Investigating the Mechanism of Activation of the Mcm2-7 Replicative Helicase

Thomas Deegan

University College London
and
Cancer Research UK London Research Institute
PhD Supervisor: Dr. John Diffley

A thesis submitted for the degree of
Doctor of Philosophy
University College London
September 2014

Declaration

I Thomas Deegan confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Abstract

DNA replication initiation is co-ordinated with progression through the cell cycle via a two-step mechanism. The first step, termed origin licensing, involves the assembly of a pre-replicative complex (pre-RC) at origins of replication, in which the Mcm2-7 replicative helicase is loaded onto DNA in an inactive form. Origin unwinding and DNA synthesis is only initiated during the second step of this process, origin firing, which requires the recruitment of multiple 'firing factors', such as Sld3/7 and Cdc45, as well as the activity of two essential cell cycle regulated kinases, cyclin-dependent kinase (CDK) and Dbf4-dependent kinase (DDK).

Whilst studies of the mechanism of origin licensing have been greatly furthered by the availability of an *in vitro* pre-RC assembly assay, the subsequent activation of the replicative helicase has not been well characterised, and how the firing factor proteins catalyse the complex topological changes required for Mcm2-7 helicase activation is unknown.

In this study, I used an *in vitro* biochemical approach to investigate the function of DDK and Sld3/7 in origin firing. I reconstituted the recruitment of Sld3/7 to the pre-RC *in vitro*, which is thought to be the first step during Mcm2-7 helicase activation. I observed recruitment of Sld3/7 to be dependent on phosphorylation of the loaded Mcm2-7 complex by DDK, and was subsequently able to map the Mcm2-7 binding activity of Sld3/7 to a central domain of Sld3. By isolating a number of point mutants in Sld3 that were specifically defective in Mcm2-7 binding, I showed that the Sld3-Mcm2-7 interaction is essential for replicative helicase activation. Furthermore, I showed that the central portion of Sld3 contains a Cdc45 interacting site, which is also required for efficient replication initiation.

Subsequently, I showed that Sld3 can interact with both Mcm6 and Mcm4 in a phosphorylation-dependent manner. Sld3 itself was shown to contain a novel phosphopeptide binding activity, and can interact with numerous phosphorylated residues throughout the N-terminal half of Mcm6. Elimination of these

phosphorylation sites resulted in defects in both Sld3/7 recruitment to the pre-RC and replicative helicase activation.

Thus, the novel DDK-dependent Sld3-Mcm2-7 interaction described in this study helps to explain the function of DDK during the early stages of origin firing. Additionally, the observation that Sld3 can interact directly with phosphorylated residues on Mcm2-7 indicates that Sld3 is a reader of DDK activity. As both an essential CDK substrate and a DDK reader, Sld3 thus functions as a point of intersection for the activities of CDK and DDK during Mcm2-7 helicase activation.

Acknowledgement

Firstly, I would like to thank my supervisor, Dr. John Diffley, for the guidance and inspiration he has consistently offered to me throughout my PhD. I have benefited greatly from the unique perspective John takes on most scientific problems, and very much hope to carry the lessons learnt under John's supervision with me for the remainder of my time in scientific research. Thanks should also go to the members of my thesis committee, Drs. Simon Boulton, Helle Ulrich and Peter Cherepanov, for the useful advice and discussion they have provided on numerous occassions over the past four years.

Within the London Research Institute, thanks should go to all members of the Protein Analysis and Proteomics group, for performing the mass spectrometry analysis presented in this thesis, as well as to Nicola O'Reilly and the Peptide Synthesis Laboratory, for producing the large number of peptides used in this study. Additonally, I would like to thank Ali and Namita in Fermentation Services, for growing up the many hundreds of litres of yeast needed during this project.

Special thanks are extended to all previous and current members of the Diffley lab, who have provided a relaxed, fun and stimulating environment within which to complete my PhD. I would like to particularly thank Max Douglas, who took me under his wing when I didn't know which way round to hold a pipette, and taught me how to think creatively about scientific problems. Additionally, I thank Stephanie Carter and Anne Early for the seemingly endless patience and support they have shown whilst working next to me for the majority of my time at Clare Hall.

Thanks go to my family and friends, many of whom have kept me on an even keel when I've flown too close to the edge over the past four years. Special thanks go to my sister, Rachel, for blazing a trail as the first Dr. Deegan, and to my Dad, for always reminding me to work less!

Finally, I thank Rach, my wife, for convincing me to do a PhD when I didn't know where to turn, for spending many an evening listening to my agonising over

experiments and, most importantly, for making me feel strong and free in the face of adversity.

This thesis is dedicated to the memory of my Mum, Angela, who first encouraged my love of science as a child, and remains a source of inspiration to this day.

Table of Contents

Abstract **3**
Table of Contents **7**
Table of figures **11**
List of tables **13**
Abbreviations **14**
Chapter 1. Introduction **17**
1.1 Overview **17**
1.2 Origins of Replication **20**
1.3 Origin Licensing **22**
1.3.1 Components of the pre-RC 23
1.3.2 Mechanism of pre-RC assembly 30
1.3.3 Regulation of origin licensing 35
1.4 Origin firing **36**
1.4.1 Firing factor proteins 40
1.4.2 Regulation of origin firing 46
1.4.3 Mechanism of origin firing 50
1.5 Thesis summary **53**
Chapter 2. Materials & Methods **55**
2.1 Enzymes and Reagents **55**
2.1.1 Enzymes and proteins 55
2.1.2 Antibodies 55
2.2 Media **56**
2.2.1 Media for E. coli cells 56
2.2.2 Media for yeast cells 56
2.3 Buffers and Solutions **57**
2.3.1 Buffers for general manipulation of DNA 57
2.3.2 Buffers for general manipulation of proteins 57
2.3.3 Buffers for protein purification 57
2.3.4 Buffers for *in vitro* biochemical assays 58
2.4 Plasmids **59**
2.5 DNA Oligonucleotides **66**
2.6 Strains **68**
2.6.1 *E. Coli* strains 68
2.6.2 Yeast strains 69
2.7 Molecular Biology Methods for *E. coli* **72**
2.7.1 Transformation of *E. coli* with heat-shock method 72
2.7.2 Isolation of plasmid DNA 72
2.8 Molecular Biology Methods for Yeast **73**
2.8.1 Transformation of yeast cells 73
2.8.2 Isolation of genomic DNA from yeast 73
2.8.3 Sporulation and tetrad dissection 74
2.8.4 TCA extraction of protein 74
2.9 General Methods for Manipulation of DNA **75**
2.9.1 DNA standards 75
2.9.2 Determination of DNA concentration 75

2.9.3 Agarose gel electrophoresis 75
2.9.4 Visualisation of DNA by gel staining 75
2.9.5 Phosphorimager analysis.. 76
2.10 General Methods for Manipulation of Proteins 76
2.10.1 Molecular weight standards 76
2.10.2 SDS-PAGE 76
2.10.3 Determination of protein concentration 77
2.10.4 Coomassie blue staining.. 77
2.10.5 Silver staining... 77
2.10.6 Immunoblotting 77
2.11 Protein Purification 78
2.11.1 Purification of ORC 78
2.11.2 Purification of Cdc6.. 79
2.11.3 Purification of Mcm2-7/Cdt1.. 80
2.11.4 Purification of DDK... 82
2.11.5 Purification of Sld3/7 from *S. cerevisiae* 83
2.11.6 Purification of Sld3 and Sld7 from *E. coli* 84
2.11.7 Purification of Dpb11.. 85
2.11.8 Purification of Sld2 85
2.11.9 Purification of Cdc45.. 86
2.12 Preparation of S-phase Whole Cell Extract 87
2.13 Preparation of DNA Substrates for *In Vitro* Biochemical Assays..... 88
2.13.1 1 kb photocleavable linear DNA beads 88
2.13.2 Purification of plasmid DNA template for in vitro replication assays . 88
2.14 Protein-Protein and Protein-DNA interaction Assays....................... 89
2.14.1 *In vitro* pre-RC assembly assay 89
2.14.2 Sld3/7 recruitment assay 89
2.14.3 Replisome assembly assay 90
2.14.4 Interaction of Sld3/7 with individual Mcm subunits 91
2.14.5 Immunoprecipitation of FLAG-Sld3 from S-phase extract 92
2.14.6 Interaction of Sld3 with Mcm6 peptides 92
2.14.7 Interaction of Sld3 and Sld7 93
2.15 *In Vitro* DNA Replication Assay 93
Chapter 3. Results 1 95
3.1 Introduction 95
3.2 Results 96
3.2.1 Strain construction 96
3.2.2 Optimisation of extract conditions 96
3.2.3 Purification of Dpb11, Sld2 and Cdc45 98
3.2.4 Purification of Sld3/7 102
3.2.5 Characterisation of the Sld3/7 complex 102
3.2.6 Purification of pre-RC proteins.. 107
3.2.7 Purification of DDK.. 109
3.2.8 Reconstitution of Sld3/7 recruitment to the pre-RC *in vitro* 111
3.3 Discussion 113
3.3.1 Towards reconstitution of Mcm2-7 activation *in vitro* 113
3.3.2 Sld3/7 contains multiple copies of Sld3 114

3.3.3 DDK directly regulates Sld3/7 binding to the Mcm2-7 double hexamer 115
Chapter 4. Results 2 117
4.1 Introduction 117
4.2 Results 118
4.2.1 The N-terminus of Sld3 is required for Sld7 recruitment to the pre-RC 118
4.2.2 Mapping the Mcm2-7 binding site on Sld3 119
4.2.3 Isolation of Mcm2-7 binding mutants in Sld3 120
4.2.4 One Mcm2-7 binding site on Sld3/7 is sufficient for recruitment to the pre-RC 124
4.2.5 Mcm2-7 binding mutants of Sld3 are defective for DNA replication *in vitro* 126
4.2.6 Mcm2-7 binding mutants of Sld3 are defective for CMG assembly *in vitro* 128
4.2.7 Cdc45 is not required for Sld3 recruitment *in vitro* 130
4.2.8 Mcm2-7 binding mutations in Sld3 affect cell viability 131
4.2.9 Mutations within the Sld3-Treslin Domain of Sld3 specifically disrupt Cdc45 binding 132
4.3 Discussion 137
4.3.1 A function for Sld7 during DNA replication initiation? 137
4.3.2 Sld3 residues 251-585 are functional for both Cdc45 and Mcm2-7 binding during DNA replication initiation 138
4.3.3 Mcm2-7 binding mutants of Sld3 are recruited to the pre-RC in an S-phase extract 139
Chapter 5. Results 3 141
5.1 Introduction 141
5.2 Results 142
5.2.1 Interaction of Sld3/7 with individual Mcm subunits 142
5.2.2 Mapping the interaction between Sld3/7 and Mcm6 145
5.2.3 Sld3 interacts with multiple phosphopeptides in the N-terminal portion of Mcm6 147
5.2.4 Identification of phosphorylation sites in Mcm6 151
5.2.5 The N-terminus of Mcm6 is required for Sld3/7 binding 153
5.2.6 Mapping a second Sld3/7 binding site on Mcm6 155
5.2.7 Sld3/7 binding mutants in Mcm6 are defective for DNA replication 157
5.3 Discussion 159
5.3.1 The extreme N-terminus of Mcm6 is an Sld3/7 binding site 159
5.3.2 Implications for the function of DDK during origin firing 160
Chapter 6. Discussion 163
6.1 Introduction 163
6.2 Key Conclusions 163
6.3 The function of Sld3 during origin firing 165
6.4 Convergence of the DDK and CDK-regulated pathways of replicative helicase activation 168
6.5 The function of DDK during origin firing 170
6.6 A model for replicative helicase activation 172
Chapter 7. Appendix 175

7.1 Peptides used in Mcm6 phosphopeptide arrays 175
7.1.1 Part 1 175
7.1.2 Part 2 185
7.1.3 Part 3 200
7.2 Phosphopeptides used in peptide pulldowns 202
Reference List 203

Table of figures

Figure 1.1 Cell cycle regulation of DNA replication 18
Figure 1.2 Domain architecture and subunit organisation of the *S. cerevisiae* Mcm2-7 complex 29
Figure 1.3 Pre-RC assembly produces an Mcm2-7 double hexamer bound around dsDNA 31
Figure 1.4 A model for the mechanism of origin licensing in *S. cerevisiae*............. 34
Figure 1.5 Activation of the Mcm2-7 replicative helicase during origin firing 39
Figure 1.6 Model for multi-step assembly of the pre-initiation complex (pre-IC) 52
Figure 3.1 Construction of *S. cerevisiae* strains for firing factor purification........... 97
Figure 3.2 Solubility of firing factor proteins in cell lysates 98
Figure 3.3 Purification of Dpb11, Sld2 and Cdc45... 100
Figure 3.4 Gel filtration analysis of Cdc45 101
Figure 3.5 Purification of Sld3/7.. 102
Figure 3.6 Molecular weight estimation of Sld3/7 104
Figure 3.7 Purification of Sld3 and Sld7 after expression in *E. coli* 105
Figure 3.8 *In vitro* reconstitution of the Sld3/7 complex with purified proteins...... 107
Figure 3.9 Purification of pre-RC proteins 108
Figure 3.10 Purification of DDK 110
Figure 3.11 Reconstitution of Sld3/7 recruitment to the pre-RC *in vitro* 112
Figure 4.1 The N-terminus of Sld3 is required for Sld7 recruitment to the pre-RC 119
Figure 4.2 Sld3 residues 251-585 are necessary and sufficient for recruitment to the pre-RC 121
Figure 4.3 Purification of Sld3 Mcm2-7 binding mutants 123
Figure 4.4 Mutations in Sld3 residues 511-531 specifically disrupt binding to the pre-RC 124
Figure 4.5 One Mcm2-7 binding site on Sld3/7 is sufficient for recruitment to the pre-RC 125
Figure 4.6 Mcm2-7 binding mutants of Sld3 are defective for DNA replication *in vitro* 127

Figure 4.7 Mcm2-7 binding mutants of Sld3 are defective for CMG assembly *in vitro* 129
Figure 4.8 Sld3 recruitment to the pre-RC is independent of Cdc45 *in vitro* 131
Figure 4.9 Mcm2-7 binding mutations in Sld3 affect cell viability 132
Figure 4.10 Testing Sld3 S.T.D. mutants for binding to the pre-RC 133
Figure 4.11 Purification of Sld3 S.T.D. mutants... 134
Figure 4.12 Mutations in Sld3 residues 296-320 disrupt binding to both the pre-RC and Cdc45 135
Figure 4.13 Cdc45 binding mutants of Sld3 are defective for DNA replication *in vitro* 136
Figure 5.1 The Mcm2-7 double hexamer is disrupted by prolonged incubation in 0.5 M NaCl 143
Figure 5.2 Sld3/7 interacts with phosphorylated Mcm6 and Mcm4 144
Figure 5.3 Schematic of TEV-cleavable Mcm6TEV2.. 145
Figure 5.4 Mcm6 1-486 is necessary and sufficient for Sld3/7 binding 146
Figure 5.5 Sld3 can bind to multiple phosphopeptides in Mcm6 1-486 148
Figure 5.6 Interactions of Sld3 with 30-residue Mcm6 phosphopeptides 150
Figure 5.7 Preparation of phosphorylated Mcm2-7 for mass spectrometry analysis 151
Figure 5.8 Summary of phosphorylation sites identified in Mcm6 152
Figure 5.9 The N-terminus of Mcm6 is required for Sld3/7 binding 154
Figure 5.10 The Sld3/7-Mcm6 interaction is disrupted by mutation of multiple distinct sites in Mcm6.. 156
Figure 5.11 The interaction between Mcm6 and Sld3/7 is required for efficient DNA replication initiation *in vitro*.. 158
Figure 6.1 Sld3 interacts with multiple partner proteins during replication initiation 167
Figure 6.2 Treslin contains conserved basic residues between the S.T.D. and CDK phosphorylation sites 168
Figure 6.3 DDK and CDK-dependent regulation of origin firing converges at Sld3 169
Figure 6.4 A model for Sld3/7-dependent CMG formation 174

List of tables

Table 2-1 Enzymes used in this study 55
Table 2-2 Primary monoclonal antibodies used in this study........ 55
Table 2-3 Primary polyclonal antibodies used in this study 55
Table 2-4 Secondary antibodies used in this study 56
Table 2-5 Plasmids used in this study 59
Table 2-6 Plasmids generated in this study. 60
Table 2-7 Oligonucleotides used in this study 66
Table 2-8 *E. coli* strains used in this study 68
Table 2-9 Yeast strains used in this study. 69
Table 5-1 Summary of Mcm6 phosphopeptides that bind to Sld3 149
Table 5-2 Identification of phosphorylation sites in Mcm6 153

Abbreviations

AAA+	ATPases associated with diverse cellular activities
Abf	ARS binding factor
ACS	ARS consensus sequence
APC/C	anaphase promoting complex / cyclosome
ARS	autonomously replicating sequence
ATP	adenosine 5'-triphosphate
ATPγS	adenosine 5'-O-(3-thio)triphosphate
Bp	base pairs
BRCT	BRCA1 C-terminus
BSA	bovine serum albumin
CBP	calmodulin binding peptide
Cdc	cell division cycle
Cdh	Cdc20 homologue
CDK	cyclin-dependent kinase
Cdt	Cdc10 dependent transcript
Clb	cyclin B
CMG	Cdc45-Mcm2-7-GINS
Ctf	chromosome transmission fidelity
Dbf	dumbbell former
DDK	Dbf4-dependent kinase
DNA	deoxyribonucleic acid
dNTP	deoxynucleoside triphosphate
Dpb	DNA polymerase B possible subunit
Drc	DNA replication checkpoint protein
dsDNA	double stranded DNA
DTT	dithiothreitol
EDTA	ethylenediaminetetraacetic acid
EGTA	ethylene glycol tetraacetic acid
EM	electron microscopy
GINS	Go-Ichi-Ni-San (Sld5-Psf1-Psf2-Psf3)
GST	glutathione S-transferase

HRP	horseradish peroxidase
IgG	immunoglobulin G
IPTG	isopropyl β-D-1-thiogalactopyranoside
Kb	kilobase
MBF	multiprotein bridging factor
Mcm	minichromosome maintenance
Mec	mitosis entry checkpoint
Mrc	mediator of the replication checkpoint
MTBP	Mdm2 binding protein
NTP	nucleoside triphosphate
OB	oligosaccharide binding
OCM	ORC-Cdc6-Mcm2-7
Orc	origin recognition complex
PBS	phosphate buffered saline
PCNA	proliferating cell nuclear antigen
PEG	polyethylene glycol
P_i	inorganic phosphate
Pol	polymerase
pre-IC	pre-initiation complex
pre-LC	pre-loading complex
pre-RC	pre-replicative complex
Psf	partner of Sld5
Rad	radiation sensitive
RecQL	RecQ protein-like
Rif	RAP1 interacting factor
RLF	replication licensing factor
RPA	replication protein A
RPC	replisome progression complex
rpm	revolutions per minute
SDS-PAGE	sodium dodecyl sulphate polyacrylamide gel electrophoresis
Sld	synthetically lethal with *dpb11-1*
Sic	substrate inhibitor of cyclin dependent kinase
ssDNA	single stranded DNA
STD	Sld3/Treslin domain

TCP	TEV-CBP-protein A
TEV	tobacco etch virus
TICRR	TopBP1 interacting checkpoint and replication regulator
TopBP	topoisomerase II binding protein

Chapter 1. Introduction

1.1 Overview

An essential requirement during the proliferation of any cell is the ability to duplicate their genomes, to allow genetic information to be propagated to new daughter cells. In eukaryotic cells, the completion of genome duplication is co-ordinated with progression through the cell cycle, which in turn is driven by oscillations in the activity of the cyclin-dependent kinases (CDKs). Genome duplication occurs primarily during S (synthetic) phase of the cell cycle, before chromosome segregation and cell division ensues in the subsequent M (mitosis) phase.

The genetic information present in every living cell is stored in the DNA double helix. Notably, eukaryotic genomes are much larger than their bacterial counterparts; the DNA in the 46 chromosomes present in most of our body's cells has a combined length of approximately 2 m. Importantly, every section of these large genomes must be duplicated only once during each S-phase, as over- or under-replication can lead to gene amplification or gross chromosomal rearrangements, which in turn can instigate genome instability and tumorigenesis (Arias and Walter, 2007, Blow and Gillespie, 2008). Thus, the cell not only needs DNA replication to occur on a remarkable scale and at great frequency, but also requires that high fidelity is maintained during the genome duplication process.

Given these requirements, genome duplication is tightly regulated, primarily during the initiation phase of DNA replication. Whilst bacterial genomes can be duplicated with sufficient speed from a single site (termed an origin of replication), the larger size of eukaryotic genomes demands that replication initiates from numerous distinct sites positioned on multiple chromosomes. DNA replication initiation at these origins of replication proceeds via a conserved two-step mechanism, as depicted in Fig. 1.1 (Blow, 1993, Blow and Laskey, 1988, Diffley et al., 1994). The first step, termed origin licensing, occurs in late mitosis and G1 phase of the cell cycle, and involves the assembly of pre-replicative complexes (pre-RCs) at the origin. This step can only occur when the activity of CDK is low and that of the

Anaphase Promoting Complex / Cyclosome (APC/C) E3 ubiquitin ligase is high (Arias and Walter, 2007, Siddiqui et al., 2013). Upon progression into S-phase, CDK activity increases and the APC/C is inactivated, which facilitates the second step of replication initiation, origin firing. It is only during this second step that DNA synthesis is initiated. Importantly, the high CDK activity present during S-phase is inhibitory for origin licensing, which provides a potent block to re-replication by preventing the assembly of pre-RCs onto replication origins from which DNA synthesis has already initiated (Diffley, 2011, Arias and Walter, 2007). Metazoans have evolved additional CDK-independent mechanisms to prevent DNA re-replication (Blow and Dutta, 2005, Arias and Walter, 2007), as will be discussed in section 1.3.3.

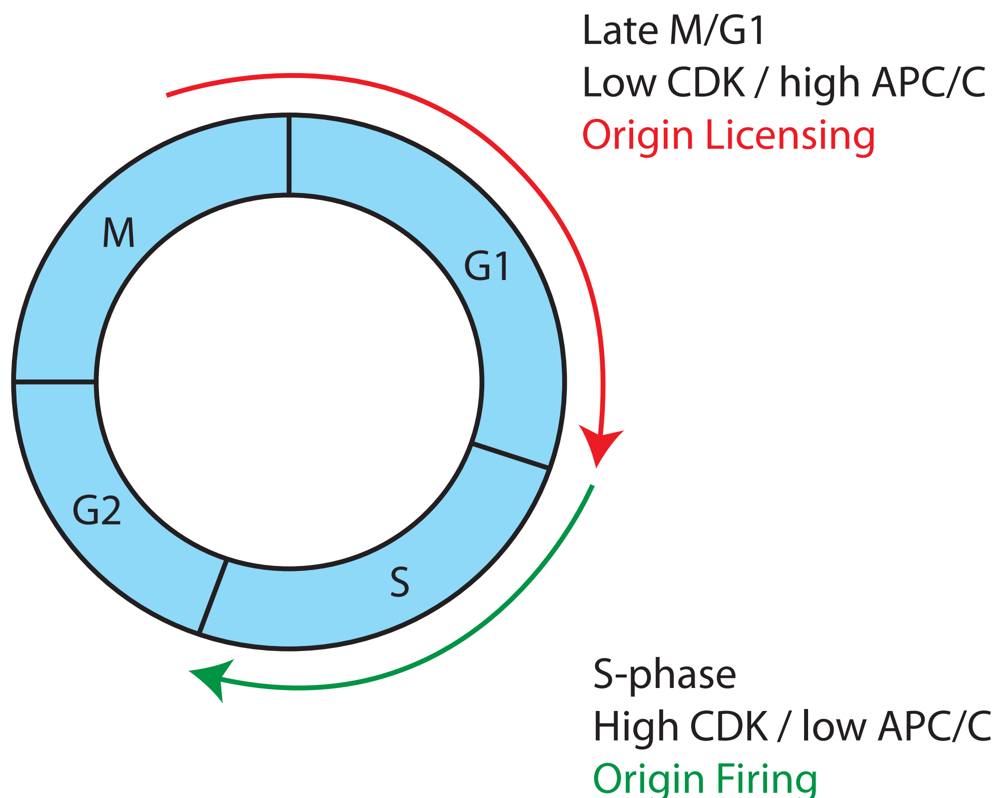


Figure 1.1 Cell cycle regulation of DNA replication

The eukaryotic cell cycle is composed of four distinct phases: G1, S, G2 and M. The initiation of DNA replication is co-ordinated with progression through the cell cycle via a two-step mechanism, which is controlled by the activities of CDK and the APC/C. Origin licensing and origin firing are mutually exclusive steps, which prevents re-replication. The details of this mechanism are first discussed in section 1.1 of the text, and are reviewed throughout the introduction.

On a molecular level, the enzyme that underpins this complex two-step mechanism is the Minichromosome-maintenance (MCM) replicative helicase (Mcm2-7). This hetero-hexameric enzyme is responsible for unwinding the two strands of the DNA double helix, to allow the replicative DNA polymerases access to the single-stranded DNA (ssDNA) template required for DNA synthesis. Origin licensing, or pre-RC assembly, involves the loading of this Mcm2-7 complex onto DNA at origins of replication. This reaction requires the initial binding of the Origin Recognition Complex (ORC) to the origin; this protein complex in turn collaborates with two other factors, Cdt1 and Cdc6, to facilitate the stable association of Mcm2-7 with DNA. The origin licensing reaction has been reconstituted *in vitro* with purified proteins; these and other studies indicate that the product of the pre-RC assembly step is a double hexamer of Mcm2-7 bound around double-stranded DNA (dsDNA) (Remus et al., 2009, Evrin et al., 2009, Gambus et al., 2011, Gillespie et al., 2001).

The Mcm2-7 double hexamer produced during pre-RC assembly is inactive as a DNA helicase. It is only upon progression into S-phase, when the activities of CDK and a second protein kinase, Dbf4-dependent kinase (DDK), increase, that the helicase function of Mcm2-7 is activated, and origin unwinding and subsequent DNA synthesis (or origin firing) commences. Origin firing requires the extensive remodelling of the Mcm2-7 double hexamer into the Cdc45-MCM-GINS (CMG) complex, which contains a single copy each of Mcm2-7, Cdc45 and the tetrameric GINS complex (Gambus et al., 2006, Costa et al., 2011, Fu et al., 2011), and functions as the active DNA helicase at replication forks (Ilves et al., 2010, Pacek et al., 2006, Moyer et al., 2006). Notably, formation of the CMG complex requires the recruitment of various ‘firing factors’ to the origin of replication (see section 1.4.1); these factors have been proposed to form a transient pre-initiation complex (pre-IC) on the Mcm2-7 double hexamer (Zou and Stillman, 1998, Tanaka and Araki, 2013). Although the identity of many firing factors is known, little is understood of how they contribute towards activation of the Mcm2-7 complex, or how the pre-IC assembles at origins of replication.

Following CMG assembly, additional proteins such as Mrc1 and Ctf4 are also recruited, forming a multi-protein assembly called the Replisome Progression Complex (RPC) or replisome (Gambus et al., 2006). Origin unwinding facilitates the

recruitment of the Polα/Primase complex to ssDNA, and primer synthesis is then initiated. At some juncture, Polε and Polδ take responsibility for DNA synthesis on the leading and lagging strand templates (Kunkel and Burgers, 2008), respectively, and bi-directional DNA synthesis can thereby be initiated from the origin of replication.

Work published in the last decade has contributed greatly towards our understanding of how CDK and DDK regulate Mcm2-7 helicase activation at the G1/S transition. Two firing factors, Sld2 and Sld3, were identified as the minimal CDK targets required for origin firing *in vivo*; phosphorylation of these two proteins facilitates their binding to a third firing factor, Dpb11 (Zegerman and Diffley, 2007, Tanaka et al., 2007, Masumoto et al., 2002). In contrast, DDK is known to phosphorylate the Mcm2-7 helicase itself. A number of genetic alleles that can bypass the requirement for DDK activity during replication initiation have been identified (Sheu and Stillman, 2010, Hardy et al., 1997, Hayano et al., 2012, Matsumoto et al., 2011). More recently, DDK has been proposed to function during the early stages of origin firing, promoting the recruitment of Cdc45 and Sld3 (in complex with another firing factor, Sld7) to origins of replication (Yabuuchi et al., 2006, Heller et al., 2011, Tanaka et al., 2011a). In spite of this, the exact mechanism by which DDK phosphorylation promotes activation of the replicative helicase is still poorly understood.

In this introduction, I will review our existing knowledge of both the regulation and mechanism of Mcm2-7 helicase loading and activation. I will primarily focus on the description of these events in *S. cerevisiae*, the model organism utilised in this study. Initially, I will briefly explain what is known about origins of replication, the chromosomal loci from which DNA replication initiates in every cell cycle.

1.2 Origins of Replication

The replicon model, proposed by Jacob and Brenner in 1963, postulated that DNA replication would begin when a trans-acting factor, termed the initiator, recognised a specific cis-acting DNA sequence in the genome, termed the replicator (Jacob et al., 1963). The initiator protein would then recruit the factors necessary for

unwinding of the double helix, replisome assembly and DNA synthesis. This model was initially proposed for simple bacterial systems, and was subsequently proved to be correct in *E. coli*, with the discovery of *oriC* (the replicator or origin of replication) in 1977 (Yasuda and Hirota, 1977), and DnaA, the initiator protein that binds to *oriC*, in 1983 (Fuller and Kornberg, 1983). The subsequent discovery of very similar mechanisms in both bacteriophage and eukaryotic viral systems established the replicon model as a paradigm for DNA replication initiation.

As mentioned in the previous section, a prerequisite for the duplication of the large genomes of eukaryotic organisms is the use of multiple origins of replication. The number of origins used is generally relational to genome size and cell cycle length (Leonard and Mechali, 2013). Thus, if all origins were to fire simultaneously, the length of S-phase would be proportional to the inter-origin distance, as opposed to overall genome size, thereby allowing for the timely completion of DNA replication before cells enter mitosis. Notably, in most eukaryotes, replication origins are not defined by a unique DNA sequence, contrary to the situation in simpler organisms. Instead, factors such as DNA topology and chromatin status appear to determine origin activity (Remus et al., 2004, Aggarwal and Calvi, 2004, Eaton et al., 2010). Additionally, origin usage can vary tremendously according to conditions such as the developmental state of a given organism. For example, in the early embryonic stages of *Xenopus* development, DNA synthesis initiates from approximately 300,000 sites as close as 5 kb apart (Blow et al., 2001), and any exogenous DNA introduced at this stage can be utilised as a site of replication initiation (Harland and Laskey, 1980).

S. cerevisiae is thus unique amongst eukaryotes, as its origins of replication share readily distinguishable common sequence elements. These origin sequences, termed Autonomously Replicating Sequences (ARSs), were initially identified as elements that could confer the ability to be maintained extrachromosomally to otherwise non-replicating plasmids (Hsiao and Carbon, 1979, Stinchcomb et al., 1979). The subsequent physical mapping of replication initiation to these sites confirmed their relevance as replication origins *in vivo* (Brewer and Fangman, 1987, Huberman et al., 1987).

Recent estimates suggest the presence of 350-400 replication origins in the *S. cerevisiae* genome (Raghuraman et al., 2001, Wyrick et al., 2001, Feng et al., 2006). In *S. cerevisiae*, individual origins are activated at a distinct and reproducible time during each S-phase, and are referred to as early or late firing origins accordingly. The most extensively characterised of these origins is ARS1, and molecular dissection of this sequence revealed the presence of four distinct elements required for replication. The first of these elements is the A element, which contains an 11 bp A/T-rich stretch called the ARS Consensus Sequence (ACS). The ACS is an essential feature of all known ARSs (Celniker et al., 1984, Van Houten and Newlon, 1990), and functions as half of a bipartite binding site for the eukaryotic initiator protein, ORC (Rowley et al., 1995, Rao and Stillman, 1995). ARS1 also contains three B-elements (B1, B2 and B3), the exact nature of which varies between different origins. Whilst individually dispensable for replication initiation, the combined deletion of all three B-elements is sufficient to render ARS1 replication-incompetent (Marahrens and Stillman, 1992). The B1 element comprises the second half of the ORC binding site in the origin (Rowley et al., 1995, Rao and Stillman, 1995), whilst the B2 element is believed to perform some function during pre-RC assembly (Wilmes and Bell, 2002). Finally, the B3 element serves as a binding site for the Abf1 transcription factor (Diffley and Stillman, 1988), and may help to generate a nucleosome-free region at the origin, required for pre-RC assembly and/or origin firing (Venditti et al., 1994, Hu et al., 1999).

1.3 Origin Licensing

The existence of a two-step mechanism for DNA replication initiation was first indicated by mammalian cell fusion experiments performed over 40 years ago (Rao and Johnson, 1970). These experiments showed that the fusion of a G1 cell with an S-phase cell resulted in the rapid initiation of DNA synthesis in the G1 nuclei, whereas G2 cells did not undergo replication when fused with S-phase cells in the same manner. The presence of some factor(s) in the S-phase cell that could promote DNA replication specifically in G1 nuclei was thus inferred. These findings were extended by experiments performed in *Xenopus* egg extracts, which showed that permeabilisation of the nuclear envelope was a prerequisite for G2 nuclei to

replicate when added to a G1 cell extract (Blow and Laskey, 1988). This implied the presence of a specific 'licensing' factor in the G1 extract, which was absolutely required for DNA to gain competence for replication initiation in S-phase. It was additionally supposed that such a factor could potentially gain access to DNA after nuclear envelope breakdown during the mitotic stage of a normal cell cycle.

The molecular details of this two-step mechanism began to emerge during the early part of the 1990s, when *in vivo* footprinting experiments performed in *S. cerevisiae* revealed the presence of two different chromatin states at origins of replication at different cell cycle stages (Diffley et al., 1994). A pre-replicative state, consisting of extensive DNase protection across the replication origin, was detected in G1 phase cells. This protection was significantly reduced in S, G2 and M phases of the cell cycle, when origins adopted a post-replicative state closely resembling the *in vitro* footprint of purified ORC and Abf1.

Subsequent experiments showed that the extensive origin protection observed in G1 was caused by the formation of the pre-replicative complex (pre-RC) (Santocanale and Diffley, 1996, Labib et al., 2001, Aparicio et al., 1997, Cocker et al., 1996). Work over the ensuing decade helped define the proteins involved in pre-RC assembly in budding yeast, culminating in the *in vitro* reconstitution of this process with purified proteins in 2009 (Remus et al., 2009, Evrin et al., 2009). It is now clear that pre-RC assembly, or origin licensing, involves the loading of the Mcm2-7 replicative helicase onto origins of replication, producing an Mcm2-7 double hexamer that is inactive as a DNA helicase. This reaction requires the combined activity of the ORC, Cdc6 and Cdt1 proteins. In this section, I will describe what is known about each of these proteins, and discuss our current understanding of the mechanism and regulation of pre-RC assembly.

1.3.1 Components of the pre-RC

1.3.1.1 ORC

The Origin Recognition Complex (ORC) was first identified in *S. cerevisiae* as an ARS-interacting factor (Bell and Stillman, 1992, Diffley and Cocker, 1992), and

genetic evidence subsequently uncovered an essential function for this protein complex in DNA replication in all eukaryotic organisms (Bell et al., 1993, Foss et al., 1993, Micklem et al., 1993, Gavin et al., 1995, Gossen et al., 1995, Carpenter et al., 1996, Muzi-Falconi and Kelly, 1995, Rowles et al., 1996). In yeast, ORC binds to origins of replication throughout the cell cycle, interacting with both the ACS and B1 elements of the ARS1 origin (Diffley and Cocker, 1992, Rowley et al., 1995, Diffley et al., 1995, Bell and Stillman, 1992, Rao and Stillman, 1995).

Whilst scORC is a sequence-specific DNA binding protein, this activity is not conserved in other eukaryotes (Vashee et al., 2003), consistent with the absence of specific origin sequences in these organisms. In *Drosophila*, for example, ORC displays a 30-fold higher affinity for negatively supercoiled DNA over relaxed DNA *in vitro* (Remus et al., 2004), whereas human ORC can interact with a WD40-repeat containing protein call ORCA (Shen et al., 2010), which in turn binds specifically to methylated histones.

ORC is a hexameric protein complex consisting of Orc1-6 proteins. As with many DNA replication initiator proteins, five out of the six ORC subunits (Orc1-5) exhibit homology to the AAA+ superfamily of ATPases. Cdc6 and Mcm2-7 also belong to this family of ATPases (see sections 1.3.1.2 and 1.3.1.4), which often exist as oligomeric assemblies in which ATPase active sites are formed at the interface between neighbouring subunits. Of the ORC subunits, only Orc1 and Orc5 have been shown to bind ATP (Klemm et al., 1997), and only Orc1 displays ATP hydrolysis activity *in vitro* (Klemm et al., 1997). ATP hydrolysis by Orc1 is dependent on the arginine finger of its neighbouring ORC subunit, Orc4 (Bowers et al., 2004). ATP binding by ORC is required for its binding to origins of replication (Bell and Stillman, 1992), which in turn is essential for pre-RC assembly.

The exact function of the ATPase activity of ORC during origin licensing remains an area of active research. Multiple Mcm2-7 complexes are loaded at each replication origin *in vivo*, and it was proposed that preventing ATP hydrolysis by ORC might block the iterative loading of Mcm2-7 (Bowers et al., 2004). Subsequent studies suggested a role for ORC ATP hydrolysis in driving the release of Cdt1 during Mcm2-7 loading (Fernandez-Cid et al., 2013), although more recent data has

indicated that an ATPase-dead ORC mutant exhibits no detectable defects in pre-RC assembly *in vitro* (Coster et al., 2014).

1.3.1.2 Cdc6

Cell Division Cycle 6 (Cdc6) was initially identified in a budding yeast screen for mutants that exhibited defects in progression through the cell cycle (Hartwell et al., 1973). A function for this protein in DNA replication was subsequently indicated by defects in minichromosome maintenance in a Cdc6 temperature-sensitive mutant (Hogan and Koshland, 1992).

Cdc6 was later shown to be a *bona fide* component of the pre-RC (Cocker et al., 1996, Santocanale and Diffley, 1996), and is now known to be essential for Mcm2-7 loading onto chromatin in all eukaryotes (Donovan et al., 1997, Coleman et al., 1996, Williams et al., 1997). Genetic interactions between an *ORC5-1* allele and *CDC6* suggested that Cdc6 might collaborate with ORC during pre-RC assembly (Liang et al., 1995), and this was subsequently confirmed by the observed physical interaction between the ORC and Cdc6 proteins (Speck et al., 2005).

Cdc6 is a AAA+ ATPase, and binds to and hydrolyses ATP *in vitro*. Whilst ATP binding by Cdc6 is critical for its interaction with ORC (and therefore its association with replication origins) (Speck et al., 2005, Perkins and Diffley, 1998, Mizushima et al., 2000), the exact function of ATP hydrolysis by Cdc6 during pre-RC assembly remains contentious. Genetic data as well as experiments performed using an extract-based system for *in vitro* pre-RC assembly indicated that ATP hydrolysis by Cdc6 is essential for Mcm2-7 loading (Perkins and Diffley, 1998, Weinreich et al., 1999, Randell et al., 2006). However, Cdc6 ATPase mutants can support productive Mcm2-7 loading in fully reconstituted pre-RC assembly reactions (Coster et al., 2014, Kang et al., 2014)

Interestingly, the ATPase activity of Cdc6 has recently been shown to be required for a novel ‘proofreading’ activity (see section 1.3.2.2) (Coster et al., 2014, Kang et al., 2014), which releases Mcm2-7 from DNA under conditions that are not

competent for productive origin licensing (Frigola et al., 2013). It is thus possible that the requirement for Cdc6 ATPase activity during Mcm2-7 loading in extracts is reflective of the accumulation of non-productive helicase loading intermediates in these systems (Randell et al., 2006), perhaps leading to a greater reliance on ‘proofreading’ than is observed in more efficient reconstituted systems.

1.3.1.3 Cdt1

Cdt1 (Cdc10 Dependent Transcript 1) was first discovered in the fission yeast *S. pombe* as a gene that is targetted by the transcription factor Cdc10 and is required for S-phase entry (Hofmann and Beach, 1994). Despite the relatively weak conservation of the primary sequence of Cdt1 throughout evolution, homologues have been identified in *S. cerevisiae, Xenopus*, *Drosophila* and human cells, and Cdt1 is essential for Mcm2-7 loading in each of these organisms (Maiorano et al., 2000, Tanaka and Diffley, 2002, Nishitani et al., 2000, Nishitani et al., 2001, Whittaker et al., 2000).

Cdt1 contains a conserved pair of winged helix domains at its C-terminus, which are required for binding to the Mcm2-7 complex (Zhang et al., 2010, Ferenbach et al., 2005). In *S. cerevisiae*, Cdt1 forms a constitutive complex with Mcm2-7, and this Cdt1/Mcm2-7 heptamer is transported into the nucleus during late mitosis / G1 phase of the cell cycle to facilitate pre-RC formation (Tanaka and Diffley, 2002). Notably, metazoan Cdt1 does not form a complex with Mcm2-7 away from chromatin, and the subtle differences in Cdt1 function observed between different model systems may be reflective of this difference in Cdt1/Mcm2-7 complex formation (Gillespie et al., 2001).

In *S. cerevisiae*, Cdt1 recruitment to origins is dependent on both ORC and Cdc6 (Randell et al., 2006, Remus et al., 2009). Indeed, Cdt1 can interact with Orc6, and this interaction was deemed to be responsible for the initial recruitment of Mcm2-7 during pre-RC assembly (Chen et al., 2007, Chen and Bell, 2011, Takara and Bell, 2011, Semple et al., 2006). Interestingly, an essential, conserved C-terminal domain of Mcm3 has since been shown to direct Mcm2-7 recruitment to ORC/Cdc6,

although Cdt1 was required for the stable association of Mcm2/4/6 with the rest of the Mcm2-7 ring in these experiments (Frigola et al., 2013). Additionally, Cdt1 takes part in an essential interaction with the C-terminus of Mcm6 (Yanagi et al., 2002). This domain has been shown to be inhibitory for Mcm2-7 recruitment and loading, and this inhibition is alleviated by the presence of Cdt1 (Fernandez-Cid et al., 2013), although the molecular explanation for this relationship remains unclear.

1.3.1.4 Mcm2-7

The Minichromosome Maintenance (MCM) proteins were initially discovered in a number of screens for mutants that were defective in cell cycle progression (Moir et al., 1982, Hennessy et al., 1991, Coxon et al., 1992) or the maintenance of extrachromosomal plasmids (Maine et al., 1984, Yan et al., 1991, Takahashi et al., 1994) in yeast. It was subsequently shown that Mcm2, Mcm3, Mcm5 and Mcm7 all localise to the nucleus and bind to chromatin during late mitosis and G1 phase (Hennessy et al., 1991, Yan et al., 1991, Dalton and Whitbread, 1995), thereby exhibiting similar behaviour to the replication licensing factor (RLF) described in *Xenopus* egg extracts (see section 1.3). A *Xenopus* homologue of Mcm3 was subsequently identified and shown to be important for origin licensing (Chong et al., 1995, Kubota et al., 1995, Madine et al., 1995), and evidence was later obtained that all six Mcm2-7 proteins were required for RLF activity in *Xenopus* egg extracts (Thommes et al., 1997).

Several studies showed that Mcm2-7 proteins form a complex that is loaded onto chromatin in an ORC and Cdc6-dependent manner (Romanowski et al., 1996, Rowles et al., 1996, Donovan et al., 1997). These experiments also showed that ORC and Cdc6 are not required for the chromatin association of Mcm2-7 once it has been loaded (Donovan et al., 1997, Rowles et al., 1999), suggestive of the production of a stably bound Mcm2-7 complex during pre-RC assembly. The stable association of Mcm2-7 with DNA was subsequently confirmed by other studies (Seki and Diffley, 2000, Edwards et al., 2002), and can now be explained by the production of an Mcm2-7 double hexamer, which is topologically bound around

dsDNA, during pre-RC assembly *in vitro* (Remus et al., 2009, Evrin et al., 2009, Gambus et al., 2011).

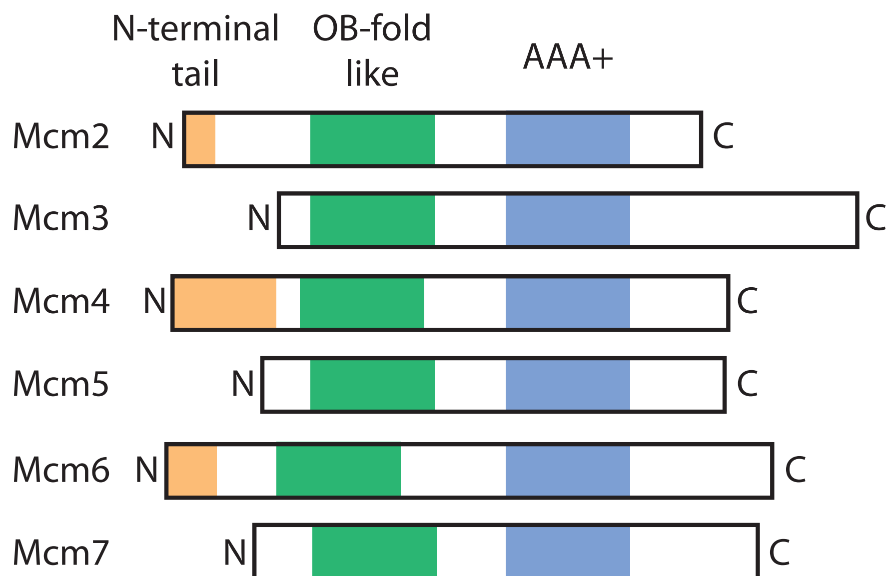
The Mcm2-7 complex is composed of 6 essential subunits, Mcm2-7. All 6 of the proteins share common N-terminal OB-fold (typically involved in binding to nucleic acids) and C-terminal AAA+ ATPase domains, as well as N and C-terminal extensions that are conserved amongst eukaryotes (Fig. 1.2A). Structurally, the Mcm2-7 complex forms a toroidal heterohexamer with a positively charged central channel (Davey et al., 2003, Remus et al., 2009, Bochman et al., 2008, Costa et al., 2011), which has recently been shown to be capable of binding ssDNA (Froelich et al., 2014). As mentioned previously, the Mcm2-7 heterohexamer is constitutively bound to Cdt1 in *S. cerevisiae*.

Individual Mcm subunits interact in a defined order within the Mcm2-7 ring (Fig. 1.2B) (Davey et al., 2003, Remus et al., 2009, Bochman et al., 2008, Costa et al., 2011). As with most AAA+ proteins, the ATPase active sites in Mcm2-7 are located at the interfaces between different Mcm subunits, with one subunit contributing the Walker A (required for ATP binding) and B motifs, and its neighbour supplying the Arginine finger motif (required for ATP hydrolysis). The ATPase function of a number of Mcm active sites is required for S-phase progression and cell viability (Schwacha and Bell, 2001). Recent biochemical analysis of Mcm ATPase mutants suggests that this ATPase activity is critical for Mcm2-7 loading, catalysing Cdt1 release at an early stage during pre-RC assembly (Coster et al., 2014, Kang et al., 2014)

Conclusive evidence for the function of Mcm2-7 as a component of the replicative helicase was not forthcoming in the years following the initial characterisation of the MCM proteins. Initially, genetic analysis indicated that MCM proteins are not required for the completion of chromosome replication once origin activation has occurred (Hennessy et al., 1991, Nasmyth and Nurse, 1981), and an Mcm4/6/7 subcomplex was shown to exhibit very poor DNA helicase activity *in vitro* (Ishimi, 1997). Subsequently, however, DNA unwinding in *Xenopus* extracts was shown to be entirely dependent on the presence of Mcm2-7 (Walter and Newport, 2000), and archaeal relatives of the eukaryotic MCM proteins were shown to function as 3’-5’

DNA helicases *in vitro* (Chong et al., 2000, Shechter et al., 2000, Kelman et al., 1999). Notably, the development of conditional degron alleles for Mcm2-7 allowed their requirement for replication fork progression throughout S-phase to be demonstrated (Labib et al., 2000).

A



B

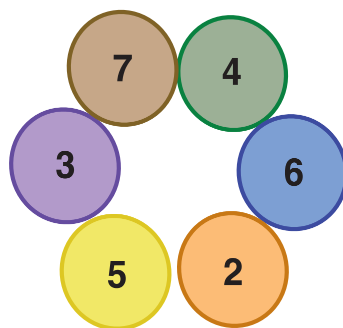


Figure 1.2 Domain architecture and subunit organisation of the *S. cerevisiae* Mcm2-7 complex

A. Cartoon depicting the domain organisation in Mcm2-7. All subunits share a conserved N-terminal OB-fold like domain (green) and C-terminal AAA+ ATPase domain (blue). Mcm2, Mcm4 and Mcm6 have extended N-terminal tails (orange), which are known to be targetted by DDK. Adapted from Randell *et al.*, 2010. **B.** Cartoon depicting the subunit organisation in the Mcm2-7 ring. The ring contains a single copy of each subunit, which are arranged in a distinct order, as depicted.

The functional replicative helicase present at replication forks is now known to be the CMG complex, which contains not only Mcm2-7, but also the Cdc45 and GINS proteins. This complex was initially isolated from *Drosophila* embryo extracts (Moyer et al., 2006), and was shown to exhibit much higher ATP-dependent DNA helicase activity than Mcm2-7 alone (Ilves et al., 2010). Subsequent studies were able to localise Mcm2-7, Cdc45 and GINS to sites of DNA unwinding in *Xenopus* extracts (Pacek et al., 2006), and all three proteins were detected as core components of the Replisome Progression Complex (RPC) (Gambus et al., 2006), thus confirming a role for the CMG complex as the eukaryotic replicative helicase. The remodelling of the Mcm2-7 double hexamer into the CMG complex now forms the basis of any models proposed for Mcm2-7 helicase activation during origin firing, as is described in section 1.4.

1.3.2 Mechanism of pre-RC assembly

Our understanding of the pre-RC assembly reaction has been extended greatly in recent years by experiments utilising a reconstituted system for Mcm2-7 loading *in vitro* (Evrin et al., 2009, Remus et al., 2009). Briefly, purified *S. cerevisiae* Mcm2-7/Cdt1, Cdc6 and ORC are incubated with linear origin-containing DNA fragments bound to beads. In the presence of ATP, ORC, Cdt1 and Cdc6 collaborate to load the Mcm2-7 complex onto DNA, producing a salt-resistant Mcm2-7 double hexamer bound around dsDNA (Fig. 1.3). When the poorly hydrolysable ATP analogue ATPγS is included in these reactions, all pre-RC components are specifically recruited, but Mcm2-7 is not stably loaded onto DNA. The next section(s) will describe our current knowledge of the mechanism of Mcm2-7 loading, drawing primarily on studies using this *in vitro* system. A model of origin licensing is depicted in Fig. 1.4.

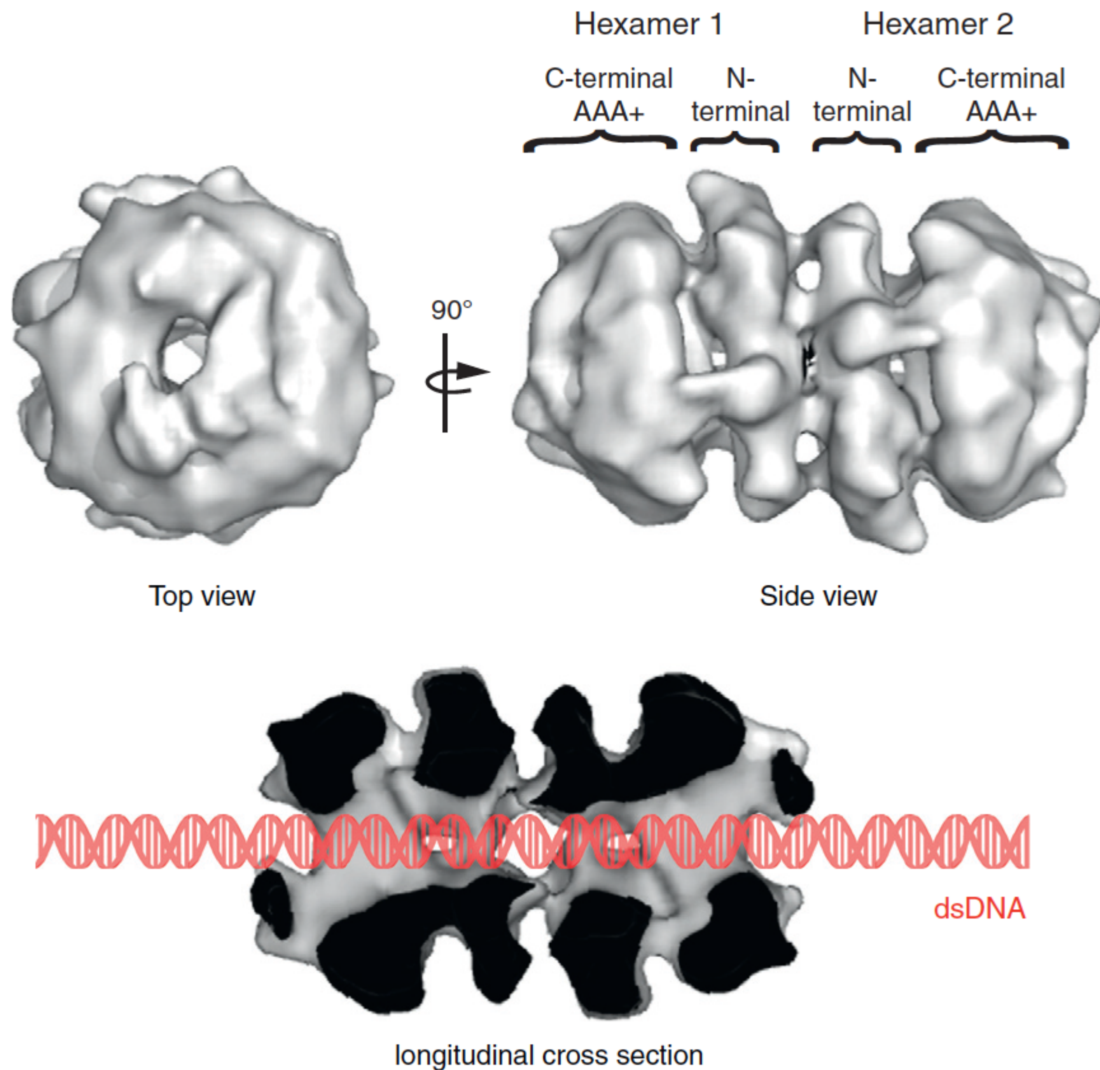


Figure 1.3 Pre-RC assembly produces an Mcm2-7 double hexamer bound around dsDNA

3D reconstruction of the Mcm2-7 double hexamer viewed from different angles. In the bottom image, dsDNA was modelled into the central channel of a longitudinal cross section of the structure. Figure taken from Boos *et al*., 2012. Original images were adapted from Remus *et al*., 2009.

1.3.2.1 Mcm2-7 recruitment

The first step in pre-RC assembly is the ATP-dependent binding of ORC to DNA at origins of replication (Fig. 1.4, step 1). Subsequently, ATP-bound Cdc6 is recruited to the ORC-DNA complex (Fig. 1.4, step 2), inducing a conformational change in ORC that stabilises its interaction with DNA (Speck et al., 2005). The Mcm2-7/Cdt1 complex is recruited to ORC/Cdc6 via a conserved C-terminal domain in Mcm3 (Fig.

1.4, step 3b), and this recruitment is independent of ATP hydrolysis (Frigola et al., 2013). Mcm2-7 binding to both ATP and Cdt1 is required to maintain the integrity of the Mcm2-7 ring at this point (Fig. 1.4, step 3a) (Coster et al., 2014), although this stabilisation is not dependent on the aforementioned Orc6-Cdt1 interaction (Frigola et al., 2013). Recent biochemical and structural analyses of pre-RC assembly reactions performed in ATPγS have indicated the presence of a single copy each of ORC, Cdc6, Cdt1 and Mcm2-7 bound to DNA (Fig. 1.4, step 4), forming a complex that is thought to represent an early intermediate in the helicase loading process (Sun et al., 2013, Fernandez-Cid et al., 2013).

Interestingly, numerous previous reports have indicated that the Mcm2-Mcm5 interface, or Mcm2/5 gate, can be a position of discontinuity in the Mcm2-7 ring (Costa et al., 2011, Bochman et al., 2008, Davey et al., 2003). The recently solved structure of the aforementioned helicase loading intermediate does indeed contain an opening between the N-termini of Mcm2 and Mcm5 (Sun et al., 2013), and the presence of some unresolved electron density within the central channel of the Mcm2-7 ring suggests that the single Mcm2-7 hexamer may be at least partially loaded onto DNA at this juncture. Furthermore, recently published work indicates that artificially locking the Mcm2/5 gate in a closed conformation is sufficient to prevent pre-RC assembly (Samel et al., 2014), indicating that this interface may be a point of DNA entry during Mcm2-7 loading.

1.3.2.2 Quality control during pre-RC assembly

The binding of Mcm3 to ORC/Cdc6 activates the ATPase activity of ORC/Cdc6. When all the pre-RC components are present, and ORC has not been phosphorylated by CDK (see section 1.3.3), productive loading of Mcm2-7 and formation of the Mcm2-7 double hexamer can proceed (see section 1.3.2.3). Notably, however, when these criteria are not met, ATP hydrolysis can lead to the release of non-productive helicase loading intermediates from DNA (Fig. 1.4, step 5a). This recently described proofreading activity is believed to prevent the accumulation of 'dead-end' complexes during pre-RC assembly (Frigola et al.,

2013), as has been shown to be dependent on the ATPase activity of Cdc6 (Coster et al., 2014, Kang et al., 2014).

1.3.2.3 ATP-dependent Mcm2-7 double hexamer formation

Following the initial recruitment of pre-RC components to DNA, one of the earliest steps during active Mcm2-7 loading is the release of Cdt1 (Fernandez-Cid et al., 2013), which is dependent on the ATPase activity of Mcm2-7 itself (Coster et al., 2014, Kang et al., 2014). The resultant ORC/Cdc6/Mcm2-7 (OCM) complex contains a single copy of each of its protein components (Fernandez-Cid et al., 2013) (Fig. 1.4, step 5b). The absence of any DNA-bound Mcm2-7 single hexamers at the end of reactions performed in ATP led to the conclusion that the loading of the two Mcm2-7 hexamers during pre-RC assembly must somehow be concerted (Evrin et al., 2009, Remus et al., 2009). How the reaction proceeds from the OCM intermediate to allow recruitment and loading of the second Mcm2-7 hexamer is the subject of ongoing investigation (Yardimci and Walter, 2014) (Fig. 1.4, step 6). If Mcm2-7 loading is indeed a concerted process, the observation that the C-terminus of Mcm3 is required for the loading of both Mcm2-7 hexamers must indicate the involvement of more than one ORC/Cdc6 complex, or the presence of multiple Mcm2-7 binding sites on a single ORC/Cdc6 (Frigola et al., 2013). Notably, the production of salt-resistant Mcm2-7 double hexamers requires the ATPase activity of Mcm2-7 (Coster et al., 2014, Kang et al., 2014), and involves the release of Cdc6 (Fig. 1.4, step 7).

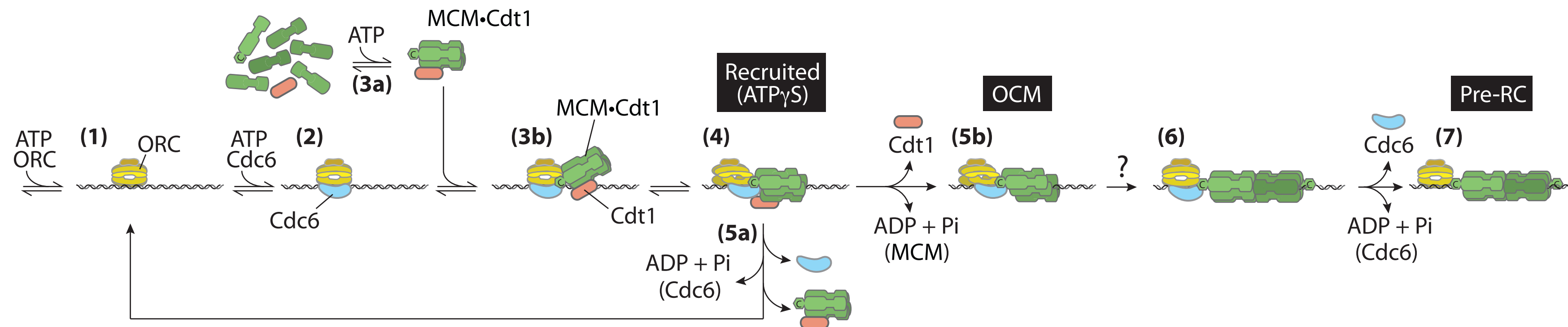


Figure 1.4 A model for the mechanism of origin licensing in *S. cerevisiae*

Details of the model are discussed in section 1.3.2 of the text. The '?' before stage 6 refers to the loading of the second Mcm2-7 hexamer, which is discussed in the text. The protein names in brackets refer to the components responsible for ATP hydrolysis at a given stage. Adapted from Coster *et al.*, 2014.

1.3.3 Regulation of origin licensing

Pre-RC assembly is restricted to late mitosis and G1 phase of the cell cycle. The low CDK and high APC/C activities present during these times are inhibitory to origin firing; the temporal separation of origin licensing and origin firing activities prevents re-replication of DNA, and also ensures that a sufficient number of pre-RCs are assembled onto chromatin before cells are committed to S-phase entry (Diffley, 2004).

The mechanisms by which CDKs inhibit pre-RC assembly are now well understood. In *S. cerevisiae*, CDKs target each pre-RC component individually; the presence of multiple redundant pathways to prevent origin licensing provides a robust block against any potential re-replication. S-phase CDKs (Clb5 and Clb6-CDK) directly phosphorylate ORC on its Orc2 and Orc6 subunits (Nguyen et al., 2001), and CDK can also directly interact with ORC in a Clb5-dependent manner (Weinreich et al., 2001, Wilmes et al., 2004). CDK-phosphorylated Orc6 cannot interact with Cdt1, which prevents pre-RC assembly (Chen and Bell, 2011). Furthermore, recent evidence also suggests that phosphorylation of ORC leads to the proofreading activity described in section 1.3.2.2, thereby preventing Mcm2-7 loading during S-phase (Frigola et al., 2013).

In budding yeast, CDKs target Cdc6 via two distinct mechanisms. Firstly, Cdc6 is allosterically inactivated during M-phase via direct binding to the mitotic cyclin Clb2 (Mimura et al., 2004). Additionally, CDK-mediated phosphorylation of Cdc6 targets it for degradation by the 26S proteasome, via the activity of the SCF^{Cdc4} ubiquitin ligase (Drury et al., 2000). Finally, the Mcm3 subunit of Mcm2-7 is also phosphorylated by CDK, which promotes Mcm2-7/Cdt1 nuclear export (Labib et al., 1999, Liku et al., 2005, Nguyen et al., 2000), thereby physically separating the soluble pool of the replicative helicase from replication origins outside of G1 phase of the cell cycle.

Although the two-step mechanism for replication initiation is conserved amongst eukaryotes, the mechanisms for preventing pre-RC assembly outside of S-phase

are somewhat different between metazoans and yeast. While both Cdc6 and ORC are phosphorylated by S-phase CDKs in metazoans (Petersen et al., 1999, Siddiqui et al., 2013), the primary regulatory target for pre-RC assembly appears to be Cdt1. Cdt1 is bound by an inhibitory protein called geminin, and can only interact with Mcm2-7 and promote origin licensing when geminin is degraded by the APC/C in late mitosis and G1 phase (Wohlschlegel et al., 2000, Tada et al., 2001, McGarry and Kirschner, 1998). Additionally, Cdt1 interacts directly with chromatin bound PCNA in S-phase in both *S. pombe* and metazoans (Arias and Walter, 2006, Guarino et al., 2011). This direct interaction with PCNA triggers ubiquitylation of Cdt1 by the CRL4-Cdt2 ubiquitin ligase, which in turn targets Cdt1 for degradation (Arias and Walter, 2005, Guarino et al., 2011, Senga et al., 2006). There is also a CDK-dependent pathway for Cdt1 degradation in human cells, which requires the SCF^{Skp2} complex (Li et al., 2003, Nishitani et al., 2006). Notably, deregulation of Cdt1 alone is sufficient to trigger re-replication in metazoans (Davidson et al., 2006, Fujita, 2006), consistent with its role as a major regulatory target in these organisms.

1.4 Origin firing

Upon progression into S-phase, the activities of both CDK and DDK increase, and these protein kinases in combination with numerous ‘firing factor’ proteins function to activate the Mcm2-7 replicative helicase, thereby triggering origin unwinding and DNA synthesis. These firing factors are thought to form a transient pre-initiation complex (pre-IC) at the origin of replication (Zou and Stillman, 1998, Tanaka and Araki, 2013), although the exact nature of this complex, and how DDK and CDK catalyse its formation, remain poorly understood. Indeed, compared with the detailed mechanistic information that has emerged for the pre-RC assembly step in recent years, relatively little is known about the mechanism of Mcm2-7 activation at the G1/S transition. Interestingly, studies performed in *Xenopus* egg extracts indicate that ORC and Cdc6 are not required for any steps in replication initiation downstream of origin licensing (Rowles et al., 1999), and models for the mechanism of origin firing thus focus on how firing factors alone remodel the Mcm2-7 double hexamer into the active CMG replicative helicase.

Biochemical studies performed in the last five years have defined the configuration of the Mcm2-7 complex both before and after origin firing has occurred. As was discussed in section 1.3, the product of pre-RC assembly is a double hexamer of Mcm2-7, in which the two hexamers are connected via their N-termini, with the C-terminal AAA+ ATPase domains facing outwards (Fig. 1.3) (Remus et al., 2009). An Mcm2-7 double hexamer has also been detected at licensed origins in *Xenopus* egg extracts (Gambus et al., 2011), consistent with a conserved mechanism of pre-RC assembly in eukaryotes. The evidence that this complex encircles dsDNA is three-fold. Firstly, Mcm2-7 double hexamers have been shown to freely 'slide' on DNA substrates without the need for ATP hydrolysis. If the Mcm2-7 complex was bound around a single DNA strand in the pre-RC, ATP hydrolysis would be required to facilitate the obligatory DNA unwinding that would occur as Mcm2-7 translocated on ssDNA. Secondly, visualisation of DNA-bound Mcm2-7 complexes by electron microscopy (EM) following rotary shadowing did not provide any proof of ssDNA. Finally, similar EM analyses of Mcm2-7 complexes loaded onto circular DNA templates appeared to show DNA in a totally relaxed conformation, inconsistent with the presence of any DNA unwinding at this stage (Remus et al., 2009).

It was initially supposed that the double hexameric form of the replicative helicase might be functional for DNA unwinding upon CDK and DDK activation (Remus et al., 2009, Evrin et al., 2009), acting in an analogous manner to the RuvAB Holliday junction branch migrating enzyme from *E. coli*. However, recent evidence suggests that extensive remodelling of the Mcm2-7 complex must occur upon S-phase entry. In 2011, Fu *et al.* examined the mechanism of Mcm2-7 helicase function in *Xenopus* egg extracts (Fu et al., 2011). In this study, the authors monitored the progression of the replisome on plasmid templates in which site-specific biotin-streptavidin roadblocks were engineered on either the leading or lagging strand template. If the helicase translocated on dsDNA, then replisome progression should be prevented, irrespective of which strand the roadblock is placed upon. Interestingly, however, only roadblocks on the leading strand template impeded helicase progression, consistent with Mcm2-7 complex functioning as a 3' to 5' ssDNA translocase on the leading strand template.

Whilst this study indicates that the Mcm2-7 complex must transition from a dsDNA to ssDNA binding mode during origin firing, additional work, again performed *in vitro* using *Xenopus* egg extracts, has shown that the active replicative helicase likely contains only a single Mcm2-7 hexamer. Yardimci *et al.* performed single molecule analysis on linear bacteriophage λ DNA that was tethered at both ends to a glass slide (Yardimci et al., 2010). Noticeably, this double tethering of the DNA substrate had no discernable effect on replication efficiency compared with singly tethered DNA, consistent with sister replisomes uncoupling from one another once bi-directional replication has been initiated at an origin of replication. In addition to this, the CMG complex purified from *Drosophila* contains only a single copy each of Mcm2-7, Cdc45 and GINS (Costa et al., 2011), and the RPC purified from *S. cerevisiae* contains only a single copy of Mcm4 (Gambus et al., 2006).

Thus, it is now clear that a number of steps must occur to remodel the Mcm2-7 double hexamer into the CMG complex during origin firing (Fig. 1.5). Initial origin melting within the double hexamer should allow for the production of ssDNA, and each Mcm2-7 ring must then open and re-close around the single stranded leading strand template. Additionally, Cdc45 and GINS must stably associate with Mcm2-7, and the Mcm2-7 double hexamer must separate, producing two Cdc45/Mcm2-7/GINS complexes (with 1:1:1 stoichiometry) that can form the basis for complete replisome assembly. Finally, the ATPase 'motor' of the Mcm2-7 complex must be activated to allow DNA unwinding to proceed. As previously mentioned, the Mcm2/5 interface has been shown to be a point of discontinuity in the Mcm2-7 ring, and may be the point at which ssDNA is extruded from the Mcm2-7 central channel during helicase activation. Notably, this interface is 'locked' in a closed conformation by the combined binding of Cdc45 and GINS in the CMG complex (Costa et al., 2011), potentially indicating how the Mcm2-7 ring re-closes around the leading strand template once extrusion of a single DNA strand has occurred.

Whilst we have no understanding of the precise order of these predicted steps, numerous studies have facilitated the identification of the 'firing factor' proteins that are required to catalyse these complex topological changes. In *S. cerevisiae*, the known firing factors are Sld3, Sld7, Dpb11, Sld2, and Mcm10, as well as CDK and

DDK, and Cdc45 and GINS themselves (Boos et al., 2012, Tanaka and Araki, 2013). Notably, in addition to is role in DNA synthesis on the leading strand template, the replicative DNA polymerase, Polε, also functions during Mcm2-7 helicase activation. In the following section, I will discuss what is known about each of these individual proteins, and discuss what little we understand about how they assemble at replication origins during Mcm2-7 helicase activation.

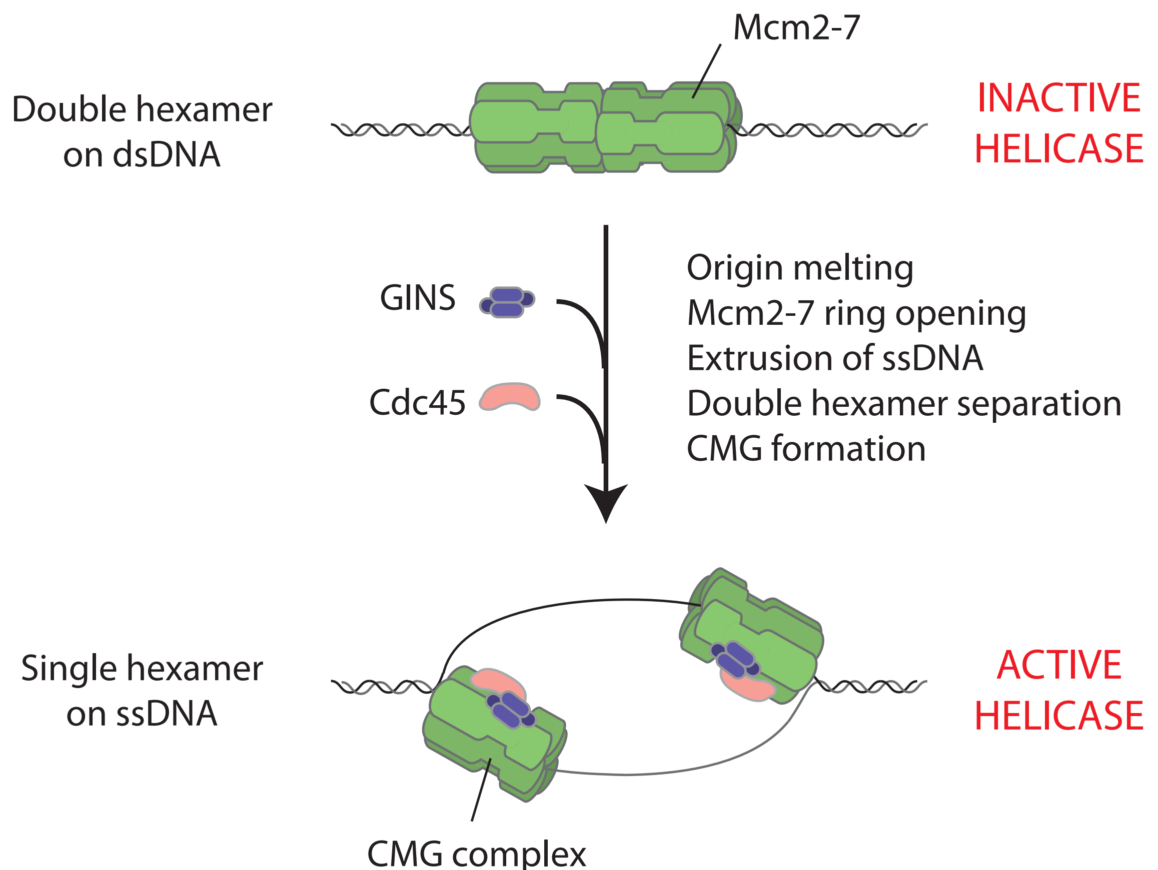


Figure 1.5 Activation of the Mcm2-7 replicative helicase during origin firing

The Mcm2-7 complex is loaded onto DNA as an inactive double hexamer, which is bound around dsDNA. The active CMG complex is bound around the single-stranded leading strand template. In order to achieve this transition, a number of steps must occur, which are listed on the right. For more details, see section 1.4 of the text.

1.4.1 Firing factor proteins

1.4.1.1 Sld3/7

The *SLD3* gene was initially isolated in an *S. cerevisiae* screen for factors exhibiting synthetic lethality with a temperature sensitive allele of *DPB11* (Kamimura et al., 1998). Numerous other firing factors were also isolated in this screen, as will be discussed in subsequent sections. The requirement for Sld3 for the chromatin loading of Cdc45 and the establishment of replication forks at origins of replication was subsequently shown (Kamimura et al., 2001, Kanemaki and Labib, 2006). Interestingly, early studies indicated that, as with a number of other firing factors, Sld3 does not travel with replication forks away from origins (Kanemaki and Labib, 2006), and must therefore be released from DNA at some point during Mcm2-7 helicase activation.

Sld3 is an essential substrate of CDK (Zegerman and Diffley, 2007, Tanaka et al., 2007). Phosphorylation of Sld3 facilitates its binding to Dpb11, as will be discussed in section 1.4.2.1. Furthermore, the recruitment of Sld3 to origins has been shown to be dependent on the activity of DDK in both budding and fission yeast (Tanaka et al., 2011a, Heller et al., 2011, Yabuuchi et al., 2006), suggestive of a role for Sld3 as a central hub for the regulation of origin firing, a notion that will be examined extensively in this thesis. A recent study has shown that Sld3 forms a stable complex with a second firing factor, Sld7, throughout the cell cycle (Tanaka et al., 2011b). Although Sld7 is not essential, cells lacking this protein do show significant defects in S-phase progression, consistent with a function for Sld7 in origin firing.

A homologue of Sld3, called Treslin/TICRR, has been identified in metazoans (Sanchez-Pulido et al., 2010, Kumagai et al., 2010, Sansam et al., 2010). Treslin is a substrate of CDK in these organisms (Kumagai et al., 2010, Boos et al., 2011), and is also required for the chromatin binding of Cdc45 (Kumagai et al., 2011), indicating that Sld3 function is conserved throughout evolution. Recently, Treslin has been shown to constitutively bind to a partner protein, MTBP, throughout the

cell cycle (Boos et al., 2013); whether or not MTBP is a functional analogue of Sld7 is yet to be determined.

1.4.1.2 Cdc45

Like Cdc6, Cdc45 was initially identified in a screen for mutants that are defective in progression through the cell cycle in yeast (Moir et al., 1982). Cdc45 was subsequently shown to interact genetically with members of both the ORC and Mcm2-7 complexes (Hopwood and Dalton, 1996, Dalton and Hopwood, 1997, Zou et al., 1997), and Cdc45 mutants were shown to be defective in minichromosome maintenance (Hardy, 1997, Hopwood and Dalton, 1996), consistent with a role for Cdc45 in DNA replication.

Early studies indicated that Cdc45 moves away from replication origins with the Mcm2-7 complex upon origin firing (Aparicio et al., 1997), and is required for the progression of individual replication forks (Tercero et al., 2000). This observation can now be explained by the finding that Cdc45 is a core component of the CMG replicative helicase, and is thus present at replication forks. Cdc45 has recently been shown to contain some weak homology to an archaeal ssDNA exonuclease called RecJ (Krastanova et al., 2012, Sanchez-Pulido and Ponting, 2011). Although Cdc45 lacks the conserved residues required for exonuclease activity, the protein can bind to ssDNA (Bruck and Kaplan, 2013, Szambowska et al., 2014). It has been suggested that the Cdc45-ssDNA interaction may be important for replisome function upon exposure to hydroxyurea (Bruck and Kaplan, 2013). Cdc45 could also conceivably bind to the single-stranded lagging strand template that is not encircled by Mcm2-7 during unperturbed replisome progression. This model has gained some support by the observation that the binding of Cdc45 and GINS to Mcm2-7 helps to create a secondary ‘channel’, separate to that which is present in the isolated Mcm2-7 complex, within the CMG replicative helicase (Costa et al., 2011).

The recruitment of Cdc45 to origins of replication was initially shown to be CDK-dependent (Zou and Stillman, 1998, Jares and Blow, 2000), but recent work has

demonstrated that two modes of pre-RC binding likely exist for this protein. Cdc45 can be detected at early firing origins in G1 phase of the cell cycle, and this recruitment is dependent on the presence of both Sld3/7 and DDK (Yabuuchi et al., 2006, Kanemaki and Labib, 2006, Heller et al., 2011, Tanaka et al., 2011a). Interestingly, however, the stable incorporation of Cdc45 into the replisome is dependent on the activity of CDK and GINS (Gambus et al., 2006, Heller et al., 2011), consistent with some transition of Cdc45 to a stronger mode of Mcm2-7 interaction occurring during the CDK-dependent step of origin firing.

