Investigating the DNA replication machinery composition and function using proteomic and chemical-biology approaches

Christopher Weekes

Imperial College London
Faculty of Medicine
Medical Research Council London Institute of Medical Sciences

A thesis submitted for the degree of Doctor of Philosophy (PhD)

January 2022

Declaration

The following work and research findings enclosed within this thesis are my own, unless explicitly stated in the text. Where assistance has been received or the work of others reproduced, it is clearly noted in the text. Information and research carried out by others that is utilised within this thesis is appropriately cited and a full bibliography is included as part of this document.

Copyright Statement

The copyright of this thesis rests with the author. Unless otherwise indicated, its contents are licensed under a Creative Commons Attribution-Non Commercial 4.0 International Licence (CC BY-NC). Under this licence, you may copy and redistribute the material in any medium or format. You may also create and distribute modified versions of the work. This is on the condition that: you credit the author and do not use it, or any derivative works, for a commercial purpose. When reusing or sharing this work, ensure you make the licence terms clear to others by naming the licence and linking to the licence text. Where a work has been adapted, you should indicate that the work has been changed and describe those changes. Please seek permission from the copyright holder for uses of this work that are not included in this licence or permitted under UK Copyright Law.

Abstract

The process of DNA replication is essential for cellular life, but must be highly regulated to ensure that damage to the genome or irregularity in copy number does not ensue. In particular, the mechanism behind the loading and activation of the eukaryotic replicative helicase, MCM2-7, onto DNA has been investigated extensively. However, despite progress afforded by the advent of cryo-EM to reveal the structures of multi-subunit protein complexes, the molecular mechanism underlying the activation of the replicative helicase has yet to be fully elucidated. In particular, the precise temporal order of initiation factor recruitment and release, and how they affect complex progression, is not known. Additionally, many *in vivo* specific factors such as chromatin factors, accessory helicases and regulatory pathways are not well understood.

This research set out to establish a more complete description of the interaction partners of the MCM2-7 helicase *in vivo* prior to, during, and after replicative helicase activation, by applying a combination of chemical-biology and proteomic techniques. Small molecule-inducible protein linkages were introduced into the MCM2-7 complex, to arrest the complex at different stages of helicase activation. This enabled investigation of transient conformational states *in vivo*.

A versatile proteomic method utilising *in vivo* cross-linking and affinity enrichment of MCM2-7 helicase interactomes followed by LC-MS/MS analysis was established, termed ChIP-MS. This technique enabled characterisation of the composition of replication complexes throughout the cell cycle, as well as in the context of the arrest states generated by the linkage constructs. Stabilisation of the *in vivo* complexes by cross-linking allowed the detection and compositional analysis of distinct sub-populations of replication complexes over multiple cell cycle stages and identified several factors and pathways, such as chromatin remodellers and accessory helicases, that may have a previously understudied or novel role in the replication process.

Acknowledgements

First, I would like to thank my principal supervisor, Professor Christian Speck, for his unfaltering support over the course of this PhD. His counsel and advice were always helpful, and his suggestions were similarly reliable. His ability to offer insight into techniques and areas less well known to him was incredibly inspiring and useful. Most importantly, Christian consistently strived to make the lab not only a productive place, but a welcome space where lab members' opinions and thoughts were always heard and valued. The lab environment being so good is testament to the effort and attention Christian puts in to help anyone with anything at any time. His enthusiasm for a field in which he has spent many years is inspirational and his dedication to adopting new techniques and furthering new forays into developing areas will remind me to remain both perseverant in what I do, but flexible enough to succeed when the nature of the challenge changes.

I would be remiss in not extending a special thanks to my supervisor Dr Luitpold "Max" Reuter. Max was my first point of call with any questions or issues that I came across, and his experience with the finer points of yeast culture and other techniques that became crucial to this project were invaluable. While Max's knowledge of the field was of great help, I have to credit him with far more than that. Max was a great mentor, colleague and friend, and made my time in the lab the highly enjoyable experience that it was. He is a fantastic, hard-working (perhaps too much so) scientist, and I wish him the very best for his future career, where I'm sure he will meet with the success he undoubtedly deserves.

A lab is no lab without the people that fill it, and the Speck lab has seen a great many fantastic scientists and brilliant people in my time there. The lab was incredibly welcoming to me and in turn to every new person that joined. The atmosphere was almost always jovial, even facing some of the most uncertain times that international science has ever faced. It is the best lab environment I have had the pleasure of experiencing, and so for the friendship, laughter and indomitable comradery, I thank; Indiana, Sarah Schneider, Yasunori, Al, Vera, Sarah Faull, Audrey, Jess, Shenaz, Lepakshi, Marina, Ric, Matt, Josh,Kyle, Owen, Cecilia, Noelia and Marta.

For the experience, support and help with my adventures with Mass Spectrometry, I thank the members of the Mass Spectrometry and Proteomics facility at the MRC-LMS. I knew very little other than the theory of Mass Spec. before I started working on this project. It is due to technical discussions and explanations from the team there that I know what I do now, and the assistance with the sample preparation and downstream analysis was crucial to the project. Special thanks to Dr. Holger Kramer, Dr. Alex Montoya Dr. Iliyana Khaneva with their patience and time.

I thank my family and friends for their unwavering support throughout this experience and all that preceded it. Exceptional thanks go to my parents, because your care for my happiness and wellbeing as well as your counsel and advice have led me to the place that I am today, and I can never thank you enough for that. Robby, my younger brother, whose infectious enthusiasm and relentless chasing at my heels has spurred me on, not least in the most literal sense, reminding me to maintain my fitness and physical wellbeing. And finally, to Gemma, whose patience and support should have long ago been exhausted; thank you for standing by my side as I experienced every high and every low of this PhD, it means so very much to me.

Abbreviations

1N	Haploid genome copy number
2N	Diploid genome copy number
4N	Tetraploid genome copy number
AAA+	ATPases Associated with diverse cellular Activities
Abf	ARS-binding factor 1
ACS	ARS consensus sequence
APC/C	Anaphase promoting complex/cyclosome
ARS	Autonomously replicating sequence
Asfl	Anti-silencing factor 1
AT-Hook	Adenine-Thymine hook
ATP	Adenosine triphosphate
BAH	Bromo-adjacent homology
BRCT	BRCA1 C-terminus
BrdU	Bromodeoxyuridine
BSA	Bovine serum albumin
CAC	Chromatin assembly complex
CAF-1	Chromatin assembly factor 1
CDC	Cell division cycle
CDK	Cyclin dependent kinase
CDT	Chromatin licensing and DNA replication factor 1
CHD1	Chromodomain helicase domain DNA binding protein 1
ChIP	Chromatin immunoprecipitation
ChIP-MS	Chromatin immunoprecipitation-mass spectrometry
ChIP-seq	ChIP coupled to genome wide sequencing
CIP box	Ctf4-interacting peptide
Clb(1-6)	CycLin B homologous protein (1-6)
Cln(1-3)	CycLin homologous protein (1-3)
CMG	Cdc45-MCM2-7-GINS
CpG island	Region of C-G dinucleotide repeats
CRAPome	Contaminant repository for affinity purification
cryo-EM	cryo-Electron microscopy
Csm3	Chromosome segregation in meiosis 3
CTE	C-terminal extension
Ctf4	Chromosome transmission fidelity 4
DAPI	4’, 6-diamidino-2-phenylindole
Dbf4	DumbBell former 4
DDC	DNA damage checkpoint
DDH2O	Double-distilled H20
DDK	Dbf4-dependent protein kinase

DH	Double hexameter
Dia2	Digs into agar 2
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
dNTP	Deoxynucleoside triphosphate
Dpb(X)	DNA polymerase B (X)
DRC	DNA replication checkpoint
DSB	Double-strand break
dsDNA	Double-stranded DNA
DTT	Dithiothreitol
DUE	DNA unwinding element
E2F	E2 factor
EB	Elution buffer
EDTA	Ethylenediaminetetraacetic acid
Elg1	Enhanced level of genomic stability 1
ER	Endoplasmic reticulum
eSPAN	Enrichment and sequencing of protein-associated nascent DNA
EtOH	Ethanol
FACS	Flourescence-activated cell sorting
FACT	Facilitates chromatin transactions
FASP	Filter aided sample preparation
FDR	False discovery rate
FKBP	FK506 binding protein
Fkh1/2	Forkhead homolog 1/2
Fob1	Fork blocking less 1
FPC	Fork protection complex
FRB	FKBP-rapamycin binding
FRET	Förster resonance energy transfer
FUN30	Function unknown now 30
G1	Gap phase 1
G2	Gap phase 2
GFP	Green fluorescent protein
GINS	Go-ichi-ni-san
GO	Gene ontology
GTPase	Guanosine triphosphate hydrolase
H2BO3	Dihydrogen borate
H2i	Helix 2 insert
H3	Histone H3
H3K46ac	Histone H3 Lysine 46 acetylation
H4	Histone H4
HCD	Higher energy collision induced dissociation
HCl	Hydrochloric acid

HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
His	Histidine
HML	Homothallic mating left
HMR	Homothallic mating right
HO endonuclease	Homothallic switching endonuclease
HR	Homologous recombination
HRP	Horseradish peroxidase
Hrq1	Homologous to RecQ protein 1
HU	Hydroxyurea
ICLs	Inter-strand cross-links
IgG	Immunoglobulin G
INO80	Inositol requiring 80
IP	Immunoprecipitation
iPond	Isolation of proteins on nascent DNA
ISWI	Initiation switch
KOAc	Potassium acetate
LB	Lithium borate buffer
LC-MS/MS	Liquid chromatography coupled tandem mass spectrometry
Leu	Leucine
LFQ	Label-free quantification
LiAc	Lithium acetate
LiCl	Lithium chloride
M phase	Mitotic phase
M-Ras	Muscle Rat sarcoma virus homolog
MCM	Minichromosome maintenance
Mec1	Mitosis entry checkpoint 1
MeCN	Acetonitrile
MEN	Mitotic exit network
MeOH	Methanol
pMET	Methionine-inducible promoter
Met	Methionine
MgAc	Magnesium acetate
MPF	Mitosis promoting factor
Mrc1	Mediator of the replication checkpoint 1
MRC-LMS	Medical research council - london institute of medical sciences
mRNA	Messenger RNA
MS	Mass spectrometry
Msh3/6	MutS homolog 3/6
MSI	Multicopy supressor of IRA1
Na-(beta)-glycerophosphate	Sodium-beta-glycerophosphate
Na-Deoxycholate	Sodium-deoxycholate
NaAc	Sodium acetate

NaCl	Sodium chloride
NAD	Nicotinamide adenine dinucleotide
NaF	Sodium fluoride
NaN3	Sodium azide
NaOH	Sodium hydroxide
Nat5	N-terminal acetyltransferase
NC	Nitrocellulose
NCC	Nascent chromatin capture
ncRNA	Non-coding RNA
NER	Nucleotide excision repair
NH4OAc	Ammonium acetate
NHEJ	Non-homologous end joining
Nma11	Nuclear mediator of apoptosis 11
Npl4	Nuclear protein localisation 4
NTD	N-terminal domain
NTE	N-terminal extension
OCCM	ORC-Cdc6-Cdt1-MCM2-7
OCM	ORC-Cdc6-MCM2-7
OD600	Optical density at 600 nm
OGRE	Origin G-rich repeated elements
ORC	Origin recognition complex
PAP	Peroxidase anti-peroxidase
PBS	Phosphate-buffered saline
Pcc1	Polarised growth chromatin-associated controller
PCNA	Proliferating cell nuclear antigen
PCR	Polymerase chain reaction
Pdr11	Pleiotropic drug resistance 11
Pif1	Petite integration frequency 1
PIP-seq	Protein interaction profiling coupled sequencing
Pob3	Pol1 binding 3
Pol(alpha, delta, epsilon)	Polymerase (alpha, delta, epsilon))
PP1	Protein phosphatase 1
pre-LC	Pre-loading complex
Pri1/2	DNA primase 1/2
Ps1(beta)	Pre-sensor one beta finger
Psf1/2/3	Partner of Sld5 1/2/3
qPCR	Quantitative PCR
Rad(X)	Radiation sensitive (X)
Rb	Retinoblastoma protein
pre-RC	Pre-replicatiion complex
Rdh54	Rad54 homolog
rDNA	Ribosomal DNA

RENT	Regulator of nuclear silencing and telophase
RFA	Replication factor A
RFC	Replication factor C
Rif1	Rap1-interacting factor
Rlf2	Rap1 protein localisation factor
RNA	Ribonucleic acid
RNase	Ribonuclease
RNR	Ribonucleotide reductase
RPA	Replication protein A
RPC	Replisome progression complex
Rpd3	Reduced potassium dependency 3
Rrm3	rDNA recombination mutation 3
Rtt	Regulator of Ty1 transposition
S phase	Synthesis phase
S-CDK	S-phase cyclin dependent kinase
S0	Minimal fold change setting
SAC	Spindle assembly checkpoint
SCF	Skp1-Cullin1-F-box
SDC	Synthetic complete (glucose) media
SDS	Sodium dodecyl sulfate
SDS-PAGE	SDS-polyacrylamide gel electrophoresis
Sen1	Splicing endonuclease 1
SGC	Synthetic complete (galactose) media
Sgs1	Slow growth suppressor 1
SILAC	Stable isotope labelling in culture
Sir(X)	Silent information regulator (X)
SIRT	Sirtuin 1 gene
Skp1	Suppressor of kinetochore protein mutant 1
Sld2/3/5/7	Synthetically lethal with Dpb11-1
SLE	Systemic lupus erythematosus
Smt3	Suppressor of Mif two
Snf2	Sucrose non-fermenting 2
SORB	Sorbitol
SP3	Single-pot, solid-phase-enhanced sample preparation
SPOC	Spindle position checkpoint
Spt16	Suppressor of Ty 16
SRC	Synthetic complete (raffinose) media
SSB	ss-DNA binding
ssDNA	Single-stranded DNA
SUMO	Small ubiquitin like modifier
Swi4/6	Switching deficient 4/6
TAE	Tris-acetate EDTA

TAP	Tandem affinity purification
TBS-T	Tris-buffered saline with Tween
TCA	Trichloroacetic acid
TE	Tris-EDTA
Ter	Termination site
TFA	Triflouroacetic acid
TFIIB	Transcription factor IIB
TMT	Tandem mass tag
Tof1	Topoisomerase 1-interacting factor
Top1	Topoisomerase 1
Top2	Topoisomerase 2
TrEMBL	Translated EMBL
Tris	Tris(hydroxymethyl)-aminomethane
tRNA	Transfer RNA
Trp	Tryptophan
TUS	Terminus utilisation substance
UA	Urea buffer
Ufd1	Ubiquitin degradation fusion protein 1
Ura	Uracil
WH	Winged-helix
WHD	Winged-helix domain
WT	Wild type
YPD	Yeast extract peptone dextrose
YPG	Yeast extract peptone galactose
ZnF	Zinc finger
ZnOAc	Zinc acetate

Table of Contents

DECLARATION **2**

COPYRIGHT STATEMENT **3**

ABSTRACT **4**

ACKNOWLEDGEMENTS **5**

ABBREVIATIONS **6**

1 INTRODUCTION **19**

1.1 THE EUKARYOTIC CELL CYCLE 20

1.1.1 The phases of the cell cycle *20*

1.1.2 Kinase activity in the regulation of cell cycle processes *22*

1.1.3 Checkpoints *24*

1.2 FEATURES OF DNA REPLICATION 29

1.2.1 Once and Only once *29*

1.2.2 Origins of replication *31*

1.3 INITIATION OF DNA REPLICATION: LICENSING 37

1.3.1 The Origin Recognition Complex *37*

1.3.2 Cdc6 ..... *39*

1.3.3 The Replicative Helicase MCM2-7 ..... *40*

1.3.4 Cdt1/Tah11 ..... *43*

1.3.5 Loading of the Replicative Helicase *43*

..... *46*

1.4 INITIATION OF DNA REPLICATION: FIRING 46

1.4.1 Kinase activity drives the activation mechanism *47*

1.4.2 Sld3/7 and Cdc45: Earliest initiation factors *48*

1.4.3 Pre-LC components Dpb11, GINS, Polε and Sld2: Intermediate initiation factors *51*

1.4.4 Mcm10: Final initiation factor *55*

1.4.5 Helicase activation summary *56*

1.5 THE COMPOSITION, STRUCTURE AND MECHANISM OF THE ACTIVE REPLISOME 57

1.5.1 The Core CMG complex *58*

1.5.2 The Eukaryotic DNA-directed DNA Polymerases *63*

1.5.3 Replication Protein A (RPA) *67*

1.5.4 Sliding clamp and clamp loader *68*

1.5.5 Structural component Ctf4 ..... *70*

1.5.6 Fork protection complex *71*
1.5.7 Topoisomerases *72*
1.5.8 Nucleosome removal/reestablishment *73*
1.6 TERMINATION OF REPLICATION 74
1.7 DNA DAMAGE 77
1.8 ADDITIONAL PROTEINS 78
1.8.1 Rrm3 ..... *78*
1.8.2 Fun30 ..... *82*
1.8.3 Hrq1 ..... *86*
1.8.4 The Sir proteins *88*
1.8.5 CAF-1 complex *90*
1.9 MASS SPECTROMETRY IDENTIFICATION AND QUANTIFICATION OF PROTEINS 91
1.10 CROSS-LINKING APPROACHES TO STABILISE PROTEIN AND PROTEIN-DNA COMPLEXES 93
1.11 RAPAMYCIN-DEPENDENT LINKAGES FOR STUDY OF MCM2-7 COMPLEX 96
1.12 PROJECT AIMS 98
2 MATERIALS AND METHODS 101
2.1 MATERIALS 101
2.1.1 Plasmids *101*
2.1.2 Oligonucleotides *101*
2.1.3 Yeast strains *101*
2.1.4 Antibodies *101*
2.1.5 Media *101*
2.2 METHODS 103
2.2.1 Bacterial transformation *103*
2.2.2 Bacterial culture, plasmid extraction and purification *103*
2.2.3 Subcloning of plasmid DNA *104*
2.2.4 Yeast transformation *105*
2.2.5 Yeast cell cycle synchronisation *106*
2.2.6 Yeast cell cycle analysis by flow cytometry *108*
2.2.7 Mating type testing and HO-nuclease induced switching *108*
2.2.8 Viability assays on agar plates *109*
2.2.9 Chromatin immunoprecipitation (ChIP-qPCR) of replication proteins *109*
2.2.10 Native immunoprecipitation of Psf2-TAP and Cdc45-FLAG *113*
2.2.11 ChIP-MS analysis of replication complexes *118*
2.2.12 Sample preparation for Mass spectrometry *120*
2.2.13 Mass spectrometry analysis and MaxQuant processing *122*

2.2.14 Downstream analysis of Mass spectrometry data 126
2.2.15 Quick extraction of cellular protein by bead beating 128
2.2.16 Trichloroacetic acid protein extraction 129
2.2.17 SDS-PAGE 130
2.2.18 Western blotting 131
2.2.19 Silver staining of proteins in SDS-PAGE gels 131
2.2.20 Genomic DNA extraction 133
2.2.21 Quantification of DNA replication by BrdU incorporation 133
2.2.22 PCR amplification of single colony genomic DNA 134
2.2.23 PCR amplification of integration cassettes 135
2.2.24 Agarose gel electrophoresis of DNA 137
2.2.25 Confocal microscopy of DAPI-stained cells 137

3 RESULTS 138

3.1 ARRESTING COMPLEXES PRIOR TO CMG FORMATION VIA PROTEIN ENGINEERING 138
3.1.1 DNA exit gate-blocking constructs: Intra-hexamer gates 138
3.1.2 Hexamer separation-blocking construct: Inter-hexamer latch 152
3.2 NATIVE IP OF REPLISOME COMPONENT PSF2 172
3.2.1 Single-step IP using Psf2-TAP and magnetic IgG-coupled beads 172
3.2.2 Optimisation of IgG-bound beads to reduce background proteins 175
3.2.3 Wash step optimisation 178
3.2.4 Investigation of pull-down complex identity 180
3.3 CHIP-MS TO INVESTIGATE COMPLEX COMPOSITION AND INTERACTORS 193
3.4 THE COMPLEX COMPOSITION AND INTERACTOME OF ACTIVE CMG *(CDC45 CHIP-MS)* 202
3.5 COMPLEX COMPOSITION AND INTERACTORS DURING STAGES OF REPLICATION INITIATION *(SLD2/3/DPB11 IPS)* 214
3.6 CHANGING COMPLEX COMPOSITION AND INTERACTOME OF MCM2-7 IN THE CELL CYCLE (*MCM4 IP TIMEPOINTS)* 216
3.6.1 Initial validation of Mcm4-FLAG ChIP-MS 216
3.6.2 Mcm4-FLAG ChIP-MS: from loading to termination? 224
3.6.3 Mcm4-FLAG ChIP-MS full time-course in quadruplicate 231
3.7 CHIP-MS TO INVESTIGATE REPLICATION COMPLEXES ARRESTED DURING INITIATION 267
3.7.1 Mcm4-FLAG ChIP-MS in the context of the DNA exit gate constructs 268
3.7.2 Mcm4-FLAG ChIP-MS in the context of the inter-hexamer latch construct 297
3.8 CHIP-MS VALIDATION, TROUBLESHOOTING AND OPTIMISATION 299
3.9 REPLISOME-ASSOCIATED PROTEINS OBSERVED IN CHIP-MS 307
3.9.1 Fun30 311

3.9.2 Rrm3 ..... *317*
3.9.3 Sir3 ..... *321*
..... *322*
3.9.4 Hrq1 ..... *323*
3.9.5 Cac2 ..... *325*
4 CONCLUSIONS AND PERSPECTIVES 327
4.1 DESIGN AND INVESTIGATION OF NOVEL CHEMICAL-BIOLOGY CONSTRUCTS TO INVESTIGATE KEY STEPS IN REPLICATION ACTIVATION 327
4.2 OPTIMISATION AND ADAPTATION OF A CHIP-MS PROTOCOL TO INVESTIGATE LABILE COMPLEXES AND TRANSIENT INTERACTORS OF LARGE *IN VIVO* PROTEIN COMPLEXES 333
4.3 IMPACT AND ADVANTAGES OF THE CHIP-MS TECHNIQUE 336
4.4 CONSIDERATIONS OF THE CHIP-MS METHOD 341
4.5 FUTURE PERSPECTIVES 345
4.6 SUMMARY 348
5 APPENDICES 349
5.1 APPENDIX 1 – PLASMIDS 349
5.2 APPENDIX 2 – OLIGONUCLEOTIDES 350
5.3 APPENDIX 3 – YEAST STRAINS 354
5.4 APPENDIX 4 – ANTIBODIES 357
5.5 SUPPLEMENTARY FIGURES 358
6 BIBLIOGRAPHY 359

TABLE OF FIGURES

FIGURE 1: OVERVIEW OF THE S.CEREVISIAE CELL CYCLE PHASES AND CHECKPOINTS 21
FIGURE 2: FIGURE 2: SCHEMATIC OVERVIEW OF THE MECHANISM ENSURING THAT DNA REPLICATION CANNOT BE RE-INITIATED ON ALREADY REPLICATED DNA WITHIN THE SAME S PHASE 31
FIGURE 3: OVERVIEW OF THE CONSERVED RECOGNITION SEQUENCES IN YEAST ARS....33
FIGURE 4: OVERVIEW OF THE LARGE SCALE DOMAIN AND STRUCTURAL FEATURES OF THE MCM2-7 COMPLEX 41
FIGURE 5: SUMMARY OF THE KEY STEPS IN LOADING OF THE MCM2-7 HELICASE AT ORIGINS OF REPLICATION 46
FIGURE 6: PROTEIN-PROTEIN INTERACTIONS OF THE PRE-LC COMPLEX DURING REPLICATION INITIATION 54
FIGURE 7: SUMMARY OF THE PROCESS OF HELICASE ACTIVATION 57
FIGURE 8: OVERVIEW OF THE STRUCTURE OF THE ACTIVE FORK 58
FIGURE 9: PROPOSED MECHANISMS FOR DNA UNWINDING OF THE MCM2-7 HELICASE IN EUKARYOTES 60
FIGURE 10: SUBUNIT COMPOSITION OF THE THREE EUKARYOTIC DNA POLYMERASES AS THEY ARE IN S.CEREVISIAE 64
FIGURE 11: ALPHAFOLD STRUCTURAL PREDICTION OF S.CEREVISIAE RRM3 80
FIGURE 12: ALPHAFOLD STRUCTURAL PREDICTION OF S.CEREVISIAE FUN30 84
FIGURE 13: ALPHAFOLD STRUCTURAL PREDICTION OF S.CEREVISIAE HRQ1....86
FIGURE 14: SCHEMATIC SHOWING THE RATIONALE BEHIND THE CHEMICAL BIOLOGY CONSTRUCT USED IN THIS PROJECT....97
FIGURE 15: OVERVIEW OF THE CHIP-MS METHOD AS DEVELOPED AND PERFORMED IN THIS STUDY 118
FIGURE 16: OVERVIEW OF CORE REPLICATION FACTOR BEHAVIOUR IN WILD TYPE CHIP-QPCR ANALYSIS 141
FIGURE 17: CHIP-QPCR OF REPLICATION FACTORS ORIGIN ASSOCIATION IN THE C-TERMINAL INTRA-HEXAMER LINKAGE CONSTRUCT 146
FIGURE 18: CHIP-QPCR ANALYSIS OF REPLICATION FACTOR ORIGIN ASSOCIATION IN THE CONTEXT OF THE N-TERMINAL INTRA-HEXAMER LINKAGE CONSTRUCT 149
FIGURE 19: HYPOTHESISED ORDER OF EVENTS DURING REPLICATION INITIATION BASED ON INTRA-HEXAMER LINKAGE CONSTRUCTS 152
FIGURE 20: SEQUENCES AND DESIGN PRINCIPLES OF THE MCM4-FRB MCM5-FKBP INTER-HEXAMER LATCH 154
FIGURE 21: SUBCLONING AND INTEGRATION PCR FOR INSERTION OF FRB AND LINKERS....156
FIGURE 22: PHENOTYPIC EFFECTS OF THE RAPAMYCIN-DEPENDENT INTER-HEXAMER LATCH ON CELLS 159
FIGURE 23: ANALYSIS OF BRDU INCORPORATION AND REPLICATION DYNAMICS IN THE INTER-HEXAMER LATCH STRAIN....163
FIGURE 24: CHIP-QPCR ANALYSIS OF REPLICATION FACTOR ORIGIN ASSOCIATION IN THE CONTEXT OF THE INTER-HEXAMER LATCH 165
FIGURE 25: CHIP-QPCR ANALYSIS OF SLD3-FLAG IN THE CONTEXT OF THE THREE RAPAMYCIN-DEPENDENT LINKAGE CONSTRUCTS168
FIGURE 26: ORDERING OF STRUCTURAL REARRANGEMENTS IN REPLICATION INITIATION....171
FIGURE 27: INITIAL IP EXPERIMENTS USING THE PSF2-TAP CONSTRUCT....173
FIGURE 28: OPTIMISATION OF THE IGG-COUPLED MAGNETIC BEADS....177
FIGURE 29: REDUCING WASH VOLUME AND REPLACING TUBES HAS NO OBSERVABLE NEGATIVE EFFECT ON IP SPECIFICITY 179
FIGURE 30: IPS OF PSF2-TAP TAGGED OR CONTROL STRAINS IN THE CONTEXT OF THE INTRA-HEXAMER LINKAGES 180
FIGURE 31: IMPACT OF ADDITION OF RAPAMYCIN TO DNA-GATE STRAINS CONTAINING PSF2-TAP ON IP EFFICIENCY AND GINS BINDING....182
FIGURE 32: INITIAL MASS SPECTROMETRY ANALYSIS OF PSF2-TAP PULL-DOWNS 185

Figure 33: Analysis of all detected proteins in non rapamycin-treated Psf2-TAP pull-downs 188
Figure 34: Analysis based upon all proteins detected in pull-down samples with and without intra-hexamer linkage binding....189
Figure 35: Cross-linking immunoprecipitation of Psf2-TAP and Cdc45-FLAG....197
Figure 36: Analysis of all detected proteins in cross-linking pull-downs 201
Figure 37: Quality control for Cdc45-FLAG ChIP-MS triplicate 205
Figure 38: Enriched proteins in the Cdc45-FLAG pull-down....208
Figure 39: Visualisation of enriched proteins in Cdc45-FLAG pull-down 212
Figure 40: Trial ChIP-MS of Mcm4-FLAG from cells at two different stages of the cell cycle 218
Figure 41: Trial ChIP-MS of Mcm4-FLAG in two separate cell cycle stages....222
Figure 42: Initial tests of a pMet-Cdc20-induced arrest 226
Figure 43: Analysis of the mass spectrometry data generated from the Mcm4-FLAG ChIP-MS extended time-course 229
Figure 44: Analysis of Mcm4-FLAG ChIP-MS time-course in quadruplicate....234
Figure 45: Principal component analysis of all time-course samples submitted to mass spectrometry 236
Figure 46: Histograms of protein intensities in time-course mass spectrometry samples 238
Figure 47: Volcano plots showing the relationship between factor enrichment in the tagged IP versus untagged control 241
Figure 48: Enrichment analysis of proteins in Mcm4-FLAG ChIP-MS across different time-points 245
Figure 49: Gene ontology enrichment analysis for the first three time-points....251
Figure 50: Enrichment analysis of a core set of replication-associated factors. 252
Figure 51: Bar plots displaying the averaged normalised intensities for individual proteins at each time-point....255
Figure 52: Quantification of MCM2-7 subunits across all time-points and replicates 256
Figure 53: Quantification of average LFQ intensity over the time-course for a range of factors....261
Figure 54: Profile plots of factors of interest in Mcm4-FLAG ChIP-MS time-course 266
Figure 55: Quality control for the large-scale Mcm4-FLAG ChIP-MS 270
Figure 56: Histograms of protein LFQ intensity distributions in intra-hexamer linkage construct Mcm4-FLAG ChIP-MS....273
Figure 57: Enrichment analysis of proteins in Mcm4-FLAG ChIP-MS in context ofintra-hexamer linkage constructs 278
Figure 58: Plots showing the enrichment of a manually curated set of replication factors across different constructs at 0 minutes post alpha-factor release. 279
Figure 59: Plots showing the enrichment of a manually curated set of replication factors across different constructs at 30 minutes post alpha-factor release. 280
Figure 60: Enrichment analysis of proteins in Mcm4-FLAG ChIP-MS in intra-hexamer linkage constructs....283
Figure 61: Analysis of the most enriched proteins in multiple pairwise comparisons 292
Figure 62: Analysis of the comparative ChIP-MS and native IP samples....302
Figure 63: Individual protein abundance analysis across test conditions of ChIP-MS and regular IP 304

Figure 64: Confirmation and viability of strains generated to investigate factors of interest 310
Figure 65: ChIP-qPCR analysis of Fun30-FLAG association with origins 312
Figure 66: Phenotypic effects of the deletion of FUN30 on cells....314
Figure 67: Overview of previously proposed Fun30/SMARCAD1 functionality....316
Figure 68: ChIP-qPCR evaluation of Rrm3 origin association 317
Figure 69: Phenotypic effects of the deletion of RRM3 on cells 319
Figure 70: ChIP-qPCR evaluation of Sir3 origin association 321
Figure 71: Phenotypic effects of the deletion of SIR3 on cells....322
Figure 72: ChIP-qPCR evaluation of Hrq1 origin association 323
Figure 73: Phenotypic effects of the deletion of HRQ1 on cells 324
Figure 74: ChIP-qPCR evaluation of Cac2 origin association....325
Figure 75: Phenotypic effects of the deletion of CAC2 on cells 326
Figure 76: Model of ordering of events in replication initiation. 330
Figure 77: Models of Fun30 function in S.cerevisiae....341
Figure S1: Supplementary figure showing the proposed ssDNA exit channel observed in structural studies alongside the insertion point of the FKBP linkers on Mcm5 for the latch construct. 358

1 Introduction

Deoxyribonucleic acid (DNA) is the core information-carrying component of living organisms. Information flow from this molecule allows the production of every other synthesised cellular component. This includes the most common functional element of a living cell, proteins. The cell's DNA encodes the nucleic acid sequences that can be translated into proteins' amino acid sequences. The DNA also encodes the information of the vast number of non-protein cellular effectors in the form of ribonucleic acid (RNA), such as ribosomal RNA (rRNA), transfer RNA (tRNA) and non-coding regulatory RNAs (ncRNA). The maintenance and protection of this most crucial molecular storage medium is of paramount importance in all living organisms.

In order to grow and survive, all organisms must carry out cellular division. A fundamental necessity of this process is the ability to replicate DNA to produce viable and further divisible progenitor cells. The entire genomic complement of a cell must be copied with high fidelity and efficiency between each cellular division. Further to this, sufficient controls are necessary, with particular importance in multi-cellular organisms, to ensure that the genome is replicated once and only once per cell cycle to maintain correct ploidy levels.

Over the past few decades, much progress has been made in the understanding of the mechanisms by which organisms replicate their genome, with considerable effort focused on understanding the structural and mechanistic details of the replicative protein complex, as well as how that protein complex is assembled and activated. A wealth of biochemical data, and the advent of high-resolution cryo-EM structural studies have brought the field closer to understanding many of the finer aspects of this process, but many features of DNA replication are still being elucidated at a rapid rate.

1.1 The Eukaryotic Cell Cycle

1.1.1 The phases of the cell cycle

As part of the regulation of cellular division, eukaryotic cells follow a program of transcription, protein production/degradation and signalling that divides the activities of a cell into stages, known as the cell cycle (Figure 1). The cell cycle is split into several main phases; the "gap" phases G_1 and G_2, historically identified by their adjacency to the main observable large-scale events of the cell cycle contained within the other phases, "mitotic" M phase, the period marked by physical separation of sister chromatids followed by physical separation of daughter cells, and "synthesis" S phase, during which replication of genomic DNA is carried out. The evidence of this split, phased structure of the cell cycle was first shown by monitoring of radio-labelled phosphate uptake into growing plant root cells, and it was noted that DNA replication occupied a distinct phase to that of cellular division, with the addition of the two G phases further segregating those processes (Howard and Pelc, 1953).

This leaves us with cells containing varied complements of genomic DNA throughout their cycle. Post-mitosis, in G_1, cells contain just a single complement of their genome, whether that be a single copy of each chromosome for naturally haploid cells (1N), or two parental-derived copies in naturally diploid cells (2N). Upon completion of S phase (G_2), in the vast majority of cases, cells have successfully replicated their genomic DNA, and possess double their initial DNA complement (2N or 4N dependant on the starting ploidy). Progressing through M phase, which itself is subdivided into prophase, metaphase, anaphase and telophase; the cells' chromatin condenses and duplicated chromosomes are segregated between the two designated daughter cells, before being separated when the cell fully divides during cytokinesis. Evidence of these macromolecular events was first seen in early microscopic analysis of tissues (Flemming, 1882).

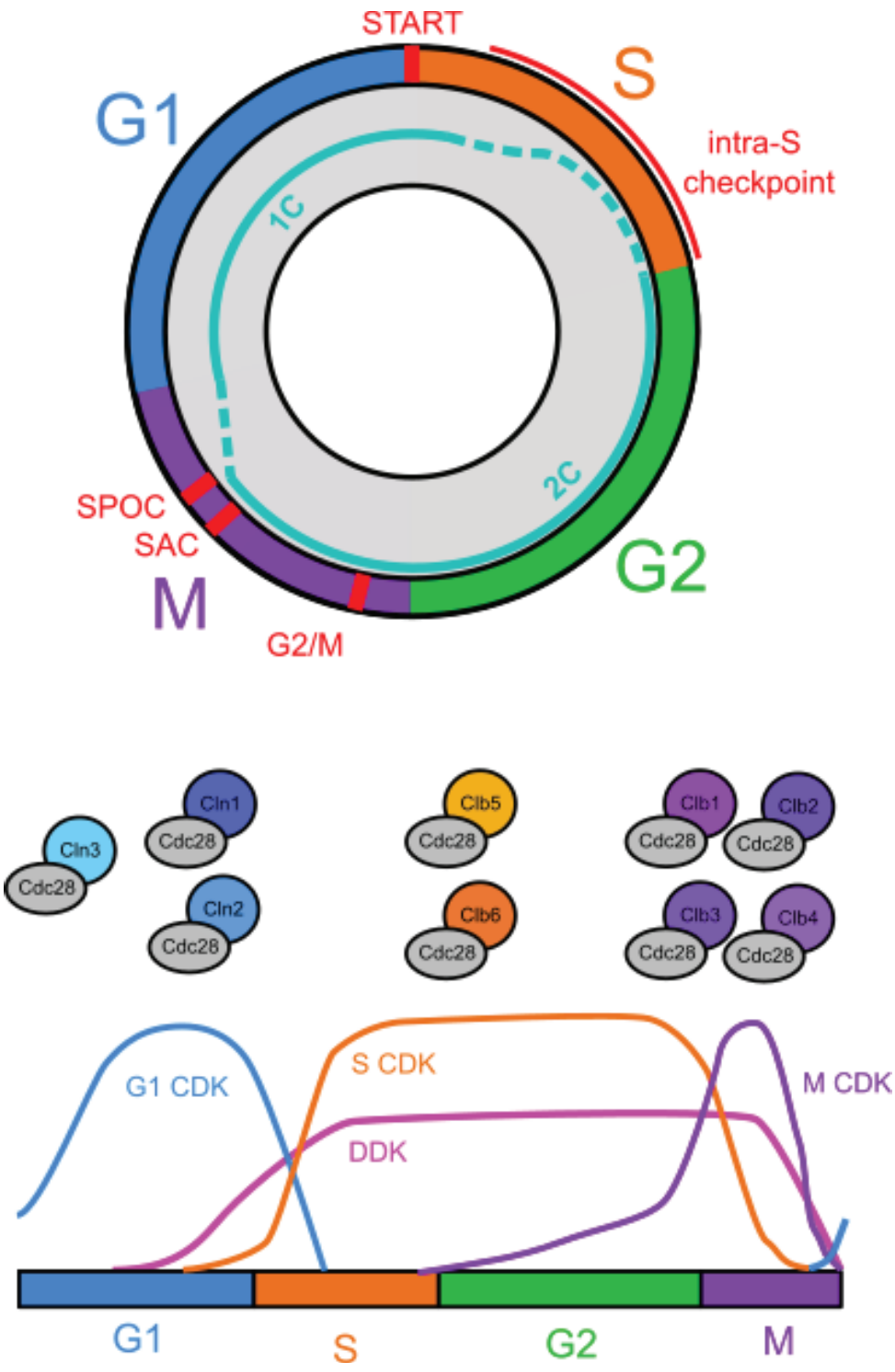


Figure 1: Overview of the S.cerevisiae cell cycle phases and checkpoints, and the activities of the main cyclin dependent kinase-cyclin complexes during the cell cycle. Cell DNA content doubles over the course of S phase, and is reduced back to a 1C level in haploid cells following cell division at the end of mitosis. The four major checkpoints in yeast being the START transition, the G2/M DNA damage checkpoint, the spindle assembly checkpoint (SAC) and spindle position checkpoint (SPOC) with the addition of the intra-S phase DNA damage checkpoint which slows but does not arrest S phase (Matellán & Monje-Casas, 2020). The Cln3, Cln1 and Cln2 cyclins are required for G1 CDK (Cdc28) activity and the transition to S phase, and DDK and S-phase CDK are responsible for the control of many steps in replication initiation. Finally, the activity of mitotic cyclins with CDK drives progression and exit from mitosis. Figure adapted from (Bardin and Amon, 2001; Morgan, 2006; Bloom and Cross, 2007; Hocheger, Takeda and Hunt, 2008)(Bardin & Amon, 2001; Bloom & Cross, 2007; Hochegger, Takeda & Hunt, 2008).

1.1.2 Kinase activity in the regulation of cell cycle processes

1.1.2.1 Cyclins and Cyclin-dependent Kinases (CDKs)

Several experiments in the 1960s and 1970s showed that nuclear events could be controlled by the cytoplasmic fraction, in a series of model organisms (Rusch et al., 1966; Rao and Johnson, 1970; Masui & Markert, 1971). It was shown that there was a factor that did not require protein synthesis that could be activated in the cytoplasm and induce maturation, coined maturation promoting factor (MPF). Later, isolation of the MPF factor provided two polypeptides of size 45 and 32 kDa protein kinase activity was associated with the active purified fraction was shown (Lohka, Hayes & Maller, 1988).

Two further independent lines of research contributed to the current molecular understanding of cell cycle control. An influential screen into cell cycle control mutants identified multiple different cell division cycle mutants (cdc), that led to cellular morphologies indicative of cell cycle arrests or defects (Hartwell, Culotti & Reid, 1970; Hartwell et al., 1973). The identified genes defined a sequential regulatory mechanism which could be halted at certain steps by interference in specific gene functions. Of particular interest were the mutant *cdc28* alleles, which, in addition to exhibiting cell cycle arrest at two different phases of the cell cycle that was unique among the *CDC* genes, also exhibited a flexibility in cell fate, even at the terminal arrest (Hartwell et al., 1974; Piggott, Rai & Carter, 1982). Cells arrested with temperature-sensitive *cdc28* alleles can enter the meiotic division process, suggesting that *CDC28* controls the decision point of cell cycle entry, termed “Start” (Hartwell et al., 1974). As a confirmation of this principle, and that the mechanisms underlying cell cycle control are well conserved in eukaryotes, similar screens carried out in *S.pombe* identified many comparable *cdc* mutants (Nurse, Thuriaux & Nasmyth, 1976). *cdc2*$^{+}$, a dominant mutation with a small-cell phenotype, mapped together with recessive mutations that arrested in G1 and mitosis (Nurse & Bissett, 1981; Nurse & Thuriaux, 1980; Nurse, 1975). It was subsequently found that budding yeast

CDC28 could complement a fission yeast *cdc2* mutation, and that the two proteins shared significant homology (Beach, Durkacz & Nurse, 1982; Hindley & Phear, 1984). From these studies the presence of a cell-cycle control regulator, conserved between yeasts, was confirmed.

The second line of research began when the discovery of proteins with variant abundance dependent on cell-cycle were discovered (Evans et al., 1983; Rosenthal, Hunt & Ruderman, 1980). It was found that injection of these so-called cyclins from clams was able to induce *Xenopus* oocytes to enter M phase, much like MPF would (Swenson, Farrell & Ruderman, 1986). A group of findings soon after linked MPF to a human *cdc2*$^{+}$/*CDC28* homologue by cross-immunoreactivity (Lee & Nurse, 1987; Gautier et al., 1988), and *cdc13*$^{+}$, a genetically linked gene to *cdc2*$^{+}$, was shown to be related to cyclin (Hagan, Hayles & Nurse, 1988). The functional effects of these combined proteins, as well as their conservation between genetically distanced organisms, and the presence of multiple different cyclins in most organisms, laid out a common mechanism for cell cycle control (Murray & Kirschner, 1989). The core component cyclin-dependent kinase (CDK) requires the presence of a specific cyclin which directs its activity, and often, a certain posttranslational modification state, with activating and inhibitory phosphorylation common (Nurse, 2002). The active CDK is able to phosphorylate a number of targets dependent on its cyclin partner, allowing for temporal-specific activation/deactivation of cellular processes to bring about cell cycle events (Swaffer et al., 2016)(Figure 1). In G1 phase, yeast Cdc28 is bound to Cln1, Cln2 or Cln3, and this activates phase specific activity for G1. To proceed to, and during S phase, binding of Cdc28 to Clb5 and Clb6 is required, and finally, Cdc28 association with Clb1, Clb2, Clb3 and Clb4 is required for mitotic processes. Once anaphase promoting complex (APC/C) activation occurs, these last four are degraded to allow cytokinesis, cyclin usage reviewed in Bloom and Cross, 2007 (Bloom & Cross, 2007).

1.1.2.2 Dbf4-dependent kinase (DDK) in DNA replication control

The other main kinase that is highly important for DNA replication in yeast, is the Dbf4-dependent kinase (DDK) (Chapman & Johnston, 1989). This kinase, comprised of the kinase subunit Cdc7, and regulatory subunit Dbf4, phosphorylates subunits of MCM2-7, namely Mcm4, Mcm6 and Mcm2, dependent on their presence as a double hexamer (Sheu & Stillman, 2006; Francis et al., 2009; Sheu & Stillman, 2010; Sun et al., 2014; Larasati & Duncker, 2017). This phosphorylation is critical for assembly of the first stages of helicase activation as will be discussed below (Labib, 2010; Larasati & Duncker, 2017). The regulation of DDK is mainly by the abundance of the regulatory subunit Dbf4, not unlike the regulation of Cdc28, which has highest levels in S phase, rising from G1 levels through the G1/S transition, and is eventually degraded as a function of the anaphase promoting complex (APC/C) (Oshiro et al., 1999; Labib, 2010; Larasati & Duncker, 2017).

1.1.3 Checkpoints

The feedback mechanisms that regulate the cell cycle are intricate and complex, but mutagenesis experiments in fission and budding yeasts were able to identify specific cell cycle steps that mutant cells would arrest in, which are the result of checkpoint activity (Murray, 1995). There are four main sections of the cell cycle that show checkpoint activity in budding yeast; the G1/S checkpoint, the intra-S DNA damage checkpoint, the G2/M checkpoint, and the mitotic or spindle assembly checkpoint (SAC), reviewed in: (Turner, Ewald & Skotheim, 2012; Barnum & O'Connell, 2014; Rhind & Russell, 1998; Lara-Gonzalez, Westhorpe & Taylor, 2012). The network of genes and feedback loops involved in these checkpoint mechanisms is complex, though their general function and activities are summarised below.

1.1.3.1 START

The yeast START decision point, known in mammals as the restriction point (Pardee, 1974), is the moment of commitment of the cell to entering S phase and cellular division (Turner, Ewald & Skotheim, 2012; Johnson & Skotheim, 2013). In yeast, this functional switch is mediated by the amount of cyclin Cln3 in cells, which goes on to regulate levels of transcription of downstream cycle effectors (Tyers, Tokiwa & Futcher, 1993; Wijnen, Landman & Futcher, 2002). Cdk-Cln3 activity phosphorylates and inactivates the transcriptional repressor Whi5, resulting in activation of transcriptional activator Swi4/Swi6, and resultant transcription of Cln1 and Cln2, which further advance cell cycle progression, as described above (de Bruin et al., 2004; Costanzo et al., 2004; Wijnen, Landman & Futcher, 2002; Tyers, Tokiwa & Futcher, 1993). The production, or lack thereof, of Cln3 is a key part of the cellular machinery that drives or withholds progression of cell cycle, and depends on a range of cell size and growth related factors (Schneider et al., 2004; Lei & Tu, 2013). Additionally, upstream cyclin-dependent inhibitor proteins such as Far1 impose a brake on CDK-Cln3 activity, and in yeast, Far1 provides the input for pheromone induced G1 arrest . Overall, this circuit of transcriptional control is analogous to that present in metazoans with the CDK4/Cyclin D – Rb – E2F proteins, despite little sequence similarity in the proteins themselves, indicating an evolutionary pressure for such a system regulating cell behaviour in response to growth rate and nutrient conditions (Johnson & Skotheim, 2013).

1.1.3.2 Intra-S checkpoint

The intra-S phase checkpoint is slightly different to the other checkpoints, in that it does not cause a halt of the cell cycle progression between phases. Instead, the activity of this checkpoint slows the rate of DNA replication seemingly to allow for repair of damage, but does not halt the process completely (Rhind & Russell, 2000; Willis & Rhind, 2009). The intra-S checkpoint

actually appears to be comprised of two inter-linked pathways of DNA damage response, the DNA damage checkpoint (DDC), which can function outside of S phase as well, and the DNA replication checkpoint (DRC), which is specific to stalled forks and fork structures. The intra-S phase checkpoint involves, primarily, the checkpoint kinases Mec1 and Rad53, both of which are involved in the DDC and DRC pathway, making the intra-S checkpoint activity more of a specialised function of the general DNA damage checkpoint (Paulovich & Hartwell, 1995; Rhind & Russell, 2000; Osborn, Elledge & Zou, 2002; Willis & Rhind, 2009). Whilst both DDC and DRC converge on Rad53 (Branzei & Foiani, 2006, 2009; Bacal et al., 2018; García-Rodríguez et al., 2018), they show slightly differing behaviour for their activation and their downstream effects (Branzei & Foiani, 2009; Galanti & Pfander, 2018). In S phase, the DDC seems to rely on the mediation of Rad9 for activation of Rad53 by Mec1 as a result of DNA lesions on nascent DNA in a fork-independent manner (Vialard et al., 1998; Sun et al., 1998; Schwartz et al., 2002; García-Rodríguez et al., 2018; Bacal et al., 2018), in addition to the activity of other damage associated factors such as the 9-1-1 clamp and Dpb11 (Pardo, Crabbé & Pasero, 2017; Pfander & Diffley, 2011). By contrast, the DRC activation of Rad53 is mediated by the phosphorylation of replication fork-bound Mrc1 by Mec1 (Alcasabas et al., 2001; Osborn & Elledge, 2003; Bacal et al., 2018; García-Rodríguez et al., 2018), and requires the presence of the Ctf18-containing version of the RFC clamp (Crabbé et al., 2010; Gellon et al., 2011; Kubota et al., 2011; García-Rodríguez et al., 2015; Pardo, Crabbé & Pasero, 2017). It has been suggested that the two pathways recognise different contexts of DNA damage in S phase, one at, and one behind, the fork, and have different sensitivities and temporal activation, allowing DRC to activate rapidly and trigger Rad53 activity within S phase, and DDC to have a slower, more sustained activation that pervades into later phases (Pardo, Crabbé & Pasero, 2017; Galanti & Pfander, 2018).

The downstream targets of the effector kinases and the individual responses to different types of DNA damage are very extensive, and are beyond the scope of this thesis, but the functions of the intra-S checkpoint specifically with respect to replication are relevant. The checkpoint is thought to induce a range of effects via downstream signalling (Iyer & Rhind, 2017). These include: transcriptional regulation such as upregulation of ribonucleotide reductase (RNR) genes which upregulates dNTP synthesis (Huang, Zhou & Elledge, 1998; Jaehnig et al., 2013), inhibition of further origin firing globally by Rad53 phosphorylation of Sld3 and DDK (Zegerman & Diffley, 2010; Lopez-Mosqueda et al., 2010; McIntosh & Blow, 2012; Yekezare, Gó mez-González & Diffley, 2013), activation of dormant origins at a local level to rescue stalled forks (Yekezare, Gómez-González & Diffley, 2013), and regulation of fork progression rate (Paulovich & Hartwell, 1995b; Willis & Rhind, 2009), stalled fork stability (Sogo, Lopes & Foiani, 2002; Tercero, Longhese & Diffley, 2003; Cotta-Ramusino et al., 2005; Segurado & Diffley, 2008; Cortez, 2015) and restart (Lopes, Foiani & Sogo, 2006). Though the requirement of checkpoint kinase activity to stabilise association of replisome components is disputed (de Piccoli et al., 2012). these activities are thought to protect the cell against DNA damage and help complete DNA replication (Jossen & Bermejo, 2013; Zegerman & Diffley, 2009; Iyer & Rhind, 2017).

1.1.3.3 G2/M checkpoint

In *S. cerevisiae,* the occurrence of DNA damage causes a G2/M arrest at the metaphase to anaphase transition. This is distinct from metazoans, which possess an arrest checkpoint entirely within G2 phase (Matellán & Monje-Casas, 2020). The main players in this arrest are the checkpoint kinases Chk1 and Rad53 (Sanchez et al., 1999; Harrison & Haber, 2006; Liang & Wang, 2007; Valerio-Santiago, de los Santos-Velázquez & Monje-Casas, 2013), which are activated by DNA damage in a similar manner to in the intra-S checkpoint, making the G2/M

checkpoint somewhat of an extension of DNA damage checkpoint activity outside of S phase, and the polo-like kinase Cdc5 (Cheng, Hunke & Hardy, 1998; de Bettignies & Johnston, 2003; Botchkarev & Haber, 2018). The activity and downstream effectors of these kinases is outside of the purview of this discussion, but they prevent completion of anaphase and exit from mitosis via many of the same targets as the spindle assembly checkpoint (SAC) and spindle position checkpoint (SPOC), such as the APC/C complex and the mitotic exit network (MEN) (Sanchez et al., 1999; de Bettignies & Johnston, 2003; Harrison & Haber, 2006; Liang & Wang, 2007; Matellán & Monje-Casas, 2020).

1.1.3.4 Spindle assembly checkpoint

The spindle assembly checkpoint (SAC) is the key method by which unattached kinetochores are signalled to the cell cycle machinery (Lara-Gonzalez, Westhorpe & Taylor, 2012), and arrests cells during mitosis (Figure 1). The details are unnecessary for discussion in this project, but the use of Cdc20-related cycle arrest is used to generate a mitotic arrest, so worth very briefly covering. In the simplest terms, because there are multiple layers and sub-cycles of regulation in this process (Lara-Gonzalez, Westhorpe & Taylor, 2012), the inhibitory signal created by unattached kinetochores prevents Cdc20 from activating the APC/C complex, preventing anaphase and mitotic exit prior to the attachment of all kinetochores to microtubules. (Dawson et al., 1993; Sigrist et al., 1995; Kramer et al., 1998; Uhlmann et al., 2000; Fang, Yu & Kirschner, 1998). The inhibition of the APC/C activation and subsequent mitotic progression can be artificially achieved by preventing the expression of the Cdc20 protein (Yeong et al., 2000) (see materials and methods).

1.1.3.5 Spindle position checkpoint

The other main checkpoint in budding yeast during mitosis is the spindle position checkpoint (Matellán & Monje-Casas, 2020). Because of fixed bud orientation in *S.cerevisiae* this additional step is needed to ensure that chromosomes migrate successfully through the bud

neck and into the daughter cell (Caydasi, Ibrahim & Pereira, 2010; Matellán & Monje-Casas, 2020). This process is mainly regulated by Kin4, but feeds into the regulatory mitotic exit network, also utilised for the spindle assembly and DNA damage checkpoints in preventing mitosis (D'Aquino et al., 2005; Pereira & Schiebel, 2005; Caydasi, Ibrahim & Pereira, 2010; Matellán & Monje-Casas, 2020).

1.2 Features of DNA replication

With the discovery that DNA formed a double-helical structure, and that the information-carrying component - the nitrogenous bases – formed two complementary strands, a potential mechanism for the replication of DNA was proposed (WATSON & CRICK, 1953). This semi-conservative model of replication, where a single strand of parental DNA is used as a template for synthesis of a complementary daughter strand, was confirmed using heavy nucleotide incorporation and gradient fractionation (Meselson & Stahl, 1958).

At its most basic level, then, for DNA to be replicated, the two complementary strands must be separated, followed by synthesis of a daughter strand. The first protein discovered to carry out either of these functions was the polymerase I isolated from *E.coli,* which was able to catalyse the addition of nucleotides to DNA template (LEHMAN et al., 1958). Later, the first enzyme with DNA unwinding properties was purified, also from *E. coli* (ABDEL-MONEM, DÜRWALD & HOFFMANN-BERLING, 1976). Over the course of many years, research has uncovered many more proteins involved with this fairly "simple" biological process.

1.2.1 Once and Only once

In addition to ensuring complete replication, cells must also prevent over-replication. Early work using *Xenopus* egg extracts showed that after a single round of replication, re-replication

does not occur unless the nuclear membrane was permeabilised (Blow & Laskey, 1988). This led to the hypothesis of a replication licensing factor (RLF) that was bound to chromatin and displaced by replication. It would then license replication again during natural nuclear envelope breakdown during mitosis. This licensing step is temporally distinct from replication activation, and separation of these steps prevents re-replication (Diffley & Stillman, 1988)(Figure 2). Assays developed to determine the nature of the licensing process identified two components required for licensing to occur, RLF-M and RLF-B (Chong et al., 1995; Mahbubani et al., 1997). The nature of RLF-M was resolved to be the six minichromosome-maintenance proteins MCM2-7 (Thommes et al., 1997), previously noted for a role in stability of minichromosomes over multiple cell divisions (Gibson, Surosky & Tye, 1990; Maine, Sinha & Tye, 1984; Sinha, Chang & Tye, 1986). RLF-B is now known as Cdt1 and functions in the loading of MCM2-7 onto DNA (Maiorano, Moreau & Méchali, 2000; Nishitani et al., 2000; Fernandez-Cid et al., 2013; Sun et al., 2013). The remaining factors required for licensing are ORC and Cdc6, and all these factors will be discussed in greater detail below.

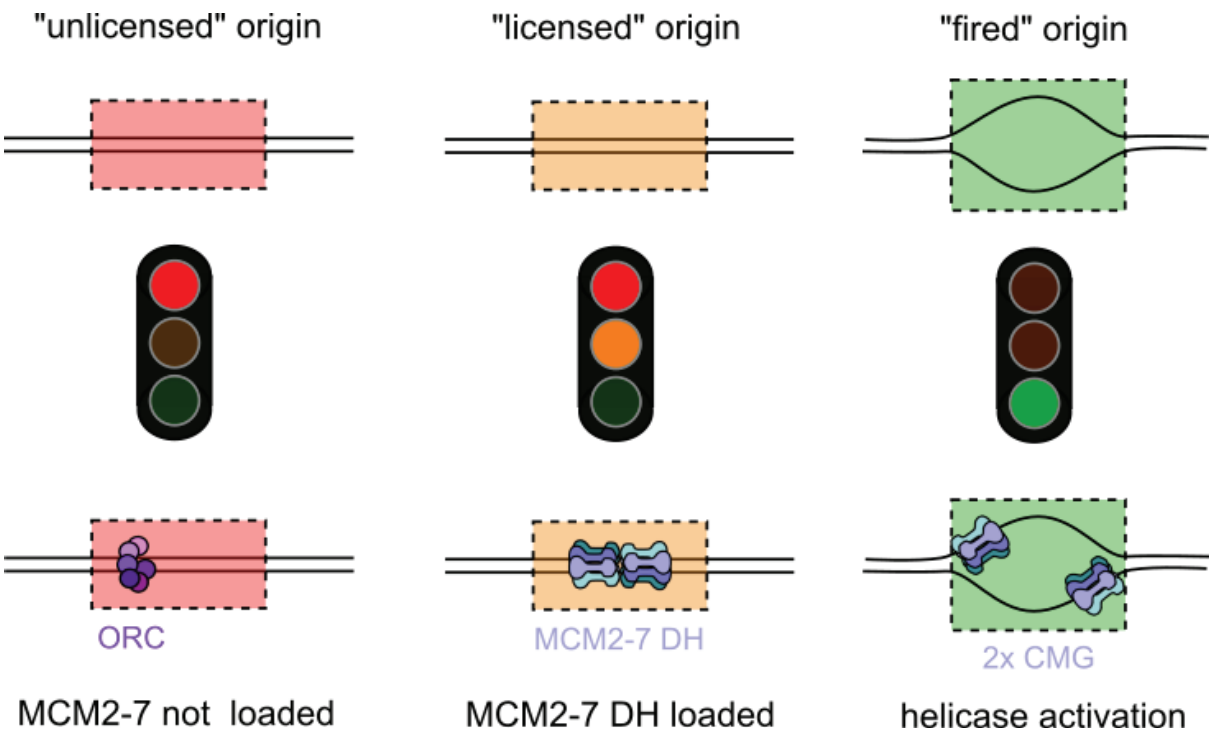


Figure 2: Figure 2: Schematic overview of the mechanism ensuring that DNA replication cannot be re-initiated on already replicated DNA within the same S phase. The separable functions of "origin licensing" and "origin firing" were found to be represented on the molecular level by helicase loading and helicase activation respectively. Only ORC is bound to origins in the "unlicensed" state in G2/M. The "licensed" state is reached in G1, when MCM2-7 is loaded as a double hexamer (DH). The "fired" state is found in S phase, once MCM2-7 helicase is associated with other replication factors to form the Cdc45-MCM2-7-GINS (CMG) core complex. The licensing step cannot occur once S phase has begun, and the firing step cannot occur in G1, ensuring prevention of re-replication (Diffley, 1996; Moyer, Lewis & Botchan, 2006; Evrin et al., 2009; Remus et al., 2009; Riera et al., 2017)

1.2.2 Origins of replication

1.2.2.1 Origin specification

Prokaryotic genomes are replicated starting from a single locus (Bird et al., 1972), with replication continuing bidirectionally around a circular genome until termination occurs, generally at termination regions (O'Donnell, Langston & Stillman, 2013; Hill, Henson & Kuempel, 1987; Massy et al., 1987; Gowrishankar, 2015). Compared to prokaryotes, eukaryotes have much larger genomes, which are encoded on multiple chromosomes. The chromosomes themselves are too large to be replicated by a single pair of forks, since the measured replication fork speed is too slow to complete synthesis within a single S phase (Yurov & Liapunova, 1977). For these reasons, a single eukaryotic origin of replication, or

even one per chromosome, would be insufficient, and instead eukaryotes have multiple origins throughout their genomes (Huberman & Riggs, 1966; Cairns, 1966; Petes & Williamson, 1975). Indeed, eukaryotes possess markedly more potential origins of replication than would be theoretically required simply considering average fork speeds (Fangman & Brewer, 1992; Petes & Williamson, 1975).

The first yeast autonomously replicating sequence (ARS), termed *ars*1, was initially discovered by fragmenting the yeast genome with restriction endonuclease action, and subsequent integration of each fragment into a bacterial plasmid (Stinchcomb, Struhl & Davis, 1979; Struhl et al., 1979). This locus improved transformation efficiency, allowing the plasmid to be inherited without having to integrate into the genome. Confirmation that, the ARS element represented an origin of replication came from examination of forks and their start locations in ARS-containing plasmids and genomic regions (Brewer & Fangman, 1987; Huberman et al., 1987, 1988).

Following the identification of subsequent ARS loci throughout the yeast genome via further recombinant plasmid generation (Chan & Tye, 1980), sequence analysis showed that there was a highly conserved 11-bp sequence across all origins identified which was AT-rich, dubbed the ARS consensus sequence (ACS) (Broach et al., 1983; Newlon, 1988). Mutagenesis on the *ARS1* locus further identified three regions acting in *cis* important for the activity of an origin. These sequences, the B1, B2 and B3 elements, are positioned closer towards or within an A rich region 3' of the ACS (Figure 3) (Bell & Labib, 2016; Marahrens & Stillman, 1992).

The ACS and B1 sequences represent the binding site of the origin recognition complex (ORC), which is an essential protein complex made up of the six proteins Orc1-6 (Bell & Stillman, 1992; Rao & Stillman, 1995; Rowley et al., 1995; Diffley & Stillman, 1988). The function and activity of ORC will be explained in much greater detail in following sections.

The B2 sequence appears to be an inverted version of the ACS, though does not show conservation at all origins (Wilmes & Bell, 2002). A recent study on the two-ORC mechanism of helicase loading suggests why this feature may be present, if not why it is absent at some origins (Miller et al., 2019).

The B3 element was found to be the binding site for the transcription factor Abf1 (Diffley & Stillman, 1988). Whilst this protein has been credited with acetylating histones nearby and aiding replication by affecting the nucleosome-free region around the origin (Venditti et al., 1994; Lipford & Bell, 2001; Buchman et al., 1988), this site is similarly not highly conserved across origins.

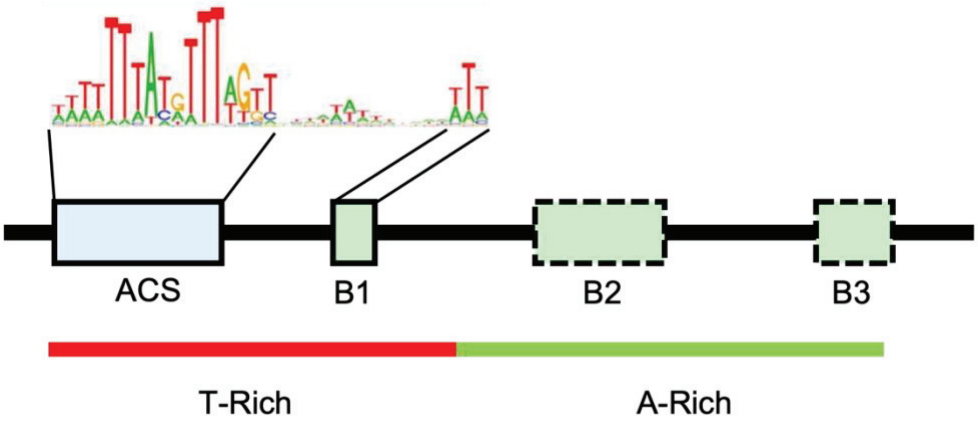


Figure 3: Overview of the conserved recognition sequences in yeast ARS . ORC binds the ACS and B1 elements (Rao & Stillman, 1995; Rowley et al., 1995). The B2, where present, resembles a reversed ACS, which may be explained by symmetrical association of two ORCs with origin sequences (Coster & Diffley, 2017; Miller et al., 2019). The B3 element is the site of Abf1 binding (Diffley & Stillman, 1988). Origin A-element sequence reproduced from Eaton et al., 2010 (Eaton et al., 2010).

In higher eukaryotes, the highly defined sequence elements and genomic locations of origins are not present. In *S.pombe*, another well studied though evolutionarily quite distant yeast (Sipiczki, 2000), origins are much larger regions, with no specific sequences that precisely define ORC binding, only AT-rich sequences bound by an Orc4 AT-Hook (Chuang & Kelly, 1999; Moon et al., 1999; Kong & DePamphilis, 2001; Lee et al., 2001; Akerman et al., 2020). Whilst the nature of origins has made the study of replication initiation very convenient in *S.cerevisiae*, this means aspects of the process will not be directly comparable with other organisms. Studies investigating origin localisation in a humans and other higher eukaryotes

reveal that use of individual origins is more flexible, and the features that enable such events are markedly less genetically defined (Remus, Beall & Botchan, 2004; Vashee et al., 2003), with few exceptions (Austin, Orr-Weaver & Bell, 1999; Bielinsky et al., 2001; Zellner et al., 2007; Abdurashidova et al., 2000). Despite this change in specification mechanism, it is still ORC binding that enables the eventual initiation in higher eukaryotes (Bell & Dutta, 2002). Many different features of metazoan origins have been proposed as possible requirements for ORC binding, but there is no one conserved and necessary feature (Méchali, 2010; Cayrou et al., 2011, 2015; Ekundayo & Bleichert, 2019; Yella et al., 2020; Akerman et al., 2020). In terms of sequence, the two main features associated with origin activity are CpG islands, generally linked to transcriptional promoters (Cadoret et al., 2008; Sequeira-Mendes et al., 2009), or AT-rich DNA unwinding elements (DUEs) (Méchali, 2010; Cayrou et al., 2011). In terms of DNA structure, an origin G-rich repeated element (OGRE) is thought to have a propensity for forming G4 quadruplexes, which has been linked to origin activity (Cayrou et al., 2012; Prorok et al., 2019), as have other topological features such as supercoiling and helical twist (Remus, Beall & Botchan, 2004; Cadoret et al., 2008; Comoglio et al., 2015; Yella et al., 2020). Finally, chromatin structure and modifications have also been shown to affect origin activity, which is associated with nucleosome-free regions and the H3K46ac mark (Cayrou et al., 2015; Prioleau & MacAlpine, 2016).

1.2.2.2 Origin timing

In *S. cerevisiae*, there is a tightly controlled process regulating the firing timing of specific origins (Raghuraman et al., 2001). This timing is separable from but linked to the efficiency of firing of a particular origin, since the later an origin fires, the more likely it is to be replicated passively, due to the passage of forks from adjacent origins past the late origin position (Aparicio, 2007). By removing the possibility of passive replication, it was shown that even very late origins are reliably fired given enough time (Santocanale, Sharma & Diffley, 1999;

Vujcic, Miller & Kowalski, 1999). One aspect of origin firing in yeast that was discovered early on, was the position effect mediated by telomeric regions on origin firing timing (Ferguson & Fangman, 1992). This suggested that the local chromatin environment, which is relatively compacted near telomeres (Grunstein, 1998), confers this property of timing to the origins present there. This was confirmed in studies that showed cells impaired in their ability to form or spread the telomeric heterochromatin, such as *SIR* protein mutants, also showed disregulation of telomeric origin firing timing (Zappulla, Sternglanz & Leatherwood, 2002; Cosgrove, Nieduszynski & Donaldson, 2002). However, the ability of chromatin and nucleosome state to influence origin timing does not appear to be limited to telomeres, and several telomere-associated proteins, including the SIR proteins and Rif1, are implicated in affecting origin timings of origins distant from the telomeres (Pappas, Frisch & Weinreich, 2004; Hayano et al., 2012; Hoggard et al., 2018). This is thought to be due to their histone deacetylase activity, a function that has been linked to origin timing in multiple studies for multiple protein complexes (Vogelauer et al., 2002; Aparicio et al., 2004; Pappas, Frisch & Weinreich, 2004; Knott et al., 2009; Yoshida et al., 2014; Aparicio, 2007). Early work also implicated the ARS-binding protein Abf1 in maintaining a nucleosome-free region at the origin (Lipford & Bell, 2001; Marahrens & Stillman, 1992).

It is unclear exactly what mechanism is causing the effects on origin timing by these factors, and at what stage in the origin licensing-firing pathway they affect. Recent work implicating Sir2 with a broad control over origin firing timing suggests that Sir2 affects the loading of MCM2-7 helicase at early and telomeric origins, and it is this that causes the preferential shift to early origin usage in *Δsir2* cells (Hoggard et al., 2020). Other studies suggest that origin firing timing is rather mediated by the later stages of replication initiation, the firing event itself, with histone deacetylase Rpd3 deletion resulting in early recruitment of firing factor Cdc45 to origins. In addition, because several of the initiating proteins are in limiting

concentration (Mantiero et al., 2011), changes in their association dynamics could have large effects on timings. Indeed, overexpression of combinations of six replication initiation proteins caused an advancement of origin firing time generally and made late origins fire earlier (Tanaka et al., 2011; McGuffee, Smith & Whitehouse, 2013; Lynch et al., 2019).

Furthermore, proteins such as Ctf19 and the Forkhead factors Fkh1 and Fkh2 have been shown to affect the recruitment of the activating replication kinase DDK to origins, which in turn affects the recruitment of Cdc45 and Sld3 firing factors, a marker of origin firing timing (Aparicio, Stout & Bell, 1999). In contrast, at late origins Rif1 is thought to recruit protein phosphatase 1 (PP1), which actively counteracts the activity of DDK at these sites, reducing binding of early replication factors (Mattarocci et al., 2014; Davé et al., 2014; Hiraga et al., 2014).

The Forkhead factors in particular appear to have an additional role in the spatial organisation of origins in the cell, which creates clusters of early firing origins that may show higher local concentrations of the limiting firing factors required for early origin firing (Pasero, Braguglia & Gasser, 1997; Lengronne et al., 2001; Heun et al., 2001). Meanwhile, regions of heterochromatin that fire later may occupy positions closer to the nuclear periphery (Aparicio, 2007).

In yeast, there are several mechanisms that seem to regulate origin firing time, that seem to act at differing times during replication initiation. Some, like Sir2, appear to affect the loading of MCM2-7 at origins (Hoggard et al., 2020), whilst others, such as the Forkhead factors, appear mainly to regulate the recruitment of activation factors during the firing process itself (Fang et al., 2017; Knott et al., 2012). This is further complicated by studies that suggest the temporal regulation of origin firing timing is established between mitosis and the end of G1 (Raghuraman, Brewer & Fangman, 1997), and very recent work that indicates disparities in MCM2-7 loading at origins can in fact explain much of the variations in origin activation

timings (Foss et al., 2021). The effect of a particular firing determinant can affect different subtypes of origin differently, as shown by the studies above that perturb timing determination, so the fine timing control for each individual origin is possibly the result of a combination of mechanisms acting at different stages of the licensing and firing processes.

In higher eukaryotes, where the mechanisms governing origin specification itself are less clearly defined, the mechanisms by which cells regulate the timing regime of origin firing is similarly less well understood than in yeast (Boos & Ferreira, 2019). However, there are some clear parallels that can be drawn between metazoans and yeast. Like yeast, determination of origin firing appears to be multi-faceted in higher eukaryotes, and perturbation of one aspect of timing determination affects certain origins more than others. Some of the individual pathways also appear to be conserved, with Rif1 showing a role in temporal regulation of origin firing in mouse and human cells, that may be related to PP1 activity or chromosomal structure (Cornacchia et al., 2012; Foti et al., 2016; Sukackaite et al., 2017). Additionally, the counterpart of the yeast initiation factor Sld3, Treslin/TICRR, has been implicated in directing replication timings via interactions with acetylated histone-binding proteins (Sansam et al., 2018), suggesting that histone acetylation, like in yeast, is an important regulator of replication timing in higher eukaryotes, and that recruitment of initiation factors may also be key in this process.

1.3 Initiation of DNA replication: Licensing

1.3.1 The Origin Recognition Complex

The origin recognition complex, ORC, is a heterohexameric complex that is well conserved across eukaryotes, with some variability in subunit size order (Gavin, Hidaka & Stillman, 1995; Tugal et al., 1998; Li & Stillman, 2012). All of the ORC subunits, save Orc6, are relatively similar in domain structure (Duncker, Chesnokov & McConkey, 2009). In addition to C-terminal winged-helix domains, important for DNA binding (Gaudier et al., 2007), they each

contain conserved AAA^+ domains (Speck & Stillman, 2005; Duncker, Chesnokov & McConkey, 2009), and appropriately, the complex shows ATP-dependent binding to origins (Bell & Stillman, 1992), in a manner reminiscent of the prokaryotic DNA replication initiator DnaA (Kawakami & Katayama, 2010). In addition, the largest subunit, Orc1, has an N-terminal bromo-adjacent homology (BAH) domain, which typically mediate protein-protein interactions (Callebaut, Courvalin & Mornon, 1999; Bell et al., 1995). In the case of Orc1, this domain appears to mediate interactions with origin-adjacent nucleosomes (de Ioannes et al., 2019). Orc6 is the most divergent, not just from the other ORC subunits in yeast, but also shows a large amount of divergence from yeast to higher eukaryotes (Dhar & Dutta, 2000; Duncker, Chesnokov & McConkey, 2009). Orc6 also appears to have an important role in the recruitment of the second MCM2-7 hexamer (Miller et al., 2019; Gupta et al., 2021)

Recent structures of ORC complexes bound to origin DNA from *D. melanogaster* and *S. cerevisiae* helped elucidate how the domain structure and three-dimensional arrangement of subunits allow DNA binding of ORC (Bleichert, Botchan & Berger, 2015; Bleichert et al., 2018; Li et al., 2018). The *Drosphila* ORC structure showed that ORC exhibits a near-planar ring conformation and in higher metazoans can adopt an autoinhibitory conformation, that blocks ring opening and engagement of DNA (Bleichert, Botchan & Berger, 2015; Bleichert, 2019; Bleichert et al., 2018). Structural information of the human ORC showed closer structural similarities to the *S.cerevisiae* complex in its active form (Sun et al., 2014), and an autoinhibitory confirmation was not observed (Tocilj et al., 2017).

The structure of the *S.cerevisiae* ORC complex bound to ARS DNA was able to explain structurally the sequence-specific nature of this binding (Li et al., 2018). Again, the AAA+ domains of the ORC subunits formed a tilted ring that in this case surrounds origin DNA. Small initation-specific motif (ISM) subdomains in all AAA+ regions of the ORC subunits except Orc4 were found to interact with DNA in a general non-specific manner, as too were the β-

hairpins of the Orc2-5 winged helices (WH), which formed close associations with the DNA, mostly with the major groove. Specific interaction that depended on DNA ACS sequence at was carried out by a basic patch in Orc1. This patch inserts into the major groove at the ACS site and interacts with thymine, and a β-hairpin of Orc4 winged helix domain (WHD). At the B1 element, a flexible loop of the Orc2 subunit, a basic patch of the Orc5 WHD, and the TFIIB domain of Orc6 interact with the DNA with slightly less expected sequence specificity (Li et al., 2018). The structure of yeast and fruit fly ORC bound to DNA shows a bend in the DNA introduced by the protein complex (Li et al., 2018; Schmidt & Bleichert, 2020). In yeast, this consisted of a 25° bend at the ACS, and another 51° bend at the B1 element. Interestingly, in *Drosophila*, the bend was seen to vary dependent on the sequence of the bound DNA. The exact function of this introduced bend is not certain, but comparison of ORC-DNA with later structures of loading intermediates involving MCM2-7 suggest it may aid that process (Yuan et al., 2017; Li et al., 2018; Yuan et al., 2020d).

1.3.2 Cdc6

The Cdc6 protein was initially shown to be required for some unknown event in early S phase (Hartwell, 1976), and subsequently identified in fission yeast as being required for S phase entry (Kelly, Nurse & Forsburg, 1993). Further work demonstrated that the presence of Cdc6 was required to extend the footprint of the nuclease-protected, and therefore protein bound, region at replication origins post-mitosis, termed the pre-RC (Diffley et al., 1994; Cocker et al., 1996; Perkins & Diffley, 1998; Speck & Stillman, 2007). Cdc6 shares many similarities in domain structure to Orc1-5 in *S.cerevisiae*. It contains a homologous AAA^+ and WH domain and shares more conserved sequence with Orc1 than Orc1 does with the other AAA^+ ORC subunits (Duncker, Chesnokov & McConkey, 2009; Bell et al., 1995). Cdc6 associates with the ORC complex bound to DNA, dependent on ATP-binding, forming a ring shaped complex (Santocanale & Diffley, 1996; Mizushima, Takahashi & Stillman, 2000; Duncker, Chesnokov

& McConkey, 2009; Speck et al., 2005; Liang, Weinreich & Stillman, 1995; Seki & Diffley, 2000; Cocker et al., 1996). ORC-Cdc6 bound to DNA extends the footprint of ORC on genuine ARS sequences, but when either not bound to DNA, or to non-origin sequences, Cdc6 ATPase activity causes disassembly of Cdc6 from the complex, preventing progression to further steps of replication initiation, namely loading of the replicative helicase (Speck & Stillman, 2007). In this way, the binding and ATPase activity of Cdc6 acts as a form of quality control for correct ORC positioning prior to helicase loading. Recent structural studies have resolved the structure of ORC-Cdc6 bound to DNA for yeast and *Drosophila* complexes (Feng et al., 2021; Schmidt & Bleichert, 2020), and reveal that, at least in yeast, the association of Cdc6 is responsible for stabilisation of ORC interaction with DNA, and the availability of several binding sites responsible for MCM2-7 recruitment. This explains why Cdc6 is required for both the extension of the ORC footprint on DNA (Speck et al., 2005), and for the ability to recruit MCM2-7 (Fernandez-Cid et al., 2013).

1.3.3 The Replicative Helicase MCM2-7

1.3.3.1 MCM structure and homology

The replicative helicase is formed from six proteins that were initially identified in screens investigating plasmid inheritance or cell cycle progression. These proteins were collated under the name Mcm, standing for minichromosome maintenance (Moir et al., 1982; Maine, Sinha & Tye, 1984; Sinha, Chang & Tye, 1986; Maine, Sinha & Tye, 1984; Thommes et al., 1992). Significant sequence similarity between six of these, and discovery of homologous genes in other organisms further defined the Mcm2-7 family of six (Dutta & Bell, 2003; Tye, 2003). The Mcm proteins contain the AAA^+ ATPase motif, and form a heterohexameric complex in a ring shape (Forsburg, 2004; Bochman & Schwacha, 2007; Davey, Indiani & O'Donnell, 2003; Tye, 2003). Initially, the activity of Mcm2-7 was only linked to origins themselves (Dutta & Bell, 2003), but the discovery of Mcm2-7 associated with non-origin DNA during S phase

established their movement with the replication fork and the hypothesis that Mcm2-7 could function as the replicative DNA helicase (Aparicio, Weinstein & Bell, 1997). This was further expanded upon with work that showed continued Mcm2-7 activity was required to complete replication after origin firing (Labib, Tercero & Diffley, 2000), and later, work that was able to demonstrate *in vitro* activity of the Mcm2-7 complex (Bochman & Schwacha, 2008), with higher and more reliable activity with the addition of other proteins known to be involved with continuation of replication beyond initiation, Cdc45 and the four GINS proteins ((Moyer, Lewis & Botchan, 2006; Kanemaki et al., 2003; Ilves et al., 2010; Tercero, Labib & Diffley, 2000; Takayama et al., 2003).

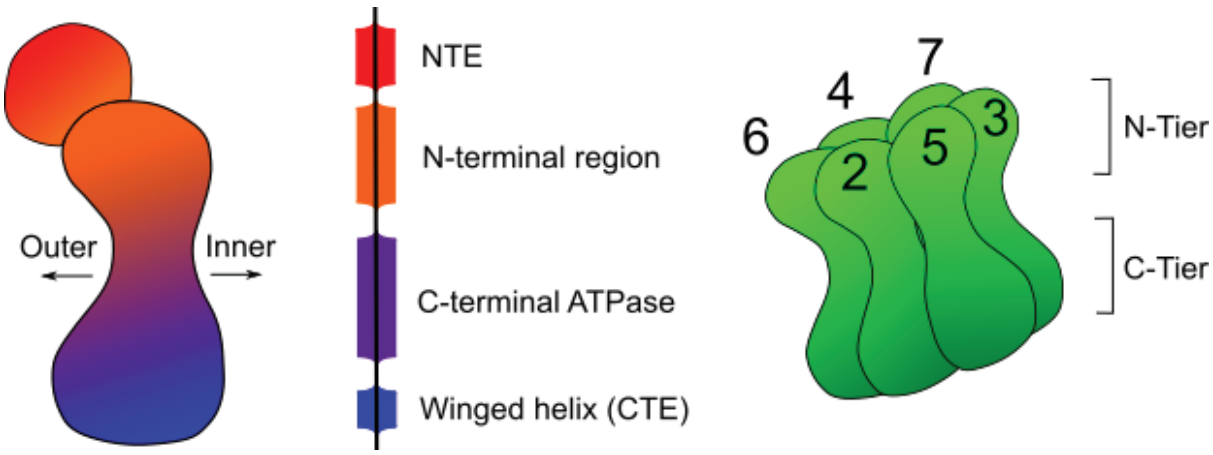


Figure 4: Overview of the large scale domain and structural features of the MCM2-7 complex .The three main domains of all six MCM2-7 subunits are the N-terminus (subdivided into N-terminal extension (NTE) and N-terminal region including a Zinc finger (ZnF) and other interaction domains), the AAA+ ATPase domain (the most conserved region) and the C-terminal extension (CTE) (containing a winged helix domain) in red-orange, purple and blue respectively (Costa & Onesti, 2009; Riera et al., 2017). The six subunits have varying lengths of N-terminal and C-terminal extensions, and are N-C terminally aligned and arranged into a ring as displayed (Davey, Indiani & O'Donnell, 2003; Bochman, Bell & Schwacha, 2008; Li et al., 2015)

The MCM proteins are comprised of three main domains, the N-terminus, the core AAA+ ATPase domain, and the C-terminal extension (Figure 4)(Costa & Onesti, 2009). The N-terminus contains domains that are involved in protein-protein interactions, both intra-hexamer and (Bochman & Schwacha, 2009), inter-hexamer (Li et al., 2015; Noguchi et al., 2017), and carry N-terminal extensions (NTEs) that have an important role in the regulation of activation

of replication as phosphorylation targets of DDK (Larasati & Duncker, 2017), and in replication-coupled chromatin assembly (Huang et al., 2015).

The AAA+ ATPase domain is the most highly conserved region, and specifically, the presence of a particular helical segment after the core ATPase domain and an insert within the core ATPase domain itself, places the MCM proteins in clade 7 of the AAA+ ATPases, the pre-sensor II insert (PSII) group. There is structural variability between the ATPase active sites of each AAA+ domain, which are actually formed at the interface between subunits, with one subunit contributing the binding pocket and the adjacent subunit providing the arginine finger for catalysis (Davey, Indiani & O'Donnell, 2003; Wendler et al., 2012; Sun et al., 2014). These sites show variable nucleotide occupancy that varies dependent on the state the complex is in, and the binding of dsDNA (Noguchi et al., 2017; Li et al., 2015; Sun et al., 2014, 2013; Yuan et al., 2017). It has been noted there appears to be cooperativity between the ATPase sites, where the nucleotide occupancy of one subunit, or the mutation of a particular ATPase site has an effect on the nucleotide binding of other sites and the overall ATPase activity of the complex (Ilves et al., 2010; Davey, Indiani & O'Donnell, 2003; Schwacha & Bell, 2001). Some of the ATPase sites were shown to have greater or lesser effects on the overall ATPase rate of the complex, suggesting that the contribution from each to the ATPase functionality of the complex is not equal (Bochman, Bell & Schwacha, 2008; Bochman & Schwacha, 2010; Bell & Botchan, 2013). There is evidence that the archael MCM complex has a rotary ATPase cycle of cooperativity, where nucleotide binding and hydrolysis occurs in a cyclical manner around the ring, and this activity drives translocation (Meagher, Epling & Enemark, 2019). However, in contrast to the archael MCM, the *S. cerevisiae* MCM2-7 does not form contacts with the lagging strand with all subunits (Yuan et al., 2020), and, while the ATP-binding pattern is consistent with a possible rotary model, it is not clear if MCM2-7 follows this pattern, despite it being common in many ring shaped ATPases (Liao et al., 2005; Thomsen & Berger, 2008),

given the mild effect of ATPase mutations on DNA helicase activity of the MCM2-7 for a some of the subunits, suggesting not all are required for full activity, as might be expected of a rotary model .

In addition to the NTD and AAA+ domain, most of the MCM2-7 proteins, aside from Mcm2 have a C-terminal extension, that forms a winged helical structure (Riera et al., 2017). Observations that the C-terminal extension (CTE) of Mcm5 appeared to block the central pore were initially explained as an exclusion mechanism to cause DNA unwinding, but this suggestion relied upon a C-terminal first translocation model, which has since been disproven. Other than this, the CTEs have been shown to provide Cdt1 and Cdc6 binding during helicase loading, and to interact with polymerase ε at the active replication fork (Jones et al., 2021), but limited other functions have been ascribed to these domains (Riera et al., 2017; Yuan et al., 2020).

1.3.4 Cdt1/Tah11

Cdt1/Tah11 was first discovered to have a role in MCM2-7 loading onto chromatin in fission yeast (Nishitani et al., 2000), and later confirmed to act similarly in budding yeast (Tanaka & Diffley, 2002; Devault et al., 2002). In G1 phase cells, MCM2-7 and Cdt1 form a heteroheptameric soluble complex together (Kawasaki et al., 2006), with a single Cdt1 binding an MCM2-7 hexamer in a manner that appears to stabilise the hexamer in an "open-ring" state, amenable for loading onto DNA (Frigola et al., 2017; Zhai et al., 2017). Cdt1 alters the conformation of the MCM2-7 hexamer ring, which allows complex formation (Fernandez-Cid et al., 2013; Zhai et al., 2017). Cdt1 itself disassociates from origins after helicase loading, and is not part of the active replisome after loading (Ticau et al., 2015, 2017).

1.3.5 Loading of the Replicative Helicase

Following work with archaeal and viral proteins (Chong et al., 2000; Costa et al., 2006; Gomez-Lorenzo et al., 2003; Li et al., 2003; Pape et al., 2003; Shechter, Ying & Gautier, 2000), it was

shown that in yeast, MCM2-7 is loaded onto origins as a double hexamer (DH) during licensing (Evrin et al., 2009; Remus et al., 2009). Investigations into the mechanism behind the loading of two hexamers revealed intermediate states and the roles of Cdt1 and Cdc6 (Figure 5). Initially, Cdc6 is recruited to chromatin-bound ORC in late M and early G1 (Weinreich, Liang & Stillman, 1999) in an ATP-dependent manner, functioning as a molecular switch to promote further pre-RC formation (Speck et al., 2005). The complex is then competent to recruit a Cdt1-bound Mcm2-7 hexamer and form an ORC/Cdc6/Cdt1/MCM2-7 (OCCM) complex (Fernandez-Cid et al., 2013; Sun et al., 2013). The ring shaped MCM2-7 hexamer is loaded onto DNA via a single DNA entry gate between subunits Mcm2 and Mcm5 (Samel et al., 2014). Loss of Cdt1 and Cdc6 through ATP hydrolysis to form an ORC/MCM2-7 intermediate (OM) (Coster et al., 2014; Kang, Warner & Bell, 2014)

seems to promote ring closure (Ticau et al., 2015) and allows recruitment of a second Cdc6, forming an ORC/Cdc6/MCM2-7 (OCM) and a second Cdt1/MCM2-7 complex (Fernandez-Cid et al., 2013; Randell et al., 2006; Ticau et al., 2015b, 2017c). This results in an MCM2-7 double hexamer (DH), with both N-terminal domains (NTDs) facing towards the other (Costa et al., 2014; Evrin et al., 2014; Gambus et al., 2011; Remus et al., 2009). Structural information obtained by cryo-EM on this double hexamer showed extensive interactions between the N-terminal domains (NTDs) of each MCM2-7 hexamer, though the full length of the N-terminal extensions of the MCM2-7 subunits were not generally resolved (Noguchi et al., 2017; Ali et al., 2017). The MCM2-7 hexamers do not stack perfectly in the double hexamer, instead they exhibit a tilt of around 14° between their axes (Li et al., 2015; Sun et al., 2014). This is because of the interface interactions being provided mostly by the Mcm3/5/7 half rings of each hexamer. This arrangement may help expose the N-terminal serine/threonine rich domains of the other half ring, Mcm2/4/6, which can then be accessed by DDK activity later during activation to perhaps initiate splitting (Noguchi et al., 2017). The two hexamers are also

rotationally offset, with the Mcm2/5 DNA entry gate (Samel et al., 2014) of each hexamer out of alignment with the other (Sun et al., 2014). The genetic structure of origins in *S.cerevisiae* posed a question that was not immediately answered when the structural details of the loading reaction intermediates were discovered. Intermediate structures clearly showed ORC bound to the C-terminus of a single MCM2-7 hexamer (Fernandez-Cid et al., 2013; Ticau et al., 2015, 2017), and structures of the OCCM showed a single ORC bound to the forming double hexamer (Sun et al., 2013; Yuan et al., 2017). Some origins in yeast appeared to show multiple copies of degenerate ORC recognition motifs (Palzkill & Newlon, 1988), and *in vitro* experiments showed that for full MCM2-7 loading activity, a single ORC site was not sufficient (Coster & Diffley, 2017). Instead, a second, if weaker, ORC site was required, suggesting a synergistic activity of the second hexamer loading, but one that still required loose association of ORC with a second site on the opposite end of the double hexamer. This fits with mutational analysis that showed that residues in Mcm3 and Mcm6 important for loading of the first hexamer, prevented second hexamer formation even once the initial hexamer was bound to DNA (Frigola et al., 2013; Fernandez-Cid et al., 2013). The question of whether the loading of two MCM2-7 hexamers required one or more ORC complexes continues to be investigated to this day. Recent time-course cryo-EM data of MCM2-7 loading intermediates using an *in vitro* reconstitution system identified that the first MCM2-7 hexamer can help recruit a second ORC via interactions with the N-terminal interface of Mcm subunits, which can then recruit the second MCM2-7 hexamer (Miller et al., 2019). This interaction of ORC with the N-terminus of MCM2-7 was recently confirmed in a FRET study, but as in an interaction in a binding mode of the initial ORC carrying out a positional switch or “flip” to the opposite side of the MCM2-7 hexamer (Gupta et al., 2021). Whether mediated *in vivo* by a second ORC complex or not, the evidence suggests that, indeed, two ORCs are required to load the two head-to-head MCM2-7 complexes.

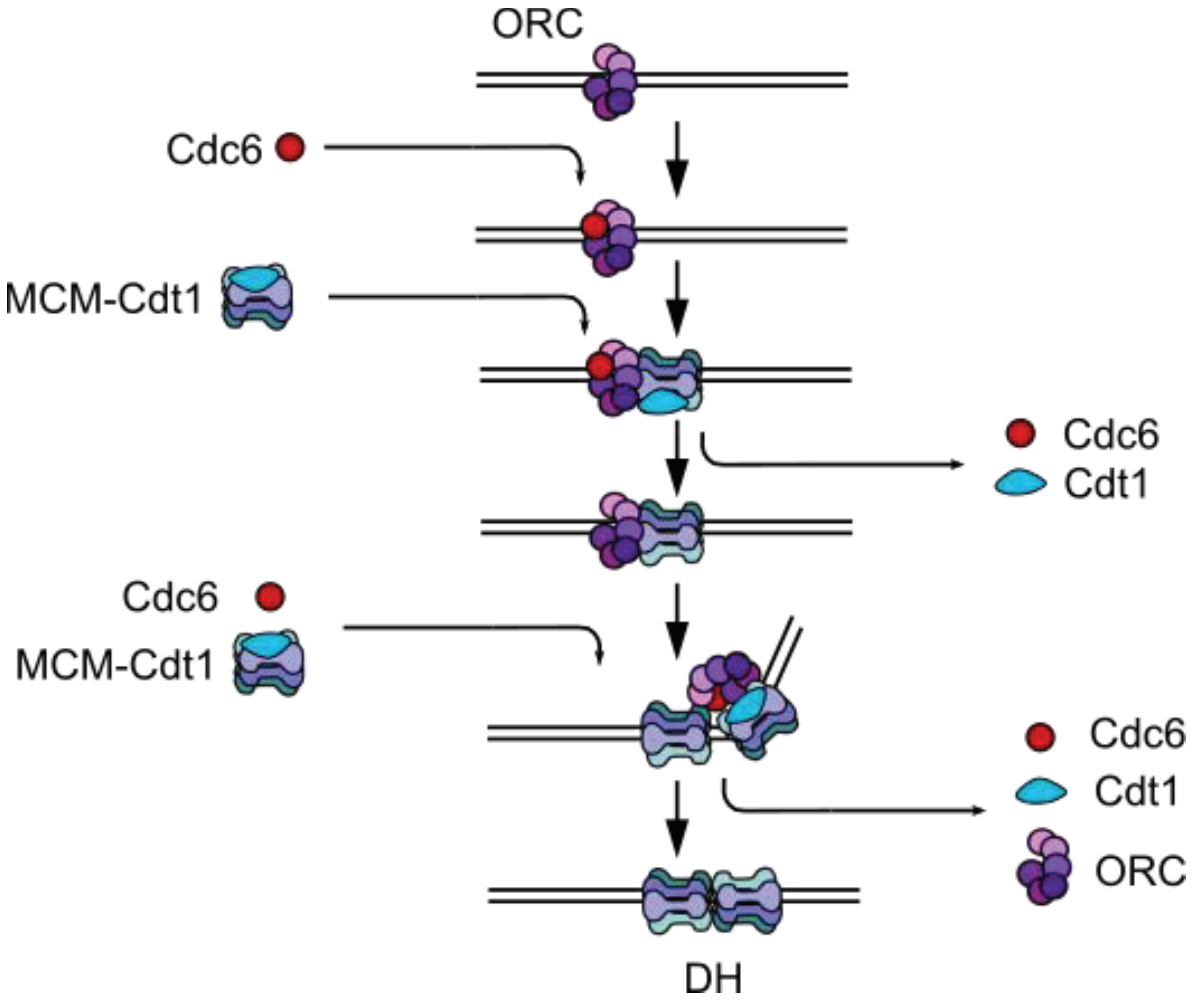


Figure 5: Summary of the key steps in loading of the MCM2-7 helicase at origins of replication . Initially, only ORC is bound. Following Cdc6 binding, ORC-Cdc6 is competent to recruit a singular complex of MCM-Cdt1 and insert DNA within the MCM2-7 central channel. Following ATP hydrolysis, Cdc6 and Cdt1 are released, resulting in MCM2-7 ring closure (Ticau et al., 2017). Release of ORC from one side of the MCM2-7 hexamer and an "ORC flip" (Miller et al., 2019; Gupta et al., 2021) followed by binding of a second Cdc6 to ORC, which allows recruitment of a second MCM2-7-Cdt1. Finally, a second round of ATP hydrolysis occurs that releases the second Cdc6 and Cdt1, resulting in stable MCM2-7 double hexamer formation (DH) at the origin. Figure adapted from and mechanism reviewed in Riera et al., 2017(Riera et al., 2017).

1.4 Initiation of DNA Replication: Firing

After DH loading origins are licensed and DNA replication activation can commence. The activation of the MCM2-7 helicase occurs at the transition from G1 to S phase and is therefore temporally separated from the loading process (Fernandez-Cid et al., 2013; Randell et al., 2006; Ticau et al., 2015, 2017). This precludes further helicase loading and potential re-replication at an activated origin (Dahmann, Diffley & Nasmyth, 1995).

Recently, the minimal yeast DNA replication initiation machinery was reconstituted *in vitro* (Yeeles et al., 2015). This narrowed down the list of factors required *in vitro* for this process to sixteen protein components: Cdt1/MCM2-7, ORC, Cdc6, DDK, CDK, Sld2, Sld3/7, Cdc45, Dpb11, GINS, Polε, Mcm10, Polα, Ctf4, RPA and TopII. However, this produces only the minimally active replication machinery, and additional factors are required for increased replicative activity (Yeeles et al., 2015), and even more factors may be required for efficiency in the same processes *in vivo*.

1.4.1 Kinase activity drives the activation mechanism

Triggering the assembly of the helicase complex requires two kinases, the Dbf4-dependent kinase (DDK) and the cyclin-dependent kinase (CDK) (Diffley, 1996; Hartwell et al., 1973; Nasmyth, 1996; Stillman, 1996). The crucial function of DDK is the phosphorylation of the MCM2-7 double hexamer, as initially identified by suppressor mutants (Francis et al., 2009; Hardy, 1997; Randell et al., 2010; Sheu & Stillman, 2006). Similarly, CDK has essential targets for this process in Sld2 and Sld3, though suppressor mutants in this case are not as effective, consistent with other essential roles of CDK (Tanaka et al., 2007; Zegerman & Diffley, 2007).

Phosphorylation of the MCM2-7 double hexamer by DDK serves as the first step in helicase activation. DDK phosphorylation is stimulated by the formation of the double hexamer, indicating that the formation of this structure is a prerequisite for further activation (Francis et al., 2009; Sun et al., 2014). Mcm4 and Mcm6 are phosphorylated by DDK at sites located on long unstructured tails (Ali et al., 2017; Sheu & Stillman, 2006, 2010). This phosphorylation drives the recruitment of Cdc45 and Sld3 to the double hexamer (Heller et al., 2011; Kamimura et al., 2001; Kanemaki & Labib, 2006; Natsume et al., 2013; Tanaka et al., 2011). *In vitro* data suggests that Sld3 recruits Cdc45 to double hexamers (Deegan, Yeeles & Diffley, 2016; Fang, Cao & Lou, 2016), something not recapitulated *in vivo,* where recruitment of Sld3 and Cdc45 is interdependent (Heller et al., 2011; Kamimura et al., 2001; Kanemaki &

Labib, 2006), and the non-essential Sld7 (Heller et al., 2011; Kamimura et al., 2001; Kanemaki & Labib, 2006) binds Sld3 conferring stability to this protein (Heller et al., 2011; Kamimura et al., 2001; Kanemaki & Labib, 2006).

The phosphorylation of Sld2 and Sld3 by CDK, similar to DDK phosphorylation of MCM2-7, also advances complex formation. Phosphorylated Sld2 and Sld3 are able to bind the BRCA1 C-terminal (BRCT) repeats in Dpb11 (Tanaka et al., 2007; Zegerman & Diffley, 2007). This induces the formation of a metastable preloading complex (pre-LC) (Muramatsu et al., 2010) comprised of Sld2 and Dpb11, along with DNA polymerase epsilon (Polε) and the four components of the go-ichi-ni-san (GINS) complex: Sld5, Psf1, Psf2 and Psf3 (Muramatsu et al., 2010). This complex is recruited to the MCM2-7 double hexamer via an interaction with phosphorylated Sld3, before Sld2, Sld3 and Dpb11 are released, forming a Cdc45/MCM2-7/GINS complex (CMG) (Heller et al., 2011; Kanemaki & Labib, 2006). The further involvement of Mcm10 with the nascent CMG seems to allow origin unwinding activity, since it facilitates the recruitment of the single-stranded DNA binding protein RPA and DNA polymerases beyond Polε, such as Polα (Heller et al., 2011; Kanemaki & Labib, 2006). The CMG is thereby activated and regarded as being the core active complex for further DNA unwinding in S phase (Gambus et al., 2006; Moyer, Lewis & Botchan, 2006).

1.4.2 Sld3/7 and Cdc45: Earliest initiation factors

1.4.2.1 Sld3/7

Sld3 was first identified in a screen for suppressors of a replication mutant (Kamimura et al., 2001). The Sld3 homolog in higher eukaryotes was non-obvious but is now widely considered to be Treslin/ticrr, due to a region of homology and similarity of function (Mueller, Keaton & Dutta, 2011; Sanchez-Pulido, Diffley & Ponting, 2010; Kumagai et al., 2010; Sansam et al., 2010; Boos et al., 2011). Sld3 is essential for viability, and shows genetic and physical

interactions with Cdc45, as well as similar genomic association patterns (Kamimura et al., 2001; Kanemaki & Labib, 2006). Sld3 and Cdc45 co-associate with origins in S-phase, and origin firing timing seems highly dependent on these early binding factors (Fang, Cao & Lou, 2016; Tanaka et al., 2011; Deegan, Yeeles & Diffley, 2016).

Sld3 is bound throughout the cell cycle by Sld7, a protein that, whilst not essential for cell viability, nonetheless promotes growth and DNA replication (Tanaka et al., 2011). Alone, Sld7 shows limited binding to MCM2-7, and is thought to stabilise Sld3, or perhaps enable higher-order dimerization of Sld3/7 that may play some role in coordinating signals across a double hexamer (Deegan, Yeeles & Diffley, 2016; Tanaka et al., 2011; Itou, Shirakihara & Araki, 2015).

1.4.2.2 Cdc45

The Cdc45 protein has a long history of association with cell cycle processes. It was first identified in a screen for cold sensitive mutants that generated a cell cycle arrest phenotype (Moir et al., 1982). It was subsequently discovered to associate with the MCM2-7 complex and promote both initiation and continued replication of DNA (Zou, Mitchell & Stillman, 1997; Hopwood & Dalton, 1996; Aparicio, Weinstein & Bell, 1997), and this was a highly conserved function, with several species' homologs seemingly functioning in the same manner (Mimura et al., 2000; Mimura & Takisawa, 1998; Miyake & Yamashita, 1998; Hopwood & Dalton, 1996). Cdc45 is essential for initial unwinding at origins (Walter & Newport, 2000; Mimura et al., 2000; Pacek & Walter, 2004), and critical for further assembly of replisome components such as polymerases (Zou & Stillman, 2000; Uchiyama, Arai & Masai, 2001; Mimura et al., 2000; Mimura & Takisawa, 1998). From compositional and structural studies it is also known that Cdc45 is a core component of the replisome as it moves through the genome (Langston et al., 2014; Costa et al., 2014; Gambus et al., 2006; Costa et al., 2011; Sun et al., 2015), and its continued association is essential for progression of replication forks (Pacek & Walter, 2004;

Tercero, Labib & Diffley, 2000). The exact timing of Cdc45 recruitment to origins and replication complexes was of some debate initially, with reports of association of the protein with chromatin in G1 (Aparicio, Stout & Bell, 1999; Aparicio, Weinstein & Bell, 1997) or other data showing association of Cdc45 only after S phase had begun (Zou & Stillman, 1998). More recent data found that Cdc45 can be found bound to origins in alpha factor arrest (Miyazawa-Onami, Araki & Tanaka, 2017). However, what is certain is that Cdc45 recruitment is dependent on both the activity of the kinase Cdc7-Dbf4 (DDK) (Jares & Blow, 2000; Deegan, Yeeles & Diffley, 2016; Heller et al., 2011; Sheu & Stillman, 2006; Tanaka et al., 2011; Yeeles et al., 2015) and co- or pre- recruitment of Sld3/7, with which Cdc45 forms a complex bound to the MCM2-7 double hexamer (Kamimura et al., 2001; Yamada, Nakagawa & Masukata, 2004; Deegan, Yeeles & Diffley, 2016; Tanaka et al., 2011), and that Cdc45 is one of the earliest factors to associate during activation (Miyazawa-Onami, Araki & Tanaka, 2017; Heller et al., 2011).

Structural work on the active replisome has revealed that, along with the GINS complex, Cdc45 occupies a position near the N terminus in the CMG, covering the Mcm2-Mcm5 gate (Costa et al., 2011). Whilst some studies have suggested that Cdc45 uses its weak DNA binding activity to guide the lagging strand (Costa et al., 2011), or in fact act as the “wedge” that separates the DNA strands (Szambowska et al., 2014), current models suggest that Cdc45, in tandem with GINS, is positioned over the DNA entry gate to control this opening in the hexamer ring, and prevent accidental ejection of the leading strand, activities which may be important during replication stress and fork stalling (Petojevic et al., 2015).

An inherent puzzle for consideration is that Cdc45 is the eukaryotic homolog of the RecJ nuclease (Sanchez-Pulido & Ponting, 2011). Whilst the Cdc45 protein has lost the catalytic nuclease function (Krastanova et al., 2012; Pellegrini, 2017), it is interesting that a protein of

such different function may have evolved to form such an essential part of the replication machinery in eukaryotes.

1.4.3 Pre-LC components Dpb11, GINS, Polε and Sld2: Intermediate initiation factors

As part of the second CDK-driven step of helicase activation, several additional factors past the initial Cdc45 and Sld3/7 are recruited to the double hexamer in a phosphorylation dependent manner (Figure 6). These factors comprise the multi-factor subcomplex known as the pre-loading complex (pre-LC) (Masumoto et al., 2002; Tanaka et al., 2007; Zegerman & Diffley, 2007).

1.4.3.1 Sld2

In addition to phosphorylating Sld3, CDK also phosphorylates the protein Sld2 (Masumoto et al., 2002). Sld2 was one of several proteins found to cause synthetic lethality with a thermosensitive Dpb11-1 mutation, and it was established that cells without functional Sld2 present were inviable (Kamimura et al., 1998). Otherwise, not much is known about the structure or function of the Sld2 protein, other than an indication that this protein might bind ssDNA when phosphorylated and possess strand annealing activity (Kanter & Kaplan, 2011), though further confirmation is required. Intriguingly, significant homology to Sld2 is found within the mammalian RecQ4 helicase-like protein, which also shows ssDNA annealing activity, in addition to its helicase activity (Macris et al., 2006; Xu & Liu, 2009; Rogers et al., 2017). Given that production of ssDNA as part of duplex unwinding is expected during the transition to active replisomes, this activity is both indicative and puzzling. Single-stranded binding can be rationalised, but the possible annealing activity has yet to be placed within any likely suggested mechanism.

1.4.3.2 Dpb11

Dpb11 is another of the factors that are recruited to the replication complex by the action of CDK (Muramatsu et al., 2010). This protein was identified initially by genetic interactions with mutations of the polymerase ε subunits Pol2 and Dpb2 (Araki et al., 1995). Dpb11 appears to function in a scaffold-like manner, and it contains four BRCT domains, which are shown to have roles in signalling, and display specificity for phosphorylated peptides (Figure 6). Of the four BRCT domains, there are two pairs (Muramatsu et al., 2010). One pair is required for binding to Sld3 for origin recruitment, and the other binds to Sld2 (Masumoto et al., 2002; Tanaka et al., 2007; Zegerman & Diffley, 2007; Muramatsu et al., 2010). An additional interaction domain is present between these pairs of domains, and it is this region that shows binding to the GINS subcomplex, specifically the Psf1 subunit (Tanaka et al., 2013), and this interaction was required for the recruitment of both proteins to origins (Takayama et al., 2003; Muramatsu et al., 2010; Miyazawa-Onami, Araki & Tanaka, 2017). Dpb11 also interacts physically with polymerase ε subunits, which potentially underlies the synthetic lethality observed with Pol2 and Dpb2 subunits .(Masumoto, Sugino & Araki, 2000; Araki et al., 1995). Outside of its role in replication initiation, Dpb11 also functions as a scaffold in processes of DNA repair (Puddu et al., 2008; Pfander & Diffley, 2011; Liu et al., 2017).

1.4.3.3 GINS

Like many of the core replication factors, GINS was also identified in screening for synthetic lethality with Dpb11 mutations (Kamimura et al., 1998; Takayama et al., 2003). In this case, the protein encoded by the SLD5 gene was found to be required for cellular viability, and three interacting partners were also identified and named Psf1, Psf2 and Psf3 (Takayama et al., 2003). As described, the initial association of GINS with origins is dependent on Dpb11 interaction (Takayama et al., 2003; Muramatsu et al., 2010; Miyazawa-Onami, Araki & Tanaka, 2017), but GINS is another initiation factor, like Cdc45, that stays associated with the

helicase over the course of replication as part of the CMG (Gambus et al., 2006), and couples the association of multiple other proteins to the complex during this time, such as Cdc45, Ctf4 and polymerase ε (Gambus et al., 2006; Costa et al., 2011; Villa et al., 2016; Simon et al., 2014; Gambus et al., 2009; Jones et al., 2021; Baretić et al., 2020; Goswami et al., 2018). Structurally, the GINS complex forms a diamond-shape tetramer, composed mainly of α-helices . Mechanistically, it is still not understood exactly the role that GINS plays in initiation or active replication, but it is likely that the positioning of the complex over the Mcm2-Mcm5 subunit interface is important. By blocking the interface most liable to open in the ring complex, GINS may be stabilising the closed ring structure, or preventing loss of leading strand DNA from the central channel should the gate open, either as part of normal translocation or due to bypass of obstruction (Costa et al., 2011; Yuan et al., 2016; Petojevic et al., 2015).

1.4.3.4 Polε

The functional role of polymerase epsilon in catalysing leading strand synthesis will be discussed below, though this protein is also a key component of the initiation process (Muramatsu et al., 2010). Polymerase ε is a large subcomplex, with its largest catalytic subunit Pol2, which shows an extended conformation, containing a polymerase domain in the N-terminus, and an extended C-terminal domain (Asturias et al., 2005; Pursell & Kunkel, 2008). Whilst the catalytic N-terminus was dispensable for cell viability, the C-terminal domain was found to have an essential function in structuring the replisome (Kesti et al., 1999; Dua, Levy & Campbell, 1999; Goswami et al., 2018). A combination of the C-terminal domain and Dpb2 was found to be sufficient for *in vitro* initiation, and so the extensive contacts that these subunits form with the MCM2-7 C-terminus, and with GINS, are likely important in the overall mechanism of replication initiation (Goswami et al., 2018). Investigation of complex formation at the origin by limiting activation using degron-mediated selective factor degradation found that although removal of Pol2 prevented activation of DNA replication, it was the loss of Dpb2

that prevented the association of other members of the pre-LC to the origin, suggesting this is the factor required for the association of pre-LC and immediate continuation of initiation (Miyazawa-Onami, Araki & Tanaka, 2017).

1.4.3.5 The Pre-LC

The combination of the above protein factors is known as the pre-LC, and is thought to assemble as a complex in response to CDK activity (Muramatsu et al., 2010). The association of this complex has been shown to be interdependent, and some evidence suggests that this complex association is "switch-like" and drives the splitting of the double hexamer in an irreversible step of replication initiation (Miyazawa-Onami, Araki & Tanaka, 2017). The splitting of the double hexamer prior to the action of the last major initiation factor, Mcm10, is

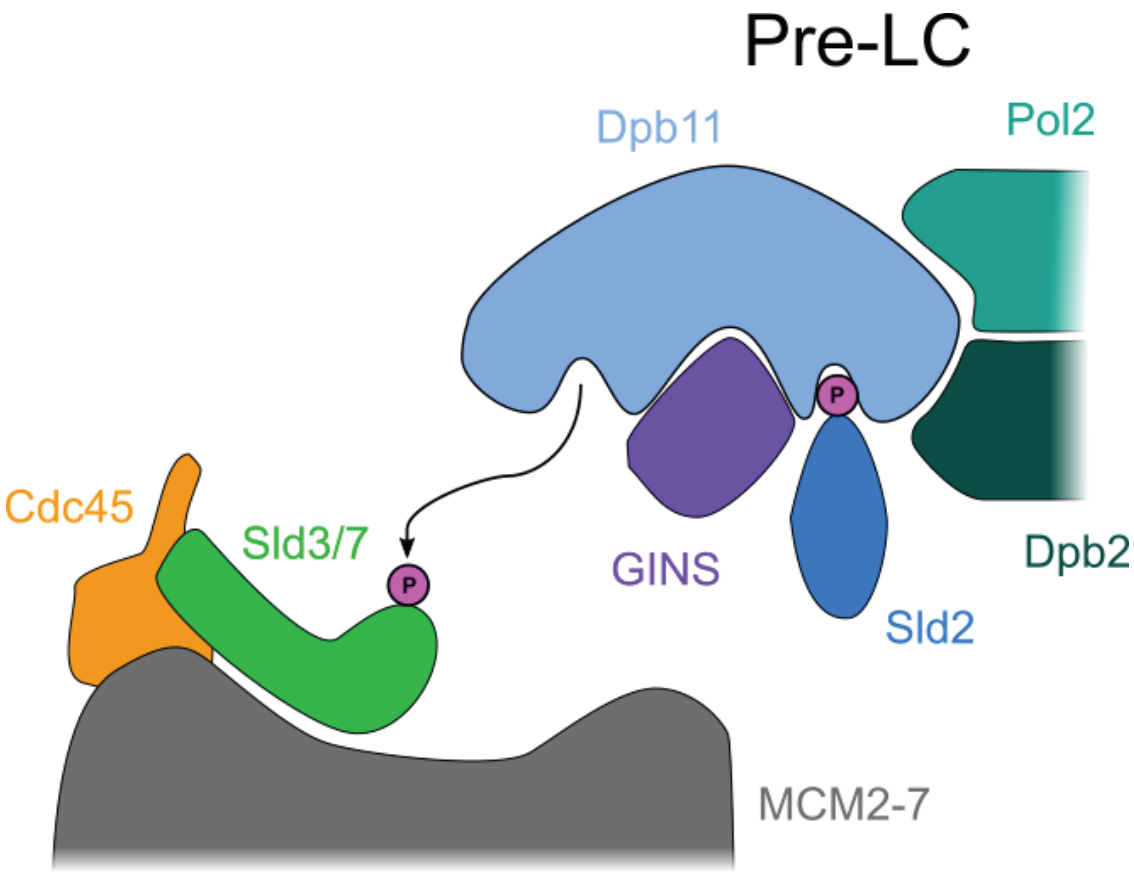


Figure 6: Protein-protein interactions of the Pre-LC complex during replication initiation . The two pairs of BRCT domains bind to phosphorylated peptides of Sld2 (within the pre-LC) and Sld3 (for pre-LC recruitment to origins). Dpb11 and GINS bind in a region between the BRCT domains. Dpb11 also shows binding to Pol2 and Dpb2 subunits of polymerase ε, and the complex associates with the replication complex during initiation in response to S-CDK phosphorylation (Araki et al., 1995; Masumoto, Sugino & Araki, 2000; Tak et al., 2006; Tanaka et al., 2007; Zegerman & Diffley, 2007; Muramatsu et al., 2010; Tanaka et al., 2013)

also suggested by particle analysis of *in vitro* initiation reactions in the absence of Mcm10, that show two highly separated hexamers after the concerted action of all previous replication factors (Douglas et al., 2018).

1.4.4 Mcm10: Final initiation factor

Mcm10 was initially identified as another factor that showed temperature sensitive alleles that affected replication and was shown to be essential for the transition into S phase and the assembly and movement of replication complexes away from origins (Dumas et al., 1982; Maine, Sinha & Tye, 1984; Solomon et al., 1992; Merchant et al., 1997; Kawasaki, Hiraga & Sugino, 2000; Gregan et al., 2003; Sawyer et al., 2004). Mcm10 shows conservation between eukaryotes in both sequence and function, though there is no homolog in prokaryotes (Du, Stauffer & Eichman, 2012). In terms of structural information, it is unclear what form Mcm10 takes *in vivo*, as evidence of multiple different homomeric and single unit forms of this protein have been produced, including a ring-shaped hexamer form, reviewed in Du, Stauffer and Eichman (Du, Stauffer & Eichman, 2012). The body of evidence now suggests that Mcm10 is a component of the travelling replisome (Ricke & Bielinsky, 2004; Das-Bradoo, Ricke & Bielinsky, 2006; Warren et al., 2009; Gambus et al., 2009; Lee et al., 2010), and is a functional element enabling replisome stimulation, the passage of replication barriers, and preventing fork regression at stalled replisomes (Lõoke, Maloney & Bell, 2017; Langston et al., 2017; Mayle et al., 2019). However, Mcm10 also has a crucial role at the initiation of DNA replication (Merchant et al., 1997; Homesley et al., 2000).

It has been shown that without Mcm10, an activated complex is able to form during replication initiation, with the association of multiple firing factors (Miyazawa-Onami et al., 2017), and a minimal amount of DNA unwinding *in vitro* (Douglas et al., 2018). These 0.6-0.7 helical turns of DNA per hexamer, is an amount easily constrained by the approximate 30 base pair footprint

of an individual hexamer, and so may not be extruded from the complex. Interestingly, several firing factors are recruited without Mcm10, and data shows that RPA does not associate with origins under these conditions (van Deursen et al., 2012; Watase, Takisawa & Kanemaki, 2012). This suggests that single-stranded DNA (ssDNA) extrusion from the helicase is limited or non-existent at this stage. In addition, it was shown that in the absence of Mcm10, both *in vivo* and *in vitro*, double hexamers appear to be separated (Miyazawa-Onami et al., 2017; Douglas et al., 2018). Drawing these conclusions from the *in vitro* data is susceptible to criticism , however, as separation was observed in samples prepared for cryo-EM under harsh washing conditions. With Mcm10 included, further unwinding occurs, and the complex becomes fully active for DNA replication (Miyazawa-Onami et al., 2017; Douglas et al., 2018).

1.4.5 Helicase activation summary

Yeeles et al.'s reconstitution of a minimal *in vitro* yeast replication assay identified many of the key factors required for DNA replication in this organism (Yeeles et al., 2015). However, whilst the entire minimal complement of replication factors required *in vitro* is now accounted for, there is still little evidence pertaining to the structural or mechanistic events that occur during the initiation process. Indeed, this system also lacks the ability to investigate additional *in vivo* regulatory or accessory factors that may be required. Instead, recent studies, some of which have been greatly aided by the reconstituted system (Douglas et al., 2018), and some pursued separately (Miyazawa-Onami, Araki & Tanaka, 2017), are currently the best insights into this unknown mechanism.

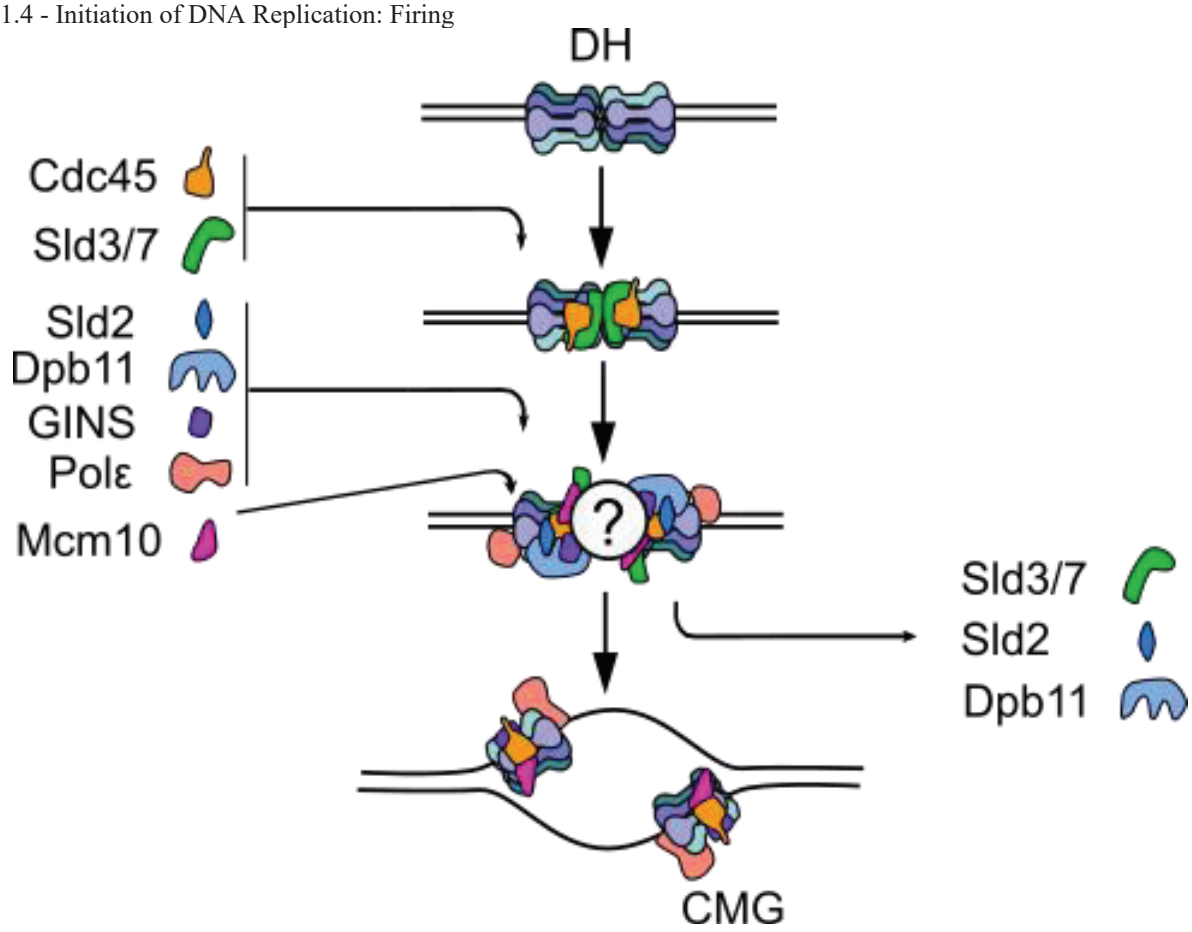


Figure 7: Summary of the process of helicase activation, based on the structural and biochemical information at the current time. Following DH loading at origins, a DDK-dependent activation step phosphorylates MCM2-7 N-termini (Labib, 2010), and causes the recruitment of Cdc45 and Sld3/7 (Heller et al., 2011; Tanaka et al., 2011; Deegan, Yeeles & Diffley, 2016). Following the association of these factors, a CDK-dependent step causes association of the pre-LC factors Sld2, Dpb11, GINS, and polε (not shown) (Masumoto et al., 2002; Tanaka et al., 2007; Zegerman & Diffley, 2007; Muramatsu et al., 2010; Miyazawa-Onami, Araki & Tanaka, 2017). This step appears to cause structural changes that separate the DH in some manner (Miyazawa-Onami, Araki & Tanaka, 2017; Douglas et al., 2018), and finally, full helicase activation seems dependent on Mcm10 association (van Deursen et al., 2012). Figure adapted from Riera et al., 2017(Riera et al., 2017).

1.5 The Composition, Structure and Mechanism of the Active Replisome

The active replisome is a large molecular machine composed of many component parts, the overall structure of the replisome at replication forks is highly conserved among eukaryotes, as are the subcomplexes and many of the individual proteins (Figure 8)(Yao & O'Donnell, 2010). The core DNA unwinding activity is carried out by the CMG complex, but there are a wide range of additional functions required as part of the replicative process, that are also associated with the replication fork (Yao & O'Donnell, 2010; Burgers & Kunkel, 2017).

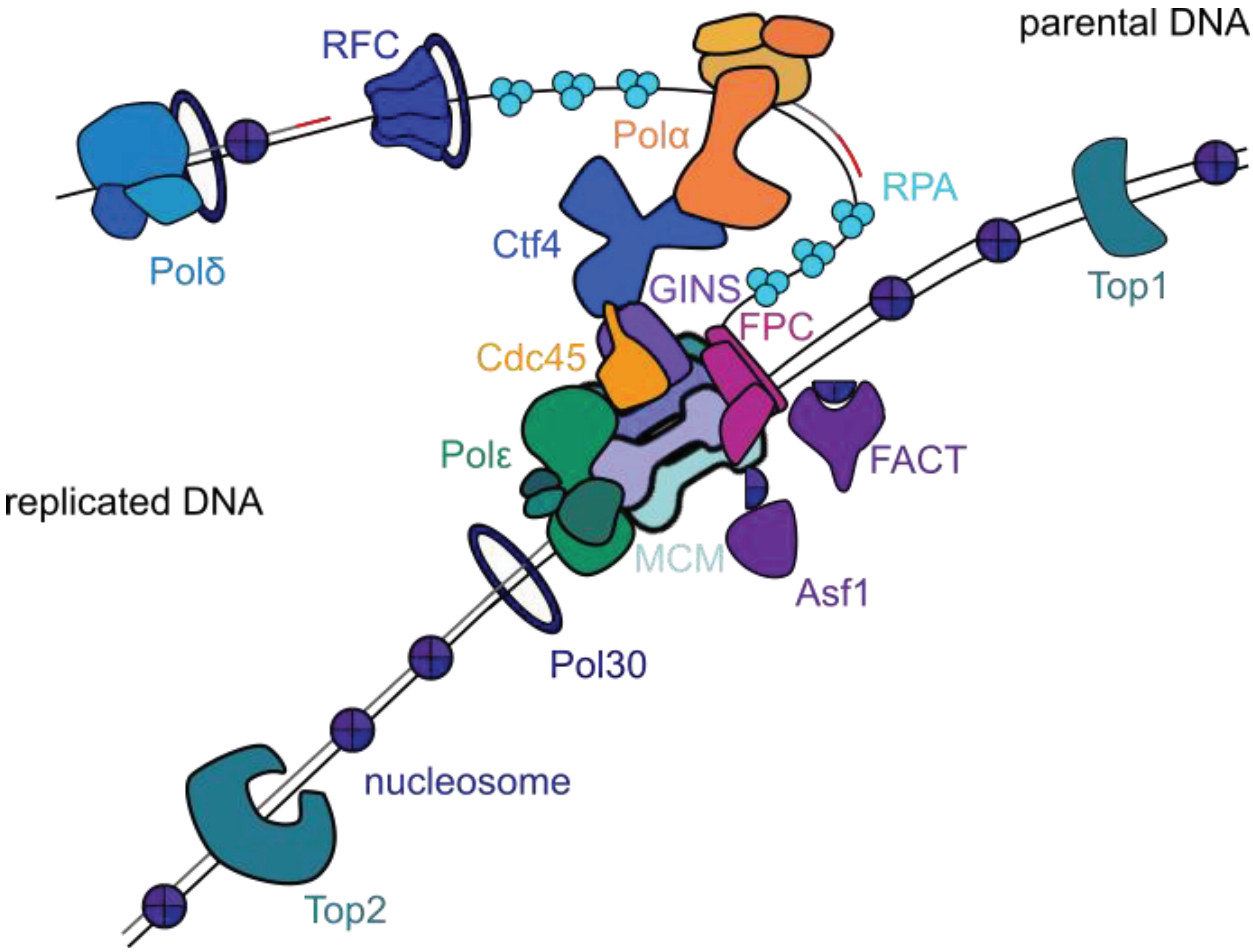


Figure 8: Overview of the structure of the active fork. The CMG complex forms the catalytic nexus of the fork, associated closely with the fork protection complex (Baretić et al., 2020) polymerase ε (Sun et al., 2015) and Ctf4 (Gambus et al., 2009; Baretić et al., 2020; Jones et al., 2021). Priming and extension of lagging strand is accomplished by polymerase α coupled to the replisome by Ctf4 (Zhu et al., 2007; Gambus et al., 2009), before polymerase δ completes Okazaki fragment synthesis (Kunkel & Burgers, 2008). On the leading strand, polymerase ε catalyses leading strand synthesis (Kunkel & Burgers, 2008). Top1 and Top2 relieve topological strain ahead of, and behind the fork, respectively (Wang, 2002). Nucleosome binding and assembly factors associate with MCM2-7 at the fork, and with a range of factors behind the fork, including the clamp and clamp loader, Pol30 (Boehm, Gildenberg & Washington, 2016) and RFC (Yao & O'Donnell, 2010), respectively . Single-stranded DNA is bound by RPA on the lagging strand where it is present (Wold, 1997).Figure adapted from Yao and O'Donnell 2010, Burgers and Kunkel 2017, with structural information from Jones et al. 2021 (Yao & O'Donnell, 2010; Burgers & Kunkel, 2017; Jones et al., 2021).

1.5.1 The Core CMG complex

As mentioned above, the DNA helicase activity of MCM2-7 is enhanced by binding of Cdc45 and the GINS complex (Ilves et al., 2010). This core of Cdc45-MCM2-7-GINS (CMG) forms a stable complex, and human and *Drosophila* CMG can be isolated as a distinct complex that possesses helicase activity (Moyer, Lewis & Botchan, 2006; Kang et al., 2012). Purification of

the yeast complex and subsequent *in vitro* analysis demonstrated that this activity is well conserved across eukaryotes (Langston et al., 2014).

The structure of the CMG core complex from both yeast and higher eukaryotes has been well established, and some very high-resolution structures are available that clearly show the binding of GINS and Cdc45 to the side of the MCM2-7, positioned over the Mcm2/Mcm5 interface (Costa et al., 2011; Yuan et al., 2016; Georgescu et al., 2017; Baretić et al., 2020; Jones et al., 2021).

Structural and biochemical studies of the CMG helicase complex have sought to establish the underlying mechanism. Most of the studies cited above that elucidated the possible mechanism of the MCM2-7 helicase in unwinding were carried out on the CMG helicase as the active form of the complex (Costa et al., 2011; Abid Ali & Costa, 2016; Abid Ali et al., 2016; Yuan et al., 2016). Whilst the full mechanistic description of DNA unwinding and translocation by the replicative helicase is still not complete, recent structural and single molecule data are getting closer to revealing the intermediate steps in the sequence of these processes.

However, it is well established that the CMG is by no means the sole component of replication forks (Yu et al., 2014; Gambus et al., 2006; Katou et al., 2003; Nedelcheva et al., 2005; Bermudez et al., 2011; Sengupta et al., 2013; Ricke & Bielinsky, 2004), and a large body of evidence shows that this functional helicase unit can also act as a scaffold and interaction partner to a much wider range of proteins that form the complex known as the replisome progression complex (RPC)(Gambus et al., 2009; Georgescu et al., 2015; O'Donnell & Li, 2016; Tanaka et al., 2009).

1.5.1.1 MCM mechanism of action as part of the CMG

Multiple different proposed mechanisms of DNA unwinding and helicase orientation have been suggested for the MCM2-7 helicase. Recent structural and single molecule analyses have started to resolve the dispute around which of these mechanisms are correct. It was suggested

early on, and later confirmed, that the likeliest conformation of the replicative helicase involved two fork complexes that moved along DNA at two individual forks that arise from the same origin (Enemark & Joshua-Tor, 2006; Botchan & Berger, 2010; Yardimci et al., 2010). Initially, evidence appeared to suggest a C-terminal-first orientation of the replicative helicase (McGeoch et al., 2005; McGeoch & Bell, 2005; Rothenberg et al., 2007; Costa et al., 2014). However, an N-terminal-first orientation was in fact more likely in yeast and higher eukaryotes (Douglas et al., 2018; Georgescu et al., 2017; Yuan et al., 2020; Baretić et al., 2020; Jones et al., 2021).

