

DNA replication initiation studied by cryo-EM

Julia Franziska Greiwe

University College London

and

The Francis Crick Institute

PhD Supervisor: Dr Alessandro Costa

A thesis submitted for the degree of

Doctor of Philosophy

University College London

December 2022

Declaration

I, Julia Franziska Greiwe 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.

Julia Greiwe

Abstract

Eukaryotic DNA replication is tightly regulated to ensure faithful duplication of genetic information before cell division. The motor of the replicative helicase is the minichromosome maintenance (MCM) complex, which unwinds duplex DNA and exposes single-stranded DNA to replicative polymerases. The MCM is loaded onto duplex DNA as an inactive double hexamer (DH), in a process named origin licensing. I solved a 3.0 Å resolution structure of the DNA-loaded *Saccharomyces cerevisiae* DH, showing that Watson–Crick base pairing remains intact. DH phosphorylation by the Dbf4-dependent kinase (DDK) represents the first step towards activation of the helicase, nucleation of DNA melting and replication fork establishment. DDK selectively targets DNA-loaded DHs, but the mechanism is unknown. Using *in vitro* reconstitution and cryo-electron microscopy (cryo-EM) I discovered that DDK recognises the three-dimensional structure of the DH by docking onto one MCM hexamer and phosphorylating the other hexamer. Truncation of the docking domain of DDK does not affect kinase function *per se* but abrogates DH phosphorylation. DDK autophosphorylation is known to reduce kinase activity. While this is true for an isolated MCM peptide, I established that DH phosphorylation is virtually unaffected, meaning selectivity for the DH is increased. When DNA damage is detected in S phase, the checkpoint kinase Rad53 targets DDK to stop origin firing. I found that Rad53 phosphorylation of DDK blocks DH engagement and helicase activation.

Little is known about DNA replication in humans. In a collaborative effort, I reconstituted human origin licensing *in vitro*. I then analysed the entire reaction by cryo-EM to identify five protein assemblies that present similarities and differences compared to helicase loading intermediates in *S. cerevisiae*. Unlike yeast DH, I found that the human DH untwists and melts duplex DNA. My results indicate that nucleation of origin DNA melting occurs *via* distinct mechanisms in yeast and humans.

Impact Statement

DNA replication is essential for the propagation of life. Errors during the duplication of the genome have devastating consequences, including genomic instability, cell death and the onset of diseases such as cancer. Understanding the mechanisms underlying DNA replication is therefore of great interest to academic researchers, pharmaceutical industry, and the public. The work presented in this dissertation provides knowledge and tools to facilitate future research projects, and forms the basis for structure-based drug design and educational purposes.

In vitro reconstitution using purified budding yeast proteins has proven invaluable to dissect and structurally characterise the molecular mechanisms of DNA replication. While the fundamental principles of this process are conserved across eukaryotes, understanding human biology and pathology requires that human proteins are studied. With my work, I contributed to the development of a biochemical toolbox to reconstitute human DNA replication. My structural analysis provided insights into the process of helicase loading and DNA melting, identifying fundamental differences between the budding yeast and human system. To disseminate the findings and enable other researchers to study the entire DNA replication reaction, we will communicate our results with the scientific community at national and international conferences and *via* publications in open-access journals and will share reagents and protocols. The data on the mechanism of helicase activation by the Dbf4-dependent kinase has already been published and presented in several international meetings, which resulted in stimulating discussions and prompted others to formulate new research questions.

I employed an unconventional approach to structural biology, in that I did not focus my analysis on purified macromolecular assemblies. I rather analysed multicomponent reactions by electron microscopy. While datasets were highly heterogeneous, I demonstrated that reaction intermediates can be captured, which would not be attainable with biochemical purification procedures. I hope to inspire colleagues to approach their research questions in a similar manner, so that the molecular mechanism of other complex biochemical pathways can be uncovered.

The Dbf4-dependent kinase is highly expressed in many cancer cells and its overexpression correlates with poor prognosis for patients. Inhibition of the kinase using ATP-competitive small-molecule inhibitors shows good results in the clinic, but these inhibitors also affect other kinases due to the conserved architecture of the active site, causing negative effects. Our structural characterisation of the selective kinase–helicase interaction can potentially serve as a basis for developing new generation inhibitors that target protein–protein interactions. We published this work in an open-access, scientific journal and the structure is available in open-access repositories (EMDB and PDB) so that pharmaceutical companies can review potential therapeutic interventions.

The public will benefit from the work described in my thesis, as research outputs can be used to demystify the process of how cells duplicate their genetic material before dividing. With the structures of reaction intermediates, we can generate a molecular movie that provides a visual aid to communicate science in an accessible manner. As such, images of a DNA molecule generated from my data have already featured in Maverick Television “Operation Ouch!”, an educational program for children.

UCL Research Paper Declaration Form

referencing the doctoral candidate's own published work(s)

Please use this form to declare if parts of your thesis are already available in another format, e.g. if data, text, or figures:

- *have been uploaded to a preprint server*
- *are in submission to a peer-reviewed publication*
- *have been published in a peer-reviewed publication, e.g. journal, textbook.*

This form should be completed as many times as necessary. For instance, if you have seven thesis chapters, two of which containing material that has already been published, you would complete this form twice.

1. For a research manuscript that has already been published (if not yet published, please skip to section 2)

a) What is the title of the manuscript?

Structural mechanism for the selective phosphorylation of DNA-loaded MCM double hexamers by the Dbf4-dependent kinase

b) Please include a link to or doi for the work

<https://doi.org/10.1038/s41594-021-00698-z>

c) Where was the work published?

Nature Structural & Molecular Biology

d) Who published the work? (e.g. OUP)

Springer Nature

e) When was the work published?

January 2022 (published online 28 December 2021)

f) List the manuscript's authors in the order they appear on the publication

Julia F. Greiwe, Thomas C. R. Miller, Julia Locke, Fabrizio Martino, Steven Howell, Anne Schreiber, Andrea Nans, John F. X. Diffley and Alessandro Costa

g) Was the work peer reviewed?

Yes

h) Have you retained the copyright?

Yes

i) Was an earlier form of the manuscript uploaded to a preprint server? (e.g. medRxiv). If 'Yes', please give a link or doi)

No

If 'No', please seek permission from the relevant publisher and check the box next to the below statement:



*I acknowledge permission of the publisher named under **1d** to include in this thesis portions of the publication named as included in **1c**.*

2. For a research manuscript prepared for publication but that has not yet been published (if already published, please skip to section 3)

a) **What is the current title of the manuscript?**

b) **Has the manuscript been uploaded to a preprint server?** (e.g. medRxiv; if 'Yes', please give a link or doi)

c) **Where is the work intended to be published?** (e.g. journal names)

d) **List the manuscript's authors in the intended authorship order**

e) **Stage of publication** (e.g. in submission)

3. For multi-authored work, please give a statement of contribution covering all authors (if single-author, please skip to section 4)

J.F.G. designed and performed all experiments, interpreted the data and wrote the manuscript with A.C. T.C.R.M. contributed to EM experiments. J.L. and A.N. contributed to cryo-EM experiments. F.M. contributed to model building. S.H. contributed to mass spectrometry analysis. A.S. contributed to peptide phosphorylation assays. J.F.X.D. contributed protein expression strains, experimental protocols and discussion. A.C. supervised the study.

4. In which chapter(s) of your thesis can this material be found?

Chapters 3.2 and 4

5. e-Signatures confirming that the information above is accurate (this form should be co-signed by the supervisor/ senior author unless this is not appropriate, e.g. if the paper was a single-author work)

Candidate

Date:

16 December 2022

Supervisor/ Senior Author (where appropriate)

Date

16 December 2022

Acknowledgements

Firstly, I would like to express my deepest gratitude to my supervisor Dr Alessandro Costa for his continuous support and guidance throughout the years. Alessandro – I am eternally grateful for your patience and your confidence in my abilities. Thank you for sharing your passion for DNA replication and cryo-EM with me. I would also like to thank my thesis committee, Drs John Diffley, Radoslav Enchev, Alan Cheung and Graeme King for their guidance.

I am forever indebted to Dr Thomas Miller, without your faith in me and my skills, I would literally not be where I am today. Thank you for being an amazing mentor and friend. To Dr Jacob Lewis, thanks for sharing your incredible knowledge, giving invaluable advice, and suffering with me through some impossible tasks. I would also like to extend gratitude to Dr Julia Locke for teaching me cryo-EM and Dr Fabrizio Martino for his model building expertise. To all former and present members of the Costa lab, thank you for your help, friendship and for creating an enjoyable working environment. I also thank Drs Sarah Henrikus, Jacob Lewis and Oliver Willhoft for reading drafts of this thesis.

Massive thank you to Dr Florian Weissmann for a highly enjoyable and rewarding collaboration. Thanks to Dr John Diffley for all the stimulating discussions. Thanks to the members of the Diffley lab, in particular Drs Allison McClure, Berta Canal and Anne Early for providing reagents, help and advice. To Dr Agostina Bertolin, thank you for instilling confidence in me, especially during the final weeks of my PhD. I would like to thank the community of researchers at the Francis Crick Institute for their contribution, including Drs Andrea Nans, Donald Benton, Andrew Purkiss, Philip Walker, Alireza Alidoust, Namita Patel and Damini Patel (Structural Biology STP), Steven Howell (Proteomics STP), Raffaella Carzaniga (Electron Microscopy STP), Peter Rosenthal, Anne Schreiber, Emma Punch, Valerie Pye and Peter Cherepanov.

Finally, I want to thank my family and friends for their endless support, listening to my worries and keeping me sane. Special thanks to my parents – without your belief and generosity none of this would have been possible.

Table of Contents

Abstract	3
Impact Statement	4
UCL Research Paper Declaration Form	6
Acknowledgements	8
Table of Contents	9
Table of figures	13
List of tables	16
Abbreviations	17
Chapter 1. Introduction	20
1.1 Eukaryotic DNA replication – a highly orchestrated process	20
1.2 Structure and function of the Mcm2–7 helicase	23
1.2.1 Structure of the Mcm2–7 helicase	23
1.2.2 ATP hydrolysis drives translocation	26
1.3 Origin licensing	27
1.3.1 Origin recognition by ORC	27
1.3.2 Origin licensing	30
1.4 Origin firing and chromatin replication	33
1.4.1 Phosphorylation of the MCM double hexamer by DDK	34
1.4.2 Recruitment of the helicase activators Cdc45 and GINS	35
1.4.3 Replication fork establishment	36
1.5 Regulation of origin licensing during the cell cycle	37
1.5.1 Importance of faithful DNA replication initiation for human health	38
1.6 Structure and function of the Dbf4-dependent kinase during origin firing	39
1.6.1 Structure and function of Cdc7–Dbf4	39
1.6.2 Regulation of DDK	42
Drf1 as an alternative activator of Cdc7	44
1.6.3 44	
1.7 Aims	46
Chapter 2. Materials & Methods	47
2.1 Reagents	47
2.1.1 <i>Saccharomyces cerevisiae</i> strains	47
2.1.2 Oligonucleotides	49
2.1.3 Plasmids	50
2.1.4 Buffers	50
2.1.5 Media for <i>S. cerevisiae</i> and <i>E. coli</i> cells	53
2.2 Generation of <i>S. cerevisiae</i> strains for protein expression	54
2.2.1 Plasmid generation	54
2.2.2 Transformation of <i>S. cerevisiae</i>	56
2.2.3 Isolation of yeast genomic DNA	57
2.2.4 Colony-PCR	57
2.2.5 Small-scale expression to select <i>S. cerevisiae</i> strains	58
2.3 Protein expression in <i>S. cerevisiae</i>	59
2.3.1 ORC and Mcm2–7–Cdt1 expression	59
2.3.2 DDK expression	59

2.4 Protein expression in <i>E. coli</i>	60
2.4.1 Cdc6 expression	60
2.4.2 Rad53 expression	61
2.4.3 Histone octamer expression	61
2.5 Expression and purification of <i>H. sapiens</i> DNA replication factors	62
2.6 Purification of <i>S. cerevisiae</i> proteins	62
2.6.1 <i>S. cerevisiae</i> ORC purification	62
2.6.2 <i>S. cerevisiae</i> Mcm2–7–Cdt1 purification	63
2.6.3 <i>S. cerevisiae</i> DDK purification	63
2.6.4 <i>S. cerevisiae</i> Cdc6 purification	64
2.6.5 <i>S. cerevisiae</i> Rad53 purification	65
2.6.6 <i>S. cerevisiae</i> histone octamer purification	66
2.7 Preparation of origin DNA templates	66
2.7.1 Amplification and purification of DNA templates	66
2.7.2 Preparation of the M. <i>Hpa</i> II–ARS1–M. <i>Hpa</i> II DNA template	67
2.7.3 Preparation of the nucleosome–Gid70–nucleosome DNA template	68
2.8 <i>In vitro</i> reconstitution	68
2.8.1 <i>S. cerevisiae</i> DH loading and phosphorylation analysed by SDS-PAGE	68
2.8.2 Reconstitution of <i>S. cerevisiae</i> DH–DDK for negative stain EM	69
2.8.3 Reconstitution of <i>S. cerevisiae</i> DH–DDK for cryo-EM	70
2.8.4 Inhibition of DH phosphorylation by Rad53 analysed by negative stain EM	70
2.8.5 Reconstitution of <i>H. sapiens</i> origin licensing for cryo-EM	71
2.8.6 Purification of the DNA-bound <i>S. cerevisiae</i> DH using streptavidin-coated magnetic beads	71
2.8.7 Purification of the DNA-bound <i>S. cerevisiae</i> DH by anion-exchange chromatography	72
2.9 Peptide phosphorylation assay	72
2.10 Mcm2–7–Cdt1 phosphorylation by DDK	73
2.11 Mass spectrometry analysis of MCM phosphorylation	73
2.12 Electron microscopy sample preparation	75
2.12.1 Negative stain grid preparation	75
2.12.2 Deposition of graphene oxide on cryo-EM grids	75
2.12.3 Cryo-EM grid preparation for <i>S. cerevisiae</i> DDK	76
2.12.4 Cryo-EM grid preparation for <i>S. cerevisiae</i> DH–DDK	76
2.12.5 Cryo-EM grid preparation for <i>H. sapiens</i> origin licensing	76
2.13 Electron microscopy data collection	77
2.13.1 Negative stain data collection	77
2.13.2 <i>S. cerevisiae</i> DDK cryo-EM data collection	77
2.13.3 <i>S. cerevisiae</i> DH–DDK cryo-EM data collection	77
2.13.4 <i>H. sapiens</i> origin licensing cryo-EM data collection	77
2.14 Electron microscopy image processing	78
2.14.1 Negative stain image processing	78
2.14.2 <i>S. cerevisiae</i> DDK cryo-EM image processing	78
2.14.3 <i>S. cerevisiae</i> DH–DDK cryo-EM image processing	79
2.14.4 Cryo-EM image processing for <i>H. sapiens</i> origin licensing in the presence of ATP	81

2.14.5 Cryo-EM image processing for initial stages of <i>H. sapiens</i> MCM loading in the presence of ATP γ S	82
2.15 EM model building	84
2.15.1 <i>S. cerevisiae</i> DH–DDK cryo-EM model building	84
2.15.2 <i>H. sapiens</i> OCCM cryo-EM model building	84
2.15.3 <i>H. sapiens</i> SH cryo-EM model building	85
2.15.4 <i>H. sapiens</i> MO cryo-EM model building	85
2.15.5 <i>H. sapiens</i> DH cryo-EM model building	85
2.15.6 Analysis of protein–DNA contacts	85
Chapter 3. Results 1 – Origin licensing in <i>S. cerevisiae</i> and <i>H. sapiens</i>	87
3.1 Introduction	87
3.2 <i>In vitro</i> reconstitution of origin licensing using <i>S. cerevisiae</i> proteins	89
3.2.1 Cryo-EM analysis of <i>S. cerevisiae</i> origin licensing	89
3.2.2 Overall architecture of the <i>S. cerevisiae</i> MCM DH	92
3.2.3 Interaction of the <i>S. cerevisiae</i> MCM DH with double-stranded DNA	93
3.2.4 ATPase state of the <i>S. cerevisiae</i> MCM DH	95
3.3 <i>In vitro</i> reconstitution of origin licensing using <i>H. sapiens</i> proteins	96
3.3.1 Reconstitution of <i>H. sapiens</i> origin licensing	96
3.3.2 Visualising initial stages of <i>H. sapiens</i> MCM loading in the presence of ATP γ S by cryo-EM	99
3.3.3 Visualising ATP-dependent <i>H. sapiens</i> DH loading by cryo-EM	102
3.4 Cryo-EM structures of <i>H. sapiens</i> OCCM and OC1M complexes	104
3.4.1 Image processing of the <i>H. sapiens</i> OCCM and OC1M complexes	104
3.4.2 Overall architecture of the <i>H. sapiens</i> OCCM and OC1M complexes	107
3.4.3 Several MCM winged helix domains establish contacts with ORC and CDC6	109
3.4.4 Protein–DNA interactions and nucleotide binding in <i>H. sapiens</i> OCCM/OC1M	110
3.5 Cryo-EM structure of <i>H. sapiens</i> single-loaded MCM hexamers	112
3.5.1 Image processing of the <i>H. sapiens</i> SH	112
3.5.2 The <i>H. sapiens</i> SH binds double-stranded DNA	116
3.5.3 ATPase state of the <i>H. sapiens</i> SH	118
3.6 Cryo-EM structure of <i>H. sapiens</i> MO	120
3.6.1 Image processing of the <i>H. sapiens</i> MO complex	120
3.6.2 Overall architecture of the <i>H. sapiens</i> MO complex	122
3.6.3 DNA and nucleotide binding by <i>H. sapiens</i> MO	123
3.6.4 <i>H. sapiens</i> MO appears to be in an inhibited conformation	125
3.7 Cryo-EM structure of the <i>H. sapiens</i> DH	128
3.7.1 Image processing of the <i>H. sapiens</i> DH	128
3.7.2 DH formation causes disruption of Watson–Crick base pairing	130
3.7.3 ATPase state of <i>H. sapiens</i> DH	135
3.8 Discussion	138
3.8.1 The <i>S. cerevisiae</i> DH is an inactive form of the MCM helicase and does not unwind DNA	138
3.8.2 <i>In vitro</i> reconstitution of origin licensing using <i>H. sapiens</i> proteins	139
3.8.3 Nucleation of DNA melting by the <i>H. sapiens</i> DH	141

3.8.4 ATP hydrolysis and DNA binding during origin licensing	143
Chapter 4. Results 2 – Structural mechanism for the selective phosphorylation of DNA-loaded DHs by the Dbf4-dependent kinase	144
4.1 Introduction	144
4.2 Structural characterisation of the isolated DDK.....	146
4.2.1 Structural characterisation of full-length <i>S. cerevisiae</i> DDK	146
4.3 Structure of DH–DDK and molecular mechanism for substrate selectivity	149
4.3.1 <i>In vitro</i> reconstitution of DH–DDK	149
4.3.2 <i>In silico</i> purification of DH–DDK.....	152
4.3.3 The catalytic core of DDK binds to the Mcm4 subunit of the MCM DH	157
4.3.4 The Dbf4 BRCT domain binds to Mcm2	161
4.3.5 DDK binds to both MCM rings in the DH to phosphorylate the MCM <i>in trans</i>	162
4.3.6 Phosphorylation of MCM–Cdt1 by DDK	165
4.3.7 DDK autophosphorylation increases substrate selectivity	168
4.4 Regulation of DDK by Rad53.....	170
4.4.1 Rad53-dependent phosphorylation in context of the DH–DDK complex	170
4.4.2 Catalytically active Rad53 reduces peptide phosphorylation efficiency of DDK.....	171
4.4.3 Phosphorylation of residues in the Dbf4 BRCT domain alone does not prevent DH phosphorylation.....	173
4.4.4 Rad53-dependent phosphorylation of DDK prevents DH–DDK interaction	175
4.5 Discussion	177
4.5.1 Selective phosphorylation of the DH by DDK	177
4.5.2 Regulation of DDK by Rad53	179
Chapter 5. Concluding summary.....	180
Chapter 6. Appendix.....	182
Reference List	189

Table of figures

Figure 1.1: Temporal separation of helicase loading and activation during the cell cycle.....	22
Figure 1.2: Structure of the Mcm2–7 hexamer.	24
Figure 1.3: Pore loop staircase in the <i>D. melanogaster</i> CMG.....	27
Figure 1.4: Origin recognition by <i>S. cerevisiae</i> ORC.	29
Figure 1.5: Origin licensing in <i>S. cerevisiae</i>	31
Figure 1.6: Origin firing.....	33
Figure 1.7: CMG formation is concomitant with disruption of the MCM dimerization interface and nucleation of DNA melting.	36
Figure 1.8: Crystal structure of the catalytic core of <i>H. sapiens</i> DDK.	40
Figure 3.1: <i>S. cerevisiae</i> DH formation analysed by cryo-EM.....	90
Figure 3.2: Processing pipeline for the DNA-loaded <i>S. cerevisiae</i> DH.	91
Figure 3.3: Structure quality of the DNA-loaded <i>S. cerevisiae</i> DH.....	92
Figure 3.4: Overall architecture of the <i>S. cerevisiae</i> DH.....	93
Figure 3.5: Protein–DNA contacts established by the <i>S. cerevisiae</i> DH.	94
Figure 3.6: Nucleotide occupancy in the <i>S. cerevisiae</i> DH.	96
Figure 3.7: <i>In vitro</i> reconstitution of origin licensing using <i>H. sapiens</i> proteins.	99
Figure 3.8: Cryo-EM analysis of <i>H. sapiens</i> MCM recruitment in the presence of ATP γ S.....	100
Figure 3.9: Processing pipeline <i>H. sapiens</i> MCM recruitment in the presence of ATP γ S.....	101
Figure 3.10: Cryo-EM analysis of <i>H. sapiens</i> MCM loading in the presence of ATP.	102
Figure 3.11: Initial processing steps of the <i>H. sapiens</i> origin licensing reaction in the presence of ATP.	103
Figure 3.12: Processing pipeline for <i>H. sapiens</i> OCCM and OC1M complexes...	105
Figure 3.13: Structure quality of <i>H. sapiens</i> OCCM and OC1M.....	106
Figure 3.14: Overall architecture of <i>H. sapiens</i> OCCM and OC1M.....	108
Figure 3.15: Interactions between MCM WHDs and ORC–CDC6/ORC in <i>H. sapiens</i> and <i>S. cerevisiae</i> OCCM/OC1M.	110
Figure 3.16: DNA and nucleotide binding by <i>H. sapiens</i> OCCM/OC1M.	111

Figure 3.17: Processing pipeline for <i>H. sapiens</i> SH (ATP dataset).....	114
Figure 3.18: Structure quality of the <i>H. sapiens</i> SH (ATP dataset).	114
Figure 3.19: Processing pipeline for <i>H. sapiens</i> SH and MO (ATP γ S dataset). ...	115
Figure 3.20: Protein–DNA contacts of the <i>H. sapiens</i> SH.....	117
Figure 3.21: Nucleotide occupancy of the <i>H. sapiens</i> SH.....	118
Figure 3.22: Structure quality of the <i>H. sapiens</i> MO complex.....	121
Figure 3.23: Overall architecture of the <i>H. sapiens</i> MO complex.....	123
Figure 3.24. DNA and nucleotide binding by <i>H. sapiens</i> MO.....	124
Figure 3.25: The <i>H. sapiens</i> MO is incompatible with recruitment of a second MCM helicase.....	126
Figure 3.26: ORC6 is dispensable for DH formation.....	127
Figure 3.27: Processing pipeline for the <i>H. sapiens</i> DH.	129
Figure 3.28: Structure quality of the DNA-loaded <i>H. sapiens</i> DH.	130
Figure 3.29: Overall structure of the <i>H. sapiens</i> DH and the Mcm7 helix-turn-helix motif that stabilises the <i>S. cerevisiae</i> DH.	131
Figure 3.30: The two MCM rings in <i>H. sapiens</i> DH are tilted and rotated against each other.....	132
Figure 3.31: Several structural elements cause DNA melting in <i>H. sapiens</i> DH. .	134
Figure 3.32: Nucleotide binding in correlation with pore loop binding.	137
Figure 4.1: Purification of full-length <i>S. cerevisiae</i> DDK.	147
Figure 4.2: <i>S. cerevisiae</i> DDK analysed by cryo-EM.	148
Figure 4.3: DH loading, phosphorylation and purification.	150
Figure 4.4: Imaging the entire DH loading and phosphorylation reaction improved decoration with DDK.	151
Figure 4.5: Image processing for the DH–DDK complex.	154
Figure 4.6: Structure quality of the DH–DDK complex.....	155
Figure 4.7: Multi-body analysis of the DH–DDK.	156
Figure 4.8: Binding of the DDK catalytic core to the DH.	159
Figure 4.9: Mcm4 is preferably phosphorylated.....	160
Figure 4.10: Resolving the Dbf4 BRCT domain bound to Mcm2.	161
Figure 4.11: Composite map of DH–DDK highlighting the bipartite binding of DDK.	162
Figure 4.12: The variant lacking the Dbf4 BRCT domain shows efficient autophosphorylation and MCM peptide phosphorylation.	164

Figure 4.13: DDK docking onto the MCM DH <i>in cis</i> is required for MCM phosphorylation <i>in trans</i>	165
Figure 4.14: Phosphorylation of MCM–Cdt1 can be achieved by increasing the DDK concentration.	166
Figure 4.15: DDK pre-autophosphorylation prevents peptide and MCM–Cdt1 phosphorylation, but not DH phosphorylation and engagement.	169
Figure 4.16: Rad53-dependent phosphorylation sites of Dbf4 in the context of the DH–DDK complex.	170
Figure 4.17: Phosphorylation of Dbf4 by Rad53 reduces peptide phosphorylation and abolishes DH phosphorylation.	172
Figure 4.18: Catalytic activity of the phospho-mimicking variant Cdc7–Dbf4 ^{T188D,S192D}	174
Figure 4.19: Phospho-mimicking mutations in the Dbf4 BRCT domain do not affect DH phosphorylation.	175
Figure 4.20: Phosphorylation of Dbf4 by Rad53 prevents DH–DDK interaction. .	176

List of tables

Table 2.1: Yeast strains	47
Table 2.2: Oligonucleotides.....	49
Table 2.3: Plasmids.....	50
Table 2.4: Buffers.....	50
Table 2.5: Media for <i>S. cerevisiae</i>	53
Table 2.6: Media for <i>E. coli</i>	54
Table 4.1: DDK-dependent phosphorylation of Mcm4 detected by mass spectrometry.	167
Appendix Table 1. Cryo-EM data collection, refinement and validation statistics for the <i>S. cerevisiae</i> DH and DH-DDK.....	182
Appendix Table 2. DDK-dependent phosphorylation sites of Mcm4 in the DH determined by mass spectrometry.	183
Appendix Table 3. DDK-dependent phosphorylation sites of Mcm4 in the DH determined by mass spectrometry (continued).....	184
Appendix Table 4. DDK-dependent phosphorylation sites of Mcm6 in the DH determined by mass spectrometry.	185
Appendix Table 5. DDK-dependent phosphorylation sites of Mcm6 in the DH determined by mass spectrometry (continued).....	186
Appendix Table 6. DDK-dependent phosphorylation sites of Mcm4 and Mcm6 in MCM-Cdt1 determined by mass spectrometry.	187
Appendix Table 7. DDK-dependent phosphorylation sites of Mcm4 and Mcm6 in MCM-Cdt1 determined by mass spectrometry (continued).....	188

Abbreviations

AAA+	ATPases associated with various cellular activities
Abf1	ARS-binding factor 1
ACS	ARS consensus sequence
ADP	adenosine diphosphate
ARS	autonomously replicating sequence
ASK	activator of the S phase kinase
ASKL1	activator of S phase kinase-like protein 1
ATP	adenosine triphosphate
BAH	bromo-adjacent homology
Bob1	bypass of block 1
BRCT	BRCA1 C-terminus
BSA	bovine serum albumin
CBP	calmodulin binding protein
Cdc	cell division cycle
CDK	cyclin dependent kinase
Cdh1	Cdc20 homolog 1
Cdt1	Cdc10 dependent transcription 1
CMG	Cdc45–MCM–GINS
CMGE	Cdc45–MCM–GINS–Pol ϵ
CTF	contrast transfer function
CV	column volume
Dbf4	dumbbell former 4
dCMGE	double Cdc45–MCM–GINS–Pol ϵ
DDK	Dbf4-dependent kinase
DH	double hexamer
DMSO	dimethyl sulfoxide
DTT	dithiothreitol
EDTA	ethylenediaminetetraacetic acid
EGTA	ethylene glycol tetraacetic acid
EM	electron microscopy
EMDB	Electron Microscopy Data Bank
FSC	Fourier Shell Correlation

GINS	go-ichi-nii-san (Japanese for 5-1-2-3)
GST	glutathione-S-transferase
h2i	helix-2-insertion
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
IDR	intrinsically disordered region
IPTG	isopropyl β -d-1-thiogalactopyranoside
KI	kinase insert
LB	lysogeny broth
MCM	minichromosome maintenance
Mec	mitosis entry checkpoint
M.HpaII	<i>HpaII</i> methyltransferase
MNase	micrococcal nuclease
Mrc1	mediator of the replication checkpoint 1
MWCO	molecular weight cut-off
dNTP	deoxyribonucleotide triphosphate
NTD	N-terminal domain / amino-terminal domain
OB-fold	oligonucleotide/oligosaccharide binding-fold
ORC	origin recognition complex
PAGE	polyacrylamide gel electrophoresis
PBS	phosphate buffered saline
PCR	polymerase chain reaction
PDB	Protein Data Bank
PEG	polyethylene glycol
PEP	posterior error probability
Pre-RC	pre-replicative complex
PS1	pre-sensor 1
Rad	radiation sensitive
Rif1	Rap-interacting factor 1
RPA	replication protein A
RPM	revolutions per minute
SDS	sodium dodecyl sulfate
Sld	synthetic lethal with Dpb11
SOC	super optimal broth with catabolite repression
TCA	trichloroacetic acid
TEG	triethyleneglycol

TFIIB	transcription factor IIB
TRP	tryptophan
WHD	winged-helix domain
Zn	Zinc

Chapter 1. Introduction

1.1 Eukaryotic DNA replication – a highly orchestrated process

All proliferating cells need to duplicate their genome before cell division to ensure faithful propagation of life. To read and duplicate the genetic information, a DNA helicase is first loaded at replication start sites, known as origins. The helicase then unwinds the DNA double helix and exposes both strands to replicative polymerases, which in turn synthesise new complementary DNA strands. Bacteria duplicate their circular genome from a single origin, whereas eukaryotes require multiple origins to replicate their genome, which is divided into multiple linear chromosomes (O'Donnell et al. 2013). As such, 400 origins have been identified in *Saccharomyces cerevisiae* and about 30,000–50,000 origins in *Homo sapiens* (Méchali et al. 2013). Replication start sites are defined DNA sequences in bacteria and *S. cerevisiae*. Most other eukaryotes define their origins by the surrounding chromatin context, which determines accessibility of the DNA to helicases, rather than the DNA sequence. Archaea share characteristics with both bacteria and eukaryotes by having multiple defined start sites on their circular genome (O'Donnell et al. 2013, Pérez-Arnaiz et al. 2020, Costa and Diffley 2022). Some archaea have evolved an additional mechanism to replicate their genome in the absence of origin sequences. This sequence-independent initiation is based on homologous recombination utilising the high copy number of chromosomes in these species (Pérez-Arnaiz et al. 2020). The observations made for the three kingdoms of life show that during evolution different mechanisms have developed to ensure duplication of the genome before cell division.

In eukaryotes, DNA replication must occur only once per cell cycle. Failure to this restriction results in accumulation of DNA damage and genomic instability (Arias and Walter 2007). A highly complex regulatory network has evolved in eukaryotes, which includes the temporal separation of helicase loading from helicase activation and DNA synthesis (Attali et al. 2021). Kinases play an important part in controlling these processes. Amongst them, the Dbf4-dependent kinase (DDK) and cyclin-dependent kinase (CDK) stimulate the recruitment of different factors to activate the helicase, while the S phase checkpoint kinase Rad53 prevents late origin firing

when DNA damage is detected (Tanaka et al. 2007, Zegerman and Diffley 2007, Labib 2010, Lopez-Mosqueda et al. 2010, Zegerman and Diffley 2010, Heller et al. 2011).

In bacteria, archaea and eukaryotes, duplex DNA is unwound by a six-subunit DNA helicase. Bacteria and most archaea have a homohexameric helicase (Pérez-Arnaiz et al. 2020). The core of the eukaryotic replicative helicase is formed from the ring-shaped minichromosome maintenance (MCM) complex, comprised of six structurally related proteins (Mcm2–7) (Dutta and Bell 1997, Forsburg 2004, Bochman and Schwacha 2009). The MCM is a molecular motor that uses ATP hydrolysis to translocate along, and unwind DNA. It also serves as a hub for the replicative polymerases and other processing factors to allow efficient DNA replication and organise chromatin on the duplicated DNA (Bell and Labib 2016, Miller and Costa 2017, Stewart-Morgan et al. 2020, Willhoft and Costa 2021).

During late mitosis and throughout the G₁ phase of the cell cycle, two MCM helicases are loaded onto origin DNA, forming a double hexamer (DH) with their amino-terminal tiers facing each other ([Figure 1.1](#)) (Evrin et al. 2009, Remus et al. 2009, Li et al. 2015, Abid Ali et al. 2017, Noguchi et al. 2017). The sequential loading of the two helicases, termed origin licensing, is coordinated by the origin recognition complex (ORC), the loading factors cell division cycle 6 (Cdc6) and Cdt1 (Fernández-Cid et al. 2013, Coster et al. 2014, Ticau et al. 2015, Coster and Diffley 2017, Frigola et al. 2017, Yuan et al. 2017, Miller et al. 2019, Yuan et al. 2020b). This process depends on the ATPase activity of the MCM (Coster et al. 2014, Kang et al. 2014). Once loaded onto duplex DNA, the helicase is found in an inactive state and does not unwind DNA, at least according to studies with *S. cerevisiae* proteins. Upon transition into the synthetic (S) phase of the cell cycle, the DH is phosphorylated by DDK, which triggers the CDK-dependent recruitment of the helicase activators Cdc45 and go-ichi-nii-san (GINS) (Sheu and Stillman 2006, Tanaka et al. 2007, Zegerman and Diffley 2007, Francis et al. 2009, Labib 2010, Muramatsu et al. 2010, Heller et al. 2011, Deegan et al. 2016). The activators stably bind to the MCM and two Cdc45-MCM-GINS (CMG) replicative helicases are generated (Moyer et al. 2006, Ilves et al. 2010, Abid Ali et al. 2016, Yuan et al. 2016, Douglas et al. 2018). CMG formation results in an initial

untwisting and initial opening of the DNA (Douglas et al. 2018, Lewis et al. 2022), which is followed by the ejection of the lagging strand DNA template from the MCM central channel upon recruitment of the Mcm10 firing factor and activation of the MCM ATPase activity (Löoke et al. 2017). While recruitment of the leading strand polymerase Pol ϵ is essential for CMG formation (Sengupta et al. 2013, Goswami et al. 2018), the Pol α /primase complex and lagging strand polymerase Pol δ are recruited at a later stage, to form two diverging replication forks, in which the CMG translocates on the leading strand.

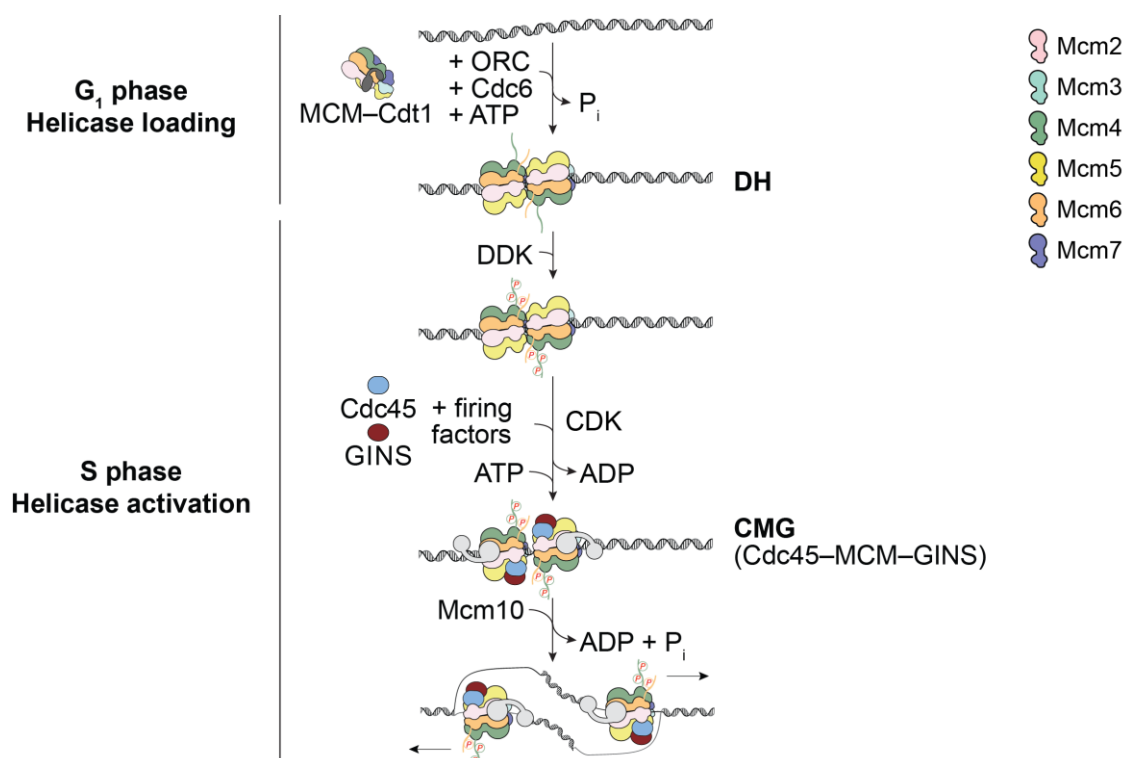


Figure 1.1: Temporal separation of helicase loading and activation during the cell cycle.

The MCM helicase is loaded onto double-stranded DNA during G₁ phase of the cell cycle. During S phase, DDK and CDK stimulate the recruitment of the helicase activators Cdc45 and GINS. Two CMG helicases are formed and the firing factor Mcm10 stimulates helicase activity and lagging strand ejection.

In this introduction, I will summarise what is known about the mechanism of MCM helicase loading and activation, obtained from studies primarily carried out in the model organism *S. cerevisiae*. I will describe the role of DDK during helicase activation and how it is regulated.

1.2 Structure and function of the Mcm2–7 helicase

1.2.1 Structure of the Mcm2–7 helicase

The Mcm2–7 helicase belongs to the AAA+ superfamily of ATPases (ATPases Associated with various cellular Activities). The members of this protein family form hexameric, ring-shaped assemblies, which use hydrolysis of nucleoside triphosphates (NTPs) for perform work, including nucleic acid and polypeptide translocation (Singleton et al. 2007, Puchades et al. 2020). All known AAA+ DNA helicases translocate with 3' to 5' direction along DNA (Singleton et al. 2007, Enemark and Joshua-Tor 2008). The six MCM subunits form a two-tiered ring in the order Mcm2–5–3–7–4–6 (Davey et al. 2003, Costa et al. 2011, Abid Ali et al. 2016) ([Figure 1.2a](#)). The two tiers are established by the N-terminal oligomerisation domains (NTD) and the AAA+ motor domains of the six subunits. In *S. cerevisiae*, all subunits apart from Mcm2 contain a C-terminal winged helix domain (WHD) appendix.

The WHDs are flexibly tethered to the ATPase domains. While WHDs are well-known for their potential to bind nucleic acids, the main role of the MCM WHDs during DNA replication is the establishment of protein–protein interactions (Brewster et al. 2008, Li et al. 2015). For example, the WHDs of Mcm3, Mcm4, Mcm6 and Mcm7 bind to the loader complex ORC–Cdc6 during helicase recruitment (discussed in section 1.3.2) (Frigola et al. 2013, Yuan et al. 2017, Yuan et al. 2020b). The Mcm6 WHD not only interacts with ORC, but also with the loading factor Cdt1 (Wei et al. , Fernández-Cid et al. 2013, Yuan et al. 2020b), and later becomes a contact point for the replication factors Mcm10 (Douglas and Diffley 2016) and Mrc1 (Komata et al. 2009), which increase replication fork speed (Lewis et al. 2017, Yeeles et al. 2017). In the active replisome, the Mcm5 WHD interacts with the leading strand polymerase Pol ϵ (Goswami et al. 2018). Together these results suggest that the WHDs play an important structural role in helicase loading as well as formation of active replication forks.

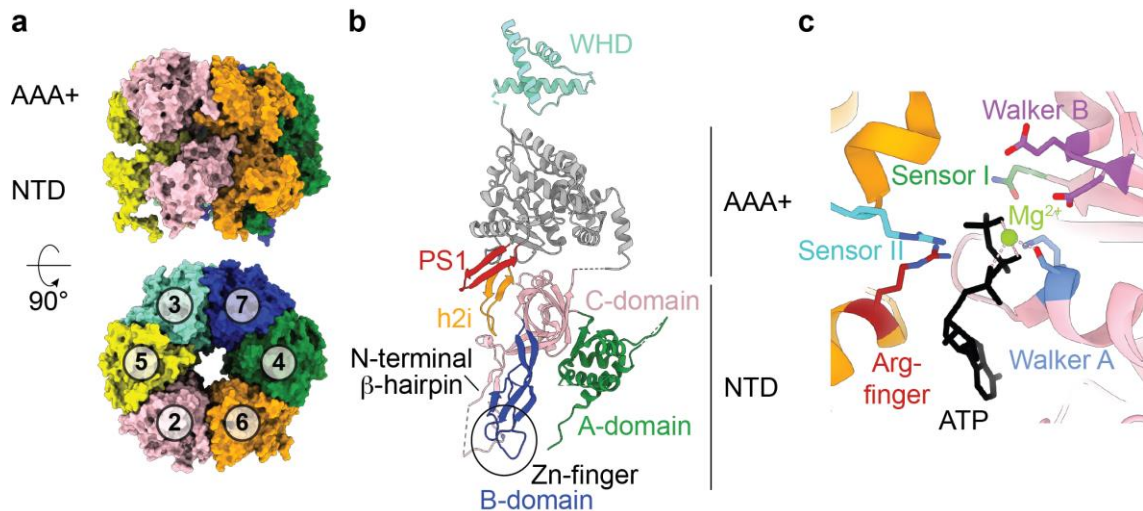


Figure 1.2: Structure of the Mcm2–7 hexamer.

a. Surface representation of the Mcm2–7 hexamer atomic model (PDB 7P30). **b.** Atomic model of the Mcm6 subunit illustrating the domain organisation and important structural elements. **c.** Atomic model of the ATPase site between Mcm2 (pink) and Mcm6 (orange) in the DH (PDB 7P30). Active site residues are highlighted.

The NTD is subdivided into three domains, called A, B and C domain (Fletcher et al. 2003) (Figure 1.2b). The A domain contains a helical bundle with four α -helices and is located on the outer perimeter of the hexameric MCM ring. This domain is thereby available for interactions with helicase activators (Costa et al. 2011, Costa et al. 2014, Abid Ali et al. 2016). The A domain can also interact with DNA (Costa et al. 2008). The role of DNA binding outside of the central MCM channel is, however, unclear. Mcm2, Mcm4 and Mcm6 have long N-terminal tails protruding from the A domain, which are important during helicase activation (Sheu and Stillman 2010, Deegan et al. 2016) as well as recycling of parental histones (Huang et al. 2015, Petryk et al. 2018). The B domain contains a Zinc (Zn) finger motif and resides at the N-terminal face of the MCM ring. During the final stages of helicase loading, the Zn finger domains of one ring become interdigitated with the Zn finger domains of the second ring to form the MCM DH (Li et al. 2015, Abid Ali et al. 2017, Noguchi et al. 2017). The importance of the Zn finger domain is highlighted by the impaired ATPase activity and DNA binding upon mutation of the Zn finger in the archaeal MCM (Poplawski et al. 2001). The third domain within the NTD, the C domain, has an OB (oligonucleotide/oligosaccharide binding)-fold formed of a β -barrel. It is not only important for oligomerisation of the MCM (Fletcher et al. 2003, Shima et al. 2007), but also for DNA binding and helicase function *via* a β -hairpin

that is projected towards the central MCM channel (McGeoch et al. 2005, Froelich et al. 2014, Abid Ali et al. 2017, Noguchi et al. 2017). This hairpin is also known as the N-terminal β -hairpin and has been implicated in splitting leading and lagging strand at the fork nexus (Baretić et al. 2020).

The AAA+ motor of the helicase, is characterised by the additional strand catalytic glutamate (ASCE) fold, which consists of a five-stranded, parallel β -sheet interspersed with several α -helices (Lyubimov et al. 2011). Elements within the ASCE fold of two adjacent MCM subunits form the ATPase site. One subunit provides the conserved Walker A and Walker B sequence motifs, which bind the phosphate groups of a nucleotide and the essential Mg^{2+} ion in the active site (Erzberger and Berger, 2006; Iyer et al., 2004; Moreau et al., 2007) (Figure 1.2c). A polar residue in proximity to the Walker A and Walker B, called sensor I, coordinates a water molecule for catalysis. These elements are joined by sensor II and an arginine finger of the adjacent MCM subunit to allow ATP hydrolysis. Therefore, the nucleotide is coordinated by Walker A, Walker B and sensor I in *cis* and the arginine finger and sensor II in *trans*. Nucleotide binding and hydrolysis are coupled to conformational changes in the MCM and transferred to two β -hairpin loops, which are inserted in the ASCE fold (Erzberger et al. 2002). The so-called pre-sensor 1 (PS1) loop resides between α -helix 3 and β -strand 4 and the helix-2-insertion (h2i) within α -helix 2. These loops are projected into the central MCM channel, where they interact with the DNA and drive helicase translocation. Mutation of these loops has significant impact on the function of the MCM. As such, mutation of the Mcm3 PS1 is lethal in *S. cerevisiae* (Lam et al. 2013, Ramey and Sclafani 2014). Mutations in the Mcm4 and Mcm5 PS1 loops result in retarded growth, while mutation of the other PS1 loops have less severe effects. Deletion of h2i loops increases DNA binding and ATPase activity, but impairs DNA unwinding (Jenkinson and Chong 2006). This finding implies that the h2i loops enable helicase translocation by destabilising other protein–DNA interactions.

1.2.2 ATP hydrolysis drives translocation

ATP hydrolysis induces conformational changes in the PS1 and h2i loops and, as such, drives helicase translocation (Jenkinson and Chong 2006, Miller et al. 2014, Eickhoff et al. 2019, Meagher et al. 2022). Studies using related AAA+ ATPases, including the papillomavirus helicase E1 (Enemark and Joshua-Tor 2006), *Saccharolobus solfataricus* MCM (Meagher et al. 2022) and *Drosophila melanogaster* CMG (Eickhoff et al. 2019, Rzechorzek et al. 2020) have shown that the pore loops are arranged in a staircase configuration, following the helical pitch of the DNA (Figure 1.3a–b). The PS1 loops are stacked on top of the h2i loops (Miller et al. 2014, Abid Ali et al. 2017). DNA-engaged subunits are usually bound to ATP, while disengaged subunits at the bottom of the staircase are devoid of a nucleotide or bound to ADP (Enemark and Joshua-Tor 2006, Goswami et al. 2018, Eickhoff et al. 2019, Rzechorzek et al. 2020, Yuan et al. 2020a) (Figure 1.3c). According to the proposed structural model for translocation, ATP hydrolysis occurs at the top of the staircase and is followed by nucleotide release and DNA disengagement. Disengaged subunits can then join the bottom of the staircase upon binding of a new ATP molecule, ready for another round of hydrolysis. ATP hydrolysis in the neighbouring subunit might also be stimulated by the conformational changes based on the nucleotide binding state, resulting in sequential catalytic activity (Lyubimov et al. 2011). These studies suggest a sequential rotary translocation mechanism, in which the DNA is pulled upwards (towards the C-terminal tier) as ATP is hydrolysed in the ATPase sites around the ring, leading to an advancement of two nucleotides per helicase subunit.

However, ATP hydrolysis by the eukaryotic MCM is not symmetric and the different ATPase sites do not contribute equally to helicase translocation (Ilves et al. 2010, Eickhoff et al. 2019). Structural and functional analysis shows that ATPase activity of Mcm3–7 is essential for helicase translocation, whereas ATP binding, but not hydrolysis at the Mcm5–3 interface is required. Furthermore, ATPase activity by Mcm4–6 seems to be dispensable. In summary, MCM helicase translocation is driven by asymmetric, likely sequential ATPase activity, which stimulates DNA engagement of the PS1 and h2i pore loops around the helicase ring.

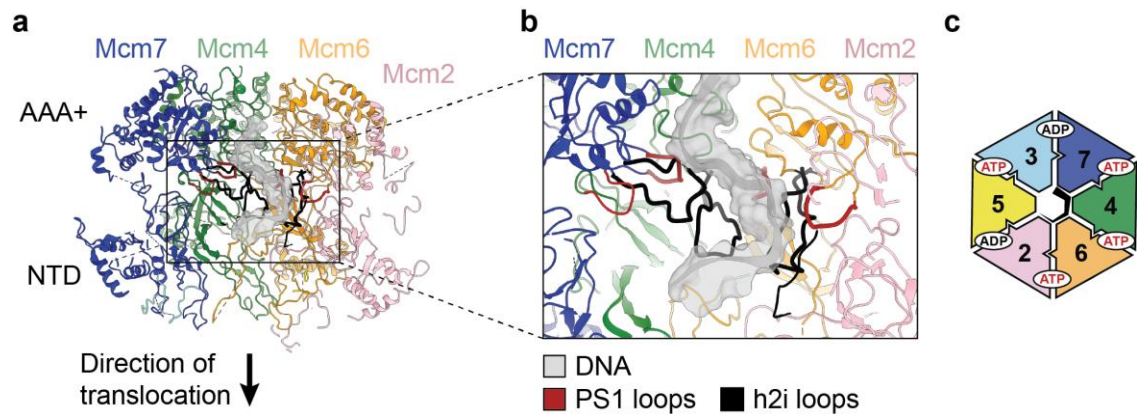


Figure 1.3: Pore loop staircase in the *D. melanogaster* CMG.

a. Cut-through view of the *D. melanogaster* CMG bound to a forked-DNA substrate (state 2A; PDB 6RAY). **b.** Zoom-in view of the PS1 and h2i pore loops that form a staircase, following the DNA helix. **c.** Cartoon representation of the nucleotide binding state of *D. melanogaster* CMG state 2A. DNA interactions with the PS1 loops of the different MCM subunits are indicated by a black line.

1.3 Origin licensing

1.3.1 Origin recognition by ORC

DNA replication is initiated from origins, which, in *S. cerevisiae*, are specific DNA sequences, known as autonomous replicating sequences (ARS) (Marahrens and Stillman 1992). *S. cerevisiae* origins contain the AT-rich ARS consensus sequences (ACS) and three additional, less conserved elements known as B1, B2 and B3 (Figure 1.4a). The helicase loader ORC binds simultaneously to the ACS and the B1 element, causing the DNA to bend, which becomes important during helicase loading (Rowley et al. 1995, Li et al. 2018). The A-rich B2 element represents an inverted ORC binding site and contributes to MCM loading (Wilmes and Bell 2002, Miller et al. 2019). The B3 element is a binding site for the Abf1 (ARS-binding factor 1) transcription factor, which influences transcription as well as replication (Diffley and Stillman 1988). Another characteristic of origins is that they usually coincide with nucleosome free regions (Eaton et al. 2010). In higher eukaryotes, which lack specific DNA sequences for ORC binding, these open chromatin structures are essential to define origins of replication (Remus et al. 2004, MacAlpine et al. 2010).

ORC is a hetero-hexamers formed of five AAA+ ATPases (Orc1–5) and an additional subunit, called Orc6, which has a transcription factor IIB (TFIIB) like fold (Liu et al. 2011, Tocilj et al. 2017, Li et al. 2018, Jaremko et al. 2020, Schmidt and Bleichert 2020) (Figure 1.4b). Each of the Orc1–5 subunits contain a N-terminal AAA+ ATPase domain, but only Orc1 and Orc4 show ATP-hydrolysis activity, and a C-terminal WHD (Coster et al. 2014, Li et al. 2018). Together, the five subunits form a crescent shape with DNA binding capacity in the central cleft (Li et al. 2018, Jaremko et al. 2020, Schmidt and Bleichert 2020). The ATPase domain of Orc1 and the Orc2 WHD form a dynamic gate, controlling DNA binding (Bleichert et al. 2018, Li et al. 2018, Jaremko et al. 2020, Schmidt and Bleichert 2020, Yuan et al. 2020b). Furthermore, Orc1 contains an N-terminal bromo-adjacent homology (BAH) domain, which facilitates ORC recruitment to the acidic patch of nucleosomes, as shown for *S. cerevisiae* Orc1 (Müller et al. 2010, De Ioannes et al. 2019), and to nucleosomes containing histone H4 tails, which are dimethylated at lysine 20 (H4K20me2), in *H. sapiens* (Kuo et al. 2012). Orc1 might therefore play an integral role in origin licensing in higher eukaryotes where sequence specificity is lacking.

Orc6 is the least conserved subunit within the ORC and its function in different species is debated, but several studies suggest its involvement in DNA replication (Prasanth et al. 2002, Thomae et al. 2011, Liu et al. 2012). *S. cerevisiae* and *D. melanogaster* Orc6 stably associates with Orc1–5, establishing contact with Orc3 via a C-terminal α helix (Bleichert et al. 2018, Li et al. 2018) (Figure 1.4b). The centrally located TFIIB-like domain within Orc6 interacts with DNA and contributes to the binding of ORC to the B1 element of origins (Li et al. 2018). Additionally, Orc6 is essential for the formation of a loading intermediate during helicase loading in *S. cerevisiae* (discussed in section 1.3.2) (Miller et al. 2019). In *H. sapiens*, ORC6 is weakly associated and its role is less clear (Dhar and Dutta 2000, Dhar et al. 2001, Vashee et al. 2001). Nevertheless, depletion of ORC6 results in origin licensing and DNA replication defects in *H. sapiens* cell lines (Prasanth et al. 2002, Stiff et al. 2013).

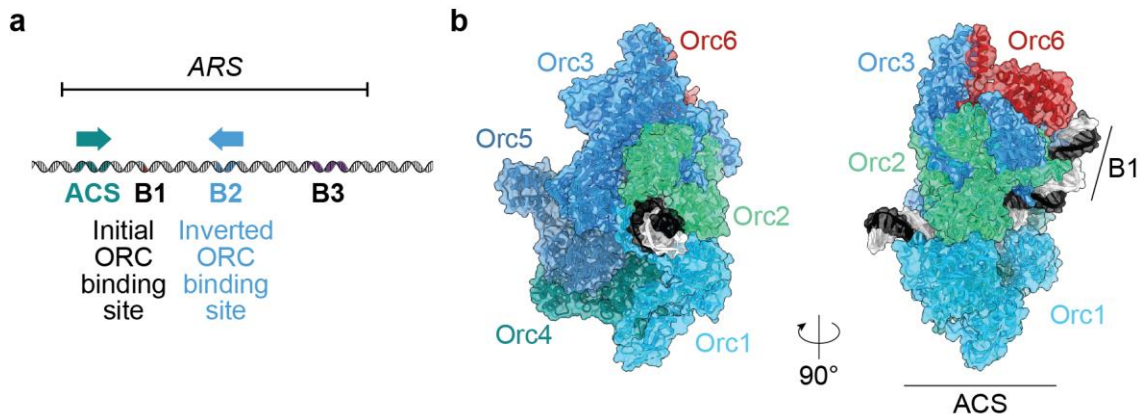


Figure 1.4: Origin recognition by *S. cerevisiae* ORC.

a. Schematic of a typical yeast origin of replication. The primary binding site for ORC is composed of the ACS and the B1 element, while the B2 element serves as a secondary binding site with inverted orientation. **b.** Atomic structure of the *S. cerevisiae* Orc1–6 complex bound to origin DNA (PDB code 5ZR1) (Li et al. 2018).

In the presence of ATP, ORC associates with origins of replication. ATP is bound to Orc1, Orc4 and Orc5, while Orc2 and Orc3 appear to be unable to recruit a nucleotide, coherent with their divergent ATPase centre (Bleichert et al. 2015). Sequence specific binding of *S. cerevisiae* ORC towards the ACS occurs due to a species-specific α -helix in Orc4 that accesses the DNA major groove, establishing direct contacts with the thymidine bases (Li et al. 2018, Hu et al. 2020). *H. sapiens* ORC4 lacks this helix insertion and only interacts with the phosphate backbone of the DNA, explaining why origin recognition does not rely on direct sequence readout in humans (Jaremko et al. 2020). In addition to the interaction with the ACS, Orc2, Orc5 and Orc6 contact the B1 element, which results in a bend in the DNA of about 76° (Li et al. 2018). This brings the DNA into an ideal position for the loading of the helicase (Yuan et al. 2017, Yuan et al. 2020b). The loading factor Cdc6, another AAA+ ATPase, nestles in between Orc1 and Orc2 and completes the ring around the DNA.

In metazoans, ORC1 and CDC6, as well as the loading factor CDT1 harbour intrinsically disordered regions (IDRs) (Parker et al. 2019, Hossain et al. 2021). These domains caused liquid condensates when *D. melanogaster* ORC was incubated with CDC6 and DNA *in vitro*. However, the significance of this effect *in vivo* remains speculative. The ORC1 and CDC6 IDRs also contain short sequence motifs that engage in cyclin-dependent protein–protein interactions and interactions

with the protein phosphatase 1 (PP1), controlling ORC1 protein levels throughout the cell cycle.

1.3.2 Origin licensing

Two copies of the hexameric MCM helicase are recruited in a stepwise manner to form a DH and allow for subsequent bidirectional replication. The Mcm2–7 hexamer is initially held in an open conformation by the loading factor Cdt1, with a gate between the Mcm2 and Mcm5 subunits (Tanaka and Diffley 2002, Frigola et al. 2017, Zhai et al. 2017). Interestingly, *H. sapiens* MCM does not require CDT1 to adopt this conformation, but still depends on its presence for recruitment to DNA (Xu et al. 2022). Instrumental in tethering Mcm2–7–Cdt1 to the DNA-bound ORC–Cdc6 loader complex are the MCM WHDs (Yuan et al. 2017, Yuan et al. 2020b). The initial contact is established between the Mcm3 WHD and Orc2–Cdc6 as well as the Mcm7 WHD and Orc1–Cdc6. Prior to DNA insertion into the MCM channel, the Orc1 WHD binds to the Mcm4 AAA+ domain and stabilises the complex. The DNA, which is held in a bent configuration, is positioned next to the Mcm2–5 gate, poised for threading into the MCM ring (Figure 1.5). The B1 element subsequently detaches from ORC and the DNA enters the MCM central channel. In the resulting ORC–Cdc6–Cdt1–MCM complex, also known as OCCM, additional contacts between the Mcm2–7–Cdt1 and ORC–Cdc6 are established. This includes interactions between the Mcm4 WHD and Orc1, the Mcm6 WHD and Orc4 and the Mcm2 AAA+ domain and Orc5. OCCM formation requires ATP binding, but not hydrolysis (Frigola et al. 2013, Coster et al. 2014, Yuan et al. 2017, Yuan et al. 2020b). Upon ATP hydrolysis by the MCM, the loading factors are released and the Mcm2–5 gate closes, so that the DNA becomes topologically entrapped by the MCM ring (Coster et al. 2014, Kang et al. 2014, Tica et al. 2015). While ORC and Cdc6 also contain ATPase function, their catalytic activity is not required for MCM loading (Chang et al., 2015; Kang et al., 2014; Randell et al., 2006; Speck and Stillman, 2007; Yuan et al., 2017). Instead, ATP hydrolysis by Cdc6 has been shown to release unproductive loading intermediates (Frigola et al. 2013, Coster et al. 2014).

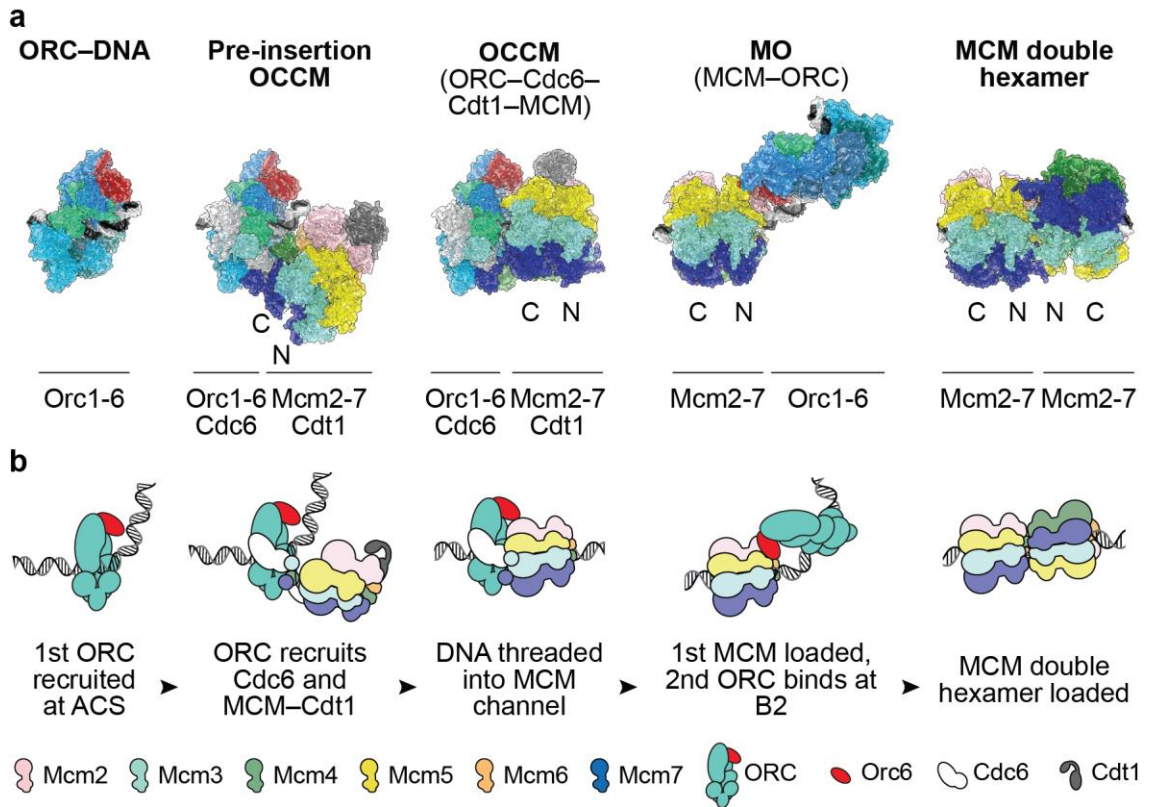


Figure 1.5: Origin licensing in *S. cerevisiae*.

a. Structures of intermediates leading to the formation of MCM double hexamers observed with *S. cerevisiae* proteins, which include ORC bound to origin DNA (PDB 5ZR1), the pre-insertion OCCM (PDB 6WGG), OCCM (PDB 5V8F), MO (PDB 6RQC), and the MCM double hexamer (PDB 6F0L). The N- and C-terminal tiers of the MCM are indicated. **b.** Cartoon representation of the origin licensing reaction.

Analysing the origin licensing reaction with *S. cerevisiae* proteins by cryo-EM, we have previously shown that ORC dissociates from the C-terminal side of the DNA-loaded MCM hexamer and that a second ORC binding event occurs on the N-terminal side, forming the MO complex (Miller et al. 2019) (Figure 1.5). This assembly coordinates the recruitment of a second Mcm2–7–Cdt1 helicase and promotes the correct orientation of the two helicases. In this configuration, ORC binds to the B2 element of the origin in an inverted orientation as compared to the first binding event at the ACS. Biochemical data has shown that this inverted configuration of the two ORC binding sites is important for MCM loading (Coster and Diffley 2017). In the MO, Orc6 mediates the interaction between Orc1–5 and the N-terminal domains of Mcm2 and Mcm5 at an approximate 90° angle. This configuration involves a degree of DNA bending, which resembles the one observed during loading of the first MCM. ORC then recruits Cdc6 and a second

Mcm2–7–Cdt1 in a similar way as observed during OCCM formation. As a result, the two MCMs face their N-terminal domains, poised to form a head-to-head double hexamer (Figure 1.5). The structural analysis of this process has shed light into the long-standing question whether a single ORC is sufficient for MCM double hexamer formation or whether two inverted ORC binding events are required (Coster and Diffley 2017, Tica et al. 2017). While the same Orc1–6 could sequentially load both MCM hexamers, ORC engagement occurs at two inverted binding sites.

Formation of the MCM DH, also known as the pre-replicative complex (pre-RC), marks the end of the origin licensing reaction (Diffley et al. 1994). The MCM rings dimerise with their Zn fingers interlocked and the Mcm2–5 gates residing at opposite sides of the DH (Li et al. 2015, Abid Ali et al. 2017, Noguchi et al. 2017). The two hexamers are stacked with a slight offset causing a 10° bend in the duplex DNA that is bound in the central channel. At the given resolution of previous cryo-EM structures, no DNA melting was observed in the *S. cerevisiae* DH. Besides the interaction of the Zn finger domains, an α -helix in the Mcm7 A domain is found in contact with the Mcm5 A domain in the opposing MCM ring. This appears to provide an additional stabilising element for the *S. cerevisiae* DH. Interestingly, this helix is significantly shorter in other species, and it remains to be investigated how DH stability and configuration is influenced by this structural feature.

The *S. cerevisiae* DH is an inactive form of the helicase as it does not unwind DNA (Li et al. 2015, Abid Ali et al. 2017, Noguchi et al. 2017, Douglas et al. 2018). Despite catalytic inactivity, the DH can passively slide on DNA (Remus et al. 2009, Gros et al. 2015, Miller et al. 2019). In accordance with biochemical data (Coster et al. 2014), cryo-EM structures indicate that the DH harbours ADP at several ATPase sites, meaning that catalytic activity is required during helicase loading. Assignment of the nucleotide binding state is, however, ambiguous due to the limited resolution of available structures. Functional characterisation revealed – similar to the effects observed for helicase translocation – that the different ATPase sites contribute to helicase loading to different extents (Coster et al. 2014). As such, it has been established that ATP binding at all sites is required for the loading process (Coster et al. 2014). On the contrary, arginine finger mutants of all subunits apart from

Mcm4 were defective in MCM loading. This suggests that ATP hydrolysis needs to occur at all ATPase sites, except for Mcm4–6. It needs to be established at which point the different sites hydrolyse ATP and whether this involves ATP-driven translocation along the DNA. Solving the DH structure as well as structures of loading intermediates at high resolution can provide significant conceptual advancement in this context.

1.4 Origin firing and chromatin replication

The activation of the MCM helicase is highly regulated and involves the activity of two kinases, DDK and CDK (Labib 2010). Stimulated by phosphorylation, multiple firing factors transiently interact with the MCM to recruit the helicase activators Cdc45 and GINS and the leading strand polymerase Pol ϵ . Eventually both MCM hexamers are activated and bidirectional replication is established (Yeeles et al. 2015) (Figure 1.6).