1.4.1.3 Dpb11

Dpb11 was first discovered as a putative subunit of DNA polymerase epsilon, as the presence of multiple copies of *DPB11* was sufficient to suppress mutations in the Pol2 and Dpb2 subunits of this enzyme (Araki et al., 1995). Whilst subsequent studies have shown that Dpb11 is not in fact a Polε subunit, Dpb11 and Polε do physically interact (Masumoto et al., 2000), and Dpb11 is required for the recruitment of both DNA Polε and Polα/Primase to origins of replication during origin firing (Masumoto et al., 2000). The homology of Dpb11 with the Cut5 protein from *S. pombe* was indicative of a second function for Dpb11 in the cell cycle checkpoint. This function was subsequently confirmed by the observation that Dpb11 is required for activation of the Mec1 kinase, which in turn phosphorylates and activates Rad53, in response to DNA damage in S-phase (Pfander and Diffley, 2011, Navadgi-Patil and Burgers, 2008, Mordes et al., 2008)

Dpb11 contains two pairs of tandem BRCA1 C-terminus (BRCT) repeats. These motifs provide phospho-dependent binding sites for both Sld3 and Sld2, and Dpb11 thus fulfils an essential role in the CDK-dependent step of origin firing (Tanaka et al., 2007, Zegerman and Diffley, 2007), as will be discussed in section 1.4.2.1. In 2010, Dpb11 was identified as a component of the pre-loading complex (pre-LC), an unstable multi-protein assembly that forms off of chromatin and also contains Sld2, Polε and GINS (Muramatsu et al., 2010). Given that it is required for the origin association of both Polε and GINS (Masumoto et al., 2000, Heller et al., 2011, Takayama et al., 2003), and interacts with Sld3, Dpb11 may function as a central

scaffold protein during origin firing, facilitating recruitment of the pre-LC to replication origins, as will be discussed in section 1.4.3.2. Like Sld3/7, Dpb11 is not a replisome component, and is likely released from origins of replication during the origin firing process.

1.4.1.4 Sld2

As was the case for Sld3 and Sld7, Sld2 was isolated in a screen for factors that interact genetically with Dpb11 (Kamimura et al., 1998). Initial studies confirmed that Sld2 physically interacts with Dpb11, and is essential for origin firing. As mentioned in section 1.4.1.3, Sld2 is a component of the pre-loading complex, and is required for the recruitment of GINS, and therefore the formation of the CMG complex, at replication origins (Heller et al., 2011). The origin association of Sld2 itself is dependent on both Dpb11 and Cdc45 (Sangrithi et al., 2005).

Sld2 is an essential CDK substrate required for DNA replication initiation (Masumoto et al., 2002). Phosphorylation of Sld2 facilitates its binding to a tandem pair of BRCT repeats near the C-terminus of Dpb11 (see section 1.4.2.1). Notably, homologues of Sld2 have been discovered in a number of other eukaryotes (Gaggioli et al., 2014). Whilst the function of Sld2 as a CDK substrate is conserved in *C. elegans* (Gaggioli et al., 2014), the interaction of the human Sld2 homologue, RecQL4, with Dpb11 appears to be CDK-independent (Matsuno et al., 2006). Interestingly, RecQL4 also contains a DEAD-box helicase domain, which is essential for DNA replication in *Xenopus* egg extracts (Sangrithi et al., 2005). The precise mechanistic requirements for the helicase function of RecQL4 during DNA replication are yet to be described.

1.4.1.5 GINS

GINS is a hetero-tetrameric protein complex, the function of which in DNA replication was first reported in three separate studies, which were published simultaneously in 2003 (Kanemaki et al., 2003, Takayama et al., 2003, Kubota et al., 2003). One of these studies made use a novel proteomic technology involving

the tagging of essential *S. cerevisiae* proteins of unknown function with a 'heat-inducible' degron, which allows for the rapid degradation of the tagged protein at 37 °C (Kanemaki et al., 2003). Three proteins, Cdc101, Cdc102 and Cdc105, were identified as essential factors required for the establishment and progression of replication forks using this method, and Cdc102 and Cdc105 were shown to physically interact with one another.

A second study, also performed using *S. cerevisiae*, identified the same factors using genetic approaches (Takayama et al., 2003). *PSF1* (*CDC101*) was discovered as a multicopy suppressor of the *SLD5* (*CDC105*) gene, which had been previously identified alongside *SLD2* and *SLD3* in a genetic screen performed in yeast (see sections 1.4.1.1 and 1.4.1.4). *PSF2* (*CDC102*) and *PSF3* were then identified as interactors of *PSF1*. The tetrameric complex of Psf1, Psf2, Psf3 and Sld5 was named GINS (Go, Ich, Nii, San are the Japanese words for five, one, two, three), and homologues were concurrently identified in *Xenopus* (Kubota et al., 2003), where the function of GINS in replication initiation was confirmed. Homologues of all the GINS subunits in human cells have since been isolated.

GINS is a component of the pre-LC (Muramatsu et al., 2010), and its recruitment to origins of replication is dependent on Sld3/7, Cdc45, Dpb11 and Sld2, and requires the activity of both CDK and DDK (Takayama et al., 2003, Heller et al., 2011, Kanemaki and Labib, 2006, Yabuuchi et al., 2006). As previously mentioned, GINS is a core component of the CMG replicative helicase, and is thus required for both the initiation and elongation stages of DNA replication. GINS is required for the stable association of Cdc45 with the Replisome Progression Complex (RPC) (Gambus et al., 2006), although the exact function of GINS during DNA unwinding is still to be determined.

1.4.1.6 Mcm10

Mcm10 was identified in the same yeast genetic screen in which the first members of the Mcm2-7 complex were isolated (Maine et al., 1984, Merchant et al., 1997). Notably, Mcm10 is not related to other Mcm proteins in its primary sequence.

Mcm10 is essential for origin firing in *S. cerevisiae* (Merchant et al., 1997), and its orthologues have also been shown to be required for DNA replication in *S. pombe* (Nasmyth and Nurse, 1981), *Xenopus* (Wohlschlegel et al., 2002) and mouse (Lim et al., 2011). In addition to its function during origin firing, Mcm10 has also been shown to travel with replication forks in yeast (Ricke and Bielinsky, 2004), *Xenopus* (Pacek et al., 2006) and human cells (Karnani and Dutta, 2011). Interestingly, however, Mcm10 only co-purifies as a component of the RPC under unphysiologically high salt concentrations (Gambus et al., 2006, van Deursen et al., 2012), and what function, if any, Mcm10 fulfils during replication elongation is still unclear.

Mcm10 can interact with a number of the subunits in the Mcm2-7 complex (Merchant et al., 1997, Homesley et al., 2000). Consistent with this, mutations in Mcm2 and Mcm7 can suppress the lethal phenotype associated with a number of temperature-sensitive Mcm10 mutants (Homesley et al., 2000, Liachko and Tye, 2005). Additionally, Mcm10 has been shown to interact directly with the loaded form of the replicative helicase during G1 phase of the cell cycle (van Deursen et al., 2012). This direct binding of Mcm10 to Mcm2-7 was shown to be independent of the activities of both DDK and CDK, despite evidence from both *in vitro* and *in vivo* studies suggesting that the origin recruitment of Mcm10 in S-phase requires the function of both of these protein kinases (Kanke et al., 2012, Heller et al., 2011). One explanation for this apparent discrepancy is the existence of two modes of Mcm10 origin recruitment; stable origin association appears to require the action of DDK and CDK, whilst the (presumably weaker) kinase-independent Mcm10-Mcm2-7 interaction observed by van Deursen *et al.* might only be detectable under less stringent experimental conditions.

The molecular function of Mcm10 during origin firing remains quite poorly characterised, although a number of recent studies indicate that Mcm10 is required for initial origin melting and polymerase recruitment (van Deursen et al., 2012, Watase et al., 2012, Kanke et al., 2012, Heller et al., 2011). These studies, performed in both budding and fission yeast, showed that whilst formation of the CMG complex proceeds normally in the absence of Mcm10, the origin recruitment of the ssDNA binding protein RPA, a marker of DNA unwinding, is greatly reduced.

Interestingly, data from one of these studies indicated that replication origin unwinding is largely dependent on the DNA-binding activity of Mcm10 (Kanke et al., 2012), which is mediated by its OB-fold and Zn-finger domains (Robertson et al., 2008, Warren et al., 2008, Eisenberg et al., 2009). Furthermore, it was shown that the stability of Polα was not affected by the absence of Mcm10 (van Deursen et al., 2012, Kanke et al., 2012), contrary to previous reports (Ricke and Bielinsky, 2004). Further biochemical investigation is required to dissect the role played by Mcm10 during Mcm2-7 activation.

1.4.2 Regulation of origin firing

In proliferating budding yeast cells, growth factors actively promote the transcription of G1 cyclins, leading to a gradual accumulation of G1 CDK activity. In turn, G1 CDKs can phosphorylate an inhibitor of the SBF and MBF transcription factors called Whi5 (Costanzo et al., 2004, de Bruin et al., 2004), which relieves the inhibition of these factors, leading to the increased transcription of both S-phase cyclins and specific replication factors (Eser et al., 2011, Spellman et al., 1998, Nasmyth and Dirick, 1991). Both G1 and S-phase CDKs can phosphorylate the CDK inhibitor, Sic1, which targets Sic1 for poly-ubiquitylation and subsequent degradation, thereby creating a positive feedback loop and further increasing the CDK activity in the cell (Feldman et al., 1997, Verma et al., 1997, Koivomagi et al., 2011). Finally, CDK can also phosphorylate the APC/C adaptor protein Cdh1 (Zachariae et al., 1998, Jaspersen et al., 1999). This prevents the binding of Cdh1 to the APC/C, thereby rendering the APC/C inactive, and allowing APC/C substrates to accumulate. One such substrate is the activating subunit of DDK, Dbf4, the levels of which are kept low during G1 phase (Weinreich and Stillman, 1999, Ferreira et al., 2000, Oshiro et al., 1999). Thus, as cells progress into S-phase, the activities of both CDK and DDK increase. Whilst CDK fulfils a function in preventing origin licensing during S-phase (see section 1.3.3), both these kinases also actively stimulate Mcm2-7 helicase activation and origin firing, as will be discussed in the next two sections.

1.4.2.1 CDK

In *S. cerevisiae*, the minimal CDK substrates required for S-phase entry are Sld2 and Sld3 (Tanaka et al., 2007, Zegerman and Diffley, 2007). Sld3 is phosphorylated at two essential sites, Ser 600 and Thr 622, whereas studies suggest that Sld2 is phosphorylated at multiple Ser/Thr, although Thr 84 is the essential site required for replication initiation (Masumoto et al., 2002). It has been hypothesised that phosphorylation of other sites in Sld2 contribute to some unknown conformational change, which in turn exposes Thr 84 for phosphorylation by CDK.

Phosphorylation of Sld2 and Sld3 facilitates their binding to a third firing factor, Dpb11. Dpb11 contains two pairs of tandem BRCT repeats (BRCT I-IV); phosphorylated Sld3 interacts with the N-terminal BRCT pair (BRCT I/II), whereas BRCT III and IV function as a binding site for phosphorylated Sld2. Importantly, the expression of a phospho-mimetic T84D Sld2 mutant, in combination with an 'SD fusion' protein, containing a non-phosphorylatable Sld3 mutant (T600A, T609A, S622A) fused directly to the C-terminal portion of Dpb11 (lacking BRCT I/II), is sufficient to bypass the requirement for CDK for DNA replication (Zegerman and Diffley, 2007). This not only indicates that Sld2 and Sld3 are the minimal CDK substrate required for S-phase entry, but also that the only function of the CDK phosphorylation of Sld3 is to promote binding to Dpb11. Phosphorylation of Sld2 has also been shown to promote the association of Polε with the pre-LC *in vitro* (Muramatsu et al., 2010). However, this effect is not dependent on Thr 84 in Sld2, suggesting that it is not an essential process.

Homologues of Sld2, Sld3 and Dpb11 exist in other eukaryotes. In fission yeast, they are Drc1, Sld3 and Cut5, respectively, and CDK phosphorylation promotes the binding of Sld3 and Drc1 to Cut5, as in *S. cerevisiae* (Nakajima and Masukata, 2002, Noguchi et al., 2002, Saka and Yanagida, 1993). The metazoan homologue of Dpb11, TopBP1, contains nine BRCT domains (BRCT 0-VIII) (Huo et al., 2010, Makiniemi et al., 2001). Treslin (Sld3) binds to BRCT I/II of TopBP1 in a CDK-dependent manner, indicating that the CDK-dependent regulation of the Sld3-Dpb11 interaction is conserved from yeast to humans (Kumagai et al., 2010,

Kumagai et al., 2011, Boos et al., 2011). BRCT IV/V of TopBP1 are homologous to BRCT III/IV of Dpb11, suggesting that this may be the binding site for RecQL4 (Sld2). However, RecQL4 can bind to both TopBP1 and chromatin in the absence of CDK activity (Matsuno et al., 2006, Sangrithi et al., 2005), and a TopBP1 fragment that contains only BRCT I-III is sufficient to support DNA replication in *Xenopus* egg extracts (Kumagai et al., 2010). Notably, the Sld2-Dpb11 interaction is CDK-dependent in *C. elegans* (Gaggioli et al., 2014); further investigation is required to establish if CDK fulfils the same minimal set of functions for S-phase entry across all eukaryotes.

1.4.2.2 DDK

Dbf4-dependent kinase (DDK) is a heterodimer of the Cdc7 catalytic subunit and Dbf4 activating subunit (Jackson et al., 1993); both of these proteins are essential for S-phase entry (Hartwell, 1973, Hereford and Hartwell, 1973, Johnston and Thomas, 1982, Chapman and Johnston, 1989, Bousset and Diffley, 1998). The interaction between Cdc7 and Dbf4 is required to stimulate the kinase activity of Cdc7 (Kitada et al., 1992, Yoon et al., 1993), and DDK activity is thus kept low during G1 phase via the APC/C-mediated degradation of Dbf4 (Weinreich and Stillman, 1999, Ferreira et al., 2000, Oshiro et al., 1999).

The first evidence that the Mcm2-7 helicase itself might be the essential target of DDK required for replication initiation came from the isolation of the *mcm5-bob1* allele, a single amino acid substitution mutant that bypasses the requirement for DDK *in vivo* (Hardy et al., 1997). DDK has since been shown to phosphorylate multiple Ser/Thr in the unstructured N-termini of Mcm2, Mcm4 and Mcm6 (Masai et al., 2006, Montagnoli et al., 2006, Sheu and Stillman, 2006, Randell et al., 2010), and exhibits a strong substrate preference for Mcm2-7 complexes that have been pre-assembled into the pre-RC (Francis et al., 2009). The reason for this substrate preference is not currently understood. However, it is known that DDK can interact with the pre-RC via the N-terminus of Dbf4 (Francis et al., 2009, Dowell et al., 1994) and, furthermore, Cdc7 and Dbf4 have been shown to interact with Mcm4 and Mcm2, respectively (Ramer et al., 2013, Bruck and Kaplan, 2009). Additionally,

the N-terminus of Mcm4 has been reported to contain a DDK binding site (Sheu and Stillman, 2006).

Although Ser/Thr residues with diverse surrounding amino acid sequences have been identified as DDK phosphorylation sites by mass spectrometry (Randell et al., 2010), DDK does display some preference for phosphorylating Ser/Thr that have either a second phosphorylated Ser/Thr or an acidic amino acid (Asp/Glu) in the +1 position (Randell et al., 2010, Masai et al., 2006, Montagnoli et al., 2006, Cho et al., 2006, Charych et al., 2008). This preference is suggestive of the existence of potential ‘priming’ kinases, which might phosphorylate specific Ser/Thr to promote DDK phosphorylation of adjacent residues. Interestingly, a recent study suggested that these priming sites conform to either S/T-P or S/T-Q consensus sequences, and Mec1 was identified as one of the kinases responsible for the phosphorylation of these sites (Randell et al., 2010).

Deleterious phenotypes have been reported for mutants of Mcm2, Mcm4 and Mcm6, which lack DDK phosphorylation sites, suggesting that phosphorylation of all three of these Mcm subunits is important for origin firing (Randell et al., 2010, Bruck and Kaplan, 2009, Sheu and Stillman, 2006). Interestingly, a genetic study published in 2010 showed that the removal of amino acids 74-174 from Mcm4 was sufficient to allow DDK-independent DNA replication *in vivo* (Sheu and Stillman, 2010). It was thus concluded that this section of Mcm4 must be inhibitory for origin firing, and that phosphorylation of Mcm2-7 by DDK somehow relieves this inhibition. This model, in addition to the existence of a DDK bypass allele in *MCM5* (see above), which is not a DDK substrate, has lead many to suppose that DDK phosphorylation must elicit some conformational change in Mcm2-7 that is required for origin firing. However, no gross structural rearrangements were visible when DDK-phosphorylated Mcm2-7 double hexamers were recently visualised by electron microscopy (On et al., 2014).

The activity of DDK has recently been shown to be required for the association of Sld3/7 and Cdc45 with origins of replication, which can occur upstream of CDK function *in vivo* and *in vitro* (Yabuuchi et al., 2006, Heller et al., 2011, Tanaka et al., 2011a). The mechanism by which DDK promotes the recruitment of these factors,

and thus origin firing, is poorly understood, and will be one of the main focuses of this study.

1.4.3 Mechanism of origin firing

In *S. cerevisiae*, we now have a good understanding of the identity of the proteins that are required for origin firing, and informative models now exist that predict the steps, which must proceed to convert the inactive Mcm2-7 double hexamer into the active CMG replicative helicase (Fig. 1.5). However, there is still relatively little known about the mechanism of Mcm2-7 helicase activation, or what functions the individual firing factors fulfil in this process. What little information we do have about this process concerns the order of the recruitment of various proteins to origins of replication, and comes primarily from chromatin binding experiments performed *in vivo*, as well as recent studies conducted using an extract-based system that recapitulates DNA replication initiation *in vitro* (Heller et al., 2011).

1.4.3.1 DDK-dependent recruitment of Cdc45 and Sld3/7

The first step in the pre-IC assembly reaction is thought to be the DDK-dependent recruitment of Sld3/7 and Cdc45 (Fig. 1.6, step 1), which can be detected at early-firing replication origins in G1 phase of the cell cycle (Yabuuchi et al., 2006, Tanaka et al., 2011a). The recruitment of Sld3 and Cdc45 to replication origins has been suggested to be inter-dependent in *S. cerevisiae* (Kamimura et al., 2001, Kanemaki and Labib, 2006), although similar studies performed in *S. pombe* suggest that Sld3 may be able to bind to chromatin in the absence of Cdc45 (Yabuuchi et al., 2006). DDK-dependent Sld3/7 and Cdc45 origin recruitment has been shown to be independent of CDK activity, consistent with studies in *Xenopus* egg extracts showing that DDK functions upstream of CDK in the origin firing process (Jares and Blow, 2000, Walter, 2000).

Mcm10 has also been detected at origins of replication in G1 phase of the cell cycle (van Deursen et al., 2012), although recent studies suggest that it may execute its function at a later step during the origin firing process (see section

1.4.1.6). Notably, initial DNA unwinding is not detectable at this stage of the reaction (Heller et al., 2011), consistent with the lack of observable ssDNA at origins of replication in G1 phase.

1.4.3.2 CDK-dependent pre-LC assembly and GINS recruitment

The increase in CDK activity at the G1/S transition is required for the binding of Sld3 and Sld2 to Dpb11 (Tanaka et al., 2007, Zegerman and Diffley, 2007, Masumoto et al., 2002), and also promotes the assembly of the pre-loading complex (pre-LC), which contains Dpb11, Sld2, GINS and Polε (Muramatsu et al., 2010). Thus, a simplistic model has emerged, whereby the CDK-dependent binding of Dpb11, a pre-LC component, to Sld3 pre-bound to the pre-RC facilitates the CDK-dependent recruitment of GINS to origins of replication (Fig. 1.6, step 2). It is as yet unclear whether any of the pre-LC components other than GINS interact directly with the Mcm2-7 complex, or whether the pre-LC is indeed an active precursor for origin firing.

The presence of Polε in the pre-LC is interesting. Despite its function as the primary replicative polymerase on the leading strand template at replication forks (Pursell et al., 2007), the N-terminal catalytic domain of the Pol2 subunit of Polε is not essential (Kesti et al., 1999, Feng and D'Urso, 2001). Surprisingly, deletion of the C-terminal portion of Pol2 is lethal in *S. cerevisiae* (Dua et al., 1999), whilst recent work suggests that the N-terminus of the Dpb2 subunit of Polε is important for proper assembly of the CMG replicative helicase (Sengupta et al., 2013). Further work is required to ascertain the exact function of Polε during Mcm2-7 activation.

At some juncture following the recruitment of GINS to the Mcm2-7 double hexamer, extensive remodelling must occur to allow two stable CMG complexes to form, each containing a single Mcm2-7 ring as well as Cdc45 and the GINS tetramer. Recently published data suggests that Mcm10 functions at a late step during origin firing to promote ATP-dependent origin unwinding (Fig. 1.6, step 3), which

seemingly occurs downstream of CMG assembly (van Deursen et al., 2012, Watase et al., 2012, Kanke et al., 2012, Heller et al., 2011).

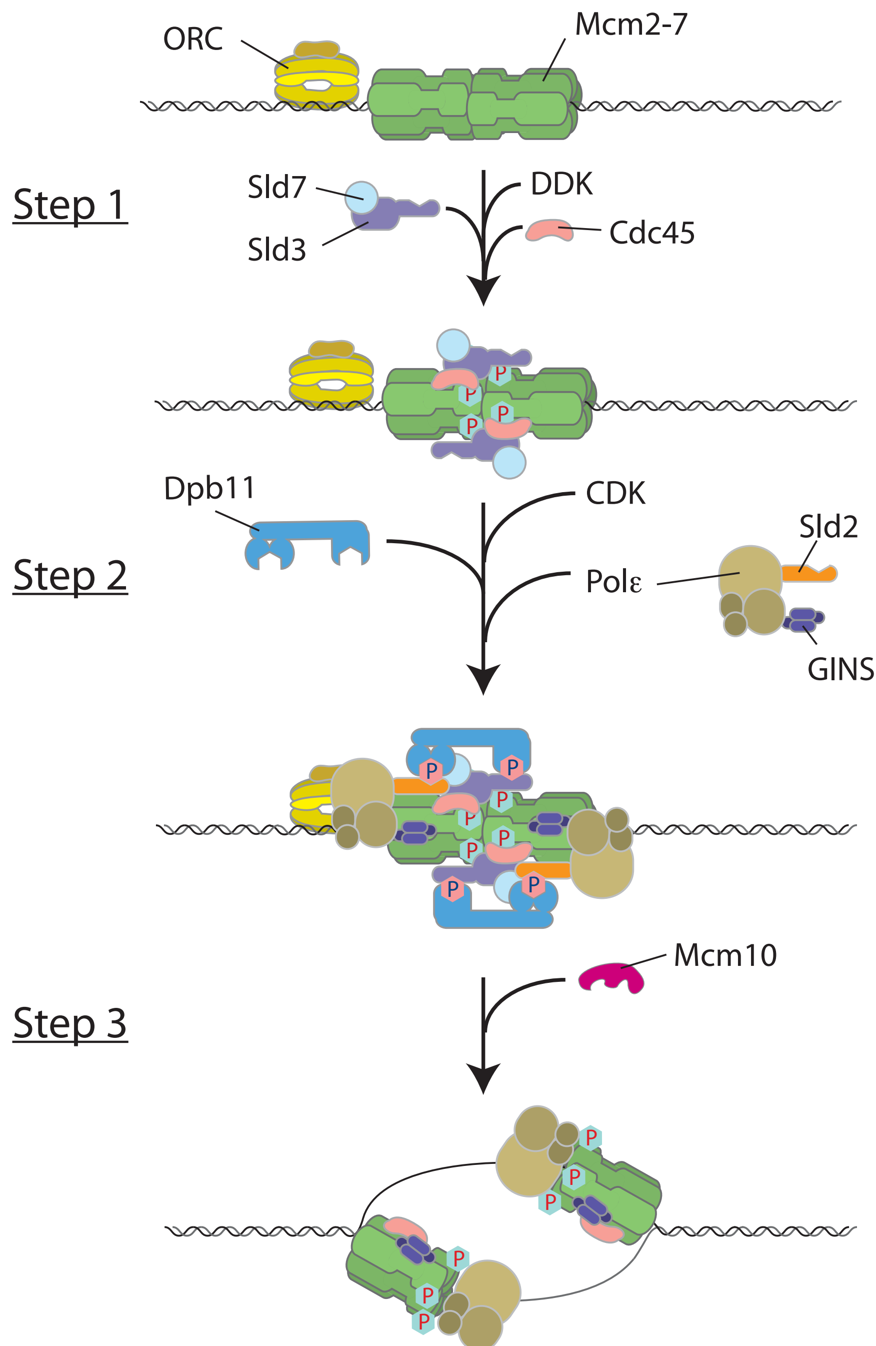


Figure 1.6 Model for multi-step assembly of the pre-initiation complex (pre-IC)

During origin firing, DDK functions upstream of CDK to promote the recruitment of a specific subset of firing factors. Mcm10 functions at a distinct step downstream of Cdc45 and GINS recruitment, and is required for origin unwinding. DDK phosphorylation sites are shown as light blue hexagons, whilst CDK phosphorylation sites are shown as pink hexagons. A full description is given in section 1.4.3 of the text.

Notably, a number of firing factors, such as Sld3, Sld2 and Mcm10, have been shown to bind to ssDNA *in vitro* (Bruck et al., 2011, Bruck and Kaplan, 2011, Robertson et al., 2008, Warren et al., 2008, Eisenberg et al., 2009), but it is not yet clear whether such interactions are important during the origin firing process. Indeed, it is not known how any of the firing factors discussed herein contribute towards origin melting, Mcm2-7 ring opening or separation of the Mcm2-7 double hexamer, let alone how other factors such as Ctf4 and Mrc1 associate to form the Replisome Progression Complex (RPC) (Gambus et al., 2006) once the active CMG replicative helicase is bound around the leading strand template.

1.5 Thesis summary

The work presented in this thesis attempts to dissect the mechanism of Mcm2-7 helicase activation, using the budding yeast *S. cerevisiae* as a model system. In recent years, our understanding of the mechanism of Mcm2-7 loading has been greatly advanced by studies performed using a reconstituted system for pre-RC assembly *in vitro*. The complete reconstitution of the entire process of DNA replication initiation using purified proteins should allow for detailed mechanistic insight about Mcm2-7 helicase activation and replisome assembly to be attained. Towards this aim, I adopted a primarily biochemical approach during this study, and initially developed protocols for the purification of many of the factors known to be required for origin firing. Interestingly, I obtained evidence that the Sld3/7 complex contains multiple copies of Sld3, which has implications for the mechanism of pre-IC assembly.

Subsequently, I predominantly focussed on understanding the functions of DDK and Sld3/7 during Mcm2-7 activation. I reconstituted the DDK-dependent recruitment of Sld3/7 to the pre-RC *in vitro*, which is believed to be the first step in pre-IC assembly. I have shown that Sld3/7 recruitment to the pre-RC is dependent on a novel phosphorylation-dependent interaction between a central domain of Sld3 and the loaded Mcm2-7 double hexamer. I showed that Sld3 itself contains a novel phosphopeptide binding activity, and can interact with a number of Mcm subunits in a phospho-dependent manner. I was subsequently able to isolate

mutants in both Sld3 and Mcm2-7 that were defective in Sld3/7 recruitment, which allowed the functional significance of this interaction for origin firing to be addressed.

In addition to providing a basis upon which future attempts to reconstitute Mcm2-7 helicase activation will be built, this study also affords significant insight into the early steps of the origin firing process. To our knowledge, the Sld3-Mcm2-7 interaction described herein is the first protein-protein interaction described that is directly regulated by DDK, and helps to explain how DDK promotes origin firing. Furthermore, the involvement of the essential CDK substrate Sld3 in a DDK-dependent pathway singles out Sld3 as a point of convergence for the activities of CDK and DDK during Mcm2-7 helicase activation.

Chapter 2. Materials & Methods

2.1 Enzymes and Reagents

2.1.1 Enzymes and proteins

Table 2-1 Enzymes used in this study

Enzyme	Source
Benzonase, ribonuclease A	Sigma-Aldrich
Restriction enzymes, λ protein phosphatase, Phusion DNA polymerase	New England Biolabs (NEB)
KOD Hot Start DNA polymerase	Novagen
T4 DNA ligase	Roche

2.1.2 Antibodies

Table 2-2 Primary monoclonal antibodies used in this study

Antibody	Source	Dilution
Anti-FLAG	Sigma-Aldrich	1:1000
Anti-Orc6	CRUK	1:1000
Anti-Myc	Santa Cruz Biotechnology	1:1000
Anti-Cdc6	CRUK	1:1000

Table 2-3 Primary polyclonal antibodies used in this study

Antibody	Source	Dilution
Anti-Mcm2 (yN-19)	Santa Cruz Biotechnology	1:1000
Ant-Mcm4 (yC-19)	Santa Cruz Biotechnology	1:1000
Anti-Mcm6 (20147)	Generated in the laboratory	1:500
Anti-Mcm7 (yN-19)	Santa Cruz Biotechnology	1:1000
Anti-Sld2	Generated in the laboratory	1:1000
Anti-Sld3	Generated in the laboratory	1:1000
Anti-Sld7	Generated in the laboratory	1:1000

Anti-Cdc45	Generated in the laboratory	1:1000
Anti-Dpb11	Generated in the laboratory	1:1000
Anti-Psf1	A gift from Dr. Karim Labib	1:3000

Table 2-4 Secondary antibodies used in this study

Antibody	Source	Dilution
Anti-goat-HRP	Stratech Scientific	1:5000
Anti-sheep-HRP	Santa Cruz Biotechnology	1:5000
Anti-mouse-HRP	Dako	1:5000
Anti-rabbit-HRP	Jackson ImmunoResearch	1:5000

2.2 Media

Ultrapure water was used in all media and solutions unless otherwise stated. All media was made by Cancer Research UK London Research Institute Cell Services

2.2.1 Media for E. coli cells

Luria Broth (LB) medium (0.5% bacto-tryptone, 0.25% bacto-yeast extract, 170 mM NaCl, pH 7.0) was used for growth in suspension. For growth on solid media, LB was supplemented with 2% agar. For selective growth, media was supplemented with ampicillin (50 μg/ml), chloramphenicol (35 μg/ml) or kanamycin (50 μg/ml). During transformation of E. coli, cells were grown in SOC medium (0.5% bacto-yeast extract, 2% bacto-tryptone, 10 mM NaCl, 2.5 mM KCl, 10 mM $MgCl_2$, 10 mM $MgSO_4$, 20 mM glucose).

2.2.2 Media for yeast cells

For growth in suspension, cells were grown in Yeast Peptone (YP) medium (1% yeast extract, 2% peptone) supplemented with 2% glucose (YPD), galactose or raffinose. For growth on solid media, 2% agar was added to YPD.

For selective growth on solid media, minimal drop-in media (2% agar, 2% glucose, 1x yeast nitrogen base) was supplemented with amino acids as required (histidine 10 mg/ml, tryptophan 2 mg/ml, leucine 10 mg/ml, adenine 5 mg/ml, uracil 2 mg/ml). For selective growth using the NatNT2 marker, cells were grown on YPD agar supplemented with 100 μg/ml Nourseothricin (LEXSY NTC, Jena Bioscience).

2.3 Buffers and Solutions

2.3.1 Buffers for general manipulation of DNA

50x TAE: 2 M Tris base, 2 M glacial acetic acid, 50 mM EDTA
6x DNA Loading Buffer: 0.15% orange G, 60% glycerol, 1 x TAE

2.3.2 Buffers for general manipulation of proteins

2x Laemmli sample buffer: 100 mM Tris-HCl pH 6.9, 4% (w/v) SDS, 20% (v/v) glycerol, 200 mM 2-mercaptoethanol. 0.1% (w/v) bromophenol blue
5x Laemmli sample buffer: 300 mM Tris-HCl pH 6.8, 20% (w/v) SDS, 50% (w/v) glycerol, 0.1% (w/v) bromophenol blue, 2.8 M 2-mercaptoethanol
TBS: 20 mM Tris-HCl pH 7.5, 150 mM NaCl
TBST: 20 mM Tris-HCl pH 7.5, 150 mM NaCl, 1 mM Tween 20
PBS: 137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 1.8 mM KH_2PO_4
Western transfer buffer: 48 mM Trizma base, 39 mM glycine, 0.0375% (w/v) SDS, 20% methanol
Blocking buffer: 5% (w/v) Marvel milk powder dissolved in TBST
Probing buffer: 2.5% (w/v) Marvel milk powder dissolved in TBST
Super blotto buffer: 0.5% (w/v) BSA, 2.5% (w/v) Marvel milk powder, 0.5% (v/v) NP-40, 1 mM Tween 20, 1 x TBS

2.3.3 Buffers for protein purification

Wash buffer I: 25 mM Hepes-KOH pH 7.6, 1 M sorbitol
Buffer A: 25 mM Hepes-KOH pH 7.6, 0.05% (v/v) NP-40, 10% (v/v) glycerol

Buffer B: 45 mM Hepes-KOH pH 7.6, 5 mM $Mg(OAc)_2$, 0.02% (v/v) NP-40, 10% (v/v) glycerol

Buffer C: 50 mM K_2HPO_4/KH_2PO_4 pH 7.6, 5 mM $MgCl_2$, 2 mM ATP, 1 mM DTT

Buffer D: 50 mM K_2HPO_4/KH_2PO_4 pH 7.6, 400 mM KOAc, 5 mM $MgCl_2$, 0.1% Triton, 15% (v/v) glycerol, 1 mM DTT

SE buffer I: 50 mM Hepes-KOH pH 7.6, 300 mM K-Glutamate, 2 mM EDTA, 800 mM sorbitol

SE buffer II: 100 mM Hepes-KOH pH 7.6, 10 mM $Mg(OAc)_2$, 800 mM sorbitol

SE buffer III: 50 mM Hepes-KOH pH 7.6, 300 mM K-Glutamate, 5 mM $Mg(OAc)_2$, 1 mM EDTA, 1 mM EGTA, 10% (v/v) glycerol

Homemade Protease Inhibitor Cocktail: 1 mM EDTA, 5 mM benzamidine-HCl, 1.5 μM pepstatin A, 0.5 mM AEBSF, 0.3 μM aprotinin, 3 mM PMSF, 2 μM leupeptin

2.3.4 Buffers for *in vitro* biochemical assays

Buffer I: 10 mM Tris-HCl pH 7.5, 1 mM EDTA, 2 M NaCl

Buffer II: 10 mM Hepes-KOH pH 7.6, 1 mM EDTA, 1 M KOAc

Buffer III: 10 mM Hepes-KOH pH 7.6, 1 mM EDTA

5x Binding Buffer (5x BB): 125 mM Hepes-KOH pH 7.6, 50 mM $Mg(OAc)_2$, 0.1% (v/v) NP-40, 25% (v/v) glycerol, 0.325 M KOAc

20x Replication Buffer (20X RPB): 800 mM Hepes-KOH pH 7.6, 160 mM $MgCl_2$

Pre-RC Wash Buffer: 45 mM Hepes-KOH pH 7.6, 5 mM $Mg(OAc)_2$, 1 mM EDTA, 1 mM EGTA, 0.02% (v/v) NP-40, 10% (v/v) glycerol

Low Salt Wash Buffer: Pre-RC Wash Buffer + 300 mM KOAc / K-Glutamate (as indicated)

High Salt Wash Buffer: Pre-RC Wash Buffer + 500 mM NaCl

Peptide Array Buffer: 25 mM Hepes-KOH pH 7.6, 10 mM $Mg(OAc)_2$, 0.02% (v/v) NP-40, 5% (v/v) glycerol, 0.5 M KOAc, 50 mM KCl, 1 mM DTT

2.4 Plasmids

Table 2-5 Plasmids used in this study

Name	Cloning vector	Insert	Use	Reference
p305BP	pBR322	*ARS305*	Template for PCR to generate biotinylated DNA fragment for *in vitro* pre-RC assembly assay	(Huang and Kowalski, 1996)
pBluescript/ARS1 WTA	N/A	N/A	Template for *in vitro* DNA replication assay	(On et al., 2014)
pAM3	pGEX-6p-1	*CDC6*	Expression of Cdc6 in *E. coli* for purification	(Frigola et al., 2013)
pJF2	pRS303	*GAL4-GAL1-10-CDT1*	Galactose-inducible expression in yeast	(Frigola et al., 2013)
pJF4	pRS305	*MCM7-GAL1-10-MCM6*	Galactose-inducible expression in yeast	(Frigola et al., 2013)
pBP83	pYM21	*3XFLAG-Nat-NT2*	Template for PCR to amplify 3xFLAG-Nat-NT2 for epitope tagging in yeast	Generated within the laboratory
pBS1539/TAP^{TCP}	N/A	N/A	Template for PCR to amplify TAP^{TCP}-URA3 for epitope tagging in yeast	(Remus et al., 2009).
pFA6a-Nat-NT2	N/A	N/A	Generation of *SLD3-Nat-NT2*	(Falbo and Shen, 2012)

			PCR cassettes for trannformation	

Table 2-6 Plasmids generated in this study.

The subscript SUP indicates that the gene's codon usage is optimised for expression.

Name	Cloning vector	Insert	Generation of insert	5' cloning site	3' cloning site
pTD2	pJF2	*DPB11*	PCR	SgrA1	NotI
pTD4	pJF2	*$CDC45_{SUP}$*	GeneArt Gene Synthesis	SgrA1	NotI
pTD5	pJF2	*$SLD2_{SUP}$*	GeneArt Gene Synthesis	SgrA1	NotI
pTD6a	pJF2	*$SLD3_{SUP}$*	GeneArt Gene Synthesis	SgrA1	NotI
pTD6b	pJF4	*$SLD7_{SUP}$*	GeneArt Gene Synthesis	SgrA1	NotI
pTD12	pJF4	*MCM6Δ2-84*	PCR	SgrA1	NotI
pTD24	pJF4	*MCM6-TEV2*	GeneArt Gene Synthesis (*MCM6-209-588)*	PshA1	SnaB1
pTDP1	pJF4	*MCM6-T150A*	GeneArt Gene Synthesis (*MCM6-1-209)*	SgrA1	PshA1
pTD29	pTDP1	*MCM6-T150A, 11S/T-A*	GeneArt Gene Synthesis (*MCM6-209-588)*	PshA1	SnaB1
pTD30	pTD12	*MCM6Δ2-*	GeneArt Gene	PshA1	SnaB1

		84, 11S/T-A	Synthesis (*MCM6-209-588)*		
pTD33	pJF4	*MCM6Δ2-84, Δ199-260*	GeneArt Gene Synthesis (*MCM6-1-588)*	SgrA1	SnaB1
pTD39	pTD12	*MCM6Δ2-84, 3S/T-A*	GeneArt Gene Synthesis (*MCM6-209-588)*	PshA1	SnaB1
pFA6a-Nat-NT2/ *SLD3 2E3*	pFA6a-Nat-NT2	*SLD3-K530E, R531E*	PCR (pGEX-6p-1/*SLD3 2E3* template)	NdeI	PvuII
pFA6a-Nat-NT2/ *SLD3 6E*	pFA6a-Nat-NT2	*SLD3-K511E, R512E, K518E, R520E, K530E, R531E*	PCR (pGEX-6p-1/*SLD3 6E* template)	NdeI	PvuII
pGEX-6p-1/*SLD3*	pGEX-6p-1	*SLD3*	PCR	BamHI	XhoI
pGEX-6p-1/*SLD7*	pGEX-6p-1	*SLD7*	PCR	BamHI	XhoI
pGEX-6p-1/*SLD3 N0*	pGEX-6p-1	*FLAG-SLD3*	PCR	BamHI	XhoI
pGEX-6p-1/*SLD3 N4*	pGEX-6p-1	*FLAG-SLD3-1-435*	PCR	BamHI	XhoI
pGEX-6p-	pGEX-	*FLAG-*	PCR	BamHI	XhoI

1/*SLD3 N5*	6p-1	*SLD3-1-585*			
pGEX-6p-1/*SLD3 C0*	pGEX-6p-1	*SLD3-FLAG*	PCR	BamHI	XhoI
pGEX-6p-1/*SLD3 C1*	pGEX-6p-1	*SLD3-586-668-FLAG*	PCR	BamHI	XhoI
pGEX-6p-1/*SLD3 C2*	pGEX-6p-1	*SLD3-436-668-FLAG*	PCR	BamHI	XhoI
pGEX-6p-1/*SLD3 C3*	pGEX-6p-1	*SLD3-326-668-FLAG*	PCR	BamHI	XhoI
pGEX-6p-1/*SLD3 C4*	pGEX-6p-1	*SLD3-251-668-FLAG*	PCR	BamHI	XhoI
pGEX-6p-1/*SLD3 C5*	pGEX-6p-1	*SLD3-133-668-FLAG*	PCR	BamHI	XhoI
pGEX-6p-1/*SLD3 M3*	pGEX-6p-1	*SLD3-251-471-FLAG*	PCR	BamHI	XhoI
pGEX-6p-1/*SLD3 M4*	pGEX-6p-1	*SLD3-251-486-FLAG*	PCR	BamHI	XhoI
pGEX-6p-1/*SLD3 M5*	pGEX-6p-1	*SLD3-251-585-FLAG*	PCR	BamHI	XhoI
pGEX-6p-1/*SLD3 2E1*	pGEX-6p-1/*SLD3 N0*	*SLD3-K511E, R512E*	GeneArt Gene Synthesis (*SLD3-492-668)*	XbaI	XhoI

pGEX-6p-1/*SLD3 2E2*	pGEX-6p-1/*SLD3 N0*	*SLD3-K518E, R520E*	GeneArt Gene Synthesis (*SLD3-492-668)*	XbaI	XhoI
pGEX-6p-1/*SLD3 2E3*	pGEX-6p-1/*SLD3 N0*	*SLD3-K530E, R531E*	GeneArt Gene Synthesis (*SLD3-492-668)*	XbaI	XhoI
pGEX-6p-1/*SLD3 4E1*	pGEX-6p-1/*SLD3 N0*	*SLD3-K511E, R512E, K518E, R520E*	GeneArt Gene Synthesis (*SLD3-492-668)*	XbaI	XhoI
pGEX-6p-1/*SLD3 4E2*	pGEX-6p-1/*SLD3 N0*	*SLD3-K511E, R512E, K530E, R531E*	GeneArt Gene Synthesis (*SLD3-492-668)*	XbaI	XhoI
pGEX-6p-1/*SLD3 4E3*	pGEX-6p-1/*SLD3 N0*	*SLD3-K518E, R520E, K530E, R531E*	GeneArt Gene Synthesis (*SLD3-492-668)*	XbaI	XhoI
pGEX-6p-1/*SLD3 6E*	pGEX-6p-1/*SLD3 N0*	*SLD3-K511E, R512E, K518E, R520E, K530E, R531E*	GeneArt Gene Synthesis (*SLD3-492-668)*	XbaI	XhoI
pGEX-6p-1/*SLD3 K267E*	pGEX-6p-1/*SLD3*	*SLD3-K267E*	GeneArt Gene Synthesis (*SLD3-215-*	BsrG1	XbaI

	C0		*492)*		
pGEX-6p-1/*SLD3 R269E*	pGEX-6p-1/*SLD3 C0*	*SLD3-R269E*	GeneArt Gene Synthesis (*SLD3-215-492)*	BsrG1	XbaI
pGEX-6p-1/*SLD3 K267E, R269E*	pGEX-6p-1/*SLD3 C0*	*SLD3-K267E, R269E*	GeneArt Gene Synthesis (*SLD3-215-492)*	BsrG1	XbaI
pGEX-6p-1/*SLD3 K272E*	pGEX-6p-1/*SLD3 C0*	*SLD3-K272E*	GeneArt Gene Synthesis (*SLD3-215-492)*	BsrG1	XbaI
pGEX-6p-1/*SLD3 K386E*	pGEX-6p-1/*SLD3 C0*	*SLD3-K386E*	GeneArt Gene Synthesis (*SLD3-215-492)*	BsrG1	XbaI
pGEX-6p-1/*SLD3 K404E*	pGEX-6p-1/*SLD3 C0*	*SLD3-K404E*	GeneArt Gene Synthesis (*SLD3-215-492)*	BsrG1	XbaI
pGEX-6p-1/*SLD3 K405E*	pGEX-6p-1/*SLD3 C0*	*SLD3-K405E*	GeneArt Gene Synthesis (*SLD3-215-492)*	BsrG1	XbaI
pGEX-6p-1/*SLD3 K404E, K405E*	pGEX-6p-1/*SLD3 C0*	*SLD3-K404E, K405E*	GeneArt Gene Synthesis (*SLD3-215-492)*	BsrG1	XbaI
pGEX-6p-1/*SLD3 K416E*	pGEX-6p-1/*SLD3 C0*	*SLD3-K416E*	GeneArt Gene Synthesis (*SLD3-215-492)*	BsrG1	XbaI

pGEX-6p-1/*SLD3 K418E*	pGEX-6p-1/*SLD3 C0*	*SLD3-K418E*	GeneArt Gene Synthesis (*SLD3-215-492)*	BsrG1	XbaI
pGEX-6p-1/*SLD3 K416E, K418E*	pGEX-6p-1/*SLD3 C0*	*SLD3-K416E, K418E*	GeneArt Gene Synthesis (*SLD3-215-492)*	BsrG1	XbaI
pGEX-6p-1/*SLD3 K301E*	pGEX-6p-1/*SLD3 C0*	*SLD3-K301E*	GeneArt Gene Synthesis (*SLD3-215-492)*	BsrG1	XbaI
pGEX-6p-1/*SLD3 3E1*	pGEX-6p-1/*SLD3 C0*	*SLD3-K303E, K304E, R305E*	GeneArt Gene Synthesis (*SLD3-215-492)*	BsrG1	XbaI
pGEX-6p-1/*SLD3 3E2*	pGEX-6p-1/*SLD3 C0*	*SLD3-K301E, R305E, R317E*	GeneArt Gene Synthesis (*SLD3-215-492)*	BsrG1	XbaI
pGEX-6p-1/*SLD3 4E4*	pGEX-6p-1/*SLD3 C0*	*SLD3-K296E, K297E, K299E, K301E*	GeneArt Gene Synthesis (*SLD3-215-492)*	BsrG1	XbaI
pGEX-6p-1/*SLD3 4E5*	pGEX-6p-1/*SLD3 C0*	*SLD3-R317E, R318E, K320E, K321E*	GeneArt Gene Synthesis (*SLD3-215-492)*	BsrG1	XbaI
pGEX-6p-1/*SLD3 8E*	pGEX-6p-1/*SLD3*	*SLD3-K296E, K299E,*	GeneArt Gene Synthesis (*SLD3-215-*	BsrG1	XbaI

	C0	*K301E, K303E, R305E, R317E, R318E, K320E*	*492)*		

2.5 DNA Oligonucleotides

Table 2-7 Oligonucleotides used in this study

Name	Sequence	Use	Purpose
Dpb11 fwd	5’-AAGCTCACCGGTGATGAAGCCCTTTCAAGGAATA-3’	5’ PCR	Construction of pTD2
Dpb11 rev	5’-CTGAGGCGGCCGCTCAAGAATCTAATTCCTTTGT-3’	3’ PCR	Construction of pTD2
Mcm6ΔN fwd	5’-GTCGGCACCGGTGATGGAAATGAACCTGCCAGAAGC-3’	5’ PCR	Construction of pTD12
Mcm6ΔN rev	5’-AGCTAGCGGCCGCTTAGCTGGAATCCTGTGGTTC-3’	3’ PCR	Construction of pTD12
Sld3 fwd	5’-CTGGACACCGGTGATGGAAACATGGGAAGTCATA-3’	5’ PCR	Construction of pGEX-6p-1/*SLD3*
Sld3 rev	5’-AGCACGCGGCCGCCTATGTGGATTCTGGAGCAAA-3’	3’ PCR	Construction of pGEX-6p-1/*SLD3*
Sld7 fwd	5’-GTCGGCACCGGTGATGTCACGGAAATTATGCACA-3’	5’ PCR	Construction of pGEX-6p-1/*SLD7*
Sld7 rev	5’-AGCTAGCGGCCGCTCATGATTTGGTAAAGAGCTT-3	3’ PCR	Construction of pGEX-6p-1/*SLD7*
N fwd	5’-AGTCGGGATCCGACTACAAAGACGATGACGACAAGGAAACATGGGAAGTCATA	5’ PCR	Construction of pGEX-6p-

	GCA-3’		1/*SLD3 N0, N4 and N5*
NR0	5’-CTGAGCTCGAGCTATGTGGATTCTG GAGCAAA-3’	3’ PCR	Construction of pGEX-6p-1/*SLD3 N0*
NR4	5’-CTGAGCTCGAGCTATGTGCTATCGT TGACCTTTTT-3’	3’ PCR	Construction of pGEX-6p-1/*SLD3 N4*
NR5	5’-CTGAGCTCGAGCTATAGGCGTATCG TTTCATTGAT-3’	3’ PCR	Construction of pGEX-6p-1/*SLD3 N5*
C rev	5’-CTGAGCTCGAGCTACTTGTCGTCAT CGTCTTTGTAGTCTGTGGATTCTGGAGC AAATAA-3’	3’ PCR	Construction of pGEX-6p-1/*SLD3 C0 – C5*
CF0	5’-AGTCGGGATCCGAAACATGGGAAGT CATAGCA-3’	5’ PCR	Construction of pGEX-6p-1/*SLD3 C0*
CF1	5’-AGTCGGGATCCCATGAACGTGTAGA CTCTGAG-3’	5’ PCR	Construction of pGEX-6p-1/*SLD3 C1*
CF2	5’-AGTCGGGATCCAATGTATCGTCACC TAATACT-3’	5’ PCR	Construction of pGEX-6p-1/*SLD3 C2*
CF3	5’-AGTCGGGATCCAAAGACAAAGGAAT CGAGAGA-3’	5’ PCR	Construction of pGEX-6p-1/*SLD3 C3*
CF4	5’-AGTCGGGATCCGATGAAAACAAAAA CAGCTCA-3’	5’ PCR	Construction of pGEX-6p-1/*SLD3 C4, M3, M4 and M5*
CF5	5’-AGTCGGGATCCAAAGACCTTAAACT	5’ PCR	Construction

	GGATATG-3’		of pGEX-6p-1/*SLD3 C5*
NR5+ FLAG	5’-CTGAGCTCGAGCTACTTGTCGTCAT CGTCTTTGTAGTCTAGGCGTATCGTTTC ATTGAT-3’	3’ PCR	Construction of pGEX-6p-1/*SLD3 M5*
NR6+ FLAG	5’-CTGAGCTCGAGCTACTTGTCGTCAT CGTCTTTGTAGTCTCTCAATGCGGGTGA AGATGG-3’	3’ PCR	Construction of pGEX-6p-1/*SLD3 M3*
NR7+ FLAG	5’-CTGAGCTCGAGCTACTTGTCGTCAT CGTCTTTGTAGTCGGGCGAAGCTATAG ATTTTCT-3’	3’ PCR	Construction of pGEX-6p-1/*SLD3 M4*
Sld3 NdeI fwd	5’-GCTGACATATGGAAACATGGGAAG TCATAGCA-3’	5’ PCR	Construction of pFA6a-Nat-NT2/ *SLD3 2E3, 6E*
Sld3 PvuII rev	5’-TACGTCAGCTGCTATGTGGATTCT GGAGC-3’	3’ PCR	Construction of pFA6a-Nat-NT2/ *SLD3 2E3, 6E*

2.6 Strains

2.6.1 *E. Coli* strains

Table 2-8 *E. coli* strains used in this study

Name	**Source**	**Use**	**Genotype**
NEB 5-alpha competent *E. coli*	NEB	Cloning	*E. coli fhuA2Δ(argF-lacZ)U169 phoA glnV44 φ80 Δ(lacZ)M15*

			gyrA96 recA1 relA1 endA1 thi-1 hsdR17
BL21-CodonPlus (DE3)-RIL	Stratagene	Protein expression	*E. coli* B F^- *ompT hsdS*($r_B^- m_B^-$) *dcm*$^+$ *Tet*^{r} *gal λ(DE3) endA Hte [argU ileY BB leuW Cam*^{r}*]*

2.6.2 Yeast strains

Table 2-9 Yeast strains used in this study.

The subscript SUP indicates that the gene's codon usage is optimised for expression.

Name	Genotype	Reference
ySDORC	*MAT**a** ade2-1 ura3-1 his3-11,15 trp1-1 leu2-3,112 can1-100, bar1::Hyg pep4::KanMX, his3::HIS3pRS303/ORC3*$_{SUP}$*, ORC4*$_{SUP}$*, trp1::TRP1pRS304/ORC5*$_{SUP}$*, ORC6*$_{SUP}$*, ura3::URA3pRS306/CBP-ORC1*$_{SUP}$*, ORC2*$_{SUP}$	(Frigola et al., 2013)
yAM33	*MAT**a** ade2-1 ura3-1 his3-11,15 trp1-1 leu2-3,112 can1-100, bar1::Hyg pep4::KanMX, his3::HIS3pRS303/CDT1,GAL4, trp1::TRP1pRS304/MCM4, MCM5, leu2::LEU2pRS305/MCM6, MCM7, ura3::URA3pRS306/MCM2, CBP-MCM3*	(Frigola et al., 2013)
yJF38	*MAT**a** ade2-1 ura3-1 his3-11,15 trp1-1 leu2-3,112 can1-100, bar1::Hyg pep4::KanMX, his3::HIS3pRS303/CDT1,GAL4, trp1::TRP1pRS304/MCM4, MCM5, leu2::LEU2pRS305/MCM6, MCM7, ura3::URA3pRS306/MCM2, 3xFLAG-MCM3*	(Frigola et al., 2013)

ySDK8	MAT**a** ade2-1 ura3-1 his3-11,15 trp1-1 leu2-3,112 can1-100, pep4::KanMX, trp1::TRP1pRS304/CDC7 $_{SUP}$, CBP-DBF4 $_{SUP}$	(On et al., 2014)
yKO3	MAT**a** ade2-1 ura3-1 his3-11,15 trp1-1 leu2-3,112 can1-100, cdc7-4, pep4:: Hyg, his3::HIS3pRS303/SLD3-13MYC, trp1::TRP1pRS304/SLD2, leu2::LEU2pRS305/ SLD7, CDC45, ura3::URA3pRS306/ DPB11	(On et al., 2014)
yJY16	MATα ade2-1 ura3-1 his3-11,15 trp1-1 leu2-3,112 can1-100, cdc7-4, pep4:: KanMX, trp1::TRP1pRS304/SLD2, ura3::URA3pRS306/ DPB11	Generated in the laboratory
yMD7	MAT**a** ade2-1 ura3-1 his3-11,15 trp1-1 leu2-3,112 can1-100, cdc7-4, pep4:: Hyg, MCM6::MCM6-3xFLAG (Nat-NT2)	Generated in the laboratory
yTD2	MAT**a** ade2-1 ura3-1 his3-11,15 trp1-1 leu2-3,112 can1-100, bar1::Hyg pep4::KanMX, his3::HIS3pRS303/DPB11-TCP, GAL4	This study
yTD4	MAT**a** ade2-1 ura3-1 his3-11,15 trp1-1 leu2-3,112 can1-100, bar1::Hyg pep4::KanMX, his3::HIS3pRS303/CDC45$_{SUP}$-TCP, GAL4	This study
yTD6	MAT**a** ade2-1 ura3-1 his3-11,15 trp1-1 leu2-3,112 can1-100, bar1::Hyg pep4::KanMX, his3::HIS3pRS303/SLD3$_{SUP}$-TCP, GAL4, leu2::LEU2pRS305/SLD7 $_{SUP}$	This study
yTD7	MAT**a** ade2-1 ura3-1 his3-11,15 trp1-1 leu2-3,112 can1-100, bar1::Hyg pep4::KanMX, his3::HIS3pRS303/CDC45$_{SUP}$-3XFLAG, GAL4	This study
yTD8	MAT**a** ade2-1 ura3-1 his3-11,15 trp1-1 leu2-3,112 can1-100, bar1::Hyg pep4::KanMX, his3::HIS3pRS303/SLD2$_{SUP}$-3XFLAG, GAL4	This study
yTD11	MAT**a** ade2-1 ura3-1 his3-11,15 trp1-1 leu2-3,112 can1-100, bar1::Hyg pep4::KanMX,	This study

	his3::HIS3pRS303/SLD3$_{SUP}$-3XFLAG, GAL4, leu2::LEU2pRS305/SLD7 $_{SUP}$	
yTD13	*MAT**a** ade2-1 ura3-1 his3-11,15 trp1-1 leu2-3,112 can1-100, cdc7-4, pep4:: Hyg, MCM6::MCM6-3xFLAG (Nat-NT2) his3::HIS3pRS303/CDT1,GAL4, trp1::TRP1pRS304/MCM4, MCM5, leu2::LEU2pRS305/MCM6Δ2-84, MCM7, ura3::URA3pRS306/MCM2, CBP-MCM3*	This study
yTD24	*MAT**a** ade2-1 ura3-1 his3-11,15 trp1-1 leu2-3,112 can1-100, cdc7-4, pep4:: Hyg, MCM6::MCM6-3xFLAG (Nat-NT2) his3::HIS3pRS303/CDT1,GAL4, trp1::TRP1pRS304/MCM4, MCM5, leu2::LEU2pRS305/MCM6TEV2, MCM7, ura3::URA3pRS306/MCM2, CBP-MCM3*	This study
yTD29	*MAT**a** ade2-1 ura3-1 his3-11,15 trp1-1 leu2-3,112 can1-100, cdc7-4, pep4:: Hyg, MCM6::MCM6-3xFLAG (Nat-NT2) his3::HIS3pRS303/CDT1,GAL4, trp1::TRP1pRS304/MCM4, MCM5, leu2::LEU2pRS305/MCM6T150A+11S/T-A, MCM7, ura3::URA3pRS306/MCM2, CBP-MCM3*	This study
yTD30	*MAT**a** ade2-1 ura3-1 his3-11,15 trp1-1 leu2-3,112 can1-100, cdc7-4, pep4:: Hyg, MCM6::MCM6-3xFLAG (Nat-NT2) his3::HIS3pRS303/CDT1,GAL4, trp1::TRP1pRS304/MCM4, MCM5, leu2::LEU2pRS305/MCM6Δ2-84+11S/T-A, MCM7, ura3::URA3pRS306/MCM2, CBP-MCM3*	This study
yTD33	*MAT**a** ade2-1 ura3-1 his3-11,15 trp1-1 leu2-3,112 can1-100, cdc7-4, pep4:: Hyg, MCM6::MCM6-3xFLAG (Nat-NT2)*	This study

	his3::HIS3pRS303/CDT1,GAL4, trp1::TRP1pRS304/MCM4, MCM5, leu2::LEU2pRS305/MCM6Δ2-84+Δ199-260, MCM7, ura3::URA3pRS306/MCM2, CBP-MCM3	
yTD39	*MAT**a** ade2-1 ura3-1 his3-11,15 trp1-1 leu2-3,112 can1-100, cdc7-4, pep4:: Hyg, MCM6::MCM6-3xFLAG (Nat-NT2) his3::HIS3pRS303/CDT1,GAL4, trp1::TRP1pRS304/MCM4, MCM5, leu2::LEU2pRS305/MCM6Δ2-84+3S/T-A, MCM7, ura3::URA3pRS306/MCM2, CBP-MCM3*	This study

2.7 Molecular Biology Methods for *E. coli*

2.7.1 Transformation of *E. coli* with heat-shock method

Chemically competent DH5α cells (NEB) were thawed rapidly at room temperature and placed on ice. 1 aliquot (50 μl) of cells was typically used for 2 transformations. The DNA or ligation reaction was mixed with the cells and left on ice for 30 min. The cells were then incubated at 42 °C for 45 s, and then placed back on ice for 2 min. Following this, the cells were resuspended in 200 μl of SOC medium and incubated at 37 °C for 1 h, with shaking at 1000 rpm. When transforming ligation reactions, 100% of the transformation mixture was plated onto LB agar plates containing the appropriate antibiotic and grown overnight at 37 °C. When re-transforming a plasmid, 10% of the transformation mixture was plated.

2.7.2 Isolation of plasmid DNA

Single *E. coli* colonies were used to inoculate 2 ml LB cultures containing the appropriate antibiotic for selection. The cultures were grown overnight at 37 °C with shaking at 250 rpm. Plasmids were isolated using the QIAprep Spin Miniprep kit (QIAGEN) according to the manufacturer's instructions. The concentration of the

isolated DNA was estimated using the Nanodrop (ND-1000 spectrophotometer, Thermo Scientific) as described in 2.9.2.

2.8 Molecular Biology Methods for Yeast

2.8.1 Transformation of yeast cells

50 ml YPD cultures were grown at 30 °C to a cell density of ~ 1 x 10^7 cells/ml. Cells were harvested by centrifugation at 3500 rpm, 2 min in an A-4-81 rotor (Eppendorf). The supernatant was removed and the cell pellet was resuspended in 25 ml of sterile distilled water, before being centrifuged at 3500 rpm for 2 min in an A-4-81 rotor. The supernatant was removed and the cell pellet was resuspended in 1 ml of sterile distilled water. The sample was transferred to a microfuge tube and cells were then pelleted at maximum speed for 10 s in an FA-45-24-11 rotor (Eppendorf) (20238g). Approximately 900 μl of sterile distilled water was added to the resultant cell pellet to give a final volume of ~1 ml.

10% of the 1 ml suspension was pelleted at maximum speed for 10 s in an FA-45-24-11 rotor. 240 μl 50% PEG 4000, 36 μl 1 M LIOAc, 5 μl 10 mg/ml salmon sperm DNA (Invitrogen) and 2-3 μg of linearised integrative plasmid DNA was added (in that order) to the resultant cell pellet and mixed thoroughly by vortexing. The mixture was incubated at 42 °C for 20 min and the cells were then pelleted at maximum speed for 10 s in an FA-45-24-11 rotor. The cell pellet was resuspended in 200 μl of sterile distilled water and plated on plates lacking the appropriate amino acids for selection. For selection using the natNT2 marker, cells were resuspended in 1 ml YPD, grown overnight at 25 °C, and subsequently plated on YPD supplemented with nourseothricin (as described in section 2.2.2).

2.8.2 Isolation of genomic DNA from yeast

Single colonies of the relevant strain were used to inoculate 4 ml YPD cultures, which were grown to saturation at 30 °C. Cells were harvested by centrifugation at 3500 rpm, 2 min in an A-4-81 rotor (Eppendorf). Cell pellets were then processed

using buffers from the QIAprep Spin Miniprep kit (QIAGEN) as follows. 250 μl of buffer P1 was added to each sample alongside 250 μl fine glass beads. Cells were then disrupted by incubating for 1 min on setting 5.5 on a FastPrep-24 cell disruptor (Zymo Research). 250 μl buffer P2 was added and samples were incubated at 60 °C for 5 min. 350 μl buffer N3 was then added and the samples were centrifuged at maximum speed for 10 min in an FA-45-24-11 rotor (Eppendorf). The supernatant was recovered and then processed and washed as in 2.7.2 according to the manufacturer's instructions. DNA was finally eluted in 50 μl buffer EB.

2.8.3 Sporulation and tetrad dissection

Single colonies were patched onto rich sporulation media and incubated at 25 °C for 2-3 days. The presence of tetrads was checked microscopically. For asci digestion, cells were incubated in 20 μl of 1 mg/ml zymolyase solution made up in 1 M sorbitol and incubated for 10 min at 37 °C. The sample was then diluted back to 1 ml with sterile distilled water and 15 μl was plated onto YPD agar. The tetrads were dissected using a tetrad dissection microscope (Singer), and then grown at 30 °C.

2.8.4 TCA extraction of protein

Cells were harvested from 1 ml of log phase yeast culture (cell density of ~ 10^7 cells/ml) by centrifugation at 3600 rpm for 3 min in an FA-45-24-11 rotor (Eppendorf) (1217g). The supernatant was removed and 200 μl of 20% (w/v) TCA was added to the cell pellet along with 200 μl of fine glass beads. The sample was vortexed vigorously for 30 s, allowed to settle, and the supernatant was then removed and retained. 200 μl of 5% (w/v) TCA was added to the glass beads and then sample was vortexed again for 5 s. The supernatant was removed and added to the supernatant retained from the previous step and the glass beads discarded. The sample was centrifuged at maximum speed in an FA-45-24-11 rotor (Eppendorf) (20238g) for 10 min. The supernatant was removed and discarded and the pellet was resuspended in 50 μl of 2x Laemmli sample buffer. 2-3 μl of 2 M Tris

base was added for neutralisation and the sample was then incubated at 95 °C for 4 min for protein denaturation.

2.9 General Methods for Manipulation of DNA

2.9.1 DNA standards

The commercially available Hyperladder I (Bioline) was used for all DNA gels in this study.

2.9.2 Determination of DNA concentration

DNA concentration was determined by measuring the absorbance of purified DNA samples at 260 nm using the Nanodrop ND-1000 spectrophotometer (Thermo Scientific). The calculations were based on a 50 μg/ml solution of dsDNA giving an absorbance of 1 at 260 nm.

2.9.3 Agarose gel electrophoresis

Agarose gels were made in TAE buffer. 1/6 of the sample volume of 6x DNA loading buffer was added to each sample. Mini-gels (8 x 10 cm) were run in TAE buffer at 80 V (8 V/cm) for 1 h in Bio-Rad horizontal gel electrophoresis apparatus for both analytical and preparative purposes. For analysis of *in vitro* DNA replication assays, mini-gels were run in TAE buffer at 100 V (10 V/cm) for 90 min in TAE buffer, and medium sized gels (15 x 10 cm) were run at 110 V (11 V/cm) for 90 min.

2.9.4 Visualisation of DNA by gel staining

For most agarose gels, SYBR Safe DNA gel stain (Life Technologies) was added to the gel before it set. For analysis of *in vitro* DNA replication assays, agarose gels were stained after running by incubation in TAE containing 0.5 μg/ml ethidium

bromide for 15 min at room temperature. Gels were then destained in TAE for 15 min before visualisation at 254 nm using a BioDoc UV transilluminator (UVP).

2.9.5 Phosphorimager analysis

Dried agarose gels were exposed for 16 h to Amersham Bioscience phosphor screens. The screens were analysed on Typhoon Trio (GE Healthcare) using ImageQuant software.

2.10 General Methods for Manipulation of Proteins

2.10.1 Molecular weight standards

The commercially available BenchMark Protein Ladder (Invitrogen) and PageRuler Plus Prestained Protein Ladder (Thermo Scientific) were used for immunoblotting, silver staining and Coomassie staining. For molecular weight estimation of purified proteins, a mix of proteins of known molecular weights (MW standard for gel filtration – Bio-Rad) was applied to the glycerol gradient or gel filtration column and used to plot a standard curve.

2.10.2 SDS-PAGE

For this study, pre-cast Bis-Tris Nu-PAGE 4-12%, Tris-Acetate Nu-PAGE 3-8% (both Invitrogen), 4-12% Criterion XT Bis-Tris or 3-8% Criterion XT Tris-Acetate (both Bio-Rad) gels were used. Gel samples were supplemented with 2x or 5x Laemmli sample buffer and incubated at 95 °C for 4 min prior to loading. Nu-PAGE 4-12% gels were run in MOPS buffer (Invitrogen) at a constant voltage of 200 V for 1h. Nu-PAGE 3-8% gels were run in Tris-Acetate (Invitrogen) buffer at a constant voltage of 150 V for 80 min. 4-12% Criterion gels were run in MOPS buffer (Bio-Rad) at a constant voltage of 200 V for 1h. Criterion 3-8% gels were run in XT-Tricine (Bio-Rad) buffer at a constant voltage of 150 V for 80 min.

2.10.3 Determination of protein concentration

The protein samples to be analysed were subjected to SDS-PAGE alongside samples of know BSA concentrations. Following coomassie staining, the intensity of the BSA bands was analysed using ImageJ software (NIH) and this was then used to plot a BSA standard curve. Using this standard curve, the concentration of protein sample of interest could be estimated.

2.10.4 Coomassie blue staining

Gels were incubated in Instant Blue staining solution (Expedeon) for 30 min following SDS-PAGE. Gels were destained in distilled water.

2.10.5 Silver staining

Gels were silver stained using the SilverQuest Silver Staining Kit (Invitrogen) according to the manufacturer's instructions.

2.10.6 Immunoblotting

SDS-PAGE was carried out as described in 2.10.2. In order to transfer proteins from the gel onto nitrocellulose membrane (Fisher), 2 pieces of Whatman gel blotting paper (Fisher) and 1 piece of nitrocellulose membrane (both cut to the size of the gel) were initially soaked in western transfer buffer. 1 piece of Whatman paper was then placed on the cathode plate, and the gel, nitrocellulose membrane and 2nd piece of Whatman paper was sequentially laid on top of this (in that order). The whole stack was then covered by the anode plate.

For NuPage gels, the Mini-Trans Blot Cell apparatus (Bio-Rad) was used. For Criterion gels, the Criterion Blotter apparatus (Bio-Rad) was used. In both instances, transfer conditions were set at a constant voltage of 80 V for 90 min at 4 °C. After the transfer, the membrane was incubated in blocking buffer for 30 min at room temperature. Primary antibody incubations were typically performed

overnight at 4 °C in probing buffer, except in the case of the anti-Cdc45 antibody, which was incubated in super blotto buffer overnight. All secondary antibody incubations were performed in probing buffer for 1 h at room temperature.

Following incubation with the secondary antibody, three consecutive 10 min washes in TBST were performed followed by application of the ECL chemiluminescence reagent (Pierce). Membranes were exposed to Amersham hyperfilm ECL (GE Healthcare) and the film was then developed using an automatic X-ray film processor (JP-33 model, Jungwon Precision Industry).

2.11 Protein Purification

2.11.1 Purification of ORC

10 L of *Saccharomyces cerevisiae* (ySDORC) were grown in YP-raffinose at 30 °C to a cell density of 4 x 10^7 cell per ml and arrested for 3 h with 100 ng/ml α-factor. Protein expression was induced by adding galactose to a final concentration of 2% for 3 h at 30 °C. Cells were harvested and washed twice with ice-cold wash buffer I, then washed once in ice-cold buffer A / 100 mM KCl / 2 mM β-mercaptoethanol. The cell pellet was resuspended with 1 volume of ice-cold buffer A / 100 mM KCl / 2 mM β-mercaptoethanol / protease inhibitors (Roche cOmplete, EDTA-free) and frozen drop-wise in liquid nitrogen (popcorn). Frozen popcorn was crushed under liquid nitrogen using a freezer mill (SPEX CertiPrep 6850 Freezer Mill) with 6 cycles of 2 min crushing at a rate setting of 15. Frozen cell powder was thawed at room temperature, then resuspended with buffer A / 100 mM KCl / 2 mM β-mercaptoethanol / protease inhibitors (Roche cOmplete, EDTA-free) and the concentration of KCl adjusted to 500 mM. The suspension was centrifuged for 1 h at 45,000 RPM at 4 °C using a Ti45 rotor (Beckman).

The clear phase was recovered and subjected to calmodulin affinity purification by adding 2 mM $CaCl_2$ and incubating the extract with 4 ml packed beads of calmodulin affinity resin (Stratagene). After 2 h rotation at 4 °C, the beads were recovered, and washed with 15 column volumes (CVs) of buffer A / 300 mM KCl / 2 mM $CaCl_2$ / 2 mM β-mercaptoethanol. Bound protein was eluted in 8 fractions of 1

column volume (CV) each of buffer A / 300 mM KCl / 2 mM EGTA / 1 mM EDTA / 2 mM β-mercaptoethanol.

Peak fractions were pooled, concentrated using an Amicon Ultra 100,000 MWCO centrifugal filter (Millipore), and fractionated over a 120 ml Superdex 200 16/60 PG column (GE Healthcare) pre-equilibrated in buffer A / 300 mM KCl / 2 mM EGTA / 1 mM EDTA / 2 mM β-mercaptoethanol. Peak fractions containing ORC were pooled and dialysed against buffer A / 150 mM KCl / 2 mM EGTA / 1 mM EDTA / 2 mM β-mercaptoethanol for 3 h. The dialysed sample was then fractionated over a 1 ml MonoQ column using a 20 ml gradient of 150 mM – 500 mM KCl. Peak fractions containing ORC were dialysed against buffer A / 300 mM K-Glutamate / 2 mM EGTA / 1 mM EDTA / 2 mM β-mercaptoethanol, concentrated and stored in aliquots at -80 °C.

2.11.2 Purification of Cdc6

The pGEX-6p-1/Cdc6 (yAM3) plasmid was transformed into BL21 CodonPlus RIL cells (Stratagene). The resultant colonies were used to inoculate a 10 ml LB / ampicillin (50 μg/ml) culture, which was grown overnight at 37 °C with shaking at 250 rpm. The following morning, the culture was diluted 1:100 in 1 L of LB / ampicillin (50 μg/ml) / chloramphenicol (35 μg/ml) and left to grow at 37 °C for ~3 h until the culture reached an OD_{600} of 0.6. Cdc6 expression was then induced by the addition of 1 mM IPTG for 5 h at 18 °C. Cells were harvested by centrifugation at 6000 rpm for 10 min in an SLA-3000 rotor (Thermo Scientific), and then washed once in ice-cold PBS.

For lysis, cells were resuspended in 50 ml ice-cold buffer C / 150 mM KOAc / 1% Triton X-100 / protease inhibitors (Roche cOmplete, EDTA-free). Cells were disrupted by passing through a French pressure cell press (2 PLUS, Constant Cell Disruption Systems) three times at 25 kpsi.

The suspension was then centrifuged at 15000 rpm for 15 min in an SS-34 rotor (Sorvall). The supernatant was transferred to a disposable gravity flow column

containing 2 ml packed bead volume of glutathione sepharose resin (GE Healthcare) that had been pre-washed in ice-cold buffer C / 150 mM KOAc / 1% Triton X-100. The column was rotated at 4 °C for 2 h. Glutathione beads and bound protein was recovered in the column and washed with 15 column volumes buffer C / 150 mM KOAc / 1% Triton X-100 / protease inhibitors (Roche cOmplete, EDTA-free) followed by 5 column volumes buffer C / 150 mM KOAc / 1% Triton X-100 without protease inhibitors. A 50% slurry was prepared using buffer C / 150 mM KOAc / 1% Triton X-100 and 50 μl preScission protease (GE Healthcare) was added. The mixture was rotated at 4 °C for 2 h. The flow-through was collected and diluted 2-fold with buffer C / 0.1% Triton X-100 to give a final KOAc concentration of 75 mM.

The diluted flow-through fraction was then applied to 2 ml hydroxyapatite pre-washed in a disposable gravity flow column with buffer C / 75 mM KOAc / 0.1% Triton X-100. The column was rotated at 4 °C for 15 min, and then washed with 5 column volumes buffer C / 75 mM KOAc / 0.1% Triton X-100. Cdc6 was finally eluted in 8 column volumes buffer D. Peak fractions containing Cdc6 were pooled and concentrated using a Centricon Plus-20 centrifugal filter (Millipore). The concentrated sample was aliquoted and stored at -80 °C.

2.11.3 Purification of Mcm2-7/Cdt1

4 L of *Saccharomyces cerevisiae* (yJF38/yAM33) were grown in YP-raffinose at 30 °C to a cell density of 4 x 10^7 cell per ml and arrested for 3 h with 100 ng/ml α-factor. Protein expression was induced by adding galactose to a final concentration of 2% for 3 h at 30 °C. Cells were harvested and washed twice with ice-cold wash buffer I, then washed once in ice-cold buffer B / 100 mM KOAc / 2 mM β-mercaptoethanol. The cell pellet was resuspended with 1 volume of ice-cold buffer B / 100 mM KOAc / 2 mM β-mercaptoethanol / protease inhibitors (Roche cOmplete, EDTA-free) and frozen drop-wise in liquid nitrogen (popcorn). Frozen popcorn was crushed under liquid nitrogen using a freezer mill (SPEX CertiPrep 6850 Freezer Mill) with 6 cycles of 2 min crushing at a rate setting of 15. Frozen cell powder was thawed at room temperature, then resuspended with buffer B / 100

mM KOAc / 2 mM β-mercaptoethanol / protease inhibitors (Roche cOmplete, EDTA-free) and the concentration of KOAc adjusted to 500 mM. The suspension was centrifuged for 1 h at 45,000 RPM at 4 °C using a Ti45 rotor (Beckman). The clear phase was recovered and dialysed for 3 h against buffer B / 100 mM KOAc / 2 mM β-mercaptoethanol, then centrifuged for 30 min at 45,000 RPM at 4 °C using a Ti45 rotor (Beckman).

When purifying Mcm2-7/Cdt1 from yAM33, the clear phase was subjected to calmodulin affinity purification by adding 2 mM $CaCl_2$ and incubating the extract with 2 ml packed beads of calmodulin affinity resin (Stratagene). After 2 h rotation at 4 °C, the beads were recovered, and washed with 15 column volumes (CVs) of buffer B / 100 mM KOAc / 2 mM $CaCl_2$ / 2 mM β-mercaptoethanol. Bound protein was eluted in 8 fractions of 1 column volume (CV) each of buffer B / 100 mM KOAc / 2 mM EGTA / 1 mM EDTA / 2 mM β-mercaptoethanol. When using yJF38, the clear phase was subjected to anti-FLAG immunoprecipitation by adding 3 ml packed bead volume of washed anti-FLAG M2 agarose (Sigma) per 50 ml of extract. After 1 h rotation at 4 °C, the beads were recovered, and washed with 15 column volumes (CVs) of buffer B / 100 mM KOAc / 2 mM β-mercaptoethanol. The washed beads were resuspended with 1 volume buffer B / 100 mM KOAc / 2 mM β-mercaptoethanol containing 1 mg/ml 3xFLAG peptide. After 30 min rotation at 4 °C the eluate fraction was collected.

For both yAM33 and yJF38, peak fractions were then pooled, concentrated using an Amicon Ultra 100,000 MWCO centrifugal filter (Millipore), and fractionated over a 120 ml Superdex 200 16/60 PG column (GE Healthcare) pre-equilibrated in buffer B/ 100 mM KOAc / 2 mM β-mercaptoethanol. Peak fractions containing Mcm2-7/Cdt1 were pooled, concentrated and stored in aliquots at -80 °C.

Purification of Mcm2-7/Cdt1 complexes containing mutant or TEV-cleavable versions of Mcm6 was performed as with yAM33, except that samples were subjected to FLAG immunoprecipitation to deplete endogenous Mcm6-3xFLAG prior to gel filtration. Briefly, peak fractions from the calmodulin affinity purification step were incubated with 0.5 ml packed bead volume of washed anti-FLAG M2 agarose (Sigma). After 1 h rotation at 4 °C, the beads were recovered, and the flow

through was collected, concentrated and subjected to gel filtration as described above.