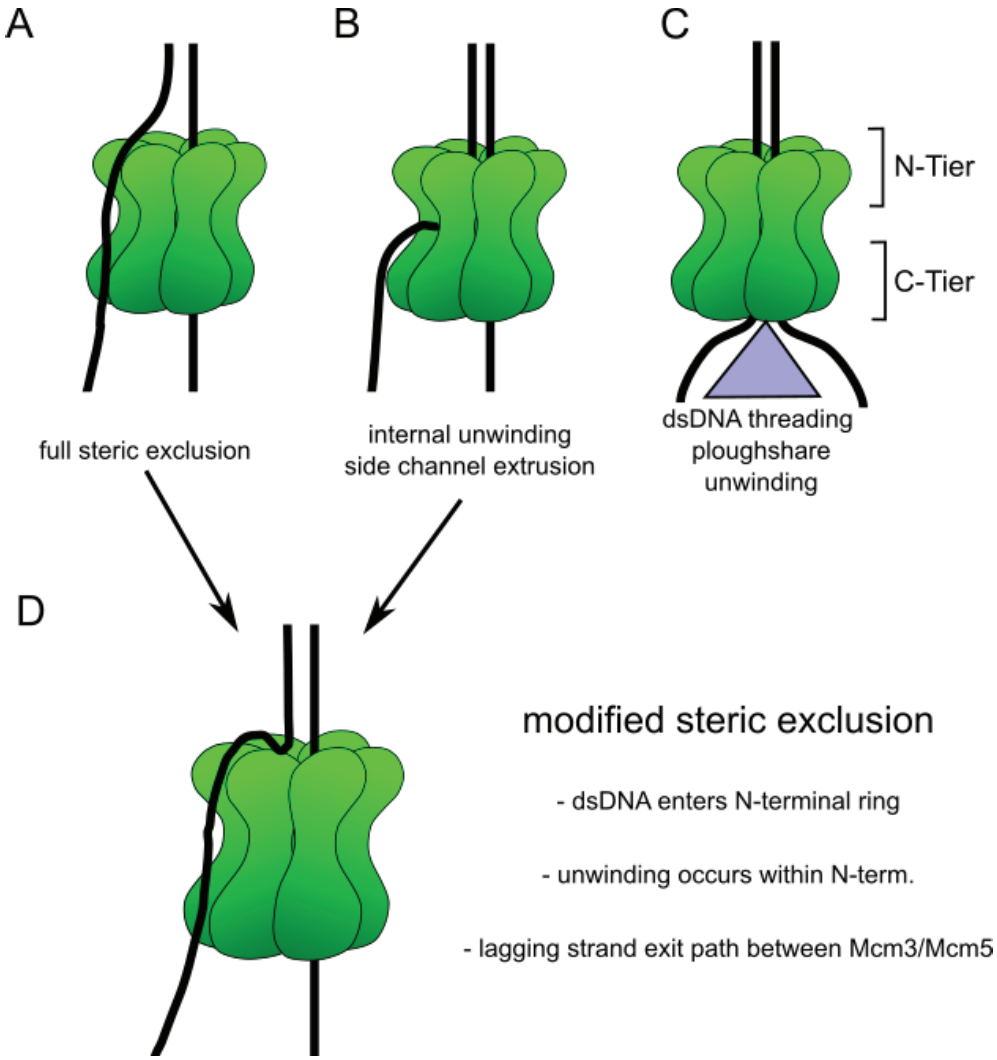


Figure 9: Proposed mechanisms for DNA unwinding of the MCM2-7 helicase in eukaryotes. A) Steric exclusion model, as evidenced by many other replicative helicases. B) Alternative internal unwinding and side channel extrusion model proposed from early CMG structures (Costa et al., 2011). C) "Ploughshare" model proposing dsDNA unwinding after MCM2-7 translocation by a separate subcomplex, initially suggested to be GINS/Cdc45 (Takahashi, Wigley & Walter, 2005). D) Most likely current model for DNA unwinding and lagging strand bypass, a modified steric exclusion model with internal unwinding taking place within the periphery of the N-terminus (Georgescu et al., 2017; O'Donnell & Li, 2018; Baretić et al., 2020).

The helicase was also suggested to encircle single-stranded lagging strand DNA (Fu et al., 2011), but the exact mechanism of DNA unwinding was still undetermined (Figure 9). The most recent analysis of structures and biochemical assays has also begun to reach a consensus, where a steric exclusion model of DNA unwinding appears to be the most likely mechanism, similar to other replicative helicases from other species (Jeong, Levin & Patel, 2004; Fu et al., 2011; Yardimci et al., 2012; Georgescu et al., 2017; O'Donnell & Li, 2018). High resolution structural data of helicase bound to the fork suggested a lagging strand exit path (Georgescu et al., 2017), and more recent structural data of the replisome bound to a forked DNA template has further elucidated this possible path of the lagging strand at the N-terminus of the helicase (Baretić et al., 2020). This structure identifies a possible separation pin in the N-terminal hairpin of Mcm7 that initiates the strand separation, and interactions between the N-terminal hairpins of Mcm6 and Mcm4 that guide the lagging strand between the zinc-finger domains of Mcm3 and Mcm5 and exiting the central channel of the helicase (Yuan et al., 2020). Overall, the evidence supports a modified steric exclusion model for MCM2-7, where the lagging strand initially enters the channel of the MCM2-7 formed by the N-terminal hairpins and zinc finger domains but is ultimately diverted after dsDNA unwinding very near to the N-terminus of the helicase (Georgescu et al., 2017; O'Donnell & Li, 2018; Baretić et al., 2020). This explains the slowing but eventual bypass of helicase past lagging strand blocks, but much reduced ability to bypass blocks on the leading strand (Fu et al., 2011; Yardimci et al., 2012; O'Donnell & Li, 2018).

Recent structural and biochemical data has shown no sign of any additional exit channel for the lagging strand from within the central channel further down to support full internal unwinding and side channel extrusion (Georgescu et al., 2017; Petojevic et al., 2015), and the tracking of the helicase on ssDNA coupled with the binding of the proposed Cdc45/GINS "ploughshare" (Takahashi, Wigley & Walter, 2005) to the side of the helicase hexamer and not

behind the mode of translocation effectively disprove the dsDNA translocation and accessory "ploughshare" hypothesis (Georgescu et al., 2017; Douglas et al., 2018).

Whilst the mechanism of DNA unwinding itself has become more settled, there is still very little conclusive evidence for how the ATPase activity of the MCM2-7 helicase is coupled to translocation. Recent studies of archaeal MCM have provided further evidence that translocation in these complexes is driven by rotary activity, and that ssDNA-MCM interactions are consistent around the hexamer ring and carried out by pre-sensor-1-β-hairpin (ps1β) and helix-2-insert (h2i) elements within the near-centre of the hexamer's length in a so-called "spiral staircase" (Meagher, Epling & Enemark, 2019). However, this study was carried out using the homo-hexameric archaeal MCM protein complex, and so may not reflect the eukaryotic MCM2-7 complex, where there are many reported differences between subunits in structure and functionality. Additionally, homo-hexameric models are not ideal since a cycling rotary process cannot be seen processing across the subunits as they are all identical (Abid Ali & Costa, 2016). Studies of MCM2-7 structural intermediates originally suggested that the limited intermediate states of MCM2-7 identifiable supported a concerted mechanism of ATPase activity induced an "inchworm" or "pumpjack" model of DNA translocation, which would cause movement of the N- and C-rings of the hexamer from a compressed to open conformation, and hence cause translocation along DNA (Costa et al., 2011; Abid Ali et al., 2016; Yuan et al., 2016). Recent single molecule experiments are consistent with these observations of asymmetry within the subunits of MCM2-7 (Burnham et al., 2019). However, the most recent cryo-EM analysis of multiple intermediates of a translocating hexamer did in fact show variability in the subunits interacting with the ssDNA in the central channel, and comparison of the different states does in fact suggest the use of a non-symmetrical rotary model for ssDNA translocation of the MCM2-7 helicase (Eickhoff et al., 2019).

The MCM2-7 helicase has a poor activity on its own, and requires association of other factors to increase its ATPase activity to levels required for DNA replication in cells (Miyazawa-Onami et al., 2017; Douglas et al., 2018). The active form of the MCM2-7 helicase is therefore the CMG complex found within the replisome progression complex.

1.5.2 The Eukaryotic DNA-directed DNA Polymerases

In eukaryotes, the machinery for the processive addition of nucleotides to single-stranded DNA (ssDNA) is carried out by three distinct and evolutionarily conserved holoenzyme polymerases. They are polymerase α, polymerase δ and polymerase ε (Figure 10). The composition of these multi-subunit complexes is very similar when comparing from *S. cerevisiae* to *H. sapiens*, with the exception of a fourth subunit of polymerase δ that is not present in budding yeast. Each of the three has different activities and roles at the active fork (Johansson & MacNeill, 2010).

The discovery of discontinuous synthesis of the lagging strand by Okazaki et al., highlighted an intrinsic asymmetry of the replication fork. The synthesis direction of the polymerases is restricted to a 5’ to 3’ direction, requiring the staggered synthesis of small fragments on the lagging strand. This, in turn, requires the synthesis and extension of primers on the lagging strand, an activity required only once on the leading strand (Bebenek & Kunkel, 2004; Rothwell & Waksman, 2005).

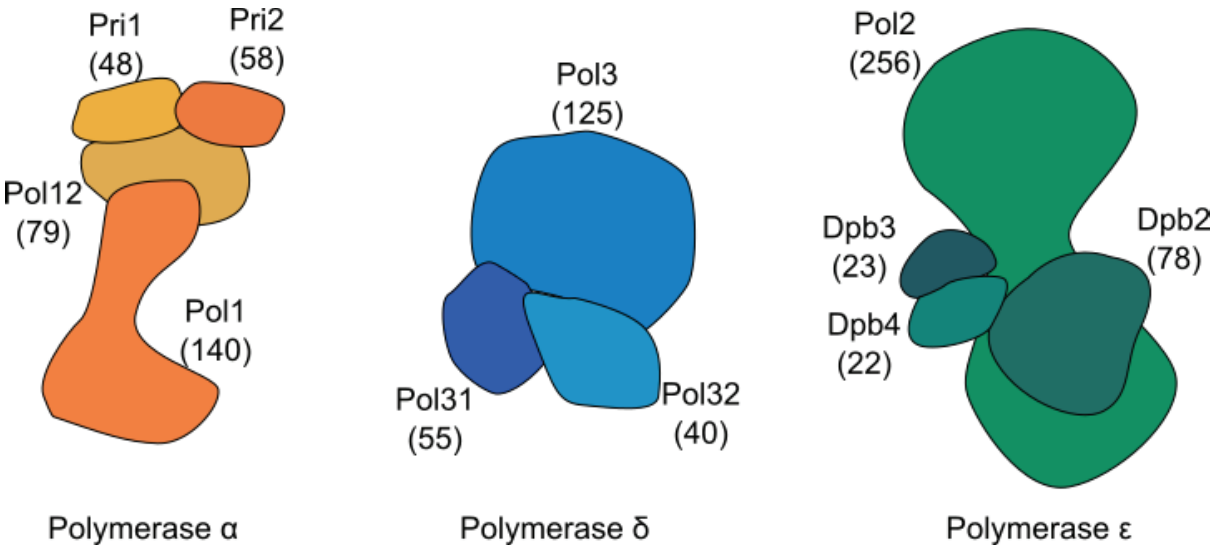


Figure 10: Subunit composition of the three eukaryotic DNA polymerases as they are in S.cerevisiae. Polymerase α is comprised of the four subunits Pol1, Pol12, Pri1 and Pri2. Polymerase δ is comprised of three subunits, Pol3, Pol31 and Pol32. Polymerase ε is the largest holoenzyme, comprising the large subunit Pol2, with the additional subunits Dpb2, Dpb3 and Dpb4. Subunit composition reviewed in Garg and Burgers, 2008 (Garg & Burgers, 2005). Structural information from Núñez-Ramirez et al., 2011; O'Donnell and Li, 2016; Zheng et al., 2020; Yuan et al., 2020 (Núñez-Ramírez et al., 2011; O'Donnell & Li, 2016; Zheng et al., 2020; Yuan et al., 2020).

1.5.2.1 Leading strand: Polymerase ε

For a considerable time, it was unclear which of polymerase ε and δ were the leading strand polymerase (Kunkel & Burgers, 2008). Both holoenzymes possess high processivity and reliable proofreading activity (Shcherbakova et al., 2003; Fortune et al., 2005; Garg & Burgers, 2005), which are necessary properties for the continuous and accurate replication of the leading strand. However, by collating data from many studies, including investigation of the mutagenic signatures of polymerase mutants deliberately made more error-prone (Pursell et al., 2007; Nick McElhinny et al., 2008), and immunodepleting each of the polymerases from *Xenopus* egg extracts (Fukui et al., 2004), a consensus was reached. Polymerase ε replicates the leading strand, whilst polymerase δ synthesises the lagging strand discontinuously (Kunkel & Burgers, 2008). Further evidence came from the high intrinsic processivity of polymerase ε, as well as a reduced strength of PCNA/Pol30 binding when compared to δ (Chilkova et al., 2007), indicating a role at the leading strand, where high processivity is more important, and PCNA/Pol30 interaction less important (Garg & Burgers, 2005). Following *in vitro* work

showed preferential binding and activity behaviour of the two polymerases on their respective strands, mediated by a range of unique structural features of each strand, such as the presence of PCNA/Pol30 and RPA on the lagging strand, and the CMG complex itself on the leading (Chilkova et al., 2007; Langston et al., 2014; Georgescu et al., 2014, 2015; Schauer & O'Donnell, 2017; Johansson, Garg & Burgers, 2004). Additionally, the use of enrichment and sequencing of protein-associated nascent DNA (eSPAN), a technique that has shown great promise in investigating nascent strand preference of many proteins (Yu, Gan & Zhang, 2018), showed categorically that this assignment was correct (Yu et al., 2014).

Structural data of the polymerase and CMGε complex further solidified the place of polymerase ε on the leading strand, with clear indications that polymerase ε is associated closely with the CMG complex (Zhou et al., 2017; Goswami et al., 2018; Yuan et al., 2020; Jones et al., 2021), which encircles leading strand ssDNA (see above). In addition, polymerase ε showed a domain not present in polymerase δ, proposed to encircle nascent dsDNA and aid processivity (Hogg et al., 2013).

However, there was a slight twist on this robust strand assignment for the two holoenzymes. Following reconstitution of the yeast origin firing system *in vitro* (Yeeles et al., 2015), it has been suggested that, following initial DNA unwinding, polymerase δ is the enzyme that catalyses the addition of new bases immediately following polymerase α activity, before "catching up" to the replisome and switching out for the polymerase ε for the main bulk of leading strand replication (Yeeles et al., 2017; Guilliam & Yeeles, 2020). This was in agreement with studies of polymerase usage genome-wide that had been conducted in *S.pombe* that also showed evidence of polymerase δ activity on the leading strand following initiation (Daigaku et al., 2015).

1.5.2.2 Lagging strand: Polymerase δ

As the data above suggests, it seems clear that, for the most part, polymerase δ is responsible for the discontinuous synthesis of lagging strands *in vivo* (Burgers & Kunkel, 2017). Polymerase α and polymerase δ were identified as the two holoenzymes responsible for aiding telomerase in extending telomeres, which involves extension of a lagging strand (Diede & Gottschling, 1999). Polymerase δ also shows several genetic or physical interactions with protein factors of the lagging strand machinery, including the Rad27 flap endonuclease and polymerase α (Garg & Burgers, 2005). Furthermore, it was shown that, in order for δ to match the processivity levels of polymerase ε, it required interaction with Pol30/PCNA, indicative of lagging strand activity (Georgescu et al., 2014, Burgers & Kunkel, 2017). The later genome wide studies and most of the *in vitro* assays confirmed this preference of polymerase δ for the lagging strand (Yu et al., 2014; Daigaku et al., 2015; Yeeles et al., 2017). However, more recent data and evidence in the context of DNA damage and replication restart indicate that polymerase δ may be specifically important in the first instance after polymerase α priming (Miyabe et al., 2015; Zhou et al., 2019; Guilliam & Yeeles, 2020).

1.5.2.3 Lagging strand: Polymerase α

Polymerase α is the primase of DNA replication, and the Pri1 subunit synthesises a short RNA primer of around 8-12 nucleotides which is then extended by the Pol1 polymerase subunit to around 20 base pairs as the start of DNA synthesis from this primer (Lehmans & Kagunisll, 1989; Johnson & O'donnell, 2005; Perera et al., 2013). In replication initiation, polymerase alpha is recruited to the replication origin following the CDK-dependent activation step and this is likely aided by the interactions shown with Mcm10 (Ricke & Bielinsky, 2004; Warren et al., 2009). Given the importance of priming on the lagging strand, where each Okazaki fragment must first be primed before extension by polymerase δ, it is unsurprising that

polymerase α appears to be associated with the replisome, and this interaction appears to be mediated by Ctf4 and also Mcm10 (Zhu et al., 2007; Gambus et al., 2009). Recent work that limited the abundance of polymerase α in the cell and observing Okazaki fragment length and abundance suggests that polymerase α dissociates and rebinds the replisome, rather than being bound to a CMG complex for the lifetime of the replication fork (Porcella et al., 2020). This also suggests that the role of Ctf4 and/or Mcm10 interactions are present to increase local effective concentrations of the complex.

The exact mechanism of priming of leading strand synthesis by polymerase α at origins unknown, but recent work has suggested that leading strand synthesis is primed from a sister fork lagging strand, extended by polymerase δ until it catches up to the helicase, and then synthesis is handed over to polymerase ε for continued leading strand synthesis (Aria & Yeeles, 2018). In this way there may be no "true" leading strand priming, just re-purposed lagging strand priming in a specific molecular context.

1.5.3 Replication Protein A (RPA)

Replication protein A (RPA) is a heterotrimeric protein comprised of Rfa1, Rfa2 and Rfa3 in yeast, that was identified as a factor required for early *in vitro* replication assays using viral helicases (Wobbe et al., 1987; Brill & Stillman, 1989). RPA shows strong levels of binding to ssDNA, and lower affinity for dsDNA (Wold, 2003). Quantification of this protein suggested it was both very abundant, and likely the principle ssDNA-binding (SSB) protein in eukaryotes (Iftode, Daniely & Borowiec, 1999; Wold, 2003). As a general sequence non-specific ssDNA binder, RPA functions in multiple cellular roles where ssDNA is present, such as in the DNA damage response and DNA replication, where it plays an essential role in both the protection of the ssDNA, and in mediating handover of the ssDNA to other proteins in these processes (Figure 8)(Wold, 1997; Fanning, Klimovich & Nager, 2006; Chen & Wold, 2014; Maréchal &

Zou, 2015). Specifically in DNA replication, RPA was found to interact with polymerase α and enhance its fidelity, as well as a role in the context of polymerase switching on the lagging strand (Dornreiter et al., 1992; Yuzhakov et al., 1999; Maga et al., 2001; Chen & Wold, 2014; Melendy & Stillman, 1993). Later , RPA was among the factors required for yeast minimal replication initiation *in vitro* (Yeeles et al., 2015). Outside of the DNA replication role, RPA plays an important role in the signalling of DNA damage and the processes of DNA repair (Coverley et al., 1991; Wold, 2003; Fanning, Klimovich & Nager, 2006; Zou & Elledge, 2003; Maréchal & Zou, 2015; Chen & Wold, 2014; Yates et al., 2018; Deshpande et al., 2017; Bantele, Lisby & Pfander, 2019).

1.5.4 Sliding clamp and clamp loader

Proliferating cell nuclear antigen (PCNA), Pol30 in yeast, is a sliding clamp that acts as a recruitment scaffold for multiple replication-associated factors and processes, with a particularly key role in enhancing the processivity of the polymerases δ and ε (Moldovan, Pfander & Jentsch, 2007; Boehm, Gildenberg & Washington, 2016). To load the ring-shaped Pol30, a clamp loading complex is needed, made up of several proteins with homology to AAA+ ATPases, most notably the replication factor C (RFC) proteins 1-5 (Yao & O'Donnell, 2012).

1.5.4.1 The sliding clamp Pol30/PCNA

Proliferating cell nuclear antigen (PCNA) was originally identified as a protein that was the target of auto-immune antibodies in cases of systemic lupus erythematosus (SLE) (Miyachi, Fritzler & Tan, 1978). Analysis of the abundance of this protein showed it was specific to replicating cells and showed a cyclical expression pattern (Bravo et al., 1981; Mathews et al., 1984). Since then, its role in DNA replication was confirmed with a combination of techniques that showed the structural similarity to the prokaryotic β DNA clamp, despite little sequence

similarity (Kong et al., 1992; Krishna et al., 1994). The clamp was required to regulate the processivity of the polymerases δ and ε (Tan et al., 1986; Prelich et al., 1987; Bravo et al., 1987), and had a wide range of interacting proteins and post-translational modification, such as sumoylation and ubiquitylation (Boehm, Gildenberg & Washington, 2016). Investigation into the specifics of DNA polymerase interaction and increase in processivity showed that Pol30 was actively involved in selection of the leading and lagging strand polymerases, enhancing processivity for polymerase δ in the lagging strand context, but favouring polymerase ε in the leading strand context (Georgescu et al., 2014). The range of factors that show interaction with PCNA/Pol30 is large, and ranges from regulatory roles such as in enhancing the degradation of Cdt1 to prevent re-loading and promoting or enhancing different repair pathways dependent on its modification state, to acting as a recruitment scaffold for nucleosome deposition, cohesion loading or the proteins required for Okazaki fragment processing (Boehm, Gildenberg & Washington, 2016; Moldovan, Pfander & Jentsch, 2007).

1.5.4.1.1 The clamp loader/unloader RFC complexes

To load and unload the stable PCNA ring requires a dedicated loader/unloader complex. Like the sliding clamps themselves, clamp loader function is present across the domains of life, reflecting their highly conserved role (Yao & O'Donnell, 2012). Clamp loaders in general form ring shaped complexes, in a right-handed spiral, with a gap that allows entry of DNA into the positively charged inner ring of the complex when ATP is bound (Bowman, O'Donnell & Kuriyan, 2004; Gaubitz et al., 2021). Once bound to PCNA, they guide the open PCNA ring to the specific primer-template interface DNA site, before ATP hydrolysis concomitant with clamp ring-closure (Gaubitz et al., 2021; Yao & O'Donnell, 2012). The RFC complex must then dissociate from DNA to allow polymerase binding to PCNA (Yao & O'Donnell, 2012). In addition to the "standard" clamp loader RFC1-5, there are alternative clamp loader complexes, such as the Ctf18-, Elg1- and Rad17- containing RFC complexes, where usually

the Rfc1 is replaced by the alternative protein subunit above (Yao & O'Donnell, 2012; Moldovan, Pfander & Jentsch, 2007). These alternative complexes are thought to have specific roles, such as the Ctf18-RFC clamp, which is thought to be preferentially responsible for leading strand clamp loading (Grabarczyk, Silkenat & Kisker, 2018). The Rad17-containing clamp loader complex loads a different clamp complex entirely, known as the 9-1-1 clamp, at sites of DNA damage (Majka, Niedziela-Majka & Burgers, 2006; Ellison & Stillman, 2003; Yao & O'Donnell, 2012; Ngo & Lydall, 2015).

Because of the inherent stability of the clamp on DNA, further activity of the RFC complexes is required to unload the PCNA/Pol30 clamp (Yao et al., 1996), and the clamp loader complexes display the ability to unload the clamps, as well as load them (Bylund & Burgers, 2005; Yao et al., 2006; Yao & O'Donnell, 2012; Moldovan, Pfander & Jentsch, 2007). As a result, both the DNA clamps and the clamp loader complexes are key components of the replication machinery.

1.5.5 Structural component Ctf4

Ctf4 was originally discovered in a screen for mutations that decreased chromosome transmission fidelity in mitosis (Spencer et al., 1990), and loss of Ctf4 caused a failure in sister chromatid cohesion (Hanna et al., 2001). It was found that this Ctf4 protein co-purified with polymerase α (Miles & Formosa, 1992), and Ctf4 was also among the proteins that were identified as being components of the replisome progression complex (Gambus et al., 2006). Further work confirmed biochemically that Ctf4 coupled polymerase α to the replisome via GINS (Gambus et al., 2009). Another study showed structurally how the Ctf4 protein assembles into a trimer by interaction of β-propellor domains that present three binding sites able to bind to multiple replication-associated cofactors. One of these cofactors being the Sld5 subunit of GINS (Simon et al., 2014). The interaction motif, Ctf4-interacting peptide (CIP box),

in Sld5 and Pol1 was subsequently shown to be present in several other proteins that are thought to have roles at the replication fork, and in Dpb2 in what appears to be an interaction additional to those already shown between polymerase ε and the CMG (Villa et al., 2016; Samora et al., 2016). Taken together, the available evidence positions Ctf4 as a general protein scaffold that recruits several separate proteins to the replisome complex (Figure 8), and can aid in positioning proteins like Pol1 for additional functions such as histone transfer (Gan et al., 2018). The tri-valent nature of Ctf4 is intriguing considering the well established two-protein axis of GINS-Ctf4-Pol1, and some data has suggested that this may be utilised for coupling multiple CMGs in a "replication factory" (Yuan et al., 2019), though could just reflect an available binding site for more transient interactions to interact without disrupting the CMG-Pol1 tether.

1.5.6 Fork protection complex

The fork protection complex consists of the proteins Csm3, Tof1 and Mrc1. Together, these proteins have been implicated in several functions relating to the stability of forks that encounter fork blocks or other stalling events (Bando et al., 2009). These proteins comprise a subcomplex of the replisome progression complex (RPC), and associate with the replisome in unperturbed replication as well as at specific fork blocks (Katou et al., 2003; Calzada et al., 2005; Gambus et al., 2006; Baretić et al., 2020). Tof1 and Mrc1 are key components of the S phase checkpoint signalling pathway in response to DNA damage (Foss, 2001; Alcasabas et al., 2001; Osborn & Elledge, 2003), but all three proteins appear to display functions relating directly to fork progression and stalling that do not depend on downstream signalling (Tourrière et al., 2005), though some of these functions appear limited to one or other of the protein factors. Whilst Tof1 and Csm3 show a unique role from Mrc1 in stalling at replication fork barriers (RFB) (Dalgaard & Klar, 2000; Mohanty, Bairwa & Bastia, 2006), Mrc1 shows a

separate role relating to modulation of the replication progression rate (Szyjka, Viggiani & Aparicio, 2005; Hodgson, Calzada & Labib, 2007). The results based on the study of mutant proteins *in vivo* is also well replicated in both reconstituted replication *in vitro* and recent structural data (Yeeles et al., 2017; Baretić et al., 2020). These data suggested that Tof1 and Csm3 bind more stably to the N-terminal tier of MCM2-7 (Figure 8), which may be related to their role in stalling at RFBs, but that Mrc1 was more flexibly tethered, with an association that seems to rely on the other members of this subcomplex (Bando et al., 2009), though was the key protein controlling replication fork rates. Overall, with functions that together modulate the rate of fork progression and stalling in response to regulated blocks and DNA damage or oxidation state perturbations (Katou et al., 2003; Labib & Hodgson, 2007; Somyajit et al., 2017), as well as the mediation of downstream damage signalling (Alcasabas et al., 2001; Osborn & Elledge, 2003; Foss, 2001), these proteins are of great importance in unperturbed and stressed replication.

1.5.7 Topoisomerases

DNA topoisomerases are enzymes that manage the topological state of DNA inside the cell. There are two classes of topoisomerase that correspond the mechanism by which they alter the topology of the DNA. Type I topoisomerases cleave a single backbone of the DNA duplex, and pass the other stand through this break to adjust the topology of the DNA (Pommier et al., 1998). Type II topoisomerases cleave both strands at one site, and pass another section of duplex DNA through this break (Champoux, 2003; Wang, 2002; Deweese & Osheroff, 2009). The type I topoisomerase in yeast, Top1, is a non-essential protein (Wang, 2003), though the type II topoisomerase, Top2, is an essential protein, likely due to this enzyme's ability to relax both directions of supercoils and compensate for the loss of Top1 (Uemura et al., 1987; Holm et al., 1985; Kim & Wang, 1989). The activity of topoisomerases is particularly important

during replication, due to the topological stress induced by the passage of the replisome along chromosomal DNA. This activity becomes even more crucial in situations of increased topological strain, such as at transcription-replication complexes (Lang & Merrikh, 2021), and in replication termination, see below. Analysis of Top1 and Top2 in the context of the fork showed that both proteins in yeast appeared to colocalise with the fork itself after initiation of DNA replication, and that these proteins play an important role at stalled replication forks (Bermejo et al., 2007; Shyian et al., 2020).

1.5.8 Nucleosome removal/reestablishment

Since the DNA inside cells is not present as a strand of dsDNA, but instead associated with histone proteins to form nucleosomes (Kornberg, 1974; Olins & Olins, 1974; Zlatanova, Leuba & van Holde, 1998), the passage of the DNA replication fork, by unwinding the dsDNA, necessarily disrupts these protein-DNA structures within about two-nucleosome lengths of the helicase. However, since a large amount of information is coded within the histone modifications within nucleosomes, ensuring that this data is transferred to each daughter DNA molecule is of great importance (Badeaux & Shi, 2013). Whilst the exact mechanism by which nucleosomes are removed from DNA ahead of the fork, and reassembled behind it, along with newly synthesised histone proteins, is not yet fully established, many proteins and complexes have been shown to be involved in this process. Key players in this process have been identified as the FACT and CAF-1 complexes, Asf1, and Rtt106 (Alabert & Groth, 2012; Budhavarapu, Chavez & Tyler, 2013; Zhang, Feng & Li, 2020; Hammond et al., 2017). In addition to these histone-chaperone specialised proteins, specific domains of Mcm2 have been shown to bind readily to histone H3, and to Asf1 through this bridging interaction (Figure 8)(Ishimi et al., 1996; Clément & Almouzni, 2015; Wang et al., 2015; Petryk et al., 2018), and loss of polymerase ε subunits Dpb3 and Dpb4 causes an impairment in inheritance of histones to the

leading strand (Yu et al., 2018). Additionally, RPA and PCNA presence on the newly replicated DNA is important for recruitment and placement of new and recycled histones (Shibahara & Stillman, 1999; Zhang, Shibahara & Stillman, 2000; Liu et al., 2017). Current models suggest that upon approach of the helicase to the nucleosome, the association with DNA is destabilised, and any or all of Mcm2, FACT or Asf1 process the nucleosome disassembly and pass histone H3-H4 dimers along either the Dpb3/4 mediated path to the leading strand, or via the Ctf4-Polα axis to the lagging strand (Zhang, Feng & Li, 2020). Behind the fork, assembly of newly synthesised histones is carried out by CAF-1, Rtt106 and FACT potentially using RPA and PCNA/RFC interactions to guide their activity (Zhang, Feng & Li, 2020).

1.6 Termination of Replication

The loading and activation of the MCM2-7 is a highly controlled process, as it needs to be to prevent over-replication (Tada & Blow, 1998; Shreeram & Blow, 2003). However, this means that a limited number of MCM2-7, and by extension, CMG, complexes are present on the genome in S phase. Whilst these complexes are removed at the end of replication (Liang & Stillman, 1997), this process must also be very tightly controlled to ensure that this removal does not occur in an unregulated manner, and potentially leave un-replicated tracts of the genome (Bailey, Moreno & Gambus, 2015).

In contrast to prokaryotes, whose circular genomes have a single defined termination zone facilitated by directionally specific helicase-blocking proteins ter/Tus (Neylon et al., 2005; Duggin et al., 2008), eukaryotes do not generally possess such defined termination sites. Instead, the majority of termination seems to occur stochastically at sites where converging forks are likely to meet (Greenfeder & Newlon, 1992; Hennion et al., 2020). There are some sites that bind specific factors that slow replication forks and increase the local termination, but overall, termination occurs roughly between origins in *S. cerevisiae,* though this can be affected

by factors such as initiation rate at those flanking origins, and in rare cases, specific sequence elements (Fachinetti et al., 2010).

There are several challenges posed by the unique situation created by replication termination. Firstly, the essential cause of the termination process is the approach of two CMG complexes towards each other, with duplex parental DNA between them. Though each CMG complex threads only the leading strand through the central MCM2-7 hexamer channel, in the current modified steric exclusion model (O'Donnell & Li, 2018; Eickhoff et al., 2019; Jones et al., 2021; Baretić et al., 2020), there remains some interaction of the lagging strand with the CMG. This might well cause a direct head-to-head collision of the CMGs, if the CMGs are significantly cross-threaded on the lagging strand due to unwinding in the N-terminal tier, which must be resolved or avoided. However, available data suggests that the bypass event causes no perceptible slowing of the two CMGs in an *in vitro* reconstituted setting. Interestingly, in the purified yeast replication system, additional factors were required to assist the unwinding of the final part of the intervening dsDNA between forks (Deegan et al., 2019). These accessory helicases are also able to bypass lagging strand blocks (Deegan et al., 2019; Schauer et al., 2020), which might reflect a role in resolving the possible steric conflict of the two CMG complexes.

Secondly, the DNA itself is a possible impediment to completion of replication. As the parental DNA duplex is unwound, the DNA ahead of the fork becomes overwound, since the supercoils built up cannot freely diffuse away along the chromatin in cells (Keszthelyi, Minchell & Baxter, 2016). This can affect the energetic viability of continued unwinding and becomes increasingly more of an issue when two forks converge towards each other (Dewar & Walter, 2017). To bypass this hurdle, fork rotation and the generation of pre-catenanes, which are then processed by class II topoisomerases, Top2 in yeast, is allowed to occur at terminating forks

(Keszthelyi, Minchell & Baxter, 2016; Dewar & Walter, 2017). This permits the topological stress to transfer from ahead of the forks, to behind them, and to be removed prior to cell division (DiNardo, Voelkel & Sternglanz, 1984).

Thirdly, the replisome, particularly the CMG core particle must be removed from DNA. Given that excessive loss of CMG from chromatin during replication would be detrimental as discussed, this must be highly regulated (Bailey, Moreno & Gambus, 2015). In yeast, targeting of replication complexes for disassembly is specific for those that are terminating. The signal conferring specificity is the polyubiquitylation of Mcm7 by the ubiquitin ligase SCF^{Dia2}, which is necessary for the disassembly of the CMG by the Cdc48 segregase (Maric et al., 2014; Moreno et al., 2014). Surprisingly, it was found that in yeast, SCF^{Dia2} travels with the replication fork, so some specific mechanism must prevent the ubiquitylation of Mcm7 until termination, though more recently, this interaction was thought to be due to post-isolation interaction introduced during experimental procedures (Morohashi, Maculins & Labib, 2009; Maculins et al., 2015). Until recently, there were multiple hypotheses as to what the specific signal unique to the terminating replisome could be that induces ubiquitylation, including structural changes in the CMG due to either the torsional strain at the termination site, or the passage onto dsDNA, or indeed the co-localisation of two oppositely orientated CMGs could provide the necessary signal (Bailey, Moreno & Gambus, 2015). Recent structural data of a terminating CMG, compared to CMG at an active fork, suggests that the path of the lagging strand obscures a Dia2-MCM interaction site, thereby preventing cullin activity until the lagging strand is lost, as occurs during termination, when the CMG passes onto dsDNA already replicated by the opposing fork (Jenkyn-Bedford et al., 2021; Baretić et al., 2020). This potential mechanism of specifying termination complex is elegant but will need to be more thoroughly investigated in future to see if this interaction domain is truly the key to SCF^{Dia2} activity and specificity during termination.

1.7 DNA damage

The response of cells to DNA damage is incredibly complex and multi-faceted, with multiple pathways of sensors, mediators and effector proteins that are activated both generally and specifically to mediate responses to protect DNA and enable repair (Lanz, Dibitetto & Smolka, 2019). Many of these pathways link directly or indirectly to the replication machinery, either due to the production and detection of DNA lesions as a result of replication activity, or to modify DNA replication processes to prevent or recover from damage (Iyer & Rhind, 2017). Whilst the mechanisms of the DNA damage response, particularly those specifically relating to stalled or collapsed forks are beyond the scope of this introduction, it should be noted that multiple studies have shown that the proteome associated with such events are distinctly different to active forks, and multiple DNA damage-related factors are recruited to the fork in the event of DNA damage causing this kind of impediment to replisome progression (Sirbu et al., 2013; Nakamura et al., 2021).

1.8 Additional proteins

1.8.1 Rrm3

Rrm3 is a factor important for genome stability (Schmidt & Kolodner, 2004; Torres, Schnakenberg & Zakian, 2004; Ooi, Shoemaker & Boeke, 2003; Szyjka, Viggiani & Aparicio, 2005; Schmidt & Kolodner, 2006) that was first associated with genome maintenance due to its effect on recombination of the rDNA repeats and *CUP1* genes (Ivessa, Zhou & Zakian, 2000) as well as the mobility of the Ty1 retrotransposon (Scholes et al., 2001). The closely related *S. cerevisiae* protein Pif1 had been found to inhibit telomerase activity by the action of 5' to 3' helicase activity (Schulz & Zakian, 1994; Zhou et al., 2000) and although Rrm3 possesses high sequence similarity to Pif1 (Bessler, Torres & Zakian, 2001), particularly in the helicase region, the activity of Rrm3 regarding telomeric regions was found to be different than for Pif1. Rrm3 was similarly found to be a 5' to 3' helicase, and deletions affected the progression of replication forks through telomeric regions. This effect was dependent on the ATPase activity of Rrm3 (Ivessa et al., 2002). Further work explained the common mechanism behind the effects on both rDNA loci and telomeric regions (Ivessa et al., 2003; Azvolinsky et al., 2006; Makovets, Herskowitz & Blackburn, 2004). Cells in which Rrm3 was deleted showed increased stalling at sites where protein-DNA complexes were present, including the mating type locus, inactive origins, centromeres, and the rDNA replication fork barrier (RFB) that is bound by the Fob1 protein (Torres, Bessler & Zakian, 2004; Ivessa et al., 2003). Analysis of chromosome duplication rates showed that, whilst Rrm3-dependent sites such as telomeres and rDNA were severely affected by Rrm3 deletion, even regions without such known sites showed delayed replication, leading to the finding that Rrm3 associated with origins of replication, and appeared to travel with the replisome away from origins, though the initial association of Rrm3 with forks at the origin in the absence of fork blocks is disputed (Calzada et al., 2005). In addition, a co-immunoprecipitation of Rrm3 with the polymerase epsilon subunit Pol2 and an

additional binding with PCNA clamp loader was detected, suggesting Rrm3 associates with replisomes *in vivo* (Azvolinsky et al., 2006; Schmidt, Derry & Kolodner, 2002). Whilst deletion of Rrm3 protein alone is not lethal and produced only limited sensitivity to DNA damaging agents such as MMS and HU (Torres, Schnakenberg & Zakian, 2004) despite constitutive hyperphosphorylation of Rad53 under non-genotoxic conditions (Ivessa et al., 2003), simultaneous deletion of Sgs1, a helicase important in recombination and repair (Frei & Gasser, 2000; Cejka et al., 2012; Versini et al., 2003; Kaliraman et al., 2001; Cobb et al., 2003), is lethal (Torres, Schnakenberg & Zakian, 2004). Strains with mutants in the DNA damage checkpoint response and the S-phase checkpoint respectively showed that whilst the DNA-damage checkpoint functionality was important for healthy growth of Rrm3-deleted cells, co-deletion of Mrc1, a key mediator of the S-phase checkpoint (Alcasabas et al., 2001), also caused synthetic lethality (Torres, Schnakenberg & Zakian, 2004). Furthermore, deletions of a range of fork integrity factors are also lethal or cause severe growth impairment when combined with Rrm3 deletion, such as Mec1, Srs2 and Rtt101 (Schmidt & Kolodner, 2004; Torres, Schnakenberg & Zakian, 2004; Luke et al., 2006). Rrm3 was also found to have an antagonistic role to components of the replication fork protection complex, Tof1 and Csm3. This complex is known to enable stable fork arrest at the rDNA Fob1-binding termination sites and other fork stalling events (Calzada et al., 2005; Tourrière et al., 2005; Katou et al., 2003), and appears to counter the activity of Rrm3 at such sites.

As a result of this, the counteraction of Rrm3 and Tof1/Csm3 has an effect on the termination of DNA replication. Without Rrm3, termination at natural Ter sites in rDNA regions and pausing at transcription-replication collisions was enhanced, and this effect was ameliorated by simultaneous deletion of Tof1 or Csm3 (Mohanty, Bairwa & Bastia, 2006). Rrm3 was also found to aid in the termination of two converging replication forks, though this appeared to be more of as a compensatory mechanism when the related Pif1 was not present (Deegan et al.,

2019). Nonetheless, the presence of Rrm3 around terminating and paused forks is well documented.

More detailed investigation into the individual domains of Rrm3 found that there appeared to be a division of function between a specific N-terminal region of Rrm3 (Figure 11), that more broadly is phosphorylated by Rad53 and contains a PCNA-interacting PIP box (Rossi et al., 2015; Azvolinsky et al., 2006), and the C-terminal helicase region (Figure 11), that is responsible for the DNA-protein block "sweepase" activity (Ivessa et al., 2003). Specifically, a small N-terminal region (186-212) seemed to confer some form of control over DNA replication in HU conditions. In yeast two-hybrid assays, this region was found to be responsible for binding of Orc5 (Syed et al., 2016; Matsuda et al., 2007), and controlled origin recruitment under HU treatment conditions, though origin association for a subset of origins

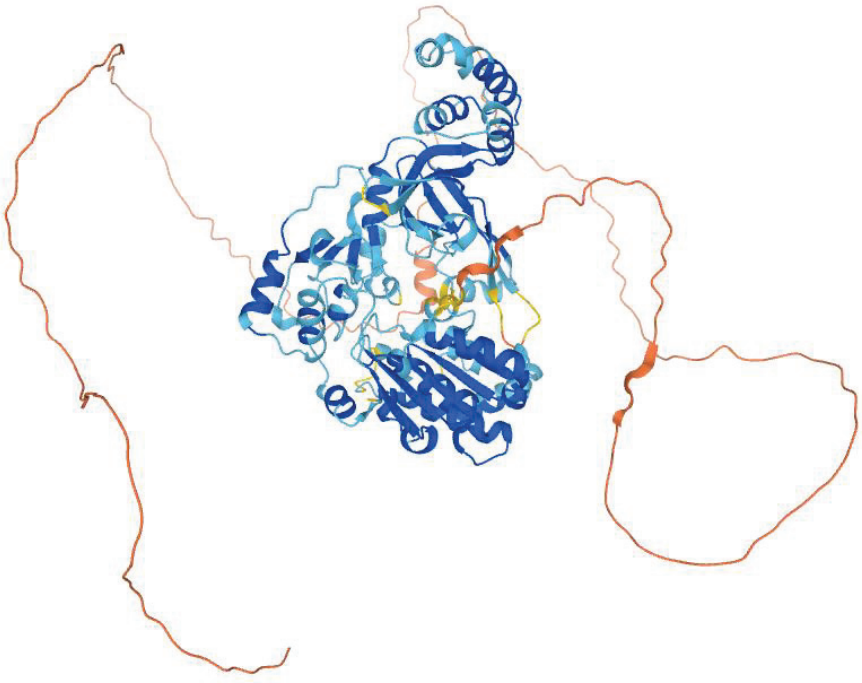


Figure 11: Alphafold structural prediction of S.cerevisiae Rrm3 (Jumper et al., 2021; Varadi et al., 2021). The high confidence C terminus (570-1131) includes the prior crystal structure of the ATPase-C domain (814-1122), and the N-terminal ATPase lobe, together forming the SNF2-like ATPase domain. The domain that displays a role in controlling origin activity (186-212) is predicted to be disordred, as is the N-terminal region additionally thought to be phosphorylated by Rad53 and have a PIP-box (1-250) (Syed et al., 2016, Bessler and Zakian, 2004). Alphafold structural prediction reproduced from Alphafold Protein Structure Database (Varadi et al., 2021).

under untreated conditions was unaffected by absence of the Orc5-interaction site. This element did show a requirement for Tof1 and Mrc1 activity when deleted in DNA damaging conditions, suggesting this region of Rrm3, too, has effects on DNA damage production and resolution (Syed et al., 2016).

It is worth noting that Rrm3, whilst not possessing sequence homologs in higher eukaryotes, has many plausible functional homologs. There are a number of 5'-3' DNA helicases encoded in metazoan species, several of which may be competent to carry out such termination-related functions (Xia et al., 2021).

In yeast, regardless, Rrm3 clearly has a role at replication forks, involving DNA-protein block bypass. Recent evidence, however, indicates that may not be the entirety of Rrm3's function in genome maintenance and DNA replication.

1.8.2 Fun30

Fun30 is a member of the SNF2 family of chromatin remodellers (Neves-Costa et al., 2009; Laurent, Yang & Carlson, 1992; Flaus et al., 2006), that was first identified to have a role in protecting chromatin segregation and genome stability from an overexpression screen investigating loss or gain of a chromosome fragment (Ouspenski, Elledge & Brinkley, 1999) The SNF2 family of proteins all contain a helicase-like ATPase domain (Figure 12) similar to that found in Snf2, the catalytic subunit of the yeast SWI/SNF complex, which has long been identified as a chromatin remodelling complex (Abrams, Neigeborn & Carlson, 1986; Laurent, Treich & Carlson, 1993; Peterson & Tamkun, 1995). There are several well-studied members of this family, with a variety of specificities regarding histone selection and remodelling activity, such as INO80, ISWI and CHD1 (Ryan & Owen-Hughes, 2011). Notably, the helicase activity of these proteins does not seem directed at DNA or DNA:RNA duplex unwinding, rather, the remodellers hydrolyse ATP to translocate DNA relative to the nucleosome (Cairns, 2007; Bowman, 2010; Clapier et al., 2017), and this is what allows the nucleosome rearrangement that each complex enacts, be it sliding, ejection, exchange or histone replacement (Cairns, 2007; Ryan & Owen-Hughes, 2011). The importance of the SNF2-like chromatin remodellers in important aspects of genome maintenance, such as enhancing DNA repair and fork stability pathways such as homologous recombination (HR), non-homologous end-joining (NHEJ), nucleotide-excision repair (NER) and the intra-S phase checkpoint (Morrison & Shen, 2009; Stadler & Richly, 2017). In addition to these proteins' varied other roles in telomere maintenance, centromere specification, chromosome segregation and transcriptional regulation, it is perhaps unsurprising that complexes containing the human homologs of these proteins are implicated in many types of cancer (Ryan & Owen-Hughes, 2011; Boginya et al., 2019; Mittal & Roberts, 2020).

Fun30 has homologs within higher eukaryotes (SMARCAD1 (Adra et al., 2000), *Etl-1* (Schoor, Schuster-Gossler & Gossler, 1993), *S.pombe* (fft3 (Neves-Costa et al., 2009)) and Neurospora (CLOCKSWITCH (Belden, Loros & Dunlap, 2007)), with a wide array of purported activities and functions.

In yeast, Fun30 has been ascribed many activities fitting its identity as an SNF2 family remodeller. Fun30 deletants demonstrated chromatin silencing defects at the HMR mating locus, telomeres, and rDNA repeats (Neves-Costa et al., 2009). More detailed *in vitro* investigation into Fun30 suggested that it was more efficient at histone replacement than nucleosome repositioning, and was found to bind both DNA and chromatin as expected of this type of protein (Awad et al., 2010). Further evidence that Fun30 acts in chromatin remodelling, on non-heterochromatin substrates, comes from experiments showing defects in centromere chromatin structure when Fun30 is deleted (Durand-Dubief et al., 2012; Strålfors et al., 2011), and from experiments showing Fun30 transcriptional repression reliant on nucleosome repositioning (Byeon et al., 2013). In keeping with a previous screen, where Fun30 was identified as a synthetically lethal genetic interactor of certain ORC mutants (Suter et al., 2004), Fun30 deletants exhibit additive issues with DNA replication when combined with temperature sensitive ORC mutants. Linking Fun30 with DNA replication processes, this study also identified camptothecin sensitivity in Fun30 mutants (Neves-Costa et al., 2009). Further complicating an already broad spectrum of activities, Fun30 has been demonstrated to act at double strand breaks to enhance DNA end resection (Siler et al., 2017; Bantele & Pfander, 2019; Eapen et al., 2012; Chen et al., 2012, 2016; Costelloe et al., 2012; Bantele et al., 2017), and, additionally, to have an end resection-independent role in reducing genotoxic stress, dependent on Rad51 and the predicted CUE ubiquitin-binding domain (Figure 12) of Fun30 itself (Neves-Costa et al., 2009; Bi et al., 2015). Fun30's reported activity at DSBs seems due to destabilisation of nucleosomes at DNA ends, to promote Exo1 and Sgs1 activity (Bantele &

Pfander, 2019; Chen et al., 2012; Eapen et al., 2012). Recruitment of Fun30 to the sites of double strand breaks is dependent on Dpb11 and the 9-1-1 clamp complex, and the activity of Fun30 at these sites relies on phosphorylation by Cdk1 in the Fun30 N terminus (Chen et al., 2016). Fun30 also has an antagonistic relationship with the end resection inhibitor Rad9 at double strand breaks, though the mechanistic nature of this relationship is unknown (Bantele & Pfander, 2019).

SMARCAD1, the human Fun30 homolog, has been found to have an additional role in deacetylation of nascent histones that are laid down immediately following replication, in a manner dependent upon an interaction with the sliding clamp PCNA, a peripheral member of the DNA replication machinery (Mermoud, Rowbotham & Varga-Weisz, 2011; Rowbotham et al., 2011). Additional activities, including the regulation of PCNA levels at replication forks

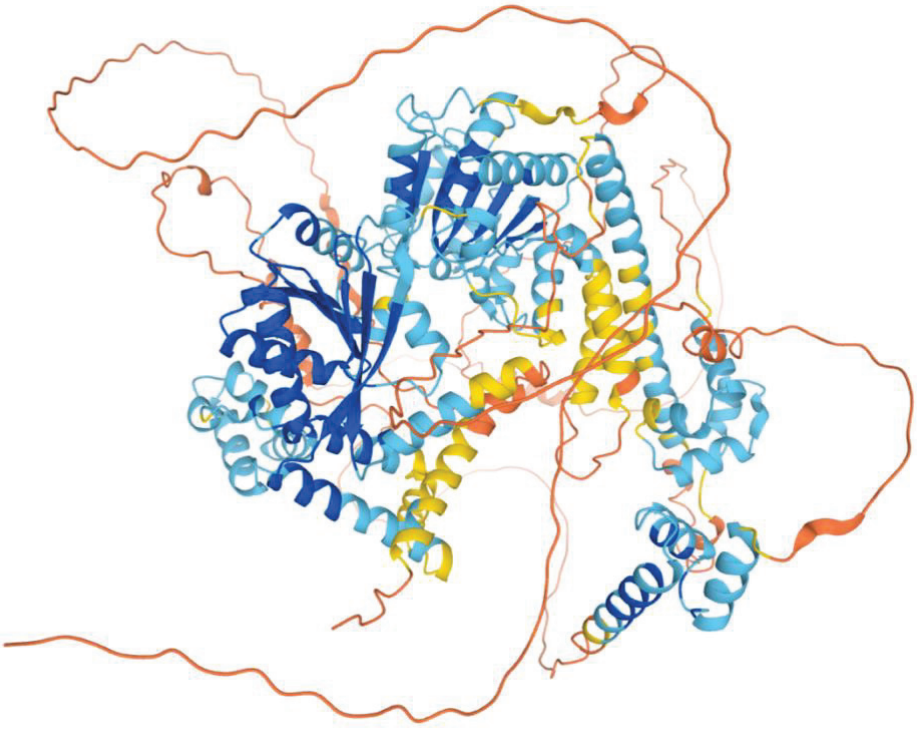


Figure 12: Alphafold structural prediction of S.cerevisiae Fun30 (Jumper et al., 2021; Varadi et al., 2021). The high confidence C terminus (570-1131) includes the prior crystal structure of the ATPase-C domain (814-1122), and the N-terminal ATPase lobe, together forming the SNF2-like ATPase domain. The reported CUE domain is also predicted to high confidence (71-116). Alphafold structural prediction reproduced from Alphafold Protein Structure Database (Varadi et al., 2021).

by the mutual interaction of these two proteins, and increasing resistance to replication stress at restarting forks by preventing 53BP1 accumulation, have also recently been attributed to this protein, providing a possible link to DNA replication (Lo et al., 2021). Association of SMARCAD1 to replication forks was shown in an isolation of proteins on nascent DNA technique (iPOND), where it showed similar levels of enrichment to both active and stalled forks (Sirbu et al., 2013).

In fission yeast, meanwhile, the Fun30 homolog Fft3 is reported to aid translational elongation of DNA-dependent RNA Polymerase II, in concert with the FACS complex (Lee et al., 2017). Here, too, Fun30/Fft3 is supposedly destabilising nucleosomes, in this context to allow unrestricted passage of the polymerase holoenzyme.

Fun30, then, has been reported to function in a wide range of chromatin remodelling contexts, that seem to vary across species. This variation in functions may seem surprising given the homology observed, but, given the wide range of specificities of SNF2 family members which share a similar functional ATPase domain, is not unreasonable (Neves-Costa et al., 2009; Ryan & Owen-Hughes, 2011; Flaus et al., 2006). Several lines of evidence, including Fun30's genetic interaction with Orc5, SMARCAD1's replisome interaction and fork stabilisation function, and Fft3's nucleosome removal ahead of a large DNA translocating holoenzyme, make Fun30 an interesting protein with respect to its possible role(s) in DNA replication.

1.8.3 Hrq1

The RecQ family of helicases are evolutionarily well conserved (Umezu, Nakayama & Nakayama, 1990; Mandell et al., 2005; Hartung & Puchta, 2006), and mutations in this gene family in humans can often lead to heritable genetic diseases (Bernstein, Gangloff & Rothstein, 2010). The RECQ4 gene is responsible for a subset of these disorders (Siitonen et al., 2009).

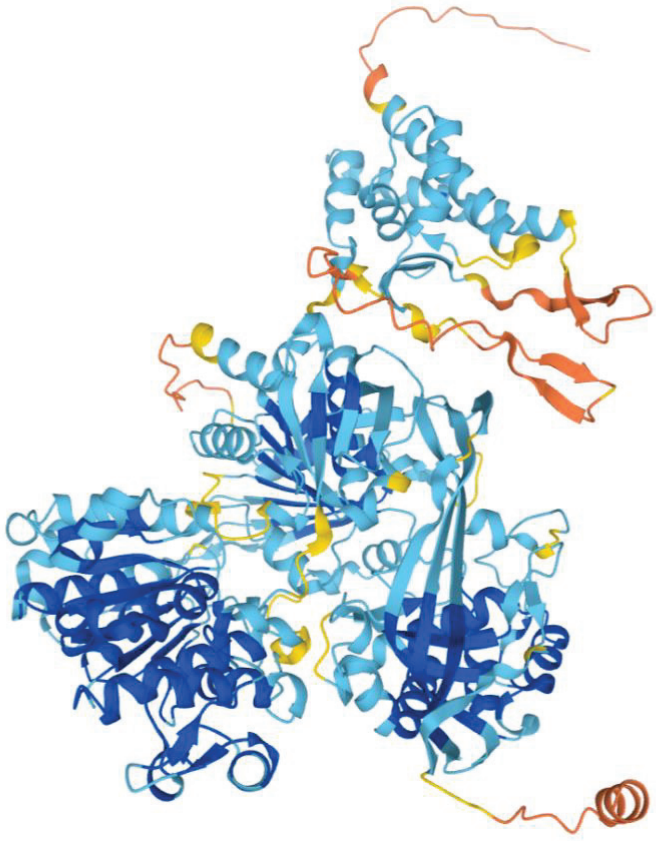


Figure 13: Alphafold structural prediction of S.cerevisiae Hrq1 (Jumper et al., 2021; Varadi et al., 2021). The N-terminal region (residues 1-299, top of image) is predicted to have some secondary structure but does not form a compact domain, and previously was considered to be natively disordered, much like the human homologue's Sld2-like N-terminus (Rogers et al., 2017).The C-terminus, containing the ATP-binding and C-terminal helicase domains (residues 299-483, lower left and residues 521-678, central domain in image) is predicted to be more structured, possessing the conserved core helicase lobe domains (Bochman et al., 2014). The RecQ homology domain (RQC)(residues 678-873, spanning all three lower regions of image) and RecQ4p/Hrq1p-conserved domain (RHCD)(residues 910-1050, lower-right region of image)(Barea, Tessaro & Bonatto, 2008) make up the remainder of the high-confidence predicted structure of the protein. Alphafold structural prediction reproduced from Alphafold Protein Structure Database (Varadi et al., 2021).

Biochemically, the RecQ proteins are 3'-5' DNA helicases (Umezu, Nakayama & Nakayama, 1990; Chu & Hickson, 2009; Gray et al., 1997). In yeast, Sgs1 was thought to be the single homologue of this gene family (Gangloff et al., 1994), but Hrq1 was found to possess

significant homology to the RecQ4 gene (Bochman et al., 2014; Lee et al., 2005; Barea, Tessaro & Bonatto, 2008). The N-terminal region of RecQ4 (Figure 13) is homologous to yeast replication factor Sld2 and retains some activity reminiscent of yeast Sld2, by binding the *Xenopus* Dpb11 homolog xCUT5, and being targeted for phosphorylation by S phase CDK activity (Matsuno et al., 2006).

As for Hrq1 itself, multiple studies have investigated its role and mechanism of action. Because RecQ4's essential Sld2-like N-terminus makes genetic manipulation of this protein difficult, Hrq1 provides a more tractable model for the helicase activity of RecQ4 *in vivo* (Rogers et al., 2017). RecQ helicases in general have been implicated in telomere maintenance, double strand break (DSB) repair and DNA inter-strand crosslink (ICL) repair (Gupta & Schmidt, 2020; Gravel et al., 2008; Rogers et al., 2020; Ghosh et al., 2012; Johnson et al., 2001; Hardy et al., 2014; Huang et al., 2001; Singh et al., 2012). Hrq1 did indeed show capability in these processes, (Bochman et al., 2014), and much research has been devoted to Hrq1's ability to aid repair of ICLs in multiple organisms where it is present (Rogers et al., 2020, 2017; Bochman et al., 2014; Röhrig et al., 2018). A recent study showed that Hrq1 is able to stimulate the exonuclease Pso2 at ICLs, perhaps by overcoming autoinhibition of Pso2 to ensure only site-specific nuclease activity (Rogers et al., 2020).

Recent efforts have been made to specify both the physical interactome of Hrq1, and the full network of genetic interactors (Rogers et al., 2020; E et al., 2020). These studies identified genetic and physical interactions with enriched GO terms for transcription, DNA repair, chromatin organisation, mitotic cell cycle and rRNA processing. Some of the strongest genetic interactors were observed with genome integrity factors as expected, but surprisingly strong physical associations were observed among transcription and rRNA processing factors (Rogers et al., 2020). The *hrq1* strain examined in this study also demonstrated increased sensitivity to high levels of caffeine, which blocks ATM/ATR activity (Kuranda et al., 2006), suggesting

these interactions are relevant to Hrq1 activity in the interaction between DNA damage and transcription. It is interesting to note, however, that caffeine has been reported to have many effects on DNA replication and repair, where more canonical Hrq1 activity is expected (Ruta & Farcasanu, 2020).

Hrq1's direct links to DNA replication are not clear, with reported roles in chromatin stability, telomere maintenance and ICL repair, but links to the homologue RecQ4, and in turn, RecQ4's region of Sld2 homology, are interesting in context of possible association of Hrq1 with the replication machinery in yeast.

1.8.4 The Sir proteins

The Sir1, Sir2, Sir3 and Sir4 proteins are regulators of silenced chromatin in yeast (Rine & Herskowitz, 1987; Aparicio, Billington & Gottschling, 1991). In *S.cerevisiae*, chromatin is not modified in the same way as other eukaryotes and lacks DNA methylation machinery (Holliday & Pugh, n.d.; Proffitt et al., 1984; Hattman et al., 1978; Singh & Klat, 1992), making heterochromatin formation that other organisms utilise with repressed regions enriched with DNA methylation unfeasible. Instead, budding yeast has regions of silenced chromatin at limited sites; telomeres (Tham & Zakian, 2002), the silent mating loci HML and HMR (Laurenson & Rine, 1992), and within the rDNA array. The telomeric silenced region requires activity of Sir2, Sir3 and Sir4 of the Sir family, and interaction with Rap1, a telomere capping protein (Rine & Herskowitz, 1987; Pardo & Marcand, 2005; Moretti et al., 1994), while at the mating type loci, Sir1 has an additional role via an interaction with Orc1, which also binds at silencers at these regions (Foss et al., 1993; Gardner, Rine & Fox, 1999). At the rDNA locus, Sir2 forms part of an entirely separate RENT complex to enact silencing (Straight et al., 1999; Shou et al., 1999).

It has been shown that Sir2 acts as an NAD^+-dependent deacetylase (Imai et al., 2000; Dutnall & Pillus, 2001), and is the most conserved of the SIR complex (Hickman, Froyd & Rusche,

2011; Brachmann et al., 1995). The deacetylase activity of Sir2 is required for silencing activity (Imai et al., 2000; Hoppe et al., 2002; Tanny et al., 1999; Tanny & Moazed, 2001).

Sir4 is the least conserved of the Sir family (Zill et al., 2010; Kellis et al., 2003; Cliften et al., 2003), and forms a native complex with Sir2, without accompanying Sir3 (Moazed et al., 1997; Ghidelli et al., 2001; Hoppe et al., 2002). Sir4 also appears to fulfil the role of Sir1 at the mating type loci at telomeres, where its binding of Rap1 is analogous to the Sir1-ORC interaction and used to recruit the Sir proteins (Hoppe et al., 2002).

Sir3 shares significant homology with Orc1 (Hickman, Froyd & Rusche, 2011). It does not possess functional ATPase activity, but maintains a BAH domain, that, like Orc1's, enables binding to nucleosomes (Bell et al., 1995; de Ioannes et al., 2019; Connelly et al., 2006; Onishi et al., 2007). This BAH domain differs enough from Orc1 to have a specificity for the unmodified histone H4 lysine K16 (Buchberger et al., 2008; de Ioannes et al., 2019), which is proposed as a method by which Sir2 is able to generate Sir3 binding sites and enable silencing (Liou et al., 2005). A second domain showing specificity for unmodified histones is reported within the AAA+ domain of Sir3, along with a Sir3-Sir4 reaction interface (Ehrentraut et al., 2011), which could explain the conservation of the integrity of the AAA+ domain relative to Orc1, without necessitating sequence conservation (Hickman, Froyd & Rusche, 2011).

Sir3 certainly seems to be responsible for the spreading of silenced heterochromatin, with many models proposed that involve extensive interactions between Sir3 and itself, Sir4, Rap1 and histones H3 and H4, leading to compacted and nuclease-resistant regions of DNA (Laura N. Rusche, Ann L. Kirchmaier & Rine, 2003; Hoppe et al., 2002; Hickman, Froyd & Rusche, 2011; Buchberger et al., 2008; Gasser & Cockell, 2001; Rusché, Kirchmaier & Rine, 2002). These types of models were tested *in vitro* and formation of heterochromatin fibres was confirmed that contain Sir2, Sir3 and Sir4. In this model, Sir3 binds to histones, before dimerising to form an extended nucleosome-Sir3 structure, with two Sir3 molecules to a

nucleosome. This is then bridged by a heterotetramer of Sir2/4 to induce further compaction (Swygert et al., 2018). This would seem to fit with initial observations that upon overexpression of Sir3, silenced regions were extended, without needing concurrent over expression of the other Sir proteins (Renauld et al., 1993).

There is also a growing body of evidence that the Sir proteins, most notably Sir2 and Sir3, act on or around origins of replication. Rescue of MCM loading mutants by deletion of Sir2 or Sir3 imply some mechanism by which the Sir proteins impede origin licensing (Pappas, Frisch & Weinreich, 2004; Crampton et al., 2008), and both proteins localise to origins and origin-proximal nucleosomes are deacetylated. The Sir3 binding and histone deacetylation levels correlate with the amount of rescue of licensing (Hoggard et al., 2018). Loss of Sir2 results in a modified temporal program of origin firing, which may be a result of altered licensing, leading to more efficient, early-firing origins becoming even more efficient, to the detriment of lower efficiency late origins (Hoggard et al., 2020). A similar effect has also been observed in humans, where the NAD^+-dependent deacetylase SIRT1, the homolog of Sir2, binds to origins and inhibits firing of "dormant" origins that are only activated under DNA damaging conditions (Utani et al., 2017). These origin-specific effects of the Sir proteins have been suggested as an explanation as to why early experiments showed chromosome instability when Sir levels were perturbed (Hoggard et al., 2018; Holmes et al., 1997).

1.8.5 CAF-1 complex

The Chromatin Assembly Complex (CAF-1), that in yeast contains the three subunits Rlf2, Cac2 and Msi1, is responsible for assembly of histones onto newly synthesised DNA at replication forks in human cells (Smith & Stillman, 1989; Krude, 1995). This activity has been demonstrated across eukaryotes, and is an essential human complex, despite being non-essential in yeast (Volk & Crispino, 2015; Hoek & Stillman, 2003; Franco & Kaufman, 2004; Tyler et al., 1999). Acting downstream of Asf1, a protein that binds the histone H3/H4 dimer

and interacts with the RFC complex in yeast (Franco et al., 2005), the CAF-1 complex is important in replication-dependent nucleosome deposition, and an interaction with Pol30 (PCNA) is necessary for continued association with the replisome (Krawitz, Kama & Kaufman, 2002; Zhang, Shibahara & Stillman, 2000).

In addition to depositing histones onto nascent DNA, CAF-1 has been implicated in other replication-independent processes. Yeast lacking components of CAF-1 show silencing defects at mating type loci and telomeres, hypothesised to be due to the instability of the silenced chromatin because of improper maintenance (Enomoto & Berman, 1998; Enomoto et al., 1997; Kaufman, Kobayashi & Stillman, 1997). Experimental evidence also exists linking CAF-1 to DNA repair pathways in a similarly replication-independent manner, such as by interaction with the KU70/80 DSB end capping complex (Gaillard et al., 1997; Green & Almouzni, 2003; Yang et al., 2017; Hoek, Myers & Stillman, 2011; Jiao et al., 2007).

1.9 Mass spectrometry identification and quantification of proteins

Mass spectrometry is an incredibly valuable tool in analysing the composition of complex samples, of both chemical and biological compounds (Finehout & Lee, 2004; Aebersold & Mann, 2003). By measuring the mass-charge ratio of fragments of a compound, the molecular weight can be determined. Mass spectrometry of biological compounds, specifically proteins, requires trypsin digest of the protein to produce many individual tryptic peptides, produced consistently by the specificity of the Trypsin enzyme (Aebersold & Mann, 2003). The masses of peptide fragments as measured in the spectrometer allows amino acid sequence information of the fragment to be obtained, as well as an absolute quantification of that peptide in the sample. Amino acid sequences can be uniquely mapped to their parent protein, and the quantification of peptides from the same protein can be used to provide a quantitative measure of protein abundance from a sample. Many techniques utilise the LC-MS/MS technique, which

involves liquid chromatography that separates the digested peptides, followed by injection and ionisation of the peptides by electrospray ionisation. The ionised peptides are detected in the first mass spectrometry spectrum (MS1), before selected peptides are fragmented further by energetic collision with a gas, and the second tandem mass spectrometry spectrum is obtained (MS/MS or MS2) (Aebersold & Mann, 2003).

The estimation of relative quantification of peptides between separate protein samples can be a challenge, and there are several techniques used to overcome this problem. The first is stable isotope labelling by amino acids in cell culture or SILAC. SILAC functions in the context of protein abundance analysis from living cells by labelling all proteins from one sample condition with a heavy labelled version of an amino acid, and the other condition with the standard weight amino acid (Ong et al., 2002). Replacement of the chosen amino acid with the heavy version in the media of one culture ensures that this heavy version is incorporated into all proteins of the cell within around five doubling times. The advantage of SILAC is that immediately after harvesting, the protein extract of the two samples being compared can be mixed and processed simultaneously. This helps reduce technical variability and allows direct comparative quantification of the protein's abundances after mass spectrometry analysis. SILAC is, however, limited in some other aspects. For complete incorporation, cells must be grown for a significant time in specialised media, and use of this media can be expensive. SILAC experiments are also most often carried out in pairs, with one condition containing heavy amino acid, and the other containing the regular amino acid (Ong et al., 2002). Some examples exist of multiplexed SILAC experiments up to 5-plexes, achieved by mixing multiple combinations of heavy amino acids (Tzouros et al., 2013), but in general SILAC experiments across multiple conditions are laborious and involve multiple pairwise comparisons. The second method, which involves the addition of different isotopic mass tags to digested peptides prior to mixing and mass spectrometry analysis, is known as tandem mass tag (TMT) labelling (Zhang & Elias,

2017). This method also requires addition of differentially heavy/light markers, but instead of in culture, these are chemically added after protease digestion. These tags are inherently more flexible, offering high sample number multiplexing, but are equally very expensive, and in the past this method has resulted in reduced precision and accuracy of peptide detection (Altelaar et al., 2013). The third method, and the one used in this study, is label-free quantification (LFQ) (Zhu, Smith & Huang, 2010; Cox et al., 2014). In this approach, samples are processed and analysed separately, and a normalisation step is applied after mass spectrometry quantification based on shared proteins and unchanging background quantification (Cox et al., 2014). Whilst this technique shares the common flaw with the TMT labelling method that protease digestion and all previous processing steps are carried out separately, and increase sample variability, the normalisation step can theoretically bring the larger variabilities back into comparability. LFQ normalisation has the advantage of being easier to set up for a large number of samples than SILAC, whilst being inexpensive compared to TMT labelling. In addition, LFQ-based approaches have been shown to produce better accuracy and precision than either other method, even though the absolute quantification of proteins is not as reliable (Li et al., 2012).

1.10 Cross-linking approaches to stabilise protein and protein-DNA complexes

The use of cross-linking in the investigation of protein-DNA interactions has been popularised by the chromatin immunoprecipitation (ChIP) protocol, which has been in stages of development and improvement since the late 1980s (Solomon, Larsen & Varshavsky, 1988), though the principle of complex stabilisation without antibody purification goes back further than that, to the 1960s (Das et al., 2004; Collas, 2010; Kuo & Allis, 1999). The isolation, identification and quantification of the protein-associated DNA was the main development in the ChIP protocol in the past couple of decades, with the advent of next generation sequencing methods allowing genome-wide analysis of protein-DNA association by ChIP-seq, as opposed

to the more locus-specific ChIP-chip or ChIP-qPCR techniques (Park, 2009; Collas, 2010). This technique has been a reliable staple method in biochemistry laboratories for many years, and so the fixation of cells with formaldehyde is a fairly common technique.

There have been several groups that have utilised the stabilisation of the formaldehyde induced cross-links to their advantage when studying more labile or transient complexes. In the early 2000s, Vasilescu et al. coupled in vivo cross-linking and affinity purification, followed by complex analysis by mass spectrometry to identify transient and easily missed protein-protein interactions of the M-Ras GTPase (Vasilescu, Guo & Kast, 2004). At the same time, Schmitt-Ulms et al. were using a similar method to attempt to study membrane protein interactomes, which are difficult to isolate from their natural state (Schmitt-Ulms et al., 2004), and Guerrero et al. developed "QTAX" to investigate the 26S proteasome interactome in yeast (Guerrero et al., 2006). Around a decade later, members of the Carroll lab were using a technique dubbed "rapid immunoprecipitation mass spectrometry endogenous proteins" (RIME) to investigate the interaction network of Estrogen-receptor-α (ER) (Mohammed et al., 2013), and Wang et al. were using "ChIP-MS" to investigate histone modifying proteins in *Drosophila* (Wang et al., 2013). Simultaneously, a thesis was submitted using a technique dubbed "proteome ChIP" (pChIP) to analyse large-scale interaction networks of differentially phosphorylated RNA polymerase II in transcription (Natarajan, 2013). A similar approach would be utilised again in the Kaiser lab in 2016 to investigate the UbR proteasome interactome (Yu et al., 2016). A FLAG-tag based "X-ChIP" protocol was also used to investigate the interactome of the mitochondrial helicase TWINKLE when overexpressed (Rajala et al., 2015). Very recently, an "XL-TAP-MS" approach was used in *Arabidopsis thaliana* to investigate the interactome of a kinase signalling module which was based on the work in the Kaiser lab (Leissing et al., 2021).

In a hybrid approach, Alabert et al. demonstrated a method termed "nascent chromatin capture" (NCC) which reversed the classical ChIP approach (Alabert et al., 2014). In this

method, newly synthesised DNA was labelled and specifically purified, and associated proteins with newly-synthesised and mature DNA were identified and compared using mass spectrometry.

These approaches have not seen much uptake within protein-protein interaction research generally. However, the re-establishment of similar techniques seems to be a result of the ability of formaldehyde cross-linking to investigate complexes where native affinity purification and mass spectrometry analysis cannot. Some of these approaches were designed around investigating very transient or labile complexes (Vasilescu, Guo & Kast, 2004; Mohammed et al., 2013; Schmitt-Ulms et al., 2004), or complexes that showed progression in a time-dependent manner and so must be fixed (Natarajan, 2013; Alabert et al., 2014), or just complexes that showed interactions that were difficult to identify under native conditions (Schmitt-Ulms et al., 2004; Mohammed et al., 2013). As such, the tagging steps and wash conditions of each of these protocols can be seen to vary widely, but the core principal is the same. Cross-linking stabilisation of protein complexes followed by purification of a protein of interest, can enable identification of interaction partners where other methods might not.

It is worth noting a disambiguation at this point. A similar technique is currently seeing popularity for structural, and also interaction, studies. This technique, often termed cross-linking mass spectrometry (XL-MS), does not generally use formaldehyde cross-linking (Yu & Huang, 2017; O'Reilly & Rappsilber, 2018; Mintseris & Gygi, 2020; Matzinger & Mechtler, 2021). Instead, the use of a mass spectrometry-cleavable crosslinker is used, which breaks down in the MS analysis pipeline in such a way that the identities of two peptides that were cross-linked can be identified. The expected length of the cross-linker can then be used to provide a constraint on the distance between those peptides at the point of fixation, providing structural information. In certain cases, this technique has been very successful, able to provide mechanistic data on cellular complexes such as ribosomes (Tüting et al., 2020). However,

although analysis techniques have been suggested that would allow this technique to work with the much more versatile formaldehyde molecule (Tayri-Wilk et al., 2020), this type of analysis attempts mainly to provide structural data over interaction data like the protocols mentioned above, and while considered carefully, was ultimately not utilised in this project.

1.11 Rapamycin-dependent linkages for study of MCM2-7 complex

MCM2-7, as shown above, is a ring-shaped hexamer of the six subunits Mcm2-7, and when loaded onto dsDNA, exists as a closed ring (Evrin et al., 2009; Remus et al., 2009; Sun et al., 2013). Whilst analysis of ATPase site mutants and complexes containing a subset of the Mcm subunits had suggested there was a unique structural discontinuation between Mcm2 and Mcm5 that might reflect a "gate" for the loading of the complex onto DNA (Bochman & Schwacha, 2007; Bochman, Bell & Schwacha, 2008; Bochman & Schwacha, 2010), this was yet to be shown conclusively. To test if the Mcm2-5 interface was indeed the possible path by which DNA was inserted into the MCM2-7 hexamer, a chemical biology approach was used (Samel et al., 2014). By creating an inducible physical block between Mcm2 and Mcm5, it was shown that not only was the interface between Mcm2 and Mcm5 a route for the passage of DNA, it was the only interface that, when blocked in such a manner, prevented helicase loading onto DNA *in vitro* and caused a cell cycle arrest in G1 *in vivo* (Samel et al., 2014). This construct utilised the FK506- and rapamycin- binding protein (FKBP) domain along with the FKBP-rapamycin binding (FRB) domain. These domains show a rapamycin-dependent interaction, forming a stable ternary complex, in which the binding of FKBP12 (the homolog gene from which the FKBP domain comes) with rapamycin bound, shows a K_D of 12±0.8 nM. The two domains also do not show any interaction in the absence of rapamycin.

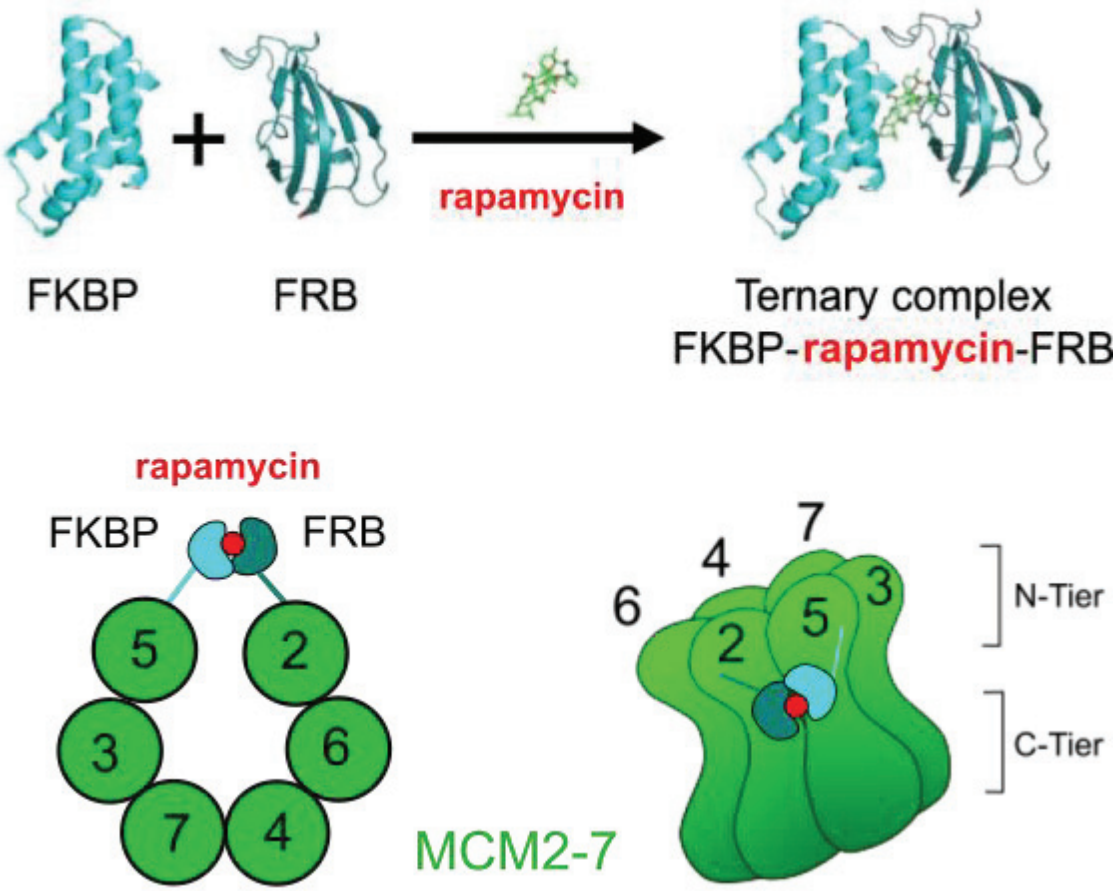


Figure 14: Schematic showing the rationale behind the chemical biology construct used in this project. S. Samel et al. (Samel et al., 2014) showed that these FRB and FKBP domains could be introduced on 20-amino acid linkers into appropriate flexible loops on the exterior of the MCM2-7 hexamer and cause a block of MCM2-7 loading, but only when introduced at the Mcm2-Mcm5 interface. This rapamycin-dependent intra-hexamer gate, when closed, prevented passage of dsDNA into the core of the hexamer, allowing loading at origins of replication. Figure adapted from Samel et al. 2014 (Samel et al., 2014).

Rapamycin is a potent antifungal and immunosuppressant that in yeast causes an irreversible cell cycle arrest in G1, that is effected by the three proteins Fpr1, Tor1 and Tor2 as part of the TOR complex (Heitman, Movva & Hall, 1991; Loewith et al., 2002). To demonstrate that the introduction of the rapamycin-dependent intra-hexamer gate linkage showed an effect on cell viability, a strain was constructed that had a mutation in Tor1 that conferred rapamycin insensitivity so that rapamycin treatment had no adverse effect in the absence of a rapamycin-dependent linkage construct (Freeman & Livi, 1996). In addition, the gene FPR1 was deleted, since it is FKBP-like and will bind FRB-rapamycin and prevent linkage across the interface (Gruber et al., 2006; Samel et al., 2014).

1.12 Project aims

As discussed above, the mechanism behind the process of DNA replication initiation is still mostly unknown. Whilst structural studies have shown the beginning and end stages of the activation reaction, in the form of the double hexamer and CMG, and a pre-MCM10 state has also been observed, the field is still awaiting further information regarding this process. With the reconstitution of the replication initiation reaction *in vitro*, there are potential avenues for structural studies into the process. However, the nature of the process may make attempts to visualise structural changes difficult. If the process is very rapid, and involves transient intermediates, then even time-resolved cryo-EM studies may struggle to obtain the necessary numbers of particles in the same conformational state to generate a high-resolution structure. Equally importantly, perhaps, is the fact that by using a reconstituted system, factors specific to the *in vivo* process may be neglected and remain undiscovered.

This project aims to investigate several steps of DNA replication *in vivo*, by a combination of methods and approaches that can overcome some of the intrinsic difficulties of studying this process.

1.12.1.1 Identification and ordering of key steps in replication initiation

Little is known about the structural changes that occur within the replication complex during activation. In particular, there is little evidence to show how two large structural changes that must occur in the transition from double hexamer to single-hexamer CMG are achieved. These are the selective extrusion of lagging strand ssDNA following initial unwinding of dsDNA, and the physical separation of the double hexamer into individual hexamers. It is uncertain whether these events are concerted or separate and the mechanisms behind them are also unknown. Because association of initiation factors is highly interdependent (Tanaka et al., 2011; Muramatsu et al., 2010), it has been difficult to determine whether any specific factor is required for any mechanistic event. So far, the limited evidence available suggests that the

association of the pre-LC complex causes disruption of the stable double hexamer complex (Miyazawa-Onami, Araki & Tanaka, 2017; Douglas et al., 2018), and a small amount of initial dsDNA unwinding in an *in vitro* system, and that the addition of Mcm10 enables further unwinding (Douglas et al., 2018). This project aims to identify and order key steps in the initiation process, and attempt to elucidate the factors involved, by using modified proteins *in vivo* to block structural events with physical linkages (Samel et al., 2014). By using multiple linkages designed to block the two gross structural changes mentioned above, complexes could potentially be arrested and studied at a crucial step in replication initiation. Comparative analyses of the effects of the different linkages aimed at each of lagging strand ssDNA and hexamer separation respectively could also provide an indication of the ordering of these two events in the initiation mechanism.

1.12.1.2 Proteomic analysis of replication complexes from loading to termination

In vivo DNA replication is a complex process, carried out and controlled by a wide array of interacting protein complexes and networks. Studies have shown in the past that affinity-purification using complex members or nascent DNA can successfully identify novel interacting and regulatory proteins (Gambus et al., 2006; Alabert et al., 2014). Mass spectrometry analysis of pull-down composition via bottom-up proteomics is an approach that is ideal for this purpose, since proteins can be detected and quantified without a priori knowledge of their identity (Richards, Eckhardt & Krogan, 2021). For this second aim of the project, use of an affinity enrichment and mass spectrometry-coupled approach was attempted to gain an unbiased identification of the DNA replication proteome. Specifically, this investigation would focus on the attempt to characterise the proteome of the DNA replication machinery not only during active replication, but during the helicase loading and helicase activation steps, which have not been specifically studied in such a interactome-wide manner before. This project aimed therefore to develop a reliable and robust approach that was able to

identify replication associated proteins, even if they formed transient or labile complexes, to identify proteins that were involved in key stages of DNA replication.

2 Materials and Methods

2.1 Materials

2.1.1 Plasmids

See Appendix 1 – Plasmids

2.1.2 Oligonucleotides

See Appendix 2 – Oligonucleotides

2.1.3 Yeast strains

See Appendix 3 – Yeast strains

2.1.4 Antibodies

See Appendix 4 – Antibodies

2.1.5 Media

Media was prepared in the lab or in media kitchen and autoclaved for sterility.

2.1.5.1 YPD and YPG

Unless otherwise specified, yeast cell cultures utilised YPD prepared as detailed here:

YPD – 1L

10 g yeast extract (Formedium yeast extract powder YEA03)
20 g peptone (Formedium peptone PEP03)
20 g glucose
HCl to pH to 5.5

For YPG, glucose was exchanged for galactose.

2.1.5.2 SDC/SRCtn/SGC minimal media

Minimal media was prepared lacking Ura, His, Trp, Leu or Met as required:

SDC – All – 1L

20 g glucose/raffinose/galactose
6.9 g yeast nitrogen base (Formedium CYN0410)
600 mg complete supplement mixture drop-out*: -His, -Leu, -Trp, -Ura (Formedium DCS1389)

Media was then supplemented with the required components (Ura/His/Trp/Leu)

* Methionine drop-out media required the use of complete supplement mixture drop-out: -Ade, -His, -Leu, -Met, -Trp, -Ura (Formedium DCS1519)

2.1.5.3 LB

LB media was prepared in the media kitchen, according to the recipe below:

LB – 1L

5 g yeast extract
10 g tryptone
10 g NacCl

2.2 Methods

2.2.1 Bacterial transformation

Chemically competent *E. coli* DH5α cells were used for standard plasmid amplification for newly ligated subcloned plasmid amplification. Cells were thawed on ice in 50 μL aliquots, mixed with the DNA for transformation, and incubated for a further 10 minutes on ice. Following heat shock at 42°C, cells were incubated on ice for a further 2 minutes, before addition of 800 μL LB buffer. Cells were incubated for at least 30 minutes, and up to 90 minutes, at 37°C with shaking at 800 rpm in a tabletop Eppendorf Thermomixer®. Finally, cells were spun down and resuspended in a smaller volume of LB, approximately 100 μL, and spread onto plates containing the selective medium.

2.2.2 Bacterial culture, plasmid extraction and purification

For growth of bacterial strains in liquid medium, LB containing the appropriate selective agent was used. Liquid culture was grown overnight at 37°C with 200 rpm shaking in an Eppendorf Excella® E25 incubator shaker. Dependent on desired yield of plasmid DNA, cells were grown in either 5 mL or 100-250 mL. Once cells had grown to an appropriate density, evaluated by eye, or after overnight incubation, cells were harvested by spinning at 4°C. Plasmid DNA was extracted from the pelleted cells using the QIAGEN QIAprep® Miniprep or QIAGEN Midiprep kits depending on the culture volume, following the manufacturer's instructions. Elution of plasmid DNA from the supplier columns in the final step was carried out using EB buffer, and for plasmids with low expected yields, the elution step was repeated with the same elution buffer, as recommended.

DNA concentration was determined using the NanoDrop™ 1000 or NanoDrop™ One spectrophotometer.

2.2.3 Subcloning of plasmid DNA

2.2.3.1 PCR reactions

For subcloning reactions, PCR was carried out using the PfuUltra Fusion II buffer and polymerase system. The reaction consists of the following components:

Reagent	Volume (µL)
PfuUltra Fusion II reaction buffer (10x)	5
dNTPs (10 mM)	5
Forward primer (10 µM)	1
Reverse primer (10 µM)	1
Plasmid DNA (1-10 ng)	1
PfuUltra Fusion II polymerase	1
Ultrapure water	36

The extension time of the PCR cycles varied dependent on the length of the region to be amplified, but an example cycling protocol is shown below.

Thermocycler amplification protocol:

95°C – 2 min.
95°C – 20 s
56°C – 20 s
72°C – 90 s
72°C – 5 min. } 32x
4°C - HOLD

PCR products were run on 1% agarose gels, and bands were visualised using a blue light transilluminator and excised from the gel. DNA was purified from gel bands using the QIAquick® gel extraction kit, following the manufacturer's instructions.

2.2.3.2 Restriction digest of DNA

Plasmid and PCR products were digested using NEB restriction enzymes and the rCutSmart™ unified buffer system. Digestion reactions were carried out at 37°C, and products run on 1% agarose gels, followed by band excision of the desired product and purification using the QIAquick® gel extraction kit.

2.2.3.3 Ligation

Ligation of purified, restriction endonuclease-cut plasmids and PCR products was carried out using the NEB Quick Ligation™ kit, following the manufacturer's instructions. Vector:insert molecular ratios were varied, but were mostly kept between 1:4 and 1:1. For low yield ligations, an additional phosphatase step was added, using the NEB Antarctic Phosphatase kit and protocol. Ligated product was taken forward directly for bacterial transformation as above.

2.2.4 Yeast transformation

Generation and transformation of competent yeast cells was carried out by an adapted protocol to that described by Michael Knop et al. (Knop et al., 1999). Yeast cells grown overnight were diluted to an OD_{600} of 0.2-0.3 in 50 mL and allowed to reach an OD_{600} of 0.5-0.7. Cells were spun at room temperature for 5 mins at 500 x g, and resuspended in 5 mL of sterile water, before being spun again. Cells were resuspended in 5 mL SORB and spun a further time. After resuspension in 360 µL SORB (100 mM LiAc, 10 mM Tris-HCl pH 8.0, 1 mM EDTA pH 8.0, 1 M Sorbitol (Merck microbiology grade)), 40 µL of pre-heated ssDNA (100°C for 10 mins,

then stored directly on ice) was added, and the cell solution was distributed in 50 µL aliquots before storage at -80°C or continuation directly to transformation step.

Competent cells prepared as above were transformed with either plasmids, restriction enzyme-cut linearised plasmids or plasmid fragments, or PCR-generated integrons designed with the rationale described by Carsten Janke et al. (Janke et al., 2004). The competent cell aliquots were thawed on ice, and the DNA added to them, along with 300 µL PEG solution (100 mM LiAc, 10 mM Tris-HCl pH 8.0, 1 mM EDTA pH 8.0, 40% PEG3350 (Sigma)). For plasmids, generally 300 ng was used per transformation, for plasmid fragments, up to 2/3 µg was used, and for PCR products, often the entire 50 µL PCR mix was added to cells. The solution was mixed and incubated at room temperature for 30 mins, or 1.5 hours if a PCR product was used. 45 µL of DMSO was then added, the solution mixed again, and cells were heat shocked for 15 mins in a 42°C water bath. Subsequently, cells were spun down at 2000 rpm for 3 mins in a tabletop micro-centrifuge, and resuspended. Cells to be selected based on auxotrophy were immediately resuspended in a small volume of minimal media and plated. Cells to be selected based on antibiotic resistance were resuspended in 800 µL YPD, and incubated at 30°C with 800 rpm shaking in a tabletop Eppendorf Thermomixer for 2-3 hours, before spinning again, and plating using a small volume of YPD for resuspension.

2.2.5 Yeast cell cycle synchronisation

2.2.5.1 Pheromone-induced arrest

Haploid yeast possess a pheromone-mediated cell cycle arrest, where cells of one mating type, typically **a**, can be arrested at START, a G1-phase state where cells are primed for mating with cells of the opposite mating type by exposure to the opposite mating type pheromone, in this case α-factor (Herskowitz, 1995; Gartner et al., 1998). Cells arrested in this way produce a projection known as a “shmoo” which is a convenient morphological readout of cell arrest. In

this state, cells have not entered S phase, and do not begin replicating their DNA, but can be synchronously and rapidly released into S phase by complete removal of the α-factor, which can be done in two quick wash steps with pheromone-free pre-warmed media.

Freely dividing cells growing at 30°C were treated with α-factor, (stock concentration 5mg/mL dissolved in 100% EtOH) to a total concentration of 5 ug/ml, and the temperature was lowered to 25°C. After one hour, the same amount of α-factor was added. After two hours of arrest, cells were checked for signs of "shmoo" formation, and released into pre-warmed media. Cells were spun down at 1,500 x g for 2 minutes at 25°C, and the pellet resuspended in a volume of pheromone-free media equal to the initial culture volume. Cells were spun again, and resuspended and left to proceed into S phase.

2.2.5.2 Nocodazole and Cdc20-depletion mediated arrest

Cdc20 regulates mitotic exit by activating the anaphase promoting complex (APC)(Dawson et al., 1993; Sigrist et al., 1995; Kramer et al., 1998; Uhlmann et al., 2000; Fang, Yu & Kirschner, 1998). By placing the sole copy of Cdc20 in a yeast cell under the control of the MET3 gene promoter, levels of Cdc20 in the cell can be controlled by the presence/absence of methionine (Mao et al., 2002), as utilised previously to cause cell cycle arrest in mitosis (Yeong et al., 2000). Addition of methionine causes cessation of Cdc20 transcription, and a lack of Cdc20 prevents mitotic exit, until Cdc20 expression is re-established when an external source of methionine is removed.

As described previously (Uhlmann et al., 2000), MET3-CDC20 cells were grown generally in minimal media lacking methionine. In these experiments, to ensure efficient eventual release, cells were arrested in mitosis using minimal media replacement with YPD media for 2.5 hours at 25°C, before washing in pre-warmed media as described for α-factor removal, using the minimal media lacking methionine to induce Cdc20 expression and release into the cell cycle.

2.2.6 Yeast cell cycle analysis by flow cytometry

To track cell populations through the cell cycle, flow cytometry with SYTOX Green DNA stain was used (Haase & Reed, 2001). 1 mL of cell culture was spun down at 8,000 x g, and the cell pellet resuspended in 70% EtOH and stored at 4°C for a minimum of o/n. Subsequently, cells were resuspended in 800 µL of 50 mM sodium citrate buffer (pH 7), and sonicated to separate cell clumps to ensure downstream single cell analysis. Cells were spun down and resuspended in 600 µL sodium citrate buffer supplemented with a total concentration of 0.25 mg/mL RNase A, and incubated with 1400 rpm shaking at 37°C for 2.5 hours in a tabletop Eppendorf Thermomixer. Cells were washed once with 800 µL sodium citrate before final resuspension in 600 µL sodium citrate supplemented with SYTOX Green (S7020, Invitrogen) to a final concentration of 2 µM. After incubation at 4°C o/n, samples were run on a BD LSRFortessa™. Downstream analysis was carried out using FlowJo v10.6.1. For quantification of cell cycle stages, cells were sorted into bins based on SYTOX intensity. Bins for each of G1, S and G2 were based on the peaks observed in cycling cell populations.

2.2.7 Mating type testing and HO-nuclease induced switching

To test for mating type, a pair of strains of mating type a and α respectively were used. These strains lack the ability to grow on minimal media alone, without complementation of their auxotrophies by mating with another strain. However, these strains can complement the typical auxotrophies present in laboratory yeast strains (e.g. ura3-1, his3-11,15, trp1-1 and leu2-3). By individually plating these test strains together with the strain requiring mating type testing on minimal media, the mating type can be determined. Only those cells able to mate can grow on SDC-All media, which provides a readout for the original mating type of the strain of interest. Induction of mating type switching in strains was carried out by transformation with a pGal-HO containing plasmid and selection for plasmid uptake by plating on appropriate media

(Herskowitz & Jensen, 1991). Colonies positive for plasmid uptake were grown overnight in minimal selective media containing raffinose as the primary sugar source (SRC), to aid in galactose induced expression of the HO endonuclease from the plasmid. The following day, cells were resuspended and grown in minimal selective media containing galactose (SGC) for six hours at 30°C to induce expression. Finally, cells were streaked on glucose-containing minimal media (SDC) plates at colony forming density. Colonies from a single cell were isolated and tested for mating type as described above.

2.2.8 Viability assays on agar plates

Viability assays were carried out to test the growth of yeast strains under normal and selective/stressor conditions. An inoculation loop of cells from the strain to be tested was resuspended thoroughly in 1 mL double distilled water (ddH_2O). This starting solution was diluted 1 in 10 into 900 μL of ddH_2O in a separate tube. This solution was the first seeding solution. This first seeding solution was sequentially diluted 1 in 10 three or four further times to generate four/five diluted yeast solutions. 5 μL of each solution was sequentially plated in a grid pattern on YPD agar or on selective media plates, and allowed to grow at 30°C for several days, until growth was observed to a sufficient level allowing comparison of strain growth over a range of cell concentrations.

2.2.9 Chromatin immunoprecipitation (ChIP-qPCR) of replication proteins

ChIP-qPCR was carried out as in prior work by L.M. Reuter (Reuter, Meinel & Sträßer, 2015). For the majority of experiments, the ChIP-qPCR approach was carried out on strains containing endogenously 5xFLAG-tagged proteins. As such, the antibodies and beads used are specific for purification of FLAG-tagged proteins. For the TAP-tagged examples, the anti-FLAG antibody and protein G beads were replaced with non-specific IgG-bound tosyl-activated magnetic beads, detailed in the ChIP procedure section below.

2.2.9.1 Cell culture and formaldehyde cell fixation

The endogenously tagged strains of *S.cerevisiae* were grown overnight in 100-300 mL appropriate media at 30°C. The following morning, the culture was diluted to an OD_{600} of 0.2-0.3 in an appropriate volume for the experimental design, with each ChIP-qPCR sample requiring 100 mL of culture. Cells were allowed to grow at 30°C until the OD_{600} was between 0.5 and 0.7. At this stage, cells were arrested by either of the two methods detailed above (Section 2.2.5.1 and 2.2.5.2). In experiments where rapamycin treatment was required, this was added in an additional 30 minutes of arrest, and also re-added to the final released S phase culture. After release into S phase at 25°C, and at the appropriate time, cells were fixed with formaldehyde treatment. 37.5% formaldehyde (Sigma) was added to a final concentration of 1%, and incubated with mild agitation for 20 minutes. After 20 minutes, the cross-linking activity was quenched with the addition of 3 M glycine solution to a final concentration of 0.25 M. After 5 minutes, cells were placed on ice, before being harvested at 4°C, washed with chilled Tris-buffered saline (1xTBS), and pellets flash frozen in liquid nitrogen, before storage at -80°C.

2.2.9.2 ChIP-qPCR procedure

On a following day, the pellets were thawed on ice and resuspended in 800 µL of chilled FA lysis buffer-low salt (50 mM HEPES-KOH pH 7.5, 150 mM NaCl, 2 mM EDTA pH 8.0, 1% Triton X-100 (v/v), 0.1% Na-Deoxycholate (w/v) and 0.1% sodium dodecyl sulphate (w/v)) which had been pre-supplemented with a protease inhibitor cocktail (SIGMAFAST™ tablets EDTA-free). After addition of an equal volume of glass beads, cells were lysed with three 45-second cycles of bead-beating using the MP FastPrep-24 5G grinder and lysis system set to a speed of 7 m/s, with 2 minute incubations on ice in between cycles. Lysates were separated from beads and diluted to a total volume of 1.3 mL with FA lysis buffer-low salt also supplemented with protease inhibitor cocktail.

The lysate was then sonicated for 45 minutes (30 seconds ON/OFF) using a Bioruptor (Diagenode) with integrated 4°C water cooling to generate chromatin fragments between 250 and 500 base pairs in length. After sonication, the lysate was clarified by two rounds of centrifugation using a benchtop centrifuge (13,000 rpm, 4°C) for 5 and 10 minutes respectively. Protein amounts in the supernatant were quantified by spectrophotometry (A_{280}), and the input amounts across all samples equalised to the lowest concentration sample, of which 1 mL was taken forward, and 100 µL was taken for chromatin size analysis, see below. 10 µL of each sample was taken as an input sample, and the rest was incubated with 4 µL anti-FLAG antibody (Sigma, F-1804) with rotation at room temperature for 1.5 hours. Following the antibody incubation, 15 µL of protein G magnetic beads (Protein G Dynabeads, 1004D) was added and incubation continued for another hour with rotation. Beads were collected on a magnetic rack and washed with several sequential washes; 2x with 800 µL FA lysis-low salt, 2x with 800 µL FA lysis-high salt (identical to low salt but with 500 mM NaCl), 2x TLEND (10 mM Tris-HCl pH 8.0, 250 mM LiCl, 1 mM EDTA pH 8.0, 0.5% NP-40 (v/v) and 0.5% sodium deoxycholate (w/v)) and 1x with Tris-EDTA (TE pH 8.0). Following these wash steps, FLAG-tagged protein-DNA complexes were eluted from the beads with 25 minutes incubation at 65°C with shaking at 1400 rpm in 130 µL elution buffer (50 mM Tris-HCl pH 7.5, 10 mM EDTA pH 8.0 and 1% sodium dodecyl sulphate (w/v)).

2.2.9.3 DNA purification and qPCR

After elution of tagged DNA-protein complexes, both input and IP samples were treated with Proteinase K (10 µL, Sigma P4850) for 2 hours at 37°C to degrade protein, and incubated at 65°C for a further 12-16 hours to de-crosslink the samples. The samples were purified using the QIAGEN QIAquick PCR Purification kit as per the manufacturer's instructions, and eluted in 140 µL elution buffer (EB, 10 mM Tris-Cl, 1 mM EDTA, pH 8.0). The DNA solutions obtained was diluted 1:4 for qPCR analysis of the relative DNA amounts in input and IP.

Analysis of quantitative data was carried out as described in (Reuter, Meinel & Sträßer, 2015). Results of quantitative data is presented with the standard deviation or standard error of the mean, where appropriate. Pairwise statistical testing was not carried out as a matter of course as standardised in the field (see e.g. (Miyazawa-Onami, Araki & Tanaka, 2017)

2.2.9.4 Chromatin size analysis

Samples taken for chromatin size analysis were similarly treated with Proteinase K and incubated at 37°C for 2 hours, then de-crosslinked overnight at 65°C. The following day, 30 µL of 4M LiCl was added to the sample, followed by 1 µL glycogen and 150 µL phenol. The samples were mixed and centrifuged in a desktop centrifuge (13,000 rpm, 10 min.) to separate the aqueous and organic phases. The supernatant was taken, and a further 150 µL of phenol added, before another centrifugation step. The supernatant was again isolated, and 100 µL chloroform added, followed by another centrifugation step (13,000 rpm, 5 min.). The supernatant of this step was added to 800 µL 100% EtOH, and placed at -20°C overnight to allow precipitation of the DNA.

The following day, the sample was centrifuged (13,000 rpm, 10 min, 4°C), and the pellet washed with cold 70% EtOH, before being resuspended in 20 µL TE pH 8.0. 5 µL RNase was added and the sample incubated at 37°C for at least 30 minutes. To visualise the size of the chromatin fragments present in the cell extract post-sonication, this sample was loaded onto a 1.5% agarose gel containing SYBR safe for electrophoresis and subsequent UV visualisation.

2.2.10 Native immunoprecipitation of Psf2-TAP and Cdc45-FLAG

Native immunoprecipitation of replication proteins was carried out using an adapted protocol of the first step of purification used by Gambus et al. to isolate GINS-containing replication complexes (Gambus et al., 2006b).

2.2.10.1 Preparation of IgG-bound magnetic beads

The preparation step of the beads was a process optimised over the course of the project, and so presented here is the most recent version of the methodology.

100 uL of bead slurry (Dynabeads® MyOne™ Tosylacttivated, Invitrogen 66501) was collected on a magnetic rack and the storage buffer removed. Beads were washed with 1 mL borate buffer pH 9.5. After washing, the beads were resuspended in 130 uL borate buffer, and 100 uL of ammonium sulfate (3 M) in borate buffer (0.1 M) to form a binding solution, that 1 mg of IgG was added to (rabbit, ChromPure 011-000-003). Beads and binding solution were incubated overnight at 37°C with vigorous shaking. The next day, beads were collected on the rack and binding solution removed. Beads were then washed with 1 mL phosphate-buffered saline (PBS, pH 7.4) and washed for one hour at 37°C with shaking. After washing, the beads were resuspended in 335 uL storage buffer.

Borate buffer 0.1M pH 9.5 – 1 L

2.18 g H_2BO_3

800 mL ddH_2O

Adjusted to pH 9.5 using 5 M NaOH

Volume adjusted to 1 L using ddH_2O

Ammonium sulfate in borate buffer – 100 mL

39.64 g ammonium sulfate

80 mL 0.1 M borate buffer pH 9.5

pH adjusted to 9.5 using NaOH or HCl

volume adjusted to 100 mL using 0.1 M borate buffer

Storage solution

0.1 M PBS pH 7.4

0.05% high quality BSA (w/v)

0.02% NaN_3

2.2.10.2 Cell culture, harvesting and extract preparation

Overnight cultures were grown at 220 rpm shaking at 30°C, before transfer to the desired media volume in High-Yield flasks following dilution to OD_{600} of around 0.2. Cells were allowed to grow until OD_{600} of around 0.5-0.6, and the alpha factor was added to arrest cells as described above. Cells were released into S phase by spinning down and media replacement, with a wash step, using pre-warmed media. At the desired time-point, cells were transferred into pre-cooled centrifuge bottles and spun down (1000 x g, 3 min, 4°C). Following a wash step in cold ddH_2O, cells were spun again, and resuspended in pre-cooled Lysis buffer, of volume equal to 1.5x the weight of the pellet. Yeast "popcorn" was prepared as described previously (Seki & Diffley, 2000), and stored at -80°C.

To prepare cell extract, the popcorn was ground in a freezer mill (Spex 6875 Freezer/Mill®) and additional Lysis buffer (1 pellet volume) and Glymix (0.25x total extract volume) added. 800 U/mL Benzonase® (Merck Millipore Novagen® Purity >99%) was added and sample incubated with rotation at 4°C for 30 minutes. Insoluble material was removed by an initial centrifugation (SS34 rotor, 14,000 rpm, 4°C, 10 min.), and further ultracentrifugation (50.2 Ti, 31,000 rpm, 4°C , 60 min.) where both the pellet and fatty surface layer were discarded.

2.2.10.3 IP binding, wash, and elution steps

For the IP step, 2.5 µL of IgG beads, prepared as described above, were added for every mg of protein as measured by a Bradford assay (for 3 L culture, approx.. 50 µL beads) . Samples were incubated at 4°C with rotation for 2 hours, before collection on a magnetic rack and removal of the flow-through. Beads were washed three times with Column wash buffer, each for 10 minutes also at 4°C. Finally, beads were washed once in Cleavage buffer, followed by elution in Cleavage buffer containing PreScission protease (in-house) at 16°C with shaking at 500 rpm for one hour. Beads were separated using the magnetic rack, and eluate collected.

2.2.10.4 Native IP-MS buffers

Lysis buffer

100 mM HEPES KOH pH7.9

50 mM KOAc

1.5 mM MgAc

100 µM EDTA

50 µM ZnOAc

Added fresh:

2 mM NaF

2 mM Na-β-glycerophosphate

1 mM DTT

4x protease inhibitors (SIGMAFAST™ tablets EDTA-free)

Glymix

100 mM HEPES KOH pH7.9

300 mM KOAc

1.5 mM MgAc

100 µM EDTA

50 µM ZnOAc

50% Glycerol v/v

0.75% NP-40 v/v

Added fresh:

2 mM NaF

2 mM Na-β-glycerophosphate

1 mM DTT

4x protease inhibitors (SIGMAFAST™ tablets EDTA-free)

Column wash buffer

100 mM HEPES KOH pH7.9

300 mM KOAc

1.5 mM MgAc

50 μM EDTA

50 μM ZnOAc

0.15% NP-40 v/v

Added fresh:

1 mM DTT

Cleavage buffer

100 mM HEPES KOH pH7.9

150 mM KOAc

1.5 mM MgAc

50 μM EDTA

50 μM ZnOAc

0.15% NP-40 v/v

Added fresh:DTT

2.2.11 ChIP-MS analysis of replication complexes

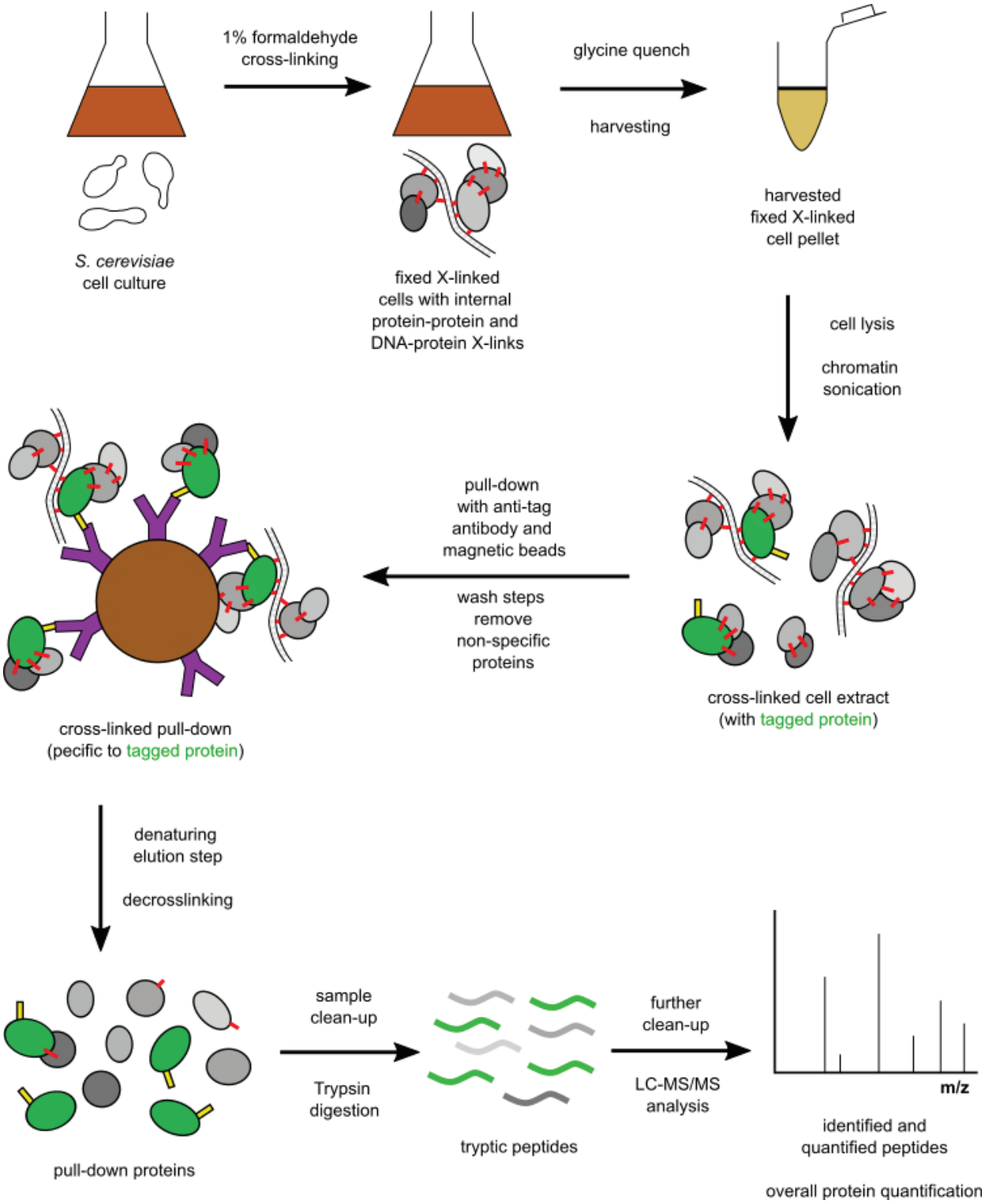


Figure 15: Overview of the ChIP-MS method as developed and performed in this study. Use of cross-linking fixes cells and stabilises protein complexes, before isolation of protein complexes using endogenously tagged protein of interest and immuno-precipitation and wash steps using magnetic beads. Isolated proteins are de-cross-linked and digested with trypsin to allow for LC-MS/MS analysis of protein identification and quantification.

Cells were grown in 600ml YPD from overnight cultures to OD600 0.6 at 220 rpm and 30°C followed by 2 hrs α-factor arrest at 25°C, 220 rpm. Cells were spun down (500 xg, 2 min, RT) washed once with pre-warmed media and released in 600 ml warm media. For samples where FRB/FKBP linkage closure was desired, media was supplemented with 100 nM rapamycin 0.5 hrs before releasing and upon release. At the desired time-point post-release, samples were crosslinked and washed as described for ChIP-qPCR. Each sample was resuspended in 1x TBS, before being spun again (8,000 xg, 4 min, 4°C) and the pellets frozen in liquid nitrogen and stored at -80 °C. Pellets were thawed on ice and resuspended in 800 uL of FA lysis (low salt) buffer complemented with a protease inhibitor cocktail, with 800 uL of glass beads added to each. Cells were lysed using an MP Fast-Prep 24 5G for 3x 45 seconds at 7 m/s with 2 min breaks on ice. Protease inhibitor was added (SIGMAFAST™ tablets EDTA-free), and sonication and centrifugation were done as described for ChIP-qPCR. Protein content of the supernatant was measured (A_{280}), and total protein amounts equated between samples by dilution, with input samples taken for quality control. The extract was then incubated with 7.2 uL anti-FLAG antibody (Sigma, F-1804) for 1.5 hrs at RT with rotation. 26 uL protein G Dynabeads (Invitrogen, 10004D) were added, and samples incubated for 1 hr. Beads were collected on a magnetic rack and subjected to washing (2x FA-lysis (low salt), 2x FA-lysis (high salt), 2x TLEND and 1x TE buffer) before combining the tubes and elution with ChIP elution buffer as done for ChIP. Samples were de-crosslinked at 95°C for 15 min and analysed by SDS-PAGE followed by silver stain.

Immunoprecipitations (IP) were processed by either filter aided sample preparation (FASP) (Wiśniewski et al., 2009), or single-pot solid-phase-enhanced sample-preparation (SP3) technology (Hughes et al., 2014, 2018).

Overall, the culture conditions, extract preparation, wash steps, elution steps, preparation for mass spectrometry and mass spectrometry analysis were dependent on the experimental design, and varied between experiments.

2.2.12 Sample preparation for Mass spectrometry

Because the protein samples prepared using the described protocols contained detergents, which cause issues with mass spectrometry analysis, the sample buffer needed to be exchanged in order for sample analysis. In addition, the trypsin digestion of proteins into the mass spectrometry-identifiable tryptic peptides needed to be carried out. Initially, the sample preparation steps were carried out by members of the MS facility at the MRC-LMS, but over the course of the project, the preparation steps were integrated into the protocol to relieve the workload of the MS facility. There were two main approaches that were used in this project, with a switch from one approach to the other at around the midway point. The first, filter aided sample preparation (FASP), carries out buffer exchange using a 10 kDa filter, that retains proteins but allows the washing of buffer and other small molecules through (Wiśniewski et al., 2009). The second, single-pot solid-phase enhanced sample preparation (SP3), carries out buffer exchange by immobilising proteins to magnetic beads by using an ethanol solution to cause hydrophilic association of proteins with the surface of the beads, allowing sequential washes not unlike the use of magnetic beads in an immunoprecipitation (Hughes et al., 2014, 2018). Owing to several factors, including issues with large amounts of detergent causing blockage of FASP columns, the overall ease of the SP3 method, and literature evidence suggesting SP3 might outperform FASP in head-to-head studies (Sielaff et al., 2017), SP3 became the sample preparation of choice for this project.

Whilst these methods were used to remove as much detergent as possible, a further cleanup step was the use of Glygen C18 reverse phase tips to reduce the amount of salts carryover.

2.2.12.1 FASP preparation

Eluted protein samples were loaded onto 30 kDa centrifugal concentrators (Millipore, MRCF0R030) and buffer exchange was carried out by centrifugation on a bench top centrifuge (10 min, 12,000 xg). Multiple buffer exchanges were performed sequentially with UA buffer (8 M urea in 100 mM Tris pH 8.5, 3x200 µl), reduction with 10 mM DTT in UA buffer (30 min, 40°C) and alkylation with 50 mM chloroacetamide in UA buffer (20 min, 25°C). This was followed by buffer exchange into UA buffer (3x100 µl) and 50 mM ammonium bicarbonate (3x100 µl). Digestion was carried out with mass spectrometry grade trypsin (Promega, V5280) using 2 & 1 µg of protease for inputs and IP's respectively (o/n, 37°C). Tryptic peptides were collected by centrifugation into a fresh collection tube (10 min, 12,000 xg) and washing of the concentrator with 0.5 M sodium chloride (50 µl, 10 min, 12,000 xg) for maximal recovery. Finally, digests were acidified to a final concentration of 0.1% TFA. Digested samples were then transferred to the Gygen C18 clean-up step detailed below.

2.2.12.2 SP3 preparation

Eluted samples were mixed with reconstitution buffer (50 mM HEPES pH 8.0, 1% sodium dodecyl sulphate (w/v) SDS, 1% Triton X-100 (v/v), 1% NP-40 (v/v), 1% Tween 20 (v/v), 1% sodium deoxycholate (w/v), 5 mM EDTA pH 8.0, 50 mM NaCl, 1% glycerol (v/v)). Proteins were reduced using 5 mM DTT, followed by alkylation using 20 mM chloroacetamide. Reduced and alkylated protein was transferred to a new tube containing 100% EtOH pre-mixed with Sera-Mag Speedbeads (GE Healthcare, 2 uL of each per sample). Reduced and alkylated protein was added, to a final concentration of 60% EtOH. Protein-bound beads were rinsed three times with 80% EtOH using a magnetic rack, and after the final wash, beads were resuspended in digestion solution (100 mM ammonium bicarbonate, pH 8.0) containing 1 ug of Trypsin Gold (Promega). After sonication to disperse the beads, digestions were incubated overnight (18 hrs) at 37°C and dried

in a centrifugal evaporator. Total protein digests were dissolved in 0.1% trifluoroacetic acid by shaking (1200 rpm) for 30 min and sonication for 10 min, followed by centrifugation (14,000 rpm, 5°C and 10 min). Digested samples were then transferred to the Gygen C18 clean-up step detailed below.

2.2.12.3 Reverse phase Glygen C18 tip cleanup

Peptide solutions from FASP or SP3 preparation steps were acidified by addition of formic acid, to a total concentration of 0.1%, then loaded onto a C18 reverse-phase desalting tip (Glygen, TT2C18.96) that had been pre-conditioned with 2 washes of organic solvent (60% MeCN, 0.1% formic acid (FA)) followed by 2 washes of aqueous solvent (1% MeCN, 0.1% formic acid (FA)). Once the peptide solution had been applied, the tip was washed with 2 further washes of 75 uL aqueous wash solution (1% MeCN, 0.1% FA), then eluted by 2 washes with 75 uL of organic wash solution (60% MeCN, 0.1% FA). The peptides were air-dried in a SpeedVac, then resuspended in 20-40 uL 0.1% triflouroacetic acid (TFA), and agitated for 30 minutes. Samples were sonicated for 10 minutes (1s ON/OFF) to fully dissolve peptides, before centrifugation to pellet any precipitates (13,000 rpm, 10 min). The majority of sample was transferred into sample injection vials, leaving a small volume containing any precipitates.

2.2.13 Mass spectrometry analysis and MaxQuant processing

Throughout the project, across experiments, there was some variation of the LC-MS/MS analysis method used. However, all methods shared common aspects, and the conditions used for the multi-replicate large-scale experiments will be specified individually. All other experiments used the same set of instruments and techniques, with slight variations. The programming and monitoring of the sample processing on the LC-MS/MS instrumentation and the analysis of the raw MS/MS files using MaxQuant was carried out by the MS facility team

(H. Kramer, A. Montoya, I. Khaneva), and the MaxQuant output files including LFQ normalisation were then taken over for analysis as described in this project.

2.2.13.1 Mass spectrometry experimental conditions and instruments

Across all samples and experiments, samples were analysed in technical duplicates, and separation prior to analysis by MS/MS was achieved using an Ultimate 3000 RSLC nano liquid chromatography system (Thermo). For the MS analysis, liquid chromatography was coupled to the MS instrument by an EASY spray source (Thermo). Samples were either analysed on a Q-Exactive or Q-Exactive HFX mass spectrometer (Thermo), depending on the experiment.

2.2.13.1.1 Cdc45 ChIP-MS (Section 3.4) mass spectrometry analysis

Having been prepared using the FASP technique, Trypsin Gold digestion, and C18 reverse phase cleanup as described, 6 uL of each sample was processed via liquid chromatography prior to MS analysis. Solutions were injected and loaded onto a trap column (Acclaim PepMap 100 C18, 100 µm × 2cm) for desalting and concentration at 8 µL/min in 2% acetonitrile, 0.1% TFA. Peptides were then eluted on-line to an analytical column (Acclaim Pepmap RSLC C18, 75 µm × 75 cm) at a flow rate of 200 nL/min. Peptides were separated using a 120 minute gradient, 4-25% of buffer B (80% acetonitrile, 0.1% FA) for 90 minutes followed by 25-45% buffer B for another 30 minutes and subsequent column conditioning and equilibration. Peptides were then analysed by the Q-Exactive mass spectrometer operating in positive polarity using a data-dependent acquisition mode. Ions for fragmentation were determined from an initial MS1 survey scan at 70,000 resolution, followed by HCD (Higher Energy Collision Induced Dissociation) of the top 12 most abundant ions at 17,500 resolution. MS1 and MS2 scan AGC targets were set to 3e6 and 5e4 for maximum injection times of 50 ms and 50 ms respectively. A survey scan m/z range of 400 – 1800 was used, normalised collision energy set to 27%, charge exclusion enabled with unassigned and +1 charge states rejected and a minimal AGC target of 1e3. Dynamic exclusion parameter set to 45 seconds.

2.2.13.1.2 Mcm4 ChIP-MS time-course (Section 3.6.3) mass spectrometry analysis

Having been prepared using the SP3 technique, Trypsin Gold digestion, and C18 reverse phase cleanup as described, 6 uL of each sample was processed via liquid chromatography prior to MS analysis. Solutions were injected and loaded onto a trap column (Acclaim PepMap 100 C18, 100 µm × 2 cm) for desalting and concentration at 8 µL/min in 2% acetonitrile, 0.1% TFA. Peptides were then eluted on-line to an analytical column (Acclaim Pepmap RSLC C18, 75 µm × 75 cm) at a flow rate of 200 nL/min. Peptides were separated using a 90 minute gradient, 4-25% of buffer B (80% acetonitrile, 0.1% FA) for 60 minutes followed by 25-45% buffer B for another 30 minutes and subsequent column conditioning and equilibration. Peptides were then analysed by the mass spectrometer operating in positive polarity using a data-dependent acquisition mode. Ions for fragmentation were determined from an initial MS1 survey scan at 70,000 resolution, followed by HCD (Higher Energy Collision Induced Dissociation) of the top 10 most abundant ions at 17,500 resolution. MS1 and MS2 scan AGC targets were set to 3e6 and 5e4 for maximum injection times of 50 ms and 100 ms respectively. A survey scan m/z range of 350 – 1800 was used, normalised collision energy set to 27%, charge exclusion enabled with unassigned and +1 charge states rejected and a minimal AGC target of 1e3. Dynamic exclusion parameter set to 45 seconds.

2.2.13.1.3 Mcm4 ChIP-MS in intra-hexamer gates (Section 3.7.1) mass spectrometry analysis

Having been prepared using the SP3 technique, Trypsin Gold digestion, and C18 reverse phase cleanup as described, 6 uL of each sample was processed via liquid chromatography prior to MS analysis. Solutions were injected and loaded onto a trap column (Acclaim PepMap 100 C18, 100 µm × 2 cm) for desalting and concentration at 8 µL/min in 2% acetonitrile, 0.1% TFA. Peptides were then eluted on-line to an analytical column (Acclaim Pepmap RSLC C18,

75 µm × 50 cm) at a flow rate of 250 nL/min. Peptides were separated using a 120 minute gradient, 1-22% of buffer B (75% acetonitrile, 5% DMSO, 0.1% FA) for 90 minutes followed by 22-42% buffer B for another 30 minutes, and subsequent column conditioning and equilibration. Peptides were analysed by the Q-Exactive HFX mass spectrometer operating in positive polarity using a data-dependent acquisition mode. Ions for fragmentation were determined from an initial MS1 survey scan at 120,000 resolution, followed by HCD (Higher Energy Collision Induced Dissociation) of the top 25 most abundant ions at 15,000 resolution. MS1 and MS2 scan AGC targets were set to 3e6 and 5e4 for maximum injection times of 25 ms and 100 ms respectively. A survey scan m/z range of 350 – 1750 was used, normalised collision energy set to 27%, charge exclusion enabled with unassigned and +1 charge states rejected and a minimal AGC target of 8e3. Dynamic exclusion parameter set to 50 seconds.

This was the only set of samples to be run using the Q-Exactive HFX instrument, all other experiments used the Q-Exactive instrument. One experiment was analysed by both instruments, and the number of proteins identified and their quantification appeared broadly similar (data not shown).

2.2.13.2 Maxquant processing

Raw MS files were analysed in the MaxQuant software (Cox & Mann, 2008; Cox et al., 2014; Tyanova, Temu & Cox, 2016). The MaxQuant analysis was conducted in a broadly similar manner across all experiments, with some minor changes to reference databases and software versions over time. As before, the precise conditions for the large-scale experiments are listed below, and all other experiments were processed very similarly.

2.2.13.2.1 Cdc45 ChIP-MS (Section 3.4) MaxQuant analysis

MaxQuant analysis was carried out as for Section 3.7.1 (see below) but with the following differences: Raw files were analysed using MaxQuant v1.6.7.0 not v1.6.10.43, and only two

missed cleavages were permitted. ‘Match between runs’ function was used with match and alignment time limits of 1 and 20 minutes respectively (inputs).

2.2.13.2.2 Mcm4 ChIP-MS time-course (Section 3.6.3) MaxQuant analysis

MaxQuant analysis was carried out as for Section 3.7.1 (see below) but with the following differences: Raw data was searched against the UniprotKB S.cerevisiae database downloaded in June 2021 instead of from March 2018. This contained 6,050 protein sequences: 6,049 from the SwissProt reviewed section and 1 from the TrEMBL unreviewed section

2.2.13.2.3 Mcm4 ChIP-MS in intra-hexamer gates (Section 3.7.1) MaxQuant analysis

Raw files were analysed using MaxQuant v1.6.10.43. Variable modifications of methionine oxidation, protein N-terminal acetylation, asparagine deamidation and cyclisation of glutamine to pryo-glutamate were permitted. Fixed modification of cysteine carbamidomethylation and up to three missed cleavages were also permitted. Label-free quantification was enabled with an LFQ minimum ratio count of 2. ‘Match between runs’ function was used with match and alignment time limits of 0.7 and 20 minutes respectively (inputs). Raw data was searched against the UniprotKB S.cerevisiae database downloaded in March 2018. This contained 6,729 protein sequences: 6,721 from the SwissProt reviewed section and 8 from the TrEMBL unreviewed section. A 1% false discovery rate (FDR) at the peptide and protein level was selected.

2.2.14 Downstream analysis of Mass spectrometry data

The data produced by MaxQuant analysis of raw MS files was analysed by two main pipelines. The Perseus software package (Tyanova et al., 2016), was used for much of the analysis of the dataset holistically, but manual analysis of LFQ values for sets of or individual protein was carried out in Microsoft Excel.

2.2.14.1 Perseus analysis

The Perseus software allows a reproducible pipeline to be used across experiments, though, as with MaxQuant, there were some differences in versions that were used across the project. The overall pipeline remained the same, however, and is described below. Analysis in MaxQuant is able to identify non-yeast proteins that are common contaminants, as well as proteins that are only identified by site. “proteins” that were identified using a reversed proteome are also indicated. These types of matches were removed at the beginning of the analysis. Following this step, the log value (base 2) of the LFQ intensity values for the detected proteins was calculated, to obtain a normal distribution of intensity values. It was at this stage that the histograms of protein abundances were produced, to determine if the data was indeed normally distributed. At this stage, if the data was shown to be lognormal, then the experimental and replicate groups were assigned. The technical replicates run on the instrument were averaged, excluding values of zero. Following technical replicate averaging, filtering of proteins that showed very sporadic and sparse detection was carried out. For a protein to be taken forward, it had to be present in at least two of the replicates in at least one experimental group. After filtering, imputation was used to replace missing values in the data, e.g. in negative control experimental groups. The built-in imputation functionality in Perseus was used for this step. In brief, Perseus calculates the width and centre of the lognormal distribution of abundances, scales the width to 0.3 and down-shifts the distribution by 1.8 standard deviations, before populating this new distribution with randomly assigned values. This imputes values for samples that do not show detection that are appropriately low. At this stage, the data processing was complete, and the various correlative tests were carried out on the data. Principal component analysis, intensity vs. intensity scatter plots, volcano plots and profile plots were all carried out on data processed in this way, using the built-in Perseus functionality. Volcano plots were produced using FDR and S0 values as indicated.

2.2.14.2 Manual analysis of peptide intensities and enrichments

As in the Perseus-based analysis, the protein matches that were identified as common contaminants, only identified by site, or as reverse proteome decoy hits were removed. Following this step, LFQ intensity values between technical reps were averaged, treating non-detections as non-integer values. In experiments containing biological replicates, the LFQ intensity of proteins across these replicates was also averaged, again treating non-detected protein intensity as non-integer values and not calculated. For enrichment values in both replicate-containing and single-rep experiments, the LFQ or average LFQ intensity of one sample (usually tagged pull-down) was divided by the LFQ or average LFQ intensity of another sample (usually untagged negative control), and enrichment given as the log of this value (base 2). For those calculations involving untagged controls where this would result in dividing by zero due to missing values, an imputation process was carried out. The lowest 10% average LFQ values in that negative control (excluding nondetected proteins) were averaged, and this value was used as the suitably low value for imputation. This is a similar approach used by the Perseus program, but lacking the introduction of random scatter. When presenting average LFQ values, an estimate of the error was obtained by calculating the standard error in the mean of the LFQ values from individual replicates. For enrichment values, the standard error in the ratio of the two samples was calculated by error addition, and then the high and low error bars were displayed with the appropriate log scaling.

2.2.15 Quick extraction of cellular protein by bead beating

To evaluate the integration and expression of FLAG-tagged constructs in endogenous loci, a quick and easy method was used to isolate cellular protein extract. A small inoculation loop was used to pick a small amount of yeast cells, enough to half fill the loop. These cells were

transferred into an 1.5 mL tube containing 40 µL 2x Laemmli buffer and enough glass beads to completely cover the bottom of the tube. Cells were vortexed at max. speed for one minute, followed by incubation at 95°C for a minute, and this process was repeated a total of four times. Extract was briefly spun at low speed to pellet debris, and 10-20 µL of the supernatant was taken for loading onto an SDS-PAGE gel.

Laemmli buffer (2x)

120 mM Tris-HCl pH 6.8

4% SDS v/v

20% Glycerol v/v

0.02% Bromothenol blue w/v

Add fresh:

100 mM DTT

2.2.16 Trichloroacetic acid protein extraction

During cellular extract preparation, proteolysis can affect the cellular proteins. This activity can introduce variability and degrade proteins. To prevent this activity, the low pH conditions of trichloroacetic acid mediated protein extraction were used, from the Dohlman lab, originally adapted from Cox et al. (Cox, Chapman & Waltert, 1997). Approximately 10 mL of OD_{600}=0.6-1.0 cells were harvested by centrifugation at 500 xg for 5 minutes at 4°C, and then washed to remove media and frozen at -80°C. 300 µL of a 10% TCA buffer was added to the cell pellet and the pellet allowed to thaw. Approximately 150 µL of glass beads was then added and the sample vortexed at max. speed for one minute, repeated a further four times, with 3 minute incubations on ice in between each vortexing. The liquid was removed to another tube also on ice, and 100 µL of the 10% TCA buffer used to wash the beads to remove any bound proteins, before this wash was pooled with the rest of the sample. The sample was then spun at max.

speed (14,000 rpm) in a desktop centrifuge cooled to 4°C to pellet precipitated proteins. The supernatant was removed, and pellet resuspended in 150 μL alkaline buffered 3% SDS solution. The sample was incubated at 95°C for 5 minutes, then cooled to room temperature. After centrifugation to pellet cell debris (14,000 rpm, 30s, room temperature), the supernatant was mixed with 2x Laemmli buffer, to a final 1x solution and boiled for 5 minutes at 95°C. 5-20 μL of this solution was then loaded onto an SDS-PAGE gel.

10% TCA buffer

10 mM Tris-HCl, pH 8.0

25 mM NH_4OAc

1 mM EDTA

100% Trichloroacetic acid (TCA) to a final percentage of 10% v/v

Alkaline 3% SDS buffer

100 mM Tris-HCl, pH 11.0

20% sodium dodecyl sulfate (SDS) to a final percentage of 3% v/v

2.2.17 SDS-PAGE

Samples were prepared for SDS-PAGE analysis by mixing 1:1 with 2x Laemmli buffer, and incubation at 95°C for 5 minutes. Samples were variably run on either 8/10/15% polyacrylamide gels prepared in-house, or on ThermoFisher NuPAGE™ 4-12% Bis-Tris gradient commercial gels. In-house polyacrylamide gels were run in Tris-Glycine buffer (3% (w/v) Tris, 14% (w/v) glycine, 1% (w/v) SDS). Commercial gels were run in NuPAGE™ MOPS running buffer. SDS-PAGE gels were run at 200 V for 45-70 minutes. Once gels were run, proteins were visualised by either silver staining or Western blotting.