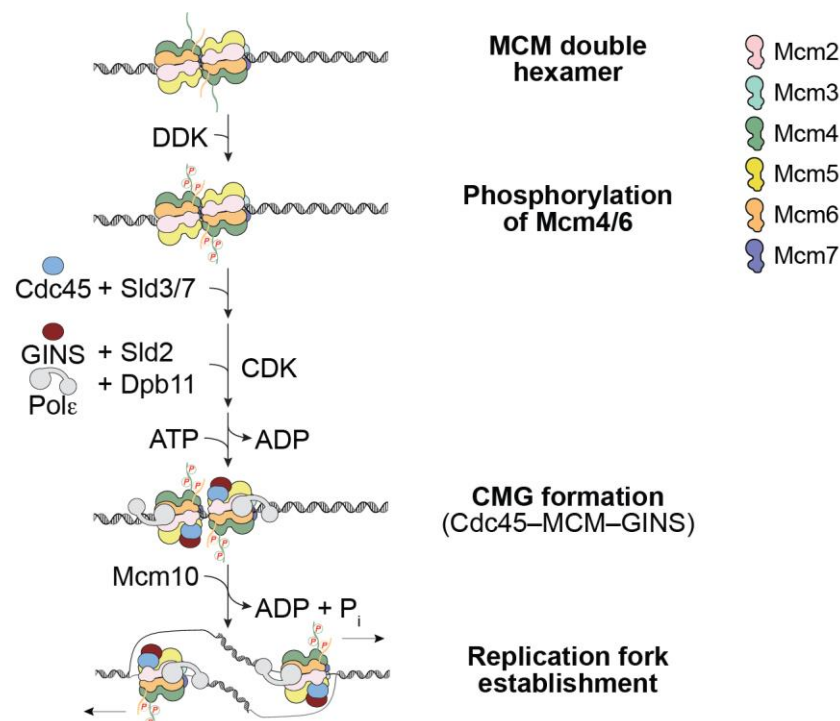


Figure 1.6: Origin firing.

Cartoon representing the activation and assembly of the CMG helicase, stimulated by DDK- and CDK-dependent phosphorylation.

1.4.1 Phosphorylation of the MCM double hexamer by DDK

DDK initiates the origin activation process by selectively phosphorylating the DNA-loaded DH (Masai et al. 2006, Sheu and Stillman 2006, Francis et al. 2009, Sheu and Stillman 2010). DDK targets multiple serine and threonine residues in the flexible N-terminal tails of Mcm4 and Mcm6, which measure over 100 amino acids in length (Labib 2010, Randell et al. 2010). MCM, in particular Mcm4, phosphorylation is essential for the recruitment of the helicase activator Cdc45 by Sld3–7 (Sheu and Stillman 2006, Deegan et al. 2016).

The mechanism for how DDK recognises the DNA-loaded DH remains unclear. Nevertheless, biochemical and genetic experiments have provided valuable insights. The N-terminal BRCT domain of the kinase activator Dbf4 interacts with Mcm2 (Francis et al. 2009, Ramer et al. 2013, Abd Wahab and Remus 2020), while a C-terminal element of Dbf4 contacts Mcm4, where phosphorylation occurs (Jones et al. 2010). Mcm2 and Mcm4 are non-adjacent subunits in the MCM, leading to the question of how DDK establishes contact with both subunits to phosphorylate the DH. Structural characterisation of the kinase in complex with its substrate will shed light onto this issue.

Another outstanding question is how DDK activates the MCM. It has been suggested that a conformational change in the MCM is responsible for activation. This is based on the *bypass of block 1 (bob1)* mutation, a point mutation of proline 87 to leucine in Mcm5, which is viable in the absence of DDK (Hardy et al. 1997). The *bob1* mutation causes a small conformational change in the A domain, which could allow for Sld3–7 and Cdc45 recruitment (Fletcher et al. 2003). Another suggestion is based on the finding that the most proximal residues of the Mcm4 N-terminal tail (residues 74–174) harbour an inhibitory function, which is relieved upon phosphorylation by DDK (Sheu and Stillman 2010). This is functionally distinct from the *bob1* mutation and highlights the fact that the exact function of DDK during origin activation is still poorly understood. Furthermore, recent studies did not reveal major conformational changes in the phosphorylated DH, apart from an apparent stabilisation of the N-terminal tails, creating a potential landing platform for the recruitment of firing factors (On et al. 2014, Abid Ali et al. 2017).

1.4.2 Recruitment of the helicase activators Cdc45 and GINS

The hetero-dimeric Sld3–7 recognises the phosphorylated DH and recruits Cdc45 (Kanemaki and Labib 2006, Deegan et al. 2016). Crystal structures of the Sld3–7 complex indicate that the C-terminal domains of two Sld7 molecules interact (at least based on the crystal packing) so that a homodimer of two Sld3–7 complexes could be formed (Itou et al. 2015). This would allow for the simultaneous recruitment of two Cdc45 molecules to the DH.

During S phase, high levels of CDK result in the phosphorylation of the firing factors Sld2 and Sld3, which in turn recruit the second helicase activator GINS. Upon phosphorylation, Sld2 associates with a BRCA1 C-terminus (BRCT) tandem repeat of Dpb11, which in turn binds GINS and the leading strand polymerase Pol ϵ . The Sld2–Dpb11–GINS–Pol ϵ is known as pre-loading complex (Muramatsu et al. 2010). A second BRCT tandem repeat in Dpb11 brings the pre-loading complex to the phosphorylated Sld3 and therefore to the DH. Upon release of ADP and binding of ATP, GINS and Cdc45 become stably associated with the MCM and two CMG helicases bound to Pol ϵ (CMGE) are formed (Douglas et al. 2018). Cdc45 and GINS bind to the Mcm2, Mcm3 and Mcm5 subunits, securing the gate through which the DNA has entered the helicase ring (Costa et al. 2011). A recent cryo-EM structure shows that the two helicases form a dimeric complex, which was termed the double CMGE (dCMGE) (Lewis et al. 2022) (Figure 1.7a). The subunit register of the two CMGE complexes is shifted by one subunit compared to the DH and the assembly shows a splayed configuration, exposing 1.5 turns of duplex DNA between the two CMGEs. Not only the interlocking of the MCM Zn finger domains is disrupted, but also the extended Mcm7 helix is disengaged from Mcm5 in the dCMGE. Formation of this complex leads to an initial untwisting of the DNA by 0.7 turns, which is associated with nucleation of DNA melting within the ATPase core of each CMG complex (Figure 1.7b). The Mcm2 h2i and PS1 loops together with a subunit-specific insertion in the N-terminal β -hairpin of Mcm6 promote and stabilise the orphan bases. An outstanding question is whether DNA untwisting is initiated at the interface of the two MCM rings upon separation and whether it is subsequently propagated into each ring in the dCMGE.

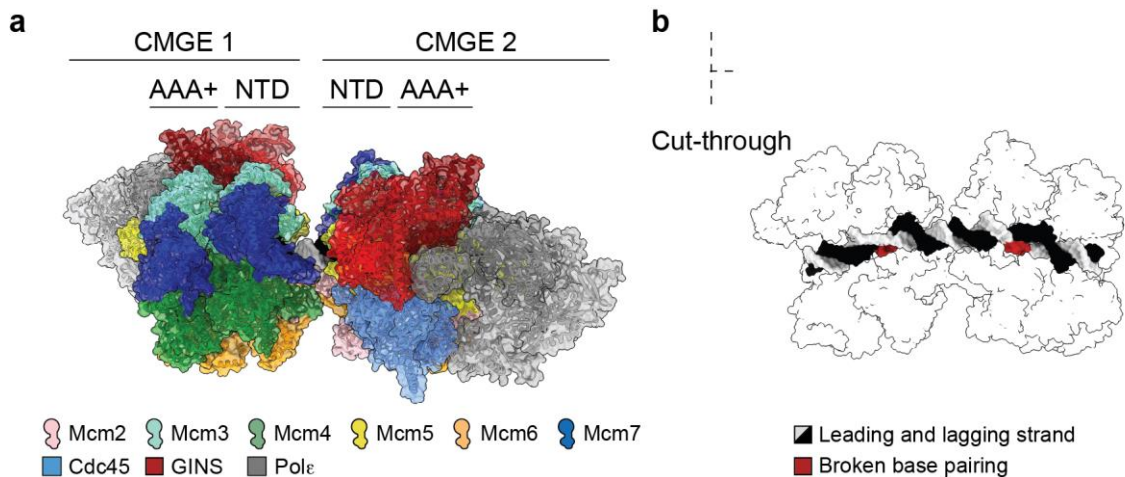


Figure 1.7: CMG formation is concomitant with disruption of the MCM dimerization interface and nucleation of DNA melting.

a. Atomic structure of the dCMGE complex (PDB 7Z13). **b.** Cut-through view of the dCMGE complex, highlighting the untwisted and melted DNA in the central channel.

1.4.3 Replication fork establishment

ATPase activity of the CMG is activated when the firing factor Mcm10 is recruited (Douglas et al. 2018). The recruitment leads to the ejection of the lagging strand and translocation along single-stranded DNA. As the CMG helicases advance along DNA with the N-terminal MCM tier at the front (Georgescu et al. 2017, Eickhoff et al. 2019), the two helicases cross their paths during this process. The symmetry of the head-to-head DH therefore ensures that two helicases are simultaneously activated to establish bidirectional replication.

The CMGE complex is at the centre of the replisome and serves as a platform for the recruitment of other replisome components (reviewed in Bell and Labib (2016)). These components include the trimeric scaffolding protein Ctf4, which docks onto GINS in the CMGE and tethers the Pol α -primase to the CMGE (Gambus et al. 2006, Gambus et al. 2009, Simon et al. 2014). Pol α -primase synthesises a DNA–RNA primer that is then extended by the leading and lagging strand polymerases Pol ϵ and Pol δ , respectively. The single-stranded DNA-binding protein replication protein A (RPA) protects the unwound DNA until it is replicated. Replication speed is stimulated in the presence of the polymerase-binding sliding clamp PCNA (proliferating cell nuclear antigen), the clamp loader RFC (replication factor C) and

the fork protection complex, consisting of Mrc1, Csm3 and Tof1 (Kurat et al. 2017, Lewis et al. 2017, Yeeles et al. 2017). Factors implicated in replisome interactions are also topoisomerase I, which removes DNA supercoils at the fork, and the histone chaperone FACT, which is involved in parental histone recycling.

1.5 Regulation of origin licensing during the cell cycle

Eukaryotic DNA replication is tightly regulated to ensure that the entire genome is replicated once and only once per cell cycle. Periodicity is driven by expression levels of cyclins, concomitant changes in CDK activity as well as the activity of the anaphase promoting complex or cyclosome (APC/C) (Kapuy et al. 2009, Costa and Diffley 2022). In G₁ phase, absence of S phase cyclins and presence of the CDK inhibitor Sic1 prevent kinase activity to allow loading of the MCM helicase. When cells reach a critical size, Sic1 is phosphorylated and degraded (Kapuy et al. 2009). Rising CDK activity during S phase then stimulates helicase activation and prevents further MCM loading in multiple ways. First, the loading factor Cdc6 is phosphorylated, resulting in ubiquitin-dependent protein degradation (Drury et al. 2000). Second, phosphorylation of Orc2 and Orc6 as well as binding of the S phase kinase Clb5–CDK (in *S. cerevisiae*) to a Arg–X–Leu (RXL) motif in Orc6 directly interfere with helicase loading (Wilmes et al. 2004, Chen and Bell 2011, Frigola et al. 2013). Third, phosphorylation of Mcm2 and Mcm3 results in nuclear export of MCM helicases that are not DNA-bound (Labib et al. 1999, Nguyen et al. 2000). Metazoans have developed additional mechanisms to prevent origin licensing outside of S phase. As such, not only CDC6, but also ORC1 and CDT1 are subject to ubiquitin-mediated degradation (Méndez et al. 2002, Arias and Walter 2005, Takeda et al. 2005, Nishitani et al. 2006, Zhang 2021). Additionally, phosphorylation of *H. sapiens* CDC6 by cyclin E–CDK2 stimulates its export from the nucleus (Jiang et al. 1999). Moreover, a protein called geminin binds to the loading factor CDT1, thereby preventing the MCM–CDT1 interaction and MCM loading (Wohlschlegel et al. 2000). To allow licensing at the start of the cell cycle, geminin is targeted for degradation by the APC/C during M and G₁ phase (McGarry

and Kirschner 1998). Together, these processes prevent re-replication and therefore ensure genome integrity.

Availability of nutrients and cell growth also influence the timing of DNA replication. Cells are thereby kept in G₁ phase until a critical cell size is reached (Jorgensen and Tyers 2004, Kapuy et al. 2009) and growth factors stimulate the transition into S phase (Farshadi et al. 2020). When nutrients are readily available, ribosomal components are expressed, leading to increased protein production, cell growth and cell cycle progression. For example, the metazoan retinoblastoma protein (Rb) binds to chromatin and prevents binding of the transcription factor E2F under restrictive conditions (Pack et al. 2019). Optimal growth conditions during G₁ phase stimulate association of cyclin D with CDK4 and CDK6, which promotes Rb phosphorylation and dissociation from chromatin. E2F then enables transcription of cyclin E, which together with CDK2 further phosphorylates Rb and drives cell cycle progression. Under restrictive conditions and upon detection of DNA damage, kinase inhibitors like p21, p27, p53 and p57 are upregulated to counteract the stimulatory effect of cyclins. It is the balance between stimulatory and inhibitory signals that ultimately defines whether and when a cell progresses to the next stage of the cell cycle.

1.5.1 Importance of faithful DNA replication initiation for human health

Mutations in initiation factors, like ORC1, ORC4, ORC6, CDT1, CDC6, MCM5, MCM3, MCM7, geminin, CDC45 and GINS2, are associated with the rare genetic disorder called Meier–Gorlin syndrome (Bellelli and Boulton 2021, Schmit and Bielinsky 2021, Nazarenko et al. 2022). Patients with this disease suffer from primordial dwarfism, aplastic or hypoplastic patellae and small ears. Consistently, cells harbouring pathological variants show reduced cell proliferation. Zebrafish models of the disease indicated that the reduced capacity of origin licensing results in delayed S phase progression and apoptosis (Yao et al. 2017). The corresponding proteins are rendered non-functional by a variety of gene alterations, resulting in frameshifts, missense and splicing errors. Particularly variants of ORC1 and ORC4 have severe phenotypes. Although the role of ORC6 in DNA replication

has been debated, mutations of ORC6 have also been associated with the Meier–Gorlin syndrome (Bleichert et al. 2013, Nazarenko et al. 2022).

Failure to initiate replication can also be a result of reduced MCM2–7 levels, which increases the susceptibility of cells for DNA damage in the presence of replication inhibitors (Boyer et al. 2016). Under these conditions, incomplete replication hampers chromosome segregation and leads to chromosome breakage.

Conversely, elevated levels or activity of loading factors, for example CDT1 and CDC6, are observed in different cancer cells (Abbas et al. 2013, Boyer et al. 2016, Muñoz and Méndez 2017). Thereby, excessive origin firing causes re-replication, depletion of deoxyribonucleotide triphosphates (dNTPs) and the ssDNA-binder RPA, replication fork collapse and DNA damage.

To develop therapeutic strategies for these diseases, a detailed understanding of the underlying molecular mechanisms needs to be acquired. Studies using the model organism *S. cerevisiae* allow to uncover the principles of eukaryotic DNA replication. Ultimately, however, the *H. sapiens* system needs to be analysed to tackle pathologies.

1.6 Structure and function of the Dbf4-dependent kinase during origin firing

1.6.1 Structure and function of Cdc7–Dbf4

DDK belongs to the family of serine/threonine kinases. This protein family transfers a phosphoryl group (γ -phosphate) from ATP to a phospho-acceptor, like the hydroxy group of serine or threonine. The reaction is catalysed by a divalent metal ion. Substrates are usually transiently bound to the kinase to allow fast phospho-transfer and dissociation. To increase affinity and specificity to the substrate, kinases often have additional substrate recognition sites (Biondi and Nebreda 2003). Substrate binding can also lead to allosteric conformational changes that

stimulate kinase activity. How DDK employs these principles is partially understood and will be discussed below.

DDK is a heterodimer, comprised of the Cdc7 kinase and the activating subunit dumbbell former 4 (Dbf4) in *S. cerevisiae*, also known as activator of the S-phase kinase (ASK) in higher eukaryotes. Crystal structures of *H. sapiens* DDK show that CDC7 has a bi-lobed fold typical for kinases where the active site lies in the central cleft (Hughes et al. 2012, Dick et al. 2020) (Figure 1.8a). A hinge allows opening and closing of the cleft and establishes additional contacts with the nucleotide in the active site.

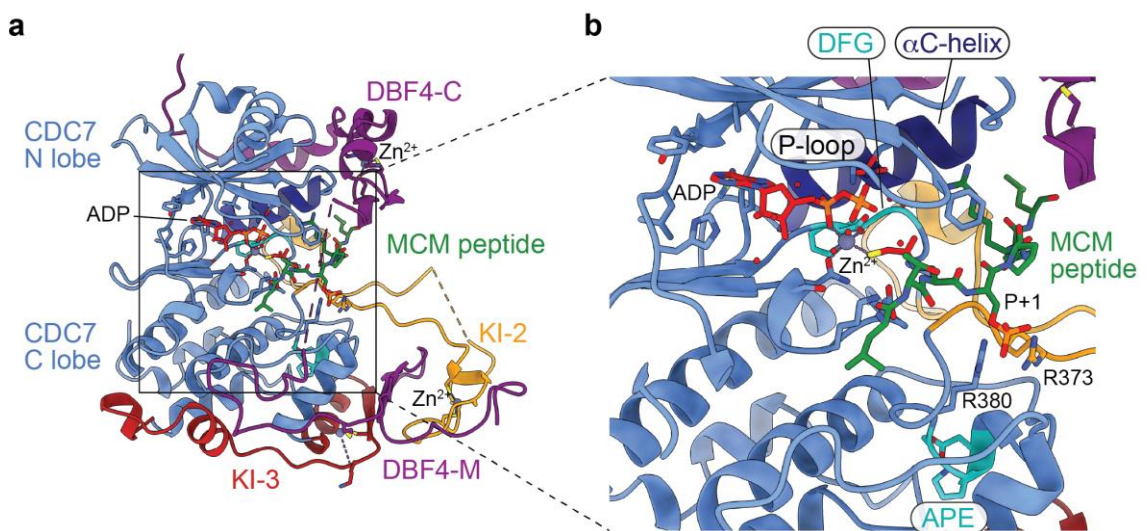


Figure 1.8: Crystal structure of the catalytic core of *H. sapiens* DDK.

a. Structure of the CDC7–DBF4 heterodimer with an MCM substrate peptide in the active site (PDB 6YA7). Important structural elements are highlighted. **b.** Zoomed-in view of the active site, showing the DFG, P-loop and αC-helix elements that coordinate the nucleotide as well as the two arginine residues (R380, R373) that bind the phosphorylated peptide at the P+1 site.

The N-terminal lobe is primarily formed of a five-stranded β-sheet and a regulatory α-helix. A glycine-rich loop between the first two β-strands, which is called the P-loop, coordinates the phosphate groups of the nucleotide in the active site and promotes ADP release after hydrolysis (Taylor and Kornev 2011) (Figure 1.8b). The regulatory αC-helix has a conserved glutamate that binds to the α and β phosphate of ATP and is important for kinase activation (Huse and Kuriyan 2002, Hughes et al. 2012).

The C-terminal lobe consists of multiple α -helices, harbouring the conserved DFG (AspPheGly) and APE (AlaProGlu) motifs, which are integral for substrate engagement and kinase function. The activation loop spanning the DFG and APE motifs contains a threonine in many kinases, which greatly stimulates catalytic activity upon autophosphorylation (Oliver et al. 2007, Wybenga-Groot et al. 2014, Gógl et al. 2019). However, glutamate is found at the equivalent position in mammalian, avian and reptile CDC7, meaning that self-activation by autophosphorylation does not occur in these species (Hughes et al. 2012). Substrate phosphorylation is enhanced by the presence of an acidic or phosphorylated residue next to the DDK target site (P+1 site) (Cho et al. 2006, Montagnoli et al. 2006, Hughes et al. 2012, Dick et al. 2020). Responsible for coordination of the negatively charged P+1 residue are two arginine residues within the activation loop and the C lobe.

The conserved kinase fold of CDC7 is interspersed with kinase insert (KI) regions, which vary in length and occurrence across species. *S. cerevisiae* Cdc7 has three inserts (KI-1, KI-2 and KI-3), whereas metazoan CDC7 only contains two (KI-2 and KI-3). These additional regions are largely unstructured and do mostly not affect *in vitro* kinase activity, hence large segments of the protein were deleted to aid crystallisation and structural characterisation (Hughes et al. 2012, Dick et al. 2020). *In vivo*, KI-2 and KI-3 are associated with nuclear import and export of the kinase (Kim and Lee 2006, Kim et al. 2007). However, one part of KI-2, which contains a Zn finger motif in metazoan CDC7, also has an activating effect on kinase activity by fixing the activation loop onto DBF4 and the C lobe. This opens the active site and allows substrate binding (Dick et al. 2020). The Zn finger within KI-2 is not conserved in yeast and it remains to be established whether a similar stabilisation of the activation loop is provided by other structural elements.

Dbf4 was named due to the dumbbell-like shape of mutant *S. cerevisiae* cells carrying a mutation in the *DBF4* gene (Johnston and Thomas 1982). The dumbbell-like phenotype is associated with an S-phase arrest, indicating the role of Dbf4 during replication initiation. The crystal structure of *H. sapiens* DDK shows that DBF4 wraps around CDC7 and stimulates kinase activity and substrate

engagement (Hughes et al. 2012). It is largely unstructured but has three conserved domains, referred to as N-terminal (N), middle (M) and C-terminal (C) domains. Important for *in vitro* kinase activity are the M and C domains, that brace the CDC7 C and N lobe, respectively (Figure 1.8a). The M domain forms a β -sheet with KI-3 and interacts with the KI-2 Zn finger in human DDK. The C domain of DBF4 contains a Zn finger motif, which is extended by an α -helix. The N domain has a BRCT or helix-BRCT motif and is dispensable for *in vitro* kinase activity (Hughes et al. 2012, Matthews et al. 2012). However, deletion of this domain causes lethality in mouse embryonic stem cells (Yamashita et al. 2005). Several interactors with the Dbf4 BRCT domain have been identified, which include Mcm2 and the checkpoint kinase Rad53 (Varrin et al. 2005, Francis et al. 2009, Matthews et al. 2012, Ramer et al. 2013, Abd Wahab and Remus 2020). The N domain thereby plays a pivotal role during DNA replication and under replication stress.

The crystal structure of human CDC7–DBF4 in complex with a substrate peptide (Dick et al. 2020) has led to a detailed description of the kinase fold and the process of peptide phosphorylation. Nevertheless, it does not explain how DDK selectively recognises the DNA-loaded DH and how the unstructured domains of CDC7 and DBF4 regulate substrate recognition and kinase activity. Structural characterisation of full-length DDK in combination with the DH will shed light onto this.

1.6.2 Regulation of DDK

The importance of tightly controlled DDK regulation is highlighted by the deleterious effects observed in cancer progression upon overexpression of CDC7, which correlates with poor prognosis for patients (Nambiar et al. 2007, Ghatalia et al. 2016). Absence of CDC7 is equally detrimental, resulting in reduced DNA replication and S phase arrest as observed in mouse embryonic stem cells (Kim et al. 1998) and embryonic lethality in a mouse model (Kim et al. 2002). Depletion of CDC7 results in senescence and apoptosis in p53-mutant cells, which makes DDK an attractive target for cancer treatment (Montagnoli et al. 2004, Suski et al. 2022).

DDK activity is influenced in multiple ways. Cdc7 is constantly expressed and associates with chromatin throughout the cell cycle in *S. cerevisiae* (Weinreich and Stillman 1999, Duncker et al. 2002). In higher eukaryotes, CDC7 is phosphorylated by the mitotic kinase CDK1, resulting in dissociation from chromatin and preventing re-replication in mitosis (Knockleby et al. 2016). Phosphorylation is counteracted by the protein phosphatase $\underline{1} \ \underline{\alpha}$ (PP1 α) to allow replication in the next cell cycle. The Cdc7 activator Dbf4 experiences fluctuations throughout the cell cycle. The oscillating levels of Dbf4 restrict DDK activity to late G₁ and S phase, which comes down to different expression levels and targeted degradation during the cell cycle (Jackson et al. 1993, Kumagai et al. 1999, Oshiro et al. 1999). Expression levels of Dbf4 are at their highest during S phase and are at their lowest at the G₂/M transition (Kumagai et al. 1999). Protein levels are further reduced due to the association of the APC/C with its adaptor protein Cdh1 during G₁ phase. In this form, the APC/C binds and degrades Dbf4 to limit DDK activity (Oshiro et al. 1999, Weinreich and Stillman 1999, Ferreira et al. 2000). During S phase, Cdh1 is phosphorylated by CDK preventing the interaction with the APC/C (Zachariae et al. 1998). This results in increasing concentrations of Dbf4 and activation of Cdc7.

Many kinases are regulated by autophosphorylation. Autophosphorylation of the activation loop, for example, stimulates catalytic activity in several cases (Oliver et al. 2007, Wybenga-Groot et al. 2014, Gógl et al. 2019). Given DDK lacks the critical threonine for self-activation within the activation loop, DDK autophosphorylation did not lead to activation, but rather a reduction in substrate peptide phosphorylation (Dick et al. 2020). How DDK autophosphorylation affects DH phosphorylation and engagement has not been studied yet.

A feedback loop is provided by the PP1, which counteracts the phosphorylation of the MCM (Davé et al. 2014, Hiraga et al. 2014, Mattarocci et al. 2014). PP1 is recruited to the MCM by the Rap-interacting factor 1 (Rif1). Rif1 itself is a DDK target and phosphorylation negatively affects the PP1–Rif1 interaction.

DNA damage or low nucleotide levels result in replication stress. To avoid genomic instability, DNA replication is slowed down. Under these conditions, the Mec1 kinase (ATR in humans) activates the checkpoint kinase Rad53 via the

phosphorylation of the mediator Mrc1 (Segurado and Tercero 2009, Pardo et al. 2017, Saldivar et al. 2017). This triggers an array of cellular responses to regulate damage-dependent transcription, cell cycle progression, dNTP levels, replication origin firing and replication fork progression (Santocanale and Diffley 1998, Zhao et al. 1998, Zegerman and Diffley 2010, Bastos de Oliveira et al. 2012, McClure and Diffley 2021). In this process, transcription of *H. sapiens* CDC7 is reduced and the protein is targeted for ubiquitin-dependent degradation (Tudzarova et al. 2016). Additionally, Dbf4 is amongst the Rad53 targets and becomes phosphorylated at multiple sites (Zegerman and Diffley 2010). Dbf4 phosphorylation abolishes DH phosphorylation and origin activation in *S. cerevisiae* (Santocanale and Diffley 1998, Lopez-Mosqueda et al. 2010, Zegerman and Diffley 2010). A recent study revealed that Rad53 and the DH compete for binding to DDK and that this competitive interaction reduces DH phosphorylation (Abd Wahab and Remus 2020). This is in contrast to work from McClure and Diffley (2021), which reports inhibition of origin firing only in the presence of catalytically active Rad53. In vertebrate, Dbf4 is also phosphorylated by ATM/ATR during the intra-S-phase checkpoint. Unlike in *S. cerevisiae*, DDK remains active under these conditions, potentially due to additional roles in stabilising and restarting stalled forks (Jones et al. 2021). Therefore, Dbf4 phosphorylation does not reduce kinase activity, or the association of Dbf4 with Cdc7, but may instead affect the interaction with kinase substrates (Lee et al. 2012).

1.6.3 Drf1 as an alternative activator of Cdc7

An alternative Cdc7 activator has been identified in *Xenopus laevis* and *H. sapiens*, which was termed Dbf4 related factor 1 (Drf1) or activator of S phase kinase-like protein 1 (ASKL1) (Montagnoli et al. 2002, Yoshizawa-Sugata et al. 2005). Sequence alignments of Dbf4 and Drf1 identified the conserved amino-terminal BRCT, M- and C-domains, which enable interaction with Cdc7 and phosphorylation of an MCM substrate. Drf1 has implications on DNA replication initiation, checkpoint response and chromosome cohesion in the early development in *X. laevis* (Takahashi and Walter 2005, Silva et al. 2006, Takahashi et al. 2008, Tsuji et al. 2008). Protein levels fluctuate during the cell cycle in a similar fashion as for

Dbf4, with highest levels being observed in S phase (Montagnoli et al. 2002). Both activators of Cdc7 thereby seem to be similarly regulated to prevent re-replication. While Dbf4 is found in most somatic cells, Drf1 is highly expressed in testis, some tumour cell lines and *X. laevis* early embryos (Montagnoli et al. 2002, Takahashi and Walter 2005). It has been shown that the Drf1–Cdc7 complex is the main form of DDK during early embryogenesis of *X. laevis*, whereas it is replaced by Dbf4–Cdc7 after the mid-blastula transition (MBT) (Takahashi and Walter 2005, Silva et al. 2006). The Cdc7–Drf1 kinase is also subject to autophosphorylation (Montagnoli et al. 2002). However, the effect on kinase function is unknown.

In *X. laevis* embryos, Drf1 is an important target in the regulation of the MBT (Collart et al. 2013, Collart et al. 2017). It is phosphorylated by the checkpoint kinase Chk1, which results in ubiquitylation by the SCF ^{β -TRCP} E3 ubiquitin ligase and Drf1 degradation (Collart et al. 2017). Limiting levels of Drf1 as well as other replication initiation factors then result in inhibition of replication initiation and lengthening of the cell cycle. Over-expression of Drf1 in the early embryo results in an increase of rapid, synchronous cell divisions at the MBT so that embryos have more but smaller cells (Collart et al. 2013). This deregulation ultimately leads to embryonic lethality. In contrast, Dbf4 is not phosphorylated by Chk1. The different expression patterns during embryogenesis and regulation by Chk1 suggest distinct roles of Drf1 and Dbf4. It is, however, not clear whether these roles are based on their function in DNA replication initiation or other cellular processes (Collart et al. 2017).

1.7 Aims

My PhD work aimed to address the molecular mechanisms of MCM loading and its targeted activation by DDK. In recent years, significant progress in the understanding of DNA replication has been made based on the *in vitro* reconstitution of the process using the model organism *S. cerevisiae*. However, it has not been addressed how DDK selectively phosphorylates the DNA-loaded MCM double hexamer, which has the symmetry to support bidirectional replication. Equally, it was unclear how the checkpoint kinase Rad53 interferes with the interaction to prevent late origin firing during the DNA damage response. I used a combination of biochemical and cryo-EM approaches to characterise MCM phosphorylation structurally and functionally in *S. cerevisiae*. Furthermore, I intended to reconstitute origin licensing with *H. sapiens* proteins to understand the conservation of molecular processes compared to the model organism. Reconstitution of *H. sapiens* DNA replication will later allow to directly test the effect of variants identified in diseases like the Meier–Grolin syndrome. Chapter 3 summarises the structural features of the licensed *S. cerevisiae* origin of replication and several licensing intermediates observed with *H. sapiens* proteins. Chapter 4 describes the molecular basis of the selective DH phosphorylation and its regulation by autophosphorylation and the checkpoint kinase Rad53.

Chapter 2. Materials & Methods

2.1 Reagents

2.1.1 *Saccharomyces cerevisiae* strains

Table 2.1: Yeast strains

Strain	Genotype (all in W303 background)	Reference
yAM33	<i>MATa</i> <i>bar1::Hyg</i> <i>pep4::KanMx</i> <i>his3::pRS303-Cdt1/Gal4 (HIS3)</i> <i>ura3::pRS306-Mcm2/Mcm3 CBP-Mcm3 (URA3)</i> <i>trp1::pRS304-Mcm4/5 (TRP1)</i> <i>leu2::pRS305-Mcm6/7 (LEU2)</i>	Coster et al. (2014)
yJF1a	<i>W303-1a</i> <i>pep4::KanMx4</i> <i>bar1::Hph-NT1</i>	Frigola et al. (2013)
yJG3	<i>MATa ade2-1 ura3-1 his3-11,15 trp1-1 leu2-3,112 can1-100</i> , <i>pep4::KanMX</i> , <i>trp1::TRP1pRS304/CDC7, CBP-DBF4^{Δ119-219}</i>	This study
yJG6	<i>MATa ade2-1 ura3-1 his3-11,15 trp1-1 leu2-3,112 can1-100</i> , <i>pep4::KanMX</i> , <i>trp1::TRP1pRS304/CDC7, CBP-DBF4^{T188D}</i>	This study
yJG10	<i>MATa ade2-1 ura3-1 his3-11,15 trp1-1 leu2-3,112 can1-100</i> , <i>pep4::KanMX</i> , <i>trp1::TRP1pRS304/CDC7, CBP-DBF4^{S192D}</i>	This study
yJG13	<i>MATa ade2-1 ura3-1 his3-11,15 trp1-1 leu2-3,112 can1-100</i> , <i>pep4::KanMX</i> , <i>trp1::TRP1pRS304/CDC7, CBP-DBF4^{T188D,S192D}</i>	This study
ySDK8	<i>MATa ade2-1 ura3-1 his3-11,15 trp1-1 leu2-3,112 can1-100</i> , <i>pep4::KanMX</i> , <i>trp1::TRP1pRS304/CDC7, CBP-DBF4</i>	On et al. (2014)

ySD- ORC	<i>MATa ade2-1 ura3-1 his3-11,15 trp1-1 leu2-3,112 can1-100</i> <i>bar1::Hyg</i> <i>pep4::KanMX</i> <i>his3::pRS303-ORC3/4 (HIS3)</i> <i>ura3::pRS306-ORC1/2 CBP-Orc1 (URA3)</i> <i>trp1::pRS304-ORC5/6 (TRP1)</i>	Frigola et al. (2013)
---------------------	---	-----------------------

2.1.2 Oligonucleotides

Table 2.2: Oligonucleotides

Name	Sequence	Purpose
AM133	CTCAACCCCTATCTCGGTCTATTCT	Confirming integration of pRS304
AM134	GTGCTGCAATGATACCGCGAG	Confirming integration of pRS304
AM139	ACAAGTTTGATTCCATTGCGGTG	Confirming integration of pRS304
pRS304 for	CAACACTGAGTAATGGTAGT	Confirming integration of pRS304
Dbf4-deltaMCM2i_F	GGAAGAAGCACTTGCTAAGACTAAGTCCGC	Deletion of Dbf4 BRCT domain
Dbf4-deltaMCM2i_R	GACAAAGTGCTTCTCCAGTTGGTTGCCATTCC	Deletion of Dbf4 BRCT domain
Dbf4-T188D_F	GAAGGACGATGATATTTTGTCTAGAGCTAAGAAGAAC	Aspartate substitution for residue T188 of Dbf4
Dbf4-T188D_R	CAAAATATCATCGTCCTTCAACAAAGTAGATGTTTTTC	Aspartate substitution for residue T188 of Dbf4
Dbf4-S192D_F	CCGATATTTTGGATAGAGCTAAGAAGAACTAC	Aspartate substitution for residue S192 of Dbf4
Dbf4-S192D_R	GCTCTATCCAAAATATCGGTGTCCTTCAAC	Aspartate substitution for residue S192 of Dbf4
Dbf4-T188D-S192D_2_F	GATGATATTTTGGATAGAGCTAAGAAGAACTACATGAAGG	Aspartate substitution for residues T188 and S192 of Dbf4
Dbf4-T188D-S192D_R	CTTAGCTCTATCCAAAATATCATCGTCCTTCAACAAAG	Aspartate substitution for residues T188 and S192 of Dbf4
MTRB F	(Des)-ATATATCC*GGCCTGTATTTTACAGATTTTATGTTTAGATCCTTTTATGC	Amplification of M.Hpal-ARSI-M.Hpal DNA construct
MTRB R	AGGGCGCC*GGAAGTGGGAAAATAGCAAATTTTCGTCAAAAATGC	Amplification of M.Hpal-ARSI-M.Hpal DNA construct
DESBIO - 177 bp (Aval) F	(Des)-CGAtagaaCTCGGGccgcctggagaatcGcggtgccg	Amplification of nucleosome-Gid70-nucleosome construct
DNACS R1	cctgcACCCCGGGGACTTTGAAGTAATAAGGAC	Amplification of nucleosome-Gid70-nucleosome construct
(Des) – Desthiobiotin TEG		
C* - 5-fluoro-2'-deoxycytosine		

2.1.3 Plasmids

Table 2.3: Plasmids

Name	Construct	Reference
pAM3	pGEX-6p-1-CDC6	Frigola et al. (2013)
	pET21b-RAD53	Gilbert et al. (2001)
pAWM7	pET21b-RAD53 ^{K227A,D339A} -6xHis	McClure and Diffley (2021)
pJG1	pRS304-CDC7-GAL-CBP-DBF4 ^{Δ119-219}	This study
pJG3	pRS304-CDC7-GAL-CBP-DBF4 ^{T188D}	This study
pJG6	pRS304-CDC7-GAL-CBP-DBF4 ^{S192D}	This study
pJG9	pRS304-CDC7-GAL-CBP-DBF4 ^{T188D,S192D}	This study
pNAN	N-ARS1-N – nucleosome-ARS1-nucleosome origin DNA template	Miller et al. (2019)
pD72	N-Gid70-N – synthetic origin DNA template containing two high-affinity ORC binding sites with 70 base pairs in between flanked by Widom 601 and Widom 603 nucleosome positioning sequences	Generated within the laboratory by Dr Thomas Miller

2.1.4 Buffers

Table 2.4: Buffers

Name	Components	Purpose
5x Laemmli buffer	10% (v/v) sodium dodecyl sulfate (SDS), 25% (v/v) glycerol, 250 mM Tris pH 6.8, 50 mM dithiothreitol (DTT)	Protein loading buffer for SDS-PAGE
Lithium acetate/TE buffer	0.1 M lithium acetate pH 7.5, 10 mM Tris-HCl pH 7.5, 1 mM EDTA pH 8.0	Yeast transformation
Lithium acetate/PEG buffer	0.1 M lithium acetate, 10 mM Tris-HCl pH 7.5, 1 mM EDTA pH 8.0, 40% (w/v) polyethylene glycol (PEG) 3,350	Yeast transformation

1x TAE buffer	40 mM Tris base, 40 mM acetic acid, 1 mM ethylenediaminetetraacetic acid (EDTA)	Yeast transformation
0.5x TAE buffer	20 mM Tris base, 10 mM acetic acid, 0.5 mM EDTA pH 8.0	DNA preparation
TE buffer	10 mM Tris-HCl pH 7.5, 1 mM EDTA pH 8.0	DNA preparation
Buffer 1	25 mM HEPES-KOH pH 7.6, 0.05% (v/v) NP-40, 10% (v/v) glycerol, 2 mM β -mercaptoethanol	<i>S. c.</i> ORC purification
Buffer 2	45 mM HEPES-KOH pH 7.6, 100 mM potassium acetate, 5 mM magnesium acetate, 0.02% (v/v) NP-40, 10% (v/v) glycerol, 2 mM β -mercaptoethanol	<i>S. c.</i> Mcm2–7–Cdt1 purification
Buffer 3	25 mM HEPES-KOH pH 7.6, 400 mM sodium chloride, 0.02% (v/v) NP-40, 10% (v/v) glycerol, 1 mM DTT	<i>S. c.</i> DDK purification
Buffer 4	25 mM HEPES-KOH pH 7.6, 0.02% (v/v) NP-40, 1 mM DTT	Anion exchange chromatography of <i>S. c.</i> DDK
Buffer 5	25 mM HEPES-KOH pH 7.6, 0.2 M potassium glutamate, 0.02% (v/v) NP-40, 1 mM DTT	Gel filtration of <i>S. c.</i> DDK
Buffer 6	50 mM potassium phosphate pH 7.6, 150 mM potassium acetate, 5 mM magnesium chloride, 2 mM ATP, 1% (v/v) triton X-100, 1 mM DTT	<i>S. c.</i> Cdc6 purification
Buffer 7	50 mM potassium phosphate pH 7.6, 5 mM magnesium chloride, 2 mM ATP, 0.1% (v/v) triton X-100, 1 mM DTT	<i>S. c.</i> Cdc6 purification
Buffer 8	25 mM HEPES-KOH pH 7.6, 10 mM magnesium acetate, 0.02% (v/v) NP-40, 100 mM potassium acetate, 10% (v/v) glycerol, 5 mM β -mercaptoethanol	Dialysis buffer for <i>S. c.</i> Cdc6

Buffer 9	25 mM HEPES-KOH pH 7.6, 300 mM sodium chloride, 0.02% (v/v) NP-40, 10% (v/v) glycerol	<i>S. c.</i> Rad53 purification
Buffer 10	20 mM HEPES pH 8.0, 0.1 mM EDTA, 10 mM β -mercaptoethanol	Histone octamer purification
Buffer 11	20 mM Tris pH 8.6, 10 mM potassium chloride, 10 mM ammonium sulfate, 2 mM magnesium chloride, 0.1% (v/v) triton X-100, 0.1 mg/ml bovine serum albumin (BSA)	PCR amplification of DNA templates
Buffer 12	50 mM Tris pH 8.0, 5 mM β -mercaptoethanol	Purification of DNA templates
Buffer 13	50 mM potassium acetate, 25 mM Tris pH 7.5, 10 mM magnesium acetate, 1 mg/ml BSA, 150 μ M S-adenosyl-methionine (NEB)	<i>Hpa</i> II methyltransferase (<i>M.Hpa</i> II) conjugation
Buffer 14	20 mM HEPES pH 7.6, 5 mM β -mercaptoethanol	Nucleosome reconstitution
Buffer 15	20 mM HEPES pH 7.6, 80 mM potassium chloride, 100 mM sodium acetate, 1 mM DTT	Final dialysis buffer during nucleosome reconstitution
Buffer A	25 mM HEPES-KOH pH 7.6, 100 mM potassium acetate, 10 mM magnesium acetate, 0.02% (v/v) NP-40, 5% (v/v) glycerol, 1 mM DTT, 5 mM ATP	<i>In vitro</i> reconstitution of <i>S. c.</i> origin licensing
Buffer B	25 mM HEPES-KOH pH 7.6, 100 mM potassium acetate, 10 mM magnesium acetate, 1 mM DTT, 5 mM ATP	EM buffer for <i>in vitro</i> reconstituted <i>S. c.</i> origin licensing
Buffer C	25 mM HEPES-KOH pH 7.6, 100 mM potassium glutamate, 10 mM magnesium acetate, 1 mM DTT, 2 mM ATP	EM buffer for <i>in vitro</i> reconstituted <i>H. s.</i> origin licensing

Buffer D	25 mM HEPES-KOH pH 7.6, 100 mM potassium acetate, 10 mM magnesium acetate, 1 mM DTT	Low-salt buffer for the purification of DNA-loaded <i>S. c.</i> DH by anion exchange chromatography
Buffer E	25 mM HEPES-KOH pH 7.6, 2 M sodium chloride, 10 mM magnesium acetate, 1 mM DTT	High-salt buffer for the purification of DNA-loaded <i>S. c.</i> DH by anion exchange chromatography
High-salt buffer	25 mM HEPES-KOH pH 7.6, 500 mM NaCl, 5 mM magnesium acetate, 0.02% (v/v) NP-40	Wash buffer for DNA-affinity purification
Low-salt buffer	25 mM HEPES-KOH pH 7.6, 300 mM sodium acetate, 5 mM magnesium acetate, 0.02% (v/v) NP-40	Wash buffer for DNA-affinity purification
Kinase assay reaction buffer	40 mM HEPES-KOH pH 7.6, 10 mM magnesium acetate, 2 mM DTT, 0.1% (v/v) NP-40, 80 µg/ml BSA, 1 mM β-glycerophosphate, 1 mM sodium fluoride, 0.1 mM ATP	Peptide phosphorylation by DDK

2.1.5 Media for *S. cerevisiae* and *E. coli* cells

Table 2.5: Media for *S. cerevisiae*

Name	Components	Purpose
Solid selectivity media	2.5% agarose, 0.67% (w/v) yeast nitrogen base without amino acids, 2% (v/v) glucose, 40 µg/ml Adenine, 40 µg/ml Uracil, 80 µg/ml Leucine, 40 µg/ml Histidine	Selection of yeast strains with DDK integrated into the TRP1 locus
Yeast Peptone (YP) medium	1% (w/v) yeast extract, 2% (w/v) peptone	Protein expression

YPD medium	1% (w/v) yeast extract, 2% (w/v) peptone, 2% (w/v) glucose	Strain propagation
YPD agar	1% (w/v) yeast extract, 2% (w/v) peptone, 2% (w/v) glucose, 2% agar	Strain propagation

Table 2.6: Media for *E. coli*

Name	Components
LB agar plates	1% (w/v) bacto tryptone, 0.5% (w/v) yeast extract, 170 mM sodium chloride, 1.5% agar
LB medium	1% (w/v) bacto tryptone, 0.5% (w/v) yeast extract, 170 mM sodium chloride
SOC medium	2% (w/v) bacto tryptone, 0.5% (w/v) yeast extract, 1 mM sodium chloride, 2.5 mM potassium chloride, 10 mM magnesium chloride, 10 mM magnesium sulphate, 3.6% (w/v) glucose

2.2 Generation of *S. cerevisiae* strains for protein expression

2.2.1 Plasmid generation

The parental plasmid, pRS304/Cdc7-Gal-CBP-Dbf4, was used for the generation of mutant DDK constructs. It encodes for wild type *S. cerevisiae* Cdc7 and N-terminally calmodulin binding protein (CBP)-tagged Dbf4 under the control of a bidirectional galactose inducible promoter (On et al. 2014). In the Δ BRCT construct, an internal truncation was achieved by deleting codons for amino acids 119–219. Additional constructs were produced, yielding aspartate substitutions for residues T188 and S192 of Dbf4, either in isolation or in combination. The In-Fusion® HD Cloning Kit (Takara Bio USA) was used to introduce these changes. 2 ng of the parental plasmid was mixed with 0.3 μ M of forward and reverse primers (see Table 2.2), 12.5 μ l CloneAmp HiFi PCR Premix (Takara Bio USA) and water was added, resulting in a total volume of 25 μ l. The mix was subjected to polymerase chain reaction (PCR), which typically consisted of one cycle of denaturation at 98°C for 1 minute, followed by 35 cycles of denaturation at 98°C for 10 seconds, primer annealing at 55–58°C for 15 seconds and elongation at 72°C for 45 seconds (5-10

seconds per kilobase). Replication of the DNA template was finished off with a polishing step at 72°C for 5 minutes.

PCR products were separated by gel electrophoresis using a 1% (w/v) agarose gel in TAE buffer (40 mM Tris Base, 20 mM acetic acid, 50 mM EDTA). DNA was visualised using a 1:50,000 dilution of SYBR Safe DNA Gel Stain (Invitrogen) mixed into agarose during gel preparation. Bands corresponding in size to the linearised plasmid of interest were excised from the gel and the DNA was isolated using the NucleoSpin Gel and PCR Clean-Up kit (Takara Bio USA) according to the manufacture's protocol.

For the In-Fusion Cloning procedure, 100 ng of the purified, linearised vector were mixed with 2 µl of the In-Fusion HD Enzyme Premix (Takara Bio USA) and water was added, obtaining a total reaction volume of 10 µl. The two ends of the linearised vectors were ligated by incubation at 50°C for 15 minutes.

Stellar Competent Cells (Takara Bio USA) were transformed with 2.5 µl of the ligation reaction using a 30-minute incubation on ice, followed by heat shock at 42°C for 45 seconds. Cells were kept on ice for 1–2 minutes and grown at 37°C for 1 hour, after addition of 500 µl SOC medium. 70 µl of the cell suspension were plated on LB agar plates supplemented with 100 µg/ml ampicillin and incubated overnight at 37°C.

Single colonies were screened for the correct change in DNA sequence by extracting plasmid DNA from 5-ml bacterial cultures using the QIAprep Spin Miniprep protocol (Qiagen). 1 µg of the extracted plasmid DNA was digested using a combination of the restriction enzymes *AscI* and *XhoI* (5 units each) in CutSmart Buffer (NEB). Fragments were analysed on a 1% (w/v) agarose gel in TAE buffer and positive plasmids were confirmed by Sanger sequencing.

2.2.2 Transformation of *S. cerevisiae*

To generate yeast strains overexpressing the DDK variants, yeast strain yJF1a (Frigola et al. 2013) was transformed with the respective plasmid according to the following procedure. 2 µg of plasmid DNA were linearised using 20 units of restriction enzyme *Pml*I at 37°C for 1 hour in a total volume of 10 µl 1x CutSmart buffer (NEB). The restriction enzyme was then inactivated by incubation at 65°C for 20 minutes.

50 ml of YPD medium (YP + 2% (w/v) glucose) were inoculated with the parental strain yJF1a and grown overnight at 30°C, shaking at 180 rpm. Cells were then diluted to 2×10^6 cells/ml in fresh YPD medium and grown for 4.5 hours to a final concentration of 1×10^7 cells/ml. 10 ml of the cell suspension were harvested by centrifugation at 2,851 $\times g$ for five minutes. The supernatant was removed and cells were washed in 10 ml of sterile distilled water. Cells were again harvested, resuspended in 1 ml of sterile distilled water and transferred to a 1.5 ml microfuge tube. Cells were pelleted by centrifugation at maximum speed (21,130 $\times g$) for five seconds. The supernatant was replaced by 1 ml of sterile 0.1 M lithium acetate/TE buffer pH 7.5. After a short wash, cells were harvested as described and resuspended in 50 µl of 0.1 M lithium acetate/TE buffer, resulting in a final concentration of 2×10^9 cells/ml. 5 µl 10 mg/ml UltraPure Salmon Sperm DNA solution (Invitrogen), 1 µg of the linearised plasmid DNA and 300 µl of lithium acetate/PEG buffer were added to the cells and mixed thoroughly by pipetting. Cells were incubated for 30 minutes at 30°C, shaking at 180 rpm. 40 µl of 100% dimethyl sulfoxide (DMSO; 10% (v/v) final concentration) were added and cells were heat shocked for 15 minutes at 42°C. Cell suspensions were mixed thoroughly and cells were pelleted by centrifugation as described above. After removal of the supernatant, cells were resuspended in 100 µl TE buffer pH 7.5 and plated onto the appropriate selective media (2.5% (w/v) agarose supplemented with yeast nitrogen base without amino acids, 2% (v/v) glucose, 40 µg/ml Adenine, 40 µg/ml Uracil, 80 µg/ml Leucine, 40 µg/ml Histidine). Cells were grown at 30°C.

2.2.3 Isolation of yeast genomic DNA

Single colonies were used to inoculate 4 ml YPD cultures to screen for successful plasmid integration. Cultures were incubated overnight at 30°C, shaking at 180 rpm. Genomic DNA was extracted using the QIAprep Spin Miniprep protocol (Qiagen) with a few adjustments: Briefly, cells were pelleted by centrifugation and resuspended in 250 µl of buffer P1. An equal volume of 425-600 µm glass beads (Sigma) was added to the cells. Cells were disrupted by vortexing at maximum speed for 1 minute. 250 µl of buffer P2 were added and the tube was heated to 60°C for five minutes. 350 µl of the neutralising buffer N3 were added, before removing solids by centrifugation at maximum speed (21,130 *xg*) for 10 to 20 minutes. The supernatant was applied to the QIAprep Amicon Ultra centrifugal filter unit to bind the DNA. The silica membrane was washed twice with 500 µl of buffer PB and twice with 750 µl of buffer PE. Genomic DNA was then eluted using 50 µl of the elution buffer EB.

2.2.4 Colony-PCR

The genomic DNA of selected colonies was analysed for correct plasmid integration using three different primer pairs:

1. AM133 + AM134 span the point of the plasmid that was used for linearisation. A product of around 2,500 base pairs indicates tandem integration of the plasmid;
2. AM133 + pRS304;
3. AM134 + AM139.

The latter two primer pairs check for integration of the plasmid into the TRP1 locus. 0.8 µl of the isolated DNA were mixed with 2.5 mM dNTPs, 10 µM of each primer, 0.2 µl *Ex Taq*® DNA Polymerase (4 units; Takara Bio USA) diluted in the *Ex Taq* buffer in a total volume of 20 µl. PCR was initiated by incubating for 10 seconds at 98°C, followed by 35 cycles of DNA denaturation at 94°C for 30 seconds, followed by primer annealing at 55°C for 30 seconds and template elongation at 72°C for 2 minutes. The PCR was finalised by incubation at 72°C for 5 minutes. Products were separated on a 1% (w/v) agarose gel in TAE buffer and visualised using

SYBR Safe DNA Gel Stain (Invitrogen), with a 1:50,000 dilution in the gel. Colonies that produced a single band with primer pairs 2 and 3 were propagated by transferring onto fresh selective medium. Strains were subsequently grown to saturation in 5 ml YPD and subsequently stored in 50% (v/v) glycerol at -80°C . Strains that additionally produced the product of 2,500 bp with the first primer pair, indicating tandem integration, were also kept.

2.2.5 Small-scale expression to select *S. cerevisiae* strains

A small-scale expression test was carried out for all newly generated yeast strains to check for protein expression levels. Briefly, 50 ml YP medium supplemented with 2% (v/v) raffinose were inoculated with the respective yeast strain and incubated for about 19 hours at 30°C with shaking at 180 rpm. Cells were then diluted in fresh medium to a final concentration of 0.5×10^7 cells/ml. When the culture reached 10^7 cells/ml (approximately 2.5 hours after dilution), protein expression was induced by the addition of galactose to a final concentration of 2% (v/v). After 2 to 3 hours at 30°C with mixing at 180 rpm, 10^8 cells were harvested by centrifugation.

Cells were mixed with 200 μl of 20% (v/v) trichloroacetic acid (TCA) and an equal volume of 425–600 μm glass beads (Sigma). Cell lysis was achieved by vortexing at maximum speed for 1 minute. The lysate was transferred to a new microfuge tube and beads were washed twice with 200 μl 5% (v/v) TCA, adding the resulting suspension to the lysate each time. Protein precipitate, which formed due to the treatment with TCA, was isolated by centrifugation at 845 $\times g$ for 10 minutes. The pellet was resuspended in 200 μl Laemmli buffer and approximately 50 μl of 1 M Tris base were added to increase the pH (colour change from red to blue). The samples were incubated at 97°C for three minutes and cleared from any precipitate by centrifugation at 845 $\times g$ for 10 minutes. 10 μl of each sample were separated using a 3–8% tris-acetate SDS-PAGE. CBP-Dbf4 was detected by Western blotting, using a 1:3,750 dilution of the Anti-Calmodulin Binding Protein Epitope Tag Antibody (Sigma Aldrich) and a 1:30,000 dilution of the Swine Anti-Rabbit Immunoglobulins/Horseradish peroxidase antibody (Dako).

2.3 Protein expression in *S. cerevisiae*

2.3.1 ORC and Mcm2–7–Cdt1 expression

S. cerevisiae strains for overexpression of yeast ORC (N-terminal CBP tag on Orc1) and Mcm2–7–Cdt1 (N-terminal CBP tag on Mcm3) (see Table 2.1) were streaked onto individual YPD agar plates and incubated at 30°C for 48 hours. 100 l large-scale expressions were set up by the Fermentation Science Technology Platform at the Francis Crick Institute. First, starter cultures were prepared in YP medium supplemented with 2% (w/v) raffinose, which were grown at 30°C and 200 rpm shaking for about 24 hours. 100 l YP medium supplemented with 2% (w/v) raffinose and 100 µg/ml ampicillin were then inoculated with the starter culture and grown at 30°C. When cultures reached $2\text{--}3 \times 10^7$ cells/ml, alpha factor was added to a final concentration of 100 ng/ml to arrest cells in the G₁ phase of the cell cycle. After three hours, expression was induced by the addition of 2% (w/v) galactose for further 3 hours. Cells were harvested and washed in Buffer 1 (ORC; 25 mM HEPES-KOH pH 7.6, 0.05% (v/v) NP-40, 10% (v/v) glycerol, 2 mM β-mercaptoethanol) and Buffer 2 (Mcm2–7–Cdt1; 45 mM HEPES-KOH pH 7.6, 100 mM potassium acetate, 5 mM magnesium acetate, 0.02% (v/v) NP-40, 10% (w/v) glycerol, 2 mM β-mercaptoethanol), respectively. Cell pellets were resuspended in the same buffers supplemented with cComplete™ EDTA-free Protease Inhibitor Cocktail tablets (Roche) at half pellet volume, drop-wise flash frozen in liquid nitrogen and crushed using a 6875D Freezer/Mill Dual Chamber Cryogenic Grinderfreezer mill (SPEX SamplePrep) at intensity 15 (six cycles of 2 min of milling with 1 min of rest). The disrupted cells were stored at –80°C until purification.

2.3.2 DDK expression

S. cerevisiae strains for overexpression of wild type and mutant DDK (see Table 2.1) were grown on YPD agar plates at 30°C for 48 hours. Individual colonies were used to inoculate 100 ml YPD medium. Starter cultures were grown to saturation at 30°C and 200 rpm shaking for 24 hours. 10 µl of the starter culture were diluted 1:10 and applied to a haemocytometer to determine cell density. 2×10^8 cells were

added to each flask containing 1 l YP medium supplemented with 2% (w/v) raffinose. A total volume of 3, 6 or 12 l were prepared for each construct, depending on the required yield. When cultures reached a density of $2\text{--}3 \times 10^7$ cells/ml, expression was induced by the addition of galactose to a final concentration of 2% (w/v). Expression was carried out for 7 to 8 hours at 30°C and 200 rpm shaking. Cells were harvested by centrifugation at 4,552 $\times g$ for 20 min at 4°C in a Beckman Coulter J6-MC Centrifuge (JS-4.2 rotor). The supernatant was discarded. Cells were washed once with Buffer 3 (25 mM HEPES-KOH pH 7.6, 400 mM sodium chloride, 0.02% (v/v) NP-40, 10% (v/v) glycerol, 1 mM DTT), resuspended in Buffer 3 supplemented with cOmplete™ EDTA-free Protease Inhibitor Cocktail tablets (Roche) at half pellet volume and drop-wise flash frozen in liquid nitrogen. Frozen cells were crushed in a 6875D Freezer/Mill Dual Chamber Cryogenic Grinderfreezer mill (SPEX SamplePrep) at intensity 15 (six cycles of 2 min of milling with 1 min of rest). Cell powder was stored at –80°C until purification.

2.4 Protein expression in *E. coli*

2.4.1 Cdc6 expression

The expression plasmid for N-terminally glutathione-S-transferase (GST)-tagged *S. cerevisiae* Cdc6 (pAM3) was transformed into BL21 (DE3) RIL *E. coli* cells. 100 ng plasmid DNA were mixed with 50 μ l chemically competent cells, incubated on ice for 2 minutes. A heat shock of 1 minute at 42°C was followed by 2 minutes incubation on ice before adding 1 ml SOC medium and incubation at 37°C and shaking at 180 rpm for 1 hour. Cells were streaked onto LB agar plates supplemented with 100 μ g/ml ampicillin and 30 μ g/ml chloramphenicol and grown overnight at 37°C. A single colony was used to inoculate 50 ml LB medium supplemented with ampicillin and chloramphenicol and incubated overnight at 37°C and 200 rpm. A 3 l expression culture was set up and grown at 37°C to an optical density 600 (OD₆₀₀) of 0.6. Cultures were cooled to 18°C and expression of Cdc6 was induced by the addition of IPTG to a final concentration of 0.5 mM. After 5 hours at 18°C and 200 rpm shaking, cells were harvested by centrifugation at 4,552 $\times g$ for 10 min at room temperature in a Beckman Coulter J6-MC Centrifuge (JS-4.2

rotor). The supernatant was discarded, cells were flash frozen in liquid nitrogen and stored at -20°C until purification.

2.4.2 Rad53 expression

BL21 (DE3) RIL *E. coli* cells were transformed as described for Cdc6, using plasmids for expression of wild type Rad53 or kinase dead Rad53 (see Table 2.3), which both had a C-terminal histidine tag with six histidines. Single colonies were used to inoculate starter cultures of 50 ml LB medium supplemented with 100 $\mu\text{g/ml}$ ampicillin and 30 $\mu\text{g/ml}$ chloramphenicol. Starter cultures were grown overnight at 37°C and 200 rpm shaking. The next day, three 1 l cultures were inoculated with the starter culture and grown at 37°C to an OD_{600} of 0.4. At this point expression was induced by the addition of IPTG to a final concentration of 1 mM. After 2 hours, cells were harvested by centrifugation at 4,552 $\times g$ for 10 minutes at room temperature and subsequently flash frozen in liquid nitrogen. Pellets were stored at -20°C until purification.

2.4.3 Histone octamer expression

All four histones were co-expressed in BL21 (DE3) RIL *E. coli* cells. Therefore, cells were simultaneously transformed with plasmids pCDFDuet containing *S. cerevisiae* H2A (*NcoI NotI*), H2B (*NdeI XhoI*) and pETDuet containing H3 (*NcoI NotI*), H4 (*NdeI XhoI*) (Saravanan et al. 2012). 50 μl cells were mixed with 100 ng of each plasmid DNA and transformation was carried out as described for Cdc6. Transformed cells were streaked onto LB agar plates supplemented with 100 $\mu\text{g/ml}$ ampicillin, 30 $\mu\text{g/ml}$ chloramphenicol and 50 $\mu\text{g/ml}$ spectinomycin. A single colony was used to inoculate 50 ml LB supplemented with ampicillin, chloramphenicol and spectinomycin. The starter culture was grown at 37°C overnight and expanded to a total volume of 12 l the next day. When OD_{600} of 0.6 was reached, IPTG was added to a final concentration of 0.5 mM and expression was continued for four hours at 37°C , mixing at 200 rpm. Cells were harvested by centrifugation at 4,552 $\times g$ for 10 minutes and frozen in liquid nitrogen and stored at -20°C until purification.

2.5 Expression and purification of *H. sapiens* DNA replication factors

H. sapiens DNA replication factors were expressed and purified by Dr Florian Weissmann, a postdoctoral fellow in the group of Dr John Diffley. He used the baculovirus-insect cell expression system for all proteins and protein complexes apart from ORC6, which was expressed in *E. coli* cells. The amino-terminal IDRs of ORC1 (residues 1–390), CDC6 (residues 1–142) and CDT1 (residues 1–166) were omitted from the constructs to prevent phase separation and protein aggregation.

2.6 Purification of *S. cerevisiae* proteins

2.6.1 *S. cerevisiae* ORC purification

Freezer milled cells, which had overexpressed Orc1–5 containing Orc1 with an N-terminal CBP-tag, equivalent to 20 l yeast culture were thawed on ice and resuspended in an equal volume of Buffer 1 (25 mM HEPES-KOH pH 7.6, 0.05% (v/v) NP-40, 10% (v/v) glycerol, 2 mM β -mercaptoethanol) supplemented with 0.1 M potassium chloride and cOmplete™ EDTA-free Protease Inhibitor Cocktail tablets (Roche). Potassium chloride was added to yield a final concentration of 0.5 M, before isolating the soluble fraction by ultracentrifugation at 235,418 $\times g$ and 4 °C for one hour using an Optima L-100 XP Ultracentrifuge (Beckman Coulter). Calcium chloride was added to the cleared lysate to a final concentration of 2 mM, followed by 5 ml of Calmodulin Affinity Resin (Agilent; bed volume) pre-equilibrated in Buffer 1 supplemented with 0.4 M potassium chloride and 2 mM calcium chloride. Bead-binding was carried out at 4°C for two hours. Resin was washed with 80 ml of Buffer 1 supplemented with 0.4 M potassium chloride and 2 mM calcium chloride. CBP-tagged ORC was eluted using Buffer 1 supplemented with 0.4 M potassium chloride, 2 mM EDTA and 2 mM EGTA. Peak fractions were pooled and concentrated on a 100,000-molecular weight cut-off (MWCO) Amicon Ultra centrifugal filter unit. The sample was then separated on a Superdex 200 16/600 gel filtration column (Cytiva) in Buffer 1 supplemented with 0.15 M potassium chloride. Peak fractions were pooled and concentrated on an Amicon Ultra centrifugal filter unit. Protein concentration was determined according to the

Bradford method using the Protein Assay Dye Reagent Concentrate (BioRad).

Aliquots were flash frozen in liquid nitrogen and stored at -80°C .

2.6.2 *S. cerevisiae* Mcm2–7–Cdt1 purification

Freezer milled cells, which had overexpressed the Mcm2–7–Cdt1 complex with an N-terminal CBP-tag on Mcm3, equivalent to 12 l yeast culture was thawed and resuspended in Buffer 2 (45 mM HEPES-KOH pH 7.6, 100 mM potassium acetate, 5 mM magnesium acetate, 0.02% (v/v) NP-40, 10% (v/v) glycerol, 2 mM β -mercaptoethanol) supplemented with protease inhibitors. The lysate was cleared from cell debris and lipids by ultracentrifugation for one hour at 235,418 $\times g$ and 4°C using an Optima L-100 XP Ultracentrifuge (Beckman Coulter). The soluble fraction was supplemented with 2 mM calcium chloride and incubated with 1 ml of Calmodulin Affinity Resin (Agilent; bed volume) pre-equilibrated in Buffer 2 supplemented with 2 mM calcium chloride. After two hours at 4°C , resin was washed with 80 ml of Buffer 2 supplemented with 2 mM calcium chloride. Elution was obtained by the addition of Buffer 2 supplemented with 2 mM EDTA and 2 mM EGTA. Mcm2–7–Cdt1-containing fractions were pooled, concentrated on a 100,000-MWCO Amicon Ultra centrifugal filter unit and separated on a Superdex 200 16/600 gel filtration column (Cytiva) in Buffer 1 supplemented with 0.15 M potassium chloride. Peak fractions were pooled and concentrated. After determination of the protein concentration, aliquots were flash frozen in liquid nitrogen and stored at -80°C .

2.6.3 *S. cerevisiae* DDK purification

Purification was carried out as described before (On et al. 2014), with an additional anion exchange chromatography step to increase sample purity.

Freezer milled cells equivalent to 6 l yeast culture was thawed and resuspended in an equal volume of Buffer 3 (25 mM HEPES-KOH pH 7.6, 400 mM sodium chloride, 0.02% (v/v) NP-40, 10% (v/v) glycerol, 2 mM DTT) supplemented with EDTA-free protease inhibitors. The soluble fraction was isolated by

ultracentrifugation at 235,418 xg and 4°C for one hour using an Optima L-100 XP Ultracentrifuge (Beckman Coulter). The cleared lysate was supplemented with 2 mM calcium chloride and incubated for 3 hours with 2.5 ml of Calmodulin Affinity Resin (Agilent; bed volume) to isolate DDK *via* the N-terminal CBP-tag on Dbp4. Buffer 3 supplemented with 2 mM calcium chloride was used to wash beads extensively. Heat shock proteins were removed by incubation with 20 ml of Buffer 3 supplemented with 2 mM calcium chloride, 1 mM ATP and 10 mM magnesium acetate for 15 minutes at 4°C. The supernatant was replaced by the same buffer and incubation with ATP was repeated once more. Beads were again washed extensively with Buffer 3 supplemented with 2 mM calcium chloride and treated for 1 hour at 4°C with 2,800 units of λ phosphatase (NEB). Excess λ phosphatase was removed by washing with 100 ml of Buffer 3 supplemented with 2 mM calcium chloride. Buffer 3 supplemented with 2 mM EDTA was used to elute DDK. Subsequently, the solution was diluted in Buffer 3 lacking sodium chloride to reduce the salt concentration to 0.1 M. The sample was then separated on a 5 ml HiTrap heparin anion exchange column (Cytiva) using a 10 CV gradient from 0.1 M to 1 M sodium chloride in Buffer 4 (25 mM HEPES-KOH pH 7.6, 0.02% (v/v) NP-40, 1 mM DTT). Peak fractions were concentrated using a 30,000-MWCO Amicon Ultra centrifugal filter unit and subjected to size exclusion chromatography using a Superdex 200 16/600 gel filtration column (Cytiva) in Buffer 5 (25 mM HEPES-KOH pH 7.6, 0.2 M potassium glutamate, 0.02% (v/v) NP-40, 1 mM DTT). Target fractions were pooled, concentrated, flash frozen in liquid nitrogen and stored at –80°C.