2.11.4 Purification of DDK

2 L of *Saccharomyces cerevisiae* (ySDK8) were grown in YP-raffinose at 30 °C to a cell density of 2 x 10^7 cell per ml. Protein expression was induced by adding galactose to a final concentration of 2% for 6 h at 30 °C. Cells were harvested and washed twice with ice-cold wash buffer I, then washed once in ice-cold buffer A / 400 mM NaCl / 2 mM β-mercaptoethanol. The cell pellet was resuspended with 1 volume of ice-cold buffer A / 400 mM NaCl / 2 mM β-mercaptoethanol / protease inhibitors (Roche cOmplete, EDTA-free) and frozen drop-wise in liquid nitrogen (popcorn). Frozen popcorn was crushed under liquid nitrogen using a freezer mill (SPEX CertiPrep 6850 Freezer Mill) with 6 cycles of 2 min crushing at a rate setting of 15. Frozen cell powder was thawed at room temperature, then resuspended with buffer A / 400 mM NaCl / 2 mM β-mercaptoethanol / protease inhibitors (Roche cOmplete, EDTA-free) and the concentration of NaCl adjusted to 500 mM. The suspension was centrifuged for 1 h at 45,000 RPM at 4 °C using a Ti45 rotor (Beckman).

The clear phase was recovered and subjected to calmodulin affinity purification by adding 2 mM $CaCl_2$ and incubating the extract with 2 ml packed beads of calmodulin affinity resin (Stratagene). After 2 h rotation at 4 °C, the beads were recovered, and washed with 15 column volumes (CVs) of buffer A / 400 mM NaCl / 2 mM $CaCl_2$ / 2 mM β-mercaptoethanol. Bound protein was eluted in 8 fractions of 1 column volume (CV) each of buffer A / 400 mM NaCl / 2 mM EGTA / 1 mM EDTA / 2 mM β-mercaptoethanol. Peak fractions were pooled and dephosphorylated by the addition of 200 μg lambda phosphatase (4 °C, 16 h).

Subsequently, the sample was concentrated using an Amicon Ultra 30,000 MWCO centrifugal filter (Millipore), and fractionated over a 24 ml Superdex 200 10/300 column (GE Healthcare) pre-equilibrated in buffer A / 400 mM KCl / 2 mM EGTA / 1 mM EDTA / 2 mM β-mercaptoethanol. Peak fractions containing DDK were pooled

and dialysed against buffer A / 200 mM K-Glutamate / 2 mM β-mercaptoethanol. The dialysed sample was concentrated and stored in aliquots at -80 °C.

2.11.5 Purification of Sld3/7 from *S. cerevisiae*

10 L of *Saccharomyces cerevisiae* (yTD6) were grown in YP-raffinose at 30 °C to a cell density of 4×10^7 cell per ml and arrested for 3 h with 100 ng/ml α-factor. Protein expression was induced by adding galactose to a final concentration of 2% for 3-4 h at 30 °C. Cells were harvested and washed twice with ice-cold wash buffer I, then washed once in ice-cold buffer A / 500 mM KCl / 1 mM DTT. The cell pellet was resuspended with 1 volume of buffer A / 500 mM KCl / 1 mM DTT / homemade protease inhibitor cocktail and frozen drop-wise in liquid nitrogen (popcorn). Frozen popcorn was crushed under liquid nitrogen using a freezer mill (SPEX CertiPrep 6850 Freezer Mill) with 6 cycles of 2 min crushing at a rate setting of 15. Frozen cell powder was thawed at room temperature, then resuspended with 1 volume of buffer A / 500 mM KCl / 1 mM DTT / homemade protease inhibitor cocktail. The suspension was centrifuged for 1 h at 45,000 RPM at 4 °C using a Ti45 rotor (Beckman).

When purifying Sld3/7 from yTD6, the clear phase was recovered and subjected to IgG affinity purification using 1 ml packed beads of IgG sepharose 6 Fast Flow slurry (GE Healthcare) per 10 ml of extract. IgG beads were rotated at 4 °C for 45 min with the extract, recovered, and washed with 15 column volumes (CVs) of buffer A / 500 mM KCl / 1 mM DTT / homemade protease inhibitor cocktail. The washed IgG beads were resuspended with 1 volume buffer A / 500 mM KCl / 1 mM DTT and bound protein was eluted by adding 0.6 mg HIS_6-TEV protease per 1 ml of packed IgG beads. After 2 h rotation at 4 °C the eluate fraction was collected, and incubated with 1 ml packed Ni^{2+}-NTA beads (Qiagen) for 5 min on ice to deplete HIS_6-TEV protease. When purifying Sld3/7 from yTD11, the clear phase was subjected to anti-FLAG immunoprecipitation by adding 3 ml packed bead volume of washed anti-FLAG M2 agarose (Sigma). After 1 h rotation at 4 °C, the beads were recovered, and washed with 15 column volumes (CVs) of buffer A / 500 mM KCl / 1 mM DTT / homemade protease inhibitor cocktail. The washed

beads were resuspended with buffer A / 500 mM KCl / 1 mM DTT containing 1 mg/ml 3 x FLAG peptide. After 30 min rotation at 4 °C the eluate fraction was collected. This elution was repeated two more times to give a final eluate fraction of 9 ml.

For both yTD6 and yTD11, the resulting fraction was concentrated using an Amicon Ultra 30,000 MWCO centrifugal filter (Millipore), and fractionated over a 24 ml Superdex 200 10/300 GL column (GE Healthcare) pre-equilibrated in buffer A / 500 mM KCl / 1 mM DTT. Peak fractions containing Sld3/7 were pooled, concentrated and stored in aliquots at -80 °C.

2.11.6 Purification of Sld3 and Sld7 from *E. coli*

The relevant pGEX-6p-1/Sld3 or pGEX-6p-1/Sld7 plasmid (see Table 2-6) was transformed into BL21 CodonPlus RIL cells (Stratagene). The resultant colonies were used to inoculate a 50 ml LB / ampicillin (50 μg/ml) culture, which was grown overnight at 37 °C with shaking at 250 rpm. The following morning, the culture was diluted 1:10 in 250 ml of LB / ampicillin (50 μg/ml) / chloramphenicol (35 μg/ml) and left to grow at 25 °C for ~3-4 h until the culture reached an OD_{600} of 0.7. Expression was then induced by the addition of 1 mM IPTG for 4 h at 25 °C. Cells were harvested by centrifugation at 3500 rpm for 15 min in an SLA-3000 rotor (Thermo Scientific), and then washed once in ice-cold PBS.

For lysis, the cell pellets were resuspended in 10 ml buffer A / 500 mM KCl / 1 mM DTT / homemade protease inhibitor cocktail and lysozyme was added to a final concentration of 500 μg/ml. The mixture was left on ice for 20 min, and then sonicated for 2 min (25 s on, 5 s off) at setting 5 on a Sonicator Ultrasonic Processor XL (Misonix). The suspension was then centrifuged at 15000 rpm for 15 min in an SS-34 rotor (Sorvall). The supernatant was subjected to GST affinity purification by incubation with 0.8 ml packed bead volume of glutathione sepharose resin (GE Healthcare) that had been pre-washed in buffer A / 500 mM KCl / 1 mM DTT. The sample was rotated at 4 °C for 2 h. Glutathione beads and bound protein was recovered in a disposable gravity flow column and washed with 15 column

volumes buffer A / 500 mM KCl / 1 mM DTT / homemade protease inhibitor cocktail followed by 5 column volumes buffer A / 500 mM KCl / 1 mM DTT without protease inhibitors. A 50% slurry was prepared using buffer A / 500 mM KCl / 1 mM DTT and 50 μl preScission protease (GE Healthcare) was added. The mixture was rotated at 4 °C overnight. When preparing GST-Sld3, the tagged protein was eluted in 6 fractions of 1 column volume (CV) each of buffer A / 500 mM KCl / 1 mM DTT / 10 mM glutathione.

The following morning, the flow-through was collected and one further column volume of buffer A / 500 mM KCl / 1mM DTT was flowed over the column to elute any remaining cleaved protein. The two flow-through fractions were pooled, concentrated using an Amicon Ultra 30,000 MWCO centrifugal filter (Millipore), and fractionated over a 24 ml Superdex 200 10/300 GL column (GE Healthcare) pre-equilibrated in buffer A / 500 mM KCl / 1 mM DTT. Peak fractions containing Sld3 or Sld7 were pooled, concentrated and stored in aliquots at -80 °C.

Purification of FLAG-tagged wild type or mutant versions of Sld3 or Sld3 fragments was performed as described above except the final gel filtration step was omitted.

2.11.7 Purification of Dpb11

Dpb11 was purified as described for Sld3/7 in 2.11.5 barring two notable differences. Firstly, yTD2 was used in place of yTD6 in the initial culture. Secondly, for Dpb11, the final 24 ml Superdex 200 10/300 GL column was run in buffer A / 300 mM K-Glutamate / 1 mM DTT.

2.11.8 Purification of Sld2

10 L of *Saccharomyces cerevisiae* (yTD8) were grown in YP-raffinose at 30 °C to a cell density of 4 x 10^7 cell per ml and arrested for 3 h with 100 ng/ml α-factor. Protein expression was induced by adding galactose to a final concentration of 2% for 3-4 h at 30 °C. Cells were harvested and washed twice with ice-cold wash buffer I, then washed once in ice-cold buffer A / 500 mM KCl / 1 mM DTT. The cell

pellet was resuspended with 1 volume of buffer A / 500 mM KCl / 1 mM DTT / homemade protease inhibitor cocktail and frozen drop-wise in liquid nitrogen (popcorn). Frozen popcorn was crushed under liquid nitrogen using a freezer mill (SPEX CertiPrep 6850 Freezer Mill) with 6 cycles of 2 min crushing at a rate setting of 15. Frozen cell powder was thawed at room temperature, then resuspended with 1 volume of buffer A / 500 mM KCl / 1 mM DTT / homemade protease inhibitor cocktail. The suspension was centrifuged for 1 h at 45,000 RPM at 4 °C using a Ti45 rotor (Beckman).

The clear phase was recovered and subjected to anti-FLAG immunoprecipitation by adding 5 ml packed bead volume of washed anti-FLAG M2 agarose (Sigma) per 50 ml of extract. After 1 h rotation at 4 °C, the beads were recovered, and washed with 15 column volumes (CVs) of buffer A / 500 mM KCl / 1 mM DTT / homemade protease inhibitor cocktail. The washed beads were resuspended with buffer A / 500 mM KCl / 1 mM DTT containing 1 mg/ml 3xFLAG peptide. After 30 min rotation at 4 °C the eluate fraction was collected.

2.11.9 Purification of Cdc45

10 L of *Saccharomyces cerevisiae* (yTD4 or yTD7) were grown in YP-raffinose at 30 °C to a cell density of 4 x 10^7 cell per ml and arrested for 3 h with 100 ng/ml α-factor. Protein expression was induced by adding galactose to a final concentration of 2% for 3-4 h at 30 °C. Cells were harvested and washed twice with ice-cold wash buffer I, then washed once in ice-cold buffer A / 500 mM KCl / 1 mM DTT. The cell pellet was resuspended with 1 volume of buffer A / 500 mM KCl / 1 mM DTT / homemade protease inhibitor cocktail and frozen drop-wise in liquid nitrogen (popcorn). EDTA was omitted from the protease inhibitor cocktail throughout when using yTD4. Frozen popcorn was crushed under liquid nitrogen using a freezer mill (SPEX CertiPrep 6850 Freezer Mill) with 6 cycles of 2 min crushing at a rate setting of 15. Frozen cell powder was thawed at room temperature, then resuspended with 1 volume of buffer A / 500 mM KCl / 1 mM DTT / homemade protease inhibitor cocktail. The suspension was centrifuged for 1 h at 45,000 RPM at 4 °C using a Ti45 rotor (Beckman).

When using yTD4, the clear phase was recovered and subjected to calmodulin affinity purification by adding 2 mM $CaCl_2$ and incubating the extract with 2 ml packed beads of calmodulin affinity resin (Stratagene). After 2 h rotation at 4 °C, the beads were recovered, and washed with 15 column volumes (CVs) of buffer A / 500 mM KCl / 2 mM $CaCl_2$ / 1 mM DTT / homemade protease inhibitor cocktail. Bound protein was eluted in 8 fractions of 1 column volume (CV) each of buffer A / 500 mM KCl / 2 mM EGTA / 1 mM ETDA / 1 mM DTT. When using yTD7, the clear phase was recovered and subjected to anti-FLAG immunoprecipitation by adding 3 ml packed bead volume of washed anti-FLAG M2 agarose (Sigma) per 50 ml of extract. After 1 h rotation at 4 °C, the beads were recovered, and washed with 15 column volumes (CVs) of buffer A / 500 mM KCl / 1 mM DTT / homemade protease inhibitor cocktail. The washed beads were resuspended with buffer A / 500 mM KCl / 1 mM DTT containing 1 mg/ml 3xFLAG peptide. After 30 min rotation at 4 °C the eluate fraction was collected.

2.12 Preparation of S-phase Whole Cell Extract

4 L of *Saccharomyces cerevisiae* (yKO3) were grown in YP-raffinose at 25 °C to a cell density of 1 x 10^7 cell per ml and galactose was then added to a final concentration of 2% for 2 h. Cells were shifted to 37 °C and then grown for a further 5 h at 37 °C. Cells were harvested by centrifugation at 4000 rpm for 15 min, and then washed twice in SE buffer I / 3 mM DTT. The cell pellet was resuspended in 0.3 volumes of SE buffer II / 1.5 M K-Glutamate / 5 mM DTT / protease inhibitors (Roche cOmplete, EDTA-free) and frozen drop-wise in liquid nitrogen (popcorn). Frozen popcorn was crushed under liquid nitrogen using a freezer mill (SPEX CertiPrep 6850 Freezer Mill) with 6 cycles of 2 min crushing at a rate setting of 15. Frozen cell powder was transferred to an ultracentrifuge tube and allowed to thaw completely for ~ 1-2 h on ice.

The thawed lysate was centrifuged at 50,000 rpm for 1 h at 4 °C in a Ti70 rotor (Beckman). The clear phase was recovered and dialysed against SE buffer III / 3 mM DTT / protease inhibitors (Roche cOmplete, EDTA-free) for 3.5 h at 4 °C. The

dialysed extract was subsequently recovered and centrifuged for 30 min at 90,000 rpm in a TLA 100.3 rotor (Beckman). The supernatant was recovered, aliquoted and stored at -80.

2.13 Preparation of DNA Substrates for *In Vitro* Biochemical Assays

2.13.1 1 kb photocleavable linear DNA beads

A 1048 bp linear DNA fragment containing the ARS305 DNA replication origin was amplified by PCR using primers ARS305-F-PC-Bio and ARS305-R and p305bp as a template. The product of this PCR reaction had a biotin group at the 5’ end, which was coupled to the DNA via a photocleavable linker. 16 x 50 μl PCR reactions were set up and the products were purified to a final volume of 400 μl using the High Pure PCR Purification Kit (Roche).

The beads from 400 μl slurry of streptavidin-coated M-280 dynabeads (Invitrogen) were isolated on a magnetic rack and the supernatant removed. The beads were washed twice in 500 μl buffer I, and then resuspended in 400 μl buffer I. 400 μl purified PCR product from the first step was added to this suspension and the mixture was rotated overnight at 4 °C.

The following morning, the beads were washed twice in 500 μl buffer II and then twice more in 500 μl buffer III. Finally, the beads were resuspended in 200 μl buffer III and stored at 4 °C.

2.13.2 Purification of plasmid DNA template for in vitro replication assays

Plasmid DNA was purified from *E. coli* using the QIAGEN Plasmid Plus Maxiprep kit (QIAGEN) according to the manufacturer’s instructions.

2.14 Protein-Protein and Protein-DNA interaction Assays

2.14.1 *In vitro* pre-RC assembly assay

2 μl of photocleavable DNA beads were used per 40 μl reaction. Reactions were assembled on ice in 25 mM Hepes-KOH pH 7.6, 10 mM $Mg(OAc)_2$, 0.02% (v/v) NP-40, 5% glycerol, 0.1 M KOAc, 1 mM DTT and 5 mM ATP. Purified ORC, Cdc6 and Mcm2-7/Cdt1 were added to this mixture (in that order) to final concentrations of 50 nM, 50 nM and 100 nM, respectively. Samples were incubated for 30 min at 30 °C with agitation at 1000 rpm. Following the incubation, beads were isolated on a magnetic rack, the supernatant was removed, and each sample was then washed in 400 μl of low salt wash buffer by vortexing. The beads were re-isolated on the magnetic rack and washed in 400 μl of either low salt or high salt wash buffer (as indicated). The wash buffer was then removed and the beads and bound proteins were resuspended in 20 μl of low salt wash buffer. Protein-DNA complexes were then eluted from the beads via UVA irradiation at 330 nm for 10 min. The 20 μl eluate fraction was supplemented with 5 x Laemmli sample buffer and incubated at 95 °C for 4 min for subsequent analysis by SDS-PAGE.

2.14.2 Sld3/7 recruitment assay

Pre-RC assembly reactions were first set up as described in 2.14.1. After 30 min incubation at 30 °C, each 40 μl sample was split into 2 x 20 μl aliquots, and the beads were then isolated on a magnetic rack. The supernatant was removed and replaced with a 20 μl reaction mixture containing 25 mM Hepes-KOH pH 7.6, 10 mM $Mg(OAc)_2$, 0.02% (v/v) NP-40, 5% glycerol, 0.1 M KOAc, 1 mM DTT and 5 mM ATP. Purified DDK was added to a final concentration of 50 nM and the samples were incubated for 30 min at 25 °C with agitation at 1000 rpm. Beads were re-isolated and the supernatant removed. A 40 μl reaction mixture containing 25 mM Hepes-KOH pH 7.6, 10 mM $Mg(OAc)_2$, 0.02% (v/v) NP-40, 5% glycerol, 0.5 M KOAc, 50 mM KCl and 1 mM DTT was added, and purified Sld3/7 was then added to a final concentration of 5 nM. Samples were incubated at 25 °C for 5 min with agitation at 1000 rpm. Beads were isolated on a magnetic rack, the supernatant

removed, and each sample was then subjected to 2 x 400 μl washes in pre-RC wash buffer / 0.5 M KOAc. The wash buffer was then removed and the beads and bound proteins were resuspended in 10 μl of pre-RC wash buffer / 0.5 M KOAc. Protein-DNA complexes were then eluted from the beads via UVA irradiation at 330 nm for 10 min. The 10 μl eluate fraction was supplemented with 10 μl 2x Laemmli sample buffer and incubated at 95 °C for 4 min for subsequent analysis by SDS-PAGE.

2.14.3 Replisome assembly assay

Pre-RC assembly reactions were first set up as described in 2.14.1, except the incubation was performed at 25 °C for 30 min. Each 40 μl sample was split into 2 x 20 μl aliquots, and the beads were then isolated on a magnetic rack. The supernatant was removed and replaced with a 15 μl reaction mixture containing 25 mM Hepes-KOH pH 7.6, 10 mM $Mg(OAc)_2$, 0.02% (v/v) NP-40, 5% glycerol, 0.1 M K-Glutamate, 1 mM DTT and 5 mM ATP. Purified DDK was added to a final concentration of 100 nM and the samples were incubated for 30 min at 25 °C with agitation at 1000 rpm. Beads were re-isolated and the supernatant removed. A 40 μl reaction mixture containing 40 mM Hepes-KOH pH 7.6, 8 mM $MgCl_2$, 40 mM creatine phosphate, 10 μg creatine phosphokinase, 1 mM DTT and 5 mM ATP was added. 500 μg yKO3 S-phase extract was then added and SE buffer III / 3 mM DTT was used to adjust the final concentration of K-Glutamate to 150 mM. Samples were incubated at 25 °C for 15 min with agitation at 1000 rpm. Beads were isolated on a magnetic rack, the supernatant removed, and each sample was then subjected to 3 x 175 μl washes in low salt wash buffer. The wash buffer was then removed and the beads and bound proteins were resuspended in 20 μl 2x Laemmli sample buffer and incubated at 95 °C for 4 min for subsequent analysis by SDS-PAGE.

When using extract made from yJY16 (as in Fig. 4.8), endogenous Cdc45 was initially depleted by incubation with Sld3-FLAG beads. 60 pmol FLAG-Sld3 was initially coupled to 37.5 μl magnetic anti-FLAG resin (Sigma), and 2.5 μl of this Sld3-coupled resin was used to deplete 1 mg of yJY16 extract. 2 x 1 h rounds of

depletion with fresh Sld3-FLAG beads were routinely performed. Recombinant FLAG-Sld3, Sld7 and Cdc45 were added to the depleted extract to final concentrations of 50 nM, 50 nM and 100 nM, respectively. Extract was then used in the replisome assembly reaction as described above.

2.14.4 Interaction of Sld3/7 with individual Mcm subunits

A 120 μl pre-RC assembly reaction (using 6 μl of photocleavable DNA beads) was prepared per sample. Reactions were performed exactly as in 2.14.1, with all reaction components scaled up 3-fold. After the pre-RC assembly step, each sample was subjected to sequential 400 μl wash steps using low salt wash buffer, high salt wash buffer and pre-RC wash buffer / 0.1 M KOAc (in that order). Beads were then resuspended in a 120 μl reaction mixture containing 25 mM Hepes-KOH pH 7.6, 10 mM $Mg(OAc)_2$, 0.02% (v/v) NP-40, 5% glycerol, 0.1 M KOAc, 1 mM DTT and 5 mM ATP. Purified DDK was added to a final concentration of 50 nM and the samples were incubated for 30 min at 25 °C with agitation at 1000 rpm. Beads were re-isolated, the supernatant removed and each sample washed in 400 μl high salt wash buffer. The wash buffer was removed and each beads sample was resuspended in 65 μl high salt wash buffer and incubated on ice overnight.

The following morning, the beads were isolated on a magnetic rack and a 60 μl supernatant fraction was recovered from each sample. A 10 μl input fraction was removed at this point, and 75 μl pre-RC wash buffer was added to the remaining 50 μl sample to adjust the concentration of NaCl to 200 mM. The resultant 125 μl sample was then added to 2.5 μl magnetic anti-FLAG beads (Sigma), which had been pre-coupled to Sld3-3xFLAG/Sld7 in high salt wash buffer and then washed in pre-RC wash buffer / 0.2 M NaCl. Samples were incubated for 1 h at 8 °C with agitation at 1100 rpm. Next, beads were isolated on a magnetic rack, the supernatant removed, and each sample was then subjected to 2 x 200 μl washes in pre-RC wash buffer / 0.2 M NaCl. In early experiments, the beads and bound proteins were then resuspended in 10 μl 2x Laemmli sample buffer and incubated at 95 °C for 4 min for subsequent analysis by SDS-PAGE. Later, input and beads samples were routinely dephosphorylated with lambda phosphatase (1 h, 30 °C)

prior to the addition of Laemmli sample buffer to prevent the masking of epitopes by phosphorylation during immunoblotting.

2.14.5 Immunoprecipitation of FLAG-Sld3 from S-phase extract

20 μl reactions were assembled containing 40 mM Hepes-KOH pH 7.6, 8 mM $MgCl_2$, 40 mM creatine phosphate, 10 μg creatine phosphokinase, 1 mM DTT and 5 mM ATP. 250 μg yKO3 S-phase extract, which had been depleted of Sld3-13Myc (see 2.15), was then added and SE buffer III / 3 mM DTT was used to adjust the final concentration of K-Glutamate to 150 mM. 0.5 pmol wild type or mutant FLAG-Sld3 and 0.5 pmol Sld7 were then added and the reaction incubated on ice for 15 min. The reaction mixture was then added to 2.5 μl magnetic anti-FLAG beads (Sigma), which had been pre-washed in SE buffer III / 3 mM DTT. Samples were incubated for 1 h on ice with occasional agitation. Next, beads were isolated on a magnetic rack, the supernatant removed, and each sample was subjected to 2 x 200 μl washes in pre-RC wash buffer / 0.15 M K-Glutamate. The beads were then resuspended in 10 μl 2x Laemmli sample buffer and incubated at 95 °C for 4 min for subsequent analysis by SDS-PAGE.

2.14.6 Interaction of Sld3 with Mcm6 peptides

N-terminally biotinylated peptides and peptide arrays were synthesised by the Peptide Synthesis Laboratory at the CRUK London Research Institute. Peptides bound to membranes were initially solubilised by washing in 50% EtOH / 10% acetic acid for 1 h. The membranes were washed briefly in TBST, then incubated at 4 °C overnight in TBST / 3% BSA. The following morning, membranes were washed briefly in Peptide Array Buffer, and then incubated with 40 ml Peptide Array Buffer / 3% BSA containing FLAG-Sld3 at a final concentration of 2 nM for 1 h at room temperature. Membranes were subsequently washed for 3 x 3 min in Peptide Array Buffer, then probed with anti-FLAG antibody (Sigma) for 3 h at room temperature. Three consecutive 10 min washes in Peptide Array Buffer were then performed followed by application of the ECL chemiluminescence reagent (Pierce).

N-terminally biotinylated peptides were initially coupled to streptavidin-coated M-280 dynabeads (Invitrogen) by incubation of 40 pmol of peptide with 4 μl beads for 30 min at 20 °C with agitation at 1250 rpm. Peptide-coupled beads were subsequently washed in Peptide Array Buffer. An 80 μl reaction mixture containing Peptide Array Buffer and FLAG-Sld3 at a final concentration of 5 nM was then added, and the reaction incubated at 25 °C for 10 min with agitation at 1000 rpm. Beads were isolated on a magnetic rack, the supernatant removed, and each sample was then subjected to 2 x 200 μl washes in Peptide Array Buffer. The beads were then resuspended in 20 μl 2x Laemmli sample buffer and incubated at 95 °C for 4 min for subsequent analysis by SDS-PAGE.

2.14.7 Interaction of Sld3 and Sld7

A 40 μl reaction was prepared containing 10 pmol each of full-length or truncated Sld3-FLAG, GST-Sld3 and Sld7. The concentration of KCl was adjusted to 0.2 M by the addition of 60 μl buffer A, and samples incubated on ice for 30 min. BSA was added to a final concentration of 100 μg/ml, and the mixture then incubated with 5 μl magnetic anti-FLAG resin (Sigma) pre-washed in buffer A / 0.2 M KCl for 30 min on ice. Beads were then washed twice with 500 μl of buffer A / 0.2 M KCl. Bound proteins were eluted either by boiling in 2x Laemmli sample buffer, or by incubation with buffer A / 0.2 M KCl containing 1 mg/ml 3xFLAG peptide for 30 min at 8 °C.

2.15 *In Vitro* DNA Replication Assay

Pre-RC assembly reactions were performed on soluble plasmid DNA, with 1 μg of template DNA used per 25 μl reaction. Reactions were assembled on ice in 25 mM Hepes-KOH pH 7.6, 10 mM $Mg(OAc)_2$, 0.02% (v/v) NP-40, 5% glycerol, 0.1 M K-Glutamate, 2 mM DTT and 5 mM ATP. Purified ORC, Cdc6 and Mcm2-7/Cdt1 were added to this mixture (in that order) to final concentrations of 75 nM, 75 nM and 150 nM, respectively. Samples were incubated for 20 min at 25 °C. The Pre-RC assembly reaction was subsequently split into 5 μl aliquots, and the volume

adjusted to 15 μl using the same conditions as in the first step. Purified DDK was added to a final concentration of 100 nM, and reactions were incubated for 30 min at 25 °C. Next, a reaction mixture containing 40 mM Hepes-KOH pH 7.6, 8 mM $MgCl_2$, 2 mM DTT, 5 mM ATP, 100 μM dATP/dTTP/dGTP, 200 μM CTP/GTP/UTP, 5 μCi ^{32}P-α-dCTP, 40 mM creatine phosphate and 10 μg creatine phosphokinase was added to the 15 μl DDK reaction. 1 mg of yKO3 S-phase extract was added to the reaction and the final volume adjusted to 50 μl. Samples were incubated for 45 min at 25 °C. Replication was quenched by the addition of 20 mM EDTA, and proteins were then denatured and digested by the addition of 0.5% SDS and 20 μg proteinase K. Samples were incubated for 30 min at 37 °C with agitation at 1000 rpm. 55 μl phenol/chloroform was added to the reaction and phenol/chloroform extraction was performed. Unincorporated radioactive dCTP was subsequently removed from the aqueous layer by passing through an Illustra Microspin S-400 HR column (GE Healthcare) according to manufacturer's instructions. Finally, RNA was digested by the addition of 2 units of RNase A (Sigma) for 15 min at 30 °C. Samples were analysed by gel electrophoresis and autoradiography.

For depletion of Sld3-13Myc, yKO3 extract was incubated with anti-Myc magnetic beads (Origene) from an equal volume of slurry as extract. 2 x 1 h depletions were performed on ice with fresh anti-Myc beads added each time.

Chapter 3. Results 1

3.1 Introduction

In G1 phase of the cell cycle, the Mcm2-7 replicative helicase is loaded onto DNA at origins of replication as a double hexamer, forming a pre-RC (Remus et al., 2009, Evrin et al., 2009, Gambus et al., 2011). In the pre-RC, the Mcm2-7 complex encircles dsDNA and is inactive as a helicase (Remus et al., 2009, Evrin et al., 2009). DNA synthesis is only initiated at the origin upon progression into S-phase, when the activities of CDK and DDK increase and the Mcm2-7 double hexamer is converted into its active form, the CMG complex (Boos et al., 2012). This activation of the replicative helicase requires not only the activities of the aforementioned protein kinases, but also a whole host of specific ‘firing factors’, namely Sld3, Sld7, Cdc45, Dpb11, Sld2, GINS, Polε and Mcm10 (Boos et al., 2012, Tanaka and Araki, 2013). Whilst many years of *in vivo* studies have confirmed the absolute requirement of these proteins for origin firing, relatively little is known about how they contribute mechanistically towards the formation of the active replicative helicase.

The availability of a reconstituted system that recapitulates origin firing *in vitro* with purified proteins would be highly beneficial for our understanding of the Mcm2-7 activation process. The first step towards this was to reconstitute the Mcm2-7 loading reaction *in vitro* (Remus et al., 2009, Evrin et al., 2009). Importantly, it has recently been shown that pre-RCs assembled *in vitro* can support DNA replication in an S-phase extract (On et al., 2014, Gros et al., 2014), consistent with the formation of an active precursor for DNA replication during the pre-RC assembly step. Mass spectrometry analysis of the proteins recruited to the loaded Mcm2-7 complex in this S-phase extract did not detect any new factors that had not been previously shown to be involved in replication initiation (On et al., 2014). We thus decided to proceed with attempts to reconstitute Mcm2-7 activation with the minimal set of proteins mentioned above.

The ability to purify sufficient amounts of the firing factor proteins is a prerequisite for the progress of this project. In this chapter, I will discuss progress made in

purifying a number of the proteins required for Mcm2-7 activation, and will describe the characterisation of the *in vitro* behaviour of one of these factors, the Sld3/7 complex. I will finally show how I was able to reconstitute the recruitment of Sld3/7 to the pre-RC *in vitro*, forming a complex that may represent an early intermediate in the helicase activation process.

3.2 Results

3.2.1 Strain construction

Separate *S. cerevisiae* strains were initially constructed in which second copies of *SLD3, SLD7, CDC45, DPB11* or *SLD2* were placed under the control of the *GAL1,10* galactose-inducible promoter. Evidence suggests that Sld3 and Sld7 form a constitutive complex throughout the cell cycle (Tanaka et al., 2011b), and *SLD3 and SLD7* were thus co-expressed from separate promoters in the same strain. The *GAL1,10* promoter has been used extensively in the past to overexpress proteins for subsequent purification, and we aimed to further increase expression levels through optimisation of the codon usage in the genes of interest. The codon-optimised genes were expressed to markedly higher levels than the wild type versions (data not shown), allowing for improved yields during protein purification. Protein expression was confirmed by immunoblot analysis, using antibodies raised against the firing factors of interest (Fig. 3.1).

3.2.2 Optimisation of extract conditions

For purification, yeast cells expressing Sld3/7, Sld2, Dpb11 or Cdc45 were arrested in G1 with the α-factor mating pheromone, allowing for the purification of a homogenous protein sample that should not carry any S-phase specific post-translational modifications. Modifications required for replication initiation, such as phosphorylation by DDK and CDK, could then be performed *in vitro* using purified proteins.

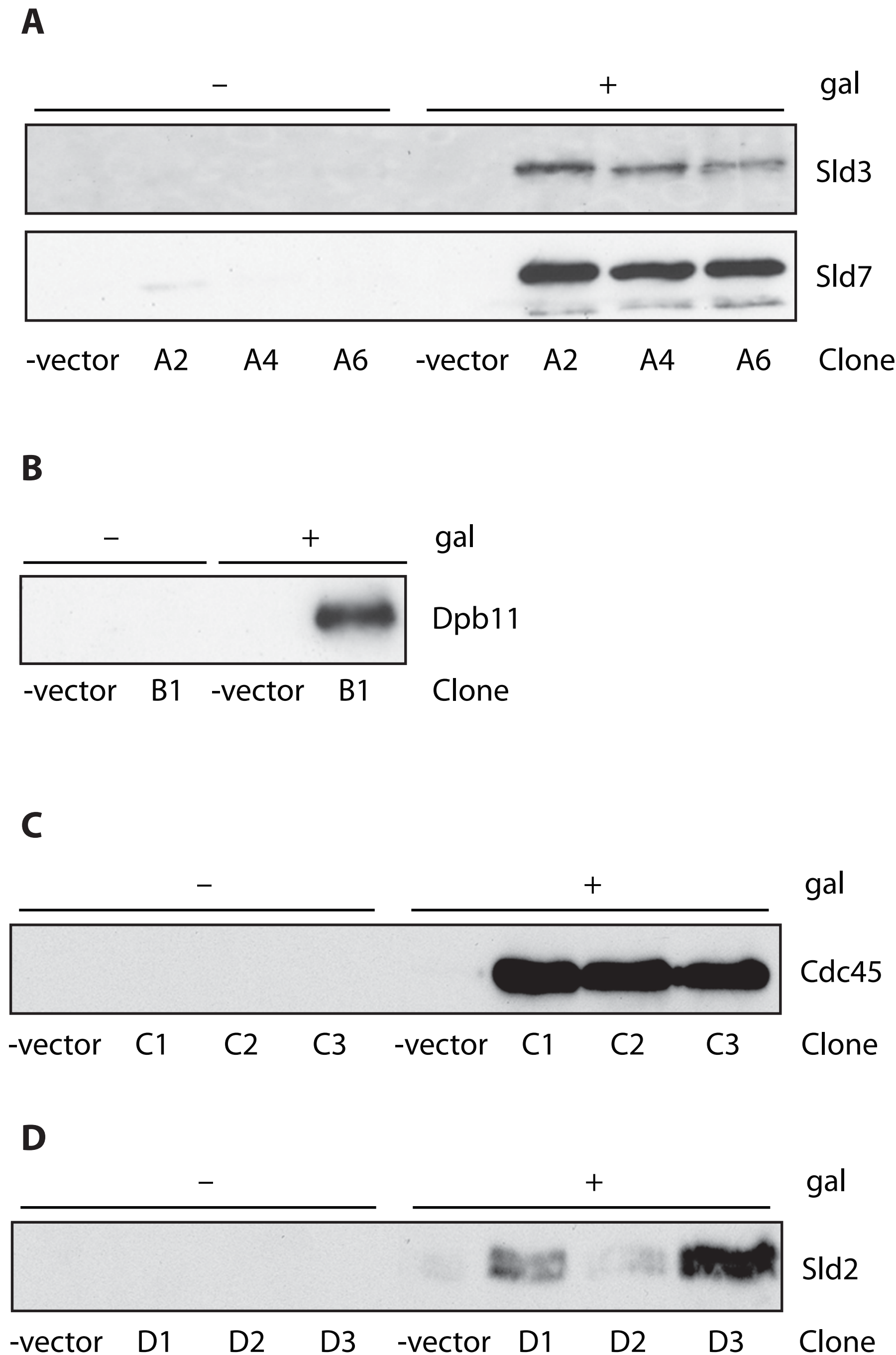


Figure 3.1 Construction of *S. cerevisiae* strains for firing factor purification

Immunoblot analysis showing galactose-inducible overexpression of Sld3/Sld7 (**A**), Dpb11 (**B**), Cdc45 (**C**) and Sld2 (**D**). Cultures of individual transformant clones were grown in YP-raffinose to log phase, arrested in G1 with α-factor, and galactose added to 2% for 3 h. Samples taken before and after galactose addition were TCA extracted, and analysed by immunoblot. The parental strain (-vector) was included as a control.

Having confirmed protein expression under these conditions, I next set out to optimise conditions for protein extraction and purification. As these proteins had not been extensively characterised before, I first tested their solubility in a cell lysate after overexpression. Notably, Sld3, Sld7 and Sld2 were not extracted at salt concentrations lower than 0.5 M KCl (Fig. 3.2A and D). Cells were therefore routinely lysed at 0.5 M KCl for purification of Sld3/7 and Sld2. Dpb11 was soluble at salt concentrations as low 0.2 M KCl (Fig. 3.2B), and Cdc45 was soluble under all salt concentrations trialled (Fig. 3.2C).

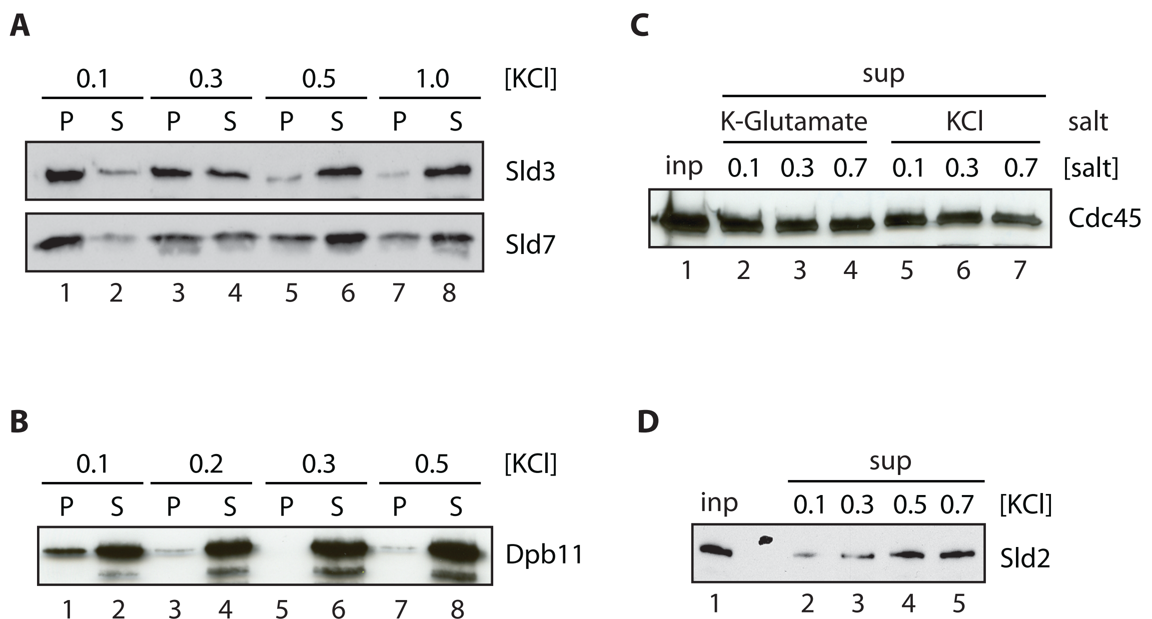


Figure 3.2 Solubility of firing factor proteins in cell lysates

Immunoblot analysis of Sld3-TCP, Sld7 (both **A**), Dpb11-TCP (**B**), Cdc45-TCP (**C**) and Sld2-TCP (**D**) under different salt concentrations. Powdered extract from G1-arrested yeast cells overexpressing the respective proteins was thawed and resuspended, and the salt concentration adjusted. Lysates were centrifuged at 42 krpm for 1 h in a Ti45 rotor, and samples taken of the pellet (P) and supernatant (S) fractions. In C and D, an input sample (inp) was taken before centrifugation, and subsequently compared with the supernatant (sup) samples.

3.2.3 Purification of Dpb11, Sld2 and Cdc45

For purification, Dpb11, Sld2 and Cdc45 were tagged at their C-termini with either a TCP (Dpb11, Cdc45) or 3xFLAG tag (Sld2, Cdc45). The TCP tag contains both calmodulin binding peptide (CBP) and protein A, as well as a cleavage site for TEV

protease. The position of the TEV cleavage sequence between the C-terminus of the tagged protein and the CBP section of the tag allows for complete removal of the tag during purification.

For Dpb11, initial attempts to purify the protein via calmodulin affinity chromatography proved unsuccessful (data not shown), perhaps indicating that the CBP section of the TCP tag was occluded by another part of the protein. I therefore purified Dpb11-TCP using an IgG sepharose column (Fig. 3.3A), which binds the protein A epitope in the TCP tag. Bound Dpb11-TCP was eluted from the IgG column by cleavage with TEV protease, allowing highly purified untagged Dpb11 to be recovered. Dpb11 was subsequently subjected to size exclusion chromatography (Fig. 3.3B), eluting from a gel filtration column as a single well-defined peak close to the 158 kDa molecular weight marker, higher than the predicted molecular weight of Dpb11 (87 kDa).

Sld2 was purified via anti-FLAG immunoprecipitation (Fig. 3.3C), and a relatively pure sample was obtained after this first affinity purification step. The Sld2 purification protocol has subsequently been developed by a colleague in the lab (Dr. Joe Yeeles), and highly pure Sld2 can now be obtained via a further ion exchange chromatography step.

Cdc45 could be purified by either calmodulin affinity chromatography or anti-FLAG immunoprecipitation (Fig. 3.3D and E). Interestingly, when either TCP or 3xFLAG-tagged Cdc45 was subsequently subjected to size exclusion chromatography (Fig. 3.4), the protein did not elute from the gel filtration column as a well-defined peak, with at least a portion of the protein eluting directly after the void volume before the 670 kDa molecular weight marker. This behaviour was refractory to any changes in buffer conditions, which I attempted (Fig. 3.4A and B and data not shown) and is consistent with aggregation of the protein into species of a molecular weight much

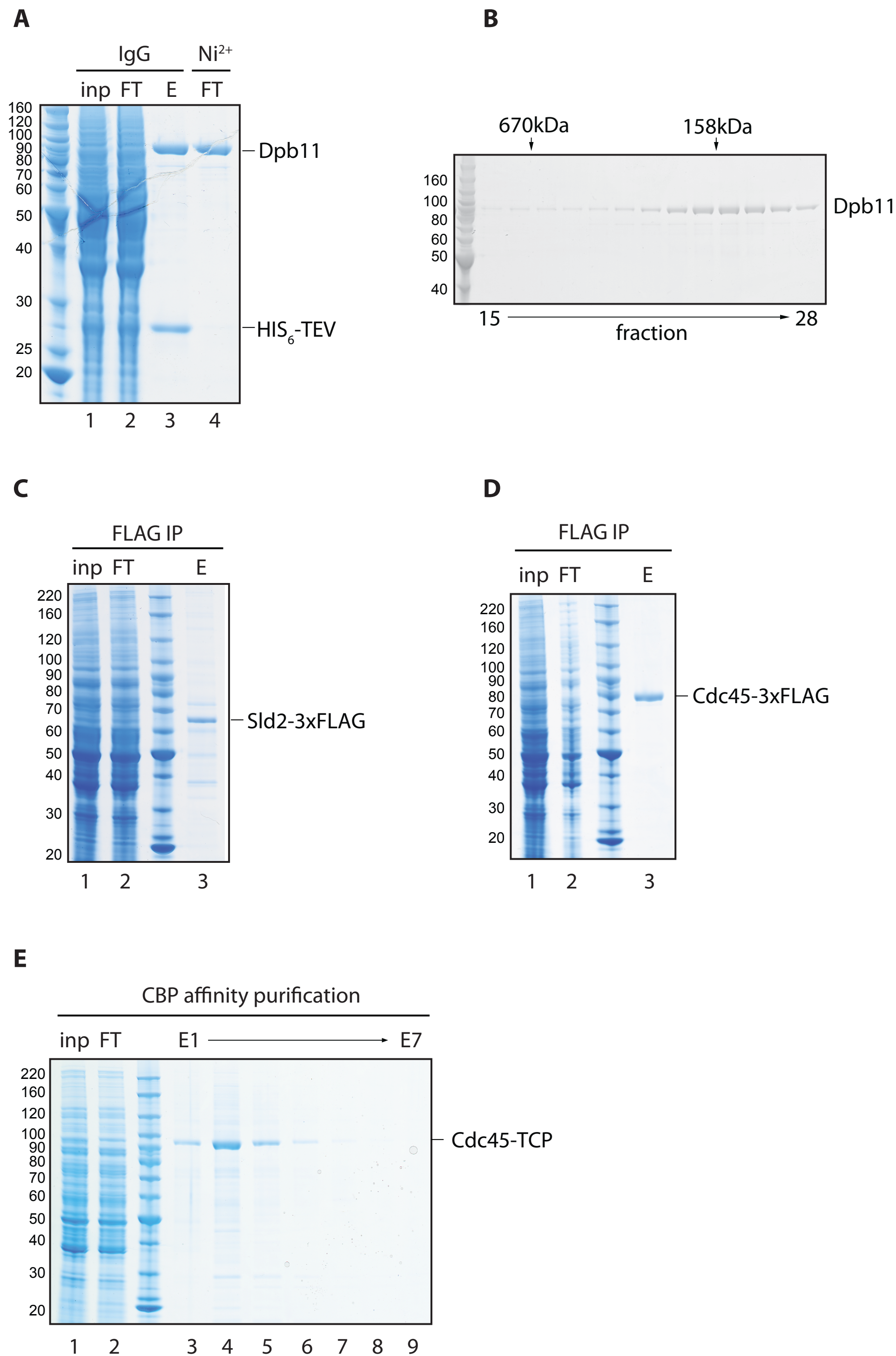


Figure 3.3 Purification of Dpb11, Sld2 and Cdc45

Purified Dpb11 (**A, B**), Sld2-3xFLAG (**C**), Cdc45-3xFLAG (**D**) and Cdc45-TCP (**E**) analysed by SDS-PAGE and coomassie staining. Input (inp), flow-through (FT) and eluate (E) fractions were collected for each purification. **A**. Dpb11 was purified using IgG and Ni-NTA beads, and **B**. gel filtration. Elution positions of molecular weight standards are shown. **C**. Sld2 was purified using anti-FLAG affinity resin. **D**. Cdc45 was purified using anti-FLAG affinity resin or **E**. calmodulin beads. Molecular weight markers (kDa) are shown to the left of each gel.

higher than the 75 kDa predicted for Cdc45. Subsequently, a colleague in the lab (Dr. Joe Yeeles) observed that the sequences of the N- and C-termini of Cdc45 are highly conserved between different species, and are not predicted to be surface-exposed. Thus, the tagging of the C-terminus of Cdc45 may be detrimental for protein folding and/or function. Indeed, Dr. Yeeles has since been able to purify an active form of Cdc45 through the use of an internal tag.

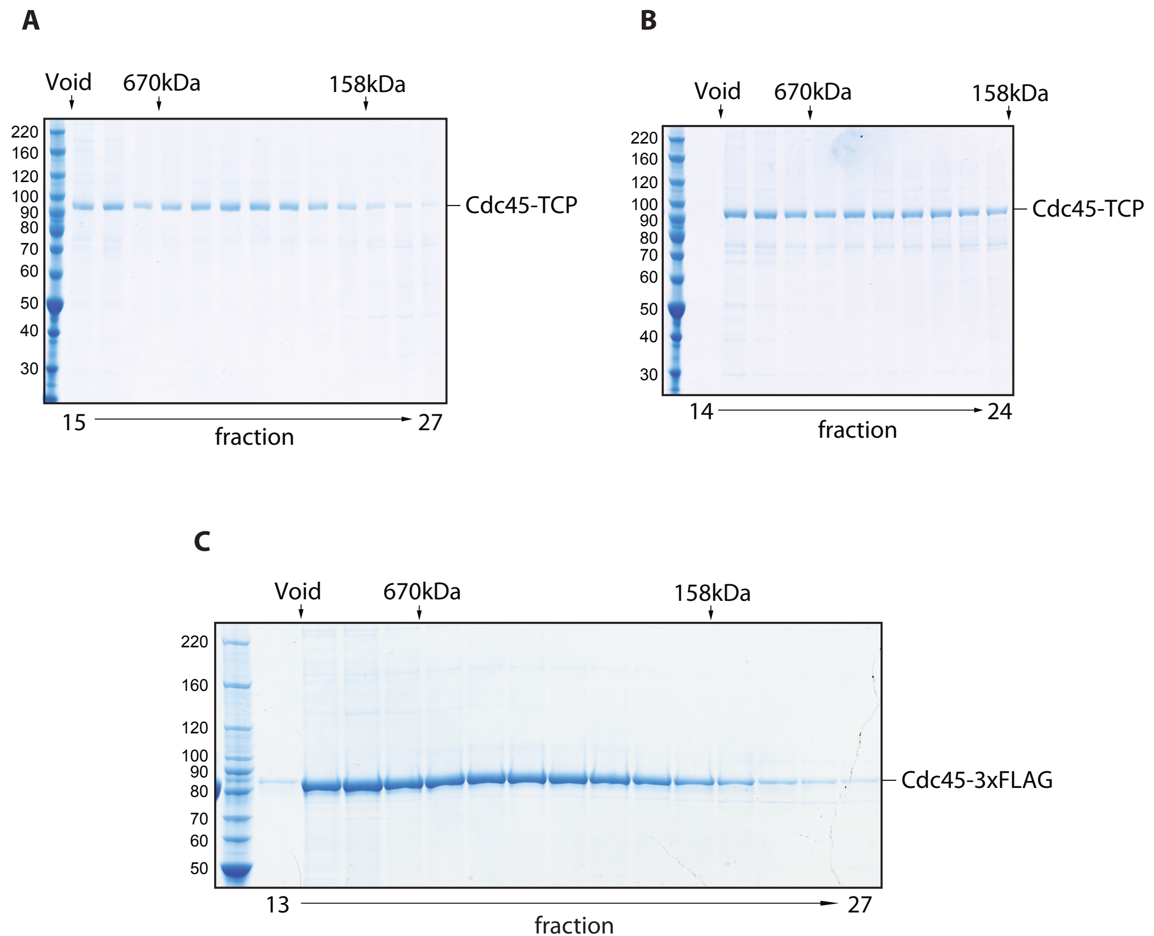


Figure 3.4 Gel filtration analysis of Cdc45

Purified Cdc45-TCP (**A, B**) and Cdc45-3xFLAG (**C**) analysed by SDS-PAGE and coomassie staining. Peak eluate fractions from affinity chromatography (Fig. 3.3D and E) were concentrated and loaded onto a Superdex 200 10/300 gel filtration column. **A**. Gel filtration was performed under 0.5 M KCl. **B.** 0.2 M K-Glutamate was used in place of 0.5 M KCl. **C**. Gel filtration of Cdc45-3xFLAG under the same conditions as in A. Molecular weight markers (kDa) are shown to the left of each gel. Elution positions of molecular weight standards are shown.

3.2.4 Purification of Sld3/7

As Sld3 and Sld7 had previously been reported to form a stable complex (Tanaka et al., 2011b), I set out to purify the Sld3/7 complex from a strain in which both proteins were overexpressed and Sld3 was TCP-tagged at the C-terminus. After IgG affinity purification, Sld3 co-eluted with a second polypeptide, which ran at approximately 29 kDa on an SDS-PAGE gel (Fig. 3.5A), consistent with the predicted molecular weight of Sld7. Sld3 subsequently co-eluted with this 29 kDa protein from a gel filtration column (Fig. 3.5B), and the presence of Sld7 was confirmed by immunoblotting using an antibody raised against Sld7 (Fig. 3.6A). The fact that these proteins co-elute under 0.5 M KCl indicates that they form a very stable complex, even at high salt concentrations.

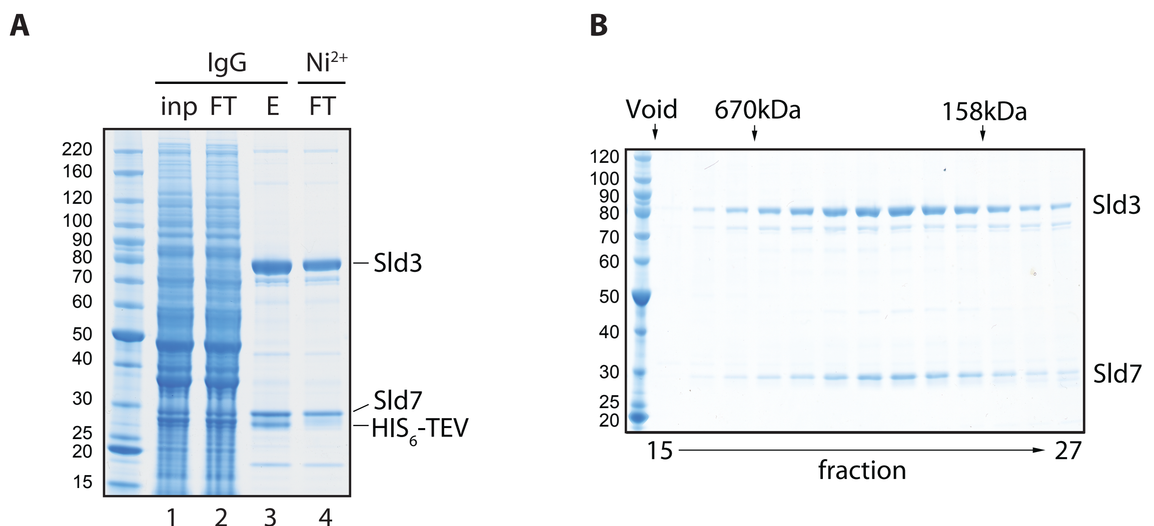


Figure 3.5 Purification of Sld3/7

A. Sld3/7 was purified using IgG and Ni-NTA (Ni^{2+}) beads. Input (inp), flow-through (FT) and eluate (E) fractions were collected. **B.** Gel filtration of the Sld3/7 complex. The Ni-NTA FT fraction from A was loaded onto a Superdex 200 10/300 gel filtration column. Elution positions of molecular weight standards are shown. Molecular weight markers (kDa) are shown to the left of each gel. Proteins were analysed by SDS-PAGE and coomassie staining.

3.2.5 Characterisation of the Sld3/7 complex

The Sld3/7 complex eluted from a gel filtration column between the positions of the 670 kDa and 158 kDa molecular weight markers (Fig. 3.5B, 3.6A) Comparison with

known standards (Fig. 3.6B) allowed the Stokes radius of the complex to be estimated at 68 Å.

Whilst gel filtration analysis can be used to give a fairly accurate approximation of molecular weight for proteins that have average hydration levels and are roughly spherical, molecular weight can be more precisely estimated using the Siegel and Monty formula (Siegel and Monty, 1966). This formula combines Stokes radius data obtained from size exclusion chromatography with sedimentation coefficients derived from density gradient centrifugation. The Sld3/7 complex migrated between the 44 kDa and 158 kDa markers on a 15-35% glycerol gradient (Fig. 3.6C), and the sedimentation coefficient of the complex was estimated at 5.8 S by comparison with known standards (Fig. 3.6D). I was subsequently able to use the Siegel and Monty formula (Equation 1) to obtain an approximate molecular weight of 163 kDa for the Sld3/7 complex. Given the predicted molecular weights of Sld3 and Sld7 (77 kDa and 29.5 kDa, respectively), it would appear likely that the Sld3/7 complex contains more than one copy of one or both of its subunits.

$$Ma = \frac{6\pi\eta Nas}{1-v\rho}$$

η = viscosity of the medium
a = Stokes radius
s = sedimentation coefficient
v = partial specific volume
ρ = density of the medium
Ma = molecular weight

Equation 1 The Siegel and Monty formula (Siegel and Monty, 1966).

To further investigate the subunit stoichiometry of Sld3/7, I next reconstituted the Sld3/7 complex from its individual components *in vitro*. To do this, I developed bacterial expression systems for the purification of either Sld3 or Sld7 after overexpression in *E. coli*. *SLD3* or *SLD7* were cloned into the pGEX-6p1 vector, which introduced a GST tag at the N-terminus of either protein. Both GST-Sld3 and GST-Sld7 were visible as major bands in the *E. coli* cell lysate after 4 h expression (Fig. 3.7A and 3.7D, lane 2), and both proteins were purified using GST-affinity chromatography (Fig. 3.7A and 3.7D). Sld3 was eluted from the glutathione beads

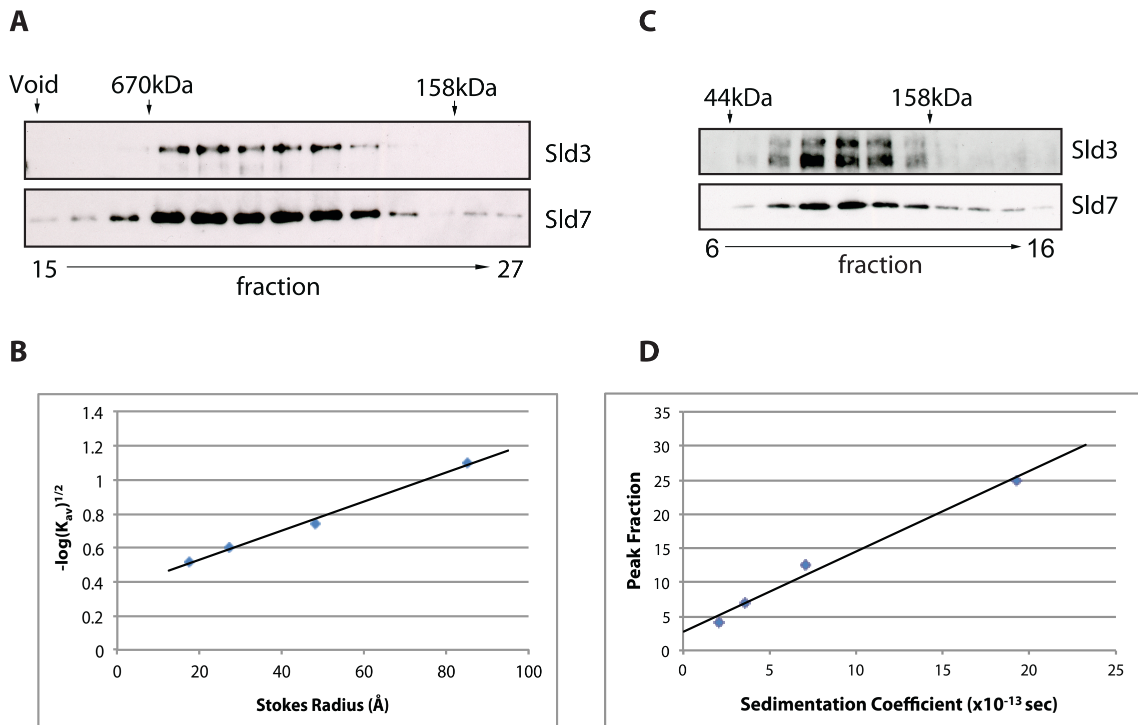


Figure 3.6 Molecular weight estimation of Sld3/7

A. Size exclusion chromatography of Sld3/7. Purified Sld3/7 was loaded onto a Superdex 200 10/300 gel filtration column, and fractions analysed by immunoblot. **B**. Protein standards (myoglobulin, ovalbumin, g-globulin, thyroglobulin) of known Stokes Radius were analysed by size exclusion chromatography as in A, and a standard curve plotted. Kav = (Ve – V0) / (Vt – V0), where Ve is elution volume, V0 is the void volume of the column (8 ml), and Vt is the total volume of the column (24 ml). **C**. Sld3/7 was loaded onto a 3.4 ml 15-35% glycerol gradient, and centrifuged in a SW55 Ti swinging bucket rotor at 42 krpm for 16 h. The gradient was fractionated and samples analysed by immunoblot. Positions of molecular weight standards are shown. **D**. The protein standards used in B were loaded onto a 15-35% glycerol gradient as in C, and their migration on the gradient was used to plot a standard curve.

via preScission-protease mediated removal of the GST tag (Fig. 3.7A, lane 5), and was subsequently purified further by gel filtration (Fig. 3.7B). Sld3 eluted from the gel filtration column as a single peak close to the position of the 670 kDa marker, much bigger than the expected molecular weight of Sld3. Alternatively, GST-Sld3 still containing the GST tag could be obtained via glutathione elution (Fig. 3.7C).

Purified Sld7 could also be obtained by removal of the GST tag (Fig. 3.7D, lane 5) and, as with Sld3, Sld7 was then purified further by gel filtration (Fig. 3.7E).

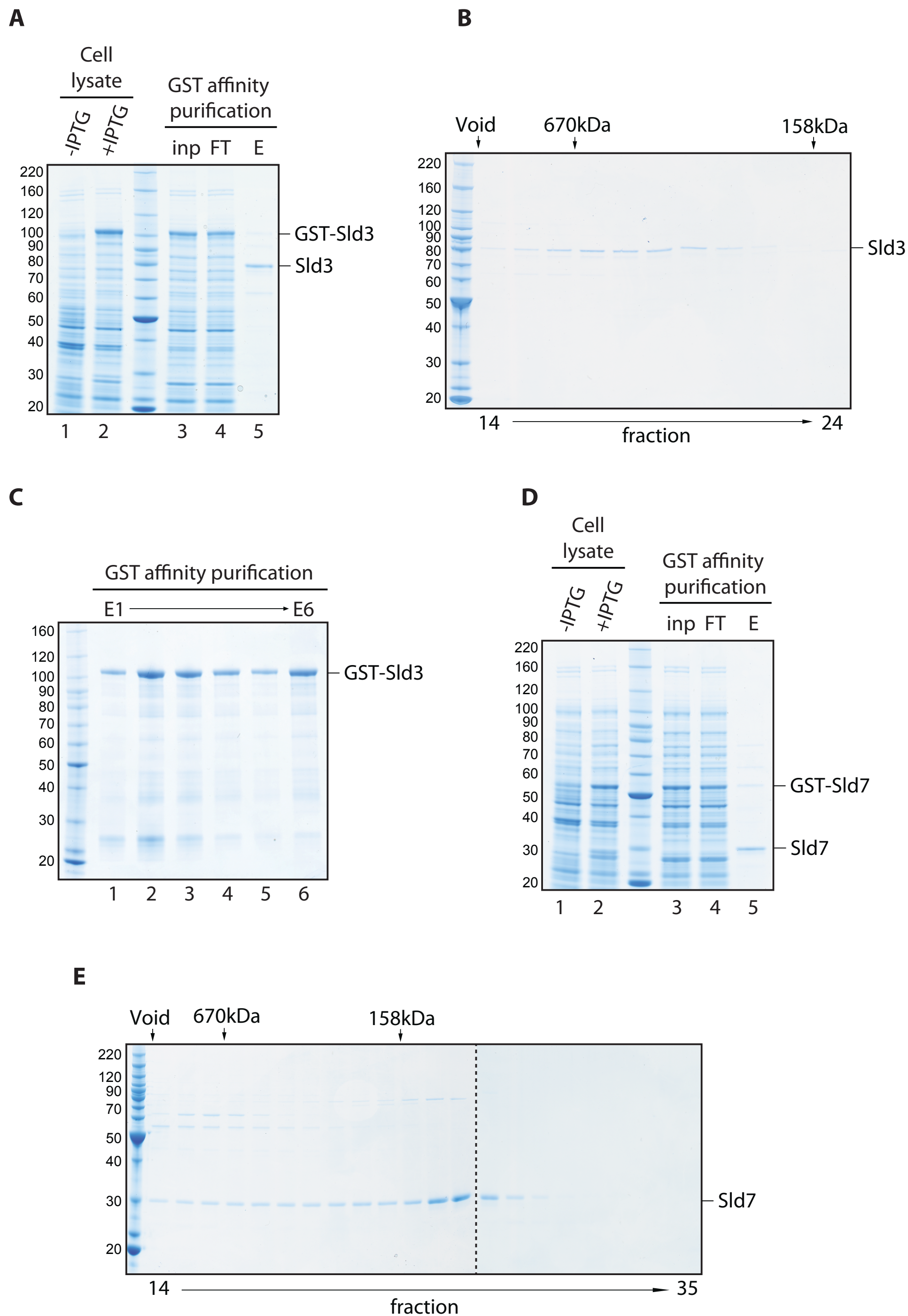


Figure 3.7 Purification of Sld3 and Sld7 after expression in *E. coli*

A. Sld3 was purified using glutathione beads. Cell lysate fractions were taken before and after induction with IPTG, and Input (inp), flow-through (FT) and eluate (E) fractions were collected during affinity purification. **B.** Gel filtration of Sld3. The eluate fraction from A was loaded onto a Superdex 200 10/300 gel filtration column. Elution positions of molecular weight standards are shown. **C.** As in A, but elution was performed in buffer containing 20 mM glutathione. **D.** Sld7 was purified using glutathione beads as for Sld3 in A. **E.** Gel filtration of Sld7 performed as for Sld3 in B. Molecular weight markers (kDa) are shown to the left of each gel. Proteins were analysed by SDS-PAGE and coomassie staining.

Sld7 eluted from the gel filtration column across a broad range of fractions, potentially indicating the existence of multiple conformationally distinct populations of Sld7 within the eluate fraction from the initial affinity purification step. Peak fractions 25-28 were pooled and used in subsequent experiments.

To test for the presence of multiple Sld3 molecules within the Sld3/7 complex, I next attempted to reconstitute the Sld3/7 complex using a mixture of GST-Sld3, Sld3-FLAG and Sld7 (for this experiment, a single FLAG tag was introduced at the C-terminus of Sld3 by PCR, and Sld3-FLAG was purified as shown in Fig. 3.7A). Purified GST-Sld3, Sld3-FLAG and Sld7 were incubated together *in vitro*, and proteins bound to Sld3-FLAG were then isolated by FLAG immunoprecipitation. Notably, both GST-Sld3 and Sld7 were enriched in the FLAG IP sample in an Sld3-FLAG-dependent manner (Fig. 3.8, compare lanes 7 and 8). The interaction of Sld3-FLAG with GST-Sld3 was independent of Sld7 (Fig. 3.8, compare lanes 7 and 9), and the interaction of Sld3-FLAG with Sld7 was independent of GST-Sld3 (Fig. 3.8, compare lanes 7 and 10). These data could be consistent with either the formation of separate Sld3-FLAG/Sld7 and Sld3-FLAG/GST-Sld3 complexes during the reaction, or, more likely, the formation of an Sld3/7 complex containing multiple copies of Sld3.

Yeast two-hybrid data from a previous report suggests that the N-terminus of Sld3 is required for Sld7 binding (Tanaka et al., 2011b). To test the requirement for the N-terminus of Sld3 for the interaction with Sld7 *in vitro*, I purified two Sld3 fragments lacking either the first 132 or 250 residues of the protein, and included these fragments in the experiment in place of full-length Sld3 (Fig. 3.8, lanes 5 and 6). Removal of the N-terminal 250 amino acids of Sld3 resulted in less GST-Sld3 and Sld7 being bound to the Sld3-FLAG beads at the end of the reaction (Fig. 3.8, lane 12), consistent with this section of Sld3 being required for both Sld7 binding and Sld3 multimerisation. The small amount of Sld7 observed in lane 12 may be bound via the N-terminus of the full-length GST-Sld3 present in the reaction.

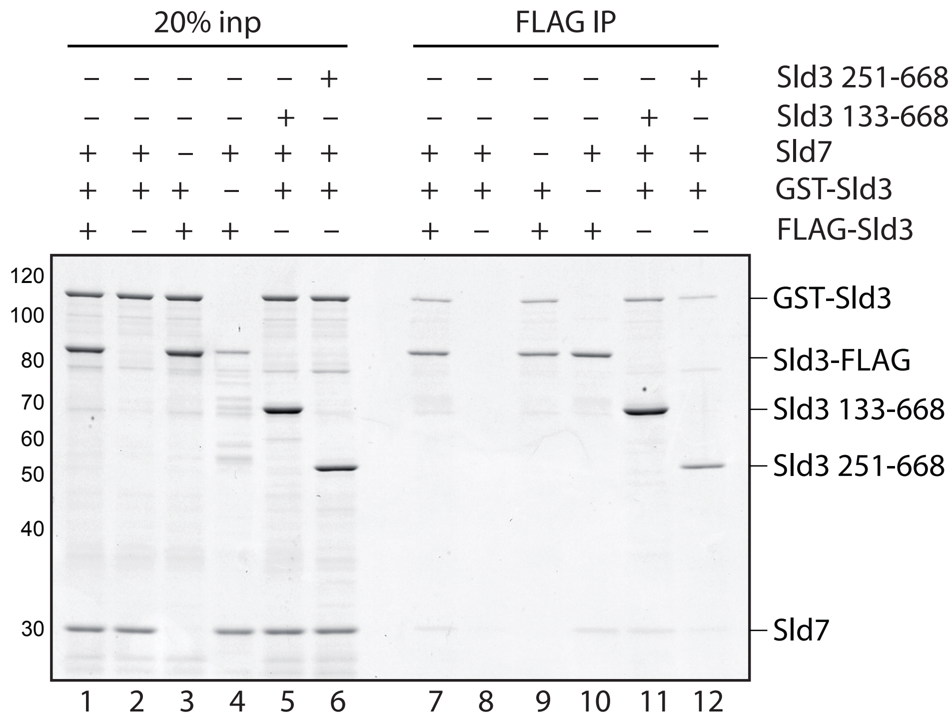


Figure 3.8 *In vitro* reconstitution of the Sld3/7 complex with purified proteins

Purified full-length or truncated Sld3-FLAG, GST-Sld3 and Sld7 were mixed together as indicated and isolated via FLAG immunoprecipitation. Bead-bound proteins were eluted by incubation with buffer containing 3xFLAG peptide, and analysed alongside input (inp) samples by SDS-PAGE and coomassie staining. Positions of proteins are indicated in the right. Molecular weight markers are indicated to the left.

3.2.6 Purification of pre-RC proteins

A prerequisite for the reconstitution of replication initiation steps downstream of pre-RC assembly was to first be able to reconstitute the Mcm2-7 loading reaction *in vitro*, as had been demonstrated previously (Remus et al., 2009, Evrin et al., 2009). Given this, I purified the ORC, Cdc6 and Mcm2-7/Cdt1 licensing proteins to homogeneity (Fig. 3.9A, B and C). The Mcm2-7/Cdt1 heptamer and ORC hexamer were both purified from yeast cells arrested in G1 with the mating pheromone α-factor. Cdc6 was purified following overexpression in *E. coli*.

To test the activity of the purified licensing proteins, *in vitro* pre-RC assembly reactions were performed, as described previously (Remus et al., 2009, Evrin et al., 2009). ORC, Cdc6 and Mcm2-7/Cdt1 were incubated with an origin-containing linear DNA fragment coupled to magnetic beads.

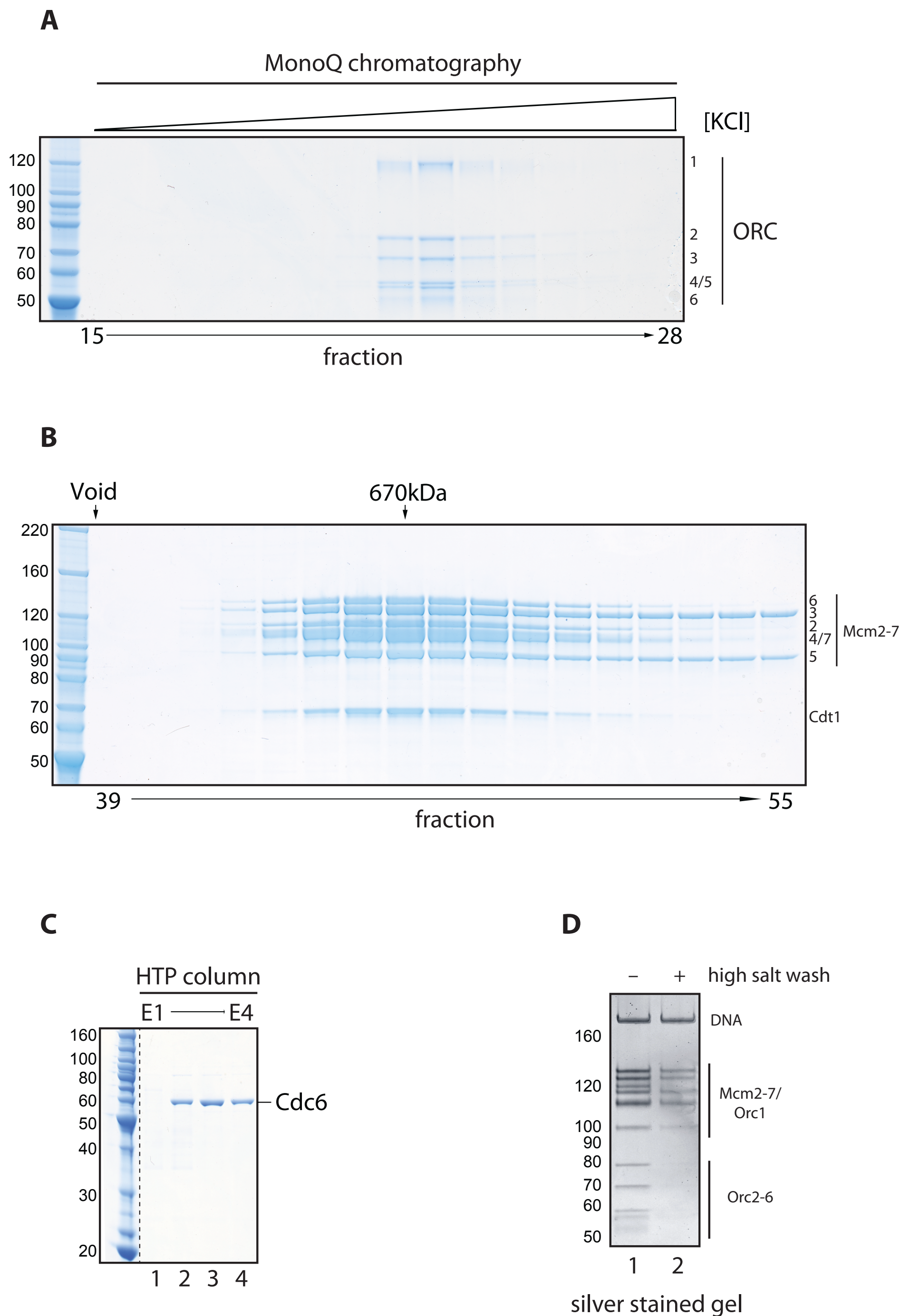


Figure 3.9 Purification of pre-RC proteins

A. Purified ORC analysed by SDS-PAGE and coomassie staining. The final ion exchange chromatography step of the purification is shown. **B**. Purified Mcm2-7/Cdt1 analysed by SDS-PAGE and coomassie staining. The final gel filtration step of the purification is shown. **C.** Purified Cdc6 from *E. coli* analysed by SDS-PAGE and coomassie staining. Elution fractions from the final hydroxyapatite (HTP) column are shown. **D.** *In vitro* pre-RC Mcm2-7 loading assay. Proteins bound to DNA were visualised by silver staining. All reactions were performed in the presence of ATP, and washed with either low salt wash (lane 1) or high salt wash (lane 2) as indicated.

In the presence of ATP, Mcm2-7 is loaded around double-stranded DNA. This loaded form of the Mcm2-7 complex is topologically linked to the DNA and is thus resistant to challenge with a high salt wash, which is sufficient to remove ORC, Cdc6 and Cdt1 from the DNA (Fig. 3.9D). Although not shown, in the presence of the slowly-hydrolysable ATP analogue ATPγS, Mcm2-7 is ‘recruited’ to DNA in a form that is not resistant to high salt washing.

3.2.7 Purification of DDK

The activity of DDK is absolutely required for replication initiation *in vivo* and *in vitro* (Bousset and Diffley, 1998, Heller et al., 2011, Hartwell et al., 1974, Hereford and Hartwell, 1974, Patterson et al., 1986, Buck et al., 1991), and evidence suggests that phosphorylation of the loaded Mcm2-7 complex by DDK is the most upstream event after pre-RC assembly during Mcm2-7 activation (Heller et al., 2011, Yabuuchi et al., 2006, Jares and Blow, 2000, On et al., 2014, Gros et al., 2014). I therefore purified DDK from a yeast strain, in which Cdc7 and Dbf4 are both overexpressed. The DDK recovered after the initial affinity chromatography step (Fig. 3.10A) was phosphorylated on both subunits, as shown by the mobility shift in an SDS-PAGE gel upon treatment with lambda phosphatase (Fig. 3.10B). This hyper-phosphorylation has recently been shown to decrease the kinase activity of DDK (On et al., 2014), and the phosphatase treatment was therefore performed as a standard during DDK purifications.

As the loaded form of the Mcm2-7 complex has previously been shown to be the preferred substrate for DDK *in vitro* (Francis et al., 2009), I used Mcm2-7 pre-assembled into the pre-RC as a substrate to test the kinase activity of purified DDK (Fig. 3.10D). Mcm6 was phosphorylated as DDK was titrated into the reaction, resulting in the formation of a slower migrating form of the protein on an SDS-PAGE gel (Fig 3.10D, lanes 5 and 6). This indicates that purified DDK has protein kinase activity *in vitro*.