2.2.18 Western blotting

Transfer of proteins from polyacrylamide gel to nitrocellulose membrane (Amersham Protran™ 0.45 μm NC), was done using a wet transfer method. Polyacrylamide gel and nitrocellulose membrane were assembled into a sandwich, between sheets of Whatman™ paper, in transfer buffer (25 mM Tris, 200 mM Glycine, 40% MeOH v/v, 1% SDS v/v). Transfer was carried out with a current of 400 mA over 75 minutes at 4°C. Nitrocellulose membrane was then stained with Ponceau stain (1% Ponceau S w/v, 1% v/v acetic acid) to ensure protein transfer and evaluate loading consistency in overall protein concentration. The membrane was washed several times in ultrapure water (VWR, AnalaR NORMAPUR) to destain, and then blocked for 45 minutes with 5% (w/v) milk in TBS-T (20 mM Tris, 150 mM HCl, 0.1% Tween w/v). Primary antibodies were incubated with the membrane in the same 5% (w/v) milk solution with rocking for either 2 hours at room temperature, or overnight at 4°C. After primary antibody binding, the membrane was washed three times for 5 minutes with TBS-T. For antibodies requiring a secondary HRP-coupled antibody, a second incubation with this antibody was carried out in 5% (w/v) milk in TBS-T for an hour at room temperature with rocking. The membrane was washed three more times in TBS-T, before incubation with a chemiluminescent substrate and enhancer solution. Dependent on the expected abundance of the target protein and the efficiency of the antibodies used, this substrate was one of ThermoFisher SuperSignal™ West Dura, Pico or Femto. Chemiluminescence was visualised using the ImageQuant™ LAS 4000 or the ImageQuant™ 800 imaging system.

2.2.19 Silver staining of proteins in SDS-PAGE gels

Visualisation of all abundant proteins in a sample was carried out using SDS-PAGE, as above, followed by a modified silver staining protocol . After gel electrophoresis, the polyacrylamide

gel was incubated in a fixation solution, solution 1, overnight, or for a minimum of 30 minutes with shaking. Incubation overnight was preferable to reduce background and enhance sharpness of bands. After incubation with solution 1, the gel was rehydrated with incubation in ultrapure water (VWR, AnalaR NORMAPUR), in two 15 minute washes with shaking. Following the wash and rehydration, the gel was incubated in chilled solution 2 for 90 seconds with shaking, before two 20 second washes in ultrapure water. Chilled solution 3 was added and gel was incubated for 15 minutes with shaking. Following a brief rinse first in ultrapure water, and then a small amount of chilled solution 4, the gel was developed in solution 4, for around 2-5 minutes, or until the desired intensity of protein bands. To stop further development, 15 mL of solution 5 was added and incubated with shaking until effervescence had ceased. The gel was transferred into water and could then be scanned and, if required, dehydrated for record keeping.

Solution 1 – 250 mL

100 mL methanol

33.75 mL 37.5% formaldehyde

116.5 mL ultrapure water

Solution 2 – 250 mL

0.05 g sodium thiosulphate

250 mL ultrapure water

Solution 3 – 250 mL

0.25 g silver nitrate

250 mL ultrapure water

Solution 4 – 250 mL

7.5 g sodium carbonate

125 μL 37.5% formaldehyde

245 mL water

5 mL Solution 2

Solution 5 – 15 mL

7.2 g citric acid

15 mL ultrapure water

2.2.20 Genomic DNA extraction

For extraction of a large amount and relatively pure sample of yeast chromatin, a phenol-chloroform extraction approach was used. A 10-50 mL overnight culture of yeast cells, ideally with an OD_{600} greater than 1, was centrifuged to pellet the cells (500 xg, 3 min.). Cells were washed in 10 mL of ddH_2O, and spun down again. Cells were resuspended in a mix of ddH_2O (500 µL) and TSNTE buffer (2% Triton X-100, 1% SDS, 100 mM NaCl, 10 mM Tris HCl pH 8.0, 1 mM EDTA) (200 µL). The solution was transferred into a 1.5 mL tube, and 300 µL of glass beads, and 300 µL of phenol was added. Cells were vortexed for 3 minutes at max speed, and centrifuged in a desktop centrifuge (13,000 rpm, 5 min.). The supernatant was taken and added to an equal volume (approx. 700 µL) of chloroform, and mixed. The centrifugation was repeated, and the supernatant again taken, and added to 1.2 mL of pre-chilled 100% EtOH, and incubated at -20°C for a minimum of 30 minutes, and often overnight. The sample was then centrifuged again (13,000 rpm, 30 min., 4°C), the supernatant discarded, and the pellet air-dried by incubation at 65°C for 10 minutes. The pellet was resuspended in 400 µL TE, and 2 µL of RNase (200 µg) added, followed by an incubation at 37°C for an hour. Following the degradation of RNA, 40 µL of a 3M NaAc solution (pH 5.2) was added. 800 µL of 100% EtOH was then added, and the sample centrifuged again (13,000 rpm, 10 min.), and washed with 80% EtOH. The dried pellet was resuspended in 30 µLTE, and genomic DNA was ready for further analysis. If visualised by agarose gel electrophoresis, 2 µL was loaded from this sample.

2.2.21 Quantification of DNA replication by BrdU incorporation

Cells that were competent for BrdU incorporation were grown and arrested as described above. 15 minutes prior to S phase release, BrdU was added to the culture to a final concentration of 40 µg/mL. Cells were released into S phase, with BrdU addition to the new media. At the

desired time-points, 30 mL of cells were were harvested by centrifugation (1800 xg, 4°C, 2 min.) in the presence of 0.1% NaN_3. The cell pellet was resuspended in 10 mL of cold TE with 0.1% NaN_3 and kept on ice. Cells were centrifuged again once all time-points were collected, and genomic DNA extracted as described above.

Genomic DNA concentrations were quantified by Nanodrop™ spectrophotometer and concentrations equalised across samples, before samples were boiled for 10 min. at 95°C. After rapid cooling on ice, 1 μg of ssDNA per sample was spotted onto a membrane (Amersham Hybond-N+). Once dried, the membrane was soaked in denaturation solution (0.5 M NaOH, 1.5 M NaCl), and washed twice for 5 min. in neutralisation solution (1 M NaCl, 0.5 M Tris-HCl pH 7.0). DNA was crosslinked by UV irradiation (0.12 J) and the membrane was blocked for 30 min. (1x PBS, 5% milk, 0.1% Tween). Membrane was blotted with anti-BrdU for 1.5 hours (1x PBS, 0.1% Tween) and washed three times for 5 min. (1x PBS, 0.1% Tween). The membrane was incubated with anti-mouse secondary for 1 hour (1xTBS-T, 1:10,000) and antibody binding visualised as for Western blotting.

2.2.22 PCR amplification of single colony genomic DNA

For rapid confirmation of the integration of a genetic sequence, the genomic DNA extraction from cells can be accomplished in a shorter time and a less involved protocol whilst retaining enough efficiency to gain a positive or negative signal in a PCR reaction. A single inoculation loop of cells, which is possible to obtain even from a single colony, was added to 100 μL of lithium acetate buffer (0.2 M LiAc, 1% SDS, in ddH_2O). The sample was then incubated at 65°C for 5 minutes before the addition of 300 μL of 100% EtOH. The sample was vortexed and centrifuged in a desktop centrifuge (13,000 rpm, 5 min.), and the supernatant discarded. The pellet was resuspended in 50 μL ultrapure water, and centrifuged again (13,000 rpm, 60 s)

to pellet cellular debris. This was then used as the template DNA for PCR, for which 1 µL was used:

Colony PCR reaction mix:

Reagent	Volume (µL)
DreamTaq Green Master Mix (2x)	5
Forward primer (10 µM)	0.2
Reverse primer (10 µM)	0.2
Genomic DNA	1
Ultrapure water	3.6

Thermocycler amplification protocol:

95°C – 2 min.
95°C – 20 s
56°C – 20 s
72°C – 60 s
} 32x
72°C – 5 min.
4°C - HOLD

The PCR was conducted using DreamTaq Green PCR Master Mix (Thermo) for these analytical PCR reactions, instead of the Pfu II used for the amplification of DNA for transformation protocols.

2.2.23 PCR amplification of integration cassettes

Creation of endogenous-tagging constructs by PCR amplification was conducted as described previously (Janke et al., 2004). Briefly, primers were synthesised that contained two regions of homology. The first (inner) region corresponded to sequence on a template plasmid (such as pCS762 for 5xFLAG and HygroR marker) to amplify both the in-frame 5xFLAG sequence and

the marker gene. The second (outer) region contained approximately 50 base pairs of homology to the 3' end of the gene to be tagged, and the intergenic region downstream of the targeted gene.

For tagging construct reactions, PCR was carried out using the PfuUltra Fusion II buffer and polymerase system. The reaction consists of the following components:

Reagent	Volume (µL)
PfuUltra Fusion II reaction buffer (10x)	5
dNTPs (10 mM)	5
Forward primer (10 µM)	1
Reverse primer (10 µM)	1
Plasmid DNA (1-10 ng)	1
PfuUltra Fusion II polymerase	1
Ultrapure water	36

The extension time of the PCR cycles were reduced compared to the subcloning of MCM2-7 subunits.

Thermocycler amplification protocol:

95°C – 2 min.
95°C – 20 s
56°C – 20 s
72°C – 45 s
} 32x
72°C – 5 min.
4°C – HOLD

5 μL of the completed PCR reaction volume was loaded with 1 μL of 6x Purple Gel Loading Dye (NEB) and run on an agarose gel as below. The remaining volume was added to competent yeast cells as described above for a general yeast transformation.

2.2.24 Agarose gel electrophoresis of DNA

To visualise and separate DNA molecules by size, one-dimensional agarose gel electrophoresis was used. A 1% agarose gel was prepared using 1x Tris-acetate-EDTA (TAE) buffer, and SYBR Safe intercalation agent was added at 1:10,000 concentration from stock. DNA samples were mixed with Purple Gel Loading Dye from NEB (NEB #B7025) in a 5:1 ratio, and loaded on the agarose gel. The gel was run at a voltage of 120 V in 1x TAE buffer for 20-30 minutes dependent on the DNA size.

2.2.25 Confocal microscopy of DAPI-stained cells

Confocal microscopy was used to investigate cellular morphology, and the staining of cells with the nuclear stain 4', 6-diamidino-2-phenylindole (DAPI) was used to visualise the arrangement and positioning of DNA within yeast cells. An overnight culture of cells was diluted and left to divide for two hours at 30°C, before the culture was split, with half treated with 1 μM rapamycin and the other left untreated. Cells were incubated at 25°C for a further 2 hours, before cells were fixed in 100% EtOH for an hour. Cells were then centrifuged in a desktop centrifuge (2,000 rpm, 3 min.) and washed in 1 mL PBS pH 7.4, before centrifugation a second time. A number of cells equal to <500 μL of culture were resuspended in 200 μL PBS containing 1:2000 DAPI. 5 μL of this solution was combined with 5 μL of mounting media (0.1 M Tris-HCl, 90% glycerol) and mounted on a glass slide. The cells were later observed on a Leica SP5 confocal microscope at 40x magnification.

3 Results

3.1 Arresting complexes prior to CMG formation via protein engineering

Previously in the group, modification of the MCM2-7 genes had been used to investigate the passage of DNA into and out of the ring-shaped hexamer. These lines of enquiry made use of the rapamycin-induced dimerisation domains, FRB and FKBP, together with protein linkers, to form a drug-inducible linkage between two adjacent Mcm subunits within the same hexamer (see Section 1.11). In this study, and in work carried out in parallel by another researcher in the lab, these gates between adjacent subunits were tested for their effect on the progression of replication activation and structural and compositional changes of the replication complex. An additional construct was also designed to bridge Mcm subunits between hexamers, in order to investigate a separate aspect of replication initiation.

3.1.1 DNA exit gate-blocking constructs: Intra-hexamer gates

3.1.1.1 Effect of gate-blocking on progression of replication *in vivo*

The existing Mcm2-FRB Mcm5-FKBP DNA gate construct had been shown to block MCM2-7 loading onto DNA, and work carried out in the lab (pers. comm., L.M. Reuter) showed that this construct halted origin firing *in vivo* in a rapamycin-dependent manner. This first construct had the FRB and FKBP integrated towards the N-termini of the Mcm2 and Mcm5 subunits, and an additional construct was tested in parallel that had these integrations towards the C-termini instead. The majority of the work characterising and testing these constructs was carried out by Reuter LM. prior to and during the period of this project, and a subset of the data collected will be presented here to better explain the context and enhance overall understanding of this collaborative work. This project later carried out compositional analysis of the arrest complexes generated by these constructs, so presentation of data regarding their activity is informative.

To understand the impact of the linkage on origin-specific complex formation, ChIP-qPCR analysis was carried out. For replication factors, the level of origin association at any given time can inform about the position and activation stage of the replicative complex. By tagging various replication factors of interest with the same penta-FLAG tag, an immuno-precipitation can be carried out in an identical manner across a range of proteins. For example, quantifying the origin DNA that is pulled-down in an Mcm4-FLAG IP from cells released from a G1 arrest, can show the presence of the MCM2-7 hexamer at that origin, and the approximate time-point when an origin fires and the replication fork leaves the origin. Figure 16 shows the results of such an experiment, showing the ChIP signals for the early-activating origin ARS305, a region of non-origin DNA positioned nine kilobases away from that origin, ARS305+9, and for the late-activating origin ARS501, carried out by L.M. Reuter.

MCM2-7 could clearly be seen to localise to both origins of replication in the G1-like state of alpha factor-induced arrest. The level of MCM2-7 associated with the ARS305+9 region was low in the arrest state, as the MCM2-7 hexamer should be located only at the origins, where it was loaded, since it is not yet activated. For replication factors at the alpha arrest, the +9 region can be considered a background binding level, as it is far enough from the origin to not receive any signal from origin-binding activity, and does not have any other origins nearby. Over the course of the release from G1, the levels of MCM2-7 at each origin began to decrease. Consistent with origin activation timings, the ARS305 region began losing MCM2-7 association sooner, between fifteen and thirty minutes, whilst the ARS501 region began losing MCM2-7 association later, around thirty to forty-five minutes (Figure 16A). Around thirty minutes, it was also possible to see a small but reproducible increase in MCM2-7 association at the ARS305+9 region, representing the movement of the helicase through the genome. This signal was unsurprisingly weak, as these complexes move fast, and the ChIP-qPCR is performed on a heterogenous population of cells. For MCM2-7, the key pattern was a stable

association in G1, progressing through to a loss of signal as the complexes moved away from origins as a result of activation.

Performing the same pull-down, but instead using Cdc45-FLAG, should give information more specific to the association of this early initiation factor, which in turn should reflect origin activation. This experiment, also performed by L.M. Reuter, showed similarities and important differences to the previous Mcm4-FLAG pull-down. In this case, Cdc45-FLAG association began low, during the G1 arrest (Figure 16B). However, as cells committed to S phase, and the activation of kinases begins, Cdc45 began to associate more strongly with origin DNA, indicating formation of an interaction with the MCM2-7 hexamer. After peaking at around thirty minutes, the levels of Cdc45 at ARS305 began to drop, reflecting CMG formation and movement away from the origin. Similarly to Mcm4, a brief peak in signal could be seen at the ARS305+9 region at thirty minutes, representing active CMG passing through this site. ARS501 demonstrates later activation again, with a delayed association of Cdc45, and does not show a drop in signal within the time-course tested. This may reflect a general effect visible in this experiment, where signal reduction at origins for Cdc45 was slightly later than for MCM2-7, possibly indicating a slight delay in origin firing induced by the tagging process on Cdc45. Regardless, the dynamics of Cdc45 binding showed that the timing of helicase activation can be approximated by association of the initiation factor Cdc45 with origin DNA.

Cell cycle analysis of the gate strains treated with rapamycin showed clear rapamycin-dependent cell cycle arrest in early S phase, seemingly prior to the synthesis of DNA, and this was confirmed by a BrdU incorporation assay (data not shown). Because these constructs were designed to affect the extrusion of ssDNA from the MCM2-7 hexamer during activation, the effect on replication factor association dynamics of these constructs was investigated.

This large body of work was carried out by Reuter LM. (unpublished), and is displayed here for context.

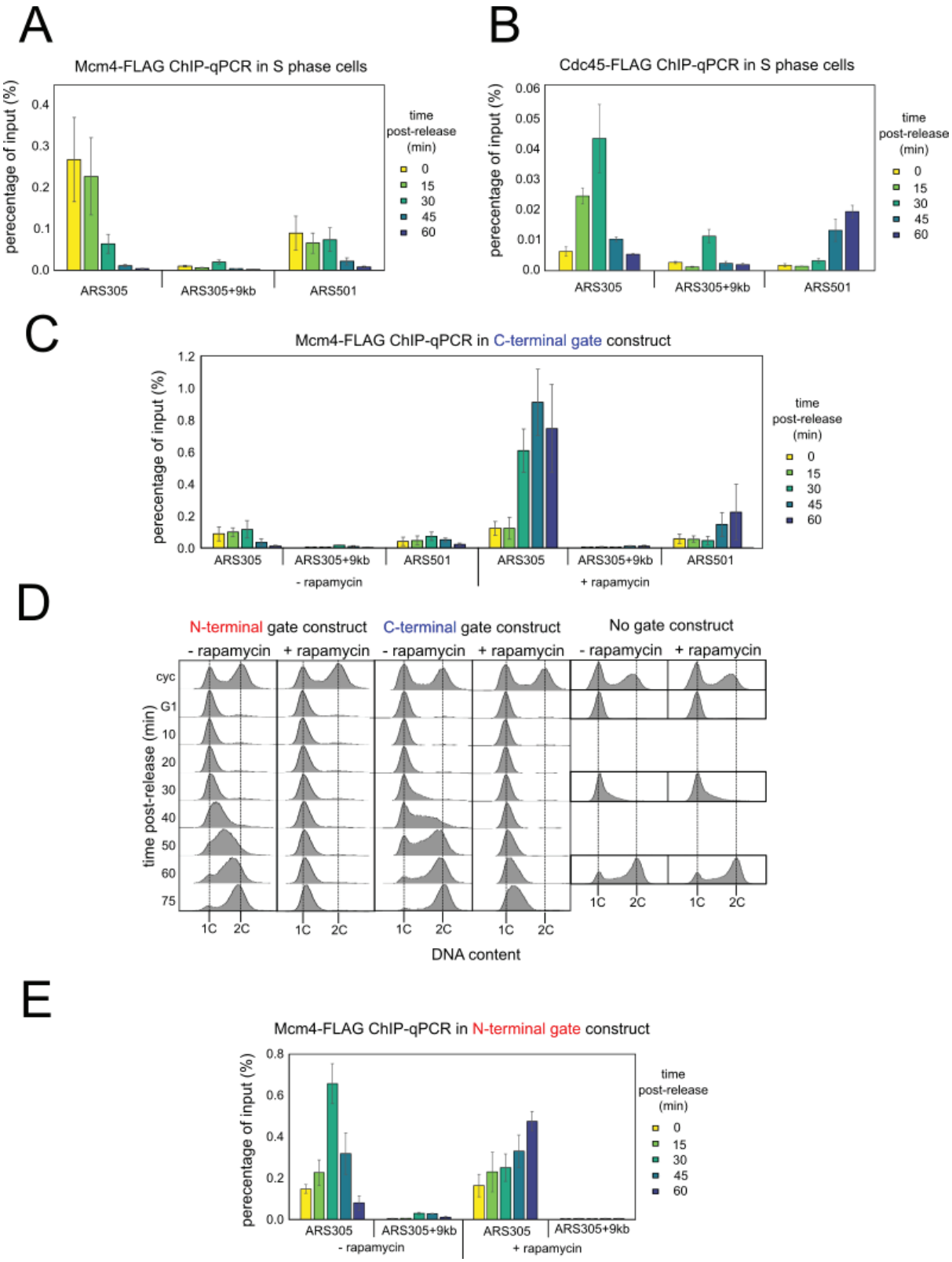


Figure 16: Overview of core replication factor behaviour in wild type ChIP-qPCR analysis of cells released into S phase from alpha factor arrest. A) ChIP-qPCR analysis of Mcm4 origin and non-origin DNA. Early and late origin ARS305 and ARS501 were tested for Mcm4-FLAG association, as well as non-origin locus ARS305+9kb. MCM2-7 association with origins falls as S phase progresses. B) As for A, but instead tagging and pulling down Cdc45-FLAG. Cdc45 origin association can be seen to rise and then fall, coupled with helicase activation. C) Testing Mcm4-FLAG ChIP-qPCR in the context of the C-terminal rapamycin-inducible ssDNA gate. Untreated cells show similar Mcm4-FLAG association dynamics to wild type cells, but rapamycin treatment results in a stable and increased association of Mcm4-FLAG with origins. D) Flow cytometry analysis of cells containing the N- or C-terminal ssDNA gate constructs or with no gate construct for comparison of overall replication dynamics at exemplary timepoints, with and without rapamycin treatment. Cells were synchronised in a G1 state by alpha factor treatment before release. E) As for C, but in the context of the N-terminal gate construct. Contrary to the other gate construct, the N-terminal gate shows an Mcm4-FLAG delayed release phenotype in the absence of rapamycin. In the presence of rapamycin, this construct shows a stable association of MCM2-7 with origins similar to the C-terminal gate. All above experiments were performed by LM. Reuter and the results presented here by their kind permission.

The C-terminal gate construct displayed an interesting phenotype by ChIP-qPCR analysis of Mcm4-FLAG pull-down (Figure 16C). Without rapamycin, MCM2-7 progression was similar to wild type cells, and showed an initial association which then decreased upon S phase progression (compare Figure 16, panels C and A, noting difference in y-axis scale). However, upon closure of the gate with rapamycin, added towards the end of the alpha arrest, MCM2-7 origin association was markedly different. In this case, loss of MCM2-7 from the origin region was not observed, and the presence of MCM2-7 at both the early and late origin was detected until sixty minutes post-release. During this period, the pull-down efficiency of the origin DNA actually increased dramatically. No signal was observed for the ARS305+9 locus after rapamycin treatment at any timepoint, indicating that no or very few active replication complexes reached this region. Coupled with the ChIP signal detected at origins and the corresponding flow cytometry data (Figure 16D), this suggested a stable arrest of MCM2-7 helicase complexes at or very close to origins.

The N-terminal gate construct showed a more complicated phenotype than the C-terminal construct. Even in the absence of rapamycin, the dynamics of Mcm4 origin association are perturbed. Following rapamycin treatment, the association of Mcm4 with origins is maintained stably past sixty minutes post-release. There is also a slight increase in pull-down efficiency over time, but to a much lesser extent than the C-terminal construct. In the absence of rapamycin, the construct causes a peak in Mcm4 origin association at 0 minutes post-G1, followed by a reduction in association towards the end of time-course. This suggests that even without closure of the gate, the linkers and FRB/FKBP domains are causing a phenotypic effect. Because of their insertion position in the N-terminal ring, it is possible that the constructs cause steric hindrance of Cdc45, Sld3/7 or GINS binding, which have been shown to bind at or towards the N terminus of MCM2-7(Sun et al., 2015; Deegan, Yeeles & Diffley, 2016; Georgescu et al., 2017). For Cdc45 and GINS, particularly, their binding across the 2/5

interface puts their binding very close to the FRB/FKBP N-terminal insertion sites. If the presence of the construct causes inefficiencies in these factors' binding, it could cause a delayed activation, and also therefore delay MCM2-7 movement away from the origin DNA. This activation delay is also hinted in the flow cytometry data (Figure 16D), where even without rapamycin, the N-terminal construct displays a slower rate of DNA replication than the untreated C-terminal construct (Figure 16D).

An interesting observation with both constructs, but particularly for the rapamycin treated C-terminal gate, is that Mcm4-FLAG origin DNA association levels could be seen to increase above their initial levels at G1 arrest over the time-course. Usually, in a ChIP-qPCR experiment, this would be taken as an indication that the amount of MCM2-7 proteins associated with the origin DNA had increased. Due to the strict controls preventing the loading of additional MCM2-7 hexamers during S phase discussed in section 1.2.1, this increase in signal must have originated from MCM2-7 complexes already present on DNA. Instead of justifying the increase with an increase in number of MCM2-7 proteins at origins, another possibility was that DNA pull-down was improved. ChIP protocols utilise the cross-linking and fixation properties of formaldehyde to form protein-DNA cross-links and kill cells prior to pelleting and freezing. This cross-linking is essential in these protocols to enable specific pull-down of DNA under reasonably strict washing conditions. The covalent linkages resulting from formaldehyde treatment stably link proteins to the DNA locus they associate with throughout the experiment. Then, heat treatment combined with SDS is able to reverse the cross-links, and proteinase K degrades protein, leaving only the enriched pull-down DNA. The ability of formaldehyde to link the protein of interest to specific DNA loci will clearly affect the observed pull-down efficiency. In proteins, formaldehyde reacts with nucleophilic side chains of several amino acids (Metz et al., 2004), and in DNA, formaldehyde reacts with nucleophilic groups in the nitrogenous bases (Hoffman et al., 2015). The reaction of formaldehyde with nucleophilic

groups in the DNA is much enhanced if the hydrogen bonds between base pairs are broken, meaning that single-stranded DNA presents a much more effective substrate for cross-linking than double-stranded DNA (McGhee & Hippel, 1975). Since the constructs used in these experiments were specifically designed to support initial DNA unwinding, but block escape of ssDNA from the helicase, it is plausible that in the arrest state, at least some of the DNA associated with the complex will be in single stranded form. This could have a knock-on effect in terms of pull-down efficiency, making the DNA pull-down more effective, and resulting in greater IP efficiencies in the stalled complex. This effect is particularly evident in the C-terminal gate construct, suggesting production of a significant amount of ssDNA that continues association with the complex in its arrested state. Efforts were made to quantify the scale of this effect by investigating protein-DNA cross-linking *in vitro* as part of a Master's project, but results were inconclusive. However, this effect is persistent when dealing with the C-terminal gate construct in particular in multiple stages of this project. It has been referred to as the "hyper-crosslinking" effect. In later work, L.M. Reuter showed that ssDNA was enriched at origins in the C-terminal intra-hexamer linkage by KMnO4 ChIP PIP-seq (L.M. Reuter, manuscript submitted).

The above data appeared to show that the MCM2-7-containing replicative complex could be stalled at origins of replication at some point during the process of activation. Other than the design of the gates themselves, which hypothetically blocks ssDNA extrusion, it was not clear at what stage the arrest occurred, and whether any proteins other than MCM2-7 itself were associated with the complex. In order to test this, sequential tagging of multiple replication factors of interest was carried out by LM. Reuter, and following successful integration of penta-FLAG tags into the endogenous copies of said replication factors, L.M. Reuter continued to carry out ChIP-qPCR based analysis.

The Mcm4-FLAG ChIP-qPCR data indicated that the C-terminal gate caused less perturbation of the normal replication process in untreated cells, and simultaneously had the larger effect on MCM2-7 association with origin DNA when treated with rapamycin. The increase in association with origin DNA after rapamycin treatment was observed with many other replication factors, including some factors that are normally transiently associated with the replication complex only during activation (Figure 17).

ChIP-qPCR of Cdc45 in the C-terminal gate construct without the addition of rapamycin produced very similar results to the wild type cells containing only Cdc45-FLAG (compare Figure 16 panel B to Figure 17 panel A), whilst also displaying the same stable efficiency increase upon rapamycin treatment observed for Mcm4 in the C-terminal gate. This effect also continued to follow normal origin activation times, with association with ARS501 only increasing later, but also showing this large increase in association.

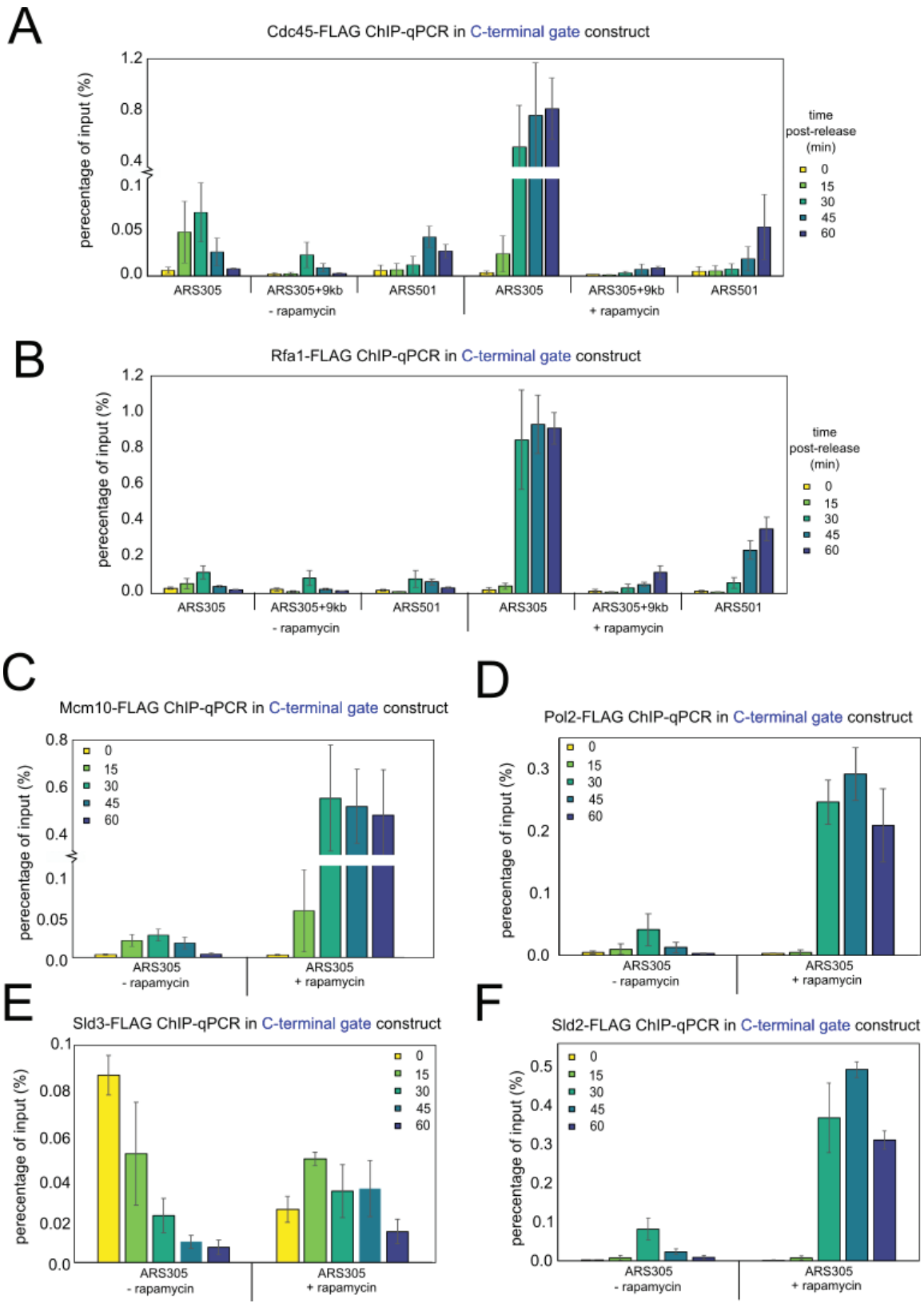


Figure 17: ChIP-qPCR of replication factors origin association in the C-terminal intra-hexamer linkage construct. A) ChIP-qPCR of Cdc45-FLAG at the early origin ARS305, the late origin ARS501, and non-origin ARS305+9kb DNA. Samples were released from alpha-factor arrest into S phase with and without the addition of 100 nM rapamycin for up to 75 minutes. Error bars indicate standard deviation for at least three individual replicates. B) As in A, but for Rfa1-FLAG C) As in A, but for Mcm10-FLAG. D) As in A, but for Pol2-FLAG. E) As in D, but for Sld3-FLAG. F) As in E, but for Sld2-FLAG. All above experiments were performed by LM. Reuter and the results presented here by their kind permission.

Consistent with the hypothesis that it was single-stranded DNA production that was responsible for the greatly increased pull-down efficiency, the association dynamics of RPA (Rfa1, 2 and 3) appeared to match exactly with the appearance of this “hyper-crosslinking” effect. In non-rapamycin treated cells, RPA association peaked at thirty minutes for ARS305 (Figure 17B). This protein can be used to approximate the maximum level of single-stranded DNA presence on average across the cell population. At thirty minutes, then, if single stranded DNA is produced and origins usually fire, most cells are likely to begin encountering difficulties with DNA synthesis in the context of an intra-hexamer linkage. Indeed, at thirty minutes and beyond in rapamycin treated cells, linkage between Mcm2 and Mcm5 induced a very strong and stable enrichment of origin DNA in Rfa1 IP. This indicates that single-stranded DNA was produced in the presence of the closed gate between fifteen and thirty minutes, and that the Mcm2/5 linkage prevented loss of this signal, as in untreated cells. This loss was likely due to movement of the complex away from the origin, since RPA is present near normal replisomes.

Other replication factors, both specific for initiation, and components of the active fork, showed a similar behaviour (Figure 17C, D and F). Mcm10, likely the last replication factor to associate with the replisome before full activation (van Deursen et al., 2012; Douglas et al., 2018), showed a rapamycin-specific stable and increased association with origin DNA. So too did the largest subunit of polymerase ε, Pol2, and Sld2, a transient initiation factor that arrives to the complex along with GINS and Dpb11 after kinase phosphorylation.

One factor that does not follow this trend is Sld3. In untreated cells, Sld3 appeared to associate with origin DNA in the G1-phase arrest, suggesting some level of pre-association with the double hexamer in the alpha-factor induced state (Figure 17E). It also showed a more rapid dissociation from the complex than other factors, suggesting Sld3 is released from the complex earlier than some of the other transient initiation factors. In rapamycin treated cells, Sld3 loss may have been slightly delayed, but there was certainly no continued association with origin

DNA for as long as some of the other factors. In addition, there was no sign of the markedly increased origin DNA pull-down efficiency observed with other factors. This suggested that either Sld3 dissociates prior to the appearance of the "hyper-crosslinking" effect, or that during this effect, Sld3 is not affected in the same way. It was hypothesised that this could mean that Sld3 dissociation occurs prior to the production of single stranded DNA, and prior to the arrest stage induced by the C-terminal gate construct.

LM. Reuter repeated this analysis for the N-terminal intra-hexamer linkage construct (Figure 18). Attempts to generate an endogenously tagged Cdc45 in this strain proved unsuccessful, potentially because of a synthetic lethal interaction between the tagged Cdc45, and the N-terminal linkage construct. This could also potentially relate to the observed phenotype of the N-terminal construct even without rapamycin. If the construct already causes inefficiencies in the initiation process, then tagging of another protein in this mechanism could further compromise replication. Alternatively, if the construct caused structural interference with the binding site of Cdc45, the addition of extra peptides to Cdc45, even if relatively small, could prevent Cdc45 binding to the double hexamer entirely. In place of Cdc45-FLAG, an antibody specific for Cdc45 was used for pull-down instead. As a result, the efficiency of this IP was much lower than for the tagged Cdc45 (Figure 18A). Despite this, the untreated dynamics of Cdc45 origin association resemble the wild type, and Cdc45 associates and disassociates consistent with the formation and activation of the CMG. Rapamycin treatment does cause stable association of Cdc45 up to sixty minutes, but there is only a minor increase in signal, the "hyper crosslinking" seen in the C-terminal gate is not observed in this construct for Cdc45. The ARS305+9 locus also sees a peak of Cdc45 association in untreated, but not rapamycin-treated cells, indicating an arrest of the complex at origins after rapamycin treatment, as with the C-terminal gate.

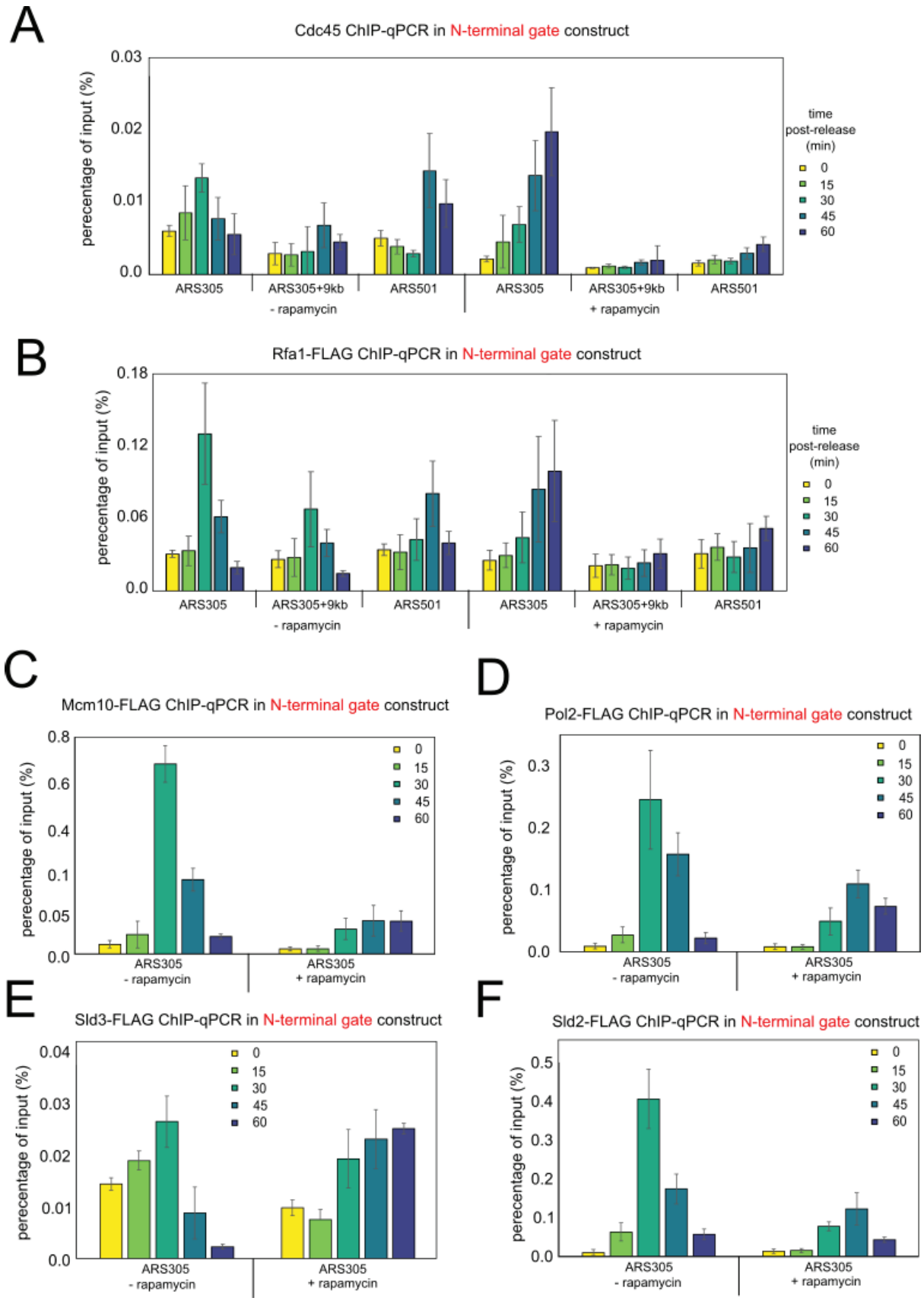


Figure 18: ChIP-qPCR analysis of replication factor origin association in the context of the N-terminal intra-hexamer linkage construct. A) ChIP-qPCR of Cdc45 at the early origin ARS305, the late origin ARS501, and non-origin ARS305+9kb DNA, using anti-Cdc45 antibody. Samples were released from alpha-factor arrest into S phase with and without the addition of 100 nM rapamycin for up to 75 minutes. Error bars indicate standard deviation for at least three individual replicates. B) As in A, but for Rfa1-FLAG C) As in A, but for Mcm10-FLAG. D) As in A, but for Pol2-FLAG. E) As in D, but for Sld3-FLAG. F) As in E, but for Sld2-FLAG. All above experiments were performed by LM. Reuter and the results presented here by their kind permission.

To investigate why this large increase in IP efficiency was not observed in this construct, ChIP-qPCR of Rfa1 was carried out (Figure 18B). Here, too, untreated cells behaved normally, with an increase in RPA origin association followed by a decrease, and rapamycin-treated cells showed a stable association, with very little increase in IP efficiency. This result showed that RPA was recruited, indicating the presence of some ssDNA, yet no induction of the "hyper-crosslinking" effect. Perhaps the amount of ssDNA produced was not sufficient to cause this effect, or it had some other cause. For the replication factors that arrive later on during initiation, Mcm10, Pol2, and Sld2, this N-terminal construct caused very similar effects on origin association Figure 18C,D,F). Untreated cells showed a clear rise and fall in factor origin association, but rapamycin-treated cells showed only a very small level of recruitment of these factors to origins. The association of Mcm10 was especially weak, and given that Mcm10 appears to associate at the very tail end of the initiation process, this suggested that the N-terminal construct struggles to progress to the stage of initiation that usually recruits Mcm10 when the gate is closed with rapamycin.

In addition to the lack of strong association of late-stage replication factors, and the absence of a "hyper-crosslinking" effect, the N-terminal gate also showed differences in the origin association dynamics of Sld3 (Figure 18E). Contrary to the C-terminal construct, in the N-terminal gate Sld3 showed a stable association with the origin, rather than a reduction in association over time. This potentially indicated again that the N-terminal gate did not progress as far through the activation process as the C-terminal gate, and whatever structural or compositional change caused the disassociation of Sld3 had not yet occurred. It is worth mentioning that the N-terminal gate did show slightly different association dynamics of Sld3 without rapamycin than the C-terminal gate. In the C-terminal gate, Sld3 showed maximum origin association at the start of the time-course, but the N-terminal gate showed an increased association then a decrease, suggesting that Sld3 had not been able to maximally associate with

the double hexamer during alpha arrest. Given Sld3 is also thought to bind the N-terminal tier of MCM2-7 (Deegan, Yeeles & Diffley, 2016), this could be an example of the N-terminal gate causing interference with the association of one of the earliest initiation factors, and this could underlie the slight delay seen in replication initiation in this construct even without rapamycin gate closure.

Based on the ChIP-qPCR data obtained by L.M. Reuter from these strains, of which a subset is shown here, and further experiments of theirs investigating initial unwinding and precise ChIP-Exo data of MCM2-7 localisation (L.M. Reuter, manuscript under preparation), a hypothesis was suggested to explain the activities of the gate constructs. Some of this unpublished data suggested that the initial DNA unwinding occurs at the interface between the two hexamers of the DH. This would explain why an N-terminal ssDNA gate might arrest the complex at an earlier stage than a C-terminal ssDNA gate. If the ssDNA extrusion also begins at the hexamers' interface, and then progresses along the length of each hexamer, then a gate positioned at the N-terminus will block ssDNA exit earlier than a C-terminal gate. Given the "hyper-crosslinking" effect observed only in the C-terminal gate is strongly suspected to be due to improved cross-linking to ssDNA. If the N-terminal gate produces less single-stranded DNA, and the single-stranded DNA that is produced is still mainly contained within the MCM2-7 hexamer, then there is less single-stranded DNA free to cross-link with the wider replication complex and other factors. The temporal progression suggested by the two constructs, with the continuous origin association of Sld3, and reduced association of late-stage factors in the N-terminal construct, would agree with the extrusion of ssDNA starting at the N-terminus then further opening of the Mcm2/5 interface down to the C-terminus. Possible complex remodelling and structural changes may occur during this process, resulting in association of late-stage initiation factors such as Pol2, Sld2 and Mcm10, followed by Sld3 loss (Figure 19).

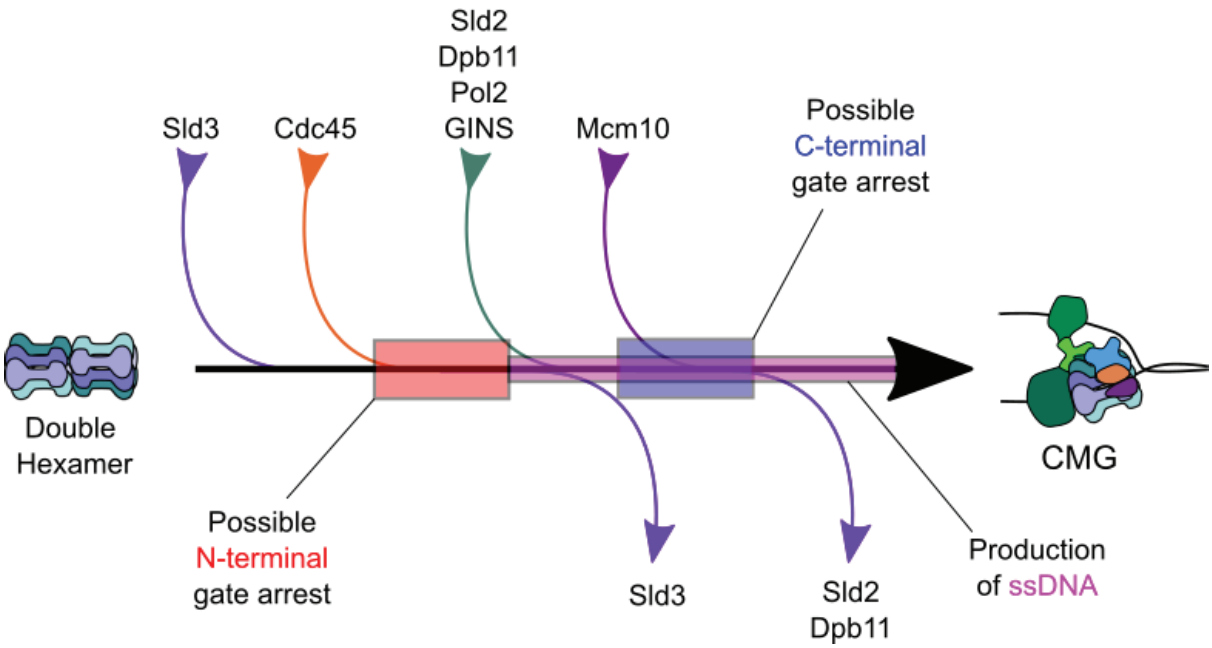


Figure 19: Hypothesised order of events during replication initiation based on intra-hexamer linkage constructs. The N-terminal gate arrests whilst Sld3 is still present, but before large amounts of ssDNA is produced, and late-firing factors Sld2, Dpb11, Pol2 and Mcm10 do not show strong origin association. The C-terminal gate arrests after Sld3 has disassociated and late-stage initiation factors are present. The "hyper-crosslinking" effect suggests presence of significant ssDNA in and around the complex.

3.1.2 Hexamer separation-blocking construct: Inter-hexamer latch

The data above, generated by LM. Reuter, showed the potential for utilising the rapamycin-dependent FRB/FKBP domains to physically arrest complexes by preventing DNA extrusion required for replication fork assembly. Previously, the introduction of the FRB, FKBP and linkers into MCM2-7 focused on preventing the extrusion of ssDNA through the Mcm2-Mcm5 gate. During discussions surrounding the effects of the ssDNA exit gates on complex progression, it was considered whether there were any other structural changes that could be effectively blocked by the use of these dimerisation domains.

3.1.2.1 Design and integration of the latch construct

Before replication initiation commences, the MCM2-7 helicase is present at origins in the form of a head-to-head double hexamer. Once activation has completed, the MCM2-7 helicase forms part of a larger replisome that unwinds DNA, and moves through the genome. In terms of

structure, the MCM2-7 in the replisome is a single hexamer, with the leading strand contained within the central channel, and the lagging strand excluded, and the helicase progresses in a 3' to 5' direction. The two gross structural rearrangements between these two states are the transition from encircling double-stranded DNA to ssDNA via strand extrusion, and the separation of the two hexamers into individual, bidirectional fork complexes. The constructs above were designed to interfere with the first of these rearrangements, and so it was considered what other structural changes might occur during activation. Because the hexamers are loaded in the DH in a head-to-head manner, they must at some point separate and pass each other in order to proceed in an N-terminal first direction. Either hexamers separate completely, then extrude ssDNA, and later pass by on opposite strands, or the hexamers slip or twist laterally past each other, and extrude ssDNA in a concerted motion to allow the continued movement of the hexamers past each other. Regardless of the exact mechanism used, which remains unknown, to achieve this passing by, the interface between the two hexamers must be disrupted at some stage in the process.

By using the FRB/FKBP domains, it was proposed that the two hexamers might be able to be secured together, in an attempt to prevent hexamer movement relative to the other. Because of the arrangement of the subunits in the hexamers, and their twist-orientation, Mcm4 from one hexamer is positioned adjacent to Mcm5 from the other hexamer, and this pairing is repeated for the subunits across the other side of the complex. By integrating an FRB domain and linkers in the N-terminus of Mcm4, and an FKBP domain and linkers in the N-terminal domain of Mcm5, two separate "latches" could be generated on either side of the double hexamer (Figure 20). Upon rapamycin addition, dimerisation of these domains would create stable links between the two hexamers, and may prevent hexamer movement relative to each other. If these links are stable enough to prevent separation or slipping/twisting, then this construct may be able to arrest replication initiation during helicase activation.

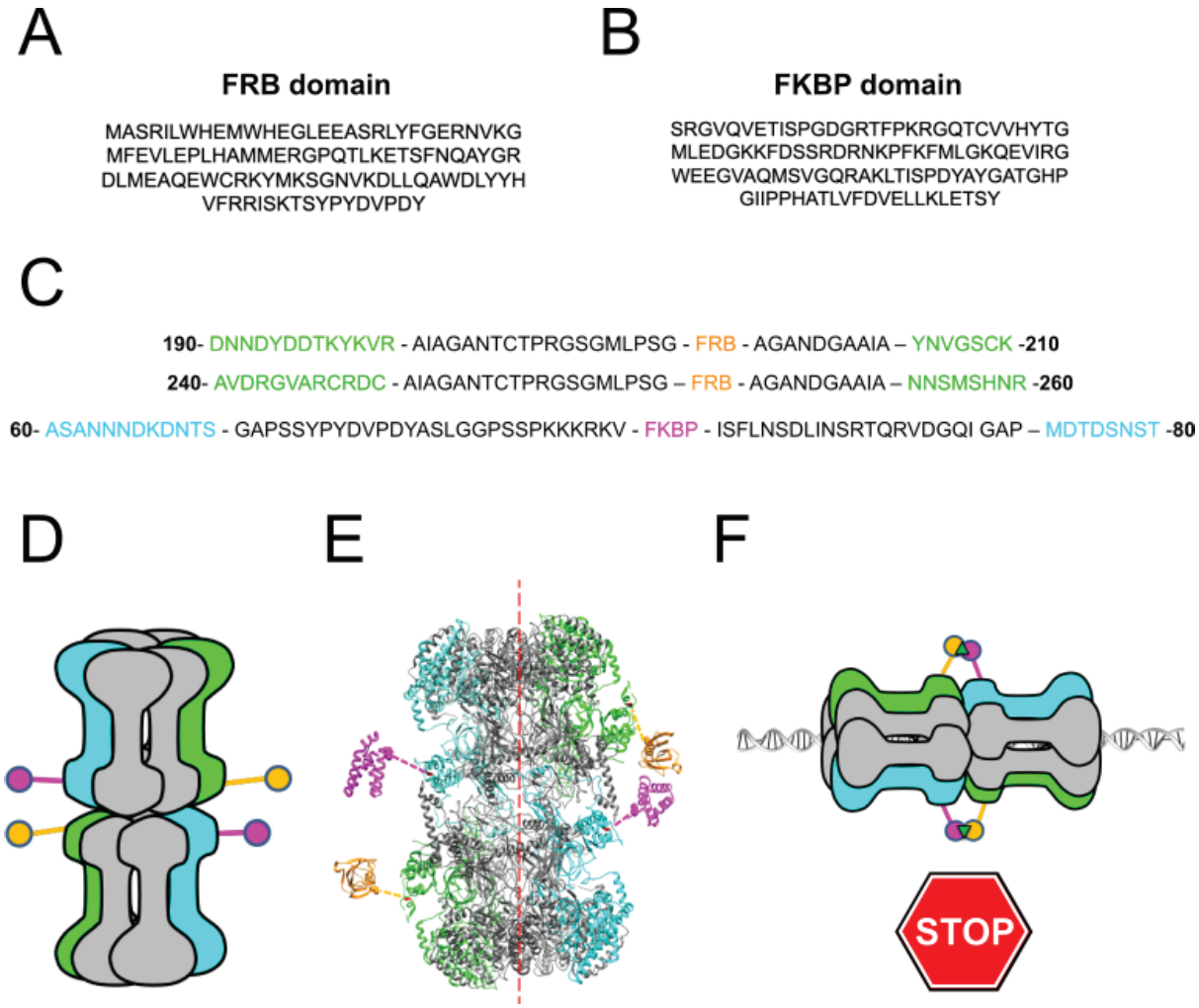


Figure 20: Sequences and design principles of the Mcm4-FRB Mcm5-FKBP inter-hexamer latch. A) Peptide sequence of the FRB domain used, identical to that used in gate strains. B) Peptide sequence of the FKBP domain used, identical to that used in gate strains. C) integration sites and linker peptide sequences used for integrations. Two sites on Mcm4 were tested, with insertion after amino acid 203 and 252. Both insertion sites are in flexible loops on the outer surface of the N-terminal tier. For Mcm5, the FKBP construct from the N-terminal gate was used. D) Schematic of the double hexamer with approximate positions of the subunits and dimerisation domains. Mcm4 = green, Mcm5 = blue, FRB = orange, FKBP = pink. E) Structure of MCM2-7 double hexamer (Noguchi et al., 2017) combined with the structures of the FRB and FKBP domains (Choi et al., 1996). Highlighted in red are the insertion sites for the Mcm4-252 insertion, or the closest structurally resolved residue in the structure. Left of the red line considers the rapamycin-absent condition, whilst right of the red line suggests the structure post-rapamycin addition. F) Schematic of the double "latch" induced upon rapamycin (green triangle) binding. The two links are approximately opposed across the double hexamer.

Cells containing the Mcm5-FKBP construct immediately after amino acid 72 had previously been shown to be viable, as this was the insertion site and insert used in one side of the N-terminal ssDNA gate construct. Because this was also very near to the N-terminus of Mcm5, this insertion construct was chosen for the Mcm5 side of the inter-hexamer latch. For the Mcm4 insertion site, two possible sites were chosen and tested simultaneously. One site was

immediately after amino acid 203, and the other immediately after amino acid 252. These sites (Figure 20C, E) had previously been identified as being in flexible loops on the outer surface of the MCM2-7 double hexamer structure (Samel et al., 2014), and so could tolerate insertion of large intervening sequences of peptides without causing disruption of MCM2-7 folding and function. Additionally, the insertion site on Mcm5 is sufficiently distant to the recently proposed ssDNA exit channel (Baretić et al., 2020) so as not to be likely to directly interfere with this mechanistic process (Supp. Fig. 1).

To modify the endogenous Mcm4 gene, FRB linkers and flanking sequences of MCM4 were subcloned into the pCS80 plasmid. The pCS80 plasmid, derived from the pFA6a precursor, contained a His3MX6 selection marker, which the integration sequence was inserted upstream of. PCR amplification of this intermediate plasmid generated PCR integration cassettes that were used to transform competent yeast cells already containing the Mcm5-FKBP inserted after amino acid 72. This strain also contained a tor1-1 mutation which enabled treatment with rapamycin without the TORC1-inhibition that the drug would normally engender, and a deletion of the FPR1 gene, the yeast equivalent of mammalian FKBP, to prevent rapamycin sequestration (Barbet et al., 1996; Schmelzle & Hall, 2000; Evans, Burgess & Gray, 2014).
Two such intermediate plasmids were constructed, one for the FRB insertion after amino acid 203 in Mcm4, and one for insertion after amino acid 252.

PCR amplification of genomic DNA extracted from colonies resulting from the transformation was carried out, with primers specific for sequences within endogenous MCM4 (Figure 21D). The insertion of FRB was detectable as an increase in PCR product size since the primers flanked both of the MCM4 sites targeted. Clones that were positive for the MCM4-FRB insertion for both insertion sites were identified and selected.

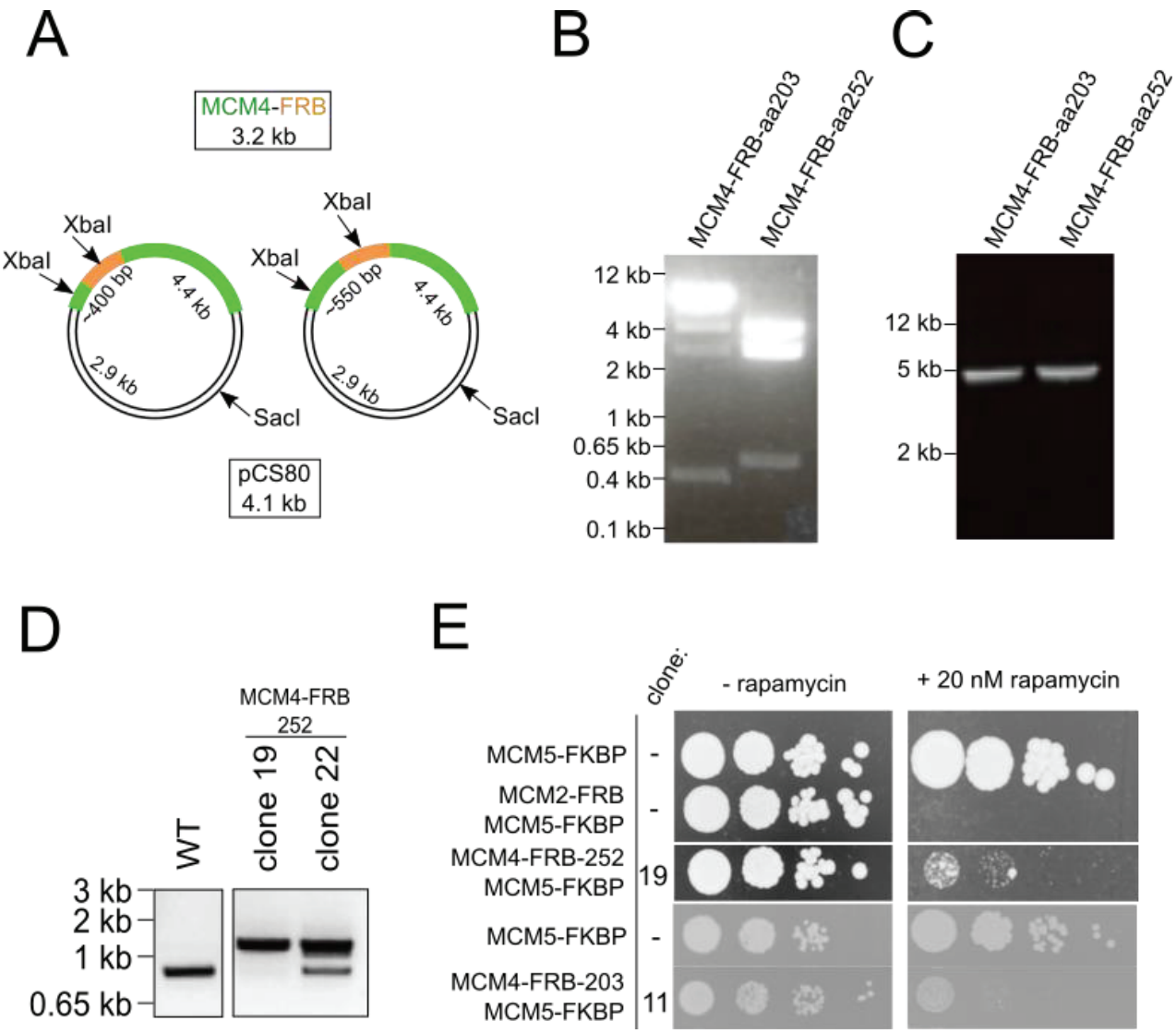


Figure 21: Subcloning and integration PCR for insertion of FRB and linkers into two candidate sites in endogenous MCM4. A) plasmid map of the intermediate stage plasmid, MCM4-FRB in pCS80. B) Analytical digest of plasmids in A, using restriction enzymes presented. Both plasmids produce the correct bands, though the MCM4-FRB-aa203 shows a higher proportion of linearised plasmid. C) PCR amplification of an integration sequence for endogenous integration of MCM4 containing the FRB insertion. D) PCR on colonies using MCM4 internal primers identifies successful transformants. Mcm2-FRB Mcm5-FKBP gate strain used a negative "WT" control. PCR of the intermediate plasmid used as a positive control (data not shown). A high band indicates successful integration of FRB and linkers into MCM4. E) Viability assay of selected clones on YPD and YPD containing rapamycin. The strain used for the transformation (MCM5-FKBP) and the N-terminal ssDNA exit gate strain (MCM2-FRB MCM5-FKBP) were used as negative and positive controls for rapamycin sensitivity respectively.

Several clones carrying the MCM4-FRB insert after amino acid 252 were identified to have a reduction in viability when growing with rapamycin present (Figure 21E). Most of these displayed normal unperturbed growth on normal media compared to strains containing the MCM5-FKBP alone, and clone 19 was selected for further work based on these characteristics and a clearly positive PCR signal for the integration. Compared to the N-terminal ssDNA exit

gate strain, the loss of viability was less extreme, as the N-terminal ssDNA exit gate abolished growth entirely when grown in the presence of rapamycin, whereas clone 19 of the latch construct still showed some growth in the presence of rapamycin. This was not confined to that single clone, indicating a less severe effect on growth in the latch construct compared to the previously studied ssDNA exit gates.

In the case of the Mcm4 construct containing the FRB insertion after amino acid 203, clones exhibited a rapamycin-dependent loss of viability (Figure 21E, lower panel). However, whilst difficult to see in the above figure, cells carrying these constructs also showed a relatively lower viability under untreated conditions than the parental strain (Figure 21E, lowest two lines, compare the second dot). This indicated that inserting the FRB and linker construct at this site of Mcm4 was having a deleterious effect that was not related to the FRB-FKBP linkage and so this site was not progressed further.

The strain containing MCM5-FKBP, tor1-1 and *Δfpr1* used for the transformation was mating type alpha, making it insensitive to alpha-factor. Selected clones were further transformed with an HO endonuclease and expression of this endonuclease was driven to induce mating type switching. Cells were plated at colony forming density, and individual colonies tested for their mating type and selection markers of the other integrations and deletions they carried (data not shown). This process resulted in strains carrying the MCM4-FRB MCM5-FKBP latch that were now the correct mating type to respond to alpha factor.

3.1.2.2 Effect of latch linkage on progression of replication *in vivo*

It had been shown that the latch construct caused a rapamycin-dependent effect on cell viability on plates. The loss of viability was clear, but the effect was less severe than for the previous constructs designed and produced in the lab. Nonetheless, this was a promising sign for the activity of the inter-hexamer latch, and the loss of viability was often observed to be near total on rapamycin-containing plates that were fresher, and potentially containing higher concentrations of rapamycin (Figure 22A), though always less extreme than the intra-hexamer gate constructs.

However, viability assays on plates subject the cells to rapamycin-induced FRB/FKBP dimerisation at all stages of the cell cycle and S phase. Indeed, the intra-hexamer gates were originally designed to prevent MCM2-7 loading, and this activity should similarly prove as lethal as preventing helicase activation. It was possible that the latch construct caused some effect that was independent of helicase activation entirely, and this was the cause of the lethality.

Flow cytometry analysis was able to investigate how rapamycin treatment in the inter-hexamer latch strain affected the cell cycle. Initially, the addition of rapamycin was performed on cycling cells, to observe if the treatment caused synchronisation of cells at any particular stage where the latch construct was causing toxicity (Figure 22B). Cells not treated with rapamycin maintained a steady ratio of G1, S and G2/M phase cells throughout the experiment. Rapamycin-treated cells, however, did show evidence of cell cycle defects. There was a slight increase in the proportion of cells undergoing S phase 2-3 hours after rapamycin addition, and after 5 hours, a reduction of G1-phase cells was observed. This indicated that cells were

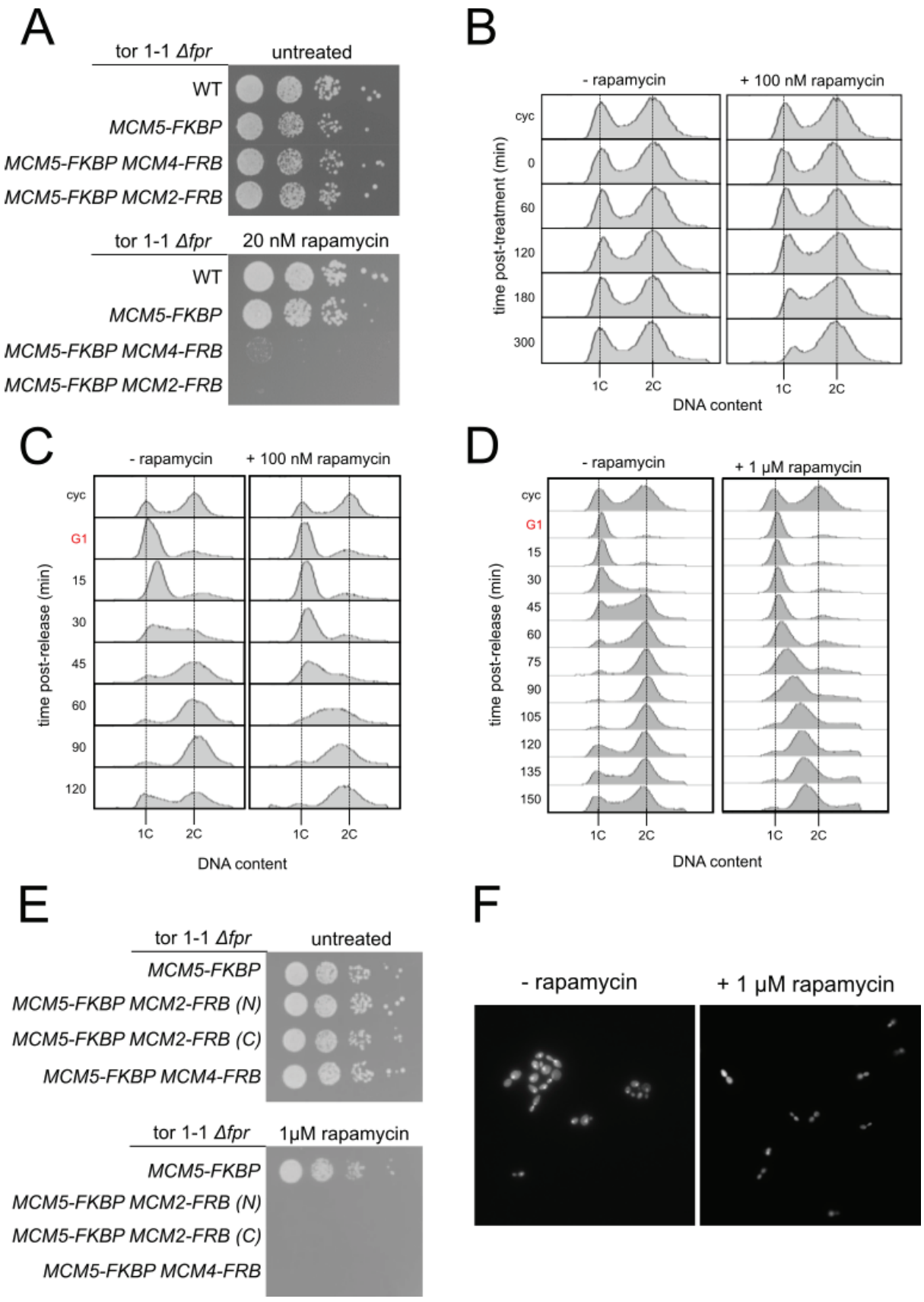


Figure 22: Phenotypic effects of the rapamycin-dependent inter-hexamer latch on cells A) Representative viability assay of the final mating-type switched MCM4-FRB MCM5-FKBP latch strain on fresh rapamycin-containing plates. The strain was compared against a strain containing only MCM5-FKBP, the MCM2-FRB MCM5-FKBP N-terminal ssDNA gate, and a strain containing no FRB/FKBP inserts. Clear loss of viability was demonstrated in the latch strain. B) Flow cytometry analysis of the inter-hexamer latch strain with and without rapamycin treatment over five hours. C) As for B, but analysing cells released into S phase in the presence or absence of rapamycin after alpha-factor synchronisation, and for only two hours. D) As for C, but with the addition of 1μM rapamycin, showing improved rapamycin-dependent initiation delay. E) Representative viability assay of the final mating-type switched MCM4-FRB MCM5-FKBP latch strain on 1 μM rapamycin-containing plates. 1 μM rapamycin is only lethal to strains with a complete gate or latch construct. F) Representative 40x confocal microscopy image of DAPI-stained cycling cells containing the inter-hexamer latch with and without rapamycin addition.

possibly unable to properly complete DNA replication, and therefore unable to correctly undergo mitosis. In order to understand the effect of the inter-hexamer latch on helicase activation, the latch strain cells were therefore arrested in G1 by alpha-factor treatment, before being released into S phase with and without the presence of rapamycin (Figure 22C). Whilst a slight delay in replication initiation was observed, the latch construct did not fully prevent DNA replication, and cells were able to steadily increase their DNA content to nearly double that of the G1 cells. The rapamycin-treated cells did still show issues during mitosis, as by 3 hours post-release, untreated cells showed a clear G1 population, whilst rapamycin-treated cells were enriched for the G2/M state (data not shown).

A failure to complete replication at a few loci across the genome could explain the DNA content changes observed. Most of the genome was replicated as shown by the near-C2 DNA content, but if a few loci remained un-replicated, this may prevent proper chromosomal segregation and cells would not divide. Usually, if replication forks stall or collapse, back-up origins are able to fire to replicate the un-replicated region. However, if the latch was able to interfere with some origins firing but not all, in a stochastic manner, it could cause the effect seen here. An overall lower number of active forks could leave difficult to replicate regions of chromatin susceptible to fork stalling without rescue. It was also noted that the viability of cells on rapamycin plates was dependent on the freshness of the plates, so flow cytometry was repeated with higher levels of rapamycin (Figure 22D). In this case, a very clear delay in initiation was observed, which lasted for 60 minutes, when the DNA content increased and progressed to a near-C2 level. Importantly, this level of rapamycin treatment did not have an effect on cells that did not contain the latch construct, but did contain the tor1-1 and *Δfpr1* mutations (Figure 22E).

The difference between these constructs in terms of their arrest ability can be rationalised by comparing the structural changes they prevent. It can be hypothesised that the extrusion of

ssDNA is a process that relies on gate opening and initial DNA unwinding, and after this has happened, specific contacts may be maintained with the leading strand, whilst the lagging strand is free to diffuse out of the central channel. This could be seen as a low-force process. For the inter-hexamer latch, however, the forces involved could be very different. Each hexamer contains six fully-functional ATPase motors. If the hexamer separation relies upon the forces generated by these motors, the latch clasp interaction, whilst comparatively stable, may prove inadequate. Similarly, if the arrested intermediate created by the latch is able to orient itself, due to molecular motion, in a way that enables ATPase activation with the forces acting against the latch clasp, then the interaction may not hold. Increasing the effective concentration of rapamycin may improve the binding occupancy of the FRB/FKBP dimer, and hence improved stability over time of the construct arrest. However, if the forces of ATPases are opposing the FRB/FKBP latch, then eventually the motors may win out, causing loss of the arrest. This may not occur for all helicases in the cell, but enough escape to induce replication of DNA on a magnitude that is observable by flow cytometry, and eventually leads to the entire genome being replicated. Some hard to replicate regions will be disproportionately affected by the lower numbers of forks, and so may remain unreplicated, causing the mitotic arrest seen even at lower rapamycin concentrations.

The cells themselves were observed morphologically by light microscopy (data not shown) and by confocal microscopy using DAPI staining (Figure 22F). Untreated cells resembled the classical progression of budding yeast, with smaller buds being generated by mother cells. Over time, clusters of yeast cells probably derived from one or few initial cells can be seen that maintain loose association (Figure 22F, top left). However, almost all rapamycin-treated cells were seen to be present as dyads or tetrads, joined at the budding neck. The pairs of cells joined in this manner can clearly be seen in the DAPI staining also. From the DAPI staining at this magnification, it is difficult to tell if the nuclei are entirely segregated, but if replication did not

proceed fully as was hypothesised, the presence of large DNA bridges or an increased level of ultra-fine bridges (UFBs) might be visible at higher magnification.

Flow cytometry analysis shows total DNA content of the cell, but there are more sensitive methods to determine the incorporation of bases into newly synthesised DNA. By using the thymidine analogue bromodeoxyuridine (BrdU), this incorporation can be directly observed. An inter-hexamer latch strain was crossed with a strain able to efficiently transport BrdU into the cell, and successfully incorporate it during DNA replication. Cells were released from alpha-factor arrest in the presence of BrdU, and cells harvested over a time-course. Genomic DNA was extracted and equal concentrations loaded onto membrane-bound “dots”, before being probed with an anti-BrdU antibody (Figure 23).

As expected, the untreated cells entered S phase normally, and incorporated large amounts of BrdU (Figure 23B). The rapamycin-treated cells displayed very little incorporation of BrdU into genomic DNA, up to 60 minutes post-release. Surprisingly, the cells kept in alpha-factor arrest were seen to incorporate a small amount of BrdU, especially towards the end of the time-course. This was despite the flow cytometry analysis showing very little visible DNA content increase. The cells used here had already been arrested for two-and-a-half hours in alpha factor prior to release. It is possible, despite re-addition of the pheromone, the response to alpha factor was attenuated and some cells entered S phase. DNA incorporation levels with the closed inter-hexamer latch were markedly lowered until at least 60 minutes. Due to these observations, flow cytometry analysis of cell cycle progression became the preferred method of monitoring replication initiation and progression in this project.

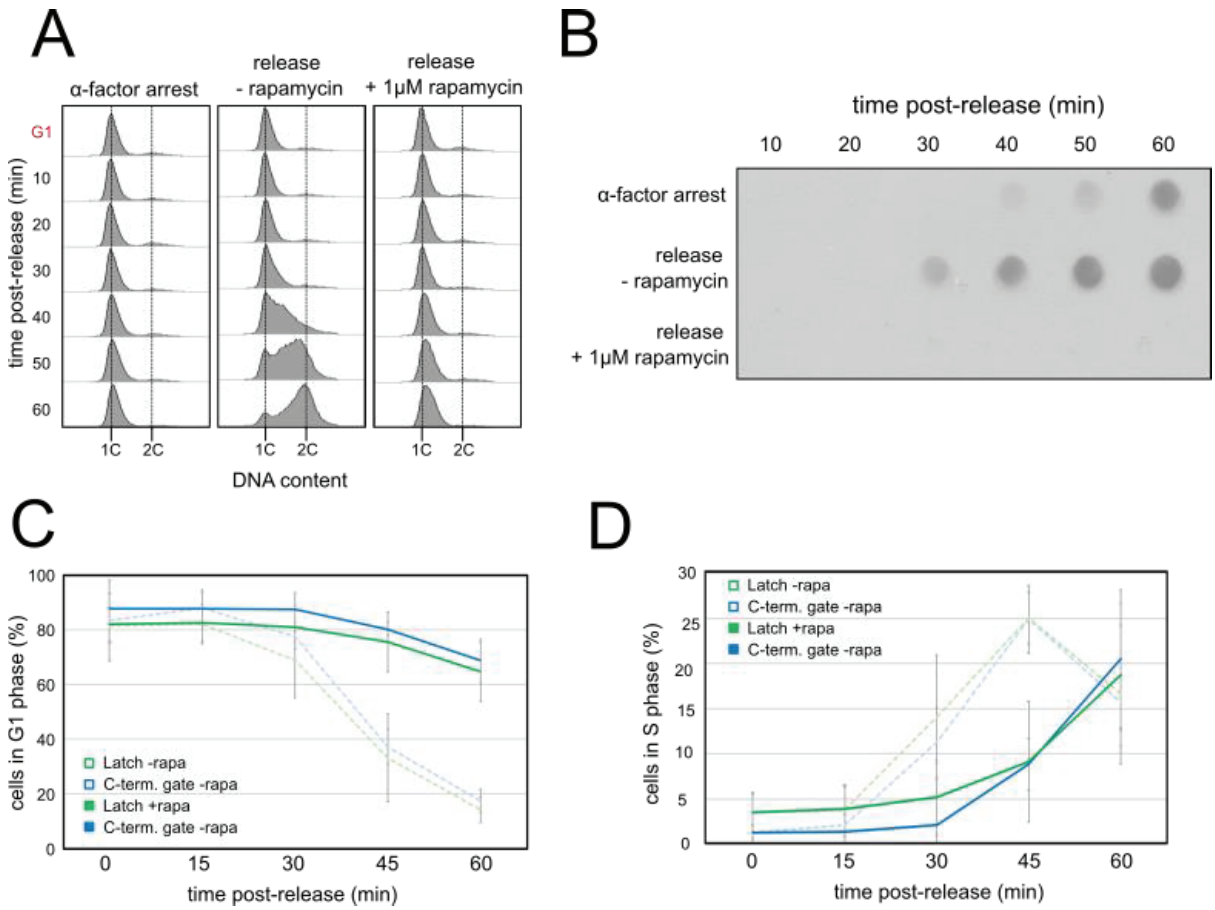


Figure 23: Analysis of BrdU incorporation and replication dynamics in the inter-hexamer latch strain. A) Flow cytometry showing the DNA content profiles of a negative control kept in alpha-factor induced arrest, latch-containing cells released into S phase without the presence of rapamycin, and cells released into S phase with 1μM rapamycin. B) Dot blot showing the relative amounts of BrdU incorporated into genomic DNA at the time-points indicated. Genomic DNA concentration was normalised before membrane binding of an equal volume of DNA-containing solution. C) Flow cytometry time-courses were compared across several replicates of α-factor arrested cells released into either the latch-mediated arrest (n=8) or the C-gate-mediated arrest (n=3). Cells were sorted into bins representing cell cycle stages based on the DNA content as measured by SYTOX Green. Latch strains were arrested using 1 μM rapamycin, and gate strains arrested using 100 nM rapamycin. Shown here is the percentage of total cells in the G1 bin averaged over several time courses. Error bars represent the standard deviation in the mean. D) As for C, plotting the percentage of total cells sorted into the S phase bin.

Whilst the arrest generated by the inter-hexamer latch seemed to be less stable than that generated by the intra-hexamer gates, by increasing the rapamycin concentration, a significant delay in the early stages of replication could be achieved. As such, the latch construct was investigated to show whether the arrest functioned similarly to the intra-hexamer gates. ChIP-qPCR analysis of factors could demonstrate if this construct arrested replisomes at origins, and what the composition of such arrested replisomes would be. Quantification of the arrest of the latch construct versus the previous C-gate construct arrest showed that, when comparing the

changes in cell numbers in each cell cycle phase over the course of the arrest, the latch appeared to show similar replication dynamics to the C-term gate. The rate of exit of cells from G1 (Figure 23C), and the rate of entry of cells into S phase (Figure 23D), as determined by flow cytometry, showed no significant difference between the two constructs over sixty minutes post-release. This was true both in the absence of rapamycin (i.e. normal release from α-factor), and in the presence of appropriate levels of rapamycin to generate an arrest in the respective construct.

The two constructs also showed a similar time of arrest over uninhibited entry to S phase. Based on coarse quantitative measurements of the cell cycle progression in these constructs with and without rapamycin, the latch arrest appears similar to the C-gate arrest. Monitoring the entrance into S phase by the number of cells lost from the G1 population as measured by DNA content from flow cytometry data, both constructs had a similar delay over untreated cells when rapamycin was added. The lag time for 15% of the initial G1 cells to exit into S (modelled as the threshold for clear S phase entry for a proportion of the population) was on average 19.9 minutes longer in the rapamycin treated C-terminal gate over the untreated control, and on average 23.7 minutes longer in the rapamycin-treated latch over the untreated cells, though this difference between constructs was not significant (analysis carried out on the data shown in Figure 23C and D).

The intra-hexamer gate constructs showed clear maintenance of Mcm4 association with origin DNA past the time when in untreated cells, the replisome complex had moved away from origins (Figure 16C, E). The C-terminal gate also had the additional effect of increased Mcm4 origin association at 30 minutes and beyond, something which the N-terminal gate lacked, whilst still showing stable association. Because the latch contains an endogenous MCM4 which has a selection marker present already, for the insertion of the FRB, Mcm2 was chosen to be tagged instead with penta-FLAG.

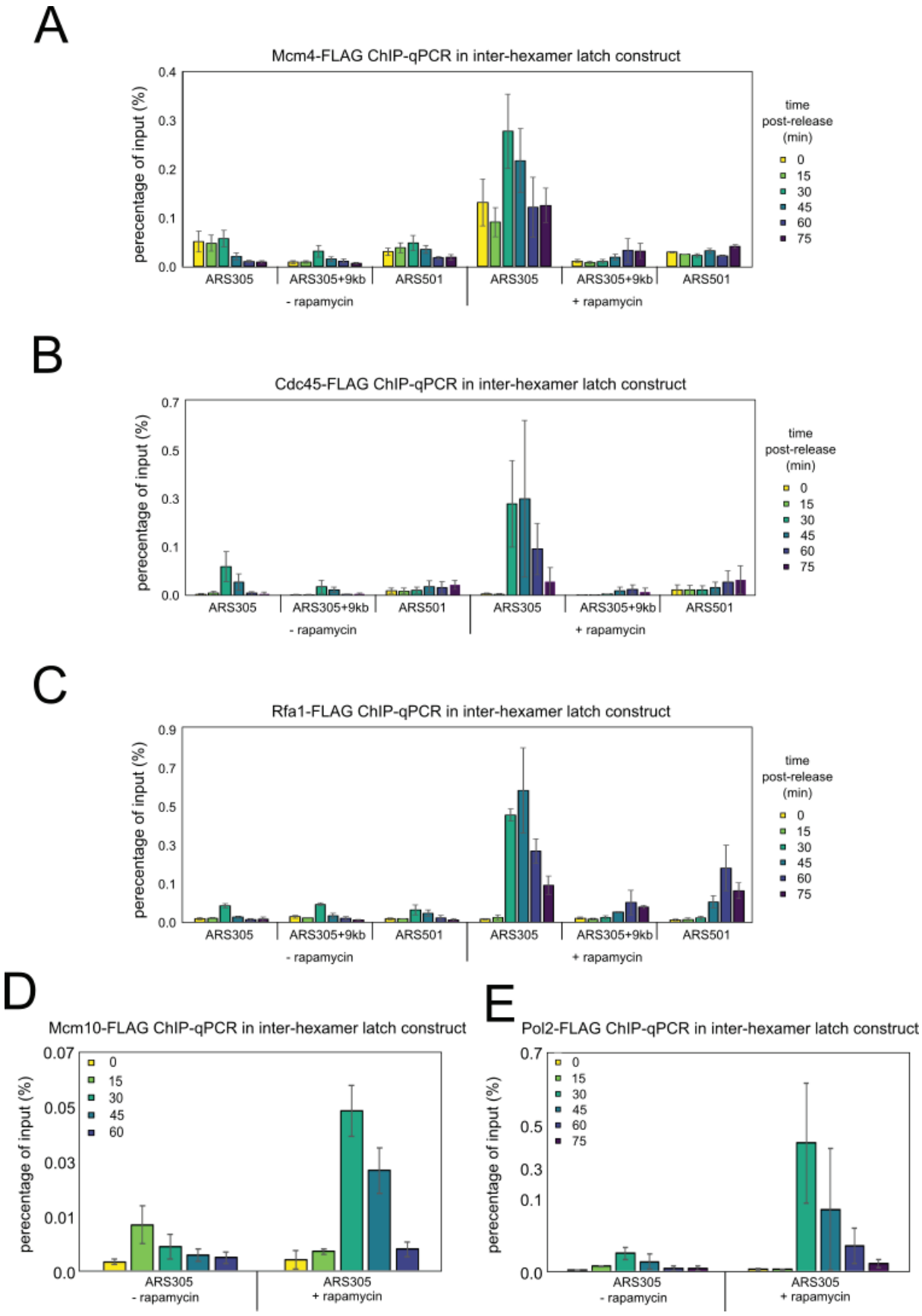


Figure 24: ChIP-qPCR analysis of replication factor origin association in the context of the inter-hexamer latch. A) ChIP-qPCR of Mcm4-FLAG at the early origin ARS305, the late origin ARS501, and non-origin ARS305+9kb DNA. Samples were released from alpha-factor arrest into S phase with and without the addition of 1 μM rapamycin for up to 75 minutes. Error bars for Mcm2 indicate standard error of the mean (s.e.m.) for six individual replicates. B) As for A, but for Cdc45-FLAG, and error bars represent standard deviation (s.d.) of three individual replicates. C) As for B, but for Rfa1-FLAG. D) As for C, but for Mcm10-FLAG. E) As for D, but for Pol2-FLAG. Error bars represent the standard deviation (s.d.) of two replicates.

The origin-association dynamics of Mcm2-FLAG in the latch closely resemble that of Mcm4 in the C-terminal gate and WT strains for the untreated cells (Figure 24A, Figure 17A). The protein maintains stable association with the ARS305 origin DNA until 30 minutes, before helicase action and replisome movement decrease the association beyond this time. In rapamycin-treated cells, Mcm2-FLAG maintains a more-or-less stable association with the origin DNA up to 75 minutes. However, in contrast to the C-terminal gate, the increase in signal is modest. Because of the flow cytometry analysis that shows escape from the latch-induced arrest beginning between 45 and 60 minutes (Figure 22D), it has to be considered that the slight fall-off in pull-down efficiency and overall lower efficiency levels might have been due to the escape of some replisomes from the arrest. Regardless, the ChIP-qPCR data showed a strong association of MCM2-7 with origins at least until the 45-60 minute mark, where some escape from the latch construct-mediated arrest was perhaps expected (Figure 24A). In general for the inter-hexamer latch, the variability of ChIP efficiency between replicates was greater than that observed in the other two gate constructs. The reason for this was likely two-fold. Firstly, because the end-point arrest generated in the latch construct was less stable overall, as evidenced by the flow cytometry, and also highly dependent on rapamycin concentration, the level and immediacy of “escape” from the arrest varied from experiment to experiment. This effect was occasionally visible in flow cytometry analysis (data not shown), and this had a knock-on impact for ChIP-qPCR analysis. In particular, the later time-points were harder to reproduce. The second reason was the development of the experimental handling. Some of the earlier replicates suffer from an inherently higher variability due to technical variance. This variability was greatly reduced over time, through several optimisation methods. One method involving conducting all of the replicates for one replication factor on successive days, which greatly improved reliability, possibly by reducing reagent or cell culture variabilities.

To build up a more complete picture of the arrested complex, ChIP-qPCR was repeated with several other tagged replication factors. The origin association dynamics of Cdc45 were very informative, since this protein showed an increase in association with origin DNA around when origin activation was predicted to begin, and because Cdc45 forms part of the CMG complex, it gives a second marker for replisome movement away from origins. In the inter-hexamer latch, it was promising that the penta-FLAG tagged Cdc45 was tolerated in combination with the latch construct. It was a concern that, in part because it shared one of the insertion sites, the latch construct might share the rapamycin-independent effects on replication initiation impairment. For Cdc45-FLAG association dynamics in untreated cells, the latch again resembled both wild-type and the C-terminal gate cells (Figure 24B, Figure 16, Figure 17). The same correlation between strains was seen for the late origin ARS501 without rapamycin addition. With rapamycin treatment, Cdc45-FLAG in the latch strain showed different dynamics to either the C- or N-terminal gate constructs. The efficiency of pull-down was increased in the latch after 30 minutes compared to the untreated sample, but not to the same extent as for the C-terminal gate. However, it did show an increased efficiency of pull-down, unlike the N-terminal gate. In the inter-hexamer latch, most likely due to replisomes "escaping" the rapamycin-mediated arrest, the association of Cdc45-FLAG with the origin began to fall after 45 minutes, and this correlated with the flow cytometry indication of the timing of the "escape". From the Cdc45-FLAG analysis, then, it seemed that the latch caused a similar arrest phenotype to the C-terminal gate, with a slight "hyper-crosslinking" effect, but the arrest is shorter-lived as the complex escapes the physical block.

To test if the increase in efficiency did correlate with single-stranded DNA production as with the C-terminal exit gate arrest, the association dynamics of Rfa1-FLAG were tested in the context of the inter-hexamer latch (Figure 24C). The results showed a very similar profile to Cdc45, where in untreated cells, the dynamics were identical to wild type and the C-terminal

gate strain. The effects in treated cells were also similar to the previous factor, with an increase in efficiency compared to the untreated cells, that was followed by an overall loss of RPA origin association from 45 minutes onwards. This suggested that, in contrast to the N-terminal gate, the latch arrest blocked the initiating complex at a stage when single-stranded DNA was indeed present. Again, the observed increases in pull-down efficiency post-30 minutes could then be due to the increased "hyper-crosslinking" suggested above in the presence of ssDNA.

The most interesting observation was made with Sld3-FLAG. In both previous constructs, the

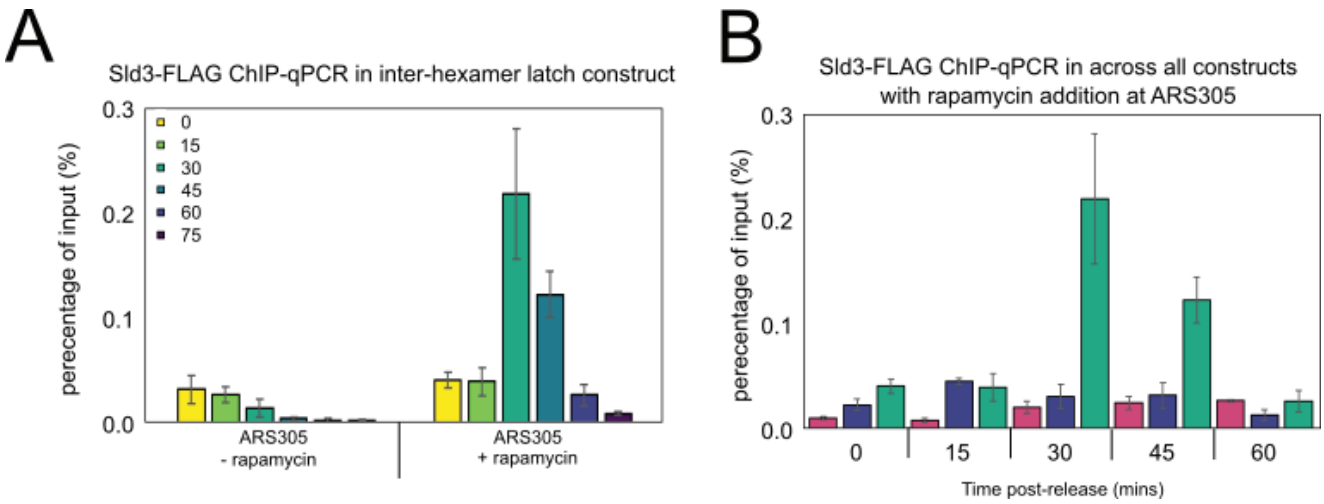


Figure 25: ChIP-qPCR analysis of Sld3-FLAG in the context of the three rapamycin-dependent linkage constructs. A) ChIP-qPCR of Sld3-FLAG in the context of the inter-hexamer latch construct for the ARS305 origin locus. Cells were released from alpha-factor arrest with and without the presence of rapamycin, and samples harvested at the time points displayed. Error bars represent th standard deviation (s.d.) from 3 independent experiments. B) Amalgamation of Sld3-FLAG ChIP-qPCR data from all three constructs, at the ARS305 origin sequence, at various time-points. Pink shows the N-terminal intra-hexamer gate construct data, blue shows the C-terminal intra-hexamer gate construct data, and green the inter-hexamer latch data. Error bars represent the standard deviation (s.d.) from 3 independent experiments.

Sld3 behaviour had not been particularly informative. The N-terminal construct showed association and slow accumulation of Sld3, presumably prior to the production of ssDNA. The C-terminal construct showed a loss of Sld3 association, and, more tellingly, Sld3 did not show any effect from the "hyper-crosslinking" observed at and beyond 30 minutes. In contrast, the latch construct showed association of Sld3 at least as long as some of the other replication factors, and a clear increase in ChIP efficiency at 30 minutes suggested that Sld3-FLAG was undergoing the "hyper-crosslinking" effect (Figure 25). This in turn suggested that Sld3 binding to the initiating complex was coincident with the production of single stranded DNA.

Whilst previous arrest stages generated by the other constructs implied that Sld3 release from the complex occurred prior to ssDNA generation, the latch construct was able to delay Sld3 release until ssDNA production. Based on this data, it was hypothesised that the latch construct interferes with a process that results in or requires the disassociation of Sld3, and that whilst this process usually occurs prior to ssDNA production, completion of the Sld3-releasing process is not fully necessary for the ssDNA production.

The latch construct was designed to prevent inter-hexamer relative motion, including separation, twisting and slipping past each other. Sld3 appears to bind the MCM2-7 double hexamer near the interface, and the tetramer of Sld3/7 observed by crystal structures could suggest that two Sld3/7 dimers, each bound to a separate hexamer, could bridge the hexamers in the DH. The fact that latch-mediated arrest appears to prevent the release of Sld3 could indicate that the release of Sld3 is dependent on double hexamer splitting. If this is the case, and so far the data lacks definitive proof of the latch's physical and structural function, then the combined information of all three constructs may be able to suggest an order to the two large-scale physical rearrangements required for helicase activation.

The early arrest, and minimal amount of ssDNA produced in the N-terminal gate-mediated intermediate, might suggest that initial unwinding occurs near the N-terminus and the DH interface, and blocking the exit channel of the ssDNA in this region prevents much further unwinding. This suggestion of the initial unwinding site is also indicated by other genomic approaches used by L.M. Reuter to study this process (L.M. Reuter, manuscript submitted). After initial unwinding, it is likely that the extrusion progresses from the N-terminus to the C-terminus along the 2/5 DNA gate. This progression causes structural changes that enable the stable recruitment of later-stage replication factors, for example Sld2, Dpb11, GINS and Mcm10. By the time the arrest intermediate generated by the C-terminal gate construct is reached, the early initiation factor Sld3 has disassociated, and many of the factors that form the

replisome progression complex are present, making the complex resemble a fully active helicase complex, but for the remaining Sld2 and Dpb11 proteins, that do not show association with a fully active RPC.

The latch construct arrest shows many of these same late-stage factors associating, with the key difference of Sld3 association. Because of the suggested mode of action of the latch, a hypothesis was proposed where initial and continued unwinding allows the first separation of the hexamers, and this process is carried out by, requires the disassociation of, or causes the release of Sld3. In the closed latch, ssDNA extrusion progresses towards a later stage, but the correct separation of the hexamers may be impaired, and this key structural difference implicates Sld3's involvement in that rearrangement.

ChIP-qPCR was only able to focus on the DNA association of a single replication factor in any given experiment, and use the efficiency of origin DNA pull-down as a proxy for the quantitative binding of a factor to the replication activation complex. Mass spectrometry offered an alternative way to quantitatively investigate factor binding to initiation complexes, without the possible biases introduced by using pulled-down DNA as the read-out of the assay.

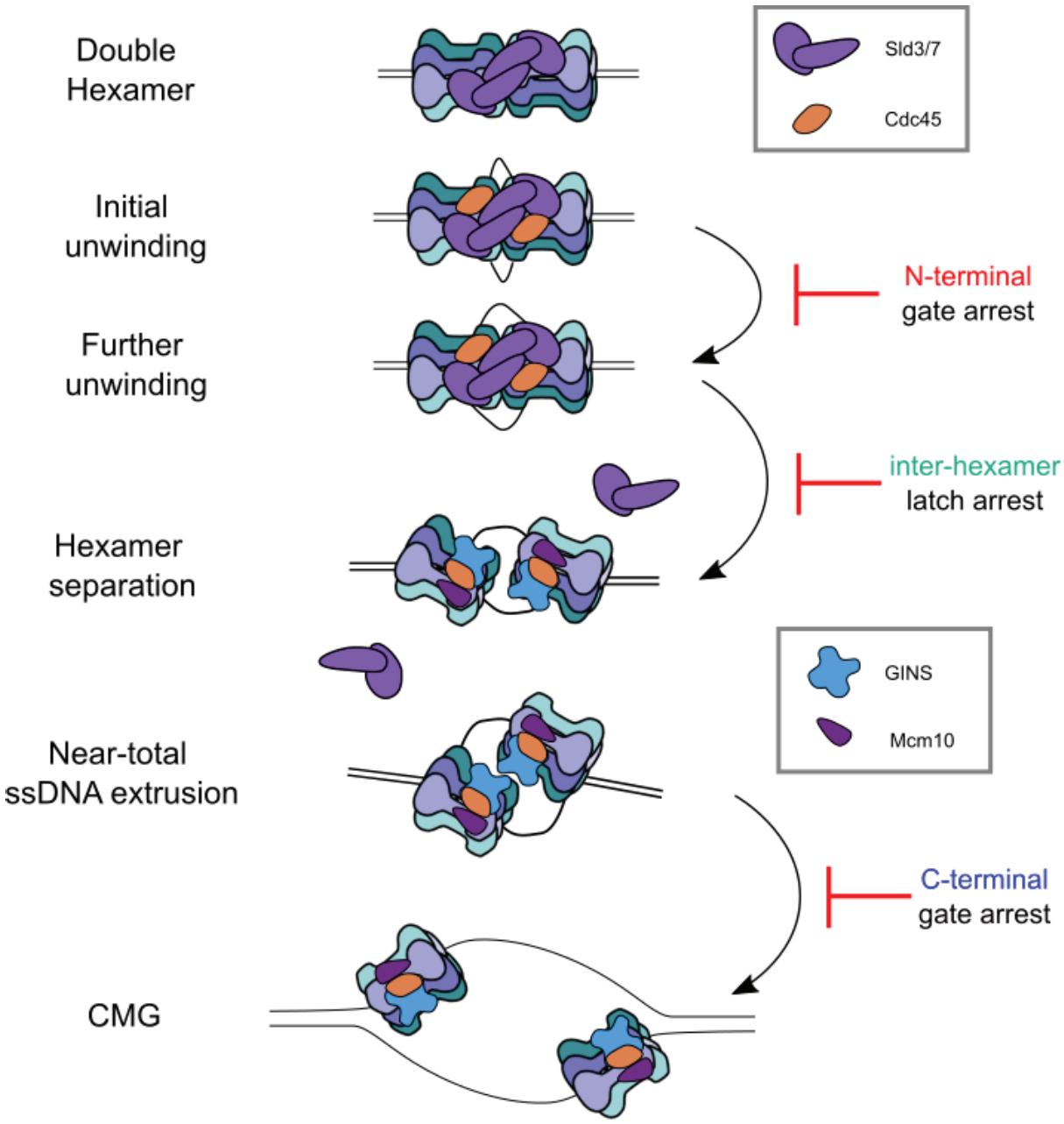


Figure 26: Ordering of structural rearrangements in replication initiation using replication factor association dynamics from two intra-hexamer single-stranded DNA exit gate constructs and an inter-hexamer latch construct. Association or disassociation with origin DNA of these factors, and their interaction with a "hyper-crosslinking" effect, led to proposal of this hypothesis. The block of the most N-terminal region of the ssDNA exit gate arrests complexes early in the initiation process. Blocking the C-terminal region of the 2/5 gate arrests relatively late, and association of RPC complex members is detectable. The inter-hexamer latch arrests between these constructs, and more closely resembles the C-terminal intermediate, but with the retention of Sld3, possibly due to a relationship of this protein with hexamer separation. Sld3/7, Cdc45, GINS and Mcm10 are displayed, whilst other late stage factors are excluded for clarity, though ChIP-qPCR data suggests their arrival at approximately similar times to GINS and Mcm10. Single-stranded DNA exit gate data collected by LM. Reuter, and inter-hexamer latch collected by C. Weekes. Hypothesis proposal and data analysis resulted from discussions between the researchers previously mentioned and C. Speck.

3.2 Native IP of replisome component Psf2

3.2.1 Single-step IP using Psf2-TAP and magnetic IgG-coupled beads

The use of immunoprecipitation to investigate the eukaryotic replisome has been utilised previously, which found a large number of the factors that we now consider core replisome components (Gambus et al., 2006, 2009). Recent advances in Liquid Chromatography coupled Mass Spectroscopy (LC-MS) mean that samples of greater complexity can be processed than ever before (Cox & Mann, 2008; Yu & Huang, 2017). Because of this, we attempted to reproduce the findings of previous work on active replisomes, before adapting the technique to investigate more specific complexes during the initiation of DNA replication. Previously, two subunits of the GINS complex, Sld5 and Psf2 were tagged endogenously with a TAP tag and used to co-purify a "Replisome Progression Complex" (RPC) (Gambus et al., 2006). In this case, instead of using an IgG Sepharose column, magnetic beads were utilised, which tend to have an advantage over agarose beads and matrix columns for the smaller sample sizes that were being targeted in this approach (Safarik & Safarikova, 2004). Significantly smaller volumes of cell culture were used in this study than previously, enabled by the ability to submit an entire immunoprecipitated eluate for LC-MS analysis, whilst retaining the sensitivity and detection capabilities to identify and quantify low-abundance peptides.

The temporal program of origin activation in yeast is well documented, and some of the earliest firing origins are already activated by twenty minutes post-release (Raghuraman et al., 2001; Yabuki, Terashima & Kitada, 2002; Alvino et al., 2007). From experiments carried out by L.M.Reuter (Section 3.1.1), M. Miyazawa-Onami et al. (Miyazawa-Onami, Araki & Tanaka, 2017) and A. Gambus et al. (Gambus et al., 2006), it is clear that GINS associates with origins and replication complexes at this time. As such, cells containing TAP-tagged Psf2 were arrested in G1 using α-factor, and released at 25°C for twenty minutes before harvesting as

described previously (Seki & Diffley, 2000; Gambus et al., 2006) to pull-down replication complexes.

It was observed that the TAP-tag IP was successful in isolating protein bands that were not present, or much less abundant, in the negative control (NC) strain that did not contain a tagged Psf2 (Figure 27, A and B). The silver staining visualisation method is sensitive enough to detect small changes in protein amounts (Blum, Beier & Gross, 1987), and the detection of IP-specific single protein bands can be seen (Figure 27, B, comparing lane 1 and 2). From this technique, however, identification of the IP-specific bands was not possible, though the presence of specific IP bands was a promising sign. Some of these groups of IP-specific bands were at molecular weights corresponding to known replisome components, such as for GINS and MCM (Figure 27, lanes A7 and B2), though this needed confirmation.

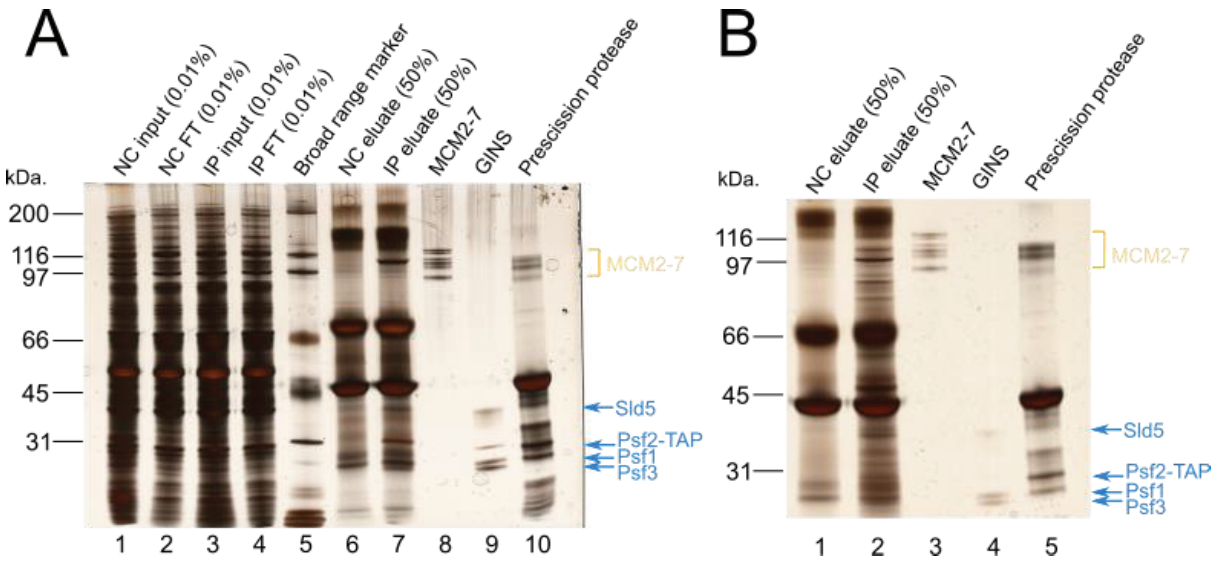


Figure 27: Initial IP experiments using the Psf2-TAP construct and magnetic bead-based purification. A) IP experiment showing inputs and flowthroughs (FT) alongside IP eluates and purified protein complexes. PreScission protease was used in the elution step and is included separately for reference. B) A second replicate IP, showing clearer definition of IP-specific protein bands. MCM2-7 and GINS complex members indicated by yellow and blue markers respectively. MCM2-7 expected Mw: 99 kDa, 107 kDa, 105 kDa, 86 kDa, 113 kDa, 95 kDa. GINS expected Mw: 34 kDa (Sld5), 31 kDa (Psf2+TAP), 24 kDa (Psf1), 22 kDa (Psf3)

These IP tests did exhibit some potential drawbacks of both the IP protocol used, and the silver staining visualisation method. For this process, a single step IP was used, and the elution step required an increase in temperature, reducing conditions in the form of DTT, and addition of PreScission protease. It was clear that, in the negative control (NC), significant protein was

detected (Figure 27, A lane 6 and B lane 1). Four main issues existed in the negative control in terms of protein presence. First, the lab-purified protease itself contributed a large amount of non-specific bands, at both low and high molecular weight, with a particularly strong band around 45 kDa (Figure 27, lane 10). The higher and lower molecular weight bands are likely either purification impurities, or degradation products, with the degradation products only explaining the lower molecular weight bands. The appearance of these protein bands was fairly consistent, but importantly seemed to affect negative control and tagged samples similarly. Other than concealing or disguising other gel bands in their proximity, the bands due to the protease were not considered much of a concern. Similarly, the elution conditions appeared to cause release of some heavy and light chain IgG from the beads, and these produced the band around 66 kDa, as well as a second band overlapping with the protease band. Again, these were shared in the control and IP sample, and so were only an inconvenience. At a height between the 116 and 200 kDa. marker, there was a third dense protein band (Figure 27, A lanes 6 and 7, B lanes 1 and 2). The source of this band was subsequently found to be due to bovine serum albumin (BSA) introduced during preparation of the magnetic beads. This band was more inconvenient since it was quite close to high molecular weight proteins of interest, for example MCM2-7, and quite high in intensity. More concerning was the general protein background that was visible in the control samples. Compared to results obtained using a two-step approach with Sepharose beads and Calmodulin resin (Gambus et al., 2006), this background was much higher. This could be explained by the reduction in purification expected from a single- versus a two-step method, but could also be due to increased non-specific binding introduced with the use of the magnetic beads, which may have different surface interactions than their agarose counterparts.

This experiment also highlights some issues with the silver staining method in general. As mentioned previously, identification of specific proteins when the level of purification does not

show only a few major bands is often impossible. In addition, the quantification of silver staining can pose challenges (Winkler et al., 2007; Westermeier & Marouga, 2005). The development reaction is not self-limiting, and so consistency of staining can pose a challenge across gels, as can be seen by the relatively lower development visible in Figure 27B than in Figure 27A. Furthermore, there can be inconsistencies within the gel as well, since silver staining is reliant on certain functional groups which could lead to some proteins staining less well than others, and high intensity bands often cause a “halo” or “crater” effect where the centre of the band stains yellow/orange rather than grey/black (Westermeier & Marouga, 2005), which can clearly be seen in Figure 27 in many places. Since quantification in this study was intended to be done using mass spectrometry approaches, this did not pose significant problems. A further issue is that because of the sensitivity of the silver stain, and its low dynamic range (Westermeier & Marouga, 2005), the loading of inputs and flowthroughs has to be carefully adjusted so as not to be overdeveloped. The higher complexity of these samples, since they are whole cell extracts, also means that evaluation of the silver stain with regard to inputs and flowthroughs is limited to overall protein levels and pattern matching of highly expressed protein bands to ensure comparable processing of samples during harvesting and cell lysis.

3.2.2 Optimisation of IgG-bound beads to reduce background proteins

Before progressing with further analysis of the pulldown proteins, it was decided to attempt to reduce as much of the non-specific protein from the eluate as possible. This optimisation began with the identification of the high molecular weight protein bands that in some experiments were causing visual obstruction of IP-specific protein bands of interest (Figure 27B). Mock IPs were carried out on beads alone, with no cell extract added (Figure 28A,B). The elution step was carried out in standard elution buffer (see materials and methods) with and without the

addition of PreScission protease, since the elution buffer had relatively high levels of detergent, and contained the reducing agent DTT, which could affect the bead-bound antibody. In both cases, the high molecular weight band was detected in both the eluate and bound to the beads after elution, and the protease had only a small effect on release of this high molecular weight contaminant from the beads (Figure 28, A).

During coupling of the IgG antibody to the bead surface, buffers containing BSA were used to help prevent bead aggregation and potentially block nonspecific interactions with the bead by antibody and other cell extract proteins later on. By varying the levels of BSA used, it was clear that the BSA was responsible for a band at 66 kDa, as well as the higher molecular weight bands, which are likely aggregates of multiple BSA proteins (Figure 28, B). Removing BSA from these buffers heavily reduced these bands, but did not prevent their appearance entirely, possibly due to either the beads or IgG themselves having BSA present in their storage buffers. It was apparent, however, that stable IgG coupling was slightly reduced when BSA was omitted (Figure 28, B, comparing lanes 1 and 3 to lane 7). A test IP was nonetheless carried out with beads prepared without BSA (Figure 28, C). The levels of IP-specific proteins were lower than when using BSA-treated beads, and the non-specific background was less of a non-specific "smear" but showed clearly defined non-specific protein bands (compare Figure 27 and Figure 28, panel C). During this process, the behaviour of these beads was noted to differ from previous immunoprecipitations. Resuspension and washing of these beads was more challenging than previously, and they gathered against the side wall of the tube when on a magnetic rack in a much more dense manner. Microscopic examination of these beads observed a greater degree of aggregation with these beads, which could explain both the reduction in IgG coupling, and subsequent reduction in IP-specific protein binding.

During these initial IPs, the amounts of IgG used to couple to the surface of the beads was conservative, and it was considered that if IgG concentration during bead preparation was

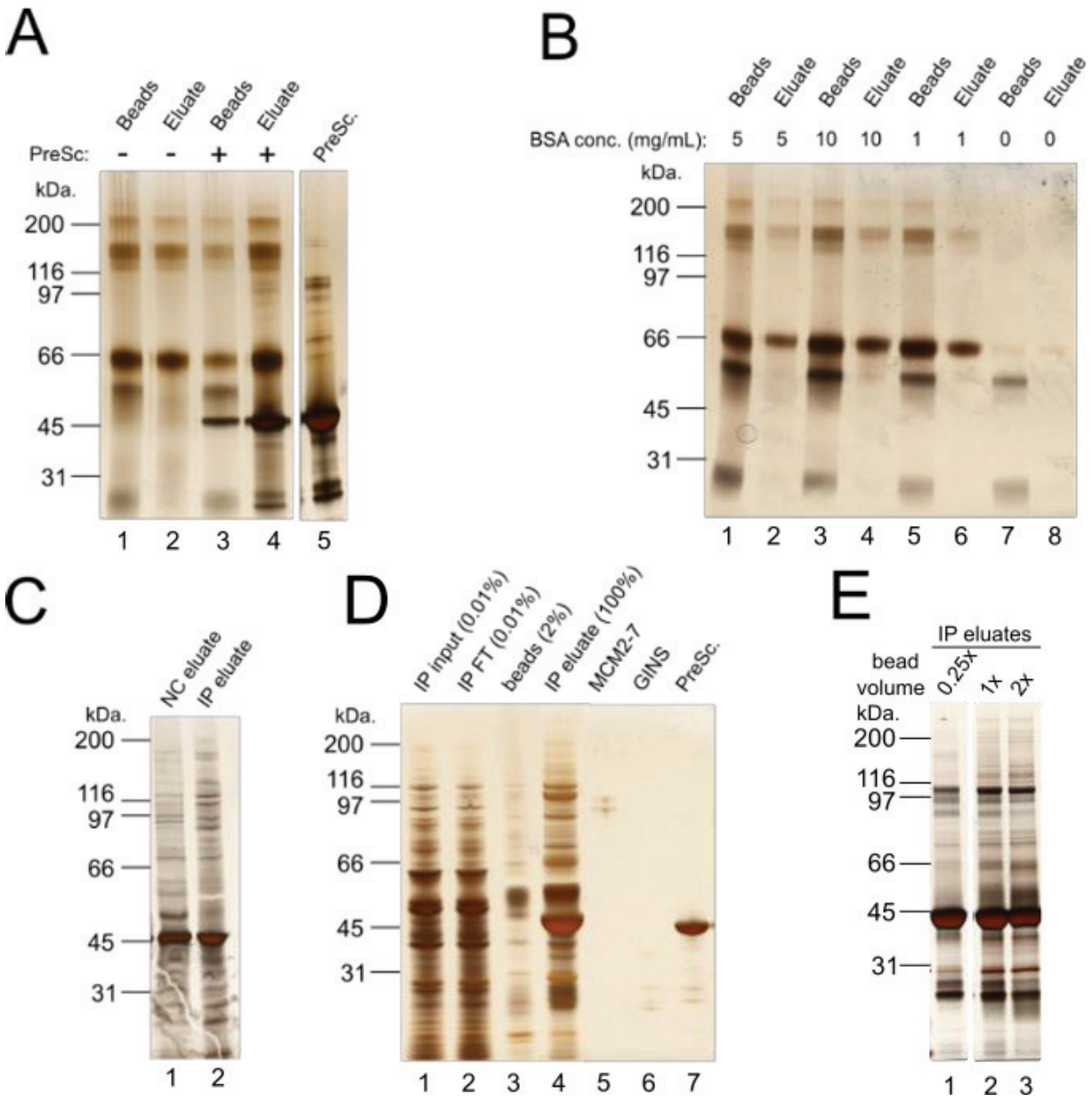


Figure 28: Optimisation of the IgG-coupled magnetic beads. A) Mock IP of IgG-coupled beads, washed as in the regular protocol, and eluted in elution buffer containing or omitting the PreScission (PreSc.)protease itself. The high molecular weight contaminant is present in all samples, irrespective of protease treatment, and is not present in the PreScission protease solution itself. B) Mock IP carried out on beads prepared under different IgG-coupling solution BSA contents. The high molecular weight contaminant presence on beads and in eluate is clearly dependent on BSA content during preparation steps. C) Test IP using beads prepared without the addition of BSA to solutions during the IgG-coupling preparative step. In both negative control (NC) and tagged (IP) strain extract eluates, the background is relatively high, and contains many strong non-specific protein bands. The IP also contains few IP-specific bands and these are not well enriched over the background. D) Test IP using beads with a higher ratio of IgG to beads. The BSA bands are much reduced, and previously observed IP-specific bands are much more evident.E) Test IP with varying ratios of beads to protein extract. Reduction in beads to 0.25x caused significant loss of IP-specific bands, whilst doubling bead amounts yielded no visible improvement.

limiting, this could leave more surface area for BSA to bind non-specifically. To test this, an IP was carried out using BSA levels as described, but with a tenfold increase in the amount of mouse IgG added (increased amount as stated in materials and methods). This had quite an exceptional effect on the IP. The IP-specific proteins were visible and possibly even present at

greater levels than before, and despite BSA presence in prior steps, the level of BSA visible in the eluate was significantly reduced (Figure 28, D).

Due to the presence of IP-specific bands, the next step was to both ensure that the IP was as specific as possible to pull down replisome components, and to begin looking at this complex in the context of cells arrested during replication using the constructs described in Section 3.1. It seemed quite likely that the IP was successfully isolating the GINS complex, due to the consistent presence of a high intensity band at approximately 31 kDa, which matched reasonably closely to the expected molecular weight of Psf2 when modified with the TAP tag we used, once it has been cleaved to leave the relatively small V5 tag. This band in clearly visible in Figure 28D as a light orange band at the 31 kDa marker, and in Figure 27A as a dark brown band at the same location.

3.2.3 Wash step optimisation

It was possible that some of the non-specific proteins carried through into the eluate was because of proteins coating the surface of the larger 15 mL Falcon tubes used in the wash steps. Additionally, the volume of beads used was small compared to the wash tube, making the loss of beads during aspiration greater than desired. The IP was repeated as before, but wash steps were carried out in 1.5 mL low-protein binding tubes, and beads were transferred into new tubes after each wash step. During trials of smaller wash volumes in 1.5 mL tubes, samples of the wash solutions after removal from the beads were kept to observe how clean the flow-through was considering the smaller wash volume. It was determined from this experiment that the washes were likely enough to ensure a level of specificity to the pulldown (Figure 29, panel A), and that the eluate itself contained proteins that were members of both the immediate target complex GINS, which the tagged protein is a stable component of (Figure 29B, lane 8), and proteins of the larger replisome complex of the CMG (Figure 29C, lane 8). Of what IP-specific protein was bound to the beads, only very minimal amounts were detected in the washes,

indicating a stable association (Figure 29, panels B and C, lanes 4-7). In addition, it appeared that most of the bound proteins of interest were eluted from the beads upon treatment with the PreScission protease (Figure 29, panels B and C, lane 9). The IP protocol used to generate these samples was therefore taken forward to investigate the composition of the pull-down by mass spectrometry approaches.

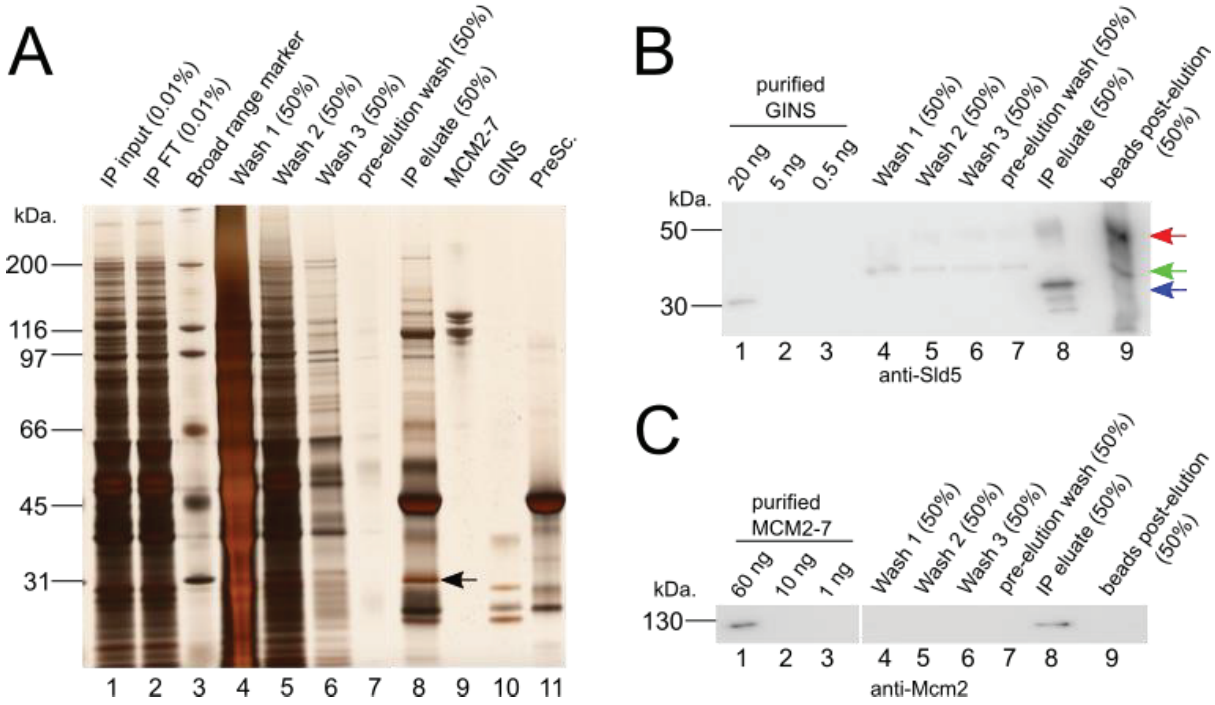


Figure 29: Reducing wash volume and replacing tubes has no observable negative effect on IP specificity. A) Single-stage IP of Psf2-TAP using low protein-binding Eppendorf tubes for wash steps with tube replacement. Initial washes contain large amounts of protein (lanes 4 and 5), but subsequent wash steps and the pre-elution buffer exchange wash contain much lower protein amounts. The elution has relatively low background, and several previously identified IP-specific bands can be identified, including what is likely Psf2 itself at the 31 kDa marker level (marked with arrow). B) Western blotting samples of the same IP with an anti-Sld5 antibody (#42, Genscript (108-Sld5)) show that SLd5 is present in the elution at a much greater rate than in the wash flow-throughs (blue arrow). There is some cross-reactivity with the rabbit light chain IgG visible due to use of an anti-rabbit secondary (red arrow), and also some possible reactivity with the protein A present in the Psf2-TAP prior to elution (green arrow). C) Western blotting the samples with anti-Mcm2 antibody (#5, CSHL in-house (49-Mcm2)) shows presence of Mcm2 in the IP (expected size 107 kDa), and successful removal from the beads, with no detectable losses in wash steps following IP incubation

3.2.4 Investigation of pull-down complex identity

Once a sufficient pull-down of replication complexes was established, the aim of this approach was to combine a purification protocol with the intra-hexamer linkage constructs allowing arrest and study of the replication complex during initiation. As such, the tagged Psf2-TAP would be combined with Mcm2/5 intra-hexamer gate constructs, allowing purification of rapamycin-arrested complexes that might represent an intermediate step in activation.

3.2.4.1 Validation of IP function in DNA-exit gate strains

As shown above (Figure 28D), pulldown of GINS via a TAP tag gives multiple IP-specific bands, some of which can be speculatively assigned to match the molecular weights of expected complex members, for example the four GINS subunits themselves. The presence of Mcm2 was also confirmed in the elution (Figure 29C). To confirm that strains carrying both the TAP

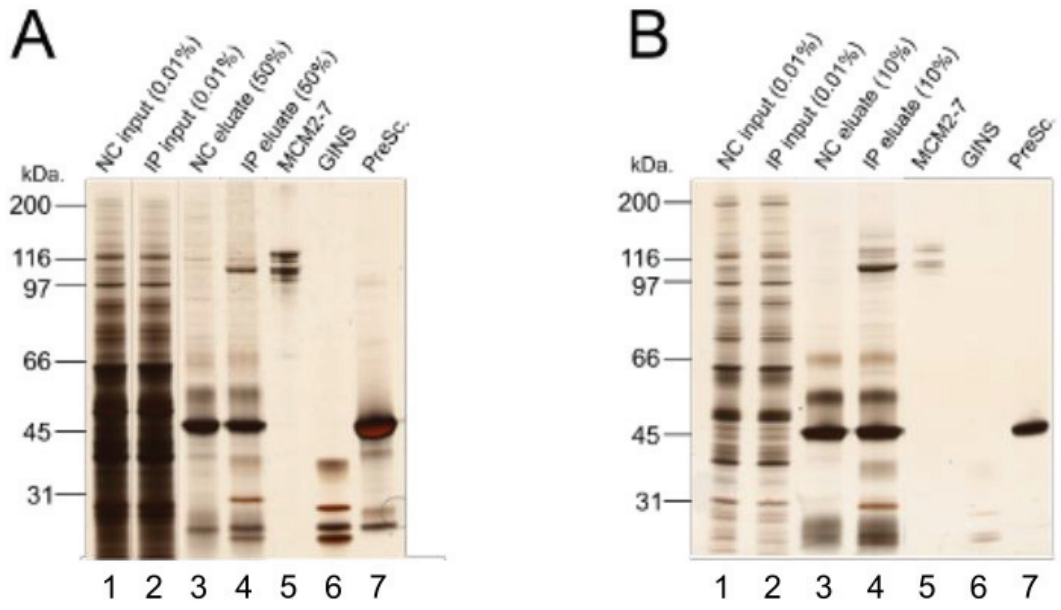


Figure 30: IPs of Psf2-TAP tagged or control strains in the context of the intra-hexamer linkages. A) IP of Psf2-TAP from strains containing the N-terminal gate, but not treated with rapamycin. GINS can be seen in an IP-specific manner, but the intensity of higher molecular weight IP-specific bands is low. B) IP of Psf2-TAP from strains containing the C-terminal gate, also not treated with rapamycin. GINS is visible in an IP-specific manner (see comparison purified GINS for reference), and higher molecular weight IP-specific bands are also visible. In both panels, lanes containing other conditions have been omitted for visual clarity.

tagged Psf2 and the DNA-exit gate constructs allowed efficient pull-down, the IP was repeated using these strains for analysis by mass spectrometry (Figure 30).

The C-terminal gate construct IP showed clear enrichment of certain protein bands, including higher molecular weight bands running at or around the running level of the MCM2-7 proteins (Figure 30, panel B), with many other IP-specific bands also visible, suggesting pull-down of a quite large associated complex, possibly resembling the replisome progression complex (RPC). However, the N-terminal gate construct did not show as clear an enrichment of protein factors in the tagged IP sample (Figure 30A, lanes 3 vs. 4).

This finding may reflect the fact that the binding sites for several proteins involved in activation are found towards the N-terminal tier of the MCM2-7 hexamer, including GINS (Sun et al., 2015; Yuan et al., 2016; Deegan, Yeeles & Diffley, 2016; Georgescu et al., 2017), and it is possible that the insertion of the rapamycin-dependent linkages, even in the absence of rapamycin, and despite any major impact on cellular viability (pers. comm. L.M. Reuter), may affect the stability of GINS binding in the context of a pull-down.

To test the ability of this pull-down to isolate an intermediate complex generated by the N- and C-terminal intra-hexamer linkage constructs, the pull-down was repeated with and without rapamycin.

Both of the DNA exit gate constructs were tested for their compatibility with the Psf2-TAP IP. In the case of the N-terminal gate, which showed already reduced purification of high molecular weight bands, the addition of rapamycin further reduced the binding of these bands of interest (Figure 31A, lane 3 vs. 2). Instead, the IP-specific bands were limited to the lower-

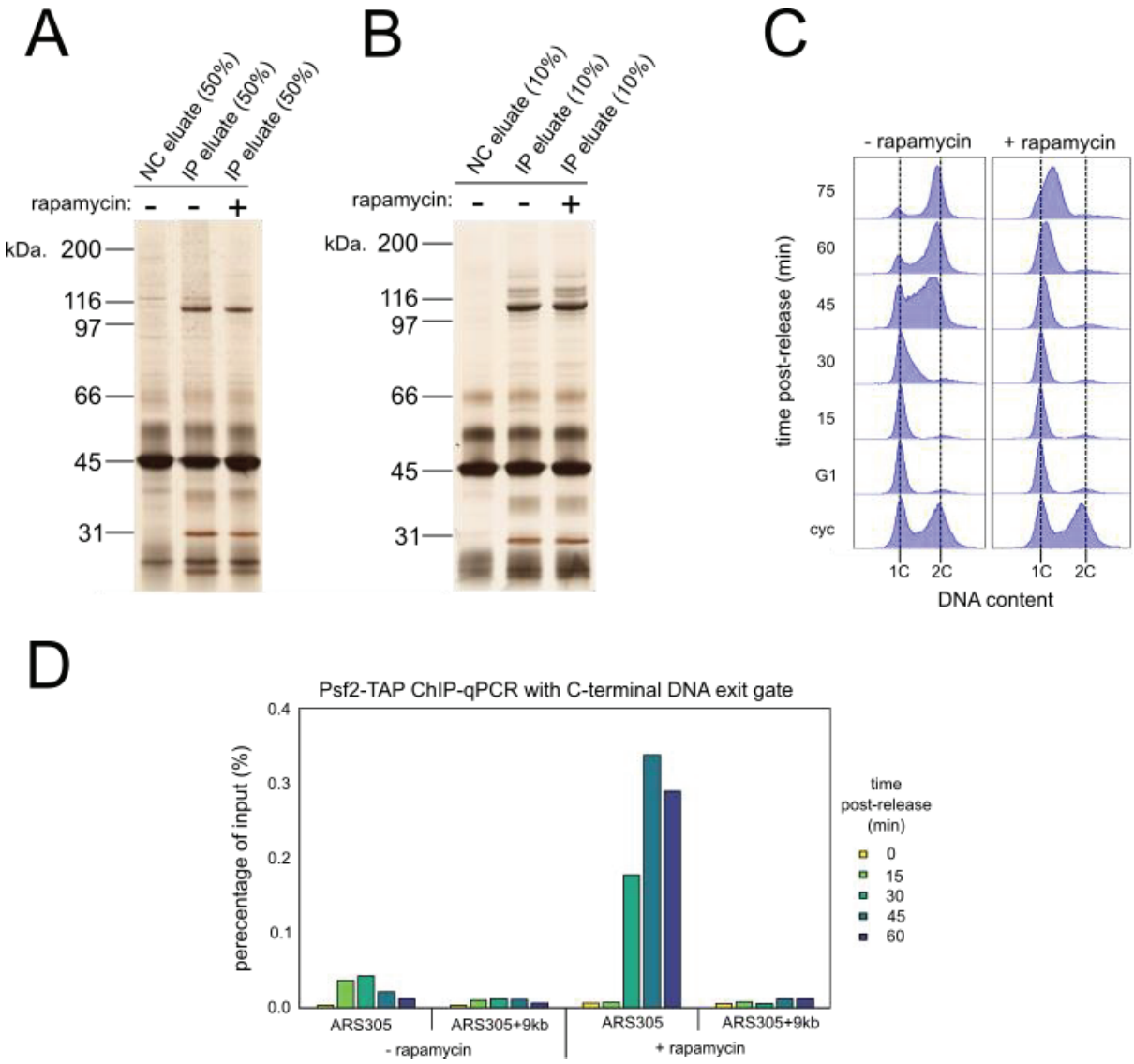


Figure 31: Impact of addition of rapamycin to DNA-gate strains containing Psf2-TAP on IP efficiency and GINS binding. A) expanded version of Fig. 10A above, showing IP of Psf2-TAP in the context of the N-terminal DNA exit gate construct with and without the addition of rapamycin. Rapamycin addition causes a general reduction of high molecular weight IP-dependent proteins. B) As in A, but in the context of the C-terminal gate construct, and an expanded version of Fig. 30B instead. Here, rapamycin addition causes no detectable change in the number or intensity of the IP-dependent protein bands. C) Flow cytometry analysis of C-terminal DNA exit gate containing cells released from α-factor, monitoring change in DNA content in the presence/absence of rapamycin-mediated gate obstruction. Rapamycin addition causes significant arrest of cells and acute prevention of S phase entry D) ChIP-qPCR analysis of Psf2 association with origin DNA in vivo following α-factor arrest in the presence/absence of rapamycin. ARS305 is an early-activating origin, and ARS305+9kb acts as a negative control for origin DNA, that can also detect the passage of mobile replisomes during S phase progression.

molecular weight bands suspected to be GINS, and a single high molecular weight band (later suggested to be Ctf4). Based on the work of L. Reuter (Section 3.1.1), this effect can be explained by the early stage at which the N-terminal DNA gate construct caused arrest of complex progression. The N-terminal gate, prevented the association of many late-stage replication factors, including GINS (Section 3.1.1.1). By preventing GINS association in this

way, rapamycin treatment would preclude the formation of a complex containing GINS bound to the replisome, and prevent pull-down of potential replication complexes.