2.6.4 *S. cerevisiae* Cdc6 purification

The cell pellet of 3 l bacterial culture expressing Cdc6 with an N-terminal GST-tag was thawed and resuspended in 30 ml Buffer 6 (50 mM potassium phosphate pH 7.6, 150 mM potassium acetate, 5 mM magnesium chloride, 2 mM ATP, 1% (v/v) triton X-100, 1 mM DTT) supplemented with protease inhibitor tablets. Cell lysis was achieved by sonication. The soluble fraction was isolated by centrifugation for 20 minutes at 58,545 xg and 4°C. The lysate was applied twice to 2 ml of pre-equilibrated Glutathione Sepharose 4B (GE; bed volume) on a gravity

chromatography column. Beads were washed with 40 ml of Buffer 6 supplemented with protease inhibitors, before washing with further 40 ml of Buffer 6. 2 ml of Buffer 6 were used to resuspend beads and 50 μ l of PreScission protease (Cytiva) were added to cleave the affinity tag off in an overnight incubation at 4°C. The eluate was afterwards diluted in Buffer 7 (50 mM potassium phosphate pH 7.6, 5 mM magnesium chloride, 2 mM ATP, 0.1% (v/v) triton X-100, 1 mM DTT) to reduce the potassium acetate concentration to 75 mM. 2 g of Bio-Gel HTP Hydroxyapatite (Bio-Rad) powder were hydrated with Buffer 7 supplemented with 75 mM potassium acetate and added for 15 minutes to the eluate. The resin was washed with 10 ml of Buffer 7 supplemented with 75 mM potassium acetate and 15% glycerol, followed by 10 ml of Buffer 7 supplemented with 150 mM potassium acetate and 15% glycerol. Cdc7 was eluted in Buffer 7 supplemented with 400 mM potassium acetate and 15% (v/v) glycerol. The eluate was dialysed against Buffer 8 (25 mM HEPES-KOH pH 7.6, 10 mM magnesium acetate, 0.02% (v/v) NP-40, 100 mM potassium acetate, 10% (v/v) glycerol, 5 mM β -mercaptoethanol) for about 20 hours. After dialysis, the sample was concentrated using a 30,000-MWCO Amicon Ultra centrifugal filter unit, flash frozen in liquid nitrogen and stored at –80°C.

2.6.5 *S. cerevisiae* Rad53 purification

The cell pellet of 2 l bacterial culture of the respective C-terminally 6xHis-tagged Rad53 variant were resuspended in 20 ml of Buffer 9 (25 mM HEPES-KOH pH 7.6, 300 mM sodium chloride, 0.02% (v/v) NP-40, 10% (v/v) glycerol) supplemented with protease inhibitor tablets per 1 l of cell culture. Cells were lysed by sonication and cell debris was removed by centrifugation at 58,545 $\times g$ and 4°C for 20 minutes. The lysate was supplemented with 10 mM imidazole and incubated for 1 hour at 4°C with 4 ml Ni-NTA Agarose (Qiagen; bed volume). A wash step with Buffer 9 supplemented with 10 mM imidazole was carried out, before Rad53 was eluted using 200 mM imidazole in Buffer 9. Target fractions were concentrated using a 30,000-MWCO Amicon Ultra centrifugal filter unit and subjected to size exclusion chromatography using a Superdex 200 increase 10/300 gel filtration column (Cytiva) equilibrated in Buffer 9. Peak fractions were pooled, concentrated, flash frozen and stored at –80°C.

2.6.6 *S. cerevisiae* histone octamer purification

The cell pellet from 12 l bacterial culture was thawed on ice and resuspended in 30 ml of Buffer 10 (20 mM HEPES pH 8.0, 0.1 mM EDTA, 10 mM β -mercaptoethanol) supplemented with 100 mM sodium chloride, protease inhibitors and 0.2 mg/ml lysozyme. Cell lysis was achieved by sonication and the soluble fraction was separated by centrifugation at 58,545 $\times g$ and 4°C for 20 minutes. The lysate was cleared from small particles using a 0.45 μ m syringe filter, before subjecting it to a 5 ml HiTrap heparin anion exchange chromatography column (Cytiva). The column was washed with Buffer 10 supplemented with 500 mM sodium chloride until the UV reading reached baseline. Histone octamers were eluted using a 16 CV gradient from 0.5 to 2 M sodium chloride in Buffer 10. Fractions containing all four histones were concentrated using a 30,000-MWCO Amicon Ultra centrifugal filter unit and separated on a Superdex 200 16/600 gel filtration column (Cytiva) in Buffer 10 supplemented with 2 M sodium chloride. Fractions containing stoichiometric amounts of all H2A/H2B/H3/H4 were pooled and concentrated to 2 mg/ml. For short-term storage, histone octamers were kept at 4°C, while long-term storage was achieved by adding glycerol in a final concentration of 50% (v/v) and flash-freezing in liquid nitrogen. Before nucleosome reconstitution, glycerol was removed by repeating the gel filtration using a Superdex 200 increase 10/300 gel filtration column (Cytiva).

2.7 Preparation of origin DNA templates

2.7.1 Amplification and purification of DNA templates

A 168 base pair linear DNA construct containing the *S. cerevisiae* *ARS1* origin, flanked by *HpaII* methyltransferases (*M.HpaII*, NEB), was used for all experiments with *S. cerevisiae* proteins. To permit reversible pulldown assays, a desthiobiotin–triethyleneglycol (TEG) moiety was added to the 5' *ARS1* end, using a chemically modified oligonucleotide (Integrated DNA Technologies) for PCR amplification of the DNA template. A synthetic origin containing two yeast-specific high-affinity ORC binding sites with 70 base pairs in between flanked by Widom 601 and

Widom 603 nucleosome positioning sequences on either side (nucleosome–Gid70–nucleosome) was used for *H. sapiens* origin licensing reactions. Widom sequences have a repeating motif of A/T-rich and G/C-rich elements, promoting efficient positioning of nucleosomes (Lowary and Widom 1998).

Both DNA templates were amplified by PCR using primers specified in Table 2.2. For the *M.HpaII*–*ARS1*–*M.HpaII* DNA construct, primers were engineered with the suicide substrate 5-fluoro-2'-deoxycytosine (C*) in the *M.HpaII* recognition sequence (CC*GG) to entrap *M.HpaII* on the DNA.

For large-scale amplification, a total of 4 ml reactions were set up for each construct in Buffer 11 (20 mM Tris pH 8.6, 10 mM potassium chloride, 10 mM ammonium sulfate, 2 mM magnesium chloride, 0.1% (v/v) triton X-100, 0.1 mg/ml BSA), using 0.25 mM dNTP mix, 0.5 µM of each primer (see Table 2.2), 10 ng template DNA (see Table 2.3) and 100 units of Pfu polymerase. A 3-step PCR was carried out using the following protocol: First, the polymerase was activated by incubation at 95°C for 2 minutes. This was followed by 35 cycles of DNA denaturation at 95°C for 25 seconds, primer annealing at 56°C (*M.HpaII*–*ARS1*–*M.HpaII* construct) / 64°C (nucleosome–Gid70–nucleosome construct) for 20 seconds and elongation at 70°C for 10 seconds. Third, DNA replication was finalised by incubation at 70°C for 30 seconds.

PCR products were purified by anion exchange chromatography using a 1 ml Resource Q column (Cytiva) with a gradient from Buffer 12 (50 mM Tris pH 8.0, 5 mM β-mercaptoethanol) to Buffer 12 supplemented with 2 M sodium chloride. DNA was concentrated by ethanol precipitation and resuspended in TE buffer.

2.7.2 Preparation of the *M.HpaII*–*ARS1*–*M.HpaII* DNA template

To conjugate methyltransferase to the *ARS1*-containing DNA template, *M.HpaII* and DNA were mixed in a 3:1 molar ratio and incubated overnight at 30°C in Buffer 13 (50 mM potassium acetate, 25 mM Tris pH 7.5, 10 mM magnesium acetate, 1 mg/ml BSA, 150 µM S-adenosyl-methionine (NEB)). Products were isolated using

40 CV gradient from 0 to 2 M sodium chloride in Buffer 12 on a 1-ml Resource Q column (Cytiva). Peak fractions were pooled, concentrated on a 30,000-MWCO Amicon Ultra centrifugal filter unit. Conjugation efficiency was assessed by separation on a 2% agarose gel, 4% native PAGE or by SDS-PAGE on a 3–8% Criterion™ XT Tris-Acetate Protein Gel (Bio-Rad). Aliquots were flash frozen in liquid nitrogen and stored at –80°C.

2.7.3 Preparation of the nucleosome–Gid70–nucleosome DNA template

Dr Thomas Miller reconstituted the nucleosome–Gid70–nucleosome DNA template that was used for *H. sapiens* origin licensing experiments. The protocol for nucleosome assembly was adapted from Luger et al. (1999). Briefly, histone octamers and the DNA template were mixed in a 2:1 molar ratio in Buffer 14 (20 mM HEPES pH7.6, 5 mM β-mercaptoethanol) supplemented with 2 M sodium chloride, with a final DNA concentration of 0.5–0.7 mg/ml. The mixture was dialysed against 100 ml of Buffer 14 supplemented with 2 M sodium chloride and the salt concentration was gradually reduced by addition of 1.9 l of Buffer 14 supplemented with 50 mM sodium chloride over a course of approximately 16 hours. Finally, refolded nucleosomes were dialysed for two hours against Buffer 15 (20 mM HEPES pH 7.6, 80 mM potassium chloride, 100 mM sodium acetate, 1 mM DTT). Successful nucleosome assembly was assessed by 4% native PAGE in 0.5x TAE buffer (20 mM Tris base, 10 mM acetic acid, 0.5 mM EDTA pH 8.0).

2.8 *In vitro* reconstitution

2.8.1 *S. cerevisiae* DH loading and phosphorylation analysed by SDS-PAGE

5.5 nM ORC was incubated with 2.25 nM of the *M.HpaII*–ARS1–*M.HpaII* DNA template and 5.5 nM Cdc6 for 10 minutes at 30°C with mixing at 1,250 rpm in Buffer A (25 mM HEPES-KOH pH 7.6, 100 mM potassium acetate, 10 mM magnesium acetate, 0.02% (v/v) NP-40, 5% (v/v) glycerol, 1 mM DTT, 5 mM ATP). Mcm2–7–Cdt1 was added to a final concentration of 7 nM, bringing the total

volume of the reaction to 40 μ l. The incubation was continued for 30 minutes. DNA-bound complexes were coupled to 2 μ l of M280 streptavidin paramagnetic beads (Invitrogen) *via* the desthiobiotin–TEG moiety on the DNA for 30 minutes. Beads were washed three times with high-salt buffer (25 mM HEPES-KOH pH 7.6, 500 mM NaCl, 5 mM magnesium acetate, 0.02% (v/v) NP-40), twice with Buffer A lacking ATP and then resuspended in 40 μ l of Buffer A. DDK (wild type or variant) was added at the indicated concentration. Phosphorylation was carried out for 15 minutes at 30°C. Beads were washed with low-salt buffer (25 mM HEPES-KOH pH 7.6, 300 mM sodium acetate, 5 mM magnesium acetate, 0.02% (v/v) NP-40) unless indicated differently. 1,000 units of micrococcal nuclease (MNase, NEB) were used to elute protein complexes within 10 minutes at 37°C. Eluates were separated by SDS-PAGE and visualised by silver staining (SilverQuest Silver Stain, Bio-Rad).

For time course experiments, DNA-loaded MCM double hexamers were reconstituted and isolated as described above. DDK was added to a final concentration of 3 nM. At indicated timepoints, an equal volume of STOP buffer (50 mM HEPES-KOH pH 7.5, 160 mM EDTA, 0.02% (v/v) NP-40) was added to prevent further phosphorylation. Beads were then washed with low-salt buffer, protein–DNA complexes were eluted using MNase and analysed by SDS-PAGE and silver staining.

2.8.2 Reconstitution of *S. cerevisiae* DH–DDK for negative stain EM

To analyse the binding of wild type or mutant DDK to MCM DH, 100 nM of ORC, Cdc6 and Mcm2–7–Cdt1 were mixed with 50 nM of the M.*HpaII*–ARS1–M.*HpaII* DNA template in a total volume of 10 μ l of Buffer B (25 mM HEPES-KOH pH 7.6, 100 mM potassium acetate, 10 mM magnesium acetate, 1 mM DTT, 5 mM ATP). After 30 minutes at 30°C and 1,250 rpm shaking, 150 nM DDK was added. 15 minutes later, reactions were diluted to a final protein concentration of 20 ng/ μ l and imaged using negative stain EM as described below.

2.8.3 Reconstitution of *S. cerevisiae* DH-DDK for cryo-EM

A larger phosphorylation reaction was prepared to allow freezing of multiple cryo-EM grids. Here, 96 nM ORC, Cdc6 and 48 nM *M.HpaII*–*ARS1*–*M.HpaII* DNA were incubated in 45 µl Buffer B (25 mM HEPES-KOH pH 7.6, 100 mM potassium acetate, 10 mM magnesium acetate, 1 mM DTT, 5 mM ATP), with mixing at 1,250 rpm for 10 minutes at 30°C. Mcm2–7–Cdt1 was added to a final concentration of 96 nM for 60 minutes, followed by DDK at 289 nM. After 50 minutes, the reaction was diluted 1:2 and used for cryo-EM grid preparation.

2.8.4 Inhibition of DH phosphorylation by Rad53 analysed by negative stain EM

For each condition, DH were loaded onto *M.HpaII*–*ARS1*–*M.HpaII* DNA by incubating 100 nM DNA with 200 nM of ORC, Cdc6 and Mcm2–7–Cdt1 in a total volume of 5 µl of Buffer B. The reaction was carried out for 30 minutes at 30°C, shaking at 1,250 rpm. In parallel, 300 nM DDK was co-incubated in an equimolar ratio with wild type or catalytically dead Rad53 in 7.5 µl of Buffer B. After 15 minutes, 5 µl of the DDK–Rad53 mixture and 5 µl of the MCM loading reaction were mixed and left to react for 15 minutes. Samples were diluted to a total protein concentration of 20 ng/µl and analysed by negative stain EM.

Phosphorylation efficiency was assessed using DNA-affinity purification. Here, 2 µl of the reaction were bound to 2 µl of M280 streptavidin paramagnetic beads (Invitrogen) and 15 µl of Buffer A lacking ATP. After 30 minutes at 30°C and mixing at 1,250 rpm, beads were washed twice with low-salt buffer and subjected to DNA digestion by 1,000 units of MNase (NEB) at 37°C for 10 minutes. Samples were separated by SDS-PAGE using 3–8% Criterion™ XT Tris-Acetate Protein Gels and visualised by silver staining (SilverQuest Silver Stain, Bio-Rad).

2.8.5 Reconstitution of *H. sapiens* origin licensing for cryo-EM

H. sapiens DH formation reactions were carried out in collaboration with Dr Florian Weissmann. First, 45 nM of the nucleosome–Gid70–nucleosome DNA template were mixed with 120 nM ORC1–5, 120 nM ORC6, 150 nM CDC6, 150 nM CDT1 and 60 nM MCM2–7 in a total volume of 35 μ l Buffer C (25 mM HEPES-KOH pH 7.6, 100 mM potassium glutamate, 10 mM magnesium acetate, 1 mM DTT, 2 mM ATP). After 30 minutes at 37°C and 1,250 rpm constant mixing, the undiluted reaction was used for cryo-EM grid preparation.

To capture early loading intermediates the reaction was set up as described above, with two exceptions. *i.* ATP γ S was used instead of ATP (to capture DNA loading intermediates that depend on DNA binding but not necessarily hydrolysis) and *ii.* DNA concentration was increased to 70 nM to reduce the number of MCM particles bound to the same DNA template.

2.8.6 Purification of the DNA-bound *S. cerevisiae* DH using streptavidin-coated magnetic beads

110 nM ORC was incubated with 45 nM of the *M.HpaII*–ARS1–*M.HpaII* DNA template (containing a desthiobiotin–TEG at the 5' end) and 110 nM Cdc6 for 10 minutes at 30°C with mixing at 1,250 rpm in Buffer A. Mcm2–7–Cdt1 was added to a final concentration of 140 nM, bringing the total volume to 40 μ l. The incubation was continued for 30 minutes. DNA-bound complexes were coupled to 4 μ l of M280 streptavidin paramagnetic beads (Invitrogen) *via* the desthiobiotin–TEG on the DNA for 30 minutes. Beads were washed three times with high-salt buffer, once with low-salt buffer and then resuspended in 20 μ l of low-salt buffer, which was supplemented with 2 mM biotin. Beads were incubated in this buffer for 60 minutes at 30°C to elute DNA-bound complexes. Samples were then analysed using negative stain EM.

2.8.7 Purification of the DNA-bound *S. cerevisiae* DH by anion-exchange chromatography

The MCM loading reaction was carried out as described in paragraph 2.8.6, except that the reaction volume was increased to 400 μ l, while keeping all concentrations the same. Furthermore, incubation with Mcm2–7–Cdt1 was extended to 60 minutes. The entire reaction was then subjected to a MonoQ 1.6/5 anion exchange column. A 40 CV gradient from Buffer D (25 mM HEPES-KOH pH 7.6, 100 mM potassium acetate, 10 mM magnesium acetate, 1 mM DTT) to Buffer E (25 mM HEPES-KOH pH 7.6, 2 M sodium chloride, 10 mM magnesium acetate, 1 mM DTT) was used for sample separation. Undiluted peak fractions were analysed by negative stain EM.

2.9 Peptide phosphorylation assay

The protocol for peptide phosphorylation by DDK was adapted from previously reported studies (Hughes et al. 2010, Hughes et al. 2012, Dick et al. 2020). Two N-terminally biotinylated peptides containing residues 35–47 of *H. sapiens* MCM2 (³⁵TDALTSSPGRDLP⁴⁷) were used; one peptide was phosphorylated at Ser41, one was unphosphorylated. Kinase activity is greatly enhanced by phosphorylation of the site preceding (P+1) the DDK target site as shown before (Hughes et al. 2012). For each experiment, 27.5 μ g peptide was incubated with 2.8 nM DDK and 16.5 μ Ci of [γ -³²P]ATP (5,000 Ci/mmol stock concentration) in 137.5 μ l of kinase assay reaction buffer (40 mM HEPES-KOH pH 7.6, 10 mM magnesium acetate, 2 mM DTT, 0.1% (v/v) NP-40, 80 μ g/ml BSA, 1 mM β -glycerophosphate, 1 mM sodium fluoride, 0.1 mM ATP). Reactions were incubated at 30°C. 25 μ l samples of the reaction were removed at indicated timepoints (5, 15, 45, 60, 90 minutes). Further phosphorylation was inhibited by denaturing DDK with the addition of guanidine hydrochloride to a final concentration of 2.5 M. 12.5 μ l of each sample were spotted onto SAM2 biotin-capture membranes (Promega). Unbound components were removed by washing the membrane three times with 100 ml of 2 M sodium chloride for 2 minutes, four times with 2 M sodium chloride in PBS, twice with distilled water

and once with 95% (v/v) ethanol for 15 seconds. After air-drying for 10–15 minutes, radioactivity was detected by phosphorescence imaging with a Typhoon FLA 9500 (Cytiva).

Experiments analysing kinase activity after DDK autophosphorylation were carried out by incubating 50 nM DDK in 20 μ l Buffer A for 15 minutes at 30°C. In the negative control condition ATP was omitted. The pre-treated DDK was used in the peptide phosphorylation assay as described above.

To test for the effect of Rad53 on peptide phosphorylation by DDK, DDK and Rad53 were mixed at equimolar ratio (50 nM) in 20 μ l Buffer A. After 15 minutes at 30°C, the protein mixture was added to the peptide and [γ -³²P]ATP as above.

2.10 Mcm2–7–Cdt1 phosphorylation by DDK

1 μ M DDK was either incubated for 15 minutes at 30°C with mixing at 1,250 rpm in 4 μ l Buffer B to induce autophosphorylation or in Buffer B lacking ATP. Buffer B and Mcm2–7–Cdt1 were added, thereby doubling the reaction volume and resulting in final concentrations of 25 nM for Mcm2–7–Cdt1 and 500 nM for DDK. After 15 minutes, phosphorylation reactions were diluted with 10 μ l water, terminated by the addition of 6 μ l 5x Laemmli sample buffer and denaturation at 97°C for 10 minutes. Samples were separated by SDS-PAGE using 3–8% Criterion™ XT Tris-Acetate Protein Gels (Bio-Rad). Proteins were visualised by silver staining (SilverQuest Silver Stain, Bio-Rad).

2.11 Mass spectrometry analysis of MCM phosphorylation

Phosphorylation of both Mcm2–7–Cdt1 and DNA-loaded DH by DDK was analysed by mass spectrometry. For the former, 100 nM Mcm2–7–Cdt1 was treated with 2

μ M wild type DDK in 15 μ l of Buffer A. The reaction was incubated for 30 minutes at 30°C and mixing at 1,250 rpm.

To analyse DH phosphorylation, 2.25 nM DNA, 5.5 nM ORC and 5.5 nM Cdc6 were incubated in 1 ml of Buffer A for 10 minutes at 30°C with mixing at 1,250 rpm. Mcm2–7–Cdt1 was added to a final concentration of 13 nM. After 30 minutes, DNA-bound complexes were coupled to 42 μ l of M280 streptavidin paramagnetic beads (Invitrogen; volume of the suspension) by mixing for 30 minutes. Beads were washed three times with 1 ml of high-salt buffer and once with low-salt buffer. For elution, the DNA was digested by 4,000 units of MNase (NEB) for 15 minutes at 37°C.

Both samples were run 7 mm into a 12% SDS-PAGE gel and visualised by InstantBlue™ Protein Stain (Expedeon). Gel slices containing all protein-containing bands were treated with 1 ml of extraction solvent (50% (v/v) acetonitrile, 100 mM ammonium bicarbonate, 5 mM DTT) overnight at 4°C to remove the Coomassie protein stain. Subsequently, Dr Steven Howell from the Proteomics Science Technology Platform at the Francis Crick Institute carried out in-gel digestion with trypsin and analysed samples by liquid chromatography-tandem mass spectrometry. A 44-minute binary gradient with an Evosep nanoHPLC coupled to an Orbitrap Lumos Tribrid mass spectrometer (Thermo Scientific) was used. Peptides were fragmented by higher-energy collisional dissociation and detected in the ion trap using the vendor's 'universal' data-dependent acquisition method. Maxquant (<https://www.maxquant.org/>) was used to search data against the uniprot *S. cerevisiae* FASTA database. Perseus (<https://maxquant.net/perseus/>) was used for visualisation. I analysed the data in collaboration with Dr Steven Howell.

2.12 Electron microscopy sample preparation

2.12.1 Negative stain grid preparation

Two types of grids were used interchangeably: 300-mesh copper grids with a thin continuous layer of carbon (EM Resolutions, C300Cu100) and 400-mesh copper carbon grids (Agar Scientific). Grids were glow-discharged at 45 mA for 30 seconds using a 100x glow discharger (EMS) or at 25 mA for 1 minute using a GloQube® Plus Glow Discharge System (Quorum) with equivalent results. 4 µl of the sample was applied to freshly glow-discharged grids for 1 minute for *S. cerevisiae* DH-DDK reactions and 2 minutes for *H. sapiens* origin licensing reactions.

For experiments on *S. cerevisiae* DH-DDK, grids were stained by gently stirring grids on four 30 µl droplets of 2% (w/v) uranyl acetate solution for 10 seconds each before blotting away excess stain. Negative stain grids of *S. cerevisiae* DDK were prepared in the same manner, applying 4 µl of the isolated protein at a concentration of 18 ng/µl.

A modified staining protocol was used for later experiments on *H. sapiens* origin licensing reactions to reduce the required amount of uranyl acetate solution. For this procedure, most of the sample was removed by blotting with a filter paper and replaced by 4 µl of 2% (w/v) uranyl acetate solution. The staining solution was quickly removed, equivalent to a quick washing step. The final staining was achieved by the addition of another droplet of 4 µl 2% (w/v) uranyl acetate solution and incubation for 40 seconds. Excess stain was removed using filter paper.

2.12.2 Deposition of graphene oxide on cryo-EM grids

10 µl graphene oxide flake dispersion (Sigma) was diluted in 80 µl water and aggregates were removed by centrifugation at 500 *xg* for 1 minute. UltrAuFoil R1.2/1.3 300-mesh grids (Quantifoil) were glow discharged at 40 mM for 5 minutes using a GloQube® Plus Glow Discharge System (Quorum), before applying 3 µl graphene oxide dispersion. After three minutes, excess liquid was removed using

filter paper. The grids were washed by gently picking up three 20 µl droplets of water (two on the front and one on the back of the grid) with blotting away liquid in between each step. Grids were left to dry at room temperature for 1-2 hours.

2.12.3 Cryo-EM grid preparation for *S. cerevisiae* DDK

4 µl of 40 ng/µl DDK applied to UltrAuFoil R1.2/1.3 300-mesh grids (Quantifoil) coated with graphene oxide (see section 2.12.2). After 30 seconds at room temperature and 90% humidity, excess sample was removed by double-side blotting for 4 seconds and a blot force of 0 using a Vitrobot Mark IV (FEI). Grids were immediately plunge frozen in liquid ethane.

2.12.4 Cryo-EM grid preparation for *S. cerevisiae* DH-DDK

Lacey grids (400 mesh) with an ultrathin layer of carbon (Agar Scientific) were glow discharged for 1 minute at 45 mA using a 100x glow discharger (EMS). 4 µl of the 1:2 diluted MCM loading and phosphorylation reaction were applied for 30 seconds at room temperature and 90% humidity. Excess liquid was removed by double-side blotting with blot force -1 for 3 seconds and grids were plunge frozen in liquid ethane using a Vitrobot Mark IV (FEI).

2.12.5 Cryo-EM grid preparation for *H. sapiens* origin licensing

4 µl of the undiluted (ATP) / 3:1 diluted (ATPyS) reaction were applied to UltrAuFoil R1.2/1.3 300-mesh grids (Quantifoil) coated with graphene oxide (see section 2.12.2) for 60 seconds at room temperature and 90% humidity in a Vitrobot Mark IV (FEI). Grids were double-side plotted with blot force 0 for 5 seconds and immediately plunge frozen in liquid ethane.

2.13 Electron microscopy data collection

2.13.1 Negative stain data collection

A FEI Tecnai G2 Spirit transmission electron microscope operated at 120 keV, equipped with a 2K x 2K GATAN UltraScan 1000 CCD camera, was used to collect negative-stain micrographs. Images were collected at a nominal magnification of 30,000x, yielding a pixel size of 3.45 Å at the specimen level, and a defocus range of -0.6 to -1.4 µm.

2.13.2 *S. cerevisiae* DDK cryo-EM data collection

Data were acquired in counting mode on a Titan Krios transmission electron microscope, operated at 300 keV and equipped with a K2 Summit direct electron detector and a BioQuantum energy filter (Gatan Inc.). 9,469 movies with 28 frames/movie were collected at 165,000x magnification, yielding a pixel size of 0.84 Å at specimen level. The total electron dose was 62.2 e⁻/Å². A defocus range of -2.0 to -4.5 µm was used.

2.13.3 *S. cerevisiae* DH-DDK cryo-EM data collection

18,135 movies with 30 frames/movie were collected on a Titan Krios transmission electron microscope operated at 300 keV with a total dose of 51.3 e⁻/Å². The K2 Summit direct electron detector (Gatan Inc.) was operated in counting mode together with a BioQuantum energy filter. Images were recorded with 130,000x nominal magnification, yielding a pixel size of 1.08 Å. Defocus values between -2.0 and -4.1 µm were applied. Further collection details are found in Appendix **Error! Reference source not found.**

2.13.4 *H. sapiens* origin licensing cryo-EM data collection

Datasets for *H. sapiens* origin licensing in the presence of ATP_γS and ATP, respectively, were imaged on a Titan Krios transmission electron microscope with a

K2 Summit direct electron detector and BioQuantum energy filter. Data were collected in counting mode with 32 frames/movie, a total electron dose of $49.28 \text{ e}^-/\text{\AA}^2$, and a defocus range from -1.0 to -2.5 μm . A total of 7,227 movies were collected for the reaction in the presence of ATP and a total of 31,569 movies for the reaction in the presence of ATP γ S.

2.14 Electron microscopy image processing

2.14.1 Negative stain image processing

All negative-negative stain data were processed using RELION-3.1 (Zivanov et al. 2020). Gctf v1.06 (Zhang 2016) was used to estimate contrast transfer function (CTF) parameters. Particles were picked semi-automatically using Topaz v0.2.5 (Bepler et al. 2019), extracted with a 128-pixel box and re-scaled to 64 pixels. Particles were classified with a circular, 280 $\text{\AA}$ mask and CTF correction restricted to only flipping phases. 2D class averages were grouped manually based on recognisable complexes (*i.e.* ORC, MCM loading intermediates, nucleosomes). Particles contributing to the different complexes were sub-classified, using a circular mask of 320 $\text{\AA}$ for OCCM, MO and DH and 250 $\text{\AA}$ for ORC and nucleosomes.

2.14.2 *S. cerevisiae* DDK cryo-EM image processing

Beam-induced motion was corrected for all movies using MotionCor2 with 5 x 5 patches (Zheng et al. 2017). A total of 2,967,226 particles were picked semi-automatically from dose-weighted micrographs using crYOLO v1.40 (Wagner et al. 2019). For particle picking, a box of 128 pixel and a selectivity threshold of 0.1 was chosen. Further processing was carried out in RELION-3.0.7 (Zivanov et al. 2018). CTF parameters were estimated on non-dose-weighted movie sums using Gctf v1.18 (Zhang 2016). Micrographs with anisotropic Thon rings or strong ice rings observed in Fourier space were discarded. Particles were extracted with a 240-pixel box and simultaneously rescaled to 60 pixels, yielding a pixel size of 3.36 $\text{\AA}$.

Two rounds of reference-free 2D classification with a 140-Å mask were carried out. The resulting particles (399,580) were extracted without rescaling and subjected to another round of 2D classification.

2.14.3 *S. cerevisiae* DH-DDK cryo-EM image processing

Movies were corrected for beam-induced motion using MotionCor2 with 5 x 5 patches (Zheng et al. 2017). Dose-weighted movie sums were used for semi-automatic particle picking in crYOLO v1.40 (Wagner et al. 2019), applying a box size of 200 pixel and a selectivity threshold of 0.4. CTF parameters were determined for each micrograph using Gctf v1.18 (Zhang 2016) and micrographs with anisotropic Thon-rings or strong ice rings were discarded. A total of 3,529,085 particles were extracted from dose-weighted micrographs with a 480-pixel box, rescaled to 120 pixel (4.32 Å/pixel), and subjected to three rounds of reference-free 2D classification in RELION-3.1 (Zivanov et al. 2020). 294,802 DH particles were identified. A 3D reference was generated by ab initio reconstruction in cryoSPARC v2.8 (Punjani et al. 2017), followed by homogenous refinement. The reference was filtered to 60 Å and used for 3D classification in RELION. To obtain all well-behaved DH particles from this dataset, three independent classifications were carried out in parallel with three classes each. The first classification with a 400 Å mask yielded 220,385 particles, which refined to 4.4 Å (imposing C2 symmetry) after re-extraction at the original pixel size of 1.08 Å/pixel. The remaining 74,417 particles were subjected to reference-free 2D classification, resulting in 3,673 well-resolved DH particles that were added to the previously selected particles. The second and third 3D classification used a 350 Å mask and refinement imposing C1 and C2 symmetry, respectively. Particles contributing to recognisable DH 3D classes were selected. Particle sets of all selections were compared and all unique particles were combined, resulting in a total of 238,620 DH particles. The combined particle set was refined to 3.8 Å using a mask of 400 Å and imposing C2 symmetry. The refinement was followed by two rounds of CTF refinement (Zivanov et al. 2020), correcting for per-micrograph and per-particle defocus, respectively. Particles were then subjected to Bayesian polishing (Zivanov et al. 2019) and six rounds of CTF refinement (correcting for 1. per-particle defocus plus beam tilt; 2.

per-particle defocus plus per-micrograph B-factor and beam tilt; 3. per-particle defocus plus astigmatism, B-factor, beam tilt and trefoil; 4. anisotropic magnification). 4,472 particles that were closer than 160 pixels to the edge of a micrograph were removed. The cleaned particle set was subjected to another round of Bayesian polishing and CTF refinement (correcting for per-particle defocus, astigmatism, B-factor, beam tilt, trefoil and 4th order aberrations). This resulted in a structure of 3.0 Å resolution (imposing C2 symmetry). Density modification was carried out using ResolveCryoEM (Terwilliger et al. 2020) in Phenix v1.19.2 (Liebschner et al. 2019).

Symmetry expansion in RELION was used to improve the resolution of DDK in complex with the DH. Thereby, the set of particles contributing the 3.0-Å resolution DH structure was duplicated and simultaneously rotated around the C2 symmetry axis. The signal of the DH and one of the DDK molecules was removed and particles were separated by focused 3D classification without alignment, using a mask around the other DDK molecule and the regularisation parameter T set to 20. 149,876 particles with signal for DDK were recovered. Signal subtraction of the DH was reverted, while the signal of the second DDK molecule was removed again. The DH-DDK structure was refined to 3.3 Å average resolution.

To improve the Mcm2 interacting part of DDK, signal of the flexible MCM ATPase tiers and the second DDK was subtracted from symmetry-expanded particles. A mask encompassing the entire DDK (catalytic core and Dbf4 BRCT domain) was used for focused 3D classification without alignment (T=20). 128,200 DDK particles in complex with the N-terminal tiers of the MCM were refined, resulting in an average resolution of 3.4 Å. Interpretability of the Mcm2-bound density of DDK was further enhanced by LAFTER filtering (Ramlal et al. 2019). For visualisation purposes, the filtered Mcm2-bound density (Dbf4 BRCT domain) was isolated and combined with the DH-DDK map within UCSF Chimera v1.14 (Pettersen et al. 2004).

Multi-body refinement in RELION (Nakane et al. 2018) with soft masks around the entire DDK and the DH (extend: 15 pixels, soft-edge: 12 pixels) was used to assess relative movement of the two subcomplexes. The standard deviation of a Gaussian

prior on the Euler angles was set to 10 and the translational offsets for both bodies was set to 2 pixels.

2.14.4 Cryo-EM image processing for *H. sapiens* origin licensing in the presence of ATP

RELION-4.0b-GPU (Kimanius et al. 2021) and cryoSPARC v3.3.2 were used for image processing at different stages of the processing pipeline as indicated. Movies were corrected for beam-induced motion using RELION's own implementation with 5 x 5 patches. Subsequent processing was carried out within the cryoSPARC software. CTF parameters of motion corrected micrographs were estimated using CTFFIND v4.1.10 (Rohou and Grigorieff 2015). Particles were picked semi-automatically using Topaz v0.2.4 (Bepler et al. 2019), extracted with a 440-pixel box size, and re-scaled to 110 pixel. Micrographs were curated based on a CTF fit resolution of 2.57–4.50 Å, CTF fit cross-correlation of 0.07–0.27 and median pick score of 20.18–43.55, resulting in a total of 3,589 micrographs with 970,326 particles. Three rounds of reference-free 2D classification were carried out, and 213,807 particles were used for ab initio reconstruction and subsequent heterogeneous refinement with four classes. The resulting 49,485 DH particles were re-extracted without downscaling and subjected to two rounds of 2D classification. 19,049 particles, which contributed to high-resolution DH 2D class averages, were used for ab initio reconstruction. The structure was refined (homogeneous, non-uniform and local) to 3.1 Å, imposing C2 symmetry.

DH particles were re-extracted and refined to 3.5 Å in RELION. Duplicate particles were removed and particles were then re-grouped to improve signal-to-noise-ratio during subsequent alignments. 3D classification without alignment (320 Å mask, regularisation parameter T=4) was used to isolate 15,874 DH particles. Particles were subjected to Bayesian polishing (Zivanov et al. 2019), followed by 2D classification without alignment and 3D refinement imposing C2 symmetry. Three rounds of CTF refinement (Zivanov et al. 2020) were carried out (1. per-particle defocus, per-micrograph astigmatism; 2. per-particle defocus, per-particle astigmatism, beamtilt; 3. per-particle defocus, per-particle astigmatism, beamtilt,

trefoil, 4th order aberrations), resulting in a 3.3-Å resolution DH structure. The same particle set was refined to 3.1 Å in cryoSPARC using homogeneous, followed by non-uniform refinement.

The same dataset also contained single MCM hexamers (SH) that encircled duplex DNA. Some density was observed for ORC that was bound to the N-terminal side of the MCM ring (so called MO complex). 60,711 SH particles (mixed with some MO particles) were isolated by heterogeneous refinement in cryoSPARC.

Heterogeneous refinement was then used to separate 37,396 SH from 9,355 MO. The SH structure was further refined (homogeneous, non-uniform, local) to 3.4 Å. The resulting set of particles was re-extracted and refined in RELION, giving rise to a 4.1-Å SH structure. Particles were subjected to CTF refinement (per-particle defocus, per-micrograph astigmatism), Bayesian polishing and another round of CTF refinement (per-particle defocus, per-particle astigmatism, beamtilt), which improved the resolution to 3.5 Å. 3D classification identified 25,069 high-resolution SH particles, which refined to 3.4 Å. The same particle set was further refined in cryoSPARC (homogeneous, non-uniform, local) bringing the resolution to 3.2 Å.

2.14.5 Cryo-EM image processing for initial stages of *H. sapiens* MCM loading in the presence of ATP γ S

Beam-induced motion was carried out using RELION's implementation with 5 x 5 patches in RELION-4.0b-GPU (Kimanius et al. 2021). CTFFIND v4.1.13 (Rohou and Grigorieff 2015) was used to estimate CTF parameters on non-dose weighted micrographs. A total of 1,334,277 particles were picked from motion corrected micrographs by Topaz v0.2.4 (Bepler et al. 2019). Two rounds of reference-free 2D classification resulted in a set of 629,241 MCM-containing particles (i. e. OCCM, SH, MO). To separate the different complexes, particles were subjected to 3D classification using 3D references for SH, OCCM, MO and a spherical volume low-pass filtered to 20 Å. This gave rise to 114,995 SH (mixed with some MO), 170,792 OCCM and 203,088 MO particles. The particle sets of the different complexes were re-extracted using a 400-pixel box without rescaling. Particles were refined individually in cryoSPARC (homogeneous, non-uniform, local) to 3.6 Å for SH, 4.0

Å for OCCM and 3.7 Å for MO. Each structure was then subjected to Bayesian polishing in RELION (Zivanov et al. 2019).

For OCCM, three rounds of CTF refinement (1. per-particle defocus, per-micrograph astigmatism; 2. anisotropic magnification; 3. per-particle defocus, per-particle astigmatism, beamtilt) followed by another round of Bayesian polishing were carried out, resulting in a final resolution of 4.2 Å. Focused 3D classification without alignment, imposing a mask around CDC6 and ORC1 allowed to isolate 100,567 particles that had clearly defined density for the two components. This set of particles was refined in RELION and cryoSPARC. Local refinement with a soft mask around the ORC–CDC6 subcomplex was then carried out. 3D classification without alignment in cryoSPARC allowed to isolate 34,116 OCCM particles with well resolved CDC6 density and 49,771 particles that lacked CDC6. Both complexes were subjected to local refinement using a mask encompassing the entire complex, resulting in maps with an average resolution of 3.8 and 4.1 Å, respectively.

As SH and MO complexes both contain a DNA-loaded MCM hexamer, the two particle sets were refined together using a mask around the MCM, giving rise to a 3.8-Å SH structure. CTF parameters were optimised (1. per-particle defocus, per-micrograph astigmatism; 2. anisotropic magnification) and another round of Bayesian polishing was carried out, which improved the resolution to 3.7 Å. 3D classification with 3D references for MO and SH, low-pass filtered to 30 Å, separated 182,341 MO particles from 135,742 SH particles. Refinement (homogeneous, non-uniform, followed by local refinement) in cryoSPARC generated a 3.4-Å resolution map of the SH. MO particles were subjected to homogeneous refinement, followed by non-uniform refinement in cryoSPARC giving rise to a 3.6-Å resolution map. Local refinement of the MO using soft masks around MCM and ORC, respectively, resulted in a 3.5-Å SH map and a 4.0-Å ORC map.

2.15 EM model building

2.15.1 *S. cerevisiae* DH-DDK cryo-EM model building

The atomic model of the MCM DH (PDB entry 6EYC), which had been re-refined with Isolde (Croll 2018), was refined against the cryo-EM density of this study using Namdinator (Kidmose et al. 2019). Further adjustments and DNA building were carried out in Coot v0.9-pre (Casañal et al. 2020). Homology models of the DDK catalytic core were generated using HHPRED (Söding et al. 2005) and I-Tasser (Yang and Zhang 2015), based on human DDK PDB entry 6YA7 (Dick et al. 2020). Rigid-body docking within UCSF Chimera v1.14 (Pettersen et al. 2004) was used to fit the Dbf4 BRCT domain (PDB entry 3QBZ (Matthews et al. 2012)) into our cryo-EM density. The combined atomic model was real-space refined in Phenix v1.19.2 (DiMaio et al. 2013).

2.15.2 *H. sapiens* OCCM cryo-EM model building

The atomic model of *H. sapiens* OCCM was assembled by rigid-body docking using UCSF ChimeraX v1.4 (Goddard et al. 2018, Pettersen et al. 2021). First, the atomic model of *H. sapiens* ORC1–5–DNA (PDB code 7JPS) (Jaremko et al. 2020) was docked into the 3.8-Å cryo-EM map of the OCCM. Second, the model of the ORC2 WHD (residues 472–575) was extracted from the open ORC1–5 complex (PDB code 7JPR) (Jaremko et al. 2020) and docked into the map, matching the position of the homologous domain in *S. cerevisiae* OCCM (PDB code 5V8F) (Yuan et al. 2017). Third, models of the individual MCM subunits, CDT1 and CDC6 generated by AlphaFold (Evans et al. 2022) were added to the model. To fit the CDT1 model into the map, the structured domains (residues 167–387, 418–440 and 441–546) were separated and overlayed with Cdt1 in the yeast OCCM. Each domain was then rigid-body docked into the cryo-EM density. All structures were combined into one PDB file and completed by the addition of idealised B-form duplex DNA. The model was adjusted and nucleotides within the ATPase sites were added using Coot v0.9.8.1 EL (Casañal et al. 2020).

2.15.3 *H. sapiens* SH cryo-EM model building

Atomic models of the hexameric human MCM2–7 amino- and carboxy-terminal tiers were generated using AlphaFold-Multimer (Evans et al. 2022). The assemblies were docked into the 3.2-Å cryo-EM map using rigid-body docking within UCSF ChimeraX v1.4 (Goddard et al. 2018, Pettersen et al. 2021). Idealised B-form duplex DNA was built into all structures. Nucleotide molecules were modelled in the ATPase sites and the entire model was adjusted using Coot v0.9.8.1 EL (Casañal et al. 2020).

2.15.4 *H. sapiens* MO cryo-EM model building

The atomic structure of the SH (see paragraph 2.15.3) was rigid-body docked into the 3.5 Å locally refined map of the MCM using UCSF ChimeraX v1.4 (Goddard et al. 2018, Pettersen et al. 2021). The *H. sapiens* ORC1–5 complex in its open conformation (PDB code 7JPR) was docked into the 4.0 Å map of ORC after local refinement. The two TFIIB-like ORC6 domains (residues 1–94 and 95–190) were extracted from the atomic model of human ORC6 generated by AlphaFold (Varadi et al. 2022). Each domain was rigid-body docked into the cryo-EM map before combining all models into one PDB file. Coot v0.9.8.1 EL (Casañal et al. 2020) was used to adjust the structure.

2.15.5 *H. sapiens* DH cryo-EM model building

To obtain an atomic structure of the *H. sapiens* DH, the model of the SH (see paragraph 2.15.3) was rigid-body docked into both MCM densities in the DH cryo-EM map using UCSF ChimeraX v1.4 (Goddard et al. 2018, Pettersen et al. 2021). The entire structure was adjusted in Coot v0.9.8.1 EL (Casañal et al. 2020).

2.15.6 Analysis of protein–DNA contacts

To generate maps of protein–DNA contacts, the atomic structures of the different complexes were subjected to analysis on the DNAProDB server (Sagendorf et al.

2017, Sagendorf et al. 2020) and manual inspection in Coot v0.9.8.1 EL (Casañal et al. 2020).

Chapter 3. Results 1 – Origin licensing in *S. cerevisiae* and *H. sapiens*

3.1 Introduction

DNA replication has been studied extensively using the model organism *S. cerevisiae* (reviewed in Costa and Diffley (2022)). Our molecular understanding of the eukaryotic replisome has greatly improved in the last decade, thanks to *in vitro* reconstitutions recapitulating the DNA replication process in the test tube, using purified *S. cerevisiae* proteins (Evrin et al. 2009, Remus et al. 2009, Yeeles et al. 2015). These studies have allowed to identify the set of proteins essential for DNA replication and allowed us to describe the sequence of events leading to origin firing. Furthermore, structural studies using a pre-assembled CMG helicase on forked DNA substrates have provided a first structural framework to understand how the eukaryotic replicative helicase translocates along single-stranded DNA, in a process that is driven by ATP hydrolysis (Abid Ali et al. 2016, Georgescu et al. 2017, Goswami et al. 2018, Eickhoff et al. 2019, Baretić et al. 2020).

The helicase is loaded in an inactive state, in a process that is also known as origin licensing. First, the origin recognition complex (ORC) binds to origin DNA during late mitosis and throughout the G₁ phase of the cell cycle. In *S. cerevisiae*, ORC recognises specific DNA sequences, while there is no sequence specificity driving ORC recruitment in higher eukaryotes, where origins are loosely defined as a stretch of naked DNA flanked by nucleosomes (Eaton et al. 2010, Masai et al. 2010, Mechali 2010). The six-subunit, crescent shaped ORC engages with the DNA in an ATP-binding, but not hydrolysis dependent manner (Coster et al. 2014). The loading factor Cdc6 associates with ORC and completes the ring (Schmidt and Bleichert 2020, Feng et al. 2021). ORC causes the DNA to bend, which correctly positions the DNA for recruitment and subsequent loading of a first MCM helicase (Li et al. 2018). The hexameric MCM ring, which is held in an open conformation by the loading factor Cdt1, is recruited by ORC and Cdc6 *via* the C-terminal winged-helix domains of the MCM. The complex consisting of ORC, Cdc6, Cdt1 and MCM is known as OCCM (Sun et al. 2013, Yuan et al. 2017, Miller et al. 2019, Yuan et

al. 2020b). Upon ATP-hydrolysis by the MCM complex, the helicase topologically entraps double-stranded DNA in its central channel and the three loading factors are released (Ticau et al. 2015, Ticau et al. 2017). A second ORC-binding event occurs at the N-terminal side of the MCM ring, mediated by the Orc6 subunit (Miller et al. 2019). In this configuration, ORC retains the ability to bend DNA as described above. A second MCM hexamer is thereafter recruited by ORC, again involving interactions with the C-terminal face of the incoming MCM–Cdt1 complex. Thus, the elements involved in the recruitment of the first MCM ring are also required for the recruitment of the second ring (Frigola et al. 2013, Miller et al. 2019). The result of the origin licensing reaction is a head-to-head MCM double hexamer (DH) that encircles double-stranded DNA (Evrin et al. 2009, Remus et al. 2009, Li et al. 2015, Abid Ali et al. 2017, Noguchi et al. 2017). The *S. cerevisiae* DH is an inactive form of the helicase, which can neither hydrolyse ATP nor unwind DNA. The previous structures of the unmodified and phosphorylated DH provided an initial view of protein–DNA interactions as well as the nucleotides bound in the ATPase sites (Abid Ali et al. 2017, Noguchi et al. 2017). However, the structures were limited in resolution and suffered from poor DNA occupancy. To obtain a detailed understanding of how the phosphorylated DH interacts with DNA and the role of ATP binding and hydrolysis in origin licensing, I reconstituted the loading and phosphorylation reaction of the MCM helicase onto DNA, using purified *S. cerevisiae* proteins and imaged the entire reaction by cryo-EM. By determining the nucleotide occupancy in the DH ATPase sites, I describe the end point of the ATP-hydrolysis driven origin licensing reaction with implications for our understanding of helicase recruitment.

In this chapter, I also describe my efforts to understand the structural basis for human MCM loading onto double-stranded DNA. This work was carried out in collaboration with Dr Florian Weissmann from Dr John Diffley's laboratory. Our aim was to identify both conserved as well as divergent features of the helicase loading reaction across species.

3.2 *In vitro* reconstitution of origin licensing using *S. cerevisiae* proteins

3.2.1 Cryo-EM analysis of *S. cerevisiae* origin licensing

I used purified *S. cerevisiae* proteins and an *ARS1* origin DNA template to reconstitute origin licensing *in vitro* (On et al. 2014, Yeeles et al. 2015, Miller et al. 2019) and determine the structure of the phosphorylated MCM double hexamer, at a resolution that would allow identifying protein residues involved in DNA binding (Figure 3.1a). To prevent the helicase from sliding off the DNA substrate, *HpaII* methyltransferases were attached to each end of the DNA (Miller et al. 2019). The entire loading reaction was applied to lacey carbon grids, coated with a thin layer of carbon, and plunge frozen in liquid ethane, in collaboration with Dr Julia Locke in our laboratory. Data were acquired on a Titan Krios electron microscope equipped with a K2 Summit direct electron detector and a GIF Quantum energy filter. Particles were semi-automatically picked from motion corrected micrographs, which had undergone CTF estimation (Figure 3.1b). Reference-free 2D classification in RELION revealed the presence of the loading-competent MCM–Cdt1 complex, ORC bound to DNA, the loading intermediate MO and the DH (Figure 3.1c).

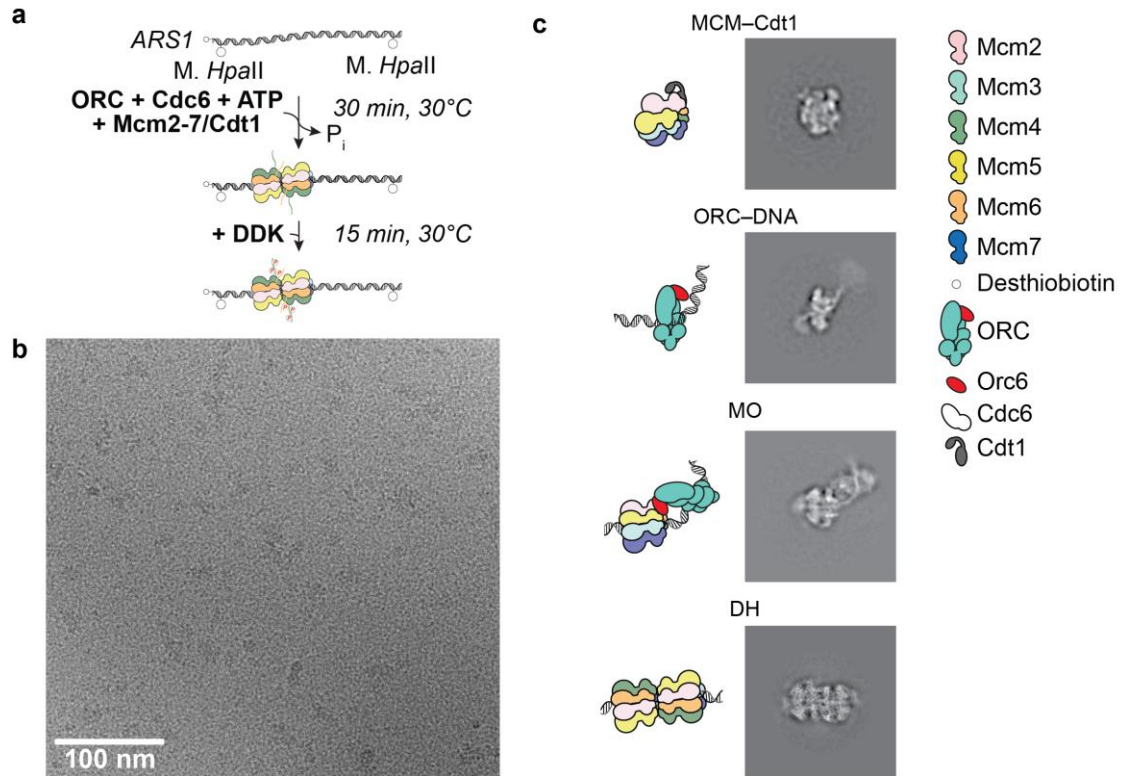


Figure 3.1: *S. cerevisiae* DH formation analysed by cryo-EM.

a. Cartoon representation of MCM double hexamer formation and phosphorylation reconstituted *in vitro* using purified *S. cerevisiae* proteins and a DNA template containing the *S. cerevisiae* origin ARS1, capped by *HpaII* methyltransferases. **b.** Representative cryo-EM micrograph of the entire MCM loading and phosphorylation reaction. **c.** 2D classes and cartoon representation of observed complexes.

In this study, I focused on the structure of the DNA-loaded DH. I therefore isolated DH particles in three rounds of reference-free 2D classification, followed by further 3D and 2D classifications (Figure 3.2). A total of 238,620 DH particles were selected and subjected to CTF refinement (Zivanov et al. 2020) and Bayesian polishing (Zivanov et al. 2019) until no further improvement was observed. The DH structure was refined to 3.1 Å, after imposing C2 symmetry (Figure 3.3). Density modification in Phenix (Liebschner et al. 2019, Terwilliger et al. 2020) improved the resolution to 3.0 Å.

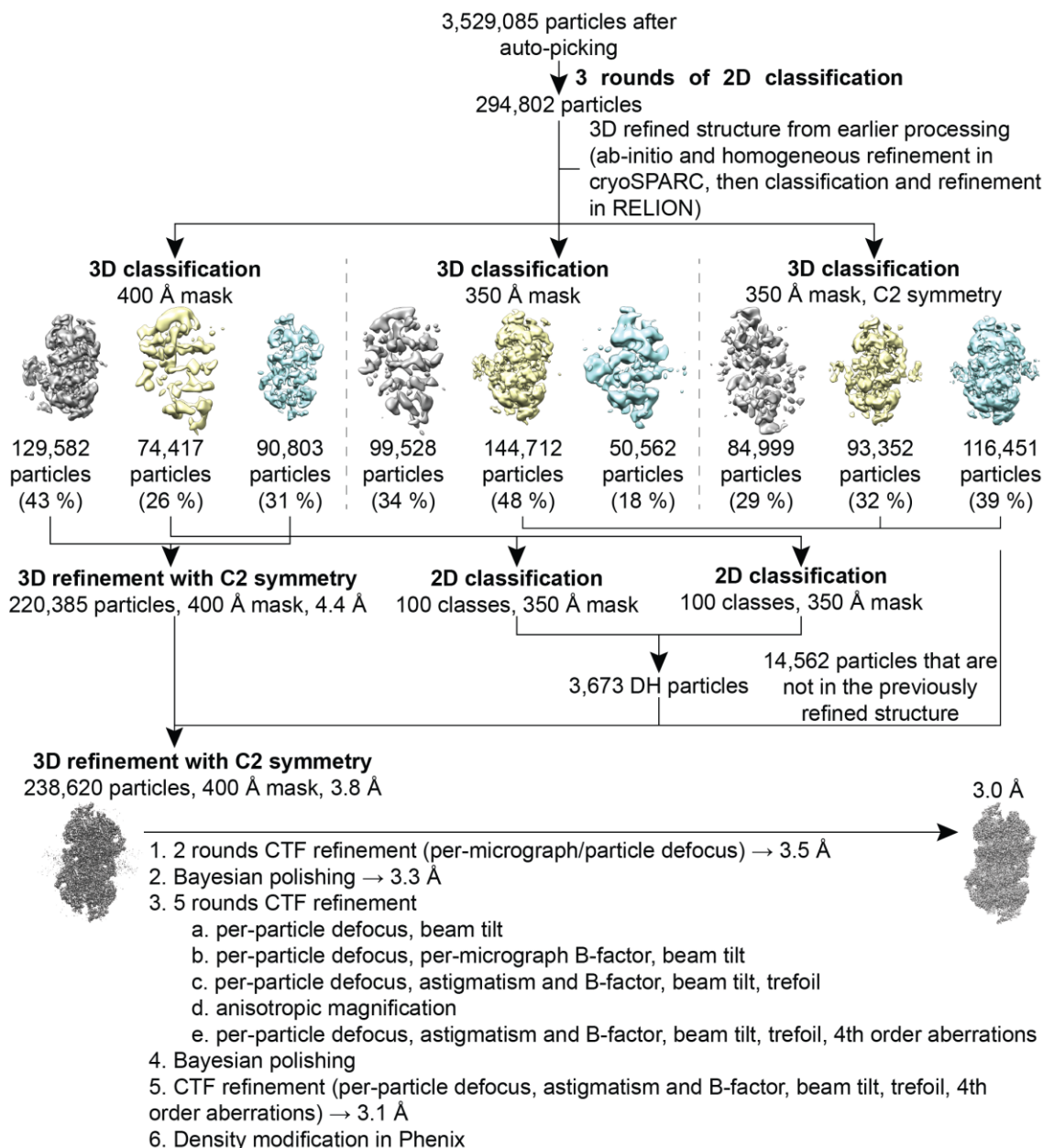


Figure 3.2: Processing pipeline for the DNA-loaded *S. cerevisiae* DH.
Overview of the image processing procedure for the *S. cerevisiae* DH.

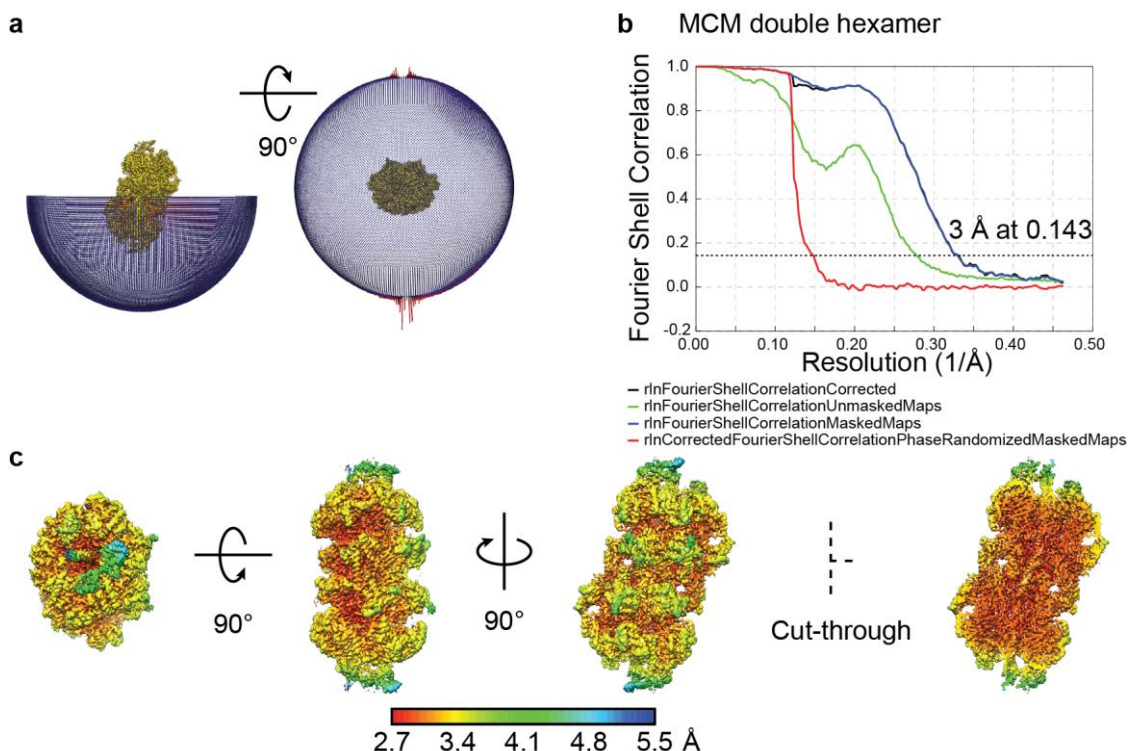


Figure 3.3: Structure quality of the DNA-loaded *S. cerevisiae* DH.

a. Angular distribution of the C2-symmetric DH. **b.** Resolution of the DH estimated using gold-standard Fourier Shell Correlation. **c.** Three rotated views and one cut-through view of the DH 3D structure, colour-coded according to the local resolution.

3.2.2 Overall architecture of the *S. cerevisiae* MCM DH

To assess the molecular details, I refined an atomic structure based on the new cryo-EM map. To achieve this, the atomic model of the MCM DH (PDB entry 6EYC) was docked into the cryo-EM density and refined using Namdinator (Kidmose et al. 2019), Coot v0.9-pre (Casañal et al. 2020) and real-space refinement in Phenix v1.19.2 (Liebschner et al. 2019) ([Appendix Table 1](#)).

As observed in previous studies, the two MCM hexamers dimerise *via* their N-terminal domains, forming a head-to-head double hexamer (Li et al. 2015, Abid Ali et al. 2017, Noguchi et al. 2017). The C-terminal AAA+ tiers cap the complex from both sides. In this structure, we also observe density for the winged helix domains (WHD) of Mcm5, Mcm6 and Mcm7, which are on top of the AAA+ tiers. The two MCM hexamers interlock through interactions between their Zn finger domains in the NTD. The MCM DH is further stabilised by a helix-turn-helix motif in the Mcm7

is consistent with the notion that the DH is an inactive form of the MCM replicative helicase motor, which does not unwind duplex DNA.

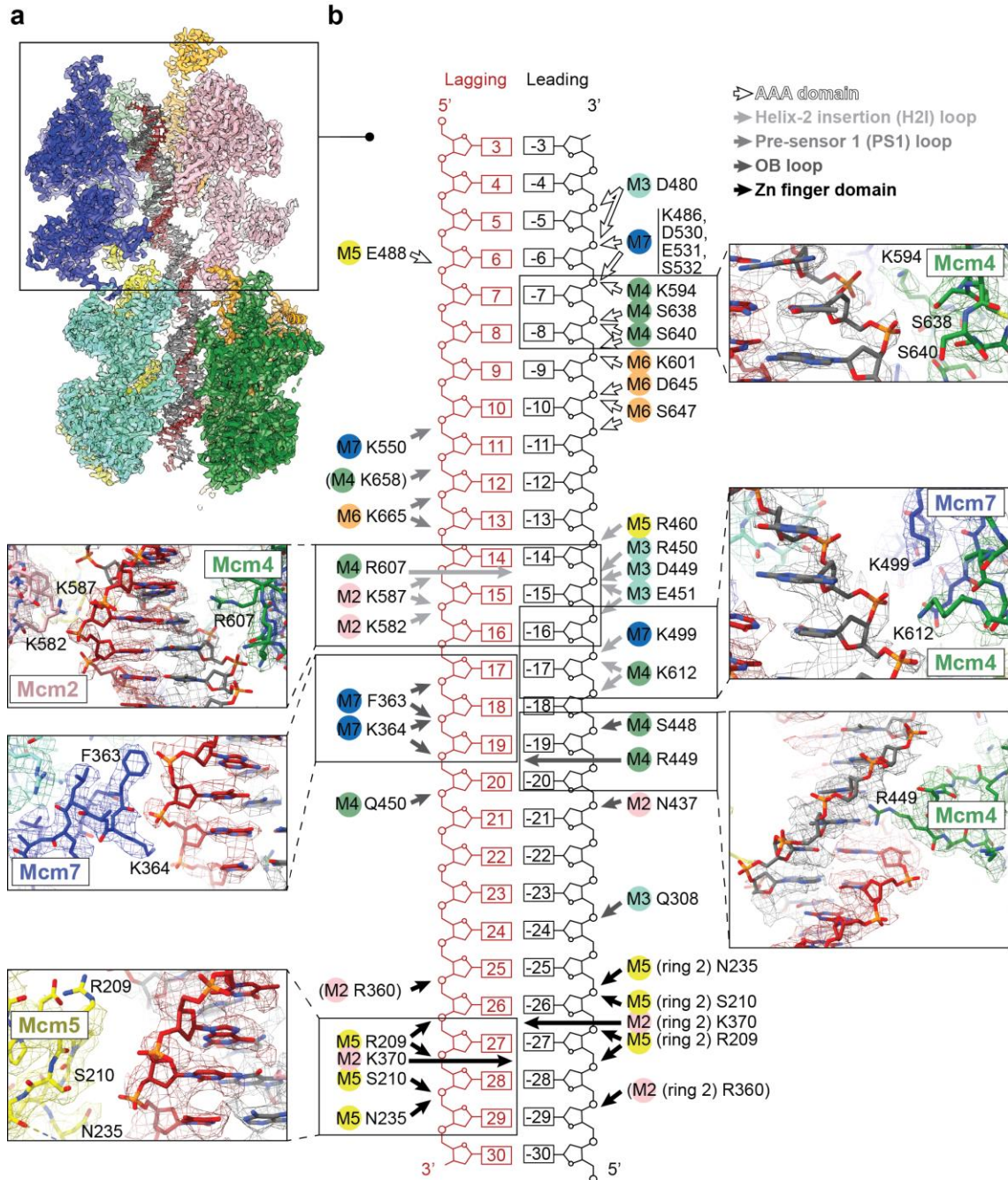


Figure 3.5: Protein–DNA contacts established by the *S. cerevisiae* DH.

a. Cut-through view of the DH, highlighting the bent double-stranded DNA in the central channel. **b.** Amino acids involved in DNA engagement. These residues mostly contact the phosphate backbone of the DNA. Inserts show details of the cryo-EM density with the atomic model.

3.2.4 ATPase state of the *S. cerevisiae* MCM DH

The ATPase sites found at the interface between two neighbouring MCM subunits, are described at a resolution that allows assigning each of the nucleotide states with a high degree of confidence. Most ATPase sites in the DH are occupied by ADP, and hence represent a post-catalytic state (Figure 3.6). The nucleotide density in the Mcm7–4 ATP hydrolysis centre is weaker than the density in the other sites, likely indicating partial nucleotide occupancy. The Mcm4–6 site does not contain a nucleotide, which correlates with biochemical and *in vivo* evidence indicating that an arginine finger mutant of Mcm4 (deficient for ATP hydrolysis) supports MCM loading *in vitro* as well as cell viability (Coster et al. 2014, Kang et al. 2014). The ATPase site formed between the Mcm6 and Mcm2 subunits is occupied by ATP. This finding is unexpected, given that mutation of the Mcm6 arginine finger, which is believed to be defective in ATP hydrolysis (but not ATP binding) prevents MCM loading (Coster et al. 2014, Kang et al. 2014). We speculate that ADP must have been released after MCM loading and a new ATP molecule associated. Alternatively, mutation of the Mcm6 arginine finger could alter the structure of the active site and prevent ATP binding, which could in turn cause the defect in helicase loading.

In summary, the structure of the *S. cerevisiae* DH reveals that most ATPase sites in MCM are in a post-catalytic state, compatible with the notion that ATP hydrolysis by the MCM is important for helicase loading. The structure also illustrates that *S. cerevisiae* MCM helicase loading does not promote duplex DNA melting, given that DNA remains in a duplex B-form configuration with continuous base pairing throughout the length of the MCM DH.