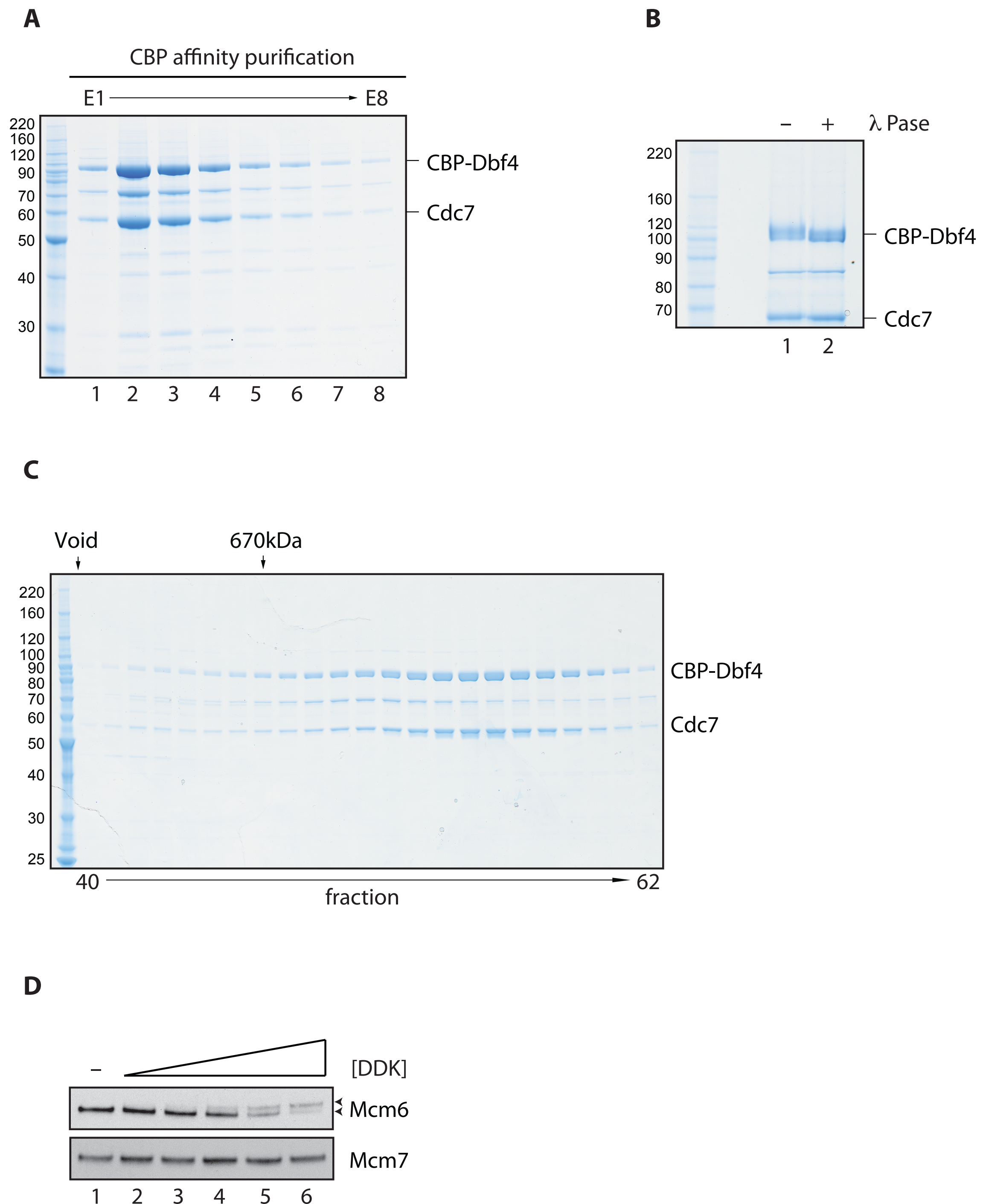


Figure 3.10 Purification of DDK

Purified DDK analysed by SDS-PAGE and coomassie staining (**A, B, C**). DDK was subjected to sequential calmodulin affinity chromatography (**A**), lambda phosphatase treatment (**B**) and gel filtration (**C**) steps. Molecular weight markers (kDa) are shown to the left of each gel. **D**. Purified DDK was tested for *in vitro* kinase activity using loaded Mcm2-7 as a substrate. DDK was added to the final concentrations of 6.25, 12.5, 25, 50 and 100 nM, and samples were then analysed by SDS-PAGE and immunoblot.

3.2.8 Reconstitution of Sld3/7 recruitment to the pre-RC *in vitro*

Data from both *in vivo* and *in vitro* studies suggests that recruitment of the Sld3/7 complex to the pre-RC is required for Cdc45 recruitment (Heller et al., 2011, Kanemaki and Labib, 2006, Tanaka et al., 2011a, Kamimura et al., 2001), and is upstream of the CDK-dependent recruitment of Dpb11, Sld2, GINS and Polε in the form of the pre-loading complex (Muramatsu et al., 2010, Heller et al., 2011, Yabuuchi et al., 2006). Notably, this recruitment of Sld3/7 has been shown to be somehow dependent on the activity of DDK (Heller et al., 2011, Tanaka et al., 2011a, Yabuuchi et al., 2006). To test the hypothesis that this DDK-dependence was due to the direct regulation of an interaction between Sld3/7 and a pre-RC component, I attempted to reconstitute binding of Sld3/7 to pre-RCs assembled *in vitro*.

Pre-RCs were first assembled onto linear DNA coupled to beads *in vitro*, and then phosphorylated with DDK. Following the removal of the DDK reaction mixture, purified Sld3/7 was added, as is described in Fig. 3.11A. Recruitment of proteins to the DNA was monitored by immunoblotting or silver staining. In the complete reaction, when all pre-RC components and DDK were present (Fig. 3.2B, lane 1 and Fig. 3.2C, lane 3), the Mcm2-7 complex was phosphorylated, as indicated by the formation of a slower migrating form of Mcm4 in an SDS-PAGE gel, and Sld3 and Sld7 were both recruited to the DNA. Strikingly, if DDK was omitted from the reaction (Fig. 3.2B, lane 2 and Fig. 3.2C, lane 2), Sld3/7 recruitment was reduced to background levels, and this was also observed upon omission of ORC, Cdc6 or Mcm2-7/Cdt1 (Fig. 3.2B, lanes 3-5). The DDK-independent binding of Sld3/7 to the pre-RC was increased in the presence of lower salt concentrations (Fig. 3.2E, compare lanes 1 and 3), and the Sld3/7 recruitment step was therefore performed under 0.5 M salt as a standard. These data suggests that recruitment of Sld3/7 to the pre-RC is directly regulated by DDK phosphorylation of one or more pre-RC component.

Sld3/7 was not recruited when Mcm2-7/Cdt1 was omitted from the reaction (Fig. 3.2B, lane 5) despite the presence of both ORC and Cdc6 on the DNA. This is consistent with an interaction between Sld3/7 and the Mcm2-7 complex being

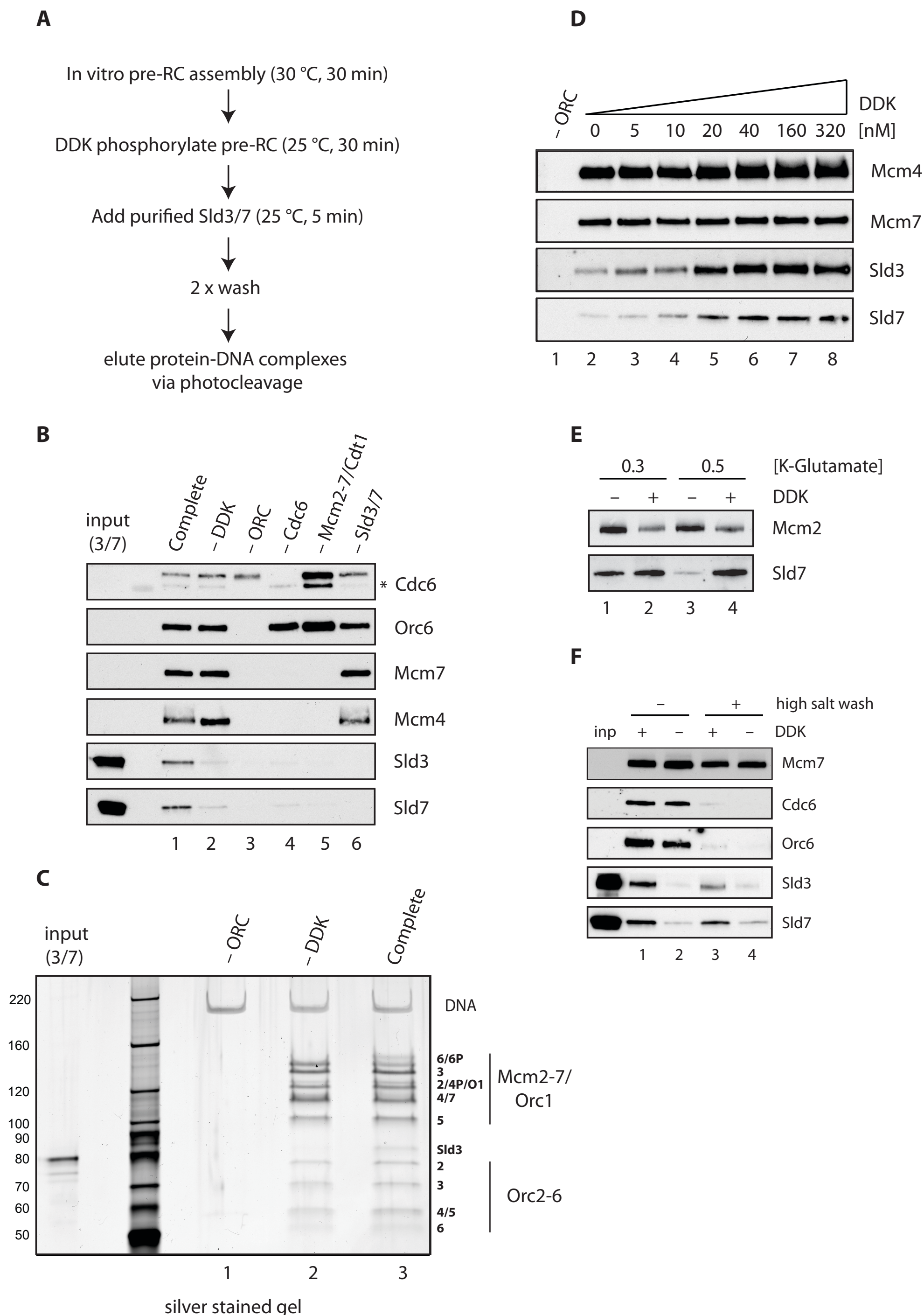


Figure 3.11 Reconstitution of Sld3/7 recruitment to the pre-RC *in vitro*

A. Schematic of experimental setup for Sld3/7 recruitment reaction. DNA-bound samples were analysed by either immunoblot (**B**) or silver staining (**C**) with the indicated proteins omitted. The lower band (*) on the anti-Cdc6 immunoblot in B is Orc6. Visible bands are annotated on the right in C. ‘P’ indicates the position of proteins after phosphorylation **D**. Immunoblot analysis of Sld3/7 recruitment reactions with increasing amounts of DDK. **E**. Immunoblot analysis of Sld3/7 recruitment reactions performed at either 0.3 M or 0.5 M salt. **F**. Reactions were performed as described in A, with a high salt wash step included before addition of Sld3/7 as indicated. Samples were analysed by immunoblot.

required for Sld3/7 recruitment. To test this premise, I included a high salt wash before adding Sld3/7 in the recruitment reaction, which results in the removal of ORC and Cdc6 from the DNA, leaving behind only loaded Mcm2-7 complexes. The loaded Mcm2-7 complexes alone were sufficient for DDK-dependent recruitment of Sld3/7 to the pre-RC (Fig. 3.2F, lane 3), indicating that ORC and Cdc6 are not required for this interaction.

3.3 Discussion

3.3.1 Towards reconstitution of Mcm2-7 activation *in vitro*

The early work presented in this chapter described the purification of numerous firing factors known to be required for the activation of the Mcm2-7 replicative helicase. In addition to the Sld3/7 complex (Fig. 3.5), which I went on to characterise extensively (Fig. 3.6 and 3.8), I also established purification strategies for Dpb11, Cdc45 and Sld2 (Fig. 3.3), all based upon galactose-inducible overexpression in yeast (Fig. 3.1). These strategies have subsequently been developed by colleagues within the lab. The purification of these proteins represents significant progress towards the reconstitution of DNA replication initiation *in vitro*.

With the aim of understanding how the Mcm2-7 complex is activated, I also purified the Cdc6, ORC and Cdt1/Mcm2-7 licensing factors (Fig. 3.9), and showed that these purified proteins were competent for pre-RC assembly *in vitro* (Fig. 3.9D), as had been previously described (Remus et al., 2009, Evrin et al., 2009). Additionally, I was also able to reconstitute quantitative phosphorylation of the loaded Mcm2-7 complex with purified DDK (Fig. 3.10D). These DDK-phosphorylated Mcm2-7 complexes can support the replication of a plasmid template in an S-phase extract (On et al., 2014, Gros et al., 2014), and will therefore likely represent the basis for future reconstitution efforts.

3.3.2 Sld3/7 contains multiple copies of Sld3

In section 3.2.4 of this chapter, I described the purification of the Sld3/7 complex following overexpression in yeast. This complex remained stable during gel filtration under 0.5 M KCl (Fig. 3.5B), consistent with previous data indicating that Sld3 and Sld7 are stably bound to one another throughout the cell cycle (Tanaka et al., 2011b). Through a combination of gel filtration and density gradient centrifugation (Fig. 3.6), I was able to estimate the molecular weight of the complex at 163 kDa, indicative of the presence of more than one copy of Sld3 and/or Sld7 in the complex. The estimated Stokes Radius of Sld3/7 was 68 Å; this relatively large Stokes radius for a complex of this size suggests that Sld3/7 adopts an extended conformation.

Subsequently, I was able to reconstitute the Sld3/7 complex *in vitro* from individually purified subunits, and showed that a complex can be formed containing more than one copy of Sld3 (Fig. 3.8, lane 7). Interestingly, the abundance of the second copy of Sld3 in the complex was reduced upon deletion of residues 1-250 of Sld3 (Fig. 3.8, compare lanes 7 and 12). This data is suggestive of a novel role for this section of the protein, previously shown to be required for Sld7 binding (Tanaka et al., 2011b), in Sld3 multimerisation. The Sld7 binding activity previously ascribed to this section of Sld3 was shown to reside within residues 91-121 (Tanaka et al., 2011b). The precise location of the Sld3 multimerisation function is still to be determined, but the fact that both Sld7 and a second copy of Sld3 could be simultaneously bound to Sld3-beads (Fig. 3.8, lane 7) could suggest that the Sld3 and Sld7 binding sites in Sld3 do not overlap. Whether or not the complex contains more than one copy of Sld7 is yet to be examined; determination of the Sld3 binding site on Sld7 would allow for a clearer understanding of the exact architecture of the Sld3/7 complex.

One notable limitation of this analysis is the relatively broad peaks observed for Sld3/7 during gel filtration (Fig. 3.6B), making determination of an exact elution volume very difficult. One explanation for this behaviour may be the existence of multiple different multimeric forms of Sld3/7 in the purified sample. Alternative means of molecular weight estimation, such as analytical ultracentrifugation or

multi-angle light scattering, could potentially be used to obtain a more accurate estimate of the size of the complex.

In spite of these limitations, the notion that Sld3/7 likely contains two Sld3 molecules is intriguing. During DNA replication initiation, an Mcm2-7 double hexamer must be converted into a pair of divergent bi-directional replication forks, each containing only a single Mcm2-7 hexamer in the form of the CMG complex (Gambus et al., 2006, Yardimci et al., 2010, Costa et al., 2011, Boos et al., 2012, Raschle et al., 2008). Sld3 recruitment to the pre-RC has been previously shown to be upstream of the recruitment of all other known firing factors (Heller et al., 2011, Yabuuchi et al., 2006). The presence of two Sld3 molecules (in the form of a single Sld3/7 complex) bound to the pre-RC would allow for the CDK-dependent recruitment of two copies of Dpb11, and potentially, two copies of GINS. Thus, the stoichiometry of the Sld3/7 complex could conceivably determine the number of molecules of all other firing factors that are subsequently recruited, leading to the downstream formation of two CMG complexes from a single Mcm2-7 double hexamer (discussed further in section 6.6).

3.3.3 DDK directly regulates Sld3/7 binding to the Mcm2-7 double hexamer

In the final part of this chapter, I described a reconstituted system for studying the recruitment of the Sld3/7 complex to the pre-RC *in vitro* (Fig. 3.11). In addition to the requirement for ORC, Cdc6 and Mcm2-7 (Fig. 3.11B, lanes 3-5), the recruitment of Sld3/7 in this system was dependent on the presence of DDK (Fig. 3.11B, compare lanes 1 and 2). Furthermore, I was able to show that the loaded Mcm2-7 complex is both necessary and sufficient for Sld3/7 recruitment (Fig. 3.11B, lane 5 and 3.11F, lane 3). Although consistent with previous data suggesting that DDK is required for Sld3/7 recruitment to replication origins *in vivo* (Heller et al., 2011, Tanaka et al., 2011a), this is the first evidence of a direct role for DDK in regulating a protein-protein interaction between Sld3/7 and the pre-RC.

Although the Mcm2, 4 and 6 subunits of Mcm2-7 are themselves phosphorylation targets for DDK (Masai et al., 2006, Montagnoli et al., 2006, Sheu and Stillman,

2006, Randell et al., 2010), the essential function of this kinase in DNA replication can be bypassed by alleles in a number of genes that have not been shown to be DDK substrates, namely *mcm5-bob1* (Hardy et al., 1997) and, in *S. pombe*, by deletion of *mrc1* or *rif1* (Hayano et al., 2012, Matsumoto et al., 2011). Based upon these observations, it has previously been proposed that DDK might promote origin firing by eliciting some conformational change in the Mcm2-7 double hexamer, which is required for its activation. The binding of Sld3/7 to the Mcm2-7 complex described in this chapter is the first example of a DDK-regulated interaction involving proteins required for DNA replication initiation, and is suggestive of a novel function for DDK during activation of the replicative helicase. Investigating the mechanism of Sld3/7 recruitment, and examining how phosphorylation of the Mcm2-7 complex by DDK promotes this event, will provide the basis of the rest of this thesis.

Chapter 4. Results 2

4.1 Introduction

The requirement for Sld3 during DNA replication initiation was first described over 10 years ago (Kamimura et al., 2001). Originally, the *SLD3* gene was isolated in a screen for factors exhibiting synthetic lethality with the *dpb11-1* allele (Kamimura et al., 1998), and was subsequently shown to be required for the chromatin association of Cdc45 in S-phase (Kamimura et al., 2001, Kanemaki and Labib, 2006). Since then, it has become increasingly clear that Sld3 has a central role in the regulation of replication origin firing.

In 2007, two papers showed that, in yeast, Sld3 is one of two critical CDK substrates required for the G1/S transition (Tanaka et al., 2007, Zegerman and Diffley, 2007). Phosphorylation of Sld3 and Sld2 generates binding sites for another protein, Dpb11, via its tandem BRCT repeats, and formation of the Sld3-Dpb11-Sld2 complex is required for activation of the loaded Mcm2-7 helicase. This function of Sld3 has since been shown to be conserved with its human homologue, Treslin (Kumagai et al., 2010, Kumagai et al., 2011, Boos et al., 2011).

In addition to its function as an essential CDK substrate, Sld3 is also a substrate of the Rad53 DNA damage checkpoint kinase (Lopez-Mosqueda et al., 2010, Zegerman and Diffley, 2010). Phosphorylation of Sld3 by Rad53 in S-phase acts to block Sld3 binding to both Dpb11 and Cdc45, thereby preventing further origin firing in the presence of DNA damage. Most recently, Sld3 has been shown to form a constitutive complex throughout the cell cycle with a second protein involved in replication initiation, Sld7 (Tanaka et al., 2011b).

In the previous chapter, I presented evidence that the Sld3/7 complex likely contains multiple copies of Sld3. I also described a reconstituted system for studying the recruitment of Sld3/7 to the pre-RC *in vitro*, and showed data supporting the notion that DDK directly regulates binding of Sld3/7 to the loaded Mcm2-7 complex. In the following section, I will present work showing that Sld3/7 recruitment to the pre-RC is mediated via a distinct domain of Sld3, which also

contains a Cdc45 binding site. By isolating a number of Sld3 mutants specifically defective in pre-RC or Cdc45 binding, I have examined the functional significance of these interactions for replication initiation. This work indicates that Sld3, an essential CDK substrate, has an important function in the DDK-dependent regulation of replication initiation, thereby providing a point of intersection for the activities of DDK and CDK during origin firing.

4.2 Results

4.2.1 The N-terminus of Sld3 is required for Sld7 recruitment to the pre-RC

As a first step towards understanding how the Sld3/7 complex is recruited to the pre-RC during replication initiation, I first examined the individual pre-RC binding activities of both Sld3 and Sld7. To do this, I purified Sld3 and Sld7 after overexpression in *E. coli*, as described in the previous chapter. I then tested these individually purified, recombinant proteins for their ability to recruit to *in vitro* reconstituted pre-RCs in a DDK-dependent fashion, using the system described in 3.2.8.

Whilst Sld3 alone could be recruited to the pre-RC to the same levels as in the Sld3/7 complex purified from yeast (Fig. 4.1B, compare lanes 2 and 4), individually purified Sld7 exhibited no detectable pre-RC binding activity (Fig. 4.1B, lane 6). Importantly, the bacterially expressed Sld3 and Sld7 could be used to make a recombinant Sld3/7 complex *in vitro*, and this complex was competent for DDK-dependent Sld7 recruitment to the pre-RC (Fig. 4.1B, lane 8).

Consistent with previously published data (Tanaka et al., 2011b), I demonstrated the involvement of the N-terminus of Sld3 in Sld7 binding in the previous chapter (Fig. 3.8). To test the requirement for this section of Sld3 for Sld7 recruitment to the pre-RC, I next replaced full-length Sld3 with an Sld3 fragment lacking residues 1-250 in an Sld3/7 recruitment assay (Fig. 4.1C). Whilst this Sld3 fragment could bind to the pre-RC to the same levels as full-length Sld3 (Fig. 4.1C, compare lanes 7 and 9), it was not competent for Sld7 recruitment in this system (Fig. 4.1C, lane 15). This indicates that, whilst dispensable for the binding of Sld3 to the Mcm2-7 double

hexamer, the N-terminus of Sld3 is essential for Sld7 recruitment to the pre-RC, in agreement with previous data (Fig. 3.8, Tanaka et al., 2011b).

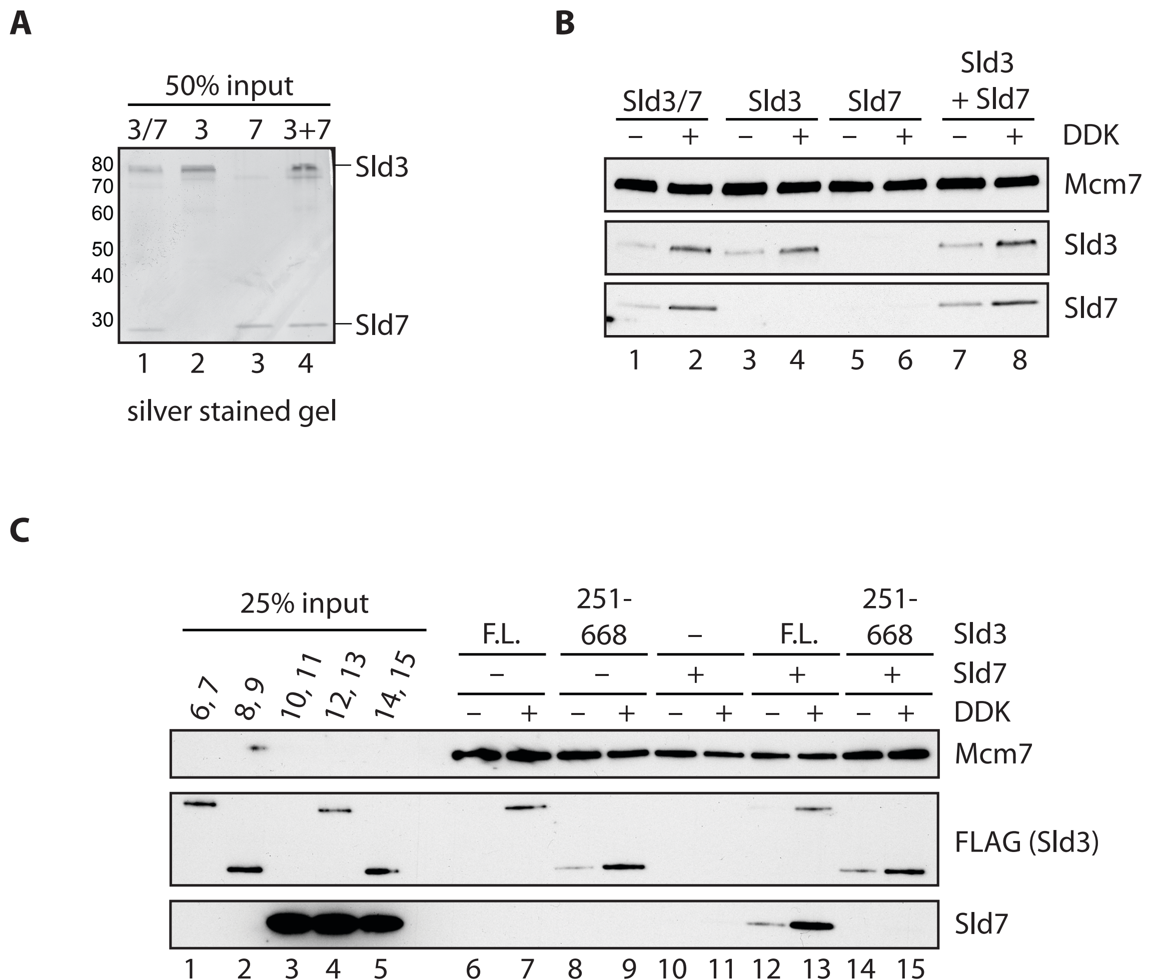


Figure 4.1 The N-terminus of Sld3 is required for Sld7 recruitment to the pre-RC

A. Purified Sld3/7, Sld3 and Sld7 analysed by silver staining. The 3+7 complex in lane 4 was made by mixing of the purified proteins in lanes 2 and 3 under 0.5 M KCl. Each lane contains 50% of the amount of input protein used in B. In **B** and **C**, indicated proteins were tested for their ability to bind to the pre-RC *in vitro*. DDK was omitted as indicated. The presence of full-length (F.L.) or truncated Sld3 is indicated in **C**. DNA-bound proteins were analysed by immunoblot.

4.2.2 Mapping the Mcm2-7 binding site on Sld3

Despite its essential role in the regulation of origin firing, Sld3 has no known enzymatic activity, and there is little functional annotation of the protein's primary sequence. Whilst protein-protein interactions with Sld7, Cdc45 and Dpb11 have

been mapped (Fig. 4.2A), the only described domain within Sld3 is the Sld3-Treslin Domain (S.T.D.), which is the section of the protein that shares sequence homology with the human Sld3 homologue, Treslin (Sanchez-Pulido et al., 2010).

Given the observation that Sld3 alone can be recruited to the pre-RC (Fig. 4.1), I next wanted to map the Mcm2-7 binding site on Sld3. To do this, I made bacterial expression constructs for the expression and purification of various Sld3 fragments. The positions of the N and C-termini or these fragments were chosen based on a secondary structure prediction performed using the Phyre2 server (Kelley and Sternberg, 2009). The fragments were all purified via an N-terminal GST tag (Fig. 4.2B), and the tag was subsequently removed via preScission protease mediated cleavage during elution from glutathione beads.

Following purification, fragments of Sld3 were tested for their ability to bind to DDK-phosphorylated loaded Mcm2-7 complexes *in vitro*. The initial analysis indicated that the first 250 residues (containing the Sld7 binding site) (Fig. 4.2C, lane 10) and last 82 residues (containing the two essential CDK phosphorylation sites) (Fig. 4.2C, lane 16) of Sld3 were dispensable for Mcm2-7 binding. Deletion of residues 1-326 (Fig. 4.2C, lane 12) or 435-668 (Fig. 4.2C, lane 18) produced Sld3 fragments that were not competent for pre-RC binding. Indeed, upon purification of a fragment containing only Sld3 251–585, I could show that this portion of the protein alone was sufficient for recruitment to the pre-RC (Fig. 4.2D, lane 8).

4.2.3 Isolation of Mcm2-7 binding mutants in Sld3

Having mapped a minimal Mcm2-7 interacting portion of Sld3/7 to residues 251-585 of Sld3, I next set out to try and isolate point mutants in Sld3 that were specifically defective in DDK-dependent pre-RC binding, with the aim of using any such mutants to examine the functional importance of this interaction for Mcm2-7 helicase activation.

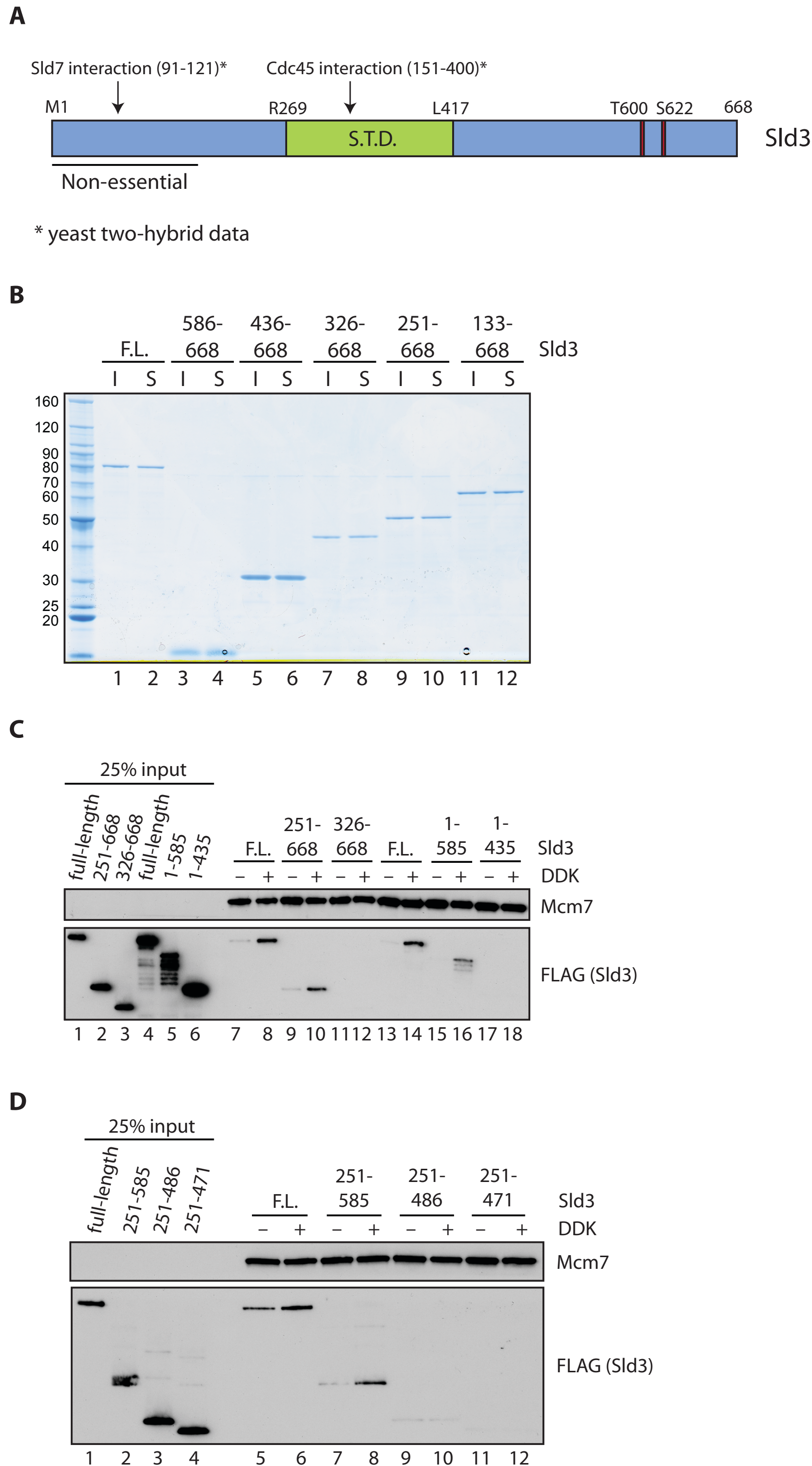


Figure 4.2 Sld3 residues 251-585 are necessary and sufficient for recruitment to the pre-RC

A. Schematic of Sld3 from S. cerevisiae. Residue numbers are shown and positions of known interaction sites for other proteins are depicted. Two essential CDK phosphorylation sites are shown by red vertical bars. The Sld3-Treslin

Domain (S.T.D.) is shown in green. **B.** Representative example of purified Sld3 fragments visualised by SDS-PAGE and coomassie staining. To test solubility, purified proteins were centrifuged at 21,000 g for 5 min, and input (I) and supernatant (S) samples taken before and after centrifugation, respectively. Molecular weight markers (kDa) are indicated to the left. In **C** and **D**, full-length (F.L.) or fragments of Sld3 were tested for their ability to bind to the pre-RC *in vitro*. DDK was omitted as indicated. Input and DNA-bound proteins were analysed by immunoblot.

Given the phosphorylation-dependence of the Sld3-Mcm2-7 interaction, we initially targetted conserved positively charged amino acids for mutagenesis, hypothesising that such residues may be involved in co-ordinating phosphorylated Ser or Thr on the Mcm2-7 complex. The observation that Sld3 251-585 can bind to the pre-RC *in vitro* but Sld3 251-486 cannot (Fig 4.2D, lanes 8 and 10) may suggest that the Mcm2-7 binding activity resides within residues 486-585 of Sld3. Given this, I initially focussed on this region of Sld3 as a target for mutagenesis, generating seven different amino acid substitution mutants in which Arg or Lys were replaced with Glu residues in full-length Sld3 (Fig. 4.3A). These mutant proteins were purified from *E. coli* alongside wild type Sld3 as previously described (Fig. 4.3B), and then tested for their ability to bind to loaded Mcm2-7 (Fig. 4.4A). Strikingly, mutation of 3 distinct pairs of Lys/Arg within residues 511-531 (K511/R512, K518/R520 or K530/R531) rendered Sld3 completely defective for DDK-dependent binding to the pre-RC (Fig. 4.4A, lanes 4, 6, 8).

As a control for the specificity of the binding defect observed, I next tested the Sld3 mutants for other protein-protein interactions, which have been previously reported for Sld3 (Fig. 4.2A). To do this, I incubated the purified Sld3 mutants in an S-phase protein extract, which has previously been used in the development of an *in vitro* DNA replication assay (see 4.2.5). Recombinant Sld3 was then re-isolated from the extract via immunoprecipitation of the FLAG tag on its N-terminus. Upon addition of wild type Sld3 to the S-phase extract, I observed a specific enrichment of both Cdc45 and Sld7 on the anti-FLAG beads at the end of the reaction (Fig. 4.4B, lane 11). The reported CDK-dependent interaction between Sld3 and Dpb11 was not observed in this system. Reassuringly, all of the mutants tested were able to bind Cdc45 and Sld7 to the same level as wild type Sld3 (Fig. 4.4B lanes 12-18),

indicating that mutations in residues 511-531 do not affect the Cdc45 or Sld7 binding activities of Sld3.

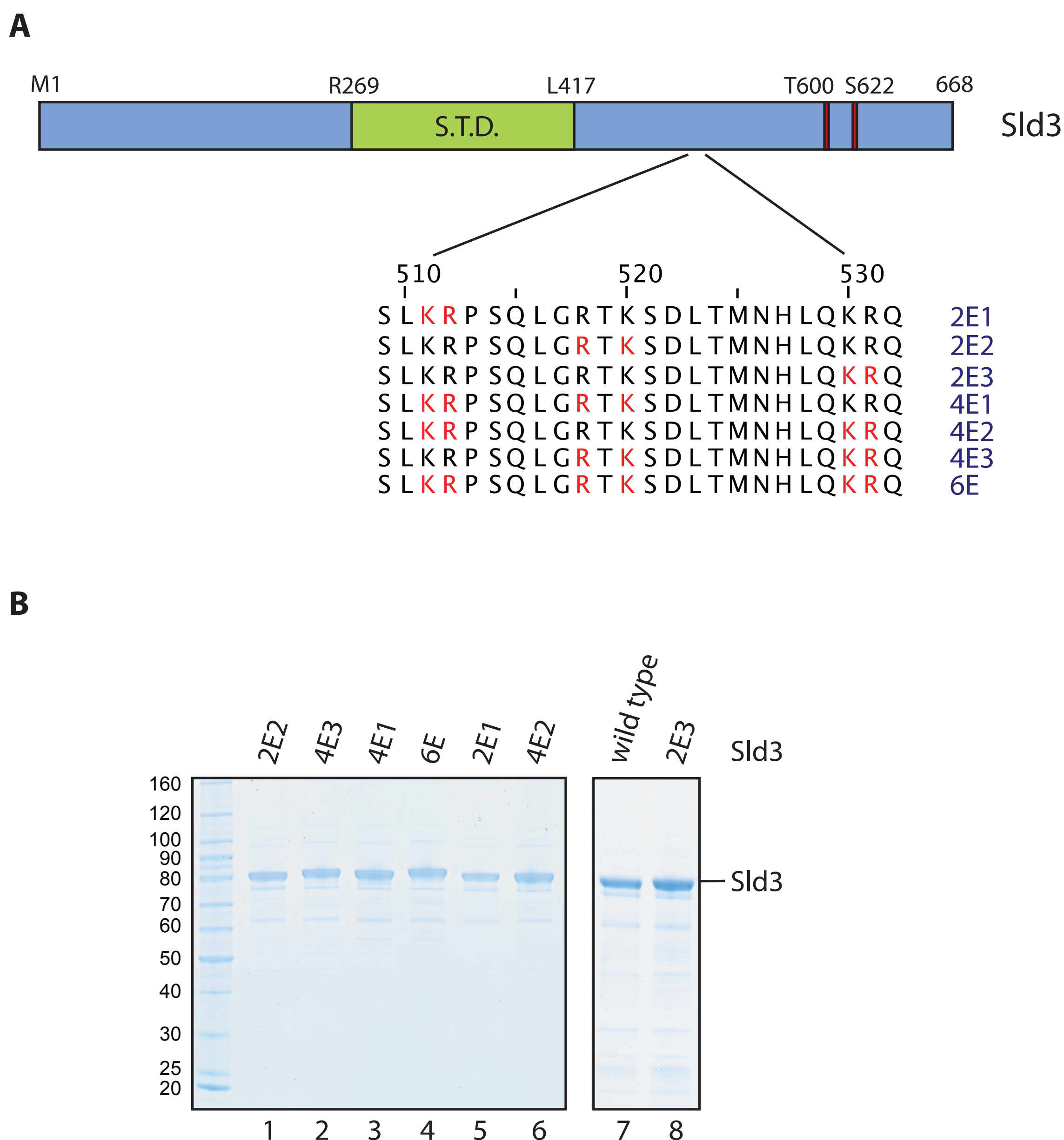


Figure 4.3 Purification of Sld3 Mcm2-7 binding mutants

A. Schematic showing position of amino acid substitution mutations in Sld3. Sequence of Sld3 509-532 is shown. Residues highlighted in red were replaced with Glu. Names of mutants are shown to the right. **B**. Purified Sld3 mutants visualised by SDS-PAGE and coomassie staining. Mutant and wild type Sld3 were purified after overexpression in *E. coli*. Molecular weight markers (kDa) are shown to the left.

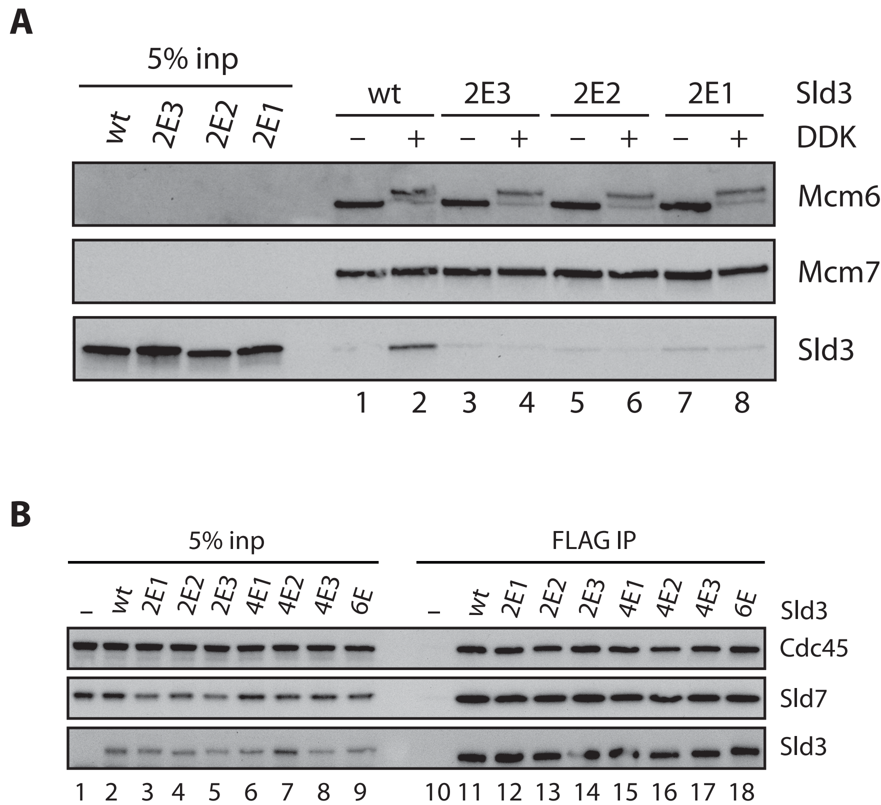


Figure 4.4 Mutations in Sld3 residues 511-531 specifically disrupt binding to the pre-RC

A. Wild type (wt) or mutant Sld3 was tested for binding to the pre-RC *in vitro*. DDK was omitted as indicated. Input and DNA-bound proteins were analysed by immunoblot. **B**. Wild type or mutant FLAG-tagged Sld3 was added to an S-phase protein extract and tested for interactions with Cdc45 and Sld7. Immunoprecipitated proteins were eluted from anti-FLAG beads by boiling in Laemmli sample buffer, and then analysed by immunoblot.

4.2.4 One Mcm2-7 binding site on Sld3/7 is sufficient for recruitment to the pre-RC

The data presented thus far shows that Sld3/7 is recruited to the pre-RC via an interaction between the central portion of Sld3 and the loaded Mcm2-7 double hexamer. Interestingly, biochemical analyses shown in the previous chapter indicates that the Sld3/7 complex likely contains two copies of Sld3. If the Mcm2-7 binding activity of only one of these Sld3 molecules was required for Sld3/7 recruitment, then the Sld3 2E3 mutant should be recruited to the pre-RC in the presence of wild type Sld3. To test this, I used untagged wild type and FLAG-tagged mutant versions of Sld3, which were distinguishable in immunoblot analysis.

The mutant and wild type Sld3 proteins were mixed together in the presence or absence of Sld7, and tested for their ability to bind to the DDK phosphorylated Mcm2-7 double hexamer, as depicted in Fig. 4.5A. In the absence of wild type Sld3, the Sld3 2E3 mutant was not recruited to the pre-RC above background levels (Fig. 4.5B, lane 4), as had been observed previously, but this recruitment defect was partially rescued upon the inclusion of wild type Sld3. Sld7 recruitment was decreased in the sample where both wild type and mutant Sld3 was present (Fig 4.5B, lane 6), which likely represents the uptake of a portion of the Sld7 pool into Sld3/7 complexes containing only mutant Sld3 molecules, which would be incompetent for pre-RC binding. Strikingly, the rescue of the recruitment defect of Sld3 2E3 was dependent upon the inclusion of Sld7 in the reaction (Fig 4.5B, compare lanes 6 and 8). Given the lack of observable pre-RC binding activity in Sld7, the simplest explanation for this observation is that, in the context of recruitment to the pre-RC, Sld7 acts as a 'bridge' between the multiple Sld3 molecules present in Sld3/7 (discussed further in section 4.3.1). Taken together, these data supports a model in which recruitment of the Sld3/7 complex only requires the Mcm2-7 binding activity of one Sld3 molecule.

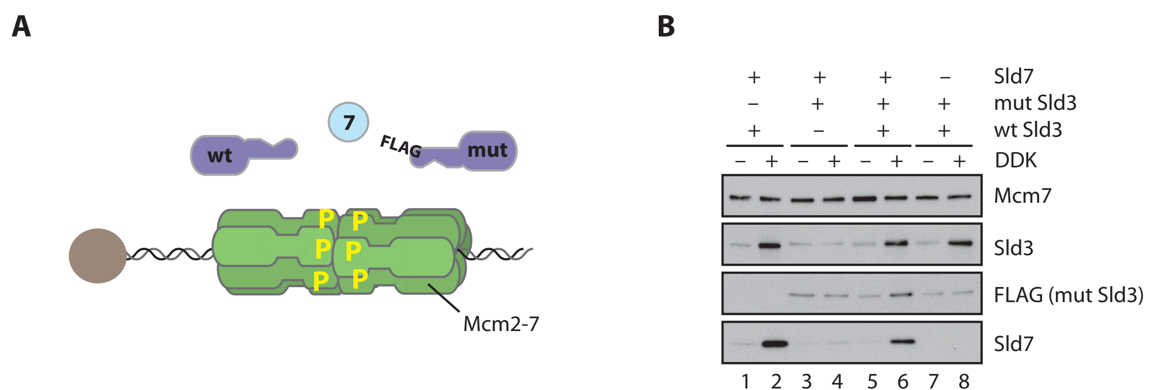


Figure 4.5 One Mcm2-7 binding site on Sld3/7 is sufficient for recruitment to the pre-RC

A. Schematic of experimental setup. Untagged wild type (wt) and FLAG-tagged Sld3 2E3 (mut) were mixed together with Sld7 (7) in various combinations, and tested for recruitment to DDK-phosphorylated pre-RCs. **B.** DNA-bound proteins were analysed by immunoblotting, with the indicated proteins omitted. Although indistinguishable from wild type Sld3 in its migration during SDS-PAGE, FLAG-tagged Sld3 could be specifically detected by anti-FLAG immunoblotting.

4.2.5 Mcm2-7 binding mutants of Sld3 are defective for DNA replication *in vitro*

I next wanted to test the functional significance of the Sld3-Mcm2-7 interaction for DNA replication initiation. To do this, I made use of a recently developed *in vitro* DNA replication assay (On et al., 2014, Gros et al., 2014). In this assay, pre-RCs are initially assembled onto a soluble plasmid DNA template using purified proteins. Following phosphorylation with purified DDK, these pre-RCs are then incubated in an S-phase protein extract, in which Sld3, Sld7, Cdc45, Sld2 and Dpb11 have all been overexpressed. This system supports the complete replication of input plasmids, resulting in the synthesis of covalently closed, circular products.

In order to test the functionality of the Sld3 Mcm2-7 binding mutants in this assay, I first developed an immunodepletion protocol to remove Sld3 from the S-phase extract. The overexpressed second copy of Sld3 present in the extract is C-terminally tagged with 13Myc, and I therefore attempted immunodepletion using anti-Myc beads (Fig. 4.6A). Repeated rounds of depletion were performed, with fresh anti-Myc beads incubated with extract for 1 h at a time. 2 x 1 h incubations with anti-Myc beads from Origene were sufficient to completely deplete Sld3-13Myc from the extract (Fig. 4.6A, lanes 6-9). A significant proportion of the Sld7 present was co-depleted with Sld3 (Fig. 4.6B), and recombinant Sld7 was therefore routinely added back to the Sld3-depleted extracts in subsequent experiments.

Depletion of Sld3-13Myc almost completely abolished the replication activity observed in the undepleted extract (Fig. 4.6C, compare lanes 1 and 2), indicating that any endogenous untagged Sld3 left after depletion of Sld3-13Myc was not sufficient to support DNA synthesis in this system. This defect was rescued by the addition of wild type recombinant Sld3 (and Sld7) to the extract (Fig. 4.6C, lane 3), and the Mcm2-7 binding mutants of Sld3 were subsequently tested. Of the three double point mutants examined, 2E1 and 2E2 both showed a partial defect in DNA replication activity compared with wild type Sld3 (Fig. 4.6C, compare lanes 4/5 with lane 3). The functional relevance of the K511/R512 and K518/R520 residues was confirmed by the absence of DNA replication observed in the 4E1 mutant (Fig. 4.6C, lane 8), in which all four of these residues are mutated. Strikingly, 2E3-

complemented extracts showed no appreciable increase in replication activity relative to the Sld3-depleted extract (Fig. 4.6C, lane 6), and this was also the case for the 4E2, 4E3 and 6E mutants (Fig. 4.6C, lanes 7, 9 and 10), indicating that binding of Sld3 to Mcm2-7 is essential for DNA replication initiation in this *in vitro* system.

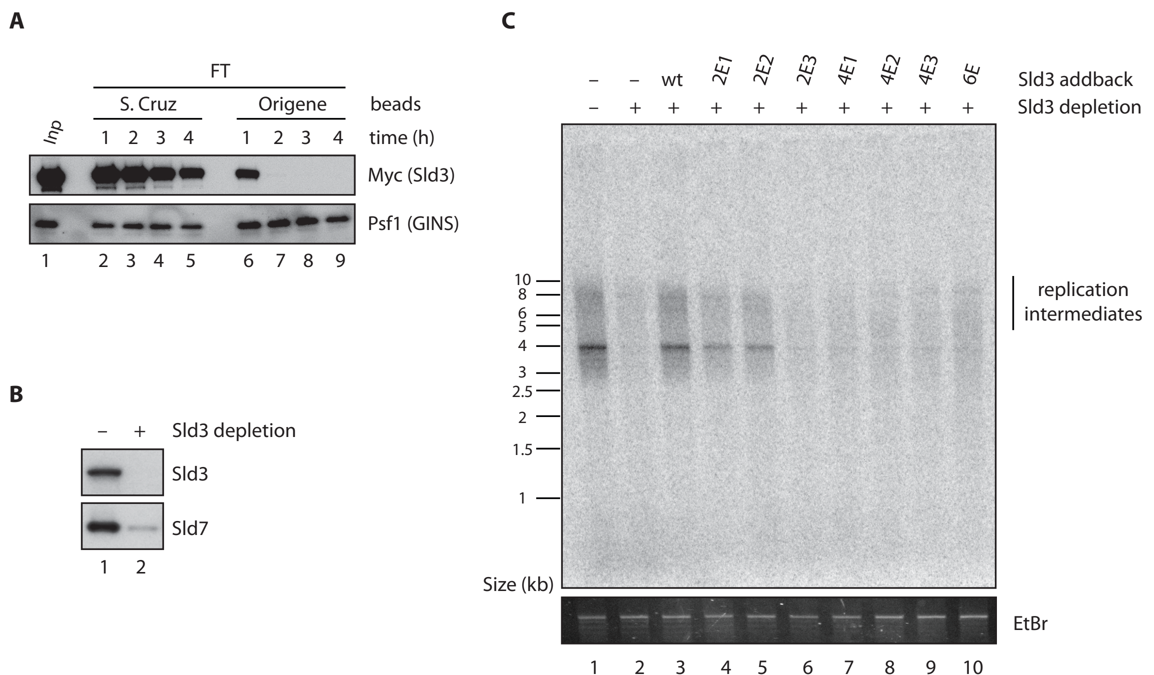


Figure 4.6 Mcm2-7 binding mutants of Sld3 are defective for DNA replication *in vitro*

A. Immunodepletion of Sld3-13Myc from S-phase extracts. Samples were subjected to repeated 1 h rounds of depletion using the beads indicated. Samples were taken at the start of the experiment and after each round of depletion, and analysed by immunoblot against Myc. Anti-Psf1 immunoblot is shown as control for the specificity of the depletion. **B**. Co-depletion of Sld7 with Sld3-13Myc. Extract samples were taken before and after Sld3 depletion and analysed by immunoblot. **C**. *In vitro* DNA replication assay. Wild type or mutant Sld3 was tested for its ability to support replication of a 3.2 kb plasmid template in an S-phase protein extract depleted of Sld3-13Myc. Recombinant proteins were added back at the same concentration as in Fig. 4.7A. Incorporation of ^{32}P-α-dCTP was monitored and samples were analysed by gel electrophoresis and autoradiography. Ethidium bromide staining is shown as a loading control.

4.2.6 Mcm2-7 binding mutants of Sld3 are defective for CMG assembly *in vitro*

To investigate the replication defects described in 4.2.5 in more detail, I next examined the activities of the Sld3 mutants in an *in vitro* replisome assembly assay. This assay is performed in a similar manner to the *in vitro* replication assay described above, except it makes use of linear DNA coupled to beads rather than a soluble plasmid template. The use of DNA beads allows one to follow the recruitment of specific replisome components to the DNA following addition of an S-phase extract to DDK-phosphorylated pre-RCs (On et al., 2014, Gros et al., 2014). I used this system to test for the recruitment of Cdc45 and the Psf1 subunit of GINS, two essential components of the CMG replicative helicase (Fig. 4.7A).

Upon addition of an undepleted extract, Sld3-13Myc, Cdc45 and GINS were all recruited to the DNA in a DDK-dependent manner (Fig. 4.7A, lanes 1 and 2) consistent with previously reported data (On et al., 2014, Gros et al., 2014). Sld3-13Myc depletion resulted in a loss of detectable Cdc45 and GINS recruitment (Fig. 4.7A, lane 3), and this could be rescued by the addition of wild type recombinant Sld3 (Fig. 4.7A, lane 5). This indicates that Cdc45 and GINS recruitment is Sld3-dependent in this system. Addition of the 2E1 and 2E2 mutants, both of which showed a partial defect in DNA replication activity (Fig. 4.6C), resulted in a partial loss of GINS recruitment (Fig. 4.7A, lanes 6 and 7). Notably, the absence of DNA replication activity observed for the 2E3, 4E1, 4E2, 4E3 and 6E mutants correlated with a severe defect in GINS recruitment in this assay (Fig. 4.7A, lanes 8-12).

Interestingly, immunoblotting against Sld3 indicated that the 2E3, 4E1, 4E2, 4E3 and 6E mutants all migrated differently to wild type Sld3 in an SDS-PAGE gel (Fig. 4.7A, lanes 8-12). This unusual migration was not observed in the input samples for this experiment (Fig. 4.7A, input samples 8-12), consistent with some post-translational modification occurring during incubation of recombinant mutant Sld3 in the S-phase extract. To test the possibility that this behaviour was caused by phosphorylation of the Sld3 mutants, I repeated the replisome assembly assay as before, this time splitting the beads in two at the end of the reaction, and subjecting half the sample to treatment with lambda phosphatase (Fig. 4.7B). All the Sld3

mutant proteins tested migrated as a single band in an SDS-PAGE gel upon phosphatase treatment.

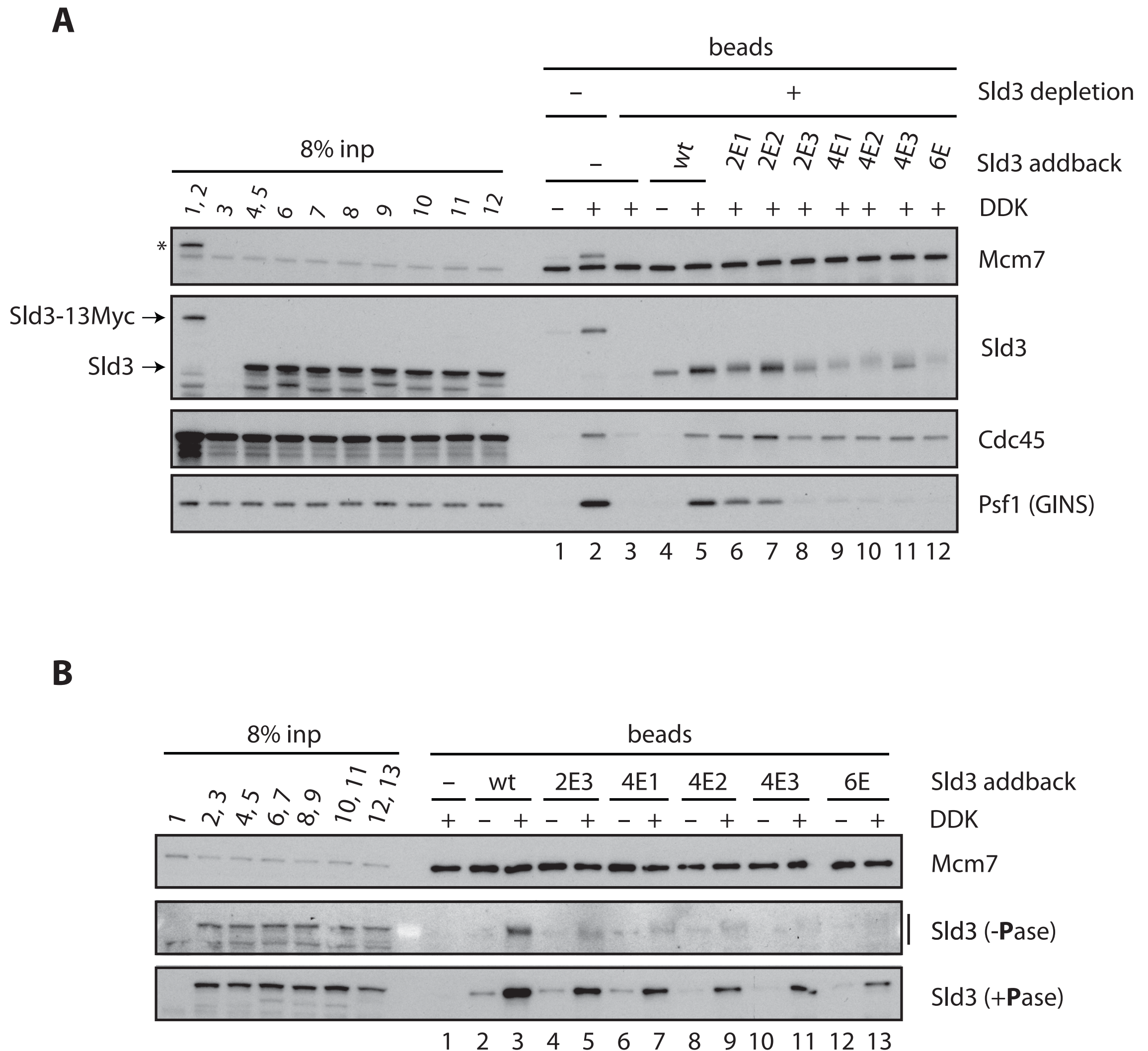


Figure 4.7 Mcm2-7 binding mutants of Sld3 are defective for CMG assembly *in vitro*

A. Pre-RCs were assembled onto linear DNA coupled to beads, treated with DDK as indicated, and an S-phase extract containing wild type or mutant Sld3 then added. DNA-bound proteins were eluted from beads by boiling in Laemmli sample buffer, and analysed alongside input samples by immunoblot. **B**. As in A, but each sample was split in two at the end of the reaction, and half the sample was then treated with lambda phosphatase.

Whilst we are yet to ascertain if this phosphorylation has any functional significance, the phosphatase treatment did reveal that all of the Sld3 mutants were being recruited to the DNA in a DDK-dependent manner in this experiment, albeit at reduced levels relative to wild type Sld3 (Fig. 4.7B). The small amount of DDK-dependent recruitment of these Sld3 mutants appears sufficient for normal levels of Cdc45 recruitment (Fig. 4.7A, compare lanes 8-12 with lane 5). Although we are still investigating the exact mechanism by which these Sld3 mutants can be recruited to the DNA in this system (see 4.2.7), it is clear that the recruitment observed is not functional for proper formation of the active CMG replicative helicase, or for replication initiation.

4.2.7 Cdc45 is not required for Sld3 recruitment *in vitro*

Previously published chromatin immunoprecipitation data has shown that the association of Sld3 with origins of replication is reduced at the restrictive temperature in a cold-sensitive *CDC45* mutant, consistent with Sld3 recruitment to the pre-RC being dependent on Cdc45 *in vivo* (Kamimura et al., 2001). Given this, the presence of Cdc45 in the S-phase extract provided a potential explanation as to why the Sld3 2E1, 2E2, 2E3, 4E1, 4E2, 4E3 and 6E mutants were recruited to DNA in this system, despite the absence of any detectable Mcm2-7 binding activity.

To test whether Sld3 recruitment was dependent upon Cdc45 in the extract system, I made use of a Cdc45 depletion protocol developed in collaboration with a colleague in the lab, Dr. Joe Yeeles. In this protocol, an S-phase extract is made from a strain derived from yKO3, in which the vector from which Cdc45 is overexpressed has been removed (yJY16). Any remaining endogenous Cdc45 can then be depleted from the extract by incubation with Sld3-FLAG beads, as described in 4.2.3. This Cdc45-depleted extract was complemented with purified recombinant Cdc45 (a gift from Dr. Yeeles), and Sld3 recruitment to loaded pre-RCs was then tested (Fig. 4.8). Notably, the absence of Cdc45 in the S-phase extract had no effect on the levels of Sld3 recruitment observed in this system (Fig. 4.8, compare lanes 4 and 6). Significantly, the purified Cdc45 added in this experiment was recruited to DNA beads in a DDK-dependent manner (Fig. 4.8,

lanes 1 and 2), and was sufficient to restore replication activity to Cdc45-depleted extracts (personal communication from Dr. Joe Yeeles). This experiment indicates that Sld3 recruitment to the pre-RC is independent of Cdc45 in this *in vitro* system.

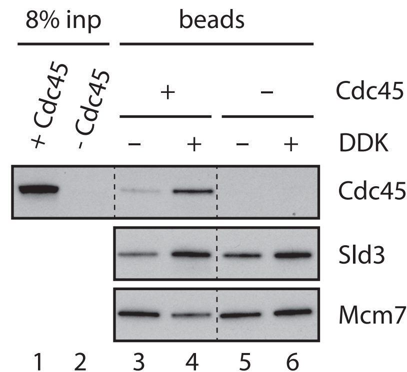


Figure 4.8 Sld3 recruitment to the pre-RC is independent of Cdc45 *in vitro*

Pre-RCs were assembled onto linear DNA coupled to beads, treated with DDK as indicated, and a Cdc45-depleted extract was then added. Purified Cdc45 was added as indicated. DNA-bound proteins were eluted by boiling in Laemmli sample buffer, and analysed by immunoblot.

4.2.8 Mcm2-7 binding mutations in Sld3 affect cell viability

In order to examine the requirement for the Mcm2-7 binding activity of Sld3 for DNA replication *in vivo*, *sld3-2E3* and *sld3-6E* were next tested for their ability to support cell growth. A single copy of *SLD3* in a diploid W303 yeast strain was replaced with a PCR cassette containing the nourseothricin-resistance marker (NAT) and *sld3-2E3* or *sld3-6E*. The resultant heterozygotes were subjected to sporulation, tetrads were dissected, and the spores were germinated on rich media.

An *SLD3*$^{+}$/*sld3-2E3* heterozygote produced tetrads in which two out of four spores exhibited a marked growth defect (Fig. 4.9A, upper panel). The presence of *sld3-2E3* in these slow-growing spores was confirmed by replica plating onto plates containing nourseothricin (Fig. 4.9A, lower panel). Thus, *sld3-2E3* can support viability, although cell growth was reduced relative to spores containing wild type *SLD3*. The inability of *sld3-2E3* to support DNA replication *in vitro* (Fig. 4.6C) likely reflects the relative inefficiency of the reconstituted DNA replication system compared with DNA replication *in vivo*.

When tetrads from an *SLD3⁺/sld3-6E* heterozygote were dissected, only two viable spores were produced (Fig. 4.9B). This indicates that *sld3-6E* is unable to support cell growth when present as a single copy, consistent with the *in vitro* DNA replication defects described in sections 4.2.5 and 4.2.6.

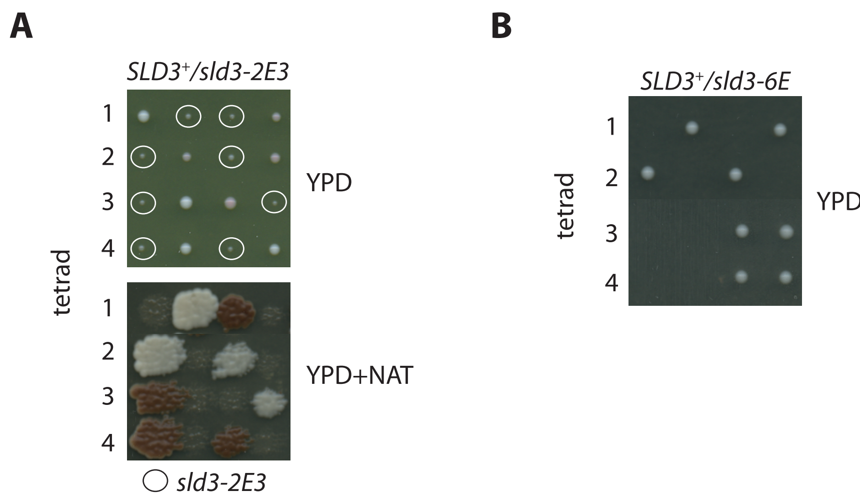


Figure 4.9 Mcm2-7 binding mutations in Sld3 affect cell viability

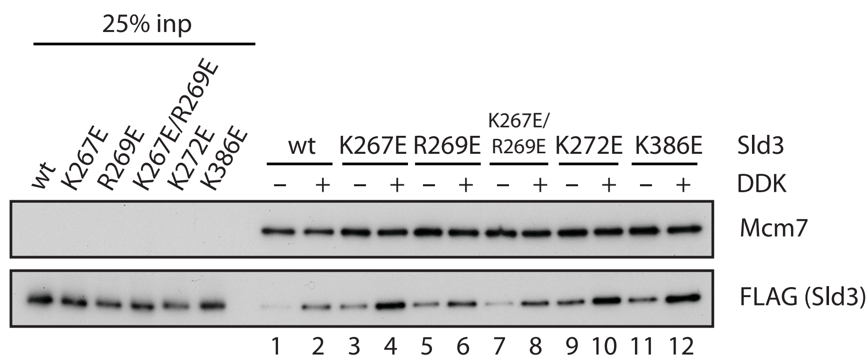
Representative tetrads from **A.** *SLD3⁺/sld3-2E3* and **B.** *SLD3⁺/sld3-6E* heterozygotes. The numbers indicate different tetrads that were dissected. The presence of *sld3-2E3* in the spores circled in A was confirmed by replica plating onto plates containing nourseothricin (NAT).

4.2.9 Mutations within the Sld3-Treslin Domain of Sld3 specifically disrupt Cdc45 binding

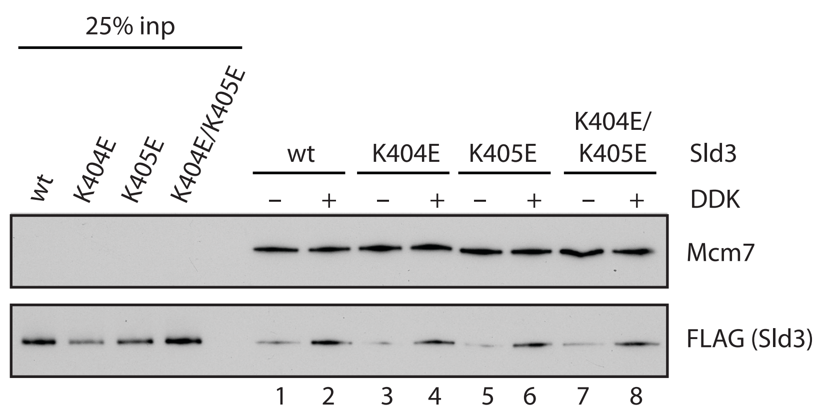
Given its position within Sld3 residues 251-585 and conservation with the human homologue of Sld3, Treslin, we thought that the Sld3-Treslin Domain (S.T.D.) of Sld3 represented an interesting additional target for mutagenesis. Whilst none of the 11 single or double point mutants initially tested showed any discernable deficiencies in pre-RC binding (Fig. 4.10), I was subsequently able to isolate another Sld3 mutant (8E) (Fig. 4.11A), which was completely defective in pre-RC binding *in vitro* (Fig. 4.12A, lane 8). This 8E mutant contained eight amino acid substitutions within residues 296-320 (Fig. 4.11A); attempts to isolate the Mcm2-7 binding defect to a smaller subset of these residues proved unsuccessful, as all

other mutants tested for this region were wild type for recruitment to the loaded Mcm2-7 complex (Fig. 4.12A and B).

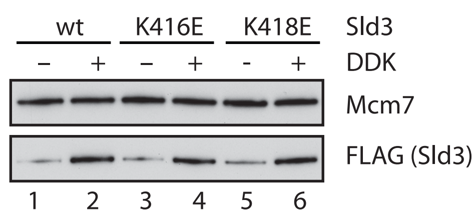
A



B



C



D

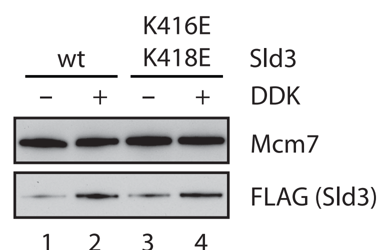
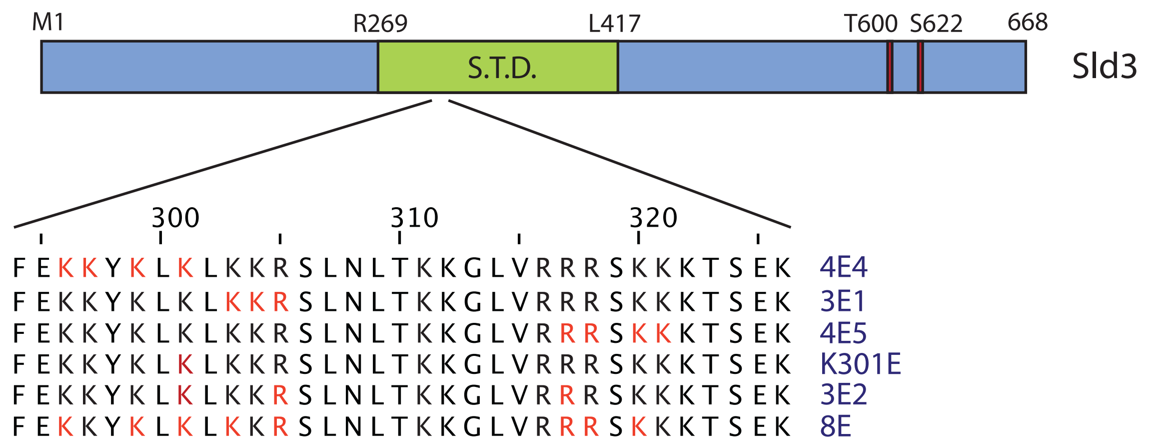


Figure 4.10 Testing Sld3 S.T.D. mutants for binding to the pre-RC

Wild type or mutant Sld3 was tested for binding to the pre-RC *in vitro*. DDK was omitted as indicated, and proteins were analysed by immunoblot.

A



B

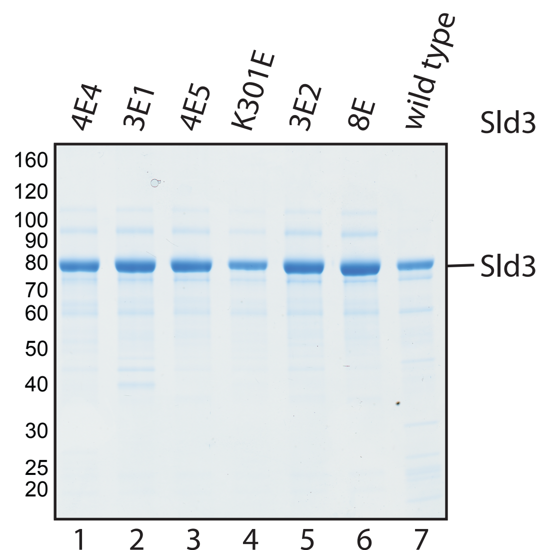


Figure 4.11 Purification of Sld3 S.T.D. mutants

A. Schematic showing position of amino acid substitution mutations in Sld3. Sequence of Sld3 294-326 is shown. Residues highlighted in red were replaced with Glu. Names of mutants are shown to the right. **B**. Sld3 mutants were purified after expression in *E. coli* and visualised by SDS-PAGE and coomassie staining. Molecular weight markers (kDa) are shown to the left.

As a control for the specificity of this Mcm2-7 binding defect, I next went on to test the Sld3 mutants for binding to Cdc45 in an S-phase extract, as previously described in 4.2.3 (Fig. 4.12C). In this instance, all of the proteins tested, which carried mutations in residues 296-320, were defective in Cdc45 binding (Fig. 4.12C, lanes 9-12). The Cdc45-binding defect was present in even those mutants containing only three (3E1) or four (4E4, 4E5) point mutations (Fig. 4.12C, lanes 9-

11), all of which were wild type for binding to the pre-RC (Fig. 4.12A and B). The position of this Cdc45 binding activity is in agreement with previous yeast two-hybrid data, which indicates that residues 151-400 of Sld3 are involved in the interaction with Cdc45 (Tanaka et al., 2011b).

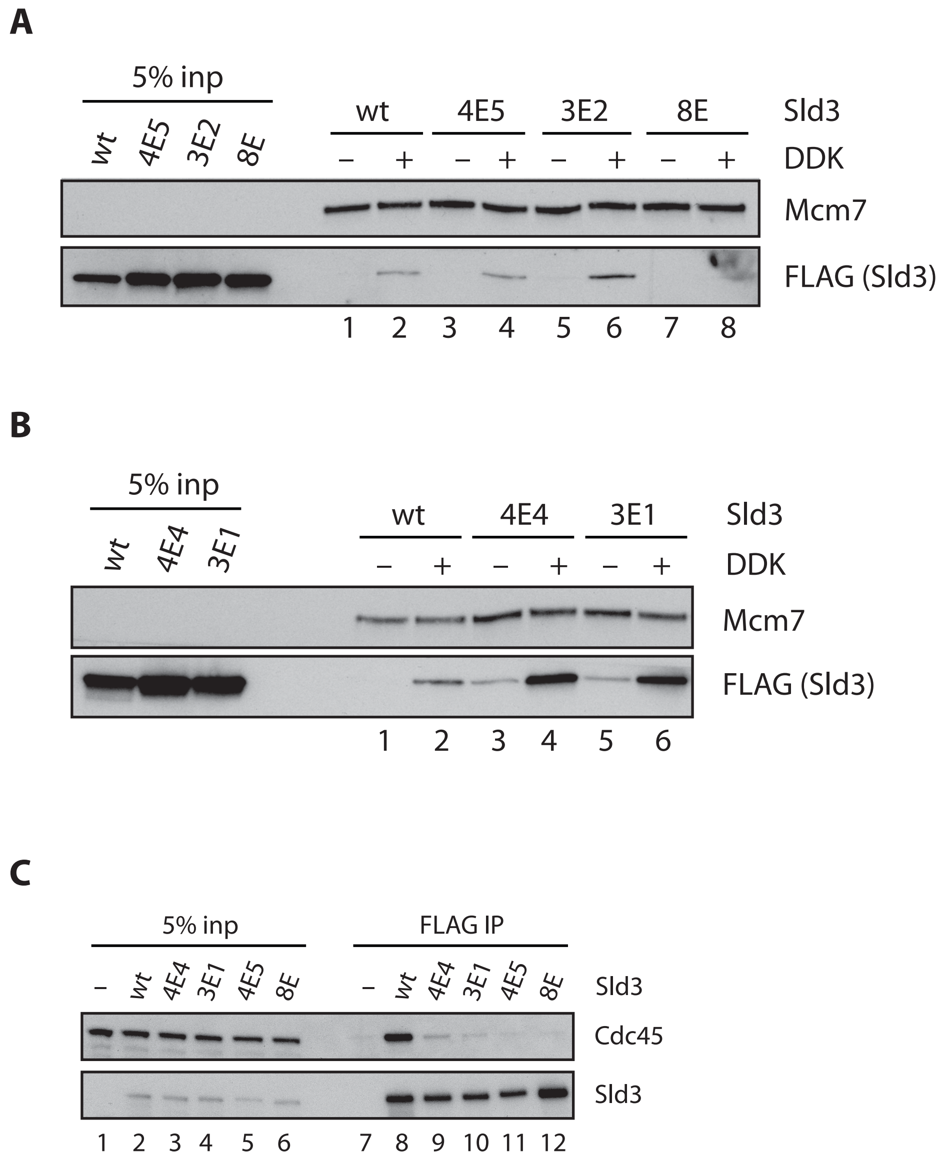


Figure 4.12 Mutations in Sld3 residues 296-320 disrupt binding to both the pre-RC and Cdc45

In **A** and **B**, wild type or mutant Sld3 was tested for binding to the pre-RC *in vitro*. DDK was omitted as indicated, and DNA-bounds proteins were analysed by immunoblot. **C**. Wild type or mutant FLAG-tagged Sld3 was added to an S-phase protein extract and tested for interactions with Cdc45. Bead-bound proteins were eluted by boiling in Laemmli sample buffer, and then analysed by immunoblot.

Having isolated a number of Sld3 mutants (3E1, 4E4 and 4E5) that exhibited an apparently specific defect in Cdc45 binding, I next wanted to test the effect of disrupting this interaction on DNA replication initiation. To do this, I made use of the *in vitro* DNA replication assay described in 4.2.5, again examining the ability of the purified Sld3 mutants to rescue the replication activity of an Sld3-depleted extract (Fig 4.13).

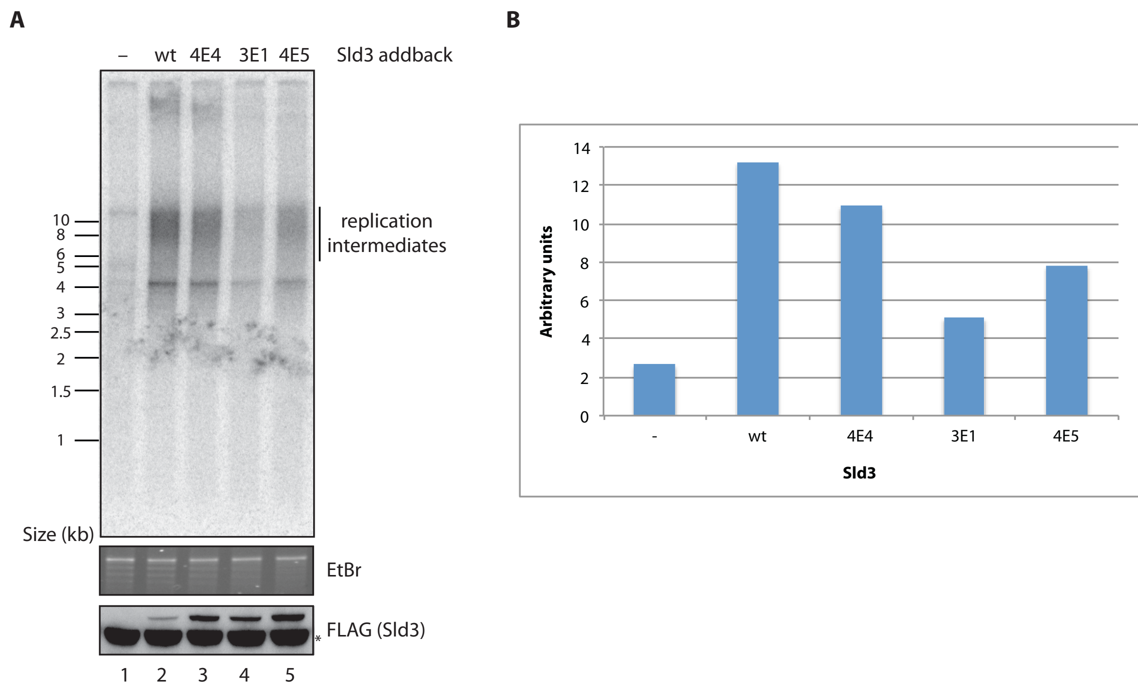


Figure 4.13 Cdc45 binding mutants of Sld3 are defective for DNA replication *in vitro*

A. *In vitro* DNA replication assay. Wild type or mutant Sld3 was tested for its ability to support replication of a 3.2 kb plasmid template in an S-phase protein extract depleted of Sld3-13Myc. Recombinant Sld3 was visualised by immunoblot. * indicates a non-specific band. **B**. Quantification of A, performed using ImageJ software.