The C-terminal gate, contrastingly, showed very little difference in the IP-specific bands observed by silver stain between rapamycin treated and non-treated cells (Figure 31B, lane 3 vs. 2). Both high molecular weight bands and low molecular weight bands appeared identical. Whilst surprising, since this construct did show differences in replication factor recruitment by ChIP-qPCR (Section 3.1.1.1), this result appeared to confirm that rapamycin treatment did not prevent GINS association with the replication complex.

For this reason, further investigation of the Psf2-TAP IP in the context of intra-hexamer linkage constructs was limited to the C-terminal gate.

ChIP-qPCR of Psf2-TAP tag was carried out to establish how the intra-hexamer linkage would affect GINS recruitment to origins and the replication complex in this strain (Figure 31D). In cells not treated with rapamycin, pull-down of Psf2 yielded ARS305 origin DNA that showed association as expected. Levels of GINS binding to origins was maximum between 15 and 30 minutes post release, which correlates well with the beginning of DNA replication as observed in flow cytometry data (Figure 31C). Levels of origin association then decrease, as GINS maintains binding with the replisome as it moves away from the origin. In the rapamycin treated cells, GINS association begins at a similar time, but is maintained for far longer at the origin, and the magnitude of origin DNA isolated is much higher, due to cross-linking effects.

Since the Psf2-TAP IP could presumably isolate GINS containing complexes in both unperturbed S phase and in a gate-arrested activation step, samples of the IP elution under these conditions was analysed by mass spectroscopy to obtain a more detailed and specific protein composition profile for each of these complexes (Figure 32).

3.2.4.2 Mass spec analysis of GINS-associated complex in non-arrested and rapamycin-arrested contexts

To obtain a list of the most enriched factors, an enrichment value was calculated by comparing label-free quantification (LFQ) normalised protein intensities between the pull-down and negative control samples. This value was then $\log_2$ transformed to generate an enrichment score. These enrichment scores were compared between replicates to observe which proteins were consistently enriched. In the nascency of the project, because the samples were run independently on the MS instrument, direct comparison of protein intensities between replicates was not attempted, due to variabilities from technical effects that might be proportionately large compared to biological variability. However, comparison of enrichment scores from paired pull-downs and their associated controls provided a reasonably reliable measure of which proteins were enriched consistently.

Silver staining of the pull-downs revealed very clear IP-specific bands, which provided confidence in the sample quality (Figure 32A). Downstream protein identification also provided enough quantitative evidence to retrospectively suggest identities for the clearest IP-specific bands, and so those bands which could be tentatively assigned an identity based on molecular weight and abundance in the mass spectrometry are marked (Figure 32A). The most enriched proteins in this experiment were the GINS proteins, Psf1-3 and Sld5 (Figure 32B). Given the tag was fused to Psf2, and these proteins form a stable complex (Takayama et al., 2003), it was expected that soluble GINS would be a large component of the pull-down. Ctf4, too, has extensive interaction with GINS (Simon et al., 2014), and has been previously shown to purify well in tandem with GINS (Gambus et al., 2009). This analysis also showed clear enrichment of the MCM2-7 helicase (Figure 32C), which provided evidence that this GINS pull-down was able to isolate active or activating replisomes from DNA.

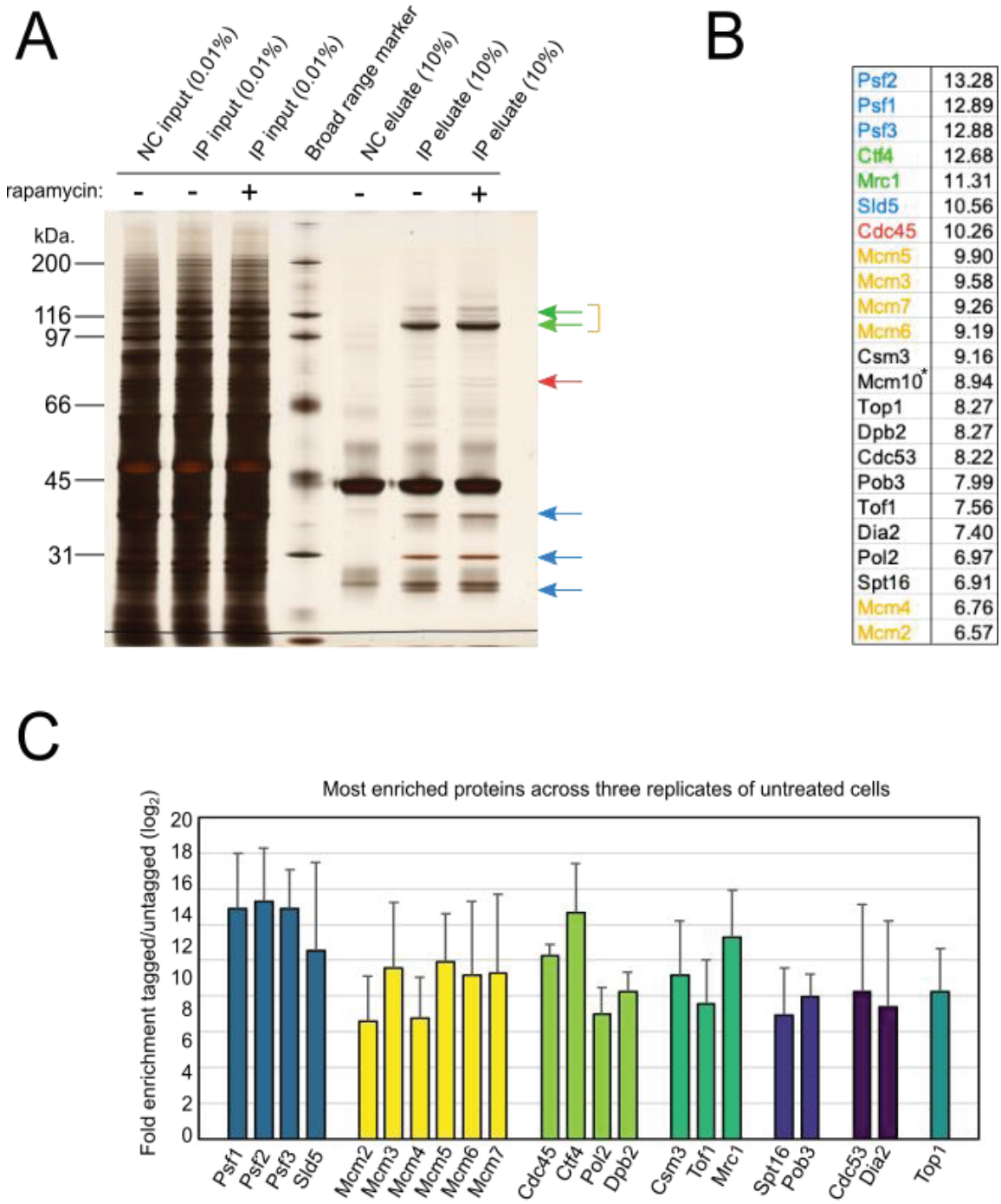


Figure 32: Initial mass spectrometry analysis of Psf2-TAP pull-downs indicates specific IP of many known replisome progression complex members. A) Representative silver stain shows clear enrichment of IP-specific proteins, from both rapamycin-treated and untreated cells. Cells were released from α-factor arrest before harvesting on ice after 40 minutes at 25℃. Based on expected molecular weights and aided by MS quantification (see panel B), plausible assignment of protein bands identity is possible. (blue arrows: GINS complex, including Sld5, top arrow, Psf2, middle arrow, and Psf1 and Psf3, lower arrow. Red arrow: bands that likely include Cdc45. Light green arrow: Ctf4. Dark green arrow: Mrc1. Orange bracket: MCM2-7 expected position.). B) Most enriched proteins over three independent replicates. Enrichment presented as mean of enrichment values in each replicate. Where proteins were not detected in the untagged control, manual imputation based on the lowest intensity proteins identified provided an enrichment value. Mcm10 was detected only in the IP of replicate 1, but with a high enough imputed enrichment to merit display. C) Mean enrichments displayed for these proteins, grouped in subcategories based on either subcomplex identity, or known core RPC identity. Error bars represent the s.d. about the mean enrichment, to display the variability in this value. Clearly enriched proteins include the GINS complex, the MCM2-7 helicase, several members of active RPC (Cdc45-Dpb2), the fork protection complex (Csm3, Tof1 and Mrc1) and other origin or replisome associated proteins.

Other members of the replisome were also among the most enriched proteins, showing consistency with previous GINS IPs (Gambus et al., 2006, 2009). Cdc45 and the two largest subunits of polymerase ε, which structural studies have shown to be closely associated with the GINS-coupled helicase, are present, as well as the three members of the fork-protection complex, Csm3, Mrc1 and Tof1, of which recent structures have shown significant interaction with the N-terminal ring of MCM2-7 (Baretić et al., 2020). Also present were members of the FACT complex, Spt16 and Pob3, and Topoisomerase I, Top1, all known to associate with GINS pull-down (Gambus et al., 2006). Interestingly, members of the SCF^{Dia2} ubiquitin ligase, Dia2 and Cdc53, were also enriched in this pull-down. These proteins are involved in the ubiquitination and subsequent disassembly of the replisome during termination (Mimura et al., 2009; Maculins et al., 2015), suggesting at the 40 minutes post-release that cells were harvested at, a not insignificant number of replisomes were undergoing termination and disassembly. It is also possible, however that this complex associated with MCM2-7 complexes during isolation, as removal from DNA may present a binding surface for this termination-associated complex.

Whilst not detected in the subsequent replicates, Mcm10 was detected in the first replicate of the Psf2-TAP pull-down. Due to its consistent absence in negative control samples, its enrichment was imputed based on the intensity of the lowest abundance proteins in the negative control, on the premise that, since mass spectrometry detection has a reasonably clear cut-off limit, since Mcm10 peptides were not detected in the negative control, its abundance is at or below the intensity measured for the lowest detected protein. By proxying negative control intensity of Mcm10 as this upper bound, an approximation of enrichment can be calculated. This should be an underestimate, since the true “intensity” of Mcm10 in the negative control is likely well below that of the lowest detected protein. Nonetheless, caution should be advised when carrying out such operations. Other proteins in this list also utilised imputation, such as

Top1 and Mrc1, but showed a much more consistent presence in tagged samples. The inconsistency in Mcm10 pull-down is likely related to previous findings that Mcm10 seems to maintain association with the complex during pull-downs at very high salt, possibly due to extensive use of hydrophobic interactions (Gambus et al., 2006). This IP was at relatively lower salt levels, hence Mcm10 association would be generally compromised.

Extending the analysis to all detected peptides in untreated pull-downs confirms previous observations from the most enriched proteins (Figure 33). The complex or complexes isolated contain many factors associated with the replisome progression complex, and reliably recruited the MCM2-7 helicase, and all members of the core CMG. In a strain where RPC formation and subsequent S phase progression was not impeded, the Psf2-TAP IP reflects an active replisome composition. To investigate what effect the prevention of ssDNA extrusion would have on the complex composition, rapamycin-treated samples were compared against these untreated samples (Figure 34).

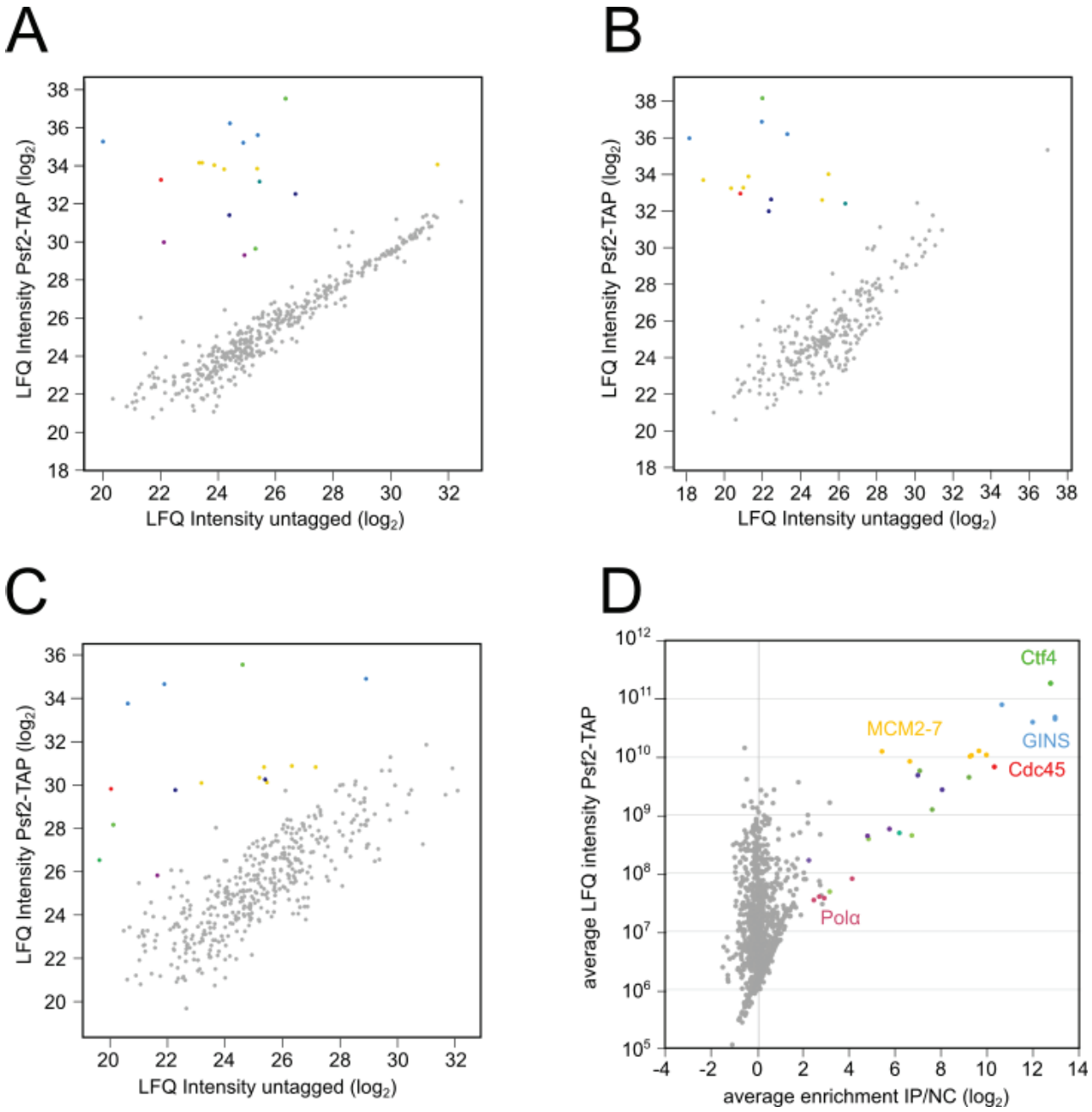


Figure 33: Analysis of all detected proteins in non rapamycin-treated Psf2-TAP pull-downs. The Perseus software from MaxQuant was used to generate scatter plots of normalised intensity in tagged vs. untagged samples. A) scatter plot of Psf2-TAP IP intensity (Y-axis) vs. untagged negative control (X-axis) for replicate 1. Colour scheme as for Fig. R6B/C: yellow/orange = MCM2-7, blue = GINS, red = Cdc45, light green = Ctf4, Pol2, Dpb2, dark green = Fork protection complex, purple = SCF^{Dia2}, dark blue = FACT, turquoise = Top1. B) as for A), but for replicate 2. C) as for A), but replicate 3. D) Plot of average enrichment across all three replicates against average intensity in the pull-down sample. Many consistently enriched proteins are observed in the top right following the colour scheme above, with the addition of Polymerase alpha (pink), and Mcm10 (purple, above polymerase alpha).

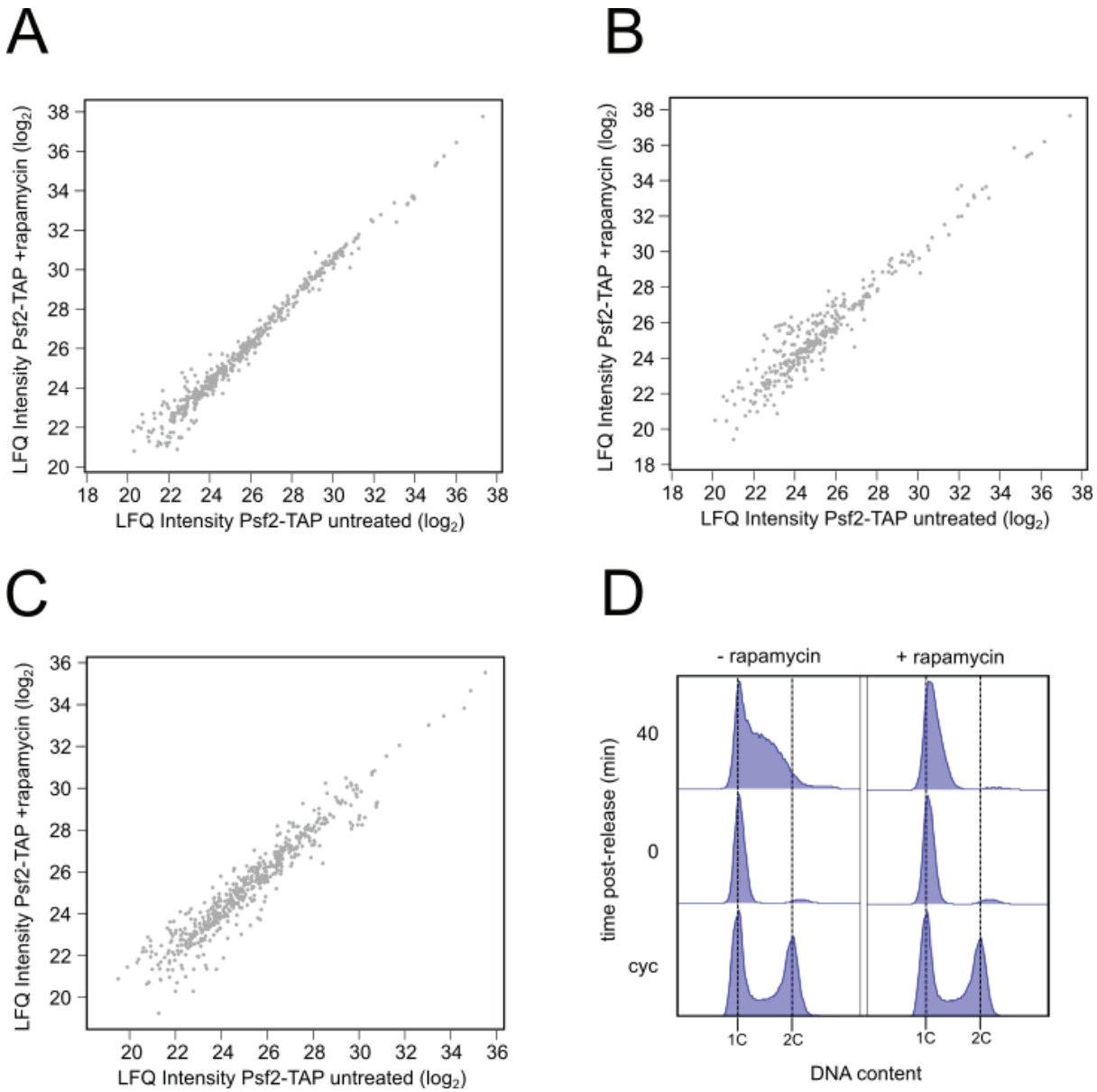


Figure 34: Analysis based upon all proteins detected in pull-down samples with and without intra-hexamer linkage binding. As in Figure 33, protein intensities for each replicate were plotted against each other. In this case, instead of tagged vs. untagged, the scatter plot shows the correlation between Psf2-TAP pull-down from rapamycin-treated cells and a Psf2-TAP pull-down from untreated cells. A) scatter plot of Psf2-TAP protein intensity in rapamycin-treated sample (Y-axis) vs. untreated sample (X-axis) for replicate 1. B) as in A), but for replicate 2. C) As in A), but for replicate 3. D) Representative flow cytometry analysis of DNA replication in cells used for the IP. Rapamycin treatment prevents normal entry into S phase and limits DNA replication.

Surprisingly, it was observed that rapamycin treatment had little effect on the overall composition of the GINS-associating complex (Figure 34, panels A-C). The protein intensities demonstrated good linear correlation, and there were no clear and consistent variants from this overall trend. This seemed to contradict previous results involving this construct (Section. 3.1.1.1). In previous ChIP-qPCR experiments, stable association of the more transient replication factors Sld2, Sld3/7 and Dpb11 can clearly be seen at origins. It had been hypothesised that these proteins in particular, since they do not form part of the active replisome, would be enriched in the stalled intermediates generated by the DNA gate strain. However, similarly to the active replisome pull-down, these proteins were not detected at all in the arrested state achieved with rapamycin. In addition, because the ChIP-qPCR was only able to investigate factors that were chosen to be tagged, the pull-down was expected to be more of an "agnostic" approach, which could identify new factors that previously had not been identified as necessary in this activation process. However, no proteins showed a consistent enough or large enough enrichment in the rapamycin-treated sample over the untreated samples to suggest as possible new interactors. Importantly, the non-rapamycin-treated cells with Psf2-TAP and the gate displayed no replication-associated phenotype of any kind, and cells showed no loss of viability or replication speed, which suggests there was no specific defect introduced by interference of the gate and the TAP-tagged Psf2-TAP.

Why, then, did the pull-down method utilised above, not reproduce the data observed in ChIP-qPCR experiments? There are several possible reasons:

Firstly, ChIP-qPCR has been known to produce spurious results. In particular, DNA loci that have a very "open" chromatin state, such as transcription initiation sites on highly expressed genes, can produce false positives in terms of DNA pull-down efficiency (Teytelman et al., 2013). Origins of replication can be adjacent to active promoters and do have a defined nucleosome-free region (ML et al., 2010). However, given the time-specific association of the

replication factors and the relatively high efficiency values observed in the ChIP-qPCR compared to negative control non-origin regions, this would seem to not be a likely cause.

Secondly, whilst ChIP-qPCR results indicate that in the C-terminal intra-hexamer gate construct GINS shows strong association with the origin, it may be in a configuration unlike that seen in the CMG. It is possible that before full activation has been achieved, and DNA extrusion has occurred, GINS is less stably bound to the replisome. This hypothesis would fit with the idea that GINS and Cdc45 cover the DNA exit gate and prevent loss of the leading strand (Petojevic et al., 2015). For the lagging strand extrusion to occur, GINS must not yet be in such a binding state as to prevent gate opening. In this case, the binding mode of GINS may not be stable enough to enable complex pull-down through sequential washes, and so any activating complexes would not be represented in the final mass spectrometry analysis. However, aside from a small de-enrichment in replicate 3, the amount of MCM2-7 and other replisome components detected in the pull-downs did not decrease as might be expected in an unstable pull-down. How would MCM2-7 be pulled-down in this instance? This is not entirely clear, but possibilities included some replisomes that escaped the rapamycin-induced arrest, or complexes that progressed in assembly despite rapamycin addition, resulting in a more stable CMG-like state.

Complete escape of replisomes seems unlikely, as FACS analysis shows a stable arrest in the rapamycin-treated strain, but the small amount of DNA synthesis observed in these strains during arrest, hypothesised to be due to mitochondrial replication which uses a minimal and distinct replisome (Milenkovic et al., 2013), could also be due to replisomes that escape the gate-mediated arrest. This possibility also assumes that there are enough of these complexes to saturate the IP beads to the same level as a cell extract containing freely replicating CMG, which given the optimisation of bead levels carried out, also weighs in against this hypothesis (Figure 34E).

The alternative hypothesis could explain this effect, but does not fully explain why the complex was able to escape rapamycin enough to release transient replication factors that are present in the ChIP-qPCR, but not enough to progress with replication. It is possible that during the harvesting and cell lysis steps, the concentration of rapamycin is lowered, and allows partial release of the gate arrest, enough for the potentially quick progression through an activation step. All these processes are carried out at 4°C, however, so cellular processes should be slowed dramatically.

Finally, an explanation of the similarity between rapamycin-treated and untreated complexes could be found in the alternate preparation methods used in these protocols. The ChIP-qPCR approach fixes cells in 1% formaldehyde, which causes cross-links between protein and DNA, and also between proteins and other proteins. This has the added benefit of causing a very immediate halt to cellular processes, as protein complexes are cross-linked and presumably fail to continue to function correctly. The replication factors' origin-association observed in the ChIP-qPCR could be dependent on transient and weak interactions, which only survive through a pull-down procedure when proteins are extensively cross-linked together. The fixation may prevent the disruption of the complex and preserve the interaction of transient replication factors. The instability of these initiation factors' association may also explain their absence from even the active CMG pull-downs. The temporal variability between early and late origins in yeast means that activation complexes should be present, albeit in small proportions, even in a cell mid-S-phase. Perhaps the lack of these initiation factors in these pull-downs and others is due to the lability of the complexes they form during the activation of DNA replication.

3.3 ChIP-MS to investigate complex composition and interactors

Immuno-precipitations have proven valuable in the past but rely on the presence of a reasonably stable complex, and optimisation of buffer conditions to ensure integrity of the complex during the pull-down steps. This was apparent even for the relatively stable replisome progression complex, where salt concentrations had to be very high in the buffers for Mcm10 association with the complex to be stable. The constructs used in this project were designed to arrest this complex prior to formation of the stable RPC, during structural rearrangements involving single-strand DNA exclusion. The proteins involved during this process are unlikely to be as stably associated as factors are with the CMG, and there are initiation factors that are ever only transiently associated with origins, e.g. Sld3/7, Sld2 and Dpb11, and so it is possible their binding modes are much less stable than factors that associate with the RPC during the main stage of S phase. Hypothetically, even if the time-point of activation were chosen precisely enough to guarantee the highest yield of activation intermediates, or a construct were used to stall all replisomes at an intermediate step, the complexes present in these samples might be so labile that they would not survive the pull-down and washing steps. In addition, there is a possibility that after cell lysis, the stalled gate construct may resolve from an unstable intermediate, into a more stable form reminiscent of the CMG. As mentioned above, this could explain why a non-gate-arrested RPC looks very similar to the arrested RPC in the C-terminal gate strain.

Other work involving the gate constructs had made extensive use of formaldehyde cross-linking during ChIP-qPCR protocols to identify replication factor origin association dynamics. Under these conditions, the rapamycin-treated gate construct displayed a marked difference from non-treated complexes. Indeed, these experiments (Figure 16, Figure 17, Figure 18 and Figure 31D) show not only that complex composition changes between the two conditions at the time-points investigated by mass-spectrometry, but that the quantitative association of

replication factors was also very different (see Figure 31D, Psf2 IP efficiency at 45 mins post release). Formaldehyde cross-linking has been used for many years as a fixative, and has been used in ChIP-qPCR protocols to stabilise the interaction between DNA-associated proteins and the genomic regions with which they bind to great success (Hoffman et al., 2015). Formaldehyde has a wide range of chemical groups that it can react with, and the same reactions that form the chemical cross-link between DNA bases and protein amino groups are also able to generate short-range protein-protein cross-links (Metz et al., 2004; Chang & Loew, 1994). Despite widespread use, only recently has there been any real information on how the formaldehyde cross-link appears chemically (Tayri-Wilk et al., 2020). This has made the use of formaldehyde in what is traditionally termed cross-linking mass spectrometry (XL-MS)(Yu & Huang, 2018), where information gained from specifically the cross-linked peptide pairs can help determine protein or complex structure, limited (Tayri-Wilk et al., 2020). However, formaldehyde has seen use in contexts that are more appropriate to our use case, where the cross-linking is utilised to stabilise weak or transient protein-protein interactions, allowing downstream mass spectrometry to identify previously overlooked complex members ((Guerrero et al., 2006, 2008; Kaake et al., 2010)(PhD thesis Nath Kedar).

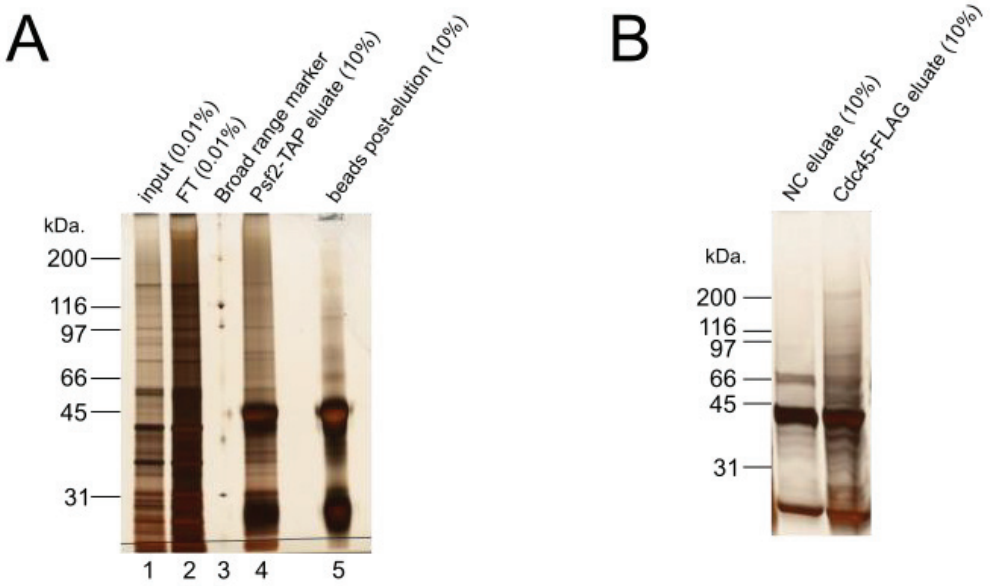
Due to the successful pull-down that is produced by ChIP-qPCR methods, this protocol was adapted so as to enable pull-down and analysis of proteins rather than DNA. Both methods harvested similarly, but the ChIP method used bead beating to lyse cells, prior to sonication. This suited the culture volumes that were being used better than the freezer milling that was previously in use for native pull-downs, and so was adopted for this protocol. In addition, because it was anticipated that formaldehyde cross-linking may produce a higher level of aggregation and “sticky” protein complexes, the wash conditions for regular ChIP was also adopted in the first instance. These in general contained higher levels of non-ionic and ionic detergents, such as NP-40, sodium deoxycholate and Triton X-100, and higher levels of salt,

mainly potassium, sodium and chloride, as opposed to the potassium, magnesium and chloride used previously. Since FA-induced cross-links were to be used to stabilise protein complexes, however, changes in salt concentration may not affect the pull-down as much as might otherwise be expected.

The introduction of formaldehyde was also hypothesised to potentially negatively affect the TAP tag function in the IP. Since the protein A domain of the tag requires an intact three-dimensional tertiary structure to effectively bind the constant regions of IgG on the beads, the action of formaldehyde cross-linking this domain may interfere with this step. Because of this, an alternative tag was considered. At the same time, GINS was reconsidered as an ideal pull-down target. ChIP-qPCR data and other literature (Miyazawa-Onami, Araki & Tanaka, 2017) showed that other replication factors associated at earlier steps, such as Cdc45. Cdc45 also benefited from being a long-term member of the travelling replisome, and Cdc45 tagged with a penta-FLAG tag showed good ChIP-qPCR pull-down efficiency. The FLAG sequence does contain two lysines, which are a target of formaldehyde, but is a relatively small tag, and the antibody recognition, since the flag epitope is relatively short, seems less likely to depend on well folded tertiary structure that formaldehyde might be detrimental to. Another benefit is that the recognition domain of the TAP-IgG interaction, namely the protein A fragment, undergoes cross-linking, whilst in the case of the FLAG tag, the recognition side of the interaction is by the specific anti-FLAG used to coat magnetic beads, and does not get exposed to formaldehyde. To compare the two tags, the protocol was standardised as much as possible between them. Notably, this meant that elution from beads was accomplished in the same manner as for ChIP-qPCR, which utilised an elution buffer containing 1% SDS and heating at 65°C for 25 minutes, for both strains. Analysis of this eluate for the Psf2-TAP pull-down showed that this method did manage to elute a large amount of protein from the beads (Figure 29, panel A). However, in contrast to previous native pull-downs, the resultant eluate was not as easy to analyse by

silver stain. Previously, eluates had several clear and distinct IP-specific bands, with a low background (see Figure 29A for an example). In this case, the lane was more of a smear, and the previously easy to detect indicative bands of Ctf4 and GINS were not distinguishable. Similar results were seen with the Cdc45-FLAG pull-down, but comparison with the untagged control strain demonstrated that despite not forming clear bands, the protein smear observed appeared to be IP-specific (Figure 35). In subsequent MS analysis, both IPs contained high abundance of the respective tagged proteins, indicating a successful IP (Figure 35C).

The two pull-downs, using Psf2-TAP and Cdc45-FLAG respectively, along with the untagged control sample, were submitted for mass spectrometry analysis together. As previously, the lack of certain proteins in the negative control strain posed problems for calculating enrichments, particularly for some replication proteins that have limiting concentrations in the cell (Mantiero et al., 2011), and so are unlikely to be present in a negative control background proteome. In this case, a conservative imputation was performed, where any missing values in the negative control were replaced with an average of the lowest 10 percent of intensities in that negative control, to allow calculation of an enrichment value.



C

Psf2-TAP IP most enriched proteins	
PSF1	11.1914316
CSR2	10.6634564
PSF2	8.32476838
PSF3	8.19339558
SLD5	8.10372638
COQ2	7.6828889
YDC1	7.58627213
POL2	7.24051065
PER33	6.7083852
TOF1	6.11253043
DGA1	5.9013872
TVP23	5.88070014
TVP18	5.85262732
ATP6	5.32418865
PCS60	5.14465484
POM33	4.9570031
SFK1	4.95293582
CTF4	4.94474317
IRC3	4.82831175
GMH1	4.64437922
HBS1	4.50211472
AUR1	4.4766953
RPS29A	4.44228973
SOD2	4.31347305
PER1	4.30210184

Cdc45-FLAG IP most enriched proteins	
CDC45	10.2810209
PSF1	8.61322249
POL2	7.99781969
TOF1	7.62322147
PSF3	6.56769896
PSF2	5.62580878
CTF4	5.42158978
MRC1	5.18555758
SLD5	5.01133138
CTF18	4.9365004
SLD3	4.65066183
DPB2	4.63399088
PRI2	4.43173279
SEN1	4.42911633
HIR3	4.37006492
MCM2	4.0354175
DPB11	3.87449017
MCM10	3.87131287
DCC1	3.87010133
SSF1	3.77602175
MCM7	3.73906339
RXT2	3.69468767
MCM6	3.59521585
RFA2	3.50499307
SLD7	3.49522751

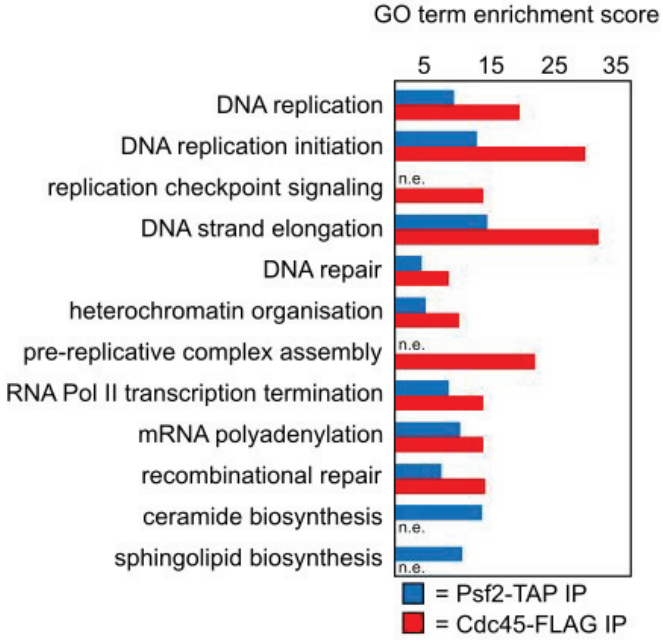


Figure 35: Cross-linking immunoprecipitation of Psf2-TAP and Cdc45-FLAG yields contrasting results. A) Silver stain of Psf2-TAP IP from 1% formaldehyde cross-linked cells. The elution appears more complex than in previous native pull-downs. B) As in A, but for Cdc45-FLAG pull-down and accompanying untagged strain control. The smear observed in Psf2-TAP (panel A, lane 4)) and Cdc45-FLAG eluates seems IP-specific, and is likely related to cross-linking. C) Most highly enriched proteins from mass spectrometry analysis of Psf2-TAP and Cdc45-FLAG are listed. Psf2-TAP IP shows strong GINS enrichment, but other replisome components are less well enriched. Cdc45-FLAG shows strong enrichment of a range of replisome components. The colour scheme follows that used previously (Figure 32) D) Gene ontology analysis (GO) using the Panther dataset, comparing the most enriched proteins (proteins that showed an log_2[enrichment] value of 2 or more) against the entire genome, identifies many enriched biological processes, of which a selection are displayed here. In general, Cdc45-FLAG pulldown is more enriched for GO terms relating to replication and replication-associated processes. Not enriched (n.e.) shows a specific GO term was not significantly enriched in this pull-down.

The proteins with the highest resulting enrichment value are listed for both pull-downs (Figure 35D). The Psf2-TAP pull-down after formaldehyde treatment maintained enrichment of the GINS complex, but other members of the DNA replication machinery were markedly absent from the most enriched proteins. This might be due to a greater proportion of soluble GINS being pulled-down compared to chromatin-bound GINS, which may suffer more extensive cross-linking due to its proximity to DNA, and particularly single-stranded DNA, which are potent targets for formaldehyde cross-linking.

In contrast, the Cdc45-FLAG pull-down showed strong enrichment of Cdc45, GINS, and many other replication proteins. The MCM2-7 complex was distributed down the list, but all members are present at a reasonable level of enrichment. Strikingly, the Cdc45 IP contained Sld3, Sld7 and Dpb11, which were not previously detected and are clear markers of an activating complex. The Psf2-TAP pull-down was also able to detect these proteins under cross-linking conditions, but at markedly lower enrichment levels (Figure 36, panels C and D).

For both pull-downs, gene ontology (GO) term enrichment analysis identified several DNA replication-related biological processes (Figure 35E). These processes extended beyond just DNA replication and DNA replication initiation, to DNA repair and replication checkpoint signalling, though it is worth noting that many of the core replication machinery proteins are also assigned these tags due to their involvement in these processes as part of their role in replication itself. However, for the Cdc45-FLAG pull-down in particular, the enrichment of genuine repair factors such as Sen1, Msh6, and Rdh54, linked to transcriptional-coupled repair, mismatch repair, and recombination respectively does confirm this link (Alzu et al., 2012; Marsischky et al., 1996; Klein, 1997). Enriched gene ontologies and factors will be discussed in greater detail in later sections.

As mentioned, the observed differences could be due to multiple factors. The use of the penta-FLAG tag was predicted to behave better under cross-linking conditions, but why cross-linking

effects seem to reduce the Psf2-TAP pull-down of chromatin-bound GINS is unclear. The Cdc45-FLAG-mediated pull-down of Sld3/7 and Dpb11 under cross-linking conditions was highly exciting, and opened up the potential ability to study the composition of replication complexes in early stages of activation. The appearance of these elusive factors might have been due to the use of the earlier-associating Cdc45, or the use of cross-linking to stabilise transient intermediate complexes, or both, and it was unclear from this experiment which of these effects had the greatest influence.

The variability in enrichment of the MCM2-7 complex subunits under cross-linking conditions where previously they clustered near the top of the enriched proteins list could also be due to multiple effects. Cross-linking may have caused increased pull-down of proteins that are normally removed during wash steps, and these proteins now are able to compete for the highest enrichments where they were not able to before. Equally, the effect on general background proteins induced by the new protocol and formaldehyde cross-linking may have had an effect on the number of highly enriched factors, simply as a result of a reduced signal-noise ratio, potentially due to increased "stickiness" and non-specific binding of cross-linked protein complexes to beads and plasticware. Certainly, the number of factors identified in negative control samples, the "background", was greater than previously seen under native IP conditions (Figure 36B). In this experiment, almost 2,400 proteins were present in even the negative control, whereas in the previous runs, the number of proteins identified in the negative control was more in the region of 300-700. This increased background might also be due to the less-specific elution step of boiling. The native IPs used PreScission protease cleavage of the tag to elute, which should only affect specifically-bound complexes. Regardless, an increased background ultimately proves helpful for reducing the amount of imputation necessary, if the signal to noise ratio can be appropriately managed. Lastly, since formaldehyde cross-linked peptides are not recognised by mass spectrometry correctly, if a certain protein has more tryptic

peptides that are formaldehyde-modified, it will reduce the peptides available for quantification of that protein by MS. This variability may drive the variance in enrichment seen even for members of the same complex as for MCM2-7.

Despite the increased background, the pull-downs did retain the ability to isolate and enrich specific factors (Figure 35C-E). When imputation was carried out in the Perseus software for MS analysis, the negative control could clearly be seen to have more factors needing imputation, without reaching an inappropriate level of imputation (Figure 36A). The tagged samples reported some unique proteins over the negative control, but the reverse was not true (Figure 36B). This indicated that despite the levels of background having been generally higher in this experiment, certain proteins were enriched above the detection threshold in the tagged samples, but weren't in the negative control. In terms of enrichment rather than simply looking at presence or absence, the cross-linking conditions had mixed results. In general, enrichment values were reduced when compared to native pull-downs (compare Figure 32B to Figure 35C, and Figure 33A-C to Figure 36C,D). Despite this, there were clearly replication-associated factors enriched in these pull-downs, and scatter plots show that whilst these proteins did not necessarily make the most highly enriched factor list in the Psf2-TAP pull-down, they were still enriched at some level for the most part (Figure 36C). The Cdc45-FLAG pull-down displayed better enrichment of replication factors (Figure 36D), and aside from Sld2, which in this experiment did not enrich well, most of the replication proteins used as diagnostic markers here were enriched.

Cdc45 forms a stable interaction with the replisome at an earlier point than GINS, making it a better candidate for detection of activation intermediates in the earliest stages of DNA replication. In addition, the Cdc45-FLAG pull-down outperformed the Psf2-TAP pull-down here in terms of specificity of factor enrichment. For these reasons, for the next series of

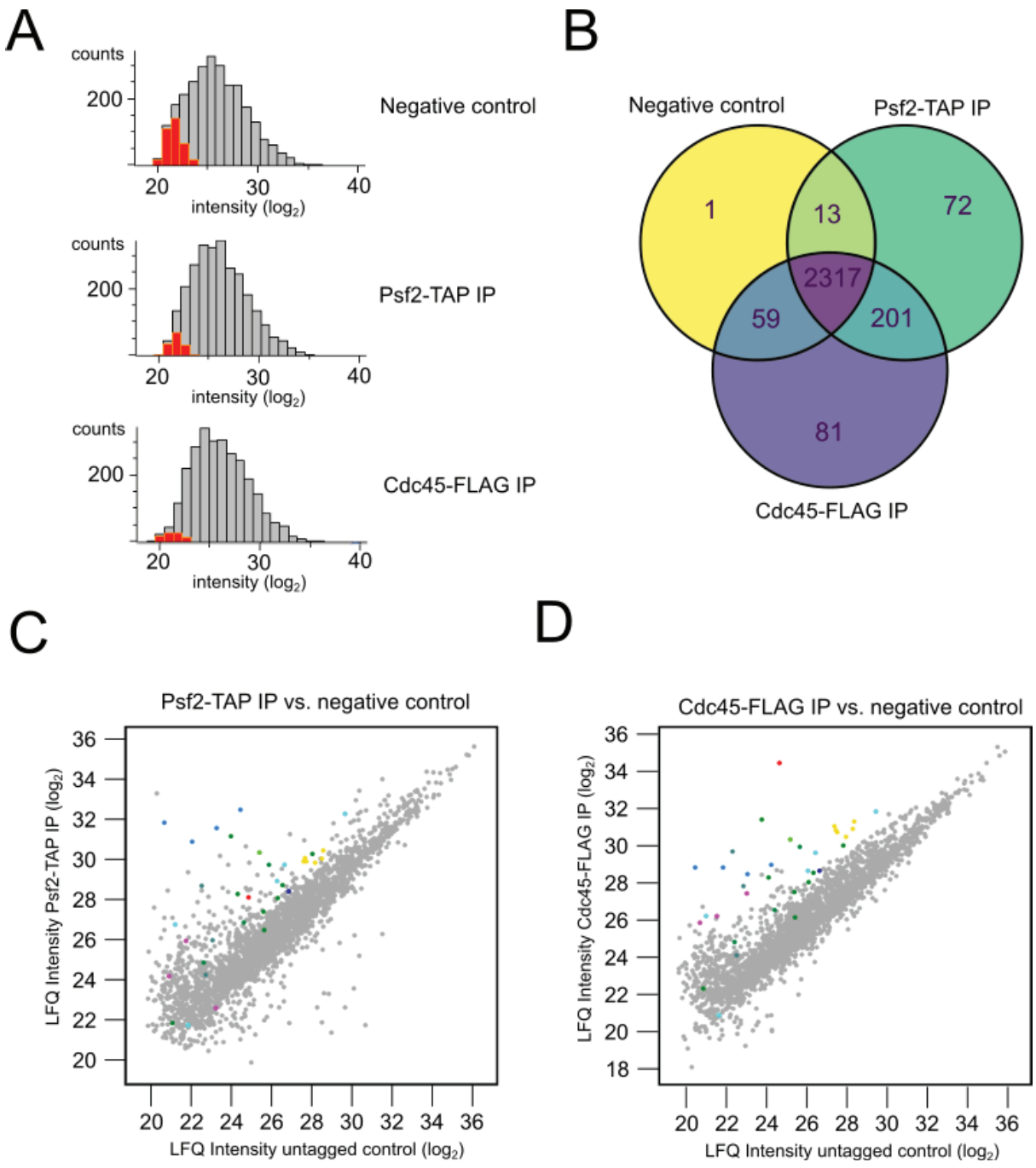


Figure 36: Analysis of all detected proteins in cross-linking pull-downs. A) Distribution of protein intensities plotted as a histogram. Imputation was carried out within the Perseus software and imputed values are displayed in red. These values do not form a large proportion of the detected proteins in any sample, and fit within the expected distribution. The negative control contains the most imputed values as expected. B) Venn diagram showing presence of detected proteins in each sample. Most proteins are detected at some level in all samples, and proteins that are unique to a sample are more prevalent in the tag-specific samples. C) Scatter plot of protein intensities compared between Psf2-TAP pull-down and the negative control. IP-enriched factors (upper left) contain very obviously the GINS complex (dark blue markers), but other replication factors are less clearly enriched. Dark blue = GINS, light blue = Dpb11, Sld2 and RPA, light green = Ctf4, dark green = fork protection complex and polymerases, red = Cdc45, pink = Sld3/7 and Mcm10. D) Scatter plot as in C, but comparing Cdc45-FLAG to the negative control. Here, a more reliable enrichment of a range of replication factors is observed, consistent with the most enriched factors listed previously (Figure 35D).

experiments, Cdc45-FLAG was used as the protein and tag of choice to investigate the composition of the replisome during and after activation.

3.4 The complex composition and interactome of active CMG *(Cdc45 ChIP-MS)*

To gain a reliable and comprehensive view of the proteins that form associations with the Cdc45-bound replisome during replication, the pull-down of the tagged Cdc45 from a formaldehyde cross-linked context was repeated three times, and the eluates would be processed in parallel for mass spectrometry analysis. To get the best comparative analysis, it is important to run individual samples as a group on the MS instrument, as the detection and quantitation can vary between runs, enough that the label-free quantification (LFQ) steps used are no longer reliable to normalise samples to each other.

In the previous trial of cross-linking conditions, the non-specific background was considered to be too high. To try and reduce the carryover of non-specific proteins, the sample tubes used during wash steps were changed to Eppendorf LoBind® tubes (catalog number: 0030108434). This appeared to have little effect on the protein bands visible in silver stain, but the use of these tubes was continued nonetheless as they are recommended for proteomics applications.

For the triplicate, 1 L cultures were released into S phase from α-factor arrest for 40 minutes at 25°C. After formaldehyde cross-linking and harvesting as for the previous ChIP-MS trial, they were lysed, sonicated, and anti-FLAG pull-down of the cell extract carried out. By superficial silver stain analysis, all three processed replicates looked similar in their overall protein content, and in the IP-specific smear observable in each IP sample (Figure 37A). Earlier work optimising levels of the anti-FLAG antibody and magnetic beads had also provided western blot examples of the presence of Cdc45-FLAG within IP eluates (Figure 37B). The comparison of molecular weight and band heights in the western blot led to the tentative assignment of a Cdc45-FLAG band in the silver stain, at around the 97 kDa marker band

(Figure 37A), and this band was consistently present in all three replicates in an IP-specific manner. In addition to the tagged pull-downs and negative controls, a small amount of input for each tagged sample was also submitted. Mass spectrometry analysis on all nine samples gave a large number of identified proteins. An average of 2954±184 were identified in the tagged samples, and an average of 1445±348 were identified in the untagged negative controls. The inputs contained on average 3510±5 proteins. Principal component analysis of the proteins identified showed that samples grouped mainly by type, with a single component that accounted for most of the variability cleanly separating the tagged and untagged samples (Figure 37C). Reflecting the variance seen from the number of identified proteins, the negative control samples displayed more variance between them than the tagged pull-downs did, though the first replicate IP did differ more from the very similar second and third IP replicates. However, given the separation between IP and negative controls, and crucially the fact that samples did not group within their replicates but within their sample type, this data was appropriate for further comparative analysis. Some of this analysis was carried out in the Perseus software package as used previously.

By principal component analysis, the samples looked acceptable, but how much did the calculated enrichment values of individual proteins vary between replicates? The differences between replicates for the enrichment value (i.e. tagged pull-down intensity over negative control intensity) of each protein was calculated, and the median difference value between the replicates was plotted (Figure 37D). The values cluster tightly about zero, indicating that the replicates were not providing drastically different enrichments, further substantiating their reproducibility.

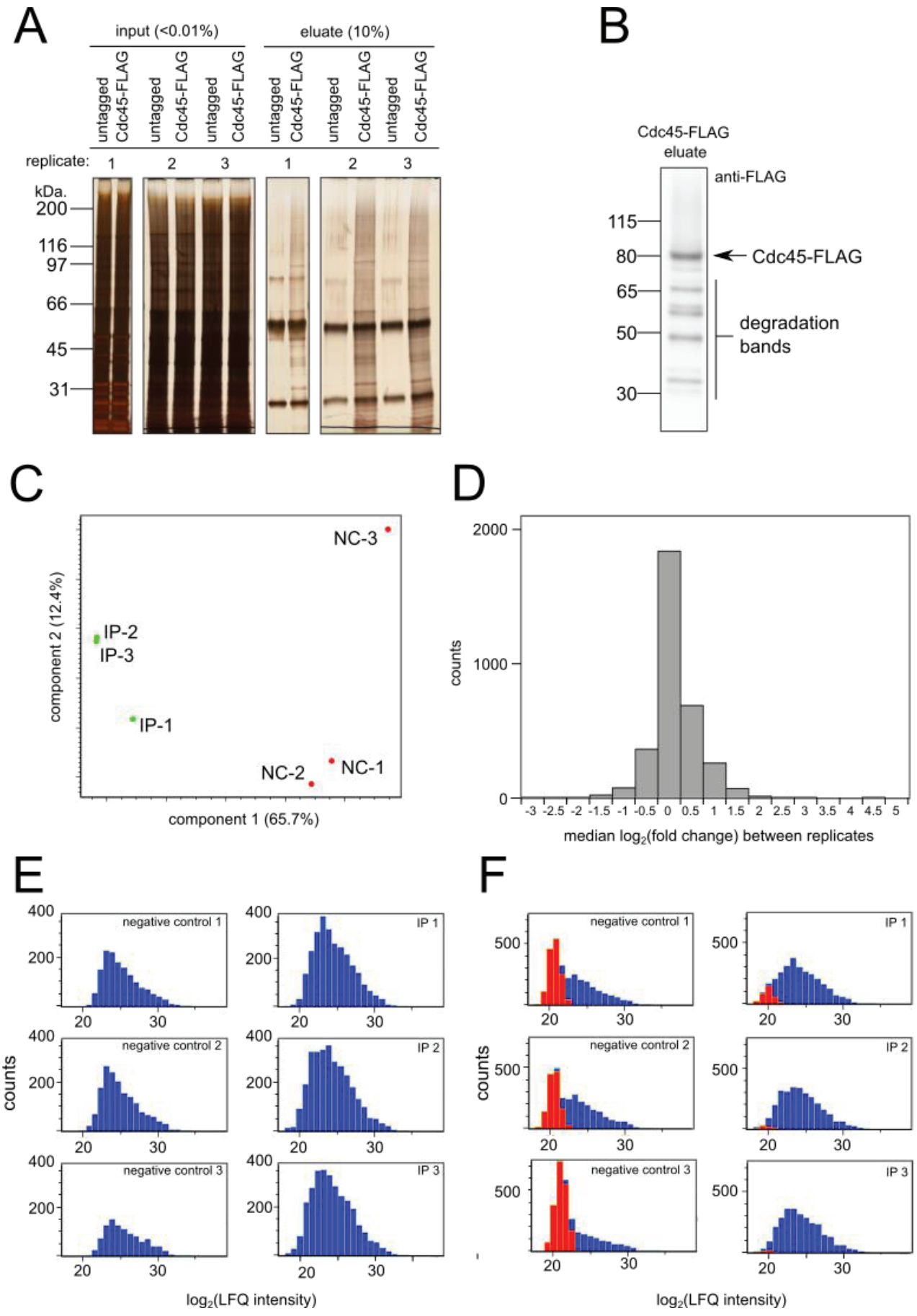

A
input (<0.01%)
eluate (10%)
untagged
Cdc45-FLAG
replicate:
1
2
3
kDa.
200
116
97
66
45
31
B
Cdc45-FLAG
eluate
anti-FLAG
115
80
65
50
30
Cdc45-FLAG
degradation
bands
C
NC-3
IP-2
IP-3
IP-1
NC-2
NC-1
component 2 (12.4%)
component 1 (65.7%)
D
2000
1000
0
counts
-3 -2.5 -2 -1.5 -1 -0.5 0 0.5 1 1.5 2 2.5 3 3.5 4 4.5 5
median log2(fold change) between replicates
E
negative control 1
negative control 2
negative control 3
IP 1
IP 2
IP 3
400
200
20
30
counts
log2(LFQ intensity)
F
500
counts
log2(LFQ intensity)

Figure 37: Quality control for Cdc45-FLAG ChIP-MS triplicate. A) Silver stains of the three replicate pull-downs carried out for this experiment. Note that replicate 1 was run on a separate gel than replicates 2 and 3, indicated by panel separation. As for previous cross-linked IPs, the tag-specific proteins are present as a smear in the lane as opposed to clear and distinct bands. B) Western blot testing for the presence of Cdc45-FLAG in eluates. A companion experiment tested the effect of bead concentration on pull-down yield. This blot is the sample containing comparable bead concentration to the main IP. Cdc45-FLAG is clearly detected by anti-FLAG, as well as several degradation bands. C) Principal component analysis carried out using the Perseus software package shows clear distinction between tagged samples (IP) and untagged controls (NC). In the largest component, x, the pull-downs are distinct from the negative controls. A secondary component segregates the controls and IPs further, showing some variation between replicates. D) Histogram showing the median difference in protein enrichment between replicates. The tight clustering about zero suggests for most proteins the replicates display good reproducibility and only small amounts of variance, enabling comparative analysis. E) Histograms displaying range of protein intensities in each individual sample. Controls have overall lower levels of proteins present. F) As in E, but after the use of imputation in Perseus. Pull-down samples are relatively unaffected, but control samples have a large number of missing values (IP-specific proteins) that do perturb the distribution.

In contrast to previous cross-linking experiments, the number of proteins in the negative control was substantially less than that for the tagged pull-downs. This would mean that there would be many missing values in the controls when enrichment values were computed. The use of imputation here can mitigate this problem. Manual imputation was carried out using the average intensity of the protein with the lowest intensity that was detected in all three negative control samples, whilst imputation was also carried out using the inbuilt function of Perseus to assign a random scatter of low intensity values to missing values. The effect of this can be seen in Figure 37, panels E and F. For the controls, with large numbers of missing values, the imputation is disruptive to the distribution, introducing lots of data. However, the imputation did not cause the introduction of many values in the tagged pull-downs, and, since a clean negative control is likely to have many proteins falling below the detection threshold that will be abundant in the tagged IP, this was considered an acceptable process for this data.

Calculating enrichment values for all proteins, and averaging across all three replicates generates similar results to the trial experiment performed previously (Figure 38A, compare to Figure 35D). Cdc45 is the most highly enriched protein, as is expected, and all six of the MCM2-7 proteins are among the most highly enriched, suggesting a reasonable amount of MCM-bound Cdc45 is purified, representing a replicative complex on chromatin. Other proteins immediately recognisable as fork-associated factors are also present, such as Rfa1 and 2, Ctf4, Tof1, and several polymerases required for DNA replication. Interestingly, other

proteins not usually associated with a non-perturbed fork are present, such as Sen1 (Alzu et al., 2012), and Fun30, which after this data was collected was shown to associate with replication forks in humans (Sirbu et al., 2013). In yeast, however, Fun30 has yet to be directly linked to the replication machinery, and its high enrichment value in this experiment suggests that it may show strong association with replication forks.

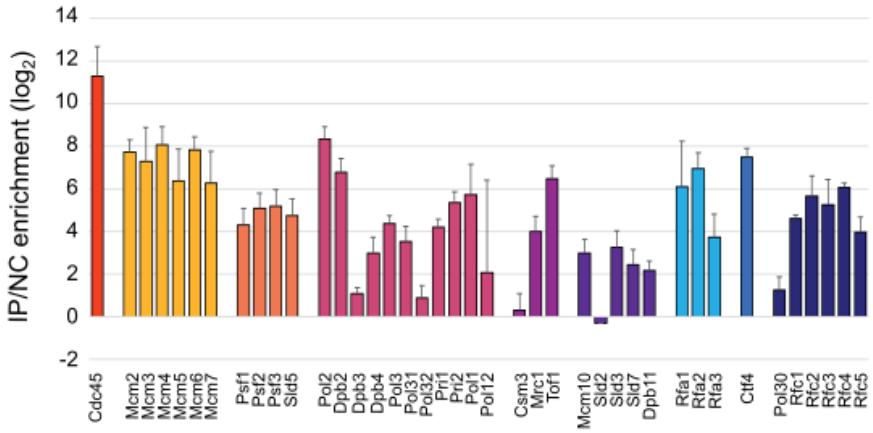
The presence of Leu3 and Ssf1, however, cannot be so easily explained. Whilst both are chromatin-associating (Friden & Schimmel, 1988; Fatica et al., 2002), they do not appear to have any clear role to play in replication in the literature or their predicted activities. Whether these proteins associate with the replisome itself, whether they are interactors of soluble Cdc45, or whether they are false positives, is difficult to say, and is a limitation of this approach in identifying novel roles of factors. Further validation should be carried out to rule out the possibility of a false positive.

Summarising the recruitment of a range of replisome-associating proteins and subcomplexes, similar to presented previously for nascent chromatin capture (NCC) (Alabert et al., 2014), shows in general a strong enrichment of a core replisome complex, with most of the major interacting subcomplexes present in a reliable manner (Figure 38B). Some proteins show lower enrichments, and for a some of these, this low enrichment is reproducible, for example Csm3 or Dpb3. Whether this is because of fewer direct contacts with Cdc45 and the core replisome or less reliable inter-subcomplex cross-links to preserve their association, this could be due to the formaldehyde-coupled technique affecting specific proteins more than others.

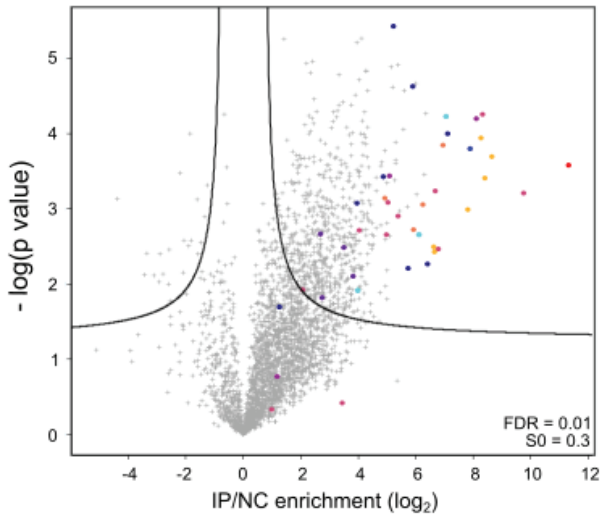
A

CDC45	11.29243
POL2	8.33536
MCM4	8.057909
LEU3	7.952485
MCM6	7.834634
MCM2	7.740507
CTF4	7.47554
MCM3	7.279753
RFA2	6.942093
DPB2	6.765392
TOF1	6.453967
MCM5	6.370085
MCM7	6.279855
RFA1	6.098937
RFC4	6.055136
POL1	5.716838
RFC2	5.655553
SEN1	5.576508
PRI2	5.357951
SSF2	5.3254
FUN30	5.280605
RFC3	5.234767

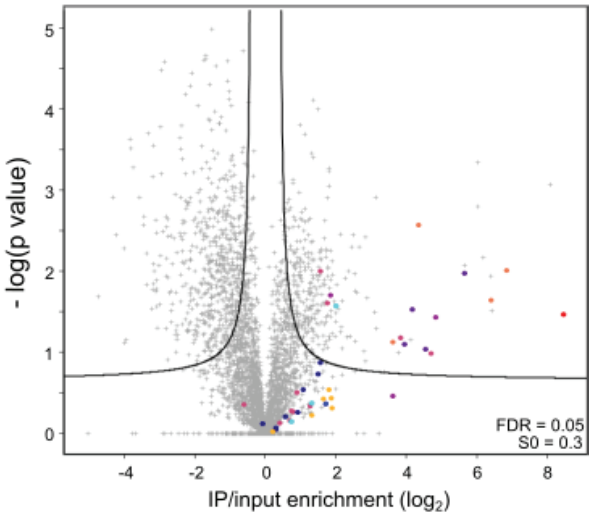
B



C



D



E

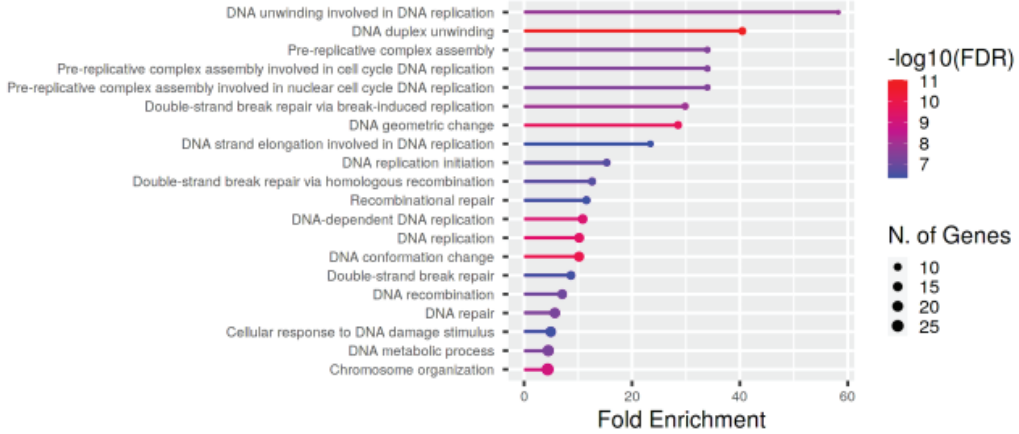


Figure 38: Enriched proteins in the Cdc45-FLAG pull-down. A) The most enriched proteins across all three replicates, with average enrichment values. Many members of the DNA replication machinery are again present. Proteins without recorded intensity in the negative control relied on an imputed value to determine an enrichment score. B) Mean enrichment scores of a set of replication proteins across all three replicates. Error bars represent the standard deviation of this value. Members of several replication-associated complexes are consistently enriched, but some members' enrichments are generally lower than others. C) Volcano plot comparing enrichment between pull-down and negative control samples. The Y axis represents reproducibility of a proteins enrichment across replicates. The false discovery rate (FDR) for the threshold line is 0.01, and the S0 value is 0.3. Imputation carried out within the Perseus software. Colour identification matches panel B. D) as in C, but for pull-down versus input samples. Imputation was not carried out on this dataset, since both sample types contained proteins that the other did not. E) A students' t-test was carried out in Perseus to identify significantly enriched proteins from the analysis presented in C, and these proteins were compared against the reference genome to identify gene ontology (GO) biological functions enriched in this subset of significant proteins (Ge et al., 2020). The twenty most enriched terms are represented here. FDR for GO enrichment analysis is 0.05.

By plotting the enrichment values of proteins in the pull-down vs. negative control against a p-value obtained from a two-sample t-test, it was possible to generate a volcano plot which can be very informative when interpreting mass spectrometry data. For this dataset, it could be seen that many of the targeted replisome-associated factors shown in Figure 38B were enriched in a reliable manner (Figure 38C, Figure 39A). It also showed that quite a large number of other proteins are similarly enriched (Figure 38C). This was interesting because non cross-linked samples processed via pull-down for mass spectrometry often show only a few, highly enriched factors, which are clearly distinguishable on such a plot that shows enrichment values against some other metric (Figure 33D). In this case, whilst the most enriched factors were clearly separable, there is a continuous "cloud" of factors across a wide range of enrichment values. However, it is plausible that this effect is directly due to the action of formaldehyde in this experiment. By forming many cross-links between proteins that either transiently interact, or share a close cellular proximity, formaldehyde could enable specific pull-down of a wide range of factors, even beyond direct interactors. Cross-linked factors could even "bridge" complexes to the replisome when in close proximity.

Because of concerns about the need for imputation in negative controls to calculate enrichment values, a comparison was also made between pull-down samples and their respective inputs (Figure 38D). In this instance, many more factors were enriched in the input sample, which is unsurprising, since it contains an entire cell extract. In addition, whilst normalisation of total

protein concentration prior to mass spectrometry processing was a key step, input samples may have been at an overall higher protein level at the point of MS analysis. They did not contain any antibody, which was something that undoubtedly affected the protein concentration readout for the eluates. Furthermore, the label-free quantification (LFQ) method requires an unchanging "background" to have maximum reliability, and whilst this is provided when comparing two samples that have both undergone pull-down, it was not comparable between an IP and the whole cell extract input. As a result, the observations from comparing pull-down and input were less clear than from when using negative controls. Replication proteins still display a level of enrichment in this comparison, but it is less strong and much less reliable.

A Student's t test was performed to compare significantly different proteins in pull-downs and negative controls. Proteins significantly enriched in the pull-down by this metric were analysed against a reference genome to assess significantly enriched gene ontology (GO) terms. The top twenty enriched biological processes are displayed in Figure 38E. The identities of these GO terms confirmed previous preliminary results carried out using only a single replicate (Figure 33E). DNA replication was enriched strongly, and, similarly to before, DNA repair processes were also supposedly enriched within this enriched subset of proteins. Again, it is worth considering that many of the normal components of the replication machinery play roles in repair similar to their normal function, and so are assigned this term without necessarily being an indicator of repair processes by themselves. However, there are many genuine repair-specific factors that are enriched in this pull-down (Figure 39B/C).

From looking at the distribution of the replication factors in the volcano plot more closely, it is possible to draw some preliminary conclusions and hypothesise about the nature of complexes pulled-down in this experiment (Figure 39A). The consistent high enrichment of the MCM2-7 proteins suggests strong association and cross-linking that most likely reflects the close interaction in the core CMG. The slightly lower enrichments of GINS are more puzzling,

since they lie close to Cdc45 in the CMG, and so should associate and cross-link very well. Perhaps the available residues for cross-linking are few at the GINS-Cdc45 interface, or, perhaps quite a large proportion of Cdc45-MCM2-7 complexes do not contain GINS, which would also be surprising, given at the time point cells were harvested, most origins of replication are expected to have fired, and so any initiating complexes should have been converted to active replisomes or been removed by the passage of forks through origins. Notably, however, Sld3/7, Dpb11 and Mcm10, do show an enrichment in this pull-down, despite all but the latter being almost exclusively associated with initiating complexes, though Dpb11 also associates with sites of DNA damage (Wang & Elledge, 2002; Puddu et al., 2008). This would suggest that, in fact, some initiating complexes remain intact at this time point, and this pull-down is able to isolate a not-insignificant number of them. They most likely represent a minor sub-population in the pull-down, reflected in the lower enrichment scores for these activation factors compared to MCM2-7 or GINS.

Whilst identification of subcomplexes containing known replication factors was of interest for future work, a main purpose of this pull-down was to identify factors that associate with the replication complex that had not been previously known to do so, or were thought to only associate upon specific events or signals. Many of the factors enriched in this pull-down are known to be associated with replication-linked activities, and some of these only loosely. Figure 39B displays some of the factors that may not have been expected to show such strong enrichment with a replication complex.

Given that the pull-down enriches several proteins linked to the DNA damage response and DNA repair, a similar analysis was performed to that for nascent chromatin capture (NCC)(Alabert et al., 2014). By identifying the yeast homologs of the genome maintenance factors used to show a replication-dependent and replication-independent bias in their study, it was shown that a similar bias exists in this data (Figure 39C). Replication-dependent damage

and checkpoint factors are more often enriched in the Cdc45 pull-down, and replication-independent factors are more often not-enriched. This does suggest that despite the promiscuous activity of formaldehyde, enriched factors do generally show association with the process of replication.

Intriguingly, reliable enrichments of certain proteins were also detected in the negative control samples over the pull-down samples (Figure 39D). This runs contrary to expectations, as it is initially counter-intuitive to understand why a protein would associate more in the untagged control than in the tagged pull-down.

Nonetheless, factors are reliably enriched in the negative control. One hypothesis is that in the absence of a genuine FLAG sequence to bind, the anti-FLAG antibody is able to bind more broadly, and so some FLAG-mimicking peptides may be able to bind in a semi-specific manner to the beads. Another is that without specific complex coating the surface of the beads, non-specific binding is increased, and if certain proteins are particularly prevalent in a non-specific background then these may become over-represented. Upon closer examination, these enriched proteins contain several members of the anaphase promoting complex (APC), and also actin. Actin filaments may fit the description of a protein that is especially able to bind surfaces non-specifically. The rationale as to why the APC complex is specifically de-enriched in the tagged pull-down, however, is not clear. It may be that an APC subunit has a “cryptic” FLAG sequence, or that this complex binds surfaces well. Otherwise, perhaps APC’s de-enrichment could be due to a factor in the specific IP that causes its degradation, as the extract is from a pre-M-phase state.

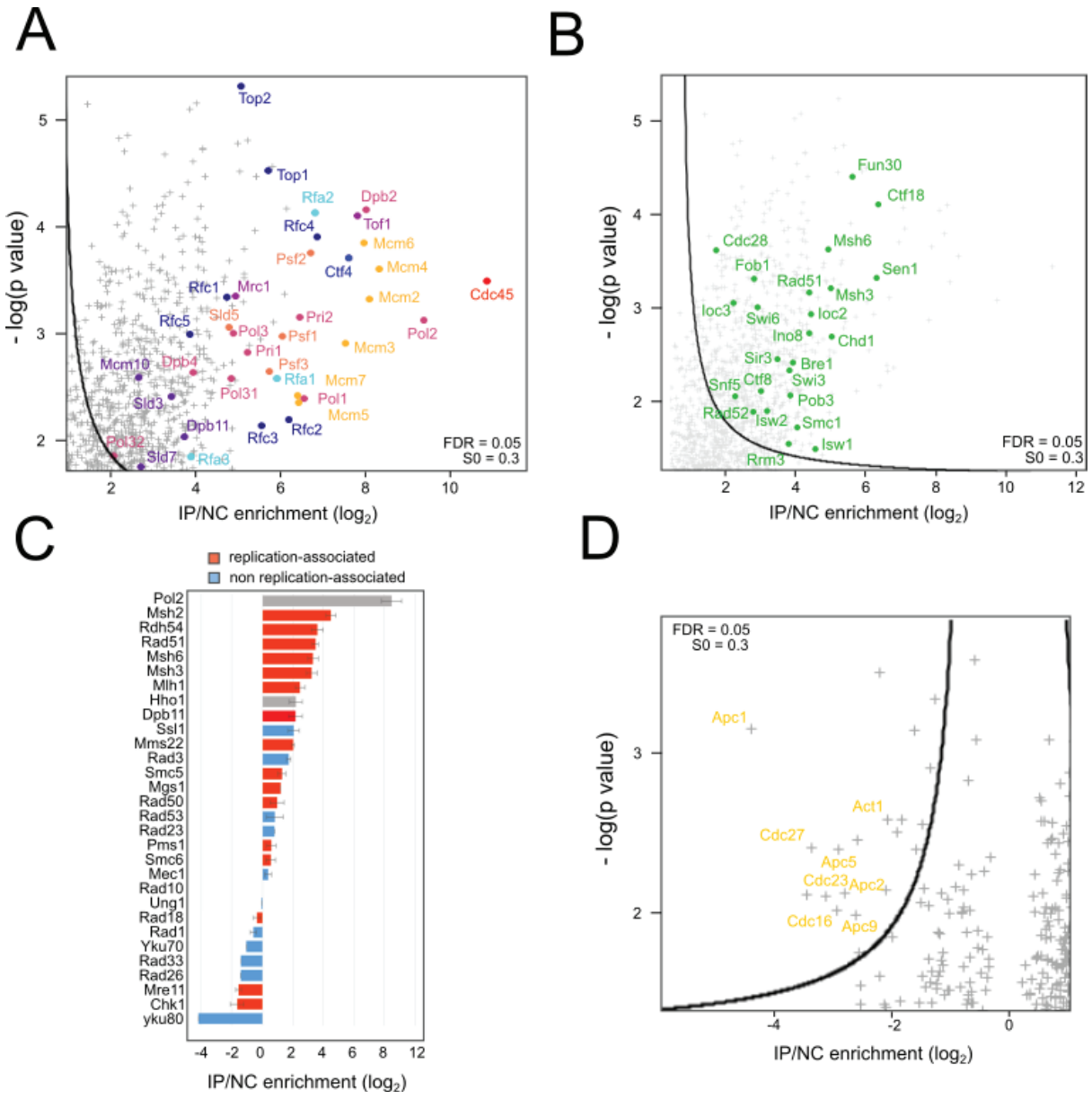


Figure 39: Visualisation of enriched proteins in Cdc45-FLAG pull-down. A) extension of Figure 38C, showing in greater detail the enriched protein factors that are core fork and initiation components. B) Additionally enriched protein factors that are either suggested to associate with replication forks from literature (e.g. Ctf18, Rrm3), form part of complexes involved in DNA damage, chromatin remodelling or other chromatin-related processes (e.g. Msh3/6, Sen1, ISW, Sir3), or possess activities that make them interesting hits with regard to potential new roles in the context of replication (e.g. Fun30). C) Enrichment values of a selection of DNA damage-associated factors in the IPs. Pol2 and Hho1 are control protein representing replisome and chromatin IP respectively. DNA damage factors are assigned to be "replication-associated" or "non-replication associated" by their homology to human factors as presented in Alabert et al., 2014, and S.cerevisiae literature. D) Additional view of the volcano plot of Figure 38C, showing surprisingly reproducibly de-enriched proteins.

The ChIP-MS of Cdc45-FLAG presented here can be considered to have been highly successful. The pull-down identified strongly enriched proteins that are known to be part of the replication machinery, and displayed good reproducibility for these factors. The pull-down also

enriched for the transient initiation factors indicative of activating complexes which this project aimed to study. These complexes appeared to be a sub-population within the complexes enriched in this experiment, reflective of a heterogenous population of Cdc45 complexes, the majority of which seemed to be associated with MCM2-7 and likely reflect active CMG complexes and associated proteins. This is fairly consistent with the cell cycle stage that the harvested cells had reached by the time of cross-linking. In addition to core factors of the CMG, this IP displayed enrichment of replication-associated complexes, such as the anticipated polymerases ε, δ and α, the FACT complex, the Pol30 DNA clamp and its clamp loader, the hub protein Ctf4, and the ssDNA binding protein RPA. It also identified many repair-associated proteins and chromatin remodellers aside from FACT, which may not have been expected to be recruited so well.

By narrowing down the heterogeneity of the replication complex present in the whole cell extract, or by making the pull-down specific for particular sub-populations of replication complexes, this approach could generate a great deal of qualitative and quantitative data on the proteins associated with various stages of replication *in vivo*.

3.5 Complex composition and interactors during stages of replication initiation *(Sld2/3/Dpb11 IPs)*

ChIP-MS had been shown to be able to isolate a heterogenous population of replication complexes by tagging of Cdc45. The presence of Sld3/7 and Dpb11 indicated that some of these complexes were isolated during an early stage of replication initiation, and this state was potentially preserved by the formaldehyde cross-linking present. The enrichment of transient replication initiation factors was significant, but not as large as the enrichments observed for some other replication proteins that form part of the stable CMG during active replication. To study the activating complexes and the interactors present at that stage, it would be ideal to only isolate those complexes that were undergoing activation.

Since Sld2, Sld3 and Dpb11 show transient binding to the MCM2-7 helicase complex that is restricted to the activation phase, these proteins could act as a selective mechanism for targeting the replisome at this stage. To test whether the Cdc45-FLAG pull-down could be easily modified to provide reproducible interactomes that show the proteins that associate with the replisome during activation *in vivo*, tagged strains containing the penta-FLAG tag fused to each of these three proteins were tested for their ability to pull-down specific complexes. To try and increase the proportion of origins undergoing activation, an earlier time point was used for these pull-downs, otherwise an identical approach was used as for Cdc45 pull downs.

These experiments showed a clear pull-down of the targeted factors, along with soluble binding partners, Sld3 associated with Sld7 and Cdc45 for example, and Sld2 and Dpb11 showed a pull-down of subunits of polymerase ε.

The tagged factors in this pull-down are thought to only associate with the replication machinery during the process of activation, except for Dpb11, which has some involvement with DNA damage resolution. This brief association was used to attempt to isolate a sub-population of activating complexes. Any MCM2-7 subunits detected in these experiments, for

example, was therefore likely to be from activating replication complexes. However, this necessarily meant that the proportion of helicase complexes that were in a state amenable to pull-down would be minimal. In the resulting pull-down experiments targeting Sld2, SLd3 or Dpb11 showed much less of a pull-down of the larger replication complex as a whole.

Given the transient association of these factors with the complex during initiation, this is likely due to an over-representation of soluble complexes of Sld2, Sld3 and Dpb11 in their respective pull-downs.

Interestingly, it is thought that these initiation factors are in limiting concentration in the cell, in order to properly control origin replication timing (Mantiero et al., 2011; Lynch et al., 2019). Sld3 especially is thought to be a key regulator of stochastic origin firing. If this is indeed the case, it might be expected that during activation, when Sld3 is in greatest “demand”, most of the Sld3 in the cell would be associated with replication origins and bound to activating replication complexes. The data from the attempted Sld3 IP would appear to contradict that expectation, suggesting that despite the limiting concentration of these firing factors, they do exist in a soluble form at a reasonable level even during the transition into S phase.

Without reducing the proportion of soluble complexes pulled-down, or increasing the proportion of pulled-down complexes that were bound to the larger replication complexes, these pull-downs were unlikely to give the same level of success as pull-downs of more consistent members of the replication complex.

3.6 Changing complex composition and interactome of MCM2-7 in the cell cycle (*MCM4 IP timepoints)*

3.6.1 Initial validation of Mcm4-FLAG ChIP-MS

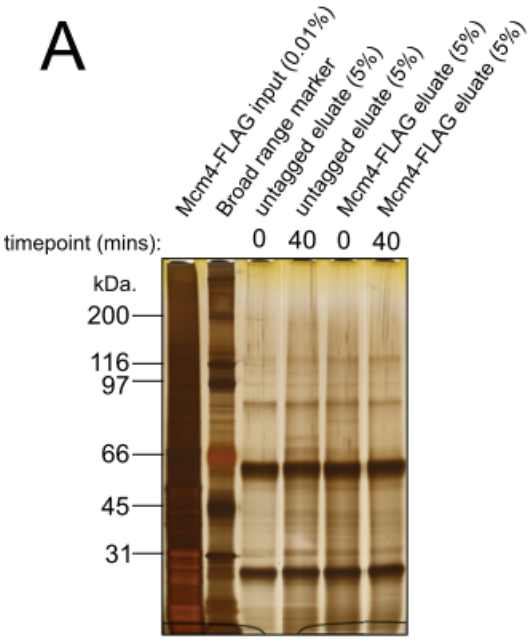
Tagging GINS limited complex pull-down to replisomes that were almost fully activated. Using Cdc45 as the tagged bait protein improved the range of complexes that could be isolated, extending the pull-down to complexes during the early stages of activation. Since the primary component of the replicative complexes was the helicase itself, and this complex is loaded onto DNA during G1, specifically targeting the MCM2-7 proteins would enable isolation of the helicase complex at any and all stages of replication. MCM proteins had already been shown to produce efficient pull-downs in the ChIP-qPCR experiments in Section X. Nonetheless, the efficiency of pull-down achievable using ChIP-MS was tested to see how it compared to the other factors used previously for ChIP-MS.

To test the ability of Mcm4-FLAG to pull-down helicase complexes at time-points prior to activation, a test IP was carried out both on S phase cells, and on cells that were arrested by alpha factor treatment. The S phase cells were released from alpha factor-arrest for 40 minutes before harvesting, in an attempt to isolate maximum amounts of active replisomes. In addition to the tagged cells, untagged negative controls at both time points were also processed, and submitted for mass spectrometry analysis (Figure 40). As common with cross-linking treated samples, the results of silver staining the pull-down eluates was inconclusive, and the presence of IP-specific bands was difficult to confirm (Figure 40A). Only one possibly IP-specific band was observed at around 80 kDa in the alpha factor-arrested pull-down, but despite this, the samples were submitted to MS. Mass spectrometry analysis was an efficient way to evaluate the entirety of the pull-down composition, both to check if the Mcm4-FLAG IP was able to

efficiently pull-down the MCM2-7 helicase, and to investigate the associated proteins at the time points chosen.

Both samples were shown to efficiently pull-down the MCM2-7 complex, and the Mcm subunits were among the most enriched proteins in each (Figure 40B,C). The sample that was prepared from alpha factor-arrested cells displayed good MCM2-7 enrichment, but other replication-associated proteins were conspicuously absent from the most enriched proteins. Only one of the "core replisome" labelled proteins used in this study was enriched enough to make the list, and surprisingly this was a member of the replication factor C (RFC) complex. This result was not surprising, as in alpha factor-mediated arrest, the MCM2-7 complex was predicted to be present as a stable double-hexamer, with activation processes yet to begin, requiring kinase activity. Many of the replisome factors studied here would not be expected to associate until later, once the helicase was active and replication forks established. In this sense, the strong enrichment of RFC3 was therefore surprising. However, care was taken with this result, since this was only a single test replicate, and statistical methods are required to ensure the rate of false positives in mass spectrometry data is kept low. Present in the most enriched list was Tah11/Cdt1, the protein required for the process of MCM2-7 loading at origins and component of pre-RC intermediates formed during this process. Since cells had been maintained in alpha factor for 2 hours prior to harvesting, it was considered unlikely that this association was due to the presence of helicases still undergoing loading. Instead, this could represent weak residual binding of Cdt1 to the double-hexamer, or association with MCM2-7 hexamers in solution, that had not been loaded, and would still associate with Cdt1. It is important to note that the tagging and isolation method used here does not preclude the isolation of such chromatin-independent MCM2-7 complexes.

A



B

MCM4	11.3296195
MCM6	10.2037141
MCM7	8.99092534
MCM3	8.51799911
GCN5	8.08509055
MCM2	7.71985706
MRPL35	7.00822391
MCM5	6.32991878
TAH11	6.32787971
RFC3	5.90830899
CKA1	5.4885773
STO1	5.17241385
LYS20	5.06720317
RPC82	4.81305112
DIS3	4.71174456
GSH1	4.68824441
BRE1	4.68799278
RRP12	4.6637601
PUS1	4.64917275
TRM8	4.6148143
GAR1	4.57744347
NRD1	4.55966983
PUS7	4.44101075
STH1	4.42973815

C

MCM4	13.7362549
POL2	11.6417851
MCM7	11.211151
MCM6	11.0188147
MCM2	10.5415886
MCM5	10.4642244
GCN5	10.0412679
MCM3	9.92647215
TOF1	9.56742144
DPB2	9.46644084
CTF4	9.06849118
CDC45	8.48821375
PSF3	8.33861369
CTF18	7.33934779
PSF2	7.28730562
TAH11	7.15259817
SLD5	7.09687112
PSF1	7.01264137
MRC1	6.27434866
TOP1	6.16821349
RFC4	5.76273867
RFA1	5.67837623
RFA3	5.63564063
POB3	5.53024661

D

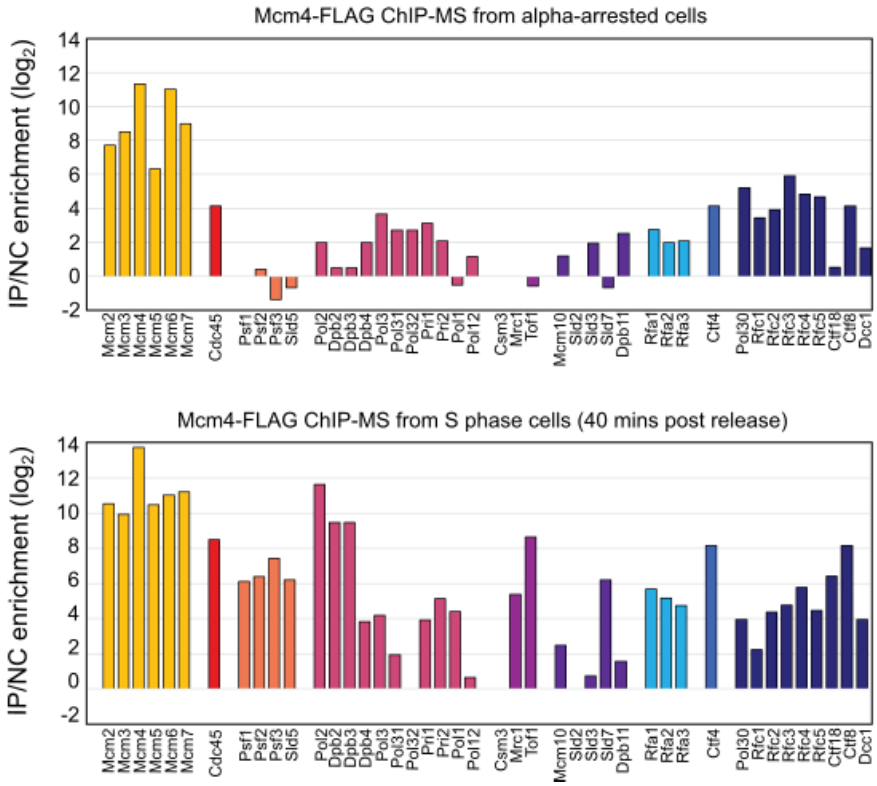


Figure 40: Trial ChIP-MS of Mcm4-FLAG from cells at two different stages of the cell cycle. Cells were arrested in alpha factor, and either harvested in alpha arrest or released into S phase for 40 minutes before harvesting. A) Silver stain of pull-down eluates sent for mass spectrometry analysis. As for most cross-linked samples, IP-specific bands are difficult to distinguish. B) The most enriched proteins in the alpha factor arrest pull-down. For missing values in the negative control, conservative manual imputation was carried out as previously, using the lowest 10% of NC proteins average intensity. The successful pull-down of MCM2-7 helicase is confirmed. C) As in B, but for the S phase time point. The enrichment of both MCM2-7 and other replisome associated factors can clearly be seen in this sample. D) Comparison of enrichment values observed for each timepoint for known replisome components. The 40 minute sample shows greater enrichment of replisome-associated factors, as expected. The RFC complex and Pol30 display a surprising level of enrichment in alpha factor arrest

By contrast, the S phase sample showed not only high efficiency MCM2-7 pull-down, but also very high efficiency pull-down of replication-associated factors (Figure 40C). Whilst Cdc45 and GINS were not as enriched as when either was tagged as the bait, they were very abundant in the pull-down sample. Polymerases, in particular polymerase ε, were also very well enriched in the S phase sample, and fork protection complex members Tof1 and Mrc1 were also present in large abundance, likely due to the close association of both the Pol2 CTD and the fork protection complex to the hexamer itself (Yuan et al., 2020; Baretić et al., 2020). Tah11/Cdt1 was still observed in this sample, lending credence to the hypothesis that non-chromatin bound MCM2-7 was also being pulled-down along with the chromatin-loaded MCM2-7.

By organising the data to specifically show the enrichments of known replication factors, two observations became apparent. Firstly, the Mcm4-FLAG IP was able to isolate a complete replisome with similar effectiveness to the Cdc45-FLAG pull-down shown above (compare Figure 40D and Figure 38B). Indeed, most components of the "core replisome" considered were present, and the same proteins that showed variable or lower enrichment in the Cdc45-FLAG IP also tended to show reduced enrichment in the Mcm4-FLAG IP. This indicated that these proteins were not being pulled-down well by this method, regardless of the bait protein used. This either meant that their association with the CMG was weaker, that they were particularly negatively affected by formaldehyde cross-linking, or generally more difficult to detect by mass spectrometry. The exact reason may vary depending on the protein, but at least for the polymerase holoenzymes and the fork protection complex, it was generally the smaller subunits that showed the lowest enrichment and therefore lowest abundance in the pull-down. This means they would generate fewer tryptic peptides for MS identification and quantification and so this could affect their detection efficiency. With the addition of formaldehyde cross-linking, the fewer tryptic peptides these smaller proteins could produce may also be negatively affected by modification that further reduces detection ability.

Secondly, whilst replisome proteins were not among the most enriched proteins in the alpha factor sample, the enrichments of certain factors and subcomplexes were higher than might have been expected (Figure 40E). Cdc45 was fourfold enriched in the pull-down sample compared to the negative control, and Ctf4 was similarly enriched. In addition, subunits of polymerase α and δ were slightly enriched, and the sliding clamp and its clamp loader complex, Pol30 and RFC, showed surprisingly high enrichment.

The mass spectrometry data was further analysed using the Perseus software, to gather a holistic understanding of the complexes isolated (Figure 41). By comparing the number of proteins identified in each sample between the tagged pull-down and untagged control, it was possible to observe how complex each sample was, and how many proteins were exclusively detected in each (Figure 41A). The alpha arrested samples showed what might be more expected of a pull-down. The overall number of proteins detected in the tagged IP was greater than for the control, but there was a set of non-IP-specific proteins that comprised a general background due to tag-independent bead association. Correspondingly, there are very few proteins exclusively detected in the negative control, since the non-specific interactions with beads were present in both samples, although there may have been a slightly greater surface of bead unbound in the negative control due to the lack of FLAG-tagged complexes binding. There are many proteins that are exclusively detected in the IP. These proteins reflect those that are co-immunoprecipitated with Mcm4-FLAG that do not form a detectable part of the non-specific background. The S phase time point, in contrast, showed a less expected distribution of proteins identified. In this sample, the negative control displayed a considerable number of exclusive proteins, and an overall large number of proteins compared to the alpha-arrested sample. It was not clear exactly why this occurred, but with the ChIP-MS technique, it had been observed that the amount of non-specific background was much higher than in the native pull-downs, even despite more stringent wash steps. Regardless, the number of proteins

detected in the IP was similar to the other time point, and there were still a number of IP exclusive proteins.

The histograms of identified proteins with abundance information showed that indeed, when imputation was applied, disproportionately more proteins had their abundance imputed in the alpha-arrested negative control (Figure 41B). This sample was the only one where the applied imputation affected the overall distribution strongly. It was therefore expected that particularly at the low-abundance range of intensities, some proteins would be reliant on imputed values that could introduce some error. Otherwise, the distributions appeared well balanced and suitable for direct comparison. Figure 41C shows the result of a direct comparison of protein abundances in the alpha-arrested IP and negative control. Reassuringly, all six subunits of MCM2-7 are clearly enriched in the IP, as was Tah11/Cdt1. Also showing enrichment were Spt16 and Pob3, components of the FACT complex. Association of FACT with the replisome was observed both in the native Psf2-TAP pull-downs carried out in this study, and in the original GINS-TAP pull-down carried out previously (Gambus et al., 2006). Interestingly, this IP confirms that this association is present prior to the full activation of the replication machinery as also observed in that work. While previously this activation-independent association had also been seen with Mcm10, this ChIP-MS IP did not detect enrichment of Mcm10 at the alpha-factor-arrest stage. The methods used to prevent or preclude activation differ in each case. This experiment looked at alpha-arrested cells, whilst previously, a Psf2 degron was used to eliminate GINS activity and prevent full activation. There will have been several differences in the state of the cell and replication complex between these experiments, most notably, the MCM2-7 will have been phosphorylated and prepared for activation in the degron experiments, just unable to fully activate due to lack of GINS. In this experiment, the MCM2-7 complex will not have been yet prepared for activation in this way.

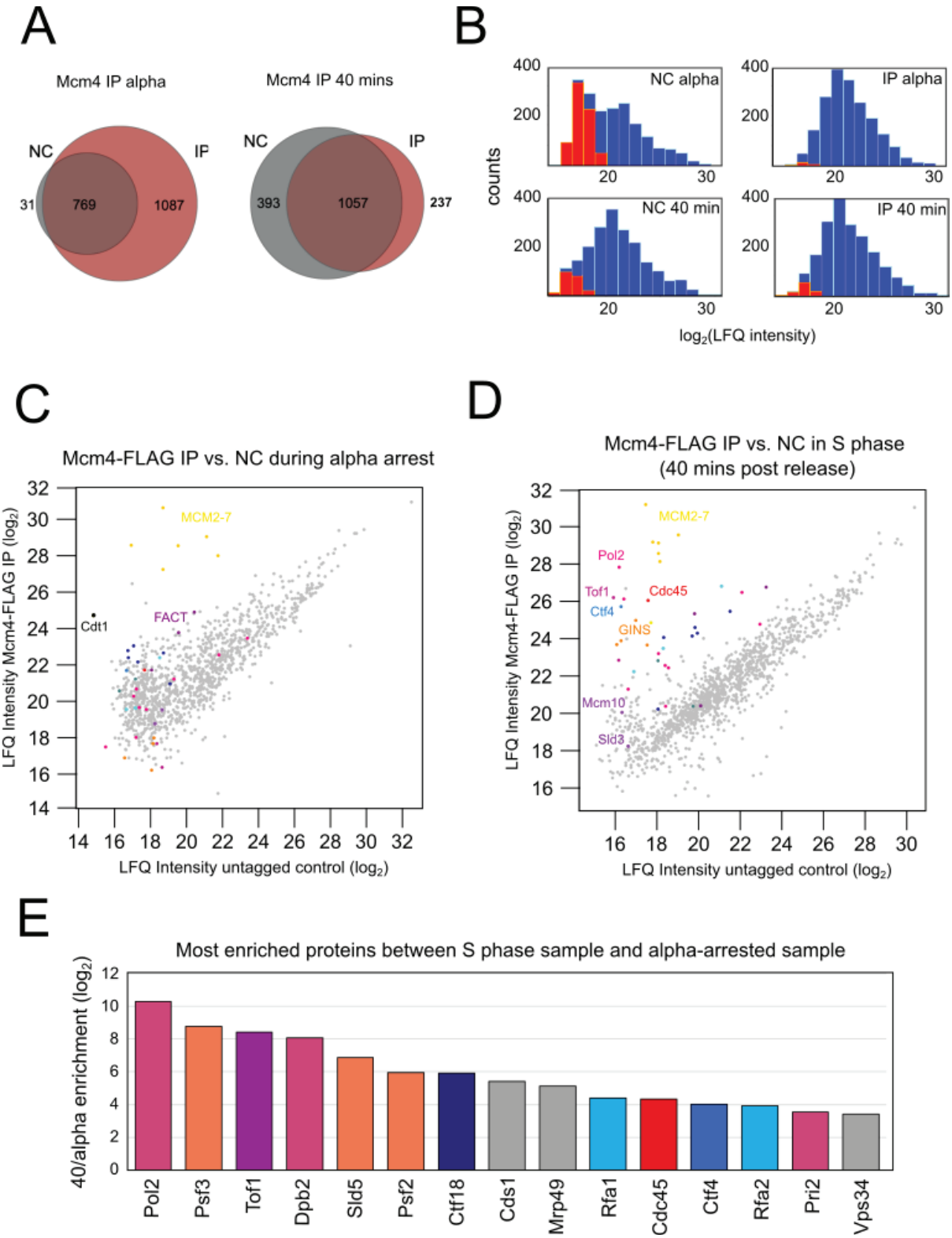


Figure 41: Trial ChIP-MS of Mcm4-FLAG in two separate cell cycle stages. A) Venn diagrams showing the proteins identified in the tagged pull-downs (red) and untagged control (grey) for both time points.The negative control sample for the 40 minutes has more total proteins and more unique proteins than ideal, but the tagged samples are comparable in proteins identified. B) Histogram of proteins identified and their intensities. The imputation affects the distribution of the negative controls, and most notably the alpha factor control. C) Scatter plot of protein abundances in the pull-down and untagged control samples from alpha-arrested cells. MCM2-7 is clearly enriched in the pull-down. D) As in C, but for the S phase sample. In addition to MCM2-7, many replisome proteins are also enriched in the pull-down. E) Visualisation of the enrichments of the most enriched proteins when comparing the S phase pull-down to the alpha-arrest pull-down. Many of the factors that increase in Mcm4 association are replisome factors.

The effects of imputation could be observed in this time point in the low-abundance proteins (Figure 41C). A cluster of proteins in the lower left of the plot showed a shift to the left, which is normally indicative of enrichment. In this case, due to the larger population shifted in this way, particularly at the low end of abundances, care must be taken with these proteins, as imputation can cause similar shifts.

The readout from the pull-down from S phase cells is much clearer (Figure 41D). Again, all six MCM2-7 subunits are enriched, along with many other factors already identified in Figure 40C. Many proteins comprising the core replisome are clearly enriched in this IP, similar to the Cdc45-FLAG IPs performed in the previous section. Comparing only the IPs from these two timepoints allowed investigation of how the complex composition changes over time (Figure 41E). The proteins with the strongest enrichment in S phase opposed to alpha-arrest are almost exclusively proteins with a known role in the replication fork, that associate during helicase activation in early S phase. A few non-replication associated proteins made this list (grey bars), where no clear explanation for their enrichment was found from a brief literature search. Because this was a single replicate, however, further data was required before any of these targets were considered for further investigation. A single replicate could reflect enrichment values that are not consistent, and there is always the risk that protein abundances detected in one sample are the result of cross-contamination. Whilst this risk was minimised as much as possible, further replicates were required to investigate exactly how the replication complex composition varied throughout its life cycle.

3.6.2 Mcm4-FLAG ChIP-MS: from loading to termination?

In the previous section, it was shown that the changing composition of the replication complex could be investigated using a ChIP-MS approach with Mcm4-FLAG as the bait protein. These initial experiments were done using the G1-like arrest induced by alpha-factor treatment of mating type a cells. The results from these experiments provided evidence for the changing complex, detectable by the ChIP-MS technique. The results also suggested that even in the alpha factor arrest, there appeared to be some association of replication complex components with the MCM2-7 helicase. Whilst activation of the helicase had not yet occurred in alpha factor, it was considered that the arrest state may superficially resemble G1, but by using this "artificial" arrest, the results may not fully represent the true progression from G1 to S. Since loss of synchronicity would both reduce the efficiency of replisome pulldown, and make time-dependent analysis impossible, to test this hypothesis, a synchronisation step was required to replace the pheromone treatment. Therefore, a methionine-promoter driven Cdc20 method was utilised to arrest cells in mitosis, as described in the methods section. Addition of methionine to cells containing this construct caused repression of the cell's sole copy of Cdc20, a protein required for APC/C activation and mitotic exit. This block enables synchronisation of cells in a mitotic state. Cells that are grown in the absence of an external source of methionine or that have methionine-containing media washed away with media lacking methionine, are able to exit mitosis and divide normally.

The MetP-Cdc20 construct was integrated into cells containing Mcm4-FLAG, and the arrest phenotype tested (Figure 42A). Treatment with both solely methionine to a final concentration of 2.5 mM, and replacement of the media with YPD was tested. Both approaches were found to cause a secure mitotic arrest after two hours of treatment, but release from the block was much more immediate and reliable in cells that were arrested using YPD.

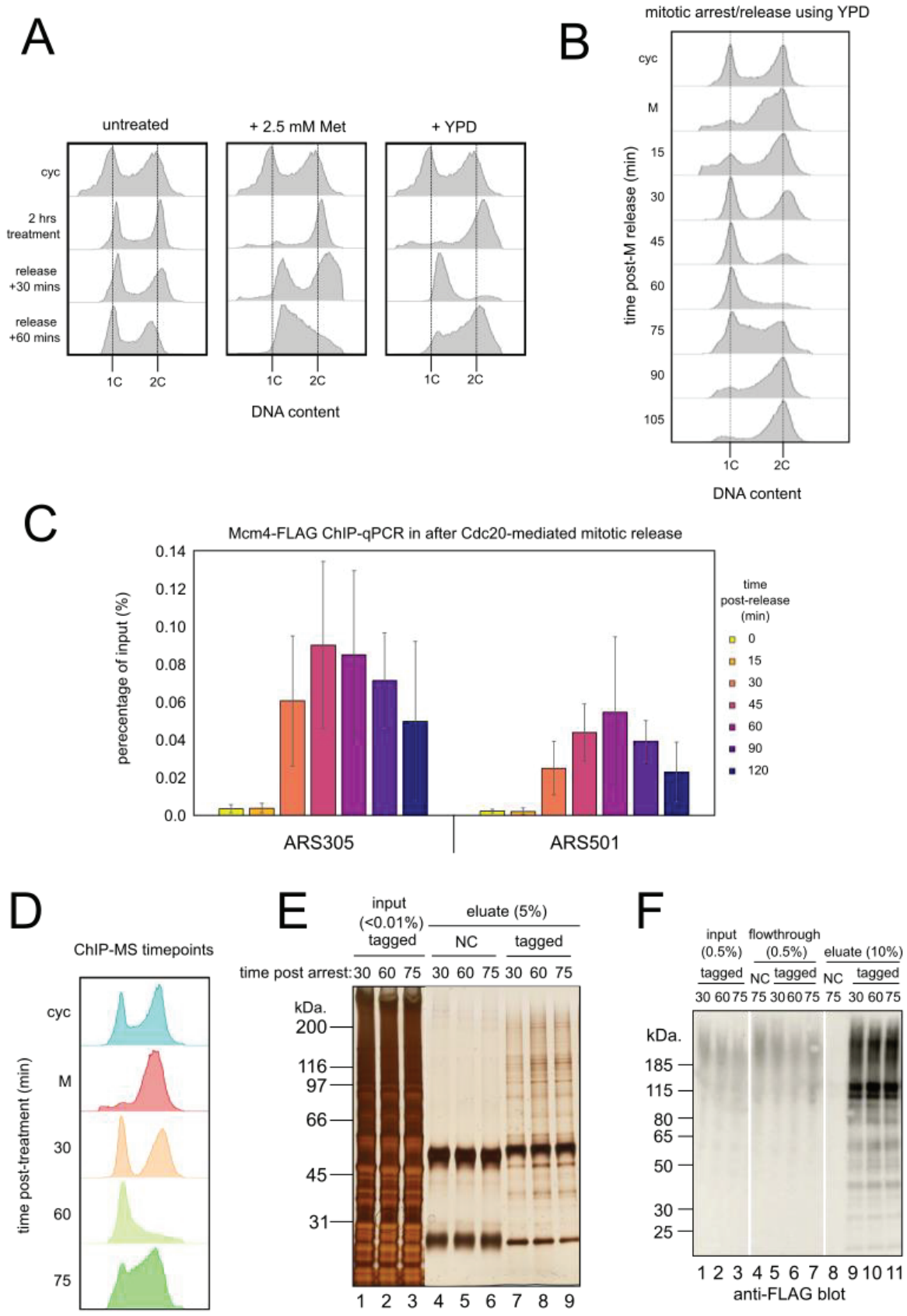

A
untreated
+ 2.5 mM Met
+ YPD
cyc
2 hrs treatment
release +30 mins
release +60 mins
1C
2C
DNA content
B
mitotic arrest/release using YPD
cyc
M
15
30
45
60
75
90
105
time post-M release (min)
1C
2C
DNA content
C
Mcm4-FLAG ChIP-qPCR in after Cdc20-mediated mitotic release
perecentage of input (%)
0.14
0.12
0.10
0.08
0.06
0.04
0.02
0.0
time post-release (min)
0
15
30
45
60
90
120
ARS305
ARS501
D
ChIP-MS timepoints
time post-treatment (min)
cyc
M
30
60
75
E
input (<0.01%) tagged
eluate (5%)
NC
tagged
time post arrest: 30 60 75 30 60 75 30 60 75
kDa.
200
116
97
66
45
31
1 2 3 4 5 6 7 8 9
F
input (0.5%) tagged
flowthrough (0.5%) NC
flowthrough (0.5%) tagged
eluate (10%) NC
eluate (10%) tagged
30 60 75 75 30 60 75 75 30 60 75
kDa.
185
115
80
65
50
30
25
1 2 3 4 5 6 7 8 9 10 11
anti-FLAG blot

Figure 42: Initial tests of a pMet-Cdc20-induced arrest followed by release and the effect on cell cycle event timings. A) Flow cytometry data of DNA content as measured by SYTOX incorporation in cells containing the sole copy of Cdc20 under the copy of the Met promoter under different media conditions followed by release from said media conditions. B) Flow cytometry analysis of the DNA content as measured by SYTOX incorporation in pMet-Cdc20 cells arrested in YPD, then released into media lacking Methionine for the indicated time. C) ChIP-qPCR analysis of Mcm4-FLAG origin association over the course of a release from mitotic arrest mediated by the pMet-Cdc20-mediated arrest. Error bars represent standard deviation (s.d.) of three independent experiments (n=3). D) Flow cytometry analysis of cells taken at specific time-points post-mitotic release. Cells at 30, 60 and 75 minutes were also used for an Mcm4-FLAG ChIP-MS experiment. E) Silver stain visualisation of SDS-PAGE analysis of Mcm4-FLAG ChIP-MS samples. Input indicates cell extract immediately prior to addition of IP antibody, eluate refers to sample directly after bead elution prior to preparation for mass spectrometry analysis. F) Western blot of SDS-PAGE gel run of samples equivalent to those in E. Membrane was incubated with anti-FLAG HRP overnight at 4℃ before visualisation.

Flow cytometry analysis was carried out on cells arrested using this method, to look for potential post-release times that could be used to investigate specific MCM2-7 complexes during loading, activation or normal replication (Figure 42B). At thirty minutes post release, roughly half the cells had successfully divided, and this was chosen as a time-point to investigate the complexes involved in helicase loading, since it would be around this time that helicases would be being loaded at origins. At around sixty minutes, the first signs of DNA replication were observed, making this a reasonable time-point to investigate activating complexes at early origins. At seventy-five minutes, most of the cells in the population were actively replicating, with quite a large proportion of cells in the main body of S phase. This time-point was selected as the most likely time to produce the largest proportion of actively replicating replisomes. Post seventy-five minutes, fewer and fewer cells remained in S phase, making these timepoints interesting to study exclusively terminating complexes. Because the arrest was now more distant in time from the entry into S phase, there was more opportunity for drift of the cells away from synchronicity. In addition to the natural variation in firing times of early and late origins, this drift would likely mean that the seventy-five minute time-point would be highly heterogeneous in terms of the helicase-containing complexes present. There could easily be a sub-population of initiating complexes alongside complexes that were undergoing termination. This risk is not entirely unique to this experiment, even in alpha-factor arrested cells, there are likely to be late origins still firing, whilst replication forks produced

from early origins that are close may have met and the replisomes terminated. This approach would probably increase the overall heterogeneity, however.

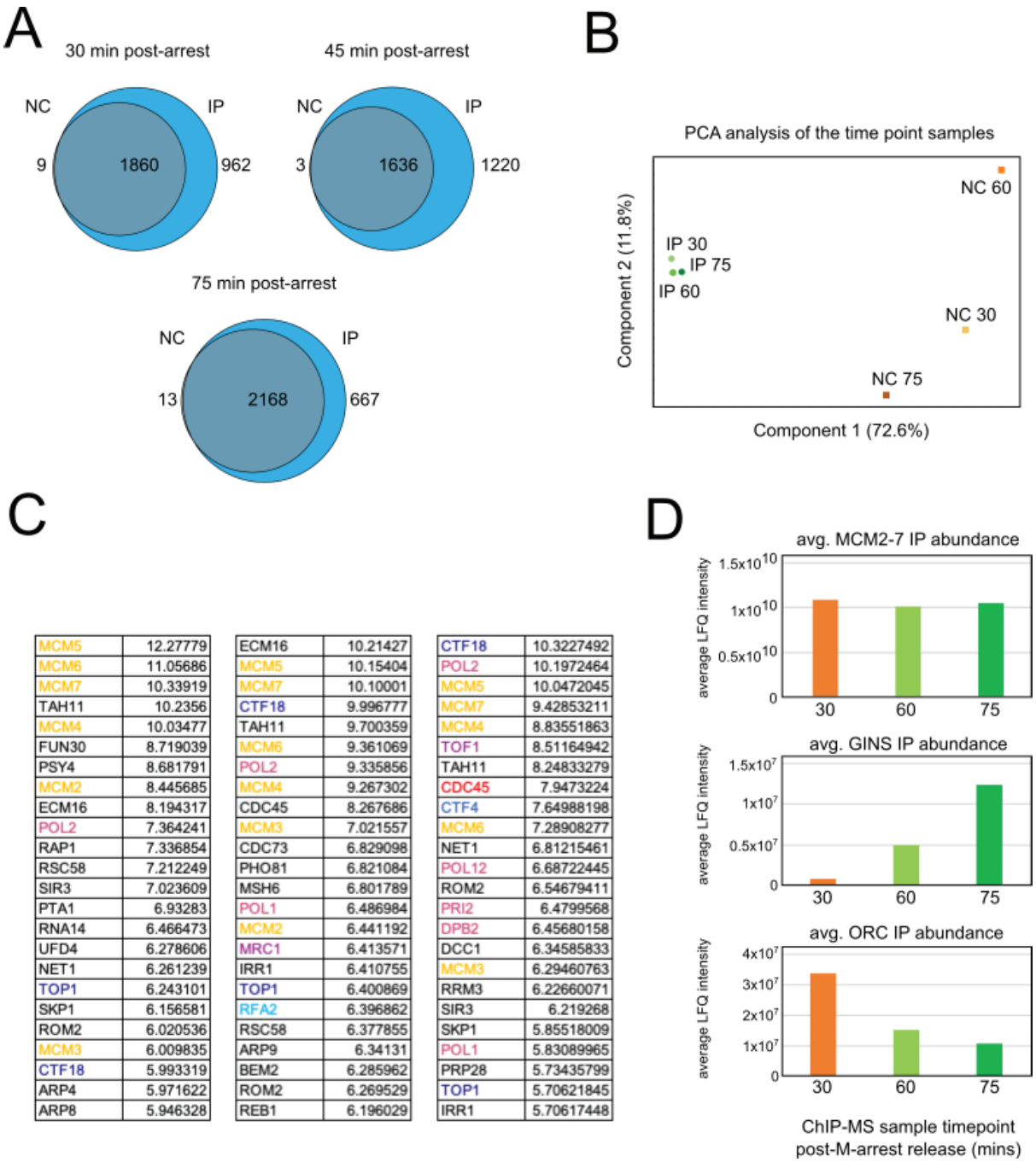
To show MCM2-7 dynamics with respect to chromatin association, ChIP-qPCR of Mcm4-FLAG was carried out over the course of a release from mitosis using the inducible Cdc20 expression (Figure 42C). Prior to 30 minutes, there was minimal association of MCM2-7 with origin DNA, indicating that helicase loading had not yet begun. The peak of MCM2-7 association depended on the origin tested, but was generally around 45 to 60 minutes. Given maximal MCM2-7 loading is expected when the maximum number of complexes are loaded at origins, which would be after loading but prior to activation, this suggested an activation time in the region of 45-60 minutes for ARS305, and 60-90 minutes for ARS501, which matches the difference in firing times observed for these origins in the literature (Santocanale, Sharma & Diffley, 1999).

As an initial test of some of the time-points chosen, ChIP-MS was carried out on mitosis-released cells after 30, 60, and 75 minutes and sent for mass spectrometry analysis (Figure 42D-F). The flow cytometry analysis in Figure 42C shows a close match to the previously observed cell cycle dynamics (compare to Figure 42B), though at the 75 minute time-point, cells were in general further progressed through S phase. The pull-downs were checked by silver staining, and showed a number of surprisingly clear IP-specific bands (Figure 42E). A western blot confirmed that the pull-down sample eluates, but not a representative negative control sample, contained a FLAG-tagged protein (Figure 42F). Surprisingly, when IP input and flowthrough samples were also blotted with anti-FLAG, the negative control sample produced equal signal to the tagged samples in these lanes (Figure 42F, lane 4). This could indicate the presence of factors in the control and tagged strains that could generate non-specific binding of anti-FLAG antibody. Regardless, the purification procedure effectively removed this signal in the negative control eluate, suggesting while it was able to bind the

blotting antibody in a non-specific manner, it did not bind to the anti-FLAG beads in a meaningful way.

Following mass spectrometry analysis, all three time-points tested appeared relatively consistent in the overall protein levels, and number of factors identified in each pull-down and untagged sample (Figure 43). In general, the untagged controls showed fewer identified proteins than their respective pull-downs, and very few unique proteins. The 75 minute time-point did show a slightly greater number of proteins in the negative control than the others, but still fewer than its pull-down pair (Figure 43A). By comparing the protein sets by principal component analysis, it was clear that the pull-downs grouped closely with each other, and were well separated than the more variable untagged controls (Figure 43B). Moving on from analysis of the protein set as a whole, the proteins which showed the most enrichment in each of the time-points were calculated and displayed in Figure 43C. The MCM2-7 subunits showed consistently strong enrichment, and were represented among the most enriched proteins in all three samples, aside from Mcm2 in the 75 minute timepoint. However, with an enrichment score of 5.54, this subunit was still very much enriched.

Most interestingly, the composition of the most enriched proteins was seen to vary across the time-points, replicating effect seen in the explorative experiment shown above (Figure 41). At 30 minutes, prior to the activation of replication, the pull-down contains much fewer replication-associated factors. Of the factors associated with replication forks, only Pol2, Top1 and Ctf18 show very strong enrichment. Instead, some of the most enriched proteins are the MCM2-7 complex itself, and the MCM2-7 binding partner involved in origin licensing, Cdt1/Tah11.



MCM5	12.27779
MCM6	11.05686
MCM7	10.33919
TAH11	10.2356
MCM4	10.03477
FUN30	8.719039
PSY4	8.681791
MCM2	8.445685
ECM16	8.194317
POL2	7.364241
RAP1	7.336854
RSC58	7.212249
SIR3	7.023609
PTA1	6.93283
RNA14	6.466473
UFD4	6.278606
NET1	6.261239
TOP1	6.243101
SKP1	6.156581
ROM2	6.020536
MCM3	6.009835
CTF18	5.993319
ARP4	5.971622
ARP8	5.946328

ECM16	10.21427
MCM5	10.15404
MCM7	10.10001
CTF18	9.996777
TAH11	9.700359
MCM6	9.361069
POL2	9.335856
MCM4	9.267302
CDC45	8.267686
MCM3	7.021557
CDC73	6.829098
PHO81	6.821084
MSH6	6.801789
POL1	6.486984
MCM2	6.441192
MRC1	6.413571
IRR1	6.410755
TOP1	6.400869
RFA2	6.396862
RSC58	6.377855
ARP9	6.34131
BEM2	6.285962
ROM2	6.269529
REB1	6.196029

CTF18	10.3227492
POL2	10.1972464
MCM5	10.0472045
MCM7	9.42853211
MCM4	8.83551863
TOF1	8.51164942
TAH11	8.24833279
CDC45	7.9473224
CTF4	7.64988198
MCM6	7.28908277
NET1	6.81215461
POL12	6.68722445
ROM2	6.54679411
PRI2	6.4799568
DPB2	6.45680158
DCC1	6.34585833
MCM3	6.29460763
RRM3	6.22660071
SIR3	6.219268
SKP1	5.85518009
POL1	5.83089965
PRP28	5.73435799
TOP1	5.70621845
IRR1	5.70617448

Figure 43: Analysis of the mass spectrometry data generated from the Mcm4-FLAG ChIP-MS extended time-course of a single replicate. A) Venn diagrams indicating the proteins identified by MS, showing shared and unique proteins in tagged (IP) and untagged (NC) samples for each of the three time-points tested. B) PCA analysis of the tagged and untagged pull-downs. C) The most enriched proteins in the pull-downs, 30 min, 45 min and 75 min post-mitosis (left to right). For missing values in the negative control, conservative manual imputation was carried out as previously, using the lowest 10% of NC proteins average intensity. D) Average LFQ values of MCM2-7, the GINS complex, and ORC1-6 over the three time-point IPs (n=1).

Over the course of this exploratory time-course, the number of replication fork associated factors represented in the most enriched proteins increased. By the 75 minute time-point, Ctf18 and Pol2 are the most enriched proteins, and many other replication factors are also strongly

enriched, for example Cdc45 and Ctf4. The maintenance of Cdt1/Tah11 among the most enriched proteins likely indicates that soluble MCM2-7 complexes are still being isolated alongside chromatin-bound MCM2-7, even at the later time-points. Several other highly enriched factors were also noted in this experiment, including Fun30, Rrm3 and Sir3. As discussed, these factors had been shown to have either loosely-associated roles in replication, or had a chromatin-associated role with no proven replication involvement. The high enrichment values observed for these proteins were of interest, and the behaviour of these factors in larger scale experiments was closely monitored going forwards. Fun30, whose human homolog's association with the replisome had not yet been published, was of particular interest, since it was known to be a chromatin remodeller, and appeared to be enriched most prior to activation, during loading. It was hypothesised that Fun30 may have a role to play in establishing a chromatin environment at origins that was permissive or restrictive to loading. This role may be specific for *in vivo* replication of chromatinised DNA, and so could implicate Fun30 as a new factor required for replication or the regulation thereof.

By applying a more quantitative approach, and observing the changes in LFQ intensity (a normalised measure of the protein abundance in the MS sample) for multi-subunit complexes involved in replication, general trends could be observed that fit with the observations based on individual factors and cell cycle state (Figure 43D). The MCM2-7 complex shows stable abundance in the pull-down, as should be expected since this stable hexamer contains the bait protein Mcm4. This indicated that the combination of loading equal protein amounts and the post-MS normalisation, was generating well-comparable samples. Confident that despite being limited to a single replicate, this experiment could still provide useful data, this analysis was extended to other factors. The average observed intensity of the GINS complex was representative of many other individual factors and subcomplexes associated with the fork in this experiment (data not shown). Indeed, the dynamics observed also fit with expectations

from cell cycle stage. At 30 minutes, prior to activation, GINS showed very little association with MCM2-7. Once activation was seen to begin around 60 minutes (Figure 42C), the abundance of GINS in the pull-down increased greatly. By 75 minutes, when many cells were undergoing the main to late stage of S phase, the association of GINS with MCM2-7 was greatest, most likely representing the presence of many CMG complexes, where GINS is stably associated with the helicase. The opposite trend was observed for ORC. ORC1-6 was seen to reduce in abundance in the pull-down over time. This fits predictions well, as during loading, the association of ORC with MCM2-7 is clear, since it physically binds MCM2-7 to load the helicase onto origin DNA. ORC abundance in the pull-down decreases over the time-course, as the helicase loading activity in cells is ceased, and MCM2-7 complexes move away from the origins where ORC generally binds. The dynamics of many other factors were also investigated, which exemplifies the strength of ChIP-MS when compared to ChIP-qPCR, since it can observe the association of many factors simultaneously, as opposed to the single factor ChIP-qPCR approach.

The success of this single-replicate condensed time-course supported the progression to a much larger, multiple replicate time-course with more time-points. Running the time-course in quadruplicate would allow a much more quantitative measure of the dynamics observed of factors, whilst further time-points could explore more sectors of the cell cycle, and investigate previously targeted temporal regions with more precision.

3.6.3 Mcm4-FLAG ChIP-MS full time-course in quadruplicate

Following the success of the single-replicate Mcm4-FLAG ChIP-MS time-course, the experiment was repeated in quadruplicate covering extended time-points. The Cdc20-mediated mitotic arrest was again utilised, and cells were released from this block into S phase. Six time-points were selected to harvest cells at: 0, 30, 45, 60, 75 and 90 minutes post release. Immediately after release represented a mitotic state, where MCM2-7 hexamers should all be

non-chromatin associated, and provide a control of sorts to remove soluble MCM2-7 interactors from comparative enrichment analyses. Thirty minutes was chosen to investigate the formation of pre-RC and MCM2-7 loading complexes prior to double hexamer formation. At forty-five minutes, DNA replication had not yet begun, but the majority of cells had undergone cytokinesis, and Mcm4-FLAG ChIP-qPCR under this release protocol showed maximum Mcm4 origin association at this time (Figure 42C). This was likely to represent a state prior to replication initiation, and be as close as possible to the alpha factor-mediated arrest previously studied. The first signs of DNA replication by flow cytometry were visible at around sixty minutes (Figure 42D), and so this time was chosen to attempt to capture the complexes undergoing the transition from double hexamer to functional helicase. At seventy-five minutes, a large proportion of cells were in the main body of S phase and so this was used to further investigate the active replisome isolated by Mcm4-FLAG with more replicates. By ninety minutes, most cells had a DNA complement of near 2N, indicating that DNA replication was nearing completion or had ended. This time-point would be used to try and investigate terminating complexes and their protein complexes. Cell cultures were grown on four successive days to reduce the variability of growth conditions, and the ChIP-MS extract preparation and wash steps were carried out on two separate days, two replicates per day, processed separately. Figure 44 panel A shows the results of SDS-PAGE analysis of the sample eluates visualised by silver stain. The presence of certain IP-specific bands and general protein smear density was visible in the tagged samples, and mostly the general protein density of tagged samples was greater than the untagged negative controls. In some of the untagged controls (e.g. replicate 4, lane 4-6), a level of density was visible that was greater than expected. To rule out the possibility of cross-contamination and confirm the efficiency of the Mcm4-FLAG pull-down, the SDS-PAGE analysis was repeated and Western blotting used to investigate the presence of FLAG epitope in the eluates (Figure 44B). All of the tagged samples

displayed reasonably high levels of FLAG epitope at the expected height of the Mcm4-FLAG protein, with the cross-linking induced smearing also visible. The quantity of Mcm4-FLAG present in the eluted samples was not entirely consistent between samples, and there was no clear temporal pattern visible over the time-course. However, in the untagged controls, there was very little FLAG epitope signal visible compared to the pull-downs. Most samples showed no signal even at greater exposure, though in one untagged control (rep. 4 NC 60 minutes) there was a faint band visible (high exposure blot not shown).

At each time-point cells were harvested, a small number of cells were taken for flow cytometry analysis (Figure 44C). An exemplar replicate release profile is shown. In contrast to previous trials of this release (Figure 42B,D), the cells proceeded into S phase faster in all four replicates, with signs of DNA replication visible at 45 minutes, and with a large number of cells in the main body of S phase by 60 minutes. This adjusted the stages of the replication complex that were being isolated in this experiment. However, the offset did allow many of the intended stages to be investigated. In this large scale experiment, then, 45 minutes appeared to be the time-point where initiating replication complexes could be isolated. 60 minutes would be the time-point where active replisomes were likely to dominate, and whilst 75 and 90 minutes would likely appear more similar than previously, 75 minutes would likely have active and terminating replisomes, and 90 minutes could possibly be dominated by terminating complexes with very few active replisomes remaining.

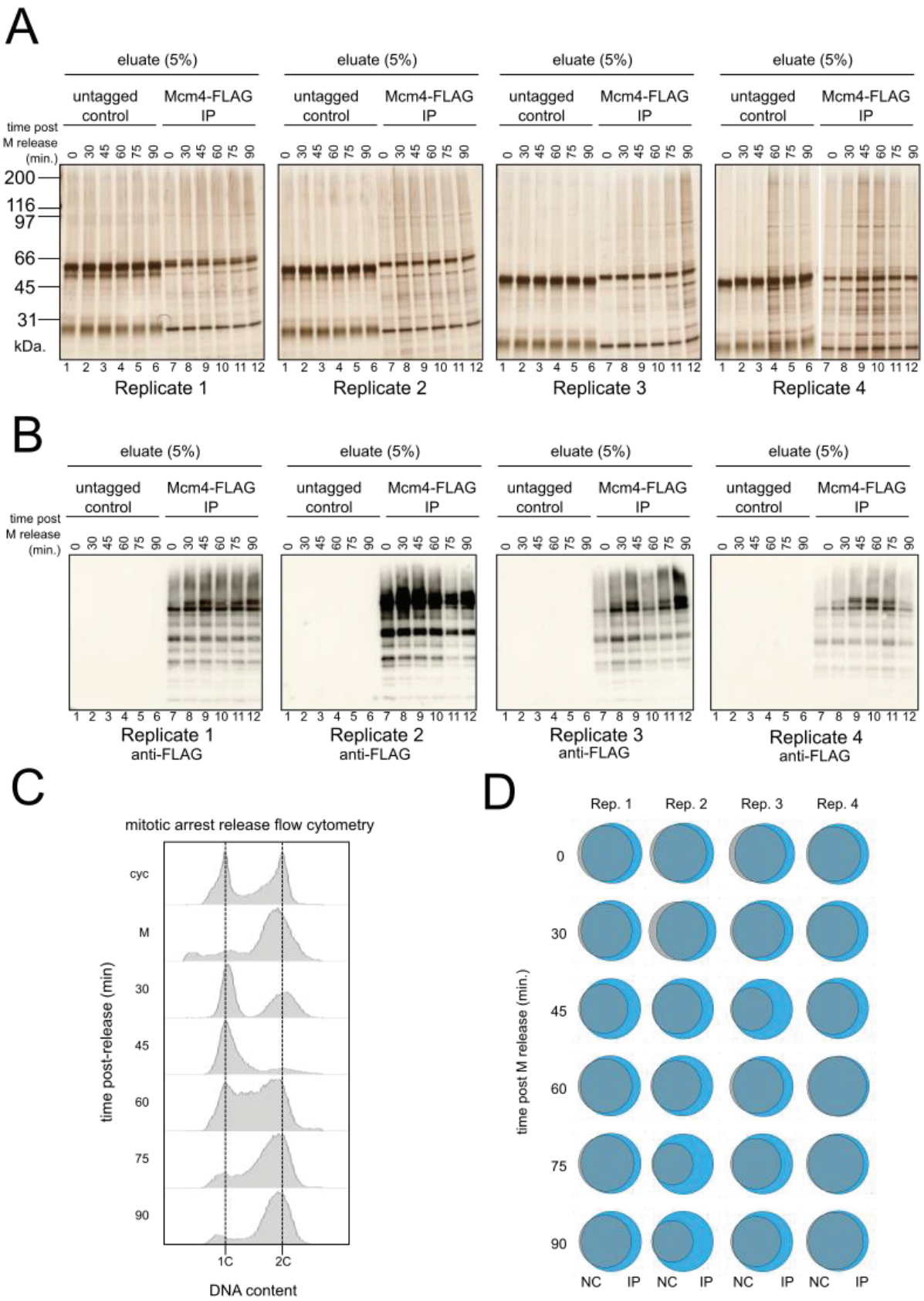


Figure 44: Analysis of Mcm4-FLAG ChIP-MS time-course in quadruplicate, intial sample quality tests. A) Silver stain analysis of SDS-PAGE gels of control and pull-down samples. Lanes in replicate 4 re-arranged from their original layout for clarity. B) Analysis of samples by Western blotting. SDS-PAGE gels as in A, were instead analysed by Western blotting to check for anti-FLAG signal in the pull-downs, and identify any cross-contamination of negative controls. C) Flow cytometry analysis of an example replicate. In general, cells released into S phase at a faster rate than previous tests D) Venn diagrams showing the results of unique protein identifications in each pair of tagged and untagged samples. Light blue regions indicate proteins uniquely identified in tagged samples, and purely grey areas indicate proteins uniquely identified in the untagged control.

As a further quality control measure, the number of unique proteins in each sample of the tagged and untagged pairs was compared across the experiment (Figure 44D). Generally, the number of unique proteins in the tagged sample was greater than that in the untagged sample. The majority of proteins detected were present in both samples in the pair, however. Some pairs showed more unique proteins in the tagged pull-down than others, but there was no clear pattern overall that caused any concern. The number of proteins detected overall in each pair was also consistently high, with an average of 2747±152 proteins detected for each IP-NC pair.

Principal component analysis was carried out on the samples following mass spectrometry analysis (Figure 45). Pool samples, made up of small volumes taken from every sample, were also run to ensure the spectrometer functionality was maintained over the course of the run, and these were included in the PCA analysis. Importantly, there was both a clear separation between the tagged samples and the untagged controls, and the samples tended to group with others of the same time-point rather than with samples of the same replicate. A slight trend also emerged in the PCA analysis, where except for 0 minutes, the other time-points showed a linear relationship along both principal components, possibly indicating a progression of the isolated complex over time. The 0 minute timepoint also associated the closest with the untagged controls, which could indicate that this sample had fewer IP-specific proteins, and overall lower complexity. It is not possible to draw too many firm conclusions from such analysis, however.

To investigate the composition of the helicase complex at each timepoint, the enrichment of proteins in pull-down over the untagged control was calculated as before. Where a protein was not detected in the untagged control, imputation was carried out manually or in the Perseus analysis software as previously.

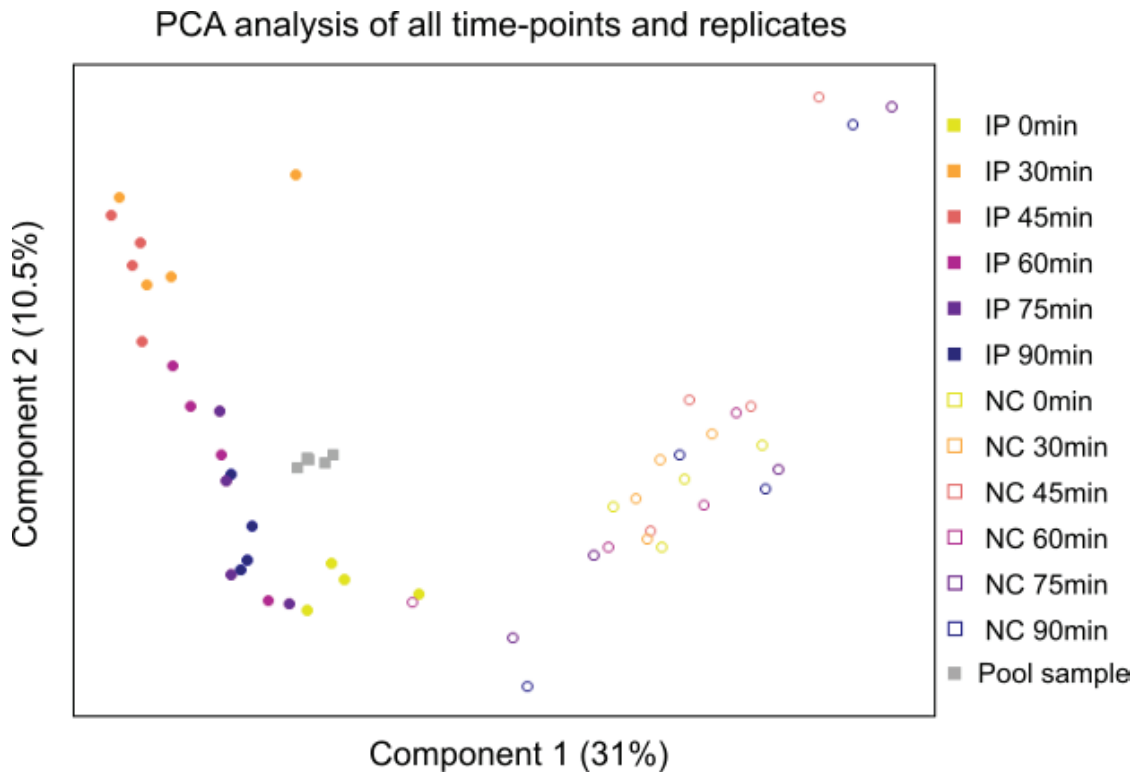


Figure 45: Principal component analysis of all time-course samples submitted to mass spectrometry, including a set of 4 "pool" samples containing a small sample volume from every sample mixed, to ensure comparability of the start and end of the instrument run. Tagged pull-down samples are indicated by filled circles, and untagged controls indicated by open circles, coloured by time-point.

