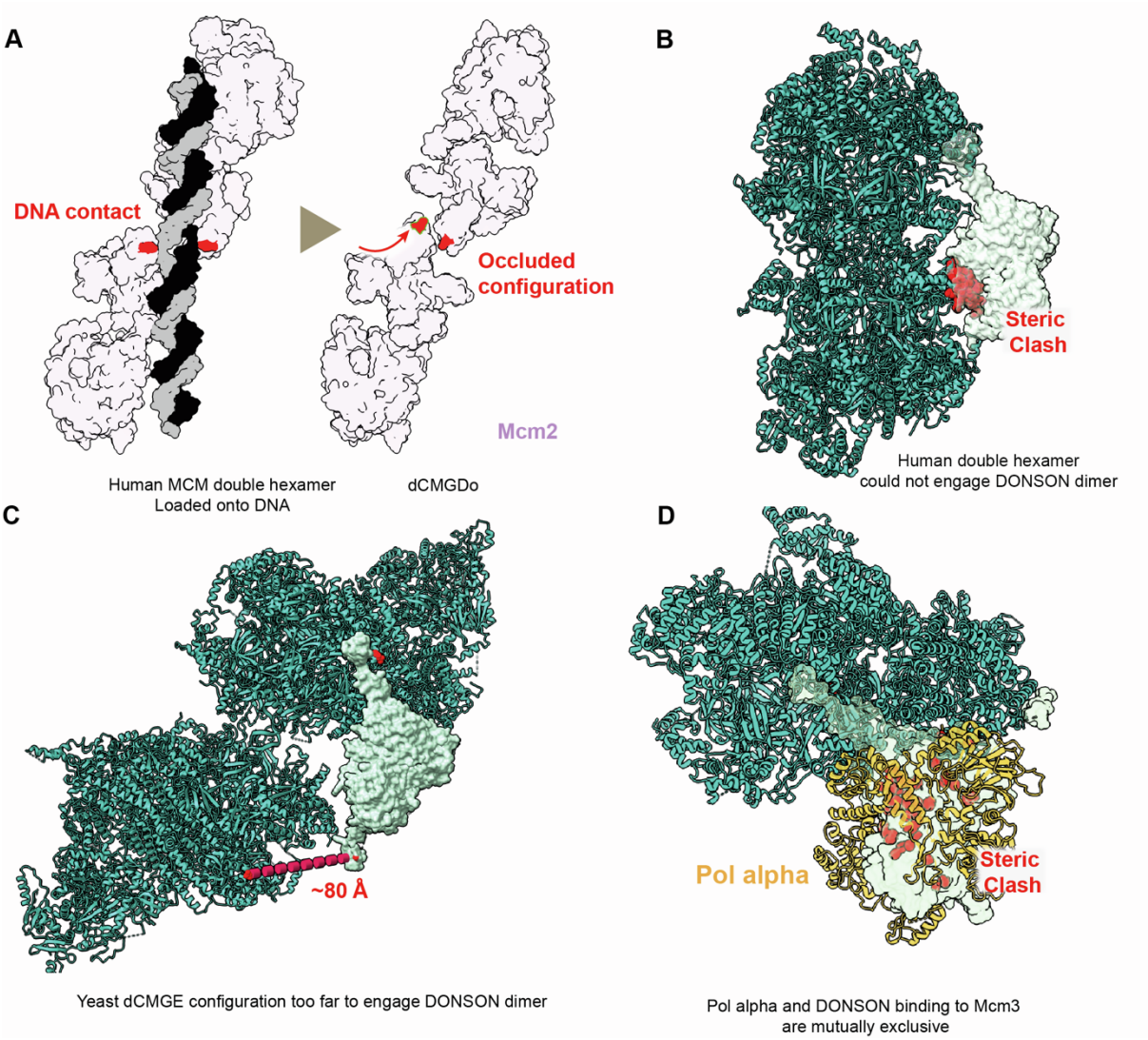
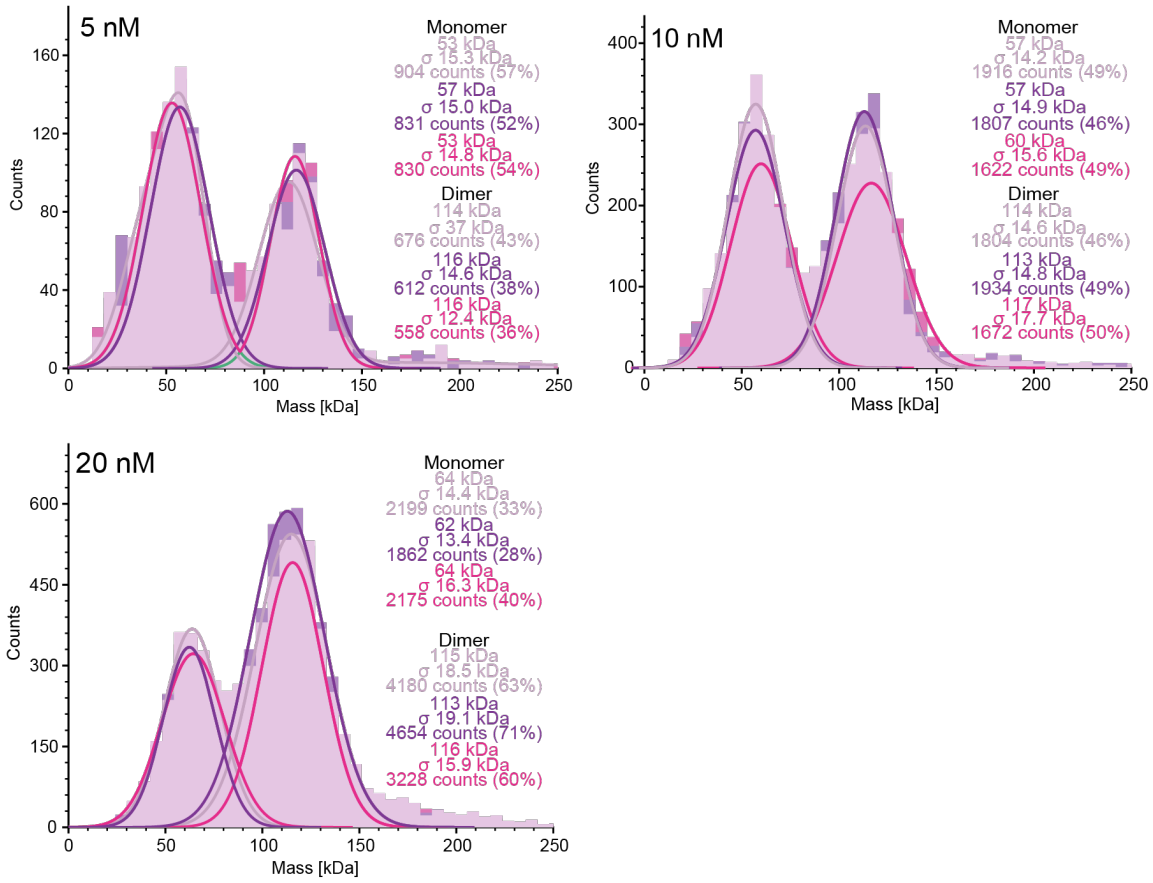