Addition of either the Sld3 3E1 or 4E5 mutants resulted in significant replication defects compared to wild type Sld3 (Fig. 4.13A, compare lanes 2 with 4 and 5), consistent with an interaction between Sld3 and Cdc45 being required for efficient replication initiation. Notably, the extract complemented with the Sld3 4E4 mutant was almost as active for DNA replication as the extract containing wild type Sld3 (Fig. 4.13A, compare lanes 2 and 3). This mutant exhibited a small amount of

Cdc45 binding activity in the previous analysis (Fig. 3.12C, lane 9), which appears sufficient to support near wild type levels of replication in this *in vitro* assay. Taken together, these data indicate that Sld3 residues 251-585 contain both Cdc45 and Mcm2-7 binding activities, both of which are important for activation of the Mcm2-7 replicative helicase.

4.3 Discussion

4.3.1 A function for Sld7 during DNA replication initiation?

sld7Δ mutant yeast cells exhibit a marked growth defect consistent with slow progression through S-phase of the cell cycle (Tanaka et al., 2011b). However, apart from its role as a constitutive binding partner of Sld3, the exact function of Sld7 during origin firing has not been well characterised. The CDK-dependent interaction between Sld3 and Dpb11 can occur *in vitro* in the absence of Sld7 (Zegerman and Diffley, 2007). Furthermore, in the first section of this chapter, I showed that Sld7 is not required for direct binding of Sld3 to the Mcm2-7 complex (Fig. 4.1B). In these experiments, Sld7 recruitment was totally dependent on the N-terminal 250 residues of Sld3 (Fig. 4.1C), consistent with previous data showing that Sld7 origin recruitment is reduced at the restrictive temperature in an *sld3-5* temperature-sensitive mutant, and that the N-terminus of Sld3 is an Sld7 binding site (Fig. 3.8, Tanaka et al., 2011b).

The data presented in Fig. 4.5 indicates that, whilst not required for the direct interaction between Sld3 and Mcm2-7, Sld7 is required for the recruitment of a second molecule of Sld3, if the Mcm2-7 binding activity of that copy of Sld3 is disrupted (Fig. 4.5B, compare lanes 6 and 8). This observation is suggestive of a novel role for Sld7 as a ‘bridge’ between multiple Sld3 molecules when Sld3/7 is recruited to the pre-RC. Notably, the multimerisation of Sld3 *per se* is not dependent on Sld7 (Fig. 3.8). One possible explanation for this discrepancy is that the conformation of Sld3 bound to the Mcm2-7 double hexamer is not competent for binding to a second molecule of Sld3. Whether or not this bridging function of Sld7 has any functional relevance is still to be examined. Although speculative at present, one intriguing possibility is that Sld7 associates with multiple Sld3

molecules, which themselves are bound to different Mcm2-7 hexamers within the pre-RC. By interacting with multiple Sld3 molecules across the interface of the Mcm2-7 double hexamer, Sld7 may co-ordinate the formation of two replisomes, each containing a single Mcm2-7 hexamer, during origin firing. To test this hypothesis, it may be informative to use DNA fibre analysis to look for bi-directional replication initiation (i.e. the co-ordinated activation of two Mcm2-7 hexamers) in the presence and absence of Sld7 *in vitro* and *in vivo*.

4.3.2 Sld3 residues 251-585 are functional for both Cdc45 and Mcm2-7 binding during DNA replication initiation

In section 4.2.2 of this chapter, I used purified fragments of Sld3 to map the Mcm2-7 binding activity of the protein to residues 251-585 (Fig. 4.2). Mutation of 3 distinct pairs of Lys/Arg within residues 511-531 specifically disrupted the Sld3-Mcm2-7 interaction (Fig. 4.4A). I subsequently showed that these mutants exhibit defects in DNA replication (Fig. 4.6C) and CMG assembly (Fig. 4.7A) *in vitro*. Furthermore, *sld3-6E* was incapable of supporting cell viability (Fig. 4.9B). Collectively, these data are consistent with a functional requirement for the novel Mcm2-7 binding activity of Sld3 during activation of the replicative helicase.

The central section of Sld3 contains the Sld3-Treslin Domain (S.T.D.), which is conserved with Treslin, the human Sld3 homologue, and I was able to show that mutations within this domain can disrupt binding to both the pre-RC and Cdc45 (Fig. 4.12). The charge reversal that these Arg/Lys to Glu mutations introduced may be consistent with the existence of a negatively charged Sld3 binding site on Cdc45. Whilst these observation could be consistent with a model in which residues 296-320 of Sld3 are involved in both Mcm2-7 and Cdc45 binding, the Cdc45 binding defects of the 3E1, 4E4 and 4E5 mutants (Fig. 4.12C), all of which were competent for pre-RC binding (Fig. 4.12A and B), argue that this region likely represents a *bona fide* Cdc45 binding site, consistent with yeast two-hybrid data (Tanaka et al., 2011b). Only extensive mutation of the S.T.D., as in the 8E mutant, can lead to a likely non-specific disruption of the Mcm2-7 binding activity of Sld3, which resides C-terminal of the S.T.D. in residues 511-531.

Thus, the central section of Sld3 is a multi-functional region, containing binding sites for both Cdc45 and Mcm2-7, and the disruption of either of these interactions results in deficient DNA replication initiation. This section of Sld3 has previously been shown to be a target of the Rad53 DNA damage checkpoint kinase (Lopez-Mosqueda et al., 2010, Zegerman and Diffley, 2010). Phosphorylation of Sld3 by Rad53 disrupts the interaction between Sld3 and Cdc45, thereby helping to prevent late origin firing in response to DNA damage in S-phase. Given the likely ionic nature of the Sld3-Mcm2-7 interaction, and the fact that a number of Rad53 phosphorylation sites are located between residues 511-531 of Sld3 (Lopez-Mosqueda et al., 2010, Zegerman and Diffley, 2010), it seems likely that Rad53 phosphorylation might also block the interaction between Sld3 and Mcm2-7. Whilst experiments to test this hypothesis are ongoing, such a mechanism would represent an additional level of redundancy in the intra-S phase checkpoint, and provide a potentially potent block to further origin firing under conditions of DNA damage.

4.3.3 Mcm2-7 binding mutants of Sld3 are recruited to the pre-RC in an S-phase extract

Despite the lack of any detectable Mcm2-7 binding activity, all of the Sld3 Mcm2-7 binding mutants discussed in this chapter were recruited to pre-RCs in a DDK-dependent manner when added to an S-phase extract (Fig. 4.7). This is unlikely to be caused by Cdc45-dependent Sld3/7 recruitment, as Cdc45 was not required for DDK-dependent Sld3/7 recruitment in this *in vitro* system (Fig. 4.8). This lack of dependency on Cdc45 is consistent with the observed pre-RC binding activity of Sld3/7 alone, as well as previously published data in *S. pombe*, which shows that Sld3 can be recruited to origins of replication that are not bound by Cdc45 *in vivo* (Yamada et al., 2004).

Given this, it seems likely that the Sld3/7 recruitment observed with these mutants may be explained by one of two reasons. Firstly, the extract-based replisome assembly assay is performed in much lower salt concentrations (0.15 M K-

Glutamate) than the Sld3/7 recruitment assay, which is performed in 0.5 M KOAc. Thus, it may be that the Mcm2-7 binding defect observed for these mutants under 0.5 M KOAc is simply not manifested at lower salt conditions. Additionally, work presented in this chapter suggests that wild type Sld3 can rescue the recruitment defect of an Mcm2-7 binding mutant of Sld3 (Fig. 4.5). The Sld3 immunodepletion protocol, which I developed, only removes the 13Myc-tagged second copy of Sld3 from the extract (Fig. 4.6A). Whilst insufficient to support any DNA replication (Fig. 4.6C), the small amount of endogenous Sld3 remaining in the extract after Sld3-13Myc depletion may be sufficient to allow for a small amount of recruitment of the exogenously added mutant Sld3. Importantly, despite the fact that some of the Sld3 mutants, such as Sld3 2E3, are recruited to DNA to near wild type levels (Fig. 4.7B), they all exhibit defects in DNA replication (Fig. 4.6C), consistent with the direct binding of Mcm2-7 by Sld3 being required for activation of the replicative helicase during replication initiation.

Chapter 5. Results 3

5.1 Introduction

The requirement for Cdc7 kinase activity during DNA replication initiation was first shown over 40 years ago (Hereford and Hartwell, 1974, Hartwell et al., 1974), and much work has since been published on how this kinase promotes replicative helicase activation upon S-phase entry. Unlike the case of CDK (Zegerman and Diffley, 2007, Tanaka et al., 2007), however, it has proved difficult to pinpoint a minimal set of functions for DDK, despite the identification of a number of alleles that overcome the inviability phenotype caused by *cdc7* or *dbf4* deletion *in vivo* (Sheu and Stillman, 2010, Hardy et al., 1997, Hayano et al., 2012, Matsumoto et al., 2011).

The Mcm2-7 complex itself is a DDK target; its Mcm2, Mcm4 and Mcm6 subunits have all been identified as DDK substrates *in vitro* and *in vivo* (Masai et al., 2006, Montagnoli et al., 2006, Sheu and Stillman, 2006, Randell et al., 2010). In *S. cerevisiae*, the majority of mapped DDK phosphorylation sites are located within the N-terminal regions of these subunits (Randell et al., 2010), which are rich in Ser/Thr residues and predicted to be in a disordered conformation. An N-terminal deletion of Mcm4 can bypass the requirement for DDK *in vivo* (Sheu and Stillman, 2010), and it was thus proposed that DDK promotes origin firing by alleviating some intrinsic inhibitory activity that resides within this section of Mcm4. It is notable that DDK activity can also be bypassed by the *mcm5-bob1* allele (Hardy et al., 1997) and, in *S. pombe*, by deletion of *mrc1* or *rif1* (Hayano et al., 2012, Matsumoto et al., 2011), none of which have been shown to be DDK substrates. This may indicate that DDK phosphorylation induces a conformational change in the Mcm2-7 complex that is required for its activation, with this conformational change somehow being constitutively adopted in the context of the aforementioned DDK bypass alleles.

In the previous two chapters I have presented data, which shows that DDK regulates recruitment of Sld3/7 to the pre-RC during Mcm2-7 helicase activation. This recruitment is mediated by a direct interaction between Sld3 and the loaded

Mcm2-7 complex, and is essential for replication initiation. In the following section, I will discuss work towards an explanation for the role of DDK in regulating this interaction. I will describe a number of novel phosphorylation-dependent binding sites for Sld3 on the Mcm2-7 complex, and show data, which illustrates the functional consequences of disrupting the binding of Sld3 to a specific Mcm subunit, Mcm6.

5.2 Results

5.2.1 Interaction of Sld3/7 with individual Mcm subunits

In order to better understand how DDK regulates binding of Sld3/7 to the Mcm2-7 complex, I first wanted to examine the interaction of Sld3/7 with individual Mcm subunits. The presumed phosphorylation-dependence of any such interactions presented us with a significant technical obstacle, namely that DDK exhibits a strong substrate preference for only those Mcm complexes that have been assembled into the pre-RC *in vitro* (Randell et al., 2010). Consistent with this, we were never able to reconstitute quantitative phosphorylation of individually expressed and purified Mcm subunits (data not shown).

Because of this issue, we decided to develop an approach that took advantage of the fact that loaded Mcm2-7 complexes have the ability to slide on dsDNA *in vitro* (Remus et al., 2009), and will dissociate from linear DNA fragments relatively rapidly if incubated in high salt buffers. We hypothesised that once dissociated from the DNA, these Mcm2-7 complexes would be significantly destabilised, potentially dissociating into individual subunits. Such a scenario would allow us to first phosphorylate Mcm2-7 in the context of the pre-RC, and then, having affected dissociation of the double hexamer, examine interactions between Sld3/7 and individual phosphorylated Mcm subunits (Fig. 5.1A).

The confirm the validity of this approach, I initially tested whether the Mcm2-7 complex remains intact after dissociation from DNA under high salt conditions. Mcm2-7 was loaded onto DNA and phosphorylated with DDK, and then incubated in a buffer containing 0.5 M NaCl overnight. Those complexes that dissociated from

the DNA beads were then subjected to calmodulin affinity chromatography (Fig. 5.1B). No Mcm subunits other than CBP-Mcm3 were recovered on the calmodulin beads during this reaction (Fig. 5.1B, lanes 3 and 6), consistent with the fact that the Mcm2-7 ring is unstable under the experimental conditions used.

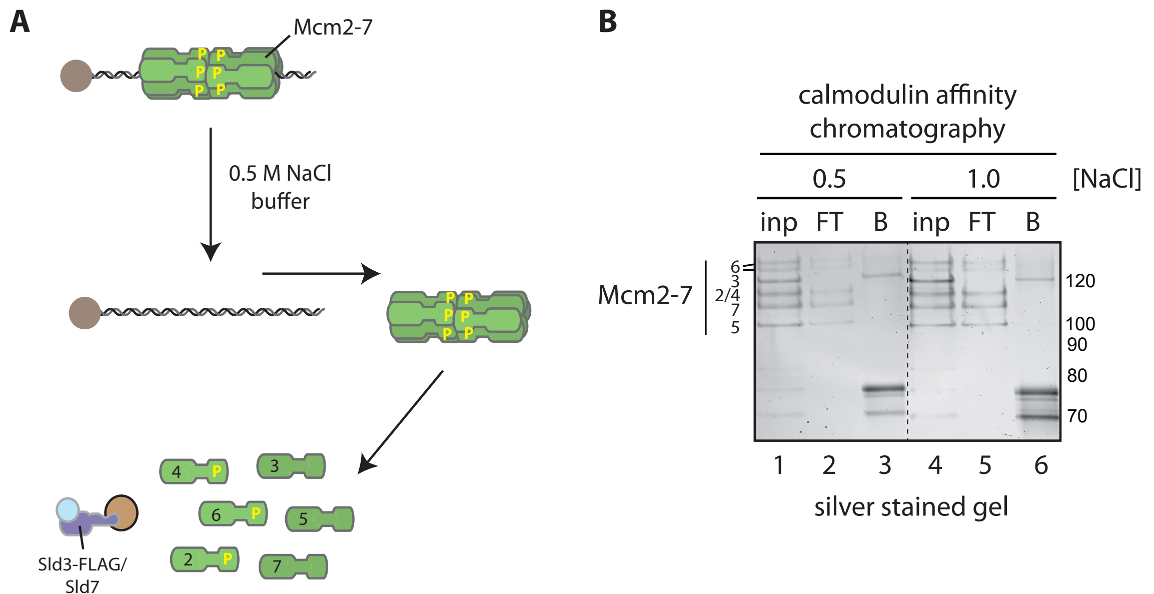


Figure 5.1 The Mcm2-7 double hexamer is disrupted by prolonged incubation in 0.5 M NaCl

A. Schematic of experimental setup. Mcm2-7 loaded onto linear DNA bound to beads was phosphorylated with DDK and then left in buffer containing 0.5 M NaCl overnight. Under these conditions, Mcm2-7 dissociates from DNA into individual Mcm subunits. In later experiments, interactions with Sld3/7 beads were examined. **B**. Phosphorylated Mcm2-7 was subjected to calmodulin affinity chromatography after dissociation from DNA under different [NaCl] as indicated. Input (inp), flow-through (FT) and beads (B) samples were collected and analysed by SDS-PAGE and silver staining. Visible bands are annotated to the left.

Subsequently, I went on to test for interactions between Sld3/7 and phosphorylated Mcm subunits (Fig. 5.2). Sld3/7 was coupled to anti-FLAG resin via a C-terminal 3xFLAG tag on Sld3, and then incubated with a sample of phosphorylated Mcm2-7 complexes that had been disrupted as described above. Whilst Mcm3 is bound non-specifically to the anti-FLAG resin under these conditions, we observed an Sld3/7-dependent enrichment of phosphorylated Mcm6 on the beads (Fig. 5.2A, lane 2). The phosphorylation-dependence of the Sld3/7-Mcm6 interaction was confirmed by either omission of DDK during the initial phosphorylation of the pre-

RC, or by treatment of the dissociated Mcm proteins with lambda phosphatase before incubation with Sld3/7 beads (Fig. 5.2B).

By performing immunoblot analysis, we observed a second, weaker DDK-dependent Sld3/7-Mcm interaction with Mcm4 (Fig. 5.2C), which was not detectable by silver staining. Notably, mutations in Sld3 that were shown to disrupt pre-RC binding in the previous chapter could be seen to disrupt binding to both Mcm4 and Mcm6 (Fig. 5.2C), consistent with these interactions being important in the context of the Sld3/7 recruitment to the pre-RC.

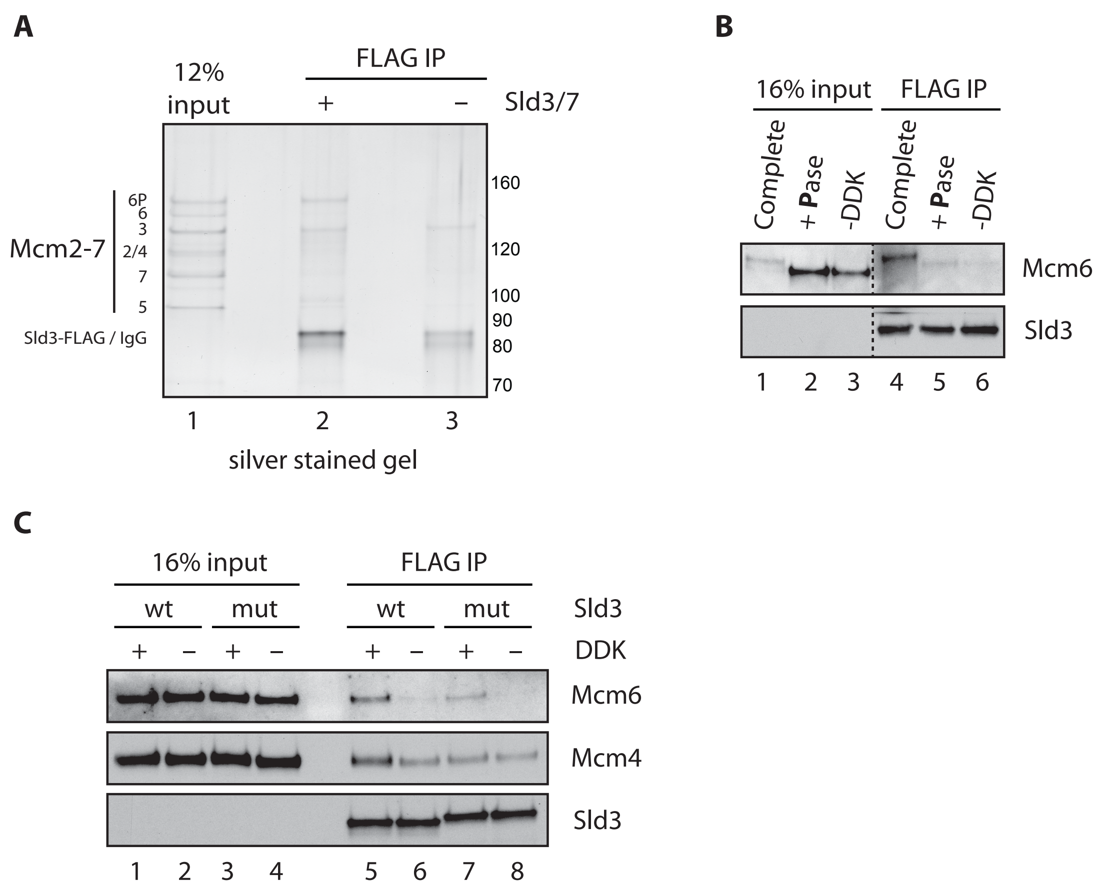


Figure 5.2 Sld3/7 interacts with phosphorylated Mcm6 and Mcm4

A. A mixture of phosphorylated Mcm2-7 subunits was incubated with Sld3/7-3xFLAG-coupled or empty anti-FLAG beads as indicated. Proteins were eluted by boiling in Laemmli sample buffer and visualised by SDS-PAGE and silver staining. Visible bands are annotated. **B**. Samples were prepared as in A and analysed by immunoblot. Lambda phosphatase / DDK were included as indicated. **C**. Mcm6 and Mcm4 were tested for interaction with wild type or 6E mutant Sld3. Samples were treated with lambda phosphatase before gel loading to prevent epitope masking of Mcm6 (as seen in B, lane 1).

5.2.2 Mapping the interaction between Sld3/7 and Mcm6

As a first step towards identifying the phosphorylation sites on Mcm6 that are important for Sld3/7 binding, I next set out to determine the region of Mcm6 to which Sld3/7 binds. In order to do this, I engineered a cleavage site for TEV protease into Mcm6 between Ala 497 and Asn 498, in a region of the protein that is poorly conserved between species (Fig. 5.3). The aim was to proteolytically cleave Mcm6 following its dissociation from DNA, and then examine the binding of Sld3/7 to the resultant two fragments. A 9Myc tag was introduced at the N-terminus of Mcm6 in this construct, allowing detection by immunoblot. The C-terminal fragment produced after cleavage could be detected using an antibody raised against the extreme C-terminus of Mcm6.

Mcm6TEV2

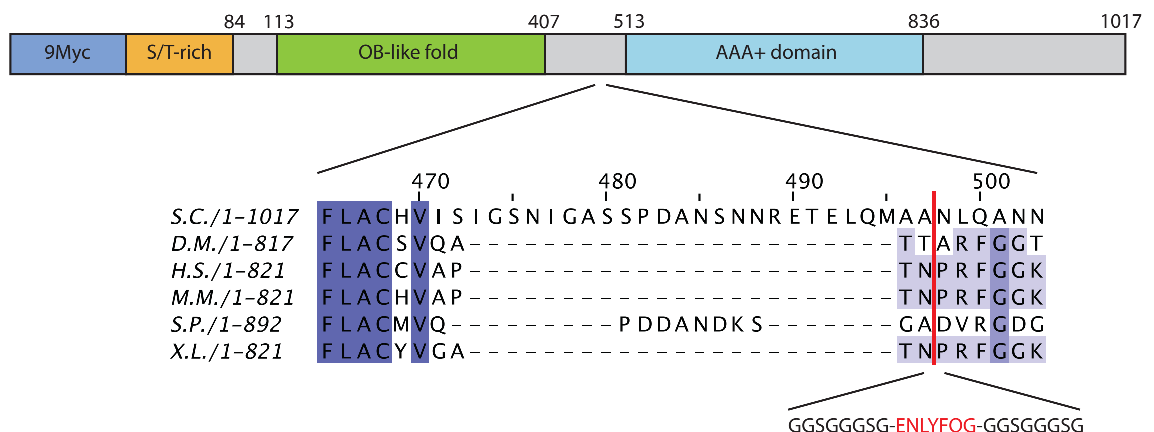


Figure 5.3 Schematic of TEV-cleavable Mcm6TEV2

The domain architecture of Mcm6 and alignment of Mcm6 residues 465-503. The alignment was generated in Jalview using Clustal (Sievers et al., 2011). Mcm6 from various eukaryotic species is included (S.C., *Saccharomyces cerevisiae*, D.M., *Drosophila melanogaster*, H.S., *Homo sapiens*, M.M., *Mus musculus*, S.P., *Schizosaccharomyces pombe*, X.L., *Xenopus laevis*). Residue numbers above the alignment correspond to *S. cerevisiae* Mcm6. The red vertical line indicates the position where the TEV cleavage site (red) and linker sequences (black) were inserted.

After overexpression in yeast, an Mcm2-7/Cdt1 heptamer containing the TEV-cleavable version of Mcm6 (Mcm6TEV2) was purified to homogeneity (Fig. 5.4A). As the ATG codon at the start of Mcm6 had not been destroyed during cloning of the 9Myc tag, some untagged TEV-cleavable Mcm6 was present in the final

sample (Fig. 5.4A, lane 1). As a control for whether the insertion of the TEV cleavage site into Mcm6 had affected the activity of the Mcm2-7 complex, the Mcm6TEV2 complex was initially tested for pre-RC assembly *in vitro* (Fig. 5.4B). Reassuringly, the Mcm6TEV2 complex could be recruited to DNA and loaded into a salt resistant form in a manner that was indistinguishable from a wild type Mcm2-7 complex (Fig. 5.4B, compare lanes 2 and 4).

The addition of TEV protease to reactions during the overnight incubation in 0.5 M NaCl buffer resulted in complete cleavage of Mcm6 into N and C-terminal fragments (Fig. 5.4C). Upon incubation with Sld3/7 beads, the Mcm6 N-terminal fragment was specifically bound in a manner that was sensitive to phosphatase treatment or the omission of DDK (Fig. 5.4D, compare lanes 5 with 6 and 7), whereas Mcm6 487-1017 was never detected on the beads at the end of the reaction (Fig. 5.4D, lane 5). This indicates that Mcm6 1-486 is both necessary and sufficient for DDK-dependent Sld3/7 binding.

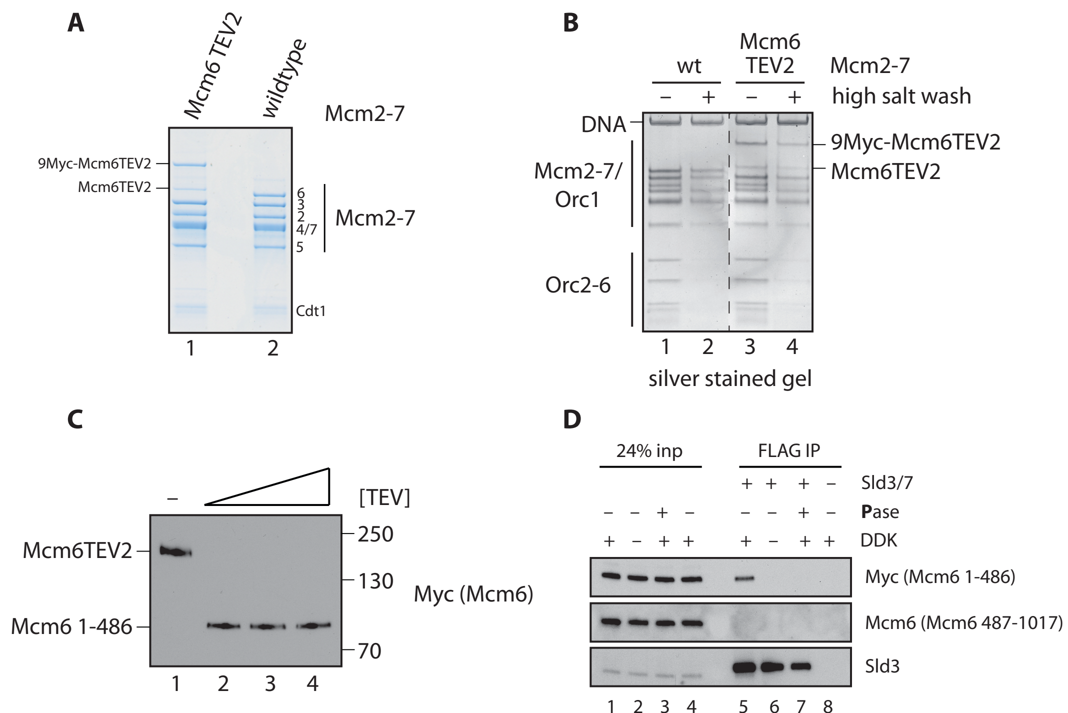


Figure 5.4 Mcm6 1-486 is necessary and sufficient for Sld3/7 binding

A. Purified Mcm2-7/Cdt1 visualised by SDS-PAGE and coomassie staining. The presence of TEV-cleavable Mcm6 is indicated. **B**. Mcm2-7 loading assay performed with wild type or TEV-cleavable Mcm2-7. Samples were analysed by

silver staining. **C**. Purified TEV protease was added to TEV-cleavable Mcm2-7 dissociated from DNA under 0.5 M NaCl, and the products analysed by SDS-PAGE and immunoblotting. Full-length Mcm6 and cleavage products are annotated. Molecular weight markers are indicated on the right. **D**. N and C-terminal fragments of TEV-cleavable Mcm6 produced after proteolysis were tested for interaction with Sld3/7 beads. Proteins were omitted as indicated. Immunoblotting was performed against Myc, Mcm6 and Sld3.

5.2.3 Sld3 interacts with multiple phosphopeptides in the N-terminal portion of Mcm6

In an attempt to identify the Mcm6 phosphorylation sites that are involved in Sld3/7 binding, I next analysed the interaction between Mcm6 and Sld3/7 using a phosphopeptide array screen. A membrane was spotted with 18-residue peptides covering the entire sequence of Mcm6. A two-residue start increment was used between peptides, with unphosphorylated and singly phosphorylated versions of each peptide spotted next to one another. This array was then probed with full-length FLAG-tagged Sld3, and bound Sld3 was detected using an anti-FLAG antibody (Fig. 5.5). Given the observation that Mcm6 1-486 is sufficient for the interaction with Sld3/7, I initially focussed on Sld3-bound phosphopeptides located in this section of Mcm6.

The results from the peptide array are summarised in table 5-1. A total of 18 potential sites were identified in peptides that interacted with Sld3 in a phosphorylation-dependent manner. These sites were positioned throughout the N-terminal portion of Mcm6. Notably, in a number of cases, the residue immediately downstream of the phosphorylated Ser/Thr was either an additional Ser/Thr (as with Ser 65 and Ser 226) or a Glu residue (as with Thr 259 and Thr 480), consistent with the predicted consensus sequence for DDK phosphorylation sites (Randell et al., 2010).

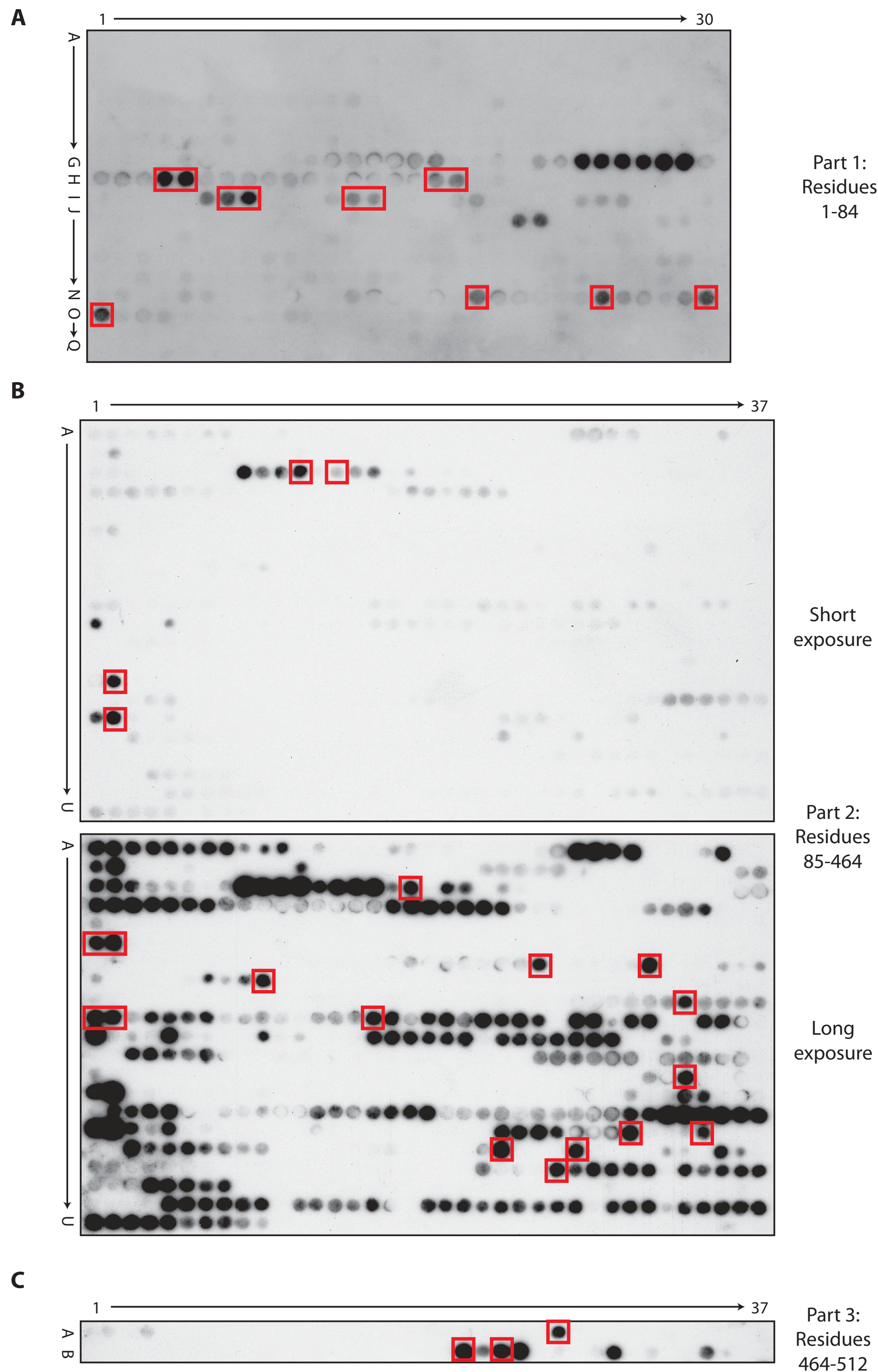


Figure 5.5 Sld3 can bind to multiple phosphopeptides in Mcm6 1-486

Peptide array assay for interaction of full-length Sld3 with Mcm6. 18-residue peptides covering Mcm6 residues 1-486 were spotted on 3 separate membranes (**A, B, C**) and incubated with purified FLAG-Sld3. Interactions were detected with an anti-FLAG antibody. Position references are given to the left and top of each array. Phosphopeptides exhibiting strong interactions with Sld3 (summarised in Table 5-1) are highlighted in red.

Site	Sequence	Array	Total Peptides	Peptide References
S37	SSGLDSQIGSR	Part 1	4	H4 (*G30*), H17 (*H13*), I7 (*I6*), I13 (*I12*)
S41	DSQIGSRLHFP	Part 1	4	H5 (*G30*), H18 (*H13*), I8 (*I6*), I14 (*I12*)
S65	PFVNDSTQFSS	Part 1	3	N19 (*N17*), N25 (*N23*), N30 (*N29*)
S66	PFVNDSTQFSS	Part 1	1	O1 (*N29*)
T150	IYDLNTIYIDY	Part 2	3	C12 (*C11*), C14 (*C13*), C18 (*C17*)
S226	DDMNGSSLPRD	Part 2	1	F1 (*E37*)
S227	DDMNGSSLPRD	Part 2	1	F2 (*E37*)
T247	TSAMATRSITT	Part 2	3	G25 (*G20*), G31 (*G27*), H10 (*H7*)
S255	ITTSTSPEQTE	Part 2	1	J1 (*I35*)
T259	TSPEQTERVFQ	Part 2	2	I33 (*I27*), J2 (*I35*)
T272	FFNLPTVHRIR	Part 2	1	J16 (*J13*)
S324	DNVEQSFKYTE	Part 2	2	M33 (*M32*), N2 (*N1*)
S351	LNVTRSRFLDW	Part 2	1	P2 (*P1*)
T376	GSMPRTLDVIL	Part 2	2	P30 (*P27*), P34 (*P31*)
T398	DRCKFTGVEIV	Part 2	2	Q23 (*Q22*), Q27 (*Q26*)
T423	SSTLDTRGISK	Part 2	1	R26 (*R22*)
S475	SPDANSNNRET	Part 3	1	A26 (*A21*)
T480	SNNRETELQMA	Part 3	2	B21 (*B20*), B23 (*B22*)

Table 5-1 Summary of Mcm6 phosphopeptides that bind to Sld3

Table showing phosphorylation sites in Mcm6 that interact with Sld3 when incorporated into 18-residue peptides. The site is shown in the first column, and the local sequence surrounded the site (highlighted in red) is shown in column 2. The array on which the peptide was spotted is shown in column 3. The number of Sld3-interacting peptides that contained the site is shown in column 4. The position of the Sld3-interacting peptides on the array is shown in column 5. The position of the corresponding unphosphorylated peptides is shown in italics in brackets after each peptide reference.

To validate the interactions observed on the peptide array, I next designed a set of phosphorylated 30-residue peptides containing a subset of the sites observed to interact with Sld3. The peptides were coupled to streptavidin beads via an N-terminal biotin group, and these beads were then tested for interaction with Sld3 (Fig. 3.6). Of the 15 phosphorylated peptides tested, four showed significant binding to Sld3 (Fig. 5.6A and B). The interaction between Sld3 and S351P was subsequently shown to be independent of phosphorylation (data not shown). Sld3 binding to peptides containing phosphorylated Thr 150, Ser 255 or Thr 259 was reduced upon the addition of lambda phosphatase (Fig. 5.6 C and D, compare lanes 3 and 4), or when the phosphopeptide was replaced with the corresponding unphosphorylated version (Fig. 5.6 C and D, compare lanes 2 and 3). These data not only indicate that Sld3 can interact directly with multiple phosphorylation sites in the N-terminal half of Mcm6, but also uncover a novel phosphopeptide binding activity resident within Sld3.

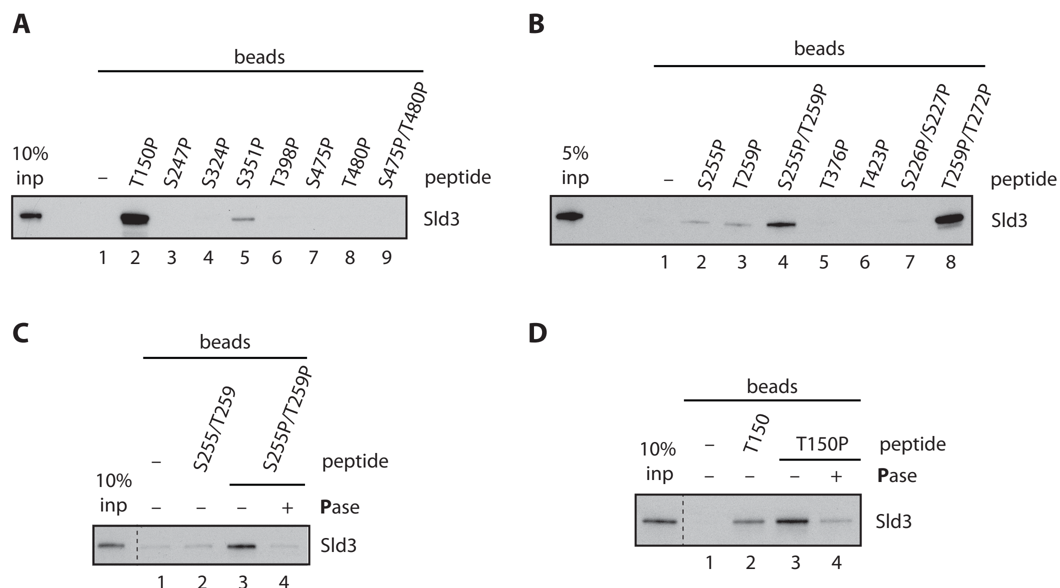


Figure 5.6 Interactions of Sld3 with 30-residue Mcm6 phosphopeptides

30-residue phosphorylated peptides were bound to beads and used to pulldown Sld3. Bound proteins were eluted by boiling in Laemmli buffer and analysed by immunoblot. In **C** and **D** (lane 4), peptide-coupled beads were dephosphorylated with lambda phosphatase before incubation with Sld3.

5.2.4 Identification of phosphorylation sites in Mcm6

The finding that the Sld3/7 binding site on Mcm6 is located within the first 487 residues of the protein is consistent with a previous report showing that this region of Mcm6 is phosphorylated by DDK *in vitro* (Randell et al., 2010). Notably, this region of Mcm6 also contains a number of phosphopeptides, which Sld3 can interact with (see table 5-1). In collaboration with the Protein Analysis and Proteomics Facility at the LRI we performed our own mass spectrometry analysis to map phosphorylation sites on Mcm6. Mcm2-7 was loaded onto DNA and phosphorylated with DDK to a level that was saturating for Sld3/7 recruitment (Fig. 5.7). The mass spectrometry analysis was performed twice using either AspN or trypsin for proteolytic digestion, and 76% peptide coverage of Mcm6 was achieved across the two experiments.

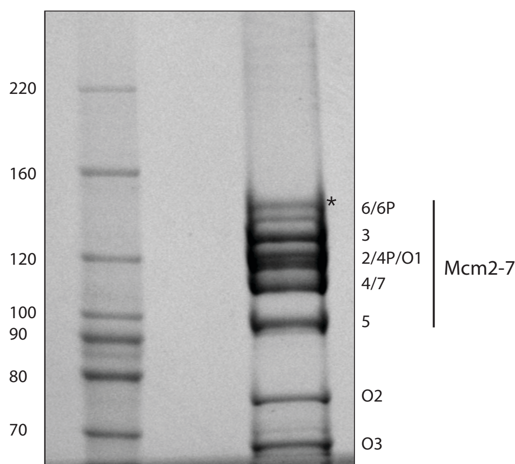


Figure 5.7 Preparation of phosphorylated Mcm2-7 for mass spectrometry analysis

40 x 40 μl pre-RC assembly assays were prepared and each sample was phosphorylated with DDK. After elution from DNA beads via photocleavage, samples were TCA precipitated and separated by SDS-PAGE, then stained with coomassie. Bands corresponding to Mcm6 (*) were excised from the gel and digested by trypsin or AspN. Positions of known proteins are annotated to the right (O refers to ORC subunits). Molecular weight markers are shown on the left.

The results of the mass spectrometry analysis are summarised in table 5-2. All but one of the phosphorylation sites detected clustered in two *S. cerevisiae* specific regions of Mcm6 (Fig. 5.8), as was observed in Randell, 2010. 11 potential sites

were localised to the extreme N-terminus of the protein, which is highly Ser/Thr rich and is predicted to be disordered in secondary structure prediction (Fig. 5.8). In addition to this, four phosphorylated sites were detected between residues 226 – 251, again in a region of Mcm6 that is not conserved in other species.

```
                  .........10 .........20 .........30 .........40 .........50
(PRED) Mcm6Hs   ---------- ---------- ---------- ---------- ----------
(PRED) Mcm6Mm   ---------- ---------- ---------- ---------- ----------
(PRED) Mcm6Dr   ---------- ---------- ---------- ---------- ----------
(PRED) Mcm6Xl   ---------- ---------- ---------- ---------- ----------
(PRED) Mcm6Dm   ---------- ---------- ---------- ---------- ----------
(PRED) Mcm6Sc   MSSPFPADTP SSNRPSNSSP PPSSIGAGFG SSSGLDSQIG SRLHFPSSSQ
(PRED) Mcm6Sp   ---------- ---------- ---------- --MSSLASQG NNASTPAY-R

                  .........60 .........70 .........80 .........90 .........100
(PRED) Mcm6Hs   ---------- ---------- ---------- ----MDLAAA AEPGAGSQHL
(PRED) Mcm6Mm   ---------- ---------- ---------- ----MDLAAA AEPGAGSQHP
(PRED) Mcm6Dr   ---------- ---------- ---------- ---------- MEPGVEAAGV
(PRED) Mcm6Xl   ---------- ---------- ---------- ----MDLVDP SQSAAAAAGT
(PRED) Mcm6Dm   ---------- ---------- ---------- ----MDVADA QVGQLR----
(PRED) Mcm6Sc   PHVSNSQTGP FVNDSTQFSS QRLQTDGSAT NDMEGNEPAR SFKSRALNHV
(PRED) Mcm6Sp   YGFQTSEVGD RPTVSM---- --PSQPESSA MLEDDGMVKR KPFAAIGEAI

                  .........110 .........120 .........130 .........140.........150
(PRED) Mcm6Hs   E-VRDEVAEK CQKLFLDFLE EFQ------- ----SSDGEI KYLQLAEELI
(PRED) Mcm6Mm   E-VRDEVAEK CQKLFLDFLE EFQ------- ----GSDGEI KYLQFAEELI
(PRED) Mcm6Dr   H-VRDELSEK CQKLFLEFLE EFQ------- ----DKNGDA LYLSDAQELI
(PRED) Mcm6Xl   QLVKDEVAEK CQKLFQDFLE EFR------- ----GSDGEL KYQSDAEELI
(PRED) Mcm6Dm   --VKDEVGIR AQKLFQDFLE EFK------- -----EDGEI KYTRPAASLE
(PRED) Mcm6Sc   KKVDDVTGEK VREAFEQFLE DFSVQ----- -STDTGEVEK VYRAQIEFMK
(PRED) Mcm6Sp   PKVIDTTGES VREAFEEFLL SFSDDRVAGG DALPSASQEK YYVQQIHGLA

                  .........160 .........170 .........180 .........190.........200
(PRED) Mcm6Hs   RPERNTLVVS FVDLEQFN-Q QLSTTIQEEF YRVYPYLCRA LKTFVKDRKE
(PRED) Mcm6Mm   RPERNTLVVS FADLEQFN-Q QLSTTIQEEF YRVYPYLCRA LKTFVKDRKE
(PRED) Mcm6Dr   RPERNTLTVS FSNIEHYN-Q QLATTIQEEY YRVYPFLCRA VRHFARDHGN
(PRED) Mcm6Xl   RPERNTLLVS FVDLEQFN-Q QLATTIQEEF YRVYPYLCRA VKAFARDHGN
(PRED) Mcm6Dm   SPDRCTLEVS FEDVEKYD-Q NLATAIIEEY YHIYPFLCQS VSNYVKDRIG
(PRED) Mcm6Sc   IYDLNTIYID YQHLSMRENG ALAMAISEQY YRFLPFLQKG LRRVVRKYAP
(PRED) Mcm6Sp   MYEIHTVYVD YKHLTSY-ND VLALAIVEQY YRFSPFLLRA LQKLIEKFEP

                  .........210 .........220 .........230 .........240.........250
(PRED) Mcm6Hs   IPLA------ ---------- ---------- ---------- ----------
(PRED) Mcm6Mm   IPFA------ ---------- ---------- ---------- ----------
(PRED) Mcm6Dr   IPPS------ ---------- ---------- ---------- ----------
(PRED) Mcm6Xl   IPQN------ ---------- ---------- ---------- ----------
(PRED) Mcm6Dm   LKTQ------ ---------- ---------- ---------- ----------
(PRED) Mcm6Sc   ELLNTSDSLK RSEGDEGQAD EDEQQDDDMN GSSLPRDSGS SAAPGNGTSA
(PRED) Mcm6Sp   EYYRSSLSRE NASLSPNFKA SD-------- ---------- ----------

                  .........260 .........270 .........280 .........290.........300
(PRED) Mcm6Hs   ---------- ------KDFY VAFQDLPTRH KIRELTSSRI GLLTRISGQV
(PRED) Mcm6Mm   ---------- ------KDFY VAFQDLPTRH KIRELTSSRI GLLTRISGQV
(PRED) Mcm6Dr   ---------- ------KEFY VAFSEFPSRQ KIRELSTVRI GTLLRISGQV
(PRED) Mcm6Xl   ---------- ------KEFY VAFQELPTRH KIRELTTPRI GSLLRISGQV
(PRED) Mcm6Dm   ---------- ------KDCY VAFTEVPTRH KVRDLTTSKI GTLIRISGQV
(PRED) Mcm6Sc   MATRSITTST SPEQTERVFQ ISFFNLPTVH RIRDIRSEKI GSLLSISGTV
(PRED) Mcm6Sp   ---------- ------KTFA LAFYNLPFRS TIRDLRTDRI GRLTTITGTV
```

Figure 5.8 Summary of phosphorylation sites identified in Mcm6

Secondary structure prediction for Mcm6 residues 1-294. Phosphorylated residues identified by mass spectrometry are shown in yellow. The prediction was performed using DPSS (Kabsch and Sander, 1983) and PSIPRED (McGuffin et al., 2000). Mcm6 from various eukaryotic species is included (Sc, *Saccharomyces cerevisiae*, Dm, *Drosophila melanogaster*, Dr, *Danio rerio,* Hs, *Homo sapiens*, Mm, *Mus musculus*, Sp, *Schizosaccharomyces pombe*, Xl, *Xenopus laevis*). Predicted α-helices are shown in red and β-strands are shown in blue.

Peptide Sequence	Protease Used	Modified Residues	Site Probabilities
DTPSSNRPSNSSPPPSSIGA GFGSSSGL	AspN	S31, S32, S33	S31 (49.8%) S32 (49.8%) S33 (0.5%)
DSQIGSRLHFPSSSQPHVSN SQTGPFVN	AspN	S37, S48, S49, S56	S37 (25%) S48 (25%) S49 (25%) S56 (25%)
DSTQFSSQRLQT	AspN	S69, S70	S69 (50%) S70 (50%)
DSTQFSSQRLQTDGSATN	AspN	S78, T80	S78 (50%) T80 (50%)
DDMNGSSLPR	AspN	S226, S227	S226 (50%) S227 (50%)
DDGTGSGVITSEPPA	AspN	T870	T870
SITTSTSPEQTER	Trypsin	S249, T251	S249 (50%) T251 (50%)

Table 5-2 Identification of phosphorylation sites in Mcm6

Table showing Mcm6 peptides identified as being modified with one or more phosphate. The peptide sequence is shown in the first column, with modified residues coloured red. Two different proteases were used for proteolysis, as indicated in column 2. Where the sites of phosphorylation were ambiguous, the probabilities for each site are shown in column 4.

5.2.5 The N-terminus of Mcm6 is required for Sld3/7 binding

Thus far, I have presented evidence that the extreme N-terminus of Mcm6 is not only positioned within the region of Mcm6 that is sufficient for Sld3/7 binding, but is also highly phosphorylated and contains a number of phosphopeptides that bind Sld3 directly. Given these observations, I next set out to test the idea that the extreme N-terminus of Mcm6 is an Sld3/7 binding site. To do this, I generated a truncated form of Mcm6 in which the first 84 residues had been deleted. An Mcm2-7/Cdt1 complex containing this truncated version of Mcm6 (6ΔN) was purified after overexpression in yeast (Fig. 5.9A), and was indistinguishable from a wild type

Mcm2-7/Cdt1 complex in its ability to be assembled into the pre-RC *in vitro* (Fig. 5.9B). Strikingly, Mcm6 lacking the N-terminal 84 amino acids showed a distinct defect in DDK-dependent Sld3/7 binding compared with the full-length protein (Fig. 5.9C, compare lanes 12-15 with 17-20). This binding defect could be partially suppressed under high concentrations of DDK (Fig. 5.9C, lanes 19 and 20), indicating that Sld3/7 can bind to Mcm6 lacking the N-terminus, albeit at much reduced levels compared to the wild type. Notably, when using a concentration of DDK under which Sld3/7 cannot bind to Mcm6ΔN (50 nM), deletion of the N-terminus of Mcm6 was sufficient to significantly disrupt recruitment of Sld3/7 to the pre-RC (Fig. 5.9D), consistent with the Mcm6-Sld3/7 interaction being important for binding of Sld3/7 to the Mcm2-7 double hexamer.

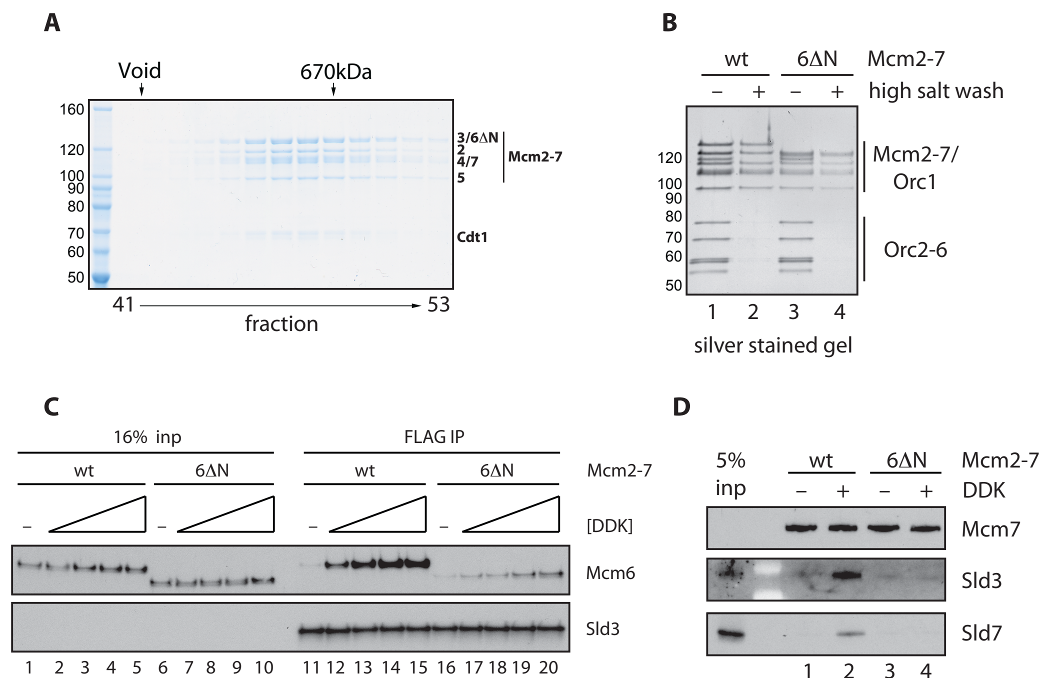


Figure 5.9 The N-terminus of Mcm6 is required for Sld3/7 binding

A. Purified Mcm2-7/Cdt1 complex lacking the first 84 residues of Mcm6 (6ΔN) visualised by SDS-PAGE and coomassie staining. The final gel filtration step of the purification is shown, with the elution positions of molecular weight markers indicated. **B**. Mcm2-7 loading performed with wild type or mutant (6ΔN) Mcm2-7. High salt wash was included as indicated. Proteins were visualised by silver staining. **C**. Sld3/7 binding assay performed with wild type Mcm6 or Mcm6ΔN across a range of DDK concentrations. DDK was included at 0, 25, 50, 100 and 200 nM. Proteins were eluted from beads by boiling in Laemmli sample buffer, and visualised by immunoblot. **D**. Sld3/7 recruitment assay performed using loaded Mcm2-7 complexes containing full-length Mcm6 or Mcm6ΔN. DDK was included at 50 nM as indicated. Samples were analysed by immunoblot.

5.2.6 Mapping a second Sld3/7 binding site on Mcm6

The observation that Sld3/7 can still bind weakly to Mcm6ΔN implied that phosphorylation sites other than those located in the first 84 amino acids of Mcm6 may be involved in Sld3/7 binding. To address this point, I returned to the phosphopeptide array screen, aiming to test the effect of mutating additional sites (other than those in the extreme N-terminus of Mcm6) on the Sld3/7-Mcm6 interaction.

To do this, I generated a yeast strain that expressed an Mcm6 allele, which both lacked the first 84 residues and also had 11 Ser/Thr (all of which interacted with Sld3 on the peptide array – see table 5-1) mutated to Ala between residues 237-486 (6ΔN + 11S/T-A). Further strains were made in which these 11 S/T-A mutations were combined with mutation of Thr 150 to Ala (T150A + 11S/T-A), as this site interacted with Sld3 in the peptide binding analysis (Fig. 5.6A and D), and in which deletion of the N-terminus was combined with mutation of three selected sites (Ser 255, Thr 259 and Thr 272) that interacted with Sld3 when phosphorylated in 30-residue peptides (6ΔN + 3S/T-A) (Fig. 5.6). A final mutant was analysed in which both the *S. cerevisiae* specific regions of Mcm6 that were shown to be phosphorylated in the mass spectrometry analysis (Fig. 5.8) were deleted from the protein (6ΔN + Δ199-260).

Upon testing the Mcm6 mutants for interaction with Sld3/7, we again observed significantly less Sld3/7 binding in the absence of the Mcm6 N-terminus (Fig. 5.10B). The small amount of DDK-dependent Sld3/7 binding recovered with 6ΔN under high DDK concentrations was lost in the combined 6ΔN + 11S/T-A mutant (Fig. 5.10B, lanes 20, 21), indicating that some or all of the phosphopeptide interactions detected in the peptide array are involved in Mcm6 binding to Sld3/7. The loss of Mcm6 binding in these mutants had no effect on the amount of Mcm4 recovered (Fig. 5.10B), indicating that the DDK-dependent Sld3/7-Mcm4 interaction observed is not caused by complex formation between Mcm4 and Mcm6.

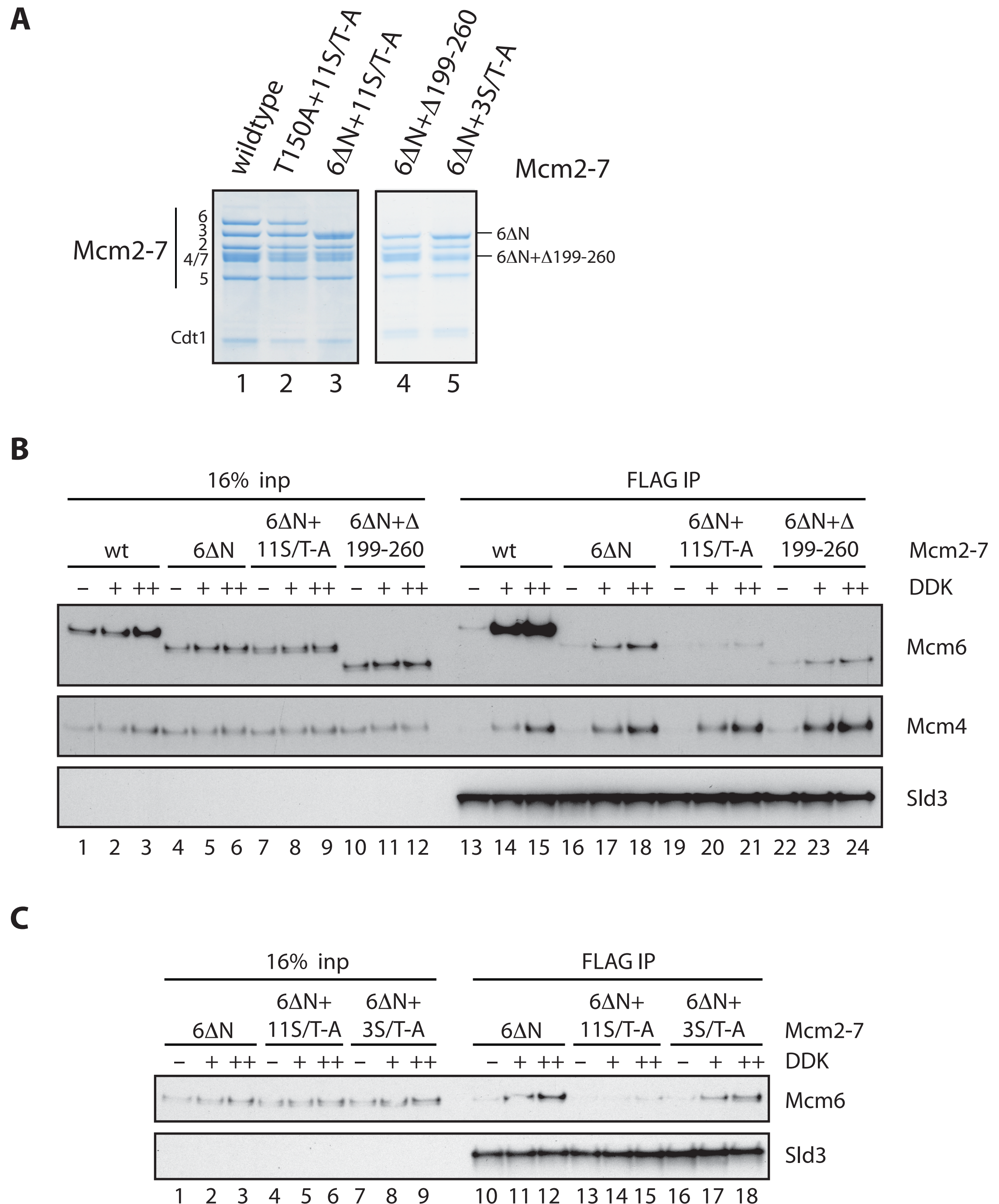


Figure 5.10 The Sld3/7-Mcm6 interaction is disrupted by mutation of multiple distinct sites in Mcm6

A. Purified Mcm2-7/Cdt1 mutants visualised by SDS-PAGE and coomassie staining. The presence of mutant Mcm6 is indicated. **B** and **C**. Sld3/7 binding assay performed with wild type or mutant Mcm6 across a range of DDK concentrations. DDK was included at 0, 50 (+) and 250 (++) nM. Proteins were eluted from beads by boiling in Laemmli sample buffer, and visualised by immunoblot against Mcm4, Mcm6 and Sld3.

The additive effect on Sld3/7 binding was not observed upon combining mutation of Ser 255, Thr 259 and Thr 272 with an N-terminal deletion of Mcm6 (6ΔN + 3S/T-A) (Fig. 5.10C), suggesting that some or all of the eight sites (Ser 247, Ser 324, Ser 351, Thr 376, Thr 398, Thr 423, Ser 475, Thr 480) in the 11 S/T-A mutant that did not bind Sld3 in the context of 30-residue peptides (Fig. 5.6A and B) are likely to be involved in the interaction between Sld3/7 and Mcm6. The lack of Sld3 binding observed for these sites in the phosphopeptide pulldown experiments may have been because the peptides folded in such a way as to occlude the phosphorylated Ser/Thr. Combined deletion of Mcm6 residues 2-84 and 199-260 was also not sufficient to disrupt Sld3/7 binding to the same extent as the 6ΔN + 11S/T-A mutant (Fig. 5.10B, compare lanes 21 and 24), indicating that phosphorylation sites other than those detected in the mass spectrometry analysis must be involved in the Sld3/7-Mcm6 interaction in this *in vitro* assay.

5.2.7 Sld3/7 binding mutants in Mcm6 are defective for DNA replication

Having isolated a number of Mcm6 mutants that prevented Sld3/7 binding, I next wanted to examine the functional consequences of disrupting this interaction on DNA replication. To do this, I utilised the *in vitro* replication assay introduced in the previous chapter. As a control for the activity of the Mcm2-7 complex in the presence of mutant Mcm6, I first tested the T150A + 11S/T-A and 6ΔN + 11S/T-A Mcm2-7 complexes for pre-RC assembly *in vitro* (Fig. 5.11A). The extensive mutation of the N-terminal portion of Mcm6 did not affect the activity of the Mcm2-7 complex in this assay, as all mutants could be loaded onto DNA in a salt resistant manner (Fig. 5.11A, lanes 2, 4, 6 and 8).

Loaded Mcm2-7 complexes containing mutant Mcm6 were then tested for their ability to support DNA replication in an S-phase extract *in vitro* (Fig. 5.11B). Notably, all the mutants tested showed reduced replication activity compared with wild type Mcm2-7 (Fig. 5.11B and C). Deletion of the Mcm6 N-terminus was alone sufficient to reduce replication activity by 30-40% (Fig. 5.11C). Although, the T150A + 11S/T-A mutant alone showed mild defects in DNA replication (Fig. 5.11B, lanes 8 and 9), the combined 6ΔN + 11S/T-A mutant did not show significantly reduced activity

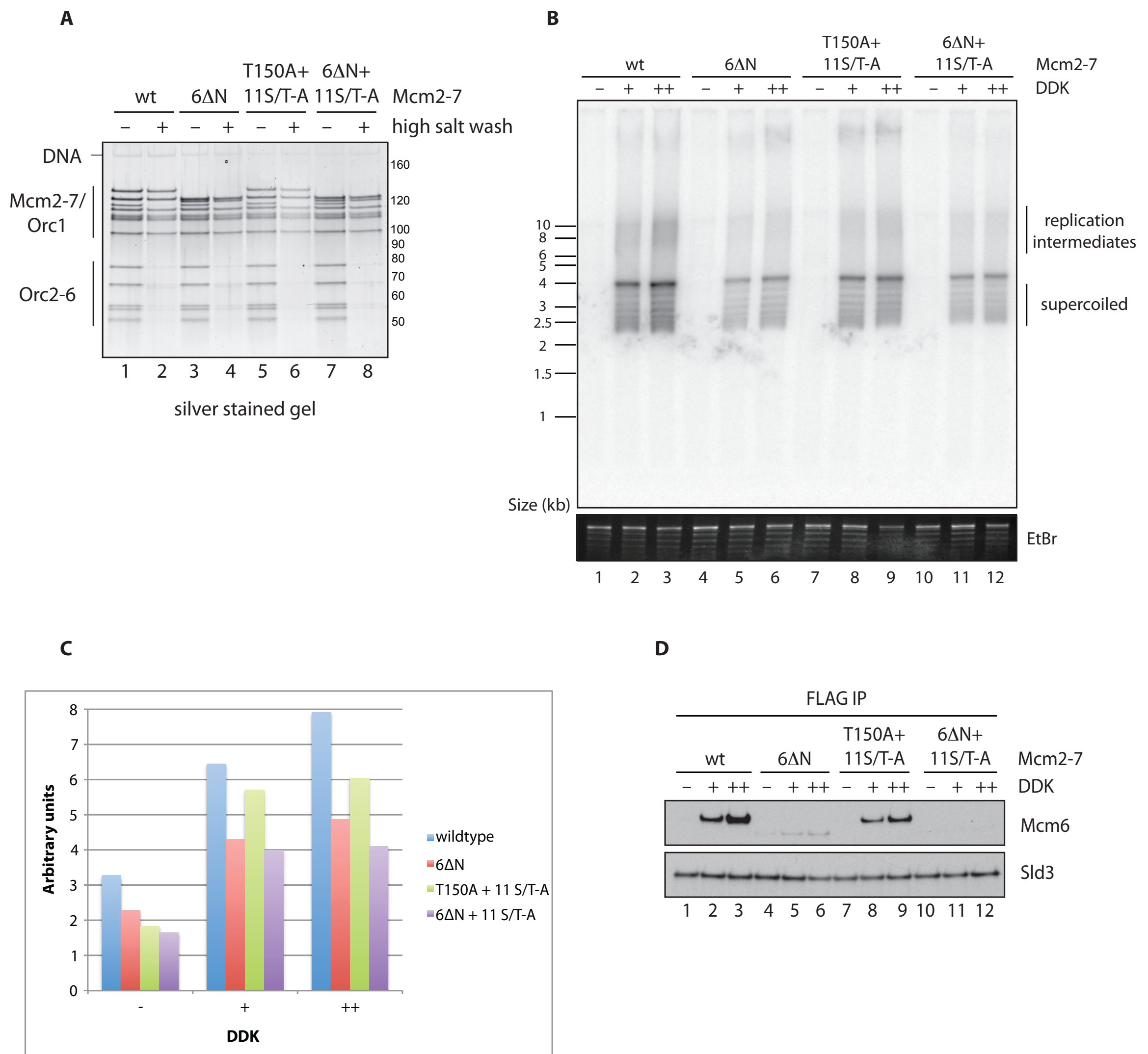


Figure 5.11 The interaction between Mcm6 and Sld3/7 is required for efficient DNA replication initiation *in vitro*

A. Mcm2-7 loading assay performed with wild type or mutant Mcm2-7, as indicated. Proteins were analysed by silver staining. **B**. *In vitro* DNA replication assay. Wild type or mutant Mcm2-7 was loaded onto a 3.2 kb plasmid and tested for ability to support DNA replication in an S-phase protein extract. DDK was included at concentrations of 0, 33 (+) and 133 (++) nM. DNA synthesis was monitored by incorporation of ^{32}P-α-dCTP and samples were analysed by gel electrophoresis and autoradiography. The identity of DNA species is shown on the right. Ethidium bromide staining of the gel is shown as a loading control. **C**. Quantification of B, performed using ImageJ software. **D**. Sld3/7 binding assay performed with wild type or mutant Mcm6. DDK was included at 0, 50 (+) and 250 (++) nM. Bead-bound proteins were eluted by boiling in Laemmli sample buffer and visualised by SDS-PAGE and immunoblotting.

relative to the 6ΔN complex (Fig. 5.11B, compare purple and red bars). The pattern of DNA replication activity exhibited in this assay broadly correlates with the Sld3/7 binding activity of these mutants (Fig. 5.11D), with the T150A + 11S/T-A mutant showing a mild Sld3/7 binding defect (Fig. 5.11D, compare lanes 3 and 9), and 6ΔN and 6ΔN + 11S/T-A showing stronger effects, as discussed previously. Whilst these data do suggest that DDK-dependent binding of Sld3/7 to Mcm6 is important for efficient DNA replication initiation, it is clear that disrupting this interaction alone is not terminal for Mcm2-7 helicase activation. The partial replication defects observed in the absence of the extreme N-terminus of Mcm6 are consistent with previously reported data (Randell et al., 2010), which showed that an Mcm6 mutant containing 22 S/T-A mutations in this region exhibited a slow growth phenotype *in vivo*.

5.3 Discussion

5.3.1 The extreme N-terminus of Mcm6 is an Sld3/7 binding site

In the first part of this chapter, I showed that Sld3 can interact directly with the N-terminal half of phosphorylated Mcm6 (Fig. 5.4C). Subsequently, I used a phosphopeptide array screen to show that Sld3 can bind to multiple phosphorylated Ser/Thr throughout this region (Fig. 5.5), uncovering a novel phosphopeptide binding activity in Sld3. Of the Sld3-bound phosphopeptides detected in the peptide array analysis, four were located within the first 84 amino acids of Mcm6 (Table 5-1), and consistent with previously published data (Randell et al., 2010), multiple phosphorylation sites were detected in this region by mass spectrometry (Table 5-2). Furthermore, deletion of the residues 1-84 of Mcm6 was sufficient to disrupt Sld3/7 binding to both Mcm6 alone and the Mcm2-7 double hexamer (Fig. 5.9C and D). Taken together, these data indicate that the extreme N-terminus of Mcm6 represents a *bona fide* DDK-dependent binding site for Sld3/7 in the pre-RC.

I am yet to determine if the phosphorylation-dependence of this interaction can be attributed to a small subset of phosphorylated Ser/Thr within the N-terminus of Mcm6. It is possible that mutation of Ser 37, Ser 41, Ser 65 and Ser 66 alone, all of which interacted with Sld3 as phosphopeptides (Table 5-1), would be sufficient to

disrupt the Mcm6-Sld3 interaction. Alternatively, the highly phosphorylated nature of the Mcm6 N-terminus may allow Sld3 to interact with Mcm6 via a large number of redundant phosphopeptides. Whilst sufficient to show that Sld3 is a phosphopeptide binding protein, one significant limitation of the peptide array analysis was the lack of any doubly phosphorylated peptides, especially given that DDK has previously been reported to have a preference for sites contain two adjacent Ser/Thr *in vitro* (Randell et al., 2010). Indeed, it is possible that Sld3 can interact with many more phosphopeptides other than the four sites mentioned above, but only when two neighbouring Ser/Thr are modified.

Consistent with the presence of multiple phosphorylation sites being a functional feature of the Mcm6 N-terminus, previous work indicates that only the wholesale elimination of 22 phosphorylation sites from this domain can produce a deleterious phenotype *in vivo* (Randell et al., 2010). Any mutant tested that contained less than this number of mutations was indistinguishable from the wild type. In addition to this, data in the previous chapter showed that mutation of six separate Arg/Lys in Sld3 residues 511-531 can disrupt binding to the pre-RC (Fig. 4.4A), consistent with the presence of multiple phospho-binding sites within this small section of Sld3. If Sld3 does indeed bind to Mcm6 via multiple redundant phosphorylation sites, this would potentially produce an interaction of high avidity (discussed further in section 6.5). High specificity might be achieved through some other means, such as strict requirements for the relative conformations of the Sld3/7 and Mcm2-7 complexes during Sld3/7 recruitment.

5.3.2 Implications for the function of DDK during origin firing

Although sufficient to significantly reduce the recruitment of Sld3/7 to the pre-RC, deletion of the N-terminal 84 residues of Mcm6 resulted in only a partial loss of activity in an *in vitro* replication assay (Fig. 5.11B). By isolating the 6ΔN + 11S/T-A mutant, which was entirely defective for Sld3/7 binding even under high DDK concentrations (Fig. 5.10B), I was able to test the hypothesis that the residual replication activity observed for the 6ΔN Mcm2-7 complex was attributable to the small amount of Sld3/7 binding still possible in the absence of the Mcm6 N-

terminus (Fig. 5.9C). Notably, the 6ΔN + 11S/T-A mutant was still able to support some replication in an S-phase extract (Fig. 5.11B), indicating that Sld3/7 binding to Mcm6 *per se* is not essential for Mcm2-7 helicase activation. Given the fact that DDK *is* essential for DNA replication *in vivo* and *in vitro*, there must therefore be phosphorylation sites elsewhere in the Mcm2-7 complex that are also important for origin firing.

Consistent with this idea, I showed that Sld3/7 can interact directly with phosphorylated Mcm4 (Fig. 5.2C). This interaction was markedly reduced by mutations in residues 511-531 of Sld3 (Fig. 5.2C), consistent with the same site on Sld3 being involved in phosphorylation-dependent binding to both Mcm4 and Mcm6. Although yet to be examined, it is tempting to speculate that an interaction between Sld3 and Mcm4 could allow for some replication initiation in the absence of a cognate Sld3 binding site on Mcm6. Although this putative Sld3-Mcm4 interaction was not sufficient for Sld3/7 recruitment to the pre-RC using purified proteins (Fig. 5.9D), it may be that the less stringent buffer conditions used in the extract-based *in vitro* replication assay allowed for some Sld3/7 recruitment, and therefore some replication initiation, in the absence of the Mcm6 extreme N-terminus.

The notion that Mcm6 is not the only DDK substrate required for DNA replication initiation is consistent with the previously reported slow growth phenotypes for phosphorylation site mutants in the N-termini of Mcm2 and Mcm4 (Randell et al., 2010). One can envisage how phosphorylation of Mcm4 might provide a second binding site on the loaded Mcm2-7 complex for Sld3 511-531, and it is possible that combining the deletion of the N-terminus of Mcm6 with an Mcm4 phosphorylation site mutant would lead to an additive reduction in Sld3/7 recruitment to the pre-RC, thereby preventing replication initiation. Importantly, the combination of Mcm4 and Mcm6 mutants, both of which contain multiple alanine substitutions at DDK phosphorylation sites, has previously been shown to be incapable of supporting cell growth (Randell et al., 2010). It is pertinent to note that the Sld3-Mcm2-7 interaction itself is essential for replication initiation (Fig. 4.6C), and it should thus be possible to isolate an Sld3-binding mutant in Mcm2-7 that causes an equally deleterious phenotype as that observed for the Sld3 mutants discussed in the previous chapter.

It is possible to bypass the essential function of DDK in replication initiation by deletion of an internal domain of Mcm4 (Sheu and Stillman, 2010) or with a single point mutation in Mcm5 (Hardy et al., 1997), as previously mentioned. Given the involvement of the N-terminus of Mcm6 in DDK-dependent Sld3/7 recruitment to the pre-RC (Fig. 5.9D), it is difficult to predict how these DDK bypass mutants function for proper Mcm2-7 helicase activation. One possibility is that the Mcm2-7 double hexamers produced in these mutants have a higher than normal affinity for some firing factor, such as Cdc45, which is normally recruited downstream of Sld3/7 binding to Mcm2-7. Thus, rather than bypassing the requirement for DDK for the Sld3/7-Mcm2-7 interaction, these mutants might actually bypass the requirement for Sld3 itself during replication initiation, thereby allowing for formation of the active CMG complex even in the absence of DDK-dependent Sld3/7 recruitment to the pre-RC.

Chapter 6. Discussion

6.1 Introduction

During the origin licensing step of DNA replication initiation, the Mcm2-7 replicative helicase is loaded onto DNA at origins of replication, forming a pre-RC (Blow, 1993, Blow and Laskey, 1988, Diffley et al., 1994). The product of this pre-RC assembly step is a double hexamer of Mcm2-7 bound around dsDNA (Remus et al., 2009, Evrin et al., 2009, Gambus et al., 2011), in which the Mcm2-7 complex is inactive as a helicase. Upon progression into S-phase, the replicative helicase is activated by the combined activities of two protein kinases, DDK and CDK, and a whole host of specific firing factors (Boos et al., 2012). In *S. cerevisiae*, the known firing factors are Sld3, Sld7, Cdc45, Dpb11, Sld2, Polε, GINS and Mcm10, and these are believed to assemble into a transient pre-initiation complex (pre-IC) on the pre-RC (Tanaka and Araki, 2013, Zou and Stillman, 1998), leading to extensive remodelling of the Mcm2-7 replicative helicase into its active form, the CMG complex. The CMG complex itself contains a single copy each of Mcm2-7, Cdc45 and the tetrameric GINS complex (Gambus et al., 2006, Costa et al., 2011, Fu et al., 2011), and forms the basis for the assembly of the Replisome Progression Complex (RPC) and the establishment of two divergent replication forks from a single Mcm2-7 double hexamer. There is currently little understanding of the precise mechanism of replicative helicase activation, or how individual firing factors contribute towards this process.

6.2 Key Conclusions

In this study, I adopted a primarily biochemical approach to attempt to dissect the mechanism of Mcm2-7 helicase activation, principally focussing on the function of DDK and the Sld3/7 complex during this process. I purified the Sld3/7 complex after overexpression in yeast (Fig. 3.5), and was subsequently able to reconstitute the complex from individually purified Sld3 and Sld7 (Fig. 3.8), obtaining evidence that Sld3/7 is likely to contain at least two copies of Sld3 in the process (Fig. 3.6 and 3.8). Given its position upstream of the recruitment of all other known firing

factors (Yabuuchi et al., 2006, Heller et al., 2011), the stoichiometry of the Sld3/7 complex could have considerable implications for the mechanism of pre-IC assembly (see section 6.6).

Previous studies have indicated a function for DDK in promoting the recruitment of Sld3/7 to replication origins (Yabuuchi et al., 2006, Heller et al., 2011, Tanaka et al., 2011a), which is thought to be one of the earliest events during conversion of the pre-RC into two RPCs. I was able to reconstitute the recruitment of Sld3/7 to the pre-RC *in vitro* (Fig. 3.11), and discovered a requirement for prior phosphorylation of the loaded Mcm2-7 complex by DDK during this reaction. This is the first example of a DDK-regulated protein-protein interaction, and implies a novel function for both Sld3/7 and DDK during replication initiation.