Figure 46 shows the distribution of intensity values of detected proteins for each sample. Protein values that were imputed are shown in red. For the tagged pull-down samples, the relative number of imputed values was generally low, and the resultant introduced low intensity values filled out, rather than disrupted, the lognormal distribution of the protein intensities. The 0 minute timepoint showed the most need of imputation, which reflected the fact that at this time-point, the MCM2-7 complex was not chromatin bound, and so possibly had a much reduced set of associated proteins. This could have led to more of the low cellular abundance, replication specific proteins not being detected and showing up as missing values due to their lack of MCM2-7 interaction at this time. The untagged control samples showed on the whole a much increased number of imputed values. The introduction of these values did disrupt the smooth lognormal distribution. However, since these were untagged negative controls meant

to show the non-specific background, a lower number of proteins were expected to be detected in these samples. In this case, imputation was still occurring on the lower end of protein intensities, reducing the chances of inappropriately high artificial enrichments being produced. Some of the untagged control samples displayed greater numbers of missing values that required imputation than others. This indicated that the non-specific background was slightly variable over all samples. This could be due to variability introduced at multiple points, but is most likely a result of slight variance in wash steps. This variability was introduced despite all due care being taken to carry out this step across all four replicates within a narrow time frame, utilising the same buffers, and standardising the technique.

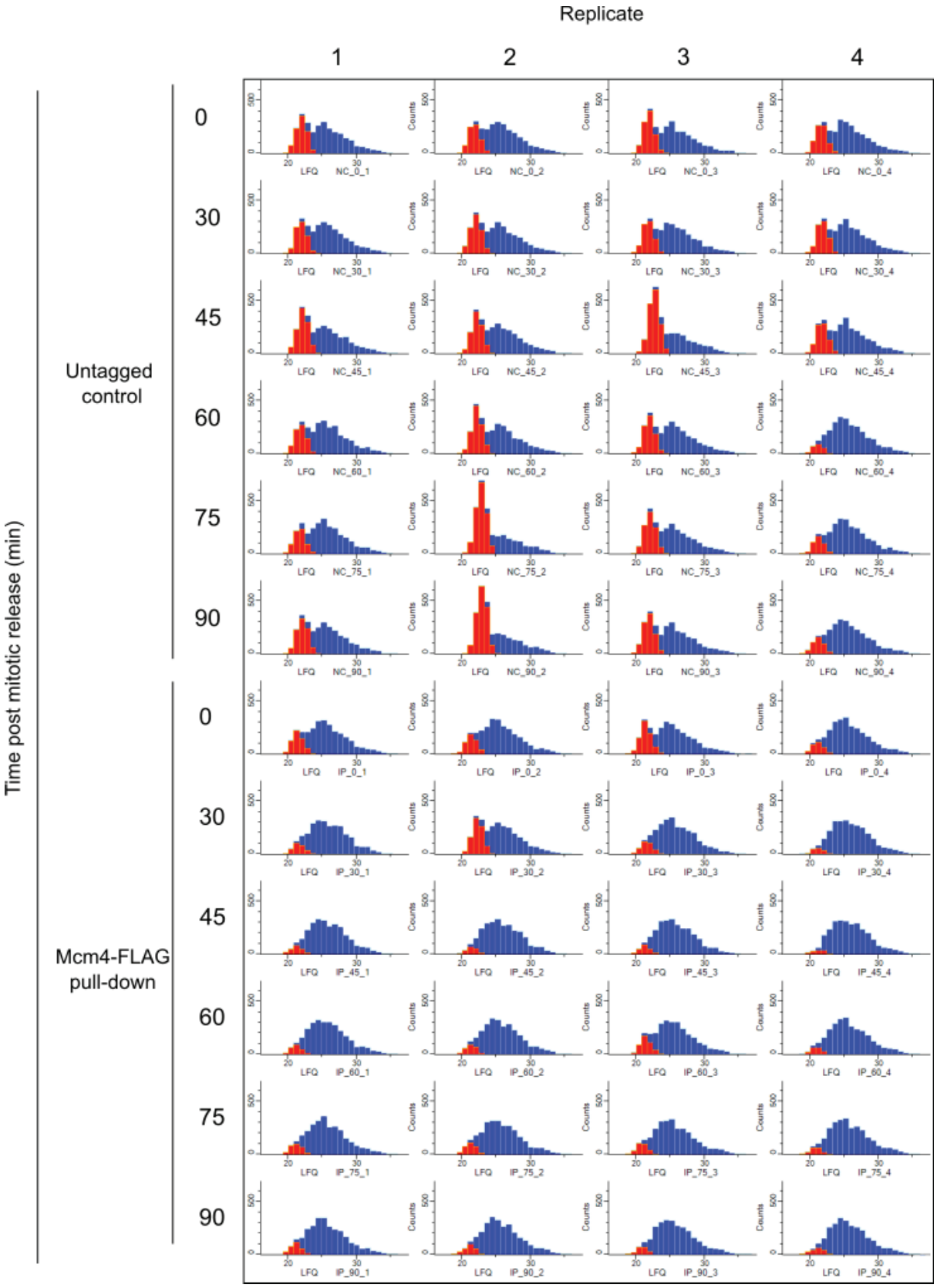


Figure 46: Histograms of protein intensities in time-course mass spectrometry samples. Proteins with measured intensities are shown in blue, and imputed proteins are indicated in red. Imputed proteins fall at the lower end of the intensity scale, and are more prevalent in the untagged controls, which have a greater number of missing values due to the lack of tag-specific pull-down.

Despite minor issues with imputation specifically in the untagged controls, the pull-down data was of sufficient quality to proceed to more detailed analysis of protein composition and dynamics across the time-course. As before, an informative method of displaying enrichment data in mass spectrometry is a volcano plot, combining the enrichment values in one dimension with the significance values of a permutation based two-sample t-test on the other. Performing this analysis on the six time-points collected when comparing the tagged pull-down to the untagged negative control generated the plots shown in Figure 47.

In all cases, the six MCM2-7 subunits were among the most reliably enriched proteins (Figure 47A-F, yellow points). This confirmed that the pull-down was correctly targeting and enriching for this complex at every stage of the time-course. Whilst the exact positioning of the subunits was variable from time-point to timepoint, reflecting changes in both the calculated enrichment score and the p-value, the overall position towards the top right of the chart was consistent.

Another relatively consistent factor present in the pull-down was the MCM2-7 binding partner Cdt1. As discussed above, this factor binds hexamers prior to and during loading onto DNA, and therefore likely represents the amount of soluble non-chromatin associated MCM2-7. The enrichment of Cdt1 was highest at the 0 minute time-point (Figure 47A), but nonetheless remained high throughout the time-course, indicating reasonable amounts of soluble MCM2-7 present. It was proposed that the level of Cdt1 in the IP generally could be used as a proxy measure for the amount of soluble vs. chromatin-bound MCM2-7.

The enrichment of other factors was more variable across the time-course, however. A notable protein that displayed variable enrichment was the origin recognition complex (ORC). The complex, excluding Orc3, was not enriched at the 0 minute timepoint, but by 30 minutes, five out of the six subunits were significantly enriched in the pull-down (Figure 47B, dark green).

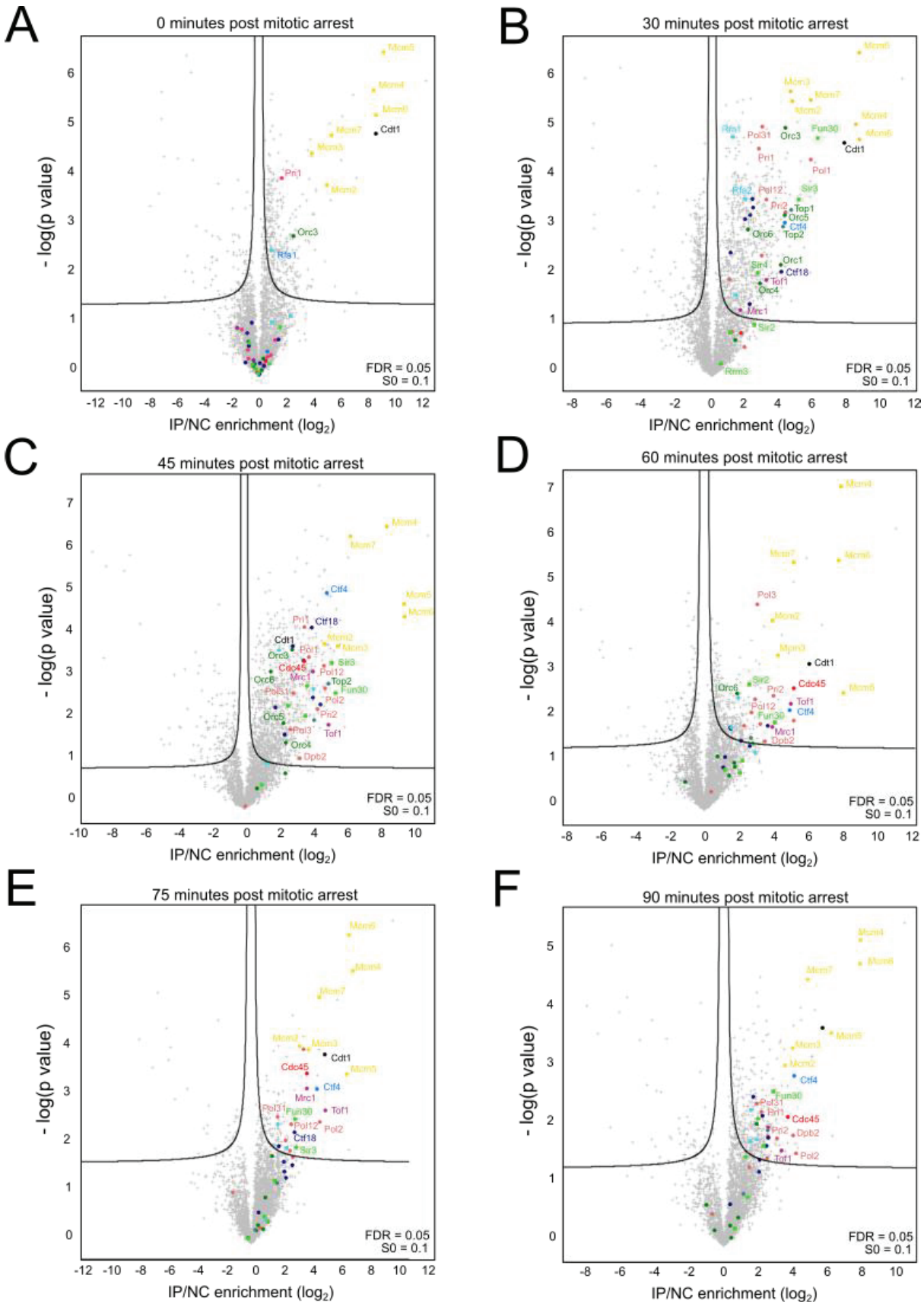

A
0 minutes post mitotic arrest
B
30 minutes post mitotic arrest
C
45 minutes post mitotic arrest
D
60 minutes post mitotic arrest
E
75 minutes post mitotic arrest
F
90 minutes post mitotic arrest
- log(p value)
IP/NC enrichment (log2)
FDR = 0.05
S0 = 0.1

Figure 47: Volcano plots showing the relationship between factor enrichment in the tagged IP versus untagged control and p-value as calculated by a two-sample t-test for each of the time-points. For all plots the false discovery rate (FDR) was set to 0.05, and the S0 value set to 0.1. A manually curated subset of factors associated with the replication complex are highlighted as coloured filled circles, and significantly enriched members of this subset are labelled. The factors are colour-coded as follows: Red = Cdc45, Yellow = MCM2-7, Pink = Polymerases, Dark green = ORC1-6, Turquoise = Topoisomerases 1/2, Light blue = RPA, Blue = Ctf4, Dark blue = Clamp and clamp loader complex, Purple = Fork protection complex, Black = Cdt1, Light green = additional protein factors of interest.

This increased association of ORC1-6 in the pull-down indicated that a subset of the pulled-down MCM2-7 had become chromatin-associated. ORC is consistently DNA-bound, and binds to origins of replication strongly in G1. This was a clear indication that helicase loading was being carried out at the 30 minute time-point. ORC1-6 enrichment could be seen to decrease at further time-points, indicating that the association of ORC and the MCM2-7 complex was reduced. This could be the result of MCM2-7 helicases activating and moving away from the origins. An alternative hypothesis suggested by data collected by L.M. Reuter in the lab would suggest that once MCM2-7 is loaded as a double hexamer, ORC is displaced from origins (manuscript submitted). If this is indeed the case, the reduced ORC1-6 association would then reflect completion of helicase loading and double hexamer formation. By the 60 minute time-point, ORC1-6 enrichment is much reduced compared to 30 minutes, either suggesting movement of helicases away from origins, or completion of loading and double hexamer formation. Flow cytometry data from this experiment favoured the first hypothesis, since in this case, replication had clearly begun by 60 minutes (Figure 44C), but it is possible that loading was still occurring at the 45 minute time-point and ORC association with MCM2-7 does indeed decrease immediately following double hexamer formation.

An interesting result was observed when looking at factors typically representative of fully active replication complexes, such as polymerases, RPA, Ctf4 and the fork protection complex. Generally these factors did not show significant enrichment at 0 minutes, prior to the beginning of helicase loading, but some of them began to show significant enrichment as early as 30 minutes (Figure 47B). Flow cytometry data indicated that this was a significant period before replication activation (Figure 44C). The presence of close association of these factors with

Mcm4 at this time-point was therefore unexpected. Whilst all three of the replicative polymerases showed significant enrichment of at least one of their subunits, since they are stable complexes, it would be expected that they would display similar enrichments. The enrichment of polymerase ε is therefore the least certain at this time, with only Dpb4 being significantly enriched. However, polymerase α shows significant enrichment of all four subunits at this very early time-point. It is hard to conclusively show that this is not as a result of early initiation events in some cells, but given time before replication was visible by flow cytometry, and the magnitude of the enrichment, this enrichment appears to occur prior to activation for a reasonable proportion of helicases. Together, the association of certain replication factors with the helicase prior to initiation indicated some pre-association of factors. This need not be direct pre-association with the MCM2-7 helicase itself, since the cross-linking step allows retention of factors that might be indirectly linked by intermediate factors or by DNA. However, it does indicate that certain replication proteins, may be associated with the origins of replication and/or MCM2-7 earlier than might have been expected.

At 45 minutes post mitotic release, the complex composition appeared to build upon what had been observed at 30 minutes (Figure 47C). Notably, Cdc45 was significantly enriched at this time, implying that helicase activation, driven by kinase activity, was underway. Almost all of the polymerase subunits detected in the mass spectrometry were now significantly enriched (aside from Dpb4), and Ctf4 showed very strong enrichment. Because cells in this experiment had proceeded more quickly to S phase than previous trials (Figure 44C), this could correspond to the activation of replisomes.

After 45 minutes, though, the trend that had been observed did not appear to continue. Instead of more replication factors associating with the complex as might be expected of a replisome transitioning to a fully functional state, few factors showed an increase in enrichment over previous time-points, and the number of factors significantly enriched dropped markedly. The

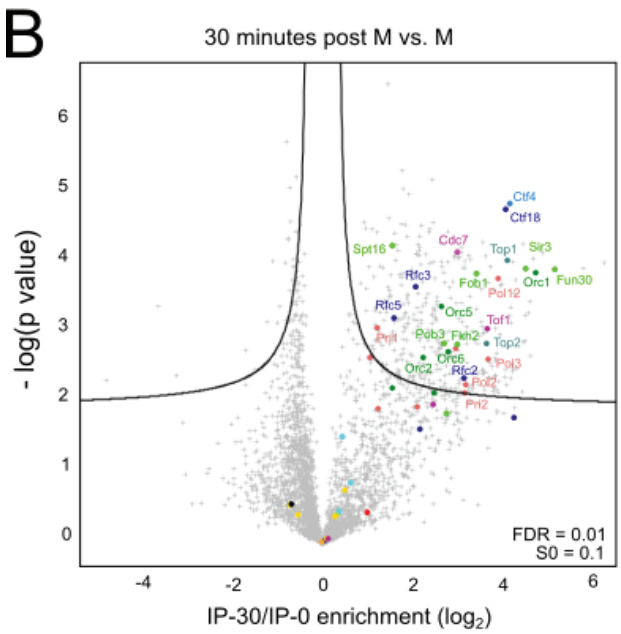
flow cytometry data indicated that at 60 minutes, a large proportion of the cell population were in S phase (Figure 44), which should have meant a large number of active replisomes. MCM2-7 enrichment remained high, but other, more peripheral, replication factors were seen to be much less reliably enriched at this and later time-points (Figure 47D-F). A reduction in enrichment of replication factors was expected at the later time-points as termination events occurred, reducing the number of active replisomes, but 60 minutes was considered too early for termination events to have much of an effect.

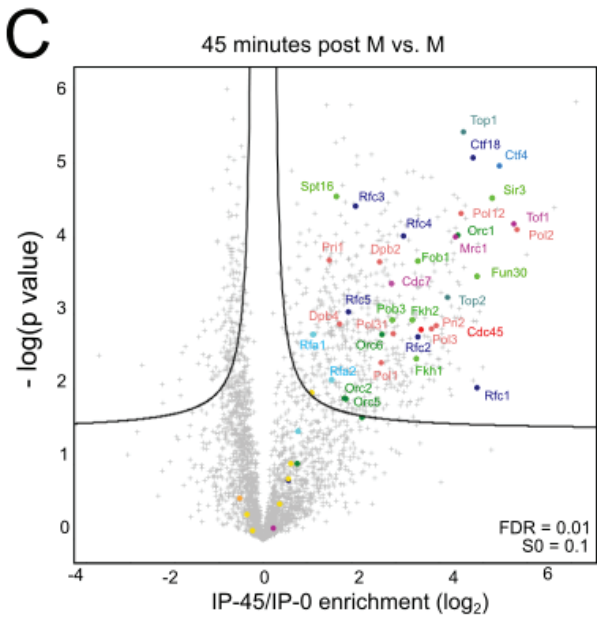
Despite the unexpected results from the latter three time-points, further investigations were carried out into both the composition of proteins in the pull-downs at each time-point, and also how the time-points compared to each other in this regard (Figure 48, Figure 49, Figure 50).

By simply ranking the detected proteins by enrichment values as done for previous experiments, a few interesting observations could be made. Firstly, the consistency of MCM2-7 and Cdt1/Tah11 enrichment could be seen across the time-points (Fig. Figure 48A). The number of replication associated factors in the most enriched proteins could be seen to increase over the time-course, reflecting the progression from unloaded MCM2-7 hexamers, through to stable double hexamer, and finally to active replisomes, possibly ending with terminating complexes. However, there were overall fewer expected replication factors making the most enriched protein list than previous experiments. The presence of certain protein factors in this list appears to chart the progression of the complex over time. Ctf4, Cdc45 and Tof1 all show greatest enrichment values at 60 minutes, when the helicase complex is part of an active replisome, and the enrichment values decline as the time-course progresses, potentially indicating loss of replisome complexes to termination events. As mentioned previously, a surprisingly high enrichment of Ctf4 was observed at 45 minutes, when only a small amount of DNA replication was detected (Figure 44C).

A

0 min		30 min		45 min		60 min		75 min		90 min	
AIR1	10.57671	AIR1	10.0039	AIR1	9.183896	AIR1	9.723939	AIR1	9.587325	AIR1	9.841744
MCM5	8.871952	MCM5	8.425967	MCM5	8.94683	MCM4	8.00444	MCM4	7.818896	MCM4	8.203649
MCM4	8.284834	MCM4	8.30077	MCM6	8.820648	MCM6	7.77107	MCM6	7.53418	MCM6	7.974918
MCM6	8.230289	MCM6	8.225218	MCM4	8.352832	MCM5	6.772272	MCM5	6.478201	MCM5	6.000811
TAH11	8.220623	TAH11	7.629196	MCM7	6.201183	TAH11	5.931971	TAH11	5.278257	TAH11	5.409407
DAL81	6.448083	IMG2	5.800765	TAH11	5.931971	MCM7	5.174997	CTF4	5.202788	PRP3	4.940909
IMG2	6.083918	MCM7	5.631928	MCM3	5.218703	TAH11	4.993068	MCM7	5.153045	MCM7	4.900706
SAT4	5.373078	FUN30	5.562986	FKH1	5.036692	IMG2	4.88968	SAT4	4.992757	DAL81	4.871984
MCM7	5.163638	POL1	5.329745	TAF6	4.962379	DAL81	4.679629	CTF18	4.737986	SAT4	4.798641
MCM2	5.143906	PRP3	5.280904	RSC58	4.922735	TOF1	4.585746	POL2	4.718222	IMG2	4.657017
YAT2	5.121212	SAT4	4.674514	MCM2	4.921628	MCM3	4.367381	IMG2	4.705167	OAF1	4.611765
OAF1	5.031761	CLF1	4.651258	CTF4	4.852908	SAT4	4.326339	SIR3	4.365707	RBL2	4.553927
FZO1	4.922724	MCM2	4.650737	UTP20	4.830104	OAF1	4.324399	MCM3	4.311258	CTF4	4.420838
XKS1	4.059232	SIR3	4.592239	PRP3	4.719168	CDC45	4.119712	TOF1	4.283641	MCM3	3.829965
MCM3	3.777893	OAF1	4.55723	INO80	4.566599	MCM2	4.079443	DPB2	4.072387	BUD14	3.672519
GDB1	3.526654	MCM3	4.500359	POL2	4.565005	ERR3;ERR	3.798054	MCM2	3.688687	MCM2	3.540301
SMI1	3.379212	FZO1	4.431281	FUN30	4.546299	XKS1	3.775801	CDC45	3.635557	POL2	3.52463
RIC1	3.241543	INO80	4.430202	SIR3	4.54615			PRI2	3.437867	CDC45	3.475126





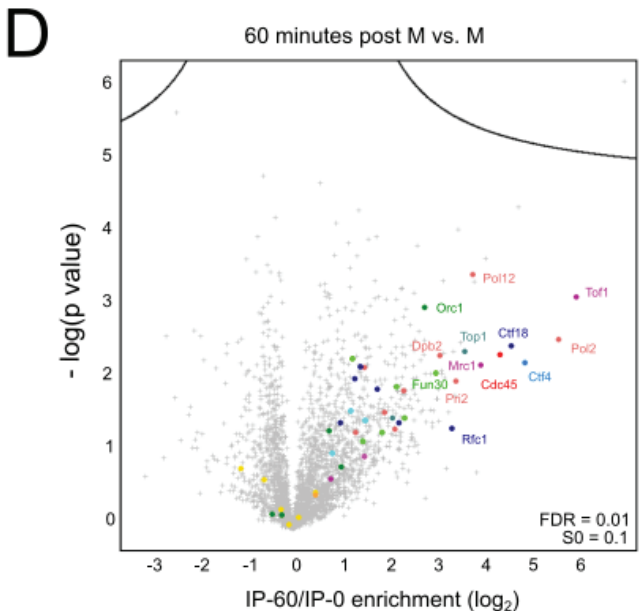


Figure 48: Enrichment analysis of proteins in Mcm4-FLAG ChIP-MS across different time-points post-mitotic-arrest. A)Tables showing the most enriched proteins when comparing tagged pull-downs against the untagged negative controls for each time-point. Proteins that are part of the manually curated set of replication factors are coloured according to the colour scheme used in prior figures. Enrichment values are the mean enrichments across four replicates. Where a protein was not detected in the untagged control, conservative imputation was used to assign a value equal to the mean of the lowest 10% of detected proteins in the untagged control. B-D) Enrichment analysis of proteins in Mcm4-FLAG ChIP-MS comparing tagged pull-downs against tagged pull-downs from different time-points. Protein enrichments between tagged samples were plotted against the p-value from a permutation-based two-sample t-test carried out in the Perseus software, using an FDR of 0.01, and an S0 value of 0.1. Proteins are coloured according to the previously used schema. B) Volcano plot showing enrichment against p-value for the comparison between Mcm4-FLAG cells at 30 minutes post-mitotic-arrest against Mcm4-FLAG cells at 0 minutes post-mitotic-arrest. C) As in B, but for 45 minutes post-mitotic-arrest against 0 minutes post-mitotic-arrest. D) As in B, but for 60 minutes post-mitotic-arrest against 0 minutes post-mitotic-arrest.

By comparing pull-down protein intensities directly and conducting a t-test between them, volcano plots could be constructed that show the protein composition changes directly between time-points (Figure 48B-C). This has advantages over the comparison of enrichment values. It is clear from multiple tests of ChIP-MS and even the replicates within this large scale experiment, that the variability of the untagged controls is large. Both the protein composition and protein quantity can be inconsistent. By comparing the pull-down samples directly, this source of variability can be removed.

By comparing the protein quantification at 30 minutes post-mitosis release to the quantification at the 0 minutes time-point (Figure 48B), the increased association of many of the factors identified from enrichment analysis (Figure 47A and B) could be confirmed. Most of the subunits of ORC1-6 showed significant enrichment at 30 minutes over 0 minutes at an FDR of 0.01, and so too did most of the polymerase α subunits, as indicated previously (Figure 47A-B). The increased Mcm4 association of Ctf4 by 30 minutes was also confirmed to be significant, as was the association of Tof1. Cdc7 was also shown to be significantly enriched at 30 minutes over 0 minutes. This is interesting, as Cdc7 is one subunit of the two-protein DDK kinase which is responsible for MCM2-7 phosphorylation to initiate initiation factor binding. Dbf4, the other subunit, was not detected in the mass spectrometry. This might indicate the potential for MCM2-7 activation from thirty minutes onwards, and possibly explain some of the association of later replication factors. Several subunits of the RFC complex were also significantly enriched at 30 minutes over 0 minutes. These proteins are

responsible for the loading of the Pol30 clamp on daughter strands behind the active replisome. The enrichment of these factors by 30 minutes is surprising, as they would potentially indicate an active replisome. While it is possible that some small proportion of complexes initiate prior to this timepoint, and because of the scarcity of these complexes, the small protein Cdc45 is not detected in the mass spectrometry, this data could also represent evidence of pre-association of certain complex members.

The same analysis carried out on the 45 minute time-point, comparing it to 0 minutes yielded more confirmatory results to the enrichment comparisons (Figure 48C). A greater number of polymerase components and Cdc45 were now significantly enriched. The particularly high enrichment of Pol2 and Cdc45 was indicative that at this time-point, some replisomes had begun to be activated. Overall, this data mirrored that of the IP/NC enrichment analysis, suggesting that the variability of untagged controls was not having a consequential effect on the visibility of the overall trends.

Comparing the 60 minute and 0 minute pull-downs directly yielded a confusing result. Under the same FDR and S0 constraints, none of the replication associated factors tested showed significant enrichment in 60 minutes over 0 minutes. This was highly surprising, but did suggest that the unexpected results of the enrichment analysis (Figure 47D) reflected the underlying data. From the PCA analysis, it could be seen that 60 minute pull-down from replicate 4 showed separation from the other three samples at this time-point. However, conducting the analysis again without this sample did not improve the analysis a great deal. Considering only three 60 minute replicates, and a relaxed FDR of 0.05, only Pol2, Tof1 and Ctf4 were significantly enriched. Enrichment of replication factors could still be observed (Figure 48D), but did not reach statistical significance.

To explain why complexes at such diverse stages of the cell cycle could appear so similar, it was considered whether the proportion of soluble MCM2-7 complexes isolated could be

underlying this effect. Since at 0 minutes, the majority of MCM2-7 complexes were soluble and not associated with the replication machinery, any sample that pulled-down a large proportion of soluble non-chromatin associated MCM2-7 complexes would on the whole resemble the 0 minute sample. The enrichment of replication factors caused by chromatin-associated MCM2-7 could be insignificant if a low enough proportion of these complexes was isolated.

The volcano plots from the 75 and 90 minute time-points appeared similar (data not shown). Termination events prior to 75 and 90 minutes could explain the loss of chromatin-bound MCM2-7 in the pull-down for these samples, but the behaviour of these samples was similar enough to that of the 60 minute sample to suggest a common explanation for these timepoints. Pol2, Tof1 and Ctf4 were tending to enrichment in even the 90 minute timepoint, but did not reach the level of significance. This suggested that replication complexes were present as a subpopulation in all three of these later time-points, but were disproportionately few compared to the soluble MCM2-7 that was being pulled-down.

Whilst it was not clear what was causing this effect in the data at these later timepoints, three hypotheses were suggested from the data and the experimental method. Firstly, the reduction in significant enrichment could be caused simply by an increase in the variability between samples. Since the later time-points were further away from the synchronous arrest stage, population drift in cell cycle stage could be responsible for increased variability of protein pull-down. This was a tempting hypothesis, but flow cytometry data showed a large proportion of cells in S phase at 60 mins (Figure 44C), and this should create an abundance of replication complexes that would be expected to cause a significant enrichment of replication proteins despite slight cell cycle heterogeneity in the cell population. Secondly, it was hypothesised that there could be a mechanism by which chromatin-bound MCM2-7 was eliminated from the pull-down in favour of the soluble MCM2-7. It was possible that a high level of cross-linking to

DNA, coupled with inefficient sonication, could result in a loss of chromatin bound MCM2-7. This matched well with the results of these later time-points, but did not easily explain why this effect did not negatively impact the 30 and 45 minute time-points as well. Whilst not conclusive, it was suggested that this difference might be due to the cross-linking to ssDNA present in the active replisome, but not in double hexamer. If excessive cross-linking and inefficient chromatin fragmentation was to blame, then MCM2-7 cross-linking to ssDNA should have been affected more, due to the preferential cross-linking of ssDNA over dsDNA. Thirdly, it has to be acknowledged that termination of replication forks does not occur uniformly at the end of S phase, but stochastically whenever two opposing forks meet. In some regions of the genome, many early-firing origins are clustered close together. It is clear that from even very shortly after release into S phase, a population of replisomes will be terminating/have terminated. In fact, most termination likely occurs quite rapidly after entry into S phase, an effect that is exacerbated by the similar initiation timings of early and late origins that was observed in the Cdc20 release (Figure 42C). Unloading of MCM2-7 from DNA during termination creates a large population of soluble pull-down targets that are likely to out-compete chromatin-bound complexes. By the later timepoints, this effect is likely to be very evident, and is therefore the most likely reason for reduced replisome pull-down in the timepoints past 60 minutes.

Regardless of the cause of the issue, the latter three timepoints of the experiment clearly showed issues with either the reliability of the pull-down, or the proportion of chromatin-bound MCM2-7 present, and so any analysis on these samples should be considered with some scepticism.

For those proteins that were significantly enriched in the tagged pull-down over the untagged control in the first three time-points, gene ontology enrichment analysis was carried out (Figure 49). All three time-points showed enrichment of GO terms relating to DNA replication, with

the enrichment significance being much greater in both the 30 and 45 minute time-points when compared to the 0 minute timepoint (Figure 49, compare panels B and C to panel A). The enrichment shown for DNA replicative processes at the 0 minute time-point is likely mainly due to the presence of the MCM2-7 and Cdt1 proteins, which are reliably enriched. At the two later time-points, the presence of multiple additional replication proteins is likely the cause of the increase in significance.

To compare how the set of replication-associated proteins behaved over the course of the time-course, the average enrichment values of this set of factors across all four replicates was plotted for all time-points (Figure 50). In comparison to previous trials of the Mcm4-FLAG ChIP-MS, the detection of these factors in the mass spectrometry was much less complete. As previous figures show, the pull-down of MCM2-7 subunits was efficient in this experiment, but the pull-down of certain other replication factors was impaired comparatively. Notably, the initiation proteins Sld3, Sld3 and Dpb11 were not detected at any time-point in any sample, and Sld7 and Mcm10 were only detected at the 45 minute time-point in the tagged sample. Most concerningly, of the GINS complex, only Psf1 was ever detected across all samples and timepoints.

As previously noted, the behaviour of Cdc45 and Ctf4 roughly followed the expected progression of the complex, increasing from very low enrichments at 0 minutes, through to high enrichment at 45 minutes, before slowly reducing in enrichment through 60, 75 and 90 minutes. Of the fork protection complex proteins, Tof1 and Mrc1 followed this trend, but their complex partner Csm3 did not show a similar trend. In this case, the levels of Csm3 detected in the tagged sample were low enough that the conservative manual imputation caused this protein to display a negative enrichment value. Psf1, the only GINS member detected, was also affected by this same issue, showing negative enrichments throughout. A trend similar to that noted for Cdc45, Ctf4, Tof1 and Mrc1 was also seen for several of the subunits of polymerase

ε, though Dpb3 and Dpb4 often were only detected at low abundances, and were also affected by imputation effects. The early association of polymerases α and δ with MCM2-7 is still visible in this visualisation, showing increases in enrichment prior to those shown by the other replication factors, highlighting the possible pre-association of these factors.

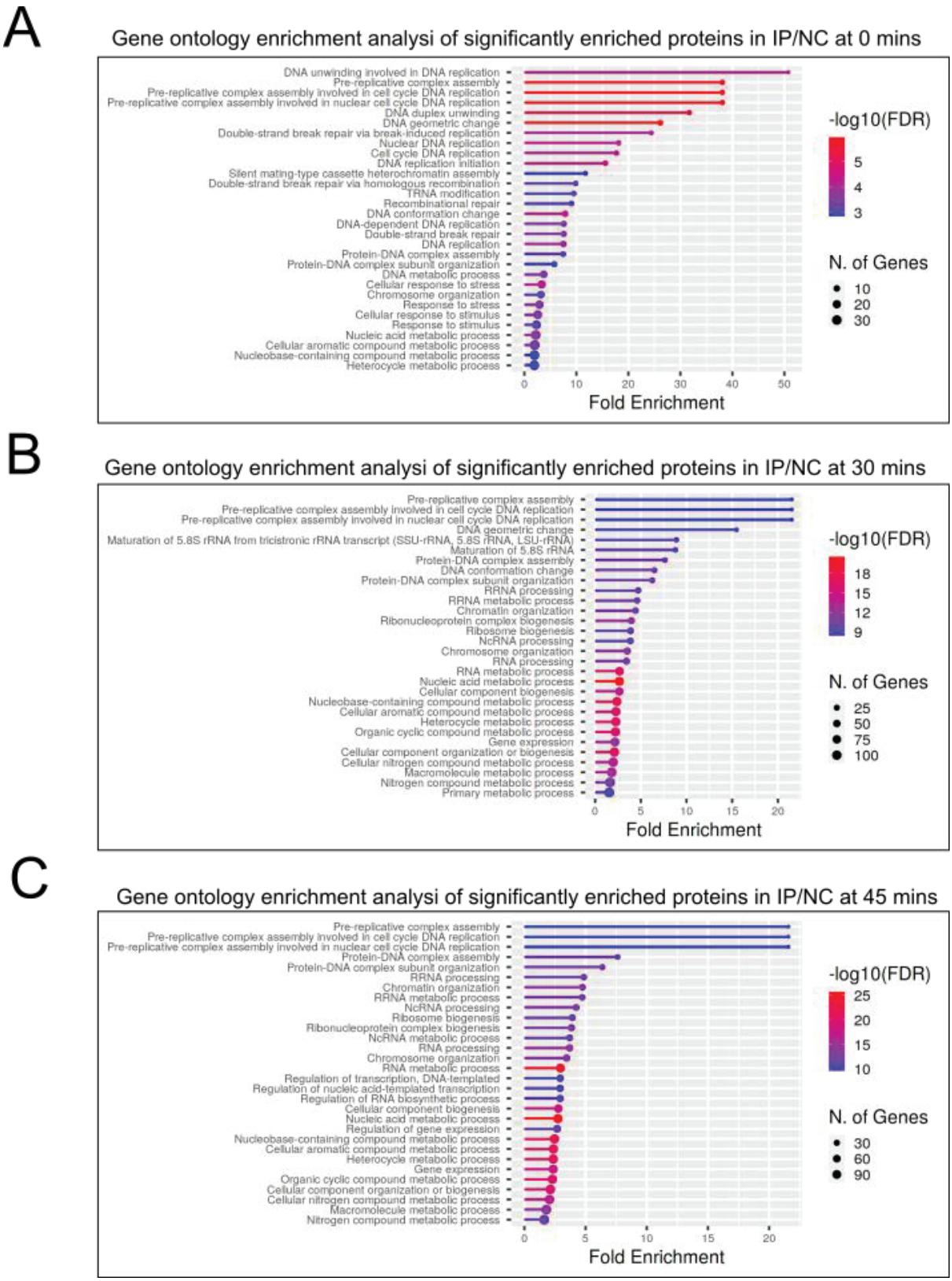


Figure 49: Gene ontology enrichment analysis for the first three time-points. T-test significantly enriched proteins between tagged and untagged pairs were compared against the total proteome to identify enriched biological processes (Ge et al., 2020). A) Analysis of the 0 minute time-point. B) Analysis of the 30 minute time-point. C) Analysis of the 45 minute time-point.

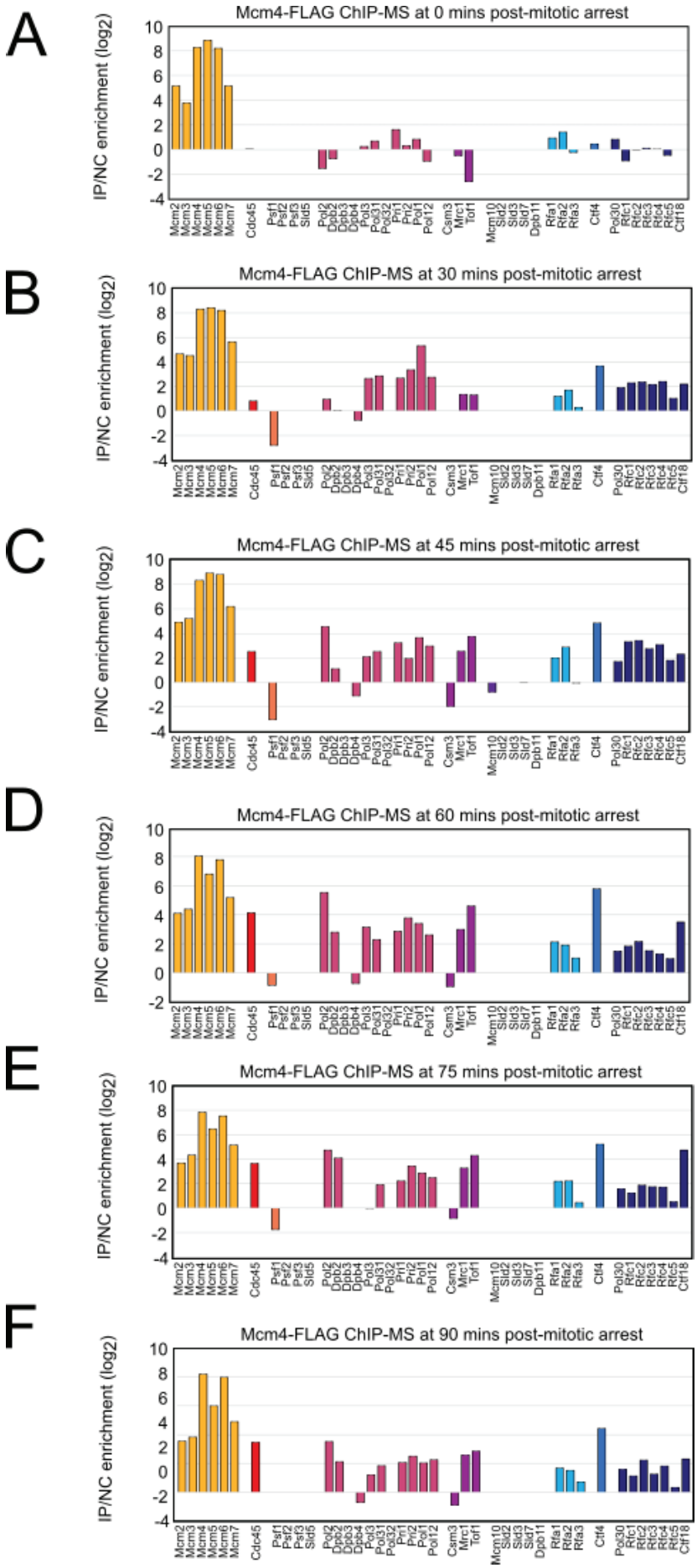


Figure 50: Enrichment analysis of a core set of replication-associated factors.

Enrichment values were calculated using average LFQ values from four replicates for each time-point. Where proteins were not detected in the untagged negative control, a conservative manual imputation was performed as previously. A) Enrichment analysis for cells arrested immediately after mitotic-arrest release. B-F) As in A, but for each time-point tested: 30, 45, 60, 75 and 90 minutes post mitotic-arrest release.

Clearly, this experiment showed different results than in previous trials. The lack of detection of the GINS complex and most of the transient initiation factors was disappointing. It was also not entirely apparent why mass spectrometry analysis of this pull-down was not able to detect these proteins when it had with consistency before. It is possible that high levels of cross-linking and insufficient sonication could explain both the issues with the later time-points and the lack of certain replication factors. An underrepresentation of chromatin-bound MCM2-7 complexes would impact the overall levels of replication proteins isolated. GINS subunits are among the smallest proteins in the replication factor set, and so are already less easy to detect by mass spectrometry, especially if present at low levels. Cross-linking artifacts preventing peptide recognition will only exacerbate this issue. The lack of detection of GINS and other factors could have been because they were at a low enough abundance in the sample to drop below the detection threshold of the MS instrument.

Because calculated enrichment values, and specifically the ones utilising imputation, were generating misleading results, it was decided to investigate complex composition and protein abundance over the time-course looking only at the LFQ normalised intensity values for individual protein factors. Whilst not as quantitative as an approach like SILAC or TMT labelling, LFQ normalisation can produce reliable normalisation of protein abundance. Comparing LFQ intensity of individual proteins should not be used in isolation, but with information gained from more statistically controlled methods displayed above, some conclusions could nonetheless be drawn from this dataset.

Firstly, the average LFQ intensities for the six MCM2-7 subunits over the time-course do decrease (Figure 51A). This would appear to fit the hypothesis that a subset of MCM2-7 complexes are not released from cross-linked chromatin, and so are “lost” from the pull-down. This trend was consistent for all subunits, so indicated a general effect for intact hexamers. In particular, the trend for Mcm7 did not display a markedly different profile. Because Mcm7 is

the subunit that is ubiquitinated by SCF^{Dia2} and then targeted by the Cdc48 segregase (Moreno et al., 2014; Maric et al., 2014, 2017), it might have been expected that Mcm7 abundance decreased if termination was occurring. Instead, Mcm7 closely matched the profiles of the other subunits, even at ninety minutes. This held true for both calculated enrichment values and LFQ intensities (Figure 52).

Calculating average LFQ intensities of other factors also yielded results consistent with the previous analysis. The abundance of Cdc45 within the pull-down samples followed the expected pattern (Figure 51B). Initially, very little Cdc45 was detected, but once MCM2-7 was loaded at origins at 30 minutes and beyond, Cdc45 abundance levels in the pull-down increased. The maximum abundance was detected at 60 minutes, which would agree with flow cytometry data suggesting that this time-point contains the most cells actively in S phase (Figure 44C). After this point, the abundance levels decreased, which would correlate with the termination of replication and disassembly of CMG complexes.

Similarly, by looking at the dynamics of ORC abundance in the pull-down, an easily explained pattern was observed (Figure 51C). Levels were low at 0 minutes, since MCM2-7 was not chromatin-associated. However, once ORC-dependent MCM2-7 loading at origins begun, the association of ORC with MCM2-7 increased. ORC abundance began to decrease immediately after its peak association at 30 minutes, due to two interlinked possible causes. Loading reactions, responsible for the direct structural interaction of ORC and MCM2-7, ceased, and MCM2-7 double hexamer occupancy at origins displaced ORC complexes from their binding sites. By the end of DNA replication as visualised by flow cytometry, ORC-MCM2-7 association as measured by ORC1-6 abundance in this IP, were once again much reduced to around the mitotic level.

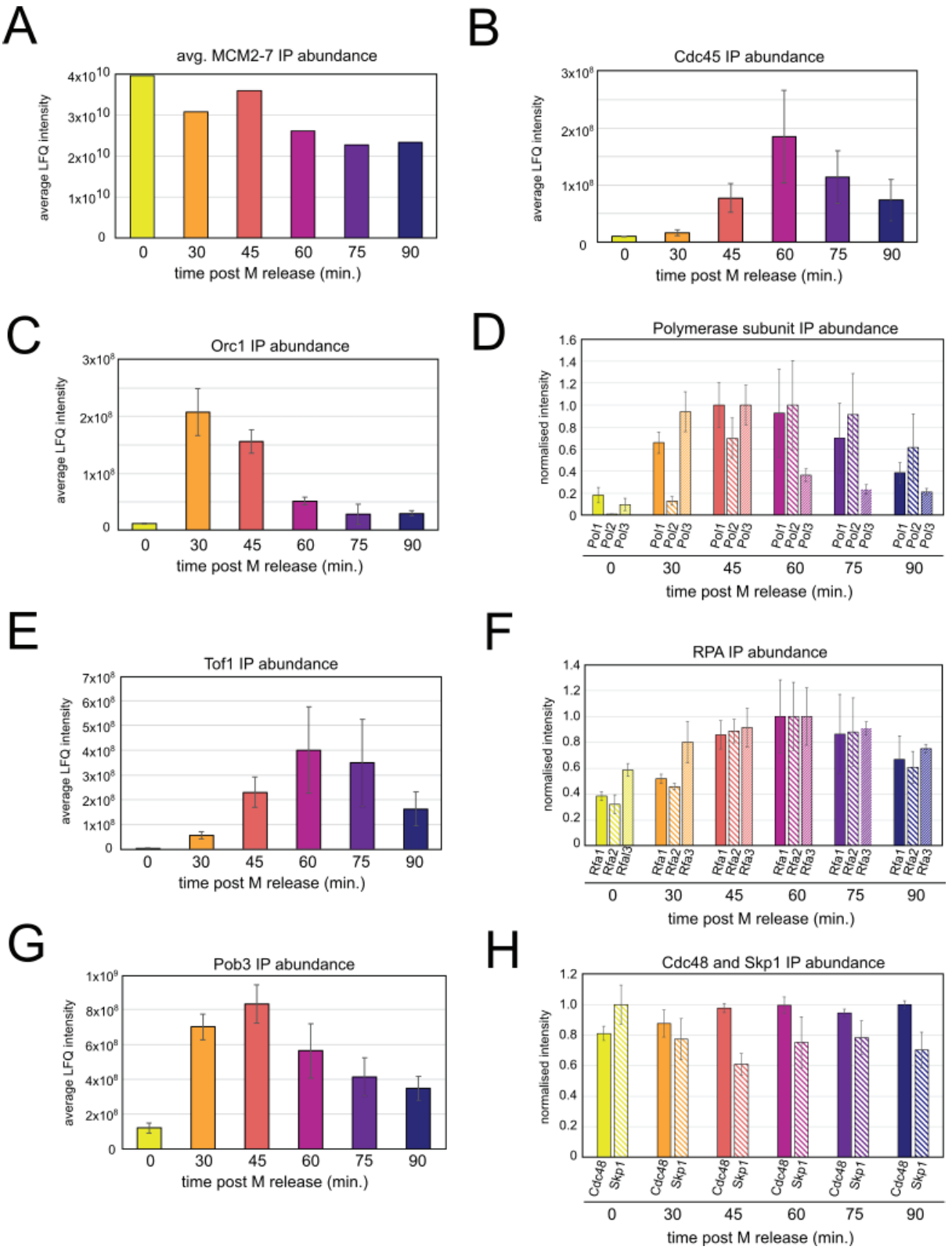


Figure 51: Bar plots displaying the averaged normalised intensities for individual proteins at each time-point. The mean of LFQ values from the four replicates was calculated. Where displayed, error bars indicate the standard error in the mean (s.e.m.). A) Plot showing the average mean LFQ intensity for the MCM2-7 subunits at each time-point. B) Plot showing the mean LFQ intensity of Cdc45 over the time-course. C) As for B, but for the origin recognition complex subunit Orc1. D) As for B, but showing the normalised mean LFQ intensity of the larger subunits of the three main replicative polymerases, Pol α, Pol ε and Pol δ respectively. Intensity means were normalised to maximum values in the time-course for visual clarity and comparison with the other polymerases. E) As for B, but for Tof1. F) As for D, but for the three subunits of the ssDNA-binding protein RPA. G) As for B, but for FACT subunit Pob3. H) As for D, but for two proteins involved in helicase ubiquitination, unloading and degradation, Cdc48 and Skp1.

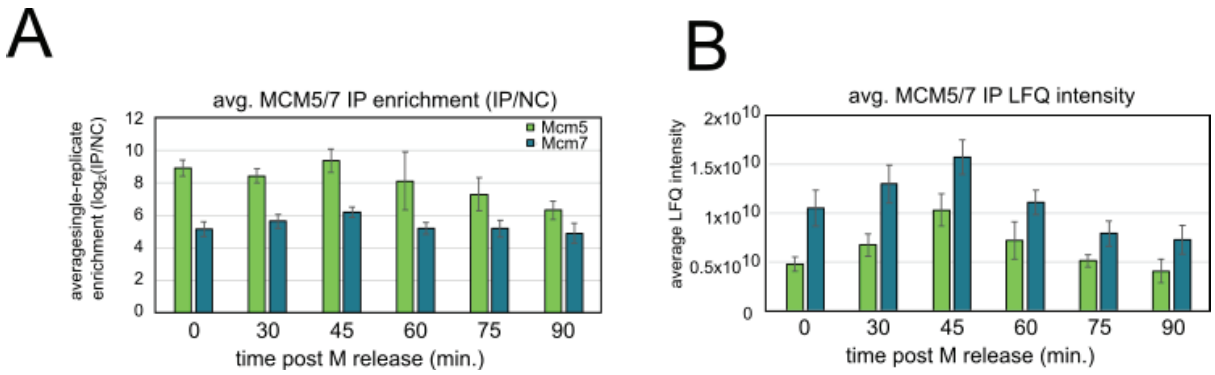


Figure 52: Quantification of MCM2-7 subunits across all time-points and replicates by two separate metrics. A) Calculation of an enrichment value for Mcm5 and Mcm7 was carried out by comparing tagged sample intensity against untagged sample LFQ intensity pairwise for four replicates individually. The resulting average enrichment value is plotted, error bars represent standard error of the mean (s.e.m.). B) The average LFQ intensity for the Mcm5 and Mcm7 subunits in the tagged IP samples across four replicates is plotted. Error bars represent standard error of the mean (s.e.m.).

The intensity variations for other replication factors over this experiment were also visualised. As a measure of polymerase association, the intensity values of the larger subunits of polymerases α, ε and δ, Pol1, Pol2 and Pol3 were compared (Figure 51D). These subunits were chosen because they showed the most reliable intensity values in this ChIP-MS experimental run, as previously mentioned, the smaller subunits displayed some issues with the detection threshold. However, the larger subunits showed robust quantification values here, and so were used as a proxy for the polymerase multi-subunit complexes associations with MCM2-7. Comparing their normalised intensities, it was possible to see a similar pattern to that of Cdc45, with some key differences. For polymerases α and δ, their peak abundance was detected at 45 minutes, suggesting this was the time-point where these proteins showed maximal association with MCM2-7. Whilst the abundance after this time followed a similar trend to Cdc45, likely correlating with the number of active forks, the majority of cells at 45 minutes had either not started or were just initiating DNA replication. Indeed, the rise in association of polymerase α and δ with MCM2-7 was seen even at 30 minutes, at a time where MCM2-7 was being or had only recently been loaded at origins. This was a clear demonstration of the "pre-association" of these factors with the MCM2-7-bound origin, even before activation of the helicase complex. In comparison, polymerase ε showed a more Cdc45-like pattern. This is likely due to

the role of polymerase ε in the activation step of DNA replication initiation, where it is recruited by kinase activity, along with GINS, Sld2 and Dpb11. It was an interesting observation that the other two polymerases show such early association, and should this data prove reproducible, it poses questions about what exactly is causing this association. Are polymerases α and δ recruited to origins prior to replication initiation? If so, what is the mechanisms behind this recruitment? Is it via direct interaction with MCM2-7 or is another factor present at origins responsible?

The fork protection complex comprises Tof1, Mrc1 and Csm3. Csm3 is a very small protein, and as such was often not detected in ChIP-MS experiments. However, Mrc1 and Tof1 were both reliably detected in Mcm4-FLAG pull-downs. Both these factors demonstrated a similar profile to Cdc45 over this time-course (Figure 51E, Mrc1 data not shown) which does suggest that these proteins' abundance in the pull-down does track the formation of an active replication complex.

Abundance investigation of RPA subunits did not yield so clear a result (Figure 51F). Not only were the raw LFQ intensity values very different between the three subunits, but the abundance of these factors was already relatively high at 0 minutes. This indicated the background noise level of this factor was high, or that there was some RPA-MCM2-7 interaction in solution, which was considered unlikely. Whilst the general trend for this factor mirrored that of other replication factors, this high background meant that whilst usually this factor can be used to reliably indicate the presence of ssDNA, in this experiment it most likely did not indicate this.

Another factor that showed a perhaps surprisingly early association with the MCM2-7 complex was the FACT complex (Figure 51G). The FACT complex is required for progression of the replication fork *in vivo* and travels with the replication fork (Zhang, Feng & Li, 2020; Kurat et al., 2017; Gambus et al., 2006). Studies in human and yeast cells indicated that the MCM2-7 FACT interaction was important for FACT recruitment to origins, and that disruption

of this interaction delayed DNA replication initiation (Han et al., 2010; Tan et al., 2006). Interestingly, in this time-course, Pob3, one subunit of FACT showed a high increase in abundance from 0 to 30 minutes. This indicated association with MCM2-7 that was in some way regulated by MCM2-7 origin association. Either FACT was already present at origins and MCM2-7 loading subsequently caused in increased association, or origin-bound MCM2-7 recruited FACT. Consistent with previous literature, the abundance of FACT remained high throughout replication, showing the same decrease over time seen for other replication factors.

Confident that by analysing the data on the individual protein abundance level yielded results that confirmed previous data, and could provide new insights, this approach was continued for a further set of factors. This experiment aimed to characterise the complex composition of the MCM2-7-containing replisome over the life cycle of the helicase. ChIP-MS had not previously been attempted to investigate terminating complexes, as work in the lab focusses mainly on initiation. A key step in termination is the ubiquitylation of Mcm7 by SCF^{Dia2}, followed by Ufd1/Npl4-guided Cdc48 mediated disruption of the replicative helicase complex (Xia, 2021; Maric et al., 2014; Moreno et al., 2014). Whilst SCF^{Dia2} was thought to travel with the replication machinery, it might have been predicted that Ufd1, Npl4 and Cdc48 might show increases in intensity towards the later time-points. Dia2 itself was not detected in this ChIP-MS run, but Skp1, a component of SCF, showed fairly constant abundance in the pull-down, even at 0 minutes (Figure 51H). Care must be taken with these results, since they do not take into account abundance in the untagged control, and indeed, the LFQ intensity levels of Skp1 in the pull-down are only roughly two-fold greater (data not shown). Similarly, Npl4 and Ufd1 showed low enrichments over the negative control, and fairly constant intensity values. Cdc48 appeared to show a slight increase in association over the course of the experiment, but it too showed minimal enrichment, suggesting the abundance observed was simply non-specific background. This would fit with the hypothesis that SCF^{Dia2} associates only once CMG is

released from chromatin, and previous interactions identified in the literature (Morohashi, Maculins & Labib, 2009; Maculins et al., 2015) were artifacts of the isolation protocol. In contrast, the ChIP-MS technique is not so susceptible to post-isolation interactions due to the temporal restriction of the activity of cross-linker, and so this data may represent the true, very minimal, levels of SCF^{Dia2} interaction with the replisome. In addition, SCF^{Dia2} regulates many processes during S phase and may have many binding partners, making the amount binding to the replisome at any given time minimal

From the enrichment analysis, several factors were identified that either had been recently linked to replication, or could feasibly have a previously unidentified role in replicative processes. A factor that showed consistently strong enrichment was Fun30, a member of the SNF2-family of chromatin remodellers. This factor was also not unique to this single experiment, it had been observed in multiple other ChIP-MS experiments, which increased the confidence in its genuine association with MCM2-7. While yeast Fun30 had previously been implicated in DNA damage pathways, a recent study of the human homolog SMARCAD1 was the first literature outside of this study to suggest that Fun30 might be a component of the replication fork (Lo et al., 2021; Costelloe et al., 2012; Terui et al., 2018; Adra et al., 2000; Bi et al., 2015; Eapen et al., 2012). To investigate how Fun30 might be functioning in the yeast system, the changes in abundance over the time-course were compared with other replication factors (Figure 53A). Interestingly, the dynamics of Fun30 did not reflect the profile of Cdc45, Tof1 and other factors associated specifically with the active fork (compare Figure 51B with Figure 53A). Instead, Fun30 showed the greatest abundance in the pull-down at the 30 minute time-point, when MCM2-7 was recently associated with origins. The increase in association from 0 minutes was very clear, so the Fun30 association with MCM2-7 was chromatin-dependent. After this timepoint, Fun30's MCM2-7 association was maintained to a certain extent, but did decline faster than other fork factors. Given Fun30's previously ascribed activity

as a chromatin remodeller (Strålfors et al., 2011; Chen et al., 2012; Awad et al., 2010; Siler et al., 2017; Bantele & Pfander, 2019), this behaviour was very interesting. This data suggested that instead of travelling directly with the fork, which may still occur to a lesser extent than for other fork factors, Fun30 may instead have a role in modifying origin chromatin environments to modulate replication-associated processes. Fun30 could potentially act at the stage of MCM2-7 loading, which may require nucleosome repositioning to accommodate the double hexamer at origins, or later, at the stage of activation, ensuring access of replication factors and remodelling chromatin to allow the generation of the forks themselves. Compared to other

replication factors, Fun30 association actually resembled ORC dynamics more than fork factors, indicating that perhaps the role in loading was more likely.

The quantification of Hrq1, another protein that had shown enrichment in ChIP-MS experiments in this study, most notably in artificial arrest states (see Section 3.7.1), followed the trend set by Fun30, displaying a clear increase in abundance from 0 to 30 minutes, before levels gradually declined over the time-course (Figure 53B). In contrast to Fun30, the possible role of Hrq1 at origins at this stage of the cell cycle is not as clear. Hrq1's role in yeast has previously been looked at in the context of inter-strand cross-link repair (Rogers et al., 2020,

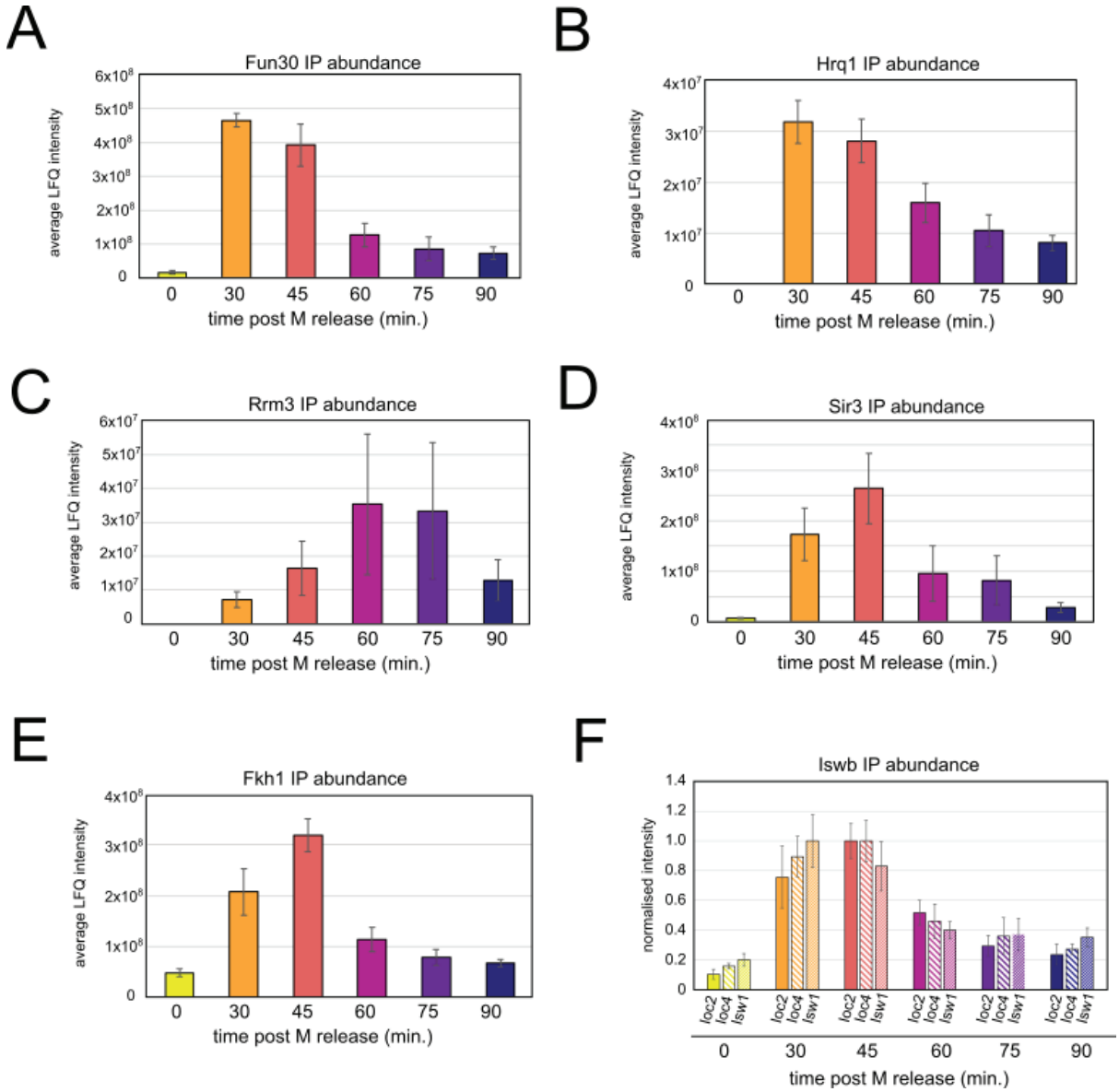


Figure 53: Quantification of average LFQ intensity over the time-course for a range of factors. Mean LFQ values from four replicates were calculated, and error bars represent the standard error of the mean (s.e.m.) of this value. A) Quantification of Fun30 intensity over the time-course. B) As in A but for Hrq1. C) As in A but for Rrm3. D) As in A but for Sir3. E) As in A but for Fkh1. F) Comparison of the normalised mean LFQ intensities of the Isw1b complex, error bars represent the standard error of the mean (s.e.m.) and values are normalised to the highest mean LFQ intensity over the time-course, separately for each factor.

2017; Bochman et al., 2014; Röhrig et al., 2018), and there is no obvious rationale as to why this activity would be required at origins prior to initiation. Regardless, the intensity values were consistent, and Hrq1 was not detected in the untagged controls in this experiment at any time-point, giving strong confidence in this protein's association.

In contrast to the previous two factors, Rrm3, another protein of interest that was investigated in more detail here, showed abundance changes that more closely resembled those of a fork factor (Figure 53C). Considering previous literature concerning this protein, this was not in fact surprising. Not only has Rrm3 been shown to associate with replication forks (Schmidt, Derry & Kolodner, 2002), it has been shown to have a role at sites of replication fork pausing where it has a counteracting activity to Tof1 (Ivessa et al., 2002; Torres, Schnakenberg & Zakian, 2004; Mohanty, Bairwa & Bastia, 2006). Whilst Rrm3's intensity changes do indicate that reports of fork association are correct, the higher variability of its abundance could well indicate that the association is not always present, and may rely on the fork stalling to induce association.

Among the factors identified as being of potential interest from this and previous experiments was the protein Sir3. Despite canonical roles in silencing telomeric chromatin (Renauld et al., 1993; Hecht et al., 1995; Buchberger et al., 2008), the Sir proteins have also been implicated in determining timing of origin activation in yeast (Crampton et al., 2008; Hoggard et al., 2018). Interestingly, Sir3 in this time-course shows an intermediate association profile (Figure 53D). It shows a large increase in abundance in the pull-down from 0 to 30 minutes, but does not peak in abundance until 45 minutes, in contrast to Fun30 and Hrq1 which show greatest abundance at 30 minutes. The abundance also decreases faster than for the for the genuine and suggested fork factors Cdc45, Tof1 and Rrm3. Of particular interest to the lab and the aims of this project, Sir3 shows the greatest abundance at the time-point where replication initiation can be seen to be occurring in the cells. It has a profile that is highly similar to that of Fkh1

(Figure 53E), a known origin timing regulator (Knott et al., 2012b; Reinapae et al., 2017; Zhang et al., 2019).

Further to these factors that appear to be involved with specification of origin timing, and have association dynamics that appear to show maximum MCM2-7-factor association at the point of S phase entry, the Isi1b complex was also investigated by this approach (Figure 53F). This complex was seen to show a robust association profile that matched Sir3 and Fkh1, potentially implicating this protein complex in similar activities. As a chromatin remodeller, Isw1b could well be involved in processes regulating chromatin accessibility to allow helicase activation (Eriksson & Clark, 2021).

The above investigation showed that despite issues within this large-scale experiment, the LFQ intensity values produced across the time-course can still be used to infer information about possible roles of replication-associated factors. With appropriate caution, especially considering the abundance in the untagged control, factors that had no previous literature implicating an involvement with replication could well be identified for future investigation to confirm.

Because the intensity profile proved useful, further analysis was done using profile plots and Pearson correlation analysis in the mass spectrometry Perseus software (Figure 54). A selected protein's profile plot was used to identify other proteins whose profiles closely matched the selected protein. Given the results in previous section, it was hypothesised that the profile plot of a protein can be used to infer association dynamics of proteins with MCM2-7. As a test of principle, this analysis was performed on Mcm2 and Cdc45. If the analysis was able to identify factors that covary with these proteins that are also known replication proteins, then the results for the other factors of interest may generate some interesting proteins for further investigation. The top ten factors identified in this manner are displayed in Figure 54 for all the factors tested. The top ten factors for Mcm2 contain four of the remaining five subunits of the MCM2-7

complex, Cdt1/Tah11, and four other proteins that from a brief literature search appear to be highly-expressed non-specific dirt (Figure 54A). Because the MCM2-7 complex contains the bait protein of the pull-down, its variance within this assay is not as informative as other factors, such as Cdc45. The top ten factors correlating with Cdc45's profile are all likely members of an active replisome, with the exception of Pdc6, which is likely a non-specific highly expressed background protein (Figure 54B, note the highest orange track starting at the left side of the graph, which does not vary too much over the time-course and shows the abundance of Pdc6). Repeating the analysis with Fun30, Rrm3 and Sir3 showed different results in each case. The majority of the proteins that best matched the Fun30 profile over the time-course were involved in either transcriptional control and post-processing, chromatin regulation or ribosomal synthesis and export (Figure 54C). Notably both Spt3 and Chd1 are members of the SAGA complex, responsible for histone acetylation as part of transcriptional activation (Dudley, Rougeulle & Winston, 1999; Sterner et al., 1999; Tran et al., 2000; Pray-Grant et al., 2005). Ino80 is a member of the ISW1 group of chromatin remodellers (Yen et al., 2012). Dst1, also known as elongation factor TFIIS, is involved in modulating polymerase pausing during transcription (Ishibashi et al., 2014). The E3 ligase Bre1 is involved with a number of chromatin related functions, including ubiquitylation of histones and the DNA damage response (Sen et al., 2013; Kim & Roeder, 2009). Dbp3 and Dbp7 are both involved in rRNA processing (Weaver, Sun & Chang, 1997; Daugeron & Linder, 1998), and Noc2 and Rrp6 both have roles in RNA nuclear export (Hilleren et al., 2001; Hierlmeier et al., 2013). These were quite a diverse range of proteins and functional roles, and it was not clear why these proteins would show similar dynamics to Fun30. Several of these proteins are chromatin associated and function in chromatin modification like Fun30 itself, so could be active in a similar role to Fun30. It was also possible that the association of Fun30 with MCM2-7 at origins occured in specific contexts, such as at sites which show localisation near to highly expressed genes. This

could cause a correlation between these genes generally associated with transcription and Fun30 itself. It is also always possible that the correlation to Fun30 is coincidental, and the matched proteins are simply background factors that by chance show variation over the time-course to match the profile of Fun30. For Fun30, though, the correlated proteins were not able to provide clear additional data.

By contrast, the matched protein profiles for Rrm3 showed many overlapping proteins with the Cdc45 set, and all the most correlated proteins were replication factors (Figure 54D). This dataset clearly and unambiguously grouped Rrm3 with replication factors by its abundance profile.

Sir3 showed a more mixed result for the most well-correlated protein profiles (Figure 54E). Some replication-associated factors were among these proteins, including Top1, Rfc2 and Spt16. This was a slightly different subset of factors to those associating with Cdc45 or Rrm3, which fit with the previous visual distinction noted for this factor (Figure 51D). The other factors, similarly to that observed with Fun30, had a range of canonical activities in nuclear processes. Chd1 and Rpd3 function in histone acetylation and deacetylation (Pray-Grant et al., 2005; Rundlett et al., 1996). Nma11 is a nuclear serine protease with a role in promoting cellular apoptosis. Bdf1 is a transcriptional activator that also shows association with DNA damage repair. Smt3 is a homolog of SUMO, and is required for signalling and chromatin segregation.

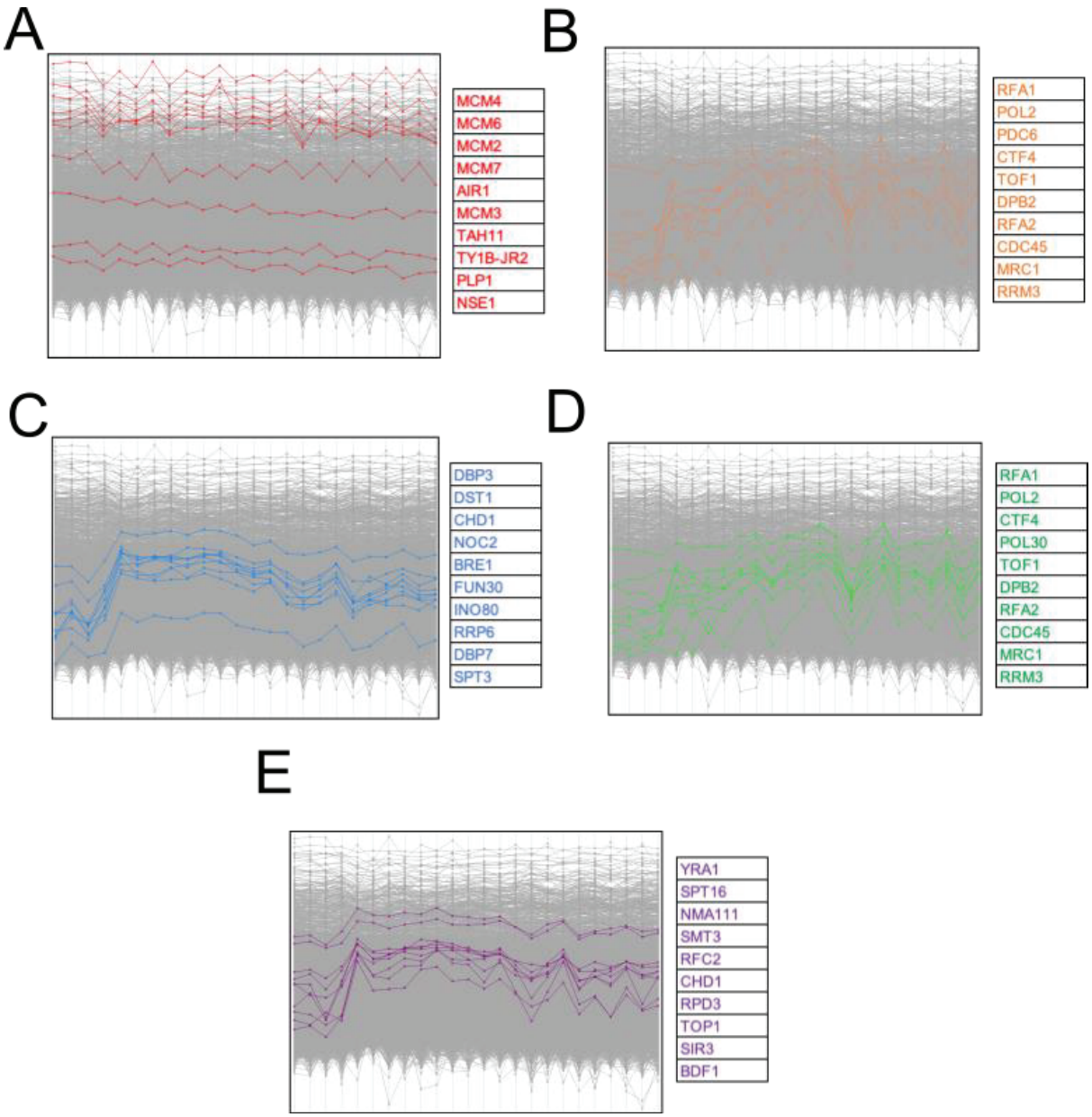


Figure 54: Profile plots of factors of interest in Mcm4-FLAG ChIP-MS time-course. Diagrams show traces of all proteins' LFQ intensities across all conditions and replicates (grey lines). Coloured are the reference profile, and the nine most similar protein profiles identified using a Pearson correlation in the Perseus software package (Tyanova et al., 2016). The list displays all ten proteins coloured, in order of their average LFQ value across all samples, most abundant first. A) Profile plot of Mcm2 and closely matched proteins by profile. B) As in A, but for Cdc45. C) As in A, but for Fun30. D) As in A, but for Rrm3. E) As in A, but for Sir3.

3.7 ChIP-MS to investigate replication complexes arrested during initiation

The ChIP-MS method had been shown to provide a comprehensive and complex interactome of replication proteins, and identified many proteins known to be associated with active forks, as well as some factors with less well characterised roles that showed association with replication complexes. Extension of this method to investigate replication complex composition over time produced an interactome that varied in time, and added temporal information to the association dynamics of both known replication factors, and of less well-categorised factors. However, detection of certain transient initiation factors such as Sld2, Sld3, Sld7 and Dpb11 was not as consistent in the large-scale time-course that was conducted. The lack of detection of other more stably associated replication factors such as GINS suggested that conditions of the ChIP-MS protocol needed further optimisation to be fully consistent from replicate to replicate. This was not the only issue facing this approach, however. Heterogeneity was prevalent in both the cell cycle stage of cell populations, and in the stages of replication complexes themselves. Cell cycle stage homogeneity could be improved given time to optimise the release of larger volumes of cells from synchronous arrest. The heterogeneity in complex composition due to a population of replication complexes at different stages of activation was less easily redressed. Two aims of this study were to investigate the dynamics of transient replication factors during initiation, and identify any novel factors specific to replication initiation *in vivo*. By utilising the rapamycin-dependent linkage constructs discussed in Section 3.1, replication complexes could be arrested at a single step across the cell and origin population. These arrested complexes would then be amenable to examination by the ChIP-MS approach, with a predicted reduction in the heterogeneity. This would potentially improve the detection of both known transient initiation factors, and increase the likelihood of identifying a novel factor in this process.

Previous work had focused on using ChIP-qPCR to analyse the origin association dynamics of replication factors in the presence of a rapamycin-dependent linkage. ChIP-qPCR has proven to be a very useful method in this lab and many others, but in the context of this use case, it exhibited two drawbacks:

Firstly, it was a very narrow approach. Each experimental run required a focus on a single replication factor. This meant that investigating association dynamics of a range of known factors was time-consuming and involved generation of multiple endogenously-tagged strains. It also meant that discovery of novel factors was not within the scope of this type of experiment.

Secondly, the readout of the ChIP-qPCR method was the abundance of particular DNA sequences using PCR quantification. This was highly useful in identifying the association of factors to origins, but the identification of cross-linking effects that seemed to be caused by the generation of ssDNA during activation meant that the exact quantification of factors was influenced heavily by the stage of complex activation. Since ChIP-MS also utilises cross-linking, some cross-linking effects may affect quantification, as observed in previous experiments. However, because the read-out of the ChIP-MS is peptide abundance, the cross-linking of DNA to protein should be a less crucial factor.

3.7.1 Mcm4-FLAG ChIP-MS in the context of the DNA exit gate constructs

At the inception of the proposal to investigate arrested complexes using ChIP-MS, the most well characterised rapamycin-dependent linkage constructs were the Mcm2-Mcm5 intra-hexamer ssDNA gates. Extensive ChIP-qPCR analysis had suggested that two versions of the gate, using the N-terminal and C-terminal insertion sites, induced arrest at slightly different stages of initiation (see Section 3.1.1). This provided two possible complex states to investigate using the ChIP-MS approach. In particular, it was anticipated that the mass spectrometry quantification of factors in pulled-down complexes could provide a more accurate measure of

the abundances of factors in each construct's arrest state. Cross-linking effects believed to be caused by production of large amounts of ssDNA in the C-terminal gate construct made comparisons between these complexes by ChIP-qPCR difficult, and so in addition to potentially identifying novel proteins involved in initiation, this method would give an overview of the two arrest states.

Because this experiment was to investigate the process of initiation, the bait protein chosen was Mcm4 endogenously tagged with a penta-FLAG tag on the C-terminus, as used in Section 3.6. Whilst Cdc45-FLAG showed an effective IP as well, the initial association of Cdc45 with the MCM2-7 hexamer was not too temporally distant from the cell cycle stages to be investigated, and so the Mcm4 bait was chosen to provide better consistency between complex stages. The use of Mcm4-FLAG also allowed analysis of the replication complex at a time-point immediately after release into S-phase as a "ground state" control, prior to initiation events.

Since initiation events during S phase were the target of the investigation, there was no need of a release from as far back as mitosis, such as for the Cdc20-mediated arrest used in Section 3.6.2 and 3.6.3. Indeed, the alpha-factor arrest and release had proven to be more reproducible in the lab (see cell cycle variability identified in Section 3.6.3), and a more well-founded technique.

For quantitative mass spectrometry analysis by label-free quantification, the experiment was replicated in quadruplicate, to ensure valid statistical analysis and as insurance if one replicate was in some way compromised during the experiment.

This experiment was carried out as a collaborative effort between L.M. Reuter and C.Weekes, and the initial steps of cell culture, arrest, release, formaldehyde cross-linking, and harvesting were carried out by L.M. Reuter with assistance from C.Weekes. The following steps through to mass spectrometry sample submission and then analysis of MS data was carried out by C.Weekes.

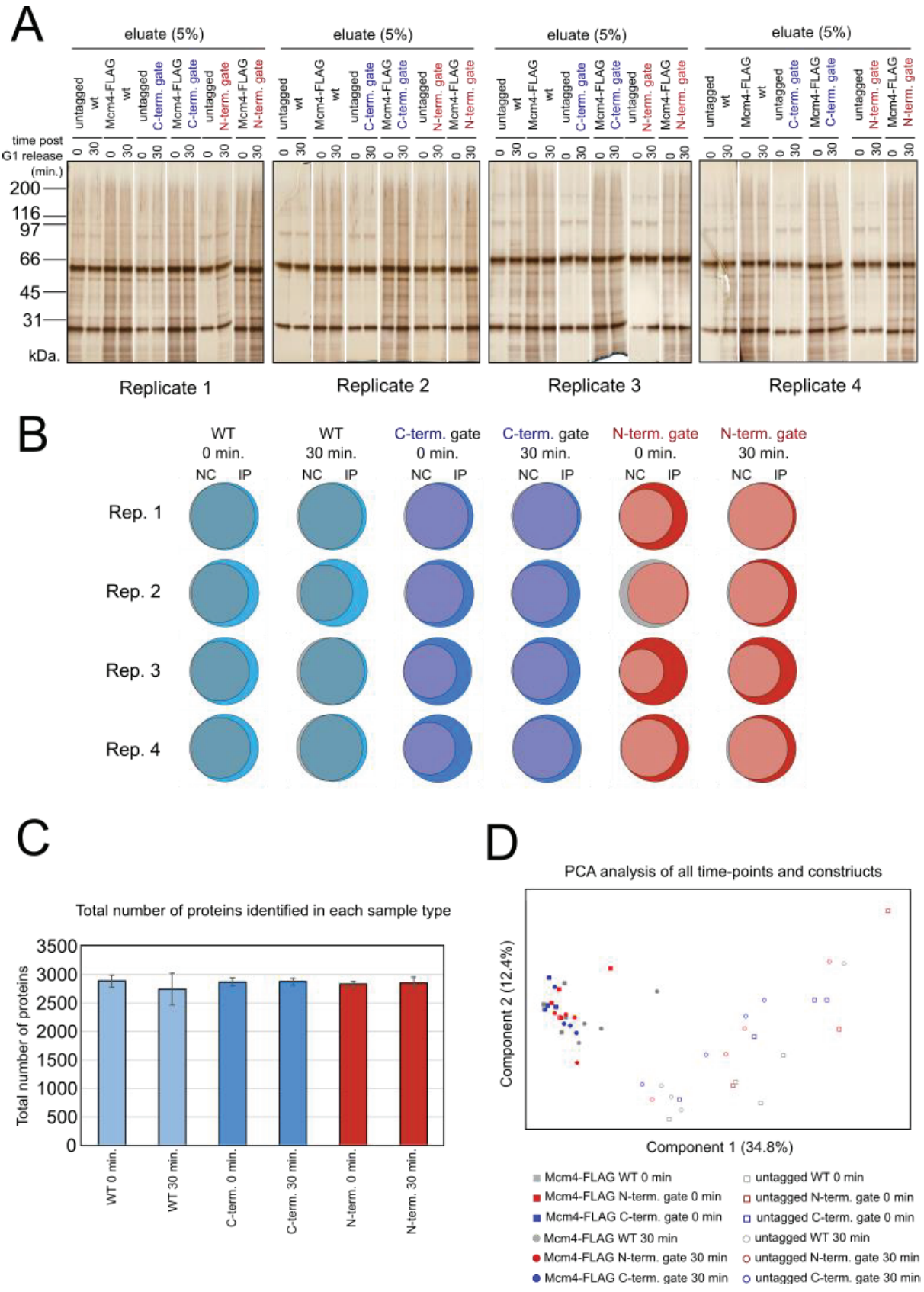


Figure 55: Quality control for the large-scale Mcm4-FLAG ChIP-MS analysis in the context of rapamycin-dependent gate linkage constructs. 600 mL of cell culture were harvested for all strains and time-points post alpha-factor mediated arrest, and processed according to the ChIP-MS method. A) Silver stain visualisation of samples after elution from beads analysed by SDS-PAGE. Note some lanes are re-arranged for visual clarity. B) Venn diagram visualisation of unique proteins in each tagged and untagged sample pair. Unique proteins in tagged IPs are displayed on the right, and unique proteins in untagged controls are displayed on the left. C) Average number of proteins in each sample by type across all four replicates, including tagged and untagged samples. Error bars displayed are one standard deviation (s.d.). D) Principal component analysis of all samples.

The eluates of all four replicates were run using SDS-PAGE, and proteins visualised by silver staining (Figure 55A). A clear distinction between tagged pull-down samples and untagged control samples was noted as previously (Section 3.6.3). As before, this analysis method does not easily identify IP-specific bands when in conjunction with cross-linking, but the intensity of the protein smear can indicate the amount of protein in each sample. There was some variability noted between the replicates, with some appearing to have slightly varying protein amounts overall, but this can be explained by the need to run multiple gels which may show variation in the development of the stain.

Following Trypsin digest, these samples were loaded and analysed by LC-MS. The tagged and untagged sample pairs were superficially compared across sample types and replicates by analysing the number of shared and unique proteins between them (Figure 55B). Most pairs show a greater number of unique proteins in the tagged pull-down sample than the untagged control sample, as expected. The amount of unique proteins versus the number of proteins shared in the pairs did vary across samples and replicates, and it could be seen that this ratio roughly tended to vary across replicates rather than sample type. This indicated a batch effect introduced during cell culture and harvesting or during the IP and washes. Only one sample showed a greater number of unique proteins in the untagged control than in the tagged pull-down.

The overall numbers of proteins identified in each sample type across the replicates was very consistent (Figure 55C). Despite the slight variations in distribution of proteins between tagged and untagged samples, overall the protein list identified remained consistent. The factor most likely to result in these variations is the sample wash steps. Care was taken to ensure that these replicates were as reproducible as possible in the manner in which they were washed, but some variability is nonetheless expected.

Principal component analysis was also conducted in the Perseus software package. Comparisons of samples in this manner showed a clear distinction between almost all of the tagged and untagged samples (Figure 55D). The untagged control for the N-terminal gate for the 0 minute timepoint in replicate 2 showed an alarming similarity to the other tagged samples rather than the other untagged controls. This was also visible in the Venn diagrams (Figure 55B), where this untagged sample actually showed more unique proteins than the tagged sample. Because the paired tagged pull-down did not group with the untagged controls, this was not the result of sample mis-identification. Instead, this result was considered due to cross-contamination, and this sample was excluded from further analysis.

As part of the pipeline of analysis in Perseus, an automated imputation was carried out as previously (Figure 46). The distribution of protein intensities in each sample with the imputed values highlighted in red is displayed in Figure 56. For the tagged pull-down samples, the number of imputed values was low, and did not often distort the distribution. For untagged controls, as identified previously, the lack of certain IP-specific factors from the general background led to a large number of missing values. This in turn led to a large amount of imputation at the low end of the intensity plot. As always, whilst allowing semi-quantitative analysis of IP-specific factors, this approach did lead to a risk of giving inaccurate measurements of enrichment values, particularly for low-abundance proteins in the tagged sample. Because missing values were expected in the untagged controls, as the non-specific background did not contain every cellular protein at high enough abundances to be detected, analysis was continued with appropriate caution, utilising these imputed values.

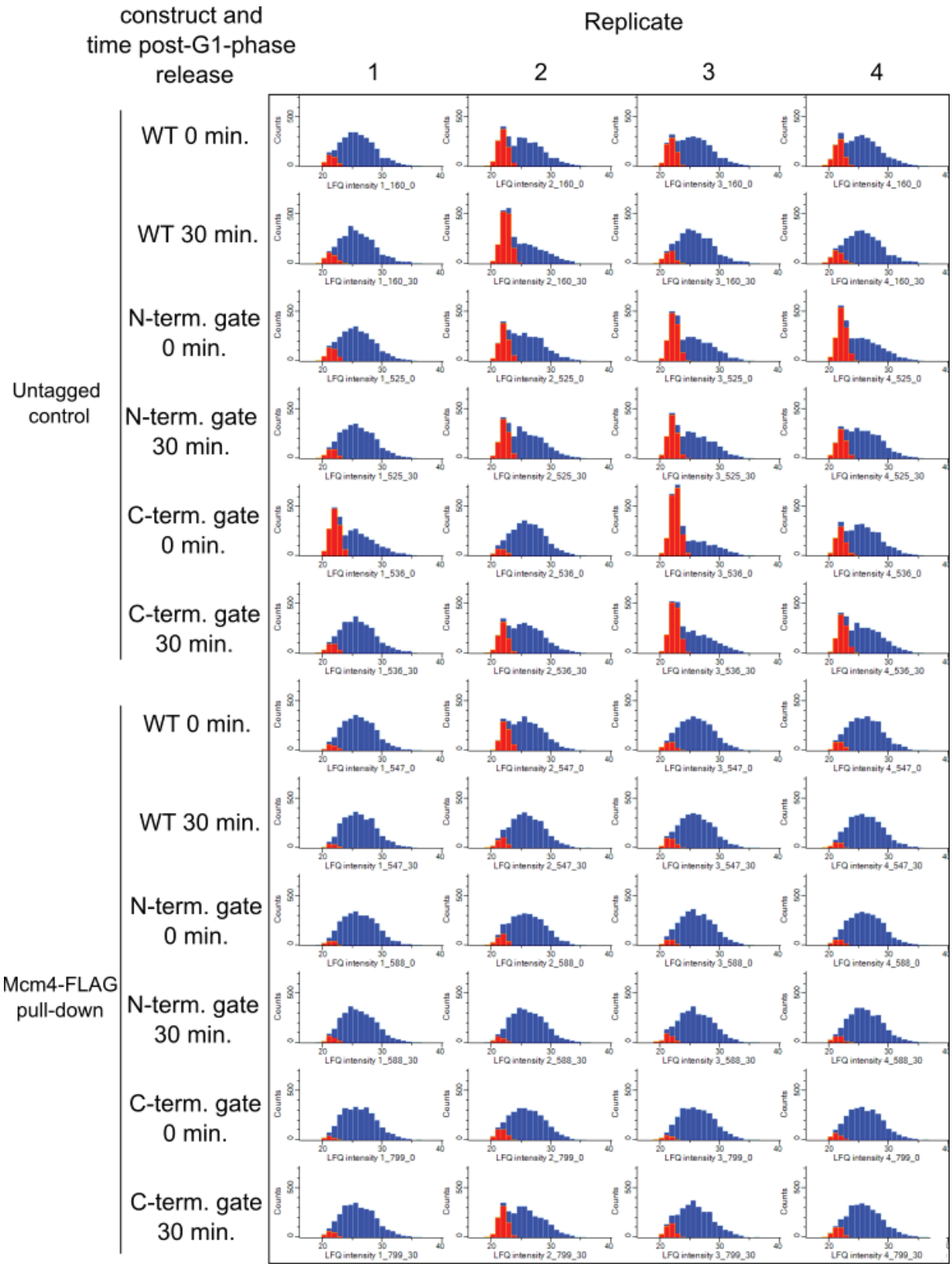


Figure 56: Histograms of protein LFQ intensity distributions in intra-hexamer linkage construct Mcm4-FLAG ChIP-MS for each individual sample. Samples are arranged according to type and replicate, with untagged controls at the top, tagged pull-downs at the bottom. Replicates progress from left to right. For proteins that were not detected in untagged control samples, but were present in 2 or more tagged pull-downs from a single condition, imputation was carried out in the Perseus software (Tyanova et al., 2016), to replace the missing value to allow for approximate quantification in later steps. Such proteins are coloured red.

Comparing the enrichment and p values of proteins between the tagged pull-downs and untagged controls, It was immediately apparent that the pull-down was successful in isolating an MCM2-7 containing complex across all samples (Figure 57). In every pull-down sample, the MCM2-7 subunits were present at a consistently high enrichment over the respective untagged control, and are therefore observed in the top right side of the volcano plots. Many other replication-associated proteins could also be seen to be enriched, dependant on the time-point and strain observed. As with previous experiments utilising Mcm4-FLAG as bait, Cdt1 was ubiquitous in the pull-downs, likely reflecting the presence of soluble MCM2-7 complexes alongside the chromatin-bound complexes of interest.

Reassuringly, in this experimental run, GINS, Sld2, Sld3, Sld7 and Dpb11 were all detected and showed consistent presence across replicates. Whatever issue prevented the identification of these factors in the Mcm4-FLAG ChIP-MS large-scale cell cycle analysis above (Section 3.6.3) was not problematic in this experiment.

In addition, attention was paid to the quantification of those factors that were of interest in previous sections. The enrichment values of Fun30, Hrq1, Rrm3 and Sir3 were monitored in each of the samples analysed, to observe if they showed any interesting changes in association when the rapamycin-inducible linkages were present.

The WT strain, which was released into S phase from alpha factor without the presence of any rapamycin-dependent linkages, showed comparative results to those seen in the prior test experiment, with greater detail and further replicates allowing more statistical analysis. At 0 minutes, immediately after release from alpha-factor mediated arrest, many factors showed statistically significant enrichment with MCM2-7 (Figure 57A). At this time-point, MCM2-7 was still expected to be in the form of a stable double hexamer, and, prior to the Mcm4-FLAG ChIP-MS experiments in this project, was expected to show little to no association with the

replication factors that are part of the replication fork. For samples that had been investigated by ChIP-qPCR, this was especially true. ChIP-qPCR of Cdc45 showed low origin DNA abundance in the pull-down at 0 minutes post G1 arrest, with only slightly greater ARS305 signal than signal of the control non-origin region ARS305+9kb (Figure 16B). By ChIP-MS enrichment analysis, however, Cdc45 was clearly significantly enriched even at 0 minutes. This was true for multiple other replication factors that were previously investigated using ChIP-qPCR and others that were not. Mcm10, which showed minimal origin DNA association at 0 minutes in ChIP-qPCR (Figure 17C), showed significant enrichment to MCM2-7 at 0 minutes by ChIP-MS. The RFC complex, Pol30, and Ctf18 also showed significant enrichment at 0 minutes, despite having roles that are specifically related to the active replication fork. The FACT complex behaved similarly, and showed modest but consistent enrichment even at 0 minutes. For Cdc45, this association at 0 minutes is surprising, but could be due to low levels of DDK activity prior to S phase entry.

Interestingly, all six subunits of the ORC complex were enriched at 0 minutes, suggesting some continued association of ORC with MCM2-7 even after double hexamer loading. Whether this is from a subset of complexes at specific origins or a general feature of the double hexamer in G1 arrest is unclear. Fun30, Hrq1 and Sir3 all showed significant enrichment at this time-point too. This further confirmed the results of the above time-course study that suggested these proteins showed MCM2-7 association prior to activation (Figure 51). Sir3 showed a particularly strong enrichment, which is interesting. Because of the Sir proteins' role at telomeres, it was considered that the Sir3-MCM2-7 association was a result of close associations limited to telomeric ARS sequences (Renauld et al., 1993; Radman-Livaja et al., 2011; Hoggard et al., 2018), and was not a general feature of origins throughout the genome. The enrichment values observed here would seem to make this possibility less likely, as a small subset of complexes would be unlikely to generate such high enrichment values. This data

rather suggested that Sir3 might have a greater role to play at origins across the genome, which fits with recent research indicating Sir protein association with origin-adjacent nucleosomes, and a replication timing control function of Sir2 (Hoggard et al., 2018, 2020).

Some factors showed a clear absence of enrichment, however, in keeping with ChIP-qPCR results. Sld2, GINS and Dpb11 all did not show significant enrichment at 0 minutes. This suggested that whilst other replication proteins might begin to associate prior to activation, the pre-LC components that have been said to function in a switch-like manner in complex progression are indeed tightly regulated and show no association with MCM2-7 at this early time-point (Muramatsu et al., 2010; Miyazawa-Onami, Araki & Tanaka, 2017).

At the 30 minute time-point, many further replication factors showed significant enrichment over the untagged control (Figure 57B). Consistent with previous data, the progression to active replisomes caused a much clearer and reproducible association of polymerase ε, GINS and all three subunits of RPA. Cdc45 showed a much greater magnitude of enrichment at 30 minutes, and so too did Tof1 and Mrc1. The third member of the fork protection complex, Csm3, was not detected consistently in the four replicates, and was removed by quality control measures in the Perseus software pipeline. Fewer ORC subunits were significantly enriched, perhaps as a result of complex movement away from origins. The enrichment of ORC still observable was possibly due to double hexamers that had not undergone activation by 30 minutes, as would be the case for many late origins. Rrm3 was clearly significantly enriched at 30 minutes, consistent with its association with active forks, and previous time-course data (Schmidt, Derry & Kolodner, 2002; Azvolinsky et al., 2006) (Figure 53C). The FACT complex continued to show highly consistent but modest enrichment, and from this visualisation it was hard to tell whether this complex showed an increased MCM2-7 association after activation than before it.

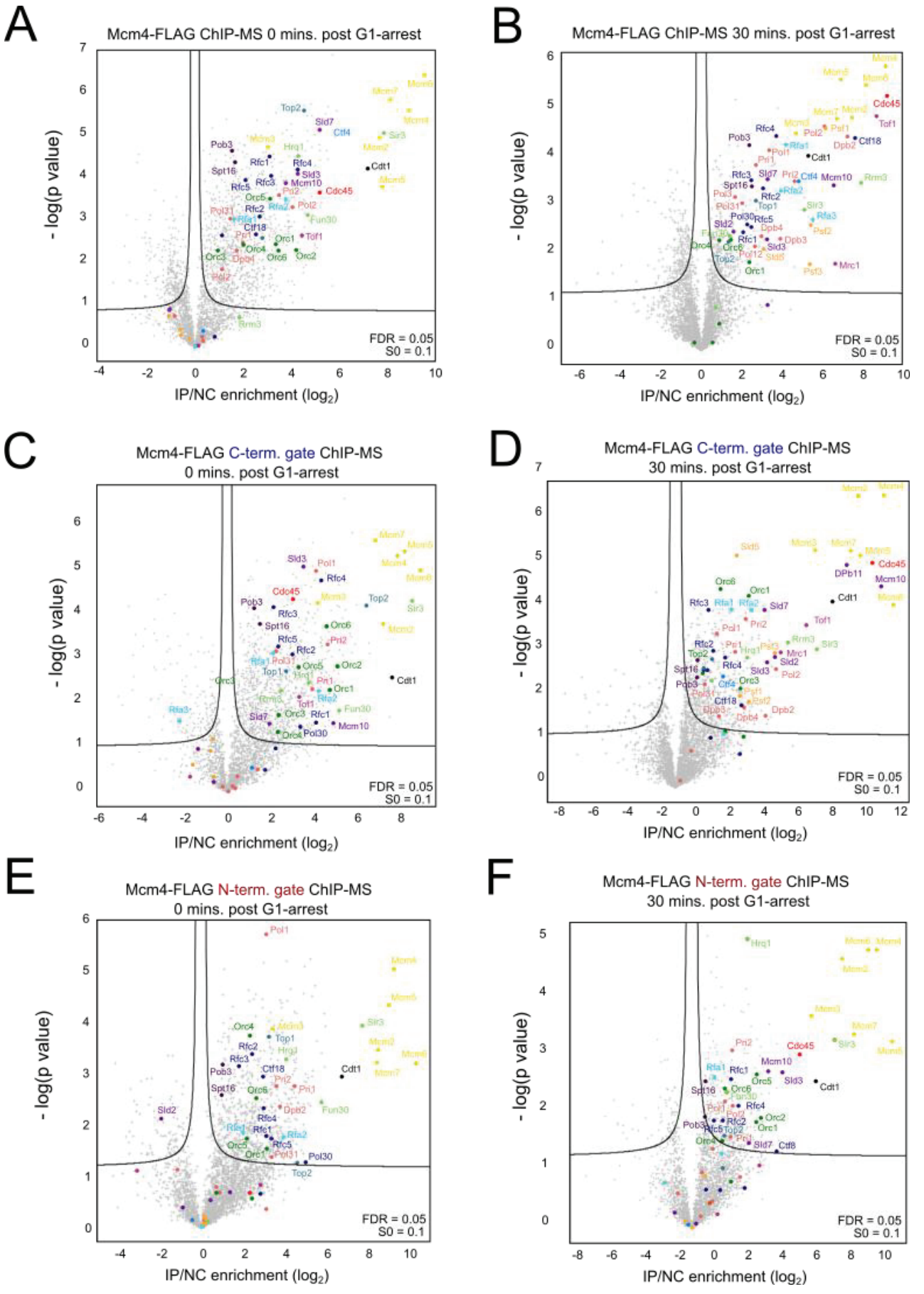

A
Mcm4-FLAG ChIP-MS 0 mins. post G1-arrest
B
Mcm4-FLAG ChIP-MS 30 mins. post G1-arrest
C
Mcm4-FLAG C-term. gate ChIP-MS
0 mins. post G1-arrest
D
Mcm4-FLAG C-term. gate ChIP-MS
30 mins. post G1-arrest
E
Mcm4-FLAG N-term. gate ChIP-MS
0 mins. post G1-arrest
F
Mcm4-FLAG N-term. gate ChIP-MS
30 mins. post G1-arrest
- log(p value)
IP/NC enrichment (log2)
FDR = 0.05
S0 = 0.1

Figure 57: Enrichment analysis of proteins in Mcm4-FLAG ChIP-MS in context ofintra-hexamer linkage constructs comparing tagged pull-downs against their respective untagged controls by volcano plot. Protein enrichments between tagged samples and untagged controls were plotted against the p-value from a permutation-based two-sample t-test carried out in the Perseus software (Tyanova et al., 2016), using an FDR of 0.05, and an S0 value of 0.1. Proteins are coloured according to the previously used schema. A) Volcano plot showing enrichment against p-value for the comparison between Mcm4-FLAG cells without a rapamycin-dependent linkage construct against the untagged negative control cells, at 0 minutes post-G1 arrest. B) As in A, but for 30 minutes post-G1 arrest. C) Volcano plot showing enrichment against p-value for the comparison between Mcm4-FLAG cells in the context of the C-terminal intra-hexamer rapamycin-dependent linkage against the untagged negative control cells, at 0 minutes post-G1 arrest. D) As in C, but for 30 minutes post-G1 arrest. E) Volcano plot showing enrichment against p-value for the comparison between Mcm4-FLAG cells in the context of the N-terminal intra-hexamer rapamycin-dependent linkage against the untagged negative control cells, at 0 minutes post-G1 arrest. F) As in E, but for 30 minutes post-G1 arrest.

Analysing the data for these same two time-points but for the rapamycin-inducible linkage constructs showed some minor differences at 0 minutes, but some more obvious differences at 30 minutes (Figure 57C-F). Overall, the rapamycin-dependent linkage constructs showed enrichment of a similar set of factors at 0 minutes to the WT, including evidence of the same pre-association of replication factors. An obvious difference was that Cdc45 did not show significant enrichment in the N-terminal rapamycin-dependent linkage construct. For the 30 minutes, differences were more apparent, including a very strong enrichment of Dpb11 in the C-terminal construct, and a general lack of enrichment of several replication factors in the N-terminal construct.

To exhibit the enrichments of a core set of replication factors more clearly, the average enrichment values of these factors in each sample type are shown in Figure 58 and Figure 59. The consistently high enrichment of the subunits of MCM2-7 could be seen at 0 (Figure 58) and 30 (Figure 59) minutes. Figure 58A, in particular, shows that in the Mcm4-FLAG ChIP-MS of a sample not containing any rapamycin-dependent linkages, all of the selected factors displayed enrichment over the untagged control. The most clearly observed differences already remarked upon, namely the lack of Cdc45 enrichment in the N-terminal construct at 0 minutes, and the strong enrichment of Dpb11 in the C-terminal construct at 30 minutes, can be clearly observed here.

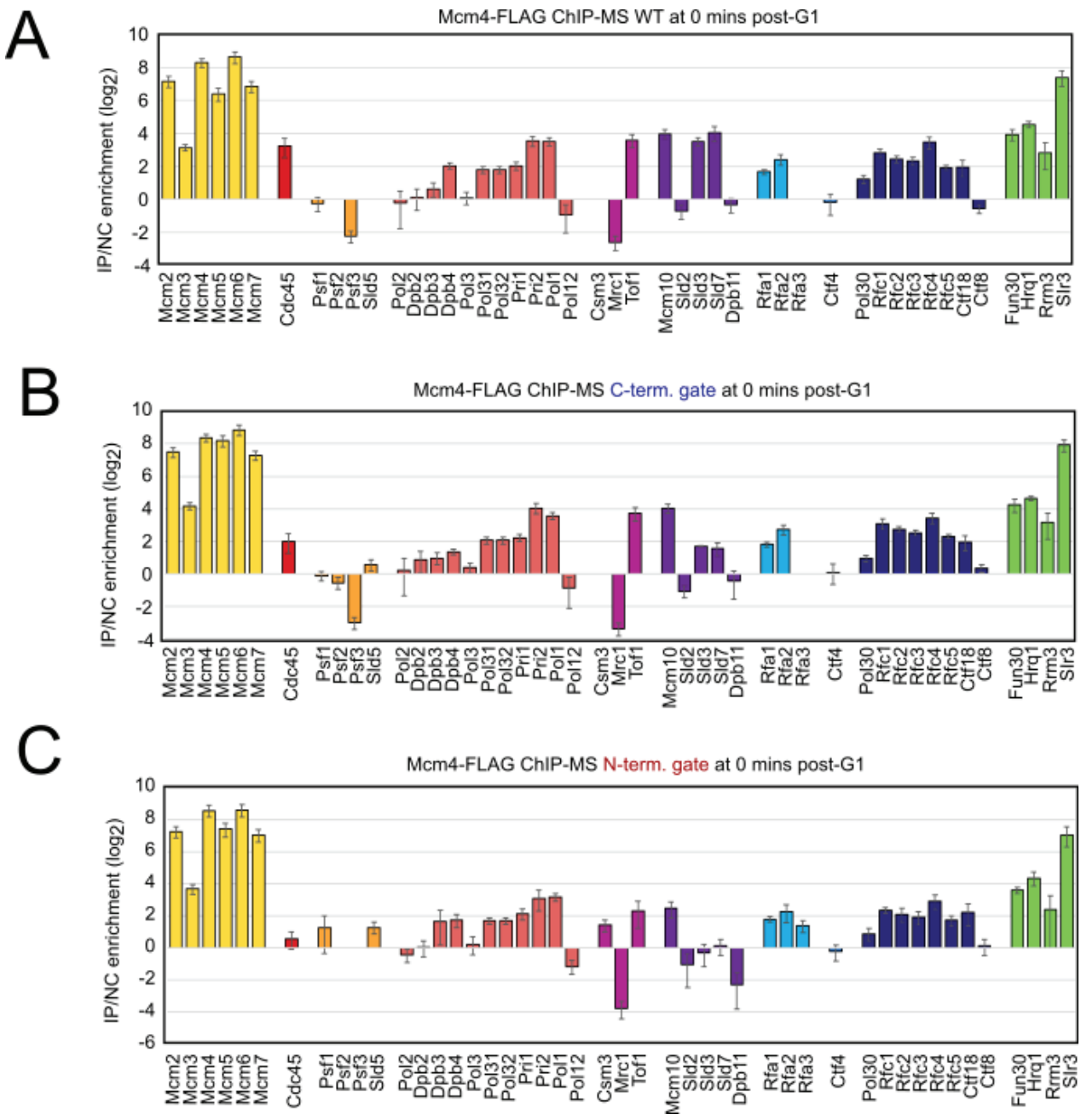


Figure 58: Plots showing the enrichment of a manually curated set of replication factors across different constructs at 0 minutes post alpha-factor release. Enrichment values were calculated using the mean LFQ intensities from tagged pull-down samples across all four replicates, and the mean LFQ intensities from a pool of untagged control samples across all four replicates, shared between the constructs. The error bars displayed are the calculated standard errors in the mean (s.e.m.) scaled appropriately for the log_2(enrichment). A) enrichment plot for the set of replication factors in the ChIP-MS of an Mcm4-FLAG strain containing no rapamycin-dependent linkages. B) As for A, but for a strain containing the C-terminal rapamycin-dependent linkage, treated with 100 nM rapamycin. C) As for B, but containing the N-terminal rapamycin-dependent linkage.

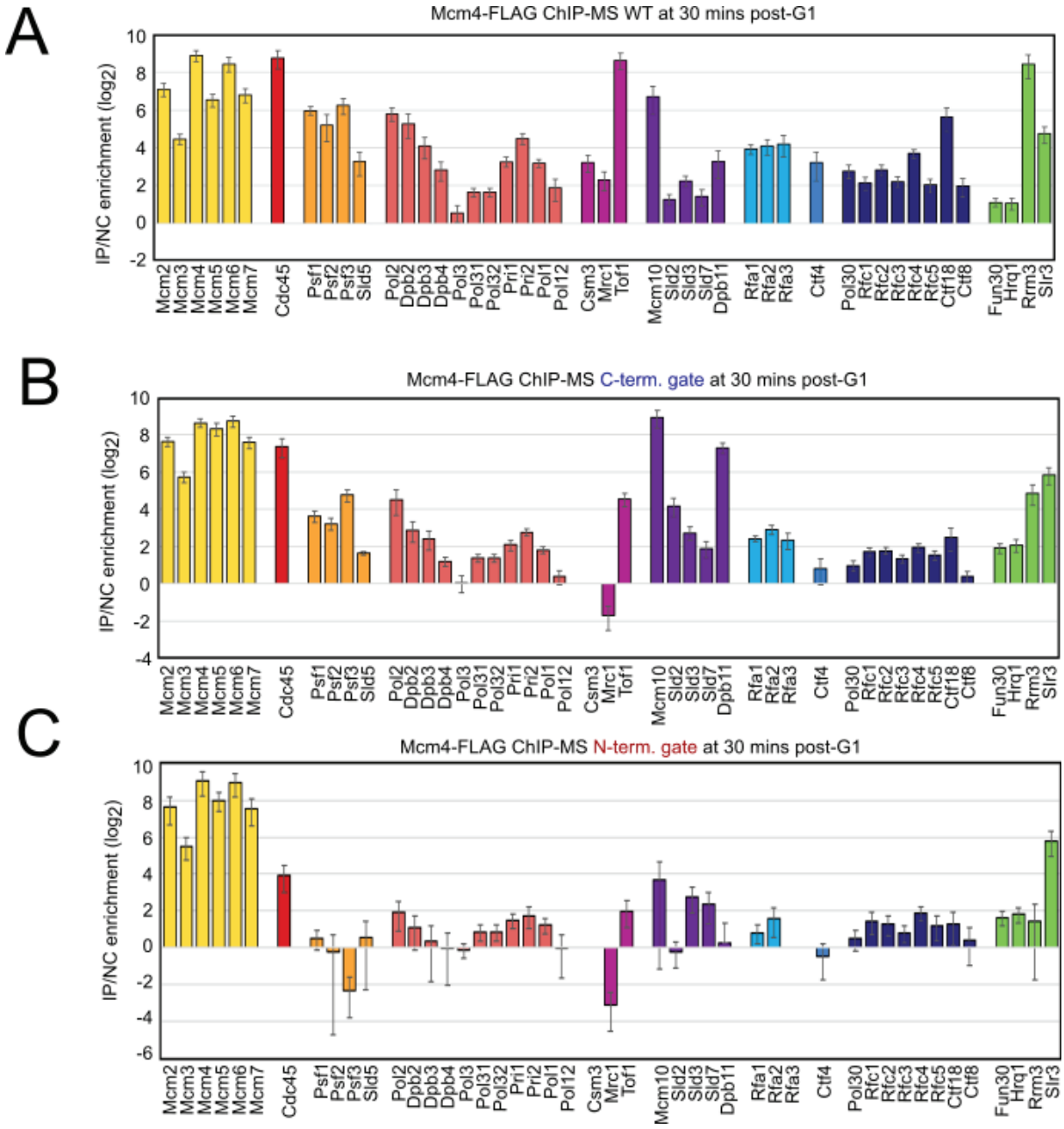


Figure 59: Plots showing the enrichment of a manually curated set of replication factors across different constructs at 30 minutes post alpha-factor release. Enrichment values were calculated using the mean LFQ intensities from tagged pull-down samples across all four replicates, and the mean LFQ intensities from a pool of untagged control samples across all four replicates, shared between the constructs. The error bars displayed are the calculated standard errors in the mean (s.e.m.) scaled appropriately for the log_2(enrichment). A) enrichment plot for the set of replication factors in the ChIP-MS of an Mcm4-FLAG strain containing no rapamycin-dependent linkages. B) As for A, but for a strain containing the C-terminal rapamycin-dependent linkage, treated with 100 nM rapamycin. C) As for B, but containing the N-terminal rapamycin-dependent linkage

Carrying out volcano-plot based analysis on pairwise comparisons of conditions in this experiment allowed identification of proteins that were significantly enriched or de-enriched dependent on the time-point and construct under investigation (Figure 60).

As a first step, this analysis was conducted within the construct types, comparing the 30 minute time-point to the 0 minute time-point (Figure 60A-C). This would show the changes in protein composition of the complexes over the process of activation from the double-hexamer state of the G1 arrest, through to the active replisome in the WT strain, and through to the arrest intermediate generated in the rapamycin-dependent linkage strains. The strain without any linkage showed the most significant changes, with a number of protein significantly enriched and de-enriched in the 30 minute time-point versus the 0 minute time-point (Figure 60A). The identities of these factors was consistent with the expected structural and compositional changes expected in the transition from double-hexamer prior to activation, through to a fully active replisome. The GINS complex, which associates as part of the pre-LC complex during activation, and then stays stably bound to the replisome (Gambus et al., 2006; Muramatsu et al., 2010), showed one of the highest enrichments in 30 minutes compared to 0 minutes.

Cdc45 and Mcm10 also showed a relative increase in abundance in the 30 minute time-point, despite showing a reasonable presence in both individual time-points (Figure 57A and B). This is likely due to an increase in the stability of these proteins' interaction with MCM2-7 following activation.

Ctf4 showed an increased association in the 30 minute time-point, and since this protein acts as a central interacting hub at the replisome, an increased association once replication begins makes sense, since it is then that its binding partner GINS has associated with the replisome, and the overall structure of the complex allows it to couple to other factors to link them with the core helicase (Simon et al., 2014; Gambus et al., 2009).

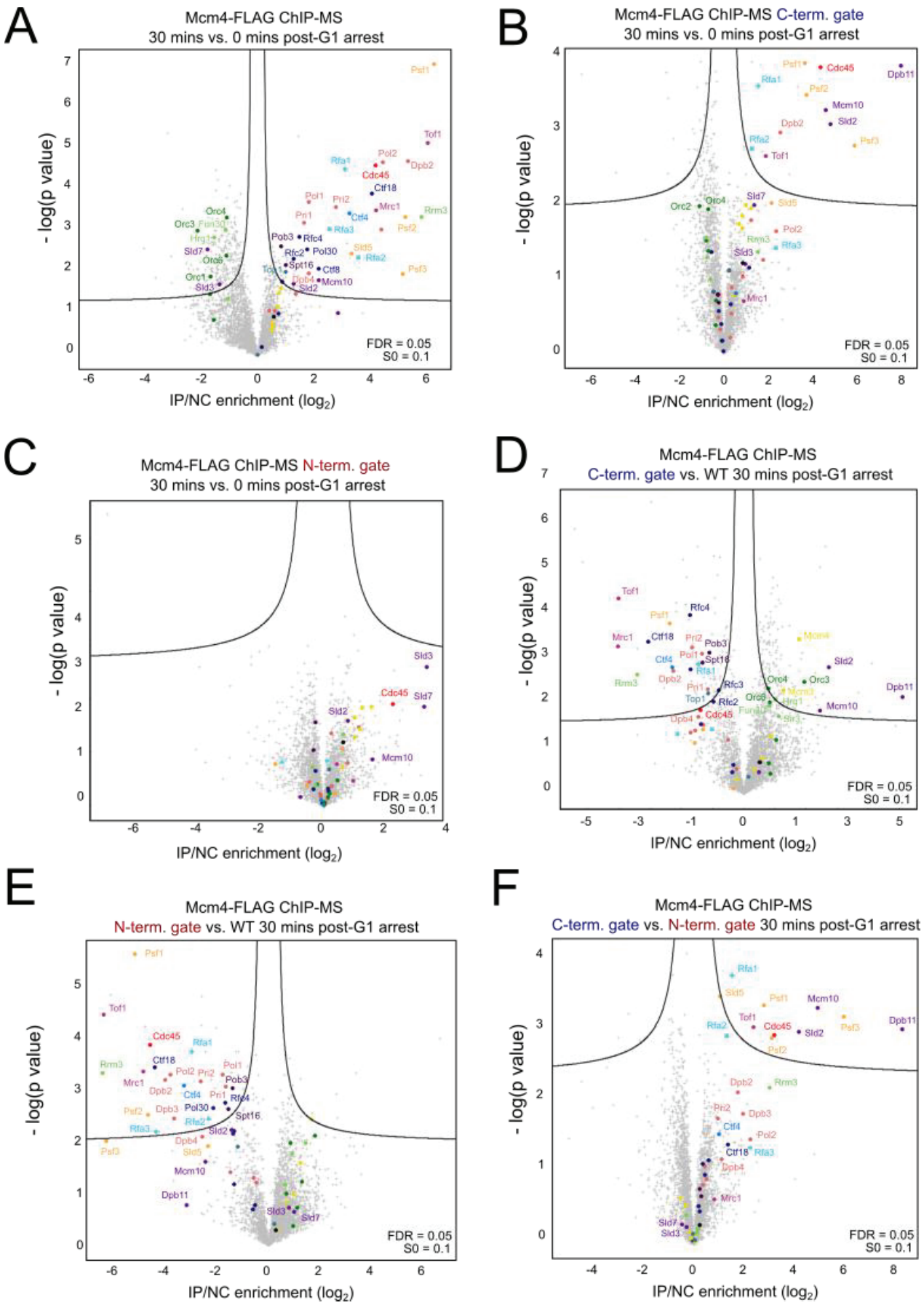

A
Mcm4-FLAG ChIP-MS
30 mins vs. 0 mins post-G1 arrest
- log(p value)
IP/NC enrichment (log2)
FDR = 0.05
S0 = 0.1
B
Mcm4-FLAG ChIP-MS C-term. gate
30 mins vs. 0 mins post-G1 arrest
- log(p value)
IP/NC enrichment (log2)
FDR = 0.05
S0 = 0.1
C
Mcm4-FLAG ChIP-MS N-term. gate
30 mins vs. 0 mins post-G1 arrest
- log(p value)
IP/NC enrichment (log2)
FDR = 0.05
S0 = 0.1
D
Mcm4-FLAG ChIP-MS
C-term. gate vs. WT 30 mins post-G1 arrest
- log(p value)
IP/NC enrichment (log2)
FDR = 0.05
S0 = 0.1
E
Mcm4-FLAG ChIP-MS
N-term. gate vs. WT 30 mins post-G1 arrest
- log(p value)
IP/NC enrichment (log2)
FDR = 0.05
S0 = 0.1
F
Mcm4-FLAG ChIP-MS
C-term. gate vs. N-term. gate 30 mins post-G1 arrest
- log(p value)
IP/NC enrichment (log2)
FDR = 0.05
S0 = 0.1

Figure 60: Enrichment analysis of proteins in Mcm4-FLAG ChIP-MS in intra-hexamer linkage constructs comparing tagged pull-downs against other tagged pull-down conditions by volcano plot. Protein enrichments between tagged samples were plotted against the p-value from a permutation-based two-sample t-test carried out in the Perseus software, using an FDR of 0.05, and an S0 value of 0.1. Proteins are coloured according to the previously used schema. A) Volcano plot showing enrichment against p-value for the comparison between Mcm4-FLAG cells without the presence of a rapamycin-dependent linkage at 30 minutes post-G1 arrest against Mcm4-FLAG without the presence of a rapamycin-dependent linkage cells at 0 minutes post-G1 arrest. B) As in A, but in the context of cells containing the C-terminal intra-hexamer rapamycin-dependent linkage C) As in B, but in the context of the N-terminal linkage D) Volcano plot showing enrichment against p-value for the comparison between Mcm4-FLAG cells in the context of the C-terminal linkage against Mcm4-FLAG cells without any rapamycin-dependent linkage present, at 30 minutes post-G1 arrest. E) As in D, but for the N-terminal linkage against the cells without any linkage. F) As in D, but for the comparison between cells containing the C-terminal intra-hexamer rapamycin-dependent linkage against cells containing the N-terminal linkage, at 30 minutes post-G1 arrest.

The detected members of the fork protection complex, Tof1 and Mrc1, as well as most of the subunits of polymerases α and ε were also significantly enriched. These protein complexes show stable association with the replication fork, and so it is likely that the transition increases the stability of their association or induces their association in the first place, as for polymerase ε.

In addition, the three subunits of RPA all show significant enrichment, which is a clear indication of the presence of ssDNA, as expected at active replication forks. Whilst two of three subunits showed enrichment in the 0 minute time-point individually, the clear enrichment here demonstrates that the active fork has significantly more RPA association. The enrichment observed at 0 minutes for RPA could perhaps be the result of ssDNA-independent recruitment to origins, and is overall unlikely to represent the real presence of ssDNA prior to activation.

Several factors are de-enriched in the transition from 0 to 30 minutes (Figure 60A). Among these proteins are the transient initiation factors Sld3 and Sld7, several subunits of ORC, Fun30 and Hrq1. Because of the pre-association observed for Sld3/7 at 0 minutes, the lower abundance of these proteins is not unexpected. Once initiation has completed, both Sld3 and Sld7 are released from MCM2-7. The 0 minute time-point suggests that Sld3 and Sld7 show some amount of pre-association, despite this having been shown to require DDK phosphorylation of MCM2-7 (Fang, Cao & Lou, 2016; Deegan, Yeeles & Diffley, 2016). If Sld3/7 are indeed pre-associated, then this was either an unstable association stabilised by

cross-linking that does not require phosphorylation of MCM2-7, or there was some level of DDK activity in the G1 arrest. Regardless, these two factors showed pre-association in this experiment, and so once activation begins, these factors will begin to be lost from MCM2-7 complexes, resulting in a greater proportion of active replisomes without Sld3/7 bound, and a lower proportion of double hexamers awaiting activation.

The loss of ORC association is more easily explained. Despite the initial ORC association at 0 minutes seeming to contradict data from the lab suggesting double hexamer formation and ORC origin binding are mutually exclusive (L.M. Reuter, manuscript in preparation), some level of interaction or association clearly remained at origins. This association does therefore seem to be confined to origins, since at 30 minutes, many more MCM2-7 complexes have been activated and so the amount of ORC association was shown to be lower at this time-point, as MCM2-7 complexes move away from origins as part of replication forks.

Previous experiments in this study (Section 3.6.3), showed that Fun30 and Hrq1 showed greatest abundance in ChIP-MS samples prior to activation. As the 30 minute time-point here should contain a lower proportion of pre-activation MCM2-7 complexes, the association of these proteins with MCM2-7 was also seen to decrease, further exemplifying that whatever the association of these factors might be, it seems either confined spatially to origins, or temporally to pre-activation MCM2-7 complexes.

A key observation and measure of reliability was that the enrichment of MCM2-7 subunits themselves when comparing between pull-downs was not significantly different. Since culture volumes were the same, the pull-down performed the same way, and protein concentrations normalised during preparation for MS submission, the level of MCM2-7 in each sample should be similar. A possible exception could be caused by production of new MCM2-7 complexes or degradation of old MCM2-7 complexes. This would be more of a consideration if more complexes were expected to be terminating, but at 30 minutes this proportion was expected to

be low. Whilst the enrichment in 30 minutes over 0 minutes was not significant, there was still a slight enrichment of MCM2-7 in the 30 minute timepoint over 0. It was not entirely clear why this occurred, but it was consistently enriched at 30 minutes over 0 minutes for all strains (Figure 60A-C).

For the C-terminal rapamycin-dependent linkage construct specifically, there were a number of replication factors that showed enrichment in the 30 minute over the 0 minute time-point (Figure 60B). Notably, among these factors were Dpb11, Sld2, Mcm10 and three of the four GINS subunits, with Sld5 trending towards significant enrichment but not quite making the threshold. These factors are indicative of late stage activation, which fits well with the ChIP-qPCR experiments studying the C-terminal linker arrest phenotype. It does indeed appear that the complex is arrested in the process of a late stage initiation step, and that failure to complete this step appropriately prevents the release of Dpb11 and Sld2. The enrichment of Mcm10 also indicates that unlike the iAID-induced depletions of pre-LC components that prevented Mcm10 association (Miyazawa-Onami, Araki & Tanaka, 2017), the C-terminal rapamycin-dependent linkage arrests at a later step, once pre-LC components have performed some function that allows for Mcm10 binding. This could be the result of the splitting that Miyazawa-Onami et al. claim is caused by the activity of the pre-LC, or of some other process that Dpb11, Sld2 and GINS are required for.

This arrested complex also showed enrichment of Cdc45 when compared to 0 minutes. As with the cells without a linkage construct, this increase in abundance could have been caused by generation of a more stable interaction, as a result of MCM2-7 phosphorylation during S-phase entry, or of some other change, perhaps the N-terminal extrusion of ssDNA through the Mcm2-Mcm5 gate that allowed Cdc45 to bind over this interface securely.

The remaining factors that showed significant enrichment in the C-terminal linkage construct at 30 minutes when compared to 0, are two of the three subunits of RPA, Tof1 and Dpb2. The

increased association of RPA correlates with the ChIP-qPCR data that suggested the arrest state of the C-terminal linkage was able to produce ssDNA (Figure 17B). The enrichment of Tof1 without a clear increase in Mrc1 was not easy to explain. Perhaps some functional change in the complex during activation allows the pre-association of Tof1 displayed in 0 minute samples to become a more stable interaction, or, perhaps the presence of ssDNA increases the propensity of this complex to cross-link from protein to DNA, and this enhances the stability of these complexes during the pull-down process. Dpb2 as a subunit of polymerase ε was easier to explain in isolation, since polymerase ε is recruited as part of the pre-LC along with the notable initiation factor Pol2, and so it makes sense that this protein was enriched. The lack of significant enrichment of the other members of polymerase ε makes this factor's enrichment alone harder to explain, but Pol2 does trend towards significant enrichment even if it does not meet the standard for significance.

In the N-terminal linkage construct, there were no replication factors that met the standard for significance in their enrichment in 30 minutes over 0 minutes. This suggested that the complex composition of the arrested intermediate in this construct was overall similar to that of the double hexamer in G1. This also fits well with the hypothesis that by preventing DNA ssDNA extrusion at the N-terminus, the activation process is halted early on.

While there weren't any replication factors significantly enriched in the N-terminal linkage strain, there were a few factors that tended towards high enrichment, but were too variable to meet the significance threshold. The three clearest factors in this group were Cdc45, Sld3 and Sld7. Because these factors are among the first to associate in response to activation signals (Zou & Stillman, 1998; Seki & Diffley, 2000; Pacek & Walter, 2004; Walter & Newport, 2000), it makes sense that they are the factors most tending towards enrichment, which could indicate the first steps of activation. Further study including more replicates would most likely be able to investigate this more fully and show if the enrichment is real.

From the comparative temporal analysis, it was clear that the cells without a linkage, and the cells with a C-terminal linkage did show clear differences in complex composition over time, whilst cells with an N-terminal linkage showed little difference in complex composition. The factors that were enriched in the C-terminal linkage were also generally enriched in the cells without a linkage, with the exception of Dpb11, which indicated that the C-terminal linkage complex could well represent a stalled intermediate state where the transient factor Dpb11 had not yet been released.

To further confirm the differences between construct arrests, direct comparison was carried out as previously between time-points, but instead between constructs at 30 minutes. By comparing the complexes generated by the constructs at this time-point, the amount of complex progression in these cells compared to the cells not hindered by rapamycin-dependent linkages could be shown (Figure 60D-F).

First, both construct-containing pull-downs were compared to the "wild type" unhindered pull-down individually. For the C-terminal linkage construct, there were enriched proteins in each of the conditions relative to the other (Figure 60D). This was important, because the process of initiation is known to not simply be a sequential association of factors that accumulate to form the final complex, but also requires the activity of transient associating factors. The constructs used here were intended to interfere with activation at a point where transient initiation factors were potentially still bound. Excitingly, the ChIP-MS confirms the indications from ChIP-qPCR with regard to the transient factors, as Dpb11 and Sld2 both show clear enrichment in the C-terminal construct-induced arrest state, over the un-arrested replisome present in the "wild type". Interestingly, Mcm10 is also enriched in this arrested complex relative to active replisomes. Although it is not entirely clear whether Mcm10 is a component of active, unperturbed replisomes, the balance of evidence over the past few years would suggest that it does interact with the active replisome post-activation (Thu & Bielinsky, 2013). Because

Mcm10 behaves similarly to Dpb11 and Sld2, which are not thought to associate with active forks (Gambus et al., 2006, 2009), this data could suggest that Mcm10 does associate much more during activation steps than with the active fork. Alternatively, the interaction of Mcm10 with the MCM2-7 could be much stronger during activation, and weaker during fork progression. The need of high salt conditions to co-purify Mcm10 with the replisome during replication could support the hypothesis that Mcm10 association during replication is maintained by weaker hydrophobic interactions than the interactions during activation.

Two subunits of ORC also showed significant enrichment in the C-terminal linkage construct at 30 minutes post-release, and the others did show enrichment that did not reach the threshold for statistical significance. The mechanism by which the rapamycin-dependent intra-hexamer linkages are thought to arrest complex progression is by prevention of proper ssDNA exit from the MCM2-7 hexamer. By restricting the motion of DNA in and around the helicase, it is likely that movement of the complex is reduced, both possibly by the inactivity of the helicase, and the constrained DNA. The ChIP-qPCR data suggested that the MCM2-7 complex maintains association with origin DNA when arrested with the intra-hexamer linkages, and the enrichment of ORC in the C-terminal linkage construct would seem to confirm that this was indeed the case here as well.

Two of the subunits of MCM2-7, Mcm4 and Mcm3, also showed enrichment in the C-terminal construct state. Whilst the other subunits show variable levels of enrichment, these two subunits are statistically enriched in this construct. This was surprising, since in both conditions, Mcm4 was used as the bait protein, and so MCM2-7 subunits should not be significantly enriched in either case. Despite this, in several of the conditions in this experiment, a general trend of MCM2-7 enrichment could be seen when comparing conditions. It’s not clear why this occurred, since overall MCM2-7 levels are expected to fairly consistent across the conditions.

Despite this bias of MCM2-7 enrichment, there are still many proteins that were enriched in the "wild type" condition not containing any linkage (Figure 60D). These factors are some of the familiar factors that are associated with active replication forks. Tof1 and Mrc1 both showed strong and consistent enrichment, and so too did GINS and Ctf4. Several members of polymerase complexes were significantly enriched, as well as members of the RFC and related complexes. Cdc45 was also enriched in this condition, but did not meet the threshold of significance. RPA showed a single subunit that was significantly enriched, Rfa1, but the other two, whilst enriched, were not significantly so.

Altogether, this group of factors confirm that the "wild type" has a greater abundance of active replisome-associated factors, indicating that the C-terminal linkage-induced complex, whilst clearly at a late stage in activation, is not yet a fully formed replication fork, able to recruit all the usual factors. The higher abundance of Cdc45 and GINS in the "wild type", despite these factors being present in the linkage construct-induced arrest (Figure 58D,Figure 59B,Figure 60B), suggests that the strength of the interaction with MCM2-7 is improved in the "wild type". Because of the C-terminal linkage causing interference of ssDNA extrusion through the Mcm2-Mcm5 interface, where Cdc45 and GINS eventually bind in the CMG (Sun et al., 2015), it is possible that the binding of these factors is not fully secured due to the blocked or improper ssDNA exclusion. The variable enrichment of the subunits of RPA, whilst not overall significant, does tend towards enrichment in the active replication complex which suggests despite ssDNA expected to be present in the C-terminal linkage-mediated arrest complex (Figure 17B), the amount of RPA associated with that complex is still less than that of a fully active fork, putting a limit on how much ssDNA could be produced in the arrest complex.

The N-terminal linkage construct shows a different result when compared directly against the "wild type" (Figure 60E). In this case, no replication factors showed significant enrichment in the construct-mediated arrest condition. Instead, only the active, un-arrested replisome showed

significant enrichment of replication factors. These factors are similar to those identified in the enrichment over the C-terminal construct, indicative of a fully functioning replisome, and indicating recruitment of these factors to the arrested construct was much lower. Among the factors that tended towards significant enrichment in the construct-mediated arrest, are Sld3, Sld7 and ORC. These factors make sense as they indicate an origin-bound complex that is early in the stage of initiation, as discussed. It is possible with improved reproducibility and additional replicates this enrichment could be significant.