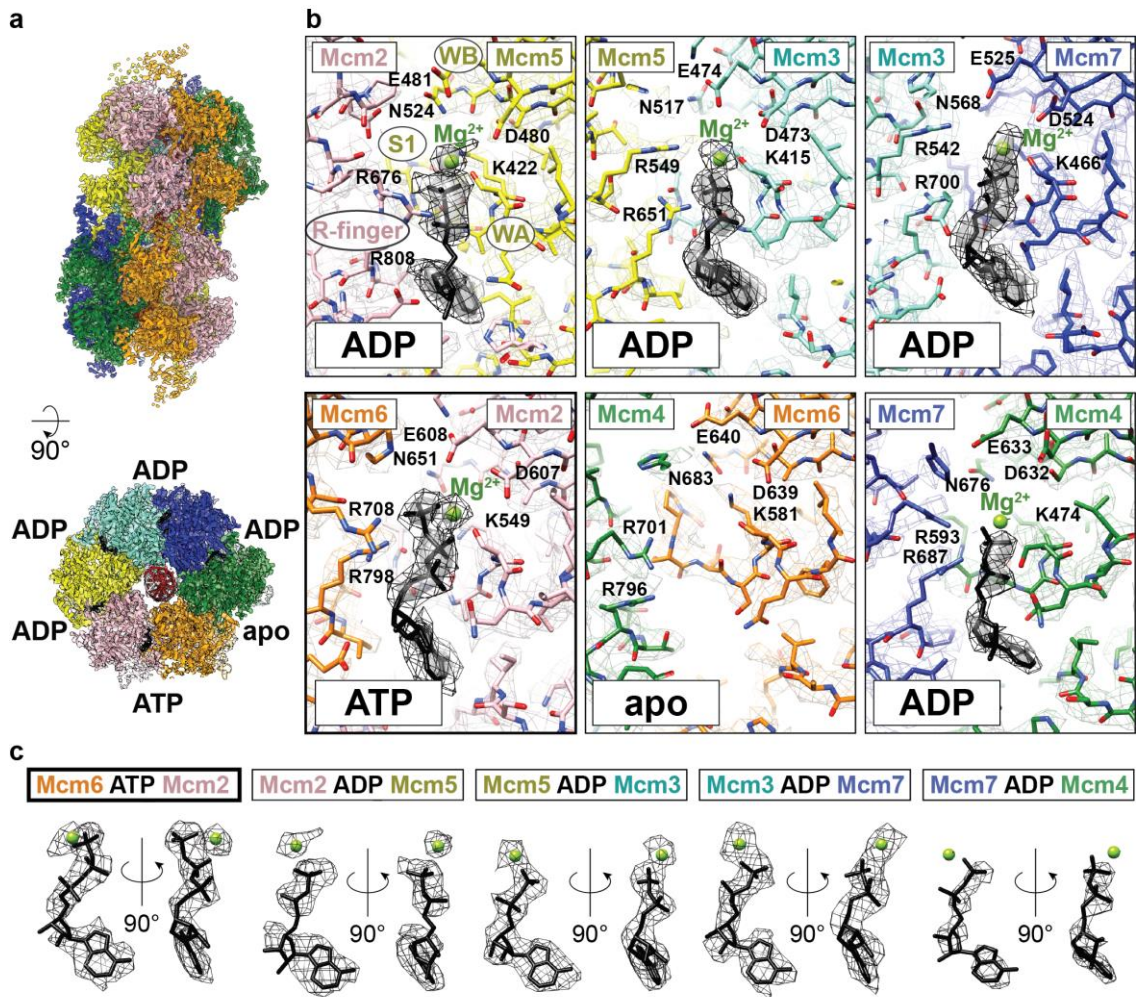


Figure 3.6: Nucleotide occupancy in the *S. cerevisiae* DH.

a. Side view of the *S. cerevisiae* DNA-loaded DH and bottom view of the ATPase sites. Bound nucleotides are shown in black. **b.** Illustration of the different ATPase sites between the MCM monomers and the nucleotides found at each site. **c.** Segmented density for each nucleotide and the coordinated magnesium ion.

3.3 *In vitro* reconstitution of origin licensing using *H. sapiens* proteins

3.3.1 Reconstitution of *H. sapiens* origin licensing

Having gained a good understanding of origin licensing in *S. cerevisiae*, we aimed to characterise the same reaction with purified *H. sapiens* proteins. Dr Florian Weissmann purified the *H. sapiens* proteins ORC1–5, ORC6, CDC6, CDT1 and MCM2–7. Compared to *S. cerevisiae*, CDT1 does not form a stable complex with

the MCM2–7 helicase and was therefore produced individually. The same is true for ORC6 and its interaction with the other five ORC subunits. Thus, ORC6 and the rest of the ORC complex were purified as separate entities. For the experiments described below, truncation variants of ORC1, CDC6 and CDT1 were designed, which lack N-terminal, intrinsically disordered regions for which a phase separation function had previously been invoked (Parker et al. 2019). These truncations were designed to prevent protein aggregation from interfering with helicase loading. Initially, we wanted to mimic the biochemical work performed with the *S. cerevisiae* system by loading MCM on long-non-roadblocked, biotinylated DNA and capturing loaded double hexamers on streptavidin coated paramagnetic beads (Yeeles et al. 2015). While this approach would have allowed high-salt washes to remove non-topologically loaded factors, it yielded inconclusive results, possibly due to the fact that *H. sapiens* DHs are more salt sensitive compared to *S. cerevisiae* counterparts. Instead, we incubated the five different protein complexes with a short DNA segment and imaged the entire reaction, adapting protocols from *S. cerevisiae* experiments (Miller et al. 2019). In particular, we have used a synthetic *S. cerevisiae* origin containing two high affinity ORC binding sites and flanked by two strong-positioning Widom sequences, which have repeats of A/T-rich and G/C-rich units, representing high affinity sites for histone octamers and promoting reproducible positioning of nucleosomes (Lowary and Widom 1998, Coster and Diffley 2017, Miller et al. 2019). As for the methyltransferase roadblocks described before, nucleosomes flanking the origin function to limit linear diffusion of the double hexamer particles. Loading reactions were carried out in the presence of ATP in the attempt to achieve MCM double hexamer formation and observe other loading intermediates. In a second experiment, ATP was replaced with the slowly hydrolysable ATP analogue, ATP γ S, to establish which loading intermediates require ATP binding but not hydrolysis. We then visualised the entire *H. sapiens* MCM loading reaction using negative stain EM (Figure 3.7a).

In the presence of ATP, we obtained 2D classes resembling ORC, nucleosomes, the MO loading intermediate, the DH and the open, loading competent MCM complex (Figure 3.7b). At first inspection, these classes appear similar to the complexes observed for *S. cerevisiae*. Different from yeast, a significant number of single MCM side views (Figure 3.7b fourth class from the top) were captured. We

speculated that the side view orientation of single MCMs might be enforced by a rigid, comparatively long stretch of duplex DNA that runs through the central MCM channel. Given DNA cannot be seen consistently in negatively stained EM images, we decided to use cryo-EM to understand whether DNA indeed runs through the central channel. In the presence of ATP γ S, the same complexes were observed as for ATP, apart from the DH species, which is coherent with the observation with *S. cerevisiae* proteins that ATP hydrolysis is required to complete MCM DH loading. One additional species was visible in ATP γ S, representing OCCM (Figure 3.7c), which recapitulates the findings in *S. cerevisiae*. We observed that ATP-hydrolysis dependency of other loading intermediates is different for human and yeast. For example, the *H. sapiens* MO was observed both in the presence of ATP as well as ATP γ S, while with *S. cerevisiae* proteins ATP hydrolysis is required to form the MO. Additionally, the density for ORC appeared smaller than the one in the *S. cerevisiae* MO, at least according to the most represented view visualised with negative stain EM (Figure 3.7d). We could not establish from these data alone whether the smaller density reflected an incomplete ORC1–5 complex or rather a different relative orientation of MCM and ORC in the *H. sapiens* MO. This issue will be addressed by a three-dimensional cryo-EM structure.

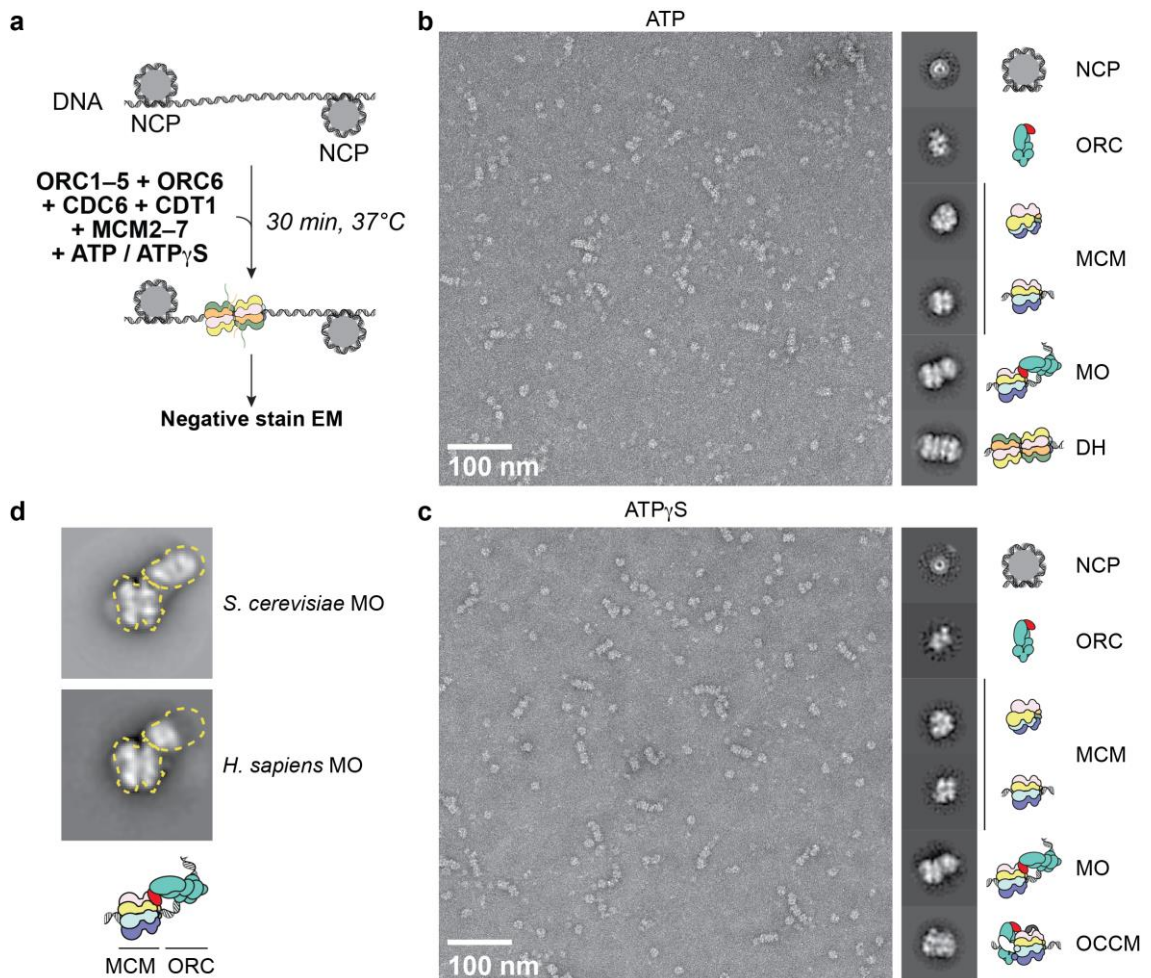


Figure 3.7: *In vitro* reconstitution of origin licensing using *H. sapiens* proteins.

a. Cartoon representation of the workflow to reconstitute origin licensing *in vitro*. Recombinantly expressed and purified *H. sapiens* proteins were incubated with origin DNA, which was capped by nucleosomes (NCP) on both ends. **b.** Representative negative stain EM micrograph of the licensing reaction performed in the presence of ATP. Representative 2D classes of the different complexes are shown on the right, including a cartoon illustration of each complex. **c.** Representative negative stain EM micrograph of helicase recruitment in the presence of ATP γ S and corresponding 2D classes. **d.** Comparison between the MO complex observed with *S. cerevisiae* and *H. sapiens* proteins. The outline of the *S. cerevisiae* MO complex (yellow, dashed line) is overlaid with both 2D classes to highlight the size difference of the ORC feature.

3.3.2 Visualising initial stages of *H. sapiens* MCM loading in the presence of ATP γ S by cryo-EM

To structurally characterise the helicase loading steps that depend on ATP binding but not hydrolysis, I reconstituted the *H. sapiens* MCM recruitment reaction in the presence of the slowly hydrolysable ATP analogue, ATP γ S. Next, the entire reaction was applied to R1.2/1.3 holey gold grids (UltraAuFoil), on which flakes of

graphene oxide had been deposited. Cryo-EM data were collected on a Titan Krios electron microscope equipped with a K2 Summit direct electron detector and a GIF Quantum energy filter (helped by Drs Donald Benton and Andrea Nans). Motion correction and CTF estimation were carried out, followed by particle picking, extraction and reference-free 2D classification in RELION. 2D classes of nucleosomes, OCCM and MO were observed (Figure 3.8). Single MCMs were also observed, which fell into two categories. One category represented open MCM rings viewed from the top, the other side views of MCM, which appeared DNA-loaded as judged from the cryo-EM density. MCM-containing particles were subjected to *ab initio* reconstruction and heterogeneous refinement in cryoSPARC, which yielded recognisable starting volumes for OCCM, single-loaded MCM hexamers (SH) and MO complexes (but not open MCMs, which were likely too poorly represented in the dataset to yield a reconstruction). These volumes were then used as input for a multi-reference 3D classification in RELION (Figure 3.9). Subsequent processing steps for each of the three complexes (OCCM, SH and MO) are discussed in paragraphs below.

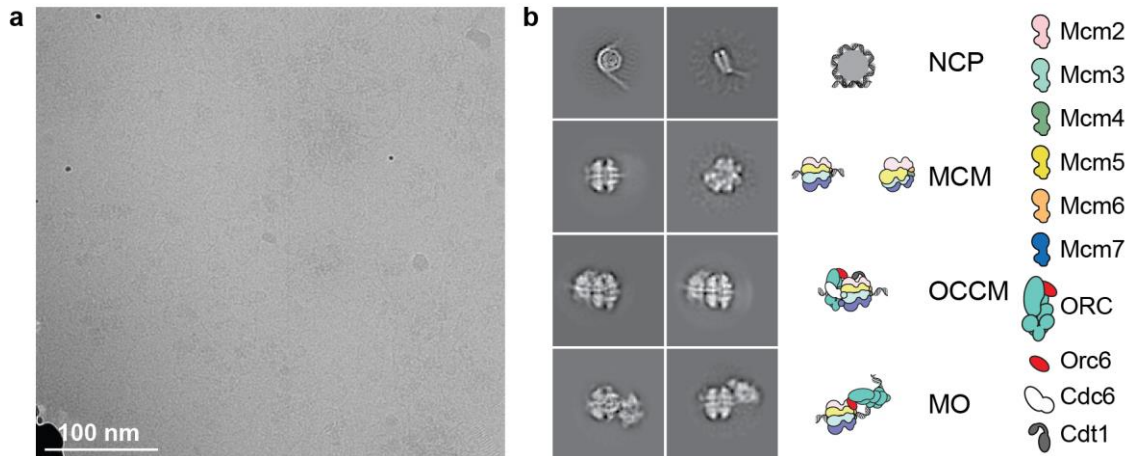


Figure 3.8: Cryo-EM analysis of *H. sapiens* MCM recruitment in the presence of ATP γ S.

a. Representative aligned movie sum (lowpass filtered). **b.** 2D classes and cartoon representation of complexes observed by cryo-EM. Side views of DNA-bound MCMs (second row, left) as well as open MCM complexes (second row, right) were observed.

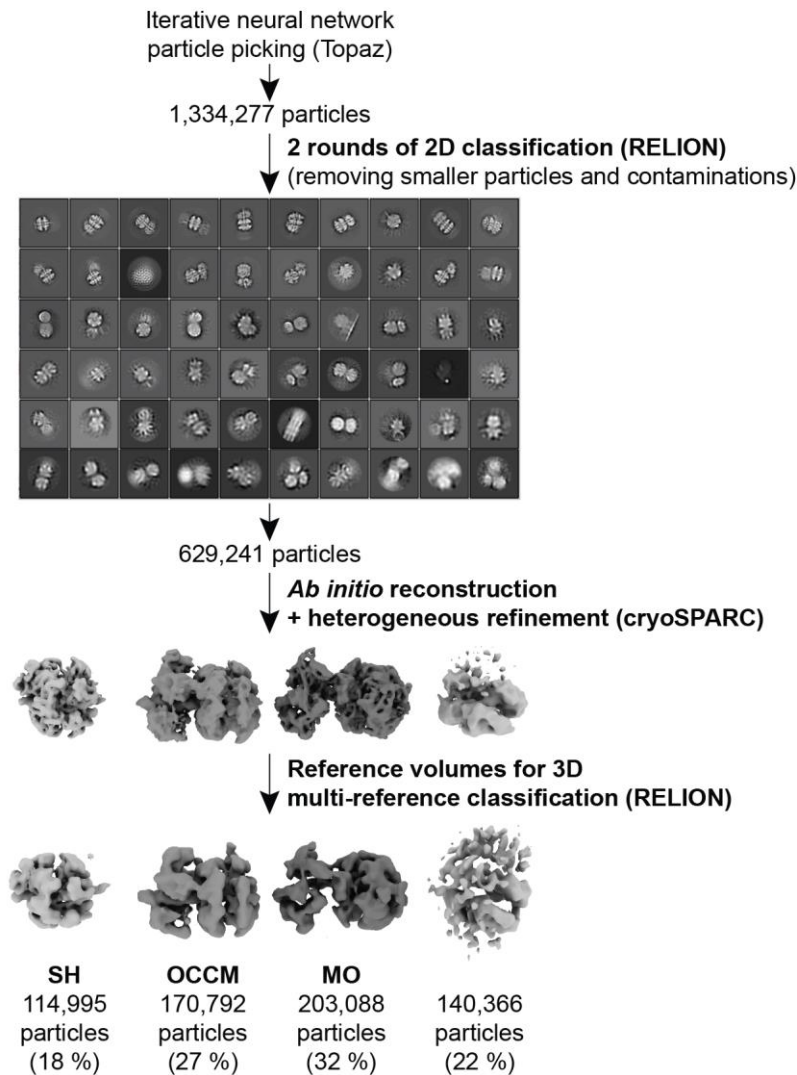


Figure 3.9: Processing pipeline *H. sapiens* MCM recruitment in the presence of ATP_γS.

Overview of the image processing procedure used to isolate single MCM hexamers, OCCM and MO complexes obtained in the presence of ATP_γS.

In summary, the cryo-EM analysis resulted in a complex resembling the *S. cerevisiae* OCCM, single MCM hexamers (SH), which were immediately recognised as DNA-bound, as well as a complex, in which ORC was bound to the N-terminal side of a DNA-loaded MCM hexamer. Structural characteristics of these complexes will be discussed in paragraphs below.

3.3.3 Visualising ATP-dependent *H. sapiens* DH loading by cryo-EM

To obtain a structure of the MCM DH and analyse any ATP-hydrolysis dependent intermediates, I visualised the loading reaction in ATP by cryo-EM. Sample preparation and data collection were carried out as described for the reaction in ATP γ S, aside from nucleotide usage. Here, motion correction and CTF estimation were followed by particle extraction and reference-free 2D classification in cryoSPARC. 2D classes of nucleosomes, single MCM hexamers (open and DNA-bound), MO and MCM double hexamers were obtained (Figure 3.10). Particles contributing to the distinct complexes were then separated using heterogeneous refinement (Figure 3.11a). The DH particles identified with this strategy were used for refinement as described below (see paragraph 3.7). Further training of the Topaz (Bepler et al. 2019) model for particle picking was instead required to improve the other 3D volumes (Figure 3.11b). This optimisation yielded particles used to generate an interpretable initial 3D model of the single-loaded MCM hexamer. Open MCM rings were also observed, which resulted in an interpretable 3D volume of a cracked open MCM ring similar but not identical to the *S. cerevisiae* MCM–Cdt1. Since subsequent processing did not push the resolution of this structure to the subnanometre range, this particle set was not considered further. Details of additional refinement steps and structural characteristics of DH, SH and MO are discussed in paragraphs below.

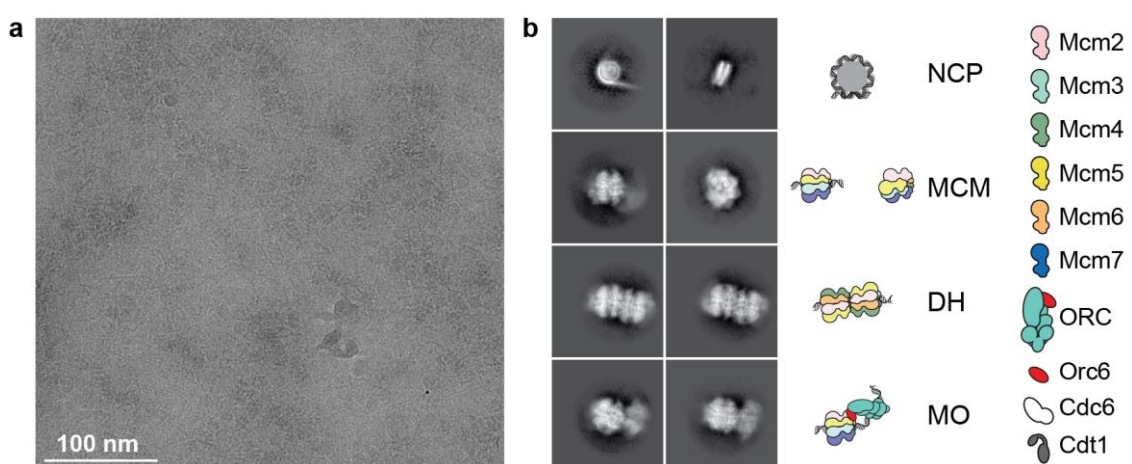


Figure 3.10: Cryo-EM analysis of *H. sapiens* MCM loading in the presence of ATP. **a.** Representative movie sum (lowpass filtered). **b.** 2D classes and cartoon representation of complexes present in the reaction.

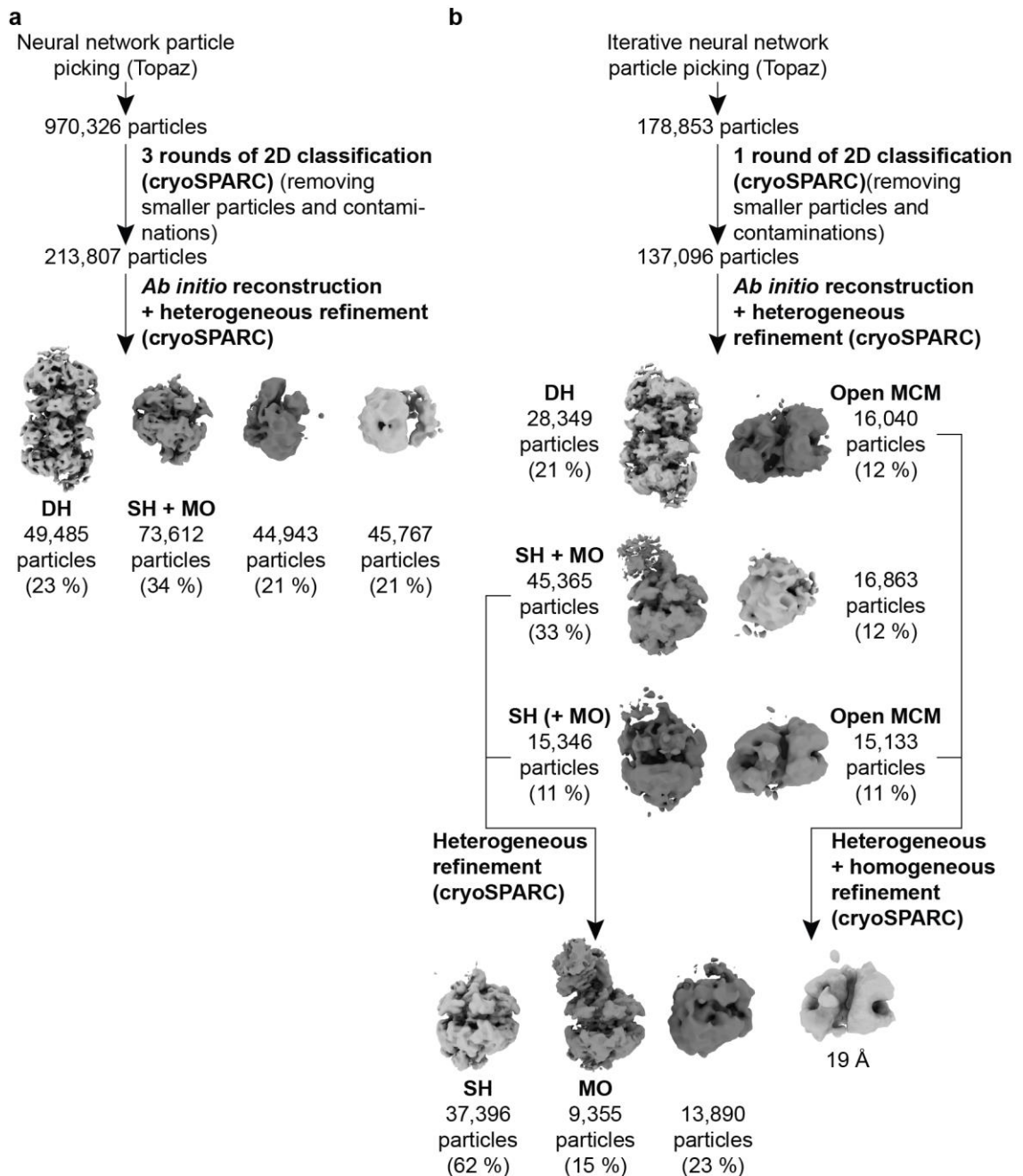


Figure 3.11: Initial processing steps of the *H. sapiens* origin licensing reaction in the presence of ATP.

a. Initial processing steps leading to the isolation of DH particles. **b.** Processing pipeline starting from an optimised neural network particle picking model, which resulted in the isolation of high-quality SH and MO particles.

3.4 Cryo-EM structures of *H. sapiens* OCCM and OC1M complexes

3.4.1 Image processing of the *H. sapiens* OCCM and OC1M complexes

According to studies with *S. cerevisiae* proteins, a key intermediate on the path to MCM DH formation is the so called OCCM complex, which reflects the recruitment of the first MCM ring and is comprised of ORC, Cdc6, Cdt1 and Mcm2–7. Formation of this complex depends on ATP binding, but not hydrolysis (Coster et al. 2014, Yuan et al. 2017, Yuan et al. 2020b). Equivalent to the previous *S. cerevisiae* studies, we have reconstituted the origin licensing reaction in the presence of the slowly hydrolysable ATP analogue ATP γ S seeking to capture the *H. sapiens* OCCM. Complexes resembling the OCCM were separated from other reaction intermediates as described before (paragraph 3.3.2) and subjected to Bayesian polishing (Zivanov et al. 2019) and CTF refinement (Zivanov et al. 2020) until no further improvement was observed. The resulting map showed only weak density in an ORC-proximal region, which is occupied by Cdc6 in the *S. cerevisiae* complex (Figure 3.12). Focused classification in RELION with a mask around this feature allowed separating a subset of 100,567 particles with more prominent density that was identified as CDC6. These particles were refined to 3.7 Å average resolution in cryoSPARC. Further separation of the full OCCM complex from a partial complex was achieved by local refinement of ORC–CDC6, followed by focused 3D classification using a mask that encompassed the entire ORC–CDC6 region. 34,116 particles containing CDC6 (OCCM) and 49,771 particles that lacked density for CDC6 were identified. As the latter structure contained ORC, CDT1 and MCM, this complex was termed OC1M. OCCM and OC1M were refined to 3.8 and 4.1 Å, respectively (Figure 3.13).

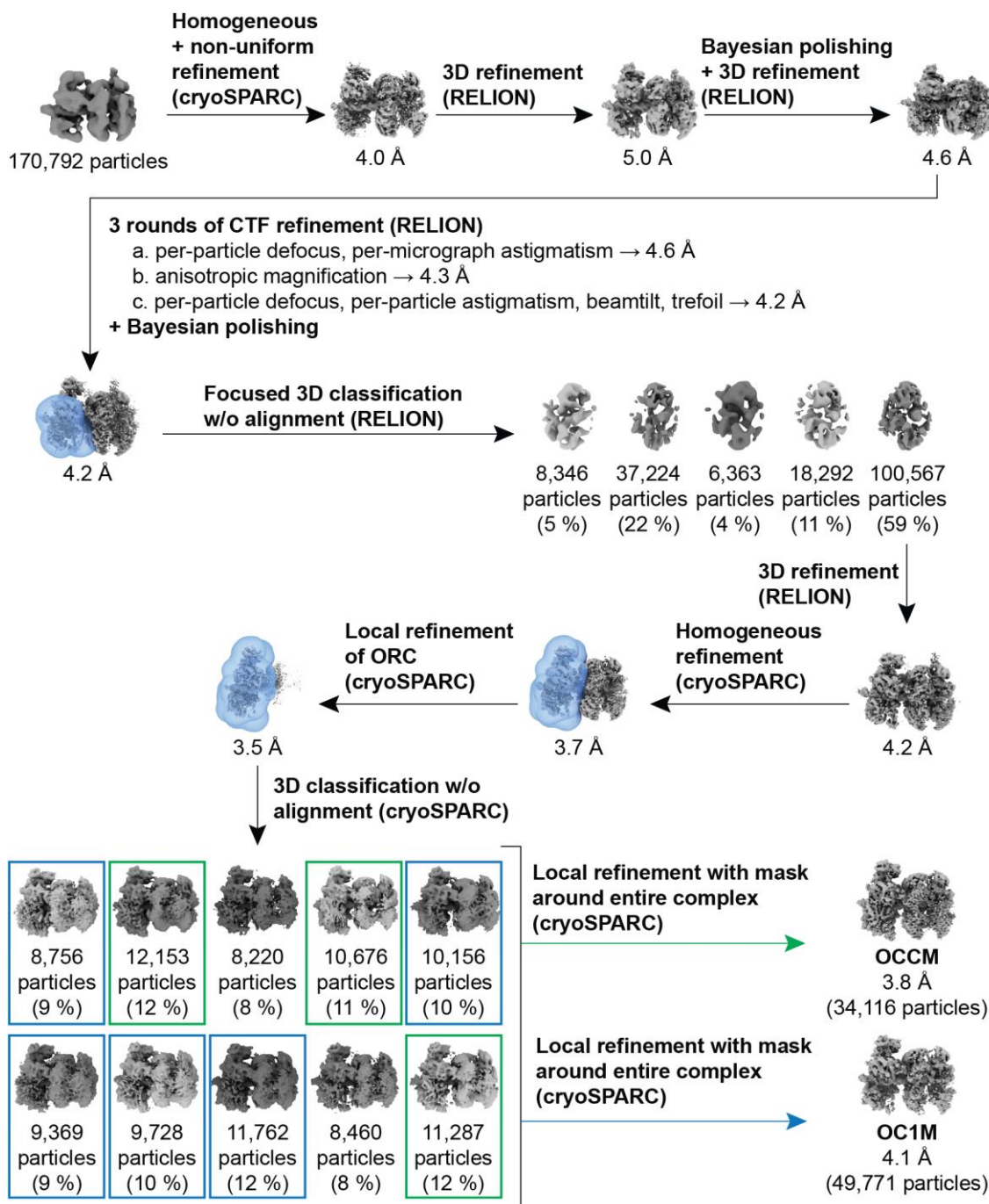


Figure 3.12: Processing pipeline for *H. sapiens* OCCM and OC1M complexes. Overview of the image processing steps to separate and refine OCCM and OC1M particles. Blue envelopes represent masks used in the subsequent processing step.

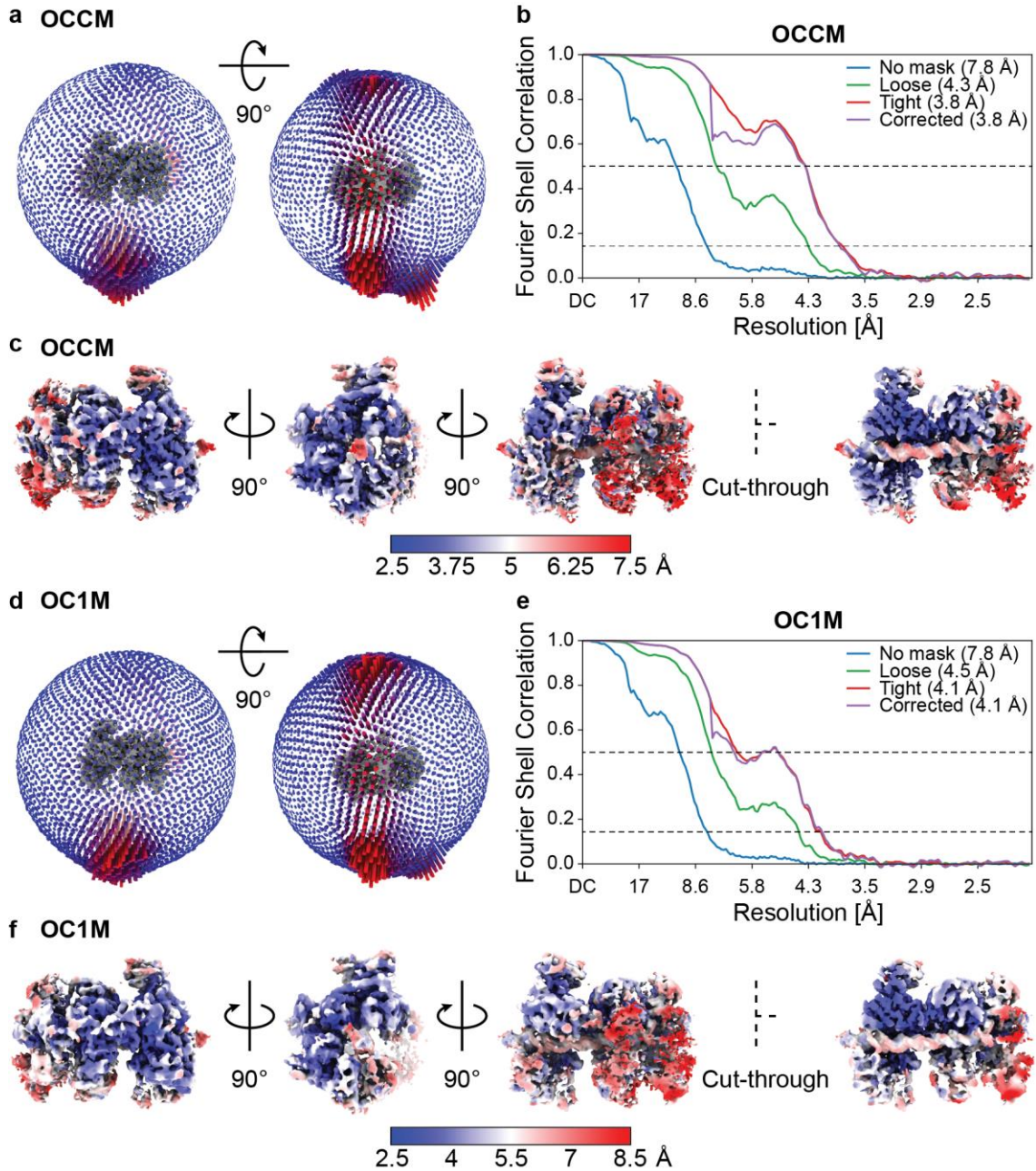


Figure 3.13: Structure quality of *H. sapiens* OCCM and OC1M.

a. Angular distribution of OCCM particles. **b.** Resolution of the OCCM estimated using gold-standard Fourier Shell Correlation. **c.** Three rotated views and one cut-through view of the OCCM 3D structure, colour-coded according to the local resolution. **d.** Angular distribution of OC1M particles. **e.** Resolution of the OC1M estimated using gold-standard Fourier Shell Correlation. **f.** Three rotated views and one cut-through view of the OC1M 3D structure, colour-coded according to the local resolution.

3.4.2 Overall architecture of the *H. sapiens* OCCM and OC1M complexes

To build atomic structures of the OCCM and OC1M complexes, I used models of *H. sapiens* ORC1–5–DNA (PDB code 7JPS), the open ORC1–5 complex (PDB code 7JPR) (Jaremko et al. 2020) as well as predicted structures of the individual MCM subunits, CDT1 and CDC6, generated using AlphaFold (Evans et al. 2022). The structures were docked into the cryo-EM map, combined into one PDB file and adjusted into the density using Coot v0.9-pre (Casañal et al. 2020). To complete the structure, idealised, B-form duplex DNA was fitted into the density after restraining Watson–Crick base pairing as well as stacking interactions.

The *H. sapiens* OCCM resembles the architecture of the *S. cerevisiae* complex (Sun et al. 2013, Yuan et al. 2017, Yuan et al. 2020b), showing the ORC1–5–CDC6 ring bound to the C-terminal face of the MCM helicase (Figure 3.14a). Double-stranded DNA is entrapped by ORC–CDC6 and is inserted into the MCM central channel. Half of the MCM hexamer, spanning MCM5–3–7, has a lower local resolution, which indicates flexibility of this region. Furthermore, the MCM2–5 gate, through which the DNA enters the MCM ring, remains slightly ajar. Work with *S. cerevisiae* proteins indicated that gate closure only occurs upon ATP hydrolysis by the MCM, which promotes the release of the loading factors (Fernández-Cid et al. 2013, Coster et al. 2014, Kang et al. 2014, Ticaú et al. 2015, Ticaú et al. 2017). Such structural transition is disallowed in our working conditions, given that ATP was swapped for a slowly hydrolysable ATP analogue. In the here presented *H. sapiens* OCCM structure, the CDT1 loading factor wraps around the MCM2, MCM6 and MCM4 subunits, with the MID and C-terminal winged helix domains of CDT1 interdigitating between the N-terminal A domains. CDT1 also binds to the MCM6 WHD, stabilising the MCM6–ORC5 interaction. In contrast to the *S. cerevisiae* OCCM, no density for ORC6 is observed in the *H. sapiens* complex.

The OC1M matches the OCCM structure, apart from the lack of CDC6 and a higher degree of flexibility in the MCM as well as ORC (Figure 3.14b). In the absence of CDC6, the ORC2 WHD can adopt several positions according to studies with *S. cerevisiae*, *H. sapiens* and *Drosophila* proteins (Bleichert et al. 2018, Li et al. 2018, Jaremko et al. 2020, Schmidt and Bleichert 2020, Yuan et al. 2020b). Consistent

with these data, the ORC2 WHD is not resolved in the OC1M structure. It has been proposed that this flexibility allows DNA insertion into a central groove of the crescent shaped ORC, as well as recruitment of CDC6 (Li et al. 2018, Jaremko et al. 2020, Schmidt and Bleichert 2020, Yuan et al. 2020b). The MCM ring also displays a larger degree of flexibility in the absence of CDC6. In fact, MCM5, MCM3 and MCM7 are resolved to lower local resolution in OC1M compared to OCCM. The OC1M has not been described to date, raising the question of where this intermediate stands in the helicase-loading reaction. Previous studies using *S. cerevisiae* proteins have shown that binding of Cdc6 to ORC is a prerequisite for the recruitment of the MCM (Fernández-Cid et al. 2013, Frigola et al. 2013, Yuan et al. 2020b). These data support the notion that OC1M is formed after OCCM assembly, and after CDC6 disengagement in our experiments. Disengagement however is understood to occur after hydrolysis of ATP, which would not be favoured when using a slowly hydrolysable ATP analogue. Thus, the possibility remains that OC1M formation might precede OCCM at least in the working conditions employed in the current study with *H. sapiens* proteins.

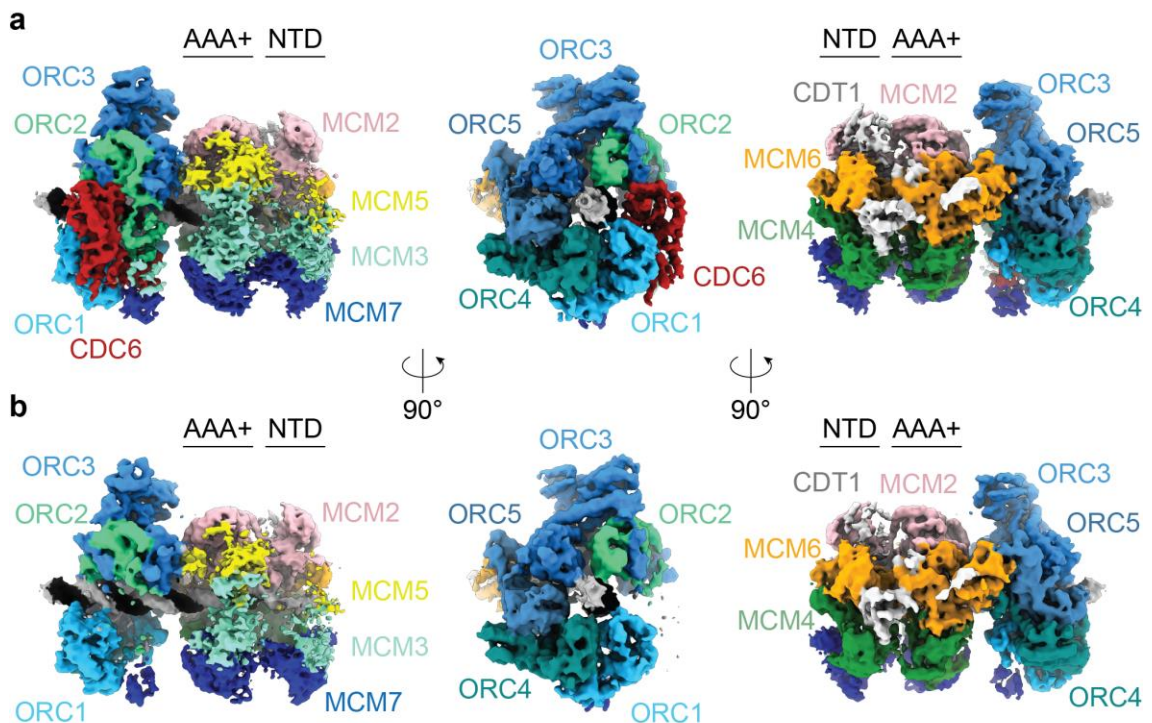


Figure 3.14: Overall architecture of *H. sapiens* OCCM and OC1M.

Iso-surface representation of the *H. sapiens* OCCM (a) and OC1M (b) complex in three different views.

3.4.3 Several MCM winged helix domains establish contacts with ORC and CDC6

Recruitment of the MCM helicase into the OCCM complex depends on the interaction of its flexibly tethered WHD with the C-terminal face of ORC–CDC6 (Yuan et al. 2017, Yuan et al. 2020b). Four MCM WHD are visible in the *H. sapiens* OCCM. First, the MCM3 WHD binds to CDC6 and the ORC2 WHD (Figure 3.15). Second, MCM7 interacts with CDC6 and ORC1. Third, MCM6 and ORC5, stabilised by CDT1, form a tight junction between the loader and the helicase. Despite the limited resolution, additional density between ORC3 and MCM5 could be assigned to the MCM5 WHD. The WHD of MCM4 is unresolved in the *H. sapiens* OCCM. MCM2 instead binds to ORC5 via its C-terminal ATPase tier. Therefore all MCM subunits, apart from MCM4, secure the helicase onto the ORC–CDC6 loading platform, allowing for the threading of the duplex DNA into the MCM ring. Similar interaction interfaces are observed in the OC1M, but conformational differences are caused by the absence of CDC6. As such, the WHD of ORC2 becomes highly dynamic, which consequently cannot be observed interacting with C-terminal MCM3.

The protein contacts that keep the *H. sapiens* OCCM together only partially match those observed in the *S. cerevisiae* complex (Yuan et al. 2017, Yuan et al. 2020b). Conserved interactions include ORC–CDC6 contacts with the WHD of MCM3, MCM7 and MCM6 and the MCM2 ATPase domain. In the *S. cerevisiae* OCCM map, density can be observed for the Mcm5 WHD like in the *H. sapiens* complex (Yuan et al. 2017), but it has not been assigned in the previous structures. A significant difference in the ORC–MCM interaction is the MCM4 WHD. This domain is proximal to Orc1 in the *S. cerevisiae* OCCM when the DNA is fully inserted into the MCM channel whereas the *H. sapiens* complex lacks this interaction. In future efforts, it will be important to biochemically and structurally validate that the MCM4 WHD is not required during helicase loading in humans.

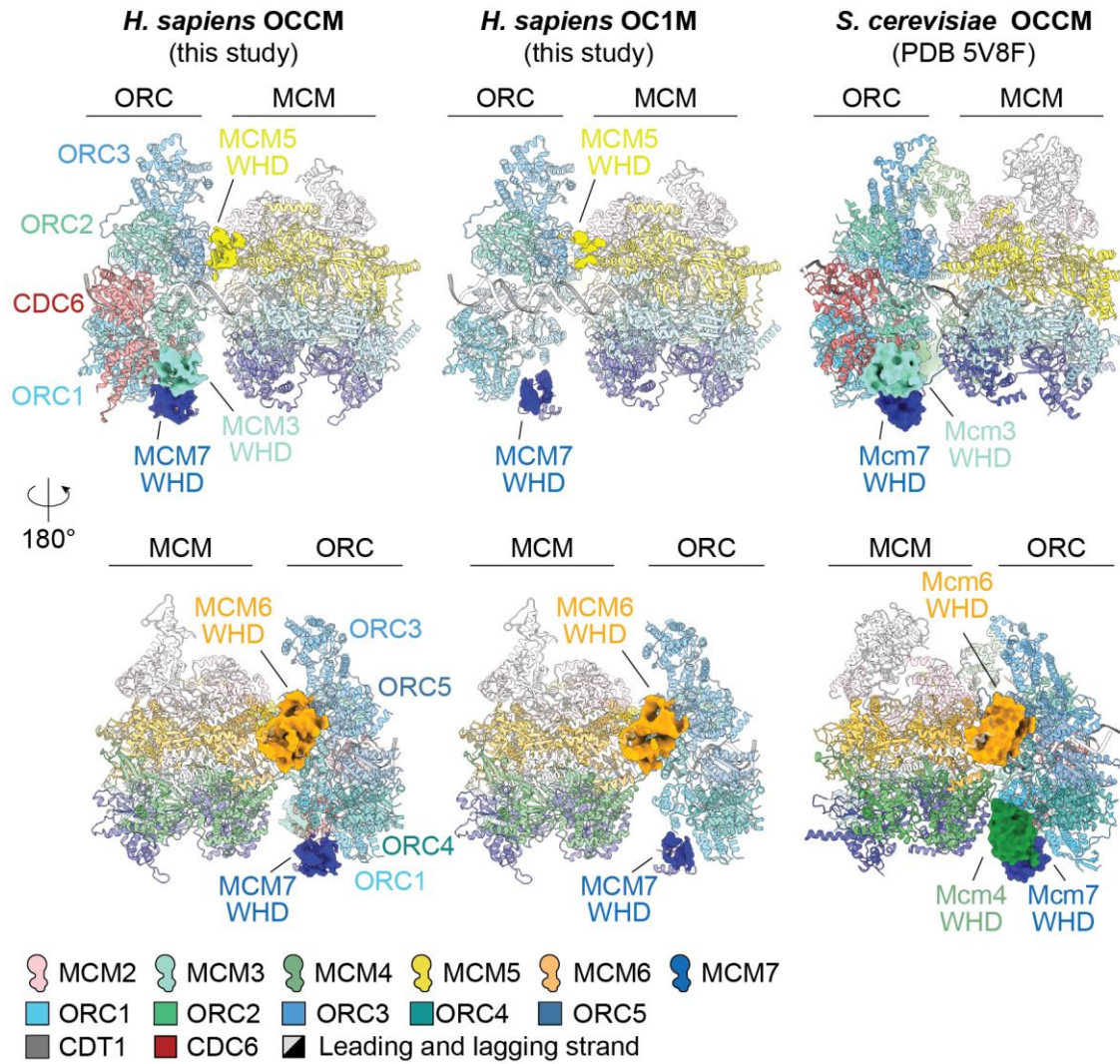


Figure 3.15: Interactions between MCM WHDs and ORC–CDC6/ORC in *H. sapiens* and *S. cerevisiae* OCCM/OC1M.

Atomic structures of *H. sapiens* and *S. cerevisiae* OCCM/OC1M with segmented cryo-EM density of the MCM winged-helix domains.

3.4.4 Protein–DNA interactions and nucleotide binding in *H. sapiens* OCCM/OC1M

The DNA found within the central MCM channel of the *H. sapiens* OCCM and OC1M complexes follows a similar path to the *S. cerevisiae* OCCM (PDB code 5V8F) (Yuan et al. 2017), however, several differences can be identified (Figure 3.16a). The DNA runs straight through the MCM to ORC in the *H. sapiens* complexes, whereas a slight bend is observed in the *S. cerevisiae* OCCM, where MCM and ORC are slightly offset. In all structures, the MCM grips the leading

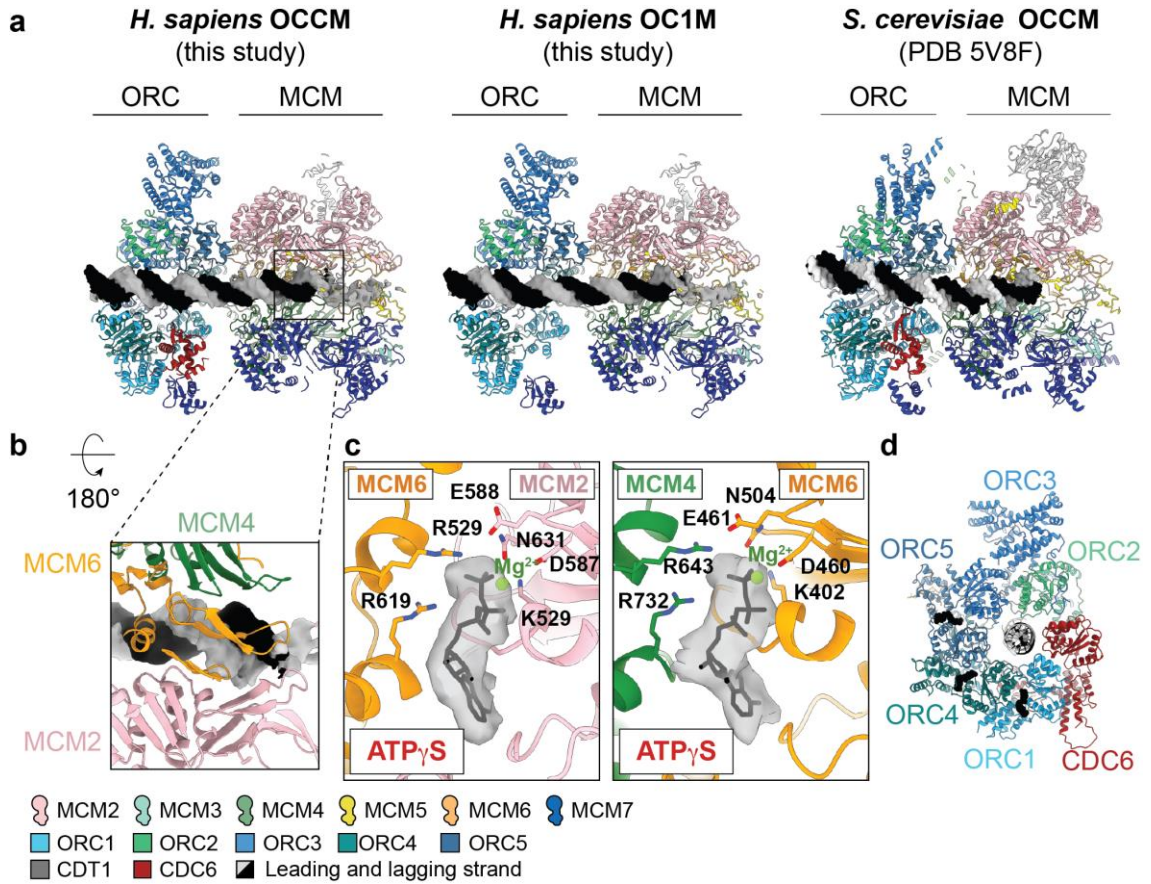


Figure 3.16: DNA and nucleotide binding by *H. sapiens* OCCM/OC1M.

a. Cut-through views illustrating the DNA bound to *H. sapiens* OCCM and OC1M and their comparison with *S. cerevisiae* OCCM. **b.** Atomic model of the MCM2–6–4 pore loops interacting with the DNA in the *H. sapiens* OCCM complex. **c.** Nucleotide density found within the MCM6–2 and MCM4–6 ATPase site. **d.** Top view of ORC in the *H. sapiens* OCCM structure showing the ATP_γS molecules bound in three of the four available ATPase sites. The limited resolution of the ORC1–CDC6 ATPase site does not allow for unambiguous assignment of the nucleotide state.

strand DNA via the PS1 ATPase pore loops of MCM2, MCM6 and MCM4 (Figure 3.16b).

While the local resolution of the MCM5–3–7–4 ATPase domains is insufficient to assign the nucleotide state of the active site, ATP_γS can clearly be identified within the MCM6–2 and MCM4–6 sites of the *H. sapiens* OCCM and OC1M (Figure 3.16c). This assignment agrees with observations made with the *S. cerevisiae* structure (Yuan et al. 2017). Together, these results indicate that ATPase subunits that are ATP_γS-bound are also DNA-engaged. This binding mode appears conserved between *S. cerevisiae* and *H. sapiens* complexes. Interestingly, a similar correlation is found in *S. cerevisiae*, *D. melanogaster*, as well as *H. sapiens* CMG

bound to an artificial fork substrate (Goswami et al. 2018, Eickhoff et al. 2019, Rzechorzek et al. 2020, Yuan et al. 2020a).

Equivalent to the *S. cerevisiae* OCCM, ATP γ S is found in the ATPase sites of ORC1, ORC4 and ORC5. The ATPase site established between ORC1 and CDC6 was solved at lower local resolution and does not allow determination of the nucleotide binding state (Figure 3.16d).

In summary, the *H. sapiens* OCCM shares the same overall architecture, the MCM–DNA interaction and nucleotide occupancy profile with its *S. cerevisiae* counterpart. Differences in the tethering of the MCM to ORC–CDC6 are observed, involving the MCM4 WHD in particular. Additionally, I have captured an OC1M complex that lacks CDC6. The role of this protein assembly in the MCM loading process remains to be established.

3.5 Cryo-EM structure of *H. sapiens* single-loaded MCM hexamers

3.5.1 Image processing of the *H. sapiens* SH

According to studies in *S. cerevisiae*, after MCM recruitment mediated by ORC, Cdc6 and Cdt1, the loading factors are released in an ATP-hydrolysis dependent manner, leaving the MCM hexamer loaded around double-stranded DNA. Upon visualisation of the loading reaction with *H. sapiens* proteins using cryo-EM, I observed DNA-loaded single hexamers (SHs) both in the presence of ATP as well as ATP γ S. This suggests that either ATP hydrolysis is not strictly required to load the MCM or that ATP γ S has been hydrolysed during the course of the reaction before plunge freezing. Extensive classification approaches were used to isolate SH particles from other reaction components (also see paragraphs 3.3.2 and 3.3.3). In total 37,396 SH particles were identified in the ATP dataset. These particles were subjected to CTF refinement and Bayesian polishing, resulting in a cryo-EM map solved to 3.5 Å resolution (Figure 3.17). 3D classification without

alignment was used to isolate 25,069 high-resolution particles, which were locally refined to 3.2 Å resolution in cryoSPARC (Figure 3.18).

A larger number of SH particles (114,995) was derived from subclassification of the ATP_γS dataset. Bayesian polishing followed by 3D refinement yielded a 3.6-Å resolution map (Figure 3.19). The ATP_γS dataset also contained 203,088 particles of DNA-loaded MO complexes, which are in all aspects identical to SHs, apart from the presence of ORC interacting with the N-terminal tier of the MCM. Given this identity, the two particle subsets were combined and subjected to CTF refinement and Bayesian polishing. The two complexes were then separated using 3D multi-reference classification in RELION. Refinement in cryoSPARC yielded a 3.5-Å resolution structure of the SH and 3.6-Å resolution structure of the MO complex. The local resolution of MCM in the MO improved to 3.5 Å after local refinement, while ORC was refined to 4.0 Å.

No difference was detected between the SH maps solved from the ATP and the ATP_γS datasets. For the structure analysis reported below, the higher resolution ATP map was used.

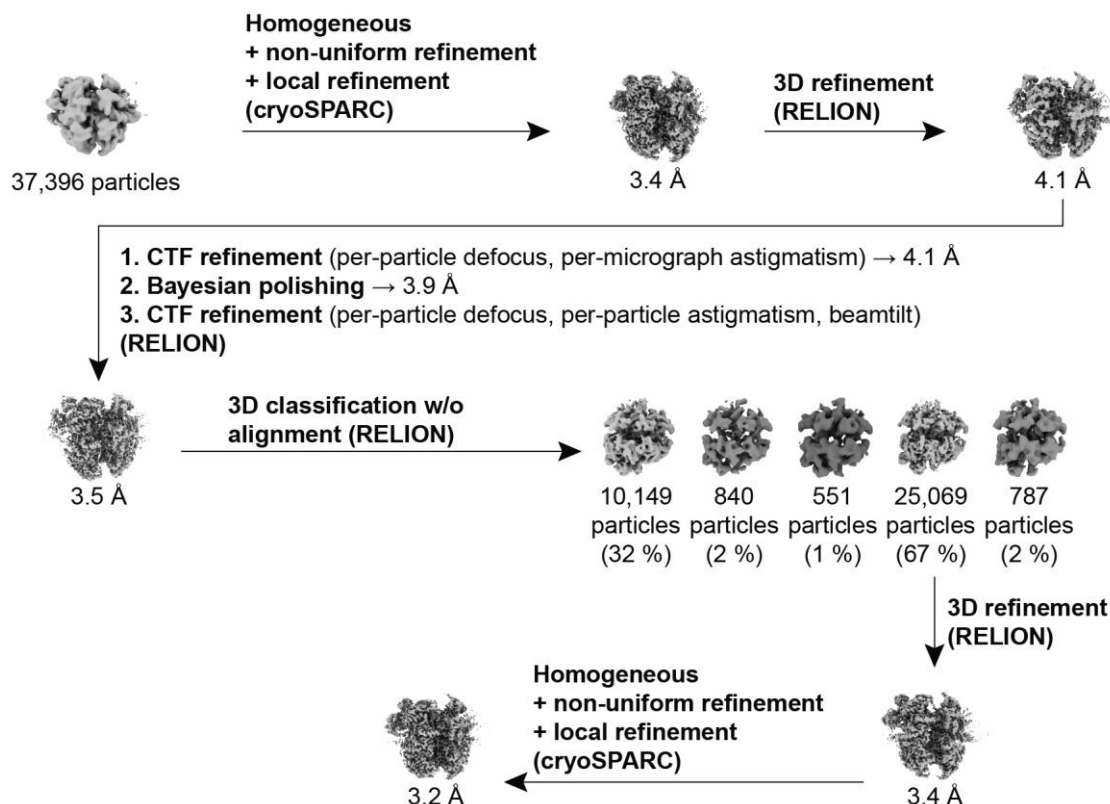


Figure 3.17: Processing pipeline for *H. sapiens* SH (ATP dataset).

Overview of the refinement strategy for the single-loaded MCM hexamer in the ATP dataset.

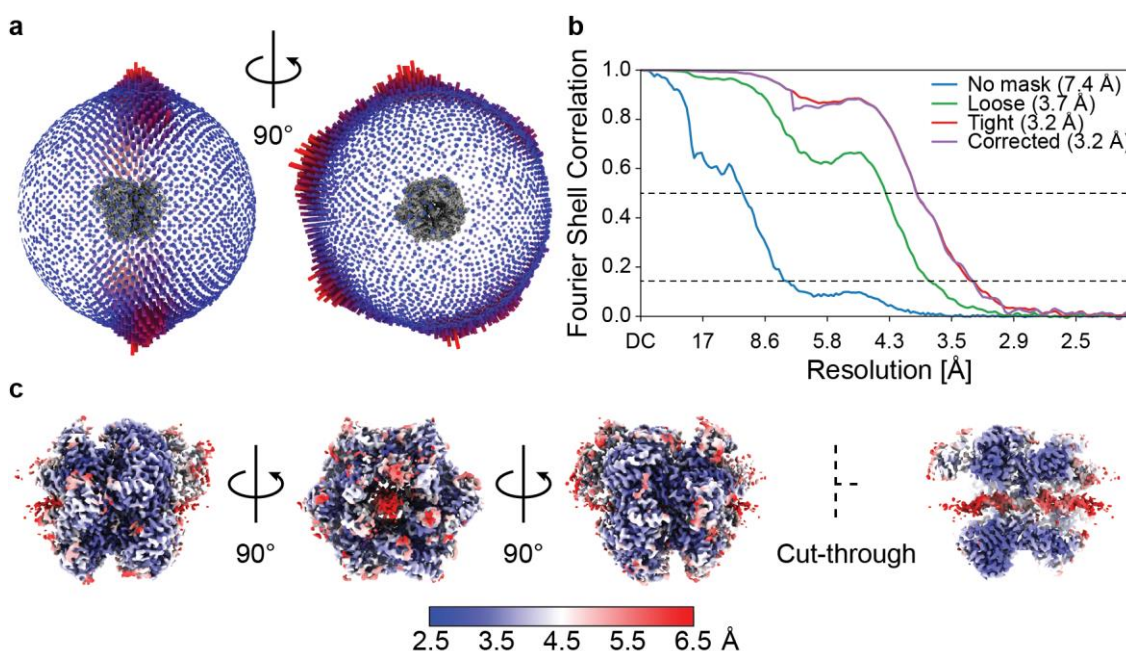


Figure 3.18: Structure quality of the *H. sapiens* SH (ATP dataset).

a. Angular distribution of SH particles viewed from the side and the C-terminal face. **b.** Resolution estimation according to gold-standard Fourier Shell Correlation. **c.** Surface view and cut-through view of the *H. sapiens* SH coloured according to the local resolution.

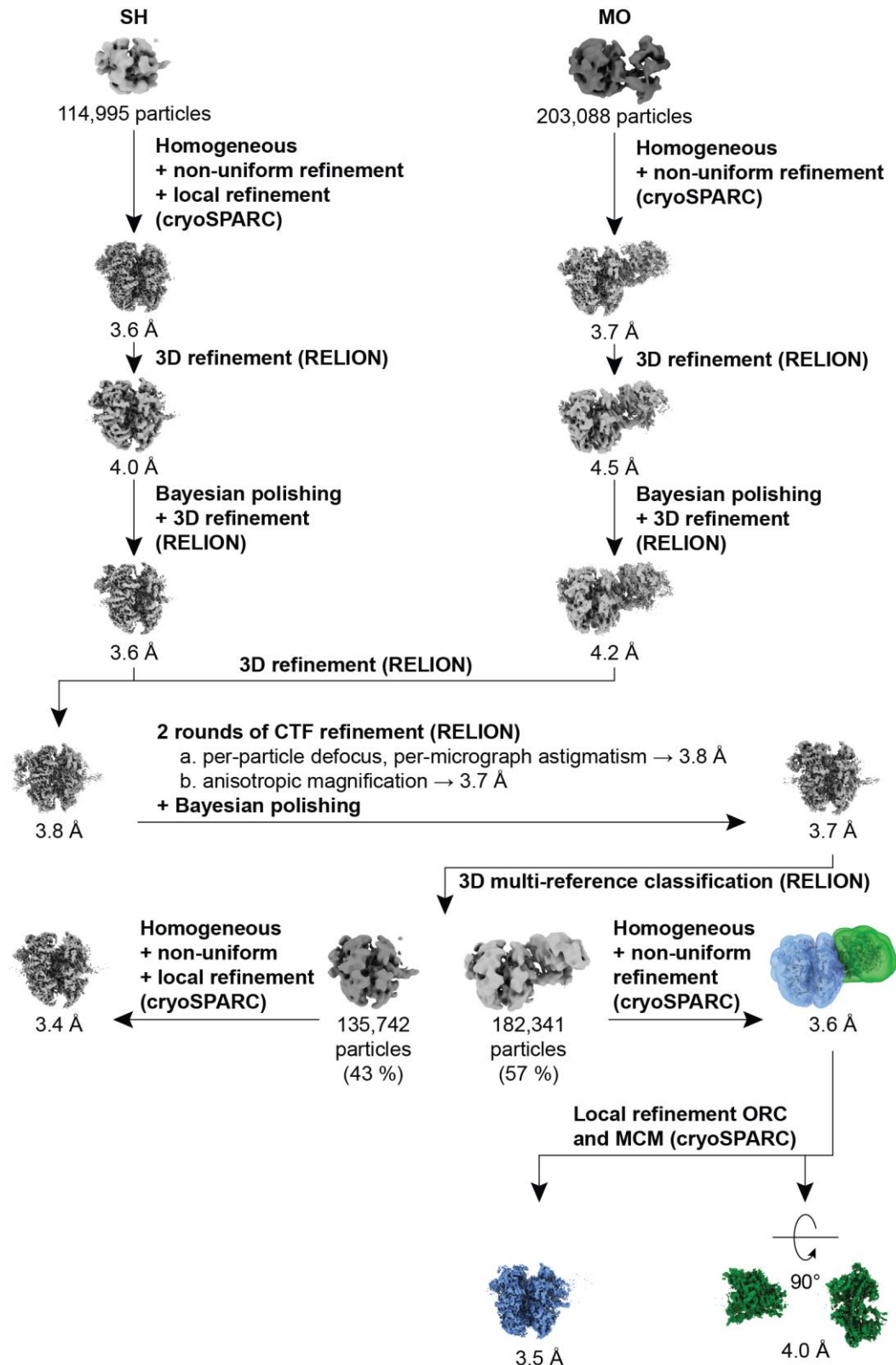


Figure 3.19: Processing pipeline for *H. sapiens* SH and MO (ATP γ S dataset). Outline of the image processing steps leading from the initial 3D classes to refined structures of SH and MO.

3.5.2 The *H. sapiens* SH binds double-stranded DNA

To obtain an atomic structure of the SH, models of the hexameric *H. sapiens* MCM2–7 amino- and carboxy-terminal tiers were generated using AlphaFold-Multimer (Evans et al. 2022). The assemblies were docked into the 3.2-Å cryo-EM map, combined, and idealised B-form duplex DNA was added. The entire model was adjusted using Coot v0.9-pre (Casañal et al. 2020).

The overall structure of the SH resembles the one of a single MCM ring in a DH. The six MCM subunits thereby form a closed ring around the DNA, with the AAA+ and N-terminal tiers arranged in a parallel configuration (Figure 3.20a). The duplex DNA has a slight bend at the interface between the AAA+ tier to the NTD but retains Watson–Crick base pairing (Figure 3.20b). This DNA distortion appears to be introduced by the ATPase PS1 loops of MCM7, MCM4 and MCM6 as well as the h2i loop of MCM2, which grip the leading strand (Figure 3.20c). As the pore loops follow the pitch of the DNA helix, they are arranged in a staircase configuration with MCM7 positioned at 3'-end of the leading strand. Additional protein–DNA contacts are established at various points of the MCM hexamer and include interactions with the lagging strand as well as with major and minor grooves (MCM4 K549, MCM5 R280, MCM5 R282). The last point of contact towards the N-terminal face of the MCM is R195 of the MCM5 subunit, which engages the phosphate backbone of the leading strand.