I subsequently mapped the Mcm2-7 binding activity of Sld3/7 to the central portion of Sld3 (Fig. 4.2), and isolated a number of amino acid substitution mutants in Sld3 that specifically disrupted recruitment to the pre-RC (Fig. 4.4). Addition of these mutant proteins to an S-phase extract resulted in severe defects in DNA replication initiation (Fig. 4.6), likely arising from deficient assembly of the CMG complex in the absence of the Mcm2-7 binding activity of Sld3 (Fig. 4.7). Notably, Mcm2-7 binding mutants of Sld3 were also unable to support cell viability (Fig. 4.9). During the course of mutagenising the central portion of Sld3, I also isolated a number of mutants that were specifically defective in Cdc45 binding (Fig. 4.12), consistent with previous mapping of the Cdc45 interaction site on Sld3 (Tanaka et al., 2011b). These mutants also exhibited replication defects *in vitro* (Fig. 4.13), indicating that the central portion of Sld3 is functional for both Mcm2-7 and Cdc45 binding.

The final section of this study was dedicated to understanding the role of DDK in promoting Sld3/7 recruitment. I initially showed that Sld3 can interact with both Mcm6 and Mcm4 in a phosphorylation-dependent manner (Fig. 5.2C), and subsequently discovered multiple phosphopeptides throughout the N-terminal half of Mcm6 to which Sld3 can bind (Fig. 5.5). The phosphopeptide binding activity of Sld3 was suggestive of a direct interaction with phosphorylated Ser/Thr on Mcm2-7 during Sld3/7 recruitment. This notion was strengthened by the observation that the binding of Sld3/7 to the pre-RC was largely dependent on the heavily

phosphorylated extreme N-terminus of Mcm6 (Fig. 5.9D), a domain which itself contains a number of phosphopeptides to which Sld3 can bind (Fig. 5.5A). Notably, deletion of this section of Mcm6 only resulted in a partial defect in DNA replication initiation *in vitro* (Fig. 5.11), consistent with the previously reported *in vivo* phenotype of an Mcm6 mutant lacking DDK phosphorylation sites in this domain (Randell et al., 2010). This observation implies that Sld3/7 may be able to interact with Mcm subunits other than Mcm6 in the pre-RC, and that any such interactions are sufficient to support a sub-optimal level of CMG formation and origin firing.

6.3 The function of Sld3 during origin firing

The requirement for Sld3 during DNA replication initiation was first revealed over 10 years ago and, since then, Sld3 has been shown to fulfil important functions in both the CDK and Rad53-dependent regulation of origin firing (Tanaka et al., 2007, Zegerman and Diffley, 2007, Zegerman and Diffley, 2010, Lopez-Mosqueda et al., 2010, Kamimura et al., 2001). CDK phosphorylates Sld3 at two conserved sites located near its C-terminus, creating a binding site for a second protein, Dpb11, and this interaction is essential for replication initiation (Tanaka et al., 2007, Zegerman and Diffley, 2007). Additionally, in response to DNA damage, Sld3 is phosphorylated by the Rad53 DNA damage checkpoint kinase at multiple Ser/Thr positioned throughout its C-terminal half. This not only blocks the Sld3-Dpb11 interaction, but also inhibits binding of Sld3 to Cdc45, thereby preventing further origin firing in the presence of DNA damage (Lopez-Mosqueda et al., 2010, Zegerman and Diffley, 2010).

Work presented in this study contributes significantly to our understanding of Sld3 function during Mcm2-7 helicase activation. In chapter 3, I showed that Sld3 binding to Sld7 is dependent on the N-terminus of Sld3 (Fig. 3.8), consistent with previously published data (Tanaka et al., 2011b). Notably, Sld3 multimerisation was also reduced in the absence of the N-terminal 250 residues (Fig. 3.8). Thus, the N-terminal portion of Sld3 appears to be important for interactions within Sld3/7, thereby aiding the establishment of the overall architecture of the complex.

Sld3 residues 251-585, which represent the minimal portion of Sld3 required for recruitment to the pre-RC, contain both Mcm2-7 and Cdc45 binding activities (Fig. 4.4 and 4.12). This central portion of Sld3 includes the Sld3 Treslin Domain (S.T.D.), which is conserved between Sld3 and its higher eukaryotic homologue, Treslin (Sanchez-Pulido et al., 2010). The interaction between Sld3 and Cdc45 was dependent on a highly basic patch of amino acids within the S.T.D. itself (Fig. 4.12), whereas Mcm2-7 binding could be abolished by mutation of six separate basic residues C-terminal of the S.T.D. (Fig. 4.4), all of which are conserved between different fungal species (Fig. 6.1).

Thus, in spite of the absence of any known enzymatic or protein-protein interaction domains, Sld3 is a vital component of the replicative helicase activating machinery. It seems likely that Sld3 functions as a central scaffold protein, interacting with Sld7, Cdc45, Mcm2-7 and Dpb11 during origin firing (Fig. 6.1). Notably, the binding sites for these various proteins are located in distinct sections of Sld3 (Fig. 6.1); whether or not Sld3 can interact with all its binding partners simultaneously, or whether the release of certain factors is co-ordinated with binding to others, represents an interesting avenue for future investigation.

The function of Sld3 as a CDK substrate is conserved in Treslin, and an interaction between Treslin and Cdc45 has also been reported (Boos et al., 2011, Kumagai et al., 2010, Kumagai et al., 2011). Despite its location within the S.T.D., the basic stretch of residues shown to be required for Cdc45 binding in this study is not conserved with Treslin (Sanchez-Pulido et al., 2010); it will be interesting to determine if this binding activity is localised to a different region of the protein in higher eukaryotes.

The Mcm2-7 binding activity of Sld3 is located between the S.T.D and the two essential CDK target sites (Fig. 6.1). Upon performing a multiple sequence alignment of Treslin from various higher metazoan species, it is clear that there are a number of highly conserved residues within this region (Fig. 6.2). Two of the most conserved areas within this region are conspicuous by the presence of a number of basic amino acids, analogous to the Mcm2-7 binding residues described for *S. cerevisiae* Sld3. Any DDK dependency of Treslin recruitment to replication origins

is yet to be examined in higher eukaryotes, but if, as with other Sld3 functions, this activity is conserved between yeast and humans, the regions highlighted in figure 6.2 would undoubtedly represent good candidate Mcm2-7 binding sites.

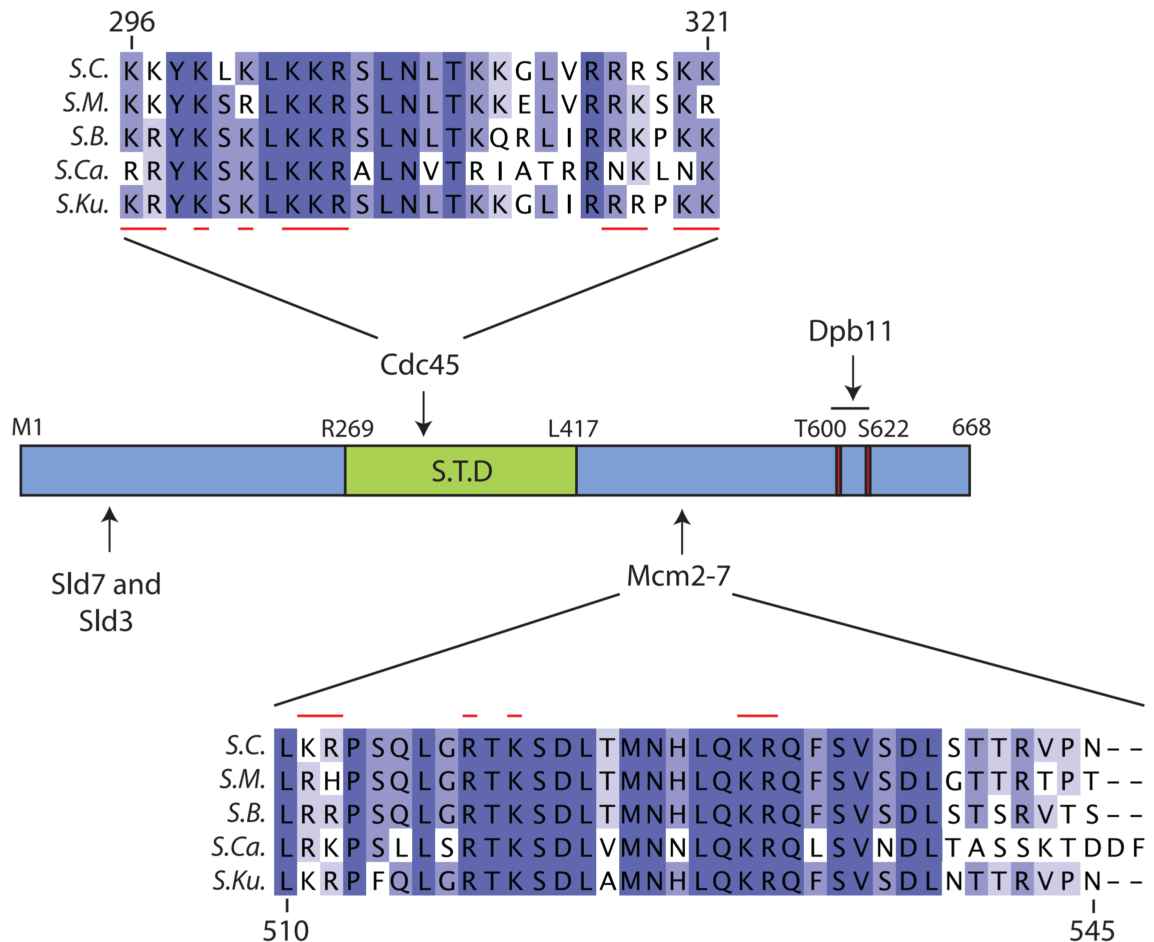


Figure 6.1 Sld3 interacts with multiple partner proteins during replication initiation

Schematic of *S. cerevisiae* Sld3 showing positions of interacting regions for other replication proteins. Two essential CDK target sites are shown as vertical red bars. The conserved Sld3 Treslin Domain (S.T.D.) is shown in green. Sequence alignments, generated in Jalview using Clustal (Sievers et al., 2011), are shown for Cdc45 and Mcm2-7 interacting regions, with residues essential for these interactions marked by red lines. Sld3 from various fungal species is included (S.C., *Saccharomyces cerevisiae*, S.M., *Saccharomyces mikatae,* S.B., *Saccharomyces bayanus,* S.Ca., *Saccharomyces castellii,* S.Ku., *Saccharomyces kudriavzevii*). Residue numbers correspond to *S. cerevisiae* Sld3.

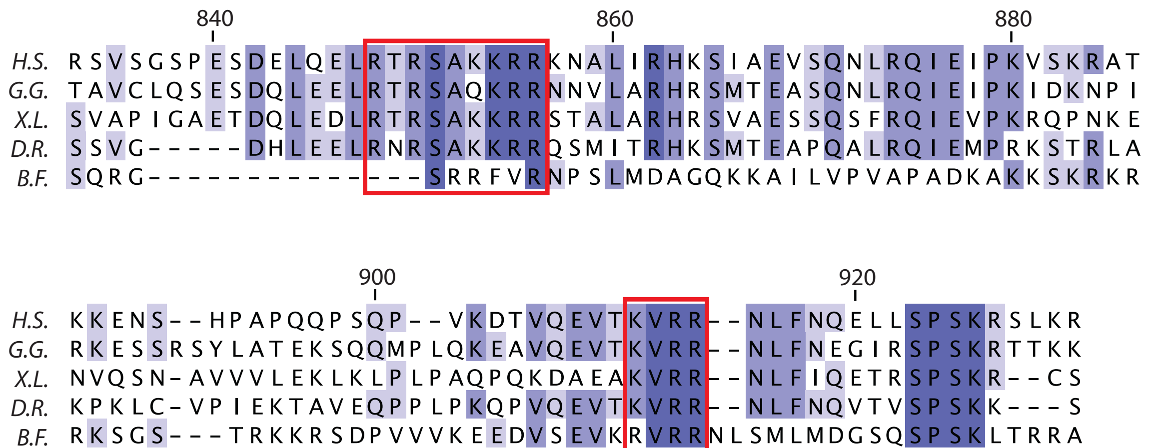


Figure 6.2 Treslin contains conserved basic residues between the S.T.D. and CDK phosphorylation sites

Sequence alignment of Treslin generated in Jalview. For *Homo sapiens* Treslin, the S.T.D. ends at residue 792, and CDK consensus sites are Thr 969 and Ser 1001. Treslin from various eukaryotic species is included (H.S., *Homo sapiens*, G.G., *Gallus gallus*, X.L., *Xenopus laevis*, D.R., *Danio rerio*, B.F., *Branchiostoma floridae*). Residue numbers above the alignment correspond to *Homo sapiens* Treslin. Conserved basic residues are indicated in red boxes.

6.4 Convergence of the DDK and CDK-regulated pathways of replicative helicase activation

The observation that Sld3 is recruited to the pre-RC in a DDK-dependent manner denotes a function for Sld3 as a reader of the phosphorylation ‘status’ of Mcm2-7. The phosphopeptide binding activity ascribed to Sld3 in this study likely indicates that this reader function is direct, and that Sld3 binds directly to phosphorylated peptides within Mcm2-7 (Fig. 5.5 and 5.6).

Phosphorylation of Mcm2-7 by DDK is likely to be the most upstream step during replicative helicase activation, and previous work also indicates that DDK must function before CDK for productive origin firing to occur (Yabuuchi et al., 2006, Jares and Blow, 2000, Heller et al., 2011). Given this, the first step during pre-IC assembly is likely to be DDK-dependent recruitment of unphosphorylated Sld3/7 to the pre-RC, as has been detected at early-firing replication origins as early as G1 phase of the cell cycle (Tanaka et al., 2011a).

Previous work has shown that DDK itself specifically binds to and phosphorylates only those Mcm2-7 complexes that have been assembled into the pre-RC (Randell et al., 2010). This specificity for the DNA-bound form of the replicative helicase was proposed to spatially constrain the helicase activation process, preventing inappropriate pre-IC assembly on Mcm2-7 that is free in solution. The involvement of Sld3, an essential CDK substrate, as a reader of DDK activity could extend this spatial restriction to CDK-regulated processes; Sld3 pre-bound to Mcm2-7 could be phosphorylated by CDK, and formation of the CDK-dependent Sld3-Dpb11-Sld2 complex would thus only occur at the origin of replication. The downstream recruitment of GINS, potentially as part of the pre-loading complex (Muramatsu et al., 2010), would then allow for productive assembly of the CMG replicative helicase and subsequent formation of the Replisome Progression Complex (Gambus et al., 2006). Notably, even if CDK dependent binding of Sld3 to Dpb11 were to occur before the recruitment of Sld3 to the pre-RC, Sld3 would still provide a point of convergence for the CDK and DDK-dependent pathways through subsequent DDK-dependent binding to Mcm2-7 (Fig. 6.3).

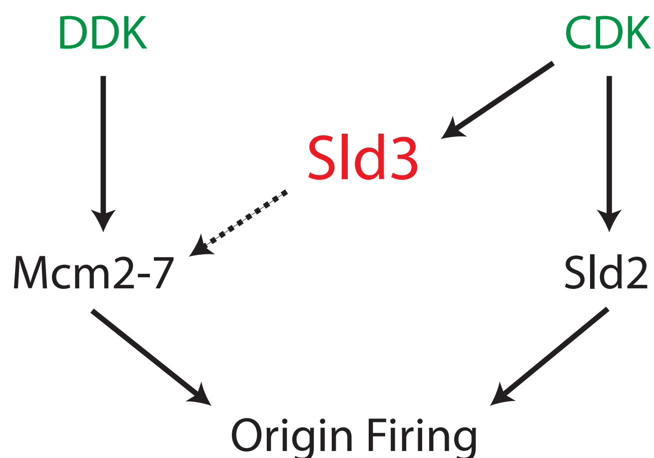


Figure 6.3 DDK and CDK-dependent regulation of origin firing converges at Sld3

Sld2 and Sld3 are the two essential CDK substrates required for origin firing. DDK phosphorylates Mcm2-7, promoting an interaction between Mcm2-7 and Sld3. Sld3, as an essential CDK substrate and a reader of the DDK pathway, thereby allows the integration of CDK and DDK activities during replication initiation.

Previous work in *S. cerevisiae* has indicated that Cdc45 can also be recruited to the pre-RC in a manner that is independent of CDK activity but dependent on DDK (Sheu and Stillman, 2006, Heller et al., 2011, Tanaka et al., 2011a), suggestive of a role for Cdc45 as a second DDK reader. Work presented in this study, however, suggests that, whilst Sld3 alone can bind to Mcm2-7 (Fig. 3.11 and 4.8), DDK-dependent recruitment of Cdc45 is downstream of Sld3/7 (Fig. 4.7). This sequential DDK-dependent recruitment of Sld3/7 followed by Cdc45 to the pre-RC has since been reconstituted with purified proteins (Dr. Joe Yeeles), lending weight to the argument that, of the known firing factors, Sld3 is the only protein that can directly read DDK phosphorylation of Mcm2-7 during replicative helicase activation.

6.5 The function of DDK during origin firing

The work presented in this study indicates that one essential function of DDK during replication initiation is to catalyse the binding of Sld3 to the loaded Mcm2-7 complex. Sld3/7 recruitment is at least partly dependent on the N-terminal tail of Mcm6, which is phosphorylated by DDK *in vitro* (Fig. 5.8 and 5.9). Deletion of this section of Mcm6, however, only partially reduces the ability of Mcm2-7 to support DNA replication (Fig. 5.11B and C), indicating that phosphorylation-dependent interactions with other Mcm subunits must be involved in Sld3/7 recruitment. Consistent with this, Sld3 can also bind to Mcm4 in a phosphorylation-dependent manner (Fig. 5.2C). The precise nature of this interaction, as well as its functional significance, is still to be determined. However, deleterious phenotypes have been reported for mutants of Mcm4 in which DDK phosphorylation sites have been removed (Randell et al., 2010).

The protein kinase activity of DDK is atypical in a number of ways. Firstly, the specificity of DDK towards the Mcm2-7 double hexamer appears to be primarily determined by the conformation of this substrate (Randell et al., 2010). Cdc7 has been reported to bind Mcm4, whilst Dbf4 can interact with Mcm2 (Ramer et al., 2013, Bruck and Kaplan, 2009). The relative position of these Mcm subunits in the loaded Mcm2-7 double hexamer compared with Mcm2-7 that is in solution is one potential explanation for this. Despite its absolute specificity for Mcm2-7 complexes

loaded onto DNA, DDK is somewhat promiscuous *in vitro*, phosphorylating *S. cerevisiae* Mcm2, Mcm4 and Mcm6 at many different sites throughout their N-termini, with only moderate specificity for any single consensus sequence (Randell et al., 2010). Additionally, many of the DDK phosphorylation sites that have been mapped are not conserved between yeast and humans, despite the fact that DDK is absolutely essential for S-phase entry across all eukaryotes.

Notwithstanding the lack of conservation of individual DDK target sites between different species, the N-termini of Mcm2, Mcm4 and Mcm6 in higher eukaryotes are relatively rich in Ser/Thr residues, similar to yeast (Masai et al., 2006). Notably, only the comprehensive elimination of multiple phosphorylation sites from the N-termini of Mcm2, Mcm4 or Mcm6 is sufficient to cause defects in DNA replication in *S. cerevisiae* (Randell et al., 2010). This observation, coupled with the fact that the highly phosphorylated N-terminus of Mcm6 is a *bona fide* Sld3 binding site (Fig. 5.9), suggests that poly-phosphorylation may well be a functionally significant feature of DDK activity.

If this is true, then DDK-catalysed phosphorylation would be operating in a very similar way to another post-translational modification, SUMOylation, during the DNA damage response (Psakhye and Jentsch, 2012). In that situation, DNA damage leads to a wave of SUMOylation of multiple DNA repair proteins at many different sites. Just as DDK specifically targets a conformationally unique form of Mcm2-7 in the pre-RC, specificity in the SUMO-dependent DNA damage response is achieved by the arrangement of DNA repair proteins and the SUMO modifying enzymes around the site of damage. The numerous SUMOylated sites created during this response are said to act as a ‘glue’, stabilising multiple protein-protein contacts, and producing interactions of high avidity.

Although it was previously hypothesised that extensive modification of the N-termini of Mcm2, Mcm4 and Mcm6 might promote separation of the Mcm2-7 double hexamer during origin firing (Remus and Diffley, 2009), this has since been shown to not be the case (On et al., 2014). Alternatively, this phosphorylation may generate multiple redundant phosphopeptide binding sites for Sld3 on Mcm2-7. Indeed, Sld3 itself contains multiple Mcm2-7 binding residues (Fig. 4.4), and there

are multiple Sld3 molecules within the Sld3/7 complex (Fig. 3.8), consistent with a mode of binding involving numerous phosphorylated Ser/Thr. Although further investigation of the interactions between Sld3 and Mcm4/Mcm6 is required to prove the existence of this system, one can envisage how such a mechanism would produce an additive effect on the affinity of the final complex. Thus, a stable Sld3/7-Mcm2-7 complex could form as the first step during pre-IC assembly, which in turn would facilitate the downstream recruitment of other firing factors, whilst reducing the possibility of any non-productive release of Sld3 from the origin before Mcm2-7 helicase activation has occurred.

6.6 A model for replicative helicase activation

During activation of the replicative helicase, a single Mcm2-7 double hexamer must be converted into a pair of bi-directional replication forks, each containing a single CMG complex made up of one copy each of Mcm2-7, Cdc45 and GINS (Gambus et al., 2006, Costa et al., 2011, Fu et al., 2011). Two molecules each of Cdc45 and the GINS tetramer must therefore be recruited to each double hexamer, which imposes strict stoichiometric constraints on the recruitment of other firing factors.

Notably, the recruitment of both Cdc45 and GINS to the pre-RC is dependent on Sld3 (Fig. 4.7A and Heller et al., 2011). Whilst the initial recruitment of Cdc45 is CDK-independent and likely occurs via a direct interaction with Sld3 (Kanemaki and Labib, 2006, Heller et al., 2011, Gros et al., 2014), GINS recruitment requires both the activity of CDK and the presence of Dpb11 and Sld2 (Kanemaki and Labib, 2006, Heller et al., 2011, Takayama et al., 2003). Given that GINS is a component of the pre-loading complex (pre-LC), which also contains Dpb11 and Sld2 (Muramatsu et al., 2010), the likeliest explanation for CDK-dependent GINS recruitment is that phosphorylation-dependent binding of Dpb11 to Sld3 facilitates pre-LC recruitment to the pre-RC.

Assuming that each Sld3 molecule can only interact with a single copy of Cdc45, and that each Sld3-Dpb11 interaction is only sufficient to recruit a single GINS tetramer, the simplest way to recruit the two Cdc45 and GINS molecules required

for replicative helicase activation would be to initially recruit two Sld3 molecules to the Mcm2-7 double hexamer. Work presented in this thesis supports the idea that the Sld3/7 contains two molecules of Sld3 (Fig. 3.6 and 3.8). Although possible, the presence of multiple Sld3/7 complexes would potentially allow for the futile recruitment of more than two Cdc45 and/or GINS molecules to the pre-RC. Thus, to fulfil the stoichiometric requirements of the helicase activation process, the first step in pre-IC assembly should be the recruitment of a single Sld3/7 complex to each Mcm2-7 double hexamer.

Sld3 itself can interact with multiple Mcm subunits in a phosphorylation-dependent manner (Fig. 5.2C). Importantly, mutations in residues 511-531 of Sld3 disrupt binding to both Mcm4 and Mcm6 (Fig. 5.2C), and are sufficient to abolish DNA replication activity *in vitro* (Fig. 4.6C). In contrast, Mcm6 mutants, which disrupt Sld3 binding, only cause a partial DNA replication defect (Fig. 5.11B), consistent with Sld3 binding to both Mcm4 and Mcm6 being important during Mcm2-7 activation.

Given these predictions, it is possible to devise a provisional model for assembly of the pre-IC, as depicted in figure 6.4. In this model, a single Sld3/7 molecule is bound to the Mcm2-7 double hexamer, with each Sld3 molecule concurrently interacting with phosphorylated residues at the N-termini of both Mcm4 and Mcm6. Sld7 may connect the two Mcm-bound Sld3 molecules, enabling Sld3/7 to bridge the hexamer-hexamer interface within the double hexamer.

The high-resolution structure of a dimeric form of the CMG complex has recently been determined by electron microscopy (Costa et al., 2014). In this structure, the two Mcm2-7 hexamers are arranged in a similar manner to the Mcm2-7 double hexamer, with their N-terminal regions facing towards each other in a head-to-head arrangement. Although the physiological relevance of this dimeric assembly is unclear, the structure does allow one to visualise the potential relative arrangement of individual Mcm subunits within the Mcm2-7 double hexamer. The two Mcm hexamers within the CMG dimer are staggered relative to one another, positioning Mcm4/6 of one hexamer opposite the Mcm2 and Mcm5 subunits of the other, as is shown in figure 6.4. Importantly, a previously solved structure of the CMG complex

indicates that Cdc45 and GINS contact Mcm2 and Mcm5, respectively (Costa et al., 2011). Given the aforementioned dependence of Cdc45 and GINS recruitment on Sld3/7, one can envisage how an Sld3 molecule bound to Mcm4/6 on one hexamer might direct the recruitment of Cdc45 and GINS to Mcm2/5 on the opposite side. Although Sld3 might also be able recruit Cdc45 to the same Mcm hexamer to which Sld3 itself is bound, we favour the mechanism depicted in figure 6.4**,** as it potentially allows for the co-ordinate formation of two CMG complexes from a single Mcm2-7 double hexamer, thereby facilitating bi-directional replication initiation from the pre-RC.

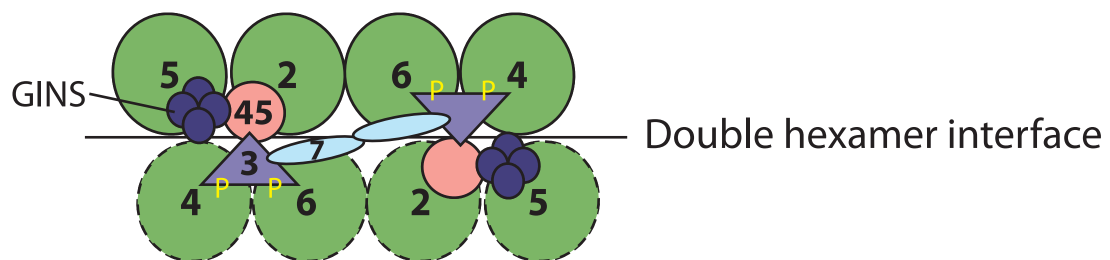


Figure 6.4 A model for Sld3/7-dependent CMG formation

For full details see text in section 6.6. Mcm subunits are shown in green and numbered, and subunits from different Mcm2-7 hexamers are outlined with either complete or dashed lines. GINS, Cdc45 (45), Sld3 (3) and Sld7 (7) are also shown. P represents phosphorylated Ser/Thr on Mcm4/6.

Although speculative at present, this model provides a basic framework within which to develop our understanding of the Mcm2-7 activation process. Work presented herein contributes towards our knowledge of the early DDK-dependent steps of pre-IC assembly, yet we still have little understanding of how Dpb11 and Sld2 direct recruitment of GINS and Polε to the pre-RC. Furthermore, nothing is known about how the recruitment of these various firing factors is co-ordinated with double hexamer separation or the melting of origin DNA, or how complete replisome assembly and DNA synthesis is subsequently initiated. The development of a fully reconstituted system for DNA replication initiation will greatly aid future attempts to further dissect this intriguing process.

Chapter 7. Appendix

7.1 Peptides used in Mcm6 phosphopeptide arrays

7.1.1 Part 1

Nr.	Pos.	Mol. Weight	Sequence
1	A 1	1879.1	M-S-S-P-F-P-A-D-T-P-S-S-N-R-P-S-N-S
2	A 2	2049.1	M-pS-S-P-F-P-A-D-T-P-S-S-N-R-P-S-N-S
3	A 3	2049.1	M-S-pS-P-F-P-A-D-T-P-S-S-N-R-P-S-N-S
4	A 4	2049.2	M-S-S-P-F-P-A-D-pT-P-S-S-N-R-P-S-N-S
5	A 5	2049.1	M-S-S-P-F-P-A-D-T-P-pS-S-N-R-P-S-N-S
6	A 6	2049.1	M-S-S-P-F-P-A-D-T-P-S-pS-N-R-P-S-N-S
7	A 7	2049.1	M-S-S-P-F-P-A-D-T-P-S-S-N-R-P-pS-N-S
8	A 8	2049.1	M-S-S-P-F-P-A-D-T-P-S-S-N-R-P-S-N-pS
9	A 9	1835	S-S-P-F-P-A-D-T-P-S-S-N-R-P-S-N-S-S
10	A10	2005	pS-S-P-F-P-A-D-T-P-S-S-N-R-P-S-N-S-S
11	A11	2005	S-pS-P-F-P-A-D-T-P-S-S-N-R-P-S-N-S-S
12	A12	2005.1	S-S-P-F-P-A-D-pT-P-S-S-N-R-P-S-N-S-S
13	A13	2005	S-S-P-F-P-A-D-T-P-pS-S-N-R-P-S-N-S-S
14	A14	2005	S-S-P-F-P-A-D-T-P-S-pS-N-R-P-S-N-S-S
15	A15	2005	S-S-P-F-P-A-D-T-P-S-S-N-R-P-pS-N-S-S
16	A16	2005	S-S-P-F-P-A-D-T-P-S-S-N-R-P-S-N-pS-S
17	A17	2005	S-S-P-F-P-A-D-T-P-S-S-N-R-P-S-N-S-pS
18	A18	1845	S-P-F-P-A-D-T-P-S-S-N-R-P-S-N-S-S-P
19	A19	2015	pS-P-F-P-A-D-T-P-S-S-N-R-P-S-N-S-S-P
20	A20	2015.1	S-P-F-P-A-D-pT-P-S-S-N-R-P-S-N-S-S-P
21	A21	2015	S-P-F-P-A-D-T-P-pS-S-N-R-P-S-N-S-S-P
22	A22	2015	S-P-F-P-A-D-T-P-S-pS-N-R-P-S-N-S-S-P
23	A23	2015	S-P-F-P-A-D-T-P-S-S-N-R-P-pS-N-S-S-P
24	A24	2015	S-P-F-P-A-D-T-P-S-S-N-R-P-S-N-pS-S-P
25	A25	2015	S-P-F-P-A-D-T-P-S-S-N-R-P-S-N-S-pS-P
26	A26	1855	P-F-P-A-D-T-P-S-S-N-R-P-S-N-S-S-P-P
27	A27	2025.1	P-F-P-A-D-pT-P-S-S-N-R-P-S-N-S-S-P-P
28	A28	2025	P-F-P-A-D-T-P-pS-S-N-R-P-S-N-S-S-P-P
29	A29	2025	P-F-P-A-D-T-P-S-pS-N-R-P-S-N-S-S-P-P
30	A30	2025	P-F-P-A-D-T-P-S-S-N-R-P-pS-N-S-S-P-P
31	B 1	2025	P-F-P-A-D-T-P-S-S-N-R-P-S-N-pS-S-P-P
32	B 2	2025	P-F-P-A-D-T-P-S-S-N-R-P-S-N-S-pS-P-P
33	B 3	1855	F-P-A-D-T-P-S-S-N-R-P-S-N-S-S-P-P-P
34	B 4	2025.1	F-P-A-D-pT-P-S-S-N-R-P-S-N-S-S-P-P-P
35	B 5	2025	F-P-A-D-T-P-pS-S-N-R-P-S-N-S-S-P-P-P
36	B 6	2025	F-P-A-D-T-P-S-pS-N-R-P-S-N-S-S-P-P-P
37	B 7	2025	F-P-A-D-T-P-S-S-N-R-P-pS-N-S-S-P-P-P
38	B 8	2025	F-P-A-D-T-P-S-S-N-R-P-S-N-pS-S-P-P-P
39	B 9	2025	F-P-A-D-T-P-S-S-N-R-P-S-N-S-pS-P-P-P
40	B10	1794.9	P-A-D-T-P-S-S-N-R-P-S-N-S-S-P-P-P-S
41	B11	1965	P-A-D-pT-P-S-S-N-R-P-S-N-S-S-P-P-P-S

42	B12	1964.9	P-A-D-T-P-pS-S-N-R-P-S-N-S-S-P-P-P-S
43	B13	1964.9	P-A-D-T-P-S-pS-N-R-P-S-N-S-S-P-P-P-S
44	B14	1964.9	P-A-D-T-P-S-S-N-R-P-pS-N-S-S-P-P-P-S
45	B15	1964.9	P-A-D-T-P-S-S-N-R-P-S-N-pS-S-P-P-P-S
46	B16	1964.9	P-A-D-T-P-S-S-N-R-P-S-N-S-pS-P-P-P-S
47	B17	1964.9	P-A-D-T-P-S-S-N-R-P-S-N-S-S-P-P-P-pS
48	B18	1784.9	A-D-T-P-S-S-N-R-P-S-N-S-S-P-P-P-S-S
49	B19	1955	A-D-pT-P-S-S-N-R-P-S-N-S-S-P-P-P-S-S
50	B20	1954.9	A-D-T-P-pS-S-N-R-P-S-N-S-S-P-P-P-S-S
51	B21	1954.9	A-D-T-P-S-pS-N-R-P-S-N-S-S-P-P-P-S-S
52	B22	1954.9	A-D-T-P-S-S-N-R-P-pS-N-S-S-P-P-P-S-S
53	B23	1954.9	A-D-T-P-S-S-N-R-P-S-N-pS-S-P-P-P-S-S
54	B24	1954.9	A-D-T-P-S-S-N-R-P-S-N-S-pS-P-P-P-S-S
55	B25	1954.9	A-D-T-P-S-S-N-R-P-S-N-S-S-P-P-P-pS-S
56	B26	1954.9	A-D-T-P-S-S-N-R-P-S-N-S-S-P-P-P-S-pS
57	B27	1827	D-T-P-S-S-N-R-P-S-N-S-S-P-P-P-S-S-I
58	B28	1997.1	D-pT-P-S-S-N-R-P-S-N-S-S-P-P-P-S-S-I
59	B29	1997	D-T-P-pS-S-N-R-P-S-N-S-S-P-P-P-S-S-I
60	B30	1997	D-T-P-S-pS-N-R-P-S-N-S-S-P-P-P-S-S-I
61	C 1	1997	D-T-P-S-S-N-R-P-pS-N-S-S-P-P-P-S-S-I
62	C 2	1997	D-T-P-S-S-N-R-P-S-N-pS-S-P-P-P-S-S-I
63	C 3	1997	D-T-P-S-S-N-R-P-S-N-S-pS-P-P-P-S-S-I
64	C 4	1997	D-T-P-S-S-N-R-P-S-N-S-S-P-P-P-pS-S-I
65	C 5	1997	D-T-P-S-S-N-R-P-S-N-S-S-P-P-P-S-pS-I
66	C 6	1769	T-P-S-S-N-R-P-S-N-S-S-P-P-P-S-S-I-G
67	C 7	1939.1	pT-P-S-S-N-R-P-S-N-S-S-P-P-P-S-S-I-G
68	C 8	1939	T-P-pS-S-N-R-P-S-N-S-S-P-P-P-S-S-I-G
69	C 9	1939	T-P-S-pS-N-R-P-S-N-S-S-P-P-P-S-S-I-G
70	C10	1939	T-P-S-S-N-R-P-pS-N-S-S-P-P-P-S-S-I-G
71	C11	1939	T-P-S-S-N-R-P-S-N-pS-S-P-P-P-S-S-I-G
72	C12	1939	T-P-S-S-N-R-P-S-N-S-pS-P-P-P-S-S-I-G
73	C13	1939	T-P-S-S-N-R-P-S-N-S-S-P-P-P-pS-S-I-G
74	C14	1939	T-P-S-S-N-R-P-S-N-S-S-P-P-P-S-pS-I-G
75	C15	1739	P-S-S-N-R-P-S-N-S-S-P-P-P-S-S-I-G-A
76	C16	1909	P-pS-S-N-R-P-S-N-S-S-P-P-P-S-S-I-G-A
77	C17	1909	P-S-pS-N-R-P-S-N-S-S-P-P-P-S-S-I-G-A
78	C18	1909	P-S-S-N-R-P-pS-N-S-S-P-P-P-S-S-I-G-A
79	C19	1909	P-S-S-N-R-P-S-N-pS-S-P-P-P-S-S-I-G-A
80	C20	1909	P-S-S-N-R-P-S-N-S-pS-P-P-P-S-S-I-G-A
81	C21	1909	P-S-S-N-R-P-S-N-S-S-P-P-P-pS-S-I-G-A
82	C22	1909	P-S-S-N-R-P-S-N-S-S-P-P-P-S-pS-I-G-A
83	C23	1699	S-S-N-R-P-S-N-S-S-P-P-P-S-S-I-G-A-G
84	C24	1869	pS-S-N-R-P-S-N-S-S-P-P-P-S-S-I-G-A-G
85	C25	1869	S-pS-N-R-P-S-N-S-S-P-P-P-S-S-I-G-A-G
86	C26	1869	S-S-N-R-P-pS-N-S-S-P-P-P-S-S-I-G-A-G
87	C27	1869	S-S-N-R-P-S-N-pS-S-P-P-P-S-S-I-G-A-G
88	C28	1869	S-S-N-R-P-S-N-S-pS-P-P-P-S-S-I-G-A-G
89	C29	1869	S-S-N-R-P-S-N-S-S-P-P-P-pS-S-I-G-A-G
90	C30	1869	S-S-N-R-P-S-N-S-S-P-P-P-S-pS-I-G-A-G
91	D 1	1759.1	S-N-R-P-S-N-S-S-P-P-P-S-S-I-G-A-G-F

92	D 2	1929.1	pS-N-R-P-S-N-S-S-P-P-P-S-S-I-G-A-G-F
93	D 3	1929.1	S-N-R-P-pS-N-S-S-P-P-P-S-S-I-G-A-G-F
94	D 4	1929.1	S-N-R-P-S-N-pS-S-P-P-P-S-S-I-G-A-G-F
95	D 5	1929.1	S-N-R-P-S-N-S-pS-P-P-P-S-S-I-G-A-G-F
96	D 6	1929.1	S-N-R-P-S-N-S-S-P-P-P-pS-S-I-G-A-G-F
97	D 7	1929.1	S-N-R-P-S-N-S-S-P-P-P-S-pS-I-G-A-G-F
98	D 8	1729.1	N-R-P-S-N-S-S-P-P-P-S-S-I-G-A-G-F-G
99	D 9	1899.1	N-R-P-pS-N-S-S-P-P-P-S-S-I-G-A-G-F-G
100	D10	1899.1	N-R-P-S-N-pS-S-P-P-P-S-S-I-G-A-G-F-G
101	D11	1899.1	N-R-P-S-N-S-pS-P-P-P-S-S-I-G-A-G-F-G
102	D12	1899.1	N-R-P-S-N-S-S-P-P-pS-S-I-G-A-G-F-G
103	D13	1899.1	N-R-P-S-N-S-S-P-P-P-S-pS-I-G-A-G-F-G
104	D14	1702.1	R-P-S-N-S-S-P-P-P-S-S-I-G-A-G-F-G-S
105	D15	1872.1	R-P-pS-N-S-S-P-P-P-S-S-I-G-A-G-F-G-S
106	D16	1872.1	R-P-S-N-pS-S-P-P-P-S-S-I-G-A-G-F-G-S
107	D17	1872.1	R-P-S-N-S-pS-P-P-P-S-S-I-G-A-G-F-G-S
108	D18	1872.1	R-P-S-N-S-S-P-P-P-pS-S-I-G-A-G-F-G-S
109	D19	1872.1	R-P-S-N-S-S-P-P-P-S-pS-I-G-A-G-F-G-S
110	D20	1872.1	R-P-S-N-S-S-P-P-P-S-S-I-G-A-G-F-G-pS
111	D21	1633	P-S-N-S-S-P-P-P-S-S-I-G-A-G-F-G-S-S
112	D22	1803	P-pS-N-S-S-P-P-P-S-S-I-G-A-G-F-G-S-S
113	D23	1803	P-S-N-pS-S-P-P-P-S-S-I-G-A-G-F-G-S-S
114	D24	1803	P-S-N-S-pS-P-P-P-S-S-I-G-A-G-F-G-S-S
115	D25	1803	P-S-N-S-S-P-P-P-pS-S-I-G-A-G-F-G-S-S
116	D26	1803	P-S-N-S-S-P-P-P-S-pS-I-G-A-G-F-G-S-S
117	D27	1803	P-S-N-S-S-P-P-P-S-S-I-G-A-G-F-G-pS-S
118	D28	1803	P-S-N-S-S-P-P-P-S-S-I-G-A-G-F-G-S-pS
119	D29	1623	S-N-S-S-P-P-P-S-S-I-G-A-G-F-G-S-S-S
120	D30	1793	pS-N-S-S-P-P-P-S-S-I-G-A-G-F-G-S-S-S
121	E 1	1793	S-N-pS-S-P-P-P-S-S-I-G-A-G-F-G-S-S-S
122	E 2	1793	S-N-S-pS-P-P-P-S-S-I-G-A-G-F-G-S-S-S
123	E 3	1793	S-N-S-S-P-P-P-pS-S-I-G-A-G-F-G-S-S-S
124	E 4	1793	S-N-S-S-P-P-P-S-pS-I-G-A-G-F-G-S-S-S
125	E 5	1793	S-N-S-S-P-P-P-S-S-I-G-A-G-F-G-pS-S-S
126	E 6	1793	S-N-S-S-P-P-P-S-S-I-G-A-G-F-G-S-pS-S
127	E 7	1793	S-N-S-S-P-P-P-S-S-I-G-A-G-F-G-S-S-pS
128	E 8	1593	N-S-S-P-P-P-S-S-I-G-A-G-F-G-S-S-S-G
129	E 9	1763	N-pS-S-P-P-P-S-S-I-G-A-G-F-G-S-S-S-G
130	E10	1763	N-S-pS-P-P-P-S-S-I-G-A-G-F-G-S-S-S-G
131	E11	1763	N-S-S-P-P-P-pS-S-I-G-A-G-F-G-S-S-S-G
132	E12	1763	N-S-S-P-P-P-S-pS-I-G-A-G-F-G-S-S-S-G
133	E13	1763	N-S-S-P-P-P-S-S-I-G-A-G-F-G-pS-S-S-G
134	E14	1763	N-S-S-P-P-P-S-S-I-G-A-G-F-G-S-pS-S-G
135	E15	1763	N-S-S-P-P-P-S-S-I-G-A-G-F-G-S-S-pS-G
136	E16	1592.1	S-S-P-P-P-S-S-I-G-A-G-F-G-S-S-S-G-L
137	E17	1762.1	pS-S-P-P-P-S-S-I-G-A-G-F-G-S-S-S-G-L
138	E18	1762.1	S-pS-P-P-P-S-S-I-G-A-G-F-G-S-S-S-G-L
139	E19	1762.1	S-S-P-P-P-pS-S-I-G-A-G-F-G-S-S-S-G-L
140	E20	1762.1	S-S-P-P-P-S-pS-I-G-A-G-F-G-S-S-S-G-L
141	E21	1762.1	S-S-P-P-P-S-S-I-G-A-G-F-G-pS-S-S-G-L

142	E22	1762.1	S-S-P-P-P-S-S-I-G-A-G-F-G-S-pS-S-G-L
143	E23	1762.1	S-S-P-P-P-S-S-I-G-A-G-F-G-S-S-pS-G-L
144	E24	1620.1	S-P-P-P-S-S-I-G-A-G-F-G-S-S-S-G-L-D
145	E25	1790.1	pS-P-P-P-S-S-I-G-A-G-F-G-S-S-S-G-L-D
146	E26	1790.1	S-P-P-P-pS-S-I-G-A-G-F-G-S-S-S-G-L-D
147	E27	1790.1	S-P-P-P-S-pS-I-G-A-G-F-G-S-S-S-G-L-D
148	E28	1790.1	S-P-P-P-S-S-I-G-A-G-F-G-pS-S-S-G-L-D
149	E29	1790.1	S-P-P-P-S-S-I-G-A-G-F-G-S-pS-S-G-L-D
150	E30	1790.1	S-P-P-P-S-S-I-G-A-G-F-G-S-S-pS-G-L-D
151	F 1	1620.1	P-P-P-S-S-I-G-A-G-F-G-S-S-S-G-L-D-S
152	F 2	1790.1	P-P-P-pS-S-I-G-A-G-F-G-S-S-S-G-L-D-S
153	F 3	1790.1	P-P-P-S-pS-I-G-A-G-F-G-S-S-S-G-L-D-S
154	F 4	1790.1	P-P-P-S-S-I-G-A-G-F-G-pS-S-S-G-L-D-S
155	F 5	1790.1	P-P-P-S-S-I-G-A-G-F-G-S-pS-S-G-L-D-S
156	F 6	1790.1	P-P-P-S-S-I-G-A-G-F-G-S-S-pS-G-L-D-S
157	F 7	1790.1	P-P-P-S-S-I-G-A-G-F-G-S-S-S-G-L-D-pS
158	F 8	1651.1	P-P-S-S-I-G-A-G-F-G-S-S-S-G-L-D-S-Q
159	F 9	1821.1	P-P-pS-S-I-G-A-G-F-G-S-S-S-G-L-D-S-Q
160	F10	1821.1	P-P-S-pS-I-G-A-G-F-G-S-S-S-G-L-D-S-Q
161	F11	1821.1	P-P-S-S-I-G-A-G-F-G-pS-S-S-G-L-D-S-Q
162	F12	1821.1	P-P-S-S-I-G-A-G-F-G-S-pS-S-G-L-D-S-Q
163	F13	1821.1	P-P-S-S-I-G-A-G-F-G-S-S-pS-G-L-D-S-Q
164	F14	1821.1	P-P-S-S-I-G-A-G-F-G-S-S-S-G-L-D-pS-Q
165	F15	1667.2	P-S-S-I-G-A-G-F-G-S-S-S-G-L-D-S-Q-I
166	F16	1837.2	P-pS-S-I-G-A-G-F-G-S-S-S-G-L-D-S-Q-I
167	F17	1837.2	P-S-pS-I-G-A-G-F-G-S-S-S-G-L-D-S-Q-I
168	F18	1837.2	P-S-S-I-G-A-G-F-G-pS-S-S-G-L-D-S-Q-I
169	F19	1837.2	P-S-S-I-G-A-G-F-G-S-pS-S-G-L-D-S-Q-I
170	F20	1837.2	P-S-S-I-G-A-G-F-G-S-S-pS-G-L-D-S-Q-I
171	F21	1837.2	P-S-S-I-G-A-G-F-G-S-S-S-G-L-D-pS-Q-I
172	F22	1627.2	S-S-I-G-A-G-F-G-S-S-S-G-L-D-S-Q-I-G
173	F23	1797.2	pS-S-I-G-A-G-F-G-S-S-S-G-L-D-S-Q-I-G
174	F24	1797.2	S-pS-I-G-A-G-F-G-S-S-S-G-L-D-S-Q-I-G
175	F25	1797.2	S-S-I-G-A-G-F-G-pS-S-S-G-L-D-S-Q-I-G
176	F26	1797.2	S-S-I-G-A-G-F-G-S-pS-S-G-L-D-S-Q-I-G
177	F27	1797.2	S-S-I-G-A-G-F-G-S-S-pS-G-L-D-S-Q-I-G
178	F28	1797.2	S-S-I-G-A-G-F-G-S-S-S-G-L-D-pS-Q-I-G
179	F29	1627.2	S-I-G-A-G-F-G-S-S-S-G-L-D-S-Q-I-G-S
180	F30	1797.2	pS-I-G-A-G-F-G-S-S-S-G-L-D-S-Q-I-G-S
181	G 1	1797.2	S-I-G-A-G-F-G-pS-S-S-G-L-D-S-Q-I-G-S
182	G 2	1797.2	S-I-G-A-G-F-G-S-pS-S-G-L-D-S-Q-I-G-S
183	G 3	1797.2	S-I-G-A-G-F-G-S-S-pS-G-L-D-S-Q-I-G-S
184	G 4	1797.2	S-I-G-A-G-F-G-S-S-S-G-L-D-pS-Q-I-G-S
185	G 5	1797.2	S-I-G-A-G-F-G-S-S-S-G-L-D-S-Q-I-G-pS
186	G 6	1696.3	I-G-A-G-F-G-S-S-S-G-L-D-S-Q-I-G-S-R
187	G 7	1866.3	I-G-A-G-F-G-pS-S-S-G-L-D-S-Q-I-G-S-R
188	G 8	1866.3	I-G-A-G-F-G-S-pS-S-G-L-D-S-Q-I-G-S-R
189	G 9	1866.3	I-G-A-G-F-G-S-S-pS-G-L-D-S-Q-I-G-S-R
190	G10	1866.3	I-G-A-G-F-G-S-S-S-G-L-D-pS-Q-I-G-S-R
191	G11	1866.3	I-G-A-G-F-G-S-S-S-G-L-D-S-Q-I-G-pS-R

192	G12	1696.3	G-A-G-F-G-S-S-S-G-L-D-S-Q-I-G-S-R-L
193	G13	1866.3	G-A-G-F-G-pS-S-S-G-L-D-S-Q-I-G-S-R-L
194	G14	1866.3	G-A-G-F-G-S-pS-S-G-L-D-S-Q-I-G-S-R-L
195	G15	1866.3	G-A-G-F-G-S-S-pS-G-L-D-S-Q-I-G-S-R-L
196	G16	1866.3	G-A-G-F-G-S-S-S-G-L-D-pS-Q-I-G-S-R-L
197	G17	1866.3	G-A-G-F-G-S-S-S-G-L-D-S-Q-I-G-pS-R-L
198	G18	1776.3	A-G-F-G-S-S-S-G-L-D-S-Q-I-G-S-R-L-H
199	G19	1946.3	A-G-F-G-pS-S-S-G-L-D-S-Q-I-G-S-R-L-H
200	G20	1946.3	A-G-F-G-S-pS-S-G-L-D-S-Q-I-G-S-R-L-H
201	G21	1946.3	A-G-F-G-S-S-pS-G-L-D-S-Q-I-G-S-R-L-H
202	G22	1946.3	A-G-F-G-S-S-S-G-L-D-pS-Q-I-G-S-R-L-H
203	G23	1946.3	A-G-F-G-S-S-S-G-L-D-S-Q-I-G-pS-R-L-H
204	G24	1852.4	G-F-G-S-S-S-G-L-D-S-Q-I-G-S-R-L-H-F
205	G25	2022.4	G-F-G-pS-S-S-G-L-D-S-Q-I-G-S-R-L-H-F
206	G26	2022.4	G-F-G-S-pS-S-G-L-D-S-Q-I-G-S-R-L-H-F
207	G27	2022.4	G-F-G-S-S-pS-G-L-D-S-Q-I-G-S-R-L-H-F
208	G28	2022.4	G-F-G-S-S-S-G-L-D-pS-Q-I-G-S-R-L-H-F
209	G29	2022.4	G-F-G-S-S-S-G-L-D-S-Q-I-G-pS-R-L-H-F
210	G30	1892.4	F-G-S-S-S-G-L-D-S-Q-I-G-S-R-L-H-F-P
211	H 1	2062.4	F-G-pS-S-S-G-L-D-S-Q-I-G-S-R-L-H-F-P
212	H 2	2062.4	F-G-S-pS-S-G-L-D-S-Q-I-G-S-R-L-H-F-P
213	H 3	2062.4	F-G-S-S-pS-G-L-D-S-Q-I-G-S-R-L-H-F-P
214	H 4	2062.4	F-G-S-S-S-G-L-D-pS-Q-I-G-S-R-L-H-F-P
215	H 5	2062.4	F-G-S-S-S-G-L-D-S-Q-I-G-pS-R-L-H-F-P
216	H 6	1832.3	G-S-S-S-G-L-D-S-Q-I-G-S-R-L-H-F-P-S
217	H 7	2002.3	G-pS-S-S-G-L-D-S-Q-I-G-S-R-L-H-F-P-S
218	H 8	2002.3	G-S-pS-S-G-L-D-S-Q-I-G-S-R-L-H-F-P-S
219	H 9	2002.3	G-S-S-pS-G-L-D-S-Q-I-G-S-R-L-H-F-P-S
220	H10	2002.3	G-S-S-S-G-L-D-pS-Q-I-G-S-R-L-H-F-P-S
221	H11	2002.3	G-S-S-S-G-L-D-S-Q-I-G-pS-R-L-H-F-P-S
222	H12	2002.3	G-S-S-S-G-L-D-S-Q-I-G-S-R-L-H-F-P-pS
223	H13	1862.3	S-S-S-G-L-D-S-Q-I-G-S-R-L-H-F-P-S-S
224	H14	2032.3	pS-S-S-G-L-D-S-Q-I-G-S-R-L-H-F-P-S-S
225	H15	2032.3	S-pS-S-G-L-D-S-Q-I-G-S-R-L-H-F-P-S-S
226	H16	2032.3	S-S-pS-G-L-D-S-Q-I-G-S-R-L-H-F-P-S-S
227	H17	2032.3	S-S-S-G-L-D-pS-Q-I-G-S-R-L-H-F-P-S-S
228	H18	2032.3	S-S-S-G-L-D-S-Q-I-G-pS-R-L-H-F-P-S-S
229	H19	2032.3	S-S-S-G-L-D-S-Q-I-G-S-R-L-H-F-P-pS-S
230	H20	2032.3	S-S-S-G-L-D-S-Q-I-G-S-R-L-H-F-P-S-pS
231	H21	1862.3	S-S-G-L-D-S-Q-I-G-S-R-L-H-F-P-S-S-S
232	H22	2032.3	pS-S-G-L-D-S-Q-I-G-S-R-L-H-F-P-S-S-S
233	H23	2032.3	S-pS-G-L-D-S-Q-I-G-S-R-L-H-F-P-S-S-S
234	H24	2032.3	S-S-G-L-D-pS-Q-I-G-S-R-L-H-F-P-S-S-S
235	H25	2032.3	S-S-G-L-D-S-Q-I-G-pS-R-L-H-F-P-S-S-S
236	H26	2032.3	S-S-G-L-D-S-Q-I-G-S-R-L-H-F-P-pS-S-S
237	H27	2032.3	S-S-G-L-D-S-Q-I-G-S-R-L-H-F-P-S-pS-S
238	H28	2032.3	S-S-G-L-D-S-Q-I-G-S-R-L-H-F-P-S-S-pS
239	H29	1903.3	S-G-L-D-S-Q-I-G-S-R-L-H-F-P-S-S-S-Q
240	H30	2073.3	pS-G-L-D-S-Q-I-G-S-R-L-H-F-P-S-S-S-Q
241	I 1	2073.3	S-G-L-D-pS-Q-I-G-S-R-L-H-F-P-S-S-S-Q

242	I 2	2073.3	S-G-L-D-S-Q-I-G-pS-R-L-H-F-P-S-S-S-Q
243	I 3	2073.3	S-G-L-D-S-Q-I-G-S-R-L-H-F-P-pS-S-S-Q
244	I 4	2073.3	S-G-L-D-S-Q-I-G-S-R-L-H-F-P-S-pS-S-Q
245	I 5	2073.3	S-G-L-D-S-Q-I-G-S-R-L-H-F-P-S-S-pS-Q
246	I 6	1913.3	G-L-D-S-Q-I-G-S-R-L-H-F-P-S-S-S-Q-P
247	I 7	2083.3	G-L-D-pS-Q-I-G-S-R-L-H-F-P-S-S-S-Q-P
248	I 8	2083.3	G-L-D-S-Q-I-G-pS-R-L-H-F-P-S-S-S-Q-P
249	I 9	2083.3	G-L-D-S-Q-I-G-S-R-L-H-F-P-pS-S-S-Q-P
250	I10	2083.3	G-L-D-S-Q-I-G-S-R-L-H-F-P-S-pS-S-Q-P
251	I11	2083.3	G-L-D-S-Q-I-G-S-R-L-H-F-P-S-S-pS-Q-P
252	I12	1993.3	L-D-S-Q-I-G-S-R-L-H-F-P-S-S-S-Q-P-H
253	I13	2163.3	L-D-pS-Q-I-G-S-R-L-H-F-P-S-S-S-Q-P-H
254	I14	2163.3	L-D-S-Q-I-G-pS-R-L-H-F-P-S-S-S-Q-P-H
255	I15	2163.3	L-D-S-Q-I-G-S-R-L-H-F-P-pS-S-S-Q-P-H
256	I16	2163.3	L-D-S-Q-I-G-S-R-L-H-F-P-S-pS-S-Q-P-H
257	I17	2163.3	L-D-S-Q-I-G-S-R-L-H-F-P-S-S-pS-Q-P-H
258	I18	1979.2	D-S-Q-I-G-S-R-L-H-F-P-S-S-S-Q-P-H-V
259	I19	2149.2	D-pS-Q-I-G-S-R-L-H-F-P-S-S-S-Q-P-H-V
260	I20	2149.2	D-S-Q-I-G-pS-R-L-H-F-P-S-S-S-Q-P-H-V
261	I21	2149.2	D-S-Q-I-G-S-R-L-H-F-P-pS-S-S-Q-P-H-V
262	I22	2149.2	D-S-Q-I-G-S-R-L-H-F-P-S-pS-S-Q-P-H-V
263	I23	2149.2	D-S-Q-I-G-S-R-L-H-F-P-S-S-pS-Q-P-H-V
264	I24	1951.2	S-Q-I-G-S-R-L-H-F-P-S-S-S-Q-P-H-V-S
265	I25	2121.2	pS-Q-I-G-S-R-L-H-F-P-S-S-S-Q-P-H-V-S
266	I26	2121.2	S-Q-I-G-pS-R-L-H-F-P-S-S-S-Q-P-H-V-S
267	I27	2121.2	S-Q-I-G-S-R-L-H-F-P-pS-S-S-Q-P-H-V-S
268	I28	2121.2	S-Q-I-G-S-R-L-H-F-P-S-pS-S-Q-P-H-V-S
269	I29	2121.2	S-Q-I-G-S-R-L-H-F-P-S-S-pS-Q-P-H-V-S
270	I30	2121.2	S-Q-I-G-S-R-L-H-F-P-S-S-S-Q-P-H-V-pS
271	J 1	1978.2	Q-I-G-S-R-L-H-F-P-S-S-S-Q-P-H-V-S-N
272	J 2	2148.2	Q-I-G-pS-R-L-H-F-P-S-S-S-Q-P-H-V-S-N
273	J 3	2148.2	Q-I-G-S-R-L-H-F-P-pS-S-S-Q-P-H-V-S-N
274	J 4	2148.2	Q-I-G-S-R-L-H-F-P-S-pS-S-Q-P-H-V-S-N
275	J 5	2148.2	Q-I-G-S-R-L-H-F-P-S-S-pS-Q-P-H-V-S-N
276	J 6	2148.2	Q-I-G-S-R-L-H-F-P-S-S-S-Q-P-H-V-pS-N
277	J 7	1937.2	I-G-S-R-L-H-F-P-S-S-S-Q-P-H-V-S-N-S
278	J 8	2107.2	I-G-pS-R-L-H-F-P-S-S-S-Q-P-H-V-S-N-S
279	J 9	2107.2	I-G-S-R-L-H-F-P-pS-S-S-Q-P-H-V-S-N-S
280	J10	2107.2	I-G-S-R-L-H-F-P-S-pS-S-Q-P-H-V-S-N-S
281	J11	2107.2	I-G-S-R-L-H-F-P-S-S-pS-Q-P-H-V-S-N-S
282	J12	2107.2	I-G-S-R-L-H-F-P-S-S-S-Q-P-H-V-pS-N-S
283	J13	2107.2	I-G-S-R-L-H-F-P-S-S-S-Q-P-H-V-S-N-pS
284	J14	1952.1	G-S-R-L-H-F-P-S-S-S-Q-P-H-V-S-N-S-Q
285	J15	2122.1	G-pS-R-L-H-F-P-S-S-S-Q-P-H-V-S-N-S-Q
286	J16	2122.1	G-S-R-L-H-F-P-pS-S-S-Q-P-H-V-S-N-S-Q
287	J17	2122.1	G-S-R-L-H-F-P-S-pS-S-Q-P-H-V-S-N-S-Q
288	J18	2122.1	G-S-R-L-H-F-P-S-S-pS-Q-P-H-V-S-N-S-Q
289	J19	2122.1	G-S-R-L-H-F-P-S-S-S-Q-P-H-V-pS-N-S-Q
290	J20	2122.1	G-S-R-L-H-F-P-S-S-S-Q-P-H-V-S-N-pS-Q
291	J21	1996.1	S-R-L-H-F-P-S-S-S-Q-P-H-V-S-N-S-Q-T

292	J22	2166.1	pS-R-L-H-F-P-S-S-S-Q-P-H-V-S-N-S-Q-T
293	J23	2166.1	S-R-L-H-F-P-pS-S-S-Q-P-H-V-S-N-S-Q-T
294	J24	2166.1	S-R-L-H-F-P-S-pS-S-Q-P-H-V-S-N-S-Q-T
295	J25	2166.1	S-R-L-H-F-P-S-S-pS-Q-P-H-V-S-N-S-Q-T
296	J26	2166.1	S-R-L-H-F-P-S-S-S-Q-P-H-V-pS-N-S-Q-T
297	J27	2166.1	S-R-L-H-F-P-S-S-S-Q-P-H-V-S-N-pS-Q-T
298	J28	2166.2	S-R-L-H-F-P-S-S-S-Q-P-H-V-S-N-S-Q-pT
299	J29	1966.1	R-L-H-F-P-S-S-S-Q-P-H-V-S-N-S-Q-T-G
300	J30	2136.1	R-L-H-F-P-pS-S-S-Q-P-H-V-S-N-S-Q-T-G
301	K 1	2136.1	R-L-H-F-P-S-pS-S-Q-P-H-V-S-N-S-Q-T-G
302	K 2	2136.1	R-L-H-F-P-S-S-pS-Q-P-H-V-S-N-S-Q-T-G
303	K 3	2136.1	R-L-H-F-P-S-S-S-Q-P-H-V-pS-N-S-Q-T-G
304	K 4	2136.1	R-L-H-F-P-S-S-S-Q-P-H-V-S-N-pS-Q-T-G
305	K 5	2136.2	R-L-H-F-P-S-S-S-Q-P-H-V-S-N-S-Q-pT-G
306	K 6	1809.9	L-H-F-P-S-S-S-Q-P-H-V-S-N-S-Q-T-G
307	K 7	1979.9	L-H-F-P-pS-S-S-Q-P-H-V-S-N-S-Q-T-G
308	K 8	1979.9	L-H-F-P-S-pS-S-Q-P-H-V-S-N-S-Q-T-G
309	K 9	1979.9	L-H-F-P-S-S-pS-Q-P-H-V-S-N-S-Q-T-G
310	K10	1979.9	L-H-F-P-S-S-S-Q-P-H-V-pS-N-S-Q-T-G
311	K11	1979.9	L-H-F-P-S-S-S-Q-P-H-V-S-N-pS-Q-T-G
312	K12	1980	L-H-F-P-S-S-S-Q-P-H-V-S-N-S-Q-pT-G
313	K13	1793.8	H-F-P-S-S-S-Q-P-H-V-S-N-S-Q-T-G-P
314	K14	1963.8	H-F-P-pS-S-S-Q-P-H-V-S-N-S-Q-T-G-P
315	K15	1963.8	H-F-P-S-pS-S-Q-P-H-V-S-N-S-Q-T-G-P
316	K16	1963.8	H-F-P-S-S-pS-Q-P-H-V-S-N-S-Q-T-G-P
317	K17	1963.8	H-F-P-S-S-S-Q-P-H-V-pS-N-S-Q-T-G-P
318	K18	1963.8	H-F-P-S-S-S-Q-P-H-V-S-N-pS-Q-T-G-P
319	K19	1963.9	H-F-P-S-S-S-Q-P-H-V-S-N-S-Q-pT-G-P
320	K20	1803.9	F-P-S-S-S-Q-P-H-V-S-N-S-Q-T-G-P-F
321	K21	1973.9	F-P-pS-S-S-Q-P-H-V-S-N-S-Q-T-G-P-F
322	K22	1973.9	F-P-S-pS-S-Q-P-H-V-S-N-S-Q-T-G-P-F
323	K23	1973.9	F-P-S-S-pS-Q-P-H-V-S-N-S-Q-T-G-P-F
324	K24	1973.9	F-P-S-S-S-Q-P-H-V-pS-N-S-Q-T-G-P-F
325	K25	1973.9	F-P-S-S-S-Q-P-H-V-S-N-pS-Q-T-G-P-F
326	K26	1974	F-P-S-S-S-Q-P-H-V-S-N-S-Q-pT-G-P-F
327	K27	1803.9	F-P-S-S-S-Q-P-H-V-S-N-S-Q-T-G-P-F
328	K28	1973.9	F-P-pS-S-S-Q-P-H-V-S-N-S-Q-T-G-P-F
329	K29	1973.9	F-P-S-pS-S-Q-P-H-V-S-N-S-Q-T-G-P-F
330	K30	1973.9	F-P-S-S-pS-Q-P-H-V-S-N-S-Q-T-G-P-F
331	L 1	1973.9	F-P-S-S-S-Q-P-H-V-pS-N-S-Q-T-G-P-F
332	L 2	1973.9	F-P-S-S-S-Q-P-H-V-S-N-pS-Q-T-G-P-F
333	L 3	1974	F-P-S-S-S-Q-P-H-V-S-N-S-Q-pT-G-P-F
334	L 4	1755.8	P-S-S-S-Q-P-H-V-S-N-S-Q-T-G-P-F-V
335	L 5	1925.8	P-pS-S-S-Q-P-H-V-S-N-S-Q-T-G-P-F-V
336	L 6	1925.8	P-S-pS-S-Q-P-H-V-S-N-S-Q-T-G-P-F-V
337	L 7	1925.8	P-S-S-pS-Q-P-H-V-S-N-S-Q-T-G-P-F-V
338	L 8	1925.8	P-S-S-S-Q-P-H-V-pS-N-S-Q-T-G-P-F-V
339	L 9	1925.8	P-S-S-S-Q-P-H-V-S-N-pS-Q-T-G-P-F-V
340	L10	1925.9	P-S-S-S-Q-P-H-V-S-N-S-Q-pT-G-P-F-V
341	L11	1772.8	S-S-S-Q-P-H-V-S-N-S-Q-T-G-P-F-V-N

342	L12	1942.8	pS-S-S-Q-P-H-V-S-N-S-Q-T-G-P-F-V-N
343	L13	1942.8	S-pS-S-Q-P-H-V-S-N-S-Q-T-G-P-F-V-N
344	L14	1942.8	S-S-pS-Q-P-H-V-S-N-S-Q-T-G-P-F-V-N
345	L15	1942.8	S-S-S-Q-P-H-V-pS-N-S-Q-T-G-P-F-V-N
346	L16	1942.8	S-S-S-Q-P-H-V-S-N-pS-Q-T-G-P-F-V-N
347	L17	1942.9	S-S-S-Q-P-H-V-S-N-S-Q-pT-G-P-F-V-N
348	L18	1800.8	S-S-Q-P-H-V-S-N-S-Q-T-G-P-F-V-N-D
349	L19	1970.8	pS-S-Q-P-H-V-S-N-S-Q-T-G-P-F-V-N-D
350	L20	1970.8	S-pS-Q-P-H-V-S-N-S-Q-T-G-P-F-V-N-D
351	L21	1970.8	S-S-Q-P-H-V-pS-N-S-Q-T-G-P-F-V-N-D
352	L22	1970.8	S-S-Q-P-H-V-S-N-pS-Q-T-G-P-F-V-N-D
353	L23	1970.9	S-S-Q-P-H-V-S-N-S-Q-pT-G-P-F-V-N-D
354	L24	1800.8	S-Q-P-H-V-S-N-S-Q-T-G-P-F-V-N-D-S
355	L25	1970.8	pS-Q-P-H-V-S-N-S-Q-T-G-P-F-V-N-D-S
356	L26	1970.8	S-Q-P-H-V-pS-N-S-Q-T-G-P-F-V-N-D-S
357	L27	1970.8	S-Q-P-H-V-S-N-pS-Q-T-G-P-F-V-N-D-S
358	L28	1970.9	S-Q-P-H-V-S-N-S-Q-pT-G-P-F-V-N-D-S
359	L29	1970.8	S-Q-P-H-V-S-N-S-Q-T-G-P-F-V-N-D-pS
360	L30	1814.8	Q-P-H-V-S-N-S-Q-T-G-P-F-V-N-D-S-T
361	M 1	1984.8	Q-P-H-V-pS-N-S-Q-T-G-P-F-V-N-D-S-T
362	M 2	1984.8	Q-P-H-V-S-N-pS-Q-T-G-P-F-V-N-D-S-T
363	M 3	1984.9	Q-P-H-V-S-N-S-Q-pT-G-P-F-V-N-D-S-T
364	M 4	1984.8	Q-P-H-V-S-N-S-Q-T-G-P-F-V-N-D-pS-T
365	M 5	1984.9	Q-P-H-V-S-N-S-Q-T-G-P-F-V-N-D-S-pT
366	M 6	1814.8	P-H-V-S-N-S-Q-T-G-P-F-V-N-D-S-T-Q
367	M 7	1984.8	P-H-V-pS-N-S-Q-T-G-P-F-V-N-D-S-T-Q
368	M 8	1984.8	P-H-V-S-N-pS-Q-T-G-P-F-V-N-D-S-T-Q
369	M 9	1984.9	P-H-V-S-N-S-Q-pT-G-P-F-V-N-D-S-T-Q
370	M10	1984.8	P-H-V-S-N-S-Q-T-G-P-F-V-N-D-pS-T-Q
371	M11	1984.9	P-H-V-S-N-S-Q-T-G-P-F-V-N-D-S-pT-Q
372	M12	1864.9	H-V-S-N-S-Q-T-G-P-F-V-N-D-S-T-Q-F
373	M13	2034.9	H-V-pS-N-S-Q-T-G-P-F-V-N-D-S-T-Q-F
374	M14	2034.9	H-V-S-N-pS-Q-T-G-P-F-V-N-D-S-T-Q-F
375	M15	2035	H-V-S-N-S-Q-pT-G-P-F-V-N-D-S-T-Q-F
376	M16	2034.9	H-V-S-N-S-Q-T-G-P-F-V-N-D-pS-T-Q-F
377	M17	2035	H-V-S-N-S-Q-T-G-P-F-V-N-D-S-pT-Q-F
378	M18	1814.9	V-S-N-S-Q-T-G-P-F-V-N-D-S-T-Q-F-S
379	M19	1984.9	V-pS-N-S-Q-T-G-P-F-V-N-D-S-T-Q-F-S
380	M20	1984.9	V-S-N-pS-Q-T-G-P-F-V-N-D-S-T-Q-F-S
381	M21	1985	V-S-N-S-Q-pT-G-P-F-V-N-D-S-T-Q-F-S
382	M22	1984.9	V-S-N-S-Q-T-G-P-F-V-N-D-pS-T-Q-F-S
383	M23	1985	V-S-N-S-Q-T-G-P-F-V-N-D-S-pT-Q-F-S
384	M24	1984.9	V-S-N-S-Q-T-G-P-F-V-N-D-S-T-Q-F-pS
385	M25	1802.9	S-N-S-Q-T-G-P-F-V-N-D-S-T-Q-F-S-S
386	M26	1972.9	pS-N-S-Q-T-G-P-F-V-N-D-S-T-Q-F-S-S
387	M27	1972.9	S-N-pS-Q-T-G-P-F-V-N-D-S-T-Q-F-S-S
388	M28	1973	S-N-S-Q-pT-G-P-F-V-N-D-S-T-Q-F-S-S
389	M29	1972.9	S-N-S-Q-T-G-P-F-V-N-D-pS-T-Q-F-S-S
390	M30	1973	S-N-S-Q-T-G-P-F-V-N-D-S-pT-Q-F-S-S
391	N 1	1972.9	S-N-S-Q-T-G-P-F-V-N-D-S-T-Q-F-pS-S

392	N 2	1972.9	S-N-S-Q-T-G-P-F-V-N-D-S-T-Q-F-S-pS
393	N 3	1843.9	N-S-Q-T-G-P-F-V-N-D-S-T-Q-F-S-S-Q
394	N 4	2013.9	N-pS-Q-T-G-P-F-V-N-D-S-T-Q-F-S-S-Q
395	N 5	2014	N-S-Q-pT-G-P-F-V-N-D-S-T-Q-F-S-S-Q
396	N 6	2013.9	N-S-Q-T-G-P-F-V-N-D-pS-T-Q-F-S-S-Q
397	N 7	2014	N-S-Q-T-G-P-F-V-N-D-S-pT-Q-F-S-S-Q
398	N 8	2013.9	N-S-Q-T-G-P-F-V-N-D-S-T-Q-F-pS-S-Q
399	N 9	2013.9	N-S-Q-T-G-P-F-V-N-D-S-T-Q-F-S-pS-Q
400	N10	1886	S-Q-T-G-P-F-V-N-D-S-T-Q-F-S-S-Q-R
401	N11	2056	pS-Q-T-G-P-F-V-N-D-S-T-Q-F-S-S-Q-R
402	N12	2056.1	S-Q-pT-G-P-F-V-N-D-S-T-Q-F-S-S-Q-R
403	N13	2056	S-Q-T-G-P-F-V-N-D-pS-T-Q-F-S-S-Q-R
404	N14	2056.1	S-Q-T-G-P-F-V-N-D-S-pT-Q-F-S-S-Q-R
405	N15	2056	S-Q-T-G-P-F-V-N-D-S-T-Q-F-pS-S-Q-R
406	N16	2056	S-Q-T-G-P-F-V-N-D-S-T-Q-F-S-pS-Q-R
407	N17	1912.1	Q-T-G-P-F-V-N-D-S-T-Q-F-S-S-Q-R-L
408	N18	2082.2	Q-pT-G-P-F-V-N-D-S-T-Q-F-S-S-Q-R-L
409	N19	2082.1	Q-T-G-P-F-V-N-D-pS-T-Q-F-S-S-Q-R-L
410	N20	2082.2	Q-T-G-P-F-V-N-D-S-pT-Q-F-S-S-Q-R-L
411	N21	2082.1	Q-T-G-P-F-V-N-D-S-T-Q-F-pS-S-Q-R-L
412	N22	2082.1	Q-T-G-P-F-V-N-D-S-T-Q-F-S-pS-Q-R-L
413	N23	1912.1	T-G-P-F-V-N-D-S-T-Q-F-S-S-Q-R-L-Q
414	N24	2082.2	pT-G-P-F-V-N-D-S-T-Q-F-S-S-Q-R-L-Q
415	N25	2082.1	T-G-P-F-V-N-D-pS-T-Q-F-S-S-Q-R-L-Q
416	N26	2082.2	T-G-P-F-V-N-D-S-pT-Q-F-S-S-Q-R-L-Q
417	N27	2082.1	T-G-P-F-V-N-D-S-T-Q-F-pS-S-Q-R-L-Q
418	N28	2082.1	T-G-P-F-V-N-D-S-T-Q-F-S-pS-Q-R-L-Q
419	N29	1912.1	G-P-F-V-N-D-S-T-Q-F-S-S-Q-R-L-Q-T
420	N30	2082.1	G-P-F-V-N-D-pS-T-Q-F-S-S-Q-R-L-Q-T
421	O 1	2082.2	G-P-F-V-N-D-S-pT-Q-F-S-S-Q-R-L-Q-T
422	O 2	2082.1	G-P-F-V-N-D-S-T-Q-F-pS-S-Q-R-L-Q-T
423	O 3	2082.1	G-P-F-V-N-D-S-T-Q-F-S-pS-Q-R-L-Q-T
424	O 4	2082.2	G-P-F-V-N-D-S-T-Q-F-S-S-Q-R-L-Q-pT
425	O 5	1970.1	P-F-V-N-D-S-T-Q-F-S-S-Q-R-L-Q-T-D
426	O 6	2140.1	P-F-V-N-D-pS-T-Q-F-S-S-Q-R-L-Q-T-D
427	O 7	2140.2	P-F-V-N-D-S-pT-Q-F-S-S-Q-R-L-Q-T-D
428	O 8	2140.1	P-F-V-N-D-S-T-Q-F-pS-S-Q-R-L-Q-T-D
429	O 9	2140.1	P-F-V-N-D-S-T-Q-F-S-pS-Q-R-L-Q-T-D
430	O10	2140.2	P-F-V-N-D-S-T-Q-F-S-S-Q-R-L-Q-pT-D
431	O11	1930.1	F-V-N-D-S-T-Q-F-S-S-Q-R-L-Q-T-D-G
432	O12	2100.1	F-V-N-D-pS-T-Q-F-S-S-Q-R-L-Q-T-D-G
433	O13	2100.2	F-V-N-D-S-pT-Q-F-S-S-Q-R-L-Q-T-D-G
434	O14	2100.1	F-V-N-D-S-T-Q-F-pS-S-Q-R-L-Q-T-D-G
435	O15	2100.1	F-V-N-D-S-T-Q-F-S-pS-Q-R-L-Q-T-D-G
436	O16	2100.2	F-V-N-D-S-T-Q-F-S-S-Q-R-L-Q-pT-D-G
437	O17	1870	V-N-D-S-T-Q-F-S-S-Q-R-L-Q-T-D-G-S
438	O18	2040	V-N-D-pS-T-Q-F-S-S-Q-R-L-Q-T-D-G-S
439	O19	2040.1	V-N-D-S-pT-Q-F-S-S-Q-R-L-Q-T-D-G-S
440	O20	2040	V-N-D-S-T-Q-F-pS-S-Q-R-L-Q-T-D-G-S
441	O21	2040	V-N-D-S-T-Q-F-S-pS-Q-R-L-Q-T-D-G-S