Finally, the two construct-mediated arrests were directly compared (Figure 60F). Again, the N-terminal linkage construct complex showed no significantly enriched replication factors over the compared condition. However, the C-terminal construct complex showed significant enrichment of several replication factors. Unsurprisingly, this included Sld2, Dpb11 and GINS. These pre-LC members associate at a stage later than that which the N-terminal construct is able to reach, and so there is a clear enrichment in the C-terminal construct. Similarly, there is an enrichment of RPA, of which Rfa1 is the only subunit reaching statistical significance here. This agrees with ChIP-qPCR data that suggests only a slight recruitment of RPA to the N-terminal arrest, but more RPA recruitment to the C-terminal arrest (Figure 17B and Figure 18B). Tof1 was also significantly enriched, which continued to imply that this factor begins its association relatively early with the MCM2-7 complex, despite its binding position on the N-terminus of the MCM2-7 hexamer and engagement with parental DNA at the fork (Baretić et al., 2020). Finally, Cdc45 also showed significant enrichment. This was surprising, as the factors that came closest to significant enrichment in the N-terminal construct from 0 to 30 minutes were Cdc45 and Sld3/7 (Figure 60C). Sld3/7 did not show any enrichment when comparing N- and C-terminal constructs, yet Cdc45 did. This could be explained by the fact that whilst Sd3/7 are transient replication factors that are lost from the complex during activation, Cdc45 retains association post-activation. Thus, as complex progression continues

from early to late stages of activation, a progression reconstituted by these two linkage construct-induced arrests, Cdc45 association is likely to tend towards becoming more stably bound. This could be the rationale behind its increased abundance in the C-terminal linkage construct.

The investigation of the dynamics of known replication factors was only one aspect of the ChIP-MS study into these arrest states. A further goal was to observe any factors that showed enrichment with any complex that had not previously been indicated to have a role in replication, and specifically replication initiation. The un-arrested Cdc45-FLAG ChIP-MS and Mcm4-FLAG ChIP-MS time-course had already generated a list of possible interactors that associated with active replication, and the time-course had narrowed down the association of some factors to actually be prior to activation, but there was still no clear indication of any factor that specifically could be tied to initiation itself. Therefore, by calculating the enrichment values when comparing the pull-downs from this experiment over both their paired untagged controls, and each other, potential novel factors could be highlighted (Figure 61).

A

25 most enriched proteins IP/NC
WT 30 min. post G1-arrest

MCM4	9.110935
RRM3	8.501461
MCM6	8.39277
CDC45	8.244452
TOF1	8.178161
MCM2	7.542657
CTF18	7.462917
DPB2	7.160337
MCM5	7.049138
MCM7	6.920801
MCM10	6.785369
DDC1	6.656502
POL2	5.951731
PSF1	5.88462
PSF3	5.599055
ASI1	5.459609
TAH11	5.430418
DCC1	5.400707
NAT5	4.843379
SIR3	4.674882
CTF4	4.59624
PSF2	4.555116
PRI2	4.318815
MCM3	4.31313

25 most enriched proteins IP/NC
C-term. gate 30 min. post G1-arrest

MCM6	8.593328
MCM10	8.547171
MCM4	8.542756
MCM2	7.699573
MCM5	7.653056
MCM7	7.475406
CDC45	7.22625
DPB11	6.575652
PDR11	6.319201
TAH11	5.893713
MCM3	5.780278
SIR3	5.460625
TOF1	4.920185
CAC2	4.756451
RRM3	4.64059
POL2	4.303892
NAT5	4.144904
PSF3	4.033908
PTA1	3.899828
SLD7	3.871182
PCC1	3.829802
SLD3	3.741974
PSF1	3.69253
SLD2	3.451639

25 most enriched proteins IP/NC
N-term. gate 30 min. post G1-arrest

MCM6	9.134925
MCM4	8.915517
MCM5	8.435452
MCM7	7.517788
MCM2	7.220293
PDR11	6.568671
SIR3	6.347648
MCM3	5.614001
CDC45	5.547526
TAH11	5.029091
FAR1	4.60518
MCM10	4.278404
SHE3	4.005527
SLD3	3.660813
YDR132C	3.630951
PTA1	3.450106
YRR1	3.429146
PCC1	3.364162
RTC1	3.215555
RSC2	3.123414
TRP4	3.002354
PEX5	2.886923
CTI6	2.86471
YRF1-X	2.82461

B

most enriched proteins
C-term. gate vs. WT
30 min. post G1-arrest

DPB11	4.023379
SLD2	2.899921
RBD2	2.695633
MCM10	2.217211
PAP2	2.04868
APC5	1.988033
FUS1	1.960892
ASG7	1.945502
ORC3	1.905096
MCM5	1.764302
APC4	1.74963
APC2	1.655779
TOR2	1.612313
QDR2	1.601891
DBP6	1.597102
SPT7	1.565599
OPT1	1.525304
YLR413W	1.52525
URB2	1.444115
NCS6	1.434566
RNT1	1.428728
SGD1	1.414126
UTP30	1.412614
FAL1	1.410197
SLK19	1.387815

most enriched proteins
N-term. gate vs. WT
30 min. post G1-arrest

FAR1	3.613822
RBD2	3.010631
YKU70	2.772535
GAD1	2.769331
HEL1	2.754573
ADR1	2.749069
GCV1	2.637322
HIS3	2.499796
SMI1	2.407603
GDH2	2.393079
YHR033W	2.325203
DUR1,2	2.289946
YBL055C	2.182721
GPH1	2.139944
APC4	2.101753
GUT2	1.9527
RRN6	1.903539
ADH2	1.830259
PRP18	1.735464
NCS6	1.644871
YIH1	1.608523
APC2	1.589391
APC11	1.579364
LCD1	1.558724
PAP2	1.515995

C

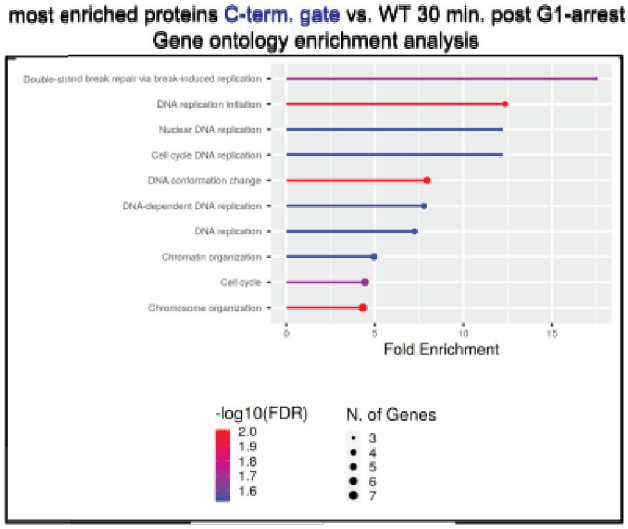


Figure 61: Analysis of the most enriched proteins in multiple pairwise comparisons. A) Tables showing the 25 most enriched proteins when comparing tagged pull-downs against the untagged negative controls for each condition at 30 minutes post G1-arrest. Proteins that are part of the manually curated set of replication factors are coloured according to the colour scheme used in prior figures. Enrichment values are the mean enrichments across four replicates. Where a protein was not detected in the untagged control, conservative imputation was used to assign a value equal to the mean of the lowest 10% of detected proteins in the untagged control. B) As in A, but comparing the two linkage construct arrest conditions against the strain without a linkage present, also at 30 minutes post G1-arrest. C) Gene ontology enrichment analysis for the subset of proteins enriched in the C-terminal gate samples versus the tagged WT samples. All proteins with a log_2(enrichment) greater than 1 were submitted for analysis using the ShinyGO (Ge et al., 2020) online tool, with all other detected proteins as the submitted background. Shown are the GO molecular function terms that were significantly enriched, with an FDR set to 0.05.

Initially, the enrichment values were simply calculated for the pull-downs over their untagged controls at 30 minutes (Figure 61A). Unsurprisingly, the most enriched proteins in the cells without a linkage construct showed a similar set of factors to the prior experiments using Mcm4-FLAG (Figure 38C). The extended list of enriched factors, which contains over 300 proteins with an enrichment of two-fold or greater, contains many protein factors that could be linked to replicative processes either by known involvement, or by involvement with adjacent process such as DNA damage signalling or repair. Three additional factors beyond the manually curated set of replication factors were present in the 25 most enriched proteins, Ddc1, Asi1 and Nat5. Ddc1 is a component of the yeast 9-1-1 sliding clamp, and is involved in the DNA damage response pathway responding to double-stranded breaks, where, interestingly, it cooperates with Dpb11 (Longhese et al., 1997; Hong & Shirleen Roeder, 2002; Wang & Elledge, 2002; Burgess, Han & Zhang, 2014). Asi1 is an E3 ubiquitin ligase that is localised in the inner nuclear membrane ER sub-compartment, and is involved in degradation of misfolded proteins there (Zargari et al., 2007; Foresti et al., 2014; Khmelinskii et al., 2014). Other than an interaction with the Cdc48 protein, that also functions in replication termination (Pantazopoulou et al., 2016; Maric et al., 2014, 2017), the possible link to replication is unclear. Not much is known about Nat5, but it is classified as one of a distinct family of N-terminal acetyltransferases based on sequence similarity to other enzymes of this type, with no substrates identified as of yet (Polevoda & Sherman, 2003). This protein was only sporadically detected in the pull-down samples, and was not present in untagged controls, so represents an uncertain hit.

The C-terminal linkage construct sample contains many of the same proteins on the list of most enriched, as the “wild type” does, and they broadly fall under the same categories of molecular function. The top 25 enriched proteins are also similar, but the increased enrichment of the transient replication factors over the “wild type” is clear, as Sld2, Sld3, Sld7 and Dpb11

all make the top 25 factors for this condition (Figure 61A). Of the top 25 factors, there are a few outside of the expected curated set. Nat5 is once again present, along with Pdr11, Cac2, Pta1 and Pcc1. Pdr11 is a drug transporter and so natively exists in a separate cellular compartment, and so is unlikely to be a genuine hit (Buechel & Pinkett, 2020). It also has no quantification in the untagged controls, so could have an inflated enrichment value due to the imputation process. Pta1 is involved with the cleavage and polyadenylation of mRNA and tRNA 3' ends as part of the yeast homolog of cleavage-polyadenylation specificity factor (O'Connor & Peebles, 1992; Zhao et al., 1999). Despite the possible proximity of transcription-related processes and DNA replication events, it is also unclear why this protein would be associating with replication machinery, particularly at a complex stalled during an initiation step. Pcc1 is a member of a conserved but enigmatic nuclear complex that has been implicated in telomere length maintenance and tRNA modification (Srinivasan et al., 2011; Daugeron et al., 2011; Hu et al., 2013). Understudied complexes present avenues for possible new functions, but the other members of the Pcc1-containing complex, Kae1, Bud32 and Cgi121 are only enriched slightly over untagged control, if they are present at all, making this hit also somewhat dubious.

Cac2 is a member of the CAF-1, the complex required for assembly of new histones onto newly replicated DNA (Smith & Stillman, 1989; Krude, 1995). This complex has a clear role in replication, and is associated with the replication fork by interactions with Pol30 (Krawitz, Kama & Kaufman, 2002; Zhang, Shibahara & Stillman, 2000a). Because of the nature of the arrested complex present in the C-terminal linkage construct sample, it was surprising that a member of this complex showed association at this stage, prior to the successful formation of replication forks. Because this associated was unexpected, and was at a stage where initiation was artificially prevented from completion, Cac2 was taken forward for further investigation with regard to a role in replication initiation.

Despite having a similar number of enriched factors, the N-terminal linkage construct most enriched proteins contained far fewer factors that are related to DNA replication or DNA repair (Figure 61A). This is likely due to the fact that the arrest state generated by this construct is very similar to the inactive double-hexamer prior to the start of activation, only a few of the initial activation factors have begun to associate at the stage it was stalled. Sld3 and Cdc45 are both enriched, being early stage initiation factors, although the association of Mcm10 is still surprising. For additional factors not part of the curated list of replication factors, it shares several with the C-terminal arrest state, such as Pdr11, Pta1 and Pcc1. Of the other factors present in this list, the most interesting hits are Far1, YDR132C, and Yrf1-X.

Perhaps most interesting as a method to identify initiation-specific factors is the comparison between the two arrest states generated by the linkage constructs, and the "wild type" 30 minute condition. This would be able to identify any factors that were highly abundant at the arrested states, but were then lost during the transition into a fully active replication complex. Because both samples were pull-downs, and neither was a negative control, imputation could lead to inflated enrichment values at a much higher rate than normal, increasing the noise in the data significantly. The most enriched proteins presented in Figure 61B are therefore only those factors that were at least detected in both conditions.

Unsurprisingly, the C-terminal construct most-enriched list is headed by Sld2 and Dpb11, as explained previously. Mcm10 is also near the top of the list. Other than the presence of Mcm4 and Orc3, which have also been discussed in the explanation of the previous figures, there are no other members of the curated list present. The N-terminal construct most-enriched list contains none of the curated replication-associated factors (Figure 61B), but does include some interesting proteins, such as the yeast Ku70 protein, which is involved in telomere maintenance and double-strand break repair (Doherty & Jackson, 2001). Other potential factors of interest included Rbd2, of which very little is known (McQuibban, Saurya & Freeman, 2003).

However, many of these factors showed differences because of noise, as they were not particularly, or even negatively enriched in the C- or N-terminal construct pull-downs themselves. For example, the APC/C subunits (e.g. Apc2) shown here actually have high de-enrichment in the pull-downs due to non-specific binding in the untagged negative controls. When controlling for enrichment within the individual pull-downs, very few factors show an enrichment of magnitude, other than the already known and categorised replication-associated proteins. Those proteins with a $\log_2$(enrichment) of 1 or higher in the C-terminal gate construct were compared against all proteins identified across the samples, and gene ontology enrichment was carried out on this subset versus the larger set of identified proteins using the ShinyGO tool (Ge et al., 2020) (Figure 61C). Few terms were found to be significantly enriched in this subset, and of those, many pertained to DNA replication or cell cycle associated processes. However, the presence of Dpb11, Sld2, Mcm10 and Mcm5 in this subset alone was enough to cause a significant enrichment of several of these GO terms. For the other factors present in this subset, then, there are likely no GO terms that show significant enrichment. For the N-terminal gate construct, no terms met the significance threshold with an FDR set to 0.05.

Despite the lack of identification of a prominent novel factors, the ChIP-MS analysis performed here in combination with the rapamycin-dependent intra-hexamer linkages between Mcm2 and Mcm5, helped elucidate the complex composition when the linkage was utilised to arrest complexes during initiation. For the most part, the data confirmed the results of the ChIP-qPCR data, but expanding the dataset to include the association dynamics of many more replication factors in the context of these initiation arrests. A small number of factors were identified that either showed more association than might previously have been expected, or showed an association which was entirely unexpected. Some of these factors were taken forward for validation and to investigate their roles in the process of replication initiation.

3.7.2 Mcm4-FLAG ChIP-MS in the context of the inter-hexamer latch construct

The Mcm4-FLAG ChIP-MS experiment in the context of the intra-hexamer gate constructs yielded a detailed picture of complex composition at the arrest state generated by these constructs. This technique was able to quickly validate the ChIP-qPCR data of replication factors, whilst also monitoring the association of a much greater range of replication factors than would be possible only by iterating ChIP-qPCR analysis over multiple target factors. As part of this research, a new construct was designed and tested (Section 3.1.2). This rapamycin-dependent inter-hexamer latch had been tested with a limited number of replication factors, and was thought to generate a unique arrest state that was distinct from either of the intra-hexamer gate-induced arrest states. This arrest state was also thought to be within the process of replication initiation. The use of ChIP-MS on this construct arrest could therefore be hoped to quickly and comprehensibly investigate the complex composition at this arrest state. A large-scale experiment was carried out in quadruplicate to characterise the protein composition of the latch construct arrest, and comparator gate construct pull-downs were included to compare this arrest state to that investigated previously.

Following analysis of the data generated in this experiment, specifically for the expected set of replication-associated proteins, it was observed that the pull-down in this experiment was not of as good quality as previously. Whilst MCM2-7 subunits were consistently highly enriched, the levels of other expected factors in the experiment were highly variable or not detected. Even the co-IP of closely-associated proteins was unexpectedly inefficient. The enrichment of Cdc45 was much lower than expected, and the GINS complex was not detected at all, which stood in stark contrast to the strong enrichments seen in previous experiments of this type. For other components of the replisome, and initiation factors, the enrichments were also disappointingly low, well below previously observed levels. This was true even in the intra-hexamer gate construct, so was not a problem uniquely introduced by the inter-hexamer

latch. Instead, the experimental conditions for this analysis had somehow caused this issue for both types of construct.

From the analysis of the data, it seemed as though the large majority of MCM2-7-containing complexes pulled-down in this experiment were not chromatin associated. The disproportionate loss of chromatin-associated MCM2-7 was something that was thought to affect the latter time-points of the Mcm4-FLAG ChIP-MS experiment previously and appears to have affected this experiment particularly severely. Again, the hypothesis is that incomplete sonication and/or improved DNA-crosslinking of the arrested replication complex resulted in a disproportionate loss of chromatin-bound replication complexes from the pull-down. Following this experiment, attention was focused on ensuring a reproducible quality of pull-downs, before attempting to repeat this experiment. Full characterisation of the inter-hexamer latch, then, is a goal that will unfortunately fall outside of the attainment of this project by the time of this thesis' completion.

3.8 ChIP-MS validation, troubleshooting and optimisation

In progressive experiments carried out as part of this project, it was observed that utilising formaldehyde cross-linking and washing using buffers of salt and detergent levels commonly used for ChIP-qPCR protocols seemed to give improved results for mass spectrometry-linked pull-downs. Whilst pull-downs under native conditions were able to isolate the core CMG very well, and also identify a range of other replisome-associated factors in addition (Section 3.2, Figure 32D), the change in protocol to the ChIP-MS method led to the identification of further replication-associated factors and in particular, several initiation factors such as Sld2, Sld3, Sld7 and Dpb11 (Section 3.3 and 3.4). It was important to confirm that this improvement in factor detection was in fact due to this change in technique, and not due to small optimisations in the general cell harvesting, sample handling or preparation for mass spectrometry analysis.

Additionally, later ChIP-MS experiments identified a prevailing issue in the experimental technique, where pull-downs were losing not correlating with previous runs, and certain known interactors were increasingly absent from the mass spectrometry identification. The possible source of this issue was thought to be an inefficiency in the release of chromatin-associated complexes from chromatin. Variability in either the formaldehyde-induced cross-linking, or in the efficiency of sonication could have resulted in large pieces of intact chromatin remaining in the sample. Over-crosslinking could compact the chromatin to a point where it would almost inevitably pellet out of solution during clarification steps, or insufficient sonication could result in large chromatin fragments that would also be lost. Overall this would result in an over-abundance of soluble bait complexes at the expense of chromatin-bound complexes, which is indeed what was observed in some of the later experiments. To test whether the current ChIP-MS protocol was an improvement over native IP, and to test if the loss of chromatin-bound protein could reduced, an experiment was carried out to directly compare the native IP method against the new cross-linking method and an optimisation to the ChIP-MS. All conditions used

the Cdc45-FLAG bait protein, so as to be comparable. Cdc45-FLAG was chosen since native IP had only been attempted with Cdc45-FLAG (data not shown) and Psf2-TAP. A regular ChIP-MS would be carried out alongside a native IP, and a modified ChIP-MS protocol with the addition of Benzonase digestion immediately prior to sonication, to potentially help with chromatin disruption.

In addition to the native IP carried out as for the Psf2-TAP pull-downs demonstrated earlier, there was also a sample that was treated the same as the ChIP-MS sample, only without the cross-linking. This was to ensure any differences were not only due to the modified preparation between native IP and ChIP-MS. This was because there were several differences between the original protocol and the ChIP-MS one (see materials and methods for full details). The most notable differences were in the method of cell lysis, the method of chromatin disruption, presence or absence of clarification, and the wash buffers. The ChIP-MS protocol used bead-beating instead of freezer-milling for cell lysis, sonication instead of Benzonase treatment for chromatin disruption, did not make use of an ultracentrifugation step, and used more stringent wash steps using buffers with a generally higher detergent concentration. Also present was a negative control sample prepared in the same manner as for the ChIP-MS, but in an untagged strain. 500 mL of cell culture grown and arrested in alpha-factor and subsequently released for 40 minutes was the input for all the conditions. After pull-down steps were complete, and isolated protein eluted from magnetic beads using boiling in the presence of 1% SDS, the samples were prepared for mass spectrometry analysis in the same manner.

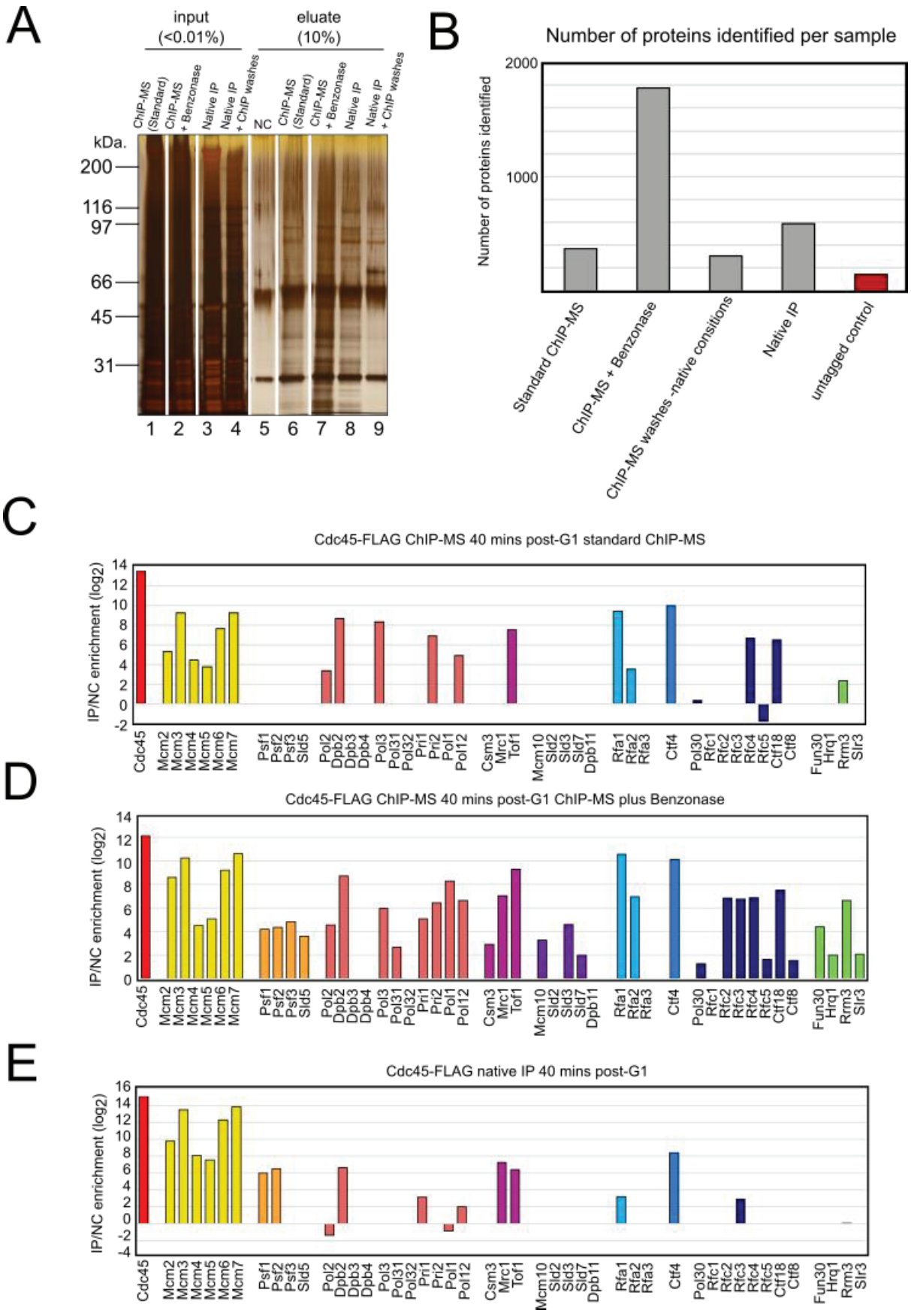

A
input (<0.01%)
eluate (10%)
ChIP-MS (Standard)
ChIP-MS + Benzonase
Native IP
Native IP + ChIP washes
NC
kDa.
200
116
97
66
45
31
1 2 3 4 5 6 7 8 9
B
Number of proteins identified per sample
Number of proteins identified
2000
1000
Standard ChIP-MS
ChIP-MS + Benzonase
ChIP-MS washes -native consitions
Native IP
untagged control
C
Cdc45-FLAG ChIP-MS 40 mins post-G1 standard ChIP-MS
IP/NC enrichment (log2)
D
Cdc45-FLAG ChIP-MS 40 mins post-G1 ChIP-MS plus Benzonase
IP/NC enrichment (log2)
E
Cdc45-FLAG native IP 40 mins post-G1
IP/NC enrichment (log2)
Cdc45 Mcm2 Mcm3 Mcm4 Mcm5 Mcm6 Mcm7 Psf1 Psf2 Psf3 Sld5 Pol2 Dpb2 Dpb3 Dpb4 Pol3 Pol31 Pol32 Pri1 Pri2 Pol1 Pol12 Csm3 Mrc1 Tof1 Mcm10 Sld2 Sld3 Sld7 Dpb11 Rfa1 Rfa2 Rfa3 Ctf4 Pol30 Rfc1 Rfc2 Rfc3 Rfc4 Rfc5 Ctf18 Ctf8 Fun30 Hrq1 Rrm3 Slr3

Figure 62: Analysis of the comparative ChIP-MS and native IP samples. A) Silver stain visualisation of samples run on SDS-PAGE. Inputs refer to cell extracts immediately prior to addition of IP antibody, and eluates refer to samples after elution from beads immediately prior to preparation for mass spectrometry submission. Sample 1: regular ChIP-MS, sample 2: Optimised ChIP-MS with addition of Benzonase®, sample 3: non-cross-linked sample processed using ChIP-MS sample prep and washes, sample 4: Native IP as in section 3.2. B) Comparison of the total number of proteins identified in each sample, based on quantified LFQ values. C-E) Enrichment analysis of the set of manually curated replication factors across three pull-down samples. Enrichment was calculated using LFQ values from pull-down samples for each protein and comparing to LFQ values in the untagged control sample. Imputation was carried out manually where missing values were present in the untagged control using the mean LFQ intensity of the lowest 10% of detected proteins in the untagged control. C) Enrichment of replication factors in the regular ChIP-MS protocol. D) Enrichment of replication factors in an optimised ChIP-MS protocol utilising Benzonase®. E) Enrichment of replication factors in an original native IP in this same experiment.

From silver stain analysis of the samples run on SDS-PAGE, clear differences can already be seen (Figure 62). The lack of cross-linking is visible in the inputs of lanes 3 and 4. Strikingly, in the eluates, lane 7, ChIP-MS with the addition of Benzonase® prior to sonication, shows a much greater protein content than the other samples. All tagged pull-down lanes also show greater protein levels than the untagged negative control. The number and clarity of pull-down specific bands in lane 8, the native IP, is less than observed previously (Section 3.2). This had been observed in the native FLAG-tag pull-down of Cdc45 in prior tests (data not shown) as the elution method of boiling from beads using SDS is less clean and specific than the use of PreScission® protease to elute target complex. In terms of total numbers of proteins identified, the most obvious effect, again, was that the Benzonase®-treated ChIP-MS sample contained many more proteins at a high enough abundance for MS detection (Figure 62B). Surprisingly, the native IP showed a greater number of detected proteins than the regular ChIP-MS protocol, indicating a more complex pull-down than expected with this sample. It is worth noting that the numbers of proteins identified in both samples, however, was lower than what was usually detected in regular ChIP-MS experiments. The untagged control demonstrated the lowest number of identified proteins, and the non-cross-linked cells processed using the ChIP-MS method otherwise showed a number of proteins lower than the regular ChIP-MS. This was actually surprisingly high, since ChIP-MS washes were thought to be much more stringent, and so in the absence of cross-linking, fewer factors were expected to be pulled-down. To investigate how abundant the proteins detected in these samples actually were, and specifically

for the replication factor curated set, enrichment values were calculated and plotted for the samples (Figure 62C-E). Firstly, the results for the regular ChIP-MS were disappointing, reflecting the issues that had been affecting this technique in several of the most recent ChIP-MS experiments. Many expected replication factors were not present in the sample, including many GINS and polymerase subunits that should have been associated with Cdc45 at the time-point tested. However, a clear difference was seen in the sample treated with Benzonase® (Figure 62D). The enrichment data for this sample resembles much more the initial ChIP-MS data obtained in Section 3.3. The native pull-down in this experiment also underperformed, with many replication factors missing, and the reason for this is not entirely clear (Figure 62E). The sample not treated with formaldehyde, but processed as a ChIP-MS sample after that point, showed enrichment of almost none of the replication-associated factors, only showing clear enrichment for Cdc45, Ctf4 and Pol3. This showed that it was clearly the introduction of cross-linking, in combination with the stringent washes of the ChIP-MS protocol, that resulted in maintained association of replication factors in the pull-down. Without it, the stringent washes removed many more associated factors, as expected.

A quick analysis of the variation in LFQ intensities across these samples was also carried out for these samples (Figure 63). In terms of Cdc45 abundance in the sample, the native IP performed the best (Figure 63A, bar 3). Since all pull-downs in this experiment relied upon the Cdc45-FLAG-anti-FLAG interaction, this was likely due to the less stringent washes removing fewer Cdc45 proteins from the beads in the native IP. Otherwise, the results were mainly as expected, though the Benzonase®-treated sample (Figure 63, bar 2) showed surprisingly low abundance of Cdc45 relative to the regular ChIP-MS sample (Figure 63, bar 1). Reassuringly, the non-cross-linked sample (Figure 63, bar 4) showed abundance of Cdc45 roughly equal to that of the regular ChIP-MS. Since the Cdc45-FLAG-beads interaction remained the same, and the washes were the same in both cases, this was expected.

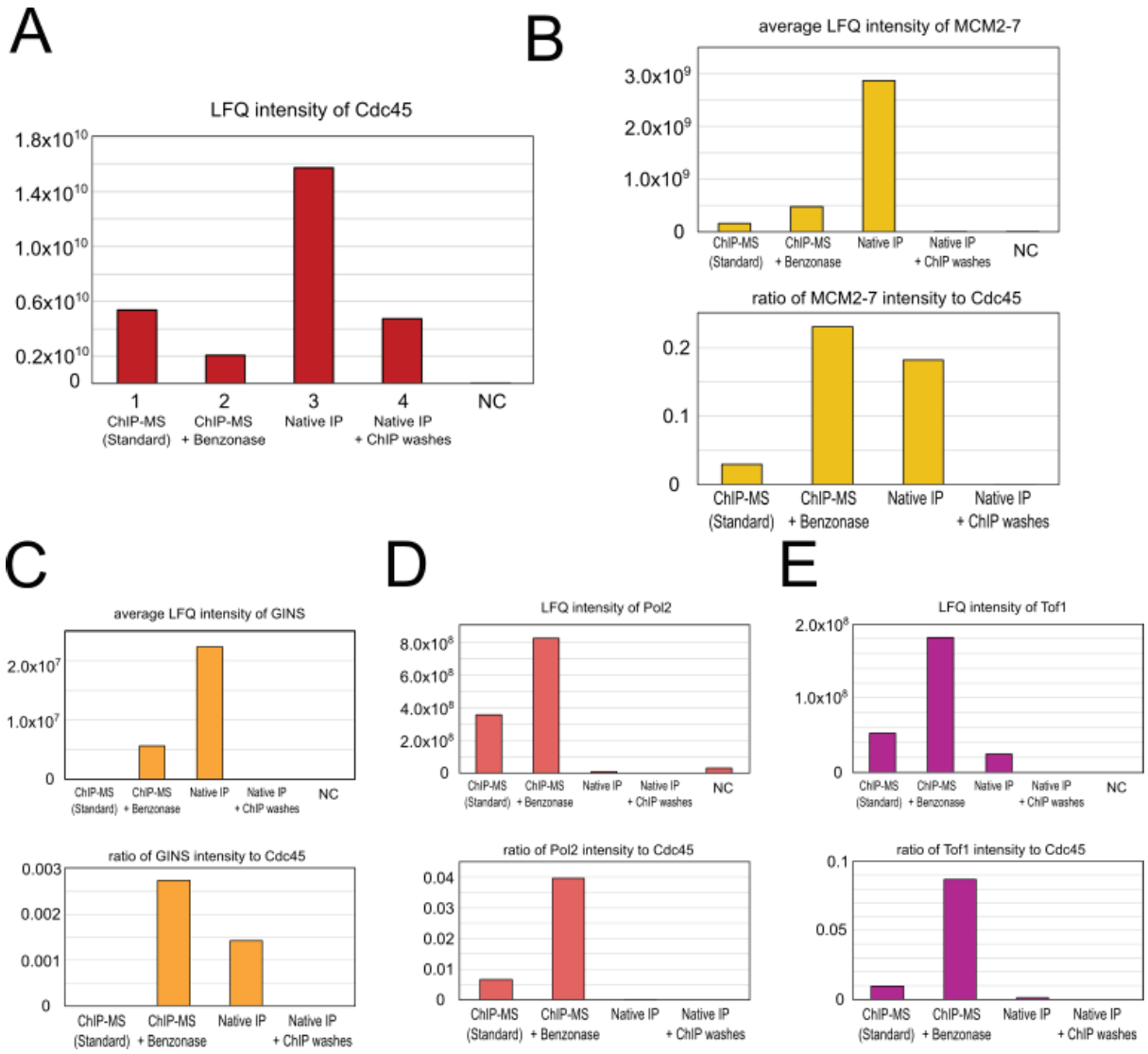


Figure 63: Individual protein abundance analysis across test conditions of ChIP-MS and regular IP. LFQ intensity values of the specified proteins are displayed relative to one another. For proteins other than Cdc45, the relative ratio of that protein to Cdc45 in the pull-down samples is displayed below. A) LFQ intensity of Cdc45 across the samples. B) Average LFQ intensity of MCM2-7 across the samples, and the ratio of this intensity to Cdc45 in the pull-down samples. C) As in B, but for the GINS complex. D) LFQ intensity of Pol2 across samples, and the ratio of Pol2 to Cdc45 in the pull-downs. E) As in D, but for Tof1.

This analysis was also repeated for other replication-associated proteins. Similarly to Cdc45, MCM2-7 showed the greatest abundance in the native IP (Figure 63B). This was also as expected, the interactions between Cdc45 and MCM2-7 are stable and are clearly not significantly disrupted in the native IP. Tellingly, this is where the Benzonase®-treated ChIP-MS sample showed an advantage over the regular ChIP-MS, indicative of loss of Cdc45-MCM2-7 complexes in the regular ChIP-MS. The non-cross-linked ChIP-MS sample showed

very little MCM2-7 abundance in comparison to the other samples, which suggested that the more stringent washes were capable of disrupting CMG complexes when cross-links were not present.

The ratio of MCM2-7 abundance to Cdc45 abundance was also calculated to give information on MCM2-7 pull-down efficiency normalised to Cdc45 amount. The native IP and Benzonase®-treated ChIP-MS performed similarly, indicating about the same proportion of Cdc45 was associated with an MCM2-7 hexamer in each pull-down.

The GINS complex showed variable presence in these pull-downs, and subunits were only detected in the Benzonase®-treated ChIP-MS and the native IP. Comparing the average abundance of the detected subunits shows a similar trend to MCM2-7 between the two samples which showed GINS presence (Figure 63C). Whilst by abundance, the native IP again performed better, the ratio of GINS to Cdc45 was slightly higher in the modified ChIP-MS protocol.

Moving on to replication factors that showed less direct association to Cdc45, but were still associated with the replisome demonstrated that in this case, ChIP-MS appeared to be superior to native IP, and that the Benzonase®-treated ChIP-MS protocol improved further and helped combat some of the recent issues. Pol2 abundance was drastically lower in the native IP than in either of the cross-linked samples, and as a result, showed a much better ratio to Cdc45 in the ChIP-MS protocols, with the modified protocol performing better (Figure 63D). Tof1 showed a similar pattern to Pol2 and several other more peripheral replication factors such as Pol1, Pol30 and RFC (data not shown). Tof1 abundance was higher in the ChIP-MS samples, and highest in the Benzonase®-treated ChIP-MS (Figure 63E). As such the protocol showing the most efficient pull-down of this protein was again the modified version of the ChIP-MS protocol, which pulled-down substantially more Tof1 relative to the Cdc45 abundance than the other samples.

This seemed to be a clear indication that somehow, it was the presence of DNA and chromatin in the regular ChIP-MS that was causing the issues in previous runs of ChIP-MS, as predicted by the bias towards soluble complexes in recent experiments. Given this data, it is likely that the sonication of the ChIP-MS samples over recent experiments was inefficient, resulting in large tracts of intact chromatin that would be lost during spin steps prior to the pull-down. Whilst data from later ChIP-MS experiments was still of enormous value, it is possible that with the addition of Benzonase®-treatment, results would become more reproducible and the isolation of chromatin-bound replication complexes would be more effective.

3.9 Replisome-associated proteins observed in ChIP-MS

Several factors showed enrichments or abundance dynamics in one or several of the replication-complex associated pull-downs, to merit further investigation of their possible role in replication. By focusing on specific factors, the aim was to hopefully validate the results of the ChIP-MS data for these factors, whilst also revealing potentially novel functionality. An initial step was to confirm whether or not FLAG-tagged versions of these factors of interest were able to show an origin-association via ChIP-qPCR methods. Additionally, deletion strains were also prepared to test if the null mutants showed any noticeable phenotype generally (Figure 64), but crucially whether they showed any specific phenotype relating to replication and replication initiation.

Fun30 was a protein that was reliably detected and strongly enriched in ChIP-MS experiments in this study. It often showed very high levels of enrichment, and analysis of the dynamics of MCM2-7 association for this factor showed an interesting pattern (Figure 53). This protein appeared to show maximum association with MCM2-7 from the point of loading at origins and showed continued association through to time-points where active replisomes were present (Figure 39), but with decreased association. This pattern indicated that Fun30 associated with MCM2-7 during or after loading, and whilst it was possibly able to travel with the active replisome, the decrease in association over time post-activation indicated this association was most likely limited to origins of replication. Because of the role of Fun30 as a nucleosome remodeller (Lee et al., 2017; Lee, Choi & Lee, 2018; Costelloe et al., 2012; Neves-Costa et al., 2009), it was considered that Fun30 may have a role in shaping the origin chromatin environment which could modulate either the loading or activation step of MCM2-7, and was therefore chosen for further investigation.

There is some evidence that in addition to helping resolve fork-stalling at replication blocks (Torres, Bessler & Zakian, 2004; Ivessa et al., 2003; CF et al., 2019), Rrm3 also interacts with

members of the replisome and may even associate with the fork in the absence of replication fork blocks, though this is disputed (Azvolinsky et al., 2006; Schmidt, Derry & Kolodner, 2002). The association of Rrm3 with MCM2-7 complexes was clear in this study, and dynamics of the association (Figure 53) as well as lower association with the rapamycin-dependent linkage constructs (Figure 60), indicated that this factor indeed preferentially associates with activated replisome complexes. The magnitude of the enrichment observed over control samples suggested that this association is possibly with a large proportion of replisomes, rather than a small subset encountering blocks, suggesting an ubiquitous interaction with active forks, without the need of fork blocks. This factor was therefore investigated further to confirm this association and identify any additional functions of this protein.

Sir3 was another protein that was reproducibly enriched in many of the ChIP-MS pull-downs in this study. Though canonically this protein is associated with establishment and spreading of heterochromatin at telomeres and the mating type locus, it and Sir2 have been linked to replication origins in their ability to compensate for MCM loading defect mutations and in the changing origin firing patterns upon Sir2 deletion (Pappas, Frisch & Weinreich, 2004; Crampton et al., 2008). Despite this previously reported indication for Sir3 association with origins, the high enrichment values reported via ChIP-MS made this protein an interesting target for further investigation.

Though Hrq1 showed varying abundance in regular ChIP-MS with either Cdc45-FLAG or MCM4-FLAG, this protein was one of the few proteins with previously suggested chromatin association that was enriched in the arrested intermediate complex generated by the intra-hexamer C-terminal gate construct (Figure 60D). Tracking the intensity of this protein over time in the time-course ChIP-MS showed that it appeared to associate with the replication complex at origins, prior to activation (Figure 53). Though there is little evidence of Hrq1's involvement with replicative processes, the RecQ4 helicase in humans contains a region

homologous to replication factor Sld2 in addition to the Hrq1 homology domain (Rogers et al., 2017). Hrq1 itself appears to mainly act in repair of interstrand cross-links independently of replication (Sanders et al., 2020; Rogers et al., 2020a, 2020b, 2020c), but its unique position as an enriched factor in an initiation-arrest complex and identity as a helicase with a link to RecQ4's Sld2 homology domain made it worthy of further study.

Cac2 is a subunit of the CAF-1 complex, which in yeast has been associated with telomeric and other heterochromatin silencing, as well as deactivation of the DNA damage checkpoint after successful repair of the damage . Despite no prior association with replication origins, Cac2 showed a high imputed enrichment value in the wild type Mcm4-FLAG ChIP-MS in G1. The association of CAF-1 with replication-coupled histone deposition is clearer in humans than in yeast, but this role could link Cac2 with a possible role in histone dynamics at origins. Whilst not well enriched in the time-course ChIP-MS, Cac2 was still considered for further experiments.

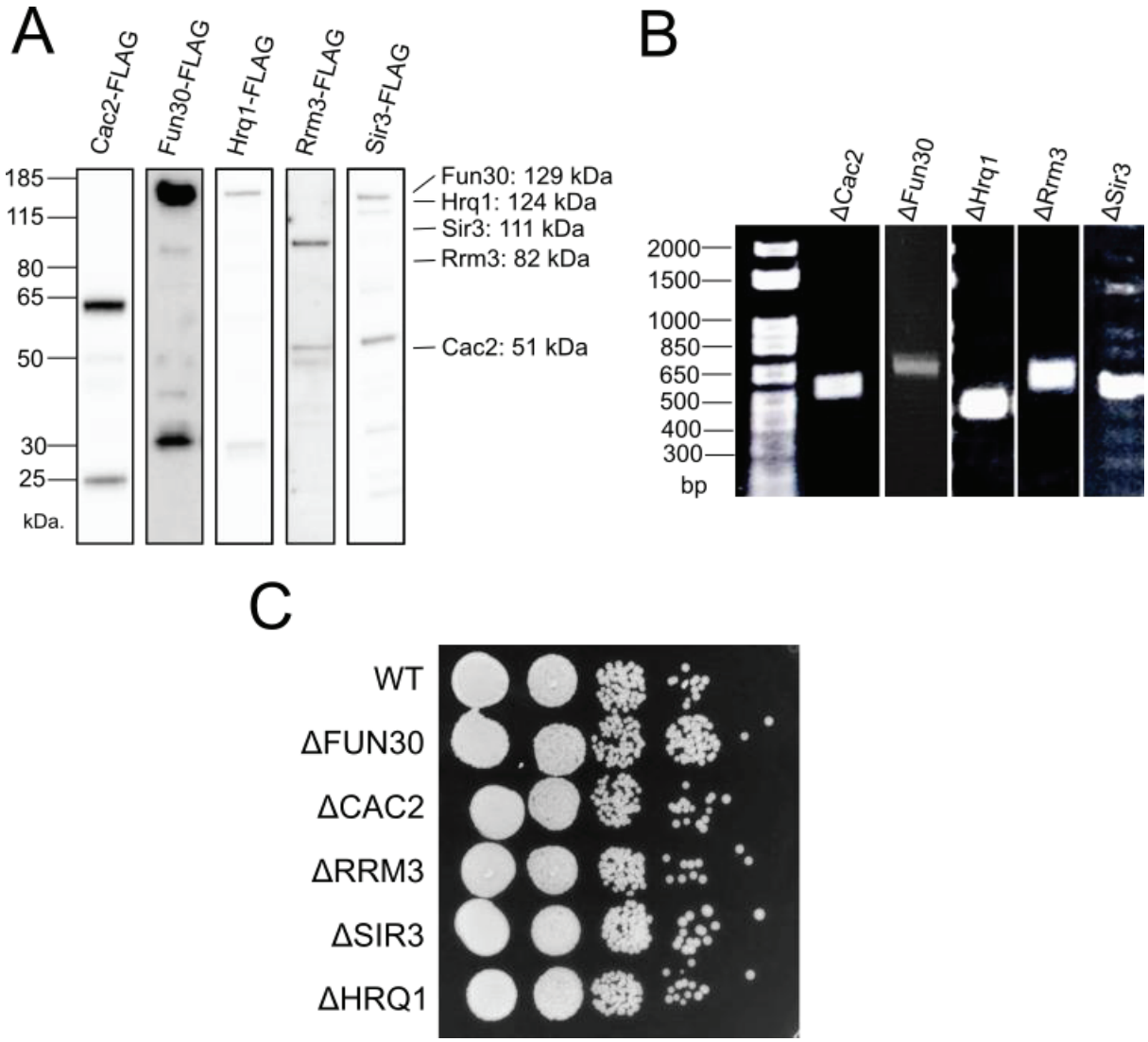


Figure 64: Confirmation and viability of strains generated to investigate factors of interest. A) Western blot of SDS-PAGE analysis of cell extracts prepared from clones from hygromycin-containing media to select for integrated penta-FLAG tags at the C-termini of the genes displayed. Membrane was treated with HRP-coupled anti-FLAG antibody. All major bands match well with the expected sizes of the FLAG-tagged proteins, within the range possible due to posttranslational modifications. Expected 5xFLAG size: approx.. 5 kDa B) Colony PCR analysis of genomic DNA isolated from positive colonies selected on hygromycin-containing media to select for deletion of the endogenous copy of the displayed genes. Primers used were #2369, which anneals to the Hygromycin resistance gene in the integration cassette, and primers that anneal to genomic DNA to the 3' of the coding region of the gene of interest. These were #2781, #2685, #2782, #2783 and #2780 corresponding to the genes CAC2, FUN30, HRQ1, RRM3 and SIR3 respectively. C) Dot-spot viability assay of the deletion mutants generated and confirmed in B. 10-fold dilutions of cells from the same initial quantity were grown on YPD plate to asses any intrinsic growth defects caused by gene deletion. "WT" is YC160, the in-house W303 base strain.

3.9.1 Fun30

To attempt to confirm the association of the Fun30 protein at replication origins and potentially at replication forks as well, ChIP-qPCR was carried out using a endogenously tagged Fun30-5xFLAG to isolate chromatin associated with this protein. Cells were synchronised in alpha-factor before being released in a manner identical to that described in section 3.1, and at specific time-points during and after release, cells were harvested and ChIP-qPCR performed. Unfortunately, initial results proved unpromising (Figure 65A). Although a general signal of origin DNA was observed at origins, it was extremely variable in initial tests, and showed indistinguishable association with the non-origin control region. This was despite showing more DNA signal than an IgG control. Whilst the ChIP-MS association of Fun30 with MCM2-7 appeared to be maximal prior to activation (Figure 53), there was still some association visible post-activation. During this time period, research implicating the Fun30 homolog SMARCAD1 in protection of replication forks from certain DNA damage was published (Lo et al., 2021). It was considered if Fun30 origin-proximal association would be increased in the presence of hydroxyurea (HU), which can both slow/stall replication forks, and induce DNA damage (Koç et al., 2004; Alvino et al., 2007; Singh & Xu, 2016). The ChIP-qPCR experiment was repeated with cells released into S phase in the presence of 200 mM HU, and under these conditions, a small increase in Fun30 origin association was seen over time, to a maximum at 45 minutes (Figure 65B). The timing of the association increase matched well with the association timings of replication factors at this origin (Figure 65C), particularly given that the HU treatment was likely to stall forks close to the origin. Notably, this time-dependent change in DNA pull-down was not replicated in the non-origin control DNA, for Fun30 or Cdc45 (Figure 65B,C). This provided evidence for the association of Fun30 with replication forks, but did not fully validate the ChIP-MS data, which suggested association with MCM2-7 mainly prior to replication initiation.

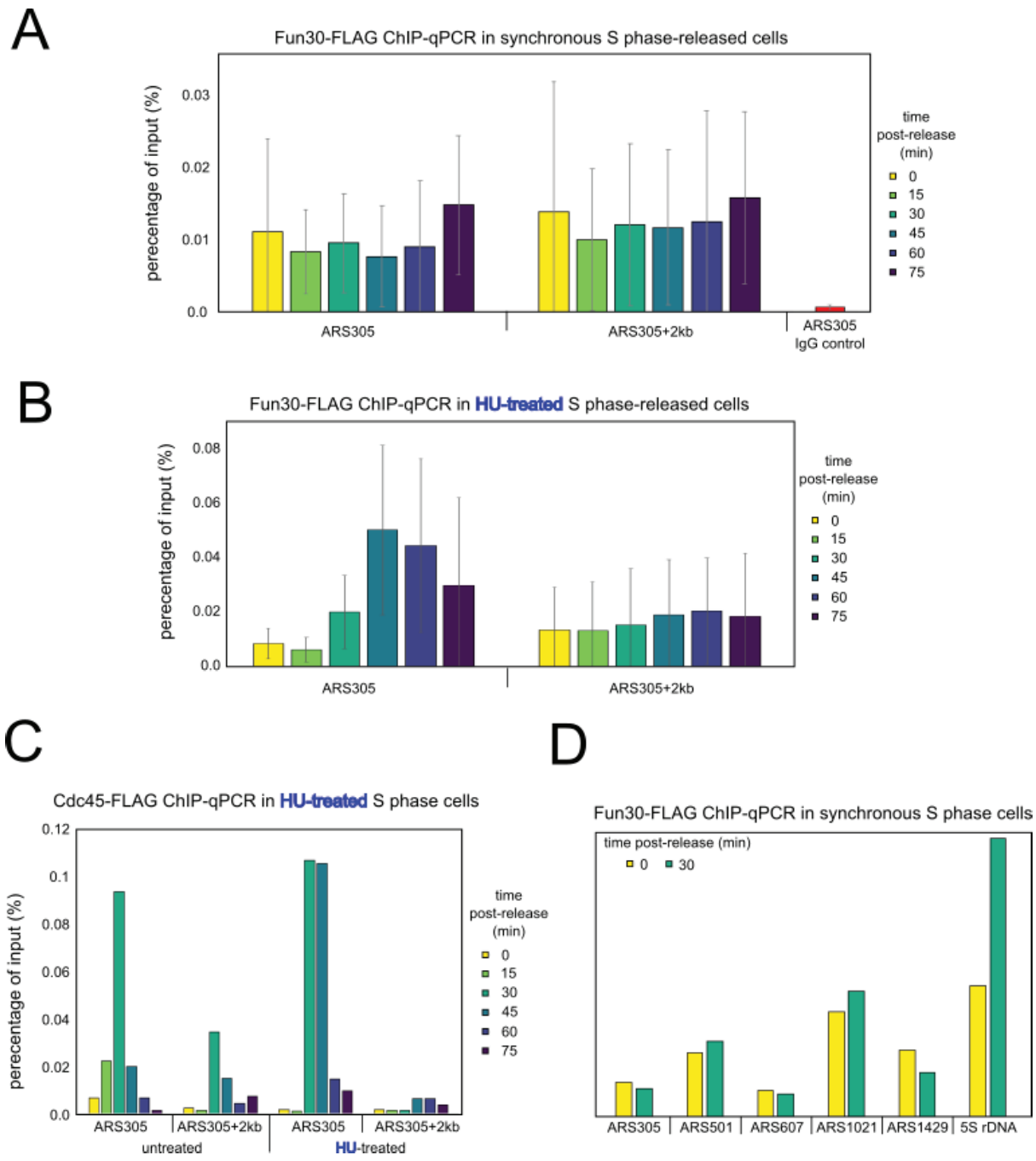


Figure 65: ChIP-qPCR analysis of Fun30-FLAG association with origins. A) ChIP-qPCR of Fun30-FLAG cells released from alpha-factor arrest. The IP efficiency of ARS305 origin DNA, and of the non-origin ARS305+2 kb DNA is shown. Also shown is the IP efficiency of ARS305 from the same strain but using a non-specific IgG antibody for the pull-down step. B) As in A, but cells were treated with 200 mM hydroxyurea (HU) to prevent nucleotide incorporation and stall replication forks. C) As in B, using a control Cdc45-FLAG strain to test the effect of the HU treatment on a known replication factor.(n=1) D) Further data from a replicate of A, testing alternative origins and genomic regions (n=1). The 5S rDNA locus is used as a control for a site of highly active transcription, and therefore high DNA accessibility.

It is possible that the strong association of Fun30 with Mcm4 could be due to an interaction at a subset of origins prior to initiation. To investigate this possibility, the initial qPCR was repeated for other replication origins and genomic regions. None of the other origins tested had markedly different pull-down efficiency in the Fun30-FLAG ChIP-qPCR, and there was no observable pattern between early (ARS305, ARS607, ARS1021) and late (ARS501, ARS1429) origins (Raghuraman et al., 2001). The pull-down of the highly expressed locus of the 5S rRNA was the most efficient of the genomic loci tested. Given Fun30 has a proposed function in histone remodelling around the RNA polymerase complex (Lee et al., 2017; Lee, Choi & Lee, 2018), this enrichment could be due to this role of Fun30. High pull-down efficiency is also observed with highly expressed regions of open chromatin in ChIP-based experiments, and so this could also reflect this effect (Teytelman et al., 2013). As with several of the factors in this section, the absence of ChIP-qPCR signal need not be taken as a definitive lack of association. A possible issue with these proteins is that the epitope tagging may disrupt the binding of the factor in the replication-associated context and prevent origin association.

To investigate whether the deletion of Fun30 would have any noticeable effect on the timing of DNA replication initiation, the progression or speed of replication itself, or the loading of MCM2-7 at origins, strains that had their endogenous copy of Fun30 deleted were created. Cells without Fun30 were then tested for any phenotypic changes relative to control strains (Figure 66). Initially the progression of cells with or without endogenous Fun30 into and through S phase was investigated using DNA content quantification by SYTOX incorporation and flow cytometry. In cells released into S phase from alpha-factor arrest, there was no discernible effect on the timing of S phase entry or the completion of S phase in *Δfun30* cells at the population level (Figure 66A). Previous studies suggested that Fun30 had a protective role in the response to genotoxic agents (Bi et al., 2015), and so the cell cycle analysis was repeated but with the presence of 200 mM of HU (Figure 66B).

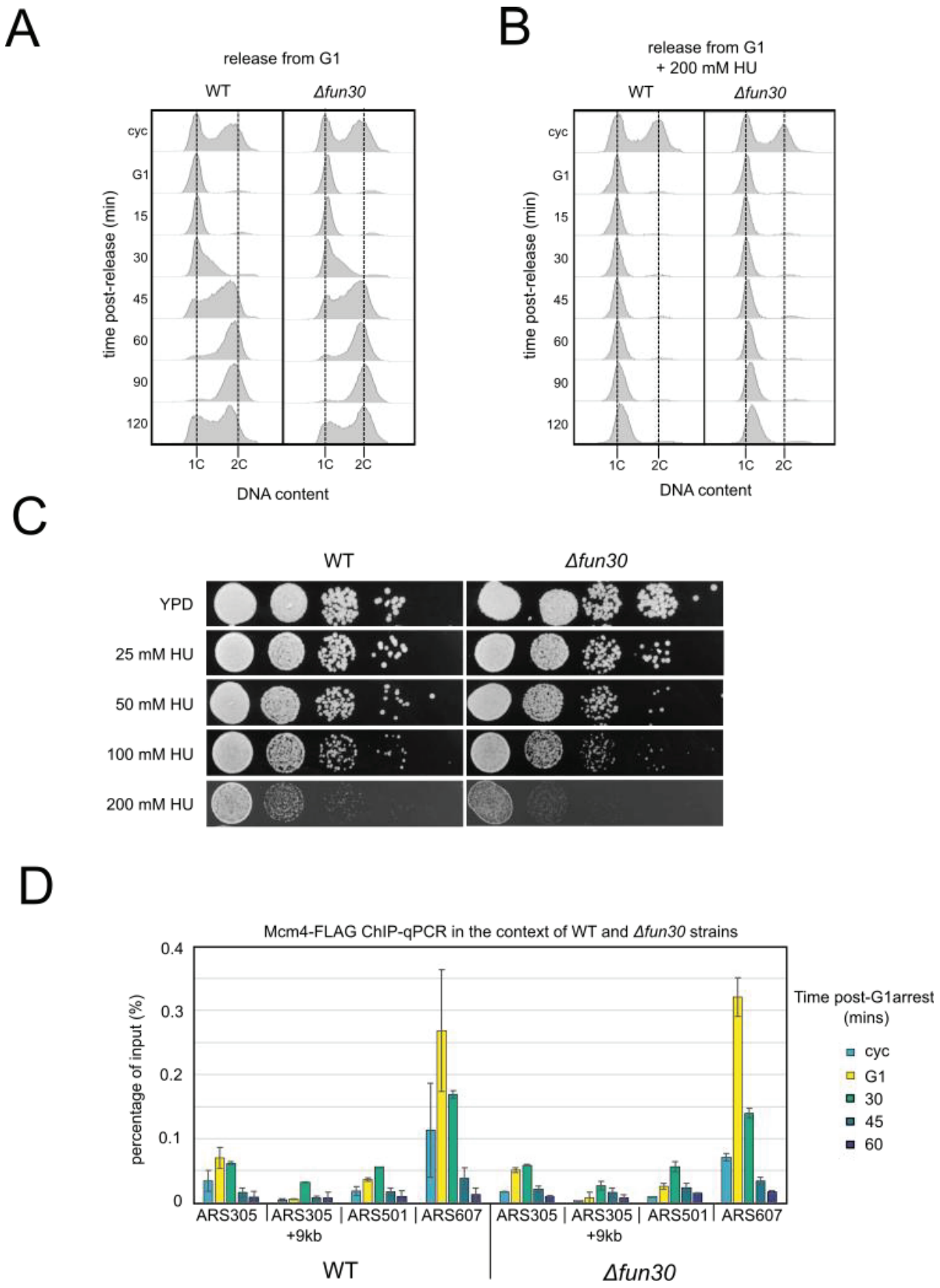


Figure 66: Phenotypic effects of the deletion of FUN30 on cells. A) DNA content changes as measured by SYTOX incorporation and flow cytometry of wild type and ⊿Fun30 cells released from alpha factor-induced G1 arrest into normal media. B) As in A, but released into S phase in the presence of 200 mM HU. C) Dot-spot viability assay of wild type (WT) and ⊿fun30 cells grown on plates containing the indicated concentrations of HU. Cells were plated in 10-fold serial dilutions. E) ChIP-qPCR analysis of Mcm4-FLAG origin association in the context of wild type or ⊿fun30 cells. Asynchronous (cyc) cells were arrested in G1 and released into S phase, and the origin association of Mcm4-FLAG compared at several stages of the time-course.

longer-term effects of HU on *Δfun30* a dot-spot viability assay was performed with increasing concentrations of HU in the plate media (Figure 66C). A very minor loss of viability was observed in the *Δfun30* cells relative to the wild type. Finally, to test whether the loss of Fun30 had any effect on the origin-associated processes leading up to and including replication initiation, Fun30 was knocked-out in the context of an Mcm4-FLAG containing strain. By carrying out an Mcm4-FLAG ChIP-qPCR in *Δfun30* cells, it would be possible to see if the association of MCM2-7 with origins in G1 was affected by the loss of Fun30, or if the timing of replisome assembly and loss of MCM2-7 association with origins was affected. Disappointingly, it was not possible from this particular assay to see any effect caused by the loss of Fun30. The levels of Mcm4 at origins were consistent in alpha factor-mediated G1 arrest between *Δfun30* and wild type cells. Similarly, the loss of MCM2-7 complexes occurred within the same time-frame, and dynamics of this loss appeared identical. This was true for both of the origins tested in this case, ARS305 and ARS607. Given time limitations, it was not possible to follow up on this particular investigation, but further experiments looking at other origins, over tighter time-points and with other tagged replication factors could all be considered to look for a more specific or more subtle effect of the loss of this factor in the processes leading up to and including DNA replication initiation.

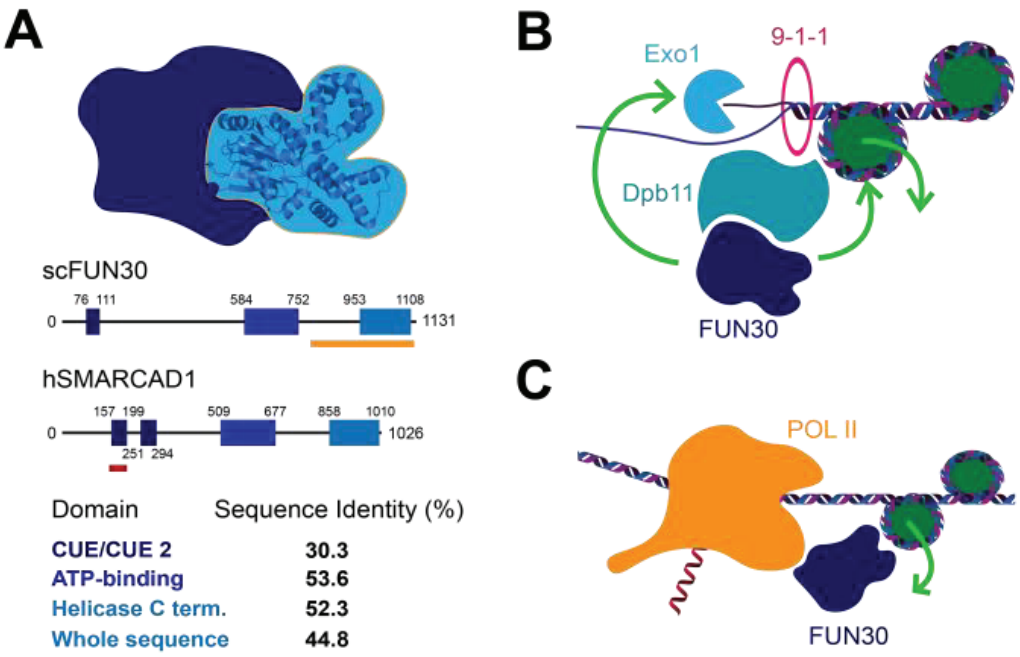


Figure 67: Overview of previously proposed Fun30/SMARCAD1 functionalityA) scFun30 and hsSMARCAD1 share significant homology, but mostly within the ATP-binding and helicase C-terminal domain (Neves-Costa et al., 2009), crystal structure by Liu and Jiang, 2017 (Liu & Jiang, 2017) orange and red bars show the regions of protein that have been structurally determined in the S. cerevisiae Fun30 *and human SMARCAD1 respectively. B) Fun30 is thought to act in nucleosome remodelling at sites of double-strand breaks (Bantele & Pfander, 2019). C) Fun30 also appears to play a similar role in histone remodelling ahead of RNA polymerase II (Lee, Choi & Lee, 2018).*

3.9.2 Rrm3

A similar strategy to that used with Fun30 was used to further investigate the replication association of Rrm3. ChIP-qPCR was carried out on strains containing endogenously tagged Rrm3-5xFLAG and origin DNA quantified (Figure 68). A clear association of Rrm3 with origin and with the non-origin but origin-proximal site ARS305+5kb was observed. This result appeared to tentatively confirm previous studies that suggested Rrm3 associates with the replication fork even in unperturbed conditions. Given this data suggested the association of Rrm3 began with forks at or very near the origin just after or perhaps even during activation, it also suggests the interaction between Rrm3 and the replisome begins early. It is worth noting that this data is the result of a single replicate, so should be interpreted with caution. Further replicates and more detailed time-courses might be able to further elucidate the stage at which Rrm3 first associates with the origin.

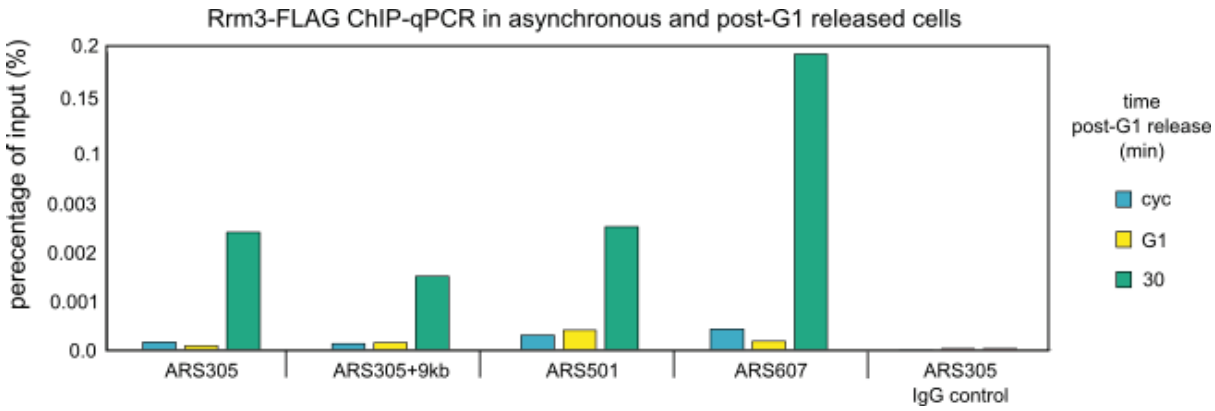


Figure 68: ChIP-qPCR evaluation of Rrm3 origin association. Cells were collected for processing from three separate conditions: asynchronous cycling cells, alpha factor-induced G1 arrest, and 30 mins post-G1 arrest. Origin DNA pull-down efficiency was compared across several origin and non-origin sites in the Rrm3-FLAG samples and a with samples from a mock IP using non-specific IgG. Represented data comes from a single test replicate.

Further experiments into a possible role of Rrm3 at origins of replication during loading or initiation were carried out using strains in which the endogenous copy of Rrm3 had been deleted (Figure 69). Firstly, wild type and *Δrrm3* cells were released from an alpha factor-induced G1 arrest into S phase (Figure 69A). Two interesting aspects were noted about the replication dynamics of the *Δrrm3* cells. Firstly, the overall completion of replication was delayed in the *Δrrm3* cells, and this likely results from the studied function of Rrm3 in overcoming replication fork blocks (Torres, Schnakenberg & Zakian, 2004; Ivessa et al., 2002, 2003; Azvolinsky et al., 2006; Torres, Bessler & Zakian, 2004). However, the entry of cells into S phase is remarkably synchronous in the *Δrrm3* population. At 45 minutes post-release, the *Δrrm3* cells have almost all started DNA replication. In contrast, there remains a significant number of cells in the wild type population that have yet to begin DNA replication and enter S phase. This effect was remarkably consistent over multiple replicates (data not shown). This effect is visible in previous experiments involving *Δrrm3* strains in the literature (Ivessa et al., 2002), but goes seemingly unremarked upon. It is unclear how the current proposed functions of Rrm3 could cause this effect, which suggests an additional role of Rrm3 in modulating S phase entry and DNA replication, possibly at a whole-cell level. Given research suggesting a role in controlling replication via interactions with ORC (Syed et al., 2016), and the data from the ChIP-MS time-course in this study suggesting an association with MCM2-7 that starts prior to full activation (Figure 53), it was considered that Rrm3 might affect aspects of helicase loading or activation.

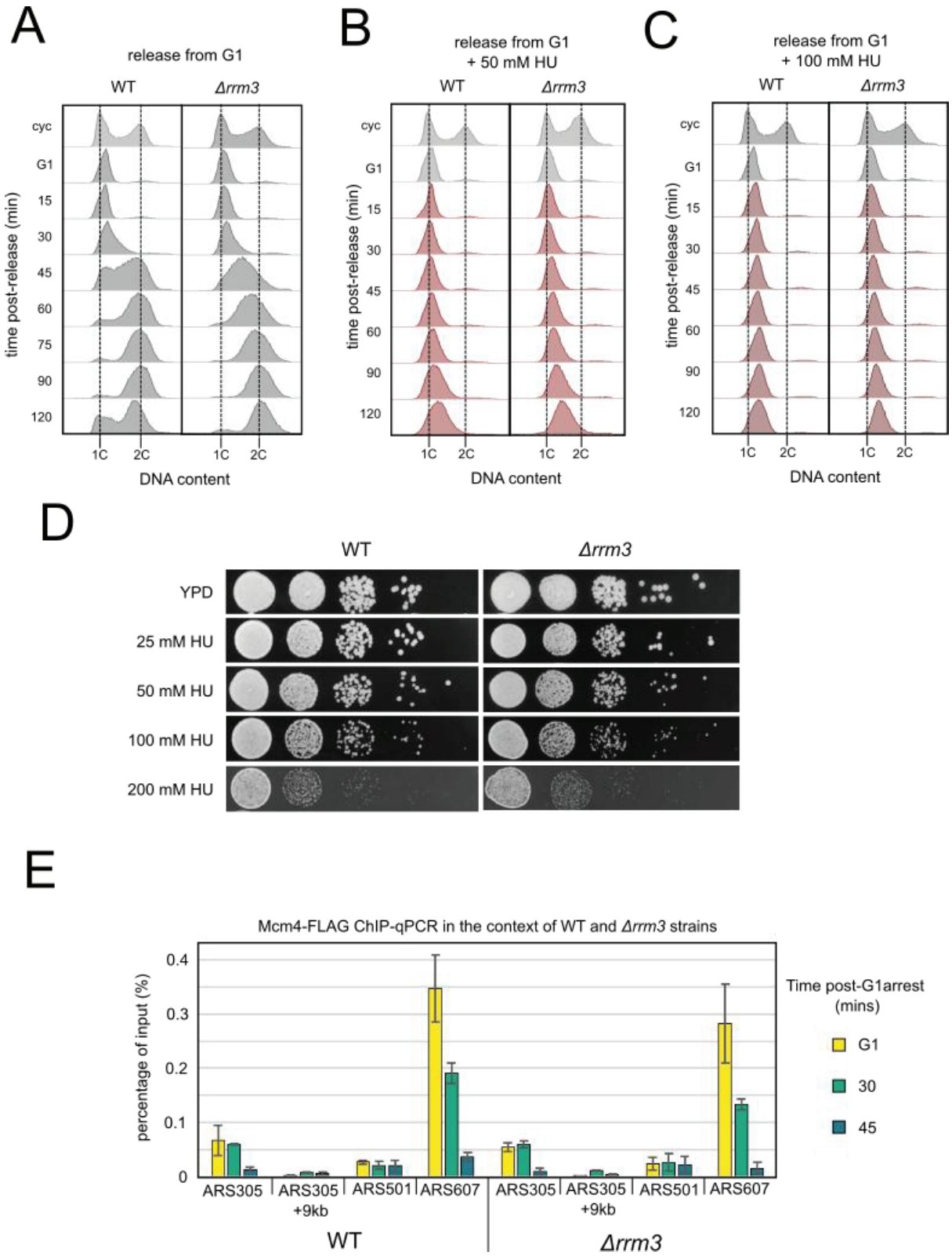


Figure 69: Phenotypic effects of the deletion of RRM3 on cells. A) DNA content changes as measured by SYTOX incorporation and flow cytometry of wild type and Δrrm3 cells released from alpha factor-induced G1 arrest into normal media. B) As in A, but released into S phase in the presence of 50 mM HU. C) As in B, but with 100 mM HU. D) Dot-spot viability assay of wild type (WT) and Δrrm3 cells grown on plates containing the indicated concentrations of HU. Cells were plated in 10-fold serial dilutions. E) ChIP-qPCR analysis of Mcm4-FLAG origin association in the context of wild type or Δrrm3 cells. Cells were arrested in G1 and released into S phase, and the origin association of Mcm4-FLAG compared.

Studies on *Δrrm3* strains show a reduced response of cells to HU during S phase due to the reduction of checkpoint activity in these cells (Syed et al., 2016). The effect of releasing *Δrrm3* cells into high concentrations of HU was tested (Figure 69B,C). In each case, *Δrrm3* still showed similar levels of replication stalling in contrast to the replication dynamics of these cells in the literature. Similarly to the published literature, however, the *Δrrm3* cells showed no loss of viability on plates with increasing concentrations of HU (Figure 69D).

Finally, the effect of loss of Rrm3 was tested, as with Fun30, on the origin association of MCM2-7 as measured by ChIP-qPCR of Mcm4-FLAG (Figure 69E). As with Fun30, at this stage, no significant differences were observed between *Δrrm3* cells and wild type cells in MCM2-7 origin-association dynamics. Roughly similar levels of MCM2-7 were associated in G1 for both strains across both origins tested, and the dynamics of replication initiation and replisome movement away from the origin were mostly indistinguishable. Again, as with the further investigation of Fun30, additional replicates and further experiments could indicate the presence of a role of Rrm3 at origins, but at present, it cannot be said that Rrm3 influences helicase loading or initiation at the time-points and origins tested.

3.9.3 Sir3

For Sir3, too, it was sought whether the ChIP-qPCR assay could show the presence of this protein at origins. Several lines of inquiry have suggested that Sir3 and the related protein Sir2 are active at nucleosomes adjacent to origins (Radman-Livaja et al., 2011; Hoggard et al., 2018). The tagging of the endogenous copy of SIR3 allowed assessment of origin binding (Figure 70). The levels of Sir3 associating with the origins DNA was comparatively small, with a similar association at the non-origin ARS305+9kb site as at the ARS305 site. For comparison, the levels of Sir3 associating with the MAT locus, a domain known for Sir protein-mediated silencing, was also quantified by qPCR, and showed around 100-fold greater association than origin DNA.

From this, and observation of previous genome-wide data, it was concluded that the Sir3 association at origins might be too subtle of an effect to pick up with ChIP-qPCR. It was also possible that the Sir3 association with origins is limited to a subset of origins, that account for most of the association of MCM and Sir3 in the ChIP-MS data. The ARS sequences in telomeric regions are obvious loci that deserve further investigation of the co-association of Sir3 and the replication machinery (Stevenson & Gottschling, 1999).

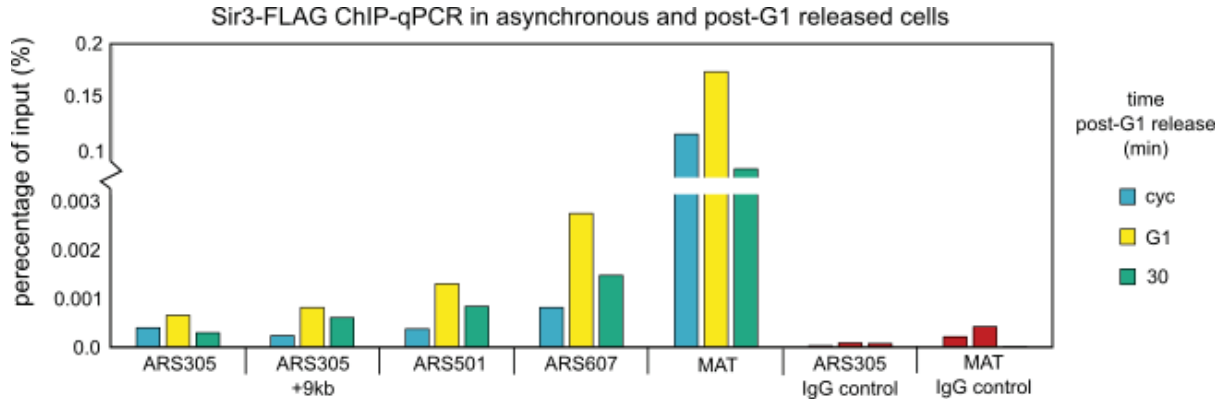


Figure 70: ChIP-qPCR evaluation of Sir3 origin association. Cells were collected for processing from three separate conditions: asynchronous cycling cells, alpha factor-induced G1 arrest, and 30 mins post-G1 arrest. Origin DNA pull-down efficiency was compared across several origin and non-origin sites in the Sir3-FLAG samples and a with samples from a mock IP using non-specific IgG. Represented data comes from a single test replicate.

Strains were also generated lacking SIR3 (Figure 71). However, due to the loss of the pheromone response due to a loss of silencing of the *HML* and *HMR* mating type loci, *Δsir3* cells could not be synchronised using α-factor. Otherwise, the *Δsir3* cells appeared normal in their response to treatment with HU, and showed no loss of viability under HU treatment compared to wild type cells

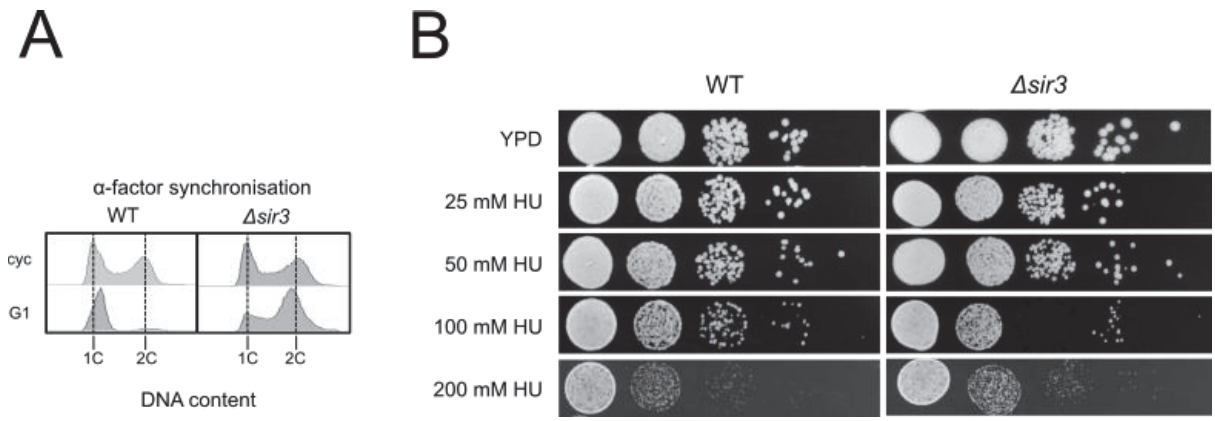


Figure 71: Phenotypic effects of the deletion of SIR3 on cells. A) Trial attempt to synchronise Δsir3 cells using α-factor. Δsir3 cells lack a functional response to α-factor. B) Dot-spot viability assay of wild type (WT) and Δsir3 cells grown on plates containing the indicated concentrations of HU. Cells were plated in 10-fold serial dilutions.

3.9.4 Hrq1

The behaviour of Hrq1-FLAG in a ChIP-qPCR assay for origin association was not promising from the results of the first initial replicate (Figure 72). Though the signal of origin DNA at all cell-cycle points tested was well above the levels of an IgG control, The ARS305+9 region, as a non-origin locus, showed significantly more signal than the ARS305 region itself. Because Hrq1 only showed reproducible association with the MCM2-7 complex when initiation was arrested by an intra-hexamer linkage, it is possible that its association with origins is too brief to visualise using ChIP-qPCR at the time-points chosen. Another possibility is that the Hrq1 protein, with links to DNA repair , is recruited to the arrested complex to resolve the block to replication initiation artificially created by the construct. Unfortunately, this possibility was not able to be tested within the time constraints of the project, but is an avenue for further investigation of why the Hrq1 protein showed association in ChIP-MS in this specific sample.

Cells without an endogenous copy were created in the same manner as for the other proteins investigated as part of this project. The DNA replication dynamics and viability of Δhrq1 cells were tested under varying concentrations of hydroxyurea

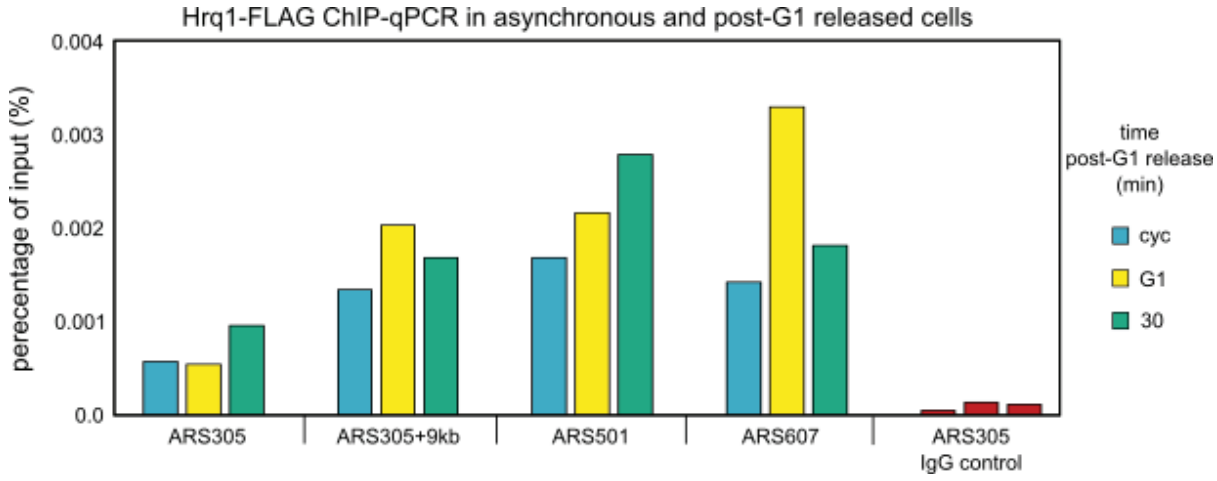


Figure 72: ChIP-qPCR evaluation of Hrq1 origin association. Cells were collected for processing from three separate conditions: asynchronous cycling cells, alpha factor-induced G1 arrest, and 30 mins post-G1 arrest. Origin DNA pull-down efficiency was compared across several origin and non-origin sites in the Hrq1-FLAG samples and a with samples from a mock IP using non-specific IgG. Represented data comes from a single test replicate.

In cell cycle analysis, *Δhrq1* did not behave any differently to the wild type cells when released into S phase with or without the presence of 50 or 100 mM HU (Figure 73). However, in a viability assay, *Δhrq1* cells showed a slight loss of viability when compared to the wild type cells when exposed to chronically higher concentrations of HU (Figure 73D). This was contrary to the evidence that *Δhrq1 S.cerevisiae* cells showed no loss of viability in previous studies. When treated with HU (Leung et al., 2014), though it did reflect similar results to studies of strains of *S.pombe* or *Arabidopsis thaliana* with deleted or mutated Hrq1 homologs (Groocock et al., 2012; Röhrig et al., 2018).

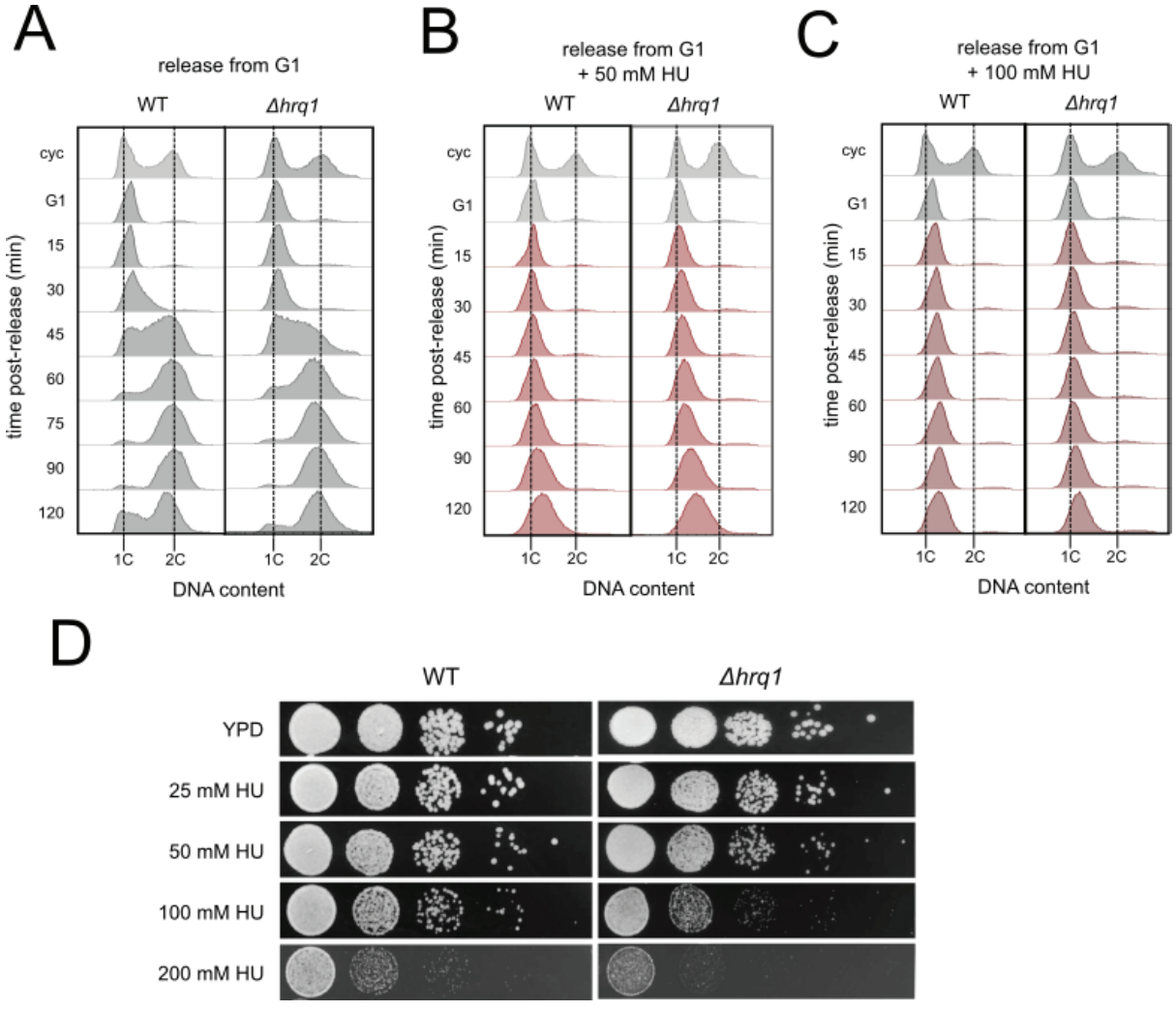


Figure 73: Phenotypic effects of the deletion of HRQ1 on cells. A) DNA content changes as measured by SYTOX incorporation and flow cytometry of wild type and Δhrq1 cells released from alpha factor-induced G1 arrest into normal media. B) As in A, but released into S phase in the presence of 50 mM HU. C) As in B, but with 100 mM HU. D) Dot-spot viability assay of wild type (WT) and Δhrq1 cells grown on plates containing the indicated concentrations of HU. Cells were plated in 10-fold serial dilutions. E) ChIP-qPCR analysis of Mcm4-FLAG origin association in the context of wild type or Δhrq1 cells. Cells were arrested in G1 and released into S phase, and the origin association of Mcm4-FLAG compared.

3.9.5 Cac2

Though it was a lower-confidence candidate overall, Cac2 was tagged and a single replicate of ChIP-qPCR was carried out to observe if there was any clear origin association of this protein with this method (Figure 74). In this case, similarly to some of the other factors, it was not clear if the DNA signal in the ChIP-qPCR truly indicated association of Cac2 with origins. Whilst the signal was much greater than that of the IgG control, the non-origin ARS305+9 kb locus showed DNA pull-down efficiencies of a similar order of magnitude to those of the origins tested. Overall the results were too close to determine any conclusive information from a single replicate, but the preliminary test indicated that in addition to a general background of DNA pull-down in ChIP-qPCR, there might be a very slight indication of an interaction in G1 with a subset of origins which might be worth further replicates and investigation.

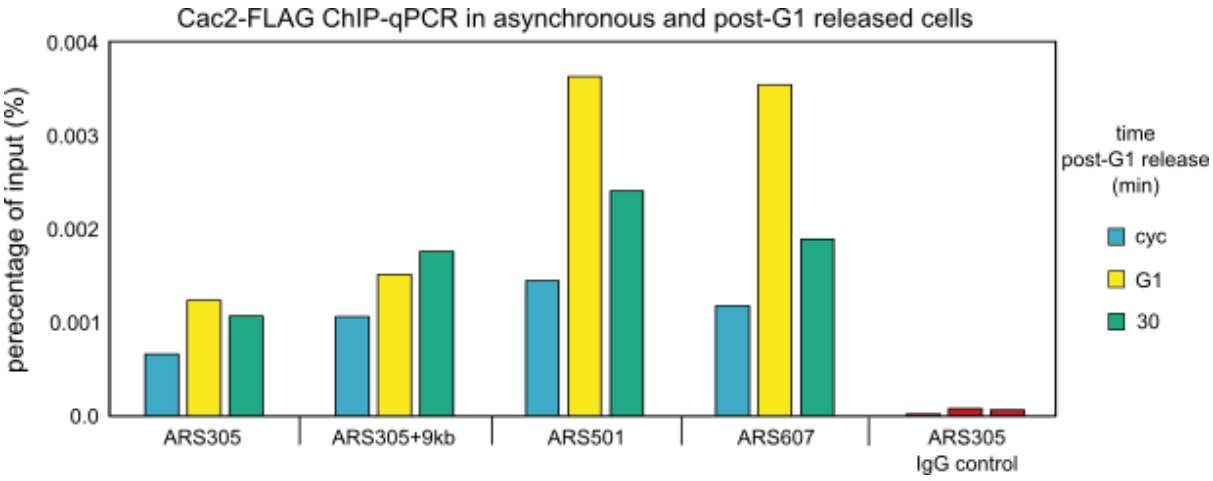


Figure 74: ChIP-qPCR evaluation of Cac2 origin association. Cells were collected for processing from three separate conditions: asynchronous cycling cells, alpha factor-induced G1 arrest, and 30 mins post-G1 arrest. Origin DNA pull-down efficiency was compared across several origin and non-origin sites in the Cac2-FLAG samples and a with samples from a mock IP using non-specific IgG. Represented data comes from a single test replicate.

A Δcac2 strain was created to observe if there were any phenotypic effects on the cell cycle or overall viability from loss of Cac2. Flow cytometry was used to analyse wild type and Δcac2 cells replication dynamics from a G1 arrest in the presence or absence of 50 mM and 100 mM HU (Figure 75A-C). There were no observable differences between wild type and Δcac2 cells

in terms of their cell cycle progress. In testing the viability of Δcac2 cells when exposed to chronic HU treatment, a slight loss of viability was observed (Figure 75D). This was surprising, as although compromising CAF-1 was deleterious in cells treated with DNA damage causing double strand breaks, treatment with HU had not been shown to cause any loss of viability in these cells (Linger & Tyler, 2005).

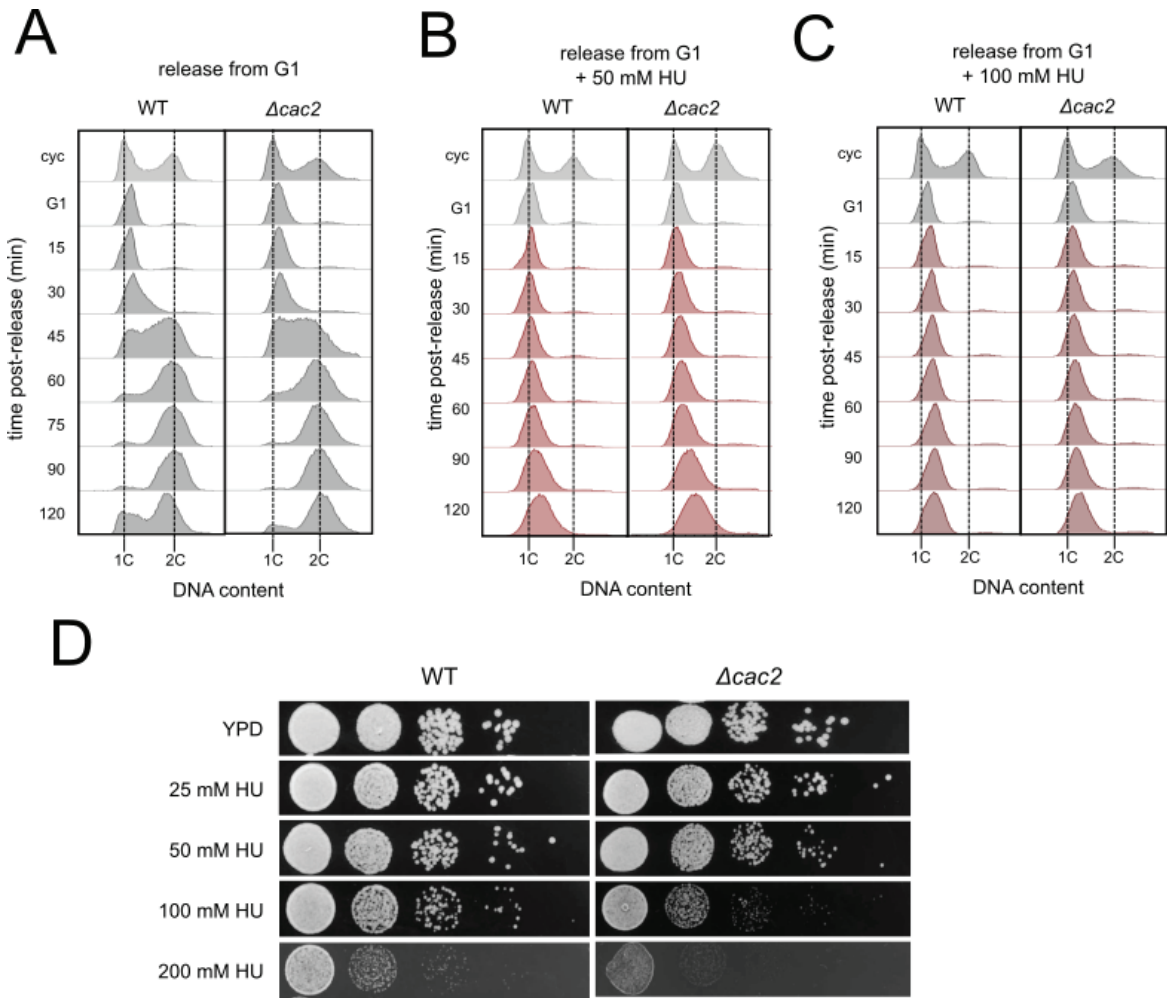


Figure 75: Phenotypic effects of the deletion of CAC2 on cells. A) DNA content changes as measured by SYTOX incorporation and flow cytometry of wild type and Δcac2 cells released from alpha factor-induced G1 arrest into normal media. B) As in A, but released into S phase in the presence of 50 mM HU. C) As in B, but with 100 mM HU. D) Dot-spot viability assay of wild type (WT) and Δcac2 cells grown on plates containing the indicated concentrations of HU. Cells were plated in 10-fold serial dilutions. E) ChIP-qPCR analysis of Mcm4-FLAG origin association in the context of wild type or Δcac2 cells. Cells were arrested in G1 and released into S phase, and the origin association of Mcm4-FLAG compared.

4 Conclusions and Perspectives

The understanding of the molecular events that led up to, and are part of, DNA replication has grown dramatically over the last few decades. In particular, the advent of cryo-EM techniques to analyse large macromolecular protein complexes assembled *in vitro* has generated a remarkable series of structural discoveries relating to the DNA replication machinery (Riera et al., 2017; Georgescu et al., 2017; Jones et al., 2021; Sun et al., 2015). However, the investigation of the same mechanisms *in vivo* can produce completely new findings and benefit from being examined in the cellular environment in which they take place. Technological advancement has not been excluded from *in vivo* work either, with next-generation sequencing combined with genomic techniques such as ChIP-Exo, and improvements in proteomic methods generating new and exciting ways to investigate the genomic binding and interactomes of protein complexes on chromatin in a quantitative manner. Despite the wealth of structural data and assay-based techniques generated by *in vitro* techniques, *in vivo* methods remain of outstanding importance to the discovering and understanding of biological mechanisms. This study used a combination of *in vivo* techniques to uncover new mechanisms of DNA replication initiation, and systematically identify the proteins associated with helicase loading, helicase activation and beyond.

4.1 Design and investigation of novel chemical-biology constructs to investigate key steps in replication activation

Three individual rapamycin-dependent linkages were investigated as part of this project. These constructs were designed to interfere with the progression of replication activation in different ways. Two constructs were investigated that utilised an intra-hexamer linkage between Mcm2 and Mcm5 subunits of the same hexamer and were shown by L.M. Reuter to successfully arrest DNA replication initiation, most plausibly by interfering with the proper extrusion of the

lagging strand DNA through the proposed DNA exit gate. This project was able to determine that the complex composition of each stage of arrest facilitated by these linkage constructs was different. This in turn implied that the stage of arrest was dependent on the positioning of the linkage along the length of the double-hexamer Mcm2-Mcm5 interface. Combined with genomic methods such as ChIP-Exo and potassium permanganate ChIP-seq (PIP-seq) (L.M. Reuter, pers. comm.) this data provided compelling evidence that the extrusion of lagging strand DNA begins at the interface of the MCM2-7 double hexamer (N-termini). The reduced association of pre-LC components, and lack of a ssDNA-associated hyper-crosslinking effect, in the N-terminal linkage arrest also implied that the initial extrusion of ssDNA only occurs at a later stage, once initiation factors that form the pre-LC attached to the helicase. The retention of pre-LC components and association of Mcm10 observed in the C-terminal intra-hexamer linkage arrest was also indicative that initial extrusion of lagging strand DNA could occur and was associated with progression of the complex to the active CMG state. The mechanistic event that caused release of the transient initiation factors Sld2 and Dpb11 was prevented in this arrest, demonstrating that the complex was a true intermediate.

The second clear structural change in initiation, other than ssDNA extrusion, is hexamer separation. It was considered if it was possible to physically block this structural rearrangement. The inter-hexamer Mcm4-Mcm5 linkage was designed to keep the hexamers closely bound and prevent lateral sliding or twisting. Testing revealed that this construct, too, was able to induce an arrest prior to the full activation of replication. Most interestingly of all, differences were observed in the arrest state generated by this inter-hexamer latch linkage construct compared to the intra-hexamer gates, most notably the stability of the arrest and the retention of the initiation factor Sld3. ChIP-qPCR results for replication factors in the context of the inter-hexamer linkage were highly informative. The complex composition state was

consistent with an intervening stage between the N-terminal intra-hexamer gate and the C-terminal intra-hexamer gate.

The components of the pre-LC complex showed strong association, but it was the retention of Sld3 that was the most elucidating. In total, the data implied that hexamer separation was required for or linked to the release of Sld3 from the complex. Considering the replication factors present in each of the arrest states generated, it was possible to suggest an order to structural events during initiation.

It appeared that *in vivo* an initial amount of DNA extrusion occurred at the N-terminal interface of the double-hexamer, prior to the CDK-dependent association of the pre-LC. This is interesting, since studies investigating complex formation in the absence of Mcm10 suggest that this factor is responsible for activation of the full unwinding activity of the replisome, and in the absence of Mcm10, splitting occurs (at least *in* vitro), and the replisome recruits all other initiation factors competently (Miyazawa-Onami, Araki & Tanaka, 2017) but does not fully extrude ssDNA (Douglas et al., 2018). This data would suggest that, although Mcm10 is still likely a more "final" initiation factor, unwinding is not a separate process that occurs in single hexamers post-splitting, but is more concerted *in vivo*.

It is likely that pre-LC association then allowed progression of the complex to hexamer separation, which it appeared Sld3 release was dependent on. Evidence of a possible role for Sld3 during later stages of activation is highly informative, since while Sld3's role in recruiting early factors to replication complexes at origins during the earliest stage of initiation (Fang, Cao & Lou, 2016; Tanaka et al., 2011; Deegan, Yeeles & Diffley, 2016), studying later events is challenging, since interference with Sld3 function arrests complexes at an early stage due to the protein's role in recruitment of Cdc45 (Miyazawa-Onami, Araki & Tanaka, 2017).

Finally, the completion of lagging strand extrusion from the C-terminal end of the MCM2-7 hexamer happened after hexamers had separated. This proposed ordering of the structural

changes during initiation suggests that ssDNA extrusion begins sometime before double hexamer separation, but is not completed until after a degree of separation of the hexamers has occurred, implying some form of concerted mechanism (Figure 76).

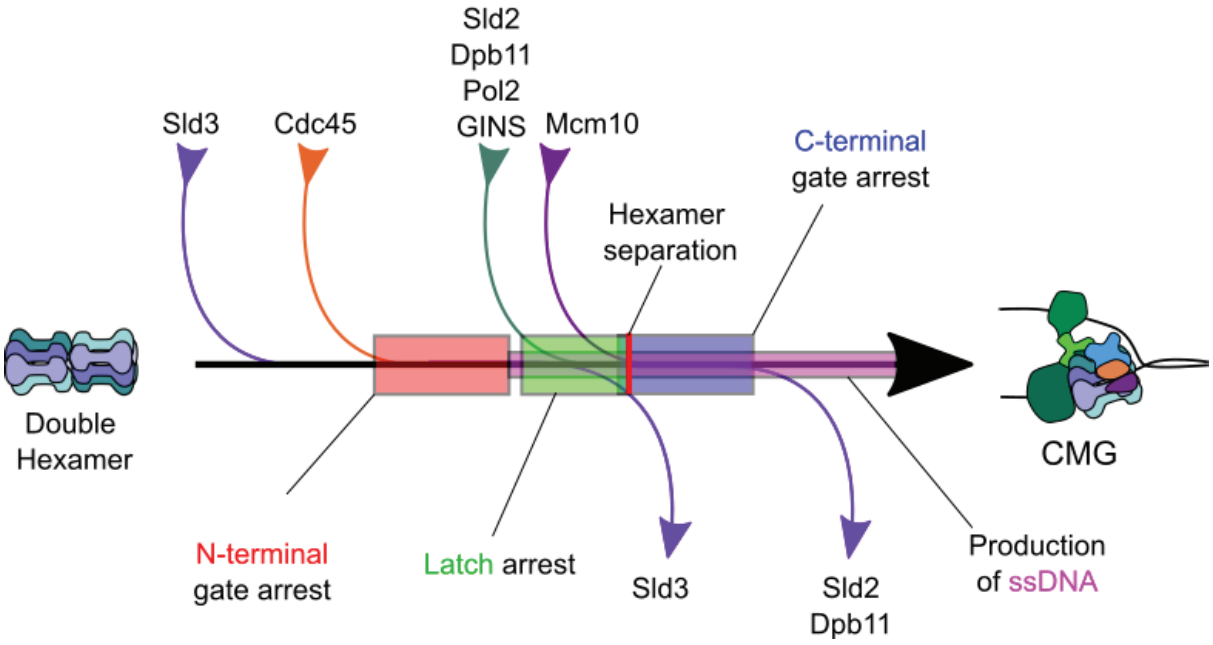


Figure 76: Model of ordering of events in replication initiation.

DNA extrusion occurs over multiple complex composition changes. Double-hexamer splitting, as prevented by the inter-hexamer latch, occurs sometime during or prior to the final extrusion of ssDNA from the C-terminus of the MCM2-7. Preventing separation prevents the release of Sld3, suggesting these two changes are linked. Pre-LC and Mcm10 are already present prior to splitting, or the failure to split does not interfere with their recruitment timings and ability.

This is a tentative model, since it is based on the *in vivo* complex characterisation shown in this project, and additional genome-wide methods (L.M. Reuter, pers. comm.), but matches well with previous *in vivo* data (Miyazawa-Onami, Araki & Tanaka, 2017), and the genetic and biochemical experiments conducted previously. The suggestion that hexamer separation is associated with Sld3 release, and occurs prior to Sld2/Dpb11 release, while completion of ssDNA extrusion is associated with Sld2/Dpb11 release, are novel findings of this collaborative study that have not been shown previously.

There are some considerations to take into account when using these constructs to draw conclusions about mechanistic events. Firstly, whilst the linkages were designed to interfere in specific structural changes, it is difficult to determine the actual mechanism by which they

mediate an arrest to the cell cycle. By far the most plausible mechanism is the physical linkage they were designed to impart, and this is further corroborated by the overall lack of a phenotype in the absence of the linkage-inducing rapamycin. However, in the absence, at present, of structural data of these modified complexes at arrest, it cannot be fully excluded that a linkage could form in another configuration, between Mcm4 and Mcm5 of the same hexamer, either by circling around the hexamer circumference, or across the N-terminal interface. This is considered unlikely, both due to the limited length of the FRB/FKBP and linker inserts, but also because of the arrangement of the double hexamer preventing a short access route between Mcm4 and Mcm5 of the same hexamer. Indeed, Mcm4 and Mcm5 were chosen because they were the subunits that showed the closest proximity to the other in the opposite hexamer, whilst being diametrically opposed to the other subunit within their own hexamer. Further testing with shorter linkers or linkers with a lower flexibility could help rule this possibility out.

Another unknown quantity was how completely the inter-hexamer latch was able to prevent relative motion of the hexamers. The construct was competent to induce an arrest, but it was not known how tightly the hexamers were held and what the margins of the linkage were like at a structural level. Were the two hexamers held so tightly that no motion was possible, or was there some room for lateral motion of the hexamers to “slip” laterally or separate a small distance? Equally, though the inserted domains were clearly able to form a linkage, were there any structural distortions that affected the activity of the replication complex? Without accurate structural information, these are still open questions when using the constructs.

In addition, though the inter-hexamer linkage was fully able to induce an arrest in replication initiation, it was found that induction of this arrest required 10-fold higher levels of rapamycin than the intra-hexamer linkages previously, and the arrest generated was not as stable over time. This was considered to reflect the fundamental differences in the structural events being inhibited by the two types of linkage. If the separation of the hexamers is driven by strong

forces, then it may be more susceptible to disassociation if rapamycin concentration is not high enough to ensure very high occupancy. The MCM2-7 helicase is an ATPase motor, and so if the separation process is in any way related to the potentially very high driving force of this motor, then this might explain the relative instability of the arrest over time. It is also possible that accessory factors might be recruited over the lifetime of the arrest, and aid the separation process. Because the activation step has been difficult to study *in vivo* up to this point, this arrest construct represents a method to test if such factors function in this context.

Experiments investigating the activation of DNA damage in the context of the intra- and inter-hexamer linkage construct arrests were inconclusive (data not shown), but it is possible that protein activity, such as accessory helicases that function in other replication-related events (Deegan et al., 2019), was able to "resolve" the impediment to helicase activation in the inter-hexamer latch in a way not possible with the intra-hexamer gates. In fact, the possible existence of such a resolution mechanism during replication initiation seemingly only recruited after some elapsed time was of great interest, and the further analysis of the inter-hexamer arrest complex composition by proteomic methods was hoped to elucidate whether such a mechanism was represented in protein factors present, see below. The finding that early and efficient origins in human replication have a distinct mutation signature due to dsDNA breaks at origins and error-prone synthesis proximal to the origin (Murat, Guilbaud & Sale, 2021), suggests that evolution of damage control mechanisms specifically focussed on origins and activating replisomes is plausible.

Regardless of the considerations, the experiments herein represent a unique method of investigating the key steps in replication initiation, and were able to corroborate previous proposals about the subcomplexes recruited in initiation, e.g. CDK-dependent pre-LC association and DDK-dependent Sld3/7 and Cdc45 association. In addition, the combination

of the linkage constructs was able to hypothesise about a proposed ordering of structural events in replication initiation.

4.2 Optimisation and adaptation of a ChIP-MS protocol to investigate labile complexes and transient interactors of large *in vivo* protein complexes

Affinity purification techniques have been applied to the process of DNA synthesis as a whole, but targeted investigation of replication initiation has not been investigated in this manner. This is most likely due to the probably transient and labile nature of such complexes. A native pull-down approach was optimised and shown to be able to isolate GINS-containing replisomes efficiently after only one step of purification that could be characterised by mass spectrometry (Section 3.2.4, Figure 32, Figure 33). However, after testing this approach in the intra-hexamer gate strains, it was shown that the isolated interactome of Psf2-TAP in this context appeared unchanged, with no observably enriched or de-enriched protein factors (Section 3.2.4, Figure 34).

To solve this issue, in an approach directly analogous to the ChIP-qPCR method, termed ChIP-MS, cells would be fixed with formaldehyde *in vivo*, and then specific complexes isolated using tagged proteins and washing steps derived from those used in ChIP-qPCR.

Use of the Cdc45-FLAG tag and the ChIP-MS technique was able to isolate and characterise a replication-associated complex from cells *in vivo* (Section 3.4), and identify the enrichment of many of the core replication factors that were indicative of the replisome progression complex and whose human homologs had been identified with alternative cross-linking based techniques in human cells (Gambus et al., 2006; Alabert et al., 2014). In particular, the comparison of the factors enriched in the ChIP-MS of replication complexes in yeast, with those identified in human cells using nascent chromatin capture (NCC) was favourable. Many of the homologues of fork components identified by Alabert et al. were also clearly enriched

in the ChIP-MS, though for the purpose of this study, the ChIP-MS method had some advantages. The NCC technique relied upon the production of nascent DNA to identify interacting factors, which is of use when comparing protein association with nascent or matured chromatin (Alabert et al., 2014). The ability of the ChIP-MS technique to identify a similar complex without the need of nascent chromatin as a pull-down target meant that this technique could be used to investigate complexes prior to the synthesis of nascent DNA, i.e. during replication initiation. Correspondingly, and very excitingly, even this initial Cdc45-FLAG ChIP-MS investigating un-arrested replication, was able to identify enrichment of Sld3, Sld7 and Dpb11 (Figure 38). This was very exciting, as the detectable presence of these transient initiation factors meant that this technique was likely capable of studying transient complexes formed during replication initiation, a key goal of the project and an aim of the replication field as a whole. In particular, this was thought to be unlikely to be due to association of these initiation factors with active replisome, due to a predicted overlap of the binding sites for Sld3 with the active replisome-associated GINS complex in the CMG (Fang, Cao & Lou, 2016).

Testing FLAG-tagging of the Mcm4 subunit as an alternative target showed competent recruitment of replication complexes, and initially a small number of time-points were tested which showed clear temporal dynamics for protein association with the MCM2-7 complex, in line with expectations from a protein's role (Section 3.6.1, 3.6.2). A large scale time-course investigating the interactome of MCM2-7 through multiple stages of the cell cycle, including G1 and several time-points of S phase was conducted (Section 3.6.3). the goal of this experiment, to investigate complex dynamics using ChIP-MS, was shown to be very much achieved, and generated novel data regarding association dynamics of a wide range of proteins, particularly for the earlier time-points investigated.

As part of ongoing attempts to characterise the arrest states of the linkage constructs, ChIP-MS was also applied in large-scale experiments to investigate the complex compositions of

arrest states of the intra-hexamer gates and the inter-hexamer latch, to varying degrees of success (Section 3.7). The data obtained for the intra-hexamer gates provided valuable information confirming aspects of complex composition, and was used to help aid understanding of these constructs and their modes of action (manuscript in preparation).

As discussed in Sections 3.6 and 3.8, indications suggested that certain issues of reliability with this technique were due to inefficient break-up of cellular chromatin, a problem potentially compounded due to formaldehyde cross-linking causing large tracts of heavily inter-linked protein-DNA complexes. It was apparent that, in the current state of the technique, this release from chromatin needed to be improved. Whilst the level of cross-linking or efficacy of sonication could have been root causes, these were standardised as much as possible during the project. Instead, by utilising DNase treatment to ensure maximum release of protein factors from chromatin, the problem might be avoided entirely. Indeed, initial tests involving Benzonase® treatment appeared highly successful and offered significant improvement. This extra release of chromatin-bound MCM2-7 complexes likely allowed them to compete against the soluble MCM2-7 complexes that were likely binding preferentially prior to full break-up of cellular chromatin. It was possible that extended tracts of DNA were acting to bridge proteins in the pull-down, and enabling enrichment of indirect interactors in the standard ChIP-MS technique. However, in the initial tests utilising Benzonase®, there were no significant populations of proteins that showed a loss of enrichment due to degradation of DNA, indicating indirect interactions mediated by DNA were not a common form of bridging interaction in the pull-down. In potential future development of this technique, this effect should be monitored closely, as it remains an important unknown regarding the ChIP-MS technique.

4.3 Impact and advantages of the ChIP-MS technique

Though native pull-downs have been applied to yeast replication complexes (Gambus et al., 2006, 2009), and cross-linked isolation of proteins linked to nascent chromatin has also been carried out , this project represents the first application of cross-linking and affinity purification targeting the yeast replication machinery, particularly that involved in the initiation of DNA replication itself. The high sensitivity of the mass spectrometry instrumentation used here also gives advantages over earlier analysis of the yeast replication machinery, and the development of label-free quantification allows comparison of samples directly without the need for SILAC or other more involved mass spectrometry techniques.

Several large-scale ChIP-MS experiments were performed as part of this study, and provide a significant amount of data regarding the interactome of the replication machinery even in unperturbed cells. The Cdc45-FLAG ChIP-MS dataset was able to show enrichment of almost all of the proteins noted as enriched in previous native pull-downs (Gambus et al., 2006, 2009), except histones and Csm3 (Figure 39). These proteins are relatively small, and so may result in fewer tryptic peptides, an issue that may be compounded by formaldehyde cross-linking. Histone proteins were detected reliably in earlier ChIP-MS experiments, however. Beyond the set of core replication factors, most interesting was the enrichment of initiation factors Sld3, Sld7 and Dpb11. These proteins are known initiation factors that associate transiently, and in the case of Sld3/7, exclusively, with replication complexes during initiation (see Section 1.4.2 and 1.4.3). The presence of these factors indicated that even in an S phase cell population, there appeared to be some complexes isolated that were actively undergoing initiation. In addition to these initiation factors, the Cdc45-FLAG ChIP-MS also showed a great number of significantly enriched factors over the negative control (Figure 39). Comparison of these listed proteins with the human homologs identified in nascent chromatin capture identified quite a number of shared hits. In particular, the Cdc45-FLAG ChIP-MS showed enrichment of several

replication-associated DNA damage factors, such as members of the mismatch repair pathway Msh3 and Msh6 (Marsischky et al., 1996), and Rad51 and Rad52, proteins linked to homologous recombination (Lee et al., 2003; Symington, Rothstein & Lisby, 2014) (Figure 39B,C). This matches well with the findings from nascent chromatin capture in human cells, where DNA damage and repair factors show association with nascent chromatin even in normal media .

In addition, many chromatin remodellers were also enriched, suggesting a close association of the replication fork with these proteins, including the FACT complex, and the protein Fun30, which until recently had not been linked to DNA replication (Alabert et al., 2014).

The data from large scale Mcm4-FLAG ChIP-MS also showed many enriched factors in unarrested replisomes in S phase (Sections 3.6.3, 3.7.1). Among them for the Mcm4-FLAG at 30 minutes are many of the same factors enriched in the Cdc45 ChIP-MS. In both sets of experiments, there were also additional proteins that were repeatedly well enriched. Among these were Fun30, Sir3 and Rrm3. All three showed consistently high enrichment values across both tagged proteins and were considered for further investigation. The most interesting aspect of the Mcm4-FLAG time-course, however, was the ability to monitor the changing MCM2-7 complex over time and the cell cycle (3.6.3). Changes in abundance of complex components were reproducible for many factors, and showed clear dynamics of association over the course of the experiment. In addition to many replisome factors that showed dynamics that matched both theoretical expectations and experimental ChIP-qPCR data, unexpected dynamics were observed for several members of the replication machinery. As an example, polymerase α and δ showed association with MCM2-7 prior to the time-point where the initiation of replication was expected to begin. Such early association was surprising, given the suggested link between Mcm10 and Ctf4 replisome association and the stabilisation of polymerase α at origins (Ricke & Bielinsky, 2004).

Given a small amount of early firing forks in the timecourse, this profile of polymerases could be representative of the initial usage of DNA polymerases at the origin, where polymerase α and δ catalyse DNA addition of the first section of unwound DNA, before a polymerase switch occurs (Aria & Yeeles, 2018). It could also be representative of a recruitment mechanism pre-initiation for these enzymes. Pol α interacts with both Mcm10 (Warren et al., 2009) and Ctf4 (Zhu et al., 2007; Gambus et al., 2009), and this interaction, and similar interactions for polymerase δ could ensure pre-initiation association for efficient initial DNA synthesis. The association of these two polymerases contrasts interestingly with the association of polymerase ε during replication. The catalytic activity of polymerase ε may not be required initially (Yeeles et al., 2017; Guilliam & Yeeles, 2020), yet the noncatalytic subunit of polymerase ε is absolutely required for the initiation process as part of the pre-LC (Kesti et al., 1999; Dua, Levy & Campbell, 1999; Goswami et al., 2018). Whether this is a quality control mechanism ensuring engagement of the main leading strand polymerase prior to full activation, or it serves some other mechanistic function will require further elucidation of the mechanism of initiation in its entirety in the future.

However, perhaps the most useful aspect of analysing replication profiles was applying comparative analysis to factors with more undetermined roles in replication (Figure 53). Fun30, Hrq1, Sir3, Forkhead factors 1 and 2, and members of the ISW1b complex all showed dynamics indicative of association with origins prior to and during helicase activation. Interestingly, by comparing the profiles of several of these factors with similar profiles in the dataset (Figure 54), it was determined that Fun30 followed a similar association dynamic to that of Rpd3, a protein known to have a role in specifying origin timing (Aparicio et al., 2004; Knott et al., 2009; Yoshida et al., 2014).

Further time-course based experiments with a higher temporal resolution might be able to determine in even finer detail the association dynamics of these factors. For now, however, the

association and calculated enrichment values are indicative of an interaction with the replication complex, and may suggest a possible aspect of their function, but validation and confirmation by other methods are still required.

When analysing the intra-hexamer gates in comparison to wild type using ChIP-MS, the method functioned as expected, and was able to show differences in the associations of certain replication factors between the arrest states. In the context of the inter-hexamer latch however, as previously mentioned, the method suffered some issues linked to release of chromatin-bound proteins.

The ChIP-MS of the intra-hexamer gate constructs confirmed several of the observations from ChIP-qPCR analysis of the arrest state, such as the increased and continued association of initiation factors Dpb11, Sld2 and Mcm10 with the complex when the C-terminal intra-hexamer gate was induced. By contrast, the association of many replication factors linked to active replication was comparatively reduced in the C-terminal intra-hexamer linkage, supporting the previously hypothesised model of the function of the construct and the nature of the arrested complex (Sections 3.1.1, 3.4, 3.7). Unfortunately, though it had been anticipated that analysis of the arrested complex might identify novel factors associating with the initiating replisome, there were few significantly enriched factors in either the C-terminal or N-terminal intra-hexamer linkage arrests (Figure 60). The most reliably enriched likely chromatin-associated factor was Hrq1, which showed enrichment in the C-terminal construct-induced arrest. Further investigation of Hrq1 association in the context of the arrested complex was not carried out, but it would be valuable to confirm this association and investigate the mechanism that may underlie it.

Overall, the use of the ChIP-MS technique as presented in this project allowed the stabilisation and isolation of a range of replication-associated complexes from yeast *in vivo*. The technique was able to identify transient interactors of these complexes, as well as interacting factors more

peripheral to the complex. Application of the approach in a time-variant manner showed the ability of this technique to detect subpopulations of replication complexes throughout the cell cycle, and analyse their compositions. This included complexes at or around the time of helicase loading, helicase activation, and active replication. Future attempts at time-resolved ChIP-MS might additionally identify and analyse complexes associated with termination, or by increasing the temporal resolution, gain a more accurate insight into the dynamics affecting replication complexes during initiation. The ChIP-MS method also identified a number of factors that have an indistinct or previously unknown role at replication complexes. As has been stated when using this kind of affinity enrichment technique coupled to mass spectrometry (Rajala et al., 2015), it is not possible to absolutely define a list of certainly replication-associated proteins, because of certain considerations that will be discussed below. However, with appropriate validation techniques, new interactions and functions can be ascribed to factors of interest. Additionally, from the high proportion of known replication factors showing enrichment in this study, it is certainly possible that many of the factors with no known function in replication, may yet be proven to have one in future research (Rajala et al., 2015).

Despite this, ChIP-MS in this study has identified a few factors that are likely to have a previously unknown role during replication or the events leading up to it, with Fun30 as a leading candidate, supported by recent research suggesting the homolog SMARCAD1 travels with replication forks in humans (Lo et al., 2021). In particular, it is interesting to hypothesise that a suggested chromatin remodeller such as Fun30 could have important roles at the origin regulating the interaction of the replication protein complexes and chromatin. Fun30 could help ensure that progression of the active replisome is unimpeded by nucleosomes, as it has been suggested to do with RNA polymerase (Lee, Choi & Lee, 2018), or it could function at the origin, ensuring expansion of a nucleosome free region to allow expansion in size of the replication complex, possibly first during loading of one, and then two, helicases, or during the

activation of the complex to allow the association of further replication factors, unwinding of DNA, and complex activation. As a non-essential protein, Fun30 may have a minor role at many origins or as part of a normal replication fork, or may have a specific but as yet unacknowledged, role during replication stress or at specific origins of replication (Figure 77). Validation of the activity of Fun30 and other proteins of interest at the fork or during helicase loading and activation has yet to be conclusively obtained in the work detailed in this study. However, further work investigating these factors with more specific assays and techniques may yet show a clear role for Fun30 and other candidate proteins at replication origins or active replication.

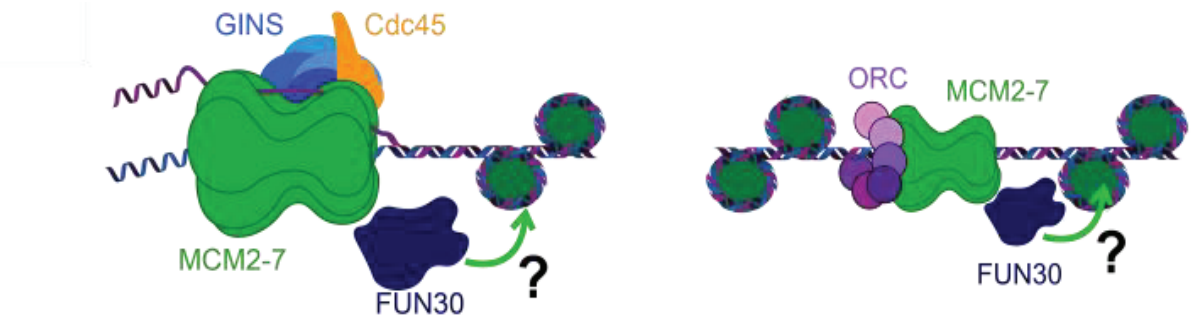


Figure 77: Models of Fun30 function in S.cerevisiae. *The data presented here and a recent study in humans (Lo et al., 2021) suggests that Fun30 associates with replication machinery, and this could be at active forks as suggested in humans (left hand side of figure), or at origins to aid in helicase loading and/or activation (right hand side of figure).*

4.4 Considerations of the ChIP-MS method

While the ChIP-MS method has shown its ability to identify enrichment of a multitude of factors, many with known links to replication and many without, there are some considerations regarding its use. Whilst the data obtained has in many cases been useful, the technique is unable to categorically state whether or not proteins are enriched because of a real association with the target complex or process, or whether they are a false positive. The very same rationale that enables the detection of transient interactors, the cross-linking, makes the technique susceptible to a potentially high false positive rate. A high nuclear protein and DNA density,

combined with the proximal activity of other DNA-associated events such as transcription and transcript processing, could possibly lead to a high rate of cross-linking to the replication machinery. Indeed, there were often large numbers of transcript processing or rRNA biogenesis factors that showed enrichment in ChIP-MS experiments. It was not entirely clear whether these were due to non-specific binding to beads or due to inadvertent pull-down due to proximity, though enrichment over untagged negative control samples would suggest the latter. RNA transcription and processing complexes are likely to have a major effect on replisome progression, so these factors may yet prove to be valid candidates for interactors or replisome-proximal factors. Regardless, the high level of false positives is a clear trend in those experiments relying on cross-linking to increase isolated complex stability. There are several approaches to combat this effect that have been attempted. Some methods use very stringent and denaturing wash buffers to ensure that only the tightest protein-protein interactions, and those stabilised by cross-links, remain (Mohammed et al., 2013; Guerrero et al., 2006). Other techniques make use of an elaborate set of negative controls or comparative conditions to best enable true interactors to be separated (Rajala et al., 2015; Natarajan, 2013).

The first approach can reduce non-specific binding false positives by separating genuine cross-linked interactors from the non-specific noise. However, if the “false-positives” are instead proteins that share a cellular compartment, and cross-link due to proximity, they will not be affected. In addition, by relying so heavily on the cross-linking to stabilise pull-down complexes, the results could be biased towards proteins that cross-link well to the tagged protein/complex, and the enrichments may be skewed towards, for example, proteins with an abundance of exposed lysine residues or that share a large interaction interface that is exposed to formaldehyde reaction from some orientation. This approach also reduces the general background that is able to be picked up in negative controls, which creates a large number of missing values. As a result, many methods utilising stringent washes operate on a rationale that

considers proteins as interaction partners only when they are detected in the pull-down, and are not detected at all in the corresponding negative control. This provides the advantage of a very binary interactor/non-interactor definition, but could lead to many missed potential interactors. In this approach, it must be accepted that many interactors could be missed, or that imputation of some form must be used to counteract the many missing values in controls.

The second approach works better at reducing the missing values, but is experimentally more challenging, since the negative control(s) must be chosen appropriately. Rajala et al. use a combination of non-crosslinked sample controls, and a control involving a tagged Luciferase that is biochemically directed towards the cellular environment from which the similarly tagged protein of interest is being isolated from (Rajala et al., 2015). This approach limits the problem of missing values, while controlling for the effect of proximity cross-linking in the cellular locale the tagged protein is in. Currently, the version of ChIP-MS used in this project uses relatively stringent washes with a high salt and high detergent content. However, it falls well short of the stringency used in other studies such as the RIME or QTAX protocols (Guerrero et al., 2006; Mohammed et al., 2013). The negative control used at present is simply an untagged strain, which does not fully control for the cellular compartment proximity cross-linking effects (Rajala et al., 2015), and such a control could be adopted in future.

Another consideration of the ChIP-MS technique used in this study, particularly when investing complex composition over time (Section 3.6.3), is that the complex being purified is quite heterogenous. Whilst this effect could be eliminated when studying initiation via the artificial complex arrests created by the intra-hexamer and inter-hexamer linkage constructs, it was still present in the time-course ChIP-MS experiment. It is inevitable that some replication complexes would be undergoing initiation at later-activated origins, whilst others would be terminating. Overall, whilst differences in factor association and therefore enrichment could be seen between these time-points, for clear statistical significance for multiple factors there

would need to be an increase in the number of replicates, or an increase in reproducibility. With the current protocol, reproducibility was controlled as much as possible, and so further replicates would be required. This would be very intensive both on processing time and instrument time, and might not be possible, especially if more time-points to achieve greater temporal resolution was required. Methods to increase signal to noise ratio could be considered to help manage this problem, such as by using a 10xFLAG tag to compensate for epitope loss due to cross-linking, and generally increase pull-down efficiency.

As with many cross-linking purifications followed by mass spectrometry, the uncertainty due to high false positives, even in the most well-controlled experiments (Rajala et al., 2015), means that protein hits require validation before they are considered "real" interactors (Guerrero et al., 2006; Wang et al., 2013; Mohammed et al., 2013; Rajala et al., 2015; Wang et al., 2017; Leissing et al., 2021), and this ChIP-MS technique is no exception. While confidence can be increased due to the identification of many well-established proteins in the replication machinery in these experiments, further work is required to properly validate some of the factors that show association with the replication machinery in this study.

The addition of formaldehyde itself to the cells might cause a perturbation in the replication machinery. Cross-linking of the entire cellular environment is not instantaneous, and formaldehyde induced cross-links could well act as impairments to DNA replication that affect the recruitment of DNA damage factors to the replication fork and to DNA-protein adducts that stall the fork (Grafstrom et al., 1983; de Graaf, Clore & McCullough, 2009; Kawanishi, Matsuda & Yagi, 2014). This might be a possible reason that DNA damage response factors such as Msh3/6 and Rad51/Rad52 were shown to be associated with replication forks in this study. Alabert et al. also showed that genome stability factors were present at unperturbed forks (Alabert et al., 2014), and follow-up research investigating factors associating with stalled or broken forks showed clear differences, suggesting the DNA damaging effects of the

formaldehyde itself cause much smaller effects than intentionally added DNA damaging agents (Nakamura et al., 2021). It is, however, a possible mechanism for the low levels of repair and damage response factors seen at forks in this study and their original work (Alabert et al., 2014). Finally, the method used for quantification of peptides in the mass spectrometry in this study was label free quantification (LFQ) in MaxQuant (Cox et al., 2014). Though this normalisation method is very robust, it is possible that stark differences between pull-down and untagged control samples could impair the ability of this software to provide accurate quantification. However, in this instance, the presence of a non-specific background in the untagged negative controls in the ChIP-MS data was thought to aid in the normalisation since the samples had a greater overlap in terms of identified proteins than if there were no non-specific background. It seems that the differences between pull-down and negative control samples did not adversely affect this normalisation, something that was observed when attempting to apply LFQ across input and pull-down samples together (data not shown). Comparing results with and without this normalisation step did not drastically change the results of any ChIP-MS experiments, and given the protein loading onto MS instruments was normalised as much as possible, this normalisation was considered effective and appropriate in this use case. Other label-free methods have also proven similarly effective on pull-down data (Choi et al., 2012).

4.5 Future perspectives

The application of chemical biology and proteomic techniques has produced many valuable insights in this study. The development of the inter-hexamer latch construct to prevent hexamer separation and the development of the ChIP-MS technique additionally provide many further lines of inquiry. The definition of the molecular mechanism that underlies the inter-hexamer latch arrest and the prevention of Sld3 release during initiation in the presence of this linkage is one that requires additional investigation. Validation that the linkage is functioning as

designed, by preventing hexamer separation or lateral movement past one another, is another requirement for building on these findings in a meaningful way. Further work on the intra-hexamer gate constructs has been carried out using genome-wide techniques to establish binding of RPA and detection of ssDNA production at a base pair resolution (L.M. Reuter, pers. comm.) and this is an approach that might be considered for the inter-hexamer latch. Additionally, the construct could be instead used in an *in vitro* context, and investigated structurally, using cryo-EM or by cross-linking mass spectrometry techniques to compare intermediate structures in replication initiation with and without the presence of the latch inter-hexamer linkage.

Regarding the ChIP-MS technique, there are many possible future directions for this method. By further optimising the protocol to ensure reproducibility and efficient release of protein complexes from chromatin, and by considering alternative negative control types and washing conditions, future experiments may be able to investigate complex composition in a range of contexts with greater precision and reliability. Improvements to temporal resolution in the time-course experiments as carried out in this study could further define the dynamics of many replication factors, and analyse replication complexes in contexts such as termination, stalled forks, or in other DNA-damaging conditions. Furthermore, in addition to several other studies (Vasilescu, Guo & Kast, 2004; Schmitt-Ulms et al., 2004; Guerrero et al., 2006; Mohammed et al., 2013; Wang et al., 2013; Natarajan, 2013; Yu et al., 2016; Rajala et al., 2015; Leissing et al., 2021), this research suggests that formaldehyde cross-linking followed by affinity purification coupled to mass spectrometry analysis can be a useful tool to investigate labile or transient complexes in a range of organisms.

Extensive modifications could also be made to the ChIP-MS technique described here, to allow for alternative functionaility. Addition of a second purification step, as originally utilised for GINS-based purification of replication complexes (Gambus et al., 2006), could ensure greater

specificity of the pull-down and increased purity of replication complexes. Secondary purification of a tagged factor unique to a specific stage of replication, e.g. Sld3, could much reduce heterogeneity in complex composition, though yields would likely be substantially decreased, requiring much larger cell cultures. Additionally, whilst mass-spectrometry cleavable cross-linkers, able to identify the locations of protein-protein cross-links and provide structural information (Yu & Huang, 2018), are not feasible *in vivo* with yeast culture due to the large volumes involved and a relatively impermeable cell wall, a combinatorial approach could be considered. Formaldehyde cross-linking *in vivo* to stabilise complexes and fix cells as used here, could be combined with *in vitro* cross-linking using specialised cross-linking reagents once cell extract was prepared. Such combinations of formaldehyde and MS-cleavable cross-linkers have been used fully *in vitro* to stabilise complexes for structural studies, but not yet used in a purification-based technique (Yu & Huang, 2018). This could expand the capabilities of the technique beyond just quantification and identification of interacting proteins, but also allow generation of structural data as well.