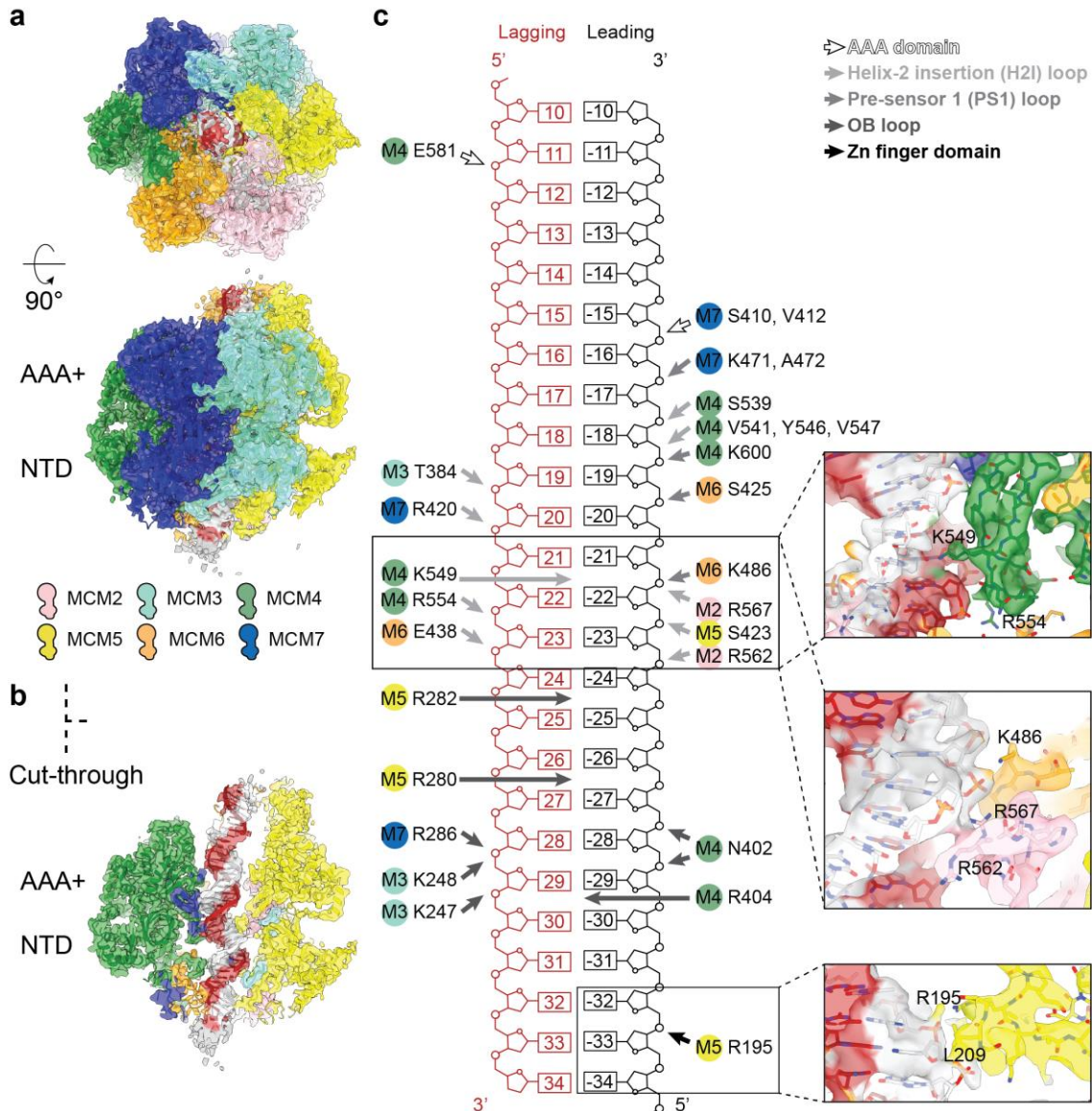


Figure 3.20: Protein–DNA contacts of the *H. sapiens* SH.

a. Iso-surface representation of the SH. The top image shows the complex viewed from the N-terminal side. The bottom image illustrates the complex from the side, with the AAA+ tier at the top and the NTD at the bottom. **b.** Cut-through view of the SH, showing duplex DNA running through the central channel. **c.** Overview of amino acids that contact DNA. Short arrows indicate interactions with the phosphate backbone, while long arrows represent interactions with bases in the major and minor grooves. The cryo-EM density and modelled interactions of selected regions of the structure are shown in the insets.

3.5.3 ATPase state of the *H. sapiens* SH

Loading of the MCM helicase and release of the loading factors has been shown to require ATP hydrolysis by the MCM (Coster et al. 2014). Establishing the nucleotide state of the six ATPase sites in the SH can start to inform the sequence of events that occur during DH loading. Inspection of the different inter-subunit interfaces in the *H. sapiens* SH revealed that the MCM2–5, MCM5–3 and MCM3–7 ATPase sites harbour ADP, while the MCM7–4 active site is devoid of nucleotide density (Figure 3.21). Conversely, the MCM6–2 and MCM4–6 sites show density compatible with ATP. These data suggest that hydrolysis occurs in at least the first three ATPase sites, whereas catalytic activity of the MCM6–2 and MCM4–6 might not be essential for loading. Biochemical and *in vivo* experiments with ATP-hydrolysis incompetent *S. cerevisiae* MCM variants revealed that activity of all sites except of the MCM4–6 site is essential for the loading of the double hexamer (Coster et al. 2014). Assuming that the SH is on the path to DH formation, the observation that ATP is bound at the MCM6–2 site suggests that catalytic activity might not be required for the loading of the first hexamer – but might become

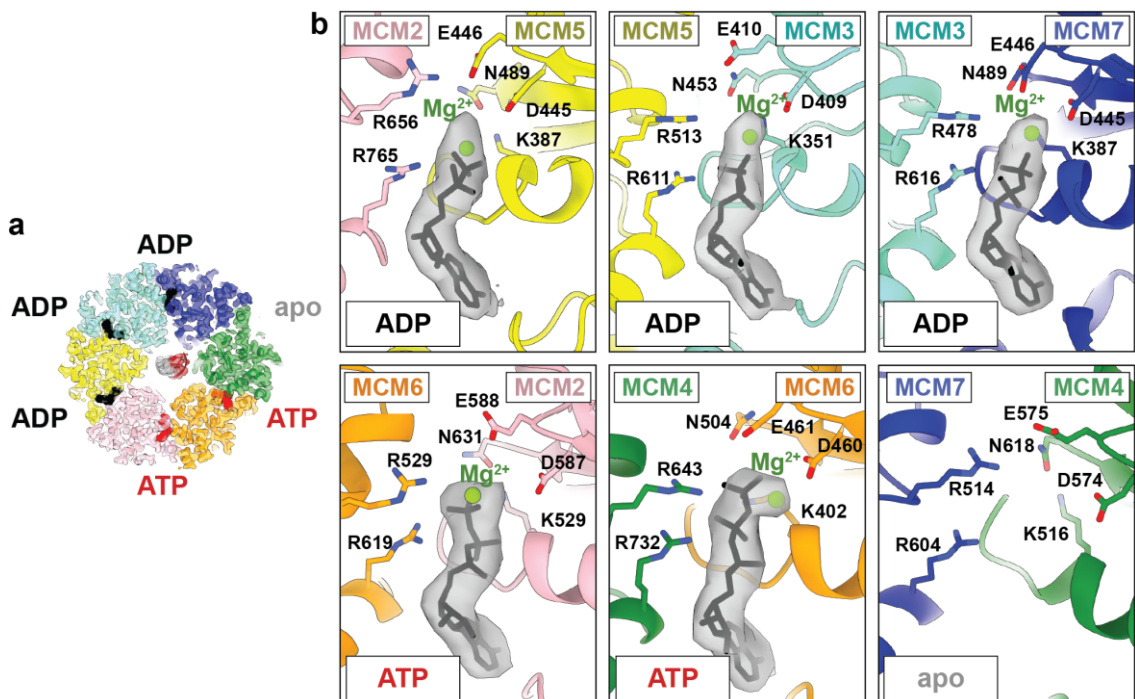


Figure 3.21: Nucleotide occupancy of the *H. sapiens* SH.

a. Bottom view of the ATPase sites. Bound ADP and ATP molecules are shown in black and red, respectively. **b.** Illustration of the different ATPase sites between the MCM monomers and the nucleotides found at each site.

important for the association with the second MCM ring. That said, the *S. cerevisiae* DH also contains ATP at the Mcm6–2 site, at least in its DDK-phosphorylated form (see paragraph 3.2.4). Hence, spent ADP might be released from the ATPase active site and substituted by ATP, ready to fire when the helicase is activated. Establishing the nucleotide binding state for the *H. sapiens* DH will further inform on any difference in nucleotide requirements to achieve full topological loading of the helicase motor with symmetry that can support bidirectional replication.

As described earlier, SHs were obtained not only when ATP was used in a full loading reaction, but also in the presence of ATP γ S. Nucleotide occupancy was found to be the same, irrespective of the nucleotide used. As nucleotide density compatible with ADP was observed in three of the six ATPase sites, we conclude that, under our experimental conditions, ATP γ S hydrolysis must have occurred, allowing for maturation from OCCM to full loading of a SH. Alternatively, the loading-competent MCM could already be ADP-bound at the MCM2–5, MCM5–3 and MCM3–7 sites, and ADP could remain bound at these sites throughout OCCM and SH formation. The OCCM structure, which was described earlier (paragraph 3.4.4), does not allow to shed light onto this issue, given that the local resolution of the pertinent sites is not sufficient to assign nucleotide occupancy.

Previous studies on the translocating *D. melanogaster* CMG helicase bound to a forked-DNA substrate have established that the nucleotide binding state influences the orientation of ATPase pore loops that interact with the DNA (Eickhoff et al. 2019). According to the derived single-stranded DNA translocation model, consecutive ATP binding and hydrolysis drive the sequential DNA engagement of neighbouring MCM subunits around the ring, which supports helicase translocation. With the MCM7–4–6–2 pore loops interacting with the DNA (Figure 3.20c) and ATP found in the latter two sites, the SH follows a similar ATPase-modulated DNA engagement pattern. It remains to be established whether ATP-hydrolysis driven MCM translocation is required during the helicase loading process in *H. sapiens*.

Unlike for the *H. sapiens* proteins described herein, previous cryo-EM analysis of the origin licensing reaction using wild type *S. cerevisiae* proteins failed to yield a

structure of the single-loaded MCM helicase. This suggests that, during helicase loading in *S. cerevisiae*, the first loaded MCM hexamer is quickly followed by a second ORC binding event, forming the MO and ultimately leading to the recruitment of a second MCM helicase. Stability of the *S. cerevisiae* SH might also be limited compared to the *H. sapiens* counterpart, resulting in fast dissociation, and might be counteracted by MO or DH formation. It remains to be established how the second MCM ring is efficiently loaded in the correct orientation to allow for the head-to-head DH formation, which contains the symmetry for bidirectional replication.

3.6 Cryo-EM structure of *H. sapiens* MO

3.6.1 Image processing of the *H. sapiens* MO complex

In *S. cerevisiae*, ORC detaches from the C-terminal MCM tier once the helicase is loaded onto DNA and a second ORC-binding event occurs, which both engages the N-terminal MCM tier as well as an inverted ORC binding site on the DNA (Miller et al. 2019). Given the direction of DNA binding, ORC then recruits a second MCM ring in the correct orientation to form a head-to-head MCM double hexamer. With *H. sapiens* proteins, we also observed ORC binding at the N-terminal side of the MCM. ORC still binds DNA in an inverted configuration compared to OCCM, but the relative orientation of ORC and MCM appears to be different (Figure 3.7d). To characterise this alternate arrangement, I aimed to solve a high-resolution structure of the *H. sapiens* MO complex. Our cryo-EM analysis showed that single-loaded MCM hexamers can form in the presence of either ATP or the slowly hydrolysable ATP γ S. In both conditions, ORC interacts with the N-terminal face of the SH forming the MO complex. A larger dataset was collected for the helicase recruitment reaction in ATP γ S, resulting in a higher number of MO particles and therefore a higher resolution structure. The lower-resolution MO of the ATP dataset showed a similar overall structure.

As described previously (Figure 3.19), Bayesian polishing and CTF refinement were carried out for the MO in combination with SH particles. MO particles were

then separated and refined to an average resolution of 3.6 Å in cryoSPARC (Figure 3.22a-c). While the MCM was well resolved in this map, interpretability of the ORC density was limited. I therefore performed local refinement on both ORC and MCM using masks encompassing only one of the two subcomplexes at a time (Figure 3.19). This procedure resulted in a significantly improved map of ORC at a resolution of 4.0 Å and a map of MCM at 3.5 Å (Figure 3.22d).

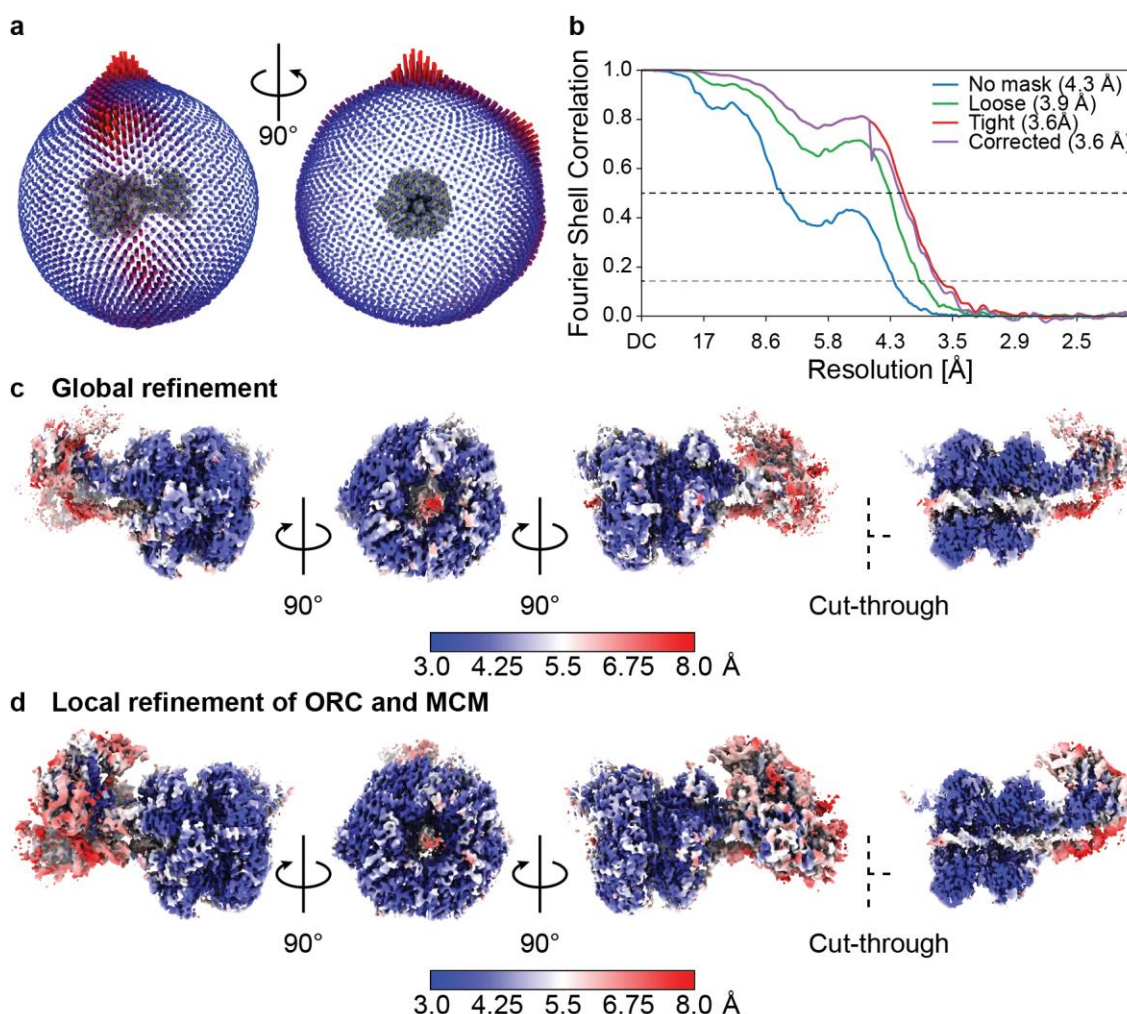


Figure 3.22: Structure quality of the *H. sapiens* MO complex.

a. Angular distribution of MO particles. **b.** Resolution estimation based on Fourier Shell Correlation. **c.** Map of the refined MO complex coloured according to the local resolution. **d.** Composite map of locally refined ORC and MCM coloured according to the local resolution.

3.6.2 Overall architecture of the *H. sapiens* MO complex

The atomic structure of the *H. sapiens* MO complex was built starting from the *H. sapiens* ORC1–5 complex in its open conformation (PDB code 7JPR) (Jaremko et al. 2020), the SH model (described in paragraph 3.5.1), and the ORC6 model generated using AlphaFold (Varadi et al. 2022). The models were docked into the cryo-EM density and adjusted using Coot v0.9-pre (Casañal et al. 2020). ORC binds to the N-terminal tier of the DNA-loaded MCM hexamer, with ORC6 bridging between ORC1–5 and MCM (Figure 3.23a). Additionally, ORC and MCM are connected via the duplex DNA, which exits the MCM central channel and is bent towards ORC. In the MO, the ORC3–4–5 subcomplex adopts an open conformation, similar to the ORC1–5 complex obtained in the absence of DNA (Jaremko et al. 2020) (Figure 3.23b). In contrast to the open ORC complex, lower local resolution of the ORC1 AAA+ domain in the MO suggests a higher degree of flexibility of this domain. Furthermore, the WHD of ORC2 is unresolved in the map, which is consistent with the *S. cerevisiae* MO (Miller et al. 2019) and the DNA-bound *H. sapiens* ORC1–5 (Jaremko et al. 2020). Together, the dynamic character of the ORC1 AAA+ domain and ORC2 WHD might allow the DNA to enter the cavity of the crescent shaped ORC, and later CDC6 to be recruited.

ORC6 sits in between ORC1–5 and MCM. The two TFIIIB-like domains that form ORC6 can be confidently docked into the cryo-EM map (Figure 3.23c), allowing for identification of the tethering points. The N-terminal domain (residues 1–94) interacts with MCM2, MCM6, ORC4 and ORC5, while the C-terminal TFIIIB-like domain (residues 95–190) bridges between MCM4–6 and ORC5. ORC6 thereby generates an extensive interaction interface, connecting the C-terminal face of ORC with the N-terminal face of MCM.

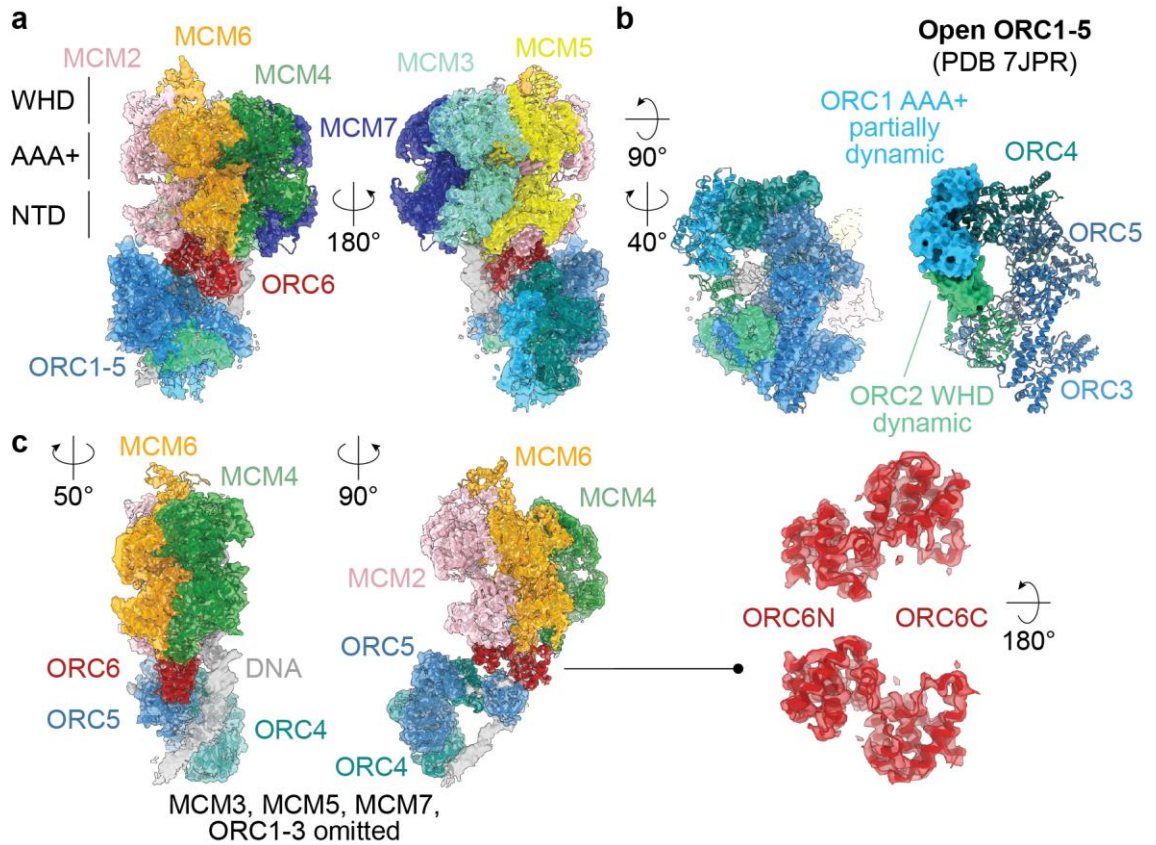


Figure 3.23: Overall architecture of the *H. sapiens* MO complex.

a. Composite map of the locally refined ORC and MCM in the *H. sapiens* MO complex, coloured by subunit. The NTD, ATPase and WH domains of the MCM ring are indicated. **b.** The *H. sapiens* ORC1–5 complex in its open conformation fits into the cryo-EM map of the MO. However, the AAA+ domain of ORC1 is poorly resolved and the ORC2 WHD is not visible, indicating a dynamic nature for these two domains. **c.** ORC6 interacts with MCM2–6–4 and ORC4–5. The N- and C-terminal TFIIB-like domains of ORC6 can be confidently docked into the cryo-EM map as shown on the right.

3.6.3 DNA and nucleotide binding by *H. sapiens* MO

The MCM in the MO resembles the structure of single-loaded hexamers observed in this dataset. In addition, the DNA in the central MCM channel follows the same path in both structures (Figure 3.24a), with the ATPase pore loops of MCM7,4,6 and 2 contacting the leading strand (Figure 3.24b and c). This protein–DNA interaction pattern differs from the *S. cerevisiae* MO, in which the same MCM subunits contact the lagging, and not the leading strand, which is instead touched by Mcm3 and Mcm5 (Miller et al. 2019).

Similar to the *H. sapiens* SH, most ATPase sites are occupied by ADP, except of the MCM6–2 and MCM4–6 sites, which contain ATP (Figure 3.24c). The MCM7–4 site harbours ADP, while the same ATPase centre is nucleotide free in the SH (Figure 3.24c). Assuming that the MO is formed after the SH loading, one could speculate that ATP binds at the free MCM7–4 site and becomes hydrolysed to support MO formation. It remains unclear whether or not ORC engagement at the N-terminal MCM side stimulates this process. The change in nucleotide occupancy does not affect the DNA-binding by the ATPase pore loops. Conversely, differences in nucleotide binding observed in *H. sapiens* and *S. cerevisiae* MO correlate with changes in DNA binding. In *S. cerevisiae*, the Mcm7–4 and Mcm4–6

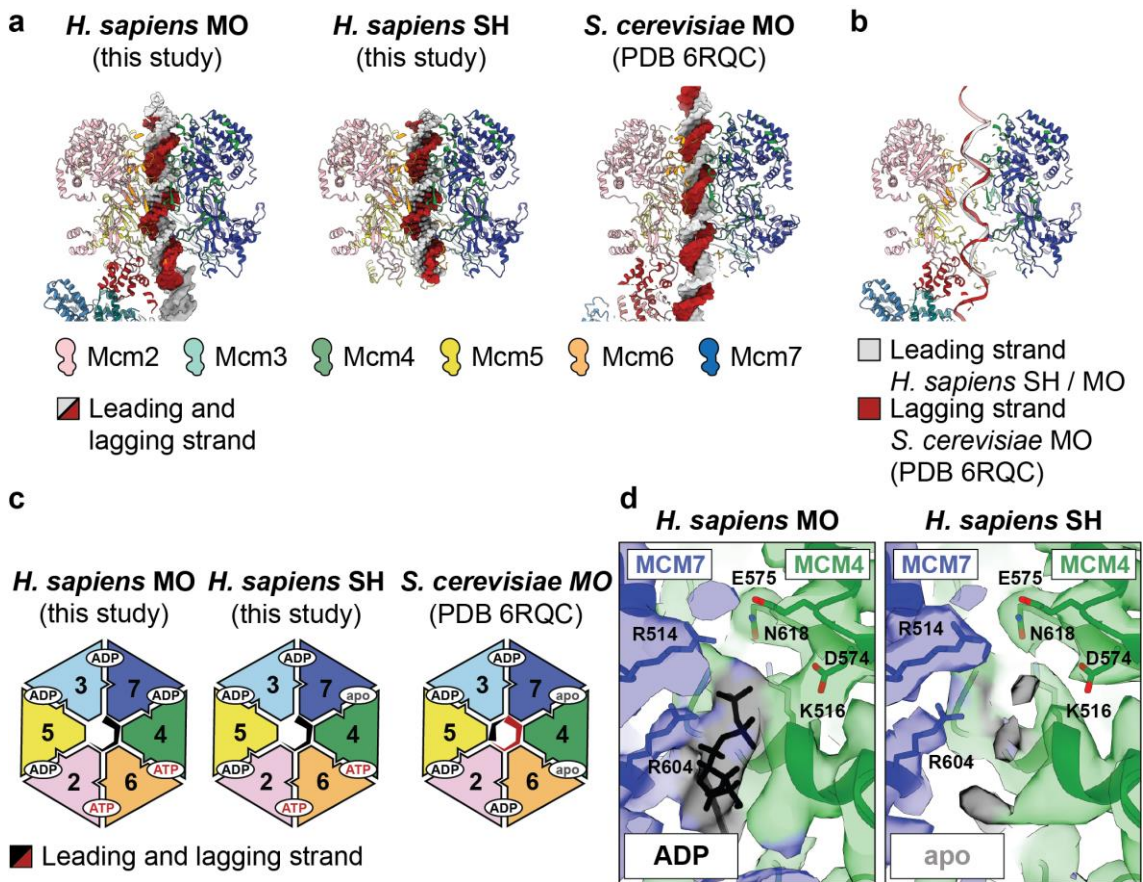


Figure 3.24. DNA and nucleotide binding by *H. sapiens* MO.

a. Atomic structures the *H. sapiens* MO and SH compared to the *S. cerevisiae* MO. *H. sapiens* MO and SH have the same DNA binding pattern, whereas the DNA is found in a different position in the *S. cerevisiae* MO. **b.** Illustration of *H. sapiens* MO omitting the lagging strand, overlaid with the lagging strand of *S. cerevisiae* MO. **c.** Cartoon showing the nucleotide binding and DNA interactions of the ATPase pore loops in the three different complexes. **d.** Cryo-EM density of the MCM7–4 ATPase site in the *H. sapiens* MO and SH.

sites are devoid of nucleotide density, while all other sites are occupied. The identity of the bound nucleotides could, however, not be established unambiguously due to a lower resolution of the cryo-EM map. In conclusion, *H. sapiens* and *S. cerevisiae* MO differ both in nucleotide occupancy and in the DNA contacts established by the MCM.

3.6.4 *H. sapiens* MO appears to be in an inhibited conformation

Significant differences in the conformation of *H. sapiens* and *S. cerevisiae* MO can be observed, which arise from protein–protein interactions that are distinct in the two complexes. *H. sapiens* ORC6 interacts with ORC4–5, whereas Orc2, 3 and 5 are contacted by Orc6 in *S. cerevisiae* (Miller et al. 2019) (Figure 3.25a). These interactions cause the C-terminal face of ORC to point towards the MCM in *H. sapiens* MO, while the same surface is solvent exposed in the *S. cerevisiae* complex. Furthermore, ORC bridges across the Mcm2 and Mcm5 subunits in *S. cerevisiae*, which enables ORC to recognise the DNA-loaded MCM helicase when it is topologically closed around the DNA, with a closed Mcm2–5 DNA gate. Instead, *H. sapiens* ORC forms a completely different interaction interface in the MO by interacting with MCM2, MCM6 and MCM4. *H. sapiens* ORC would therefore not be able to sense gate closure in the context of the MO. Despite this observation, like for *S. cerevisiae*, *H. sapiens* MO only forms when the MCM is DNA loaded.

The new conformation observed for the *H. sapiens* MO has significant implications for the mechanism of second MCM helicase recruitment. In *S. cerevisiae*, a second MCM–Cdt1 complex can bind to the solvent-exposed C-terminal face of ORC in the MO, recapitulating an interaction observed in the pre-insertion OCCM (Yuan et al. 2020b) (Figure 3.25b). Conversely, the same binding interface is occluded in the *H. sapiens* MO so that the second MCM hexamer would clash with the first loaded MCM. This complex might therefore represent a state that is not on the path to double hexamer formation, and it could rather be an inhibited configuration. Alternatively, a reconfiguration of *H. sapiens* MO could occur that allows for the recruitment of a second MCM–CDT1, similar to the *S. cerevisiae* system.

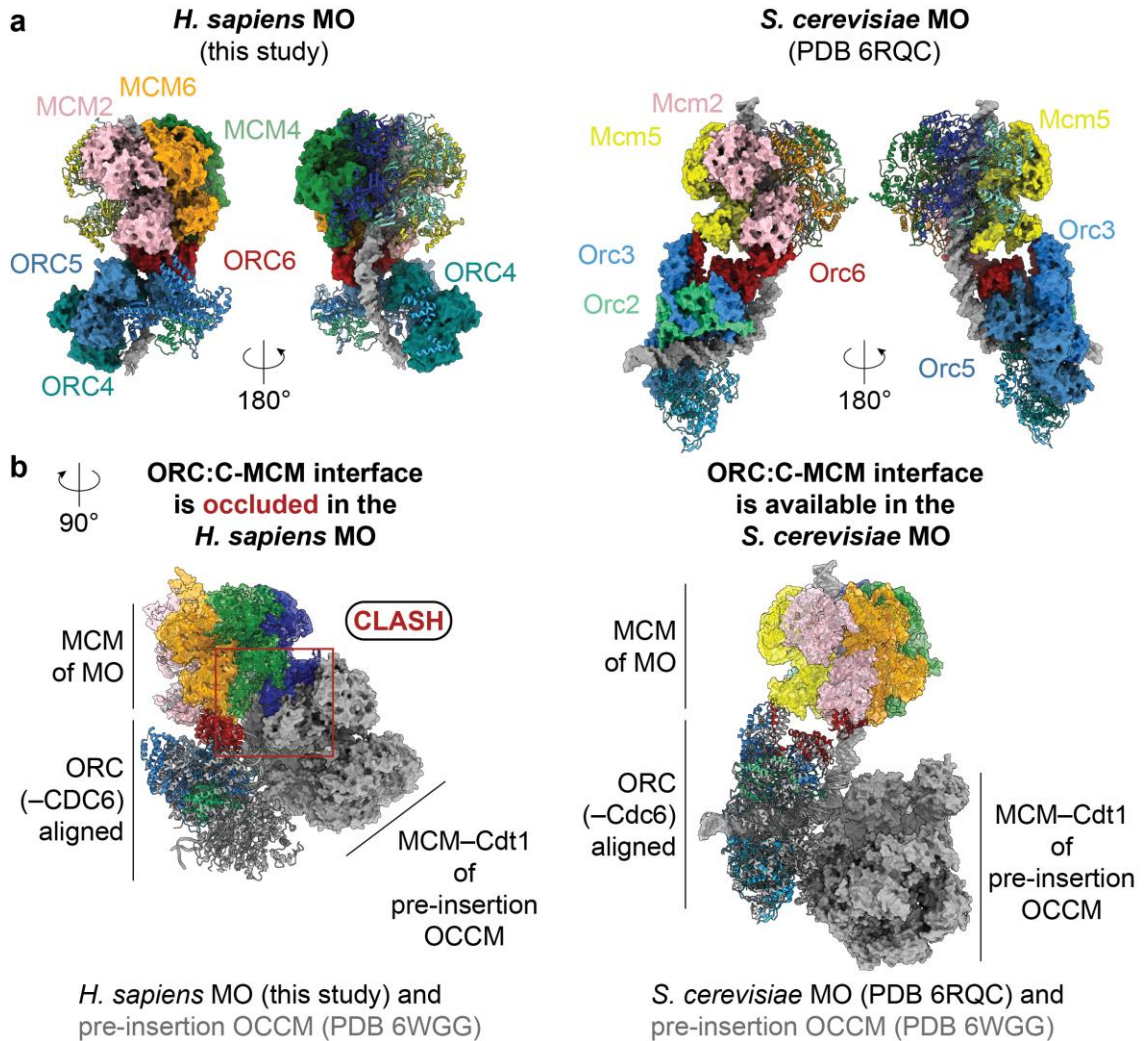


Figure 3.25: The *H. sapiens* MO is incompatible with recruitment of a second MCM helicase.

a. *H. sapiens* and *S. cerevisiae* MO structures. Subunits that interact with ORC6 are represented with surface rendering. **b.** Overlay of MO and pre-insertion OCCM models, which were aligned *via* ORC. The *H. sapiens* MO and pre-insertion OCCM complexes show a clash between the loaded and incoming MCM, which suggests that the *H. sapiens* MO complex is incompatible with loading of a second MCM ring. On the contrary, the equivalent *S. cerevisiae* complexes do not show a clash.

To corroborate the finding that *H. sapiens* ORC6 is not required for DH formation, I analysed the ATP-dependent licensing reaction in the presence and absence of ORC6, using negative stain EM (Figure 3.26a). Reactions, which were supplemented with ORC6 resulted in 2D classes of nucleosomes, ORC, SH, OCCM, MO and DH (Figure 3.26b). DHs were formed with equivalent efficiency both in the absence as well as the presence of ORC6. However, the MO intermediate could not be identified without ORC6 (Figure 3.26c). This finding confirms that ORC6 is required for MO, but not DH formation. Taken at face value,

these results indicate that *H. sapiens* DH loading onto duplex DNA does not depend on the MO intermediate. It is tempting to speculate that DHs could be formed by two opposed OCCM intermediates loading single hexamers that eventually become interlocked ([Figure 3.26d](#)).

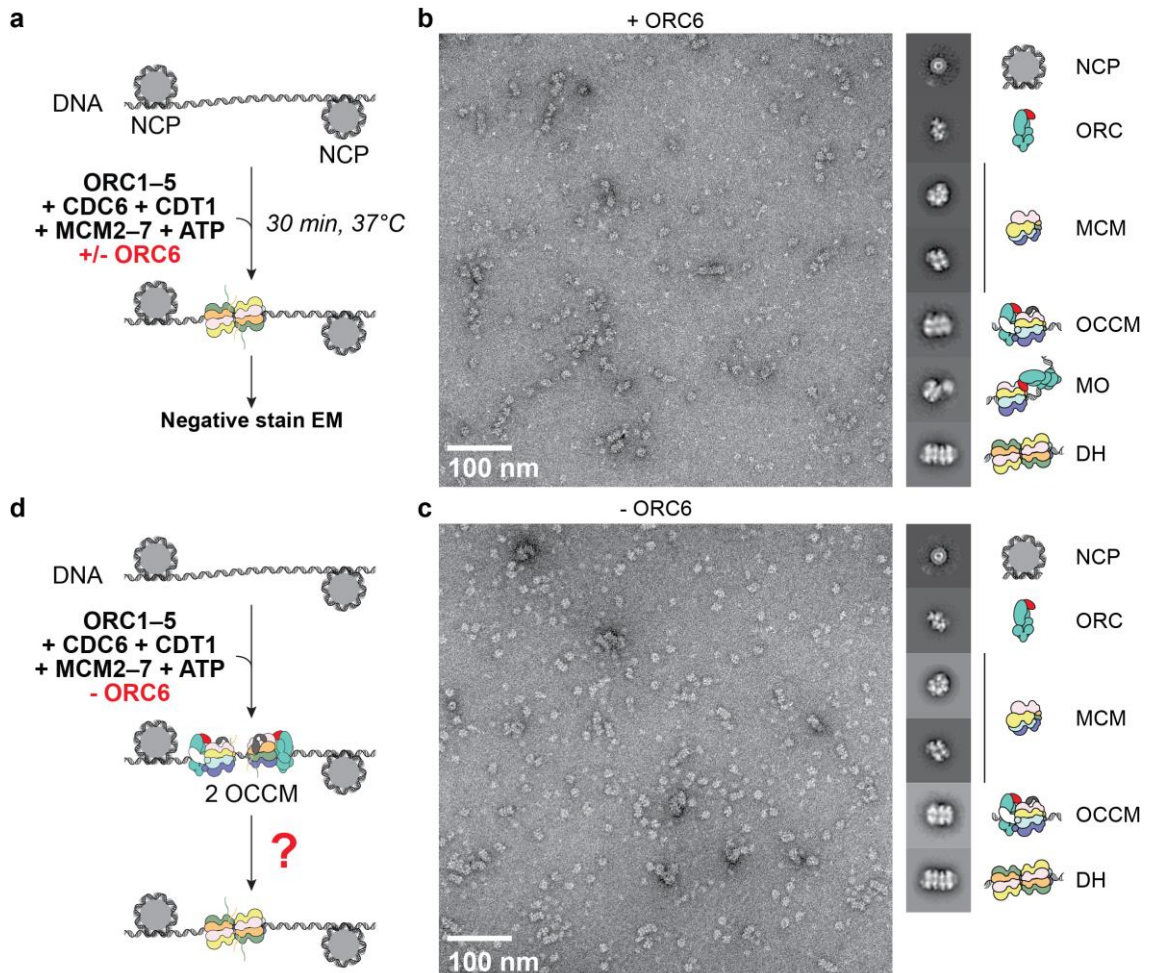


Figure 3.26: ORC6 is dispensable for DH formation.

a. Overview of the procedure to reconstitute origin licensing in the presence and absence of ORC6. **b.** Representative negative stain EM micrograph and 2D classes of an origin licensing reaction carried out with the addition of ORC6. **c.** Negative stain EM micrograph and 2D classes of a reaction without ORC6. **d.** Model of DH loading via the formation of two OCCM complexes in a head-to-head configuration.

3.7 Cryo-EM structure of the *H. sapiens* DH

3.7.1 Image processing of the *H. sapiens* DH

The ATP-dependent origin licensing reaction terminates with successful loading of the DH, encircling double-stranded DNA. My cryo-EM analysis of the origin licensing reaction yielded 49,485 DH particles, which were separated from other reaction intermediates in the dataset as described in paragraph 3.3.3. Reference-free 2D classification was then used to isolate high-quality particles that contributed to averages displaying secondary structure features (Figure 3.27). A total of 19,049 particles were used for *ab initio* reconstruction, followed by homogeneous, non-uniform and local refinement imposing C2 symmetry in cryoSPARC. The particle set was re-extracted in RELION, duplicates were removed, and 3D classification was carried out. A subset of 15,874 particles, which were refined to 3.8 Å, were subjected to Bayesian polishing (Zivanov et al. 2019) and reference-free 2D classification without alignment to select the best particles. CTF parameters were refined (Zivanov et al. 2020) in three consecutive rounds to yield a 3.3-Å resolution DH structure. The same structure was refined to 3.1 Å in cryoSPARC (Figure 3.28).

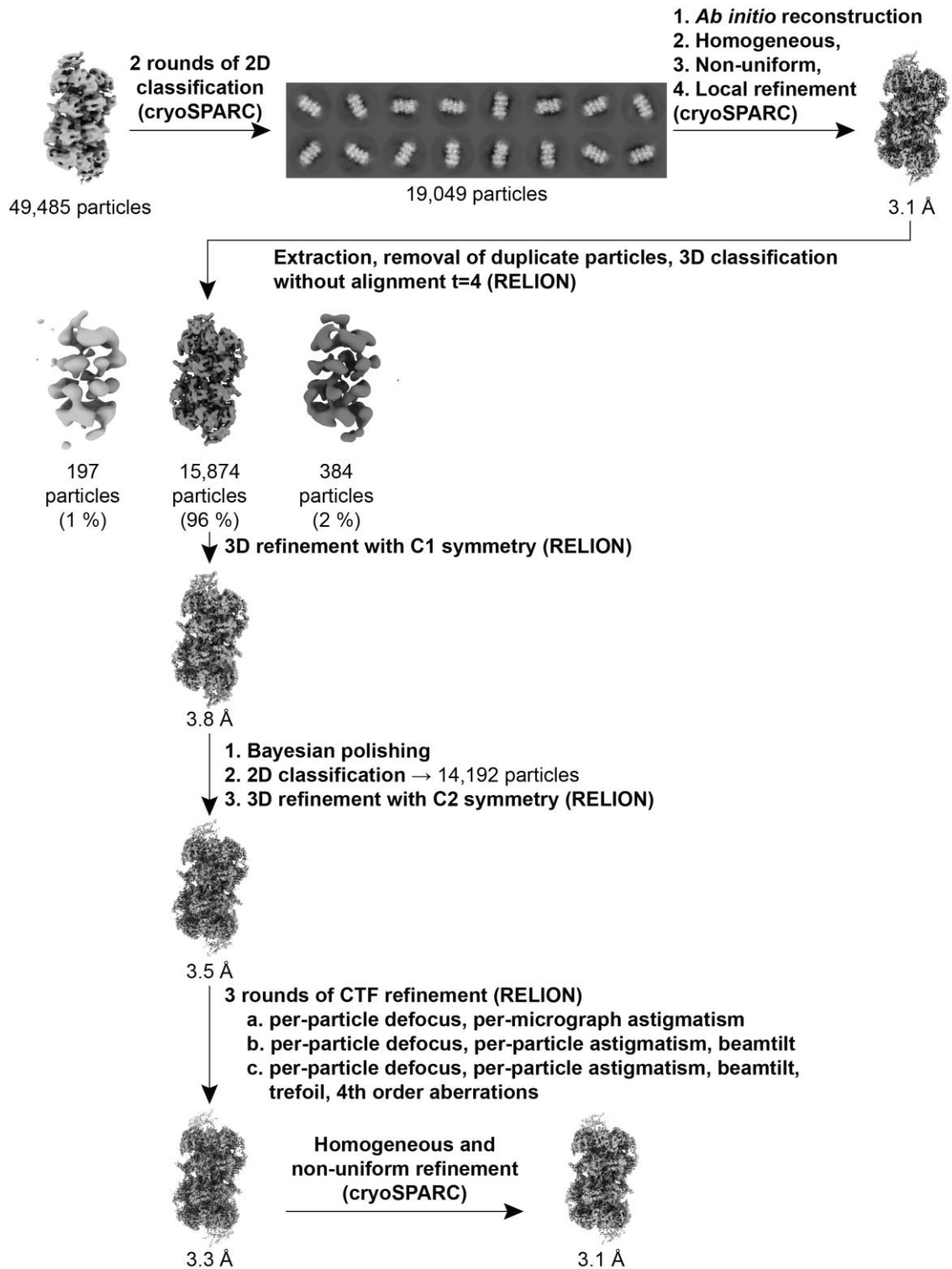


Figure 3.27: Processing pipeline for the *H. sapiens* DH.

Overview of the image processing procedure used to generate the 3.1-Å-resolution *H. sapiens* DH. Steps preceding this pipeline are described in chapter 3.3.3.

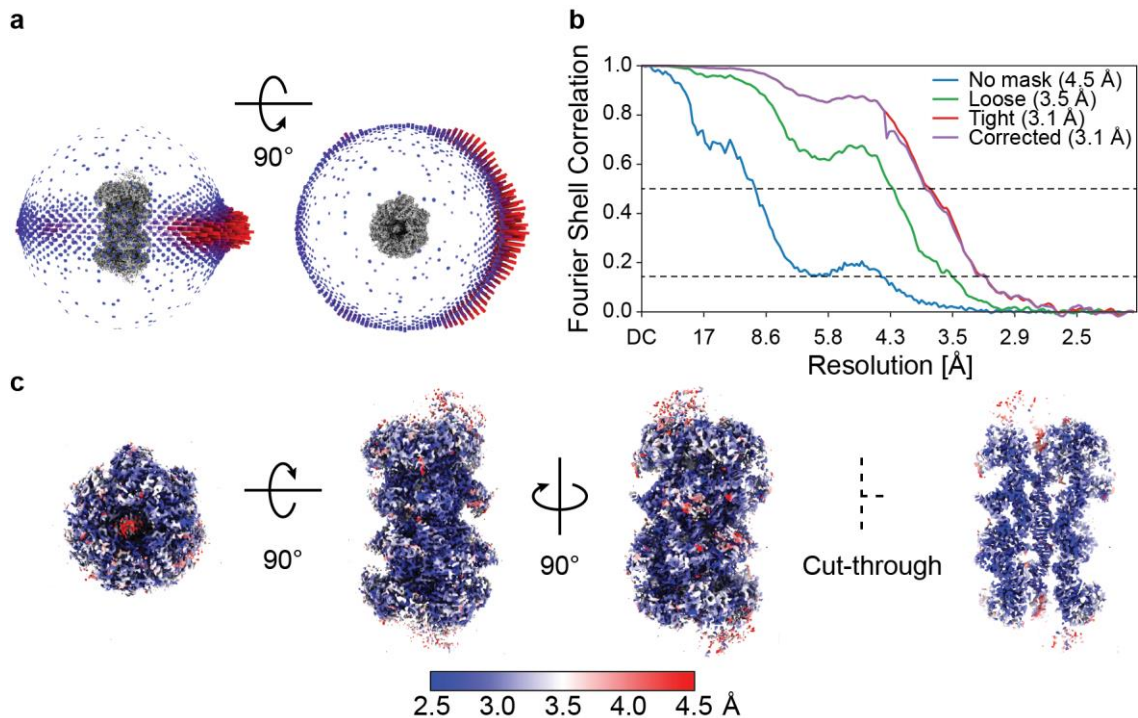


Figure 3.28: Structure quality of the DNA-loaded *H. sapiens* DH.

a. Angular distribution of the *H. sapiens* DH map. **b.** Fourier Shell Correlation plot for the 3.1-Å-resolution *H. sapiens* DH structure. **c.** Four views of the DH structure coloured according to the local resolution.

3.7.2 DH formation causes disruption of Watson–Crick base pairing

The atomic structure of the *H. sapiens* DH was obtained by docking the atomic model of the *H. sapiens* SH, which was generated in this study, into the cryo-EM density of both MCM rings. The model was inspected and adjusted in Coot v0.9-pre (Casañal et al. 2020). At a first glance, the overall architecture of the *H. sapiens* DH (Figure 3.29a) resembles that observed for *S. cerevisiae* proteins. The two MCM rings dimerise *via* their N-terminal tiers and have a lateral offset, whereby the central channels in the two MCM rings are slightly misaligned. A more detailed analysis, however, revealed significant differences in the interface between the two MCM hexamers, the protein–DNA interaction and the nucleotide state of the ATPase sites.

In *S. cerevisiae*, the DH is held together by interlocking of the N-terminal Zn finger domains of opposing MCM subunits. Additionally, a helix-turn-helix motif in Mcm7 contacts the Mcm5 subunit of the opposing MCM ring. This motif is 17 amino acids

shorter in the *H. sapiens* ortholog and other species and does not reach the opposing MCM5 subunit (Figure 3.29b and c). The double hexamer is therefore only stabilised by the interlocked Zn finger domains. When the atomic models of the *H. sapiens* and *S. cerevisiae* DH (PDB code 7P30) were overlaid via one single ring, a relative tilt of 3.4° of the second MCM hexamer away from the MCM5 subunit in the opposing ring with respect to the *S. cerevisiae* DH was observed in the *H. sapiens* complex (Figure 3.30a). Furthermore, the second MCM ring appears rotated around the DNA in respect to the *S. cerevisiae* complex. Due to the rotation and relative tilt, MCM7 from one ring gives the perception of drifting away from the juxtaposed MCM5 subunit in the second ring. Globally, the counter-clockwise rotation of one ring with respect to the other would result in an

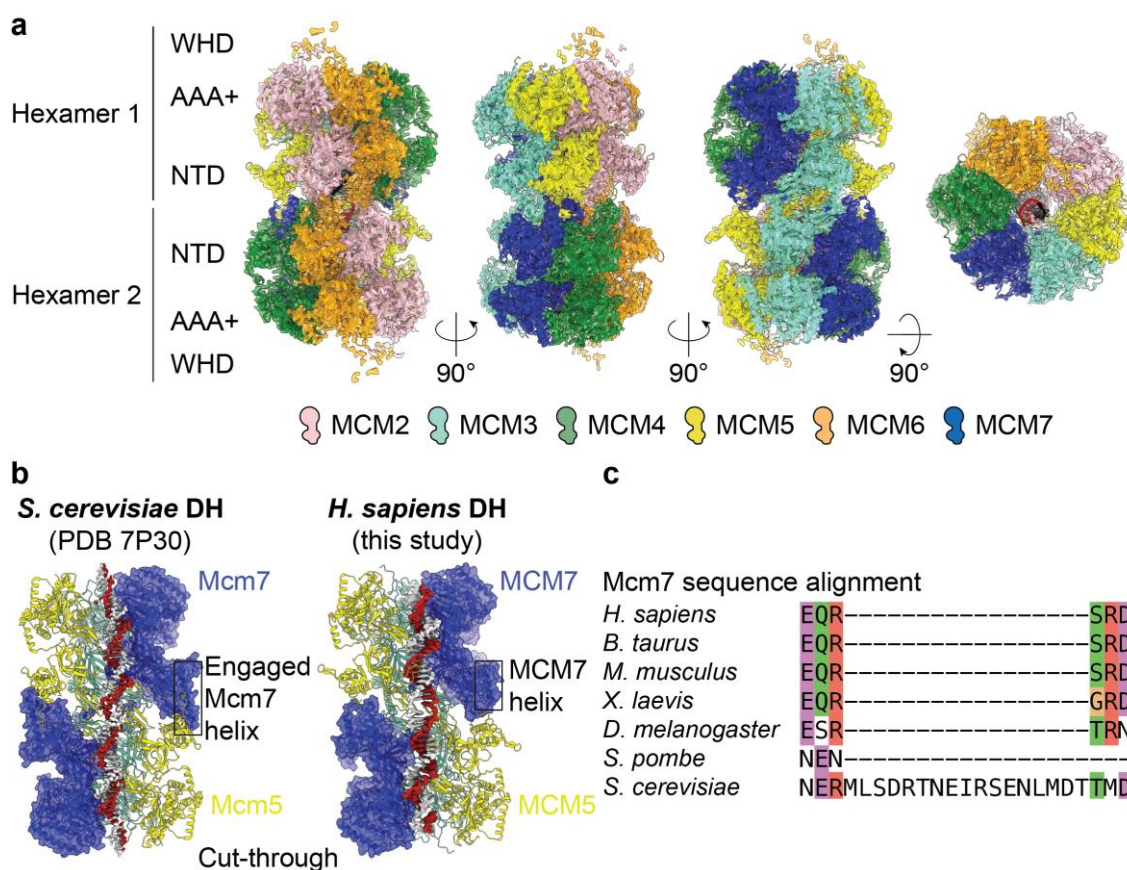


Figure 3.29: Overall structure of the *H. sapiens* DH and the MCM7 helix-turn-helix motif that stabilises the *S. cerevisiae* DH.

a. Four views of the 3.1-Å resolution *H. sapiens* DH structure. **b.** Cut-through view of the atomic model of *S. cerevisiae* (PDB 7P30) and *H. sapiens* DH, highlighting the extended MCM7 helix-turn-helix motif that interacts with MCM5 in the opposing MCM ring in the *S. cerevisiae* DH. **c.** Sequence alignment of the MCM7 helix-turn-helix motif. Residues are coloured based on physico-chemical properties and conservation (Clustal X).

underwinding effect for the duplex DNA harboured in the central channel. In this modelled rotation, the N-terminal tier (bearing N-terminal pore loops that contact DNA) and the ATPase tier (containing the PS1 β -hairpin and h2i pore loops) move *en bloc* (Figure 3.30b). Taken at face value, it appears that the longer *S. cerevisiae* Mcm7 helix-turn-helix motif might serve as a latch that locks the two MCM hexamers in a more retracted position.

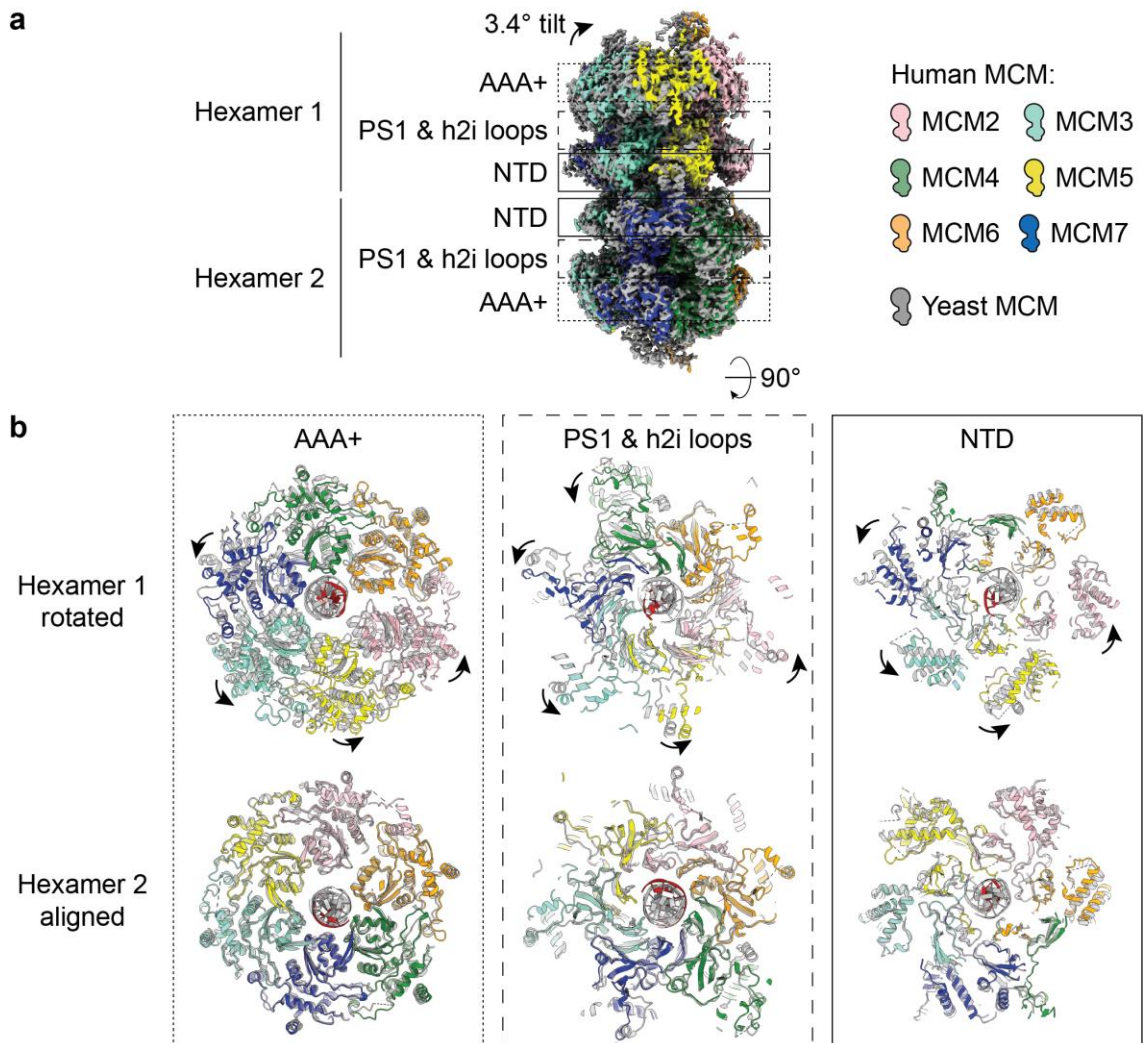


Figure 3.30: The two MCM rings in *H. sapiens* DH are tilted and rotated against each other.

a. Iso-surface representation of the *H. sapiens* DH (multicoloured) overlaid with the *S. cerevisiae* DH (EMBD-12176), which is shown in grey. ATPase- and N-terminal tiers as well as the position of the PS1 and h2i loops are highlighted. In the *H. sapiens* DH, hexamer 1 is tilted by 3.4° away from the MCM5 subunit in hexamer 2. **b.** Rotated, cut-through views of the domains indicated in a. Hexamer 2 of *H. sapiens* (multicoloured) and *S. cerevisiae* (grey) are aligned, while hexamer 1 shows a counter-clockwise rotation of the *H. sapiens* MCM compared to the *S. cerevisiae* MCM in the DH.

The apparent relative rotation of the two MCM hexamers in the *H. sapiens* and *S. cerevisiae* DH structures might explain differences in the state of duplex DNA underwinding. In fact, while the DNA in the *S. cerevisiae* DH mainly adopts B-form, with all modelled bases engaged in Watson–Crick base pairing, the DNA in the *H. sapiens* DH appears melted at the dimerization interface between the two MCM hexamers (Figure 3.31a). The two MCM rings grip the leading strand with the PS1 loops of MCM7, MCM4, MCM6 and the h2i loop of MCM2, forming a right-handed spiral with MCM7 at the 3'-end of the leading strand (Figure 3.31b). The same subunits are engaged with DNA in the *S. cerevisiae* DH (PDB code 7P30). However, in the *S. cerevisiae* structure, Mcm7, Mcm4 and Mcm6 interact with the lagging strand and not the leading strand (Figure 3.31c). Although the length of DNA in the *S. cerevisiae* and the *H. sapiens* structure is the same, the number of bases that make up the reconstructed double helix is smaller. In fact, the number of nucleotides spanning between the two MCM7 PS1 loops were found to be 30 in the *H. sapiens* structure, compared to 32 in the *S. cerevisiae* structure (Figure 3.31d). This observation corroborates the observation that the duplex DNA encircled by the *H. sapiens* DH is severely underwound.

The melted DNA at the N-terminal dimerization interface between MCM rings is stabilised by two elements contained in MCM5 from the two hexamers. First, L209 pushes two bases apart, disrupting their base pairing interaction and disturbing the base stacking with neighbouring bases from the same DNA strand (Figure 3.31e). The second element is a proximal arginine in MCM5 (R195), which stabilises the melted DNA in two ways. First, it establishes hydrogen bonds with the broken bases to effectively mimic the paired base. Second, it engages with the neighbouring base in a cation- π interaction to make up for the missing base stacking. While R195 is conserved across species, L209 is only found in metazoan and is replaced by a proline in *S. cerevisiae* (Figure 3.31f). This proline to leucine change might be a critical difference that explains why double hexamer loading is sufficient to nucleate DNA unwinding with *H. sapiens* proteins but not with *S. cerevisiae* proteins. To test this hypothesis, we are currently introducing humanising sequence changes in the *S. cerevisiae* protein.

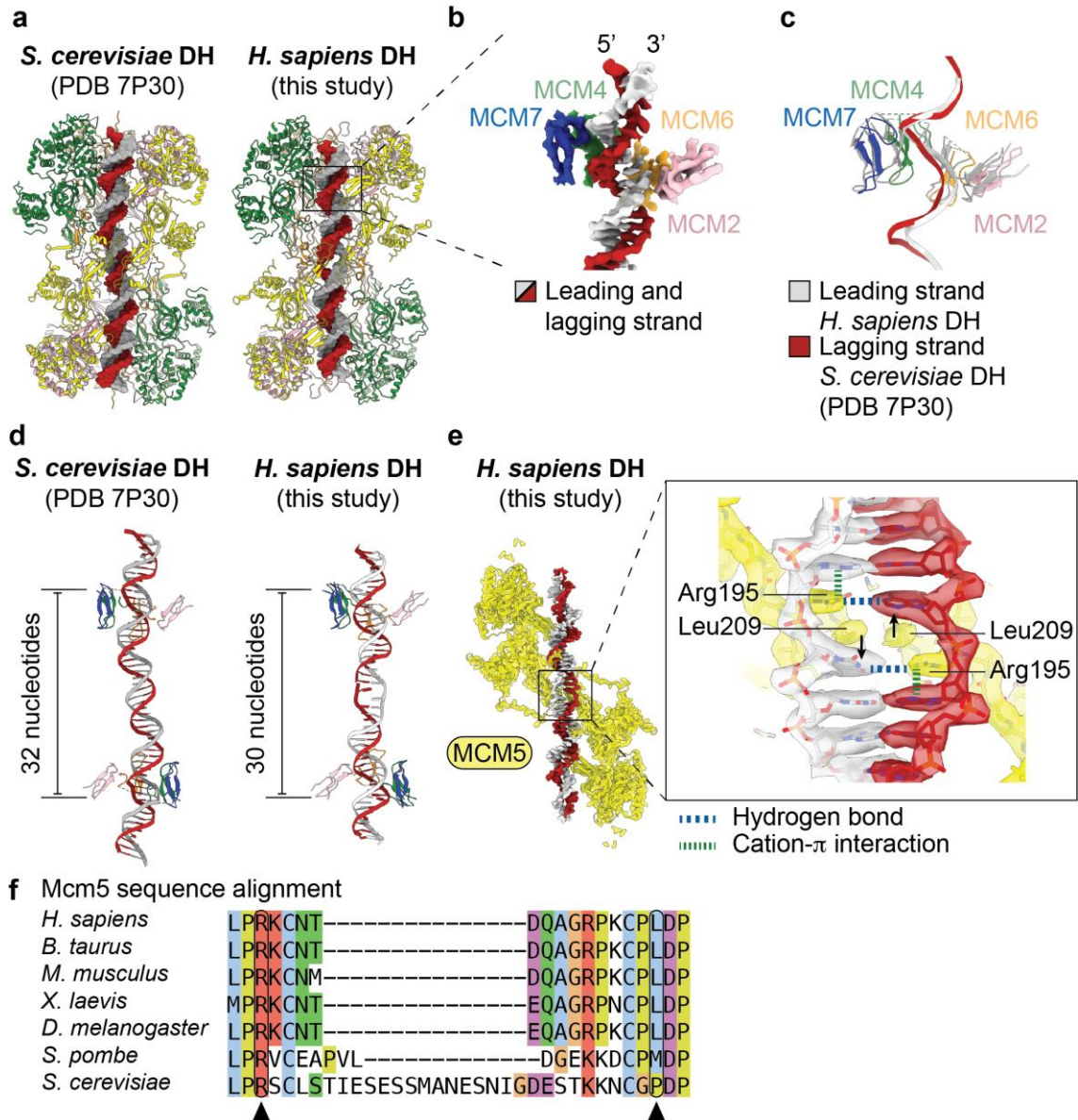


Figure 3.31: Several structural elements cause DNA melting in *H. sapiens* DH.

a. Cut-through view of the *S. cerevisiae* and *H. sapiens* DH atomic models with DNA in the central channel shown in surface representation. **b.** Segmented map of *H. sapiens* DH, showing the pore loops of MCM7, MCM4, MCM6 and MCM2 that form a right-handed spiral following the leading strand template. **c.** The leading strand bound in the *H. sapiens* DH overlays with the lagging strand that is contacted by the same MCM pore loops in the *S. cerevisiae* DH (PDB code 7P30). **d.** Atomic model the pore loops of both hexamers in the *S. cerevisiae* and *H. sapiens* DH that contact DNA. In *S. cerevisiae*, 32 nucleotides can be counted between the Mcm7 pore loops of both MCM hexamers, while only 30 nucleotides are found in the *H. sapiens* DH within the same region. **e.** Atomic model of the MCM5 subunits and the DNA in the *H. sapiens* DH, which contribute to DNA melting. The insert shows the iso-surface representation and atomic model of the melted DNA. L209 from both Mcm5 subunits push two bases apart. R195 engages in a hydrogen bond with the base and in cation- π interaction with the neighbouring base. **f.** Sequence alignment of the MCM5 region that interacts with the melted DNA. The R195 and L209 residues are highlighted with a triangle. Residues are coloured based on physico-chemical properties and conservation (Clustal X).

The SH observed in the *H. sapiens* origin licensing reaction engages in similar protein–DNA interactions as the *H. sapiens* DH. However, in the SH L209 does not disrupt the base pairing and R195 engages with the phosphate backbone of the leading strand instead of the base. This finding suggests that interaction of the two MCM rings is required to melt DNA.

3.7.3 ATPase state of *H. sapiens* DH

Structural differences between *H. sapiens* and *S. cerevisiae* DH are not only observed in the arrangement of the two MCM hexamers and the DNA, but also in the nucleotide binding state of the ATPase tier. ADP is found in sites mapping at the MCM2–5, MCM5–3 and MCM3–7 interfaces (Figure 3.32a), which matches the observation in the *S. cerevisiae* DH (Figure 3.32b). Conversely, the MCM7–4 site that contained weak nucleotide density for ADP in *S. cerevisiae*, does not harbour any nucleotide in the *H. sapiens* structure. Both MCM6–2 and MCM4–6 sites are occupied by ATP in *H. sapiens* DH, whereas the latter is free of nucleotide density in *S. cerevisiae* (Figure 3.6c). Therefore, two sites differ in their nucleotide occupancy comparing *H. sapiens* and *S. cerevisiae* proteins, namely the MCM7–4 and MCM4–6 ATPase centres. Previous work on single-stranded DNA translocation by the replication-fork engaged *D. melanogaster* CMG established that the nucleotide binding state dictates the DNA engagement state for the MCM ATPase pore loops (Eickhoff et al. 2019). In most cases, for CMG, ATP-bound MCM subunits contacted the DNA. A similar observation was made for the homodimeric and Pol- ϵ -engaged CMG (dCMGE) complex, an intermediate that follows DH formation on the path to maturation into active, fork engaged CMG (Lewis et al. 2022). The *H. sapiens* DH structure described in this chapter also appears to follow this general principle, given that the two ATP-bound subunits, MCM4 and MCM6, contain pore loop residues that are DNA engaged and form part of the pore-loop staircase contacting the leading strand DNA template. Exceptions exist, as for example the MCM7 subunit, which is at the top of the pore-loop staircase and is ATP bound in the CMG conformer “state 2A”, but nucleotide-free in the *H. sapiens* DH (Figure 3.32c). Apart from this one discrepancy in the nucleotide engagement profile, the ATPase state of the *H. sapiens* DH is consistent with the observation

that ATP binding promotes DNA engagement (Enemark and Joshua-Tor 2006, Goswami et al. 2018, Eickhoff et al. 2019, Rzechorzek et al. 2020, Yuan et al. 2020a). The *S. cerevisiae* DH, on the other hand, binds DNA differently compared to the translocating CMG. Here, all MCM subunits engage DNA and they do this in a fashion that has no impact on the DNA unwinding state (Figure 3.32c). It needs to be established whether ATP hydrolysis driven translocation has any role in bringing two hexameric helicases together, and whether it is the interlocking between two helicases that promotes DNA melting at the inter-ring interface.