442	O22	2040.1	V-N-D-S-T-Q-F-S-S-Q-R-L-Q-pT-D-G-S
443	O23	2040	V-N-D-S-T-Q-F-S-S-Q-R-L-Q-T-D-G-pS
444	O24	1842	N-D-S-T-Q-F-S-S-Q-R-L-Q-T-D-G-S-A
445	O25	2012	N-D-pS-T-Q-F-S-S-Q-R-L-Q-T-D-G-S-A
446	O26	2012.1	N-D-S-pT-Q-F-S-S-Q-R-L-Q-T-D-G-S-A
447	O27	2012	N-D-S-T-Q-F-pS-S-Q-R-L-Q-T-D-G-S-A
448	O28	2012	N-D-S-T-Q-F-S-pS-Q-R-L-Q-T-D-G-S-A
449	O29	2012.1	N-D-S-T-Q-F-S-S-Q-R-L-Q-pT-D-G-S-A
450	O30	2012	N-D-S-T-Q-F-S-S-Q-R-L-Q-T-D-G-pS-A
451	P 1	1829	D-S-T-Q-F-S-S-Q-R-L-Q-T-D-G-S-A-T
452	P 2	1999	D-pS-T-Q-F-S-S-Q-R-L-Q-T-D-G-S-A-T
453	P 3	1999.1	D-S-pT-Q-F-S-S-Q-R-L-Q-T-D-G-S-A-T
454	P 4	1999	D-S-T-Q-F-pS-S-Q-R-L-Q-T-D-G-S-A-T
455	P 5	1999	D-S-T-Q-F-S-pS-Q-R-L-Q-T-D-G-S-A-T
456	P 6	1999.1	D-S-T-Q-F-S-S-Q-R-L-Q-pT-D-G-S-A-T
457	P 7	1999	D-S-T-Q-F-S-S-Q-R-L-Q-T-D-G-pS-A-T
458	P 8	1829	D-S-T-Q-F-S-S-Q-R-L-Q-T-D-G-S-A-T
459	P 9	1828	S-T-Q-F-S-S-Q-R-L-Q-T-D-G-S-A-T-N
460	P10	1998	pS-T-Q-F-S-S-Q-R-L-Q-T-D-G-S-A-T-N
461	P11	1998.1	S-pT-Q-F-S-S-Q-R-L-Q-T-D-G-S-A-T-N
462	P12	1998	S-T-Q-F-pS-S-Q-R-L-Q-T-D-G-S-A-T-N
463	P13	1998	S-T-Q-F-S-pS-Q-R-L-Q-T-D-G-S-A-T-N
464	P14	1998.1	S-T-Q-F-S-S-Q-R-L-Q-pT-D-G-S-A-T-N
465	P15	1998	S-T-Q-F-S-S-Q-R-L-Q-T-D-G-pS-A-T-N
466	P16	1998.1	S-T-Q-F-S-S-Q-R-L-Q-T-D-G-S-A-pT-N
467	P17	1856	T-Q-F-S-S-Q-R-L-Q-T-D-G-S-A-T-N-D
468	P18	2026.1	pT-Q-F-S-S-Q-R-L-Q-T-D-G-S-A-T-N-D
469	P19	2026	T-Q-F-pS-S-Q-R-L-Q-T-D-G-S-A-T-N-D
470	P20	2026	T-Q-F-S-pS-Q-R-L-Q-T-D-G-S-A-T-N-D
471	P21	2026.1	T-Q-F-S-S-Q-R-L-Q-pT-D-G-S-A-T-N-D
472	P22	2026	T-Q-F-S-S-Q-R-L-Q-T-D-G-pS-A-T-N-D
473	P23	2026.1	T-Q-F-S-S-Q-R-L-Q-T-D-G-S-A-pT-N-D
474	P24	1886.1	Q-F-S-S-Q-R-L-Q-T-D-G-S-A-T-N-D-M
475	P25	2056.1	Q-F-pS-S-Q-R-L-Q-T-D-G-S-A-T-N-D-M
476	P26	2056.1	Q-F-S-pS-Q-R-L-Q-T-D-G-S-A-T-N-D-M
477	P27	2056.2	Q-F-S-S-Q-R-L-Q-pT-D-G-S-A-T-N-D-M
478	P28	2056.1	Q-F-S-S-Q-R-L-Q-T-D-G-pS-A-T-N-D-M
479	P29	2056.2	Q-F-S-S-Q-R-L-Q-T-D-G-S-A-pT-N-D-M
480	P30	1887.1	F-S-S-Q-R-L-Q-T-D-G-S-A-T-N-D-M-E
481	Q 1	2057.1	F-pS-S-Q-R-L-Q-T-D-G-S-A-T-N-D-M-E
482	Q 2	2057.1	F-S-pS-Q-R-L-Q-T-D-G-S-A-T-N-D-M-E
483	Q 3	2057.2	F-S-S-Q-R-L-Q-pT-D-G-S-A-T-N-D-M-E
484	Q 4	2057.1	F-S-S-Q-R-L-Q-T-D-G-pS-A-T-N-D-M-E
485	Q 5	2057.2	F-S-S-Q-R-L-Q-T-D-G-S-A-pT-N-D-M-E

7.1.2 Part 2

Nr.	Pos.	Mol. Weight	Sequence
1	A 1	2039.5	G-N-E-P-A-R-S-F-K-S-R-A-L-N-H-V-K-K
2	A 2	2209.5	G-N-E-P-A-R-pS-F-K-S-R-A-L-N-H-V-K-K
3	A 3	2209.5	G-N-E-P-A-R-S-F-K-pS-R-A-L-N-H-V-K-K
4	A 4	2082.5	E-P-A-R-S-F-K-S-R-A-L-N-H-V-K-K-V-D
5	A 5	2252.5	E-P-A-R-pS-F-K-S-R-A-L-N-H-V-K-K-V-D
6	A 6	2252.5	E-P-A-R-S-F-K-pS-R-A-L-N-H-V-K-K-V-D
7	A 7	2070.5	A-R-S-F-K-S-R-A-L-N-H-V-K-K-V-D-D-V
8	A 8	2240.5	A-R-pS-F-K-S-R-A-L-N-H-V-K-K-V-D-D-V
9	A 9	2240.5	A-R-S-F-K-pS-R-A-L-N-H-V-K-K-V-D-D-V
10	A10	2001.4	S-F-K-S-R-A-L-N-H-V-K-K-V-D-D-V-T-G
11	A11	2171.4	pS-F-K-S-R-A-L-N-H-V-K-K-V-D-D-V-T-G
12	A12	2171.4	S-F-K-pS-R-A-L-N-H-V-K-K-V-D-D-V-T-G
13	A13	2171.5	S-F-K-S-R-A-L-N-H-V-K-K-V-D-D-V-pT-G
14	A14	2024.4	K-S-R-A-L-N-H-V-K-K-V-D-D-V-T-G-E-K
15	A15	2194.4	K-pS-R-A-L-N-H-V-K-K-V-D-D-V-T-G-E-K
16	A16	2194.5	K-S-R-A-L-N-H-V-K-K-V-D-D-V-pT-G-E-K
17	A17	2064.4	R-A-L-N-H-V-K-K-V-D-D-V-T-G-E-K-V-R
18	A18	2234.5	R-A-L-N-H-V-K-K-V-D-D-V-pT-G-E-K-V-R
19	A19	2037.3	L-N-H-V-K-K-V-D-D-V-T-G-E-K-V-R-E-A
20	A20	2207.4	L-N-H-V-K-K-V-D-D-V-pT-G-E-K-V-R-E-A
21	A21	2086.3	H-V-K-K-V-D-D-V-T-G-E-K-V-R-E-A-F-E
22	A22	2256.4	H-V-K-K-V-D-D-V-pT-G-E-K-V-R-E-A-F-E
23	A23	2125.4	K-K-V-D-D-V-T-G-E-K-V-R-E-A-F-E-Q-F
24	A24	2295.5	K-K-V-D-D-V-pT-G-E-K-V-R-E-A-F-E-Q-F
25	A25	2111.3	V-D-D-V-T-G-E-K-V-R-E-A-F-E-Q-F-L-E
26	A26	2281.4	V-D-D-V-pT-G-E-K-V-R-E-A-F-E-Q-F-L-E
27	A27	2159.4	D-V-T-G-E-K-V-R-E-A-F-E-Q-F-L-E-D-F
28	A28	2329.5	D-V-pT-G-E-K-V-R-E-A-F-E-Q-F-L-E-D-F
29	A29	2131.4	T-G-E-K-V-R-E-A-F-E-Q-F-L-E-D-F-S-V
30	A30	2301.5	pT-G-E-K-V-R-E-A-F-E-Q-F-L-E-D-F-S-V
31	A31	2301.4	T-G-E-K-V-R-E-A-F-E-Q-F-L-E-D-F-pS-V
32	A32	2188.4	E-K-V-R-E-A-F-E-Q-F-L-E-D-F-S-V-Q-S
33	A33	2358.4	E-K-V-R-E-A-F-E-Q-F-L-E-D-F-pS-V-Q-S
34	A34	2358.4	E-K-V-R-E-A-F-E-Q-F-L-E-D-F-S-V-Q-pS
35	A35	2147.3	V-R-E-A-F-E-Q-F-L-E-D-F-S-V-Q-S-T-D
36	A36	2317.3	V-R-E-A-F-E-Q-F-L-E-D-F-pS-V-Q-S-T-D
37	A37	2317.3	V-R-E-A-F-E-Q-F-L-E-D-F-S-V-Q-pS-T-D
38	B 1	2317.4	V-R-E-A-F-E-Q-F-L-E-D-F-S-V-Q-S-pT-D
39	B 2	2050.2	E-A-F-E-Q-F-L-E-D-F-S-V-Q-S-T-D-T-G
40	B 3	2220.2	E-A-F-E-Q-F-L-E-D-F-pS-V-Q-S-T-D-T-G
41	B 4	2220.2	E-A-F-E-Q-F-L-E-D-F-S-V-Q-pS-T-D-T-G
42	B 5	2220.3	E-A-F-E-Q-F-L-E-D-F-S-V-Q-S-pT-D-T-G
43	B 6	2220.3	E-A-F-E-Q-F-L-E-D-F-S-V-Q-S-T-D-pT-G
44	B 7	2078.2	F-E-Q-F-L-E-D-F-S-V-Q-S-T-D-T-G-E-V
45	B 8	2248.2	F-E-Q-F-L-E-D-F-pS-V-Q-S-T-D-T-G-E-V
46	B 9	2248.2	F-E-Q-F-L-E-D-F-S-V-Q-pS-T-D-T-G-E-V

47	B10	2248.3	F-E-Q-F-L-E-D-F-S-V-Q-S-pT-D-T-G-E-V
48	B11	2248.3	F-E-Q-F-L-E-D-F-S-V-Q-S-T-D-pT-G-E-V
49	B12	2059.2	Q-F-L-E-D-F-S-V-Q-S-T-D-T-G-E-V-E-K
50	B13	2229.2	Q-F-L-E-D-F-pS-V-Q-S-T-D-T-G-E-V-E-K
51	B14	2229.2	Q-F-L-E-D-F-S-V-Q-pS-T-D-T-G-E-V-E-K
52	B15	2229.3	Q-F-L-E-D-F-S-V-Q-S-pT-D-T-G-E-V-E-K
53	B16	2229.3	Q-F-L-E-D-F-S-V-Q-S-T-D-pT-G-E-V-E-K
54	B17	2046.2	L-E-D-F-S-V-Q-S-T-D-T-G-E-V-E-K-V-Y
55	B18	2216.2	L-E-D-F-pS-V-Q-S-T-D-T-G-E-V-E-K-V-Y
56	B19	2216.2	L-E-D-F-S-V-Q-pS-T-D-T-G-E-V-E-K-V-Y
57	B20	2216.3	L-E-D-F-S-V-Q-S-pT-D-T-G-E-V-E-K-V-Y
58	B21	2216.3	L-E-D-F-S-V-Q-S-T-D-pT-G-E-V-E-K-V-Y
59	B22	2031.2	D-F-S-V-Q-S-T-D-T-G-E-V-E-K-V-Y-R-A
60	B23	2201.2	D-F-pS-V-Q-S-T-D-T-G-E-V-E-K-V-Y-R-A
61	B24	2201.2	D-F-S-V-Q-pS-T-D-T-G-E-V-E-K-V-Y-R-A
62	B25	2201.3	D-F-S-V-Q-S-pT-D-T-G-E-V-E-K-V-Y-R-A
63	B26	2201.3	D-F-S-V-Q-S-T-D-pT-G-E-V-E-K-V-Y-R-A
64	B27	2010.2	S-V-Q-S-T-D-T-G-E-V-E-K-V-Y-R-A-Q-I
65	B28	2180.2	pS-V-Q-S-T-D-T-G-E-V-E-K-V-Y-R-A-Q-I
66	B29	2180.2	S-V-Q-pS-T-D-T-G-E-V-E-K-V-Y-R-A-Q-I
67	B30	2180.3	S-V-Q-S-pT-D-T-G-E-V-E-K-V-Y-R-A-Q-I
68	B31	2180.3	S-V-Q-S-T-D-pT-G-E-V-E-K-V-Y-R-A-Q-I
69	B32	2100.3	Q-S-T-D-T-G-E-V-E-K-V-Y-R-A-Q-I-E-F
70	B33	2270.3	Q-pS-T-D-T-G-E-V-E-K-V-Y-R-A-Q-I-E-F
71	B34	2270.4	Q-S-pT-D-T-G-E-V-E-K-V-Y-R-A-Q-I-E-F
72	B35	2270.4	Q-S-T-D-pT-G-E-V-E-K-V-Y-R-A-Q-I-E-F
73	B36	2144.5	T-D-T-G-E-V-E-K-V-Y-R-A-Q-I-E-F-M-K
74	B37	2314.6	pT-D-T-G-E-V-E-K-V-Y-R-A-Q-I-E-F-M-K
75	C 1	2314.6	T-D-pT-G-E-V-E-K-V-Y-R-A-Q-I-E-F-M-K
76	C 2	2204.7	T-G-E-V-E-K-V-Y-R-A-Q-I-E-F-M-K-I-Y
77	C 3	2374.8	pT-G-E-V-E-K-V-Y-R-A-Q-I-E-F-M-K-I-Y
78	C 4	2274.8	E-V-E-K-V-Y-R-A-Q-I-E-F-M-K-I-Y-D-L
79	C 5	2261.8	E-K-V-Y-R-A-Q-I-E-F-M-K-I-Y-D-L-N-T
80	C 6	2431.9	E-K-V-Y-R-A-Q-I-E-F-M-K-I-Y-D-L-N-pT
81	C 7	2280.9	V-Y-R-A-Q-I-E-F-M-K-I-Y-D-L-N-T-I-Y
82	C 8	2451	V-Y-R-A-Q-I-E-F-M-K-I-Y-D-L-N-pT-I-Y
83	C 9	2246.9	R-A-Q-I-E-F-M-K-I-Y-D-L-N-T-I-Y-I-D
84	C10	2417	R-A-Q-I-E-F-M-K-I-Y-D-L-N-pT-I-Y-I-D
85	C11	2310.9	Q-I-E-F-M-K-I-Y-D-L-N-T-I-Y-I-D-Y-Q
86	C12	2481	Q-I-E-F-M-K-I-Y-D-L-N-pT-I-Y-I-D-Y-Q
87	C13	2319.9	E-F-M-K-I-Y-D-L-N-T-I-Y-I-D-Y-Q-H-L
88	C14	2490	E-F-M-K-I-Y-D-L-N-pT-I-Y-I-D-Y-Q-H-L
89	C15	2261.9	M-K-I-Y-D-L-N-T-I-Y-I-D-Y-Q-H-L-S-M
90	C16	2432	M-K-I-Y-D-L-N-pT-I-Y-I-D-Y-Q-H-L-S-M
91	C17	2287.8	I-Y-D-L-N-T-I-Y-I-D-Y-Q-H-L-S-M-R-E
92	C18	2457.9	I-Y-D-L-N-pT-I-Y-I-D-Y-Q-H-L-S-M-R-E
93	C19	2457.8	I-Y-D-L-N-T-I-Y-I-D-Y-Q-H-L-pS-M-R-E
94	C20	2182.6	D-L-N-T-I-Y-I-D-Y-Q-H-L-S-M-R-E-N-G
95	C21	2352.7	D-L-N-pT-I-Y-I-D-Y-Q-H-L-S-M-R-E-N-G
96	C22	2352.6	D-L-N-T-I-Y-I-D-Y-Q-H-L-pS-M-R-E-N-G

97	C23	2138.6	N-T-I-Y-I-D-Y-Q-H-L-S-M-R-E-N-G-A-L
98	C24	2308.7	N-pT-I-Y-I-D-Y-Q-H-L-S-M-R-E-N-G-A-L
99	C25	2308.6	N-T-I-Y-I-D-Y-Q-H-L-pS-M-R-E-N-G-A-L
100	C26	2125.7	I-Y-I-D-Y-Q-H-L-S-M-R-E-N-G-A-L-A-M
101	C27	2295.7	I-Y-I-D-Y-Q-H-L-pS-M-R-E-N-G-A-L-A-M
102	C28	2033.6	I-D-Y-Q-H-L-S-M-R-E-N-G-A-L-A-M-A-I
103	C29	2203.6	I-D-Y-Q-H-L-pS-M-R-E-N-G-A-L-A-M-A-I
104	C30	2021.5	Y-Q-H-L-S-M-R-E-N-G-A-L-A-M-A-I-S-E
105	C31	2191.5	Y-Q-H-L-pS-M-R-E-N-G-A-L-A-M-A-I-S-E
106	C32	2191.5	Y-Q-H-L-S-M-R-E-N-G-A-L-A-M-A-I-pS-E
107	C33	2021.5	H-L-S-M-R-E-N-G-A-L-A-M-A-I-S-E-Q-Y
108	C34	2191.5	H-L-pS-M-R-E-N-G-A-L-A-M-A-I-S-E-Q-Y
109	C35	2191.5	H-L-S-M-R-E-N-G-A-L-A-M-A-I-pS-E-Q-Y
110	C36	2090.6	S-M-R-E-N-G-A-L-A-M-A-I-S-E-Q-Y-Y-R
111	C37	2260.6	pS-M-R-E-N-G-A-L-A-M-A-I-S-E-Q-Y-Y-R
112	D 1	2260.6	S-M-R-E-N-G-A-L-A-M-A-I-pS-E-Q-Y-Y-R
113	D 2	2132.7	R-E-N-G-A-L-A-M-A-I-S-E-Q-Y-Y-R-F-L
114	D 3	2302.7	R-E-N-G-A-L-A-M-A-I-pS-E-Q-Y-Y-R-F-L
115	D 4	2091.7	N-G-A-L-A-M-A-I-S-E-Q-Y-Y-R-F-L-P-F
116	D 5	2261.7	N-G-A-L-A-M-A-I-pS-E-Q-Y-Y-R-F-L-P-F
117	D 6	2161.8	A-L-A-M-A-I-S-E-Q-Y-Y-R-F-L-P-F-L-Q
118	D 7	2331.8	A-L-A-M-A-I-pS-E-Q-Y-Y-R-F-L-P-F-L-Q
119	D 8	2162.8	A-M-A-I-S-E-Q-Y-Y-R-F-L-P-F-L-Q-K-G
120	D 9	2332.8	A-M-A-I-pS-E-Q-Y-Y-R-F-L-P-F-L-Q-K-G
121	D10	2229.9	A-I-S-E-Q-Y-Y-R-F-L-P-F-L-Q-K-G-L-R
122	D11	2399.9	A-I-pS-E-Q-Y-Y-R-F-L-P-F-L-Q-K-G-L-R
123	D12	2300.9	S-E-Q-Y-Y-R-F-L-P-F-L-Q-K-G-L-R-R-V
124	D13	2470.9	pS-E-Q-Y-Y-R-F-L-P-F-L-Q-K-G-L-R-R-V
125	D14	2340	Q-Y-Y-R-F-L-P-F-L-Q-K-G-L-R-R-V-V-R
126	D15	2340.1	Y-R-F-L-P-F-L-Q-K-G-L-R-R-V-V-R-K-Y
127	D16	2188.9	F-L-P-F-L-Q-K-G-L-R-R-V-V-R-K-Y-A-P
128	D17	2170.8	P-F-L-Q-K-G-L-R-R-V-V-R-K-Y-A-P-E-L
129	D18	2153.8	L-Q-K-G-L-R-R-V-V-R-K-Y-A-P-E-L-L-N
130	D19	2153.8	L-Q-K-G-L-R-R-V-V-R-K-Y-A-P-E-L-L-N
131	D20	2100.7	K-G-L-R-R-V-V-R-K-Y-A-P-E-L-L-N-T-S
132	D21	2270.8	K-G-L-R-R-V-V-R-K-Y-A-P-E-L-L-N-pT-S
133	D22	2270.7	K-G-L-R-R-V-V-R-K-Y-A-P-E-L-L-N-T-pS
134	D23	2117.6	L-R-R-V-V-R-K-Y-A-P-E-L-L-N-T-S-D-S
135	D24	2287.7	L-R-R-V-V-R-K-Y-A-P-E-L-L-N-pT-S-D-S
136	D25	2287.6	L-R-R-V-V-R-K-Y-A-P-E-L-L-N-T-pS-D-S
137	D26	2287.6	L-R-R-V-V-R-K-Y-A-P-E-L-L-N-T-S-D-pS
138	D27	2089.6	R-V-V-R-K-Y-A-P-E-L-L-N-T-S-D-S-L-K
139	D28	2259.7	R-V-V-R-K-Y-A-P-E-L-L-N-pT-S-D-S-L-K
140	D29	2259.6	R-V-V-R-K-Y-A-P-E-L-L-N-T-pS-D-S-L-K
141	D30	2259.6	R-V-V-R-K-Y-A-P-E-L-L-N-T-S-D-pS-L-K
142	D31	2077.6	V-R-K-Y-A-P-E-L-L-N-T-S-D-S-L-K-R-S
143	D32	2247.7	V-R-K-Y-A-P-E-L-L-N-pT-S-D-S-L-K-R-S
144	D33	2247.6	V-R-K-Y-A-P-E-L-L-N-T-pS-D-S-L-K-R-S
145	D34	2247.6	V-R-K-Y-A-P-E-L-L-N-T-S-D-pS-L-K-R-S
146	D35	2008.5	K-Y-A-P-E-L-L-N-T-S-D-S-L-K-R-S-E-G

147	D36	2178.6	K-Y-A-P-E-L-L-N-pT-S-D-S-L-K-R-S-E-G
148	D37	2178.5	K-Y-A-P-E-L-L-N-T-pS-D-S-L-K-R-S-E-G
149	E 1	2178.5	K-Y-A-P-E-L-L-N-T-S-D-pS-L-K-R-S-E-G
150	E 2	2178.5	K-Y-A-P-E-L-L-N-T-S-D-S-L-K-R-pS-E-G
151	E 3	1961.3	A-P-E-L-L-N-T-S-D-S-L-K-R-S-E-G-D-E
152	E 4	2131.4	A-P-E-L-L-N-pT-S-D-S-L-K-R-S-E-G-D-E
153	E 5	2131.3	A-P-E-L-L-N-T-pS-D-S-L-K-R-S-E-G-D-E
154	E 6	2131.3	A-P-E-L-L-N-T-S-D-pS-L-K-R-S-E-G-D-E
155	E 7	2131.3	A-P-E-L-L-N-T-S-D-S-L-K-R-pS-E-G-D-E
156	E 8	1978.3	E-L-L-N-T-S-D-S-L-K-R-S-E-G-D-E-G-Q
157	E 9	2148.4	E-L-L-N-pT-S-D-S-L-K-R-S-E-G-D-E-G-Q
158	E10	2148.3	E-L-L-N-T-pS-D-S-L-K-R-S-E-G-D-E-G-Q
159	E11	2148.3	E-L-L-N-T-S-D-pS-L-K-R-S-E-G-D-E-G-Q
160	E12	2148.3	E-L-L-N-T-S-D-S-L-K-R-pS-E-G-D-E-G-Q
161	E13	1922.2	L-N-T-S-D-S-L-K-R-S-E-G-D-E-G-Q-A-D
162	E14	2092.3	L-N-pT-S-D-S-L-K-R-S-E-G-D-E-G-Q-A-D
163	E15	2092.2	L-N-T-pS-D-S-L-K-R-S-E-G-D-E-G-Q-A-D
164	E16	2092.2	L-N-T-S-D-pS-L-K-R-S-E-G-D-E-G-Q-A-D
165	E17	2092.2	L-N-T-S-D-S-L-K-R-pS-E-G-D-E-G-Q-A-D
166	E18	1939.1	T-S-D-S-L-K-R-S-E-G-D-E-G-Q-A-D-E-D
167	E19	2109.2	pT-S-D-S-L-K-R-S-E-G-D-E-G-Q-A-D-E-D
168	E20	2109.1	T-pS-D-S-L-K-R-S-E-G-D-E-G-Q-A-D-E-D
169	E21	2109.1	T-S-D-pS-L-K-R-S-E-G-D-E-G-Q-A-D-E-D
170	E22	2109.1	T-S-D-pS-L-K-R-S-E-G-D-E-G-Q-A-D-E-D
171	E23	2109.1	T-S-D-S-L-K-R-pS-E-G-D-E-G-Q-A-D-E-D
172	E24	2008.1	D-S-L-K-R-S-E-G-D-E-G-Q-A-D-E-D-E-Q
173	E25	2178.1	D-pS-L-K-R-S-E-G-D-E-G-Q-A-D-E-D-E-Q
174	E26	2178.1	D-S-L-K-R-pS-E-G-D-E-G-Q-A-D-E-D-E-Q
175	E27	2049.1	L-K-R-S-E-G-D-E-G-Q-A-D-E-D-E-Q-Q-D
176	E28	2219.1	L-K-R-pS-E-G-D-E-G-Q-A-D-E-D-E-Q-Q-D
177	E29	2037.9	R-S-E-G-D-E-G-Q-A-D-E-D-E-Q-Q-D-D-D
178	E30	2207.9	R-pS-E-G-D-E-G-Q-A-D-E-D-E-Q-Q-D-D-D
179	E31	2039.9	E-G-D-E-G-Q-A-D-E-D-E-Q-Q-D-D-D-M-N
180	E32	1997.9	D-E-G-Q-A-D-E-D-E-Q-Q-D-D-D-M-N-G-S
181	E33	2167.9	D-E-G-Q-A-D-E-D-E-Q-Q-D-D-D-M-N-G-pS
182	E34	1954	G-Q-A-D-E-D-E-Q-Q-D-D-D-M-N-G-S-S-L
183	E35	2124	G-Q-A-D-E-D-E-Q-Q-D-D-D-M-N-G-pS-S-L
184	E36	2124	G-Q-A-D-E-D-E-Q-Q-D-D-D-M-N-G-S-pS-L
185	E37	2022.1	A-D-E-D-E-Q-Q-D-D-D-M-N-G-S-S-L-P-R
186	F 1	2192.1	A-D-E-D-E-Q-Q-D-D-D-M-N-G-pS-S-L-P-R
187	F 2	2192.1	A-D-E-D-E-Q-Q-D-D-D-M-N-G-S-pS-L-P-R
188	F 3	2038.1	E-D-E-Q-Q-D-D-D-M-N-G-S-S-L-P-R-D-S
189	F 4	2208.1	E-D-E-Q-Q-D-D-D-M-N-G-pS-S-L-P-R-D-S
190	F 5	2208.1	E-D-E-Q-Q-D-D-D-M-N-G-S-pS-L-P-R-D-S
191	F 6	2208.1	E-D-E-Q-Q-D-D-D-M-N-G-S-S-L-P-R-D-pS
192	F 7	1938.1	E-Q-Q-D-D-D-M-N-G-S-S-L-P-R-D-S-G-S
193	F 8	2108.1	E-Q-Q-D-D-D-M-N-G-pS-S-L-P-R-D-S-G-S
194	F 9	2108.1	E-Q-Q-D-D-D-M-N-G-S-pS-L-P-R-D-S-G-S
195	F10	2108.1	E-Q-Q-D-D-D-M-N-G-S-S-L-P-R-D-pS-G-S
196	F11	2108.1	E-Q-Q-D-D-D-M-N-G-S-S-L-P-R-D-S-G-pS

197	F12	1839.1	Q-D-D-D-M-N-G-S-S-L-P-R-D-S-G-S-S-A
198	F13	2009.1	Q-D-D-D-M-N-G-pS-S-L-P-R-D-S-G-S-S-A
199	F14	2009.1	Q-D-D-D-M-N-G-S-pS-L-P-R-D-S-G-S-S-A
200	F15	2009.1	Q-D-D-D-M-N-G-S-S-L-P-R-D-pS-G-S-S-A
201	F16	2009.1	Q-D-D-D-M-N-G-S-S-L-P-R-D-S-G-pS-S-A
202	F17	2009.1	Q-D-D-D-M-N-G-S-S-L-P-R-D-S-G-S-pS-A
203	F18	1764.1	D-D-M-N-G-S-S-L-P-R-D-S-G-S-S-A-A-P
204	F19	1934.1	D-D-M-N-G-pS-S-L-P-R-D-S-G-S-S-A-A-P
205	F20	1934.1	D-D-M-N-G-S-pS-L-P-R-D-S-G-S-S-A-A-P
206	F21	1934.1	D-D-M-N-G-S-S-L-P-R-D-pS-G-S-S-A-A-P
207	F22	1934.1	D-D-M-N-G-S-S-L-P-R-D-S-G-pS-S-A-A-P
208	F23	1934.1	D-D-M-N-G-S-S-L-P-R-D-S-G-S-pS-A-A-P
209	F24	1705.1	M-N-G-S-S-L-P-R-D-S-G-S-S-A-A-P-G-N
210	F25	1875.1	M-N-G-pS-S-L-P-R-D-S-G-S-S-A-A-P-G-N
211	F26	1875.1	M-N-G-S-pS-L-P-R-D-S-G-S-S-A-A-P-G-N
212	F27	1875.1	M-N-G-S-S-L-P-R-D-pS-G-S-S-A-A-P-G-N
213	F28	1875.1	M-N-G-S-S-L-P-R-D-S-G-pS-S-A-A-P-G-N
214	F29	1875.1	M-N-G-S-S-L-P-R-D-S-G-S-pS-A-A-P-G-N
215	F30	1618	G-S-S-L-P-R-D-S-G-S-S-A-A-P-G-N-G-T
216	F31	1788	G-pS-S-L-P-R-D-S-G-S-S-A-A-P-G-N-G-T
217	F32	1788	G-S-pS-L-P-R-D-S-G-S-S-A-A-P-G-N-G-T
218	F33	1788	G-S-S-L-P-R-D-pS-G-S-S-A-A-P-G-N-G-T
219	F34	1788	G-S-S-L-P-R-D-S-G-pS-S-A-A-P-G-N-G-T
220	F35	1788	G-S-S-L-P-R-D-S-G-S-pS-A-A-P-G-N-G-T
221	F36	1788.1	G-S-S-L-P-R-D-S-G-S-S-A-A-P-G-N-G-pT
222	F37	1632	S-L-P-R-D-S-G-S-S-A-A-P-G-N-G-T-S-A
223	G 1	1802	pS-L-P-R-D-S-G-S-S-A-A-P-G-N-G-T-S-A
224	G 2	1802	S-L-P-R-D-pS-G-S-S-A-A-P-G-N-G-T-S-A
225	G 3	1802	S-L-P-R-D-S-G-pS-S-A-A-P-G-N-G-T-S-A
226	G 4	1802	S-L-P-R-D-S-G-S-pS-A-A-P-G-N-G-T-S-A
227	G 5	1802.1	S-L-P-R-D-S-G-S-S-A-A-P-G-N-G-pT-S-A
228	G 6	1802	S-L-P-R-D-S-G-S-S-A-A-P-G-N-G-T-pS-A
229	G 7	1634	P-R-D-S-G-S-S-A-A-P-G-N-G-T-S-A-M-A
230	G 8	1804	P-R-D-pS-G-S-S-A-A-P-G-N-G-T-S-A-M-A
231	G 9	1804	P-R-D-S-G-pS-S-A-A-P-G-N-G-T-S-A-M-A
232	G10	1804	P-R-D-S-G-S-pS-A-A-P-G-N-G-T-S-A-M-A
233	G11	1804.1	P-R-D-S-G-S-S-A-A-P-G-N-G-pT-S-A-M-A
234	G12	1804	P-R-D-S-G-S-S-A-A-P-G-N-G-T-pS-A-M-A
235	G13	1638	D-S-G-S-S-A-A-P-G-N-G-T-S-A-M-A-T-R
236	G14	1808	D-pS-G-S-S-A-A-P-G-N-G-T-S-A-M-A-T-R
237	G15	1808	D-S-G-pS-S-A-A-P-G-N-G-T-S-A-M-A-T-R
238	G16	1808	D-S-G-S-pS-A-A-P-G-N-G-T-S-A-M-A-T-R
239	G17	1808.1	D-S-G-S-S-A-A-P-G-N-G-pT-S-A-M-A-T-R
240	G18	1808	D-S-G-S-S-A-A-P-G-N-G-T-pS-A-M-A-T-R
241	G19	1808.1	D-S-G-S-S-A-A-P-G-N-G-T-S-A-M-A-pT-R
242	G20	1636.1	G-S-S-A-A-P-G-N-G-T-S-A-M-A-T-R-S-I
243	G21	1806.1	G-pS-S-A-A-P-G-N-G-T-S-A-M-A-T-R-S-I
244	G22	1806.1	G-S-pS-A-A-P-G-N-G-T-S-A-M-A-T-R-S-I
245	G23	1806.2	G-S-S-A-A-P-G-N-G-pT-S-A-M-A-T-R-S-I
246	G24	1806.1	G-S-S-A-A-P-G-N-G-T-pS-A-M-A-T-R-S-I

247	G25	1806.2	G-S-S-A-A-P-G-N-G-T-S-A-M-A-pT-R-S-I
248	G26	1806.1	G-S-S-A-A-P-G-N-G-T-S-A-M-A-T-R-pS-I
249	G27	1694.1	S-A-A-P-G-N-G-T-S-A-M-A-T-R-S-I-T-T
250	G28	1864.1	pS-A-A-P-G-N-G-T-S-A-M-A-T-R-S-I-T-T
251	G29	1864.2	S-A-A-P-G-N-G-pT-S-A-M-A-T-R-S-I-T-T
252	G30	1864.1	S-A-A-P-G-N-G-T-pS-A-M-A-T-R-S-I-T-T
253	G31	1864.2	S-A-A-P-G-N-G-T-S-A-M-A-pT-R-S-I-T-T
254	G32	1864.1	S-A-A-P-G-N-G-T-S-A-M-A-T-R-pS-I-T-T
255	G33	1864.2	S-A-A-P-G-N-G-T-S-A-M-A-T-R-S-I-pT-T
256	G34	1864.2	S-A-A-P-G-N-G-T-S-A-M-A-T-R-S-I-T-pT
257	G35	1724.1	A-P-G-N-G-T-S-A-M-A-T-R-S-I-T-T-S-T
258	G36	1894.2	A-P-G-N-G-pT-S-A-M-A-T-R-S-I-T-T-S-T
259	G37	1894.1	A-P-G-N-G-T-pS-A-M-A-T-R-S-I-T-T-S-T
260	H 1	1894.2	A-P-G-N-G-T-S-A-M-A-pT-R-S-I-T-T-S-T
261	H 2	1894.1	A-P-G-N-G-T-S-A-M-A-T-R-pS-I-T-T-S-T
262	H 3	1894.2	A-P-G-N-G-T-S-A-M-A-T-R-S-I-pT-T-S-T
263	H 4	1894.2	A-P-G-N-G-T-S-A-M-A-T-R-S-I-T-pT-S-T
264	H 5	1894.1	A-P-G-N-G-T-S-A-M-A-T-R-S-I-T-T-pS-T
265	H 6	1894.2	A-P-G-N-G-T-S-A-M-A-T-R-S-I-T-T-S-pT
266	H 7	1740.1	G-N-G-T-S-A-M-A-T-R-S-I-T-T-S-T-S-P
267	H 8	1910.2	G-N-G-pT-S-A-M-A-T-R-S-I-T-T-S-T-S-P
268	H 9	1910.1	G-N-G-T-pS-A-M-A-T-R-S-I-T-T-S-T-S-P
269	H10	1910.2	G-N-G-T-S-A-M-A-pT-R-S-I-T-T-S-T-S-P
270	H11	1910.1	G-N-G-T-S-A-M-A-T-R-pS-I-T-T-S-T-S-P
271	H12	1910.2	G-N-G-T-S-A-M-A-T-R-S-I-pT-T-S-T-S-P
272	H13	1910.2	G-N-G-T-S-A-M-A-T-R-S-I-T-pT-S-T-S-P
273	H14	1910.1	G-N-G-T-S-A-M-A-T-R-S-I-T-T-pS-T-S-P
274	H15	1910.2	G-N-G-T-S-A-M-A-T-R-S-I-T-T-S-pT-S-P
275	H16	1910.1	G-N-G-T-S-A-M-A-T-R-S-I-T-T-S-T-pS-P
276	H17	1826.1	G-T-S-A-M-A-T-R-S-I-T-T-S-T-S-P-E-Q
277	H18	1996.2	G-pT-S-A-M-A-T-R-S-I-T-T-S-T-S-P-E-Q
278	H19	1996.1	G-T-pS-A-M-A-T-R-S-I-T-T-S-T-S-P-E-Q
279	H20	1996.2	G-T-S-A-M-A-pT-R-S-I-T-T-S-T-S-P-E-Q
280	H21	1996.1	G-T-S-A-M-A-T-R-pS-I-T-T-S-T-S-P-E-Q
281	H22	1996.2	G-T-S-A-M-A-T-R-S-I-pT-T-S-T-S-P-E-Q
282	H23	1996.2	G-T-S-A-M-A-T-R-S-I-T-pT-S-T-S-P-E-Q
283	H24	1996.1	G-T-S-A-M-A-T-R-S-I-T-T-pS-T-S-P-E-Q
284	H25	1996.2	G-T-S-A-M-A-T-R-S-I-T-T-S-pT-S-P-E-Q
285	H26	1996.1	G-T-S-A-M-A-T-R-S-I-T-T-S-T-pS-P-E-Q
286	H27	1898.1	S-A-M-A-T-R-S-I-T-T-S-T-S-P-E-Q-T-E
287	H28	2068.1	pS-A-M-A-T-R-S-I-T-T-S-T-S-P-E-Q-T-E
288	H29	2068.2	S-A-M-A-pT-R-S-I-T-T-S-T-S-P-E-Q-T-E
289	H30	2068.1	S-A-M-A-T-R-pS-I-T-T-S-T-S-P-E-Q-T-E
290	H31	2068.2	S-A-M-A-T-R-S-I-pT-T-S-T-S-P-E-Q-T-E
291	H32	2068.2	S-A-M-A-T-R-S-I-T-pT-S-T-S-P-E-Q-T-E
292	H33	2068.1	S-A-M-A-T-R-S-I-T-T-pS-T-S-P-E-Q-T-E
293	H34	2068.2	S-A-M-A-T-R-S-I-T-T-S-pT-S-P-E-Q-T-E
294	H35	2068.1	S-A-M-A-T-R-S-I-T-T-S-T-pS-P-E-Q-T-E
295	H36	2068.2	S-A-M-A-T-R-S-I-T-T-S-T-S-P-E-Q-pT-E
296	H37	1995.2	M-A-T-R-S-I-T-T-S-T-S-P-E-Q-T-E-R-V

297	I 1	2165.3	M-A-pT-R-S-I-T-T-S-T-S-P-E-Q-T-E-R-V
298	I 2	2165.2	M-A-T-R-pS-I-T-T-S-T-S-P-E-Q-T-E-R-V
299	I 3	2165.3	M-A-T-R-S-I-pT-T-S-T-S-P-E-Q-T-E-R-V
300	I 4	2165.3	M-A-T-R-S-I-T-pT-S-T-S-P-E-Q-T-E-R-V
301	I 5	2165.2	M-A-T-R-S-I-T-T-pS-T-S-P-E-Q-T-E-R-V
302	I 6	2165.3	M-A-T-R-S-I-T-T-S-pT-S-P-E-Q-T-E-R-V
303	I 7	2165.2	M-A-T-R-S-I-T-T-S-T-pS-P-E-Q-T-E-R-V
304	I 8	2165.3	M-A-T-R-S-I-T-T-S-T-S-P-E-Q-pT-E-R-V
305	I 9	2068.2	T-R-S-I-T-T-S-T-S-P-E-Q-T-E-R-V-F-Q
306	I10	2238.3	pT-R-S-I-T-T-S-T-S-P-E-Q-T-E-R-V-F-Q
307	I11	2238.2	T-R-pS-I-T-T-S-T-S-P-E-Q-T-E-R-V-F-Q
308	I12	2238.3	T-R-S-I-pT-T-S-T-S-P-E-Q-T-E-R-V-F-Q
309	I13	2238.3	T-R-S-I-T-pT-S-T-S-P-E-Q-T-E-R-V-F-Q
310	I14	2238.2	T-R-S-I-T-T-pS-T-S-P-E-Q-T-E-R-V-F-Q
311	I15	2238.3	T-R-S-I-T-T-S-pT-S-P-E-Q-T-E-R-V-F-Q
312	I16	2238.2	T-R-S-I-T-T-S-T-pS-P-E-Q-T-E-R-V-F-Q
313	I17	2238.3	T-R-S-I-T-T-S-T-S-P-E-Q-pT-E-R-V-F-Q
314	I18	2011.2	S-I-T-T-S-T-S-P-E-Q-T-E-R-V-F-Q-I-S
315	I19	2181.2	pS-I-T-T-S-T-S-P-E-Q-T-E-R-V-F-Q-I-S
316	I20	2181.3	S-I-pT-T-S-T-S-P-E-Q-T-E-R-V-F-Q-I-S
317	I21	2181.3	S-I-T-pT-S-T-S-P-E-Q-T-E-R-V-F-Q-I-S
318	I22	2181.2	S-I-T-T-pS-T-S-P-E-Q-T-E-R-V-F-Q-I-S
319	I23	2181.3	S-I-T-T-S-pT-S-P-E-Q-T-E-R-V-F-Q-I-S
320	I24	2181.2	S-I-T-T-S-T-pS-P-E-Q-T-E-R-V-F-Q-I-S
321	I25	2181.3	S-I-T-T-S-T-S-P-E-Q-pT-E-R-V-F-Q-I-S
322	I26	2181.2	S-I-T-T-S-T-S-P-E-Q-T-E-R-V-F-Q-I-pS
323	I27	2105.3	T-T-S-T-S-P-E-Q-T-E-R-V-F-Q-I-S-F-F
324	I28	2275.4	pT-T-S-T-S-P-E-Q-T-E-R-V-F-Q-I-S-F-F
325	I29	2275.4	T-pT-S-T-S-P-E-Q-T-E-R-V-F-Q-I-S-F-F
326	I30	2275.3	T-T-pS-T-S-P-E-Q-T-E-R-V-F-Q-I-S-F-F
327	I31	2275.4	T-T-S-pT-S-P-E-Q-T-E-R-V-F-Q-I-S-F-F
328	I32	2275.3	T-T-S-T-pS-P-E-Q-T-E-R-V-F-Q-I-S-F-F
329	I33	2275.4	T-T-S-T-S-P-E-Q-pT-E-R-V-F-Q-I-S-F-F
330	I34	2275.3	T-T-S-T-S-P-E-Q-T-E-R-V-F-Q-I-pS-F-F
331	I35	2130.4	S-T-S-P-E-Q-T-E-R-V-F-Q-I-S-F-F-N-L
332	I36	2300.4	pS-T-S-P-E-Q-T-E-R-V-F-Q-I-S-F-F-N-L
333	I37	2300.5	S-pT-S-P-E-Q-T-E-R-V-F-Q-I-S-F-F-N-L
334	J 1	2300.4	S-T-pS-P-E-Q-T-E-R-V-F-Q-I-S-F-F-N-L
335	J 2	2300.5	S-T-S-P-E-Q-pT-E-R-V-F-Q-I-S-F-F-N-L
336	J 3	2300.4	S-T-S-P-E-Q-T-E-R-V-F-Q-I-pS-F-F-N-L
337	J 4	2140.4	S-P-E-Q-T-E-R-V-F-Q-I-S-F-F-N-L-P-T
338	J 5	2310.4	pS-P-E-Q-T-E-R-V-F-Q-I-S-F-F-N-L-P-T
339	J 6	2310.5	S-P-E-Q-pT-E-R-V-F-Q-I-S-F-F-N-L-P-T
340	J 7	2310.4	S-P-E-Q-T-E-R-V-F-Q-I-pS-F-F-N-L-P-T
341	J 8	2310.5	S-P-E-Q-T-E-R-V-F-Q-I-S-F-F-N-L-P-pT
342	J 9	2192.4	E-Q-T-E-R-V-F-Q-I-S-F-F-N-L-P-T-V-H
343	J10	2362.5	E-Q-pT-E-R-V-F-Q-I-S-F-F-N-L-P-T-V-H
344	J11	2362.4	E-Q-T-E-R-V-F-Q-I-pS-F-F-N-L-P-T-V-H
345	J12	2362.5	E-Q-T-E-R-V-F-Q-I-S-F-F-N-L-P-pT-V-H
346	J13	2204.6	T-E-R-V-F-Q-I-S-F-F-N-L-P-T-V-H-R-I

347	J14	2374.7	pT-E-R-V-F-Q-I-S-F-F-N-L-P-T-V-H-R-I
348	J15	2374.6	T-E-R-V-F-Q-I-pS-F-F-N-L-P-T-V-H-R-I
349	J16	2374.7	T-E-R-V-F-Q-I-S-F-F-N-L-P-pT-V-H-R-I
350	J17	2245.7	R-V-F-Q-I-S-F-F-N-L-P-T-V-H-R-I-R-D
351	J18	2415.7	R-V-F-Q-I-pS-F-F-N-L-P-T-V-H-R-I-R-D
352	J19	2415.8	R-V-F-Q-I-S-F-F-N-L-P-pT-V-H-R-I-R-D
353	J20	2259.8	F-Q-I-S-F-F-N-L-P-T-V-H-R-I-R-D-I-R
354	J21	2429.8	F-Q-I-pS-F-F-N-L-P-T-V-H-R-I-R-D-I-R
355	J22	2429.9	F-Q-I-S-F-F-N-L-P-pT-V-H-R-I-R-D-I-R
356	J23	2200.7	I-S-F-F-N-L-P-T-V-H-R-I-R-D-I-R-S-E
357	J24	2370.7	I-pS-F-F-N-L-P-T-V-H-R-I-R-D-I-R-S-E
358	J25	2370.8	I-S-F-F-N-L-P-pT-V-H-R-I-R-D-I-R-S-E
359	J26	2370.7	I-S-F-F-N-L-P-T-V-H-R-I-R-D-I-R-pS-E
360	J27	2241.8	F-F-N-L-P-T-V-H-R-I-R-D-I-R-S-E-K-I
361	J28	2411.9	F-F-N-L-P-pT-V-H-R-I-R-D-I-R-S-E-K-I
362	J29	2411.8	F-F-N-L-P-T-V-H-R-I-R-D-I-R-pS-E-K-I
363	J30	2091.6	N-L-P-T-V-H-R-I-R-D-I-R-S-E-K-I-G-S
364	J31	2261.7	N-L-P-pT-V-H-R-I-R-D-I-R-S-E-K-I-G-S
365	J32	2261.6	N-L-P-T-V-H-R-I-R-D-I-R-pS-E-K-I-G-S
366	J33	2261.6	N-L-P-T-V-H-R-I-R-D-I-R-S-E-K-I-G-pS
367	J34	2090.7	P-T-V-H-R-I-R-D-I-R-S-E-K-I-G-S-L-L
368	J35	2260.8	P-pT-V-H-R-I-R-D-I-R-S-E-K-I-G-S-L-L
369	J36	2260.7	P-T-V-H-R-I-R-D-I-R-pS-E-K-I-G-S-L-L
370	J37	2260.7	P-T-V-H-R-I-R-D-I-R-S-E-K-I-G-pS-L-L
371	K 1	2092.8	V-H-R-I-R-D-I-R-S-E-K-I-G-S-L-L-S-I
372	K 2	2262.8	V-H-R-I-R-D-I-R-pS-E-K-I-G-S-L-L-S-I
373	K 3	2262.8	V-H-R-I-R-D-I-R-S-E-K-I-G-pS-L-L-S-I
374	K 4	2262.8	V-H-R-I-R-D-I-R-S-E-K-I-G-S-L-L-pS-I
375	K 5	2000.8	R-I-R-D-I-R-S-E-K-I-G-S-L-L-S-I-S-G
376	K 6	2170.8	R-I-R-D-I-R-pS-E-K-I-G-S-L-L-S-I-S-G
377	K 7	2170.8	R-I-R-D-I-R-S-E-K-I-G-pS-L-L-S-I-S-G
378	K 8	2170.8	R-I-R-D-I-R-S-E-K-I-G-S-L-L-pS-I-S-G
379	K 9	2170.8	R-I-R-D-I-R-S-E-K-I-G-S-L-L-S-I-pS-G
380	K10	1931.6	R-D-I-R-S-E-K-I-G-S-L-L-S-I-S-G-T-V
381	K11	2101.6	R-D-I-R-pS-E-K-I-G-S-L-L-S-I-S-G-T-V
382	K12	2101.6	R-D-I-R-S-E-K-I-G-pS-L-L-S-I-S-G-T-V
383	K13	2101.6	R-D-I-R-S-E-K-I-G-S-L-L-pS-I-S-G-T-V
384	K14	2101.6	R-D-I-R-S-E-K-I-G-S-L-L-S-I-pS-G-T-V
385	K15	2101.7	R-D-I-R-S-E-K-I-G-S-L-L-S-I-S-G-pT-V
386	K16	1917.6	I-R-S-E-K-I-G-S-L-L-S-I-S-G-T-V-T-R
387	K17	2087.6	I-R-pS-E-K-I-G-S-L-L-S-I-S-G-T-V-T-R
388	K18	2087.6	I-R-S-E-K-I-G-pS-L-L-S-I-S-G-T-V-T-R
389	K19	2087.6	I-R-S-E-K-I-G-S-L-L-pS-I-S-G-T-V-T-R
390	K20	2087.6	I-R-S-E-K-I-G-S-L-L-S-I-pS-G-T-V-T-R
391	K21	2087.7	I-R-S-E-K-I-G-S-L-L-S-I-S-G-pT-V-T-R
392	K22	2087.7	I-R-S-E-K-I-G-S-L-L-S-I-S-G-T-V-pT-R
393	K23	1836.4	S-E-K-I-G-S-L-L-S-I-S-G-T-V-T-R-T-S
394	K24	2006.4	pS-E-K-I-G-S-L-L-S-I-S-G-T-V-T-R-T-S
395	K25	2006.4	S-E-K-I-G-pS-L-L-S-I-S-G-T-V-T-R-T-S
396	K26	2006.4	S-E-K-I-G-S-L-L-pS-I-S-G-T-V-T-R-T-S

397	K27	2006.4	S-E-K-I-G-S-L-L-S-I-pS-G-T-V-T-R-T-S
398	K28	2006.5	S-E-K-I-G-S-L-L-S-I-S-G-pT-V-T-R-T-S
399	K29	2006.5	S-E-K-I-G-S-L-L-S-I-S-G-T-V-pT-R-T-S
400	K30	2006.5	S-E-K-I-G-S-L-L-S-I-S-G-T-V-T-R-pT-S
401	K31	2006.4	S-E-K-I-G-S-L-L-S-I-S-G-T-V-T-R-T-pS
402	K32	1848.4	K-I-G-S-L-L-S-I-S-G-T-V-T-R-T-S-E-V
403	K33	2018.4	K-I-G-pS-L-L-S-I-S-G-T-V-T-R-T-S-E-V
404	K34	2018.4	K-I-G-S-L-L-pS-I-S-G-T-V-T-R-T-S-E-V
405	K35	2018.4	K-I-G-S-L-L-S-I-pS-G-T-V-T-R-T-S-E-V
406	K36	2018.5	K-I-G-S-L-L-S-I-S-G-pT-V-T-R-T-S-E-V
407	K37	2018.5	K-I-G-S-L-L-S-I-S-G-T-V-pT-R-T-S-E-V
408	L 1	2018.5	K-I-G-S-L-L-S-I-S-G-T-V-T-R-pT-S-E-V
409	L 2	2018.4	K-I-G-S-L-L-S-I-S-G-T-V-T-R-T-pS-E-V
410	L 3	1860.3	G-S-L-L-S-I-S-G-T-V-T-R-T-S-E-V-R-P
411	L 4	2030.3	G-pS-L-L-S-I-S-G-T-V-T-R-T-S-E-V-R-P
412	L 5	2030.3	G-S-L-L-pS-I-S-G-T-V-T-R-T-S-E-V-R-P
413	L 6	2030.3	G-S-L-L-S-I-pS-G-T-V-T-R-T-S-E-V-R-P
414	L 7	2030.4	G-S-L-L-S-I-S-G-pT-V-T-R-T-S-E-V-R-P
415	L 8	2030.4	G-S-L-L-S-I-S-G-T-V-pT-R-T-S-E-V-R-P
416	L 9	2030.4	G-S-L-L-S-I-S-G-T-V-T-R-pT-S-E-V-R-P
417	L10	2030.3	G-S-L-L-S-I-S-G-T-V-T-R-T-pS-E-V-R-P
418	L11	1958.4	L-L-S-I-S-G-T-V-T-R-T-S-E-V-R-P-E-L
419	L12	2128.4	L-L-pS-I-S-G-T-V-T-R-T-S-E-V-R-P-E-L
420	L13	2128.4	L-L-S-I-pS-G-T-V-T-R-T-S-E-V-R-P-E-L
421	L14	2128.5	L-L-S-I-S-G-pT-V-T-R-T-S-E-V-R-P-E-L
422	L15	2128.5	L-L-S-I-S-G-T-V-pT-R-T-S-E-V-R-P-E-L
423	L16	2128.5	L-L-S-I-S-G-T-V-T-R-pT-S-E-V-R-P-E-L
424	L17	2128.4	L-L-S-I-S-G-T-V-T-R-T-pS-E-V-R-P-E-L
425	L18	2023.4	S-I-S-G-T-V-T-R-T-S-E-V-R-P-E-L-Y-K
426	L19	2193.4	pS-I-S-G-T-V-T-R-T-S-E-V-R-P-E-L-Y-K
427	L20	2193.4	S-I-pS-G-T-V-T-R-T-S-E-V-R-P-E-L-Y-K
428	L21	2193.5	S-I-S-G-pT-V-T-R-T-S-E-V-R-P-E-L-Y-K
429	L22	2193.5	S-I-S-G-T-V-pT-R-T-S-E-V-R-P-E-L-Y-K
430	L23	2193.5	S-I-S-G-T-V-T-R-pT-S-E-V-R-P-E-L-Y-K
431	L24	2193.4	S-I-S-G-T-V-T-R-T-pS-E-V-R-P-E-L-Y-K
432	L25	1981.3	S-G-T-V-T-R-T-S-E-V-R-P-E-L-Y-K-A-S
433	L26	2151.3	pS-G-T-V-T-R-T-S-E-V-R-P-E-L-Y-K-A-S
434	L27	2151.4	S-G-pT-V-T-R-T-S-E-V-R-P-E-L-Y-K-A-S
435	L28	2151.4	S-G-T-V-pT-R-T-S-E-V-R-P-E-L-Y-K-A-S
436	L29	2151.4	S-G-T-V-T-R-pT-S-E-V-R-P-E-L-Y-K-A-S
437	L30	2151.3	S-G-T-V-T-R-T-pS-E-V-R-P-E-L-Y-K-A-S
438	L31	2151.3	S-G-T-V-T-R-T-S-E-V-R-P-E-L-Y-K-A-pS
439	L32	2085.4	T-V-T-R-T-S-E-V-R-P-E-L-Y-K-A-S-F-T
440	L33	2255.5	pT-V-T-R-T-S-E-V-R-P-E-L-Y-K-A-S-F-T
441	L34	2255.5	T-V-pT-R-T-S-E-V-R-P-E-L-Y-K-A-S-F-T
442	L35	2255.5	T-V-T-R-pT-S-E-V-R-P-E-L-Y-K-A-S-F-T
443	L36	2255.4	T-V-T-R-T-pS-E-V-R-P-E-L-Y-K-A-S-F-T
444	L37	2255.4	T-V-T-R-T-S-E-V-R-P-E-L-Y-K-A-pS-F-T
445	M 1	2255.5	T-V-T-R-T-S-E-V-R-P-E-L-Y-K-A-S-F-pT
446	M 2	2103.4	T-R-T-S-E-V-R-P-E-L-Y-K-A-S-F-T-C-D

447	M 3	2273.5	pT-R-T-S-E-V-R-P-E-L-Y-K-A-S-F-T-C-D
448	M 4	2273.5	T-R-pT-S-E-V-R-P-E-L-Y-K-A-S-F-T-C-D
449	M 5	2273.4	T-R-T-pS-E-V-R-P-E-L-Y-K-A-S-F-T-C-D
450	M 6	2273.4	T-R-T-S-E-V-R-P-E-L-Y-K-A-pS-F-T-C-D
451	M 7	2273.5	T-R-T-S-E-V-R-P-E-L-Y-K-A-S-F-pT-C-D
452	M 8	2080.4	T-S-E-V-R-P-E-L-Y-K-A-S-F-T-C-D-M-C
453	M 9	2250.5	pT-S-E-V-R-P-E-L-Y-K-A-S-F-T-C-D-M-C
454	M10	2250.4	T-pS-E-V-R-P-E-L-Y-K-A-S-F-T-C-D-M-C
455	M11	2250.4	T-S-E-V-R-P-E-L-Y-K-A-pS-F-T-C-D-M-C
456	M12	2250.5	T-S-E-V-R-P-E-L-Y-K-A-S-F-pT-C-D-M-C
457	M13	2119.5	E-V-R-P-E-L-Y-K-A-S-F-T-C-D-M-C-R-A
458	M14	2289.5	E-V-R-P-E-L-Y-K-A-pS-F-T-C-D-M-C-R-A
459	M15	2289.6	E-V-R-P-E-L-Y-K-A-S-F-pT-C-D-M-C-R-A
460	M16	2103.6	R-P-E-L-Y-K-A-S-F-T-C-D-M-C-R-A-I-V
461	M17	2273.6	R-P-E-L-Y-K-A-pS-F-T-C-D-M-C-R-A-I-V
462	M18	2273.7	R-P-E-L-Y-K-A-S-F-pT-C-D-M-C-R-A-I-V
463	M19	2079.5	E-L-Y-K-A-S-F-T-C-D-M-C-R-A-I-V-D-N
464	M20	2249.5	E-L-Y-K-A-pS-F-T-C-D-M-C-R-A-I-V-D-N
465	M21	2249.6	E-L-Y-K-A-S-F-pT-C-D-M-C-R-A-I-V-D-N
466	M22	2065.4	Y-K-A-S-F-T-C-D-M-C-R-A-I-V-D-N-V-E
467	M23	2235.4	Y-K-A-pS-F-T-C-D-M-C-R-A-I-V-D-N-V-E
468	M24	2235.5	Y-K-A-S-F-pT-C-D-M-C-R-A-I-V-D-N-V-E
469	M25	1989.2	A-S-F-T-C-D-M-C-R-A-I-V-D-N-V-E-Q-S
470	M26	2159.2	A-pS-F-T-C-D-M-C-R-A-I-V-D-N-V-E-Q-S
471	M27	2159.3	A-S-F-pT-C-D-M-C-R-A-I-V-D-N-V-E-Q-S
472	M28	2159.2	A-S-F-T-C-D-M-C-R-A-I-V-D-N-V-E-Q-pS
473	M29	2106.4	F-T-C-D-M-C-R-A-I-V-D-N-V-E-Q-S-F-K
474	M30	2276.5	F-pT-C-D-M-C-R-A-I-V-D-N-V-E-Q-S-F-K
475	M31	2276.4	F-T-C-D-M-C-R-A-I-V-D-N-V-E-Q-pS-F-K
476	M32	2122.4	C-D-M-C-R-A-I-V-D-N-V-E-Q-S-F-K-Y-T
477	M33	2292.4	C-D-M-C-R-A-I-V-D-N-V-E-Q-pS-F-K-Y-T
478	M34	2292.5	C-D-M-C-R-A-I-V-D-N-V-E-Q-S-F-K-Y-pT
479	M35	2130.4	M-C-R-A-I-V-D-N-V-E-Q-S-F-K-Y-T-E-P
480	M36	2300.4	M-C-R-A-I-V-D-N-V-E-Q-pS-F-K-Y-T-E-P
481	M37	2300.5	M-C-R-A-I-V-D-N-V-E-Q-S-F-K-Y-pT-E-P
482	N 1	2144.4	R-A-I-V-D-N-V-E-Q-S-F-K-Y-T-E-P-T-F
483	N 2	2314.4	R-A-I-V-D-N-V-E-Q-pS-F-K-Y-T-E-P-T-F
484	N 3	2314.5	R-A-I-V-D-N-V-E-Q-S-F-K-Y-pT-E-P-T-F
485	N 4	2314.5	R-A-I-V-D-N-V-E-Q-S-F-K-Y-T-E-P-pT-F
486	N 5	2117.3	I-V-D-N-V-E-Q-S-F-K-Y-T-E-P-T-F-C-P
487	N 6	2287.3	I-V-D-N-V-E-Q-pS-F-K-Y-T-E-P-T-F-C-P
488	N 7	2287.4	I-V-D-N-V-E-Q-S-F-K-Y-pT-E-P-T-F-C-P
489	N 8	2287.4	I-V-D-N-V-E-Q-S-F-K-Y-T-E-P-pT-F-C-P
490	N 9	2116.2	D-N-V-E-Q-S-F-K-Y-T-E-P-T-F-C-P-N-P
491	N10	2286.2	D-N-V-E-Q-pS-F-K-Y-T-E-P-T-F-C-P-N-P
492	N11	2286.3	D-N-V-E-Q-S-F-K-Y-pT-E-P-T-F-C-P-N-P
493	N12	2286.3	D-N-V-E-Q-S-F-K-Y-T-E-P-pT-F-C-P-N-P
494	N13	2077.2	V-E-Q-S-F-K-Y-T-E-P-T-F-C-P-N-P-S-C
495	N14	2247.2	V-E-Q-pS-F-K-Y-T-E-P-T-F-C-P-N-P-S-C
496	N15	2247.3	V-E-Q-S-F-K-Y-pT-E-P-T-F-C-P-N-P-S-C

497	N16	2247.3	V-E-Q-S-F-K-Y-T-E-P-pT-F-C-P-N-P-S-C
498	N17	2247.2	V-E-Q-S-F-K-Y-T-E-P-T-F-C-P-N-P-pS-C
499	N18	2092.2	Q-S-F-K-Y-T-E-P-T-F-C-P-N-P-S-C-E-N
500	N19	2262.2	Q-pS-F-K-Y-T-E-P-T-F-C-P-N-P-S-C-E-N
501	N20	2262.3	Q-S-F-K-Y-pT-E-P-T-F-C-P-N-P-S-C-E-N
502	N21	2262.3	Q-S-F-K-Y-T-E-P-pT-F-C-P-N-P-S-C-E-N
503	N22	2262.2	Q-S-F-K-Y-T-E-P-T-F-C-P-N-P-pS-C-E-N
504	N23	2104.3	F-K-Y-T-E-P-T-F-C-P-N-P-S-C-E-N-R-A
505	N24	2274.4	F-K-Y-pT-E-P-T-F-C-P-N-P-S-C-E-N-R-A
506	N25	2274.4	F-K-Y-T-E-P-pT-F-C-P-N-P-S-C-E-N-R-A
507	N26	2274.3	F-K-Y-T-E-P-T-F-C-P-N-P-pS-C-E-N-R-A
508	N27	2162.3	Y-T-E-P-T-F-C-P-N-P-S-C-E-N-R-A-F-W
509	N28	2332.4	Y-pT-E-P-T-F-C-P-N-P-S-C-E-N-R-A-F-W
510	N29	2332.4	Y-T-E-P-pT-F-C-P-N-P-S-C-E-N-R-A-F-W
511	N30	2332.3	Y-T-E-P-T-F-C-P-N-P-pS-C-E-N-R-A-F-W
512	N31	2112.3	E-P-T-F-C-P-N-P-S-C-E-N-R-A-F-W-T-L
513	N32	2282.4	E-P-pT-F-C-P-N-P-S-C-E-N-R-A-F-W-T-L
514	N33	2282.3	E-P-T-F-C-P-N-P-pS-C-E-N-R-A-F-W-T-L
515	N34	2282.4	E-P-T-F-C-P-N-P-S-C-E-N-R-A-F-W-pT-L
516	N35	2099.3	T-F-C-P-N-P-S-C-E-N-R-A-F-W-T-L-N-V
517	N36	2269.4	pT-F-C-P-N-P-S-C-E-N-R-A-F-W-T-L-N-V
518	N37	2269.3	T-F-C-P-N-P-pS-C-E-N-R-A-F-W-T-L-N-V
519	O 1	2269.4	T-F-C-P-N-P-S-C-E-N-R-A-F-W-pT-L-N-V
520	O 2	2108.3	C-P-N-P-S-C-E-N-R-A-F-W-T-L-N-V-T-R
521	O 3	2278.3	C-P-N-P-pS-C-E-N-R-A-F-W-T-L-N-V-T-R
522	O 4	2278.4	C-P-N-P-S-C-E-N-R-A-F-W-pT-L-N-V-T-R
523	O 5	2278.4	C-P-N-P-S-C-E-N-R-A-F-W-T-L-N-V-pT-R
524	O 6	2151.4	N-P-S-C-E-N-R-A-F-W-T-L-N-V-T-R-S-R
525	O 7	2321.4	N-P-pS-C-E-N-R-A-F-W-T-L-N-V-T-R-S-R
526	O 8	2321.5	N-P-S-C-E-N-R-A-F-W-pT-L-N-V-T-R-S-R
527	O 9	2321.5	N-P-S-C-E-N-R-A-F-W-T-L-N-V-pT-R-S-R
528	O10	2321.4	N-P-S-C-E-N-R-A-F-W-T-L-N-V-T-R-pS-R
529	O11	2200.6	S-C-E-N-R-A-F-W-T-L-N-V-T-R-S-R-F-L
530	O12	2370.6	pS-C-E-N-R-A-F-W-T-L-N-V-T-R-S-R-F-L
531	O13	2370.7	S-C-E-N-R-A-F-W-pT-L-N-V-T-R-S-R-F-L
532	O14	2370.7	S-C-E-N-R-A-F-W-T-L-N-V-pT-R-S-R-F-L
533	O15	2370.6	S-C-E-N-R-A-F-W-T-L-N-V-T-R-pS-R-F-L
534	O16	2311.7	E-N-R-A-F-W-T-L-N-V-T-R-S-R-F-L-D-W
535	O17	2481.8	E-N-R-A-F-W-pT-L-N-V-T-R-S-R-F-L-D-W
536	O18	2481.8	E-N-R-A-F-W-T-L-N-V-pT-R-S-R-F-L-D-W
537	O19	2481.7	E-N-R-A-F-W-T-L-N-V-T-R-pS-R-F-L-D-W
538	O20	2324.8	R-A-F-W-T-L-N-V-T-R-S-R-F-L-D-W-Q-K
539	O21	2494.9	R-A-F-W-pT-L-N-V-T-R-S-R-F-L-D-W-Q-K
540	O22	2494.9	R-A-F-W-T-L-N-V-pT-R-S-R-F-L-D-W-Q-K
541	O23	2494.8	R-A-F-W-T-L-N-V-T-R-pS-R-F-L-D-W-Q-K
542	O24	2352.8	F-W-T-L-N-V-T-R-S-R-F-L-D-W-Q-K-V-R
543	O25	2522.9	F-W-pT-L-N-V-T-R-S-R-F-L-D-W-Q-K-V-R
544	O26	2522.9	F-W-T-L-N-V-pT-R-S-R-F-L-D-W-Q-K-V-R
545	O27	2522.8	F-W-T-L-N-V-T-R-pS-R-F-L-D-W-Q-K-V-R
546	O28	2260.7	T-L-N-V-T-R-S-R-F-L-D-W-Q-K-V-R-I-Q

547	O29	2430.8	pT-L-N-V-T-R-S-R-F-L-D-W-Q-K-V-R-I-Q
548	O30	2430.8	T-L-N-V-pT-R-S-R-F-L-D-W-Q-K-V-R-I-Q
549	O31	2430.7	T-L-N-V-T-R-pS-R-F-L-D-W-Q-K-V-R-I-Q
550	O32	2289.6	N-V-T-R-S-R-F-L-D-W-Q-K-V-R-I-Q-E-N
551	O33	2459.7	N-V-pT-R-S-R-F-L-D-W-Q-K-V-R-I-Q-E-N
552	O34	2459.6	N-V-T-R-pS-R-F-L-D-W-Q-K-V-R-I-Q-E-N
553	O35	2261.6	T-R-S-R-F-L-D-W-Q-K-V-R-I-Q-E-N-A-N
554	O36	2431.7	pT-R-S-R-F-L-D-W-Q-K-V-R-I-Q-E-N-A-N
555	O37	2431.6	T-R-pS-R-F-L-D-W-Q-K-V-R-I-Q-E-N-A-N
556	P 1	2246.6	S-R-F-L-D-W-Q-K-V-R-I-Q-E-N-A-N-E-I
557	P 2	2416.6	pS-R-F-L-D-W-Q-K-V-R-I-Q-E-N-A-N-E-I
558	P 3	2201.5	F-L-D-W-Q-K-V-R-I-Q-E-N-A-N-E-I-P-T
559	P 4	2371.6	F-L-D-W-Q-K-V-R-I-Q-E-N-A-N-E-I-P-pT
560	P 5	2085.3	D-W-Q-K-V-R-I-Q-E-N-A-N-E-I-P-T-G-S
561	P 6	2255.4	D-W-Q-K-V-R-I-Q-E-N-A-N-E-I-P-pT-G-S
562	P 7	2255.3	D-W-Q-K-V-R-I-Q-E-N-A-N-E-I-P-T-G-pS
563	P 8	2012.3	Q-K-V-R-I-Q-E-N-A-N-E-I-P-T-G-S-M-P
564	P 9	2182.4	Q-K-V-R-I-Q-E-N-A-N-E-I-P-pT-G-S-M-P
565	P10	2182.3	Q-K-V-R-I-Q-E-N-A-N-E-I-P-T-G-pS-M-P
566	P11	2013.3	V-R-I-Q-E-N-A-N-E-I-P-T-G-S-M-P-R-T
567	P12	2183.4	V-R-I-Q-E-N-A-N-E-I-P-pT-G-S-M-P-R-T
568	P13	2183.3	V-R-I-Q-E-N-A-N-E-I-P-T-G-pS-M-P-R-T
569	P14	2183.4	V-R-I-Q-E-N-A-N-E-I-P-T-G-S-M-P-R-pT
570	P15	1986.3	I-Q-E-N-A-N-E-I-P-T-G-S-M-P-R-T-L-D
571	P16	2156.4	I-Q-E-N-A-N-E-I-P-pT-G-S-M-P-R-T-L-D
572	P17	2156.3	I-Q-E-N-A-N-E-I-P-T-G-pS-M-P-R-T-L-D
573	P18	2156.4	I-Q-E-N-A-N-E-I-P-T-G-S-M-P-R-pT-L-D
574	P19	1957.3	E-N-A-N-E-I-P-T-G-S-M-P-R-T-L-D-V-I
575	P20	2127.4	E-N-A-N-E-I-P-pT-G-S-M-P-R-T-L-D-V-I
576	P21	2127.3	E-N-A-N-E-I-P-T-G-pS-M-P-R-T-L-D-V-I
577	P22	2127.4	E-N-A-N-E-I-P-T-G-S-M-P-R-pT-L-D-V-I
578	P23	1983.5	A-N-E-I-P-T-G-S-M-P-R-T-L-D-V-I-L-R
579	P24	2153.6	A-N-E-I-P-pT-G-S-M-P-R-T-L-D-V-I-L-R
580	P25	2153.5	A-N-E-I-P-T-G-pS-M-P-R-T-L-D-V-I-L-R
581	P26	2153.6	A-N-E-I-P-T-G-S-M-P-R-pT-L-D-V-I-L-R
582	P27	1970.5	E-I-P-T-G-S-M-P-R-T-L-D-V-I-L-R-G-D
583	P28	2140.6	E-I-P-pT-G-S-M-P-R-T-L-D-V-I-L-R-G-D
584	P29	2140.5	E-I-P-T-G-pS-M-P-R-T-L-D-V-I-L-R-G-D
585	P30	2140.6	E-I-P-T-G-S-M-P-R-pT-L-D-V-I-L-R-G-D
586	P31	1914.4	P-T-G-S-M-P-R-T-L-D-V-I-L-R-G-D-S-V
587	P32	2084.5	P-pT-G-S-M-P-R-T-L-D-V-I-L-R-G-D-S-V
588	P33	2084.4	P-T-G-pS-M-P-R-T-L-D-V-I-L-R-G-D-S-V
589	P34	2084.5	P-T-G-S-M-P-R-pT-L-D-V-I-L-R-G-D-S-V
590	P35	2084.4	P-T-G-S-M-P-R-T-L-D-V-I-L-R-G-D-pS-V
591	P36	2001.5	G-S-M-P-R-T-L-D-V-I-L-R-G-D-S-V-E-R
592	P37	2171.5	G-pS-M-P-R-T-L-D-V-I-L-R-G-D-S-V-E-R
593	Q 1	2171.6	G-S-M-P-R-pT-L-D-V-I-L-R-G-D-S-V-E-R
594	Q 2	2171.5	G-S-M-P-R-T-L-D-V-I-L-R-G-D-pS-V-E-R
595	Q 3	2056.6	M-P-R-T-L-D-V-I-L-R-G-D-S-V-E-R-A-K
596	Q 4	2226.7	M-P-R-pT-L-D-V-I-L-R-G-D-S-V-E-R-A-K

597	Q 5	2226.6	M-P-R-T-L-D-V-I-L-R-G-D-pS-V-E-R-A-K
598	Q 6	1982.5	R-T-L-D-V-I-L-R-G-D-S-V-E-R-A-K-P-G
599	Q 7	2152.6	R-pT-L-D-V-I-L-R-G-D-S-V-E-R-A-K-P-G
600	Q 8	2152.5	R-T-L-D-V-I-L-R-G-D-pS-V-E-R-A-K-P-G
601	Q 9	1996.5	L-D-V-I-L-R-G-D-S-V-E-R-A-K-P-G-D-R
602	Q10	2166.5	L-D-V-I-L-R-G-D-pS-V-E-R-A-K-P-G-D-R
603	Q11	1999.5	V-I-L-R-G-D-S-V-E-R-A-K-P-G-D-R-C-K
604	Q12	2169.5	V-I-L-R-G-D-pS-V-E-R-A-K-P-G-D-R-C-K
605	Q13	2035.5	L-R-G-D-S-V-E-R-A-K-P-G-D-R-C-K-F-T
606	Q14	2205.5	L-R-G-D-pS-V-E-R-A-K-P-G-D-R-C-K-F-T
607	Q15	2205.6	L-R-G-D-S-V-E-R-A-K-P-G-D-R-C-K-F-pT
608	Q16	1922.3	G-D-S-V-E-R-A-K-P-G-D-R-C-K-F-T-G-V
609	Q17	2179.4G	-D-S-pS-V-E-R-A-K-P-G-D-R-C-K-F-T-G-V
610	Q18	2092.4	G-D-S-V-E-R-A-K-P-G-D-R-C-K-F-pT-G-V
611	Q19	1992.4	S-V-E-R-A-K-P-G-D-R-C-K-F-T-G-V-E-I
612	Q20	2162.4	pS-V-E-R-A-K-P-G-D-R-C-K-F-T-G-V-E-I
613	Q21	2162.5	S-V-E-R-A-K-P-G-D-R-C-K-F-pT-G-V-E-I
614	Q22	2004.4	E-R-A-K-P-G-D-R-C-K-F-T-G-V-E-I-V-V
615	Q23	2174.5	E-R-A-K-P-G-D-R-C-K-F-pT-G-V-E-I-V-V
616	Q24	1931.3	A-K-P-G-D-R-C-K-F-T-G-V-E-I-V-V-P-D
617	Q25	2101.4	A-K-P-G-D-R-C-K-F-pT-G-V-E-I-V-V-P-D
618	Q26	1932.2	P-G-D-R-C-K-F-T-G-V-E-I-V-V-P-D-V-T
619	Q27	2102.3	P-G-D-R-C-K-F-pT-G-V-E-I-V-V-P-D-V-T
620	Q28	2102.3	P-G-D-R-C-K-F-T-G-V-E-I-V-V-P-D-V-pT
621	Q29	2019.3	D-R-C-K-F-T-G-V-E-I-V-V-P-D-V-T-Q-L
622	Q30	2189.4	D-R-C-K-F-pT-G-V-E-I-V-V-P-D-V-T-Q-L
623	Q31	2189.4	D-R-C-K-F-T-G-V-E-I-V-V-P-D-V-pT-Q-L
624	Q32	1918.3	C-K-F-T-G-V-E-I-V-V-P-D-V-T-Q-L-G-L
625	Q33	2088.4	C-K-F-pT-G-V-E-I-V-V-P-D-V-T-Q-L-G-L
626	Q34	2088.4	C-K-F-T-G-V-E-I-V-V-P-D-V-pT-Q-L-G-L
627	Q35	1841.2	F-T-G-V-E-I-V-V-P-D-V-T-Q-L-G-L-P-G
628	Q36	2011.3	F-pT-G-V-E-I-V-V-P-D-V-T-Q-L-G-L-P-G
629	Q37	2011.3	F-T-G-V-E-I-V-V-P-D-V-pT-Q-L-G-L-P-G
630	R 1	1820.2	G-V-E-I-V-V-P-D-V-T-Q-L-G-L-P-G-V-K
631	R 2	1990.3	G-V-E-I-V-V-P-D-V-pT-Q-L-G-L-P-G-V-K
632	R 3	1848.2	E-I-V-V-P-D-V-T-Q-L-G-L-P-G-V-K-P-S
633	R 4	2018.3	E-I-V-V-P-D-V-pT-Q-L-G-L-P-G-V-K-P-S
634	R 5	2018.2	E-I-V-V-P-D-V-T-Q-L-G-L-P-G-V-K-P-pS
635	R 6	1794.1	V-V-P-D-V-T-Q-L-G-L-P-G-V-K-P-S-S-T
636	R 7	1964.2	V-V-P-D-V-pT-Q-L-G-L-P-G-V-K-P-S-S-T
637	R 8	1964.1	V-V-P-D-V-T-Q-L-G-L-P-G-V-K-P-pS-S-T
638	R 9	1964.1	V-V-P-D-V-T-Q-L-G-L-P-G-V-K-P-S-pS-T
639	R10	1964.2	V-V-P-D-V-T-Q-L-G-L-P-G-V-K-P-S-S-pT
640	R11	1824.2	P-D-V-T-Q-L-G-L-P-G-V-K-P-S-S-T-L-D
641	R12	1994.3	P-D-V-pT-Q-L-G-L-P-G-V-K-P-S-S-T-L-D
642	R13	1994.2	P-D-V-T-Q-L-G-L-P-G-V-K-P-pS-S-T-L-D
643	R14	1994.2	P-D-V-T-Q-L-G-L-P-G-V-K-P-S-pS-T-L-D
644	R15	1994.3	P-D-V-T-Q-L-G-L-P-G-V-K-P-S-S-pT-L-D
645	R16	1869.3	V-T-Q-L-G-L-P-G-V-K-P-S-S-T-L-D-T-R
646	R17	2039.4	V-pT-Q-L-G-L-P-G-V-K-P-S-S-T-L-D-T-R