Figure S5. Functional analysis of DONSON mutants, related to Figures 3 and 4. (A) DNA replication reaction was set up in IgG- or DONSON-depleted extract optionally supplemented with 100 nM recombinant wild type or indicated DONSON mutants. Chromatin was isolated at 60 min of replication reaction. Chromatin bound factors were resolved on SDS-PAGE and immunoblotted with indicated antibodies. "No DNA" control served as chromatin specificity control. **(B)** Recombinant purified DONSON WT, Δ24 or 3A mutants were mixed and incubated with *Xenopus laevis* GINS complex expressed in *E. coli* and purified via His_6-FLAG-tagged Sld5. M2-FLAG beads were used to immunprecipitate GINS. Co-immunoprecipitating proteins were analysed by immunoblotting with indicated antibodies. **(C)** DNA replication reaction was set up in IgG- or DONSON-depleted extract optionally supplemented with 100 nM recombinant wild type or indicated DONSON mutants. Chromatin was isolated at 60 min of replication reaction. Chromatin bound factors were resolved on SDS-PAGE and immunoblotted with indicated antibodies. "No DNA" control served as chromatin specificity control.



Supplementary Figure 6. Comparison between the dCMGDo structure and previously published structures, related to Figure 5. CMG is colored in teal. DONSON is light green. **(A)** dCMGDo shuts the duplex DNA harbouring MCM channel at the N-terminal dimerisation interface. The side of the Mcm2 ZnF domain that engages duplex DNA in the human double hexamer (PDB entry TW1Y) is backed against the cognate Mcm2 ZnF element in dCMGDo. **(B)** Mcm3 engagement by DONSON, as observed in dCMGDo, is incompatible with double hexamer formation as observed with the DNA loaded human (PDB entry TW1Y), or yeast, MCM. **(C)** The double CMG configuration observed in the yeast dCMGE (7Z13) is too dilated to allow concomitant Mcm3 engagement of both protomers in the DONSON dimer. **(D)** Pol alpha (PDB entry 8B9D) and DONSON compete for the same binding site on Mcm3.



Supplementary Figure 7. Mass photometry analysis of DONSON Trp234Leu, Met463Thr (WM) mutant, related to Figure 6. *Xenopus laevis* DONSON amino acid changes homologous to patients' muations in humans destabilise the homodimerization interface. A dilution experiment highlights the tendency of the WM mutant to monomerise as concentrations drop from 20 to 5 nM.

	Protein	Molecular Weight (kDa)	Spectral Counting (SpC)
CMG	Mcm2	100	134
	Mcm3	90	127
	Mcm4	97	78
	Mcm5	82	101
	Mcm6	93	138
	Mcm7	82	72
	Cdc45	66	181
	Psf1	23	11
	Psf3	24	8
	Psf2	21	2
	Sld5	26	4
Replisome	Polα (POLA1)	165	29
	Polδ (POLD1)	125	41
	Polδ (POLD2)	49*	3*
	Polδ (POLD3)	50	11
	Polε (POLE1)	265	107
	Polε (POLE2)	60	10
	Rts (RecQ4)	169	15
	AND-1	125	69
	PCNA	29	9
	Claspin	146	8
	Timeless	149	18
	Spt16	118	37
	SSRP1	79	18
Initiation factors	Mcm10	95	28
	MTBP	96	4
	Orc1	100	4
	Orc3	81	2
	Orc5	52	4
	Rif1	257	70
	Treslin	221	33
	TopBP1	169	6
	DONSON	64	5

Supplementary Table 1. Mass spectrometry analysis of a Cdc45-containing replication intermediate, related to Figure 1. Proteins co-immunoprecipitating with Cdc45-TEV-His_{10}-$FLAG_5$ isolated from chromatin replicating in *Xenopus laevis* egg extract supplemented with aphidicolin and caffeine (as per Figure 1D) were analysed by mass spectrometry. Selected

replication factors are presented with molecular weight and total spectral count. Full data available on PRIDE. (*Xenopus tropicalis proteins).