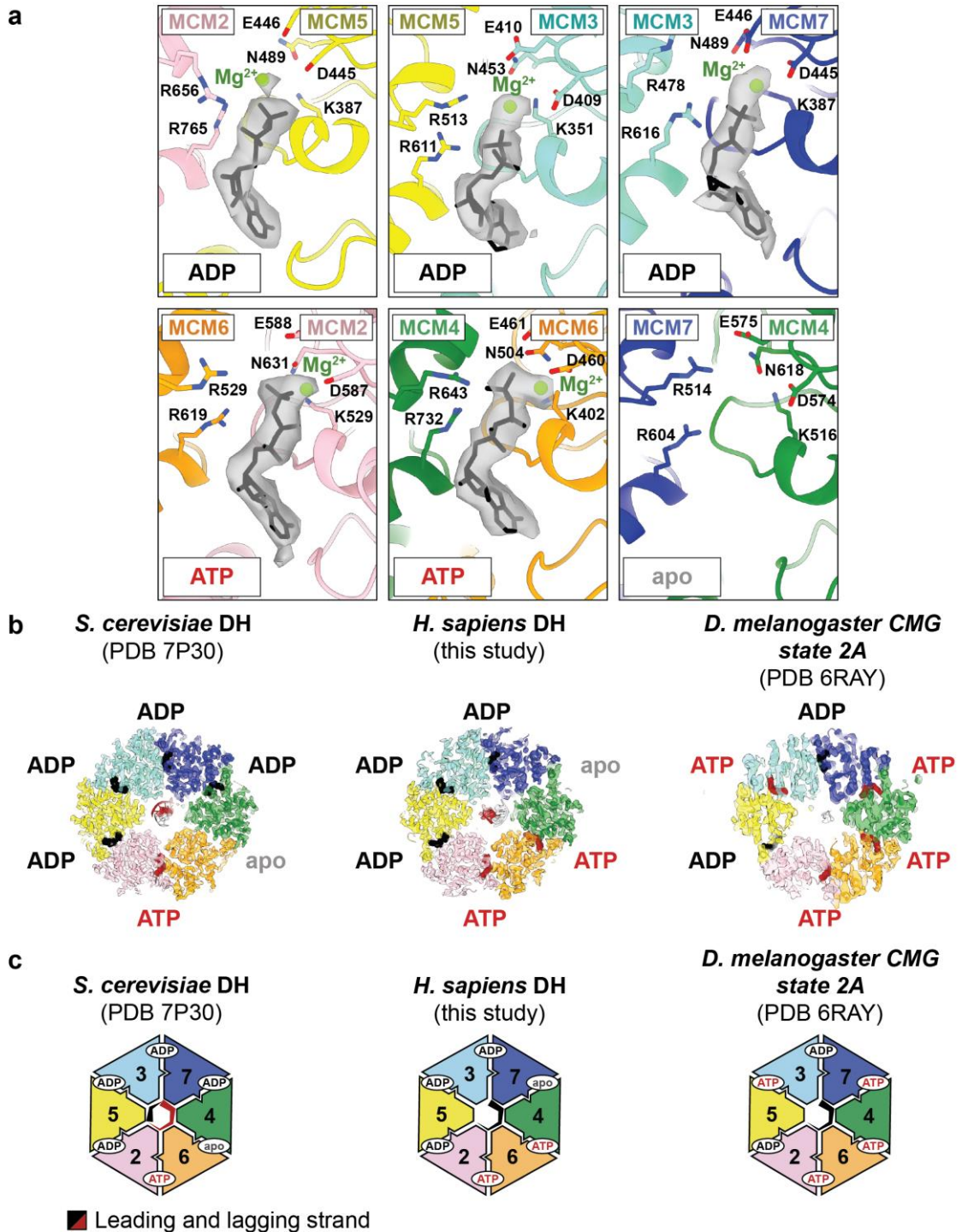


Figure 3.32: Nucleotide binding in correlation with pore loop binding.

a. ATPase sites of the *H. sapiens* DH with surface rendered nucleotides. **b.** Nucleotide occupancy of the *H. sapiens* DH compared to *S. cerevisiae* DH (PDB code 7P30) and *D. melanogaster* CMG state 2A (PDB code 6RAY). **c.** Cartoon representation of the nucleotide binding state of *H. sapiens* and *S. cerevisiae* DH as well as *D. melanogaster* CMG. DNA strands interacting with the PS1 and h2i loops of the different MCM subunits are shown (leading strand in black, lagging strand in red).

3.8 Discussion

3.8.1 The *S. cerevisiae* DH is an inactive form of the MCM helicase and does not unwind DNA

In this chapter, I have described the structure of the *S. cerevisiae* DH solved to 3.0 Å resolution. With this work, I corroborated previous findings indicating that phosphorylation does not induce significant structural changes in the loaded, inactive form of the replicative helicase (Abid Ali et al. 2017). Given the higher resolution of the here presented structure, I was able to identify specific amino acid sidechains interacting with the DNA and unambiguously assign nucleotide occupancy in the ATPase sites. Most ATPase centres are occupied by ADP, representing a post-catalytic state, while the Mcm4–6 site does not contain nucleotide density. These data are in line with biochemical and *in vivo* experiments showing that DH formation depends on ATP hydrolysis by the MCM (Coster et al. 2014). Coster et al. (2014) also established that a catalytically inactive *S. cerevisiae* Mcm4 variant has only mild loading defects, which correlates with the observation that Mcm4 does not bind any nucleotide in the loaded DH structure. In contrast, the Mcm6–2 active site harbours ATP. This observation is surprising given that an Mcm6 arginine-finger mutant, which is thought to be incompetent for ATP hydrolysis, does not allow MCM loading *in vitro* (Coster et al. 2014). A possible explanation for this finding is that ATP hydrolysis has occurred at this site, ADP has been released and a new ATP molecule bound, prior to structure determination. Alternatively, the arginine-finger variant could affect not only nucleotide hydrolysis, as the text-book view indicates, but also binding, which would thus impair helicase loading.

Duplex DNA in the central channel of the *S. cerevisiae* DH is slightly bent but maintains Watson–Crick base pairing throughout the length of the central channel. This agrees with the notion that the loaded helicase is inactive and does not contain any DNA-melting activity. Recent work from our laboratory and the laboratory of Dr John Diffley established that DNA melting is later nucleated upon recruitment of the helicase activators Cdc45 and GINS, in a process that requires release of ADP and binding of new ATP (Douglas et al. 2018, Lewis et al. 2022).

The firing factor Mcm10 is essential to trigger subsequent ATP hydrolysis, stimulate lagging strand ejection and promote origin DNA unwinding (Kanke et al. 2012, van Deursen et al. 2012, Watase et al. 2012, Douglas et al. 2018).

3.8.2 *In vitro* reconstitution of origin licensing using *H. sapiens* proteins

We have established conditions to reconstitute the origin licensing reaction *in vitro* using purified *H. sapiens* proteins. As in *S. cerevisiae*, DHs are loaded onto double-stranded DNA in an ATP-dependent manner when the MCM2–7 helicase is co-incubated with the loading factors ORC, CDC6 and CDT1. We used cryo-electron microscopy to visualise the entire origin licensing reaction and characterised not only the fully loaded DH, but also several other macromolecular assemblies, which are reminiscent of the loading intermediates visualised with *S. cerevisiae* proteins. Structures include the OCCM, an OCCM-like complex that lacks CDC6 (here described as OC1M), single MCM hexamers, which are loaded onto double-stranded DNA, as well as complexes in which ORC is bound to the N-terminal face of a single-loaded MCM ring (similar to the *S. cerevisiae* MO (Miller et al. 2019)). Despite similarities with the *S. cerevisiae* complexes, closer inspection revealed significant differences in the structural arrangement of protein domains, nucleotide state and the mode of DNA engagement. The work presented here is a snapshot of our ongoing efforts to understand the mechanism of MCM double hexamer formation in humans. Future work will include further refinement and optimisation of structure geometries.

The first captured intermediate on the path to DH formation is the OCCM complex, consisting of ORC, CDC6, CDT1 and MCM2–7 (Sun et al. 2013, Yuan et al. 2017, Yuan et al. 2020b). Here, the MCM interacts *via* the MCM3, 7 and 6 WHD and the MCM2 ATPase domain with ORC–CDC6. Unlike in the *S. cerevisiae* complex, no interaction between the MCM4 WHD and ORC is observed in the *H. sapiens* OCCM. In support of our data, loss of the MCM4 WHD domain allows for MCM loading, although to a reduced level (Guerrero-Puigdevall et al. 2021). In *S. cerevisiae* OCCM, Orc6 binds to Orc3 and DNA, which allows ORC to bend and correctly position the DNA for the loading of the MCM helicase (Yuan et al. 2020b).

In our *H. sapiens* data, no density for ORC6 was observed. Negative stain imaging of a reaction carried out in the absence of ORC6 also gave rise to 2D classes resembling the OCCM, which supports the notion that ORC6 is not required for OCCM formation in the human system (Figure 3.26c). Due to low conservation between species, the role of metazoan ORC6 during DNA replication has long been debated (Prasanth et al. 2002, Thomae et al. 2011, Liu et al. 2012). Further analysis is required to establish its exact function.

We also captured an OCCM-like complex that lacks CDC6. Working under the assumption that, like in *S. cerevisiae*, initial MCM recruitment requires CDC6 (Fernández-Cid et al. 2013, Frigola et al. 2013, Yuan et al. 2020b), I speculate that the new complex represents a maturation of the OCCM complex, where CDC6 has dissociated from MCM and ORC on the path to single hexamer loading. Alternatively, this OC1M complex could represent an off-path protein assembly.

In the next stage of the helicase loading reaction, the loading factors are released and the MCM ring closes around double-stranded DNA. Different from the *S. cerevisiae* system, we detected a significant number of DNA-loaded MCM rings (SH) working with *H. sapiens* proteins, which enabled me to determine their structure. The MCM ring in the SH adopts a similar configuration as in the DH, although no evidence for DNA melting at the MCM N-terminal face was detected. Three of the ATPase sites are occupied by ADP, corroborating the notion that loading of the MCM hexamer requires ATP hydrolysis (Coster et al. 2014).

Working with *S. cerevisiae* proteins we have previously shown that ORC supports DH formation by binding to the N-terminal face of a loaded SH, with Orc6 functioning as a bridging element (Miller et al. 2019). The resulting MO complex recruits a second MCM ring *via* interactions with the C-terminal face of the incoming helicase, *i.e.* *via* the same OCCM mechanism that supports loading of the first hexamer. With *H. sapiens* proteins we also observe ORC binding at the N-terminal side of the SH, however, the subunits bridged by *H. sapiens* ORC6 are different from the yeast counterpart. In fact, *H. sapiens* ORC6 connects ORC4–5 with MCM2–6–4, while *S. cerevisiae* Orc6 bridges between Orc2–3–5 and MCM2–5. The alternate configuration observed with *H. sapiens* proteins is incompatible

with the recruitment of a second MCM *via* the OCCM mechanism, suggesting that *H. sapiens* MO might not be a loading intermediate. Coherently, we observed that omission of ORC6 prevents MO formation but does not affect DH formation. This further supports the notion that ORC6 in the *H. sapiens* helicase loading reaction might have different roles compared to *S. cerevisiae* (Dhar and Dutta 2000, Gillespie et al. 2001, Chesnokov et al. 2003, Huijbregts et al. 2009). Conversely, *H. sapiens* cell lines depleted for ORC6 showed defects in origin licensing and DNA replication (Prasanth et al. 2002, Stiff et al. 2013). The carboxy-terminal helix of ORC6, which has been shown to mediate interaction with ORC3 both in *D. melanogaster* as well as *H. sapiens* (Bleichert et al. 2013), appears to be important, as mutations in this domain can be found in patients with the Meier-Gorlin syndrome. How the loading of two MCMs is coordinated, if not *via* the MO, remains to be investigated. It is not unlikely that higher eukaryotes have evolved additional mechanisms to load the MCM onto DNA to contribute to the licensing of the required number of origins, even when factors, like ORC6, are limiting. Therefore, two OCCMs, loaded in a head-to-head configuration, could potentially lead to DH formation.

3.8.3 Nucleation of DNA melting by the *H. sapiens* DH

We have also characterised the *H. sapiens* DNA-loaded DH to find unexpected differences in the duplex DNA structure, suggesting that the mechanism of origin unwinding contains fundamental differences compared to yeast. In the *S. cerevisiae* DH structure, DNA becomes slightly bent but maintains B-form DNA character, with Watson–Crick base pairing maintained throughout the length of the MCM channel. Previous biochemical and structural work indicated that DNA untwisting and nucleation of DNA melting is only achieved upon recruitment of the firing factors Cdc45 and GINS (Douglas et al. 2018, Lewis et al. 2022). Base flipping is thereby promoted by a nucleotide driven reconfiguration of pore loops in the ATPase domain. Only in the presence of the firing factor Mcm10, further DNA unwinding is promoted (Douglas et al. 2018). Nucleation of DNA melting by *H. sapiens* MCM is totally different, as DH formation alone is sufficient to disrupt Watson–Crick base pairing. Several structural features contribute to the DNA

untwisting activity of the DH. Each MCM hexamer interacts with one DNA strand and causes DNA to become stretched and untwisted. Concomitantly, L209 of both MCM5 subunits cause disruption of one base pair at the dimerization interface of the DH. R195 in MCM5 stabilises the broken base pairing by establishing new hydrogen bonds, and forms cation- π interactions to stabilise broken base stacking. A possible explanation for the absence of DNA melting in the *S. cerevisiae* DH is the change from leucine to proline in the *S. cerevisiae* Mcm5. Interestingly, the SH establishes the same protein–DNA interactions as the *H. sapiens* DH, with the exception of the centrally located MCM5 L209 and R195, which do not interact with the bases, but rather contact the phosphate backbone. These observations indicate that it is DH formation that causes the DNA to melt rather than the protein–DNA interactions formed within each hexamer.

One could speculate that subsequent lagging strand ejection might be facilitated by DNA melting in the *H. sapiens* DH so that more origins can be activated, even when firing factors are limited (for example in fast dividing cells). The firing factor MCM10 can indeed be limiting in higher eukaryotes as it has been shown to be short-lived in *D. melanogaster* (Christensen and Tye 2003) and protein levels are regulated by phosphorylation- and ubiquitin-dependent degradation in *H. sapiens* (Izumi et al. 2001). Contrarily, Mcm10 protein levels are constant throughout the cell cycle in *S. cerevisiae* (Homesley et al. 2000). Mcm10 is essential for DNA replication in yeast (Kanke et al. 2012, van Deursen et al. 2012, Watase et al. 2012), whereas MCM10 is not required for replication fork establishment in metazoans and rather influences replication fork stability and elongation rates (Park et al. 2008, Chadha et al. 2016). Furthermore, the melted DNA could restrict movement of the DH and thereby define origins of replication in higher eukaryotes where ORC does not bind DNA in a sequence specific manner. Accordingly, it has recently been shown that *H. sapiens* DHs are mostly found on DNA sequences, which are AT-rich in the centre of the DH, where base pairing is disrupted (Li et al. 2023). In the same study, all DHs, which had been isolated from DNase treated *H. sapiens* chromatin, were DNA-bound even though no roadblocks prevented the helicase from sliding off the DNA ends. These results indicate that *H. sapiens* DH are indeed more stationary than their *S. cerevisiae* counterparts.

3.8.4 ATP hydrolysis and DNA binding during origin licensing

It is well established that the loading of the MCM helicase depends on ATP hydrolysis (Coster et al. 2014). The set of structures I have described in this study suggests that the same ATPase dependency might also apply to the loading of the *H. sapiens* helicase, in that most sites in MCM are found in a post-catalytic state once the helicase ring is locked around DNA. Surprisingly however, the MCM6–2 and MCM4–6 ATPase sites are occupied by ATP in all structures, irrespective of ring closure, which would suggest that hydrolysis at these sites might not be required. Taken at a face value, this observation is in contrast to what is observed in *S. cerevisiae*, where an ATP-hydrolysis deficient variant of Mcm6 prevents DH formation (Coster et al. 2014). Alternatively, multiple cycles of ATP hydrolysis and nucleotide exchange might occur as part of the helicase loading mechanism.

In accordance with patterns observed for the CMG bound to a DNA fork (Goswami et al. 2018, Eickhoff et al. 2019, Rzechorzek et al. 2020, Yuan et al. 2020a), the ATP-bound MCM subunits contact the DNA *via* their ATPase pore loops in the loading intermediates and the DH assembled with *H. sapiens* proteins. In all these structures, the same MCM subunits contact the DNA.

I also found fundamental differences between the *S. cerevisiae* and *H. sapiens* MO, which imply that in the *H. sapiens* system loading of the second MCM ring might not include MO formation. An alternative mechanism could involve the encounter of two loaded SHs. We speculate that ATP-hydrolysis driven translocation of SHs along duplex DNA might support DH formation. Alternatively, passive sliding in the spatially restricted chromatin context could promote two hexamers to collide and engage in a DH.

Chapter 4. Results 2 – Structural mechanism for the selective phosphorylation of DNA-loaded DHs by the Dbf4-dependent kinase

4.1 Introduction

The Mcm2–7 helicase is loaded onto double-stranded DNA as an inactive head-to-head double hexamer (Abid Ali et al. 2017, Noguchi et al. 2017). Two helicase activators, Cdc45 and GINS, are then recruited to each hexamer to form two replicative CMG helicases, which unwind duplex DNA upon recruitment of the firing factor Mcm10. DDK initiates the activation process by phosphorylating the unstructured amino-terminal tails of Mcm4 and Mcm6, which promotes the recruitment of Cdc45 by the firing factor Sld3–7 (Francis et al. 2009, Sheu and Stillman 2010, Sheu et al. 2014, Deegan et al. 2016, Sheu et al. 2016). Importantly, only Mcm4 phosphorylation is essential for origin activation. It is known that DDK preferentially phosphorylates DNA-loaded MCMs (Francis et al. 2009), but the molecular basis is not understood. Biochemical and genetic data have provided some insights on how DDK interacts with the MCM. It has been shown that the N-terminal BRCT domain of the kinase activator Dbf4 interacts with Mcm2 (Francis et al. 2009, Ramer et al. 2013, Abd Wahab and Remus 2020), while the Dbf4 C domain interacts with Mcm4 and is essential for origin activation and progression through S phase (Jones et al. 2010). However, the interaction has not been understood at the molecular level.

Structural information of DDK is available for the catalytic core of *H. sapiens* CDC7 in complex with the M and C domains of DBF4 from X-ray crystallography (Hughes et al. 2012, Dick et al. 2020). The atomic models give important insights into the kinase fold and how a substrate peptide is engaged in the active site to be phosphorylated. Additionally, a crystal structure of the N-terminal Dbf4 domain, spanning residues 65–221, has been shown to contain a BRCT motif (Matthews et al. 2012). The BRCT domain is flexibly tethered to the core of the kinase *via* a 40-amino-acid linker. The flexible attachment possibly plays an important role in the phosphorylation of the MCM complex as Mcm2 and Mcm4, which have been

shown to interact with DDK, are not neighbouring subunits in the ring-shaped MCM hexamer. To understand the molecular mechanism of DH phosphorylation, I first sought to acquire structural information on the full length DDK complex.

Kinases are regulated in multiple ways to allow for timely phosphorylation and substrate specificity (Gógl et al. 2019, Cullati et al. 2022). For DDK, a reduction in kinase activity has been observed upon autophosphorylation (Dick et al. 2020). Due to the efficiency of autophosphorylation and mM-level cellular concentrations of ATP, DDK is presumable phosphorylated in cells. It remains unclear whether this modulation in phosphorylation activity has any effect on substrate selectivity.

DDK is also regulated when DNA damage is detected in S phase. The checkpoint kinase Rad53 thereby phosphorylates the kinase activator Dbf4 to prevent CMG helicase maturation and origin firing (Santocanale and Diffley 1998, Lopez-Mosqueda et al. 2010, Zegerman and Diffley 2010). After DNA lesions are resolved, inhibition is reversed and DNA replication can be initiated again. While 19 Rad53 phosphorylation sites on Dbf4 have been identified (Zegerman and Diffley 2010), it is unknown how this phosphorylation abrogates MCM phosphorylation by DDK and blocks origin firing. Independent of phosphorylation, Rad53 has been described to compete with the MCM for binding to DDK, therefore preventing MCM phosphorylation (Abd Wahab and Remus 2020). Further investigations are required to establish the molecular mechanism of DDK inhibition by Rad53 through phosphorylation and competitive binding.

To uncover the structural basis of origin activation by DDK and its modulation, I imaged full-length *S. cerevisiae* DDK in isolation as well as in complex with the DNA-loaded MCM double hexamer by cryo-electron microscopy. Combined with *in vitro* phosphorylation experiments and negative stain EM analysis using different DDK variants and DDK treated with Rad53, I addressed the following questions:

1. How does DDK selectively recognise MCM double hexamers?
2. How does DDK autophosphorylation affect substrate selectivity?
3. How does Rad53 interfere with DH phosphorylation by DDK to prevent late origin firing when DNA damage is detected?

4.2 Structural characterisation of the isolated DDK

4.2.1 Structural characterisation of full-length *S. cerevisiae* DDK

To investigate whether full-length *S. cerevisiae* DDK is a suitable target for single-particle cryo-EM analysis, I expressed the dimeric protein complex using the previously described yeast strain ySDK8 (On et al. 2014). This strain contains Cdc7 and N-terminally CBP-tagged Dbf4 under the control of a bidirectional galactose-inducible promoter. After cell lysis and isolation of the soluble phase, DDK was immobilised on calmodulin beads (Figure 4.1). Extensive washing with buffer first, and ATP-containing buffer later, removed unbound proteins and heat shock proteins. To reverse autophosphorylation that occurred in cells and upon incubation with ATP during the purification, DDK was treated with λ phosphatase. In the original purification protocol, the complex was then eluted and directly subjected to size exclusion chromatography. However, I found that adding a HiTrap heparin column step before size exclusion chromatography improved the purity of DDK. The purified DDK was homogeneous and monodisperse when analysed by negative stain EM (Figure 4.1e). Despite the intrinsic low resolution, reference-free 2D classification revealed a globular complex with distinct structural features, which appeared suitable for structural characterisation.

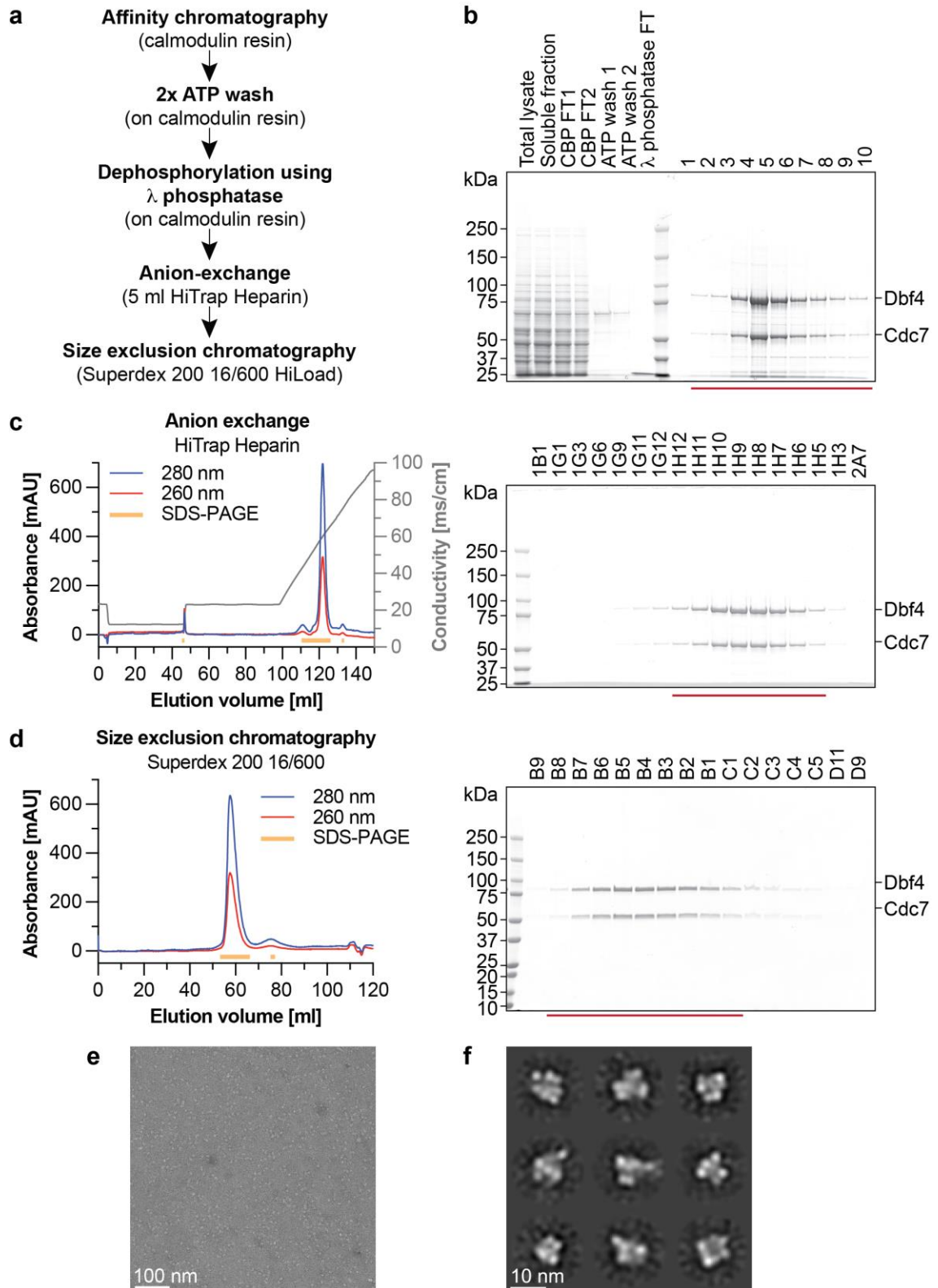


Figure 4.1: Purification of full-length *S. cerevisiae* DDK.

a. Outline of the purification protocol. **b.** Coomassie-stained SDS-PAGE gel of samples before CBP-pulldown and eluted fractions. Pooled fractions are indicated in red. **c.** Chromatogram from a 5 ml HiTrap Heparin column and Coomassie-stained SDS-PAGE of eluted fractions. **d.** Chromatogram from Superdex 200 16/600 HiLoad column

and Coomassie-stained SDS-PAGE gel of eluted fractions. **e.** Negative stained electron micrograph of the purified full-length DDK. **f.** 2D class of full-length DDK.

I then analysed the 140-kDa full-length DDK by cryo-EM. DDK was applied to 1.2/1.3 holey gold grids (UltrAuFoil), onto which a monolayer of graphene oxide flakes was deposited. Cryo grids were imaged on a Titan Krios electron microscope equipped with a K2 Summit direct electron detector and a BioQuantum energy filter (with support from Dr Andrea Nans) (Figure 4.2a). DDK particles were picked from motion-corrected movies, for which CTF parameters were estimated. Several rounds of reference-free 2D classification in RELION (Zivanov et al. 2018) led to the identification of 399,580 particles displaying high-resolution features (Figure 4.2b). Most molecules adopted a preferred orientation on the graphene oxide, which impaired isotropic three-dimensional reconstruction.

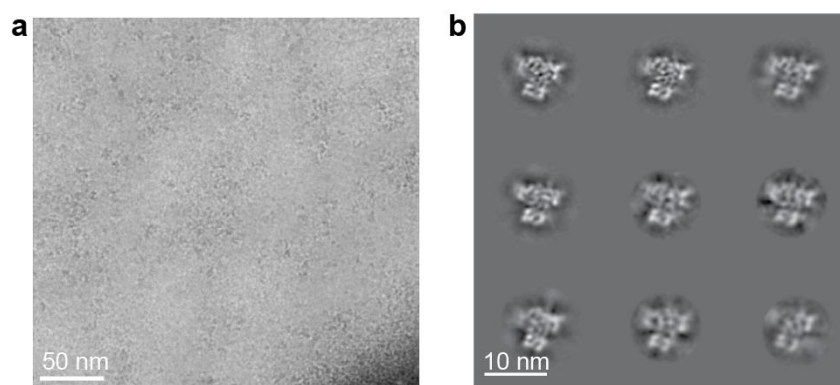


Figure 4.2: *S. cerevisiae* DDK analysed by cryo-EM.

a. Example cryo-EM micrograph of full-length *S. cerevisiae* DDK. **b.** 2D classes of DDK.

In summary, the purified full-length *S. cerevisiae* DDK adopted a globular conformation and showed high-resolution features. Knowing that full-length DDK can be analysed by cryo-EM, I next addressed how DDK recognises the MCM double hexamer substrate.

4.3 Structure of DH-DDK and molecular mechanism for substrate selectivity

4.3.1 *In vitro* reconstitution of DH-DDK

To characterise the interaction between DDK and the DNA-loaded DH, I reconstituted origin licensing *in vitro* using purified *S. cerevisiae* proteins (Yeeles et al. 2015, Miller et al. 2019) and added DDK to the reaction (On et al. 2014) (Figure 4.3a). A DNA template containing the *S. cerevisiae* ARS1 origin was utilised. Covalently-linked *Hpa*II methyltransferases at either end of the origin sequence were added to prevent the MCM helicase from sliding off the DNA. Following the origin licensing and phosphorylation reaction, DNA-bound complexes were immobilised on streptavidin paramagnetic beads via a desthiobiotin tag on the 5' end of the DNA. After washing the beads, protein complexes were eluted by treatment with micrococcal nuclease (MNase). When the beads were washed with a low-salt buffer, containing 300 mM sodium acetate, both Dbf4 and Cdc7 were observed alongside the six MCM subunits on an SDS-PAGE gel (Figure 4.3b). Mcm4 and Mcm6 migrated slower during electrophoresis, reflecting phosphorylation by DDK. Conversely, the DH-DDK interaction was lost when a high-salt buffer with 500 mM sodium chloride was used. This observation indicates that DDK efficiently interacts with and phosphorylates the DH but fails to withstand harsher purification conditions.

Aiming to improve sample homogeneity for subsequent structural characterisation, I tested two approaches to purify DNA-loaded DH. First, DNA-affinity purification and, second, anion exchange chromatography (Figure 4.3c). In the DNA-affinity purification, unbound loading factors and loading intermediates were removed by treating the immobilised DNA-bound complexes with buffer containing 500 mM sodium chloride (high-salt buffer). DNA and DNA-bound DH were then eluted using biotin, which binds to streptavidin with higher affinity than the desthiobiotin-conjugated DNA. Negative stain EM analysis of the eluate showed monodisperse DH particles on the grid (Figure 4.3d). In the second purification protocol, the licensing reaction was subjected to a MonoQ 1.6/5 anion exchange column (Figure 4.3e). A 40 CV gradient from 100 mM potassium acetate to 2 M sodium chloride

allowed for separation of DNA-loaded DH, free DNA and non-DNA bound factors. Negative stain EM analysis of the peak fraction confirmed purity of the DH (Figure 4.3f).

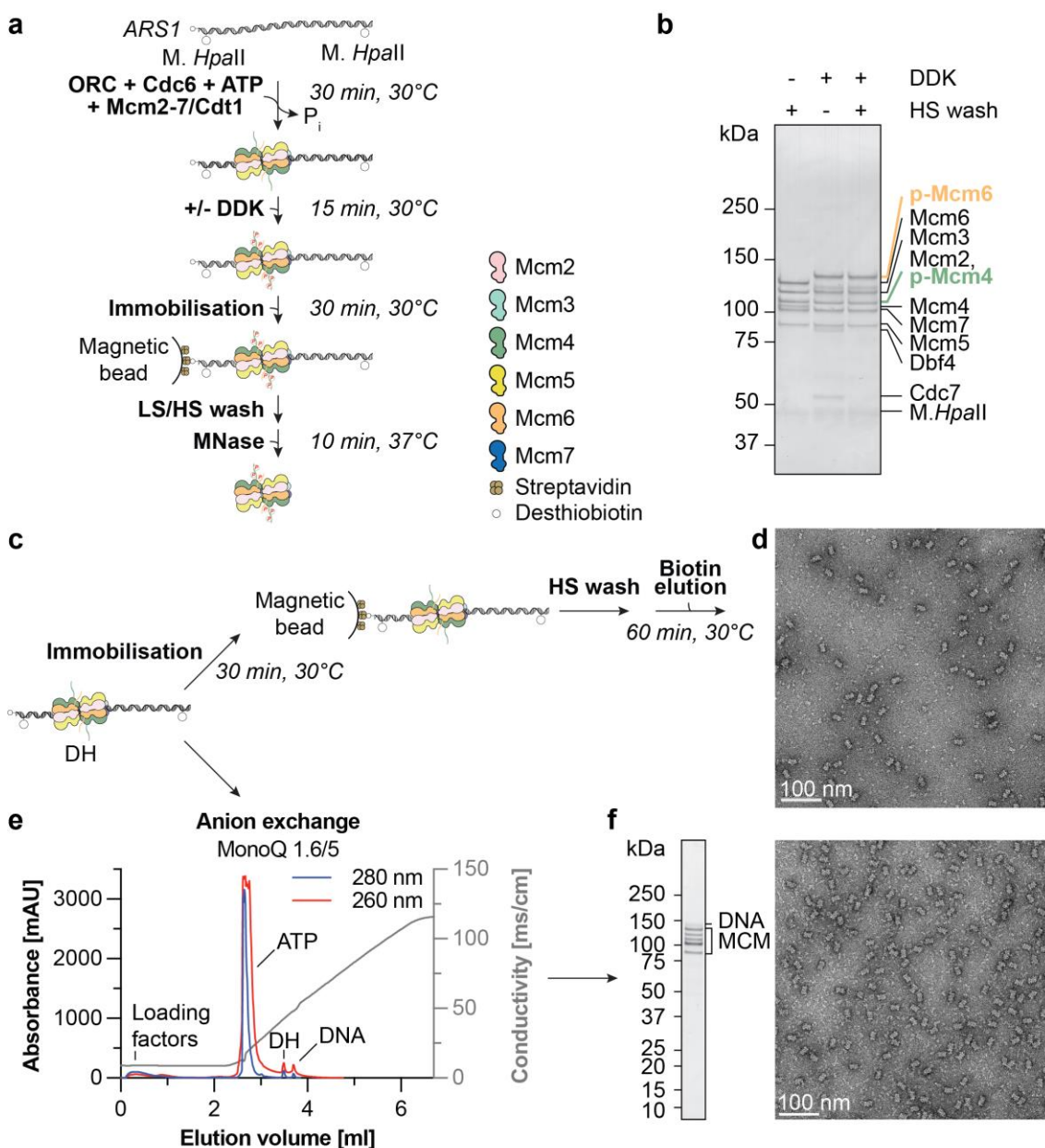


Figure 4.3: DH loading, phosphorylation and purification.

a. Overview of the DH loading and phosphorylation procedure for analysis by SDS-PAGE. **b.** Silver-stained SDS-PAGE gel of eluted DH. **c.** Overview of the two approaches to purify DNA-loaded DH. **d.** Negative stained electron micrograph showing DH after elution from streptavidin paramagnetic beads. **e.** Chromatogram of the DH loading reaction from a MonoQ 1.6/5 anion exchange column. **f.** SDS-PAGE gel and negative stained electron micrograph of the fraction containing DNA-loaded DH.

The DH, which was isolated using DNA-affinity purification or anion exchange chromatography with a subsequent dialysis into the reaction buffer, served as a substrate for DDK recruitment. The reaction was then analysed by negative stain EM ([Figure 4.4a](#)). Reference-free 2D classification revealed that additional density appeared next to the dimerization interface of the two MCM hexamers in a DH when DDK was added ([Figure 4.4b](#)). The density was rather weak and mostly visible on one side of the DH, indicating that most DHs, which contain two copies of each DDK-target subunit, were only partially occupied. Increasing DDK concentration did not enhance DDK occupancy but rather resulted in protein aggregation, which prevented visualisation of DHs.

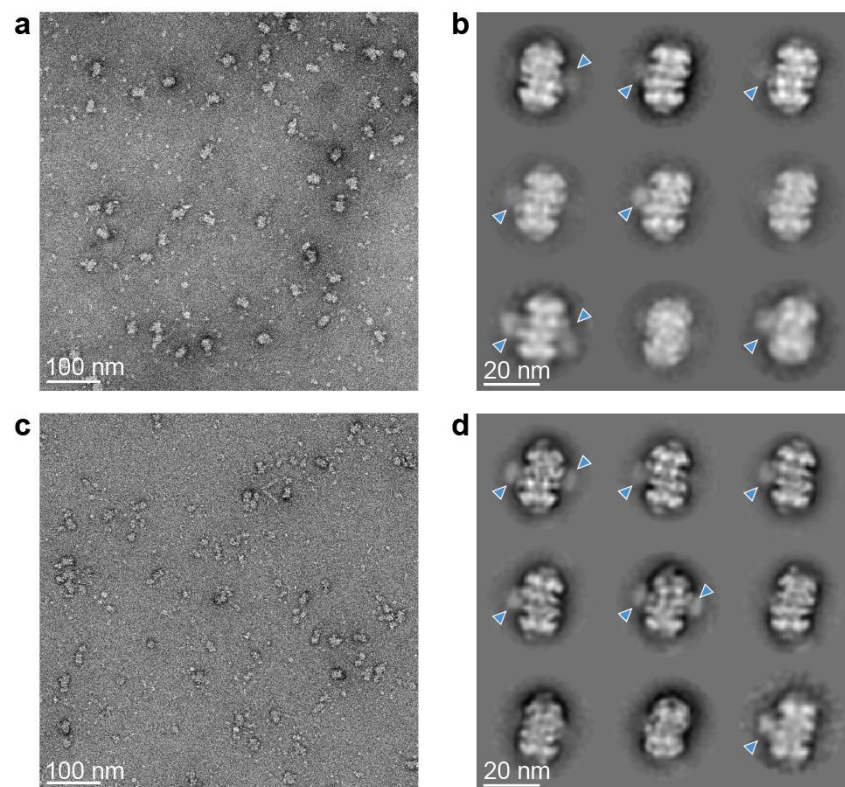


Figure 4.4: Imaging the entire DH loading and phosphorylation reaction improved decoration with DDK.

a. Negative stain electron micrograph of DNA-affinity purified DH mixed with DDK. **b.** 2D classes of DH show weak density for DDK at the interface of the two MCM hexamers. Blue arrowheads indicate DDK. **c.** Negative stain electron micrograph of the entire DH loading and phosphorylation reaction. **d.** 2D classes of DH decorated with DDK obtained from in the entire loading and phosphorylation reaction.

To gain structural insights into the mechanism of DH phosphorylation by DDK, I decided to image the entire DH loading and phosphorylation reaction by EM (Figure 4.4c). This approach resulted in a highly heterogeneous sample, as it not only contained the DH–DDK complex, but also isolated loading factors and reaction intermediates. However, we have previously shown that protein complexes that resist purification can be characterised structurally by imaging entire reactions as they occur in a test tube (Miller et al. 2019). I reasoned that a similar approach could lead to elucidate DDK docking onto and phosphorylation of DHs captured as it occurs. Reference-free 2D classification produced DH averages with DDK density on either one or two sides, which was more pronounced compared to the previous experiments (Figure 4.4d).

These data show that DNA-loaded DH can be purified to homogeneity and used as a substrate for subsequent origin firing. While it was not possible to characterise DH phosphorylation in this way, my purification strategies built the basis for studying the molecular basis of origin DNA melting upon CMG formation (Lewis et al. 2022). Furthermore, I have shown that DDK can be captured in the act of phosphorylating the DNA-loaded DH and that it is not strictly necessary to use non-hydrolysable nucleotide analogues or protein crosslinking to visualise DDK docking onto its phosphorylation substrate.

4.3.2 *In silico* purification of DH–DDK

For structural characterisation of the DH–DDK complex, the DH loading and phosphorylation reaction was applied to lacey carbon grids with a thin layer of carbon and plunge frozen in liquid ethane (in collaboration with Dr Julia Locke). Data was collected on a Titan Krios electron microscope equipped with a K2 Summit direct electron detector with a GIF Quantum energy filter. After motion correction and CTF estimation, DH particles were separated from loading intermediates in three rounds of reference-free 2D classification in RELION. DH particles were then 3D refined, subjected to Bayesian polishing and CTF refinement, as described in chapter 3.2.1. 2D classes of particles contributing to the 3.0 Å DH structure showed DDK bound on either one or two sides of the DH

(Figure 4.5a). DDK density was also observed at lower contour level (0.0017) in the 3D map (Figure 4.5b). However, the low local resolution of DDK compared to the DH core indicated partial kinase occupancy. Furthermore, particle alignment was driven by the two-fold symmetric DH, making it less likely that singly-decorated DHs would feature the DDK kinase on the same side. To overcome this issue, symmetry expansion followed by focused classification was used. The principle of symmetry expansion utilises the underlying symmetry of a structure to identify asymmetric features (Scheres 2016). Here, the set of DH particles was duplicated and rotated by 180° around the C2 symmetry axis so that DDK molecules bound to opposing sides became aligned (Figure 4.5c). 149,876 DH particles with DDK on the same side were identified using 3D classification without alignment, focussing on the DDK density. The signal of the symmetry-related DDK molecule was then subtracted from particle images to only analyse the asymmetric DH–DDK complex. The structure was refined to an average resolution of 3.3 Å, with a local resolution of approximately 5 Å for DDK (Figure 4.6).

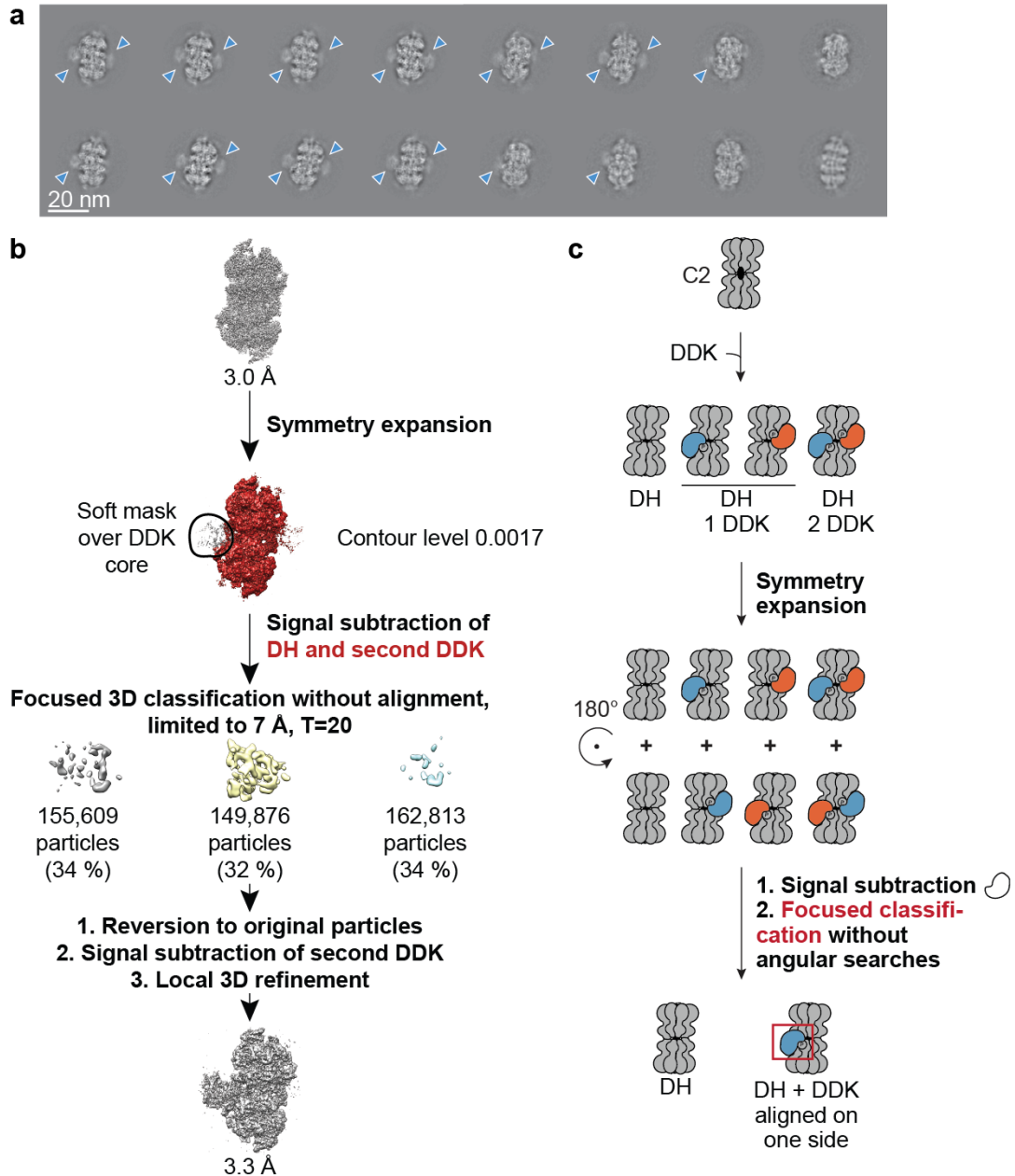


Figure 4.5: Image processing for the DH–DDK complex.

a. 2D classes of the 3.0 Å DH show DDK density at the DH dimerization interface. DDK density is indicated by blue arrowheads. **b.** Overview of the image processing procedure for the DH–DDK complex. **c.** Illustration of the symmetry expansion protocol. To align DDK molecules bound to symmetry-related sides of the DH, the dataset was duplicated and simultaneously rotated by 180° around the symmetry axis. The signal of one DDK molecule was then subtracted. Focused classification without image alignment was used to separate DDK-bound DH.

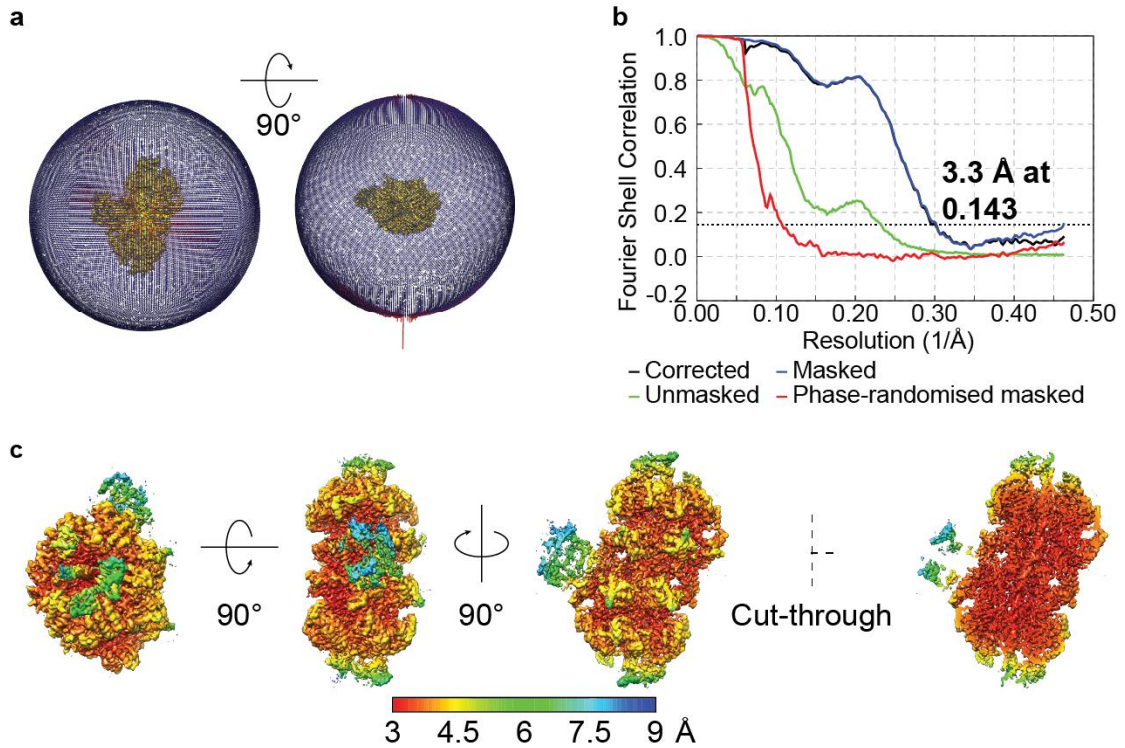


Figure 4.6: Structure quality of the DH-DDK complex.

a. Angular distribution of the DH-DDK complex. **b.** Resolution of the DH-DDK complex estimated using gold-standard Fourier Shell Correlation. **c.** Three rotated views and one cut-through view of the DH-DDK 3D structure, colour-coded according to the local resolution.

Multi-body refinement in RELION (Nakane et al. 2018) failed to improve the map further, but indicated that DDK was bound in different conformations relative to the DH (Figure 4.7). DDK exhibited a rotational movement around the perimeter of the DH and two rotations pivoting around the contact point between DDK and the DH. This structural flexibility might reflect phosphorylation of different target sites on the MCM.

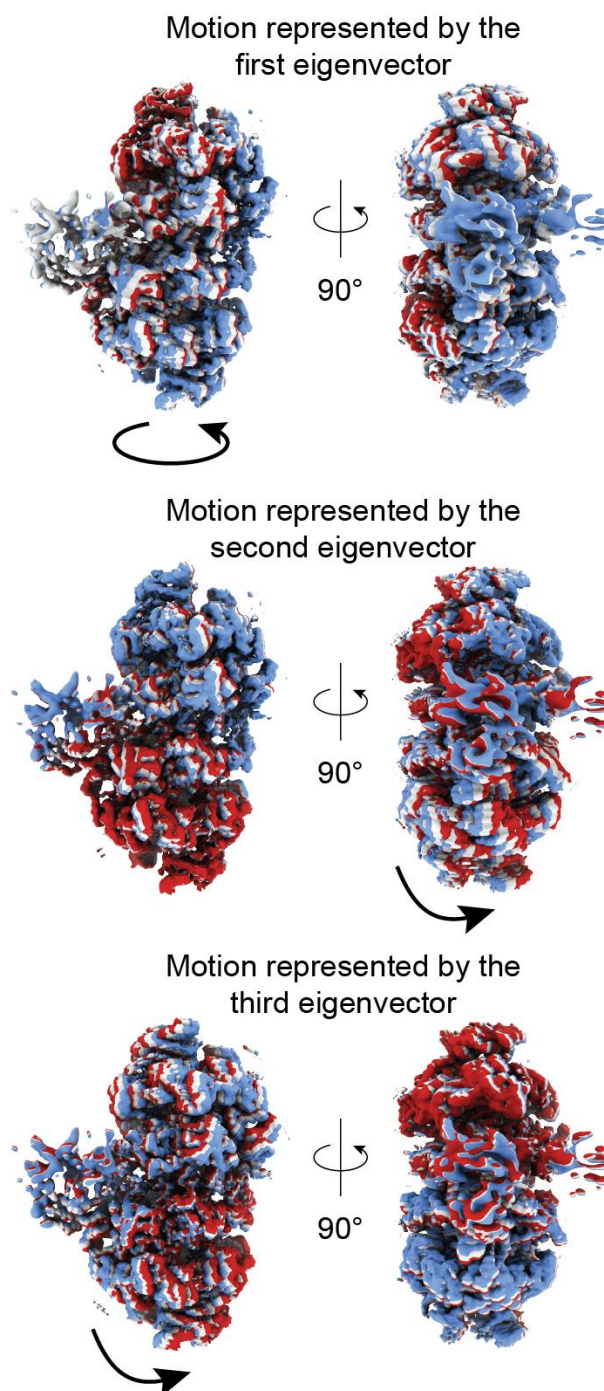


Figure 4.7: Multi-body analysis of the DH-DDK.

Illustration of the movement represented by the first three eigenvectors. The first frame is indicated in blue, the last one in red and intermediate frames in grey.

4.3.3 The catalytic core of DDK binds to the Mcm4 subunit of the MCM DH

The atomic model of the catalytic core of *H. sapiens* DDK (PDB entry 6YA7) (Dick et al. 2020) was docked into the cryo-EM map as a rigid body and served as the basis for the *S. cerevisiae* DDK homology model, which was generated using HHPRED (Söding et al. 2005) and I-Tasser (Yang and Zhang 2015). Together with Dr Fabrizio Martino, I combined the homology model with the coordinates of the DH, which were then subjected to molecular dynamics flexible fitting with Namdinator (Kidmose et al. 2019), manual adjustment in COOT (Casañal et al. 2020) and real-space refinement in Phenix (DiMaio et al. 2013) ([Appendix Table 1](#)). Assignment of the different subunits based on the DH–DDK model showed that the catalytic core of DDK interacts with Mcm4 in one of the two MCM hexamers ([Figure 4.8a](#)). It is located next to the N-terminal MCM tier, close to the dimerization interface of the DH. To analyse the interaction with Mcm4, I segmented the cryo-EM map in UCSF Chimera. This isolated view illustrated how DDK docks onto the A domain of Mcm4 ([Figure 4.8b](#)): Firstly, the Dbf4 Zn finger establishes a contact with an N-terminal α -helix in Mcm4 (residues 192–200), which agrees with the finding that residues 175–333 in Mcm4 are important for efficient phosphorylation by DDK (Sheu and Stillman 2006). Secondly, the Dbf4 M domain and the KI-2 of Cdc7 contact the Mcm4 B domain.

Comparison with previously published structures of the DH (Li et al. 2015, Abid Ali et al. 2017, Noguchi et al. 2017) indicates that the N-terminus of Mcm4 is partially stabilised upon DDK binding. We therefore extended the atomic model of Mcm4 from residue 177 to residue 155 in the DH–DDK structure. The remaining N-terminal tail, which contains several DDK phosphorylation sites, remained unresolved and no cryo-EM density corresponding to the Mcm4 substrate was observed in the Cdc7 active site. This could be the result of averaging complexes captured at different stages of the phosphorylation process, including phosphorylation of different sites as well as a post-catalytic state. To understand whether the Mcm4 N-terminal tail would be correctly positioned to enter the active site, we combined the models of human DDK and yeast DH ([Figure 4.8b](#)). I found that the N-terminal end of the Mcm4 model and the C-terminus of the co-crystallised MCM peptide in the human DDK structure are only 9 Å apart. This

agrees with the notion that the Mcm4 N-terminal tail is suitably positioned for phosphorylation.

To confirm phosphorylation of target sites in the extended Mcm4 N-terminus in our experiments, I prepared phosphorylated DH for analysis by mass spectrometry, which was subsequently carried out by Dr Steven Howell from the Proteomics Science Technology Platform at the Francis Crick Institute. Multiple phosphorylation sites, which overlapped with previously identified sites, were detected ([Figure 4.8c](#), [Table 4.1](#), [Appendix Table 2–5](#)).

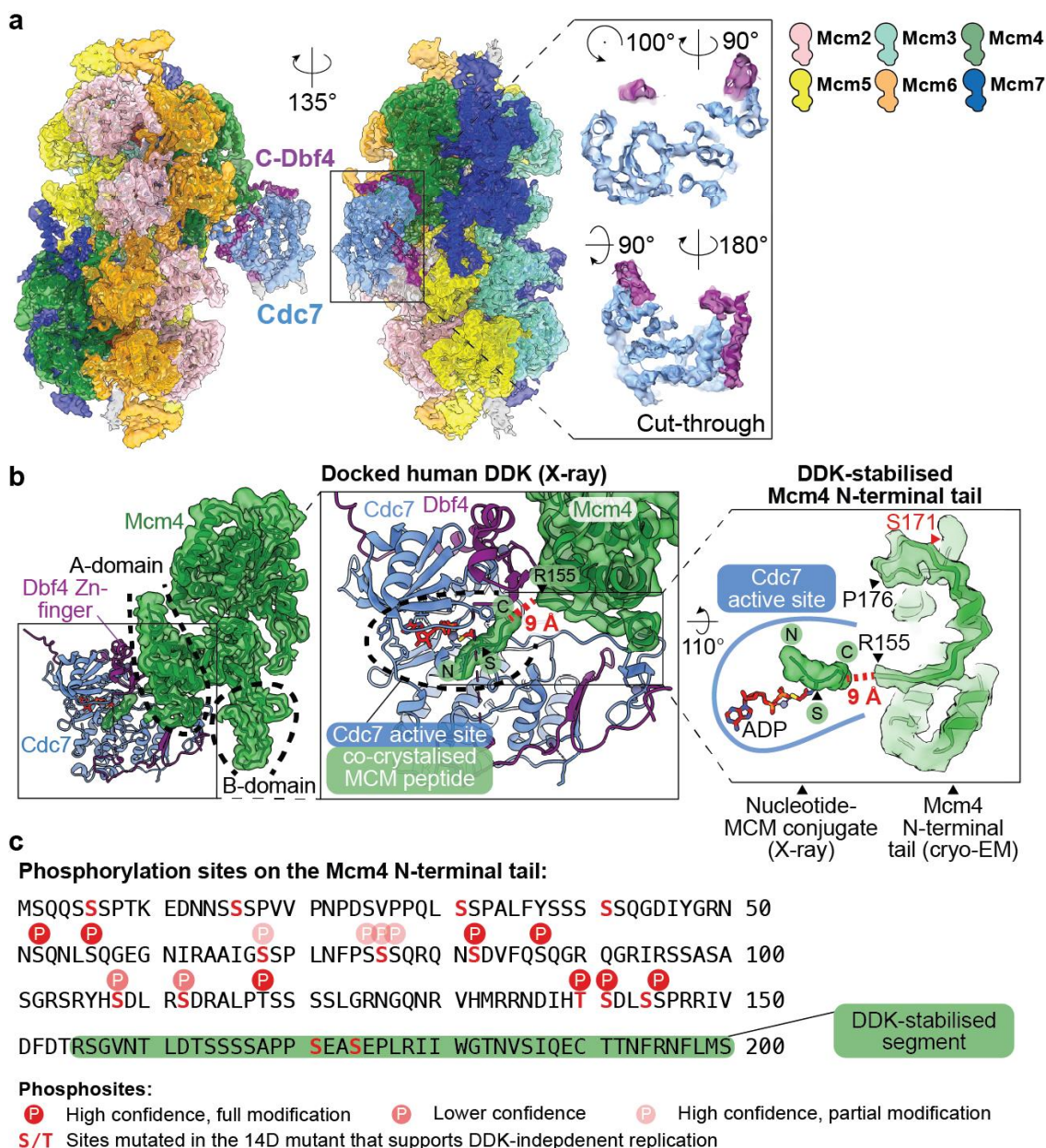


Figure 4.8: Binding of the DDK catalytic core to the DH.

a. Structure of the DH-DDK complex, showing the catalytic core of DDK (Cdc7 bound to C and M domains of Dbf4), engaged to the Mcm4 subunit of one MCM ring in the DH. Two cut-through views of the kinase core highlight how the atomic model matches cryo-EM density. **b.** DDK docks onto the Mcm4 A domain via the Dbf4 Zn finger in the C domain, and onto the Mcm4 B domain via the Dbf4 M domain. Middle: Illustration of the human DDK active site, which was co-crystallised with an MCM substrate peptide. The C-terminal end of the MCM peptide aligns with first resolved N-terminal residue (R155) of Mcm4 in the cryo-EM map. The Mcm4 N-terminal tail is therefore suitably poised for phosphorylation. Right: The N-terminal Mcm4 segment between residues P155 to R176 becomes stabilised upon DDK binding. **c.** Mass spectrometry identifies phosphorylation of S171, which maps within the DDK-stabilised N-terminal segment visible in the DH-DDK structure. Further phosphorylation sites are found upstream of the modelled N-terminal region of Mcm4. Sites labelled with “high confidence, full modification” have posterior error probability (PEP) < 0.01 and were only detected in

their phosphorylated form. Sites with PEP > 0.01 are indicated with “lower confidence”. The label “high confidence, partial modification” indicates sites, which had PEP < 0.01 and for which multiple phosphorylation states were detected. Sites shown to support DDK-independent recruitment of the firing factor Sld3, when mutated to phospho-mimicking amino acids, are highlighted (Deegan et al. 2016).

The conformation of the DH-DDK complex, with DDK bound to Mcm4, suggests a preference of DDK for phosphorylation of Mcm4 over Mcm6. To test this hypothesis, I titrated DDK against the DH and analysed the phosphorylation-dependent shifts of the MCM subunits by SDS-PAGE (Figure 4.9a). Mcm4 was mostly phosphorylated after treatment with 4 nM DDK, while complete Mcm6 phosphorylation required DDK concentrations above 16 nM. Similarly, a time course assay in the presence of 10 nM DDK resulted in complete Mcm4 phosphorylation after 10 minutes, whereas Mcm6 was only partially phosphorylated even after 30 minutes (Figure 4.9b).

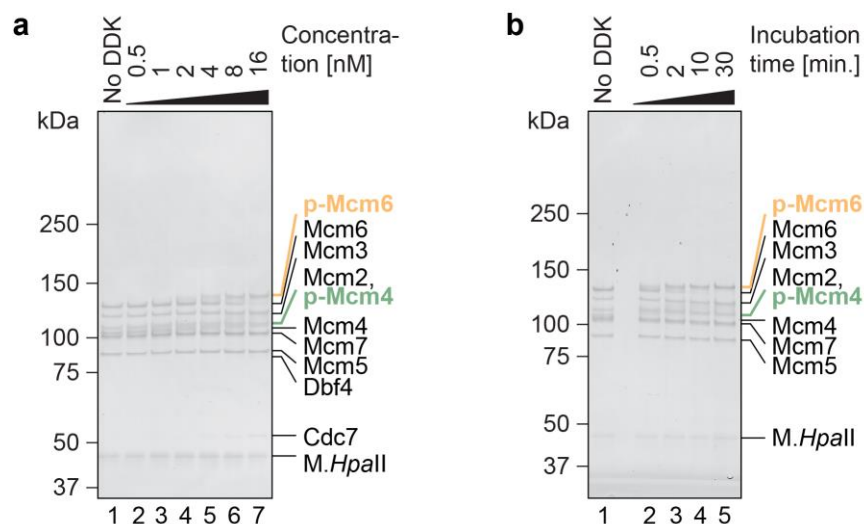


Figure 4.9: Mcm4 is preferably phosphorylated.

a. Silver-stained SDS-PAGE gel of DNA-loaded DH treated with increasing amounts of DDK. The phosphorylation-dependent shift of Mcm4 is observed prior to the shift of Mcm6. **b.** Phosphorylation of DNA-loaded DH at different timepoints after DDK addition, analysed by SDS-PAGE and silver staining. Near complete Mcm4 phosphorylation is observed at 30 seconds after DDK addition, while Mcm6 phosphorylation is completed between 10 and 30 minutes after DDK addition.

In summary, the docking of DDK onto the A and B domain of Mcm4 ensures that the substrate is suitably positioned to reach the active site. This arrangement allows for more efficient phosphorylation of Mcm4 than phosphorylation of Mcm6.

4.3.4 The Dbf4 BRCT domain binds to Mcm2

Apart from the catalytic core of DDK that binds to the Mcm4 subunit, additional, lower quality density next to the Mcm2 A domain was observed. To better resolve this density, the signal of the two C-terminal ATPase tiers of the DH was subtracted from particle images, as those domains are more flexible compared to the N-terminal domains and could distract from the Mcm2-interacting density. This was followed by focused 3D classification without image alignment using a mask around the DDK core as well as the Mcm2-interacting domain (Figure 4.10a). A class with discrete structural elements was selected and the 128,200 particles were 3D refined to 3.4 Å. The map was then denoised using LAFTER filtering (Ramlaul et al. 2019), which allowed to confidently fit the crystal structure of the yeast Dbf4 BRCT domain (PDB entry 3QBZ)(Matthews et al. 2012) into the density next to the

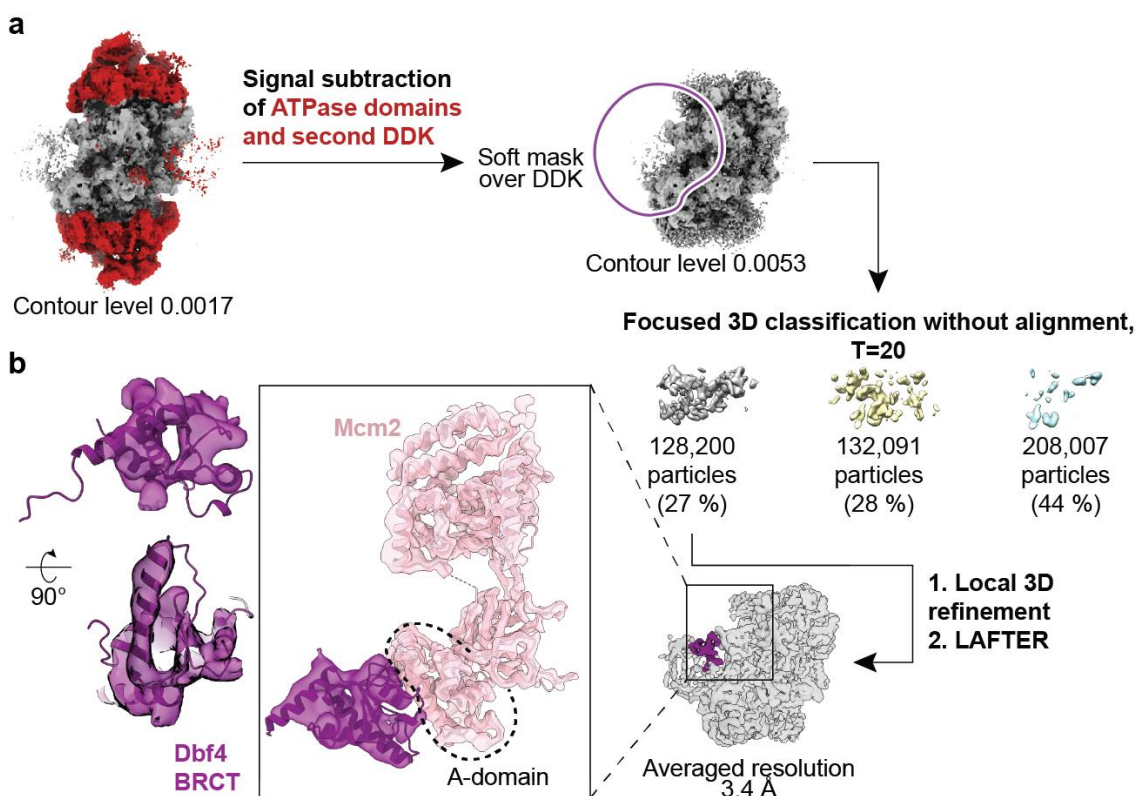


Figure 4.10: Resolving the Dbf4 BRCT domain bound to Mcm2.

a. Overview of the image processing procedure. **b.** Enlarged view of the Mcm2–Dbf4 BRCT interaction with the atomic models docked.

Mcm2 A domain (Figure 4.10b). This agrees with the biochemical evidence of Dbf4 interacting with the Mcm2 subunit (Francis et al. 2009, Ramer et al. 2013, Abd Wahab and Remus 2020).

4.3.5 DDK binds to both MCM rings in the DH to phosphorylate the MCM *in trans*

To illustrate how the full-length DDK interacts with the DH, I added the Dbf4 BRCT domain of the LAFTER-filtered map to the map of the DDK core bound to the DH. The resulting composite map and interactome map (resulting from manually splitting the two MCM rings, DDK core and Dbf4 BRCT domain) highlight that interaction with Mcm2 occurs in one MCM hexamer (referred to as *in cis*), while Mcm4 of the opposing MCM ring is targeted for phosphorylation (*in trans*) (Figure 4.11).

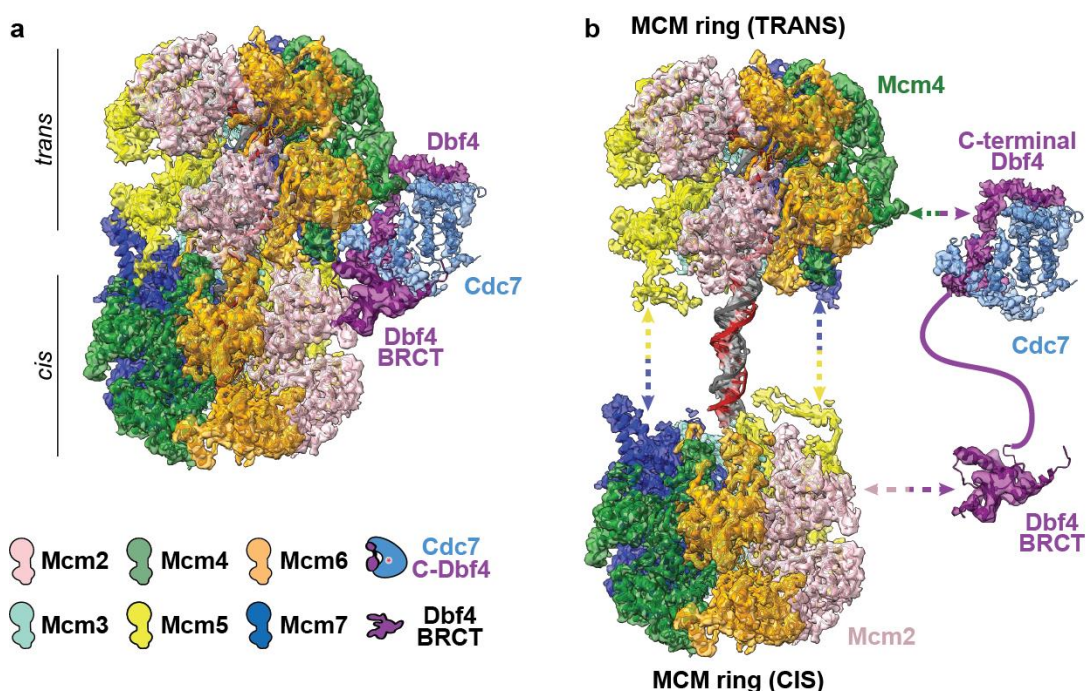


Figure 4.11: Composite map of DH-DDK highlighting the bipartite binding of DDK.

a. Composite map of the DH in complex with the DDK core and the LAFTER filtered segment of the Dbf4 BRCT domain. **b.** Manually separated components, visualising the different interactors.