647	R18	2039.3	V-T-Q-L-G-L-P-G-V-K-P-pS-S-T-L-D-T-R
648	R19	2039.3	V-T-Q-L-G-L-P-G-V-K-P-S-pS-T-L-D-T-R
649	R20	2039.4	V-T-Q-L-G-L-P-G-V-K-P-S-S-pT-L-D-T-R
650	R21	2039.4	V-T-Q-L-G-L-P-G-V-K-P-S-S-T-L-D-pT-R
651	R22	1839.4	Q-L-G-L-P-G-V-K-P-S-S-T-L-D-T-R-G-I
652	R23	2009.4	Q-L-G-L-P-G-V-K-P-pS-S-T-L-D-T-R-G-I
653	R24	2009.4	Q-L-G-L-P-G-V-K-P-S-pS-T-L-D-T-R-G-I
654	R25	2009.5	Q-L-G-L-P-G-V-K-P-S-S-pT-L-D-T-R-G-I
655	R26	2009.5	Q-L-G-L-P-G-V-K-P-S-S-T-L-D-pT-R-G-I
656	R27	1813.4	G-L-P-G-V-K-P-S-S-T-L-D-T-R-G-I-S-K
657	R28	1983.4	G-L-P-G-V-K-P-pS-S-T-L-D-T-R-G-I-S-K
658	R29	1983.4	G-L-P-G-V-K-P-S-pS-T-L-D-T-R-G-I-S-K
659	R30	1983.5	G-L-P-G-V-K-P-S-S-pT-L-D-T-R-G-I-S-K
660	R31	1983.5	G-L-P-G-V-K-P-S-S-T-L-D-pT-R-G-I-S-K
661	R32	1983.4	G-L-P-G-V-K-P-S-S-T-L-D-T-R-G-I-pS-K
662	R33	1845.3	P-G-V-K-P-S-S-T-L-D-T-R-G-I-S-K-T-T
663	R34	2015.3	P-G-V-K-P-pS-S-T-L-D-T-R-G-I-S-K-T-T
664	R35	2015.3	P-G-V-K-P-S-pS-T-L-D-T-R-G-I-S-K-T-T
665	R36	2015.4	P-G-V-K-P-S-S-pT-L-D-T-R-G-I-S-K-T-T
666	R37	2015.4	P-G-V-K-P-S-S-T-L-D-pT-R-G-I-S-K-T-T
667	S 1	2015.3	P-G-V-K-P-S-S-T-L-D-T-R-G-I-pS-K-T-T
668	S 2	2015.4	P-G-V-K-P-S-S-T-L-D-T-R-G-I-S-K-pT-T
669	S 3	2015.4	P-G-V-K-P-S-S-T-L-D-T-R-G-I-S-K-T-pT
670	S 4	1877.3	V-K-P-S-S-T-L-D-T-R-G-I-S-K-T-T-E-G
671	S 5	2047.3	V-K-P-pS-S-T-L-D-T-R-G-I-S-K-T-T-E-G
672	S 6	2047.3	V-K-P-S-pS-T-L-D-T-R-G-I-S-K-T-T-E-G
673	S 7	2047.4	V-K-P-S-S-pT-L-D-T-R-G-I-S-K-T-T-E-G
674	S 8	2047.4	V-K-P-S-S-T-L-D-pT-R-G-I-S-K-T-T-E-G
675	S 9	2047.3	V-K-P-S-S-T-L-D-T-R-G-I-pS-K-T-T-E-G
676	S10	2047.4	V-K-P-S-S-T-L-D-T-R-G-I-S-K-pT-T-E-G
677	S11	2047.4	V-K-P-S-S-T-L-D-T-R-G-I-S-K-T-pT-E-G
678	S12	1877.3	P-S-S-T-L-D-T-R-G-I-S-K-T-T-E-G-L-N
679	S13	2047.3	P-pS-S-T-L-D-T-R-G-I-S-K-T-T-E-G-L-N
680	S14	2047.3	P-S-pS-T-L-D-T-R-G-I-S-K-T-T-E-G-L-N
681	S15	2047.4	P-S-S-pT-L-D-T-R-G-I-S-K-T-T-E-G-L-N
682	S16	2047.4	P-S-S-T-L-D-pT-R-G-I-S-K-T-T-E-G-L-N
683	S17	2047.3	P-S-S-T-L-D-T-R-G-I-pS-K-T-T-E-G-L-N
684	S18	2047.4	P-S-S-T-L-D-T-R-G-I-S-K-pT-T-E-G-L-N
685	S19	2047.4	P-S-S-T-L-D-T-R-G-I-S-K-T-pT-E-G-L-N
686	S20	1837.3	S-T-L-D-T-R-G-I-S-K-T-T-E-G-L-N-S-G
687	S21	2007.3	pS-T-L-D-T-R-G-I-S-K-T-T-E-G-L-N-S-G
688	S22	2007.4	S-pT-L-D-T-R-G-I-S-K-T-T-E-G-L-N-S-G
689	S23	2007.4	S-T-L-D-pT-R-G-I-S-K-T-T-E-G-L-N-S-G
690	S24	2007.3	S-T-L-D-T-R-G-I-pS-K-T-T-E-G-L-N-S-G
691	S25	2007.4	S-T-L-D-T-R-G-I-S-K-pT-T-E-G-L-N-S-G
692	S26	2007.4	S-T-L-D-T-R-G-I-S-K-T-pT-E-G-L-N-S-G
693	S27	2007.3	S-T-L-D-T-R-G-I-S-K-T-T-E-G-L-N-pS-G
694	S28	1849.3	L-D-T-R-G-I-S-K-T-T-E-G-L-N-S-G-V-T
695	S29	2019.4	L-D-pT-R-G-I-S-K-T-T-E-G-L-N-S-G-V-T
696	S30	2019.3	L-D-T-R-G-I-pS-K-T-T-E-G-L-N-S-G-V-T

697	S31	2019.4	L-D-T-R-G-I-S-K-pT-T-E-G-L-N-S-G-V-T
698	S32	2019.4	L-D-T-R-G-I-S-K-T-pT-E-G-L-N-S-G-V-T
699	S33	2019.3	L-D-T-R-G-I-S-K-T-T-E-G-L-N-pS-G-V-T
700	S34	2019.4	L-D-T-R-G-I-S-K-T-T-E-G-L-N-S-G-V-pT
701	S35	1791.3	T-R-G-I-S-K-T-T-E-G-L-N-S-G-V-T-G-L
702	S36	1961.4	pT-R-G-I-S-K-T-T-E-G-L-N-S-G-V-T-G-L
703	S37	1961.3	T-R-G-I-pS-K-T-T-E-G-L-N-S-G-V-T-G-L
704	T 1	1961.4	T-R-G-I-S-K-pT-T-E-G-L-N-S-G-V-T-G-L
705	T 2	1961.4	T-R-G-I-S-K-T-pT-E-G-L-N-S-G-V-T-G-L
706	T 3	1961.3	T-R-G-I-S-K-T-T-E-G-L-N-pS-G-V-T-G-L
707	T 4	1961.4	T-R-G-I-S-K-T-T-E-G-L-N-S-G-V-pT-G-L
708	T 5	1777.3	G-I-S-K-T-T-E-G-L-N-S-G-V-T-G-L-R-S
709	T 6	1947.3	G-I-pS-K-T-T-E-G-L-N-S-G-V-T-G-L-R-S
710	T 7	1947.4	G-I-S-K-pT-T-E-G-L-N-S-G-V-T-G-L-R-S
711	T 8	1947.4	G-I-S-K-T-pT-E-G-L-N-S-G-V-T-G-L-R-S
712	T 9	1947.3	G-I-S-K-T-T-E-G-L-N-pS-G-V-T-G-L-R-S
713	T10	1947.4	G-I-S-K-T-T-E-G-L-N-S-G-V-pT-G-L-R-S
714	T11	1947.3	G-I-S-K-T-T-E-G-L-N-S-G-V-T-G-L-R-pS
715	T12	1777.3	S-K-T-T-E-G-L-N-S-G-V-T-G-L-R-S-L-G
716	T13	1947.3	pS-K-T-T-E-G-L-N-S-G-V-T-G-L-R-S-L-G
717	T14	1947.4	S-K-pT-T-E-G-L-N-S-G-V-T-G-L-R-S-L-G
718	T15	1947.4	S-K-T-pT-E-G-L-N-S-G-V-T-G-L-R-S-L-G
719	T16	1947.3	S-K-T-T-E-G-L-N-pS-G-V-T-G-L-R-S-L-G
720	T17	1947.4	S-K-T-T-E-G-L-N-S-G-V-pT-G-L-R-S-L-G
721	T18	1947.3	S-K-T-T-E-G-L-N-S-G-V-T-G-L-R-pS-L-G
722	T19	1817.3	T-T-E-G-L-N-S-G-V-T-G-L-R-S-L-G-V-R
723	T20	1987.4	pT-T-E-G-L-N-S-G-V-T-G-L-R-S-L-G-V-R
724	T21	1987.4	T-pT-E-G-L-N-S-G-V-T-G-L-R-S-L-G-V-R
725	T22	1987.3	T-T-E-G-L-N-pS-G-V-T-G-L-R-S-L-G-V-R
726	T23	1987.4	T-T-E-G-L-N-S-G-V-pT-G-L-R-S-L-G-V-R
727	T24	1987.3	T-T-E-G-L-N-S-G-V-T-G-L-R-pS-L-G-V-R
728	T25	1843.4	E-G-L-N-S-G-V-T-G-L-R-S-L-G-V-R-D-L
729	T26	2013.4	E-G-L-N-pS-G-V-T-G-L-R-S-L-G-V-R-D-L
730	T27	2013.5	E-G-L-N-S-G-V-pT-G-L-R-S-L-G-V-R-D-L
731	T28	2013.4	E-G-L-N-S-G-V-T-G-L-R-pS-L-G-V-R-D-L
732	T29	1921.5	L-N-S-G-V-T-G-L-R-S-L-G-V-R-D-L-T-Y
733	T30	2091.5	L-N-pS-G-V-T-G-L-R-S-L-G-V-R-D-L-T-Y
734	T31	2091.6	L-N-S-G-V-pT-G-L-R-S-L-G-V-R-D-L-T-Y
735	T32	2091.5	L-N-S-G-V-T-G-L-R-pS-L-G-V-R-D-L-T-Y
736	T33	2091.6	L-N-S-G-V-T-G-L-R-S-L-G-V-R-D-L-pT-Y
737	T34	1935.6	S-G-V-T-G-L-R-S-L-G-V-R-D-L-T-Y-K-I
738	T35	2105.6	pS-G-V-T-G-L-R-S-L-G-V-R-D-L-T-Y-K-I
739	T36	2105.7	S-G-V-pT-G-L-R-S-L-G-V-R-D-L-T-Y-K-I
740	T37	2105.6	S-G-V-T-G-L-R-pS-L-G-V-R-D-L-T-Y-K-I
741	U 1	2105.7	S-G-V-T-G-L-R-S-L-G-V-R-D-L-pT-Y-K-I
742	U 2	2025.7	V-T-G-L-R-S-L-G-V-R-D-L-T-Y-K-I-S-F
743	U 3	2195.8	V-pT-G-L-R-S-L-G-V-R-D-L-T-Y-K-I-S-F
744	U 4	2195.7	V-T-G-L-R-pS-L-G-V-R-D-L-T-Y-K-I-S-F
745	U 5	2195.8	V-T-G-L-R-S-L-G-V-R-D-L-pT-Y-K-I-S-F
746	U 6	2195.7	V-T-G-L-R-S-L-G-V-R-D-L-T-Y-K-I-pS-F

747	U 7	2009.8	G-L-R-S-L-G-V-R-D-L-T-Y-K-I-S-F-L-A
748	U 8	2179.8	G-L-R-pS-L-G-V-R-D-L-T-Y-K-I-S-F-L-A
749	U 9	2179.9	G-L-R-S-L-G-V-R-D-L-pT-Y-K-I-S-F-L-A
750	U10	2179.8	G-L-R-S-L-G-V-R-D-L-T-Y-K-I-pS-F-L-A
751	U11	2079.7	R-S-L-G-V-R-D-L-T-Y-K-I-S-F-L-A-C-H
752	U12	2249.7	R-pS-L-G-V-R-D-L-T-Y-K-I-S-F-L-A-C-H
753	U13	2249.8	R-S-L-G-V-R-D-L-pT-Y-K-I-S-F-L-A-C-H
754	U14	2249.7	R-S-L-G-V-R-D-L-T-Y-K-I-pS-F-L-A-C-H
755	U15	2048.7	L-G-V-R-D-L-T-Y-K-I-S-F-L-A-C-H-V-I
756	U16	2218.8	L-G-V-R-D-L-pT-Y-K-I-S-F-L-A-C-H-V-I
757	U17	2218.7	L-G-V-R-D-L-T-Y-K-I-pS-F-L-A-C-H-V-I
758	U18	2078.7	V-R-D-L-T-Y-K-I-S-F-L-A-C-H-V-I-S-I
759	U19	2248.8	V-R-D-L-pT-Y-K-I-S-F-L-A-C-H-V-I-S-I
760	U20	2248.7	V-R-D-L-T-Y-K-I-pS-F-L-A-C-H-V-I-S-I
761	U21	2248.7	V-R-D-L-T-Y-K-I-S-F-L-A-C-H-V-I-pS-I
762	U22	1967.6	D-L-T-Y-K-I-S-F-L-A-C-H-V-I-S-I-G-S
763	U23	2137.7	D-L-pT-Y-K-I-S-F-L-A-C-H-V-I-S-I-G-S
764	U24	2137.6	D-L-T-Y-K-I-pS-F-L-A-C-H-V-I-S-I-G-S
765	U25	2137.6	D-L-T-Y-K-I-S-F-L-A-C-H-V-I-pS-I-G-S
766	U26	2137.6	D-L-T-Y-K-I-S-F-L-A-C-H-V-I-S-I-G-pS
767	U27	1966.6	T-Y-K-I-S-F-L-A-C-H-V-I-S-I-G-S-N-I
768	U28	2136.7	pT-Y-K-I-S-F-L-A-C-H-V-I-S-I-G-S-N-I
769	U29	2136.6	T-Y-K-I-pS-F-L-A-C-H-V-I-S-I-G-S-N-I
770	U30	2136.6	T-Y-K-I-S-F-L-A-C-H-V-I-pS-I-G-S-N-I
771	U31	2136.6	T-Y-K-I-S-F-L-A-C-H-V-I-S-I-G-pS-N-I
772	U32	1830.5	K-I-S-F-L-A-C-H-V-I-S-I-G-S-N-I-G-A
773	U33	2000.5	K-I-pS-F-L-A-C-H-V-I-S-I-G-S-N-I-G-A
774	U34	2000.5	K-I-S-F-L-A-C-H-V-I-pS-I-G-S-N-I-G-A
775	U35	2000.5	K-I-S-F-L-A-C-H-V-I-S-I-G-pS-N-I-G-A
776	U36	1763.3	S-F-L-A-C-H-V-I-S-I-G-S-N-I-G-A-S-S
777	U37	1933.3	pS-F-L-A-C-H-V-I-S-I-G-S-N-I-G-A-S-S

7.1.3 Part 3

Nr.	Pos.	Mol. Weight	Sequence
778	A 1	1933.3	S-F-L-A-C-H-V-I-pS-I-G-S-N-I-G-A-S-S
779	A 2	1933.3	S-F-L-A-C-H-V-I-S-I-G-pS-N-I-G-A-S-S
780	A 3	1933.3	S-F-L-A-C-H-V-I-S-I-G-S-N-I-G-A-pS-S
781	A 4	1933.3	S-F-L-A-C-H-V-I-S-I-G-S-N-I-G-A-S-pS
782	A 5	1741.2	L-A-C-H-V-I-S-I-G-S-N-I-G-A-S-S-P-D
783	A 6	1911.2	L-A-C-H-V-I-pS-I-G-S-N-I-G-A-S-S-P-D
784	A 7	1911.2	L-A-C-H-V-I-S-I-G-pS-N-I-G-A-S-S-P-D
785	A 8	1911.2	L-A-C-H-V-I-S-I-G-S-N-I-G-A-pS-S-P-D
786	A 9	1911.2	L-A-C-H-V-I-S-I-G-S-N-I-G-A-S-pS-P-D
787	A10	1742.1	C-H-V-I-S-I-G-S-N-I-G-A-S-S-P-D-A-N
788	A11	1912.1	C-H-V-I-pS-I-G-S-N-I-G-A-S-S-P-D-A-N
789	A12	1912.1	C-H-V-I-S-I-G-pS-N-I-G-A-S-S-P-D-A-N
790	A13	1912.1	C-H-V-I-S-I-G-S-N-I-G-A-pS-S-P-D-A-N

791	A14	1912.1	C-H-V-I-S-I-G-S-N-I-G-A-S-pS-P-D-A-N
792	A15	1703.1	V-I-S-I-G-S-N-I-G-A-S-S-P-D-A-N-S-N
793	A16	1873.1	V-I-pS-I-G-S-N-I-G-A-S-S-P-D-A-N-S-N
794	A17	1873.1	V-I-S-I-G-pS-N-I-G-A-S-S-P-D-A-N-S-N
795	A18	1873.1	V-I-S-I-G-S-N-I-G-A-pS-S-P-D-A-N-S-N
796	A19	1873.1	V-I-S-I-G-S-N-I-G-A-S-pS-P-D-A-N-S-N
797	A20	1873.1	V-I-S-I-G-S-N-I-G-A-S-S-P-D-A-N-pS-N
798	A21	1761.1	S-I-G-S-N-I-G-A-S-S-P-D-A-N-S-N-N-R
799	A22	1931.1	pS-I-G-S-N-I-G-A-S-S-P-D-A-N-S-N-N-R
800	A23	1931.1	S-I-G-pS-N-I-G-A-S-S-P-D-A-N-S-N-N-R
801	A24	1931.1	S-I-G-S-N-I-G-A-pS-S-P-D-A-N-S-N-N-R
802	A25	1931.1	S-I-G-S-N-I-G-A-S-pS-P-D-A-N-S-N-N-R
803	A26	1931.1	S-I-G-S-N-I-G-A-S-S-P-D-A-N-pS-N-N-R
804	A27	1791	G-S-N-I-G-A-S-S-P-D-A-N-S-N-N-R-E-T
805	A28	1961	G-pS-N-I-G-A-S-S-P-D-A-N-S-N-N-R-E-T
806	A29	1961	G-S-N-I-G-A-pS-S-P-D-A-N-S-N-N-R-E-T
807	A30	1961	G-S-N-I-G-A-S-pS-P-D-A-N-S-N-N-R-E-T
808	A31	1961	G-S-N-I-G-A-S-S-P-D-A-N-pS-N-N-R-E-T
809	A32	1961.1	G-S-N-I-G-A-S-S-P-D-A-N-S-N-N-R-E-pT
810	A33	1889.1	N-I-G-A-S-S-P-D-A-N-S-N-N-R-E-T-E-L
811	A34	2059.1	N-I-G-A-pS-S-P-D-A-N-S-N-N-R-E-T-E-L
812	A35	2059.1	N-I-G-A-S-pS-P-D-A-N-S-N-N-R-E-T-E-L
813	A36	2059.1	N-I-G-A-S-S-P-D-A-N-pS-N-N-R-E-T-E-L
814	A37	2059.2	N-I-G-A-S-S-P-D-A-N-S-N-N-R-E-pT-E-L
815	B 1	1921.1	G-A-S-S-P-D-A-N-S-N-N-R-E-T-E-L-Q-M
816	B 2	2091.1	G-A-pS-S-P-D-A-N-S-N-N-R-E-T-E-L-Q-M
817	B 3	2091.1	G-A-S-pS-P-D-A-N-S-N-N-R-E-T-E-L-Q-M
818	B 4	2091.1	G-A-S-S-P-D-A-N-pS-N-N-R-E-T-E-L-Q-M
819	B 5	2091.2	G-A-S-S-P-D-A-N-S-N-N-R-E-pT-E-L-Q-M
820	B 6	1935.1	S-S-P-D-A-N-S-N-N-R-E-T-E-L-Q-M-A-A
821	B 7	2105.1	pS-S-P-D-A-N-S-N-N-R-E-T-E-L-Q-M-A-A
822	B 8	2105.1	S-pS-P-D-A-N-S-N-N-R-E-T-E-L-Q-M-A-A
823	B 9	2105.1	S-S-P-D-A-N-pS-N-N-R-E-T-E-L-Q-M-A-A
824	B10	2105.2	S-S-P-D-A-N-S-N-N-R-E-pT-E-L-Q-M-A-A
825	B11	1988.2	P-D-A-N-S-N-N-R-E-T-E-L-Q-M-A-A-N-L
826	B12	2158.2	P-D-A-N-pS-N-N-R-E-T-E-L-Q-M-A-A-N-L
827	B13	2158.3	P-D-A-N-S-N-N-R-E-pT-E-L-Q-M-A-A-N-L
828	B14	1975.2	A-N-S-N-N-R-E-T-E-L-Q-M-A-A-N-L-Q-A
829	B15	2145.2	A-N-pS-N-N-R-E-T-E-L-Q-M-A-A-N-L-Q-A
830	B16	2145.3	A-N-S-N-N-R-E-pT-E-L-Q-M-A-A-N-L-Q-A
831	B17	2018.2	S-N-N-R-E-T-E-L-Q-M-A-A-N-L-Q-A-N-N
832	B18	2188.2	pS-N-N-R-E-T-E-L-Q-M-A-A-N-L-Q-A-N-N
833	B19	2188.3	S-N-N-R-E-pT-E-L-Q-M-A-A-N-L-Q-A-N-N
834	B20	2079.3	N-R-E-T-E-L-Q-M-A-A-N-L-Q-A-N-N-V-Y
835	B21	2249.4	N-R-E-pT-E-L-Q-M-A-A-N-L-Q-A-N-N-V-Y
836	B22	2052.2	E-T-E-L-Q-M-A-A-N-L-Q-A-N-N-V-Y-Q-D
837	B23	2222.3	E-pT-E-L-Q-M-A-A-N-L-Q-A-N-N-V-Y-Q-D
838	B24	2065.2	E-L-Q-M-A-A-N-L-Q-A-N-N-V-Y-Q-D-N-E
839	B25	2094.2	Q-M-A-A-N-L-Q-A-N-N-V-Y-Q-D-N-E-R-D
840	B26	2092.1	A-A-N-L-Q-A-N-N-V-Y-Q-D-N-E-R-D-Q-E

841	B27	2196.2	N-L-Q-A-N-N-V-Y-Q-D-N-E-R-D-Q-E-V-F
842	B28	2196.2	Q-A-N-N-V-Y-Q-D-N-E-R-D-Q-E-V-F-L-N
843	B29	2197.3	N-N-V-Y-Q-D-N-E-R-D-Q-E-V-F-L-N-S-L
844	B30	2367.3	N-N-V-Y-Q-D-N-E-R-D-Q-E-V-F-L-N-pS-L
845	B31	2143.3	V-Y-Q-D-N-E-R-D-Q-E-V-F-L-N-S-L-S-S
846	B32	2313.3	V-Y-Q-D-N-E-R-D-Q-E-V-F-L-N-pS-L-S-S
847	B33	2313.3	V-Y-Q-D-N-E-R-D-Q-E-V-F-L-N-S-L-pS-S
848	B34	2313.3	V-Y-Q-D-N-E-R-D-Q-E-V-F-L-N-S-L-S-pS
849	B35	2125.2	Q-D-N-E-R-D-Q-E-V-F-L-N-S-L-S-S-D-E
850	B36	2295.2	Q-D-N-E-R-D-Q-E-V-F-L-N-pS-L-S-S-D-E
851	B37	2295.2	Q-D-N-E-R-D-Q-E-V-F-L-N-S-L-pS-S-D-E

7.2 Phosphopeptides used in peptide pulldowns

Name	**Phospho-residue**	**Sequence**
TD1	T150	YRAQIEFMKIYDLNTIYIDYQHLSMRENGA
TD2	T247	GSSAAPGNGTSAMATRSITTSTSPEQTERV
TD3	S324	TCDMCRAIVDNVEQSFKYTEPTFCPNPSCE
TD4	S351	SCENRAFWTLNVTRSRFLDWQKVRIQENAN
TD5	T398	DSVERAKPGDRCKFTGVEIVVPDVTQLGLP
TD6a	S475	SNIGASSPDANSNNRETELQMAANLQANNV
TD6b	T480	SNIGASSPDANSNNRETELQMAANLQANNV
TD6c	S475/T480	SNIGASSPDANSNNRETELQMAANLQANNV
TD7a	S255	AMATRSITTSTSPEQTERVFQISFFNLPTV
TD7b	T259	AMATRSITTSTSPEQTERVFQISFFNLPTV
TD7c	S255/T259	AMATRSITTSTSPEQTERVFQISFFNLPTV
TD8	T376	QENANEIPTGSMPRTLDVILRGDSVERAKP
TD9	T423	QLGLPGVKPSSTLDTRGISKTTEGLNSGVT
TD10	S226/S227	ADEDEQQDDDMNGSSLPRDSGSSAAPGNGT
TD11	T259/T272	EQTERVFQISFFNLPTVHRIRDIRSEKIGS

Reference List

AGGARWAL, B. D. & CALVI, B. R. 2004. Chromatin regulates origin activity in Drosophila follicle cells. *Nature,* 430, 372-376.
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**,** 59-69.
ARAKI, H., LEEM, S. H., PHONGDARA, A. & SUGINO, A. 1995. Dpb11, which interacts with DNApolymerase II(epsilon) in *Saccharomyces cerevisiae*, has a dual role in S-phase progression and at a cell-cycle checkpoint. *Proc. Natl Acad. Sci. U S A,* 92**,** 11791-11795.
ARIAS, E. E. & WALTER, J. C. 2005. Replication-dependent destruction of Cdt1 limits DNA replication to a single round per cell cycle in Xenopus egg extracts. *Genes Dev,* 19**,** 114-26.
ARIAS, E. E. & WALTER, J. C. 2006. PCNA functions as a molecular platform to trigger Cdt1 destruction and prevent re-replication. *Nat Cell Biol,* 8**,** 84-90.
ARIAS, E. E. & WALTER, J. C. 2007. Strength in numbers: preventing rereplication via multiple mechanisms in eukaryotic cells. *Genes Dev,* 21**,** 497-518.
BELL, S. P., KOBAYASHI, R. & STILLMAN, B. 1993. Yeast origin recognition complex functions in transcription silencing and DNA replication. *Science,* 262**,** 1844-1849.
BELL, S. P. & STILLMAN, B. 1992. ATP-dependent recognition of eukaryotic origins of DNA replication by a multiprotein complex. *Nature,* 357**,** 128-134.
BLOW, J. J. 1993. Preventing re-replication of DNA in a single cell cycle - evidence for a replication licensing factor. *J. Cell Biol.,* 122**,** 993-1002.
BLOW, J. J. & DUTTA, A. 2005. Preventing re-replication of chromosomal DNA. *Nat Rev Mol Cell Biol,* 6**,** 476-86.
BLOW, J. J. & GILLESPIE, P. J. 2008. Replication licensing and cancer--a fatal entanglement? *Nat Rev Cancer,* 8**,** 799-806.
BLOW, J. J., GILLESPIE, P. J., FRANCIS, D. & JACKSON, D. A. 2001. Replication origins in Xenopus egg extract Are 5-15 kilobases apart and are activated in clusters that fire at different times. *J Cell Biol,* 152**,** 15-25.
BLOW, J. J. & LASKEY, R. A. 1988. A role for the nuclear envelope in controlling DNA replication within the cell cycle. *Nature,* 332**,** 546-548.
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. *Mol Cell Biol,* 28**,** 5865-73.
BOOS, D., FRIGOLA, J. & DIFFLEY, J. F. X. 2012. Activation of the replicative DNA helicase: breaking up is hard to do. *Curr Opin Cell Biol,* 24**,** 423-30.
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. *Curr Biol,* 21**,** 1152-7.
BOOS, D., YEKEZARE, M. & DIFFLEY, J. F. X. 2013. Identification of a heteromeric complex that promotes DNA replication origin firing in human cells. *Science,* 340**,** 981-4.
BOUSSET, K. & DIFFLEY, J. F. X. 1998. The Cdc7 protein kinase is required for origin firing during S phase. *Genes Dev,* 12**,** 480-90.
BOWERS, J. L., RANDELL, J. C., CHEN, S. & BELL, S. P. 2004. ATP hydrolysis by ORC catalyzes reiterative Mcm2-7 assembly at a defined origin of replication. *Mol Cell,* 16**,** 967-78.

BREWER, B. J. & FANGMAN, W. L. 1987. The localization of replication origins on ARS plasmids in S. cerevisiae. *Cell,* 51**,** 463-71.

BRUCK, I., KANTER, D. M. & KAPLAN, D. L. 2011. Enabling Association of the GINS Protein Tetramer with the Mini Chromosome Maintenance (Mcm)2-7 Protein Complex by Phosphorylated Sld2 Protein and Single-stranded Origin DNA. *J Biol Chem,* 286**,** 36414-26.

BRUCK, I. & KAPLAN, D. 2009. Dbf4-Cdc7 phosphorylation of Mcm2 is required for cell growth. *J Biol Chem,* 284**,** 28823-31.

BRUCK, I. & KAPLAN, D. L. 2011. Origin single-stranded DNA releases Sld3 protein from the Mcm2-7 complex, allowing the GINS tetramer to bind the Mcm2-7 complex. *J Biol Chem,* 286**,** 18602-13.

BRUCK, I. & KAPLAN, D. L. 2013. Cdc45 protein-single-stranded DNA interaction is important for stalling the helicase during replication stress. *J Biol Chem,* 288**,** 7550-63.

BUCK, V., WHITE, A. & ROSAMOND, J. 1991. CDC7 protein kinase activity is required for mitosis and meiosis in Saccharomyces cerevisiae. *Mol Gen Genet,* 227**,** 452-7.

CARPENTER, P. B., MUELLER, P. R. & DUNPHY, W. G. 1996. Role for a Xenopus Orc2-related protein in controlling DNA replication. *Nature,* 379**,** 357-360.

CELNIKER, S. E., SWEDER, K., SRIENC, F., BAILEY, J. E. & CAMPBELL, J. L. 1984. Deletion mutations affecting autonomously replicating sequence ARS1 of Saccharomyces cerevisiae. *Mol Cell Biol,* 4**,** 2455-66.

CHAPMAN, J. W. & JOHNSTON, L. H. 1989. The yeast gene, DBF4, essential for entry into S phase is cell cycle regulated. *Exp. Cell. Res.,* 180**,** 419-428.

CHARYCH, D. H., COYNE, M., YABANNAVAR, A., NARBERES, J., CHOW, S., WALLROTH, M., SHAFER, C. & WALTER, A. O. 2008. Inhibition of Cdc7/Dbf4 kinase activity affects specific phosphorylation sites on MCM2 in cancer cells. *J Cell Biochem,* 104**,** 1075-86.

CHEN, S. & BELL, S. P. 2011. CDK prevents Mcm2-7 helicase loading by inhibiting Cdt1 interaction with Orc6. *Genes Dev,* 25**,** 363-72.

CHEN, S., DE VRIES, M. A. & BELL, S. P. 2007. Orc6 is required for dynamic recruitment of Cdt1 during repeated Mcm2-7 loading. *Genes Dev,* 21**,** 2897-907.

CHO, W. H., LEE, Y. J., KONG, S. I., HURWITZ, J. & LEE, J. K. 2006. CDC7 kinase phosphorylates serine residues adjacent to acidic amino acids in the minichromosome maintenance 2 protein. *Proc Natl Acad Sci U S A,* 103**,** 11521-6.

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. *Proc Natl Acad Sci U S A,* 97**,** 1530-1535.

CHONG, J. P. J., 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**,** 418-421.

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**,** 180-2.

COLEMAN, T. R., CARPENTER, P. B. & DUNPHY, W. G. 1996. The Xenopus Cdc6 protein is essential for the initiation of a single round of DNA replication in cell-free extracts. *Cell,* 87**,** 53-63.

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. *Nat Struct Mol Biol,* 18**,** 471-7.

COSTA, A., RENAULT, L., SWUEC, P., PETOJEVIC, T., PESAVENTO, 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*, e03273.

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, 899-913.

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. *Mol. Cell,* in press.

COXON, A., MAUNDRELL, K. & KEARSEY, S. E. 1992. Fission yeast cdc21+ belongs to a family of proteins involved in an early step of chromosome-replication. *Nucleic Acids Res,* 20, 5571-5577.

DALTON, S. & HOPWOOD, B. 1997. Characterization of Cdc47p-minichromosome maintenance complexes in *Saccharomyces cerevisiae*: identification of Cdc45p as a subunit. *Mol. Cell. Biol.,* 17, 5867-5875.

DALTON, S. & WHITBREAD, L. 1995. Cell-cycle-regulated nuclear import and export of Cdc47, a protein essential for initiation of DNA-replication in budding yeast. *Proc. Natl. Acad. Sci. U S A,* 92, 2514-2518.

DAVEY, M. J., INDIANI, C. & O'DONNELL, M. 2003. Reconstitution of the Mcm2-7p heterohexamer, subunit arrangement, and ATP site architecture. *J Biol Chem,* 278, 4491-9.

DAVIDSON, I. F., LI, A. & BLOW, J. J. 2006. Deregulated replication licensing causes DNA fragmentation consistent with head-to-tail fork collision. *Mol Cell,* 24, 433-43.

DE BRUIN, R. A., MCDONALD, W. H., KALASHNIKOVA, T. I., YATES, J., 3RD & WITTENBERG, C. 2004. Cln3 activates G1-specific transcription via phosphorylation of the SBF bound repressor Whi5. *Cell,* 117, 887-98.

DIFFLEY, J. F. X. 2004. Regulation of early events in chromosome replication. *Curr Biol,* 14, R778-86.

DIFFLEY, J. F. X. 2011. Quality control in the initiation of eukaryotic DNA replication. *Phil. Trans. R. Soc. B,* 366, 3545-3553.

DIFFLEY, J. F. X. & COCKER, J. H. 1992. Protein-DNA interactions at a yeast replication origin. *Nature,* 357, 169-172.

DIFFLEY, J. F. X., COCKER, J. H., DOWELL, S. J., HARWOOD, J. & ROWLEY, A. 1995. Stepwise assembly of initiation complexes at budding yeast replication origins during the cell cycle. *J Cell Sci Suppl,* 19, 67-72.

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, 303-16.

DIFFLEY, J. F. X. & STILLMAN, B. 1988. Purification of a yeast protein that binds to origins of DNA replication and a transcriptional silencer. *Proc Natl Acad Sci U S A,* 85, 2120-4.

DONOVAN, S., HARWOOD, J., DRURY, L. S. & DIFFLEY, J. F. X. 1997. Cdc6p-dependent loading of Mcm proteins onto pre-replicative chromatin in budding yeast. *Proc Natl Acad Sci U S A,* 94, 5611-6.

DOWELL, S. J., ROMANOWSKI, P. & DIFFLEY, J. F. X. 1994. Interaction of Dbf4, the Cdc7 protein kinase regulatory subunit, with yeast replication origins in vivo. *Science,* 265, 1243-6.

DRURY, L. S., PERKINS, G. & DIFFLEY, J. F. X. 2000. The cyclin-dependent kinase Cdc28p regulates distinct modes of Cdc6p proteolysis during the budding yeast cell cycle. *Curr Biol,* 10, 231-40.

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 epsilon and its unexpected ability to support growth in the absence of the DNA polymerase domain. *J Biol Chem,* 274, 22283-8.
EATON, M. L., GALANI, K., KANG, S., BELL, S. P. & MACALPINE, D. M. 2010. Conserved nucleosome positioning defines replication origins. *Genes Dev,* 24, 748-53.
EDWARDS, M. C., TUTTER, A. V., CVETIC, C., GILBERT, C. H., PROKHOROVA, T. A. & WALTER, J. C. 2002. MCM2-7 complexes bind chromatin in a distributed pattern surrounding the origin recognition complex in Xenopus egg extracts. *J Biol Chem,* 277, 33049-57.
EISENBERG, S., KORZA, G., CARSON, J., LIACHKO, I. & TYE, B. K. 2009. Novel DNA binding properties of the Mcm10 protein from Saccharomyces cerevisiae. *J Biol Chem,* 284, 25412-20.
ESER, U., FALLEUR-FETTIG, M., JOHNSON, A. & SKOTHEIM, J. M. 2011. Commitment to a cellular transition precedes genome-wide transcriptional change. *Mol Cell,* 43, 515-27.
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. *Proc Natl Acad Sci U S A,* 106, 20240-5.
FALBO, K. B. & SHEN, X. 2012. Function of the INO80 chromatin remodeling complex in DNA replication. *Front Biosci (Landmark Ed),* 17, 970-5.
FELDMAN, R. M., CORRELL, C. C., KAPLAN, K. B. & DESHAIES, R. J. 1997. A complex of Cdc4p, Skp1p, and Cdc53p/cullin catalyzes ubiquitination of the phosphorylated CDK inhibitor Sic1p. *Cell,* 91, 221-30.
FENG, W., COLLINGWOOD, D., BOECK, M. E., FOX, L. A., ALVINO, G. M., FANGMAN, W. L., RAGHURAMAN, M. K. & BREWER, B. J. 2006. Genomic mapping of single-stranded DNA in hydroxyurea-challenged yeasts identifies origins of replication. *Nat Cell Biol,* 8, 148-55.
FENG, W. & D'URSO, G. 2001. Schizosaccharomyces pombe cells lacking the amino-terminal catalytic domains of DNA polymerase epsilon are viable but require the DNA damage checkpoint control. *Mol Cell Biol,* 21, 4495-504.
FERENBACH, A., LI, A., BRITO-MARTINS, M. & BLOW, J. J. 2005. Functional domains of the Xenopus replication licensing factor Cdt1. *Nucleic Acids Res,* 33, 316-24.
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. *Mol Cell,* 50, 577-88.
FERREIRA, M. F., SANTOCANALE, C., DRURY, L. S. & DIFFLEY, J. F. X. 2000. Dbf4p, an essential S phase-promoting factor, is targeted for degradation by the anaphase-promoting complex. *Mol Cell Biol,* 20, 242-8.
FOSS, M., MCNALLY, F. J., LAURENSON, P. & RINE, J. 1993. Origin recognition complex (ORC) in transcriptional silencing and DNA replication in *S. cerevisiae*. *Science,* 262, 1838-1844.
FRANCIS, L. I., RANDELL, J. C., TAKARA, T. J., UCHIMA, L. & BELL, S. P. 2009. Incorporation into the prereplicative complex activates the Mcm2-7 helicase for Cdc7-Dbf4 phosphorylation. *Genes Dev,* 23, 643-54.
FRIGOLA, J., REMUS, D., MEHANNA, A. & DIFFLEY, J. F. X. 2013. ATPase-dependent quality control of DNA replication origin licensing. *Nature,* 495, 339-43.
FROELICH, C. A., KANG, S., EPLING, L. B., BELL, S. P. & ENEMARK, E. J. 2014. A conserved MCM single-stranded DNA binding element is essential for replication initiation. *Elife,* 3, e01993.

FU, Y. V., YARDIMCI, H., LONG, D. T., GUAINAZZI, A., BERMUDEZ, V. P., HURWITZ, J., VAN OIJEN, A., SCHARER, O. D. & WALTER, J. C. 2011. Selective Bypass of a Lagging Strand Roadblock by the Eukaryotic Replicative DNA Helicase. *Cell,* 146**,** 931-941.

FUJITA, M. 2006. Cdt1 revisited: complex and tight regulation during the cell cycle and consequences of deregulation in mammalian cells. *Cell Div,* 1**,** 22.

FULLER, R. S. & KORNBERG, A. 1983. Purified dnaA protein in initiation of replication at the Escherichia coli chromosomal origin of replication. *Proc Natl Acad Sci U S A,* 80**,** 5817-21.

GAGGIOLI, V., ZEISER, E., RIVERS, D., BRADSHAW, C. R., AHRINGER, J. & ZEGERMAN, P. 2014. CDK phosphorylation of SLD-2 is required for replication initiation and germline development in C. elegans. *J Cell Biol,* 204**,** 507-22.

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. *Nat Cell Biol,* 8**,** 358-66.

GAMBUS, A., KHOUDOLI, G. A., JONES, R. C. & BLOW, J. J. 2011. MCM2-7 form double hexamers at licensed origins in Xenopus egg extract. *J Biol Chem,* 286**,** 11855-64.

GAVIN, K. A., HIDAKA, M. & STILLMAN, B. 1995. Conserved initiator proteins in eukaryotes. *Science,* 270**,** 1667-71.

GILLESPIE, P. J., LI, A. & BLOW, J. J. 2001. Reconstitution of licensed replication origins on *Xenopus* sperm nuclei using purified proteins. *BMC Biochem,* 2**,** 15.

GOSSEN, M., PAK, D. T., HANSEN, S. K., ACHARYA, J. K. & BOTCHAN, M. R. 1995. A Drosophila homolog of the yeast origin recognition complex [see comments]. *Science,* 270**,** 1674-7.

GROS, J., DEVBHANDARI, S. & REMUS, D. 2014. Origin plasticity during budding yeast DNA replication in vitro. *EMBO J,* 33**,** 621-36.

GUARINO, E., SHEPHERD, M. E., SALGUERO, I., HUA, H., DEEGAN, R. S. & KEARSEY, S. E. 2011. Cdt1 proteolysis is promoted by dual PIP degrons and is modulated by PCNA ubiquitylation. *Nucleic Acids Res*.

HARDY, C. F. 1997. Identification of Cdc45p, an essential factor required for DNA replication. *Gene,* 187**,** 239-46.

HARDY, C. F., DRYGA, O., SEEMATTER, S., PAHL, P. M. & SCLAFANI, R. A. 1997. Mcm5/Cdc46-bob1 bypasses the requirement for the S phase activator Cdc7p. *Proc. Natl. Acad. Sci. USA,* 94**,** 3151-5.

HARLAND, R. M. & LASKEY, R. A. 1980. Regulated replication of DNA microinjected into eggs of Xenopus laevis. *Cell,* 21**,** 761-71.

HARTWELL, L. H. 1973. Three additional genes required for deoxyribonucleic acid synthesis in Saccharomyces cerevisiae. *J. Bacteriol.,* 115**,** 966-74.

HARTWELL, L. H., CULOTTI, J., PRINGLE, J. R. & REID, B. J. 1974. Genetic control of the cell division cycle in yeast. *Science,* 183**,** 46-51.

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**,** 267-286.

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 Dev,* 26**,** 137-50.

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**,** 80-91.

HENNESSY, K. M., LEE, A., CHEN, E. & BOTSTEIN, D. 1991. A group of interacting yeast DNA replication genes. *Genes Dev.,* 5**,** 958-969.

HEREFORD, L. M. & HARTWELL, L. H. 1973. Role of protein synthesis in the replication of yeast DNA. *Nat New Biol,* 244**,** 129-31.

HEREFORD, L. M. & HARTWELL, L. H. 1974. Sequential gene function in the initiation of *Saccharomyces cerevisiae* DNA synthesis. *J. Mol. Biol.,* 84**,** 445-461.

HOFMANN, J. F. & BEACH, D. 1994. cdt1 is an essential target of the Cdc10/Sct1 transcription factor: requirement for DNA replication and inhibition of mitosis. *EMBO J,* 13**,** 425-34.

HOGAN, E. & KOSHLAND, D. 1992. Addition of extra origins of replication to a minichromosome suppresses its mitotic loss in *cdc6* and *cdc14* mutants of *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci. USA,* 89**,** 3098-3102.

HOMESLEY, L., LEI, M., KAWASAKI, Y., SAWYER, S., CHRISTENSEN, T. & TYE, B. K. 2000. Mcm10 and the MCM2-7 complex interact to initiate DNA synthesis and to release replication factors from origins. *Genes Dev,* 14**,** 913-26.

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. *Proc. Natl. Acad. Sci. USA,* 93**,** 12309-12314.

HSIAO, C. L. & CARBON, J. 1979. High-frequency transformation of yeast by plasmids containing the cloned yeast ARG4 gene. *Proc Natl Acad Sci U S A,* 76**,** 3829-33.

HU, Y. F., HAO, Z. L. & LI, R. 1999. Chromatin remodeling and activation of chromosomal DNA replication by an acidic transcriptional activation domain from BRCA1. *Genes Dev,* 13**,** 637-42.

HUANG, R. Y. & KOWALSKI, D. 1996. Multiple DNA elements in ARS305 determine replication origin activity in a yeast chromosome. *Nucleic Acids Res,* 24**,** 816-23.

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 2 micron plasmid. *Cell,* 51**,** 473-481.

HUO, Y. G., BAI, L., XU, M. & JIANG, T. 2010. Crystal structure of the N-terminal region of human Topoisomerase IIbeta binding protein 1. *Biochem Biophys Res Commun,* 401**,** 401-5.

ILVES, I., PETOJEVIC, T., PESAVENTO, J. J. & BOTCHAN, M. R. 2010. Activation of the MCM2-7 helicase by association with Cdc45 and GINS proteins. *Mol Cell,* 37**,** 247-58.

ISHIMI, Y. 1997. A DNA helicase activity is associated with an MCM4, -6, and -7 protein complex. *J Biol Chem,* 272**,** 24508-13.

JACKSON, A. L., PAHL, P. M., HARRISON, K., ROSAMOND, J. & SCLAFANI, R. A. 1993. Cell cycle regulation of the yeast *CDC7* protein kinase by association with the *DBF4* protein. *Mol. Cell. Biol.,* 13**,** 2899-2908.

JACOB, F., BRENNER, S. & CUZIN, F. 1963. On the regulation of DNA replication in bacteria. *Cold Spring Harbor Symp. Quant. Biol.,* 28**,** 329-348.

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 Dev,* 14**,** 1528-40.

JASPERSEN, S. L., CHARLES, J. F. & MORGAN, D. O. 1999. Inhibitory phosphorylation of the APC regulator Hct1 is controlled by the kinase Cdc28 and the phosphatase Cdc14. *Curr Biol,* 9**,** 227-36.

JOHNSTON, L. H. & THOMAS, A. P. 1982. A further two mutants defective in initiation of the S phase in the yeast Saccharomyces cerevisiae. *Mol Gen Genet,* 186**,** 445-8.

KABSCH, W. & SANDER, C. 1983. Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers,* 22**,** 2577-637.

KAMIMURA, Y., MASUMOTO, H., SUGINO, A. & ARAKI, H. 1998. Sld2, which interacts with Dpb11 in *Saccharomyces cerevisiae*, is required for chromosomal DNA replication. *Mol. Cell. Biol.,* 18**,** 6102-6109.

KAMIMURA, Y., TAK, Y. S., SUGINO, A. & ARAKI, H. 2001. Sld3, which interacts with Cdc45 (Sld4), functions for chromosomal DNA replication in Saccharomyces cerevisiae. *Embo J,* 20**,** 2097-107.

KANEMAKI, M. & LABIB, K. 2006. Distinct roles for Sld3 and GINS during establishment and progression of eukaryotic DNA replication forks. *EMBO J,* 25**,** 1753-63.

KANEMAKI, M., SANCHEZ-DIAZ, A., GAMBUS, A. & LABIB, K. 2003. Functional proteomic identification of DNA replication proteins by induced proteolysis in vivo. *Nature,* 423**,** 720-4.

KANG, S., WARNER, M. D. & BELL, S. P. 2014. Multiple Functions for Mcm2-7 ATPase Motifs during Replication Initiation. *Mol Cell*.

KANKE, M., KODAMA, Y., TAKAHASHI, T. S., NAKAGAWA, T. & MASUKATA, H. 2012. Mcm10 plays an essential role in origin DNA unwinding after loading of the CMG components. *EMBO J,* 31**,** 2182-94.

KARNANI, N. & DUTTA, A. 2011. The effect of the intra-S-phase checkpoint on origins of replication in human cells. *Genes Dev,* 25**,** 621-33.

KELLEY, L. A. & STERNBERG, M. J. 2009. Protein structure prediction on the Web: a case study using the Phyre server. *Nat Protoc,* 4**,** 363-71.

KELMAN, Z., LEE, J. K. & HURWITZ, J. 1999. The single minichromosome maintenance protein of methanobacterium thermoautotrophicum DeltaH contains DNA helicase activity. *Proc Natl Acad Sci U S A,* 96**,** 14783-8.

KESTI, T., FLICK, K., KERANEN, S., SYVAOJA, J. E. & WITTENBERG, C. 1999. DNA polymerase epsilon catalytic domains are dispensable for DNA replication, DNA repair, and cell viability. *Mol Cell,* 3**,** 679-85.

KITADA, K., JOHNSTON, L. H., SUGINO, T. & SUGINO, A. 1992. Temperature-sensitive cdc7 mutations of Saccharomyces cerevisiae are suppressed by the DBF4 gene, which is required for the G1/S cell cycle transition. *Genetics,* 131**,** 21-9.

KLEMM, R. D., AUSTIN, R. J. & BELL, S. P. 1997. Coordinate Binding of ATP and Origin DNA Regulates the ATPase Activity of the Origin Recognition Complex. *Cell,* 88**,** 493-502.

KOIVOMAGI, M., VALK, E., VENTA, R., IOFIK, A., LEPIKU, M., BALOG, E. R., RUBIN, S. M., MORGAN, D. O. & LOOG, M. 2011. Cascades of multisite phosphorylation control Sic1 destruction at the onset of S phase. *Nature,* 480**,** 128-31.

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. *J Biol Chem,* 287**,** 4121-8.

KUBOTA, Y., MIMURA, S., NISHIMOTO, S.-I., TAKISAWA, H. & NOJIMA, H. 1995. Identification of the yeast MCM3-related protein as a component of *Xenopus* DNA replication licensing factor. *Cell,* 81**,** 601-609.

KUBOTA, Y., TAKASE, Y., KOMORI, Y., HASHIMOTO, Y., ARATA, T., KAMIMURA, Y., ARAKI, H. & TAKISAWA, H. 2003. A novel ring-like complex of Xenopus proteins essential for the initiation of DNA replication. *Genes Dev,* 17**,** 1141-52.

KUMAGAI, A., SHEVCHENKO, A. & DUNPHY, W. G. 2010. Treslin collaborates with TopBP1 in triggering the initiation of DNA replication. *Cell,* 140**,** 349-59.

KUMAGAI, A., SHEVCHENKO, A. & DUNPHY, W. G. 2011. Direct regulation of Treslin by cyclin-dependent kinase is essential for the onset of DNA replication. *J Cell Biol,* 193**,** 995-1007.

KUNKEL, T. A. & BURGERS, P. M. 2008. Dividing the workload at a eukaryotic replication fork. *Trends Cell Biol,* 18**,** 521-7.
LABIB, K., DIFFLEY, J. F. X. & KEARSEY, S. E. 1999. G1-phase and B-type cyclins exclude the DNA-replication factor Mcm4 from the nucleus. *Nat Cell Biol,* 1**,** 415-22.
LABIB, K., KEARSEY, S. E. & DIFFLEY, J. F. X. 2001. MCM2-7 proteins are essential components of prereplicative complexes that accumulate cooperatively in the nucleus during G1-phase and are required to establish, but not maintain, the S-phase checkpoint. *Mol Biol Cell,* 12**,** 3658-67.
LABIB, K., TERCERO, J. A. & DIFFLEY, J. F. X. 2000. Uninterrupted MCM2-7 function required for DNA replication fork progression. *Science,* 288**,** 1643-7.
LEONARD, A. C. & MECHALI, M. 2013. DNA replication origins. *Cold Spring Harb Perspect Biol,* 5**,** a010116.
LI, X., ZHAO, Q., LIAO, R., SUN, P. & WU, X. 2003. The SCF(Skp2) ubiquitin ligase complex interacts with the human replication licensing factor Cdt1 and regulates Cdt1 degradation. *J Biol Chem,* 278**,** 30854-8.
LIACHKO, I. & TYE, B. K. 2005. Mcm10 is required for the maintenance of transcriptional silencing in Saccharomyces cerevisiae. *Genetics,* 171**,** 503-15.
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**,** 667-676.
LIKU, M. E., NGUYEN, V. Q., ROSALES, A. W., IRIE, K. & LI, J. J. 2005. CDK phosphorylation of a novel NLS-NES module distributed between two subunits of the Mcm2-7 complex prevents chromosomal rereplication. *Mol Biol Cell,* 16**,** 5026-39.
LIM, H. J., JEON, Y., JEON, C. H., KIM, J. H. & LEE, H. 2011. Targeted disruption of Mcm10 causes defective embryonic cell proliferation and early embryo lethality. *Biochim Biophys Acta,* 1813**,** 1777-83.
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**,** 479-83.
MADINE, M. A., KHOO, C. Y., MILLS, A. D. & LASKEY, R. A. 1995. MCM3 complex required for cell cycle regulation of DNA replication in vertebrate cells. *Nature,* 375**,** 421-4.
MAINE, G. T., SINHA, P. & TYE, B.-K. 1984. Mutants of *S. cerevisiae* defective in the maintenance of minichromosomes. *Genetics,* 106**,** 365-385.
MAIORANO, D., MOREAU, J. & MECHALI, M. 2000. XCDT1 is required for the assembly of pre-replicative complexes in *Xenopus laevis*. *Nature,* 404**,** 622-5.
MAKINIEMI, M., HILLUKKALA, T., TUUSA, J., REINI, K., VAARA, M., HUANG, D., POSPIECH, H., MAJURI, I., WESTERLING, T., MAKELA, T. P. & SYVAOJA, J. E. 2001. BRCT domain-containing protein TopBP1 functions in DNA replication and damage response. *J Biol Chem,* 276**,** 30399-406.
MARAHRENS, Y. & STILLMAN, B. 1992. A yeast chromosomal origin of DNA replication defined by multiple functional elements. *Science,* 255**,** 817-823.
MASAI, H., TANIYAMA, C., OGINO, K., MATSUI, E., KAKUSHO, N., MATSUMOTO, S., KIM, J. M., ISHII, A., TANAKA, T., KOBAYASHI, T., TAMAI, K., OHTANI, K. & ARAI, K. 2006. Phosphorylation of MCM4 by Cdc7 kinase facilitates its interaction with Cdc45 on the chromatin. *J Biol Chem,* 281**,** 39249-61.
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**,** 651-5.

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. *Mol Cell Biol,* 20**,** 2809-17.

MATSUMOTO, S., HAYANO, M., KANOH, Y. & MASAI, H. 2011. Multiple pathways can bypass the essential role of fission yeast Hsk1 kinase in DNA replication initiation. *J Cell Biol,* 195**,** 387-401.

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 alpha in the initiation of DNA replication. *Mol Cell Biol,* 26**,** 4843-52.

MCGARRY, T. J. & KIRSCHNER, M. W. 1998. Geminin, an inhibitor of DNA replication, is degraded during mitosis. *Cell,* 93**,** 1043-53.

MCGUFFIN, L. J., BRYSON, K. & JONES, D. T. 2000. The PSIPRED protein structure prediction server. *Bioinformatics,* 16**,** 404-5.

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. *Mol Cell Biol,* 17**,** 3261-71.

MICKLEM, G., ROWLEY, A., HARWOOD, J., NASMYTH, K. & DIFFLEY, J. F. X. 1993. Yeast origin recognition complex is involved in DNA replication and transcriptional silencing. *Nature,* 366**,** 87-89.

MIMURA, S., SEKI, T., TANAKA, S. & DIFFLEY, J. F. X. 2004. Phosphorylation-dependent binding of mitotic cyclins to Cdc6 contributes to DNA replication control. *Nature,* 431**,** 1118-23.

MIZUSHIMA, T., TAKAHASHI, N. & STILLMAN, B. 2000. Cdc6p modulates the structure and DNA binding activity of the origin recognition complex in vitro. *Genes Dev,* 14**,** 1631-41.

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**,** 547-63.

MONTAGNOLI, A., VALSASINA, B., BROTHERTON, D., TROIANI, S., RAINOLDI, S., TENCA, P., MOLINARI, A. & SANTOCANALE, C. 2006. Identification of Mcm2 phosphorylation sites by S-phase-regulating kinases. *J Biol Chem,* 281**,** 10281-90.

MORDES, D. A., NAM, E. A. & CORTEZ, D. 2008. Dpb11 activates the Mec1-Ddc2 complex. *Proc Natl Acad Sci U S A,* 105**,** 18730-4.

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. *Proc Natl Acad Sci U S A,* 103**,** 10236-41.

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 Dev,* 24**,** 602-12.

MUZI-FALCONI, M. & KELLY, T. J. 1995. Orp1, a member of the Cdc18/Cdc6 family of S-phase regulators, is homologous to a component of the origin recognition complex. *Proc Natl Acad Sci U S A,* 92**,** 12475-9.

NAKAJIMA, R. & MASUKATA, H. 2002. SpSld3 is required for loading and maintenance of SpCdc45 on chromatin in DNA replication in fission yeast. *Mol Biol Cell,* 13**,** 1462-72.

NASMYTH, K. & DIRICK, L. 1991. The role of *SWI4* and *SWI6* in the activity of G1 cyclins in yeast. *Cell,* 66**,** 995-1013.

NASMYTH, K. & NURSE, P. 1981. Cell division cycle mutants altered in DNA replication and mitosis in the fission yeast Schizosaccharomyces pombe. *Mol Gen Genet,* 182**,** 119-24.

NAVADGI-PATIL, V. M. & BURGERS, P. M. 2008. Yeast DNA replication protein Dpb11 activates the Mec1/ATR checkpoint kinase. *J Biol Chem,* 283**,** 35853-9.

NGUYEN, V. Q., CO, C., IRIE, K. & LI, J. J. 2000. Clb/Cdc28 kinases promote nuclear export of the replication initiator proteins Mcm2-7. *Curr. Biol.,* 10**,** 195-205.

NGUYEN, V. Q., CO, C. & LI, J. J. 2001. Cyclin-dependent kinases prevent DNA re-replication through multiple mechanisms. *Nature,* 411**,** 1068-73.

NISHITANI, H., LYGEROU, Z., NISHIMOTO, T. & NURSE, P. 2000. The Cdt1 protein is required to license DNA for replication in fission yeast. *Nature,* 404**,** 625-8.

NISHITANI, H., SUGIMOTO, N., ROUKOS, V., NAKANISHI, Y., SAIJO, M., OBUSE, C., TSURIMOTO, T., NAKAYAMA, K. I., NAKAYAMA, K., FUJITA, M., LYGEROU, Z. & NISHIMOTO, T. 2006. Two E3 ubiquitin ligases, SCF-Skp2 and DDB1-Cul4, target human Cdt1 for proteolysis. *EMBO J,* 25**,** 1126-36.

NISHITANI, H., TARAVIRAS, S., LYGEROU, Z. & NISHIMOTO, T. 2001. The human licensing factor for DNA replication Cdt1 accumulates in G1 and is destabilized after initiation of S-phase. *J Biol Chem,* 276**,** 44905-11.

NOGUCHI, E., SHANAHAN, P., NOGUCHI, C. & RUSSELL, P. 2002. CDK Phosphorylation of Drc1 Regulates DNA Replication in Fission Yeast. *Curr Biol,* 12**,** 599-605.

ON, K. F., BEURON, F., FRITH, D., SNIJDERS, A. P., MORRIS, E. P. & DIFFLEY, J. F. X. 2014. Prereplicative complexes assembled in vitro support origin-dependent and independent DNA replication. *EMBO J,* 33**,** 605-20.

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. *Mol Cell Biol,* 19**,** 4888-96.

PACEK, M., TUTTER, A. V., KUBOTA, Y., TAKISAWA, H. & WALTER, J. C. 2006. Localization of MCM2-7, Cdc45, and GINS to the site of DNA unwinding during eukaryotic DNA replication. *Mol Cell,* 21**,** 581-7.

PATTERSON, M., SCLAFANI, R. A., FANGMAN, W. L. & ROSAMOND, J. 1986. Molecular characterization of cell cycle gene CDC7 from Saccharomyces cerevisiae. *Mol Cell Biol,* 6**,** 1590-8.

PERKINS, G. & DIFFLEY, J. F. X. 1998. Nucleotide-dependent prereplicative complex assembly by Cdc6p, a homolog of eukaryotic and prokaryotic clamp-loaders. *Mol Cell,* 2**,** 23-32.

PETERSEN, B. O., LUKAS, J., SORENSEN, C. S., BARTEK, J. & HELIN, K. 1999. Phosphorylation of mammalian CDC6 by Cyclin A/CDK2 regulates its subcellular localization. *EMBO J.,* 18**,** 396-410.

PFANDER, B. & DIFFLEY, J. F. X. 2011. Dpb11 coordinates Mec1 kinase activation with cell cycle-regulated Rad9 recruitment. *EMBO J,* 30**,** 4897-907.

PSAKHYE, I. & JENTSCH, S. 2012. Protein group modification and synergy in the SUMO pathway as exemplified in DNA repair. *Cell,* 151**,** 807-20.

PURSELL, Z. F., ISOZ, I., LUNDSTROM, E. B., JOHANSSON, E. & KUNKEL, T. A. 2007. Yeast DNA polymerase epsilon participates in leading-strand DNA replication. *Science,* 317**,** 127-30.

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**,** 115-21.

RAMER, M. D., SUMAN, E. S., RICHTER, H., STANGER, K., SPRANGER, M., BIEBERSTEIN, N. & DUNCKER, B. P. 2013. Dbf4 and Cdc7 proteins promote DNA replication through interactions with distinct Mcm2-7 protein subunits. *J Biol Chem,* 288**,** 14926-35.

RANDELL, J. C., BOWERS, J. L., RODRIGUEZ, H. K. & BELL, S. P. 2006. Sequential ATP hydrolysis by Cdc6 and ORC directs loading of the Mcm2-7 helicase. *Mol Cell,* 21**,** 29-39.

RANDELL, J. C., 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. *Mol Cell,* 40**,** 353-63.

RAO, H. & STILLMAN, B. 1995. The origin recognition complex interacts with a bipartite DNA binding site within yeast replicators. *Proc. Natl. Acad. Sci. USA,* 92**,** 2224-2228.

RAO, P. N. & JOHNSON, R. T. 1970. Mammalian cell fusion: studies on the regulation of DNA synthesis and mitosis. *Nature,* 225**,** 159-64.

RASCHLE, M., KNIPSHEER, P., ENOIU, M., ANGELOV, T., SUN, J., GRIFFITH, J. D., ELLENBERGER, T. E., SCHARER, O. D. & WALTER, J. C. 2008. Mechanism of replication-coupled DNA interstrand crosslink repair. *Cell,* 134**,** 969-80.

REMUS, D., BEALL, E. L. & BOTCHAN, M. R. 2004. DNA topology, not DNA sequence, is a critical determinant for Drosophila ORC-DNA binding. *Embo J,* 23**,** 897-907.

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**,** 719-30.

REMUS, D. & DIFFLEY, J. F. X. 2009. Eukaryotic DNA replication control: lock and load, then fire. *Curr Opin Cell Biol,* 21**,** 771-7.

RICKE, R. M. & BIELINSKY, A. K. 2004. Mcm10 regulates the stability and chromatin association of DNA polymerase-alpha. *Mol Cell,* 16**,** 173-85.

ROBERTSON, P. D., WARREN, E. M., ZHANG, H., FRIEDMAN, D. B., LARY, J. W., COLE, J. L., TUTTER, A. V., WALTER, J. C., FANNING, E. & EICHMAN, B. F. 2008. Domain architecture and biochemical characterization of vertebrate Mcm10. *J Biol Chem,* 283**,** 3338-48.

ROMANOWSKI, P., MADINE, M. A., ROWLES, A., BLOW, J. J. & LASKEY, R. A. 1996. The Xenopus origin recognition complex is essential for DNA replication and MCM binding to chromatin. *Curr Biol,* 6**,** 1416-25.

ROWLES, A., CHONG, J. P., BROWN, L., HOWELL, M., EVAN, G. I. & BLOW, J. J. 1996. Interaction between the origin recognition complex and the replication licensing system in Xenopus. *Cell,* 87**,** 287-96.

ROWLES, A., TADA, S. & BLOW, J. J. 1999. Changes in association of the Xenopus origin recognition complex with chromatin on licensing of replication origins. *J. Cell Sci.,* 112**,** 2011-2018.

ROWLEY, A., COCKER, J. H., HARWOOD, J. & DIFFLEY, J. F. X. 1995. Initiation complex assembly at budding yeast replication origins begins with the recognition of a bipartite sequence by limiting amounts of the initiator, ORC. *EMBO J,* 14**,** 2631-41.

SAKA, Y. & YANAGIDA, M. 1993. Fission yeast cut5+, required for S phase onset and M phase restraint, is identical to the radiation-damage repair gene rad4+. *Cell,* 74**,** 383-93.

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 Dev,* 28**,** 1653-66.

SANCHEZ-PULIDO, L., DIFFLEY, J. F. X. & PONTING, C. P. 2010. Homology explains the functional similarities of Treslin/Ticrr and Sld3. *Curr Biol,* 20**,** R509-10.

SANCHEZ-PULIDO, L. & PONTING, C. P. 2011. Cdc45: the missing RecJ ortholog in eukaryotes? *Bioinformatics,* 27**,** 1885-8.

SANGRITHI, M. N., BERNAL, J. A., MADINE, M., PHILPOTT, A., LEE, J., DUNPHY, W. G. & VENKITARAMAN, A. R. 2005. Initiation of DNA replication requires the RECQL4 protein mutated in Rothmund-Thomson syndrome. *Cell,* 121**,** 887-98.

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 Dev,* 24**,** 183-94.

SANTOCANALE, C. & DIFFLEY, J. F. X. 1996. ORC- and Cdc6-dependent complexes at active and inactive chromosomal replication origins in Saccharomyces cerevisiae. *Embo J,* 15**,** 6671-9.

SCHWACHA, A. & BELL, S. P. 2001. Interactions between two catalytically distinct MCM subgroups are essential for coordinated ATP hydrolysis and DNA replication. *Mol Cell,* 8**,** 1093-104.

SEKI, T. & DIFFLEY, J. F. X. 2000. Stepwise assembly of initiation proteins at budding yeast replication origins in vitro. *Proc Natl Acad Sci U S A,* 97**,** 14115-20.

SEMPLE, J. W., DA-SILVA, L. F., JERVIS, E. J., AH-KEE, J., AL-ATTAR, H., KUMMER, L., HEIKKILA, J. J., PASERO, P. & DUNCKER, B. P. 2006. An essential role for Orc6 in DNA replication through maintenance of pre-replicative complexes. *Embo J,* 25**,** 5150-8.

SENGA, T., SIVAPRASAD, U., ZHU, W., PARK, J. H., ARIAS, E. E., WALTER, J. C. & DUTTA, A. 2006. PCNA is a cofactor for Cdt1 degradation by CUL4/DDB1-mediated N-terminal ubiquitination. *J Biol Chem,* 281**,** 6246-52.

SENGUPTA, S., VAN DEURSEN, F., DE PICCOLI, G. & LABIB, K. 2013. Dpb2 integrates the leading-strand DNA polymerase into the eukaryotic replisome. *Curr Biol,* 23**,** 543-52.

SHECHTER, D. F., YING, C. Y. & GAUTIER, J. 2000. The intrinsic DNA helicase activity of Methanobacterium thermoautotrophicum delta H minichromosome maintenance protein. *J Biol Chem,* 275**,** 15049-59.

SHEN, Z., SATHYAN, K. M., GENG, Y., ZHENG, R., CHAKRABORTY, A., FREEMAN, B., WANG, F., PRASANTH, K. V. & PRASANTH, S. G. 2010. A WD-repeat protein stabilizes ORC binding to chromatin. *Mol Cell,* 40**,** 99-111.

SHEU, Y. J. & STILLMAN, B. 2006. Cdc7-Dbf4 phosphorylates MCM proteins via a docking site-mediated mechanism to promote S phase progression. *Mol Cell,* 24**,** 101-13.

SHEU, Y. J. & STILLMAN, B. 2010. The Dbf4-Cdc7 kinase promotes S phase by alleviating an inhibitory activity in Mcm4. *Nature,* 463**,** 113-7.

SIDDIQUI, K., ON, K. F. & DIFFLEY, J. F. X. 2013. Regulating DNA Replication in Eukarya. *Cold Spring Harb Perspect Biol*.

SIEVERS, F., WILM, A., DINEEN, D., GIBSON, T. J., KARPLUS, K., LI, W., LOPEZ, R., MCWILLIAM, H., REMMERT, M., SODING, J., THOMPSON, J. D. & HIGGINS, D. G. 2011. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol Syst Biol,* 7**,** 539.

SPECK, C., CHEN, Z., LI, H. & STILLMAN, B. 2005. ATPase-dependent cooperative binding of ORC and Cdc6 to origin DNA. *Nat Struct Mol Biol,* 12**,** 965-71.

SPELLMAN, P. T., SHERLOCK, G., ZHANG, M. Q., IYER, V. R., ANDERS, K., EISEN, M. B., BROWN, P. O., BOTSTEIN, D. & FUTCHER, B. 1998. Comprehensive identification of cell cycle-regulated genes of the yeast Saccharomyces cerevisiae by microarray hybridization. *Mol Biol Cell,* 9**,** 3273-97.

STINCHCOMB, D. T., STRUHL, K. & DAVIS, R. W. 1979. Isolation and characterisation of a yeast chromosomal replicator. *Nature,* 282**,** 39-43.

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. *Nat Struct Mol Biol,* 20**,** 944-51.

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 Res,* 42**,** 2308-19.

TADA, S., LI, A., MAIORANO, D., MECHALI, M. & BLOW, J. J. 2001. Repression of origin assembly in metaphase depends on inhibition of RLF- B/Cdt1 by geminin. *Nat Cell Biol,* 3**,** 107-113.

TAKAHASHI, K., YAMADA, H. & YANAGIDA, M. 1994. Fission yeast minichromosome loss mutants mis cause lethal aneuploidy and replication abnormality. *Mol Biol Cell,* 5**,** 1145-1158.

TAKARA, T. J. & BELL, S. P. 2011. Multiple Cdt1 molecules act at each origin to load replication-competent Mcm2-7 helicases. *EMBO J,* 30**,** 4885-96.

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 Dev,* 17**,** 1153-65.

TANAKA, S. & ARAKI, H. 2013. Helicase Activation and Establishment of Replication Forks at Chromosomal Origins of Replication. *Cold Spring Harb Perspect Biol*.

TANAKA, S. & DIFFLEY, J. F. X. 2002. Interdependent nuclear accumulation of budding yeast Cdt1 and Mcm2-7 during G1 phase. *Nat Cell Biol,* 4**,** 198-207.

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. *Curr Biol,* 21**,** 2055-63.

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**,** 328-32.

TANAKA, T., UMEMORI, T., ENDO, S., MURAMATSU, S., KANEMAKI, M., KAMIMURA, Y., OBUSE, C. & ARAKI, H. 2011b. Sld7, an Sld3-associated protein required for efficient chromosomal DNA replication in budding yeast. *EMBO J,* 30**,** 2019-30.

TERCERO, J. A., LABIB, K. & DIFFLEY, J. F. X. 2000. DNA synthesis at individual replication forks requires the essential initiation factor Cdc45p. *EMBO J.,* 19**,** 2082-93.

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. *EMBO J,* 16**,** 3312-9.

VAN DEURSEN, F., SENGUPTA, S., DE PICCOLI, G., SANCHEZ-DIAZ, A. & LABIB, K. 2012. Mcm10 associates with the loaded DNA helicase at replication origins and defines a novel step in its activation. *EMBO J,* 31**,** 2195-206.

VAN HOUTEN, J. V. & NEWLON, C. S. 1990. Mutational analysis of the consensus sequence of a replication origin from yeast chromosome III. *Mol Cell Biol,* 10**,** 3917-25.

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 Dev,* 17**,** 1894-908.

VENDITTI, P., COSTANZO, G., NEGRI, R. & CAMILLONI, G. 1994. ABFI contributes to the chromatin organization of Saccharomyces cerevisiae ARS1 B-domain. *Biochim Biophys Acta,* 1219**,** 677-89.

VERMA, R., ANNAN, R. S., HUDDLESTON, M. J., CARR, S. A., REYNARD, G. & DESHAIES, R. J. 1997. Phosphorylation of Sic1p by G1 Cdk Required for Its Degradation and Entry into S Phase. *Science,* 278**,** 455-460.

WALTER, J. & NEWPORT, J. 2000. Initiation of eukaryotic DNA replication: origin unwinding and sequential chromatin association of Cdc45, RPA, and DNA polymerase alpha. *Mol Cell,* 5**,** 617-27.

WALTER, J. C. 2000. Evidence for sequential action of cdc7 and cdk2 protein kinases during initiation of DNA replication in Xenopus egg extracts. *J Biol Chem,* 275**,** 39773-8.

WARREN, E. M., VAITHIYALINGAM, S., HAWORTH, J., GREER, B., BIELINSKY, A. K., CHAZIN, W. J. & EICHMAN, B. F. 2008. Structural basis for DNA binding by replication initiator Mcm10. *Structure,* 16**,** 1892-901.

WATASE, G., TAKISAWA, H. & KANEMAKI, M. T. 2012. Mcm10 plays a role in functioning of the eukaryotic replicative DNA helicase, Cdc45-Mcm-GINS. *Curr Biol,* 22**,** 343-9.

WEINREICH, M., LIANG, C., CHEN, H. H. & STILLMAN, B. 2001. Binding of cyclin-dependent kinases to ORC and Cdc6p regulates the chromosome replication cycle. *Proc Natl Acad Sci U S A,* 98**,** 11211-7.

WEINREICH, M., LIANG, C. & STILLMAN, B. 1999. The Cdc6p nucleotide-binding motif is required for loading Mcm proteins onto chromatin. *Proc. Natl. Acad. Sci. USA,* 96**,** 441-446.

WEINREICH, M. & STILLMAN, B. 1999. Cdc7p-Dbf4p kinase binds to chromatin during S phase and is regulated by both the APC and the RAD53 checkpoint pathway. *Embo J,* 18**,** 5334-5346.

WHITTAKER, A. J., ROYZMAN, I. & ORR-WEAVER, T. L. 2000. Drosophila double parked: a conserved, essential replication protein that colocalizes with the origin recognition complex and links DNA replication with mitosis and the down-regulation of S phase transcripts. *Genes Dev,* 14**,** 1765-76.

WILLIAMS, R. S., SHOHET, R. V. & STILLMAN, B. 1997. A human protein related to yeast Cdc6p. *Proc. Natl. Acad. Sci. USA,* 94**,** 142-7.

WILMES, G. M., ARCHAMBAULT, V., AUSTIN, R. J., JACOBSON, M. D., BELL, S. P. & CROSS, F. R. 2004. Interaction of the S-phase cyclin Clb5 with an 'RXL' docking sequence in the initiator protein Orc6 provides an origin-localized replication control switch. *Genes Dev,* 18**,** 981-91.

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. *Proc Natl Acad Sci U S A,* 99**,** 101-6.

WOHLSCHLEGEL, J. A., DHAR, S. K., PROKHOROVA, T. A., DUTTA, A. & WALTER, J. C. 2002. Xenopus Mcm10 binds to origins of DNA replication after Mcm2-7 and stimulates origin binding of Cdc45. *Mol Cell,* 9**,** 233-40.

WOHLSCHLEGEL, J. A., DWYER, B. T., DHAR, S. K., CVETIC, C., WALTER, J. C. & DUTTA, A. 2000. Inhibition of eukaryotic DNA replication by geminin binding to Cdt1. *Science,* 290**,** 2309-12.

WYRICK, J. J., APARICIO, J. G., CHEN, T., BARNETT, J. D., JENNINGS, E. G., YOUNG, R. A., BELL, S. P. & APARICIO, O. M. 2001. Genome-wide distribution of ORC and MCM proteins in S. cerevisiae: high-resolution mapping of replication origins. *Science,* 294**,** 2357-60.

YABUUCHI, H., YAMADA, Y., UCHIDA, T., SUNATHVANICHKUL, T., NAKAGAWA, T. & MASUKATA, H. 2006. Ordered assembly of Sld3, GINS and Cdc45 is distinctly regulated by DDK and CDK for activation of replication origins. *EMBO J,* 25**,** 4663-74.

YAMADA, Y., NAKAGAWA, T. & MASUKATA, H. 2004. A novel intermediate in initiation complex assembly for fission yeast DNA replication. *Mol Biol Cell,* 15**,** 3740-50.

YAN, H., GIBSON, S. & TYE, B. K. 1991. Mcm2 and Mcm3, two proteins important for ARS activity, are related in structure and function. *Genes Dev,* 5**,** 944-57.

YANAGI, K., MIZUNO, T., YOU, Z. & HANAOKA, F. 2002. Mouse geminin inhibits not only Cdt1-MCM6 interactions but also a novel intrinsic Cdt1 DNA binding activity. *J Biol Chem,* 277**,** 40871-80.

YARDIMCI, H., LOVELAND, A. B., HABUCHI, S., VAN OIJEN, A. M. & WALTER, J. C. 2010. Uncoupling of sister replisomes during eukaryotic DNA replication. *Mol Cell,* 40**,** 834-40.

YARDIMCI, H. & WALTER, J. C. 2014. Prereplication-complex formation: a molecular double take? *Nat Struct Mol Biol,* 21**,** 20-5.

YASUDA, S. & HIROTA, Y. 1977. Cloning and Mapping of the Replication Origin of *Escherichia coli*. *Proc. Natl. Acad. Sci. USA,* 74**,** 5458-5462.

YOON, H. J., LOO, S. & CAMPBELL, J. L. 1993. Regulation of *Saccharomyces cerevisiae* CDC7 function during the cell cycle. *Mol Biol Cell,* 4**,** 195-208.

ZACHARIAE, W., SCHWAB, M., NASMYTH, K. & SEUFERT, W. 1998. Control of cyclin ubiquitination by CDK-regulated binding of Hct1 to the anaphase promoting complex. *Science,* 282**,** 1721-4.

ZEGERMAN, P. & DIFFLEY, J. F. X. 2007. Phosphorylation of Sld2 and Sld3 by cyclin-dependent kinases promotes DNA replication in budding yeast. *Nature,* 445**,** 281-5.

ZEGERMAN, P. & DIFFLEY, J. F. X. 2010. Checkpoint-dependent inhibition of DNA replication initiation by Sld3 and Dbf4 phosphorylation. *Nature,* 467**,** 474-8.

ZHANG, J., YU, L., WU, X., ZOU, L., SOU, K. K., WEI, Z., CHENG, X., ZHU, G. & LIANG, C. 2010. The interacting domains of hCdt1 and hMcm6 involved in the chromatin loading of the MCM complex in human cells. *Cell Cycle,* 9**,** 4848-57.

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. *Mol Cell Biol,* 17**,** 553-63.

ZOU, L. & STILLMAN, B. 1998. Formation of a preinitiation complex by S-phase cyclin CDK-dependent loading of Cdc45p onto chromatin. *Science,* 280**,** 593-6.