In conclusion, the ChIP-MS technique, as developed in this project, represents a versatile and informative technique that can be applied to a range of biological questions, and has been applied in this case to provide a large amount of information regarding the changing interactome of the DNA helicase during helicase loading, helicase activation, and DNA replication.

4.6 Summary

The promising findings generated by use of the rapamycin-inducible helicase linkages, and the ChIP-MS method, will also be promising starting points for further research. The function of Fun30 could be further interrogated utilising the deletion strain generated as part of this work, and subjecting it to further genotoxic stress, monitor fork speed in the deletion strain, analyse Fun30 fork association (for example, by using eSPAN (Yu, Gan & Zhang, 2018)), and/or probe a greater number of genomic sites for the binding of Fun30 under various additional environmental and cell cycle conditions.

The association of Sld3 with the splitting of the double hexamer remains one of the most promising aspects of this work. In the absence of any structural data regarding the activation complex, this presents the first indication of the mechanistic role of Sld3 in this process. Does Sld3 have an active role in the mechanistic splitting of the hexamers, perhaps working in tandem with pre-LC components as a physical wedge? Or, does Sld3 act as a quality control component, only able to associate with double-hexamer, and, once the hexamers separate, the binding site of Sld3 is lost? With the suggestion of a function, work generating specific mutations to interrupt the function of this protein can be more targeted. This suggested function also suggests that Sld3 is likely, but not certainly, to show localisation in an activation complex that is somewhat localised to the interface of the two double hexamers, which may guide structural studies.

5 Appendices

5.1 Appendix 1 – Plasmids

Plasmid	Speck lab identifier	description	source
Mcm4-FRB (203) synthesised sequence	451	Synthesised Mcm4-FRB sequence with FRB+linkers inserted after Mcm4 amino acid 203	(Samel et al., 2014)
Mcm4-FRB (252) synthesised sequence	623	Synthesised Mcm4-FRB sequence with FRB+linkers inserted after Mcm4 amino acid 252	(Samel et al., 2014)
His-selection vector for PCR amplification	80	pFA6A-3HA-His3MX6 (pYM2) vector used for subcloning of Mcm4-FRB sequences into for PCR amplification	(Janke et al., 2004)
amplification base Mcm4-FRB (aa 252)	1095	pFA6a-Mcm4-FRB(aa 203)-stop-3HA-His3MX6 template for amplification of integration PCR product	this study (subcloning of 451 into 80)
amplification base Mcm4-FRB (aa 203)	1096	pFA6a-Mcm4-FRB(aa 252)-stop-3HA-His3MX6 template for amplification of integration PCR product	this study (subcloning of 623 into 80)
amplification template of a 5xFLAG integron (*hph*)	762	pKL255 (5xFLAG-hphNT) plasmid containing 5xFLAG and hygromycin resistance gene	gift from K.Labib
amplification template of a 5xFLAG integron (URA)	764	pKL506 (5xFLAG-URA3CP) plasmid containing 5xFLAG and hygromycin resistance gene	gift from K.Labib
HO endonuclease expression	792	pJH283 HO endonuclease under Galactose-inducible transcription, for mating type switching	gift from J.Haber (Rudin & Haber, 1988)

5.2 Appendix 2 – Oligonucleotides

Oligonucleotide	Speck lab identifier	sequence	source
Mcm4-FRB +RE site amplification fwd	1884	GCTACCCGGGATGTCTCAACAGTCTAGC TCTCCAACAA	S. Samel (Samel et al., 2014)
Mcm4-FRB +RE site amplification rev	1885	GCTATTAATTAATCAGACACGGTTATTCA GGC	S. Samel (Samel et al., 2014)
Mcm4-FRB integration primer fwd	1890	ATGTCTCAACAGTCTAGCTC	S. Samel (Samel et al., 2014)
Mcm4-FRB integration primer rev	1891	GCTATAATTAATTAGTATTTATTAATTGTTA CGCAGGGAATGATTGTAGTAGACAGCAG AATTCGAGCTCGTTTAAAC	S. Samel (Samel et al., 2014)
Cdc45 tagging fwd	2274	GTGAAGATCTTTCACCATTCCTGGAGAAG CTGACCTTGAGTGGATTGTTACGTACGCT GCAGGTCGAC	L.M. Reuter (based on Janke et al. 2004)
Cdc45 tagging/deletion rev	2275	ATATTCATATGCTGGTATATATGTACGACT AAATAATATAAATTTGATTAATCGATGAAT TCGAGCTCG	L.M. Reuter (based on Janke et al. 2004)
Mcm10 tagging fwd	2276	GGAAACTAAAGAAACTTCTGACGGTAGTG CCAGCGATCTTGAGATAATACGTACGCTG CAGGTCGAC	L.M. Reuter (based on Janke et al. 2004)
Mcm10 tagging/deletion rev	2277	CAAAAGTAATACCATTTTGGGGCCCTGAA AAGCACACCAATACTTATTTAATCGATGA ATTCGAGCTCG	L.M. Reuter (based on Janke et al. 2004)
Rfa1 tagging fwd	2278	GCTGAAGCCGACTATCTTGCCGATGAGTT ATCCAAGGCTTTGTTAGCTCGTACGCTGC AGGTCGAC	L.M. Reuter (based on Janke et al. 2004)
Rfa1 tagging/deletion rev	2279	TTTCTCATATGTTACATAGATTAAATAGTA CTTGATTATTTGATACATTAATCGATGAAT TCGAGCTCG	L.M. Reuter (based on Janke et al. 2004)
Mcm4 tagging fwd	2280	GTCCTTGGCGAGGGTGTAAGGAGATCAG TTCGCCTGAATAACCGTGTCCGTACGCTG CAGGTCGAC	L.M. Reuter (based on Janke et al. 2004)
Mcm4 tagging/deletion rev	2251	TAATTAGTATTTATTAATTGTTACGCAGGG AATGATTGTAGTAGACAGCAATCGATGAA TTCGAGCTCG	L.M. Reuter (based on Janke et al. 2004)
Pol2 tagging fwd	2285	GTATTACGGTTTTGATATATTATTGAGTTG TATTGCTGATTTGACCATACGTACGCTGC AGGTCGAC	L.M. Reuter (based on Janke et al. 2004)
Pol2 tagging/deletion rev	2286	TTTTCATGGTAAAGAGGCCATTGAACCTC GCGTTATATACTGCTTACTCAATCGATGA ATTCGAGCTCG	L.M. Reuter (based on Janke et al. 2004)
Sld2 tagging fwd	2295	CTAAAACTGCCAAAGAAAAACCGATTCTC TAATGGACGATGGGGAAGAAGGCGTACG CTGCAGGTCGAC	L.M. Reuter (based on Janke et al. 2004)
Sld2 tagging/deletion rev	2296	TAGGTTTCTATAAATTACAAATGTTTGTAT TATTTACGCCATCACGCTCAATCGATGAA TTCGAGCTCG	L.M. Reuter (based on Janke et al. 2004)
Sld3 tagging fwd	2297	ATAGCTCAAAAAGGAGAGTAAGAAGACGT TTATTTGCTCCAGAATCCACACGTACGCT GCAGGTCGAC	L.M. Reuter (based on Janke et al. 2004)

Sld3 tagging/deletion rev	2298	GTTGGGTCAACTACTTGCCAGCGAAGGT GGTCTCGCAGTTTCTTATTCTAATCGATG AATTCGAGCTCG	L.M. Reuter (based on Janke et al. 2004)
Fun30 tagging fwd	2683	TGGAGGATATAATTTATGATGAAAACTCG AAACCGAAGGGAACCAAAGAACGTACGC TGCAGGTCGAC	this study (based on Janke et al. 2004)
Fun30 tagging/deletion rev	2684	TTTGGTTTATTTTCTGCTTATCTATTTACTT TTTTACTATATTTTTATTTATTTAATCGATG AATTCGAGCTCG	this study (based on Janke et al. 2004)
Hrq1 tagging fwd	2707	GAGCTACGAAAGACGATACTCATACAAAT GAAATCATTAAAAAAGAGATACGTACGCT GCAGGTCGAC	this study (based on Janke et al. 2004)
Hrq1 tagging/deletion rev	2708	GTAGTAGAATAGAGTATTTATATTCGGTTT ACAAACTACAAATAGCGTGCTTAATCGAT GAATTCGAGCTCG	this study (based on Janke et al. 2004)
Rrm3 tagging fwd	2753	GCACCAATGAAAGAGTAAAAGATTTCTAT AAACGTTTAGAAACTTTGAAACGTACGCT GCAGGTCGAC	this study (based on Janke et al. 2004)
Rrm3 tagging/deletion rev	2754	AAGAAAAGAAAACTTCAACTAGAGTATAT GCATTTATTCGTTGCAAGTCAATCGATGA ATTCGAGCTCG	this study (based on Janke et al. 2004)
Sir3 tagging fwd	2755	TAAATTACGCCTTTTCGATGGATGAAGAA TTCAAAAATATGGACTGCATTCGTACGCT GCAGGTCGAC	this study (based on Janke et al. 2004)
Sir3 tagging/deletion rev	2756	CATAGGCATATCTATGGCGGAAGTGAAAA TGAATGTTGGTGGTCAATCGATGAATTCG AGCTCG	this study (based on Janke et al. 2004)
Cac2 tagging fwd	2759	CGATAGTAAAAAGAGGCGCATACATCCTA CGCCAGTCGATTTGCGTACGCTGCAGGT CGAC	this study (based on Janke et al. 2004)
Cac2 tagging/deletion rev	2760	GCTGGCAAACTAGATTAGGCATTCTTATG TACCGCATTAAATATATTAAAAAATCGATG AATTCGAGCTCG	this study (based on Janke et al. 2004)
Fun30 deletion fwd	2687	CAGTAAGGAACGTAAACAAGAAAAAGAGA GAAAATACGCTATAGTTGAAAACATGCGT ACGCTGCAGGTCGAC	this study (based on Janke et al. 2004)
Rrm3 deletion fwd	2771	GAACAAGCTCAAAAGTCGAGAGATTTGTT CTTATAAGACATCCCGATGCGTACGCTGC AGGTCGAC	this study (based on Janke et al. 2004)
Sir3 deletion fwd	2772	CAGGGGTTTAAGAAAGTTGTTTTGTTCTA ACAATTGGATTAGCTAAAATGCGTACGCT GCAGGTCGAC	this study (based on Janke et al. 2004)
Cac2 deletion fwd	2775	GAAGGAGGCAGAAAAGCTTAGAGCAAAG GGGCAGGACAATGCGTACGCTGCAGGTC GAC	this study (based on Janke et al. 2004)
Hrq1 deletion fwd	2776	GAGATGGTTAAGGTCGTAGAAAAGAAATG TTCATTTGAGAAGGAAAAATGCGTACGCT GCAGGTCGAC	this study (based on Janke et al. 2004)
colPCR for *hph* selection integration fwd	2369	TCGGGCGTACACAAATC	L.M. Reuter
colPCR for URA selection integration fwd	2784	CATCGTAGGTCGTGGTC	this study
colPCR for Cdc45 rev	2370	AGCCGCGCACAAAATATG	L.M. Reuter

colPCR for Mcm10 rev	2682	ATTGCAACCAAAATCAC	L.M. Reuter
colPCR for Fun30 rev	2685	CCGTGCCCATTTTAGCGGGC	this study
colPCR for Sir3 rev	2780	CTGTCCTGAAGTTTGATTAGG	this study
colPCR for Cac2 rev	2781	GTTGAAATCGTGATGACGATG	this study
colPCR for Hrq1 rev	2782	GAGGACATTAAGCGTAAAATGG	this study
colPCR for Rrm3 rev	2783	GCATACGTAGGCTTAGCTAA	this study
colPCR for Mcm4 rev	2799	AAATGTAATCATCCTTCGGG	L.M. Reuter
qPCR rRNA 5S-2-fw	2244	AAGATTGCAGCACCTGAGTTTCG	L.M. Reuter
qPCR rRNA 5S-2-rev	2245	ATGTCTGGACCCTGCCCTCATAT	L.M. Reuter
qPCR ARS305_fw	2258	AGCCTTCTTTGGAGCTCAAG	L.M. Reuter
qPCR ARS305_rev	2259	TTTGAGGAATTTCTTTTGAAGAG	L.M. Reuter
qPCR ARS305_+2fw	2260	CGATAACGGCTAAAACAATTAAA	L.M. Reuter
qPCR ARS305_+2rev	2261	GCTAGAACAGCGACAATGTTTTTG	L.M. Reuter
qPCR ARS305_+4fw	2262	GGTACAATCATAGAAAAAGGG	L.M. Reuter
qPCR ARS305_+4rev	2263	GGATTTGCTGAGTTCGCGTA	L.M. Reuter
qPCR ARS305_+6fw	2264	TGTTCTTCATTAGCGATCAGATAAGC	L.M. Reuter
qPCR ARS305_+6rev	2265	AAACAACAAACCTAAACACTCTTGGTATT	L.M. Reuter
qPCR ARS305_+9fw	2266	AACTTGGCCTTATGTAGAATTTCTT	L.M. Reuter
qPCR ARS305_+9rev	2267	AGCAATTCCACCGACCATAC	L.M. Reuter
qPCR ARS501_fw	2272	CTCATCATCATCCCCGGTA	L.M. Reuter

qPCR ARS501_rev	2273	CGTACACTAGCCCGTTGAGGT	L.M. Reuter
qPCR ARS607_fw	2688	GTGGTGATATAAACACTACATTCGC	L.M. Reuter
qPCR ARS607_rev	2689	TCCATAATATTTACCTTACGCTGGG	L.M. Reuter
qPCR ARS1021_fw	2692	GCGTGTGTTGGTTACTTAGGTT	L.M. Reuter
qPCR ARS1021_rev	2693	GGCTCTGTTTTGTTTATTTCCAGAT	L.M. Reuter
qPCR ARS1429_fw	2696	AGTTCTTGGACAAGATTGGATACAT	L.M. Reuter
qPCR ARS1429_rev	2697	TTACCTTTCAATGCGGCACG	L.M. Reuter

5.3 Appendix 3 – Yeast strains

- * indicates the use in a ChIP-MS study

Strain	Speck lab identifier	Genotype	Source
W303 WT*	YC160	*MATa ade2-1 ura3-1 his3-11,15 trp1-1 leu2-3, 112 can1-100 GAL+*	laboratory strain
Mcm2/Mcm5 C-terminal intra-hexamer linkage*	YC525	as YC160 plus: *psi+ TOR1-1 mcm2::mcm2 FRB-long (aa 712)-HisMX6 mcm5::mcm5 FKBP-long (aa 591)-kanMX6Δ Δfpr1::NAT*	L.M. Reuter (Speck lab)
Mcm2/Mcm5 N-terminal intra-hexamer linkage *	YC536	as YC160 plus: *psi+ TOR1-1 mcm2::mcm2 FRB-long (aa 221)-HisMX6 mcm5::mcm5 FKBP-long (aa 72)-kanMX6Δ Δfpr1::NAT*	L.M. Reuter (Speck Lab)
Mcm4/Mcm5 inter-hexamer linkage*	YC671	as YC160 plus: *psi+ TOR1-1 mcm4::mcm4 FRB-long (aa 252)-HisMX6 mcm5::mcm5 FKBP-long (aa 72)-kanMX6Δ Δfpr1::NAT*	this study
Mcm4/Mcm5 inter-hexamer linkage, BrdU-incorporation strain	YC729	As YC671 plus: *URA::GPD-TK5x, AUR1c::ADH1-hE*	This study
Rfa1-FLAG Mcm2/Mcm5 N-linkage	YC546	as YC536 plus: *rfa1::rfa1-5xFLAG-hph*	L.M. Reuter (Speck lab)
Mcm4-FLAG Mcm2/Mcm5 N-linkage*	YC547	as YC536 plus: *mcm4::mcm4-5xFLAG-hph*	L.M. Reuter (Speck Lab)
Mcm10-FLAG Mcm2/Mcm5 N-linkage	YC552	as YC536 plus: *mcm10::mcm10-5xFLAG-hph*	L.M. Reuter (Speck lab)

Pol2-FLAG Mcm2/Mcm5 N-linkage	YC558	as YC536 plus: p*ol2::pol2-5xFLAG-hph*	L.M. Reuter (Speck Lab)
Sld2-FLAG Mcm2/Mcm5 N-linkage	YC572	as YC536 plus: *sld2::sld2-5xFLAG-hph*	L.M. Reuter (Speck lab)
Sld3-FLAG Mcm2/Mcm5 N-linkage	YC574	as YC536 plus: sld3*::sld3-5xFLAG-hph*	L.M. Reuter (Speck Lab)
Psf2-TAP Mcm4-FLAG Mcm2/Mcm5 N-linkage	YC578	as YC536 plus: *psf2::psf2-TAP (Psf2-V5-PreSc-ProtA) mcm4::mcm4-5xFLAG-hph*	L. Karl (Speck Lab)
Mcm10-FLAG Mcm2/Mcm5 C-linkage	YC580	as YC525 plus: *mcm10::mcm10-5xFLAG-hph*	L.M. Reuter (Speck lab)
Cdc45-FLAG Mcm2/Mcm5 C-linkage	YC584	as YC525 plus: *cdc45::cdc45-5xFLAG-hph*	L.M. Reuter (Speck Lab)
Mcm4-FLAG Mcm2/Mcm5 C-linkage*	YC588	as YC525 plus: *mcm4::mcm4-5xFLAG-hph*	L.M. Reuter (Speck lab)
Pol2-FLAG Mcm2/Mcm5 C-linkage	YC590	as YC525 plus: *pol2::pol2-5xFLAG-hph*	L.M. Reuter (Speck Lab)
Rfa1-FLAG Mcm2/Mcm5 C-linkage	YC592	as YC525 plus: *rfa1::rfa1-5xFLAG-hph*	L.M. Reuter (Speck lab)
Sld2-FLAG Mcm2/Mcm5 C-linkage	YC594	as YC525 plus: *sld2::sld2-5xFLAG-hph*	L.M. Reuter (Speck Lab)
Sld3-FLAG Mcm2/Mcm5 C-linkage	YC596	as YC525 plus: *sld3::sld3-5xFLAG-hph*	L.M. Reuter (Speck Lab)
Mcm10-FLAG Mcm4/Mcm5 linkage	YC675	as YC671 plus: *mcm10::mcm10-5xFLAG-hph*	this study
Cdc45-FLAG Mcm4/Mcm5 linkage	YC676	as YC671 plus: *cdc45::cdc45-5xFLAG-hph*	this study
Sld3-FLAG Mcm4/Mcm5 linkage	YC678	as YC671 plus: *sld3::sld3-5xFLAG-hph*	this study

Pol2-FLAG Mcm4/Mcm5 linkage	YC679	as YC671 plus: *pol2::pol2-5xFLAG-hph*	this study
Rfa1-FLAG Mcm4/Mcm5 linkage	YC680	as YC671 plus: *rfa1::rfa1-5xFLAG-hph*	this study
Mcm2-FLAG Mcm4/Mcm5 linkage	YC682	as YC671 plus: *mcm2::mcm2-5xFLAG-hph*	this study
Sld2-FLAG Mcm4/Mcm5 linkage	YC683	as YC671 plus: *sld2::sld2-5xFLAG-hph*	this study
Mcm4-FLAG Mcm4/Mcm5 linkage*	YC829	as YC671 plus: *mcm4-FRB-HisMX6::mcm4-FRB-5xFLAG-hph*	this study
Mcm4-FLAG*	YC779	as YC160 plus: *mcm4::mcm4-5xFLAG-hph*	this study
Cdc45-FLAG*	YC785	as YC160 plus: *cdc45::cdc45-5xFLAG-hph*	this study
Fun30-FLAG	YC809	as YC160 plus: *fun30::fun30-5xFLAG-hph*	this study
Sir3-FLAG	YC862	as YC160 plus: *sir3::sir3-5xFLAG-hph*	this study
Cac2-FLAG	YC863	as YC160 plus: *cac2::cac2-5xFLAG-hph*	this study
Hrq1-FLAG	YC864	as YC160 plus: *hrq1::hrq1-5xFLAG-hph*	this study
Rrm3-FLAG	YC865	as YC160 plus: *rrm3::rrm3-5xFLAG-hph*	this study
pMet:Cdc20 for methionine induced mitotic arrest	YC874	as YC160 plus: cdc20::pMet-cdc20-TRP1	this study
pMet:Cdc20 Mcm4-FLAG	YC840	as YC874 plus: *mcm4::mcm4-5xFLAG-hph*	this study
Δfun30	YC803	as YC160 plus: *fun30::hph*	this study

Δsir3	YC878	as YC160 plus: *sir3::hph*	this study
Δcac2	YC876	as YC160 plus: *cac2::hph*	this study
Δhrq1	YC889	as YC160 plus: *hrq1::hph*	this study
Δrrm3	YC877	as YC160 plus: *rrm3::hph*	this study
Δfun30 Mcm4-FLAG	YC920	as YC803 plus: *mcm4::mcm4-5xFLAG-URA*	this study
Δrrm3 Mmc4-FLAG	YC912	as YC877 plus: *mcm4::mcm4-5xFLAG-URA*	this study
mating type test strain (14a)	YC533	*MATa his1*	Sträßer lab
mating type test strain (17α)	YC534	*MATα his1*	Sträßer lab

5.4 Appendix 4 – Antibodies

Antibody	Speck lab identifier	use	species	source
anti-Mcm2	5	Western blotting (primary)	mouse	CSHL in-house (49-Mcm2)
anti-Sld5	42	Western blotting (primary)	rabbit	Genscript (108-Sld5)
anti-panMCM	2	Western blotting (primary)	rabbit	CSHL in-house (CS687)
anti-FLAG-HRP	100	Western blotting (primary+HRP)	mouse	Sigma -A8592
PAP	secondary 17	Western blotting (anti-HRP)	rabbit	Sigma -P1291
anti-mouse	secondary 18	Western blotting (secondary)	goat	Sigma -A4416
anti-rabbit	secondary 3	Western blotting (secondary)	goat	Sigma -A6154
anti-BrdU	108	dot-blotting	mouse	BD Biosciences - 555627
anti-FLAG	25	ChIP-qPCR and ChIP-MS	mouse	Sigma -1804
anti-GFP	113	ChIP-qPCR	mouse	Thermo -MA5-15256
rabbit IgG	106	IP-MS and ChIP-MS	rabbit	Jackson Immune Research - 011-000-003

5.5 Supplementary Figures

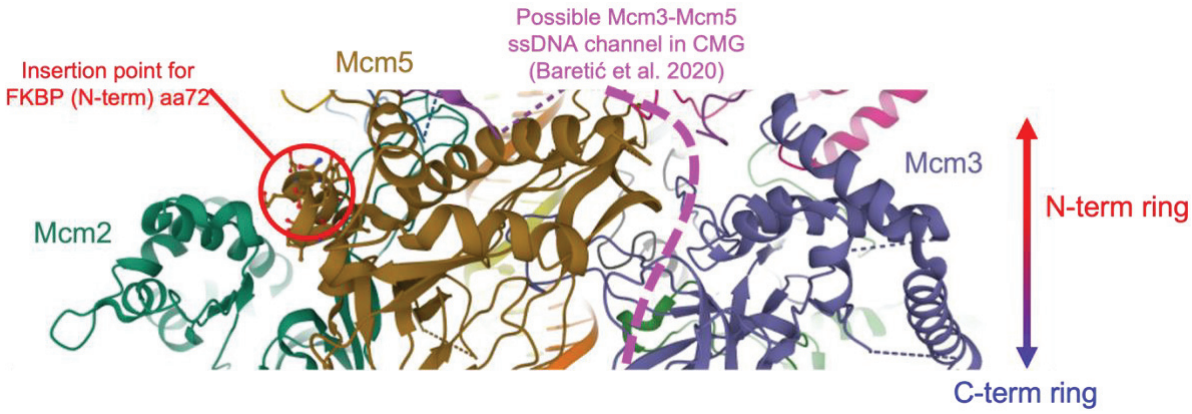


Figure S1: Supplementary figure showing the proposed ssDNA exit channel observed in recent structural studies (Baretić et al., 2020), alongside the insertion point of the FKBP linkers on Mcm5 for the latch construct.

Structural data derived from the structure of the yeast double hexamer on origin DNA (Noguchi et al., 2017)(PDB: 5BK4). Path of the exit channel drawn based on the trajectory proposed by Baretić et al. The insertion point and subsequent localisation of the inserted FKBP is far enough away from the ssDNA exit channel as to not be likely to interfere with this process directly.

6 Bibliography

Abdel-Monem, M., Dürwald, H. & Hoffman-Berling, H. (1976) Enzymic Unwinding of DNA. *European Journal of Biochemistry*. 65 (2), 441–449. doi:10.1111/j.1432-1033.1976.tb10359.x.

Abdurashidova, G., Deganuto, M., Klima, R., Riva, S., Biamonti, G., Giacca, M. & Falaschi, A. (2000) Start sites of bidirectional DNA synthesis at the human lamin B2 origin. *Science (New York, N.Y.)*. 287 (5460), 2023–2026. doi:10.1126/SCIENCE.287.5460.2023.

Abid Ali, F. & Costa, A. (2016) The MCM Helicase Motor of the Eukaryotic Replisome. *Journal of Molecular Biology*. 428 (9), 1822–1832. doi:10.1016/J.JMB.2016.01.024.

Abid Ali, F., Renault, L., Gannon, J., Gahlon, H.L., Kotecha, A., Zhou, J.C., Rueda, D. & Costa, A. (2016) Cryo-EM structures of the eukaryotic replicative helicase bound to a translocation substrate. *Nature Communications*. 7. doi:10.1038/NCOMMS10708.

Abrams, E., Neigeborn, L. & Carlson, M. (1986) Molecular analysis of SNF2 and SNF5, genes required for expression of glucose-repressible genes in Saccharomyces cerevisiae. *Molecular and Cellular Biology*. 6 (11), 3643. doi:10.1128/MCB.6.11.3643.

Adra, C.N., Donato, J.L., Badovinac, R., Syed, F., Kheraj, R., Cai, H., Moran, C., Kolker, M.T., Turner, H., Weremowicz, S., Shirakawa, T., Morton, C.C., Schnipper, L.E. & Drews, R. (2000) SMARCAD1, a Novel Human Helicase Family-Defining Member Associated with Genetic Instability: Cloning, Expression, and Mapping to 4q22–q23, a Band Rich in Breakpoints and Deletion Mutants Involved in Several Human Diseases. *Genomics*. 69 (2), 162–173. doi:10.1006/GENO.2000.6281.

Aebersold, R. & Mann, M. (2003) Mass spectrometry-based proteomics. *Nature 2003 422:6928*. 422 (6928), 198–207. doi:10.1038/nature01511.

Alabert, C., Bukowski-Wills, J.C., Lee, S.B., Kustatscher, G., Nakamura, K., de Lima Alves, F., Menard, P., Mejlvang, J., Rappsilber, J. & Groth, A. (2014a) Chromatin Dynamics During DNA Replication and Uncharacterized Replication Factors determined by Nascent Chromatin Capture (NCC) Proteomics. *Nature cell biology*. 16 (3), 281. doi:10.1038/NCB2918.

Alabert, C., Bukowski-Wills, J.-C., Lee, S.-B., Kustatscher, G., Nakamura, K., de Lima Alves, F., Menard, P., Mejlvang, J., Rappsilber, J. & Groth, A. (2014b) Nascent chromatin capture proteomics determines chromatin dynamics during DNA replication and identifies unknown fork components. *Nature cell biology*. 16 (3), 281–293. doi:10.1038/ncb2918.

Alabert, C. & Groth, A. (2012) Chromatin replication and epigenome maintenance. *Nature Reviews Molecular Cell Biology 2012 13:3*. 13 (3), 153–167. doi:10.1038/nrm3288.

Alcasabas, A.A., Osborn, A.J., Bachant, J., Hu, F., Werler, P.J.H., Bousset, K., Furuya, K., Diffley, J.F.X., Carr, A.M. & Elledge, S.J. (2001a) Mrc1 transduces signals of DNA replication stress to activate Rad53. *Nature cell biology*. 3 (11), 958–965. doi:10.1038/NCB1101-958.

Ali, F.A., Douglas, M.E., Locke, J., Pye, V.E., Nans, A., Diffley, J.F.X. & Costa, A. (2017) Cryo-EM structure of a licensed DNA replication origin. *Nature Communications*. 8 (1), 2241. doi:10.1038/s41467-017-02389-0.

Altelaar, A.F.M., Frese, C.K., Preisinger, C., Hennrich, M.L., Schram, A.W., Timmers, H.T.M., Heck, A.J.R. & Mohammed, S. (2013) Benchmarking stable isotope labeling based quantitative proteomics. *Journal of Proteomics*. 88, 14–26. doi:10.1016/J.JPROT.2012.10.009.

Alvino, G.M., Collingwood, D., Murphy, J.M., Delrow, J., Brewer, B.J. & Raghuraman, M.K. (2007) Replication in Hydroxyurea: It's a Matter of Time. *Molecular and Cellular Biology*. 27 (18), 6396–6406. doi:10.1128/MCB.00719-07/SUPPL_FILE/ALVINO_SUPPLEMENTAL.ZIP.

Alzu, A., Bermejo, R., Begnis, M., Lucca, C., Piccini, D., Carotenuto, W., Saponaro, M., Brambati, A., Cocito, A., Foiani, M. & Liberi, G. (2012) Senataxin Associates with Replication Forks to Protect Fork Integrity across RNA-Polymerase-II-Transcribed Genes. *Cell*. 151 (4), 835. doi:10.1016/J.CELL.2012.09.041.

Aparicio, J.G., Viggiani, C.J., Gibson, D.G. & Aparicio, O.M. (2004) The Rpd3-Sin3 Histone Deacetylase Regulates Replication Timing and Enables Intra-S Origin Control in Saccharomyces cerevisiae . *Molecular and Cellular Biology*. 24 (11), 4769–4780. doi:10.1128/MCB.24.11.4769-4780.2004/ASSET/1F713E3A-8F37-42E0-9E7E-213E83FDF5BD/ASSETS/GRAPHIC/ZMB0110440870007.JPEG.

Aparicio, O.M. (2007) Location, location, location: it's all in the timing for replication origins. *Natsume and Tanaka*. doi:10.1101/gad.209999.112.

Aparicio, O.M., Billington, B.L. & Gottschling, D.E. (1991) Modifiers of position effect are shared between telomeric and silent mating-type loci in S. cerevisiae. *Cell*. 66 (6), 1279–1287. doi:10.1016/0092-8674(91)90049-5.

Aparicio, O.M., Stout, A.M. & Bell, S.P. (1999) Differential assembly of Cdc45p and DNA polymerases at early and late origins of DNA replication. *Proceedings of the National Academy of Sciences*. 96 (16), 9130–9135. doi:10.1073/PNAS.96.16.9130.

Aparicio, O.M., Weinstein, D.M. & Bell, S.P. (1997) Components and Dynamics of DNA Replication Complexes in S. cerevisiae: Redistribution of MCM Proteins and Cdc45p during S Phase. *Cell*. 91 (1), 59–69. doi:10.1016/S0092-8674(01)80009-X.

Araki, H., Leem, S.H., Phongdara, A. & Sugino, A. (1995) Dpb11, which interacts with DNA polymerase II(epsilon) in Saccharomyces cerevisiae, has a dual role in S-phase progression and at a cell cycle checkpoint. *Proceedings of the National Academy of Sciences of the United States of America*. 92 (25), 11791. doi:10.1073/PNAS.92.25.11791.

Aria, V. & Yeeles, J.T.P. (2018) Mechanism of Bidirectional Leading-Strand Synthesis Establishment at Eukaryotic DNA Replication Origins. *Molecular cell*. 73 (2), 199-211.e10. doi:10.1016/J.MOLCEL.2018.10.019.

Asturias, F.J., Cheung, I.K., Sabouri, N., Chilkova, O., Wepplo, D. & Johansson, E. (2005) Structure of Saccharomyces cerevisiae DNA polymerase epsilon by cryo–electron microscopy. *Nature Structural & Molecular Biology 2005 13:1*. 13 (1), 35–43. doi:10.1038/nsmb1040.

Austin, R.J., Orr-Weaver, T.L. & Bell, S.P. (1999) Drosophila ORC specifically binds to ACE3, an origin of DNA replication control element. *Genes & development*. 13 (20), 2639–2649. doi:10.1101/GAD.13.20.2639.

Awad, S., Ryan, D., Prochasson, P., Owen-Hughes, T. & Hassan, A.H. (2010) The Snf2 Homolog Fun30 Acts as a Homodimeric ATP-dependent Chromatin-remodeling Enzyme. *The Journal of Biological Chemistry*. 285 (13), 9477. doi:10.1074/JBC.M109.082149.

Azvolinsky, A., Dunaway, S., Torres, J.Z., Bessler, J.B. & Zakian, V.A. (2006) The S. cerevisiae Rrm3p DNA helicase moves with the replication fork and affects replication of all yeast chromosomes. *Genes and Development*. 20 (22), 3104–3116. doi:10.1101/gad.1478906.

Bacal, J., Moriel-Carretero, M., Pardo, B., Barthe, A., Sharma, S., Chabes, A., Lengronne, A. & Pasero, P. (2018) Mrc1 and Rad9 cooperate to regulate initiation and elongation of DNA replication in response to DNA damage. *The EMBO journal*. 37 (21). doi:10.15252/EMBJ.201899319.

Badeaux, A.I. & Shi, Y. (2013) Emerging roles for chromatin as a signal integration and storage platform. *Nature Reviews Molecular Cell Biology 2013 14:4*. 14 (4), 211–224. doi:10.1038/nrm3545.

Bailey, R., Moreno, S.P. & Gambus, A. (2015) Termination of DNA replication forks: “Breaking up is hard to do.” *Nucleus*. 6 (3), 187. doi:10.1080/19491034.2015.1035843.

Bando, M., Katou, Y., Komata, M., Tanaka, H., Itoh, T., Sutani, T. & Shirahige, K. (2009) Csm3, Tof1, and Mrc1 Form a Heterotrimeric Mediator Complex That Associates with DNA Replication Forks. *The Journal of Biological Chemistry*. 284 (49), 34355. doi:10.1074/JBC.M109.065730.

Bantele, S.C., Ferreira, P., Gritenaite, D., Boos, D. & Pfander, B. (2017) Targeting of the Fun30 nucleosome remodeller by the Dpb11 scaffold facilitates cell cycle-regulated DNA end resection. *eLife*. 6, e21687. doi:10.7554/eLife.21687.

Bantele, S.C.S., Lisby, M. & Pfander, B. (2019) Quantitative sensing and signalling of single-stranded DNA during the DNA damage response. *Nature Communications 2019 10:1*. 10 (1), 1–12. doi:10.1038/s41467-019-08889-5.

Bantele, S.C.S. & Pfander, B. (2019) Nucleosome Remodeling by Fun30SMARCAD1 in the DNA Damage Response. *Frontiers in Molecular Biosciences*. 0, 78. doi:10.3389/FMOLB.2019.00078.

Barbet, N.C., Schneider, U., Helliwell, S.B., Stansfield, I., Tuite, M.F. & Hall, M.N. (1996) TOR controls translation initiation and early G1 progression in yeast. *Molecular biology of the cell*. 7 (1), 25–42. doi:10.1091/MBC.7.1.25.

Bardin AJ, Amon A. Men and sin: what's the difference? Nat Rev Mol Cell Biol. 2001 Nov;2(11):815-26. doi: 10.1038/35099020. PMID: 11715048.

Barea, F., Tessaro, S. & Bonatto, D. (2008) In silico analyses of a new group of fungal and plant RecQ4-homologous proteins. *Computational biology and chemistry*. 32 (5), 349–358. doi:10.1016/J.COMPBIOLCHEM.2008.07.005.

Baretić, D., Jenkyn-Bedford, M., Aria, V., Cannone, G., Skehel, M. & Yeeles, J.T.P. (2020) Cryo-EM Structure of the Fork Protection Complex Bound to CMG at a Replication Fork. *Molecular Cell*. 78 (5), 926-940.e13. doi:10.1016/J.MOLCEL.2020.04.012.

Barnum, K.J. & O'Connell, M.J. (2014) Cell Cycle Regulation by Checkpoints. *Methods in molecular biology (Clifton, N.J.)*. 1170, 29. doi:10.1007/978-1-4939-0888-2_2.

Beach, D., Durkacz, B. & Nurse, P. (1982) Functionally homologous cell cycle control genes in budding and fission yeast. *Nature*. 300 (5894), 706–709. doi:10.1038/300706a0.

Bebenek, K. & Kunkel, T.A. (2004) Functions of DNA Polymerases. *Advances in Protein Chemistry*. 69, 137–165. doi:10.1016/S0065-3233(04)69005-X.

Belden, W.J., Loros, J.J. & Dunlap, J.C. (2007) Execution of the circadian negative feedback loop in Neurospora requires the ATP-dependent chromatin-remodeling enzyme CLOCKSWITCH. *Molecular cell*. 25 (4), 587–600. doi:10.1016/J.MOLCEL.2007.01.010.

Bell, S.D. & Botchan, M.R. (2013) The Minichromosome Maintenance Replicative Helicase. *Cold Spring Harbor Perspectives in Biology*. 5 (11), 12807–12808. doi:10.1101/CSHPERSPECT.A012807.

Bell, S.P. & Labib, K. (2016) Chromosome Duplication in Saccharomyces cerevisiae. *Genetics*. 203 (3), 1027. doi:10.1534/GENETICS.115.186452.

Bell, S.P., Mitchell, J., Leber, J., Kobayashi, R. & Stillman, B. (1995) The multidomain structure of Orc1 p reveals similarity to regulators of DNA replication and transcriptional silencing. *Cell*. 83 (4), 563–568. doi:10.1016/0092-8674(95)90096-9.

Bell, S.P. & Stillman, B. (1992) ATP-dependent recognition of eukaryotic origins of DNA replication by a multiprotein complex. *Nature 1992 357:6374*. 357 (6374), 128–134. doi:10.1038/357128a0.

Bermejo, R., Doksani, Y., Capra, T., Katou, Y.M., Tanaka, H., Shirahige, K. & Foiani, M. (2007) Top1- and Top2-mediated topological transitions at replication forks ensure fork progression and stability and prevent DNA damage checkpoint activation. *Genes & development*. 21 (15), 1921–1936. doi:10.1101/GAD.432107.

Bermudez, V.P., Farina, A., Raghavan, V., Tappin, I. & Hurwitz, J. (2011) Studies on human DNA polymerase epsilon and GINS complex and their role in DNA replication. *The Journal of biological chemistry*. 286 (33), 28963–28977. doi:10.1074/JBC.M111.256289.

Bernstein, K.A., Gangloff, S. & Rothstein, R. (2010) The RecQ DNA helicases in DNA Repair. *Annual review of genetics*. 44, 393. doi:10.1146/ANNUREV-GENET-102209-163602.

Bessler, J.B., Torres, J.Z. & Zakian, V.A. (2001) *The Pif1p subfamilyof helicases: region-specific DNA helicases?* http://tcb.trends.com0962.

de Bettignies, G. & Johnston, L.H. (2003) The mitotic exit network. *Current biology : CB*. 13 (8). doi:10.1016/S0960-9822(03)00230-6.

Bi, X., Yu, Q., Siler, J., Li, C. & Khan, A. (2015) Functions of Fun30 chromatin remodeler in regulating cellular resistance to genotoxic stress. *PloS one*. 10 (3), e0121341–e0121341. doi:10.1371/journal.pone.0121341.

Bielinsky, A.K., Blitzblau, H., Beall, E.L., Ezrokhi, M., Smith, H.S., Botchan, M.R. & Gerbi, S.A. (2001) Origin recognition complex binding to a metazoan replication origin. *Current biology : CB*. 11 (18), 1427–1431. doi:10.1016/S0960-9822(01)00444-4.

Bird, R.E., Louarn, J., Martuscelli, J. & Caro, L. (1972) Origin and sequence of chromosome replication in Escherichia coli. *Journal of Molecular Biology*. 70 (3), 549–566. doi:10.1016/0022-2836(72)90559-1.

Bleichert, F. (2019) Mechanisms of replication origin licensing: a structural perspective. *Current Opinion in Structural Biology*. 59, 195–204. doi:10.1016/J.SBI.2019.08.007.

Bleichert, F., Botchan, M.R. & Berger, J.M. (2015) Crystal structure of the eukaryotic origin recognition complex. *Nature 2015 519:7543*. 519 (7543), 321–326. doi:10.1038/nature14239.

Bleichert, F., Leitner, A., Aebersold, R., Botchan, M.R. & Berger, J.M. (2018) Conformational control and DNA-binding mechanism of the metazoan origin recognition complex. *Proceedings of the National Academy of Sciences of the United States of America*. 115 (26), E5906–E5915. doi:10.1073/PNAS.1806315115/-/DCSUPPLEMENTAL.

Bloom, J. & Cross, F.R. (2007) Multiple levels of cyclin specificity in cell-cycle control. *Nature Reviews Molecular Cell Biology 2007 8:2*. 8 (2), 149–160. doi:10.1038/nrm2105.

Blow, J.J. & Laskey, R.A. (1988) A role for the nuclear envelope in controlling DNA replication within the cell cycle. *Nature*. 332, 546. https://doi.org/10.1038/332546a0.

Blum, H., Beier, H. & Gross, H.J. (1987) Improved silver staining of plant proteins, RNA and DNA in polyacrylamide gels. *ELECTROPHORESIS*. 8 (2), 93–99. doi:10.1002/ELPS.1150080203.

Bochman, M.L., Bell, S.P. & Schwacha, A. (2008) Subunit organization of Mcm2-7 and the unequal role of active sites in ATP hydrolysis and viability. *Molecular and cellular biology*. 28 (19), 5865–5873. doi:10.1128/MCB.00161-08.

Bochman, M.L., Paeschke, K., Chan, A. & Zakian, V.A. (2014) Hrq1, a homolog of the human RecQ4 helicase, acts catalytically and structurally to promote genome integrity. *Cell reports*. 6 (2), 346. doi:10.1016/J.CELREP.2013.12.037.

Bochman, M.L. & Schwacha, A. (2007) Differences in the single-stranded DNA binding activities of MCM2-7 and MCM467: MCM2 and MCM5 define a slow ATP-dependent step. *The Journal of biological chemistry*. 282 (46), 33795–33804. doi:10.1074/JBC.M703824200.

Bochman, M.L. & Schwacha, A. (2009) The Mcm Complex: Unwinding the Mechanism of a Replicative Helicase. *Microbiology and Molecular Biology Reviews*. 73 (4), 652–683. doi:10.1128/MMBR.00019-09/ASSET/89AAF25F-3730-4671-95BC-0A8EF3FB39FD/ASSETS/GRAPHIC/ZMR0040922260013.JPEG.

Bochman, M.L. & Schwacha, A. (2008) The Mcm2-7 Complex Has In Vitro Helicase Activity. *Molecular Cell*. 31 (2), 287–293. doi:10.1016/J.MOLCEL.2008.05.020.

Bochman, M.L. & Schwacha, A. (2010) The Saccharomyces cerevisiae Mcm6/2 and Mcm5/3 ATPase active sites contribute to the function of the putative Mcm2-7 'gate.' *Nucleic Acids Research*. 38 (18), 6078. doi:10.1093/NAR/GKQ422.

Boehm, E.M., Gildenberg, M.S. & Washington, M.T. (2016) The many roles of PCNA in eukaryotic DNA replication. *The Enzymes*. 39, 231. doi:10.1016/BS.ENZ.2016.03.003.

Boginya, A., Detroja, R., Matityahu, A., Frenkel-Morgenstern, M. & Onn, I. (2019) The chromatin remodeler Chd1 regulates cohesin in budding yeast and humans. *Scientific Reports 2019 9:1*. 9 (1), 1–10. doi:10.1038/s41598-019-45263-3.

Boos, D. & Ferreira, P. (2019) Origin Firing Regulations to Control Genome Replication Timing. *Genes*. 10 (3). doi:10.3390/GENES10030199.

Boos, D., Sanchez-Pulido, L., Rappas, M., Pearl, L.H., Oliver, A.W., Ponting, C.P. & Diffley, J.F.X. (2011) Regulation of DNA replication through Sld3-Dpb11 interaction is conserved from yeast to humans. *Current biology : CB*. 21 (13), 1152–1157. doi:10.1016/J.CUB.2011.05.057.

Botchan, M. & Berger, J. (2010) DNA Replication: Making Two Forks from One Prereplication Complex. *Molecular cell*. 40 (6), 860. doi:10.1016/J.MOLCEL.2010.12.014.

Botchkarev, V. v. & Haber, J.E. (2018) Functions and regulation of the Polo-like kinase Cdc5 in the absence and presence of DNA damage. *Current genetics*. 64 (1), 87. doi:10.1007/S00294-017-0727-2.

Bowman, G.D. (2010) Mechanisms of ATP-Dependent Nucleosome Sliding. *Current opinion in structural biology*. 20 (1), 73. doi:10.1016/J.SBI.2009.12.002.

Bowman, G.D., O'Donnell, M. & Kuriyan, J. (2004) Structural analysis of a eukaryotic sliding DNA clamp-clamp loader complex. *Nature*. 429 (6993), 724–730. doi:10.1038/NATURE02585.

Brachmann, C.B., Sherman, J.M., Devine, S.E., Cameron, E.E., Pillus, L. & Boeke, J.D. (1995) The SIR2 gene family, conserved from bacteria to humans, functions in silencing, cell cycle progression, and chromosome stability. *Genes & development*. 9 (23), 2888–2902. doi:10.1101/GAD.9.23.2888.

Branzei, D. & Foiani, M. (2009) The checkpoint response to replication stress. *DNA repair*. 8 (9), 1038–1046. doi:10.1016/J.DNAREP.2009.04.014.

Branzei, D. & Foiani, M. (2006) The Rad53 signal transduction pathway: Replication fork stabilization, DNA repair, and adaptation. *Experimental Cell Research*. 312 (14), 2654–2659. doi:10.1016/J.YEXCR.2006.06.012.

Bravo, R., Fey, S.J., Bellatin, J., Larsen, P.M., Arevalo, J. & Celis, J.E. (1981) Identification of a nuclear and of a cytoplasmic polypeptide whose relative proportions are sensitive to changes in the rate of cell proliferation. *Experimental cell research*. 136 (2), 311–319. doi:10.1016/0014-4827(81)90009-4.

Bravo, R., Frank, R., Blundell, P.A. & Macdonald-Bravo, H. (1987) Cyclin/PCNA is the auxiliary protein of DNA polymerase-delta. *Nature*. 326 (6112), 515–517. doi:10.1038/326515A0.

Brewer, B.J. & Fangman, W.L. (1987) The localization of replication origins on ARS plasmids in S. cerevisiae. *Cell*. 51 (3), 463–471. doi:10.1016/0092-8674(87)90642-8.

Brill, S.J. & Stillman, B. (1989) Yeast replication factor-A functions in the unwinding of the SV40 origin of DNA replication. *Nature 1989 342:6245*. 342 (6245), 92–95. doi:10.1038/342092a0.

Broach, J.R., Li, Y.Y., Feldman, J., Jayaram, M., Abraham, J., Nasmyth, K.A. & Hicks, J.B. (1983) Localization and sequence analysis of yeast origins of DNA replication. *Cold Spring Harbor symposia on quantitative biology*. 47 Pt 2 (2), 1165–1173. doi:10.1101/SQB.1983.047.01.132.

de Bruin, R.A.M., McDonald, W.H., Kalashnikova, T.I., Yates, J. & Wittenberg, C. (2004) Cln3 activates G1-specific transcription via phosphorylation of the SBF bound repressor Whi5. *Cell*. 117 (7), 887–898. doi:10.1016/J.CELL.2004.05.025.

Buchberger, J.R., Onishi, M., Li, G., Seebacher, J., Rudner, A.D., Gygi, S.P. & Moazed, D. (2008) Sir3-Nucleosome Interactions in Spreading of Silent Chromatin in Saccharomyces cerevisiae . *Molecular and Cellular Biology*. 28 (22), 6903–6918. doi:10.1128/MCB.01210-08.

Buchman, A.R., Kimmerly, W.J., Rine, J. & Kornberg, R.D. (1988) Two DNA-binding factors recognize specific sequences at silencers, upstream activating sequences, autonomously replicating sequences, and telomeres in Saccharomyces cerevisiae. *Molecular and cellular biology*. 8 (1), 210–225. doi:10.1128/MCB.8.1.210-225.1988.

Budhavarapu, V.N., Chavez, M. & Tyler, J.K. (2013) How is epigenetic information maintained through DNA replication? *Epigenetics and Chromatin*. 6 (1), 1–7. doi:10.1186/1756-8935-6-32/FIGURES/1.

Buechel, E.R. & Pinkett, H.W. (2020) Transcription factors and ABC transporters: from pleiotropic drug resistance to cellular signaling in yeast. *FEBS Letters*. 594 (23), 3943–3964. doi:10.1002/1873-3468.13964.

Burgers, P.M.J. & Kunkel, T.A. (2017) Eukaryotic DNA Replication Fork. *Annual review of biochemistry*. 86, 417. doi:10.1146/ANNUREV-BIOCHEM-061516-044709.

Burgess, R.J., Han, J. & Zhang, Z. (2014) The Ddc1-Mec3-Rad17 sliding clamp regulates histone-histone chaperone interactions and DNA replication-coupled nucleosome assembly in budding yeast. *The Journal of biological chemistry*. 289 (15), 10518–10529. doi:10.1074/JBC.M114.552463.

Burnham, D.R., Kose, H.B., Hoyle, R.B. & Yardimci, H. (2019) The mechanism of DNA unwinding by the eukaryotic replicative helicase. *Nature Communications*. 10 (1). doi:10.1038/S41467-019-09896-2.

Byeon, B., Wang, W., Barski, A., Ranallo, R.T., Bao, K., Schones, D.E., Zhao, K., Wu, C. & Wu, W.-H. (2013) The ATP-dependent chromatin remodeling enzyme Fun30 represses transcription by sliding promoter-proximal nucleosomes. *The Journal of biological chemistry*. 288 (32), 23182–23193. doi:10.1074/jbc.M113.471979.

Bylund, G.O. & Burgers, P.M.J. (2005) Replication protein A-directed unloading of PCNA by the Ctf18 cohesion establishment complex. *Molecular and cellular biology*. 25 (13), 5445–5455. doi:10.1128/MCB.25.13.5445-5455.2005.

Cadoret, J.C., Meisch, F., Hassan-Zadeh, V., Luyten, I., Guillet, C., Duret, L., Quesneville, H. & Prioleau, M.N. (2008) Genome-wide studies highlight indirect links between human replication origins and gene regulation. *Proceedings of the National Academy of Sciences of the United States of America*. 105 (41), 15837–15842. doi:10.1073/PNAS.0805208105.

Cairns, B.R. (2007) Chromatin remodeling: insights and intrigue from single-molecule studies. *Nature structural & molecular biology*. 14 (11), 989. doi:10.1038/NSMB1333.

Cairns, J. (1966) Autoradiography of HeLa cell DNA. *Journal of Molecular Biology*. 15 (1), 372-IN32. doi:10.1016/S0022-2836(66)80233-4.

Callebaut, I., Courvalin, J.C. & Mornon, J.P. (1999) The BAH (bromo-adjacent homology) domain: a link between DNA methylation, replication and transcriptional regulation. *FEBS letters*. 446 (1), 189–193. doi:10.1016/S0014-5793(99)00132-5.

Calzada, A., Hodgson, B., Kanemaki, M., Bueno, A. & Labib, K. (2005) Molecular anatomy and regulation of a stable replisome at a paused eukaryotic DNA replication fork. *Genes & development*. 19 (16), 1905–1919. doi:10.1101/GAD.337205.

Caydasi, A.K., Ibrahim, B. & Pereira, G. (2010) Monitoring spindle orientation: Spindle position checkpoint in charge. *Cell Division 2010 5:1*. 5 (1), 1–15. doi:10.1186/1747-1028-5-28.

Cayrou, C., Ballester, B., Peiffer, I., Fenouil, R., Coulombe, P., Andrau, J.C., van Helden, J. & Méchali, M. (2015) The chromatin environment shapes DNA replication origin organization and defines origin classes. *Genome Research*. 25 (12), 1873. doi:10.1101/GR.192799.115.

Cayrou, C., Coulombe, P., Puy, A., Rialle, S., Kaplan, N., Segal, E. & Méchali, M. (2012) New insights into replication origin characteristics in metazoans. *Cell cycle (Georgetown, Tex.)*. 11 (4), 658–667. doi:10.4161/CC.11.4.19097.

Cayrou, C., Coulombe, P., Vigneron, A., Stanojcic, S., Ganier, O., Peiffer, I., Rivals, E., Puy, A., Laurent-Chabalier, S., Desprat, R. & Méchali, M. (2011) Genome-scale analysis of metazoan replication origins reveals their organization in specific but flexible sites defined by conserved features. *Genome Research*. 21 (9), 1438. doi:10.1101/GR.121830.111.

Cejka, P., Plank, J.L., Dombrowski, C.C. & Kowalczykowski, S.C. (2012) Decatenation of DNA by the S. cerevisiae Sgs1-Top3-Rmi1 and RPA complex: a mechanism for disentangling chromosomes. *Molecular cell*. 47 (6), 886–896. doi:10.1016/J.MOLCEL.2012.06.032.

Chen C.F., Pohl T.J., Pott S. & Zakian V.A. (2019) Two Pif1 Family DNA Helicases Cooperate in Centromere Replication and Segregation in Saccharomyces cerevisiae. *Genetics*. 211 (1), 105–119. doi:10.1534/GENETICS.118.301710.

Champoux, J.J. (2003) DNA Topoisomerases: Structure, Function, and Mechanism. *http://dx.doi.org/10.1146/annurev.biochem.70.1.369*. 70, 369–413. doi:10.1146/ANNUREV.BIOCHEM.70.1.369.

Chan, C.S.M. & Tye, B.-K. (1980) Autonomously replicating sequences in Saccharomyces cerevisiae (chromosomal replicating segments/baker's yeast/repetitive sequences/transformation). *Biochemistry*. 77 (11), 6329–6333.

Chang, Y.-T. & Loew, G.H. (1994) Reaction Mechanisms of Formaldehyde with Endocyclic Imino Groups of Nucleic Acid Bases. *J. Am. Chem. Soc*. 116, 3548–3555. https://pubs.acs.org/sharingguidelines.

Chapman, J.W. & Johnston, L.H. (1989) The yeast gene, DBF4, essential for entry into S phase is cell cycle regulated. *Experimental cell research*. 180 (2), 419–428. doi:10.1016/0014-4827(89)90068-2.

Chen, R. & Wold, M.S. (2014) Replication protein A: single-stranded DNA's first responder: dynamic DNA-interactions allow replication protein A to direct single-strand DNA intermediates into different pathways for synthesis or repair. *BioEssays : news and reviews in molecular, cellular and developmental biology*. 36 (12), 1156–1161. doi:10.1002/BIES.201400107.

Chen, X., Cui, D., Papusha, A., Zhang, X., Chu, C.-D., Tang, J., Chen, K., Pan, X. & Ira, G. (2012) The Fun30 nucleosome remodeller promotes resection of DNA double-strand break ends. *Nature*. 489 (7417), 576–580. doi:10.1038/nature11355.

Chen, X., Niu, H., Yu, Y., Wang, J., Zhu, S., Zhou, J., Papusha, A., Cui, D., Pan, X., Kwon, Y., Sung, P. & Ira, G. (2016) Enrichment of Cdk1-cyclins at DNA double-strand breaks stimulates Fun30 phosphorylation and DNA end resection. *Nucleic acids research*. 44 (6), 2742–2753. doi:10.1093/nar/gkv1544.

Cheng, L., Hunke, L. & Hardy, C.F.J. (1998) Cell Cycle Regulation of the Saccharomyces cerevisiae Polo-Like Kinase Cdc5p . *Molecular and Cellular Biology*. 18 (12), 7360–7370. doi:10.1128/MCB.18.12.7360/ASSET/9B2E1B01-2CBF-448A-8165-ED4A02BCAB6C/ASSETS/GRAPHIC/MB1280958010.JPEG.

Chilkova, O., Jonsson, B.H. & Johansson, E. (2003) The Quaternary Structure of DNA Polymerase ε from Saccharomyces cerevisiae. *Journal of Biological Chemistry*. 278 (16), 14082–14086. doi:10.1074/JBC.M211818200.

Chilkova, O., Stenlund, P., Isoz, I., Stith, C.M., Grabowski, P., Lundström, E.B., Burgers, P.M. & Johansson, E. (2007) The eukaryotic leading and lagging strand DNA polymerases are loaded onto primer-ends via separate mechanisms but have comparable processivity in the presence of PCNA. *Nucleic acids research*. 35 (19), 6588–6597. doi:10.1093/NAR/GKM741.

Choi, H., Glatter, T., Gstaiger, M. & Nesvizhskii, A.I. (2012) SAINT-MS1: protein-protein interaction scoring using label-free intensity data in affinity purification – mass spectrometry experiments. *Journal of proteome research*. 11 (4), 2619. doi:10.1021/PR201185R.

Chong, J.P., Hayashi, M.K., Simon, M.N., Xu, R.M. & Stillman, B. (2000) A double-hexamer archaeal minichromosome maintenance protein is an ATP-dependent DNA helicase. *Proceedings of the National Academy of Sciences of the United States of America*. 97 (4), 1530–1535. doi:10.1073/pnas.030539597.

Chong, J.P., Mahbubani, H.M., Khoo, C.Y. & Blow, J.J. (1995) Purification of an MCM-containing complex as a component of the DNA replication licensing system. *Nature*. 375 (6530), 418–421. doi:10.1038/375418a0.

Chu, W.K. & Hickson, I.D. (2009) RecQ helicases: multifunctional genome caretakers. *Nature Reviews Cancer 2009 9:9*. 9 (9), 644–654. doi:10.1038/nrc2682.

Rivin C.J. & Fangman W.L. (1980) Replication fork rate and origin activation during the S phase of Saccharomyces cerevisiae. *The Journal of cell biology*. 85 (1), 108–115. doi:10.1083/JCB.85.1.108.

Clapier, C.R., Iwasa, J., Cairns, B.R. & Peterson, C.L. (2017) Mechanisms of action and regulation of ATP-dependent chromatin-remodelling complexes. *Nature Reviews Molecular Cell Biology 2017 18:7*. 18 (7), 407–422. doi:10.1038/nrm.2017.26.

Clément, C. & Almouzni, G. (2015) MCM2 binding to histones H3-H4 and ASF1 supports a tetramer-to-dimer model for histone inheritance at the replication fork. *Nature structural & molecular biology*. 22 (8), 587–589. doi:10.1038/NSMB.3067.

Cliften, P., Sudarsanam, P., Desikan, A., Fulton, L., Fulton, B., Majors, J., Waterston, R., Cohen, B.A. & Johnston, M. (2003) Finding functional features in Saccharomyces genomes by phylogenetic footprinting. *Science (New York, N.Y.)*. 301 (5629), 71–76. doi:10.1126/SCIENCE.1084337.

Cobb, J.A., Bjergbaek, L., Shimada, K., Frei, C. & Gasser, S.M. (2003) DNA polymerase stabilization at stalled replication forks requires Mec1 and the RecQ helicase Sgs1. *The EMBO journal*. 22 (16), 4325–4336. doi:10.1093/EMBOJ/CDG391.

Cocker, J.H., Piatti, S., Santocanale, C., Nasmyth, K. & Diffley, J.F.X. (1996) An essential role for the Cdc6 protein in forming the pre-replicative complexes of budding yeast. *Nature*. 379 (6561), 180–182. doi:10.1038/379180A0.

Collas, P. (2010) The current state of chromatin immunoprecipitation. *Molecular Biotechnology*. 45 (1), 87–100. doi:10.1007/S12033-009-9239-8/FIGURES/4.

Comoglio, F., Schlumpf, T., Schmid, V., Rohs, R., Beisel, C. & Paro, R. (2015) High-resolution profiling of Drosophila replication start sites reveals a DNA shape and chromatin signature of metazoan origins. *Cell reports*. 11 (5), 821–834. doi:10.1016/J.CELREP.2015.03.070.

Connelly, J.J., Yuan, P., Hsu, H.-C., Li, Z., Xu, R.-M. & Sternglanz, R. (2006) Structure and function of the Saccharomyces cerevisiae Sir3 BAH domain. *Molecular and cellular biology*. 26 (8), 3256–3265. doi:10.1128/MCB.26.8.3256-3265.2006.

Cornacchia, D., Dileep, V., Quivy, J.P., Foti, R., Tili, F., Santarella-Mellwig, R., Antony, C., Almouzni, G., Gilbert, D.M. & Buonomo, S.B.C. (2012) Mouse Rif1 is a key regulator of the replication-timing programme in mammalian cells. *The EMBO journal*. 31 (18), 3678–3690. doi:10.1038/EMBOJ.2012.214.

Cortez, D. (2015) Preventing replication fork collapse to maintain genome integrity. *DNA repair*. 32, 149–157. doi:10.1016/J.DNAREP.2015.04.026.

Cosgrove, A.J., Nieduszynski, C.A. & Donaldson, A.D. (2002) Ku complex controls the replication time of DNA in telomere regions. *Genes & Development*. 16 (19), 2485–2490. doi:10.1101/GAD.231602.

Costa, A., Ilves, I., Tamberg, N., Petojevic, T., Nogales, E., Botchan, M.R. & Berger, J.M. (2011) The structural basis for MCM2-7 helicase activation by GINS and Cdc45. *Nature structural & molecular biology*. 18 (4), 471–479. doi:10.1038/NSMB.2004.

Costa, A. & Onesti, S. (2009) Structural biology of MCM helicases. *http://dx.doi.org/10.1080/10409230903186012*. 44 (5), 326–342. doi:10.1080/10409230903186012.

Costa, A., Pape, T., van Heel, M., Brick, P., Patwardhan, A. & Onesti, S. (2006) Structural studies of the archaeal MCM complex in different functional states. *Journal of structural biology*. 156 (1), 210–219. doi:10.1016/j.jsb.2006.04.001.

Costa, A., Renault, L., Swuec, P., Petojevic, T., Pesavento, J.J., Ilves, I., MacLellan-Gibson, K., Fleck, R.A., Botchan, M.R. & Berger, J.M. (2014) DNA binding polarity, dimerization, and ATPase ring remodeling in the CMG helicase of the eukaryotic replisome. *eLife*. 3 (August2014), 1–17. doi:10.7554/ELIFE.03273.

Costanzo, M., Nishikawa, J.L., Tang, X., Millman, J.S., Schub, O., Breitkreuz, K., Dewar, D., Rupes, I., Andrews, B. & Tyers, M. (2004) CDK activity antagonizes Whi5, an inhibitor of G1/S transcription in yeast. *Cell*. 117 (7), 899–913. doi:10.1016/J.CELL.2004.05.024.

Costelloe, T., Louge, R., Tomimatsu, N., Mukherjee, B., Martini, E., Khadaroo, B., Dubois, K., Wiegant, W.W., Thierry, A., Burma, S., van Attikum, H. & Llorente, B. (2012) The yeast Fun30 and human SMARCAD1 chromatin remodellers promote DNA end resection. *Nature*. 489 (7417), 581–584. doi:10.1038/nature11353.

Coster, G. & Diffley, J.F.X. (2017) Bidirectional eukaryotic DNA replication is established by quasi-symmetrical helicase loading. *Science*. 357 (6348), 314–318. doi:10.1126/SCIENCE.AAN0063/SUPPL_FILE/AAN0063_COSTER_SM.PDF.

Coster, G., Frigola, J., Beuron, F., Morris, E.P. & Diffley, J.F.X. (2014) Origin licensing requires ATP binding and hydrolysis by the MCM replicative helicase. *Molecular cell*. 55 (5), 666–677. doi:10.1016/j.molcel.2014.06.034.

Cotta-Ramusino, C., Fachinetti, D., Lucca, C., Doksani, Y., Lopes, M., Sogo, J. & Foiani, M. (2005) Exo1 processes stalled replication forks and counteracts fork reversal in checkpoint-defective cells. *Molecular cell*. 17 (1), 153–159. doi:10.1016/J.MOLCEL.2004.11.032.

Coverley, D., Kenny, M.K., Munn, M., Rupp, W.D., Lane, D.P. & Wood, R.D. (1991) Requirement for the replication protein SSB in human DNA excision repair. *Nature*. 349 (6309), 538–541. doi:10.1038/349538A0.

Cox, J., Hein, M.Y., Luber, C.A., Paron, I., Nagaraj, N. & Mann, M. (2014) Accurate Proteome-wide Label-free Quantification by Delayed Normalization and Maximal Peptide Ratio Extraction, Termed MaxLFQ. *Molecular & Cellular Proteomics : MCP*. 13 (9), 2513. doi:10.1074/MCP.M113.031591.

Cox, J. & Mann, M. (2008) MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nature Biotechnology 2008 26:12*. 26 (12), 1367–1372. doi:10.1038/nbt.1511.

Cox, J.S., Chapman, R.E. & Waltert, P. (1997) The Unfolded Protein Response Coordinates the Production of Endoplasmic Reticulum Protein and Endoplasmic Reticulum Membrane. *Molecular Biology of the Cell*. 8, 1805–1814.

Crabbé, L., Thomas, A., Pantesco, V., de Vos, J., Pasero, P. & Lengronne, A. (2010) Analysis of replication profiles reveals key role of RFC-Ctf18 in yeast replication stress response. *Nature structural & molecular biology*. 17 (11), 1391–1397. doi:10.1038/NSMB.1932.

Crampton, A., Chang, F., Pappas, D.L., Frisch, R.L. & Weinreich, M. (2008) An ARS Element Inhibits DNA Replication through a SIR2-Dependent Mechanism. *Molecular Cell*. 30 (2), 156–166. doi:10.1016/J.MOLCEL.2008.02.019.

Kong D. & DePamphilis M.L. (2001) Site-specific DNA binding of the Schizosaccharomyces pombe origin recognition complex is determined by the Orc4 subunit. *Molecular and cellular biology*. 21 (23), 8095–8103. doi:10.1128/MCB.21.23.8095-8103.2001.

Dahmann, C., Diffley, J.F.X. & Nasmyth, K.A. (1995) S-phase-promoting cyclin-dependent kinases prevent re-replication by inhibiting the transition of replication origins to a pre-replicative state. *Current Biology*. 5 (11), 1257–1269. doi:https://doi.org/10.1016/S0960-9822(95)00252-1.

Daigaku, Y., Keszthelyi, A., Müller, C.A., Miyabe, I., Brooks, T., Retkute, R., Hubank, M., Nieduszynski, C.A. & Carr, A.M. (2015) A global profile of replicative polymerase usage. *Nature structural & molecular biology*. 22 (3), 192–198. doi:10.1038/NSMB.2962.

Dalgaard, J.Z. & Klar, A.J.S. (2000) swi1 and swi3 perform imprinting, pausing, and termination of DNA replication in S. pombe. *Cell*. 102 (6), 745–751. doi:10.1016/S0092-8674(00)00063-5.

D'Aquino, K.E., Monje-Casas, F., Paulson, J., Reiser, V., Charles, G.M., Lai, L., Shokat, K.M. & Amon, A. (2005) The protein kinase Kin4 inhibits exit from mitosis in response to spindle position defects. *Molecular cell*. 19 (2), 223–234. doi:10.1016/J.MOLCEL.2005.06.005.

Das, P.M., Ramachandran, K., VanWert, J. & Singal, R. (2004) Chromatin immunoprecipitation assay. *BioTechniques*. 37 (6), 961–969. doi:10.2144/04376RV01/ASSET/IMAGES/LARGE/FIGURE2.JPEG.

Das-Bradoo, S., Ricke, R.M. & Bielinsky, A.-K. (2006) Interaction between PCNA and Diubiquitinated Mcm10 Is Essential for Cell Growth in Budding Yeast. *Molecular and Cellular Biology*. 26 (13), 4806–4817. doi:10.1128/MCB.02062-05/ASSET/2E66945F-6447-4ABF-897F-77E30E7A0F9D/ASSETS/GRAPHIC/ZMB0130660500010.JPEG.

Daugeron, M.C., Lenstra, T.L., Frizzarin, M., el Yacoubi, B., Liu, X., Baudin-Baillieu, A., Lijnzaad, P., Decourty, L., Saveanu, C., Jacquier, A., Holstege, F.C.P., de Crécy-Lagard, V., van Tilbeurgh, H. & Libri, D. (2011) Gcn4 misregulation reveals a direct role for the evolutionary conserved EKC/KEOPS in the t6A modification of tRNAs. *Nucleic Acids Research*. 39 (14), 6148. doi:10.1093/NAR/GKR178.

Daugeron, M.C. & Linder, P. (1998) Dbp7p, a putative ATP-dependent RNA helicase from Saccharomyces cerevisiae, is required for 60S ribosomal subunit assembly. *RNA*. 4 (5), 566. doi:10.1017/S1355838298980190.

Davé, A., Cooley, C., Garg, M. & Bianchi, A. (2014) Protein phosphatase 1 recruitment by Rif1 regulates DNA replication origin firing by counteracting DDK activity. *Cell reports*. 7 (1), 53–61. doi:10.1016/J.CELREP.2014.02.019.

Davey, M.J., Indiani, C. & O'Donnell, M. (2003) Reconstitution of the Mcm2-7p Heterohexamer, Subunit Arrangement, and ATP Site Architecture *. *Journal of Biological Chemistry*. 278 (7), 4491–4499. doi:10.1074/JBC.M210511200.

Dawson, I.A., Roth, S., Akam, M. & Artavanis-Tsakonas, S. (1993) Mutations of the fizzy locus cause metaphase arrest in Drosophila melanogaster embryos. *Development*. 117 (1), 359–376. doi:10.1242/DEV.117.1.359.

Deegan, T.D., Baxter, J., Bazan, M.A.O., Yeeles, J.T.P. & Labib, K.P.M. (2019) Pif1-Family Helicases Support Fork Convergence during DNA Replication Termination in Eukaryotes. *Molecular cell*. 74 (2), 231-244.e9. doi:10.1016/j.molcel.2019.01.040.

Deegan, T.D., Yeeles, J.T. & Diffley, J.F. (2016) *Phosphopeptide binding by Sld3 links Dbf4-dependent kinase to MCM replicative helicase activation*. 35 (9), 961–973. https://onlinelibrary.wiley.com/doi/full/10.15252/embj.201593552.

Deshpande, I., Seeber, A., Shimada, K., Keusch, J.J., Gut, H. & Gasser, S.M. (2017) Structural Basis of Mec1-Ddc2-RPA Assembly and Activation on Single-Stranded DNA at Sites of Damage. *Molecular cell*. 68 (2), 431-445.e5. doi:10.1016/J.MOLCEL.2017.09.019.

van Deursen, F., Sengupta, S., Piccoli, G. de, Sanchez-Diaz, A. & Labib, K. (2012) Mcm10 associates with the loaded DNA helicase at replication origins and defines a novel step in its activation. *The EMBO journal*. 31 (9), 2195–2206. doi:10.1038/emboj.2012.69.

Devault, A., Vallen, E.A., Yuan, T., Green, S., Bensimon, A. & Schwob, E. (2002) Identification of Tah11/Sid2 as the ortholog of the replication licensing factor Cdt1 in Saccharomyces cerevisiae. *Current biology : CB*. 12 (8), 689–694. doi:10.1016/S0960-9822(02)00768-6.

Dewar, J.M. & Walter, J.C. (2017) Mechanisms of DNA replication termination. *Nature reviews. Molecular cell biology*. 18 (8), 507. doi:10.1038/NRM.2017.42.

Deweese, J.E. & Osheroff, N. (2009) The DNA cleavage reaction of topoisomerase II: wolf in sheep's clothing. *Nucleic Acids Research*. 37 (3), 738. doi:10.1093/NAR/GKN937.

Dhar, S.K. & Dutta, A. (2000) Identification and Characterization of the Human ORC6 Homolog. *Journal of Biological Chemistry*. 275 (45), 34983–34988. doi:10.1074/JBC.M006069200.

Diede, S.J. & Gottschling, D.E. (1999) Telomerase-Mediated Telomere Addition In Vivo Requires DNA Primase and DNA Polymerases α and δ. *Cell*. 99 (7), 723–733. doi:10.1016/S0092-8674(00)81670-0.

Diffley, J.F.X. (1996) Once and only once upon a time: specifying and regulating origins of DNA replication in eukaryotic cells. *Genes & development*. 10 (22), 2819–2830. doi:10.1101/GAD.10.22.2819.

Diffley, J.F.X., Cocker, J.H., Dowell, S.J. & Rowley, A. (1994) Two steps in the assembly of complexes at yeast replication origins in vivo. *Cell*. 78 (2), 303–316. doi:10.1016/0092-8674(94)90299-2.

Diffley, J.F.X. & Stillman, B. (1988) Purification of a yeast protein that binds to origins of DNA replication and a transcriptional silencer. *Proceedings of the National Academy of Sciences*. 85 (7), 2120–2124. doi:10.1073/PNAS.85.7.2120.

DiNardo, S., Voelkel, K. & Sternglanz, R. (1984) DNA topoisomerase II mutant of Saccharomyces cerevisiae: topoisomerase II is required for segregation of daughter molecules at the termination of DNA replication. *Proceedings of the National Academy of Sciences of the United States of America*. 81 (9), 2616–2620. doi:10.1073/PNAS.81.9.2616.

Doherty, A.J. & Jackson, S.P. (2001) DNA repair: How Ku makes ends meet. *Current Biology*. 11 (22), R920–R924. doi:10.1016/S0960-9822(01)00555-3.

Dornreiter, I., Erdile, L.F., Gilbert, L.U., von Winkler, D., Kelly, T.J. & Fanning, E. (1992) Interaction of DNA polymerase alpha-primase with cellular replication protein A and SV40 T antigen. *The EMBO Journal*. 11 (2), 769–776. doi:10.1002/J.1460-2075.1992.TB05110.X.

Douglas, M.E., Ali, F.A., Costa, A. & Diffley, J.F.X. (2018) The mechanism of eukaryotic CMG helicase activation. *Nature*. 555, 265. https://doi.org/10.1038/nature25787.

Du, W., Stauffer, M.E. & Eichman, B.F. (2012) Structural Biology of Replication Initiation Factor Mcm10. *Sub-cellular biochemistry*. 62, 197. doi:10.1007/978-94-007-4572-8_11.

Dua, R., Levy, D.L. & Campbell, J.L. (1999) Analysis of the Essential Functions of the C-terminal Protein/Protein Interaction Domain of Saccharomyces cerevisiae pol ε and Its Unexpected Ability to Support Growth in the Absence of the DNA Polymerase Domain *. *Journal of Biological Chemistry*. 274 (32), 22283–22288. doi:10.1074/JBC.274.32.22283.

Dua, R., Levy, D.L., Li, C.M., Snow, P.M. & Campbell, J.L. (2002) In vivo reconstitution of Saccharomyces cerevisiae DNA polymerase epsilon in insect cells. Purification and characterization. *The Journal of biological chemistry*. 277 (10), 7889–7896. doi:10.1074/JBC.M108546200.

Dudley, A.M., Rougeulle, C. & Winston, F. (1999) The Spt components of SAGA facilitate TBP binding to a promoter at a post-activator-binding step in vivo. *Genes & Development*. 13 (22), 2940. doi:10.1101/GAD.13.22.2940.

Duggin, I.G., Wake, R.G., Bell, S.D. & Hill, T.M. (2008) The replication fork trap and termination of chromosome replication. *Molecular microbiology*. 70 (6), 1323–1333. doi:10.1111/J.1365-2958.2008.06500.X.

Dumas, L.B., Lussky, J.P., McFarland, E.J. & Shampay, J. (1982) New temperature-sensitive mutants of Saccharomyces cerevisiae affecting DNA replication. *Molecular & general genetics : MGG*. 187 (1), 42–46. doi:10.1007/BF00384381.

Duncker, B.P., Chesnokov, I.N. & McConkey, B.J. (2009) The origin recognition complex protein family. *Genome Biology*. 10 (3), 214. doi:10.1186/GB-2009-10-3-214.

Durand-Dubief, M., Will, W.R., Petrini, E., Theodorou, D., Harris, R.R., Crawford, M.R., Paszkiewicz, K., Krueger, F., Correra, R.M., Vetter, A.T., Miller, J.R., Kent, N.A. & Varga-Weisz, P. (2012) SWI/SNF-Like Chromatin Remodeling Factor Fun30 Supports Point Centromere Function in S. cerevisiae. *PLOS Genetics*. 8 (9), e1002974. doi:10.1371/JOURNAL.PGEN.1002974.

Dutnall, R.N. & Pillus, L. (2001) Deciphering NAD-Dependent Deacetylases. *Cell*. 105 (2), 161–164. doi:10.1016/S0092-8674(01)00305-1.

Dutta, A. & Bell, S.P. (2003) INITIATION OF DNA REPLICATION IN EUKARYOTIC CELLS. *http://dx.doi.org/10.1146/annurev.cellbio.13.1.293*. 13, 293–332. doi:10.1146/ANNUREV.CELLBIO.13.1.293.

Sanders E., Nguyen P.A., Rogers C.M., & Bochman M.L. (2020) Comprehensive Synthetic Genetic Array Analysis of Alleles That Interact with Mutation of the Saccharomyces cerevisiae RecQ Helicases Hrq1 and Sgs1. *G3 (Bethesda, Md.)*. 10 (12), 4359–4368. doi:10.1534/G3.120.401709.

Eapen, V. v., Sugawara, N., Tsabar, M., Wu, W.-H. & Haber, J.E. (2012) The Saccharomyces cerevisiae chromatin remodeler Fun30 regulates DNA end resection and checkpoint deactivation. *Molecular and cellular biology*. 32 (22), 4727–4740. doi:10.1128/MCB.00566-12.

Ehrentraut, S., Hassler, M., Oppikofer, M., Kueng, S., Weber, J.M., Mueller, J.W., Gasser, S.M., Ladurner, A.G. & Ehrenhofer-Murray, A.E. (2011) Structural basis for the role of the Sir3 AAA+ domain in silencing: interaction with Sir4 and unmethylated histone H3K79. *Genes & Development*. 25 (17), 1835–1846. doi:10.1101/GAD.17175111.

Eickhoff, P., Kose, H.B., Martino, F., Petojevic, T., Abid Ali, F., Locke, J., Tamberg, N., Nans, A., Berger, J.M., Botchan, M.R., Yardimci, H. & Costa, A. (2019) Molecular Basis for ATP-Hydrolysis-Driven DNA Translocation by the CMG Helicase of the Eukaryotic Replisome. *Cell Reports*. 28 (10), 2673. doi:10.1016/J.CELREP.2019.07.104.

Ekundayo, B. & Bleichert, F. (2019) Origins of DNA replication. *PLOS Genetics*. 15 (9), e1008320. doi:10.1371/JOURNAL.PGEN.1008320.

Ellison, V. & Stillman, B. (2003) Biochemical characterization of DNA damage checkpoint complexes: clamp loader and clamp complexes with specificity for 5' recessed DNA. *PLoS biology*. 1 (2). doi:10.1371/JOURNAL.PBIO.0000033.

Enemark, E.J. & Joshua-Tor, L. (2006) Mechanism of DNA translocation in a replicative hexameric helicase. *Nature 2006 442:7100*. 442 (7100), 270–275. doi:10.1038/nature04943.

Enomoto, S. & Berman, J. (1998) Chromatin assembly factor I contributes to the maintenance, but not the re-establishment, of silencing at the yeast silent mating loci. *Genes & Development*. 12 (2), 219. doi:10.1101/GAD.12.2.219.

Enomoto, S., McCune-Zierath, P.D., Gerami-Nejad, M., Sanders, M.A. & Berman, J. (1997) RLF2, a subunit of yeast chromatin assembly factor-I, is required for telomeric chromatin function in vivo. *Genes & development*. 11 (3), 358–370. doi:10.1101/GAD.11.3.358.

Eriksson, P.R. & Clark, D.J. (2021) The yeast ISW1b ATP-dependent chromatin remodeler is critical for nucleosome spacing and dinucleosome resolution. *Scientific Reports 2021 11:1*. 11 (1), 1–14. doi:10.1038/s41598-021-82842-9.

Evans, S.K., Burgess, K.E.V. & Gray, J. v. (2014) Recovery from Rapamycin: Drug-Insensitive Activity of Yeast Target of Rapamycin Complex 1 (TORC1) Supports Residual Proliferation that Dilutes Rapamycin Among Progeny Cells. *Journal of Biological Chemistry*. 289 (38), 26554–26565. doi:10.1074/JBC.M114.589754.

Evans, T., Rosenthal, E.T., Youngblom, J., Distel, D. & Hunt, T. (1983) Cyclin: A protein specified by maternal mRNA in sea urchin eggs that is destroyed at each cleavage division. *Cell*. 33 (2), 389–396. doi:https://doi.org/10.1016/0092-8674(83)90420-8.

Evrin, C., Clarke, P., Zech, J., Lurz, R., Sun, J., Uhle, S., Li, H., Stillman, B. & Speck, C. (2009) A double-hexameric MCM2-7 complex is loaded onto origin DNA during licensing of eukaryotic DNA replication. *Proceedings of the National Academy of Sciences of the United States of America*. 106 (48), 20240–20245. doi:10.1073/pnas.0911500106.

Fachinetti, D., Bermejo, R., Cocito, A., Minardi, S., Katou, Y., Kanoh, Y., Shirahige, K., Azvolinsky, A., Zakian, V.A. & Foiani, M. (2010) Replication termination at eukaryotic chromosomes is mediated by Top2 and occurs at genomic loci containing pausing elements. *Molecular cell*. 39 (4), 595–605. doi:10.1016/J.MOLCEL.2010.07.024.

Fang, D., Cao, Q. & Lou, H. (2016) Sld3-MCM Interaction Facilitated by Dbf4-Dependent Kinase Defines an Essential Step in Eukaryotic DNA Replication Initiation. *Frontiers in microbiology*. 7, 885. doi:10.3389/fmicb.2016.00885.

Fang, D., Lengronne, A., Shi, D., Forey, R., Skrzypczak, M., Ginalski, K., Yan, C., Wang, X., Cao, Q., Pasero, P. & Lou, H. (2017) Dbf4 recruitment by forkhead transcription factors defines an upstream rate-limiting step in determining origin firing timing. *Genes and Development*. 31 (23–24), 2405–2415. doi:10.1101/GAD.306571.117/-/DC1.

Fang, G., Yu, H. & Kirschner, M.W. (1998) Direct Binding of CDC20 Protein Family Members Activates the Anaphase-Promoting Complex in Mitosis and G1. *Molecular Cell*. 2 (2), 163–171. doi:10.1016/S1097-2765(00)80126-4.

Fangman, W.L. & Brewer, B.J. (1992) A question of time: Replication origins of eukaryotic chromosomes. *Cell*. 71 (3), 363–366. doi:10.1016/0092-8674(92)90505-7.

Fanning, E., Klimovich, V. & Nager, A.R. (2006) A dynamic model for replication protein A (RPA) function in DNA processing pathways. *Nucleic acids research*. 34 (15), 4126–4137. doi:10.1093/NAR/GKL550.

Fatica, A., Cronshaw, A.D., Dlaki, M. & Tollervey, D. (2002) Ssf1p prevents premature processing of an early pre-60S ribosomal particle. *Molecular cell*. 9 (2), 341–351. doi:10.1016/S1097-2765(02)00458-6.

Feng, X., Noguchi, Y., Barbon, M., Stillman, B., Speck, C. & Li, H. (2021) The structure of ORC-Cdc6 on an origin DNA reveals the mechanism of ORC activation by the replication initiator Cdc6. *Nature communications*. 12 (1). doi:10.1038/S41467-021-24199-1.

Ferguson, B.M. & Fangman, W.L. (1992) A position effect on the time of replication origin activation in yeast. *Cell*. 68 (2), 333–339. doi:10.1016/0092-8674(92)90474-Q.

Fernandez-Cid, A., Riera, A., Tognetti, S., Herrera, M.C., Samel, S., Evrin, C., Winkler, C., Gardenal, E., Uhle, S. & Speck, C. (2013) An ORC/Cdc6/MCM2-7 complex is formed in a multistep reaction to serve as a platform for MCM double-hexamer assembly. *Molecular cell*. 50 (4), 577–588. doi:10.1016/j.molcel.2013.03.026.

Finehout, E.J. & Lee, K.H. (2004) An introduction to mass spectrometry applications in biological research. *Biochemistry and Molecular Biology Education*. 32 (2), 93–100. doi:10.1002/BMB.2004.494032020331.

Flaus, A., Martin, D.M.A., Barton, G.J. & Owen-Hughes, T. (2006) Identification of multiple distinct Snf2 subfamilies with conserved structural motifs. *Nucleic acids research*. 34 (10), 2887–2905. doi:10.1093/NAR/GKL295.

Flemming, W. (1882) *Zellsubstanz, Kern und Zelltheilung*.

Foresti, O., Rodriguez-Vaello, V., Funaya, C. & Carvalho, P. (2014) Quality control of inner nuclear membrane proteins by the Asi complex. *Science (New York, N.Y.)*. 346 (6210), 751–755. doi:10.1126/SCIENCE.1255638.

Forsburg, S.L. (2004) Eukaryotic MCM Proteins: Beyond Replication Initiation. *Microbiology and Molecular Biology Reviews*. 68 (1), 109–131. doi:10.1128/MMBR.68.1.109-131.2004.

Fortune, J.M., Pavlov, Y.I., Welch, C.M., Johansson, E., Burgers, P.M.J. & Kunkel, T.A. (2005) Saccharomyces cerevisiae DNA polymerase delta: high fidelity for base substitutions but lower fidelity for single- and multi-base deletions. *The Journal of biological chemistry*. 280 (33), 29980–29987. doi:10.1074/JBC.M505236200.

Foss, E.J. (2001) Tof1p regulates DNA damage responses during S phase in Saccharomyces cerevisiae. *Genetics*. 157 (2), 567–577. doi:10.1093/GENETICS/157.2.567.

Foss, E.J., Sripathy, S., Gatbonton-Schwager, T., Kwak, H., Thiesen, A.H., Lao, U. & Bedalov, A. (2021) Chromosomal Mcm2-7 distribution and the genome replication program in species from yeast to humans. *PLOS Genetics*. 17 (9), e1009714. doi:10.1371/JOURNAL.PGEN.1009714.

Foss, M., McNally, F.J., Laurenson, P. & Rine, J. (1993) Origin recognition complex (ORC) in transcriptional silencing and DNA replication in S. cerevisiae. *Science (New York, N.Y.)*. 262 (5141), 1838–1844. doi:10.1126/SCIENCE.8266071.

Foti, R., Gnan, S., Cornacchia, D., Dileep, V., Bulut-Karslioglu, A., Diehl, S., Buness, A., Klein, F.A., Huber, W., Johnstone, E., Loos, R., Bertone, P., Gilbert, D.M., Manke, T., Jenuwein, T. & Buonomo, S.C.B. (2016) Nuclear Architecture Organized by Rif1 Underpins the Replication-Timing Program. *Molecular cell*. 61 (2), 260–273. doi:10.1016/J.MOLCEL.2015.12.001.

Francis, L.I., Randell, J.C.W., Takara, T.J., Uchima, L. & Bell, S.P. (2009) Incorporation into the prereplicative complex activates the Mcm2–7 helicase for Cdc7–Dbf4 phosphorylation. *Genes & Development*. 23 (5), 643. doi:10.1101/GAD.1759609.

Franco, A.A. & Kaufman, P.D. (2004) Histone deposition proteins: links between the DNA replication machinery and epigenetic gene silencing. *Cold Spring Harbor symposia on quantitative biology*. 69, 201–208. doi:10.1101/SQB.2004.69.201.

Franco, A.A., Lam, W.M., Burgers, P.M. & Kaufman, P.D. (2005) Histone deposition protein Asf1 maintains DNA replisome integrity and interacts with replication factor C. *Genes & development*. 19 (11), 1365–1375. doi:10.1101/GAD.1305005.

Freeman, K. & Livi, G.P. (1996) Missense mutations at the FKBP12-rapamycin-binding site of TOR1. *Gene*. 172 (1), 143–147. doi:10.1016/0378-1119(96)00168-0.

Frei, C. & Gasser, S.M. (2000) The yeast Sgs1p helicase acts upstream of Rad53p in the DNA replication checkpoint and colocalizes with Rad53p in S-phase-specific foci. *Genes & Development*. 14 (1), 81. /pmc/articles/PMC316339/.

Friden, P. & Schimmel, P. (1988) LEU3 of Saccharomyces cerevisiae activates multiple genes for branched-chain amino acid biosynthesis by binding to a common decanucleotide core sequence. *Molecular and Cellular Biology*. 8 (7), 2690. doi:10.1128/MCB.8.7.2690.

Frigola, J., He, J., Kinkelin, K., Pye, V.E., Renault, L., Douglas, M.E., Remus, D., Cherepanov, P., Costa, A. & Diffley, J.F.X. (2017) Cdt1 stabilizes an open MCM ring for helicase loading. *Nature Communications 2017 8:1*. 8 (1), 1–10. doi:10.1038/ncomms15720.

Frigola, J., Remus, D., Mehanna, A. & Diffley, J.F.X. (2013) ATPase-dependent quality control of DNA replication origin licensing. *Nature*. 495 (7441), 339–343. doi:10.1038/NATURE11920.

Fu, Y. v., Yardimci, H., Long, D.T., Ho, V., Guainazzi, A., Bermudez, V.P., Hurwitz, J., van Oijen, A., Schärer, O.D. & Walter, J.C. (2011) Selective Bypass of a Lagging Strand Roadblock by the Eukaryotic Replicative DNA Helicase. *Cell*. 146 (6), 931–941. doi:10.1016/J.CELL.2011.07.045.

Fukui, T., Yamauchi, K., Muroya, T., Akiyama, M., Maki, H., Sugino, A. & Waga, S. (2004) Distinct roles of DNA polymerases delta and epsilon at the replication fork in Xenopus egg extracts. *Genes to Cells*. 9 (3), 179–191. doi:10.1111/J.1356-9597.2004.00716.X.

Gaillard, P.-H.L., Moggs, J.G., Roche, D.M.J., Quivy, J.-P., Becker, P.B., Wood, R.D. & Almouzni, G. (1997) Initiation and bidirectional propagation of chromatin assembly from a target site for nucleotide excision repair. *The EMBO Journal*. 16 (20), 6281–6289. doi:10.1093/EMBOJ/16.20.6281.

Galanti, L. & Pfander, B. (2018) Right time, right place—DNA damage and DNA replication checkpoints collectively safeguard S phase. *The EMBO Journal*. 37 (21). doi:10.15252/EMBJ.2018100681.

Gambus, A., van Deursen, F., Polychronopoulos, D., Foltman, M., Jones, R.C., Edmondson, R.D., Calzada, A. & Labib, K. (2009) A key role for Ctf4 in coupling the MCM2-7 helicase to DNA polymerase alpha within the eukaryotic replisome. *The EMBO journal*. 28 (19), 2992–3004. doi:10.1038/emboj.2009.226.

Gambus, A., Jones, R.C., Sanchez-Diaz, A., Kanemaki, M., van Deursen, F., Edmondson, R.D. & Labib, K. (2006) GINS maintains association of Cdc45 with MCM in replisome progression complexes at eukaryotic DNA replication forks. *Nature Cell Biology*. 8 (4), 358–366. doi:10.1038/ncb1382.

Gan, H., Serra-Cardona, A., Hua, X., Zhou, H., Labib, K., Yu, C. & Zhang, Z. (2018) The Mcm2-Ctf4-Polα Axis Facilitates Parental Histone H3-H4 Transfer to Lagging Strands. *Molecular Cell*. 72 (1), 140-151.e3. doi:10.1016/J.MOLCEL.2018.09.001.

Gangloff, S., McDonald, J.P., Bendixen, C., Arthur, L. & Rothstein, R. (1994) The yeast type I topoisomerase Top3 interacts with Sgs1, a DNA helicase homolog: a potential eukaryotic reverse gyrase. *Molecular and Cellular Biology*. 14 (12), 8391. doi:10.1128/MCB.14.12.8391.

García-Rodríguez, L.J., de Piccoli, G., Marchesi, V., Jones, R.C., Edmondson, R.D. & Labib, K. (2015) A conserved Polϵ binding module in Ctf18-RFC is required for S-phase checkpoint activation downstream of Mec1. *Nucleic acids research*. 43 (18), 8830–8838. doi:10.1093/NAR/GKV799.

García-Rodríguez, N., Morawska, M., Wong, R.P., Daigaku, Y. & Ulrich, H.D. (2018) Spatial separation between replisome- and template-induced replication stress signaling. *The EMBO journal*. 37 (9). doi:10.15252/EMBJ.201798369.

Gardner, K.A., Rine, J. & Fox, C.A. (1999) A region of the Sir1 protein dedicated to recognition of a silencer and required for interaction with the Orc1 protein in saccharomyces cerevisiae. *Genetics*. 151 (1), 31–44. doi:10.1093/GENETICS/151.1.31.

Garg, P. & Burgers, P.M.J. (2005) DNA polymerases that propagate the eukaryotic DNA replication fork. *Critical reviews in biochemistry and molecular biology*. 40 (2), 115–128. doi:10.1080/10409230590935433.

Gartner, A., Jovanović, A., Jeoung, D.-I., Bourlat, S., Cross, F.R. & Ammerer, G. (1998) Pheromone-Dependent G1 Cell Cycle Arrest Requires Far1 Phosphorylation, but May Not Involve Inhibition of

Cdc28-Cln2 Kinase, In Vivo. *Molecular and Cellular Biology*. 18 (7), 3681. doi:10.1128/MCB.18.7.3681.

Gasser, S.M. & Cockell, M.M. (2001) The molecular biology of the SIR proteins. *Gene*. 279 (1), 1–16. doi:10.1016/S0378-1119(01)00741-7.

Gaubitz, C., Liu, X., Pajak, J., Stone, N.P., Hayes, J.A., Demo, G. & Kelch, B.A. (2021) Cryo-EM structures reveal high-resolution mechanism of a DNA polymerase sliding clamp loader. *bioRxiv*. 2021.09.23.461575. doi:10.1101/2021.09.23.461575.

Gaudier, M., Schuwirth, B.S., Westcott, S.L. & Wigley, D.B. (2007) Structural basis of DNA replication origin recognition by an ORC protein. *Science (New York, N.Y.)*. 317 (5842), 1213–1216. doi:10.1126/SCIENCE.1143664.

Gautier, J., Norbury, C., Lohka, M., Nurse, P. & Maller, J. (1988) Purified maturation-promoting factor contains the product of a Xenopus homolog of the fission yeast cell cycle control gene cdc2+. *Cell*. 54 (3), 433–439. doi:10.1016/0092-8674(88)90206-1.

Gavin, K.A., Hidaka, M. & Stillman, B. (1995) Conserved initiator proteins in eukaryotes. *Science*. 270 (5242), 1667–1671. doi:10.1126/SCIENCE.270.5242.1667.

Ge, S.X., Jung, D., Jung, D. & Yao, R. (2020) ShinyGO: a graphical gene-set enrichment tool for animals and plants. *Bioinformatics*. 36 (8), 2628–2629. doi:10.1093/BIOINFORMATICS/BTZ931.

Gellon, L., Razidlo, D.F., Gleeson, O., Verra, L., Schulz, D., Lahue, R.S. & Freudenreich, C.H. (2011) New functions of Ctf18-RFC in preserving genome stability outside its role in sister chromatid cohesion. *PLoS genetics*. 7 (2). doi:10.1371/JOURNAL.PGEN.1001298.

Georgescu, R., Yuan, Z., Bai, L., de Luna Almeida Santos, R., Sun, J., Zhang, D., Yurieva, O., Li, H. & O'Donnell, M.E. (2017) Structure of eukaryotic CMG helicase at a replication fork and implications to replisome architecture and origin initiation. *Proceedings of the National Academy of Sciences of the United States of America*. 114 (5), E697–E706. doi:10.1073/PNAS.1620500114/-/DCSUPPLEMENTAL.

Georgescu, R.E., Langston, L., Yao, N.Y., Yurieva, O., Zhang, D., Finkelstein, J., Agarwal, T. & O'Donnell, M.E. (2014) Mechanism of asymmetric polymerase assembly at the eukaryotic replication fork. *Nature structural & molecular biology*. 21 (8), 664. doi:10.1038/NSMB.2851.

Georgescu, R.E., Schauer, G.D., Yao, N.Y., Langston, L.D., Yurieva, O., Zhang, D., Finkelstein, J. & O'Donnell, M.E. (2015) Reconstitution of a eukaryotic replisome reveals suppression mechanisms that define leading/lagging strand operation. *eLife*. 2015 (4). doi:10.7554/ELIFE.04988.

Ghidelli, S., Donze, D., Dhillon, N. & Kamakaka, R.T. (2001) Sir2p exists in two nucleosome-binding complexes with distinct deacetylase activities. *The EMBO Journal*. 20 (16), 4522–4535. doi:10.1093/EMBOJ/20.16.4522.

Ghosh, A.K., Rossi, M.L., Singh, D.K., Dunn, C., Ramamoorthy, M., Croteau, D.L., Liu, Y. & Bohr, V.A. (2012) RECQL4, the protein mutated in Rothmund-Thomson syndrome, functions in telomere maintenance. *The Journal of biological chemistry*. 287 (1), 196–209. doi:10.1074/JBC.M111.295063.

Gibson, S.I., Surosky, R.T. & Tye, B.K. (1990) The phenotype of the minichromosome maintenance mutant mcm3 is characteristic of mutants defective in DNA replication. *Molecular and cellular biology*. 10 (11), 5707–5720.

Gomez-Lorenzo, M.G., Valle, M., Frank, J., Gruss, C., Sorzano, C.O.S., Chen, X.S., Donate, L.E. & Carazo, J.M. (2003) Large T antigen on the simian virus 40 origin of replication: a 3D snapshot prior to DNA replication. *The EMBO journal*. 22 (23), 6205–6213. doi:10.1093/emboj/cdg612.

Goswami, P., Abid Ali, F., Douglas, M.E., Locke, J., Purkiss, A., Janska, A., Eickhoff, P., Early, A., Nans, A., Cheung, A.M.C., Diffley, J.F.X. & Costa, A. (2018) Structure of DNA-CMG-Pol epsilon elucidates the roles of the non-catalytic polymerase modules in the eukaryotic replisome. *Nature Communications 2018 9:1*. 9 (1), 1–13. doi:10.1038/s41467-018-07417-1.

Gowrishankar, J. (2015) End of the Beginning: Elongation and Termination Features of Alternative Modes of Chromosomal Replication Initiation in Bacteria. *PLOS Genetics*. 11 (1), e1004909. doi:10.1371/JOURNAL.PGEN.1004909.

de Graaf, B., Clore, A. & McCullough, A.K. (2009) Cellular Pathways for DNA Repair and Damage Tolerance of Formaldehyde-Induced DNA-Protein Crosslinks. *DNA repair*. 8 (10), 1207. doi:10.1016/J.DNAREP.2009.06.007.

Grabarczyk, D.B., Silkenat, S. & Kisker, C. (2018) Structural Basis for the Recruitment of Ctf18-RFC to the Replisome. *Structure (London, England : 1993)*. 26 (1), 137-144.e3. doi:10.1016/J.STR.2017.11.004.

Grafstrom, R.C., Fornace, A.J., Autrup, H., Lechner, J.F. & Harris, C.C. (1983) Formaldehyde damage to DNA and inhibition of DNA repair in human bronchial cells. *Science (New York, N.Y.)*. 220 (4593), 216–218. doi:10.1126/SCIENCE.6828890.

Gravel, S., Chapman, J.R., Magill, C. & Jackson, S.P. (2008) DNA helicases Sgs1 and BLM promote DNA double-strand break resection. *Genes & Development*. 22 (20), 2767. doi:10.1101/GAD.503108.

Gray, M.D., Shen, J.C., Kamath-Loeb, A.S., Blank, A., Sopher, B.L., Martin, G.M., Oshima, J. & Loeb, L.A. (1997) The Werner syndrome protein is a DNA helicase. *Nature genetics*. 17 (1), 100–103. doi:10.1038/NG0997-100.

Green, C.M. & Almouzni, G. (2003) Local action of the chromatin assembly factor CAF-1 at sites of nucleotide excision repair in vivo. *The EMBO Journal*. 22 (19), 5163. doi:10.1093/EMBOJ/CDG478.

Greenfeder, S.A. & Newlon, C.S. (1992) A replication map of a 61-kb circular derivative of Saccharomyces cerevisiae chromosome III. *Molecular Biology of the Cell*. 3 (9), 999. doi:10.1091/MBC.3.9.999.

Gregan, J., Lindner, K., Brimage, L., Franklin, R., Namdar, M., Hart, E.A., Aves, S.J. & Kearsey, S.E. (2003) Fission yeast Cdc23/Mcm10 functions after pre-replicative complex formation to promote Cdc45 chromatin binding. *Molecular biology of the cell*. 14 (9), 3876–3887. doi:10.1091/MBC.E03-02-0090.

Groocock, L.M., Prudden, J., Perry, J.J.P. & Boddy, M.N. (2012) The RecQ4 Orthologue Hrq1 Is Critical for DNA Interstrand Cross-Link Repair and Genome Stability in Fission Yeast. *Molecular and Cellular Biology*. 32 (2), 276. doi:10.1128/MCB.06184-11.

Gruber, S., Arumugam, P., Katou, Y., Kuglitsch, D., Helmhart, W., Shirahige, K. & Nasmyth, K. (2006) Evidence that loading of cohesin onto chromosomes involves opening of its SMC hinge. *Cell*. 127 (3), 523–537. doi:10.1016/J.CELL.2006.08.048.

Grunstein, M. (1998) Yeast heterochromatin: regulation of its assembly and inheritance by histones. *Cell*. 93 (3), 325–328. doi:10.1016/S0092-8674(00)81160-5.

Guerrero, C., Milenković, T., Pržulj, N., Kaiser, P. & Huang, L. (2008) Characterization of the proteasome interaction network using a QTAX-based tag-team strategy and protein interaction network analysis. *Proceedings of the National Academy of Sciences of the United States of America*. 105 (36), 13333. doi:10.1073/PNAS.0801870105.

Guerrero, C., Tagwerker, C., Kaiser, P. & Huang, L. (2006) An integrated mass spectrometry-based proteomic approach: quantitative analysis of tandem affinity-purified in vivo cross-linked protein complexes (QTAX) to decipher the 26 S proteasome-interacting network. *Molecular & cellular proteomics : MCP*. 5 (2), 366–378. doi:10.1074/MCP.M500303-MCP200.

Guilliam, T.A. & Yeeles, J.T.P. (2020) An updated perspective on the polymerase division of labor during eukaryotic DNA replication. *Critical reviews in biochemistry and molecular biology*. 55 (5), 469–481. doi:10.1080/10409238.2020.1811630.

Gupta, S., Friedman, L.J., Gelles, J. & Bell, S.P. (2021) A helicase-tethered ORC flip enables bidirectional helicase loading. *eLife*. 10. doi:10.7554/ELIFE.74282.

Gupta, S.V. & Schmidt, K.H. (2020) Maintenance of Yeast Genome Integrity by RecQ Family DNA Helicases. *Genes*. 11 (2). doi:10.3390/GENES11020205.

Haase, S.B. & Reed, S.I. (2001) Improved Flow Cytometric Analysis of the Budding Yeast Cell Cycle. *http://dx.doi.org/10.4161/cc.1.2.114*. 1 (2), 117–121. doi:10.4161/CC.1.2.114.

Hagan, I., Hayles, J. & Nurse, P. (1988) Cloning and sequencing of the cyclin-related cdc13+ gene and a cytological study of its role in fission yeast mitosis. *Journal of cell science*. 91 (Pt 4), 587–595.

Hammond, C.M., Strømme, C.B., Huang, H., Patel, D.J. & Groth, A. (2017) Histone chaperone networks shaping chromatin function. *Nature reviews. Molecular cell biology*. 18 (3), 141–158. doi:10.1038/NRM.2016.159.

Han, J., Li, Q., McCullough, L., Kettelkamp, C., Formosa, T. & Zhang, Z. (2010) Ubiquitylation of FACT by the Cullin-E3 ligase Rtt101 connects FACT to DNA replication. *Genes & Development*. 24 (14), 1485–1490. doi:10.1101/GAD.1887310.

Hanna, J.S., Kroll, E.S., Lundblad, V. & Spencer, F.A. (2001) Saccharomyces cerevisiae CTF18 and CTF4 are required for sister chromatid cohesion. *Molecular and cellular biology*. 21 (9), 3144–3158. doi:10.1128/MCB.21.9.3144-3158.2001.

Hardy, C.F.J. (1997) Identification of Cdc45p, an essential factor required for DNA replication. *Gene*. 187 (2), 239–246. doi:10.1016/S0378-1119(96)00761-5.

Hardy, J., Churikov, D., Géli, V. & Simon, M.N. (2014) Sgs1 and Sae2 promote telomere replication by limiting accumulation of ssDNA. *Nature communications*. 5. doi:10.1038/NCOMMS6004.

Harrison, J.C. & Haber, J.E. (2006) Surviving the breakup: the DNA damage checkpoint. *Annual review of genetics*. 40, 209–235. doi:10.1146/ANNUREV.GENET.40.051206.105231.

Hartung, F. & Puchta, H. (2006) The RecQ gene family in plants. *Journal of plant physiology*. 163 (3), 287–296. doi:10.1016/J.JPLPH.2005.10.013.

Hartwell, L.H. (1976) Sequential function of gene products relative to DNA synthesis in the yeast cell cycle. *Journal of Molecular Biology*. 104 (4), 803–817. doi:10.1016/0022-2836(76)90183-2.

Hartwell, L.H., Culotti, J., Pringle, J.R. & Reid, B.J. (1974) Genetic control of the cell division cycle in yeast. *Science (New York, N.Y.)*. 183 (4120), 46–51.

Hartwell, L.H., Culotti, J. & Reid, B. (1970) Genetic control of the cell-division cycle in yeast. I. Detection of mutants. *Proceedings of the National Academy of Sciences of the United States of America*. 66 (2), 352–359. doi:10.1073/pnas.66.2.352.

Hartwell, L.H., Mortimer, R.K., Culotti, J. & Culotti, M. (1973) Genetic Control of the Cell Division Cycle in Yeast: V. Genetic Analysis of cdc Mutants. *Genetics*. 74 (2), 267–286.

Hattman, S., Kenny, C., Berger, L. & Pratt, K. (1978) Comparative study of DNA methylation in three unicellular eucaryotes. *Journal of Bacteriology*. 135 (3), 1156. /pmc/articles/PMC222496/?report=abstract.

Hayano, M., Kanoh, Y., Matsumoto, S., Renard-Guillet, C., Shirahige, K. & Masai, H. (2012) Rif1 is a global regulator of timing of replication origin firing in fission yeast. *Genes & development*. 26 (2), 137–150. doi:10.1101/GAD.178491.111.

Hecht, A., Laroche, T., Strahl-Bolsinger, S., Gasser, S.M. & Grunstein, M. (1995) Histone H3 and H4 N-termini interact with SIR3 and SIR4 proteins: A molecular model for the formation of heterochromatin in yeast. *Cell*. 80 (4), 583–592. doi:10.1016/0092-8674(95)90512-X.

Heitman, J., Movva, N.R. & Hall, M.N. (1991) Targets for cell cycle arrest by the immunosuppressant rapamycin in yeast. *Science (New York, N.Y.)*. 253 (5022), 905–909. doi:10.1126/SCIENCE.1715094.

Heller, R.C., Kang, S., Lam, W.M., Chen, S., Chan, C.S. & Bell, S.P. (2011) Eukaryotic origin-dependent DNA replication in vitro reveals sequential action of DDK and S-CDK kinases. *Cell*. 146 (1), 80–91. doi:10.1016/J.CELL.2011.06.012.

Hennion, M., Arbona, J.M., Lacroix, L., Cruaud, C., Theulot, B., Tallec, B. le, Proux, F., Wu, X., Novikova, E., Engelen, S., Lemainque, A., Audit, B. & Hyrien, O. (2020) FORK-seq: Replication landscape of the Saccharomyces cerevisiae genome by nanopore sequencing. *Genome Biology*. 21 (1), 1–25. doi:10.1186/S13059-020-02013-3/FIGURES/7.

Herskowitz, I. (1995) MAP kinase pathways in yeast: For mating and more. *Cell*. 80 (2), 187–197. doi:10.1016/0092-8674(95)90402-6.

Herskowitz, I. & Jensen, R.E. (1991) [8] Putting the HO gene to work: Practical uses for mating-type switching. *Methods in Enzymology*. 194 (C), 132–146. doi:10.1016/0076-6879(91)94011-Z.

Hickman, M.A., Froyd, C.A. & Rusche, L.N. (2011) Reinventing heterochromatin in budding yeasts: Sir2 and the origin recognition complex take center stage. *Eukaryotic Cell*. 10 (9), 1183–1192. doi:10.1128/EC.05123-11.

Hierlmeier, T., Merl, J., Sauert, M., Perez-Fernandez, J., Schultz, P., Bruckmann, A., Hamperl, S., Ohmayer, U., Rachel, R., Jacob, A., Hergert, K., Deutzmann, R., Griesenbeck, J., Hurt, E., Milkereit, P., Baßler, J. & Tschochner, H. (2013) Rrp5p, Noc1p and Noc2p form a protein module which is part of early large ribosomal subunit precursors in S. cerevisiae. *Nucleic Acids Research*. 41 (2), 1191. doi:10.1093/NAR/GKS1056.

Hill, T.M., Henson, J.M. & Kuempel, P.L. (1987) The terminus region of the Escherichia coli chromosome contains two separate loci that exhibit polar inhibition of replication. *Proceedings of the National Academy of Sciences*. 84 (7), 1754–1758. doi:10.1073/PNAS.84.7.1754.

Hilleren, P., McCarthy, T., Rosbash, M., Parker, R. & Jensen, T.H. (2001) Quality control of mRNA 3′-end processing is linked to the nuclear exosome. *Nature 2001 413:6855*. 413 (6855), 538–542. doi:10.1038/35097110.

Hindley, J. & Phear, G.A. (1984) Sequence of the cell division gene CDC2 from Schizosaccharomyces pombe; patterns of splicing and homology to protein kinases. *Gene*. 31 (1–3), 129–134. doi:10.1016/0378-1119(84)90203-8.

Hiraga, S.I., Alvino, G.M., Chang, F.J., Lian, H.Y., Sridhar, A., Kubota, T., Brewer, B.J., Weinreich, M., Raghuraman, M.K. & Donaldson, A.D. (2014) Rif1 controls DNA replication by directing Protein Phosphatase 1 to reverse Cdc7-mediated phosphorylation of the MCM complex. *Genes & development*. 28 (4), 372–383. doi:10.1101/GAD.231258.113.

Hodgson, B., Calzada, A. & Labib, K. (2007) Mrc1 and Tof1 regulate DNA replication forks in different ways during normal S phase. *Molecular biology of the cell*. 18 (10), 3894–3902. doi:10.1091/MBC.E07-05-0500.

Hoek, M., Myers, M.P. & Stillman, B. (2011) An analysis of CAF-1-interacting proteins reveals dynamic and direct interactions with the KU complex and 14-3-3 proteins. *The Journal of biological chemistry*. 286 (12), 10876–10887. doi:10.1074/JBC.M110.217075.

Hoek, M. & Stillman, B. (2003) Chromatin assembly factor 1 is essential and couples chromatin assembly to DNA replication in vivo. *Proceedings of the National Academy of Sciences of the United States of America*. 100 (21), 12183–12188. doi:10.1073/PNAS.1635158100.

Hoffman, E.A., Frey, B.L., Smith, L.M. & Auble, D.T. (2015) Formaldehyde crosslinking: a tool for the study of chromatin complexes. *The Journal of biological chemistry*. 290 (44), 26404–26411. doi:10.1074/JBC.R115.651679.

Hogg, M., Osterman, P., Bylund, G.O., Ganai, R.A., Lundström, E.B., Sauer-Eriksson, A.E. & Johansson, E. (2013) Structural basis for processive DNA synthesis by yeast DNA polymerase ε. *Nature Structural & Molecular Biology 2013 21:1*. 21 (1), 49–55. doi:10.1038/nsmb.2712.

Hoggard, T., Müller, C.A., Nieduszynski, C.A., Weinreich, M. & Fox, C.A. (2020) Sir2 mitigates an intrinsic imbalance in origin licensing efficiency between early- and late-replicating euchromatin. *Proceedings of the National Academy of Sciences*. 117 (25), 14314–14321. doi:10.1073/PNAS.2004664117.

Hoggard, T.A., Chang, F., Perry, K.R., Subramanian, S., Kenworthy, J., Chueng, J., Shor, E., Hyland, E.M., Boeke, J.D., Weinreich, M. & Fox, C.A. (2018) Yeast heterochromatin regulators Sir2 and Sir3 act directly at euchromatic DNA replication origins. *PLOS Genetics*. 14 (5), e1007418. doi:10.1371/JOURNAL.PGEN.1007418.

Holliday R, Pugh JE. DNA modification mechanisms and gene activity during development. Science. 1975 Jan 24;187(4173):226-32. PMID: 1111098.

Holm, C., Goto, T., Wang, J.C. & Botstein, D. (1985) DNA topoisomerase II is required at the time of mitosis in yeast. *Cell*. 41 (2), 553–563. doi:10.1016/S0092-8674(85)80028-3.

Holmes, S.G., Rose, A.B., Steuerle, K., Saez, E., Sayegh, S., Lee, Y.M. & Broach, J.R. (1997) Hyperactivation of the silencing proteins, Sir2p and Sir3p, causes chromosome loss. *Genetics*. 145 (3), 605–614. doi:10.1093/GENETICS/145.3.605.

Hong, E.J.E. & Shirleen Roeder, G. (2002) A role for Ddc1 in signaling meiotic double-strand breaks at the pachytene checkpoint. *Genes & development*. 16 (3), 363–376. doi:10.1101/GAD.938102.

Hoppe, G.J., Tanny, J.C., Rudner, A.D., Gerber, S.A., Danaie, S., Gygi, S.P. & Moazed, D. (2002) Steps in Assembly of Silent Chromatin in Yeast: Sir3-Independent Binding of a Sir2/Sir4 Complex to Silencers and Role for Sir2-Dependent Deacetylation. *Molecular and Cellular Biology*. 22 (12), 4167. doi:10.1128/MCB.22.12.4167-4180.2002.

Hopwood, B. & Dalton, S. (1996) Cdc45p assembles into a complex with Cdc46p/Mcm5p, is required for minichromosome maintenance, and is essential for chromosomal DNA replication. *Proceedings of the National Academy of Sciences of the United States of America*. 93 (22), 12309–12314. doi:10.1073/PNAS.93.22.12309.

Howard, A. and Pelc, S.R. (1953) Synthesis of deoxyribonucleic acid in normal and irradiated ceils and its relation to chromosome breakage. *Heredity Suppl*. 6, 261–273.

Hu, Y., Tang, H.B., Liu, N.N., Tong, X.J., Dang, W., Duan, Y.M., Fu, X.H., Zhang, Y., Peng, J., Meng, F.L. & Zhou, J.Q. (2013) Telomerase-Null Survivor Screening Identifies Novel Telomere Recombination Regulators. *PLoS Genetics*. 9 (1), 1003208. doi:10.1371/JOURNAL.PGEN.1003208.

Huang, H., Strømme, C.B., Saredi, G., Hödl, M., Strandsby, A., González-Aguilera, C., Chen, S., Groth, A. & Patel, D.J. (2015) A unique binding mode enables MCM2 to chaperone histones H3–H4 at replication forks. *Nature Structural & Molecular Biology 2015 22:8*. 22 (8), 618–626. doi:10.1038/nsmb.3055.

Huang, M., Zhou, Z. & Elledge, S.J. (1998) The DNA Replication and Damage Checkpoint Pathways Induce Transcription by Inhibition of the Crt1 Repressor. *Cell*. 94 (5), 595–605. doi:10.1016/S0092-8674(00)81601-3.

Huang, P.H., Pryde, F.E., Lester, D., Maddison, R.L., Borts, R.H., Hickson, I.D. & Louis, E.J. (2001) SGS1 is required for telomere elongation in the absence of telomerase. *Current biology : CB*. 11 (2), 125–129. doi:10.1016/S0960-9822(01)00021-5.

Huberman, J.A. & Riggs, A.D. (1966) Autoradiography of chromosomal DNA fibers from Chinese hamster cells. *Proceedings of the National Academy of Sciences of the United States of America*. 55 (3), 599. doi:10.1073/PNAS.55.3.599.

Huberman, J.A., Spotila, L.D., Nawotka, K.A., El-Assouli, S.M. & Davis, L.R. (1987) The in vivo replication origin of the yeast 2µm plasmid. *Cell*. 51 (3), 473–481. doi:10.1016/0092-8674(87)90643-X.

Huberman, J.A., Zhu, J.G., Davis, L.R. & Newlon, C.S. (1988) Close association of a DNA replication origin and an ARS element on chromosome III of the yeast, Saccharomyces cerevisiae. *Nucleic Acids Research*. 16 (14A), 6373. doi:10.1093/NAR/16.14.6373.

Hughes, C.S., Foehr, S., Garfield, D.A., Furlong, E.E., Steinmetz, L.M. & Krijgsveld, J. (2014) Ultrasensitive proteome analysis using paramagnetic bead technology. *Molecular Systems Biology*. 10 (10), 757. doi:10.15252/MSB.20145625.

Hughes, C.S., Moggridge, S., Müller, T., Sorensen, P.H., Morin, G.B. & Krijgsveld, J. (2018) Single-pot, solid-phase-enhanced sample preparation for proteomics experiments. *Nature Protocols 2018 14:1*. 14 (1), 68–85. doi:10.1038/s41596-018-0082-x.

Iftode, C., Daniely, Y. & Borowiec, J.A. (1999) Replication protein A (RPA): the eukaryotic SSB. *Critical reviews in biochemistry and molecular biology*. 34 (3), 141–180. doi:10.1080/10409239991209255.

Ilves, I., Petojevic, T., Pesavento, J.J. & Botchan, M.R. (2010) Activation of the MCM2-7 helicase by association with Cdc45 and GINS proteins. *Molecular cell*. 37 (2), 247–258. doi:10.1016/j.molcel.2009.12.030.

Imai, S., Armstrong, C.M., Kaeberlein, M. & Guarente, L. (2000) Transcriptional silencing and longevity protein Sir2 is an NAD-dependent histone deacetylase. *Nature 2000 403:6771*. 403 (6771), 795–800. doi:10.1038/35001622.

de Ioannes, P., Leon, V.A., Kuang, Z., Wang, M., Boeke, J.D., Hochwagen, A. & Armache, K.-J. (2019) Structure and function of the Orc1 BAH-nucleosome complex. *Nature Communications 2019 10:1*. 10 (1), 1–11. doi:10.1038/s41467-019-10609-y.

Ishibashi, T., Dangkulwanich, M., Coello, Y., Lionberger, T.A., Lubkowska, L., Ponticelli, A.S., Kashlev, M. & Bustamante, C. (2014) Transcription factors IIS and IIF enhance transcription efficiency by differentially modifying RNA polymerase pausing dynamics. *Proceedings of the National Academy of Sciences of the United States of America*. 111 (9), 3419–3424. doi:10.1073/PNAS.1401611111/-/DCSUPPLEMENTAL.

Ishimi, Y., Ichinose, S., Omori, A., Sato, K. & Kimura, H. (1996) Binding of human minichromosome maintenance proteins with histone H3. *The Journal of biological chemistry*. 271 (39), 24115–24122. doi:10.1074/JBC.271.39.24115.

Itou, H., Shirakihara, Y. & Araki, H. (2015) The quaternary structure of the eukaryotic DNA replication proteins Sld7 and Sld3. *Acta crystallographica. Section D, Biological crystallography*. 71 (Pt 8), 1649–1656. doi:10.1107/S1399004715010457.

Ivessa, A.S., Lenzmeier, B.A., Bessler, J.B., Goudsouzian, L.K., Schnakenberg, S.L. & Zakian, V.A. (2003) The Saccharomyces cerevisiae Helicase Rrm3p Facilitates Replication Past Nonhistone Protein-DNA Complexes. *Molecular Cell*.12.

Ivessa, A.S., Zhou, J.Q., Schulz, V.P., Monson, E.K. & Zakian, V.A. (2002) Saccharomyces Rrm3p, a 5′ to 3′ DNA helicase that promotes replication fork progression through telomeric and subtelomeric DNA. *Genes and Development*. 16 (11), 1383–1396. doi:10.1101/gad.982902.

Ivessa, A.S., Zhou, J.-Q. & Zakian, V.A. (2000) The Saccharomyces Pif1p DNA Helicase and the Highly Related Rrm3p Have Opposite Effects on Replication Fork Progression in Ribosomal DNA Mutations in PIF1, which encodes a 5 to 3 DNA helicase (Lahaye et al., 1991), cause telomere lengthening and a large increase in the rate of forming new telo-meres (Schulz and Zakian, 1994), suggesting that Pif1p inhibits telomere replication. When the sequence of the. *Cell*.100.

Iyer, D.R. & Rhind, N. (2017) The Intra-S Checkpoint Responses to DNA Damage. *Genes*. 8 (2). doi:10.3390/GENES8020074.

Jaehnig, E.J., Kuo, D., Hombauer, H., Ideker, T.G. & Kolodner, R.D. (2013) Checkpoint kinases regulate a global network of transcription factors in response to DNA damage. *Cell reports*. 4 (1), 174–188. doi:10.1016/J.CELREP.2013.05.041.

Janke, C., Magiera, M.M., Rathfelder, N., Taxis, C., Reber, S., Maekawa, H., Moreno-Borchart, A., Doenges, G., Schwob, E., Schiebel, E. & Knop, M. (2004) A versatile toolbox for PCR-based tagging of yeast genes: new fluorescent proteins, more markers and promoter substitution cassettes. *Yeast (Chichester, England)*. 21 (11), 947–962. doi:10.1002/YEA.1142.

Jares, P. & Blow, J.J. (2000) Xenopus Cdc7 function is dependent on licensing but not on XORC, XCdc6, or CDK activity and is required for XCdc45 loading. *Genes & Development*. 14 (12), 1528. /pmc/articles/PMC316685/.

Jenkyn-Bedford, M., Jones, M.L., Baris, Y., Labib, K.P.M., Cannone, G., Yeeles, J.T.P. & Deegan, T.D. (2021) A conserved mechanism for regulating replisome disassembly in eukaryotes. *Nature 2021*. 1–5. doi:10.1038/s41586-021-04145-3.

Jeong, Y.J., Levin, M.K. & Patel, S.S. (2004) The DNA-unwinding mechanism of the ring helicase of bacteriophage T7. *Proceedings of the National Academy of Sciences*. 101 (19), 7264–7269. doi:10.1073/PNAS.0400372101.

Jiao, R., Harrigan, J.A., Shevelev, I., Dietschy, T., Selak, N., Indig, F.E., Piotrowski, J., Janscak, P., Bohr, V.A. & Stagljar, I. (2007) The Werner syndrome protein is required for recruitment of chromatin assembly factor 1 following DNA damage. *Oncogene*. 26 (26), 3811–3822. doi:10.1038/SJ.ONC.1210150.

Lee J.K., Moon K.Y., Jiang Y.. & Hurwitz J. (2001) The Schizosaccharomyces pombe origin recognition complex interacts with multiple AT-rich regions of the replication origin DNA by means of the AT-hook domains of the spOrc4 protein. *Proceedings of the National Academy of Sciences of the United States of America*. 98 (24), 13589–13594. doi:10.1073/PNAS.251530398.

Ivy J.M., Klar A.J. & Hicks J.B. (1986) Cloning and characterization of four SIR genes of Saccharomyces cerevisiae. *Molecular and cellular biology*. 6 (2), 688–702. doi:10.1128/MCB.6.2.688-702.1986.

Johansson, E. & MacNeill, S.A. (2010) The eukaryotic replicative DNA polymerases take shape. *Trends in Biochemical Sciences*. 35 (6), 339–347. doi:10.1016/J.TIBS.2010.01.004.

Johnson, A. & O'donnell, M. (2005) CELLULAR DNA REPLICASES: Components and Dynamics at the Replication Fork. *Annu. Rev. Biochem*. 74, 283–315. doi:10.1146/annurev.biochem.73.011303.073859.

Johnson, A. & Skotheim, J.M. (2013) Start and the Restriction Point. *Current opinion in cell biology*. 25 (6), 717–723. doi:10.1016/J.CEB.2013.07.010.

Johnson, F.B., Marciniak, R.A., McVey, M., Stewart, S.A., Hahn, W.C. & Guarente, L. (2001) The Saccharomyces cerevisiae WRN homolog Sgs1p participates in telomere maintenance in cells lacking telomerase. *The EMBO journal*. 20 (4), 905–913. doi:10.1093/EMBOJ/20.4.905.

Jones, M.L., Baris, Y., Taylor, M.R.G. & Yeeles, J.T.P. (2021) Structure of a human replisome shows the organisation and interactions of a DNA replication machine. *The EMBO Journal*. 40 (23). doi:10.15252/EMBJ.2021108819.

Jossen, R. & Bermejo, R. (2013) The DNA damage checkpoint response to replication stress: A Game of Forks. *Frontiers in genetics*. 4 (MAR). doi:10.3389/FGENE.2013.00026.

Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, Tunyasuvunakool K, Bates R, Žídek A, Potapenko A, Bridgland A, Meyer C, Kohl SAA, Ballard AJ, Cowie A, Romera-Paredes B, Nikolov S, Jain R, Adler J, Back T, Petersen S, Reiman D, Clancy E, Zielinski M, Steinegger M, Pacholska M, Berghammer T, Bodenstein S, Silver D, Vinyals O, Senior AW, Kavukcuoglu K, Kohli P, Hassabis D. Highly accurate protein structure prediction with AlphaFold. Nature. 2021 Aug;596(7873):583-589. doi: 10.1038/s41586-021-03819-2. Epub 2021 Jul 15. PMID: 34265844; PMCID: PMC8371605.

Kaake, R.M., Milenković, T., Pržulj, N., Kaiser, P. & Huang, L. (2010) Characterization of Cell Cycle Specific Protein Interaction Networks of the Yeast 26S Proteasome Complex by the QTAX Strategy. *Journal of proteome research*. 9 (4), 2016. doi:10.1021/PR1000175.

Kaliraman, V., Mullen, J.R., Fricke, W.M., Bastin-Shanower, S.A. & Brill, S.J. (2001) Functional overlap between Sgs1-Top3 and the Mms4-Mus81 endonuclease. *Genes & development*. 15 (20), 2730–2740. doi:10.1101/GAD.932201.

Kamimura, Y., Masumoto, H., Sugino, A. & Araki, H. (1998) Sld2, Which Interacts with Dpb11 in Saccharomyces cerevisiae , Is Required for Chromosomal DNA Replication . *Molecular and Cellular Biology*. 18 (10), 6102–6109. doi:10.1128/MCB.18.10.6102/ASSET/FA311308-784B-45CE-A57E-590727497677/ASSETS/GRAPHIC/MB1080315008.JPEG.

Kamimura, Y., Tak, Y.-S., Sugino, A. & Araki, H. (2001) Sld3, which interacts with Cdc45 (Sld4), functions for chromosomal DNA replication in Saccharomyces cerevisiae. *The EMBO Journal*. 20 (8), 2097. doi:10.1093/EMBOJ/20.8.2097.

Kanemaki, M. & Labib, K. (2006) Distinct roles for Sld3 and GINS during establishment and progression of eukaryotic DNA replication forks. *The EMBO Journal*. 25 (8), 1753. doi:10.1038/SJ.EMBOJ.7601063.

Kanemaki, M., Sanchez-Diaz, A., Gambus, A. & Labib, K. (2003) Functional proteomic identification of DNA replication proteins by induced proteolysis in vivo. *Nature 2003 423:6941*. 423 (6941), 720–725. doi:10.1038/nature01692.

Kang, S., Warner, M.D. & Bell, S.P. (2014) Multiple functions for Mcm2-7 ATPase motifs during replication initiation. *Molecular cell*. 55 (5), 655–665. doi:10.1016/j.molcel.2014.06.033.

Kang, Y.-H., Galal, W.C., Farina, A., Tappin, I. & Hurwitz, J. (2012) Properties of the human Cdc45/Mcm2-7/GINS helicase complex and its action with DNA polymerase ε in rolling circle DNA synthesis. *Proceedings of the National Academy of Sciences*. 109 (16), 6042–6047. doi:10.1073/PNAS.1203734109.

Kanter, D.M. & Kaplan, D.L. (2011) Sld2 binds to origin single-stranded DNA and stimulates DNA annealing. *Nucleic Acids Research*. 39 (7), 2580. doi:10.1093/NAR/GKQ1222.

Katou, Y., Kanoh, Y., Bando, M., Noguchi, H., Tanaka, H., Ashikari, T., Sugimoto, K. & Shirahige, K. (2003) S-phase checkpoint proteins Tof1 and Mrc1 form a stable replication-pausing complex. *Nature*. 424 (6952), 1078–1083. doi:10.1038/NATURE01900.

Kaufman, P.D., Kobayashi, R. & Stillman, B. (1997) Ultraviolet radiation sensitivity and reduction of telomeric silencing in Saccharomyces cerevisiae cells lacking chromatin assembly factor-I. *Genes & development*. 11 (3), 345–357. doi:10.1101/GAD.11.3.345.

Kawakami, H. & Katayama, T. (2010) DnaA, ORC, and Cdc6: similarity beyond the domains of life and diversity. *Biochemistry and cell biology = Biochimie et biologie cellulaire*. 88 (1), 49–62. doi:10.1139/O09-154.

Kawanishi, M., Matsuda, T. & Yagi, T. (2014) Genotoxicity of formaldehyde: Molecular basis of DNA damage and mutation. *Frontiers in Environmental Science*. 2 (SEP), 36. doi:10.3389/FENVS.2014.00036/BIBTEX.

Kawasaki, Y., Hiraga, S.I. & Sugino, A. (2000) Interactions between Mcm10p and other replication factors are required for proper initiation and elongation of chromosomal DNA replication in Saccharomyces cerevisiae. *Genes to cells : devoted to molecular & cellular mechanisms*. 5 (12), 975–989. doi:10.1046/J.1365-2443.2000.00387.X.

Kawasaki, Y., Kim, H.-D., Kojima, A., Seki, T. & Sugino, A. (2006) Reconstitution of Saccharomyces cerevisiae prereplicative complex assembly in vitro. *Genes to Cells*. 11 (7), 745–756. doi:10.1111/J.1365-2443.2006.00975.X.

Kellis, M., Patterson, N., Endrizzi, M., Birren, B. & Lander, E.S. (2003) Sequencing and comparison of yeast species to identify genes and regulatory elements. *Nature*. 423 (6937), 241–254. doi:10.1038/NATURE01644.

Kelly, T.J., Nurse, P. & Forsburg, S.L. (1993) Coupling DNA replication to the cell cycle. *Cold Spring Harbor Symposia on Quantitative Biology*. 58, 637–644. doi:10.1101/SQB.1993.058.01.071.

Kesti, T., Flick, K., Keränen, S., Syväoja, J.E. & Wittenberg, C. (1999) DNA Polymerase ε Catalytic Domains Are Dispensable for DNA Replication, DNA Repair, and Cell Viability. *Molecular Cell*. 3 (5), 679–685. doi:10.1016/S1097-2765(00)80361-5.

Keszthelyi, A., Minchell, N.E. & Baxter, J. (2016) The Causes and Consequences of Topological Stress during DNA Replication. *Genes*. 7 (12). doi:10.3390/GENES7120134.

Khmelinskii, A., Blaszczak, E., Pantazopoulou, M., Fischer, B., Omnus, D.J., Dez, G. le, Brossard, A., Gunnarsson, A., Barry, J.D., Meurer, M., Kirrmaier, D., Boone, C., Huber, W., Rabut, G., Ljungdahl, P.O. & Knop, M. (2014) Protein quality control at the inner nuclear membrane. *Nature*. 516 (7531), 410–413. doi:10.1038/NATURE14096.