The 40-amino-acid flexible linker enables DDK to establish the extended configuration, stretching from one MCM hexamer to the second one. This suggests a mechanism whereby DDK recognises the three-dimensional shape of the DH to allow selective phosphorylation. To test whether the bipartite interaction with the DH is essential for DDK activity, I generated a variant lacking residues 119–219 of Dbf4, which corresponds to the BRCT domain (Δ BRCT variant hereafter). As the BRCT domain is flexibly tethered to the Cdc7 catalytic core, an effect on the catalytic activity *per se* was not expected. To confirm this prediction, I first assessed whether Δ BRCT is competent for autophosphorylation. Both wild type DDK and Δ BRCT were incubated with ATP for up to 15 minutes and phosphorylation was visualised by SDS-PAGE and subsequent silver staining (Figure 4.12a). Samples incubated without ATP served as control. The mutant showed efficient autophosphorylation, which was comparable to wild type DDK. In both cases, phosphorylation could efficiently be reverted upon λ phosphatase treatment. To further assess kinase activity, I analysed phosphorylation of a well-characterised substrate for DDK (Montagnoli et al. 2006, Hughes et al. 2012, Dick et al. 2020). The peptide represented residues 35–47 of human MCM2, engineered with a C-terminal biotin tag, and contained the DDK target site S40. In addition to the unmodified peptide, a peptide phosphorylated at S41 (P+1) was used, which has been shown to prime the substrate for DDK phosphorylation (Montagnoli et al. 2006). To detect phosphorylation, DDK and the corresponding peptide were incubated with [γ - 32 P]ATP. Peptides were then immobilised on a SAM2 biotin-capture membrane and the radioactive signal of each sample was quantified. Both wild type and Δ BRCT showed similar phosphorylation efficiencies for the primed substrate, whereas only a weak signal was observed for the unmodified peptide in the presence of either kinase variant. These data confirmed that Δ BRCT retains kinase activity (Figure 4.12b).

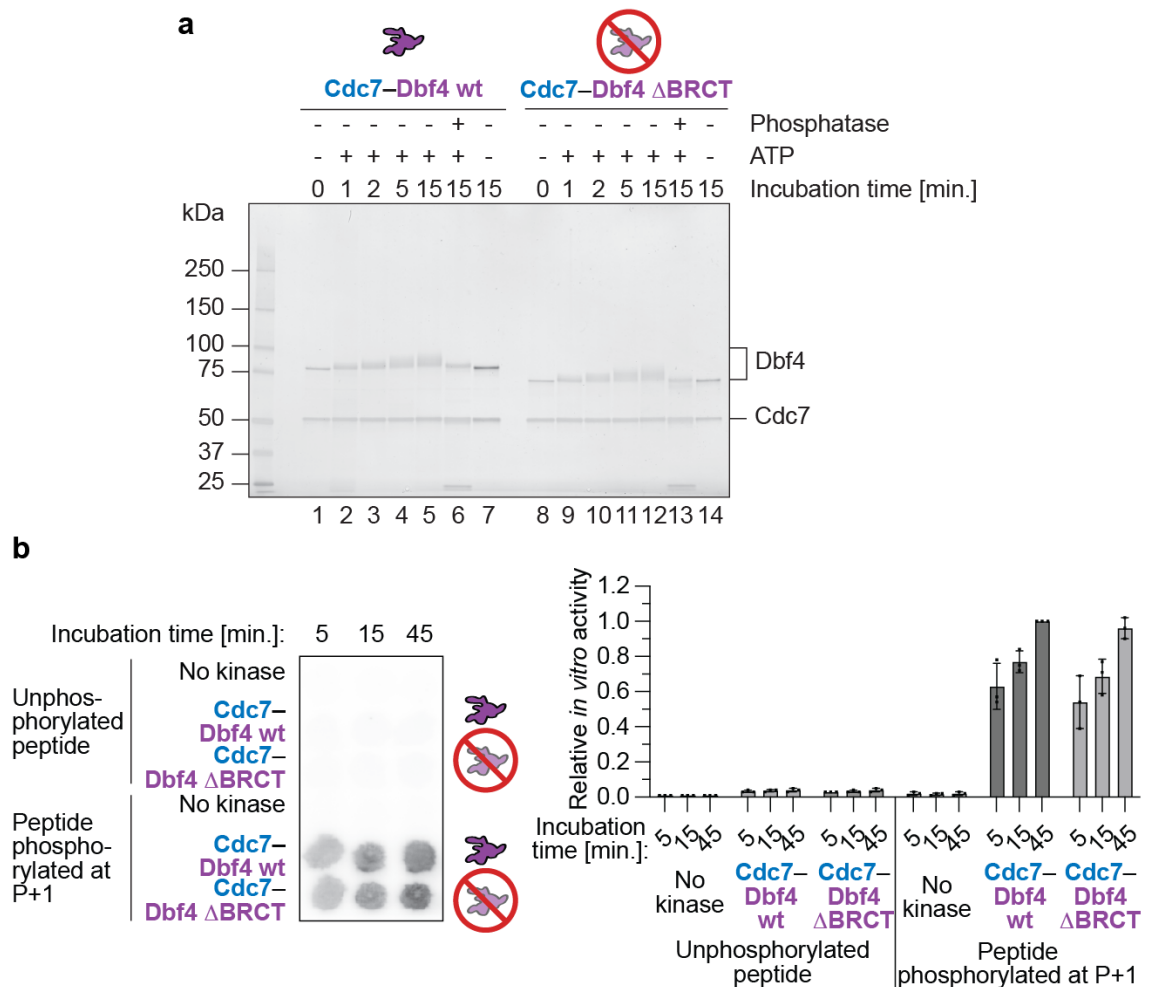


Figure 4.12: The variant lacking the Dbf4 BRCT domain shows efficient autophosphorylation and MCM peptide phosphorylation.

a. Autophosphorylation is not affected by the deletion of the Dbf4 BRCT domain. **b.** Peptide phosphorylation assay and quantification shows comparable kinase activity between wild type and Δ BRCT variant. The mean of three biological replicas is plotted and error bars show standard deviations. Reads were normalised to the 45-min time point of wild type DDK.

I then analysed how the BRCT truncation affected DH phosphorylation. For this purpose, DH were loaded onto the origin DNA template containing a desthiobiotin-TEG and immobilised on streptavidin paramagnetic beads. After treating the DH with the two DDK variants in a concentration range from 2.5–20 nM, phosphorylation of the DH was assessed by SDS-PAGE and silver staining (Figure 4.13a). Δ BRCT did not produce any phosphorylation-dependent shifts of Mcm4 or Mcm6, while the wild type kinase efficiently phosphorylated both MCM subunits. Having established that Δ BRCT was active, we asked whether the inability to phosphorylate the DH could be explained by an impaired DH-DDK interaction. To

address this question, the reaction was visualised by negative stain EM with help from Dr Thomas Miller. The analysis revealed that the Δ BRCT variant lost the ability to bind the DH (Figure 4.13b). In conclusion, the Dbf4 BRCT domain is essential for recruitment of DDK to the DH as it mediates the interaction with Mcm2 *in cis* to stimulate Mcm4 phosphorylation *in trans* (Figure 4.13c).

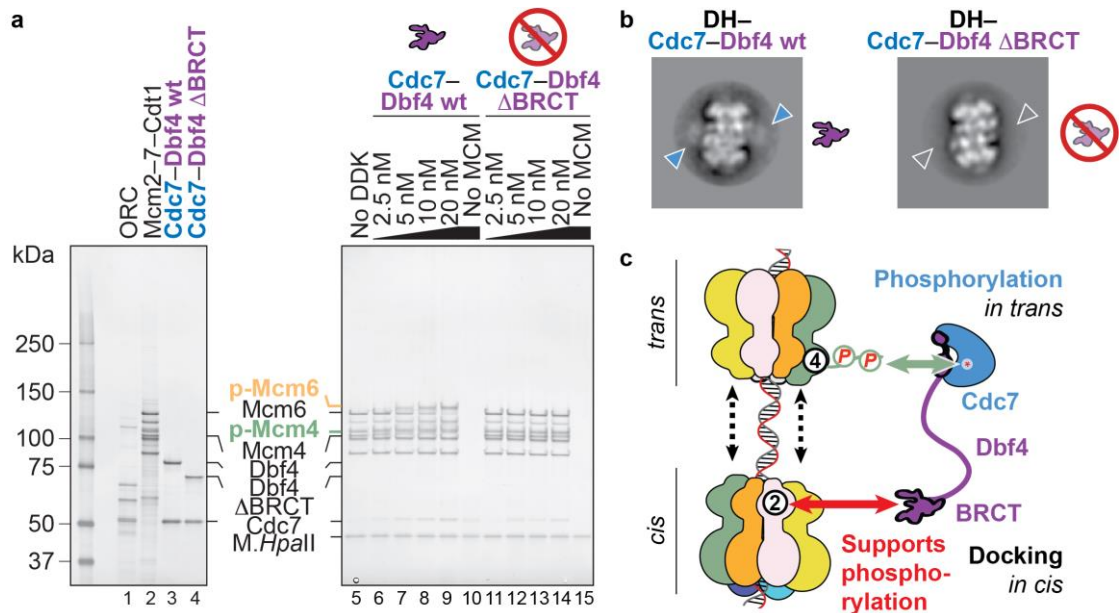


Figure 4.13: DDK docking onto the MCM DH *in cis* is required for MCM phosphorylation *in trans*.

a. DH phosphorylation was impaired for the Δ BRCT variant, when tested in the same concentration range. **b.** 2D class averages of DHs incubated with wild type DDK and Δ BRCT. **c.** Model illustrating the importance of the binding *in cis* for phosphorylation *in trans*.

4.3.6 Phosphorylation of MCM–Cdt1 by DDK

Having established that DDK needs to interact with both MCM rings in the DH to phosphorylate Mcm4 and Mcm6, we asked whether this three-dimensional arrangement is required to phosphorylate critical residues in Mcm4. For this reason, conditions that allow for phosphorylation of the loading-competent Mcm2–7–Cdt1 were established by incubation with DDK. Reactions were subsequently analysed by SDS-PAGE and silver staining (Figure 4.14). To achieve detectable phosphorylation levels of this single MCM ring substrate, DDK was used in 20-fold molar excess compared to Mcm2–7–Cdt1. This is a sevenfold higher kinase to target ratio than the ratio that was used for DH phosphorylation. The

phosphorylated Mcm2–7–Cdt1 complex was then analysed by mass spectrometry. Importantly, sites critical for origin activation were not phosphorylated under these conditions, while they are modified in the DH (Table 4.1, Appendix Table 2–7). These data show that the bipartite interaction of DDK with the DH is essential to orient the kinase correctly and to reach the target sites important for helicase activation.

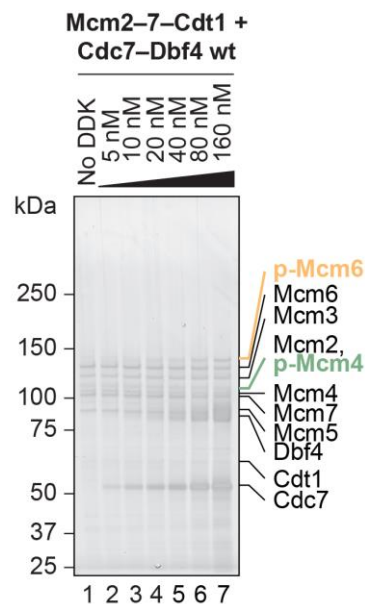


Figure 4.14: Phosphorylation of MCM–Cdt1 can be achieved by increasing the DDK concentration.

Mcm2–7–Cdt1 was co-incubated with increasing amounts of wild type DDK. Samples were separated by SDS-PAGE and visualised by silver staining.

Table 4.1: DDK-dependent phosphorylation of Mcm4 detected by mass spectrometry.

Site	Phospho- rylated in Mcm2–7– Cdt1	Localisation probability	PEP	Phospho- rylated in MCM DH	Localisation probability	PEP
S52	Yes	1.00	0.00036	Yes	1.00	5.87E-29
S56	Yes	1.00	0.00036	Yes	1.00	8.31E-20
S68/S69	No	-	2.55E-11	Majority unmodified, but some phospho- rylation detectable	0.88	1.07E-03
S75	No	-	2.55E-11	Majority unmodified, but some phospho- rylation detectable	0.99	1.07E-03
S76	No	-	2.55E-11	Majority unmodified, but some phospho- rylation detectable	0.80	1.07E-03
S77	No	-	2.55E-11	Majority unmodified, but some phospho- rylation detectable	0.81	1.38E-03
S82	No	-	4.57E-06	Yes	1.00	2.84E-05
S87	No	-	4.57E-06	Yes	1.00	3.17E-04
S108/ S112	Not detected			Only phospho- rylated peptide detected		0.044
S118	No	-	1.57E-04	Yes	1.00	4.50E-03
T140	No	-	1.99E-20	Yes	1.00	2.62E-08
S141	No	-	1.99E-20	Yes	1.00	2.62E-08
S144	No	-	1.99E-20	Yes	1.00	2.62E-08
S171	Only detected in sample without DDK			Not detected		

4.3.7 DDK autophosphorylation increases substrate selectivity

Kinases are known to regulate their activity in multiple ways. Autophosphorylation of DDK has been implicated in downregulating kinase activity (Dick et al. 2020). It was however unclear how autophosphorylation affects DH phosphorylation and substrate selectivity. For this purpose, I incubated DDK with ATP to induce autophosphorylation and analysed the effect on three different substrates: a MCM peptide, Mcm2–7–Cdt1 and DNA-loaded DH. To assess kinase activity by peptide phosphorylation, the biotinylated MCM peptide containing the priming phosphorylation at S41 was mixed with the unmodified and autophosphorylated DDK, respectively, as well as [γ - 32 P]ATP. The reaction was terminated at different timepoints, and phosphorylation of membrane-immobilised peptides was quantified (Figure 4.15a). Autophosphorylation resulted in a drop in peptide phosphorylation efficiency of 92%.

To corroborate the observation of a reduced kinase activity, Mcm2–7–Cdt1 was treated with pre-autophosphorylated and unmodified DDK, using a 1:20 substrate to kinase molar ratio. The product of each reaction was then separated by SDS-PAGE and visualised by silver staining (Figure 4.15b). I found that Mcm2–7–Cdt1 phosphorylation was completely abrogated upon DDK autophosphorylation.

The DNA-loaded DH is the third substrate that was tested. In this experiment, DH were loaded onto the short origin DNA template containing a desthiobiotin–TEG and treated with DDK. DNA-bound complexes were then isolated using streptavidin paramagnetic beads and washed with high-salt buffer. Elution was carried out by digesting DNA with micrococcal nuclease (Figure 4.15c). Surprisingly, DH phosphorylation was only marginally decreased when DDK was autophosphorylated at the beginning of the reaction. To explain the discrepancy between the effects on the three substrates, the DH phosphorylation reaction, using unmodified and autophosphorylated DDK, respectively, was imaged by negative stain EM. This experiment was carried out in collaboration with Dr Thomas Miller. 2D class averages of the DH showed decoration with DDK, regardless of the phosphorylation state of DDK (Figure 4.15d). Engagement with the DH is therefore not affected by autophosphorylation.

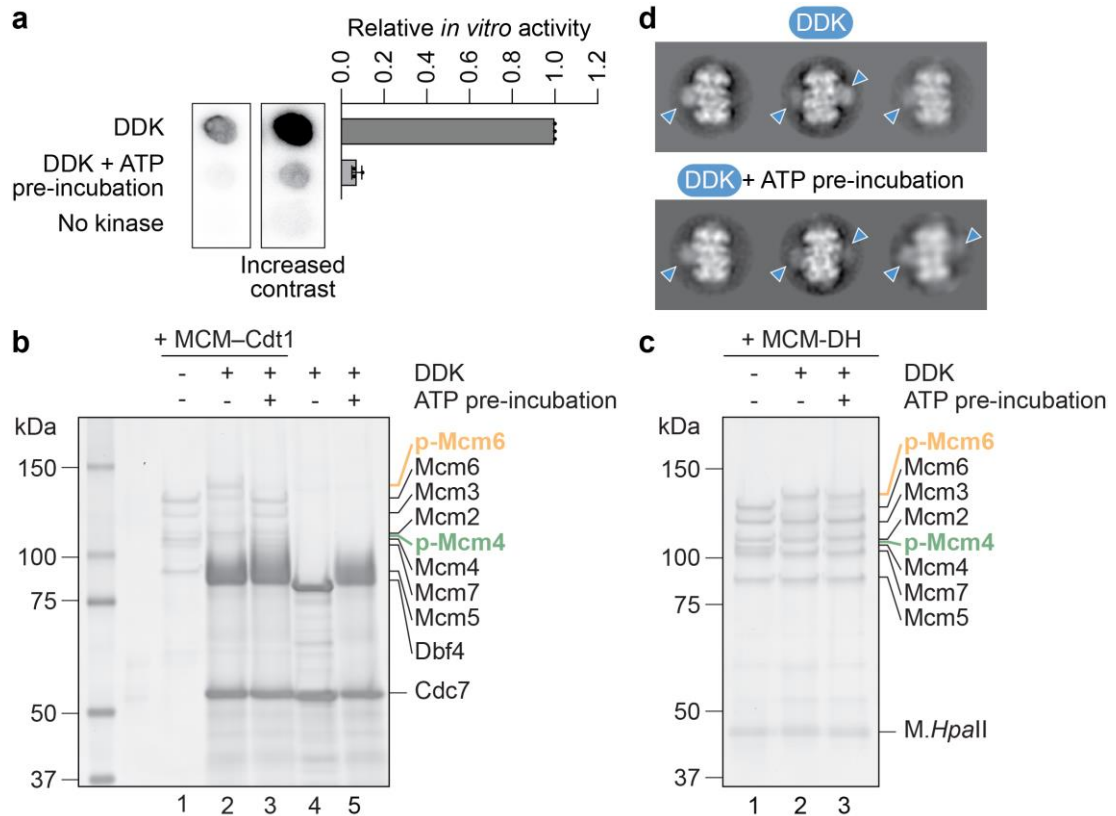


Figure 4.15: DDK pre-autophosphorylation prevents peptide and MCM-Cdt1 phosphorylation, but not DH phosphorylation and engagement.

a. MCM peptide phosphorylation drops by 92% following DDK pre-autophosphorylation. The mean of three biological replicas is plotted and error bars show standard deviations. Reads were normalised to the unphosphorylated DDK. **b.** MCM-Cdt1 phosphorylation is abrogated following DDK pre-autophosphorylation. **c.** DH phosphorylation is virtually unperturbed following DDK pre-autophosphorylation. **d.** 2D classes derived from negative stained particles indicate that DDK binding to DHs is unperturbed following DDK pre-autophosphorylation.

In conclusion, while autophosphorylation significantly reduced catalytic activity, interaction with the DH was unperturbed. Kinase inhibition is thereby overridden by the specific interaction with the DH. These data suggest that autophosphorylation promotes recognition and selective phosphorylation of the DNA-loaded DH, which contains the symmetry to support bidirectional replication.

4.4 Regulation of DDK by Rad53

4.4.1 Rad53-dependent phosphorylation in context of the DH-DDK complex

To assess how DH phosphorylation and further origin firing are impaired through phosphorylation by the checkpoint kinase Rad53, upon detection of DNA damage, I mapped the known phosphorylation sites of Dbf4 (Zegerman and Diffley 2010) on the structure of the DH-DDK complex (Figure 4.16). Many Dbf4-phosphosites reside in flexible linkers and are predicted to face the MCM. Ten of the 19 sites are found in a linker bridging the active site of the kinase and include four sites that have been identified as the minimal targets to prevent origin firing.

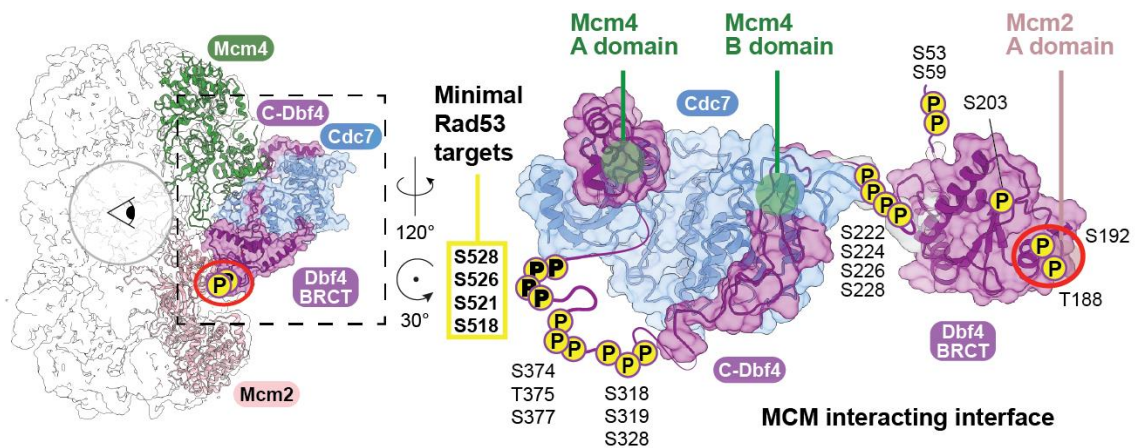


Figure 4.16: Rad53-dependent phosphorylation sites of Dbf4 in the context of the DH-DDK complex.

The 19 known phosphorylation sites in Dbf4 (Zegerman and Diffley 2010) are found in flexible linkers, predicted to face the MCM double hexamer. The sites whose mutation to alanine prevent inhibition of DDK by Rad53 (minimal Rad53 targets) reside in a linker bridging the Cdc7 active site.

Two different scenarios for how phosphorylation-dependent inhibition is achieved can be envisioned: Phosphorylation close to the Cdc7 active site could interfere with catalytic activity, either affecting phospho-transfer directly or preventing access of the substrate to the active site. Alternatively, binding of DDK to the DH could be abrogated as phosphorylation interferes with specific binding sites or electrostatic repulsion based on the large number of phosphorylation sites. The mechanism that applies remained to be uncovered.

4.4.2 Catalytically active Rad53 reduces peptide phosphorylation efficiency of DDK

To establish how Rad53 influences the catalytic activity of DDK, wild type *S. cerevisiae* Rad53 and a variant containing two point mutations (K227A and D339A), which inactivate the kinase, were recombinantly expressed in *E. coli* (with help from Dr Allison McClure and Dr Berta Canal). Both variants were purified as previously described (McClure and Diffley 2021). First, the kinase was pulled down utilising the C-terminal histidine tag. The protein was then loaded on a Superdex 200 10/300 increase gel filtration column. Peak fractions containing pure protein (as established by SDS-PAGE analysis) were pooled.

Wild type Rad53 and catalytically inactive Rad53, respectively, were then co-incubated with DDK and ATP to induce phosphorylation of Dbf4 by Rad53 (Figure 4.17a lanes 1–4). Under these conditions, DDK also showed autophosphorylation as discussed above. However, the phospho-dependent shift of Dbf4 became more pronounced in the presence of catalytically active Rad53, indicating further modification of this DDK subunit as detected by Zegerman and Diffley (2010) using mass spectrometry. Rad53 has been proposed to impair DH phosphorylation by DDK through a secondary mechanism, namely sequestration of DDK by Rad53 (Abd Wahab and Remus 2020). To investigate this alternative model and recapitulate the effect of Rad53 on DH phosphorylation by DDK that had been reported (Lopez-Mosqueda et al. 2010, Zegerman and Diffley 2010, Abd Wahab and Remus 2020), I added the DDK–Rad53 mixtures to DNA-loaded DH. Complexes bound to the DNA substrate, which contained a desthiobiotin–TEG, were isolated using streptavidin paramagnetic beads, eluted by the addition of micrococcal nuclease and analysed by SDS-PAGE (Figure 4.17a lanes 5–9). In accordance with published work (Lopez-Mosqueda et al. 2010, Zegerman and Diffley 2010, Abd Wahab and Remus 2020), no phosphorylation-dependent shift of Mcm4 or Mcm6 was observed in the presence of Rad53. I observed that this effect is strictly dependent on the catalytic activity of Rad53, given that the inactivated Rad53 variant did not show inhibition of DH phosphorylation. This contrasted with findings from Abd Wahab and Remus (2020), but corroborated results by McClure and Diffley (2021).

The DDK–Rad53 mixtures were also used as reagents in the MCM peptide phosphorylation assay. Here, only the peptide containing the priming phosphorylation at S41 (P+1) was used as the unmodified peptide exhibited low phosphorylation levels in previous experiments. Treatment of DDK with wild type Rad53 resulted in a 40% reduction in peptide phosphorylation efficiency (Figure 4.17b). Conversely, inactive Rad53 did not show this effect, indicating that phosphorylation of Dbf4 is required to inhibit DDK activity. The data also showed that the substrate peptide is specific for DDK as it is not modified by Rad53.

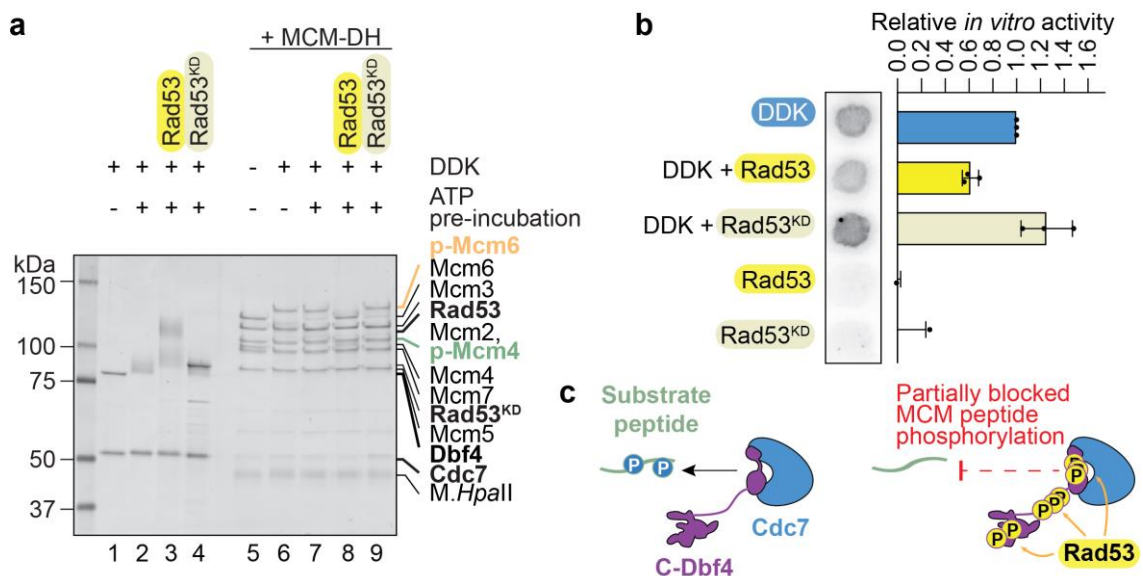


Figure 4.17: Phosphorylation of Dbf4 by Rad53 reduces peptide phosphorylation and abolishes DH phosphorylation.

a. DDK was co-incubated with ATP and Rad53 before addition to DNA-loaded DH. All samples were analysed by SDS-PAGE and silver-staining. The phosphorylation-dependent shift of Dbf4 became more pronounced in the presence of catalytically active Rad53, indicating further phosphorylation. DH phosphorylation was abolished upon phosphorylation of Dbf4 by Rad53. **b.** DDK kinase activity was probed by peptide phosphorylation. Peptide phosphorylation was reduced by 40% upon treatment of DDK with catalytically active Rad53. The mean of three biological replicas is plotted and error bars show standard deviations. Reads were normalised to the unphosphorylated DDK. **c.** Model of how phosphorylation of Dbf4 by Rad53 reduces peptide phosphorylation.

In summary, DDK kinase activity is significantly reduced upon phosphorylation of Dbf4 by Rad53. The effect could be caused either by affecting the Cdc7 active site directly or by hampering access of the substrate to the catalytic centre. The reduction in MCM-peptide phosphorylation by Rad53-phosphorylated DDK does

not account for the complete abrogation of DH phosphorylation, suggesting that other Rad53 mechanisms might exist, which interfere with MCM activation.

4.4.3 Phosphorylation of residues in the Dbf4 BRCT domain alone does not prevent DH phosphorylation

In addition to affecting the catalytic activity of DDK, phosphorylation of Dbf4 could interfere with the DH–DDK interaction. Two of the phosphorylation-sites, namely T188 and S192, are found in the Dbf4 BRCT domain, close to the interaction surface with Mcm2. As I have shown that the BRCT domain plays an essential role in the recruitment of DDK to the DH, I tested whether phosphorylation of T188 and S192 prevents the docking. For this purpose, I generated *S. cerevisiae* strains overexpressing variants of DDK that include mutations of residues T188 and S192 to aspartate, either in isolation or in combination. I used side-directed mutagenesis to change the corresponding codons in the pRS304 expression vector containing wild type Cdc7 and Dbf4 under the bidirectional inducible GAL1-10 promoter (On et al. 2014). Vectors were then linearised by restriction digest at a site within the auxotrophic TRP1 marker and transformed into the *S. cerevisiae* strain yJF1 (Frigola et al. 2013). All variants were expressed and purified in the same manner as wild type DDK.

First, catalytic activity of the double mutant (Cdc7–Dbf4^{T188D,S192D}) was assessed by autophosphorylation. DDK was incubated with ATP for up to 15 minutes and subsequently analysed by SDS-PAGE (Figure 4.18a). The variant showed the characteristic shift of Dbf4 to slower migrating species with an efficiency equivalent to the wild type protein. Thus, autophosphorylation was not influenced by the mutations in the Dbf4 BRCT domain. Additionally, kinase activity was analysed by peptide phosphorylation (Figure 4.18b). Again, the variant efficiently phosphorylated the given substrate.

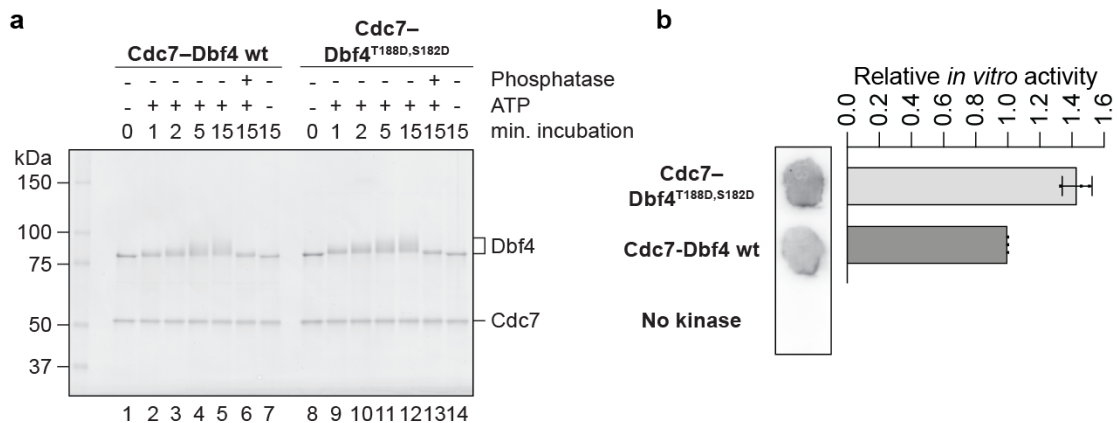


Figure 4.18: Catalytic activity of the phospho-mimicking variant Cdc7-Dbf4^{T188D,S192D}.

a. Silver-stained SDS-PAGE gel of autophosphorylated wild type DDK and Cdc7-Dbf4^{T188D,S192D}. **b.** Peptide phosphorylation by Cdc7-Dbf4^{T188D,S192D} and wild type DDK, respectively, was quantified by autoradiography. The mean of three biological replicas is plotted and error bars show standard deviations. Reads were normalised to the wild-type DDK.

All three variants were then titrated against DNA-loaded DH, immobilised on streptavidin paramagnetic beads via the desthiobiotin-TEG on the DNA, in the presence of ATP to induce DH phosphorylation. Beads were washed with a low-salt buffer (containing 300 mM sodium acetate) and eluted by DNA digestion with micrococcal nuclease. Phosphorylation efficiency was assessed by SDS-PAGE and silver staining (Figure 4.19a). All three variants induced MCM phosphorylation at a similar level when compared to wild type DDK. To test whether phosphorylation kinetics was influenced by the mutations, a time course experiment was carried out with the double mutant. Here, DH phosphorylation was terminated by the addition of 80 mM EDTA (final concentration) at timepoints between 0.5 and 30 minutes after DDK-incubation. Beads were then washed with low-salt buffer and eluted fractions were analysed by SDS-PAGE (Figure 4.19b). In this timeframe, phosphorylation of DH was observed with a comparable efficiency in the presence of either wild type or the double phospho-mimicking mutant DDK.

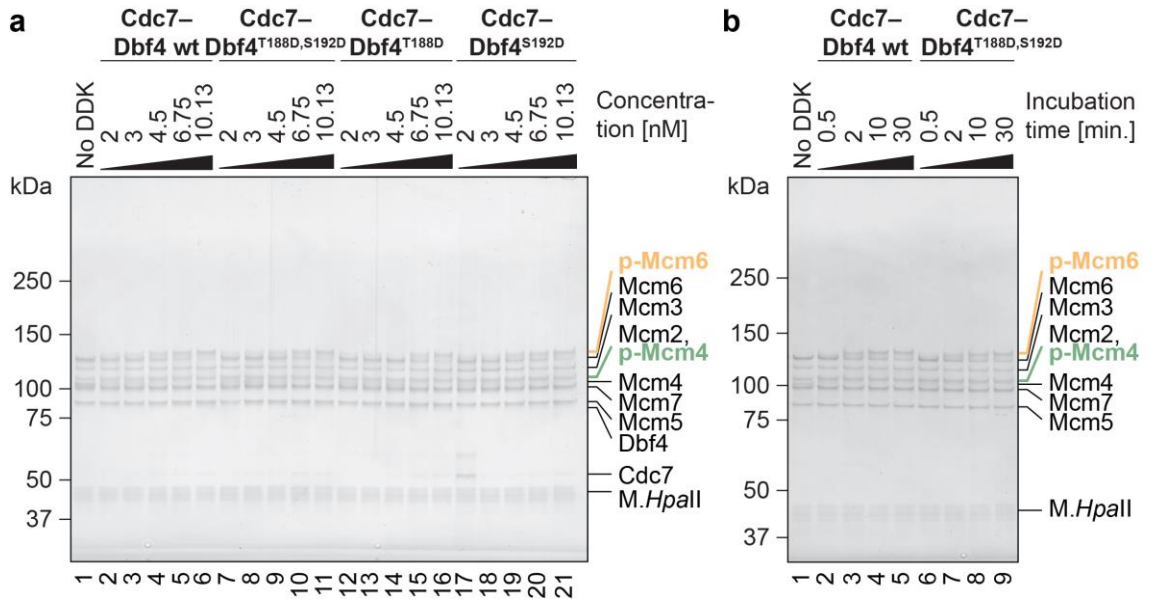


Figure 4.19: Phospho-mimicking mutations in the Dbf4 BRCT domain do not affect DH phosphorylation.

a. Silver-stained SDS-PAGE gel of DNA-loaded DH treated with increasing amounts of the different phospho-mimicking mutants showed similar phosphorylation efficiency as for the wild type DDK. **b.** A time course was carried out with the double mutant and analysed on a silver-stained SDS-PAGE gel. Phosphorylation efficiency was comparable to wild type DDK.

Therefore, while the Dbf4 BRCT domain is crucial for docking onto the Mcm2 subunit in the DH, phosphorylation of the two residues, T188 and S192, in this domain alone is insufficient to block phosphorylation of the DH. This is in line with data establishing that a different set of four residues exists, representing a minimal set of Rad53 targets, whose mutation to alanine bypass the inhibition by Rad53 (Zegerman and Diffley 2010).

4.4.4 Rad53-dependent phosphorylation of DDK prevents DH-DDK interaction

Next, we sought to investigate whether global phosphorylation of Dbf4 by Rad53 results in a loss of the DH-DDK interaction. In collaboration with Dr Thomas Miller, I co-incubated DDK with wild type and catalytically inactive Rad53, respectively, prior to addition to the MCM loading reaction. The reactions were then imaged by negative stain EM (Figure 4.20). 2D class averages of the DH showed prominent decoration with untreated DDK on either one or both sides. In the presence of

catalytically active Rad53, no DDK-engagement was observed, whereas the catalytically inactive Rad53 displayed wild type levels of DH decoration with DDK. Therefore, phosphorylation of multiple sites across Dbf4 by Rad53 as described previously (Zegerman and Diffley 2010), not only in the BRCT domain, prevents DH engagement and phosphorylation.

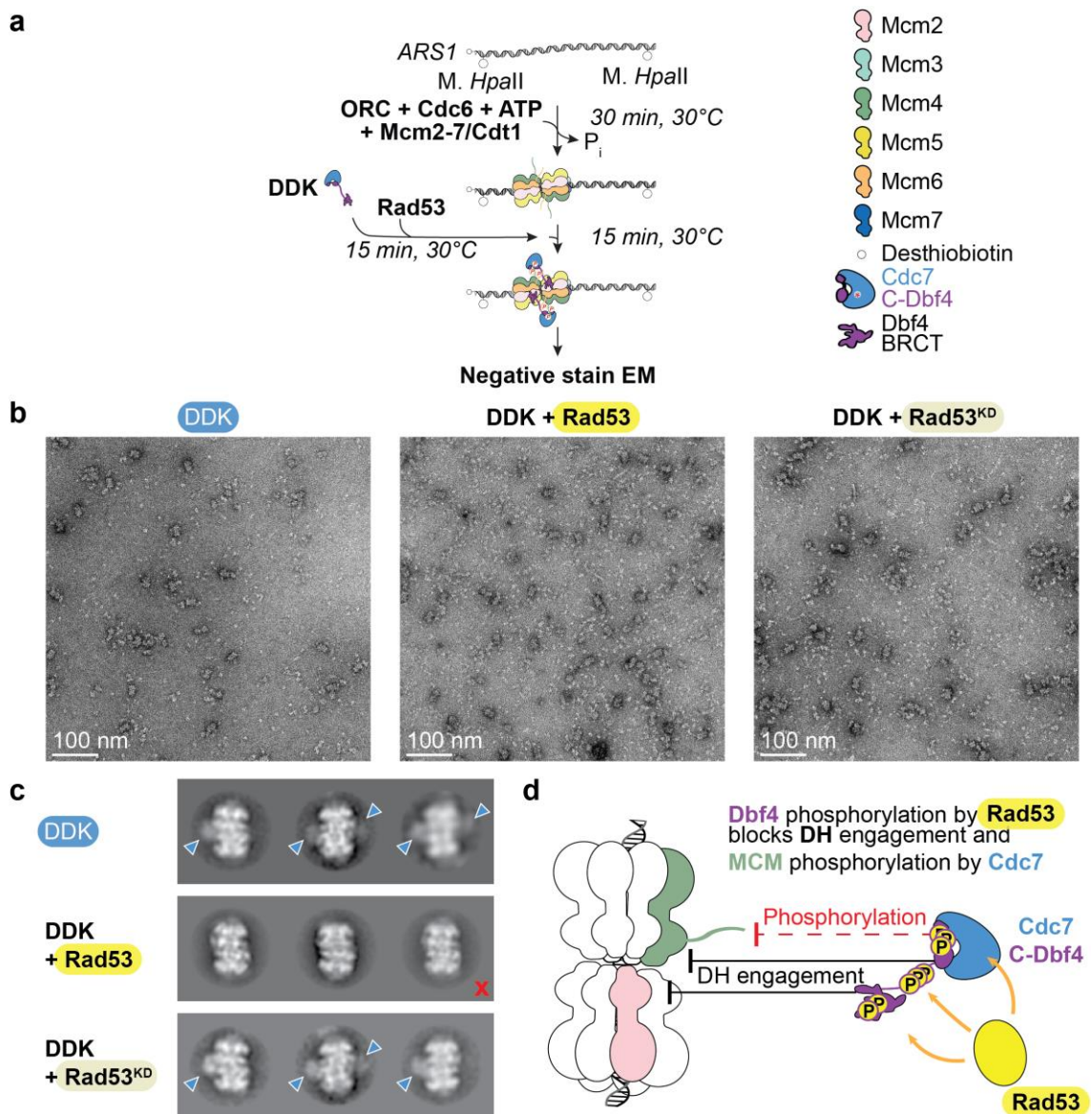


Figure 4.20: Phosphorylation of Dbf4 by Rad53 prevents DH–DDK interaction.

a. Overview of the DH–DDK reconstitution in the presence of Rad53. **b.** Representative negative stained electron micrographs of the DH loading reaction with DDK and Rad53 added. **c.** 2D classes of DH showed decoration with DDK (indicated by blue arrowheads), which was absent when DDK was treated with catalytically active Rad53. **d.** Model illustrating how phosphorylation of Dbf4 by Rad53 prevents DH engagement and phosphorylation.

4.5 Discussion

4.5.1 Selective phosphorylation of the DH by DDK

In this study, I have characterised the interaction between the *S. cerevisiae* Dbf4-dependent kinase and the DNA-loaded MCM double hexamer using *in vitro* reconstitution and cryo-EM. I thereby addressed how DDK selectively interacts with and phosphorylates the DH, which contains the symmetry to support bidirectional replication. The cryo-EM structure shows that DDK docks *via* the Dbf4 BRCT domain onto the Mcm2 subunit in one MCM ring (*in cis*), while the kinase core binds and phosphorylates the Mcm4 subunit in the opposing ring (*in trans*). I confirmed that the bipartite interaction is essential, given that an internal truncation variant of DDK, which lacks the BRCT domain, was not able to phosphorylate or bind to the DH. The structured domains of Dbf4, including the flexibly tethered BRCT domain, are conserved from yeast to human as well as in Drf1, which is an alternative Cdc7 activator in higher eukaryotes and particularly important during embryogenesis (Takahashi and Walter 2005, Silva et al. 2006, Collart et al. 2013, Collart et al. 2017). Conservation of the domain structure within Dbf4 orthologs and Drf1 suggests that the DH is recognised in the same way to initiate DNA replication.

This structural arrangement explains why DDK preferentially phosphorylates the Mcm4 subunit, which in turn is essential for subsequent origin firing (Deegan et al. 2016). It does, however, not explain how Mcm6 is phosphorylated. In parallel to our efforts, two groups have investigated the DH–DDK interaction. Both groups used chemical crosslinking prior to cryo-EM analysis (Cheng et al. 2022, Saleh et al. 2022). Crosslinking allowed to stabilise additional conformational states, including a state showing DDK tethered to the DH *via* the Dbf4 BRCT domain and the catalytic core displaced towards Mcm6. Although this structure was not solved to subnanometre resolution, which would allow atomic docking, it still suggests a mechanism by which the Mcm6 N-terminal tail substrate could reach the DDK active site and become phosphorylated.

The DH–DDK structure also explains why mutation of certain Dbf4 elements results in reduced viability. The Dbf4 Zn finger is one example. In the DH–DDK complex, this domain establishes a contact with the A domain of Mcm4. Loss of interaction would significantly reduce ability of DDK to interact with the MCM. The other element is the M domain of Dbf4, which, together with the Cdc7 KI-2 activation loop, contacts the Mcm4 B domain (Dick et al. 2020). Loss of these interaction interfaces would prevent kinase recruitment, MCM activation and therefore origin activation.

A point mutation in Mcm5 (L83P substitution), which is known as *bob1* mutation, enables *S. cerevisiae* cells to survive in the absence of DDK (Hardy et al. 1997). As DDK binds in the proximity of Mcm4 and Mcm5 residing in opposing MCM rings, the *bob1* mutation could lead to a conformational change that prevents DDK binding but enables recruitment of Sld3–7 and the helicase activator Cdc45. Indeed, structural changes in an archaeal MCM *bob1* have been observed (Fletcher et al. 2003).

DDK's role in the origin activation has also been described to relieve an auto-inhibitory function of the Mcm4 N-terminal tail (Sheu and Stillman 2010). Consistent with this notion, Cheng et al. (2022) have shown that the N-terminal Mcm4 tail is bound to the Mcm4 A domain and extends towards the Mcm6 subunit in the DH. Upon phosphorylation, the same domain becomes unstructured and the binding surface becomes available for other interactors.

In my study, I have also recapitulated previous findings, based on *H. sapiens* DDK, showing decreased DDK kinase activity upon autophosphorylation (Dick et al. 2020). Importantly, I discovered that the specific binding to the DH overrides this inhibitory effect, leading to an enhanced substrate specificity of DDK. This is an interesting example of how kinases regulate their activity and target specificity.

4.5.2 Regulation of DDK by Rad53

Upon detection of DNA damage during S phase, the checkpoint kinase Rad53 is activated and prevents further origin firing by targeting both Sld3 and DDK. I showed that phosphorylation of Dbf4 by Rad53 reduces DDK kinase activity. Furthermore, complete inhibition of DH phosphorylation occurs because the interaction between the DH and DDK is abrogated upon phosphorylation of Dbf4 by Rad53. Single and combined phospho-mimicking mutations in the Dbf4 BRCT domain did not affect DH engagement and phosphorylation. This agrees with previous work that described a different cluster of residues as the minimal set of Rad53 target sites (Zegerman and Diffley 2010). The set of residues map on the loop connecting the M and C domain of Dbf4. Cheng et al. (2022) have shown that the same loop interacts with the Mcm4 A domain, confirming the prediction we had made (Greiwe et al. 2022). Phosphorylation of the respective amino acids could therefore interfere with the binding of DDK to the DH. An additional effect of DDK phosphorylation by Rad53 could be caused by the large number of phosphorylation sites in Dbf4, as an altered surface charge of DDK upon phosphorylation by Rad53 could prevent substrate engagement.

Independent of its catalytic activity, Rad53 has been indicated to compete with the DH for binding to DDK, thereby inhibiting DH phosphorylation (Abd Wahab and Remus 2020). In our experiments, a catalytically inactive Rad53, which contained an additional point mutation compared to the variant used in the previous study, was not sufficient to prevent DH phosphorylation. We cannot rule out that the two Rad53 variants display different binding affinities to DDK, which would explain the discrepancies observed in the two studies.

Chapter 5. Concluding summary

In this thesis I have provided a structural analysis of the origin licensing reaction reconstituted *in vitro* using *H. sapiens* proteins. I identified five distinct protein assemblies, which include the OCCM, OC1M, MO, DNA-loaded single MCM hexamers and the MCM double hexamer. At a first glance, these complexes resemble the *S. cerevisiae* protein assemblies that are known to be helicase loading intermediates. However, I have identified significant differences between the two systems, as for example the conformation of the MO complex that is competent for loading the second MCM hexamer with *S. cerevisiae* but not with *H. sapiens* proteins. Negative stain EM analysis of the *H. sapiens* helicase loading reaction in conditions that prevent MO formation, in fact, show that DHs can still be loaded onto DNA. Our new structural data furthermore raises a fundamental question regarding the mechanism of DNA melting in the *H. sapiens* system compared to origin activation reconstituted with *S. cerevisiae* proteins. As such, one base pair at the dimerization interface of the two MCM hexamers is disrupted in the *H. sapiens* DH. On the contrary, *S. cerevisiae* DH loading does not promote disruption of Watson–Crick base pairing. Rather, the first evidence of DNA untwisting and melting has been identified in the downstream formation of the dCMGE origin firing intermediate (Douglas et al. 2018, Lewis et al. 2022). My structural analysis suggests that the two systems likely employ distinct mechanisms for nucleating DNA opening. First, the structural elements in MCM5 that disrupt base pairing in the *H. sapiens* DH, are conserved only in higher eukaryotes. Second, an Mcm7 helix-turn-helix motif, which interlocks with the Mcm5 A domain from the opposed ring in *S. cerevisiae*, is shorter in the *H. sapiens* ortholog, meaning that the two rings can rotate with respect to one another, contributing to DNA untwisting. In line with the hypothesis that DNA opening works differently in yeast and higher eukaryotes, nucleation of DNA melting occurs at distinct locations inside the MCM channel (at the dimerization interface in the *H. sapiens* DH and in the centre of each ATPase tier in the dCMGE). Alternatively, DNA melting could be nucleated from the dimerization interface in *S. cerevisiae* but become rapidly delocalised and stabilised by ATPase pore loops and the Mcm6 N-terminal β -hairpin insertion (Lewis et al. 2022). Time-resolved analysis of helicase

activation, both with *H. sapiens* as well as *S. cerevisiae* proteins, will be important to be able to compare and contrast DNA melting in the two systems.

I have also characterised protein–DNA interactions and the nucleotide binding state of the *S. cerevisiae* DH and uncovered the molecular mechanism of the selective phosphorylation of the DH by DDK. The symmetry of the head-to-head double hexamer appears to provide a failsafe mechanism to ensure that both helicases are activated, as they have to cross their paths to allow bidirectional replication to occur. I found that DDK recognises the three-dimensional shape of the DH, which confers substrate specificity. The DH–DDK structure does, however, not explain how both MCM hexamers are phosphorylated. In fact, 2D classes of the DH–DDK complex presented in this thesis and 3D structures by Saleh et al. (2022) and Cheng et al. (2022) show that two DDK molecules can bind simultaneously, although no interaction is detected. No conformational change in the DH is observed upon phosphorylation, which could promote the cooperative recruitment of a second DDK molecule. DDK binding therefore seems to be a stochastic event and I speculate that helicase activation might be synchronised at a later stage. Equivalent to the docking of DDK *in cis* and phosphorylation of Mcm4 *in trans*, recruitment of the helicase activator Cdc45 by Sld3–7 might occur *via* interactions across the two MCM rings. Sld3–7 could thereby bind to the phosphorylated Mcm4 N-terminal tail and deposit Cdc45 onto Mcm2 and Mcm5, which reside in direct vicinity in the opposing MCM ring. Simultaneous recruitment of two Cdc45 molecules could be achieved by dimerization of Sld3–7, as suggested by a crystal structure of an Sld7 domain, which forms symmetric dimers in the crystal lattice (Itou et al. 2015). Coordination of helicase activation could alternatively be achieved at the next step down the origin activation pathway, which is the recruitment of the second helicase activator GINS by the pre-loading complex. Future studies will need to capture the intermediates on the path to dCMGE formation to dissect the molecular mechanism leading activation of the two MCM helicase motors.

Chapter 6. Appendix

Appendix Table 1. Cryo-EM data collection, refinement and validation statistics for the *S. cerevisiae* DH and DH-DDK.

	MCM double hexamer (EMD-13176) (PDB 7P30)	DH-DDK (EMD-13211) (PDB 7P5Z)	
Data collection and processing			
Magnification	130,000		
Voltage (kV)	300		
Electron exposure (e ⁻ /Å ²)	51.3		
Energy filter slit width (eV)	20		
Defocus range (μm)	-2 to -4.1		
Pixel size (Å)	1.08		
Micrographs	18,135		
Symmetry imposed	C2	C1	
Initial particle images (no.)	3,529,085		
Final particle images (no.)	238,620	149,876	
Map resolution (Å)	2.95	3.3	
FSC threshold	0.143	0.143	
Map resolution range (Å)	2.7 – 4.5	3.0 – 8.5	
Refinement			
Initial model used (PDB code)			Model Resolution
	6F0L (DH-DNA)	6F0L (DH-DNA)	4.8
	5BK4 (DH-DNA)	5BK4 (DH-DNA)	3.9
	6EYC (DH-DNA)	6EYC (DH-DNA)	3.8
		6YA7 (Cdc7-Dbf4)	1.67
		3QBZ (Dbf4-BRCT)	2.69
Model composition			
Non-hydrogen atoms	62,909	67,425	
Protein and DNA residues	7,766	8,315	
Ligands (ATP, ADP, Mg ²⁺ , Zn ²⁺)	30	31	
B factors (Å²) (min/max/mean)			
Protein	3.37/73.99/27.47	77.38/416.67/133.20	
Nucleotide	79.86/141.97/106.37	178.21/262.13/202.32	
Ligand	8.86/93.94/27.55	102.71/268.16/126.96	
R.m.s. Deviations			
Bond lengths (Å)	0.003	0.003	
Bond angles (°)	0.551	0.568	
Validation			
MolProbity score	1.40	1.77	
Clashscore	4.34	10.77	
Poor rotamers (%)	0.06	0.03	
Ramachandran plot			
Favored (%)	96.88	96.62	
Allowed (%)	3.08	3.30	
Disallowed (%)	0.04	0.09	

Appendix Table 2. DDK-dependent phosphorylation sites of Mcm4 in the DH determined by mass spectrometry.

Gene names	Protein names	Positions within proteins	Intensity Trypsin_DH	Amino acid	Multiplicity	PEP Trypsin_DH	Number of Phospho (ST)	Localization prob	Score	Delta score	Score for localization	Mass error [ppm]	Intensity	Position in peptide	Sequence window	Score Trypsin_DH	Score diff Trypsin_DH	Localization prob Trypsin_DH	Best localization scan number	Best score scan number	Ratio modbase Trypsin_DH	Identification Type Trypsin_DH	Unique identifier
MCM4	DNA replication licensing factor	52	NaN	S	1	5.87E-29	NaN	1	255.39	180.14	221.85	-0.41231	71630000		3 YSSSSQGDYGRNNSQLSQEGNIRAAG	255.4	221.85	1	13142	11700	0.88232	By MS/MS	UID34
MCM4	DNA replication licensing factor	52	8.85511	S	2	5.87E-29	NaN	1	255.39	180.14	221.85	-0.41231	71630000		3 YSSSSQGDYGRNNSQLSQEGNIRAAG	255.4	221.85	1	13142	11700	0.88232	By MS/MS	UID134
MCM4	DNA replication licensing factor	52	NaN	S	3	5.87E-29	NaN	1	255.39	180.14	221.85	-0.41231	71630000		3 YSSSSQGDYGRNNSQLSQEGNIRAAG	255.4	221.85	1	13142	11700	0.88232	By MS/MS	UID234
MCM4	DNA replication licensing factor	56	8.5668	S	1	8.31E-20	NaN	1	221.85	190.99	221.85	-0.41231	1085100000		7 SSQGDYGRNNSQLSQEGNIRAAGSSPL	221.9	221.85	1	13142	13142	1.3366	By MS/MS	UID35
MCM4	DNA replication licensing factor	56	8.85511	S	2	8.31E-20	NaN	1	221.85	190.99	221.85	-0.41231	1085100000		7 SSQGDYGRNNSQLSQEGNIRAAGSSPL	221.9	221.85	1	13142	13142	1.3366	By MS/MS	UID135
MCM4	DNA replication licensing factor	56	NaN	S	3	8.31E-20	NaN	1	221.85	190.99	221.85	-0.41231	1085100000		7 SSQGDYGRNNSQLSQEGNIRAAGSSPL	221.9	221.85	1	13142	13142	1.3366	By MS/MS	UID235
MCM4	DNA replication licensing factor	68	NaN	S	1	0.00106	NaN	0.87816	119.37	93.934	119.37	0.44304	71914000		5 QNLSSQEGNIRAAGSSPLNFPSSSQRONSD	119.4	8.8075	0.87816	26945	26945	1.107	By MS/MS	UID26
MCM4	DNA replication licensing factor	68	6.90033	S	2	0.00106	NaN	0.87816	119.37	93.934	119.37	0.44304	71914000		5 QNLSSQEGNIRAAGSSPLNFPSSSQRONSD	119.4	8.8075	0.87816	26945	26945	1.107	By MS/MS	UID126
MCM4	DNA replication licensing factor	68	7.80594	S	3	0.00106	NaN	0.87816	119.37	93.934	119.37	0.44304	71914000		5 QNLSSQEGNIRAAGSSPLNFPSSSQRONSD	119.4	8.8075	0.87816	26945	26945	1.107	By MS/MS	UID28
MCM4	DNA replication licensing factor	75	NaN	S	1	0.00106	NaN	0.994513	119.37	93.934	96.805	0.18141	122890000		12 GNIRAAGSSPLNFPSSSQRONSDVFQSQGR	119.4	21.992	0.994513	29788	26945	1.6854	By MS/MS	UID27
MCM4	DNA replication licensing factor	75	7.65829	S	2	0.00106	NaN	0.994513	119.37	93.934	96.805	0.18141	122890000		12 GNIRAAGSSPLNFPSSSQRONSDVFQSQGR	119.4	21.992	0.994513	29788	26945	1.6854	By MS/MS	UID127
MCM4	DNA replication licensing factor	75	7.80594	S	3	0.00106	NaN	0.994513	119.37	93.934	96.805	0.18141	122890000		12 GNIRAAGSSPLNFPSSSQRONSDVFQSQGR	119.4	21.992	0.994513	29788	26945	1.6854	By MS/MS	UID227
MCM4	DNA replication licensing factor	76	NaN	S	1	0.00106	3	0.797281	113.05	80.106	96.805	0.18141	63955000		13 NIRAAGSSPLNFPSSSQRONSDVFQSQGRQ	113.1	5.9326	0.797281	29788	30024	0.9846	By MS/MS	UID28
MCM4	DNA replication licensing factor	76	NaN	S	2	0.00106	3	0.797281	113.05	80.106	96.805	0.18141	63955000		13 NIRAAGSSPLNFPSSSQRONSDVFQSQGRQ	113.1	5.9326	0.797281	29788	30024	0.9846	By MS/MS	UID128
MCM4	DNA replication licensing factor	76	7.80594	S	3	0.00106	3	0.797281	113.05	80.106	96.805	0.18141	63955000		13 NIRAAGSSPLNFPSSSQRONSDVFQSQGRQ	113.1	5.9326	0.797281	29788	30024	0.9846	By MS/MS	UID28
MCM4	DNA replication licensing factor	77	7.02198	S	1	0.00138	NaN	0.813056	107.63	77.997	60.834	-0.45528	23918000		14 IRAAGSSPLNFPSSSQRONSDVFQSQGRG	107.6	8.1439	0.813056	33085	24645	0.38817	By MS/MS	UID29

Appendix Table 3. DDK-dependent phosphorylation sites of Mcm4 in the DH determined by mass spectrometry (continued).

Gene names	Protein names	Positions within proteins	Intensity Trypsin_ DH	Amino acid	Multiplicity	PEP Trypsin_ DH	Number of Phospho (ST)	Localization prob	Score	Delta score	Score for localization	Mass error [ppm]	Intensity	Position in peptide	Sequence window	Trypsin_ n, DH	Localization prob	Best localization scan number	Ratio Trypsin_ H	Identification Trypsin_ DH	Unique identifier		
MCM4	DNA replication licensing factor MCM4	77	NaN	S	2	0.00138	NaN	0.813056	107.63	77.997	60.834	-0.45528	23918000	14	IRAAIGSSPLNFPSSSRQNSDVFSQGRQG	107.6	0.813056	33085	246.45	0.38817	By MS/MS	UID129	
MCM4	DNA replication licensing factor MCM4	77	7.12707	S	3	0.00138	NaN	0.813056	107.63	77.997	60.834	-0.45528	23918000	14	IRAAIGSSPLNFPSSSRQNSDVFSQGRQG	107.6	0.813056	33085	246.45	0.38817	By MS/MS	UID229	
MCM4	DNA replication licensing factor MCM4	82	7.86867	S	1	2.84E-05	NaN	1	158.34	123.57	155.09	0.44803	179150000	3	GSSPLNFPSSSQRQNSDVFSQGRQGRIRS	158.3	155.09	1	14770	13808	8.8759	By MS/MS	UID30
MCM4	DNA replication licensing factor MCM4	82	7.65829	S	2	2.84E-05	NaN	1	158.34	123.57	155.09	0.44803	179150000	3	GSSPLNFPSSSQRQNSDVFSQGRQGRIRS	158.3	155.09	1	14770	13808	8.8759	By MS/MS	UID130
MCM4	DNA replication licensing factor MCM4	82	7.12707	S	3	2.84E-05	NaN	1	158.34	123.57	155.09	0.44803	179150000	3	GSSPLNFPSSSQRQNSDVFSQGRQGRIRS	158.3	155.09	1	14770	13808	8.8759	By MS/MS	UID230
MCM4	DNA replication licensing factor MCM4	87	NaN	S	1	0.00032	2	1	155.09	135.34	155.09	0.44803	46316000	8	NFPSSSQRQNSDVFSQGRQGRIRSSASAS	155.1	155.09	1	14770	14770	3.0548	By MS/MS	UID31
MCM4	DNA replication licensing factor MCM4	87	7.66573	S	2	0.00032	2	1	155.09	135.34	155.09	0.44803	46316000	8	NFPSSSQRQNSDVFSQGRQGRIRSSASAS	155.1	155.09	1	14770	14770	3.0548	By MS/MS	UID131
MCM4	DNA replication licensing factor MCM4	87	NaN	S	3	0.00032	2	1	155.09	135.34	155.09	0.44803	46316000	8	NFPSSSQRQNSDVFSQGRQGRIRSSASAS	155.1	155.09	1	14770	14770	3.0548	By MS/MS	UID231
MCM4	DNA replication licensing factor MCM4	118	NaN	T	1	0.0045	2	0.998301	103.43	85.106	103.43	-0.40768	75996000	4	RSRYHSLRDLRALPTSSSLGRNGQNRVH	103.4	27.642	0.998301	166.49	2.4318	By MS/MS	UID69	
MCM4	DNA replication licensing factor MCM4	118	7.67159	T	2	0.0045	2	0.998301	103.43	85.106	103.43	-0.40768	75996000	4	RSRYHSLRDLRALPTSSSLGRNGQNRVH	103.4	27.642	0.998301	166.49	2.4318	By MS/MS	UID169	
MCM4	DNA replication licensing factor MCM4	118	NaN	T	3	0.0045	2	0.998301	103.43	85.106	103.43	-0.40768	75996000	4	RSRYHSLRDLRALPTSSSLGRNGQNRVH	103.4	27.642	0.998301	166.49	2.4318	By MS/MS	UID269	
MCM4	DNA replication licensing factor MCM4	140	6.39314	T	1	2.62E-08	NaN	1	165.89	137.66	165.89	0.07152	1609900000	6	RNGQNRVHMRNDIHTSDLSPPRIVDFDTR	165.9	63.735	1	9703	9703	1.697	By MS/MS	UID70
MCM4	DNA replication licensing factor MCM4	140	8.41577	T	2	2.62E-08	NaN	1	165.89	137.66	165.89	0.07152	1609900000	6	RNGQNRVHMRNDIHTSDLSPPRIVDFDTR	165.9	63.735	1	9703	9703	1.697	By MS/MS	UID170
MCM4	DNA replication licensing factor MCM4	140	7.91448	T	3	2.62E-08	NaN	1	165.89	137.66	165.89	0.07152	1609900000	6	RNGQNRVHMRNDIHTSDLSPPRIVDFDTR	165.9	63.735	1	9703	9703	1.697	By MS/MS	UID270
MCM4	DNA replication licensing factor MCM4	141	NaN	S	1	2.62E-08	NaN	0.999881	165.89	137.66	165.89	0.07152	676900000	7	NGQNRVHMRNDIHTSDLSPPRIVDFDTRS	165.9	47.181	0.999881	9703	1.6849	By MS/MS	UID32	
MCM4	DNA replication licensing factor MCM4	141	8.41577	S	2	2.62E-08	NaN	0.999881	165.89	137.66	165.89	0.07152	676900000	7	NGQNRVHMRNDIHTSDLSPPRIVDFDTRS	165.9	47.181	0.999881	9703	1.6849	By MS/MS	UID132	
MCM4	DNA replication licensing factor MCM4	141	7.91448	S	3	2.62E-08	NaN	0.999881	165.89	137.66	165.89	0.07152	676900000	7	NGQNRVHMRNDIHTSDLSPPRIVDFDTRS	165.9	47.181	0.999881	9703	1.6849	By MS/MS	UID232	
MCM4	DNA replication licensing factor MCM4	144	NaN	S	1	2.62E-08	NaN	0.996932	165.89	137.66	165.89	0.07152	1086400000	10	NRVHMRNDIHTSDLSPPRIVDFDTRSGVN	165.9	25.118	0.996932	9703	0.76385	By MS/MS	UID33	
MCM4	DNA replication licensing factor MCM4	144	7.86331	S	2	2.62E-08	NaN	0.996932	165.89	137.66	165.89	0.07152	1086400000	10	NRVHMRNDIHTSDLSPPRIVDFDTRSGVN	165.9	25.118	0.996932	9703	0.76385	By MS/MS	UID133	
MCM4	DNA replication licensing factor MCM4	144	7.91448	S	3	2.62E-08	NaN	0.996932	165.89	137.66	165.89	0.07152	1086400000	10	NRVHMRNDIHTSDLSPPRIVDFDTRSGVN	165.9	25.118	0.996932	9703	0.76385	By MS/MS	UID233	

Appendix Table 4. DDK-dependent phosphorylation sites of Mcm6 in the DH determined by mass spectrometry.