Kim, J. & Roeder, R.G. (2009) Direct Bre1-Paf1 Complex Interactions and RING Finger-independent Bre1-Rad6 Interactions Mediate Histone H2B Ubiquitylation in Yeast. *The Journal of Biological Chemistry*. 284 (31), 20582. doi:10.1074/JBC.M109.017442.

Kim, R.A. & Wang, J.C. (1989) Function of DNA topoisomerases as replication swivels in Saccharomyces cerevisiae. *Journal of molecular biology*. 208 (2), 257–267. doi:10.1016/0022-2836(89)90387-2.

Klein, H.L. (1997) Rdh54, a Rad54 Homologue in Saccharomyces Cerevisiae, Is Required for Mitotic Diploid-Specific Recombination and Repair and for Meiosis. *Genetics*. 147 (4), 1533. doi:10.1093/genetics/147.4.1533.

Knop, M., Siegers, K., Pereira, G., Zachariae, W., Winsor, B., Nasmyth, K. & Schiebel, E. (1999) Epitope Tagging of Yeast Genes using a PCR-based Strategy: More Tags and Improved Practical Routines. *Yeast*. 15, 963–972. doi:10.1002/(SICI)1097-0061(199907)15:10B.

Knott, S.R.V., Peace, J.M., Ostrow, A.Z., Gan, Y., Rex, A.E., Viggiani, C.J., Tavaré, S. & Aparicio, O.M. (2012) Forkhead transcription factors establish origin timing and long-range clustering in S. cerevisiae. *Cell*. 148 (1–2), 99–111. doi:10.1016/J.CELL.2011.12.012.

Knott, S.R.V., Viggiani, C.J., Tavaré, S. & Aparicio, O.M. (2009) Genome-wide replication profiles indicate an expansive role for Rpd3L in regulating replication initiation timing or efficiency, and reveal genomic loci of Rpd3 function in Saccharomyces cerevisiae. *Genes & Development*. 23 (9), 1077–1090. doi:10.1101/GAD.1784309.

Koç, A., Wheeler, L.J., Mathews, C.K. & Merrill, G.F. (2004) Hydroxyurea Arrests DNA Replication by a Mechanism That Preserves Basal dNTP Pools *. *Journal of Biological Chemistry*. 279 (1), 223–230. doi:10.1074/JBC.M303952200.

Kong, X.P., Onrust, R., O'Donnell, M. & Kuriyan, J. (1992) Three-dimensional structure of the β subunit of E. coli DNA polymerase III holoenzyme: A sliding DNA clamp. *Cell*. 69 (3), 425–437. doi:10.1016/0092-8674(92)90445-I.

Kornberg, R.D. (1974) Chromatin structure: A repeating unit of histones and DNA. *Science*. 184 (4139), 868–871. doi:10.1126/SCIENCE.184.4139.868/FORMAT/PDF.

Kramer, E.R., Gieffers, C., Hölzl, G., Hengstschläger, M. & Peters, J.-M. (1998) Activation of the human anaphase-promoting complex by proteins of the CDC20/Fizzy family. *Current Biology*. 8 (22), S1–S4. doi:10.1016/S0960-9822(07)00510-6.

Krastanova, I., Sannino, V., Amenitsch, H., Gileadi, O., Pisani, F.M. & Onesti, S. (2012) Structural and Functional Insights into the DNA Replication Factor Cdc45 Reveal an Evolutionary Relationship to the DHH Family of Phosphoesterases *. *Journal of Biological Chemistry*. 287 (6), 4121–4128. doi:10.1074/JBC.M111.285395.

Krawitz, D.C., Kama, T. & Kaufman, P.D. (2002) Chromatin Assembly Factor I Mutants Defective for PCNA Binding Require Asf1/Hir Proteins for Silencing. *Molecular and Cellular Biology*. 22 (2), 614. doi:10.1128/MCB.22.2.614-625.2002.

Krishna, T.S.R., Kong, X.P., Gary, S., Burgers, P.M. & Kuriyan, J. (1994) Crystal structure of the eukaryotic DNA polymerase processivity factor PCNA. *Cell*. 79 (7), 1233–1243. doi:10.1016/0092-8674(94)90014-0.

Krude, T. (1995) Chromatin assembly factor 1 (CAF-1) colocalizes with replication foci in HeLa cell nuclei. *Experimental cell research*. 220 (2), 304–311. doi:10.1006/EXCR.1995.1320.

Kubota, T., Hiraga, S.I., Yamada, K., Lamond, A.I. & Donaldson, A.D. (2011) Quantitative proteomic analysis of chromatin reveals that Ctf18 acts in the DNA replication checkpoint. *Molecular & cellular proteomics : MCP*. 10 (7). doi:10.1074/MCP.M110.005561.

Kumagai, A., Shevchenko, A., Shevchenko, A. & Dunphy, W.G. (2010) Treslin collaborates with TopBP1 in triggering the initiation of DNA replication. *Cell*. 140 (3), 349–359. doi:10.1016/J.CELL.2009.12.049.

Kunkel, T.A. & Burgers, P.M. (2008) Dividing the workload at a eukaryotic replication fork. *Trends in Cell Biology*. 18 (11), 521–527. doi:10.1016/J.TCB.2008.08.005.

Kuo, M.H. & Allis, C.D. (1999) In Vivo Cross-Linking and Immunoprecipitation for Studying Dynamic Protein:DNA Associations in a Chromatin Environment. *Methods*. 19 (3), 425–433. doi:10.1006/METH.1999.0879.

Kuranda, K., Leberre, V., Sokol, S., Palamarczyk, G. & François, J. (2006) Investigating the caffeine effects in the yeast Saccharomyces cerevisiae brings new insights into the connection between TOR, PKC and Ras/cAMP signalling pathways. *Molecular microbiology*. 61 (5), 1147–1166. doi:10.1111/J.1365-2958.2006.05300.X.

Kurat, C.F., Yeeles, J.T.P., Patel, H., Early, A. & Diffley, J.F.X. (2017) Chromatin Controls DNA Replication Origin Selection, Lagging-Strand Synthesis, and Replication Fork Rates. *Molecular Cell*. 65 (1), 117. doi:10.1016/J.MOLCEL.2016.11.016.

Moon K.Y., Kong D., Lee J.K., Raychaudhuri S. & Hurwitz J. (1999) Identification and reconstitution of the origin recognition complex from Schizosaccharomyces pombe. *Proceedings of the National Academy of Sciences of the United States of America*. 96 (22), 12367–12372. doi:10.1073/PNAS.96.22.12367.

Labib, K. (2010) How do Cdc7 and cyclin-dependent kinases trigger the initiation of chromosome replication in eukaryotic cells? *Genes & Development*. 24 (12), 1208. doi:10.1101/GAD.1933010.

Labib, K. & Hodgson, B. (2007) Replication fork barriers: pausing for a break or stalling for time? *EMBO Reports*. 8 (4), 346. doi:10.1038/SJ.EMBOR.7400940.

Labib, K., Tercero, J.A. & Diffley, J.F.X. (2000) Uninterrupted MCM2-7 Function Required for DNA Replication Fork Progression. *Science*. 288 (5471), 1643–1647. doi:10.1126/SCIENCE.288.5471.1643.

Lang, K.S. & Merrikh, H. (2021) Topological stress is responsible for the detrimental outcomes of head-on replication-transcription conflicts. *Cell reports*. 34 (9), 108797. doi:10.1016/J.CELREP.2021.108797.

Langston, L.D., Zhang, D., Yurieva, O., Georgescu, R.E., Finkelstein, J., Yao, N.Y., Indiani, C. & O'Donnell, M.E. (2014) CMG helicase and DNA polymerase ε form a functional 15-subunit holoenzyme for eukaryotic leading-strand DNA replication. *Proceedings of the National Academy of Sciences*. 111 (43), 15390–15395. doi:10.1073/PNAS.1418334111.

Lanz, M.C., Dibitetto, D. & Smolka, M.B. (2019) DNA damage kinase signaling: checkpoint and repair at 30 years. *The EMBO Journal*. 38 (18), e101801. doi:10.15252/EMBJ.2019101801.

Lara-Gonzalez, P., Westhorpe, F.G. & Taylor, S.S. (2012) The Spindle Assembly Checkpoint. *Current Biology*. 22 (22), R966–R980. doi:10.1016/J.CUB.2012.10.006.

Larasati & Duncker, B.P. (2017) Mechanisms Governing DDK Regulation of the Initiation of DNA Replication. *Genes*. 8 (1). doi:10.3390/GENES8010003.

Larizza, L., Roversi, G. & Volpi, L. (2010) Rothmund-Thomson syndrome. *Orphanet Journal of Rare Diseases*. 5 (1), 2. doi:10.1186/1750-1172-5-2.

Laura N. Rusche, Ann L. Kirchmaier & Rine, J. (2003) The Establishment, Inheritance, and Function of Silenced Chromatin in Saccharomyces cerevisiae. *http://dx.doi.org/10.1146/annurev.biochem.72.121801.161547*. 72, 481–516. doi:10.1146/ANNUREV.BIOCHEM.72.121801.161547.

Laurenson, P. & Rine, J. (1992) Silencers, silencing, and heritable transcriptional states. *Microbiological Reviews*. 56 (4), 543. /pmc/articles/PMC372887/?report=abstract.

Laurent, B.C., Treich, I. & Carlson, M. (1993) The yeast SNF2/SWI2 protein has DNA-stimulated ATPase activity required for transcriptional activation. *Genes & development*. 7 (4), 583–591. doi:10.1101/GAD.7.4.583.

Laurent, B.C., Yang, X. & Carlson, M. (1992) An essential Saccharomyces cerevisiae gene homologous to SNF2 encodes a helicase-related protein in a new family. *Molecular and Cellular Biology*. 12 (4), 1893. doi:10.1128/MCB.12.4.1893.

Lee, C., Liachko, I., Bouten, R., Kelman, Z. & Tye, B.K. (2010) Alternative mechanisms for coordinating polymerase alpha and MCM helicase. *Molecular and cellular biology*. 30 (2), 423–435. doi:10.1128/MCB.01240-09.

Lee, J., Choi, E.S. & Lee, D. (2018) It's fun to transcribe with Fun30: A model for nucleosome dynamics during RNA polymerase II-mediated elongation. *Transcription*. 9 (2), 108–116. doi:10.1080/21541264.2017.1356434.

Lee, J., Shik Choi, E., David Seo, H., Kang, K., Gilmore, J.M., Florens, L., Washburn, M.P., Choe, J., Workman, J.L. & Lee, D. (2017) Chromatin remodeller Fun30Fft3 induces nucleosome disassembly to facilitate RNA polymerase II elongation. *Nature Communications 2017 8:1*. 8 (1), 1–13. doi:10.1038/ncomms14527.

Lee, M.G. & Nurse, P. (1987) Complementation used to clone a human homologue of the fission yeast cell cycle control gene cdc2. *Nature*. 327 (6117), 31–35. doi:10.1038/327031a0.

Lee, S.E., Pellicioli, A., Vaze, M.B., Sugawara, N., Malkova, A., Foiani, M. & Haber, J.E. (2003) Yeast Rad52 and Rad51 recombination proteins define a second pathway of DNA damage assessment in response to a single double-strand break. *Molecular and cellular biology*. 23 (23), 8913–8923. doi:10.1128/MCB.23.23.8913-8923.2003.

Lee, W., St.Onge, R.P., Proctor, M., Flaherty, P., Jordan, M.I., Arkin, A.P., Davis, R.W., Nislow, C. & Giaever, G. (2005) Genome-wide requirements for resistance to functionally distinct DNA-damaging agents. *PLoS genetics*. 1 (2), 0235–0246. doi:10.1371/JOURNAL.PGEN.0010024.

Lehman, I.R., Bessman, M.J., Simms, E.S. & Kornberg, A. (1958) Enzymatic synthesis of deoxyribonucleic acid. I. Preparation of substrates and partial purification of an enzyme from Escherichia coli. *The Journal of biological chemistry*. 233 (1), 163–170.

Lehmans, I.R. & Kagunisll, L.S. (1989) DNA Polymerase alpha. *THE JOURNAL OF BIOLOGICAL CHEMISTRY*. 264 (8), 4268. doi:10.1016/S0021-9258(18)83733-4.

Lei, S. & Tu, B.P. (2013) Acetyl-CoA induces transcription of the key G1 cyclin CLN3 to promote entry into the cell division cycle in Saccharomyces cerevisiae. *Proceedings of the National Academy of Sciences of the United States of America*. 110 (18), 7318–7323. doi:10.1073/PNAS.1302490110/-/DCSUPPLEMENTAL/PNAS.201302490SI.PDF.

Leissing, F., Misch, N. v, Wang, X., Werner, L., Huang, L., Conrath, U. & Beckers, G.J.M. (2021) Purification of MAP–kinase protein complexes and identification of candidate components by XL–TAP–MS. *Plant Physiology*. 187 (4), 2381–2392. doi:10.1093/PLPHYS/KIAB446.

Lengronne, A., Pasero, P., Bensimon, A. & Schwob, E. (2001) Monitoring S phase progression globally and locally using BrdU incorporation in TK+ yeast strains. *Nucleic Acids Research*. 29 (7), 1433–1442. doi:10.1093/NAR/29.7.1433.

Leung, G.P., Aristizabal, M.J., Krogan, N.J. & Kobor, M.S. (2014) Conditional genetic interactions of RTT107, SLX4, and HRQ1 reveal dynamic networks upon DNA damage in S. cerevisiae. *G3: Genes, Genomes, Genetics*. 4 (6), 1059–1069. doi:10.1534/G3.114.011205/-/DC1/FIGURES1.PDF.

Li, D., Zhao, R., Lilyestrom, W., Gai, D., Zhang, R., DeCaprio, J.A., Fanning, E., Jochimiak, A., Szakonyi, G. & Chen, X.S. (2003) Structure of the replicative helicase of the oncoprotein SV40 large tumour antigen. *Nature*. 423 (6939), 512–518. doi:10.1038/nature01691.

Li, H. & Stillman, B. (2012) The origin recognition complex: a biochemical and structural view. *Sub-cellular biochemistry*. 62, 37. doi:10.1007/978-94-007-4572-8_3.

Li, N., Lam, W.H., Zhai, Y., Cheng, J., Cheng, E., Zhao, Y., Gao, N. & Tye, B.K. (2018) Structure of the origin recognition complex bound to DNA replication origin. *Nature 2018 559:7713*. 559 (7713), 217–222. doi:10.1038/s41586-018-0293-x.

Li, N., Zhai, Y., Zhang, Y., Li, W., Yang, M., Lei, J., Tye, B.-K. & Gao, N. (2015) Structure of the eukaryotic MCM complex at 3.8 Å. *Nature*. 524, 186. https://doi.org/10.1038/nature14685.

Li, Z., Adams, R.M., Chourey, K., Hurst, G.B., Hettich, R.L. & Pan, C. (2012) Systematic comparison of label-free, metabolic labeling, and isobaric chemical labeling for quantitative proteomics on LTQ orbitrap velos. *Journal of Proteome Research*. 11 (3), 1582–1590. doi:10.1021/PR200748H/SUPPL_FILE/PR200748H_SI_004.PDF.

Liang, C. & Stillman, B. (1997) Persistent initiation of DNA replication and chromatin-bound MCM proteins during the cell cycle in cdc6 mutants. *Genes & Development*. 11 (24), 3375. doi:10.1101/GAD.11.24.3375.

Liang, C., Weinreich, M. & Stillman, B. (1995) ORC and Cdc6p interact and determine the frequency of initiation of DNA replication in the genome. *Cell*. 81 (5), 667–676. doi:10.1016/0092-8674(95)90528-6.

Liang, F. & Wang, Y. (2007) DNA damage checkpoints inhibit mitotic exit by two different mechanisms. *Molecular and cellular biology*. 27 (14), 5067–5078. doi:10.1128/MCB.00095-07.

Liao, J.C., Jeong, Y.J., Kim, D.E., Patel, S.S., Oster, G. & Kim, Y.J. (2005) Mechanochemistry of t7 DNA helicase. *Journal of molecular biology*. 350 (3), 452–475. doi:10.1016/J.JMB.2005.04.051.

Linger, J. & Tyler, J.K. (2005) The Yeast Histone Chaperone Chromatin Assembly Factor 1 Protects Against Double-Strand DNA-Damaging Agents. *Genetics*. 171 (4), 1513. doi:10.1534/GENETICS.105.043000.

Liou, G.G., Tanny, J.C., Kruger, R.G., Walz, T. & Moazed, D. (2005) Assembly of the SIR Complex and Its Regulation by O-Acetyl-ADP-Ribose, a Product of NAD-Dependent Histone Deacetylation. *Cell*. 121 (4), 515–527. doi:10.1016/J.CELL.2005.03.035.

Lipford, J.R. & Bell, S.P. (2001) Nucleosomes Positioned by ORC Facilitate the Initiation of DNA Replication. *Molecular Cell*. 7 (1), 21–30. doi:10.1016/S1097-2765(01)00151-4.

Liu, S., Xu, Z., Leng, H., Zheng, P., Yang, J., Chen, K., Feng, J. & Li, Q. (2017a) RPA binds histone H3-H4 and functions in DNA replication-coupled nucleosome assembly. *Science (New York, N.Y.)*. 355 (6323), 415–420. doi:10.1126/SCIENCE.AAH4712.

Liu, Y., Cussiol, J.R., Dibitetto, D., Sims, J.R., Twayana, S., Weiss, R.S., Freire, R., Marini, F., Pellicioli, A. & Smolka, M.B. (2017b) TOPBP1Dpb11 plays a conserved role in homologous recombination DNA repair through the coordinated recruitment of 53BP1Rad9. *The Journal of Cell Biology*. 216 (3), 623. doi:10.1083/JCB.201607031.

Lo, C.S.Y., van Toorn, M., Gaggioli, V., Dias, M.P., Zhu, Y., et al. (2021) SMARCAD1-mediated active replication fork stability maintains genome integrity. *Science Advances*. 7 (19). doi:10.1126/SCIADV.ABE7804.

Loewith, R., Jacinto, E., Wullschleger, S., Lorberg, A., Crespo, J.L., Bonenfant, D., Oppliger, W., Jenoe, P. & Hall, M.N. (2002) Two TOR complexes, only one of which is rapamycin sensitive, have distinct roles in cell growth control. *Molecular cell*. 10 (3), 457–468. doi:10.1016/S1097-2765(02)00636-6.

Lohka, M.J., Hayes, M.K. & Maller, J.L. (1988) Purification of maturation-promoting factor, an intracellular regulator of early mitotic events. *Proceedings of the National Academy of Sciences*. 85 (9), 3009 LP – 3013. doi:10.1073/pnas.85.9.3009.

Longhese, M.P., Paciotti, V., Fraschini, R., Zaccarini, R., Plevani, P. & Lucchini, G. (1997) The novel DNA damage checkpoint protein ddc1p is phosphorylated periodically during the cell cycle and in response to DNA damage in budding yeast. *The EMBO journal*. 16 (17), 5216–5226. doi:10.1093/EMBOJ/16.17.5216.

Lõoke, M., Maloney, M.F. & Bell, S.P. (2017) Mcm10 regulates DNA replication elongation by stimulating the CMG replicative helicase. *Genes & Development*. 31 (3), 291–305. doi:10.1101/GAD.291336.116.

Lopes, M., Foiani, M. & Sogo, J.M. (2006) Multiple mechanisms control chromosome integrity after replication fork uncoupling and restart at irreparable UV lesions. *Molecular cell*. 21 (1), 15–27. doi:10.1016/J.MOLCEL.2005.11.015.

Lopez-Mosqueda, J., Maas, N.L., Jonsson, Z.O., Defazio-Eli, L.G., Wohlschlegel, J. & Toczyski, D.P. (2010) Damage-induced phosphorylation of Sld3 is important to block late origin firing. *Nature*. 467 (7314), 479–483. doi:10.1038/NATURE09377.

Lu, K., Ye, W., Zhou, L., Collins, L.B., Chen, X., Gold, A., Ball, L.M. & Swenberg, J.A. (2010) Structural characterization of formaldehyde-induced cross-links between amino acids and deoxynucleosides and their oligomers. *Journal of the American Chemical Society*. 132 (10), 3388–3399. doi:10.1021/JA908282F/SUPPL_FILE/JA908282F_SI_001.PDF.

Luke, B., Versini, G., Jaquenoud, M., Zaidi, I.W., Kurz, T., Pintard, L., Pasero, P. & Peter, M. (2006) The cullin Rtt101p promotes replication fork progression through damaged DNA and natural pause sites. *Current biology : CB*. 16 (8), 786–792. doi:10.1016/J.CUB.2006.02.071.

Lynch, K.L., Alvino, G.M., Kwan, E.X., Brewer, B.J. & Raghuraman, M.K. (2019) The effects of manipulating levels of replication initiation factors on origin firing efficiency in yeast. *PLOS Genetics*.

Macris, M.A., Krejci, L., Bussen, W., Shimamoto, A. & Sung, P. (2006) Biochemical characterization of the RECQ4 protein, mutated in Rothmund-Thomson syndrome. *DNA Repair*. 5 (2), 172–180. doi:10.1016/J.DNAREP.2005.09.005.

Maculins, T., Nkosi, P.J., Nishikawa, H. & Labib, K. (2015) Tethering of SCFDia2 to the Replisome Promotes Efficient Ubiquitylation and Disassembly of the CMG Helicase. *Current Biology*. 25 (17), 2254. doi:10.1016/J.CUB.2015.07.012.

Maga, G., Frouin, I., Spadari, S. & Hübscher, U. (2001) Replication protein A as a "fidelity clamp" for DNA polymerase alpha. *The Journal of biological chemistry*. 276 (21), 18235–18242. doi:10.1074/JBC.M009599200.

Mahbubani, H.M., Chong, J.P.J., Chevalier, S., Thömmes, P. & Blow, J.J. (1997) Cell Cycle Regulation of the Replication Licensing System: Involvement of a Cdk-dependent Inhibitor. *The Journal of Cell Biology*. 136 (1), 125. doi:10.1083/JCB.136.1.125.

Maine, G.T., Sinha, P. & Tye, B.-K. (1984) MUTANTS OF S. CEREVISIAE DEFECTIVE IN THE MAINTENANCE OF MINICHROMOSOMES. *Genetics*. 106 (3).

Maiorano, D., Moreau, J. & Méchali, M. (2000) XCDT1 is required for the assembly of pre-replicative complexes in Xenopus laevis. *Nature*. 404, 622. https://doi.org/10.1038/35007104.

Majka, J., Niedziela-Majka, A. & Burgers, P.M.M.J. (2006) The checkpoint clamp activates Mec1 kinase during initiation of the DNA damage checkpoint. *Molecular cell*. 24 (6), 891–901. doi:10.1016/J.MOLCEL.2006.11.027.

Makovets, S., Herskowitz, I. & Blackburn, E.H. (2004) Anatomy and Dynamics of DNA Replication Fork Movement in Yeast Telomeric Regions. *Molecular and Cellular Biology*. 24 (9), 4019. doi:10.1128/MCB.24.9.4019-4031.2004.

van Maldergem, L., Siitonen, H.A., Jalkh, N., Chouery, E., de Roy, M., Delague, V., Muenke, M., Jabs, E.W., Cai, J., Wang, L.L., Plon, S.E., Fourneau, C., Kestilä, M., Gillerot, Y., Mégarbané, A. & Verloes, A. (2006) Revisiting the craniosynostosis-radial ray hypoplasia association: Baller-Gerold syndrome caused by mutations in the RECQL4 gene. *Journal of medical genetics*. 43 (2), 148–152. doi:10.1136/JMG.2005.031781.

Mandell, J.G., Goodrich, K.J., Bähler, J. & Cech, T.R. (2005) Expression of a RecQ Helicase Homolog Affects Progression through Crisis in Fission Yeast Lacking Telomerase. *Journal of Biological Chemistry*. 280 (7), 5249–5257. doi:10.1074/JBC.M412756200.

Mantiero, D., MacKenzie, A., Donaldson, A. & Zegerman, P. (2011) Limiting replication initiation factors execute the temporal programme of origin firing in budding yeast. *The EMBO Journal*. 30 (23), 4805. doi:10.1038/EMBOJ.2011.404.

Mao, X., Hu, Y., Liang, C. & Lu, C. (2002) MET3 promoter: a tightly regulated promoter and its application in construction of conditional lethal strain. *Current microbiology*. 45 (1), 37–40. doi:10.1007/S00284-001-0046-0.

Marahrens, Y. & Stillman, B. (1992) A yeast chromosomal origin of DNA replication defined by multiple functional elements. *Science (New York, N.Y.)*. 255 (5046), 817–823. doi:10.1126/SCIENCE.1536007.

Maréchal, A. & Zou, L. (2015) RPA-coated single-stranded DNA as a platform for post-translational modifications in the DNA damage response. *Cell research*. 25 (1), 9–23. doi:10.1038/CR.2014.147.

Maric, M., Maculins, T., de Piccoli, G. & Labib, K. (2014) Cdc48 and a ubiquitin ligase drive disassembly of the CMG helicase at the end of DNA replication. *Science (New York, N.Y.)*. 346 (6208), 1253596. doi:10.1126/SCIENCE.1253596.

Maric, M., Mukherjee, P., Tatham, M.H., Hay, R. & Labib, K. (2017) Ufd1-Npl4 Recruit Cdc48 for Disassembly of Ubiquitylated CMG Helicase at the End of Chromosome Replication. *Cell Reports*. 18 (13), 3033–3042. doi:10.1016/J.CELREP.2017.03.020.

Marsischky, G.T., Filosi, N., Kane, M.F. & Kolodner, R. (1996) Redundancy of Saccharomyces cerevisiae MSH3 and MSH6 in MSH2-dependent mismatch repair. *Genes & development*. 10 (4), 407–420. doi:10.1101/GAD.10.4.407.

Massy, B. de, Béjar, S., Louarn, J., Louarn, J.M. & Bouché, J.P. (1987) Inhibition of replication forks exiting the terminus region of the Escherichia coli chromosome occurs at two loci separated by 5 min. *Proceedings of the National Academy of Sciences*. 84 (7), 1759–1763. doi:10.1073/PNAS.84.7.1759.

Masui, Y. & Markert, C.L. (1971) Cytoplasmic control of nuclear behavior during meiotic maturation of frog oocytes. *The Journal of experimental zoology*. 177 (2), 129–145. doi:10.1002/jez.1401770202.

Masumoto, H., Muramatsu, S., Kamimura, Y. & Araki, H. (2002) S-Cdk-dependent phosphorylation of Sld2 essential for chromosomal DNA replication in budding yeast. *Nature*. 415 (6872), 651–655. doi:10.1038/NATURE713.

Masumoto, H., Sugino, A. & Araki, H. (2000) Dpb11 controls the association between DNA polymerases alpha and epsilon and the autonomously replicating sequence region of budding yeast. *Molecular and cellular biology*. 20 (8), 2809–2817. doi:10.1128/MCB.20.8.2809-2817.2000.

Matellán, L. & Monje-Casas, F. (2020) Regulation of Mitotic Exit by Cell Cycle Checkpoints: Lessons From Saccharomyces cerevisiae. *Genes*. 11 (2). doi:10.3390/GENES11020195.

Mathews, M.B., Bernstein, R.M., Franza, B.R. & Garrels, J.I. (1984) Identity of the proliferating cell nuclear antigen and cyclin. *Nature*. 309 (5966), 374–376. doi:10.1038/309374A0.

Matsuda, K., Makise, M., Sueyasu, Y., Takehara, M., Asano, T. & Mizushima, T. (2007) Yeast two-hybrid analysis of the origin recognition complex of Saccharomyces cerevisiae: interaction between subunits and identification of binding proteins. *FEMS yeast research*. 7 (8), 1263–1269. doi:10.1111/J.1567-1364.2007.00298.X.

Matsuno, K., Kumano, M., Kubota, Y., Hashimoto, Y. & Takisawa, H. (2006) The N-Terminal Noncatalytic Region of Xenopus RecQ4 Is Required for Chromatin Binding of DNA Polymerase α in the Initiation of DNA Replication. *Molecular and Cellular Biology*. 26 (13), 4843. doi:10.1128/MCB.02267-05.

Mattarocci, S., Shyian, M., Lemmens, L., Damay, P., Altintas, D.M., Shi, T., Bartholomew, C.R., Thomä, N.H., Hardy, C.F.J. & Shore, D. (2014) Rif1 controls DNA replication timing in yeast through the PP1 phosphatase Glc7. *Cell reports*. 7 (1), 62–69. doi:10.1016/J.CELREP.2014.03.010.

Matzinger, M. & Mechtler, K. (2021) Cleavable Cross-Linkers and Mass Spectrometry for the Ultimate Task of Profiling Protein-Protein Interaction Networks in Vivo. *Journal of proteome research*. 20 (1), 78–93. doi:10.1021/ACS.JPROTEOME.0C00583.

Mayle, R., Langston, L., Molloy, K.R., Zhang, D., Chait, B.T. & O'Donnell, M.E. (2019) Mcm10 has potent strand-annealing activity and limits translocase-mediated fork regression. *Proceedings of the National Academy of Sciences of the United States of America*. 116 (3), 798–803. doi:10.1073/PNAS.1819107116/VIDEO-1.

McGeoch, A.T. & Bell, S.D. (2005) Eukaryotic/archaeal primase and MCM proteins encoded in a bacteriophage genome. *Cell*. 120 (2), 167–168. doi:10.1016/J.CELL.2004.11.031/ATTACHMENT/8D5AC668-5931-4783-89B1-DC9DBD8F3980/MMC3.PDF.

McGeoch, A.T., Trakselis, M.A., Laskey, R.A. & Bell, S.D. (2005) Organization of the archaeal MCM complex on DNA and implications for the helicase mechanism. *Nature Structural & Molecular Biology 2005 12:9*. 12 (9), 756–762. doi:10.1038/nsmb974.

McGhee, J.D. & Hippel, P.H. von (1975) Formaldehyde as a probe of DNA structure. I. Reaction with exocyclic amino groups of DNA bases. *Biochemistry*. 14 (6), 1281–1296. doi:10.1021/bi00677a029.

McGuffee, S.R., Smith, D.J. & Whitehouse, I. (2013) Quantitative, genome-wide analysis of eukaryotic replication initiation and termination. *Molecular cell*. 50 (1), 123–135. doi:10.1016/J.MOLCEL.2013.03.004.

McIntosh, D. & Blow, J.J. (2012) Dormant origins, the licensing checkpoint, and the response to replicative stresses. *Cold Spring Harbor perspectives in biology*. 4 (10). doi:10.1101/CSHPERSPECT.A012955.

McQuibban, G.A., Saurya, S. & Freeman, M. (2003) Mitochondrial membrane remodelling regulated by a conserved rhomboid protease. *Nature*. 423 (6939), 537–541. doi:10.1038/NATURE01633.

Meagher, M., Epling, L.B. & Enemark, E.J. (2019) DNA translocation mechanism of the MCM complex and implications for replication initiation. *Nature Communications 2019 10:1*. 10 (1), 1–13. doi:10.1038/s41467-019-11074-3.

Méchali, M. (2010) Eukaryotic DNA replication origins: many choices for appropriate answers. *Nature Reviews Molecular Cell Biology 2010 11:10*. 11 (10), 728–738. doi:10.1038/nrm2976.

Mellacheruvu, D., Wright, Z., Couzens, A.L., Lambert, J.P., St-Denis, N.A., et al. (2013) The CRAPome: a contaminant repository for affinity purification–mass spectrometry data. *Nature Methods 2013 10:8*. 10 (8), 730–736. doi:10.1038/nmeth.2557.

Merchant, A.M., Kawasaki, Y., Chen, Y., Lei, M. & Tye, B.K. (1997) A lesion in the DNA replication initiation factor Mcm10 induces pausing of elongation forks through chromosomal replication origins in Saccharomyces cerevisiae. *Molecular and cellular biology*. 17 (6), 3261–3271. doi:10.1128/MCB.17.6.3261.

Mermoud, J.E., Rowbotham, S.P. & Varga-Weisz, P.D. (2011) Keeping chromatin quiet: how nucleosome remodeling restores heterochromatin after replication. *Cell cycle (Georgetown, Tex.)*. 10 (23), 4017–4025. doi:10.4161/CC.10.23.18558.

Meselson, M. & Stahl, F.W. (1958) The replication of DNA in Escherichia coli. *Proceedings of the National Academy of Sciences*. 44 (7), 671–682. doi:10.1073/pnas.44.7.671.

Metz, B., Kersten, G.F.A., Hoogerhout, P., Brugghe, H.F., Timmermans, H.A.M., de Jong, A., Meiring, H., ten Hove, J., Hennink, W.E., Crommelin, D.J.A. & Jiskoot, W. (2004) Identification of Formaldehyde-induced Modifications in Proteins: REACTIONS WITH MODEL PEPTIDES. *Journal of Biological Chemistry*. 279 (8), 6235–6243. doi:10.1074/JBC.M310752200.

Milenkovic, D., Matic, S., kühl, I., Ruzzenente, B., Freyer, C., Jemt, E., Park, C.B., Falkenberg, M. & Larsson, N.G. (2013) TWINKLE is an essential mitochondrial helicase required for synthesis of nascent D-loop strands and complete mtDNA replication. *Human Molecular Genetics*. 22 (10), 1983. doi:10.1093/HMG/DDT051.

Miles, J. & Formosa, T. (1992) Protein affinity chromatography with purified yeast DNA polymerase alpha detects proteins that bind to DNA polymerase. *Proceedings of the National Academy of Sciences of the United States of America*. 89 (4), 1276. doi:10.1073/PNAS.89.4.1276.

Miller, T.C.R., Locke, J., Greiwe, J.F., Diffley, J.F.X. & Costa, A. (2019) Mechanism of head-to-head MCM double-hexamer formation revealed by cryo-EM. *Nature*. 575 (7784), 704. doi:10.1038/S41586-019-1768-0.

Mimura, S., Komata, M., Kishi, T., Shirahige, K. & Kamura, T. (2009) SCF(Dia2) regulates DNA replication forks during S-phase in budding yeast. *The EMBO journal*. 28 (23), 3693–3705. doi:10.1038/EMBOJ.2009.320.

Mimura, S., Masuda, T., Matsui, T. & Takisawa, H. (2000) Central role for cdc45 in establishing an initiation complex of DNA replication in Xenopus egg extracts. *Genes to cells : devoted to molecular & cellular mechanisms*. 5 (6), 439–452. doi:10.1046/J.1365-2443.2000.00340.X.

Mimura, S. & Takisawa, H. (1998) Xenopus Cdc45-dependent loading of DNA polymerase alpha onto chromatin under the control of S-phase Cdk. *The EMBO journal*. 17 (19), 5699–5707. doi:10.1093/EMBOJ/17.19.5699.

Mintseris, J. & Gygi, S.P. (2020) High-density chemical cross-linking for modeling protein interactions. *Proceedings of the National Academy of Sciences of the United States of America*. 117 (1), 93–102. doi:10.1073/PNAS.1902931116/-/DCSUPPLEMENTAL.

Mittal, P. & Roberts, C.W.M. (2020) The SWI/SNF complex in cancer — biology, biomarkers and therapy. *Nature Reviews Clinical Oncology 2020 17:7*. 17 (7), 435–448. doi:10.1038/s41571-020-0357-3.

Miyabe, I., Mizuno, K., Keszthelyi, A., Daigaku, Y., Skouteri, M., Mohebi, S., Kunkel, T.A., Murray, J.M. & Carr, A.M. (2015) Polymerase δ replicates both strands after homologous recombination–dependent fork restart. *Nature Structural & Molecular Biology 2015 22:11*. 22 (11), 932–938. doi:10.1038/nsmb.3100.

Miyachi, K., Fritzler, M.J. & Tan, E.M. (1978) Autoantibody to a Nuclear Antigen in Proliferating Cells. *The Journal of Immunology*. 121 (6).

Miyake, S. & Yamashita, S. (1998) Identification of sna41 gene, which is the suppressor of nda4 mutation and is involved in DNA replication in Schizosaccharomyces pombe. *Genes to cells : devoted to molecular & cellular mechanisms*. 3 (3), 157–166. doi:10.1046/J.1365-2443.1998.00177.X.

Miyazawa-Onami, M., Araki, H. & Tanaka, S. (2017) Pre-initiation complex assembly functions as a molecular switch that splits the Mcm2-7 double hexamer. *EMBO Reports*. 18 (10), 1752. doi:10.15252/EMBR.201744206.

Mizushima, T., Takahashi, N. & Stillman, B. (2000) Cdc6p modulates the structure and DNA binding activity of the origin recognition complex in vitro. *Genes & Development*. 14 (13), 1631. /pmc/articles/PMC316732/.

Eaton M.L., Galani K., Kang S., Bell S.P. & MacAlpine D.M. (2010) Conserved nucleosome positioning defines replication origins. *Genes & development*. 24 (8), 748–753. doi:10.1101/GAD.1913210.

Moazed, D., Kistler, A., Axelrod, A., Rine, J. & Johnson, A.D. (1997) Silent information regulator protein complexes in Saccharomyces cerevisiae: A SIR2/SIR4 complex and evidence for a regulatory domain in SIR4 that inhibits its interaction with SIR3. *Proceedings of the National Academy of Sciences*. 94 (6), 2186–2191. doi:10.1073/PNAS.94.6.2186.

Mohammed, H., D'Santos, C., Serandour, A.A., Ali, H.R., Brown, G.D., et al. (2013) Endogenous Purification Reveals GREB1 as a Key Estrogen Receptor Regulatory Factor. *Cell Reports*. 3 (2), 342–349. doi:10.1016/J.CELREP.2013.01.010.

Mohanty, B.K., Bairwa, N.K. & Bastia, D. (2006) The Tof1p-Csm3p protein complex counteracts the Rrm3p helicase to control replication termination of Saccharomyces cerevisiae. *Proceedings of the National Academy of Sciences of the United States of America*. 103 (4), 897–902. doi:10.1073/PNAS.0506540103.

Moir, D., Stewart, S.E., Osmond, B.C. & Botstein, D. (1982) Cold-Sensitive Cell-Division-Cycle Mutants Of Yeast: Isolation, Properties, And Pseudoreversion Studies. *Genetics*. 100 (4).

Moldovan, G.L., Pfander, B. & Jentsch, S. (2007) PCNA, the Maestro of the Replication Fork. *Cell*. 129 (4), 665–679. doi:10.1016/J.CELL.2007.05.003/ATTACHMENT/9E31002C-DDA5-4C30-949B-7D62517ACC5B/MMC1.PDF.

Moreno, S.P., Bailey, R., Campion, N., Herron, S. & Gambus, A. (2014) Polyubiquitylation drives replisome disassembly at the termination of DNA replication. *Science*. 346 (6208), 477–481. doi:10.1126/SCIENCE.1253585/SUPPL_FILE/PRIEGOMORENO.SM.PDF.

Moretti, P., Freeman, K., Coodly, L. & Shore, D. (1994) Evidence that a complex of SIR proteins interacts with the silencer and telomere-binding protein RAP1. *Genes & Development*. 8 (19), 2257–2269. doi:10.1101/GAD.8.19.2257.

Morgan, D. 2006. Cell cycle: Principles of Control. New Science Press: London

Morohashi, H., Maculins, T. & Labib, K. (2009) The Amino-Terminal TPR Domain of Dia2 Tethers SCFDia2 to the Replisome Progression Complex. *Current Biology*. 19 (22), 1943–1949. doi:10.1016/J.CUB.2009.09.062/ATTACHMENT/21849697-E46A-401E-BC15-01391B57306E/MMC1.PDF.

Morrison, A.J. & Shen, X. (2009) Chromatin remodelling beyond transcription: the INO80 and SWR1 complexes. *Nature reviews. Molecular cell biology*. 10 (6), 373. doi:10.1038/NRM2693.

Moyer, S.E., Lewis, P.W. & Botchan, M.R. (2006) Isolation of the Cdc45/Mcm2–7/GINS (CMG) complex, a candidate for the eukaryotic DNA replication fork helicase. *Proceedings of the National Academy of Sciences*. 103 (27), 10236–10241. doi:10.1073/PNAS.0602400103.

Mueller, A.C., Keaton, M.A. & Dutta, A. (2011) DNA replication: Mammalian Treslin-TopBP1 interaction mirrors yeast Sld3-Dpb11. *Current biology : CB*. 21 (16), R638. doi:10.1016/J.CUB.2011.07.004.

Muramatsu, S., Hirai, K., Tak, Y.S., Kamimura, Y. & Araki, H. (2010) CDK-dependent complex formation between replication proteins Dpb11, Sld2, Pol (epsilon}, and GINS in budding yeast. *Genes & development*. 24 (6), 602–612. doi:10.1101/GAD.1883410.

Murat, P., Guillaume, G., Sale, J. (2021) Replication origins drive genetic and phenotypic mutation in humans. bioRxiv doi: https://doi.org/10.1101/2021.12.16.472898

Murray, A.W. (1995) The genetics of cell cycle checkpoints. *Current Opinion in Genetics & Development*. 5 (1), 5–11. doi:10.1016/S0959-437X(95)90046-2.

Murray, A.W. & Kirschner, M.W. (1989) Dominoes and clocks: the union of two views of the cell cycle. *Science*. 246 (4930), 614–621. doi:10.1126/science.2683077.

Nakamura, K., Kustatscher, G., Alabert, C., Imhof, A., Rappsilber, J., Groth Correspondence, A., Hö, M., Forne, I., Vö Lker-Albert, M., Satpathy, S., Beyer, T.E., Mailand, N., Choudhary, C. & Groth, A. (2021) Proteome dynamics at broken replication forks reveal a distinct ATM-directed repair response suppressing DNA double-strand break ubiquitination. *Molecular Cell*. 81, 1084-1099.e6. doi:10.1016/j.molcel.2020.12.025.

Nasmyth, K. (1996) Viewpoint: putting the cell cycle in order. *Science (New York, N.Y.)*. 274 (5293), 1643–1645.

Natarajan, K.N. (2013) *Unravelling the proteome of chromatin bound RNA polymerase II using Proteome-ChIP in murine stem cells*. doi:10.25560/14704.

Natsume, T., Müller, C.A., Katou, Y., Retkute, R., Gierliński, M., Araki, H., Blow, J.J., Shirahige, K., Nieduszynski, C.A. & Tanaka, T.U. (2013) Kinetochores coordinate pericentromeric cohesion and early DNA replication by Cdc7-Dbf4 kinase recruitment. *Molecular cell*. 50 (5), 661–674. doi:10.1016/J.MOLCEL.2013.05.011.

Nedelcheva, M.N., Roguev, A., Dolapchiev, L.B., Shevchenko, A., Taskov, H.B., Shevchenko, A., Stewart, A.F. & Stoynov, S.S. (2005) Uncoupling of unwinding from DNA synthesis implies regulation of MCM helicase by Tof1/Mrc1/Csm3 checkpoint complex. *Journal of molecular biology*. 347 (3), 509–521. doi:10.1016/j.jmb.2005.01.041.

Neves-Costa, A., Will, W.R., Vetter, A.T., Miller, J.R. & Varga-Weisz, P. (2009a) The SNF2-Family Member Fun30 Promotes Gene Silencing in Heterochromatic Loci. *PLOS ONE*. 4 (12), e8111. doi:10.1371/JOURNAL.PONE.0008111.

Newlon, C.S. (1988) Yeast chromosome replication and segregation. *Microbiological Reviews*. 52 (4), 568. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC373164/.

Neylon, C., Kralicek, A. v., Hill, T.M. & Dixon, N.E. (2005) Replication termination in Escherichia coli: structure and antihelicase activity of the Tus-Ter complex. *Microbiology and molecular biology reviews : MMBR*. 69 (3), 501–526. doi:10.1128/MMBR.69.3.501-526.2005.

Ngo, G.H.P. & Lydall, D. (2015) The 9-1-1 checkpoint clamp coordinates resection at DNA double strand breaks. *Nucleic Acids Research*. 43 (10), 5017–5032. doi:10.1093/NAR/GKV409.

Nick McElhinny, S.A., Gordenin, D.A., Stith, C.M., Burgers, P.M.J. & Kunkel, T.A. (2008) Division of labor at the eukaryotic replication fork. *Molecular cell*. 30 (2), 137–144. doi:10.1016/J.MOLCEL.2008.02.022.

Nishitani, H., Lygerou, Z., Nishimoto, T. & Nurse, P. (2000) The Cdt1 protein is required to license DNA for replication in fission yeast. *Nature*. 404, 625. https://doi.org/10.1038/35007110.

Noguchi, Y., Yuan, Z., Bai, L., Schneider, S., Zhao, G., Stillman, B., Speck, C. & Li, H. (2017) Cryo-EM structure of Mcm2-7 double hexamer on DNA suggests a lagging-strand DNA extrusion model.

Proceedings of the National Academy of Sciences of the United States of America. 114 (45), E9529–E9538. doi:10.1073/pnas.1712537114.

Nurse, P. (2002) Cyclin dependent kinases and cell cycle control (nobel lecture). *Chembiochem : a European journal of chemical biology*. 3 (7), 596–603. doi:10.1002/1439-7633(20020703)3:7<596::AID-CBIC596>3.0.CO;2-U.

Nurse, P. (1975) Genetic control of cell size at cell division in yeast. *Nature*. 256 (5518), 547–551. doi:10.1038/256547a0.

Nurse, P. & Bissett, Y. (1981) Gene required in G1 for commitment to cell cycle and in G2 for control of mitosis in fission yeast. *Nature*. 292 (5823), 558–560. doi:10.1038/292558a0.

Nurse, P. & Thuriaux, P. (1980) Regulatory genes controlling mitosis in the fission yeast Schizosaccharomyces pombe. *Genetics*. 96 (3), 627–637.

Nurse, P., Thuriaux, P. & Nasmyth, K. (1976) Genetic control of the cell division cycle in the fission yeast Schizosaccharomyces pombe. *Molecular and General Genetics MGG*. 146 (2), 167–178. doi:10.1007/BF00268085.

O'Connor, J.P. & Peebles, C.L. (1992) PTA1, an essential gene of Saccharomyces cerevisiae affecting pre-tRNA processing. *Molecular and cellular biology*. 12 (9), 3843–3856. doi:10.1128/MCB.12.9.3843-3856.1992.

O'Donnell, M., Langston, L. & Stillman, B. (2013) Principles and Concepts of DNA Replication in Bacteria, Archaea, and Eukarya. *Cold Spring Harbor Perspectives in Biology*. 5 (7), a010108. doi:10.1101/CSHPERSPECT.A010108.

O'Donnell, M. & Li, H. (2016) The Eukaryotic Replisome Goes Under the Microscope. *Current Biology*. 26 (6), R247–R256. doi:10.1016/J.CUB.2016.02.034.

O'Donnell, M.E. & Li, H. (2018) The ring-shaped hexameric helicases that function at DNA replication forks. *Nature Structural & Molecular Biology 2018 25:2*. 25 (2), 122–130. doi:10.1038/s41594-018-0024-x.

Olins, A.L. & Olins, D.E. (1974) Spheroid chromatin units (v bodies). *Science (New York, N.Y.)*. 183 (4122), 330–332. doi:10.1126/SCIENCE.183.4122.330.

Ong, S.E., Blagoev, B., Kratchmarova, I., Kristensen, D.B., Steen, H., Pandey, A. & Mann, M. (2002) Stable Isotope Labeling by Amino Acids in Cell Culture, SILAC, as a Simple and Accurate Approach

to Expression Proteomics *. *Molecular & Cellular Proteomics*. 1 (5), 376–386. doi:10.1074/MCP.M200025-MCP200.

Onishi, M., Liou, G.G., Buchberger, J.R., Walz, T. & Moazed, D. (2007) Role of the conserved Sir3-BAH domain in nucleosome binding and silent chromatin assembly. *Molecular cell*. 28 (6), 1015–1028. doi:10.1016/J.MOLCEL.2007.12.004.

Ooi, S.L., Shoemaker, D.D. & Boeke, J.D. (2003) DNA helicase gene interaction network defined using synthetic lethality analyzed by microarray. *Nature genetics*. 35 (3), 277–286. doi:10.1038/NG1258.

O'Reilly, F.J. & Rappsilber, J. (2018) Cross-linking mass spectrometry: methods and applications in structural, molecular and systems biology. *Nature Structural & Molecular Biology 2018 25:11*. 25 (11), 1000–1008. doi:10.1038/s41594-018-0147-0.

Osborn, A.J. & Elledge, S.J. (2003) Mrc1 is a replication fork component whose phosphorylation in response to DNA replication stress activates Rad53. *Genes & development*. 17 (14), 1755–1767. doi:10.1101/GAD.1098303.

Osborn, A.J., Elledge, S.J. & Zou, L. (2002) Checking on the fork: the DNA-replication stress-response pathway. *Trends in Cell Biology*. 12 (11), 509–516. doi:10.1016/S0962-8924(02)02380-2.

Oshiro, G., Owens, J.C., Shellman, Y., Sclafani, R.A. & Li, J.J. (1999) Cell cycle control of Cdc7p kinase activity through regulation of Dbf4p stability. *Molecular and cellular biology*. 19 (7), 4888–4896. doi:10.1128/MCB.19.7.4888.

Ouspenski, I.I., Elledge, S.J. & Brinkley, B.R. (1999) New yeast genes important for chromosome integrity and segregation identified by dosage effects on genome stability. *Nucleic Acids Research*. 27 (15), 3001–3008. doi:10.1093/NAR/27.15.3001.

Pacek, M. & Walter, J.C. (2004) A requirement for MCM7 and Cdc45 in chromosome unwinding during eukaryotic DNA replication. *The EMBO journal*. 23 (18), 3667–3676. doi:10.1038/SJ.EMBOJ.7600369.

Palzkill, T.G. & Newlon, C.S. (1988) A yeast replication origin consists of multiple copies of a small conserved sequence. *Cell*. 53 (3), 441–450. doi:10.1016/0092-8674(88)90164-X.

Pantazopoulou, M., Boban, M., Foisner, R. & Ljungdahl, P.O. (2016) Cdc48 and Ubx1 participate in a pathway associated with the inner nuclear membrane that governs Asi1 degradation. *Journal of cell science*. 129 (20), 3770–3780. doi:10.1242/JCS.189332.

Pape, T., Meka, H., Chen, S., Vicentini, G., van Heel, M. & Onesti, S. (2003) Hexameric ring structure of the full-length archaeal MCM protein complex. *EMBO reports*. 4 (11), 1079–1083. doi:10.1038/sj.embor.embor7400010.

Pappas, D.L., Frisch, R. & Weinreich, M. (2004) The NAD+-dependent Sir2p histone deacetylase is a negative regulator of chromosomal DNA replication. *Genes & Development*. 18 (7), 769–781. doi:10.1101/GAD.1173204.

Pardee, A.B. (1974) A restriction point for control of normal animal cell proliferation. *Proceedings of the National Academy of Sciences of the United States of America*. 71 (4), 1286–1290. doi:10.1073/PNAS.71.4.1286.

Pardo, B., Crabbé, L. & Pasero, P. (2017) Signaling pathways of replication stress in yeast. *FEMS Yeast Research*. 17 (2), 101. doi:10.1093/FEMSYR/FOW101.

Pardo, B. & Marcand, S. (2005) Rap1 prevents telomere fusions by nonhomologous end joining. *The EMBO Journal*. 24 (17), 3117–3127. doi:10.1038/SJ.EMBOJ.7600778.

Park, P.J. (2009) ChIP-Seq: advantages and challenges of a maturing technology. *Nature reviews. Genetics*. 10 (10), 669. doi:10.1038/NRG2641.

Pasero, P., Braguglia, D. & Gasser, S.M. (1997) ORC-dependent and origin-specific initiation of DNA replication at defined foci in isolated yeast nuclei. *Genes & development*. 11 (12), 1504–1518. doi:10.1101/GAD.11.12.1504.

Paulovich, A.G. & Hartwell, L.H. (1995) A checkpoint regulates the rate of progression through S phase in S. cerevisiae in response to DNA damage. *Cell*. 82 (5), 841–847. doi:10.1016/0092-8674(95)90481-6.

Pellegrini, L. (2017) Structural insights into Cdc45 function: was there a nuclease at the heart of the ancestral replisome? *Biophysical Chemistry*. 225, 10. doi:10.1016/J.BPC.2016.11.011.

Pereira, G. & Schiebel, E. (2005) Kin4 kinase delays mitotic exit in response to spindle alignment defects. *Molecular cell*. 19 (2), 209–221. doi:10.1016/J.MOLCEL.2005.05.030.

Perkins, G. & Diffley, J.F.X. (1998) Nucleotide-Dependent Prereplicative Complex Assembly by Cdc6p, a Homolog of Eukaryotic and Prokaryotic Clamp-Loaders. *Molecular Cell*. 2 (1), 23–32. doi:10.1016/S1097-2765(00)80110-0.

Peterson, C.L. & Tamkun, J.W. (1995) The SWI-SNF complex: a chromatin remodeling machine? *Trends in biochemical sciences*. 20 (4), 143–146. doi:10.1016/S0968-0004(00)88990-2.

Petes, T.D. & Williamson, D.H. (1975) Fiber autoradiography of replicating yeast DNA. *Experimental Cell Research*. 95 (1), 103–110. doi:10.1016/0014-4827(75)90614-X.

Petojevic, T., Pesavento, J.J., Costa, A., Liang, J., Wang, Z., Berger, J.M. & Botchan, M.R. (2015) Cdc45 (cell division cycle protein 45) guards the gate of the Eukaryote Replisome helicase stabilizing leading strand engagement. *Proceedings of the National Academy of Sciences*. 112 (3), E249–E258. doi:10.1073/PNAS.1422003112.

Petryk, N., Dalby, M., Wenger, A., Stromme, C.B., Strandsby, A., Andersson, R. & Groth, A. (2018) MCM2 promotes symmetric inheritance of modified histones during DNA replication. *Science (New York, N.Y.)*. 361 (6409), 1389–1392. doi:10.1126/SCIENCE.AAU0294.

Pfander, B. & Diffley, J.F.X. (2011) Dpb11 coordinates Mec1 kinase activation with cell cycle-regulated Rad9 recruitment. *The EMBO journal*. 30 (24), 4897–4907. doi:10.1038/EMBOJ.2011.345.

de Piccoli, G., Katou, Y., Itoh, T., Nakato, R., Shirahige, K. & Labib, K. (2012) Replisome stability at defective DNA replication forks is independent of S phase checkpoint kinases. *Molecular cell*. 45 (5), 696–704. doi:10.1016/J.MOLCEL.2012.01.007.

Piggott, J.R., Rai, R. & Carter, B.L.A. (1982) A bifunctional gene product involved in two phases of the yeast cell cycle. *Nature*. 298 (5872), 391–393. doi:10.1038/298391a0.

Polevoda, B. & Sherman, F. (2003) N-terminal Acetyltransferases and Sequence Requirements for N-terminal Acetylation of Eukaryotic Proteins. *Journal of Molecular Biology*. 325 (4), 595–622. doi:10.1016/S0022-2836(02)01269-X.

Pommier, Y., Pourquier, P., Fan, Y. & Strumberg, D. (1998) Mechanism of action of eukaryotic DNA topoisomerase I and drugs targeted to the enzyme. *Biochimica et Biophysica Acta (BBA) - Gene Structure and Expression*. 1400 (1–3), 83–106. doi:10.1016/S0167-4781(98)00129-8.

Porcella, S.Y., Koussa, N.C., Tang, C.P., Kramer, D.N., Srivastava, P. & Smith, D.J. (2020) Separable, Ctf4-mediated recruitment of DNA Polymerase α for initiation of DNA synthesis at replication origins and lagging-strand priming during replication elongation. *PLOS Genetics*. 16 (5), e1008755. doi:10.1371/JOURNAL.PGEN.1008755.

Pray-Grant, M.G., Daniel, J.A., Schieltz, D., Yates, J.R. & Grant, P.A. (2005) Chd1 chromodomain links histone H3 methylation with SAGA- and SLIK-dependent acetylation. *Nature 2004 433:7024*. 433 (7024), 434–438. doi:10.1038/nature03242.

Prelich, G., Tan, C.K., Kostura, M., Mathews, M.B., So, A.G., Downey, K.M. & Stillman, B. (1987) Functional identity of proliferating cell nuclear antigen and a DNA polymerase-delta auxiliary protein. *Nature*. 326 (6112), 517–520. doi:10.1038/326517A0.

Prioleau, M.N. & MacAlpine, D.M. (2016) DNA replication origins—where do we begin? *Genes & Development*. 30 (15), 1683. doi:10.1101/GAD.285114.116.

Proffitt, J.H., Davie, J.R., Swinton, D. & Hattman, S. (1984) 5-Methylcytosine is not detectable in Saccharomyces cerevisiae DNA. *Molecular and Cellular Biology*. 4 (5), 985–988. doi:10.1128/MCB.4.5.985-988.1984.

Prorok, P., Artufel, M., Aze, A., Coulombe, P., Peiffer, I., Lacroix, L., Guédin, A., Mergny, J.L., Damaschke, J., Schepers, A., Ballester, B. & Méchali, M. (2019) Involvement of G-quadruplex regions in mammalian replication origin activity. *Nature Communications 2019 10:1*. 10 (1), 1–16. doi:10.1038/s41467-019-11104-0.

Puddu, F., Granata, M., Nola, L. di, Balestrini, A., Piergiovanni, G., Lazzaro, F., Giannattasio, M., Plevani, P. & Muzi-Falconi, M. (2008) Phosphorylation of the Budding Yeast 9-1-1 Complex Is Required for Dpb11 Function in the Full Activation of the UV-Induced DNA Damage Checkpoint. *Molecular and Cellular Biology*. 28 (15), 4782. doi:10.1128/MCB.00330-08.

Pursell, Z.F., Isoz, I., Lundström, E.B., Johansson, E. & Kunkel, T.A. (2007) Yeast DNA polymerase epsilon participates in leading-strand DNA replication. *Science (New York, N.Y.)*. 317 (5834), 127–130. doi:10.1126/SCIENCE.1144067.

Pursell, Z.F. & Kunkel, T.A. (2008) DNA Polymerase ε: A Polymerase of Unusual Size (and Complexity). *Progress in Nucleic Acid Research and Molecular Biology*. 82, 101–145. doi:10.1016/S0079-6603(08)00004-4.

Radman-Livaja, M., Ruben, G., Weiner, A., Friedman, N., Kamakaka, R. & Rando, O.J. (2011) Dynamics of Sir3 spreading in budding yeast: secondary recruitment sites and euchromatic localization. *The EMBO Journal*. 30 (6), 1012. doi:10.1038/EMBOJ.2011.30.

Raghuraman, M.K., Brewer, B.J. & Fangman, W.L. (1997) Cell cycle-dependent establishment of a late replication program. *Science (New York, N.Y.)*. 276 (5313), 806–809. doi:10.1126/SCIENCE.276.5313.806.

Raghuraman, M.K., Winzeler, E.A., Collingwood, D., Hunt, S., Wodicka, L., Conway, A., Lockhart, D.J., Davis, R.W., Brewer, B.J. & Fangman, W.L. (2001) Replication Dynamics of the Yeast Genome. *Science*. 294 (5540), 115–121. doi:10.1126/SCIENCE.294.5540.115.

Rajala, N., Hensen, F., Wessels, H.J.C.T., Ives, D., Gloerich, J. & Spelbrink, J.N. (2015) Whole Cell Formaldehyde Cross-Linking Simplifies Purification of Mitochondrial Nucleoids and Associated Proteins Involved in Mitochondrial Gene Expression. *PLOS ONE*. 10 (2), e0116726. doi:10.1371/JOURNAL.PONE.0116726.

Randell, J.C.W., Bowers, J.L., Rodriguez, H.K. & Bell, S.P. (2006) Sequential ATP hydrolysis by Cdc6 and ORC directs loading of the Mcm2-7 helicase. *Molecular cell*. 21 (1), 29–39. doi:10.1016/j.molcel.2005.11.023.

Randell, J.C.W., Fan, A., Chan, C., Francis, L.I., Heller, R.C., Galani, K. & Bell, S.P. (2010) Mec1 is one of multiple kinases that prime the Mcm2-7 helicase for phosphorylation by Cdc7. *Molecular cell*. 40 (3), 353–363. doi:10.1016/J.MOLCEL.2010.10.017.

Rao, H. & Stillman, B. (1995) The origin recognition complex interacts with a bipartite DNA binding site within yeast replicators. *Proceedings of the National Academy of Sciences*. 92 (6), 2224–2228. doi:10.1073/PNAS.92.6.2224.

Rao PN, Johnson RT. (1970) Mammalian Cell Fusion : Studies on the Regulation of DNA Synthesis and Mitosis. *Nature*. 225 (5228), 159–164. doi:10.1038/225159a0.

Reinapae, A., Jalakas, K., Avvakumov, N., Lõoke, M., Kristjuhan, K. & Kristjuhan, A. (2017) Recruitment of Fkh1 to replication origins requires precisely positioned Fkh1/2 binding sites and concurrent assembly of the pre-replicative complex. *PLOS Genetics*. 13 (1), e1006588. doi:10.1371/JOURNAL.PGEN.1006588.

Remus, D., Beall, E.L. & Botchan, M.R. (2004) DNA topology, not DNA sequence, is a critical determinant for Drosophila ORC-DNA binding. *The EMBO journal*. 23 (4), 897–907. doi:10.1038/SJ.EMBOJ.7600077.

Remus, D., Beuron, F., Tolun, G., Griffith, J.D., Morris, E.P. & Diffley, J.F.X. (2009) Concerted loading of Mcm2-7 double hexamers around DNA during DNA replication origin licensing. *Cell*. 139 (4), 719–730. doi:10.1016/j.cell.2009.10.015.

Renauld, H., Aparicio, O.M., Zierath, P.D., Billington, B.L., Chhablani, S.K. & Gottschling, D.E. (1993) Silent domains are assembled continuously from the telomere and are defined by promoter distance and strength, and by SIR3 dosage. *Genes & development*. 7 (7A), 1133–1145. doi:10.1101/GAD.7.7A.1133.

Reuter, L.M., Meinel, D.M. & Sträßer, K. (2015) The poly(A)-binding protein Nab2 functions in RNA polymerase III transcription. *Genes & Development*. 29 (14), 1565. doi:10.1101/GAD.266205.115.

Rhind, N. & Russell, P. (2000) Checkpoints: It takes more than time to heal some wounds. *Current biology : CB*. 10 (24), R908. doi:10.1016/S0960-9822(00)00849-6.

Rhind, N. & Russell, P. (1998) Mitotic DNA damage and replication checkpoints in yeast. *Current opinion in cell biology*. 10 (6), 749. doi:10.1016/S0955-0674(98)80118-X.

Richards, A.L., Eckhardt, M. & Krogan, N.J. (2021) Mass spectrometry-based protein–protein interaction networks for the study of human diseases. *Molecular Systems Biology*. 17 (1), e8792. doi:10.15252/MSB.20188792.

Ricke, R.M. & Bielinsky, A.K. (2004) Mcm10 regulates the stability and chromatin association of DNA polymerase-alpha. *Molecular cell*. 16 (2), 173–185. doi:10.1016/J.MOLCEL.2004.09.017.

Riera, A., Barbon, M., Noguchi, Y., Reuter, L.M., Schneider, S. & Speck, C. (2017) From structure to mechanism—understanding initiation of DNA replication. *Genes & Development*. 31 (11), 1073. doi:10.1101/GAD.298232.117.

Rine, J. & Herskowitz, I. (1987) Four Genes Responsible for a Position Effect on Expression From HML and HMR in Saccharomyces cerevisiae. *Genetics*. 116 (1).

Rogers, C.M., Lee, C.-Y., Parkins, S., Buehler, N.J., Wenzel, S., Martínez-Márquez, F., Takagi, Y., Myong, S. & Bochman, M.L. (2020a) The yeast Hrq1 helicase stimulates Pso2 translesion nuclease activity and thereby promotes DNA interstrand crosslink repair. *The Journal of Biological Chemistry*. 295 (27), 8945. doi:10.1074/JBC.RA120.013626.

Rogers, C.M., Sanders, E., Nguyen, P.A., Smith-Kinnaman, W., Mosley, A.L. & Bochman, M.L. (2020b) The Genetic and Physical Interactomes of the Saccharomyces cerevisiae Hrq1 Helicase. *G3: Genes, Genomes, Genetics*. 10 (12), 4347–4357. doi:10.1534/G3.120.401864.

Rogers, C.M., Simmons, R.H., Fluhler Thornburg, G.E., Buehler, N.J. & Bochman, M.L. (2020c) Fanconi anemia-independent DNA inter-strand crosslink repair in eukaryotes. *Progress in Biophysics and Molecular Biology*. 158, 33–46. doi:10.1016/J.PBIOMOLBIO.2020.08.005.

Rogers, C.M., Wang, J.C.-Y., Noguchi, H., Imasaki, T., Takagi, Y. & Bochman, M.L. (2017) Yeast Hrq1 shares structural and functional homology with the disease-linked human RecQ4 helicase. *Nucleic Acids Research*. 45 (9), 5217. doi:10.1093/NAR/GKX151.

Röhrig, S., Dorn, A., Enderle, J., Schindele, A., Herrmann, N.J., Knoll, A. & Puchta, H. (2018) The RecQ-like helicase HRQ1 is involved in DNA crosslink repair in Arabidopsis in a common pathway

with the Fanconi anemia-associated nuclease FAN1 and the postreplicative repair ATPase RAD5A. *New Phytologist*. 218 (4), 1478–1490. doi:10.1111/NPH.15109.

Rosenthal, E.T., Hunt, T. & Ruderman, J. V (1980) Selective translation of mRNA controls the pattern of protein synthesis during early development of the surf clam, Spisula solidissima. *Cell*. 20 (2), 487–494. doi:https://doi.org/10.1016/0092-8674(80)90635-2.

Rossi, S.E., Ajazi, A., Carotenuto, W., Foiani, M. & Giannattasio, M. (2015) Rad53-Mediated Regulation of Rrm3 and Pif1 DNA Helicases Contributes to Prevention of Aberrant Fork Transitions under Replication Stress. *Cell Reports*. 13 (1), 80. doi:10.1016/J.CELREP.2015.08.073.

Rothenberg, E., Trakselis, M.A., Bell, S.D. & Ha, T. (2007) MCM forked substrate specificity involves dynamic interaction with the 5′-tail. *Journal of Biological Chemistry*. 282 (47), 34229–34234. doi:10.1074/JBC.M706300200/ATTACHMENT/71812BDF-6DCE-49E0-B6D7-291823448F4D/MMC1.PDF.

Rothwell, P.J. & Waksman, G. (2005) Structure and mechanism of DNA polymerases. *Advances in Protein Chemistry*. 71, 401–440. doi:10.1016/S0065-3233(04)71011-6.

Rowbotham, S.P., Barki, L., Neves-Costa, A., Santos, F., Dean, W., Hawkes, N., Choudhary, P., Will, W.R., Webster, J., Oxley, D., Green, C.M., Varga-Weisz, P. & Mermoud, J.E. (2011) Maintenance of silent chromatin through replication requires SWI/SNF-like chromatin remodeler SMARCAD1. *Molecular cell*. 42 (3), 285–296. doi:10.1016/J.MOLCEL.2011.02.036.

Rowley, A., Cocker, J.H., Harwood, J. & Diffley, J.F. (1995) Initiation complex assembly at budding yeast replication origins begins with the recognition of a bipartite sequence by limiting amounts of the initiator, ORC. *The EMBO Journal*. 14 (11), 2631. /pmc/articles/PMC398377/?report=abstract.

Rudin, N. & Haber, J.E. (1988) Efficient repair of HO-induced chromosomal breaks in Saccharomyces cerevisiae by recombination between flanking homologous sequences. *Molecular and cellular biology*. 8 (9), 3918–3928. doi:10.1128/MCB.8.9.3918-3928.1988.

Rundlett, S.E., Carmen, A.A., Kobayashi, R., Bavykin, S., Turner, B.M. & Grunstein, M. (1996) HDA1 and RPD3 are members of distinct yeast histone deacetylase complexes that regulate silencing and transcription. *Proceedings of the National Academy of Sciences of the United States of America*. 93 (25), 14503–14508. doi:10.1073/PNAS.93.25.14503.

Rusch, H.P., Sachsenmaier, W., Behrens, K. & Gruter, V. (1966) Synchronization of mitosis by the fusion of the plasmodia of Physarum polycephalum. *The Journal of cell biology*. 31 (1), 204–209. doi:10.1083/jcb.31.1.204.

Rusché, L.N., Kirchmaier, A.L. & Rine, J. (2002) Ordered nucleation and spreading of silenced chromatin in Saccharomyces cerevisiae. *Molecular biology of the cell*. 13 (7), 2207–2222. doi:10.1091/MBC.E02-03-0175.

Ruta, L.L. & Farcasanu, I.C. (2020) Saccharomyces cerevisiae and Caffeine Implications on the Eukaryotic Cell. *Nutrients*. 12 (8), 1–22. doi:10.3390/NU12082440.

Chuang R.Y. & Kelly T.J. (1999) The fission yeast homologue of Orc4p binds to replication origin DNA via multiple AT-hooks. *Proceedings of the National Academy of Sciences of the United States of America*. 96 (6), 2656–2661. doi:10.1073/PNAS.96.6.2656.

Ryan, D.P. & Owen-Hughes, T. (2011) Snf2-family proteins: chromatin remodellers for any occasion. *Current opinion in chemical biology*. 15 (5), 649. doi:10.1016/J.CBPA.2011.07.022.

Safarik, I. & Safarikova, M. (2004) Magnetic techniques for the isolation and purification of proteins and peptides. *BioMagnetic Research and Technology 2004 2:1*. 2 (1), 1–17. doi:10.1186/1477-044X-2-7.

Samel, S.A., Fernandez-Cid, A., Sun, J., Riera, A., Tognetti, S., Herrera, M.C., Li, H. & Speck, C. (2014) A unique DNA entry gate serves for regulated loading of the eukaryotic replicative helicase MCM2-7 onto DNA. *Genes & development*. 28 (15), 1653–1666. doi:10.1101/gad.242404.114.

Samora, C.P., Saksouk, J., Goswami, P., Wade, B.O., Singleton, M.R., Bates, P.A., Lengronne, A., Costa, A. & Uhlmann, F. (2016) Ctf4 Links DNA Replication with Sister Chromatid Cohesion Establishment by Recruiting the Chl1 Helicase to the Replisome. *Molecular Cell*. 63 (3), 371. doi:10.1016/J.MOLCEL.2016.05.036.

Sanchez, Y., Bachant, J., Wang, H., Hu, F., Liu, D., Tetzlaff, M. & Elledge, S.J. (1999) Control of the DNA damage checkpoint by chk1 and rad53 protein kinases through distinct mechanisms. *Science (New York, N.Y.)*. 286 (5442), 1166–1171. doi:10.1126/SCIENCE.286.5442.1166.

Sanchez-Pulido, L., Diffley, J.F.X. & Ponting, C.P. (2010) Homology explains the functional similarities of Treslin/Ticrr and Sld3. *Current Biology*. 20 (12), R509–R510. doi:10.1016/J.CUB.2010.05.021.

Sanchez-Pulido, L. & Ponting, C.P. (2011) Cdc45: the missing RecJ ortholog in eukaryotes? *Bioinformatics (Oxford, England)*. 27 (14), 1885–1888. doi:10.1093/BIOINFORMATICS/BTR332.

Sanders, E., Nguyen, P.A., Rogers, C.M. & Bochman, M.L. (2020) Comprehensive Synthetic Genetic Array Analysis of Alleles That Interact with Mutation of the Saccharomyces cerevisiae RecQ Helicases Hrq1 and Sgs1. *G3: Genes|Genomes|Genetics*. 10 (12), 4359. doi:10.1534/G3.120.401709.

Sansam, C.G., Pietrzak, K., Majchrzycka, B., Kerlin, M.A., Chen, J., Rankin, S. & Sansam, C.L. (2018) A mechanism for epigenetic control of DNA replication. *Genes and Development*. 32 (3–4), 224–229. doi:10.1101/GAD.306464.117/-/DC1.

Sansam, C.L., Cruz, N.M., Danielian, P.S., Amsterdam, A., Lau, M.L., Hopkins, N. & Lees, J.A. (2010) A vertebrate gene, ticrr, is an essential checkpoint and replication regulator. *Genes & development*. 24 (2), 183–194. doi:10.1101/GAD.1860310.

Santocanale, C. & Diffley, J.F. (1996) ORC- and Cdc6-dependent complexes at active and inactive chromosomal replication origins in Saccharomyces cerevisiae. *The EMBO Journal*. 15 (23), 6671. /pmc/articles/PMC452491/?report=abstract.

Santocanale, C., Sharma, K. & Diffley, J.F.X. (1999) Activation of dormant origins of DNA replication in budding yeast. *Genes & development*. 13 (18), 2360–2364. doi:10.1101/GAD.13.18.2360.

Sawyer, S.L., Cheng, I.H., Chai, W. & Tye, B.K. (2004) Mcm10 and Cdc45 cooperate in origin activation in Saccharomyces cerevisiae. *Journal of molecular biology*. 340 (2), 195–202. doi:10.1016/J.JMB.2004.04.066.

Schauer, G.D. & O'Donnell, M.E. (2017) Quality control mechanisms exclude incorrect polymerases from the eukaryotic replication fork. *Proceedings of the National Academy of Sciences of the United States of America*. 114 (4), 675–680. doi:10.1073/PNAS.1619748114/-/DCSUPPLEMENTAL.

Schauer, G.D., Spenkelink, L.M., Lewis, J.S., Yurieva, O., Mueller, S.H., van Oijen, A.M. & O'Donnell, M.E. (2020) Replisome bypass of a protein-based R-loop block by Pif1. *Proceedings of the National Academy of Sciences of the United States of America*. 117 (48), 30354–30361. doi:10.1073/PNAS.2020189117/-/DCSUPPLEMENTAL.

Schmelzle, T. & Hall, M.N. (2000) TOR, a central controller of cell growth. *Cell*. 103 (2), 253–262. doi:10.1016/S0092-8674(00)00117-3.

Schmidt, J.M. & Bleichert, F. (2020) Structural mechanism for replication origin binding and remodeling by a metazoan origin recognition complex and its co-loader Cdc6. *Nature Communications*. 11 (1). doi:10.1038/S41467-020-18067-7.

Schmidt, K.H., Derry, K.L. & Kolodner, R.D. (2002) Saccharomyces cerevisiae RRM3, a 5′ to 3′ DNA Helicase, Physically Interacts with Proliferating Cell Nuclear Antigen. *Journal of Biological Chemistry*. 277 (47), 45331–45337. doi:10.1074/JBC.M207263200.

Schmidt, K.H. & Kolodner, R.D. (2004) Requirement of Rrm3 Helicase for Repair of Spontaneous DNA Lesions in Cells Lacking Srs2 or Sgs1 Helicase. *Molecular and Cellular Biology*. 24 (8), 3213. doi:10.1128/MCB.24.8.3213-3226.2004.

Schmidt, K.H. & Kolodner, R.D. (2006) Suppression of spontaneous genome rearrangements in yeast DNA helicase mutants. *Proceedings of the National Academy of Sciences of the United States of America*. 103 (48), 18196–18201. doi:10.1073/PNAS.0608566103.

Schmitt-Ulms, G., Hansen, K., Liu, J., Cowdrey, C., Yang, J., DeArmond, S.J., Cohen, F.E., Prusiner, S.B. & Baldwin, M.A. (2004) Time-controlled transcardiac perfusion cross-linking for the study of protein interactions in complex tissues. *Nature Biotechnology 2004 22:6*. 22 (6), 724–731. doi:10.1038/nbt969.

Schneider, B.L., Zhang, J., Markwardt, J., Tokiwa, G., Volpe, T., Honey, S. & Futcher, B. (2004) Growth rate and cell size modulate the synthesis of, and requirement for, G1-phase cyclins at start. *Molecular and cellular biology*. 24 (24), 10802–10813. doi:10.1128/MCB.24.24.10802-10813.2004.

Scholes DT, Banerjee M, Bowen B, Curcio MJ. Multiple regulators of Ty1 transposition in Saccharomyces cerevisiae have conserved roles in genome maintenance. Genetics. 2001 Dec;159(4):1449-65. doi: 10.1093/genetics/159.4.1449. PMID: 11779788; PMCID: PMC1461915.

Schoor, M., Schuster-Gossler, K. & Gossler, A. (1993) The Etl-1 gene encodes a nuclear protein differentially expressed during early mouse development. *Developmental Dynamics*. 197 (3), 227–237. doi:10.1002/AJA.1001970307.

Schulz, V.P. & Zakian, V.A. (1994) The Saccharomyces /WI DNA Helicase Inhibits Telomere Elongation and De Novo Telomere Formation. *Cell*.76.

Segurado, M. & Diffley, J.F.X. (2008) Separate roles for the DNA damage checkpoint protein kinases in stabilizing DNA replication forks. *Genes & development*. 22 (13), 1816–1827. doi:10.1101/GAD.477208.

Seki, T. & Diffley, J.F.X. (2000) Stepwise assembly of initiation proteins at budding yeast replication origins in vitro. *Proceedings of the National Academy of Sciences of the United States of America*. 97 (26), 14115–14120. doi:10.1073/PNAS.97.26.14115.

Sen, R., Lahudkar, S., Durairaj, G. & Bhaumik, S.R. (2013) Functional Analysis of Bre1p, an E3 Ligase for Histone H2B Ubiquitylation, in Regulation of RNA Polymerase II Association with Active Genes and Transcription in Vivo. *The Journal of Biological Chemistry*. 288 (14), 9619. doi:10.1074/JBC.M113.450403.

Sengupta, S., van Deursen, F., de Piccoli, G. & Labib, K. (2013) Dpb2 integrates the leading-strand DNA polymerase into the eukaryotic replisome. *Current biology : CB*. 23 (7), 543–552. doi:10.1016/J.CUB.2013.02.011.

Sequeira-Mendes, J., Díaz-Uriarte, R., Apedaile, A., Huntley, D., Brockdorff, N. & Gómez, M. (2009) Transcription initiation activity sets replication origin efficiency in mammalian cells. *PLoS genetics*. 5 (4). doi:10.1371/JOURNAL.PGEN.1000446.

Shcherbakova, P. v., Pavlov, Y.I., Chilkova, O., Rogozin, I.B., Johansson, E. & Kunkel, T.A. (2003) Unique error signature of the four-subunit yeast DNA polymerase epsilon. *The Journal of biological chemistry*. 278 (44), 43770–43780. doi:10.1074/JBC.M306893200.

Shechter, D.F., Ying, C.Y. & Gautier, J. (2000) The intrinsic DNA helicase activity of Methanobacterium thermoautotrophicum delta H minichromosome maintenance protein. *The Journal of biological chemistry*. 275 (20), 15049–15059. doi:10.1074/jbc.M000398200.

Sheu, Y.J. & Stillman, B. (2006) Cdc7-Dbf4 phosphorylates MCM proteins via a docking site-mediated mechanism to promote S phase progression. *Molecular cell*. 24 (1), 101–113. doi:10.1016/J.MOLCEL.2006.07.033.

Sheu, Y.J. & Stillman, B. (2010) The Dbf4-Cdc7 kinase promotes S phase by alleviating an inhibitory activity in Mcm4. *Nature*. 463 (7277), 113. doi:10.1038/NATURE08647.

Shibahara, K.I. & Stillman, B. (1999) Replication-Dependent Marking of DNA by PCNA Facilitates CAF-1-Coupled Inheritance of Chromatin. *Cell*. 96 (4), 575–585. doi:10.1016/S0092-8674(00)80661-3.

Shou, W., Seol, J.H., Shevchenko, A., Baskerville, C., Moazed, D., Susan Chen, Z.W., Jang, J., Shevchenko, A., Charbonneau, H. & Deshaies, R.J. (1999) Exit from mitosis is triggered by Tem1-dependent release of the protein phosphatase Cdc14 from nucleolar RENT complex. *Cell*. 97 (2), 233–244. doi:10.1016/S0092-8674(00)80733-3.

Shreeram, S. & Blow, J.J. (2003) The Role of the Replication Licensing System in Cell Proliferation and Cancer. *Progress in cell cycle research*. 5, 287. /pmc/articles/PMC3604915/.

Shyian, M., Albert, B., Zupan, A.M., Ivanitsa, V., Charbonnet, G., Dilg, D. & Shore, D. (2020) Fork pausing complex engages topoisomerases at the replisome. *Genes & development*. 34 (1–2), 87–98. doi:10.1101/GAD.331868.119/-/DC1.

Sielaff, M., Kuharev, J., Bohn, T., Hahlbrock, J., Bopp, T., Tenzer, S. & Distler, U. (2017) Evaluation of FASP, SP3, and iST Protocols for Proteomic Sample Preparation in the Low Microgram Range. *Journal of Proteome Research*. 16 (11), 4060–4072. doi:10.1021/ACS.JPROTEOME.7B00433/SUPPL_FILE/PR7B00433_SI_009.XLSX.

Sigrist, S., Jacobs, H., Stratmann, R. & Lehner, C.F. (1995) Exit from mitosis is regulated by Drosophila fizzy and the sequential destruction of cyclins A, B and B3. *The EMBO Journal*. 14 (19), 4827–4838. doi:10.1002/J.1460-2075.1995.TB00164.X.

Siitonen, A.H., Sotkasiira, J., Biervliet, M., Benmansour, A., Capri, Y., et al. (2009) The mutation spectrum in RECQL4 diseases. *European journal of human genetics : EJHG*. 17 (2), 151–158. doi:10.1038/EJHG.2008.154.

Siitonen, H.A., Kopra, O., Kääriäinen, H., Haravuori, H., Winter, R.M., Säämänen, A.M., Peltonen, L. & Kestilä, M. (2003) Molecular defect of RAPADILINO syndrome expands the phenotype spectrum of RECQL diseases. *Human molecular genetics*. 12 (21), 2837–2844. doi:10.1093/HMG/DDG306.

Siler, J., Xia, B., Wong, C., Kath, M. & Bi, X. (2017) Cell cycle-dependent positive and negative functions of Fun30 chromatin remodeler in DNA damage response. *DNA Repair*. 50, 61–70. doi:10.1016/J.DNAREP.2016.12.009.

Simon, A.C., Zhou, J.C., Perera, R.L., van Deursen, F., Evrin, C., Ivanova, M.E., Kilkenny, M.L., Renault, L., Kjaer, S., Matak-Vinkovi, D., Labib, K., Costa, A. & Pellegrini, L. (2014) A Ctf4 trimer couples the CMG helicase to DNA polymerase α in the eukaryotic replisome. *Nature 2014 510:7504*. 510 (7504), 293–297. doi:10.1038/nature13234.

Singh, A. & Xu, Y.J. (2016) The Cell Killing Mechanisms of Hydroxyurea. *Genes*. 7 (11). doi:10.3390/GENES7110099.

Singh, D.K., Ghosh, A.K., Croteau, D.L. & Bohr, V.A. (2012) RecQ helicases in DNA double strand break repair and telomere maintenance. *Mutation research*. 736 (1–2), 15–24. doi:10.1016/J.MRFMMM.2011.06.002.

Singh, J. & Klat, A.J.S. (1992) *Active genes in budding yeast display enhanced in vivo accessibility to foreign DNA methylases: a novel in vivo probe for chromatin structure of yeast.*

Sinha, P., Chang, V. & Tye, B.K. (1986) A mutant that affects the function of autonomously replicating sequences in yeast. *Journal of Molecular Biology*. 192 (4), 805–814. doi:10.1016/0022-2836(86)90030-6.

Sipiczki, M. (2000) Where does fission yeast sit on the tree of life? *Genome Biology*. 1 (2), reviews1011.1. doi:10.1186/GB-2000-1-2-REVIEWS1011.

Sirbu, B.M., McDonald, W.H., Dungrawala, H., Badu-Nkansah, A., Kavanaugh, G.M., Chen, Y., Tabb, D.L. & Cortez, D. (2013) Identification of Proteins at Active, Stalled, and Collapsed Replication Forks Using Isolation of Proteins on Nascent DNA (iPOND) Coupled with Mass Spectrometry. *The Journal of Biological Chemistry*. 288 (44), 31458. doi:10.1074/JBC.M113.511337.

Smith, S. & Stillman, B. (1989) Purification and characterization of CAF-I, a human cell factor required for chromatin assembly during DNA replication in vitro. *Cell*. 58 (1), 15–25. doi:10.1016/0092-8674(89)90398-X.

Sogo, J.M., Lopes, M. & Foiani, M. (2002) Fork reversal and ssDNA accumulation at stalled replication forks owing to checkpoint defects. *Science (New York, N.Y.)*. 297 (5581), 599–602. doi:10.1126/SCIENCE.1074023.

Solomon, M.J., Larsen, P.L. & Varshavsky, A. (1988) Mapping proteinDNA interactions in vivo with formaldehyde: Evidence that histone H4 is retained on a highly transcribed gene. *Cell*. 53 (6), 937–947. doi:10.1016/S0092-8674(88)90469-2.

Solomon, N.A., Wright, M.B., Chang, S., Buckley, A.M., Dumas, L.B. & Gaber, R.F. (1992) Genetic and molecular analysis of DNA43 and DNA52: two new cell-cycle genes in Saccharomyces cerevisiae. *Yeast (Chichester, England)*. 8 (4), 273–289. doi:10.1002/YEA.320080405.

Somyajit, K., Gupta, R., Sedlackova, H., Neelsen, K.J., Ochs, F., Rask, M.B., Choudhary, C. & Lukas, J. (2017) Redox-sensitive alteration of replisome architecture safeguards genome integrity. *Science (New York, N.Y.)*. 358 (6364), 797–802. doi:10.1126/SCIENCE.AAO3172.

Bell S.P. & Dutta A. (2002) DNA replication in eukaryotic cells. *Annual review of biochemistry*. 71, 333–374. doi:10.1146/ANNUREV.BIOCHEM.71.110601.135425.

Bell S.P., Mitchell J., Leber J., Kobayashi R. & Stillman B. (1995) The multidomain structure of Orc1p reveals similarity to regulators of DNA replication and transcriptional silencing. *Cell*. 83 (4), 563–568. doi:10.1016/0092-8674(95)90096-9.

Speck, C., Chen, Z., Li, H. & Stillman, B. (2005) ATPase-dependent, cooperative binding of ORC and Cdc6p to origin DNA. *Nature structural & molecular biology*. 12 (11), 965. doi:10.1038/NSMB1002.

Speck, C. & Stillman, B. (2007) Cdc6 ATPase ACTIVITY REGULATES ORC-Cdc6 STABILITY AND THE SELECTION OF SPECIFIC DNA SEQUENCES AS ORIGINS OF DNA REPLICATION. *The Journal of biological chemistry*. 282 (16), 11705. doi:10.1074/JBC.M700399200.

Spencer, F., Gerring, S.L., Connelly, C. & Hieter, P. (1990) Mitotic chromosome transmission fidelity mutants in Saccharomyces cerevisiae. *Genetics*. 124 (2), 237–249. doi:10.1093/GENETICS/124.2.237.

Srinivasan, M., Mehta, P., Yu, Y., Prugar, E., Koonin, E. v., Karzai, A.W. & Sternglanz, R. (2011) The highly conserved KEOPS/EKC complex is essential for a universal tRNA modification, t6A. *The EMBO Journal*. 30 (5), 873. doi:10.1038/EMBOJ.2010.343.

Stadler, J. & Richly, H. (2017) Regulation of DNA Repair Mechanisms: How the Chromatin Environment Regulates the DNA Damage Response. *International Journal of Molecular Sciences*. 18 (8). doi:10.3390/IJMS18081715.

Sterner, D.E., Grant, P.A., Roberts, S.M., Duggan, L.J., Belotserkovskaya, R., Pacella, L.A., Winston, F., Workman, J.L. & Berger, S.L. (1999) Functional Organization of the Yeast SAGA Complex: Distinct Components Involved in Structural Integrity, Nucleosome Acetylation, and TATA-Binding Protein Interaction. *Molecular and Cellular Biology*. 19 (1), 86. doi:10.1128/MCB.19.1.86.

Stevenson, J.B. & Gottschling, D.E. (1999) Telomeric chromatin modulates replication timing near chromosome ends. *Genes & Development*. 13 (2), 146. doi:10.1101/GAD.13.2.146.

Stillman, B. (1996) Cell cycle control of DNA replication. *Science (New York, N.Y.)*. 274 (5293), 1659–1664.

Stinchcomb, D.T., Struhl, K. & Davis, R.W. (1979) Isolation and characterisation of a yeast chromosomal replicator. *Nature 1979 282:5734*. 282 (5734), 39–43. doi:10.1038/282039a0.

Straight, A.F., Shou, W., Dowd, G.J., Turck, C.W., Deshaies, R.J., Johnson, A.D. & Moazed, D. (1999) Net1, a Sir2-associated nucleolar protein required for rDNA silencing and nucleolar integrity. *Cell*. 97 (2), 245–256. doi:10.1016/S0092-8674(00)80734-5.

Strålfors, A., Walfridsson, J., Bhuiyan, H. & Ekwall, K. (2011) The FUN30 chromatin remodeler, Fft3, protects centromeric and subtelomeric domains from euchromatin formation. *PLoS genetics*. 7 (3), e1001334–e1001334. doi:10.1371/journal.pgen.1001334.

Struhl, K., Stinchcomb, D.T., Scherer, S. & Davis, R.W. (1979) High-frequency transformation of yeast: autonomous replication of hybrid DNA molecules. *Proceedings of the National Academy of Sciences of the United States of America*. 76 (3), 1035–1039. doi:10.1073/PNAS.76.3.1035.

Sukackaite, R., Cornacchia, D., Jensen, M.R., Mas, P.J., Blackledge, M., Enervald, E., Duan, G., Auchynnikava, T., Köhn, M., Hart, D.J. & Buonomo, S.B.C. (2017) Mouse Rif1 is a regulatory subunit of protein phosphatase 1 (PP1). *Scientific reports*. 7 (1). doi:10.1038/S41598-017-01910-1.

Sun, J., Evrin, C., Samel, S.A., Fernandez-Cid, A., Riera, A., Kawakami, H., Stillman, B., Speck, C. & Li, H. (2013) Cryo-EM structure of a helicase loading intermediate containing ORC-Cdc6-Cdt1-MCM2-7 bound to DNA. *Nature structural & molecular biology*. 20 (8), 944–951. doi:10.1038/nsmb.2629.

Sun, J., Fernandez-Cid, A., Riera, A., Tognetti, S., Yuan, Z., Stillman, B., Speck, C. & Li, H. (2014) Structural and mechanistic insights into Mcm2–7 double-hexamer assembly and function. *Genes & Development*. 28 (20), 2291. doi:10.1101/GAD.242313.114.

Sun, J., Shi, Y., Georgescu, R.E., Yuan, Z., Chait, B.T., Li, H. & O'Donnell, M.E. (2015) The architecture of a eukaryotic replisome. *Nature structural & molecular biology*. 22 (12), 976–982. doi:10.1038/NSMB.3113.

Suter, B., Tong, A., Chang, M., Yu, L., Brown, G.W., Boone, C. & Rine, J. (2004) The Origin Recognition Complex Links Replication, Sister Chromatid Cohesion and Transcriptional Silencing in Saccharomyces cerevisiae. *Genetics*. 167 (2), 579–591. doi:10.1534/GENETICS.103.024851.

Swaffer, M.P., Jones, A.W., Flynn, H.R., Snijders, A.P. & Nurse, P. (2016) CDK Substrate Phosphorylation and Ordering the Cell Cycle. *Cell*. 167 (7), 1750-1761.e16. doi:https://doi.org/10.1016/j.cell.2016.11.034.

Swenson, K.I., Farrell, K.M. & Ruderman, J. V (1986) The clam embryo protein cyclin A induces entry into M phase and the resumption of meiosis in Xenopus oocytes. *Cell*. 47 (6), 861–870. doi:10.1016/0092-8674(86)90801-9.

Swygert, S.G., Senapati, S., Bolukbasi, M.F., Wolfe, S.A., Lindsay, S. & Peterson, C.L. (2018) SIR proteins create compact heterochromatin fibers. *Proceedings of the National Academy of Sciences*. 115 (49), 12447–12452. doi:10.1073/PNAS.1810647115.

Syed, S., Desler, C., Rasmussen, L.J. & Schmidt, K.H. (2016) A Novel Rrm3 Function in Restricting DNA Replication via an Orc5-Binding Domain Is Genetically Separable from Rrm3 Function as an

ATPase/Helicase in Facilitating Fork Progression. *PLoS Genetics*. 12 (12). doi:10.1371/JOURNAL.PGEN.1006451.

Symington, L.S., Rothstein, R. & Lisby, M. (2014) Mechanisms and regulation of mitotic recombination in Saccharomyces cerevisiae. *Genetics*. 198 (3), 795–835. doi:10.1534/GENETICS.114.166140.

Szambowska, A., Tessmer, I., Kursula, P., Usskilat, C., Prus, P., Pospiech, H. & Grosse, F. (2014) DNA binding properties of human Cdc45 suggest a function as molecular wedge for DNA unwinding. *Nucleic Acids Research*. 42 (4), 2308–2319. doi:10.1093/NAR/GKT1217.

Szyjka, S.J., Viggiani, C.J. & Aparicio, O.M. (2005) Mrc1 is required for normal progression of replication forks throughout chromatin in S. cerevisiae. *Molecular cell*. 19 (5), 691–697. doi:10.1016/J.MOLCEL.2005.06.037.

Tada, S. & Blow, J.J. (1998) The replication licensing system. *Biological chemistry*. 379 (8–9), 941. doi:10.1515/bchm.1998.379.8-9.941.

Takahashi, T.S., Wigley, D.B. & Walter, J.C. (2005) Pumps, paradoxes and ploughshares: mechanism of the MCM2-7 DNA helicase. *Trends in biochemical sciences*. 30 (8), 437–444. doi:10.1016/J.TIBS.2005.06.007.

Takayama, Y., Kamimura, Y., Okawa, M., Muramatsu, S., Sugino, A. & Araki, H. (2003) GINS, a novel multiprotein complex required for chromosomal DNA replication in budding yeast. *Genes & Development*. 17 (9), 1153. doi:10.1101/GAD.1065903.

Tan, B.C.M., Chien, C.T., Hirose, S. & Lee, S.C. (2006) Functional cooperation between FACT and MCM helicase facilitates initiation of chromatin DNA replication. *The EMBO Journal*. 25 (17), 3975. doi:10.1038/SJ.EMBOJ.7601271.

Tan, C.-K., Castillo, C., So, A.G. & Downey, K.M. (1986) *THE JOURNAL OP BIOLOGICAL CHEMISTRY An Auxiliary Protein for DNA Polymerase-6 from Fetal Calf Thymus**. 261 (26), 12310–12316. doi:10.1016/S0021-9258(18)67240-0.

Tanaka, H., Kubota, Y., Tsujimura, T., Kumano, M., Masai, H. & Takisawa, H. (2009) Replisome progression complex links DNA replication to sister chromatid cohesion in Xenopus egg extracts. *Genes to cells : devoted to molecular & cellular mechanisms*. 14 (8), 949–963. doi:10.1111/J.1365-2443.2009.01322.X.

Tanaka, S. & Diffley, J.F.X. (2002) Interdependent nuclear accumulation of budding yeast Cdt1 and Mcm2-7 during G1 phase. *Nature cell biology*. 4 (3), 198–207. doi:10.1038/NCB757.

Tanaka, S., Komeda, Y., Umemori, T., Kubota, Y., Takisawa, H. & Araki, H. (2013) Efficient Initiation of DNA Replication in Eukaryotes Requires Dpb11/TopBP1-GINS Interaction. *Molecular and Cellular Biology*. 33 (13), 2614–2622. doi:10.1128/MCB.00431-13/SUPPL_FILE/ZMB999100022SO1.PDF.

Tanaka, S., Nakato, R., Katou, Y., Shirahige, K. & Araki, H. (2011a) Origin association of Sld3, Sld7, and Cdc45 proteins is a key step for determination of origin-firing timing. *Current biology : CB*. 21 (24), 2055–2063. doi:10.1016/J.CUB.2011.11.038.

Tanaka, S., Umemori, T., Hirai, K., Muramatsu, S., Kamimura, Y. & Araki, H. (2007) CDK-dependent phosphorylation of Sld2 and Sld3 initiates DNA replication in budding yeast. *Nature*. 445 (7125), 328–332. doi:10.1038/NATURE05465.

Tanaka, T., Umemori, T., Endo, S., Muramatsu, S., Kanemaki, M., Kamimura, Y., Obuse, C. & Araki, H. (2011d) Sld7, an Sld3-associated protein required for efficient chromosomal DNA replication in budding yeast. *The EMBO Journal*. 30 (10), 2019. doi:10.1038/EMBOJ.2011.115.

Tanny, J.C., Dowd, G.J., Huang, J., Hilz, H. & Moazed, D. (1999) An enzymatic activity in the yeast Sir2 protein that is essential for gene silencing. *Cell*. 99 (7), 735–745. doi:10.1016/S0092-8674(00)81671-2.

Tanny, J.C. & Moazed, D. (2001) Coupling of histone deacetylation to NAD breakdown by the yeast silencing protein Sir2: Evidence for acetyl transfer from substrate to an NAD breakdown product. *Proceedings of the National Academy of Sciences of the United States of America*. 98 (2), 415. doi:10.1073/PNAS.031563798.

Tayri-Wilk, T., Slavin, M., Zamel, J., Blass, A., Cohen, S., Motzik, A., Sun, X., Shalev, D.E., Ram, O. & Kalisman, N. (2020) Mass spectrometry reveals the chemistry of formaldehyde cross-linking in structured proteins. *Nature Communications 2020 11:1*. 11 (1), 1–9. doi:10.1038/s41467-020-16935-w.

Tercero, J.A., Labib, K. & Diffley, J.F.X. (2000) DNA synthesis at individual replication forks requires the essential initiation factor Cdc45p. *The EMBO journal*. 19 (9), 2082–2093. doi:10.1093/EMBOJ/19.9.2082.

Tercero, J.A., Longhese, M.P. & Diffley, J.F.X. (2003) A central role for DNA replication forks in checkpoint activation and response. *Molecular cell*. 11 (5), 1323–1336. doi:10.1016/S1097-2765(03)00169-2.

Terui, R., Nagao, K., Kawasoe, Y., Taki, K., Higashi, T.L., Tanaka, S., Nakagawa, T., Obuse, C., Masukata, H. & Takahashi, T.S. (2018) Nucleosomes around a mismatched base pair are excluded via an Msh2-dependent reaction with the aid of SNF2 family ATPase Smarcad1. *Genes & development*. 32 (11–12), 806–821. doi:10.1101/gad.310995.117.

Teytelman, L., Thurtle, D.M., Rine, J. & van Oudenaarden, A. (2013) Highly expressed loci are vulnerable to misleading ChIP localization of multiple unrelated proteins. *Proceedings of the National Academy of Sciences of the United States of America*. 110 (46), 18602–18607. doi:10.1073/PNAS.1316064110/-/DCSUPPLEMENTAL.

Tham, W.-H. & Zakian, V.A. (2002) Transcriptional silencing at Saccharomyces telomeres: implications for other organisms. *Oncogene 2002 21:4*. 21 (4), 512–521. doi:10.1038/sj.onc.1205078.

Thommes, P., Fett, R., Schray, B., Burkhart, R., Barnes, M., Kennedy, C., Brown, N.C. & Knippers, R. (1992) Properties of the nuclear P1 protein, a mammalian homologue of the yeast Mcm3 replication protein. *Nucleic acids research*. 20 (5), 1069–1074. doi:10.1093/NAR/20.5.1069.

Thommes, P., Kubota, Y., Takisawa, H. & Blow, J.J. (1997) The RLF-M component of the replication licensing system forms complexes containing all six MCM/P1 polypeptides. *The EMBO journal*. 16 (11), 3312–3319. doi:10.1093/emboj/16.11.3312.

Thomsen, N.D. & Berger, J.M. (2008) Structural frameworks for considering microbial protein- and nucleic acid-dependent motor ATPases. *Molecular microbiology*. 69 (5), 1071–1090. doi:10.1111/J.1365-2958.2008.06364.X.

Thu, Y.M. & Bielinsky, A.K. (2013) Enigmatic roles of Mcm10 in DNA replication. *Trends in biochemical sciences*. 38 (4), 184. doi:10.1016/J.TIBS.2012.12.003.

Ticau, S., Friedman, L.J., Champasa, K., Corrêa, I.R., Gelles, J. & Bell, S.P. (2017) Mechanism and Timing of Mcm2–7 Ring Closure During DNA Replication Origin Licensing. *Nature structural & molecular biology*. 24 (3), 309. doi:10.1038/NSMB.3375.

Ticau, S., Friedman, L.J., Ivica, N.A., Gelles, J. & Bell, S.P. (2015) Single-Molecule Studies of Origin Licensing Reveal Mechanisms Ensuring Bidirectional Helicase Loading. *Cell*. 161 (3), 513–525. doi:10.1016/J.CELL.2015.03.012.

Tocilj, A., On, K.F., Yuan, Z., Sun, J., Elkayam, E., Li, H., Stillman, B. & Joshua-Tor, L. (2017) Structure of the active form of human origin recognition complex and its ATPase motor module. *eLife*. 6. doi:10.7554/eLife.20818.

Torres, J.Z., Bessler, J.B. & Zakian, V.A. (2004) Local chromatin structure at the ribosomal DNA causes replication fork pausing and genome instability in the absence of the S. cerevisiae DNA helicase Rrm3p. *Genes & Development*. 18 (5), 498. doi:10.1101/GAD.1154704.

Torres, J.Z., Schnakenberg, S.L. & Zakian, V.A. (2004) Saccharomyces cerevisiae Rrm3p DNA Helicase Promotes Genome Integrity by Preventing Replication Fork Stalling: Viability of rrm3 Cells Requires the Intra-S-Phase Checkpoint and Fork Restart Activities. *Molecular and Cellular Biology*. 24 (8), 3198. doi:10.1128/MCB.24.8.3198-3212.2004.

Tourrière, H., Versini, G., Cordón-Preciado, V., Alabert, C. & Pasero, P. (2005) Mrc1 and Tof1 promote replication fork progression and recovery independently of Rad53. *Molecular cell*. 19 (5), 699–706. doi:10.1016/J.MOLCEL.2005.07.028.

Tran, H.G., Steger, D.J., Iyer, V.R. & Johnson, A.D. (2000) The chromo domain protein Chd1p from budding yeast is an ATP-dependent chromatin-modifying factor. *The EMBO Journal*. 19 (10), 2323. doi:10.1093/EMBOJ/19.10.2323.

Tugal, T., Zou-Yang, X.H., Gavin, K., Pappin, D., Canas, B., Kobayashi, R., Hunt, T. & Stillman, B. (1998) The Orc4p and Orc5p Subunits of the Xenopus and Human Origin Recognition Complex Are Related to Orc1p and Cdc6p. *Journal of Biological Chemistry*. 273 (49), 32421–32429. doi:10.1074/JBC.273.49.32421.

Turner, J.J., Ewald, J.C. & Skotheim, J.M. (2012) Cell size control in yeast. *Current Biology*. 22 (9), R350. doi:10.1016/J.CUB.2012.02.041.

Tüting, C., Iacobucci, C., Ihling, C.H., Kastritis, P.L. & Sinz, A. (2020) Structural analysis of 70S ribosomes by cross-linking/mass spectrometry reveals conformational plasticity. *Scientific Reports 2020 10:1*. 10 (1), 1–13. doi:10.1038/s41598-020-69313-3.

Tyanova, S., Temu, T. & Cox, J. (2016) The MaxQuant computational platform for mass spectrometry-based shotgun proteomics. *Nature Protocols 2016 11:12*. 11 (12), 2301–2319. doi:10.1038/nprot.2016.136.

Tyanova, S., Temu, T., Sinitcyn, P., Carlson, A., Hein, M.Y., Geiger, T., Mann, M. & Cox, J. (2016) The Perseus computational platform for comprehensive analysis of (prote)omics data. *Nature Methods 2016 13:9*. 13 (9), 731–740. doi:10.1038/nmeth.3901.

Tye, B.K. (2003) MCM Proteins in DNA Replication. *http://dx.doi.org/10.1146/annurev.biochem.68.1.649*. 68, 649–686. doi:10.1146/ANNUREV.BIOCHEM.68.1.649.

Tyers, M., Tokiwa, G. & Futcher, B. (1993) Comparison of the Saccharomyces cerevisiae G1 cyclins: Cln3 may be an upstream activator of Cln1, Cln2 and other cyclins. *The EMBO Journal*. 12 (5), 1955. doi:10.1002/j.1460-2075.1993.tb05845.x.

Tyler, J.K., Adams, C.R., Chen, S.R., Kobayashi, R., Kamakaka, R.T. & Kadonaga, J.T. (1999) The RCAF complex mediates chromatin assembly during DNA replication and repair. *Nature*. 402 (6761), 555–560. doi:10.1038/990147.

Tzouros, M., Golling, S., Avila, D., Lamerz, J., Berrera, M., Ebeling, M., Langen, H. & Augustin, A. (2013) Development of a 5-plex SILAC Method Tuned for the Quantitation of Tyrosine Phosphorylation Dynamics. *Molecular & Cellular Proteomics : MCP*. 12 (11), 3339. doi:10.1074/MCP.O113.027342.

Uchiyama, M., Arai, K. & Masai, H. (2001) Sna41goa1, a novel mutation causing G1/S arrest in fission yeast, is defective in a CDC45 homolog and interacts genetically with polalpha. *Molecular genetics and genomics : MGG*. 265 (6), 1039–1049. doi:10.1007/S004380100499.

Uemura, T., Ohkura, H., Adachi, Y., Morino, K., Shiozaki, K. & Yanagida, M. (1987) DNA topoisomerase II is required for condensation and separation of mitotic chromosomes in S. pombe. *Cell*. 50 (6), 917–925. doi:10.1016/0092-8674(87)90518-6.

Uhlmann, F., Wernic, D., Poupart, M.-A., Koonin, E. v & Nasmyth, K. (2000) Cleavage of Cohesin by the CD Clan Protease Separin Triggers Anaphase in Yeast. *Cell*. 103 (3), 375–386. doi:10.1016/S0092-8674(00)00130-6.

Umezu, K., Nakayama, K. & Nakayama, H. (1990) Escherichia coli RecQ protein is a DNA helicase. *Proceedings of the National Academy of Sciences*. 87 (14), 5363–5367. doi:10.1073/PNAS.87.14.5363.

Utani, K., Fu, H., Jang, S.-M., Marks, A.B., Smith, O.K., Zhang, Y., Redon, C.E., Shimizu, N. & Aladjem, M.I. (2017) Phosphorylated SIRT1 associates with replication origins to prevent excess replication initiation and preserve genomic stability. *Nucleic Acids Research*. 45 (13), 7807. doi:10.1093/NAR/GKX468.

Valerio-Santiago, M., de los Santos-Velázquez, A.I. & Monje-Casas, F. (2013) Inhibition of the mitotic exit network in response to damaged telomeres. *PLoS genetics*. 9 (10). doi:10.1371/JOURNAL.PGEN.1003859.

Varadi M, Anyango S, Deshpande M, Nair S, Natassia C, Yordanova G, Yuan D, Stroe O, Wood G, Laydon A, Žídek A, Green T, Tunyasuvunakool K, Petersen S, Jumper J, Clancy E, Green R, Vora A, Lutfi M, Figurnov M, Cowie A, Hobbs N, Kohli P, Kleywegt G, Birney E, Hassabis D, Velankar S.

AlphaFold Protein Structure Database: massively expanding the structural coverage of protein-sequence space with high-accuracy models. Nucleic Acids Res. 2021 Nov 17:gkab1061. doi: 10.1093/nar/gkab1061. Epub ahead of print. PMID: 34791371.

Vashee, S., Cvetic, C., Lu, W., Simancek, P., Kelly, T.J. & Walter, J.C. (2003) Sequence-independent DNA binding and replication initiation by the human origin recognition complex. *Genes & development*. 17 (15), 1894–1908. doi:10.1101/GAD.1084203.

Vasilescu, J., Guo, X. & Kast, J. (2004) Identification of protein-protein interactions using in vivo cross-linking and mass spectrometry. *Proteomics*. 4 (12), 3845–3854. doi:10.1002/PMIC.200400856.

Venditti, P., Costanzo, G., Negri, R. & Camilloni, G. (1994) ABFI contributes to the chromatin organization of Saccharomyces cerevisiae ARS1 B-domain. *Biochimica et Biophysica Acta (BBA) - Gene Structure and Expression*. 1219 (3), 677–689. doi:10.1016/0167-4781(94)90227-5.

Versini, G., Comet, I., Wu, M., Hoopes, L., Schwob, E. & Pasero, P. (2003) The yeast Sgs1 helicase is differentially required for genomic and ribosomal DNA replication. *The EMBO journal*. 22 (8), 1939–1949. doi:10.1093/EMBOJ/CDG180.

Villa, F., Simon, A.C., Ortiz Bazan, M.A., Kilkenny, M.L., Wirthensohn, D., Wightman, M., Matak-Vinkovíc, D., Pellegrini, L. & Labib, K. (2016) Ctf4 Is a Hub in the Eukaryotic Replisome that Links Multiple CIP-Box Proteins to the CMG Helicase. *Molecular Cell*. 63 (3), 385. doi:10.1016/J.MOLCEL.2016.06.009.

Vogelauer, M., Rubbi, L., Lucas, I., Brewer, B.J. & Grunstein, M. (2002) Histone Acetylation Regulates the Time of Replication Origin Firing. *Molecular Cell*. 10 (5), 1223–1233. doi:10.1016/S1097-2765(02)00702-5.

Volk, A. & Crispino, J.D. (2015) The role of the chromatin assembly complex (CAF-1) and its p60 subunit (CHAF1b) in homeostasis and disease. *Biochimica et biophysica acta*. 1849 (8), 979. doi:10.1016/J.BBAGRM.2015.05.009.

Vujcic, M., Miller, C.A. & Kowalski, D. (1999) Activation of Silent Replication Origins at Autonomously Replicating Sequence Elements near the HML Locus in Budding Yeast. *Molecular and Cellular Biology*. 19 (9), 6098. doi:10.1128/MCB.19.9.6098.

Walter, J. & Newport, J. (2000) Initiation of eukaryotic DNA replication: origin unwinding and sequential chromatin association of Cdc45, RPA, and DNA polymerase alpha. *Molecular cell*. 5 (4), 617–627. doi:10.1016/S1097-2765(00)80241-5.

Wang, C.I., Alekseyenko, A.A., Leroy, G., Elia, A.E.H., Gorchakov, A.A., Britton, L.M.P., Elledge, S.J., Kharchenko, P. v., Garcia, B.A. & Kuroda, M.I. (2013) ChIP-mass spectrometry captures protein interactions and modified histones associated with dosage compensation in Drosophila. *Nature structural & molecular biology*. 20 (2), 202. doi:10.1038/NSMB.2477.

Wang, H. & Elledge, S.J. (2002) Genetic and physical interactions between DPB11 and DDC1 in the yeast DNA damage response pathway. *Genetics*. 160 (4), 1295–1304. doi:10.1093/GENETICS/160.4.1295.

Wang, H., Wang, M., Yang, N. & Xu, R.M. (2015) Structure of the quaternary complex of histone H3-H4 heterodimer with chaperone ASF1 and the replicative helicase subunit MCM2. *Protein & cell*. 6 (9), 693–697. doi:10.1007/S13238-015-0190-0.

Wang, J.C. (2002) Cellular roles of DNA topoisomerases: a molecular perspective. *Nature Reviews Molecular Cell Biology 2002 3:6*. 3 (6), 430–440. doi:10.1038/nrm831.

Wang, J.C. (2003) DNA TOPOISOMERASES. *https://doi.org/10.1146/annurev.bi.65.070196.003223*. 65, 635–692. doi:10.1146/ANNUREV.BI.65.070196.003223.

Wang, X., Cimermancic, P., Yu, C., Schweitzer, A., Chopra, N., et al. (2017) Molecular Details Underlying Dynamic Structures and Regulation of the Human 26S Proteasome. *Molecular & cellular proteomics : MCP*. 16 (5), 840–854. doi:10.1074/MCP.M116.065326.

Warren, E.M., Huang, H., Fanning, E., Chazin, W.J. & Eichman, B.F. (2009) Physical interactions between Mcm10, DNA, and DNA polymerase α. *Journal of Biological Chemistry*. 284 (36), 24662–24672. doi:10.1074/JBC.M109.020438/ATTACHMENT/6736DED2-317C-48CC-8714-04A35FFCE433/MMC1.PDF.

Watase, G., Takisawa, H. & Kanemaki, M.T. (2012) Mcm10 plays a role in functioning of the eukaryotic replicative DNA helicase, Cdc45-Mcm-GINS. *Current biology : CB*. 22 (4), 343–349. doi:10.1016/J.CUB.2012.01.023.

Watson, J.D. & Crick, F.H.C. (1953) Molecular Structure of Nucleic Acids: A Structure for Deoxyribose Nucleic Acid. *Nature*. 171 (4356), 737–738. doi:10.1038/171737a0.

Weaver, P.L., Sun, C. & Chang, T.H. (1997) Dbp3p, a putative RNA helicase in Saccharomyces cerevisiae, is required for efficient pre-rRNA processing predominantly at site A3. *Molecular and Cellular Biology*. 17 (3), 1354. doi:10.1128/MCB.17.3.1354.

Weinreich, M., Liang, C. & Stillman, B. (1999) The Cdc6p nucleotide-binding motif is required for loading mcm proteins onto chromatin. *Proceedings of the National Academy of Sciences of the United States of America*. 96 (2), 441–446.

Wendler, P., Ciniawsky, S., Kock, M. & Kube, S. (2012) Structure and function of the AAA+ nucleotide binding pocket. *Biochimica et biophysica acta*. 1823 (1), 2–14. doi:10.1016/J.BBAMCR.2011.06.014.

Westermeier, R. & Marouga, R. (2005) Protein Detection Methods in Proteomics Research. *Bioscience Reports*. 25 (1–2), 19–32. doi:10.1007/S10540-005-2845-1.

Wijnen, H., Landman, A. & Futcher, B. (2002) The G(1) cyclin Cln3 promotes cell cycle entry via the transcription factor Swi6. *Molecular and cellular biology*. 22 (12), 4402–4418. doi:10.1128/MCB.22.12.4402-4418.2002.

Willis, N. & Rhind, N. (2009) Regulation of DNA replication by the S-phase DNA damage checkpoint. *Cell Division*. 4 (1), 1–10. doi:10.1186/1747-1028-4-13/COMMENTS.

Wilmes, G.M. & Bell, S.P. (2002) The B2 element of the Saccharomyces cerevisiae ARS1 origin of replication requires specific sequences to facilitate pre-RC formation. *Proceedings of the National Academy of Sciences of the United States of America*. 99 (1), 101. doi:10.1073/PNAS.012578499.

Winkler, C., Denker, K., Wortelkamp, S. & Sickmann, A. (2007) Silver-and Coomassie-staining protocols: Detection limits and compatibility with ESI MS. *Electrophoresis*. 28 (12), 2095–2099. doi:10.1002/ELPS.200600670.

Wiśniewski, J.R., Zougman, A., Nagaraj, N. & Mann, M. (2009) Universal sample preparation method for proteome analysis. *Nature methods*. 6 (5), 359–362. doi:10.1038/NMETH.1322.

Wobbe, C.R., Weissbach, L., Borowiec, J.A., Dean, F.B., Murakami, Y., Bullock, P. & Hurwitz, J. (1987) Replication of simian virus 40 origin-containing DNA in vitro with purified proteins. *Proceedings of the National Academy of Sciences*. 84 (7), 1834–1838. doi:10.1073/PNAS.84.7.1834.

Wold, M.S. (1997) Replication protein A: a heterotrimeric, single-stranded DNA-binding protein required for eukaryotic DNA metabolism. *Annual review of biochemistry*. 66, 61–92. doi:10.1146/ANNUREV.BIOCHEM.66.1.61.

Xia, Y. (2021) The Fate of Two Unstoppable Trains After Arriving Destination: Replisome Disassembly During DNA Replication Termination. *Frontiers in Cell and Developmental Biology*. 9, 1772. doi:10.3389/FCELL.2021.658003/BIBTEX.

Xu, X. & Liu, Y. (2009) Dual DNA unwinding activities of the Rothmund–Thomson syndrome protein, RECQ4. *The EMBO Journal*. 28 (5), 568. doi:10.1038/EMBOJ.2009.13.

Yabuki, N., Terashima, H. & Kitada, K. (2002) Mapping of early firing origins on a replication profile of budding yeast. *Genes to cells : devoted to molecular & cellular mechanisms*. 7 (8), 781–789. doi:10.1046/J.1365-2443.2002.00559.X.

Yamada, Y., Nakagawa, T. & Masukata, H. (2004) A Novel Intermediate in Initiation Complex Assembly for Fission Yeast DNA Replication. *https://doi.org/10.1091/mbc.e04-04-0292*. 15 (8), 3740–3750. doi:10.1091/MBC.E04-04-0292.

Yang, C., Sengupta, S., Hegde, P.M., Mitra, J., Jiang, S., Holey, B., Sarker, A.H., Tsai, M.-S., Hegde, M.L. & Mitra, S. (2017) Regulation of oxidized base damage repair by chromatin assembly factor 1 subunit A. *Nucleic Acids Research*. 45 (2), 739. doi:10.1093/NAR/GKW1024.

Yao, N., Turner, J., Kelman, Z., Todd Stukenberg, P., Dean, F., Shechter, D., Pan, Z.Q., Hurwitz, J. & O'Donnell, M. (1996) Clamp loading, unloading and intrinsic stability of the PCNA, beta and gp45 sliding clamps of human, E. coli and T4 replicases. *Genes to cells : devoted to molecular & cellular mechanisms*. 1 (1), 101–113. doi:10.1046/J.1365-2443.1996.07007.X.

Yao, N.Y., Johnson, A., Bowman, G.D., Kuriyan, J. & O'Donnell, M. (2006) Mechanism of proliferating cell nuclear antigen clamp opening by replication factor C. *The Journal of biological chemistry*. 281 (25), 17528–17539. doi:10.1074/JBC.M601273200.

Yao, N.Y. & O'Donnell, M. (2010) SnapShot: The Replisome. *Cell*. 141 (6), 1088. doi:10.1016/J.CELL.2010.05.042.

Yao, N.Y. & O'Donnell, M. (2012) The RFC Clamp Loader: Structure and Function. *Sub-cellular biochemistry*. 62, 259. doi:10.1007/978-94-007-4572-8_14.

Yardimci, H., Loveland, A.B., Habuchi, S., van Oijen, A.M. & Walter, J.C. (2010) Uncoupling of Sister Replisomes during Eukaryotic DNA Replication. *Molecular Cell*. 40 (5), 834–840. doi:10.1016/J.MOLCEL.2010.11.027.

Yardimci, H., Wang, X., Loveland, A.B., Tappin, I., Rudner, D.Z., Hurwitz, J., van Oijen, A.M. & Walter, J.C. (2012) Bypass of a protein barrier by a replicative DNA helicase. *Nature 2012 492:7428*. 492 (7428), 205–209. doi:10.1038/nature11730.

Yates, L.A., Aramayo, R.J., Pokhrel, N., Caldwell, C.C., Kaplan, J.A., Perera, R.L., Spies, M., Antony, E. & Zhang, X. (2018) A structural and dynamic model for the assembly of Replication Protein A on

single-stranded DNA. *Nature Communications 2018 9:1*. 9 (1), 1–14. doi:10.1038/s41467-018-07883-7.

Yeeles, J.T.P., Deegan, T.D., Janska, A., Early, A. & Diffley, J.F.X. (2015) Regulated Eukaryotic DNA Replication Origin Firing with Purified Proteins. *Nature*. 519 (7544), 431. doi:10.1038/NATURE14285.

Yeeles, J.T.P., Janska, A., Early, A. & Diffley, J.F.X. (2017) How the Eukaryotic Replisome Achieves Rapid and Efficient DNA Replication. *Molecular cell*. 65 (1), 105–116. doi:10.1016/j.molcel.2016.11.017.

Yekezare, M., Gó mez-González, B. & Diffley, J.F.X. (2013) Controlling DNA replication origins in response to DNA damage - inhibit globally, activate locally. *Journal of cell science*. 126 (Pt 6), 1297–1306. doi:10.1242/JCS.096701.

Yella, V.R., Vanaja, A., Kulandaivelu, U. & Kumar, A. (2020) Delving into Eukaryotic Origins of Replication Using DNA Structural Features. *ACS Omega*. 5 (23), 13601–13611. doi:10.1021/ACSOMEGA.0C00441/SUPPL_FILE/AO0C00441_SI_001.PDF.

Yen, K., Vinayachandran, V., Batta, K., Koerber, R.T. & Pugh, B.F. (2012) Genome-wide nucleosome specificity and directionality of chromatin remodelers. *Cell*. 149 (7), 1461. doi:10.1016/J.CELL.2012.04.036.

Yeong, F.M., Lim, H.H., Padmashree, C.G. & Surana, U. (2000) Exit from Mitosis in Budding Yeast: Biphasic Inactivation of the Cdc28-Clb2 Mitotic Kinase and the Role of Cdc20. *Molecular Cell*. 5 (3), 501–511. doi:10.1016/S1097-2765(00)80444-X.

Yoshida, K., Bacal, J., Desmarais, D., Padioleau, I., Tsaponina, O., Chabes, A., Pantesco, V., Dubois, E., Parrinello, H., Skrzypczak, M., Ginalski, K., Lengronne, A. & Pasero, P. (2014) The Histone Deacetylases Sir2 and Rpd3 Act on Ribosomal DNA to Control the Replication Program in Budding Yeast. *Molecular Cell*. 54 (4), 691–697. doi:10.1016/J.MOLCEL.2014.04.032.

Yu, C., Gan, H., Han, J., Zhou, Z.-X., Jia, S., Chabes, A., Farrugia, G., Ordog, T. & Zhang, Z. (2014) Strand-Specific Analysis Shows Protein Binding at Replication Forks and PCNA Unloading from Lagging Strands when Forks Stall. *Molecular cell*. 56 (4), 551. doi:10.1016/J.MOLCEL.2014.09.017.

Yu, C., Gan, H., Serra-Cardona, A., Zhang, L., Gan, S., Sharma, S., Johansson, E., Chabes, A., Xu, R.M. & Zhang, Z. (2018) A mechanism for preventing asymmetric histone segregation onto replicating DNA strands. *Science (New York, N.Y.)*. 361 (6409), 1386–1389. doi:10.1126/SCIENCE.AAT8849.

Yu, C., Gan, H. & Zhang, Z. (2018) Strand-Specific Analysis of DNA Synthesis and Proteins Association with DNA Replication Forks in Budding Yeast. *Methods in Molecular Biology*. 1672, 227–238. doi:10.1007/978-1-4939-7306-4_17.

Yu, C. & Huang, L. (2018) Cross-Linking Mass Spectrometry (XL-MS): an Emerging Technology for Interactomics and Structural Biology. *Analytical chemistry*. 90 (1), 144. doi:10.1021/ACS.ANALCHEM.7B04431.

Yu, C., Yang, Y., Wang, X., Guan, S., Fang, L., Liu, F., Walters, K.J., Kaiser, P. & Huang, L. (2016) Characterization of Dynamic UbR-Proteasome Subcomplexes by In vivo Cross-linking (X) Assisted Bimolecular Tandem Affinity Purification (XBAP) and Label-free Quantitation. *Molecular & cellular proteomics : MCP*. 15 (7), 2279–2292. doi:10.1074/MCP.M116.058271.

Yuan, Z., Bai, L., Sun, J., Georgescu, R., Liu, J., O'Donnell, M.E. & Li, H. (2016) Structure of the eukaryotic replicative CMG helicase and pumpjack motion. *Nature structural & molecular biology*. 23 (3), 217. doi:10.1038/NSMB.3170.

Yuan, Z., Georgescu, R., Bai, L., Zhang, D., Li, H. & O'Donnell, M.E. (2020a) DNA unwinding mechanism of a eukaryotic replicative CMG helicase. *Nature communications*. 11 (1). doi:10.1038/S41467-020-14577-6.

Yuan, Z., Georgescu, R., Schauer, G.D., O'Donnell, M.E. & Li, H. (2020) Structure of the polymerase ε holoenzyme and atomic model of the leading strand replisome. *Nature Communications 2020 11:1*. 11 (1), 1–11. doi:10.1038/s41467-020-16910-5.

Yuan, Z., Riera, A., Bai, L., Sun, J., Nandi, S., Spanos, C., Chen, Z.A., Barbon, M., Rappsilber, J., Stillman, B., Speck, C. & Li, H. (2017) Structural basis of Mcm2-7 replicative helicase loading by ORC-Cdc6 and Cdt1. *Nature structural & molecular biology*. 24 (3), 316–324. doi:10.1038/nsmb.3372.

Yuan, Z., Schneider, S., Dodd, T., Riera, A., Bai, L., Yan, C., Magdalou, I., Ivanov, I., Stillman, B., Li, H. & Speck, C. (2020d) Structural mechanism of helicase loading onto replication origin DNA by ORC-Cdc6. *Proceedings of the National Academy of Sciences of the United States of America*. 117 (30), 17747–17756. doi:10.1073/PNAS.2006231117.

Yurov, Y.B. & Liapunova, N.A. (n.d.) The Units of DNA Replication in the Mammalian Chromosomes: Evidence for a Large Size of Replication Units. *Chromosoma (Berl.)*. 60, 253–267.

Yuzhakov, A., Kelman, Z., Hurwitz, J. & O'Donnell, M. (1999) Multiple competition reactions for RPA order the assembly of the DNA polymerase delta holoenzyme. *The EMBO journal*. 18 (21), 6189–6199. doi:10.1093/EMBOJ/18.21.6189.

Zappulla, D.C., Sternglanz, R. & Leatherwood, J. (2002) Control of Replication Timing by a Transcriptional Silencer. *Current Biology*. 12 (11), 869–875. doi:10.1016/S0960-9822(02)00871-0.

Zargari, A., Boban, M., Heessen, S., Andréasson, C., Thyberg, J. & Ljungdahl, P.O. (2007) Inner nuclear membrane proteins Asi1, Asi2, and Asi3 function in concert to maintain the latent properties of transcription factors Stp1 and Stp2. *The Journal of biological chemistry*. 282 (1), 594–605. doi:10.1074/JBC.M609201200.

Zegerman, P. & Diffley, J.F.X. (2010) Checkpoint-dependent inhibition of DNA replication initiation by Sld3 and Dbf4 phosphorylation. *Nature*. 467 (7314), 474–478. doi:10.1038/NATURE09373.

Zegerman, P. & Diffley, J.F.X. (2009) DNA replication as a target of the DNA damage checkpoint. *DNA Repair*. 8 (9), 1077–1088. doi:10.1016/J.DNAREP.2009.04.023.

Zegerman, P. & Diffley, J.F.X. (2007) Phosphorylation of Sld2 and Sld3 by cyclin-dependent kinases promotes DNA replication in budding yeast. *Nature*. 445 (7125), 281–285. doi:10.1038/NATURE05432.

Zellner, E., Herrmann, T., Schulz, C. & Grummt, F. (2007) Site-specific interaction of the murine pre-replicative complex with origin DNA: assembly and disassembly during cell cycle transit and differentiation. *Nucleic acids research*. 35 (20), 6701–6713. doi:10.1093/NAR/GKM555.

Zhai, Y., Cheng, E., Wu, H., Li, N., Yung, P.Y.K., Gao, N. & Tye, B.-K. (2017) Open-ringed structure of the Cdt1–Mcm2–7 complex as a precursor of the MCM double hexamer. *Nature Structural & Molecular Biology 2017 24:3*. 24 (3), 300–308. doi:10.1038/nsmb.3374.

Zhang, H., Petrie, M. v., He, Y., Peace, J.M., Chiolo, I.E. & Aparicio, O.M. (2019) Dynamic relocalization of replication origins by fkh1 requires execution of ddk function and cdc45 loading at origins. *eLife*. 8. doi:10.7554/ELIFE.45512.

Zhang, L. & Elias, J.E. (2017) Relative Protein Quantification Using Tandem Mass Tag Mass Spectrometry. *Methods in Molecular Biology*. 1550, 185–198. doi:10.1007/978-1-4939-6747-6_14.

Zhang, W., Feng, J. & Li, Q. (2020) The replisome guides nucleosome assembly during DNA replication. *Cell & Bioscience 2020 10:1*. 10 (1), 1–14. doi:10.1186/S13578-020-00398-Z.

Zhang, Z., Shibahara, K. & Stillman, B. (2000) PCNA connects DNA replication to epigenetic inheritance in yeast. *Nature 2000 408:6809*. 408 (6809), 221–225. doi:10.1038/35041601..

Zhao, J., Kessler, M., Helmling, S., O'Connor, J.P. & Moore, C. (1999) Pta1, a component of yeast CF II, is required for both cleavage and poly(A) addition of mRNA precursor. *Molecular and cellular biology*. 19 (11), 7733–7740. doi:10.1128/MCB.19.11.7733.

Zhou, J., Monson, E.K., -C Teng, S., Schulz, V.P. & Zakian, V.A. (2000) Pif1p Helicase, a Catalytic Inhibitor of Telomerase in Yeast. *J. Holdsworth, Plant J*.5. https://www.science.org.

Zhou, J.C., Janska, A., Goswami, P., Renault, L., Ali, F.A., Kotecha, A., Diffley, J.F.X. & Costa, A. (2017) CMG-Pol epsilon dynamics suggests a mechanism for the establishment of leading-strand synthesis in the eukaryotic replisome. *Proceedings of the National Academy of Sciences of the United States of America*. 114 (16), 4141–4146. doi:10.1073/PNAS.1700530114/-/DCSUPPLEMENTAL.

Zhu, W., Smith, J.W. & Huang, C.M. (2010) Mass Spectrometry-Based Label-Free Quantitative Proteomics. *Journal of Biomedicine and Biotechnology*. 2010. doi:10.1155/2010/840518.

Zhu, W., Ukomadu, C., Jha, S., Senga, T., Dhar, S.K., Wohlschlegel, J.A., Nutt, L.K., Kornbluth, S. & Dutta, A. (2007) Mcm10 and And-1/CTF4 recruit DNA polymerase α to chromatin for initiation of DNA replication. *Genes & Development*. 21 (18), 2288. doi:10.1101/GAD.1585607.

Zill, O.A., Scannell, D., Teytelman, L. & Rine, J. (2010) Co-Evolution of Transcriptional Silencing Proteins and the DNA Elements Specifying Their Assembly. *PLoS Biology*. 8 (11), 1000550. doi:10.1371/JOURNAL.PBIO.1000550.

Zlatanova, J., Leuba, S.H. & van Holde, K. (1998) Chromatin Fiber Structure: Morphology, Molecular Determinants, Structural Transitions. *Biophysical Journal*. 74 (5), 2554–2566. doi:10.1016/S0006-3495(98)77963-9.

Zou, L. & Elledge, S.J. (2003) Sensing DNA damage through ATRIP recognition of RPA-ssDNA complexes. *Science (New York, N.Y.)*. 300 (5625), 1542–1548. doi:10.1126/SCIENCE.1083430.

Zou, L., Mitchell, J. & Stillman, B. (1997) CDC45, a novel yeast gene that functions with the origin recognition complex and Mcm proteins in initiation of DNA replication. *Molecular and Cellular Biology*. 17 (2), 553. doi:10.1128/MCB.17.2.553.

Zou, L. & Stillman, B. (2000) Assembly of a complex containing Cdc45p, replication protein A, and Mcm2p at replication origins controlled by S-phase cyclin-dependent kinases and Cdc7p-Dbf4p kinase. *Molecular and cellular biology*. 20 (9), 3086–3096. doi:10.1128/MCB.20.9.3086-3096.2000.

Zou, L. & Stillman, B. (1998) Formation of a Preinitiation Complex by S-phase Cyclin CDK-Dependent Loading of Cdc45p onto Chromatin. *Science*. 280 (5363), 593–596. doi:10.1126/SCIENCE.280.5363.593.