Gene names	Protein names	Positions within proteins	Intensity Trypsin_ DH	Amino acid	Multiplicity	PEP Trypsin_ DH	Number of Phospho (ST)	Score	Delta score	Score for localization	Mass error [ppm]	Intensity	Position in peptide	Sequence window	Score Trypsin_0_DH	Score diff Trypsin_0_DH	Localization prob Trypsin_0_DH	Best localization scan number	Best score scan number	Ratio modbase Trypsin_0_H	Identification type Trypsin_0_H	Unique identifier
MCM6	DNA replication licensing factor	232	8.47233	S	1	5.3E-10	NaN	178.51	159.68	77.795	-4.0444	351910000	2	EQQDDDMNGSSLPRDSSAAPGNGTSAMAT	178.5	5.8661	0.721237	7991	10694	0.37354	By MS/MS	UID56
MCM6	DNA replication licensing factor																					
MCM6	DNA replication licensing factor	1016	7.56843	S	1	1.3E-04	1	124.88	93.34	80.596	-0.19538	70739000	20	IHPNCEVLDQLEFPQDSS	124.9	1.0304	0.559038	31798	31983	0.049068	By MS/MS	UID57
MCM6	DNA replication licensing factor																					
MCM6	DNA replication licensing factor	1017	7.44408	S	1	1.3E-04	1	124.88	93.34	124.88	-0.34787	27802000	21	HPNCEVLDQLEFPQDSS	124.9	0	0.5	31983	31983	0.036851	By MS/MS	UID58
MCM6	DNA replication licensing factor																					
MCM6	DNA replication licensing factor	78	NaN	S	1	1.59E-23	NaN	231.38	219.92	190.37	-0.95633	47052000	6	NDSTQFSSQRLQTDGSAATNDMEGNEPARSK	231.4	39.091	0.999877	9475	9889	0.56612	By MS/MS	UID59
MCM6	DNA replication licensing factor																					
MCM6	DNA replication licensing factor	226	7.92268	S	1	3.83E-26	1	210.33	180.97	127.88	0.74117	65927000	22	GQADEDEQDDDMNGSSLPRDSSAAPGNG	210.3	13.604	0.958208	10851	12867	0.050975	By MS/MS	UID60
MCM6	DNA replication licensing factor																					
MCM6	DNA replication licensing factor	249	7.87366	S	1	2.17E-27	1	232.6	202.6	232.6	0.19894	74759000	1	SSAAPGNGTSAMATRSITTSPEQTERVFQ	232.6	14.696	0.967193	10491	10491	1.4544	By MS/MS	UID61
MCM6	DNA replication licensing factor																					
MCM6	DNA replication licensing factor	242	NaN	T	1	0.00064	2	112.58	89.093	112.58	-0.91701	0	12	SLPRDSSAAPGNGTSAMATRSITTSPE	112.6	8.8481	0.854008	6908	6908	0	By MS/MS	UID95
MCM6	DNA replication licensing factor																					
MCM6	DNA replication licensing factor	75	7.29647	T	1	1.6E-23	NaN	231.38	219.92	231.38	-0.39182	10839000	3	PRVNDSTQFSSQRLQTDGSAATNDMEGNEPAR	231.4	112.51	1	9889	9889	0.54413	By MS/MS	UID96
MCM6	DNA replication licensing factor																					
MCM6	DNA replication licensing factor	80	NaN	T	1	0.00013	2	136.11	119.36	136.11	-0.29094	23440000	8	STQFSSQRLQTDGSAATNDMEGNEPARSKSR	136.1	7.1548	0.840431	9756	9756	0.29503	By MS/MS	UID97
MCM6	DNA replication licensing factor																					
MCM6	DNA replication licensing factor	438	6.88947	T	1	0.00072	1	112.96	86.139	112.96	0.58991	7753000	10	TRGISKTTTEGLNSGVTLRSLGVRLTYKIS	113	17.597	0.984008	23649	23649	NaN	By MS/MS	UID98
MCM6	DNA replication licensing factor																					
MCM6	DNA replication licensing factor	232	6.83385	S	2	5.3E-10	NaN	178.51	159.68	77.795	-4.0444	351910000	2	EQQDDDMNGSSLPRDSSAAPGNGTSAMAT	178.5	5.8661	0.721237	7991	10694	0.37354	By MS/MS	UID156
MCM6	DNA replication licensing factor																					
MCM6	DNA replication licensing factor	1016	NaN	S	2	0.00013	1	124.88	93.34	80.596	-0.19538	70739000	20	IHPNCEVLDQLEFPQDSS	124.9	1.0304	0.559038	31798	31983	0.049068	By MS/MS	UID157
MCM6	DNA replication licensing factor																					
MCM6	DNA replication licensing factor	1017	NaN	S	2	0.00013	1	124.88	93.34	124.88	-0.34787	27802000	21	HPNCEVLDQLEFPQDSS	124.9	0	0.5	31983	31983	0.036851	By MS/MS	UID158
MCM6	DNA replication licensing factor																					
MCM6	DNA replication licensing factor	78	7.65298	S	2	1.6E-23	NaN	231.38	219.92	190.37	-0.95633	47052000	6	NDSTQFSSQRLQTDGSAATNDMEGNEPARSK	231.4	39.091	0.999877	9475	9889	0.56612	By MS/MS	UID159
MCM6	DNA replication licensing factor																					
MCM6	DNA replication licensing factor	226	NaN	S	2	3.8E-26	1	210.33	180.97	127.88	0.74117	65927000	22	GQADEDEQDDDMNGSSLPRDSSAAPGNG	210.3	13.604	0.958208	10851	12867	0.050975	By MS/MS	UID160

Appendix Table 5. DDK-dependent phosphorylation sites of Mcm6 in the DH determined by mass spectrometry (continued).

Gene names	Protein names	Positions within proteins	Intensity Trypsin_ DH	Amino acid	Multiplicity	PEP Trypsin_ DH	Number of Phospho (ST)	Localization prob.	Score	Delta score	Score for localization	Mass error [ppm]	Intensity	Position in peptide	Sequence window	Score Trypsin_ DH	Score diff Trypsin_ DH	Localization prob. Trypsin_ DH	Best scan number	Best score mod/base H	Ratio	Identification Type Trypsin_ DH	Unique identifier
MCM6	DNA replication licensing factor	226 NaN	NaN	S	2	3.8E-26	1	0.958208	210.33	180.97	127.88	0.74117	659270000	22	GQADEDEQQDDDMNGSSLPRDGSAAAPGNG	210.3	13.604	0.958208	10851	12867	0.050979	By MS/MS	UID160
MCM6	DNA replication licensing factor	249 NaN	NaN	S	2	2.2E-27	1	0.967183	232.6	202.6	232.6	0.19894	747590000	1	SSAAPGNGTSAATRSITTSISHEQTERVFQ	232.6	14.696	0.967183	10491	10491	1.4544	By MS/MS	UID161
MCM6	DNA replication licensing factor	242 NaN	NaN	T	2	0.00064	2	0.854008	112.58	89.093	112.58	-0.91701	0	12	SLPRDGSSSAAPGNGTSAATRSITTSISPE	112.6	8.8481	0.854008	6908	6908	0	By MS/MS	UID195
MCM6	DNA replication licensing factor	75 7.36996	T	2	1.6E-23	NaN	1	1	231.38	219.92	231.38	-0.38182	108390000	3	PFVNDSTQFSSQRLQTDGSATNDMEGNEPAR	231.4	112.51	1	9889	9889	0.54413	By MS/MS	UID196
MCM6	DNA replication licensing factor	80 7.36996	T	2	1.27E-04	2	0.840431	0.840431	136.11	119.36	136.11	-0.29094	23440000	8	STQFSSQRLQTDGSATNDMEGNEPARSKSR	136.1	7.1548	0.840431	9756	9756	0.29503	By MS/MS	UID197
MCM6	DNA replication licensing factor	438 NaN	NaN	T	2	7.16E-04	1	0.984008	112.96	86.139	112.96	0.58991	7753000	10	TRGISKTTEGLNSGVTGLSLGVRDLTYKIS	113	17.897	0.984008	23649	23649	NaN	By MS/MS	UID198
MCM6	DNA replication licensing factor	232 NaN	NaN	S	3	5.32E-10	NaN	0.721237	176.51	159.68	77.795	-4.0444	351810000	2	EQQDDMGSSSLPRDGSAAAPGNGTSAAT	178.5	5.8661	0.721237	7991	10694	0.37354	By MS/MS	UID256
MCM6	DNA replication licensing factor	1016 NaN	NaN	S	3	0.00013	1	0.559038	124.88	93.34	80.596	-0.19538	707390000	20	IHPNCEVLDQLERQDSS	124.9	1.0304	0.559038	31798	31983	0.049068	By MS/MS	UID257
MCM6	DNA replication licensing factor	1017 NaN	NaN	S	3	0.00013	1	0.5	124.88	93.34	124.88	-0.34787	278020000	21	IHPNCEVLDQLERQDSS	124.9	0	0.5	31983	31983	0.036851	By MS/MS	UID258
MCM6	DNA replication licensing factor	78 NaN	NaN	S	3	1.6E-23	NaN	0.998877	231.38	219.92	190.37	-0.95633	470520000	6	NDSTQFSSQRLQTDGSATNDMEGNEPARSK	231.4	39.091	0.998877	9475	9889	0.56612	By MS/MS	UID259
MCM6	DNA replication licensing factor	226 NaN	NaN	S	3	3.8E-26	1	0.958208	210.33	180.97	127.88	0.74117	659270000	22	GQADEDEQQDDDMNGSSLPRDGSAAAPGNG	210.3	13.604	0.958208	10851	12867	0.050979	By MS/MS	UID260
MCM6	DNA replication licensing factor	249 NaN	NaN	S	3	2.2E-27	1	0.967183	232.6	202.6	232.6	0.19894	747590000	1	SSAAPGNGTSAATRSITTSISHEQTERVFQ	232.6	14.696	0.967183	10491	10491	1.4544	By MS/MS	UID261
MCM6	DNA replication licensing factor	242 NaN	NaN	T	3	0.00064	2	0.854008	112.58	89.093	112.58	-0.91701	0	12	SLPRDGSSSAAPGNGTSAATRSITTSISPE	112.6	8.8481	0.854008	6908	6908	0	By MS/MS	UID295
MCM6	DNA replication licensing factor	75 NaN	NaN	T	3	1.59E-23	NaN	1	231.38	219.92	231.38	-0.38182	108390000	3	PFVNDSTQFSSQRLQTDGSATNDMEGNEPAR	231.4	112.51	1	9889	9889	0.54413	By MS/MS	UID296
MCM6	DNA replication licensing factor	80 NaN	NaN	T	3	1.27E-04	2	0.840431	136.11	119.36	136.11	-0.29094	23440000	8	STQFSSQRLQTDGSATNDMEGNEPARSKSR	136.1	7.1548	0.840431	9756	9756	0.29503	By MS/MS	UID297
MCM6	DNA replication licensing factor	438 NaN	NaN	T	3	7.16E-04	1	0.984008	112.96	86.139	112.96	0.58991	7753000	10	TRGISKTTEGLNSGVTGLSLGVRDLTYKIS	113	17.897	0.984008	23649	23649	NaN	By MS/MS	UID298

Appendix Table 6. DDK-dependent phosphorylation sites of Mcm4 and Mcm6 in MCM-Cdt1 determined by mass spectrometry.

Gene names	Protein names	Positions within proteins	Intensity MCM-Cdt1_MinusDDK 1	Intensity MCM-Cdt1_MinusDDK 2	Intensity MCM-Cdt1_MinusDDK 3	Intensity MCM-Cdt1_PlusDDK 1	Intensity MCM-Cdt1_PlusDDK 2	Intensity MCM-Cdt1_PlusDDK 3	Intensity MCM-Cdt1_MinusDDK	Anno acid	Intensity MCM-Cdt1_MinusDDK	Intensity 1	Intensity 2	Intensity 3	PEP	Score	Delta score	Score for localization	Mass error [ppm]	Number of Phospho (ST)	Localization prob MCM-Cdt1_MinusDDK	
MCM4	DNA replication licensing factor MCM4	52	NaN	NaN	7.640541	NaN	7.640541	NaN	7.640541061	S	NaN	7.640541	NaN	0.0003625	151.07	122.29	151.07	-0.42214		2	NaN	
	DNA replication licensing factor MCM4	56	NaN	NaN	7.640541	NaN	7.640541	NaN	7.640541061	S	NaN	7.640541	NaN	0.0003625	151.07	122.29	151.07	-0.42214		2	NaN	
	DNA replication licensing factor MCM4	878	NaN	7.067517	NaN	NaN	7.067517	NaN	7.067517202	S	NaN	7.06752	NaN	0.0012946	81.974	46.434	81.974	-0.03315		1	NaN	
	DNA replication licensing factor MCM4	414	NaN	NaN	NaN	NaN	NaN	NaN	NaN	S	NaN	NaN	NaN	0.0017646	80.406	57.087	80.406	0.65143		1	NaN	
MCM6	DNA replication licensing factor MCM6	9	NaN	NaN	NaN	NaN	NaN	NaN	NaN	T	NaN	NaN	NaN	0.0046207	65.504	42.136	65.504	-0.25325	NaN		0.145385	
	DNA replication licensing factor MCM6	16	7.131041	NaN	NaN	NaN	NaN	NaN	7.13104093	S	7.13104093	7.13104	NaN	1.667E-12	114.06	98.539	114.06	0.19324		1	0.639828	
	DNA replication licensing factor MCM6	18	NaN	NaN	NaN	NaN	NaN	NaN	NaN	S	NaN	NaN	NaN	0.0083609	53.493	35.588	53.493	1.5095	NaN		0.221948	
	DNA replication licensing factor MCM6	19	NaN	NaN	NaN	NaN	NaN	NaN	NaN	S	NaN	NaN	NaN	0.0083609	53.493	35.588	53.493	1.5095	NaN		0.221948	
MCM6	DNA replication licensing factor MCM6	31	NaN	NaN	NaN	NaN	NaN	NaN	NaN	S	NaN	NaN	NaN	0.0043738	72.746	57.929	72.746	-3.2532	NaN		NaN	
	DNA replication licensing factor MCM6	226	NaN	8.146128	NaN	NaN	8.146128036	NaN	8.146128036	S	NaN	8.45367	8.45367	NaN	4.644E-48	286.21	247.41	256.06	0.4588		1	NaN
	DNA replication licensing factor MCM6	232	NaN	7.967159	6.808002	NaN	7.967159	6.808002	7.996266437	S	NaN	7.99627	7.96716	6.808	8.067E-06	170.11	129.64	68.261	-1.6615	NaN		NaN
	DNA replication licensing factor MCM6	234	NaN	NaN	6.808002	NaN	NaN	6.808002	6.808001543	S	NaN	6.808	NaN	6.808	8.067E-06	170.11	129.64	68.261	-1.6615		2	NaN
MCM6	DNA replication licensing factor MCM6	249	7.396252	NaN	NaN	NaN	NaN	NaN	7.39625167	S	7.39625167	7.39625	NaN	0.0037495	176.87	122.92	176.87	0.063887		1	0.648498	
	DNA replication licensing factor MCM6	251	NaN	7.153083	NaN	NaN	7.153083	NaN	7.153082804	T	NaN	7.15308	NaN	0.0107847	118.37	74.164	118.37	1.0174		1	NaN	
	DNA replication licensing factor MCM6	1016	6.205096	NaN	NaN	8.145942	NaN	NaN	8.14594187	S	6.20509605	8.89978	8.89978	NaN	8.155E-14	116.13	96.496	110.64	0.47541		1	0
	DNA replication licensing factor MCM6	1017	6.205096	NaN	NaN	8.145942	NaN	NaN	8.14594187	S	6.20509605	8.20515	8.20515	NaN	8.155E-14	116.13	96.496	110.64	-0.2609		1	0

Appendix Table 7. DDK-dependent phosphorylation sites of Mcm4 and Mcm6 in MCM-Cdt1 determined by mass spectrometry (continued).

Gene names	Protein names	Positions within proteins	Score diff MCM-Cdt1_Minus DDK	PEP MCM-Cdt1_Minus DDK	Score MCM-Cdt1_Minus DDK	Localization prob MCM-Cdt1_Minus DDK	Score diff MCM-Cdt1_Minus DDK	PEP MCM-Cdt1_Minus DDK	Score MCM-Cdt1_Minus DDK	Localization prob	Position in peptide	Sequence window	Best localization evidence id	Best localization scan number	Best score scan number	Identification type
MCM4	DNA replication licensing factor MCM4	52	NaN	NaN	NaN	1	151.07	0.000362	151.07	1	3	YSSSSQGDYGRNNSQNL	98	3509	6251	By MS/MS
MCM4	DNA replication licensing factor MCM4	56	NaN	NaN	NaN	1	151.07	0.000362	151.07	1	7	SSQGDYGRNNSQNL	99	3509	6251	By MS/MS
MCM4	DNA replication licensing factor MCM4	878	NaN	NaN	NaN	0.786503	5.81285	0.001295	81.974	0.786503	11	EDLSREIMNVLKDQASDSMSFNLIKQINEH	96	875	28108	By MS/MS
MCM4	DNA replication licensing factor MCM4	414	NaN	NaN	NaN	0.931002	11.7877	0.001765	80.406	0.931002	16	LQETPDVFDGQTPHSLSLCVDELVDSCRA	97	2868	38115	By MS/MS
MCM6	DNA replication licensing factor MCM6	9	0.431043	0.00462071	65.504	NaN	NaN	NaN	NaN	0.145385	8	_____MSSFPFPADTPSSNRPNSNSPPSS	551	4873	22138	
MCM6	DNA replication licensing factor MCM6	16	6.44169	1.6672E-12	114.06	NaN	NaN	NaN	NaN	0.639828	15	MSSFPFPADTPSSNRPNSNSPPSSGAGFGS	289	4872	22113	By MS/MS
MCM6	DNA replication licensing factor MCM6	18	0	0.00836093	53.493	NaN	NaN	NaN	NaN	0.221948	17	SPFPADTPSSNRPNSNSPPSSGAGFGSS	290	4871	21341	
MCM6	DNA replication licensing factor MCM6	19	0	0.00836093	53.493	NaN	NaN	NaN	NaN	0.221948	18	FPADTPSSNRPNSNSPPSSGAGFGSSG	291	4871	21341	
MCM6	DNA replication licensing factor MCM6	31	NaN	NaN	NaN	0.393106	2.38835	0.004374	72.746	0.393106	30	SNSSPPSSGAGFGSSGLDSQIGSRLHFP	292	4874	24673	
MCM6	DNA replication licensing factor MCM6	228	NaN	NaN	NaN	0.963944	14.2708	4.64E-48	286.21	0.963944	21	GQADEQDDDMNGSSLPRDSGSSAAPGNG	287	4418	6625	By MS/MS
MCM6	DNA replication licensing factor MCM6	232	NaN	NaN	NaN	0.979445	14.7416	8.07E-06	170.11	0.979445	2	EQQDDDMNGSSLPRDSGSSAAPGNGT SAMAT	285	694	2674	By MS/MS
MCM6	DNA replication licensing factor MCM6	234	NaN	NaN	NaN	0.6191	2.0001	8.07E-06	170.11	0.6191	4	QDDDMNGSSLPRDSGSSAAPGNGT SAMATRS	286	694	2674	By MS/MS
MCM6	DNA replication licensing factor MCM6	249	2.66093	0.00374945	176.87	NaN	NaN	NaN	NaN	0.648498	1	SSAAPGNGT SAMATRSITTSISPEQTERVFQ	288	4682	4869	By MS/MS
MCM6	DNA replication licensing factor MCM6	251	NaN	NaN	NaN	0.959842	15.2316	0.010785	118.37	0.959842	3	AAPGNGT SAMATRSITTSISPEQTERVFQIS	550	4683	4827	By MS/MS
MCM6	DNA replication licensing factor MCM6	1016	0	NaN	NaN	0.572446	1.26744	8.16E-14	116.13	0.572446	20	IHPNCEVLDQLEPQDSS	283	5535	27173	By MS/MS
MCM6	DNA replication licensing factor MCM6	1017	0	NaN	NaN	0.5	0	8.16E-14	116.13	0.5	32	IHPNCEVLDQLEPQDSS	284	573	26084	By MS/MS

Reference List

- Abbas, T., M. A. Keaton and A. Dutta (2013). "Genomic instability in cancer." Cold Spring Harbor Perspectives in Biology **5**(3): a012914.
- Abd Wahab, S. and D. Remus (2020). "Antagonistic control of DDK binding to licensed replication origins by Mcm2 and Rad53." eLife **9**: e58571.
- Abid Ali, F., M. E. Douglas, J. Locke, V. E. Pye, A. Nans, J. F. X. Diffley and A. Costa (2017). "Cryo-EM structure of a licensed DNA replication origin." Nature Communications **8**(1): 2241.
- Abid Ali, F., L. Renault, J. Gannon, H. L. Gahlon, A. Kotecha, J. C. Zhou, D. Rueda and A. Costa (2016). "Cryo-EM structures of the eukaryotic replicative helicase bound to a translocation substrate." Nature Communications **7**: 10708.
- Arias, E. E. and J. C. Walter (2005). "Replication-dependent destruction of Cdt1 limits DNA replication to a single round per cell cycle in *Xenopus* egg extracts." Genes & Development **19**(1): 114-126.
- Arias, E. E. and J. C. Walter (2007). "Strength in numbers: preventing rereplication via multiple mechanisms in eukaryotic cells." Genes & Development **21**(5): 497-518.
- Attali, I., M. R. Botchan and J. M. Berger (2021). "Structural Mechanisms for Replicating DNA in Eukaryotes." Annual Review of Biochemistry **90**: 77-106.
- Baretić, D., M. Jenkyn-Bedford, V. Aria, G. Cannone, M. Skehel and J. T. P. Yeeles (2020). "Cryo-EM Structure of the Fork Protection Complex Bound to CMG at a Replication Fork." Molecular Cell **78**(5): 926-940 e913.
- Bastos de Oliveira, F. M., M. R. Harris, P. Brazauskas, R. A. de Bruin and M. B. Smolka (2012). "Linking DNA replication checkpoint to MBF cell-cycle transcription reveals a distinct class of G1/S genes." EMBO Journal **31**(7): 1798-1810.
- Bell, S. P. and K. Labib (2016). "Chromosome duplication in *Saccharomyces cerevisiae*." Genetics **203**(3): 1027-1067.
- Bellelli, R. and S. J. Boulton (2021). "Spotlight on the Replisome: Aetiology of DNA Replication-Associated Genetic Diseases." Trends in Genetics **37**(4): 317-336.
- Bepler, T., A. Morin, M. Rapp, J. Brasch, L. Shapiro, A. Noble and B. Berger (2019). "Positive-unlabeled convolutional neural networks for particle picking in cryo-electron micrographs." Nature Methods **16**(11): 1153-1160.
- Biondi, R. M. and A. R. Nebreda (2003). "Signalling specificity of Ser/Thr protein kinases through docking-site-mediated interactions." Biochemical Journal **372**(Pt 1): 1-13.
- Bleichert, F., M. Balasov, I. Chesnokov, E. Nogales, M. R. Botchan and J. M. Berger (2013). "A Meier-Gorlin syndrome mutation in a conserved C-terminal helix of Orc6 impedes origin recognition complex formation." eLife **2**: e00882.
- Bleichert, F., M. R. Botchan and J. M. Berger (2015). "Crystal structure of the eukaryotic origin recognition complex." Nature **519**(7543): 321-326.
- Bleichert, F., A. Leitner, R. Aebersold, M. R. Botchan and J. M. Berger (2018). "Conformational control and DNA-binding mechanism of the metazoan origin recognition complex." Proceedings of the National Academy of Sciences of the United States of America **115**(26): E5906-E5915.
- Bochman, M. L. and A. Schwacha (2009). "The Mcm complex: unwinding the mechanism of a replicative helicase." Microbiology and Molecular Biology Reviews **73**(4): 652-683.
- Boyer, A. S., D. Walter and C. S. Sorensen (2016). "DNA replication and cancer: From dysfunctional replication origin activities to therapeutic opportunities." Seminars in Cancer Biology **37-38**: 16-25.

- Brewster, A. S., G. Wang, X. Yu, W. B. Greenleaf, J. M. Carazo, M. Tjajadi, M. G. Klein and X. S. Chen (2008). "Crystal structure of a near-full-length archaeal MCM: functional insights for an AAA+ hexameric helicase." Proceedings of the National Academy of Sciences of the United States of America **105**(51): 20191-20196.
- Casañal, A., B. Lohkamp and P. Emsley (2020). "Current developments in Coot for macromolecular model building of Electron Cryo-microscopy and Crystallographic Data." Protein Science **29**(4): 1069-1078.
- Chadha, G. S., A. Gambus, P. J. Gillespie and J. J. Blow (2016). "Xenopus Mcm10 is a CDK-substrate required for replication fork stability." Cell Cycle **15**(16): 2183-2195.
- Chen, S. and S. P. Bell (2011). "CDK prevents Mcm2-7 helicase loading by inhibiting Cdt1 interaction with Orc6." Genes & Development **25**(4): 363-372.
- Cheng, J., N. Li, Y. Huo, S. Dang, B.-K. Tye, N. Gao and Y. Zhai (2022). "Structural Insight into the MCM double hexamer activation by Dbf4-Cdc7 kinase." Nature Communications **13**(1): 1396.
- Chesnokov, I. N., O. N. Chesnokova and M. Botchan (2003). "A cytokinetic function of *Drosophila* ORC6 protein resides in a domain distinct from its replication activity." Proceedings of the National Academy of Sciences of the United States of America **100**(16): 9150-9155.
- Cho, W. H., Y. J. Lee, S. I. Kong, J. Hurwitz and J. K. Lee (2006). "CDC7 kinase phosphorylates serine residues adjacent to acidic amino acids in the minichromosome maintenance 2 protein." Proceedings of the National Academy of Sciences of the United States of America **103**(31): 11521-11526.
- Christensen, T. W. and B. K. Tye (2003). "*Drosophila* Mcm10 Interacts with Members of the Prereplication Complex and Is Required for Proper Chromosome Condensation." Molecular Biology of the Cell **14**(6): 2206-2215.
- Collart, C., G. E. Allen, C. R. Bradshaw, J. C. Smith and P. Zegerman (2013). "Titration of Four Replication Factors is Essential for the *Xenopus laevis* Midblastula Transition." Science **341**(6148): 893-896.
- Collart, C., J. C. Smith and P. Zegerman (2017). "Chk1 Inhibition of the Replication Factor Drf1 Guarantees Cell-Cycle Elongation at the *Xenopus laevis* Mid-blastula Transition." Developmental Cell **42**(1): 82-96 e83.
- Costa, A. and J. F. X. Diffley (2022). "The Initiation of Eukaryotic DNA Replication." Annual Review of Biochemistry **91**: 107-131.
- Costa, A., I. Ilves, N. Tamberg, T. Petojevic, E. Nogales, M. R. Botchan and J. M. Berger (2011). "The structural basis for MCM2-7 helicase activation by GINS and Cdc45." Nature Structural & Molecular Biology **18**(4): 471-477.
- Costa, A., L. Renault, P. Swuec, T. Petojevic, J. J. Pesavento, I. Ilves, K. MacLellan-Gibson, R. A. Fleck, M. R. Botchan and J. M. Berger (2014). "DNA binding polarity, dimerization, and ATPase ring remodeling in the CMG helicase of the eukaryotic replisome." eLife **3**: e03273.
- Costa, A., G. van Duinen, B. Medagli, J. Chong, N. Sakakibara, Z. Kelman, S. K. Nair, A. Patwardhan and S. Onesti (2008). "Cryo-electron microscopy reveals a novel DNA-binding site on the MCM helicase." EMBO Journal **27**(16): 2250-2258.
- Coster, G. and J. F. X. Diffley (2017). "Bidirectional eukaryotic DNA replication is established by quasi-symmetrical helicase loading." Science **357**(6348): 314-318.
- Coster, G., J. Frigola, F. Beuron, E. P. Morris and J. F. Diffley (2014). "Origin Licensing Requires ATP Binding and Hydrolysis by the MCM Replicative Helicase." Molecular Cell **55**(5): 666-677.
- Croll, T. I. (2018). "ISOLDE: a physically realistic environment for model building into low-resolution electron-density maps." Acta Crystallographica Section D **74**(6): 519-530.

- Cullati, S. N., A. Chaikuad, J. S. Chen, J. Gebel, L. Tesmer, R. Zhubi, J. Navarrete-Perea, R. X. Guillen, S. P. Gygi, G. Hummer, V. Dotsch, S. Knapp and K. L. Gould (2022). "Kinase domain autophosphorylation rewires the activity and substrate specificity of CK1 enzymes." Molecular Cell **82**(11): 2006-2020 e2008.
- Davé, A., C. Cooley, M. Garg and A. Bianchi (2014). "Protein phosphatase 1 recruitment by Rif1 regulates DNA replication origin firing by counteracting DDK activity." Cell Reports **7**(1): 53-61.
- Davey, M. J., C. Indiani and M. O'Donnell (2003). "Reconstitution of the Mcm2-7p heterohexamer, subunit arrangement, and ATP site architecture." Journal of Biological Chemistry **278**(7): 4491-4499.
- De Ioannes, P., V. A. Leon, Z. Kuang, M. Wang, J. D. Boeke, A. Hochwagen and K.-J. Armache (2019). "Structure and function of the Orc1 BAH-nucleosome complex." Nature Communications **10**(1): 2894.
- Deegan, T. D., J. T. Yeeles and J. F. Diffley (2016). "Phosphopeptide binding by Sld3 links Dbf4-dependent kinase to MCM replicative helicase activation." EMBO Journal **35**(9): 961-973.
- Dhar, S. K., L. Delmolino and A. Dutta (2001). "Architecture of the human origin recognition complex." Journal of Biological Chemistry **276**(31): 29067-29071.
- Dhar, S. K. and A. Dutta (2000). "Identification and characterization of the human ORC6 homolog." Journal of Biological Chemistry **275**(45): 34983-34988.
- Dick, S. D., S. Federico, S. M. Hughes, V. E. Pye, N. O'Reilly and P. Cherepanov (2020). "Structural Basis for the Activation and Target Site Specificity of CDC7 Kinase." Structure **28**(8): 954-962 e954.
- Diffley, J. F. and B. Stillman (1988). "Purification of a yeast protein that binds to origins of DNA replication and a transcriptional silencer." Proceedings of the National Academy of Sciences of the United States of America **85**(7): 2120-2124.
- Diffley, J. F. X., J. H. Cocker, S. J. Dowell and A. Rowley (1994). "Two steps in the assembly of complexes at yeast replication origins in vivo." Cell **78**: 303-316.
- DiMaio, F., N. Echols, J. J. Headd, T. C. Terwilliger, P. D. Adams and D. Baker (2013). "Improved low-resolution crystallographic refinement with Phenix and Rosetta." Nature Methods **10**(11): 1102-1104.
- Douglas, M. E., F. A. Ali, A. Costa and J. F. X. Diffley (2018). "The mechanism of eukaryotic CMG helicase activation." Nature **555**(7695): 265-268.
- Douglas, M. E. and J. F. X. Diffley (2016). "Recruitment of Mcm10 to sites of replication initiation requires direct binding to the minichromosome maintenance (MCM) complex." Journal of Biological Chemistry **291**(11): 5879-5888.
- Drury, L. S., G. Perkins and J. F. Diffley (2000). "The Cyclin Dependent Kinase Cdc28p Regulates Distinct Modes of Cdc6p Proteolysis during the Budding Yeast Cell Cycle." Current Biology **10**(5): 231-240.
- Duncker, B. P., K. Shimada, M. Tsai-Pflugfelder, P. Pasero and S. M. Gasser (2002). "An N-terminal domain of Dbf4p mediates interaction with both origin recognition complex (ORC) and Rad53p and can deregulate late origin firing." Proceedings of the National Academy of Sciences of the United States of America **99**(25): 16087-16092.
- Dutta, A. and S. P. Bell (1997). "Initiation of DNA replication in eukaryotic cells." Annual Reviews in Cell and Developmental Biology **13**: 293-332.
- Eaton, M. L., K. Galani, S. Kang, S. P. Bell and D. M. MacAlpine (2010). "Conserved nucleosome positioning defines replication origins." Genes & Development **24**(8): 748-753.
- Eickhoff, P., H. B. Kose, F. Martino, T. Petojevic, F. Abid Ali, J. Locke, N. Tamberg, A. Nans, J. M. Berger, M. R. Botchan, H. Yardimci and A. Costa (2019). "Molecular Basis for ATP-Hydrolysis-Driven DNA Translocation by the CMG Helicase of the Eukaryotic Replisome." Cell Reports **28**(10): 2673-2688 e2678.

- Enemark, E. J. and L. Joshua-Tor (2006). "Mechanism of DNA translocation in a replicative hexameric helicase." Nature **442**(7100): 270-275.
- Enemark, E. J. and L. Joshua-Tor (2008). "On helicases and other motor proteins." Current Opinion in Structural Biology **18**(2): 243-257.
- Erzberger, J. P., M. M. Pirruccello and J. M. Berger (2002). "The structure of bacterial DnaA: implications for general mechanisms underlying DNA replication initiation." EMBO Journal **21**(18): 4763-4773.
- Evans, R., M. O'Neill, A. Pritzel, N. Antropova, A. Senior, T. Green, A. Žídek, R. Bates, S. Blackwell, J. Yim, O. Ronneberger, S. Bodenstein, M. Zielinski, A. Bridgland, A. Potapenko, A. Cowie, K. Tunyasuvunakool, R. Jain, E. Clancy, P. Kohli, J. Jumper and D. Hassabis (2022). "Protein complex prediction with AlphaFold-Multimer." BioRxiv.
- Evrin, C., P. Clarke, J. Zech, R. Lurz, J. Sun, S. Uhle, H. Li, B. Stillman and C. Speck (2009). "A double-hexameric MCM2-7 complex is loaded onto origin DNA during licensing of eukaryotic DNA replication." Proceedings of the National Academy of Sciences of the United States of America **106**(48): 20240-20245.
- Farshadi, E., G. T. J. van der Horst and I. Chaves (2020). "Molecular Links between the Circadian Clock and the Cell Cycle." Journal of Molecular Biology **432**(12): 3515-3524.
- Feng, X., Y. Noguchi, M. Barbon, B. Stillman, C. Speck and H. Li (2021). "The structure of ORC–Cdc6 on an origin DNA reveals the mechanism of ORC activation by the replication initiator Cdc6." Nature Communications **12**(1): 3883.
- Fernández-Cid, A., A. Riera, S. Tognetti, M. C. Herrera, S. Samel, C. Evrin, C. Winkler, E. Gardenal, S. Uhle and C. Speck (2013). "An ORC/Cdc6/MCM2-7 Complex is Formed in a Multistep Reaction to Serve as a Platform for MCM Double-Hexamer Assembly." Molecular Cell **50**(4): 577-588.
- Ferreira, M. G., C. Santocanale, L. S. Drury and J. F. X. Diffley (2000). "Dbf4p, an essential S phase promoting factor, is targeted for degradation by the Anaphase Promoting Complex." **20**: 242-248.
- Fletcher, R. J., B. E. Bishop, R. P. Leon, R. A. Sclafani, C. M. Ogata and X. S. Chen (2003). "The structure and function of MCM from archaeal *M. Thermoautotrophicum*." Nature Structural Biology **10**(3): 160-167.
- Forsburg, S. L. (2004). "Eukaryotic MCM proteins: beyond replication initiation." Microbiology and Molecular Biology Reviews **68**(1): 109-131.
- Francis, L. I., J. C. Randell, T. J. Takara, L. Uchima and S. P. Bell (2009). "Incorporation into the prereplicative complex activates the Mcm2-7 helicase for Cdc7-Dbf4 phosphorylation." Genes & Development **23**(5): 643-654.
- Frigola, J., J. He, K. Kinkelin, V. E. Pye, L. Renault, M. E. Douglas, D. Remus, P. Cherepanov, A. Costa and J. F. X. Diffley (2017). "Cdt1 stabilizes an open MCM ring for helicase loading." Nature Communications **8**: 15720.
- Frigola, J., D. Remus, A. Mehanna and J. F. X. Diffley (2013). "ATPase-dependent quality control of DNA replication origin licensing." Nature **495**(7441): 339-343.
- Froelich, C. A., S. Kang, L. B. Epling, S. P. Bell and E. J. Enemark (2014). "A conserved MCM single-stranded DNA binding element is essential for replication initiation." eLife **3**: e01993.
- Gambus, A., R. C. Jones, A. Sanchez-Diaz, M. Kanemaki, F. van Deursen, R. D. Edmondson and K. Labib (2006). "GIN5 maintains association of Cdc45 with MCM in replisome progression complexes at eukaryotic DNA replication forks." Nat Cell Biol **8**: 358-366.
- Gambus, A., F. van Deursen, D. Polychronopoulos, M. Foltman, R. C. Jones, R. D. Edmondson, A. Calzada and K. Labib (2009). "A key role for Ctf4 in coupling the MCM2-7 helicase to DNA polymerase alpha within the eukaryotic replisome." EMBO Journal **28**: 2992-3004.

- Georgescu, R., Z. Yuan, L. Bai, R. de Luna Almeida Santos, J. Sun, D. Zhang, O. Yurieva, H. Li and M. E. O'Donnell (2017). "Structure of eukaryotic CMG helicase at a replication fork and implications to replisome architecture and origin initiation." Proceedings of the National Academy of Sciences of the United States of America **114**(5): E697-E706.
- Ghatalia, P., E. S. Yang, B. N. Lasseigne, R. C. Ramaker, S. J. Cooper, D. Chen, S. Sudarshan, S. Wei, A. S. Guru, A. Zhao, T. Cooper, D. L. Della Manna, G. Naik, R. M. Myers and G. Sonpavde (2016). "Kinase Gene Expression Profiling of Metastatic Clear Cell Renal Cell Carcinoma Tissue Identifies Potential New Therapeutic Targets." PLOS ONE **11**(8): e0160924.
- Gilbert, C. S., C. M. Green and N. F. Lowndes (2001). "Budding yeast Rad9 is an ATP-dependent Rad53 activating machine." Molecular Cell **8**(1): 129-136.
- Gillespie, P. J., A. Li and J. J. Blow (2001). "Reconstitution of licensed replication origins on *Xenopus* sperm nuclei using purified proteins." BMC Biochemistry **2**: 15.
- Goddard, T. D., C. C. Huang, E. C. Meng, E. F. Pettersen, G. S. Couch, J. H. Morris and T. E. Ferrin (2018). "UCSF ChimeraX: Meeting modern challenges in visualization and analysis." Protein Science **27**(1): 14-25.
- Gógl, G., A. P. Kornev, A. Reményi and S. S. Taylor (2019). "Disordered Protein Kinase Regions in Regulation of Kinase Domain Cores." Trends in Biochemical Sciences **44**(4): 300-311.
- Goswami, P., F. Abid Ali, M. E. Douglas, J. Locke, A. Purkiss, A. Janska, P. Eickhoff, A. Early, A. Nans, A. M. C. Cheung, J. F. X. Diffley and A. Costa (2018). "Structure of DNA-CMG-Pol epsilon elucidates the roles of the non-catalytic polymerase modules in the eukaryotic replisome." Nature Communications **9**(1): 5061.
- Greiwe, J. F., T. C. R. Miller, J. Locke, F. Martino, S. Howell, A. Schreiber, A. Nans, J. F. X. Diffley and A. Costa (2022). "Structural mechanism for the selective phosphorylation of DNA-loaded MCM double hexamers by the Dbf4-dependent kinase." Nature Structural & Molecular Biology **29**(1): 10-20.
- Gros, J., C. Kumar, G. Lynch, T. Yadav, I. Whitehouse and D. Remus (2015). "Post-licensing Specification of Eukaryotic Replication Origins by Facilitated Mcm2-7 Sliding along DNA." Molecular Cell **60**(5): 797-807.
- Guerrero-Puigdevall, M., N. Fernandez-Fuentes and J. Frigola (2021). "Stabilisation of half MCM ring by Cdt1 during DNA insertion." Nature Communications **12**(1): 1746.
- Hardy, C. F., O. Dryga, S. Seematter, P. M. Pahl and R. A. Sclafani (1997). "Mcm5/Cdc46-bob1 bypasses the requirement for the S phase activator Cdc7p." Proceedings of the National Academy of Sciences of the United States of America **94**(7): 3151-3155.
- Heller, R. C., S. Kang, W. M. Lam, S. Chen, C. S. Chan and S. P. Bell (2011). "Eukaryotic Origin-Dependent DNA Replication *in vitro* Reveals Sequential Action of DDK and S-CDK Kinases." Cell **146**(1): 80-91.
- Hiraga, S., G. M. Alvino, F. Chang, H. Y. Lian, A. Sridhar, T. Kubota, B. J. Brewer, M. Weinreich, M. K. Raghuraman and A. D. Donaldson (2014). "Rif1 controls DNA replication by directing Protein Phosphatase 1 to reverse Cdc7-mediated phosphorylation of the MCM complex." Genes & Development **28**(4): 372-383.
- Homesley, L., M. Lei, Y. Kawasaki, S. Sawyer, T. Christensen and B. K. Tye (2000). "Mcm10 and the MCM2-7 complex interact to initiate DNA synthesis and to release replication factors from origins." Genes & Development **14**(8): 913-926.
- Hossain, M., K. Bhalla and B. Stillman (2021). "Multiple, short protein binding motifs in ORC1 and CDC6 control the initiation of DNA replication." Molecular Cell **81**(9): 1951-1969 e1956.

- Hu, Y., A. Tareen, Y. J. Sheu, W. T. Ireland, C. Speck, H. Li, L. Joshua-Tor, J. B. Kinney and B. Stillman (2020). "Evolution of DNA replication origin specification and gene silencing mechanisms." Nature Communications **11**(1): 5175.
- Huang, H., C. B. Stromme, G. Saredi, M. Hodl, A. Strandsby, C. Gonzalez-Aguilera, S. Chen, A. Groth and D. J. Patel (2015). "A unique binding mode enables MCM2 to chaperone histones H3-H4 at replication forks." Nature Structural & Molecular Biology **22**(8): 618-626.
- Hughes, S., F. Elustondo, A. Di Fonzo, F. G. Leroux, A. C. Wong, A. P. Snijders, S. J. Matthews and P. Cherepanov (2012). "Crystal structure of human CDC7 kinase in complex with its activator DBF4." Nature Structural & Molecular Biology **19**(11): 1101-1107.
- Hughes, S., V. Jenkins, M. J. Dar, A. Engelman and P. Cherepanov (2010). "Transcriptional co-activator LEDGF interacts with Cdc7-activator of S-phase kinase (ASK) and stimulates its enzymatic activity." Journal of Biological Chemistry **285**(1): 541-554.
- Huijbregts, R. P., A. Svitin, M. W. Stinnett, M. B. Renfrow and I. Chesnokov (2009). "Drosophila Orc6 Facilitates GTPase Activity and Filament Formation of the Septin Complex." Molecular Biology of the Cell **20**(1): 270-281.
- Huse, M. and J. Kuriyan (2002). "The conformational plasticity of protein kinases." Cell **109**(3): 275-282.
- Ilves, I., T. Petojevic, J. J. Pesavento and M. R. Botchan (2010). "Activation of the MCM2-7 helicase by association with Cdc45 and GINS proteins." Molecular Cell **37**(2): 247-258.
- Itou, H., Y. Shirakihara and H. Araki (2015). "The quaternary structure of the eukaryotic DNA replication proteins Sld7 and Sld3." Acta Crystallographica Section D Biological Crystallography **71**(Pt 8): 1649-1656.
- Izumi, M., F. Yatagai and F. Hanaoka (2001). "Cell cycle-dependent proteolysis and phosphorylation of human Mcm10." Journal of Biological Chemistry **276**(51): 48526-48531.
- Jackson, A. L., P. M. Pahl, K. Harrison, J. Rosamond and R. A. Sclafani (1993). "Cell cycle regulation of the yeast CDC7 protein kinase by association with the DBF4 protein." Molecular and Cellular Biology **13**(5): 2899-2908.
- Jaremko, M. J., K. F. On, D. R. Thomas, B. Stillman and L. Joshua-Tor (2020). "The dynamic nature of the human origin recognition complex revealed through five cryoEM structures." eLife **9**: e58622.
- Jenkinson, E. R. and J. P. Chong (2006). "Minichromosome maintenance helicase activity is controlled by N- and C-terminal motifs and requires the ATPase domain helix-2 insert." Proceedings of the National Academy of Sciences of the United States of America **103**(20): 7613-7618.
- Jiang, W., N. J. Wells and T. Hunter (1999). "Multistep regulation of DNA replication by Cdk phosphorylation of HsCdc6." Proceedings of the National Academy of Sciences of the United States of America **96**(11): 6193-6198.
- Johnston, L. H. and A. P. Thomas (1982). "The isolation of new DNA synthesis mutants in the yeast *Saccharomyces cerevisiae*." Molecular & General Genetics **186**(3): 439-444.
- Jones, D. R., A. A. Prasad, P. K. Chan and B. P. Duncker (2010). "The Dbf4 motif C zinc finger promotes DNA replication and mediates resistance to genotoxic stress." Cell Cycle **9**(10): 2018-2026.
- Jones, M. J. K., C. Gelot, S. Munk, A. Koren, Y. Kawasoe, K. A. George, R. E. Santos, J. V. Olsen, S. A. McCarroll, M. G. Frattini, T. S. Takahashi and P. V. Jallepalli (2021). "Human DDK rescues stalled forks and counteracts checkpoint inhibition at unfired origins to complete DNA replication." Molecular Cell **81**(3): 426-441.e428.

- Jorgensen, P. and M. Tyers (2004). "How cells coordinate growth and division." Current Biology **14**(23): R1014-1027.
- Kanemaki, M. and K. Labib (2006). "Distinct roles for Sld3 and GINS during establishment and progression of eukaryotic DNA replication forks." EMBO Journal **25**(8): 1753-1763.
- Kang, S., M. D. Warner and S. P. Bell (2014). "Multiple Functions for Mcm2-7 ATPase Motifs during Replication Initiation." Molecular Cell **55**(5): 655-665.
- Kanke, M., Y. Kodama, T. S. Takahashi, T. Nakagawa and H. Masukata (2012). "Mcm10 plays an essential role in origin DNA unwinding after loading of the CMG components." EMBO Journal **31**(9): 2182-2194.
- Kapuy, O., E. He, S. Lopez-Aviles, F. Uhlmann, J. J. Tyson and B. Novak (2009). "System-level feedbacks control cell cycle progression." FEBS Letters **583**(24): 3992-3998.
- Kidmose, R. T., J. Juhl, P. Nissen, T. Boesen, J. L. Karlsen and B. P. Pedersen (2019). "Namdinator - automatic molecular dynamics flexible fitting of structural models into cryo-EM and crystallography experimental maps." International Union of Crystallography Journal **6**(Pt 4): 526-531.
- Kim, B. J., S. Y. Kim and H. Lee (2007). "Identification and Characterization of Human Cdc7 Nuclear Retention and Export Sequences in the Context of Chromatin Binding." Journal of Biological Chemistry **282**(41): 30029-30038.
- Kim, B. J. and H. Lee (2006). "Importin- β Mediates Cdc7 Nuclear Import by Binding to the Kinase Insert II Domain, Which Can Be Antagonized by Importin- α ." Journal of Biological Chemistry **281**(17): 12041-12049.
- Kim, J. M., K. Nakao, K. Nakamura, I. Saito, M. Katsuki, K. Arai and H. Masai (2002). "Inactivation of Cdc7 kinase in mouse ES cells results in S-phase arrest and p53-dependent cell death." EMBO Journal **21**(9): 2168-2179.
- Kim, J. M., N. Sato, M. Yamada, K. Arai and H. Masai (1998). "Growth Regulation of the Expression of Mouse cDNA and Gene Encoding a Serine/Threonine Kinase Related to *Saccharomyces cerevisiae* CDC7 Essential for G1/S Transition. Structure, chromosomal localization, and expression of mouse gene for *S. cerevisiae* Cdc7-related kinase." Journal of Biological Chemistry **273**(36): 23248-23257.
- Kimanius, D., L. Dong, G. Sharov, T. Nakane and S. H. W. Scheres (2021). "New tools for automated cryo-EM single-particle analysis in RELION-4.0." Biochemical Journal **478**(24): 4169-4185.
- Knockleby, J., B. J. Kim, A. Mehta and H. Lee (2016). "Cdk1-mediated phosphorylation of Cdc7 suppresses DNA re-replication." Cell Cycle **15**(11): 1494-1505.
- Komata, M., M. Bando, H. Araki and K. Shirahige (2009). "The direct binding of Mrc1, a checkpoint mediator, to Mcm6, a replication helicase, is essential for the replication checkpoint against methyl methanesulfonate-induced stress." Molecular and Cellular Biology **29**(18): 5008-5019.
- Kumagai, H., N. Sato, M. Yamada, D. Mahony, W. Seghezzi, E. Lees, K. Arai and H. Masai (1999). "A novel growth- and cell cycle-regulated protein, ASK, activates human Cdc7-related kinase and is essential for G1/S transition in mammalian cells." Molecular and Cellular Biology **19**: 5083-5095.
- Kuo, A. J., J. Song, P. Cheung, S. Ishibe-Murakami, S. Yamazoe, J. K. Chen, D. J. Patel and O. Gozani (2012). "The BAH domain of ORC1 links H4K20me2 to DNA replication licensing and Meier-Gorlin syndrome." Nature **484**(7392): 115-119.
- Kurat, C. F., J. T. P. Yeeles, H. Patel, A. Early and J. F. X. Diffley (2017). "Chromatin Controls DNA Replication Origin Selection, Lagging-Strand Synthesis, and Replication Fork Rates." Molecular Cell **65**(1): 117-130.

- Labib, K. (2010). "How do Cdc7 and cyclin-dependent kinases trigger the initiation of chromosome replication in eukaryotic cells?" Genes & Development **24**(12): 1208-1219.
- Labib, K., J. F. X. Diffley and S. E. Kearsey (1999). "G1-phase and B-type cyclins exclude the DNA-replication factor Mcm4 from the nucleus." Nature Cell Biology **1**(7): 415-422.
- Lam, S. K., X. Ma, T. L. Sing, B. H. Shilton, C. J. Brandl and M. J. Davey (2013). "The PS1 hairpin of Mcm3 is essential for viability and for DNA unwinding in vitro." PLOS ONE **8**(12): e82177.
- Lee, A. Y., T. Chiba, L. N. Truong, A. N. Cheng, J. Do, M. J. Cho, L. Chen and X. Wu (2012). "Dbf4 is direct downstream target of ataxia telangiectasia mutated (ATM) and ataxia telangiectasia and Rad3-related (ATR) protein to regulate intra-S-phase checkpoint." Journal of Biological Chemistry **287**(4): 2531-2543.
- Lewis, J. S., M. H. Gross, J. Sousa, S. S. Henrikus, J. F. Greiwe, A. Nans, J. F. X. Diffley and A. Costa (2022). "Mechanism of replication origin melting nucleated by CMG helicase assembly." Nature **606**(7916): 1007-1014.
- Lewis, J. S., L. M. Spenkelink, G. D. Schauer, F. R. Hill, R. E. Georgescu, M. E. O'Donnell and A. M. van Oijen (2017). "Single-molecule visualization of *Saccharomyces cerevisiae* leading-strand synthesis reveals dynamic interaction between MTC and the replisome." Proceedings of the National Academy of Sciences of the United States of America **114**: 10630-10635.
- Li, J., J. Dong, W. Wang, D. Yu, X. Fan, Y. C. Hui, C. S. K. Lee, W. H. Lam, N. Alary, Y. Yang, Y. Zhang, Q. Zhao, C.-L. Chen, B.-K. Tye, S. Dang and Y. Zhai (2023). "The human pre-replication complex is an open complex." Cell **186**(1): 98-111.e121.
- Li, N., W. H. Lam, Y. Zhai, J. Cheng, E. Cheng, Y. Zhao, N. Gao and B. K. Tye (2018). "Structure of the origin recognition complex bound to DNA replication origin." Nature **559**(7713): 217-222.
- Li, N., Y. Zhai, Y. Zhang, W. Li, M. Yang, J. Lei, B. K. Tye and N. Gao (2015). "Structure of the eukaryotic MCM complex at 3.8 Å." Nature **524**(7564): 186-191.
- Liebschner, D., P. V. Afonine, M. L. Baker, G. Bunkoczi, V. B. Chen, T. I. Croll, B. Hintze, L. W. Hung, S. Jain, A. J. McCoy, N. W. Moriarty, R. D. Oeffner, B. K. Poon, M. G. Prisant, R. J. Read, J. S. Richardson, D. C. Richardson, M. D. Sammito, O. V. Sobolev, D. H. Stockwell, T. C. Terwilliger, A. G. Urzhumtsev, L. L. Videau, C. J. Williams and P. D. Adams (2019). "Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix." Acta crystallographica. Section D, Structural biology **75**(Pt 10): 861-877.
- Liu, C., R. Wu, B. Zhou, J. Wang, Z. Wei, B. K. Tye, C. Liang and G. Zhu (2012). "Structural insights into the Cdt1-mediated MCM2-7 chromatin loading." Nucleic Acids Res **40**: 3208-3217.
- Liu, S., M. Balasov, H. Wang, L. Wu, I. N. Chesnokov and Y. Liu (2011). "Structural analysis of human Orc6 protein reveals a homology with transcription factor TFIIB." Proceedings of the National Academy of Sciences of the United States of America **108**(18): 7373-7378.
- Löoke, M., M. F. Maloney and S. P. Bell (2017). "Mcm10 regulates DNA replication elongation by stimulating the CMG replicative helicase." Genes & Development **31**(3): 291-305.
- Lopez-Mosqueda, J., N. L. Maas, Z. O. Jonsson, L. G. Defazio-Eli, J. Wohlschlegel and D. P. Toczyski (2010). "Damage-induced phosphorylation of Sld3 is important to block late origin firing." Nature **467**(7314): 479-483.
- Lowary, P. T. and J. Widom (1998). "New DNA sequence rules for high affinity binding to histone octamer and sequence-directed nucleosome positioning." Journal of Molecular Biology **276**(1): 19-42.

- Luger, K., T. J. Rechsteiner and T. J. Richmond (1999). "Preparation of nucleosome core particle from recombinant histones." Methods in Enzymology **304**: 3-19.
- Lyubimov, A. Y., M. Strycharska and J. M. Berger (2011). "The nuts and bolts of ring-translocase structure and mechanism." Current Opinion in Structural Biology **21**(2): 240-248.
- MacAlpine, H. K., R. Gordân, S. K. Powell, A. J. Hartemink and D. M. MacAlpine (2010). "*Drosophila* ORC localizes to open chromatin and marks sites of cohesin complex loading." Genome Research **20**(2): 201-211.
- Marahrens, Y. and B. Stillman (1992). "A yeast chromosomal origin of DNA replication defined by multiple functional elements." Science **255**(5046): 817-823.
- Masai, H., S. Matsumoto, Z. You, N. Yoshizawa-Sugata and M. Oda (2010). "Eukaryotic chromosome DNA replication: where, when, and how?" Annual Review of Biochemistry **79**: 89-130.
- Masai, H., C. Taniyama, K. Ogino, E. Matsui, N. Kakusho, S. Matsumoto, J. M. Kim, A. Ishii, T. Tanaka, T. Kobayashi, K. Tamai, K. Ohtani and K. Arai (2006). "Phosphorylation of MCM4 by Cdc7 kinase facilitates its interaction with Cdc45 on the chromatin." Journal of Biological Chemistry **281**(51): 39249-39261.
- Mattarocci, S., M. Shyian, L. Lemmens, P. Damay, D. M. Altintas, T. Shi, C. R. Bartholomew, N. H. Thoma, C. F. Hardy and D. Shore (2014). "Rif1 controls DNA replication timing in yeast through the PP1 phosphatase Glc7." Cell Reports **7**(1): 62-69.
- Matthews, L. A., D. R. Jones, A. A. Prasad, B. P. Duncker and A. Guarné (2012). "*Saccharomyces cerevisiae* Dbf4 has unique fold necessary for interaction with Rad53 kinase." Journal of Biological Chemistry **287**(4): 2378-2387.
- McClure, A. W. and J. F. X. Diffley (2021). "Rad53 checkpoint kinase regulation of DNA replication fork rate via Mrc1 phosphorylation." eLife **10**.
- McGarry, T. J. and M. W. Kirschner (1998). "Geminin, an inhibitor of DNA replication, is degraded during mitosis." Cell **93**: 1043-1053.
- McGeoch, A. T., M. A. Trakselis, R. A. Laskey and S. D. Bell (2005). "Organization of the archaeal MCM complex on DNA and implications for the helicase mechanism." Nature Structural & Molecular Biology **12**(9): 756-762.
- Meagher, M., A. Myasnikov and E. J. Enemark (2022) "Two Distinct Modes of DNA Binding by an MCM Helicase Enable DNA Translocation." International Journal of Molecular Sciences **23** DOI: 10.3390/ijms232314678.
- Mechali, M. (2010). "Eukaryotic DNA replication origins: many choices for appropriate answers." Nature Reviews Molecular Cell Biology **11**(10): 728-738.
- Méchal, M., K. Yoshida, P. Coulombe and P. Pasero (2013). "Genetic and epigenetic determinants of DNA replication origins, position and activation." Current Opinion in Genetics & Development **23**(2): 124-131.
- Méndez, J., X. H. Zou-Yang, S. Y. Kim, M. Hidaka, W. P. Tansey and B. Stillman (2002). "Human origin recognition complex large subunit is degraded by ubiquitin-mediated proteolysis after initiation of DNA replication." Molecular Cell **9**(3): 481-491.
- Miller, J. M., B. T. Arachea, L. B. Epling and E. J. Enemark (2014). "Analysis of the crystal structure of an active MCM hexamer." eLife **3**: e03433.
- Miller, T. C. and A. Costa (2017). "The architecture and function of the chromatin replication machinery." Current Opinion in Structural Biology **47**: 9-16.
- Miller, T. C. R., J. Locke, J. F. Greiwe, J. F. X. Diffley and A. Costa (2019). "Mechanism of head-to-head MCM double-hexamer formation revealed by cryo-EM." Nature **575**(7784): 704-710.
- Montagnoli, A., R. Bosotti, F. Villa, M. Rialland, D. Brotherton, C. Mercurio, J. Berthelsen and C. Santocanale (2002). "Drf1, a novel regulatory subunit for human Cdc7 kinase." EMBO Journal **21**(12): 3171-3181.

- Montagnoli, A., P. Tenca, F. Sola, D. Carpani, D. Brotherton, C. Albanese and C. Santocanale (2004). "Cdc7 inhibition reveals a p53-dependent replication checkpoint that is defective in cancer cells." Cancer Research **64**: 7110-7116.
- Montagnoli, A., B. Valsasina, D. Brotherton, S. Troiani, S. Rainoldi, P. Tenca, A. Molinari and C. Santocanale (2006). "Identification of Mcm2 phosphorylation sites by S-phase-regulating kinases." Journal of Biological Chemistry **281**(15): 10281-10290.
- Moyer, S. E., P. W. Lewis and M. R. Botchan (2006). "Isolation of the Cdc45/Mcm2-7/GINS (CMG) complex, a candidate for the eukaryotic DNA replication fork helicase." Proceedings of the National Academy of Sciences of the United States of America **103**(27): 10236-10241.
- Müller, P., S. Park, E. Shor, D. J. Huebert, C. L. Warren, A. Z. Ansari, M. Weinreich, M. L. Eaton, D. M. MacAlpine and C. A. Fox (2010). "The conserved bromo-adjacent homology domain of yeast Orc1 functions in the selection of DNA replication origins within chromatin." Genes & Development **24**(13): 1418-1433.
- Muñoz, S. and J. Méndez (2017). "DNA replication stress: from molecular mechanisms to human disease." Chromosoma **126**(1): 1-15.
- Muramatsu, S., K. Hirai, Y. S. Tak, Y. Kamimura and H. Araki (2010). "CDK-dependent complex formation between replication proteins Dpb11, Sld2, Pol ϵ , and GINS in budding yeast." Genes & Development **24**(6): 602-612.
- Nakane, T., D. Kimanius, E. Lindahl and S. H. W. Scheres (2018). "Characterisation of molecular motions in cryo-EM single-particle data by multi-body refinement in RELION." eLife **7**: e36861.
- Nambiar, S., A. Mirmohammadsadegh, M. Hassan, R. Mota, A. Marini, A. Alaoui, A. Tannapfel, J. H. Hegemann and U. R. Hengge (2007). "Identification and functional characterization of ASK/Dbf4, a novel cell survival gene in cutaneous melanoma with prognostic relevance." Carcinogenesis **28**(12): 2501-2510.
- Nazarenko, M. S., I. V. Viakhireva, M. Y. Skoblov, E. V. Soloveva, A. A. Sleptcov and L. P. Nazarenko (2022). "Meier-Gorlin Syndrome: Clinical Misdiagnosis, Genetic Testing and Functional Analysis of ORC6 Mutations and the Development of a Prenatal Test." International Journal of Molecular Sciences **23**(16).
- Nguyen, V. Q., C. Co, K. Irie and J. J. Li (2000). "Clb/Cdc28 kinases promote nuclear export of the replication initiator proteins Mcm2-7." Current Biology **10**(4): 195-205.
- Nishitani, H., N. Sugimoto, V. Roukos, Y. Nakanishi, M. Saijo, C. Obuse, T. Tsurimoto, K. I. Nakayama, K. Nakayama, M. Fujita, Z. Lygerou and T. Nishimoto (2006). "Two E3 ubiquitin ligases, SCF-Skp2 and DDB1-Cul4, target human Cdt1 for proteolysis." EMBO Journal **25**(5): 1126-1136.
- Noguchi, Y., Z. Yuan, L. Bai, S. Schneider, G. Zhao, B. Stillman, C. Speck and H. Li (2017). "Cryo-EM structure of Mcm2-7 double hexamer on DNA suggests a lagging-strand DNA extrusion model." Proceedings of the National Academy of Sciences of the United States of America **114**(45): E9529-E9538.
- O'Donnell, M., L. Langston and B. Stillman (2013). "Principles and concepts of DNA replication in bacteria, archaea, and eukarya." Cold Spring Harbor Perspectives in Biology **5**(7).
- Oliver, A. W., S. Knapp and L. H. Pearl (2007). "Activation segment exchange: a common mechanism of kinase autophosphorylation?" Trends in Biochemical Sciences **32**(8): 351-356.
- On, K. F., F. Beuron, D. Frith, A. P. Snijders, E. P. Morris and J. F. Diffley (2014). "Prereplicative complexes assembled in vitro support origin-dependent and independent DNA replication." EMBO Journal **33**(6): 605-620.

- Oshiro, G., J. C. Owens, Y. Shellman, R. A. Sclafani and J. J. Li (1999). "Cell cycle control of Cdc7p kinase activity through regulation of Dbf4p stability." Molecular and Cellular Biology **19**: 4888-4896.
- Pack, L. R., L. H. Daigh and T. Meyer (2019). "Putting the brakes on the cell cycle: mechanisms of cellular growth arrest." Current Opinion in Cell Biology **60**: 106-113.
- Pardo, B., L. Crabbe and P. Pasero (2017). "Signaling pathways of replication stress in yeast." FEMS Yeast Research **17**(2).
- Park, J. H., S. W. Bang, Y. Jeon, S. Kang and D. S. Hwang (2008). "Knockdown of human MCM10 exhibits delayed and incomplete chromosome replication." Biochemical and Biophysical Research Communications **365**(3): 575-582.
- Parker, M. W., M. Bell, M. Mir, J. A. Kao, X. Darzacq, M. R. Botchan and J. M. Berger (2019). "A new class of disordered elements controls DNA replication through initiator self-assembly." eLife **8**: e48562.
- Pérez-Arnaiz, P., A. Dattani, V. Smith and T. Allers (2020). "*Haloflex volcanii* – a model archaeon for studying DNA replication and repair." Open Biology **10**(12): 200293.
- Petryk, N., M. Dalby, A. Wenger, C. B. Stromme, A. Strandsby, R. Andersson and A. Groth (2018). "MCM2 promotes symmetric inheritance of modified histones during DNA replication." Science **361**(6409): 1389-1392.
- Pettersen, E. F., T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng and T. E. Ferrin (2004). "UCSF Chimera – a visualization system for exploratory research and analysis." Journal of Computational Chemistry **25**(13): 1605-1612.
- Pettersen, E. F., T. D. Goddard, C. C. Huang, E. C. Meng, G. S. Couch, T. I. Croll, J. H. Morris and T. E. Ferrin (2021). "UCSF ChimeraX: Structure visualization for researchers, educators, and developers." Protein Science **30**(1): 70-82.
- Poplawski, A., B. Grabowski, S. E. Long and Z. Kelman (2001). "The zinc finger domain of the archaeal minichromosome maintenance protein is required for helicase activity." Journal of Biological Chemistry **276**(52): 49371-49377.
- Prasanth, S. G., K. V. Prasanth and B. Stillman (2002). "Orc6 involved in DNA replication, chromosome segregation, and cytokinesis." Science **297**(5583): 1026-1031.
- Puchades, C., C. R. Sandate and G. C. Lander (2020). "The molecular principles governing the activity and functional diversity of AAA+ proteins." Nature Reviews Molecular Cell Biology **21**(1): 43-58.
- Punjani, A., J. L. Rubinstein, D. J. Fleet and M. A. Brubaker (2017). "cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination." Nature Methods **14**(3): 290-296.
- Ramer, M. D., E. S. Suman, H. Richter, K. Stanger, M. Spranger, N. Bieberstein and B. P. Duncker (2013). "Dbf4 and Cdc7 proteins promote DNA replication through interactions with distinct Mcm2-7 protein subunits." Journal of Biological Chemistry **288**(21): 14926-14935.
- Ramey, C. J. and R. A. Sclafani (2014). "Functional conservation of the pre-sensor one beta-finger hairpin (PS1-hp) structures in mini-chromosome maintenance proteins of *Saccharomyces cerevisiae* and archaea." G3 **4**(7): 1319-1326.
- Ramlal, K., C. M. Palmer and C. H. S. Aylett (2019). "A Local Agreement Filtering Algorithm for Transmission EM Reconstructions." Journal of Structural Biology **205**(1): 30-40.
- Randell, J. C., A. Fan, C. Chan, L. I. Francis, R. C. Heller, K. Galani and S. P. Bell (2010). "Mec1 is one of multiple kinases that prime the Mcm2-7 helicase for phosphorylation by Cdc7." Molecular Cell **40**(3): 353-363.
- Remus, D., E. L. Beall and M. R. Botchan (2004). "DNA topology, not DNA sequence, is a critical determinant for *Drosophila* ORC–DNA binding." EMBO Journal **23**(4): 897-907.

- Remus, D., F. Beuron, G. Tolun, J. D. Griffith, E. P. Morris and J. F. Diffley (2009). "Concerted loading of Mcm2-7 double hexamers around DNA during DNA replication origin licensing." Cell **139**(4): 719-730.
- Rohou, A. and N. Grigorieff (2015). "CTFFIND4: Fast and accurate defocus estimation from electron micrographs." Journal of Structural Biology **192**(2): 216-221.
- Rowley, A., J. H. Cocker, J. Harwood and J. F. Diffley (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 Journal **14**(11): 2631-2641.
- Rzechorzek, N. J., S. W. Hardwick, V. A. Jatikusumo, D. Y. Chirgadze and L. Pellegrini (2020). "CryoEM structures of human CMG-ATPyS-DNA and CMG-AND-1 complexes." Nucleic Acids Research **48**(12): 6980-6995.
- Sagendorf, J. M., H. M. Berman and R. Rohs (2017). "DNAproDB: an interactive tool for structural analysis of DNA-protein complexes." Nucleic Acids Research **45**(W1): W89-W97.
- Sagendorf, J. M., N. Markarian, H. M. Berman and R. Rohs (2020). "DNAproDB: an expanded database and web-based tool for structural analysis of DNA-protein complexes." Nucleic Acids Research **48**(D1): D277-D287.
- Saldivar, J. C., D. Cortez and K. A. Cimprich (2017). "The essential kinase ATR: ensuring faithful duplication of a challenging genome." Nature Reviews Molecular Cell Biology **18**(10): 622-636.
- Saleh, A., Y. Noguchi, R. Aramayo, M. E. Ivanova, K. M. Stevens, A. Montoya, S. Sunidhi, N. L. Carranza, M. J. Skwark and C. Speck (2022). "The structural basis of Cdc7-Dbf4 kinase dependent targeting and phosphorylation of the MCM2-7 double hexamer." Nature Communications **13**(1): 2915.
- Santocanale, C. and J. F. Diffley (1998). "A Mec1- and Rad53-dependent checkpoint controls late-firing origins of DNA replication." Nature **395**(6702): 615-618.
- Saravanan, M., J. Wuerges, D. Bose, E. A. McCormack, N. J. Cook, X. Zhang and D. B. Wigley (2012). "Interactions between the nucleosome histone core and Arp8 in the INO80 chromatin remodeling complex." Proceedings of the National Academy of Sciences of the United States of America **109**(51): 20883-20888.
- Scheres, S. H. W. (2016). Chapter Six - Processing of Structurally Heterogeneous Cryo-EM Data in RELION. Methods in Enzymology. R. A. Crowther, Academic Press. **579**: 125-157.
- Schmidt, J. M. and F. Bleichert (2020). "Structural mechanism for replication origin binding and remodeling by a metazoan origin recognition complex and its co-loader Cdc6." Nature Communications **11**(1): 4263.
- Schmit, M. and A. K. Bielinsky (2021). "Congenital Diseases of DNA Replication: Clinical Phenotypes and Molecular Mechanisms." International Journal of Molecular Sciences **22**(2).
- Segurado, M. and J. A. Tercero (2009). "The S-phase checkpoint: targeting the replication fork." Biology of the Cell **101**(11): 617-627.
- Sengupta, S., F. van Deursen, G. de Piccoli and K. Labib (2013). "Dpb2 integrates the leading-strand DNA polymerase into the eukaryotic replisome." Current Biology **23**(7): 543-552.
- Sheu, Y.-J., J. B. Kinney, A. Lengronne, P. Pasero and B. Stillman (2014). "Domain within the helicase subunit Mcm4 integrates multiple kinase signals to control DNA replication initiation and fork progression." Proceedings of the National Academy of Sciences of the United States of America **111**: E1899-E1908.
- Sheu, Y.-J., J. B. Kinney and B. Stillman (2016). "Concerted activities of Mcm4, Sld3, and Dbf4 in control of origin activation and DNA replication fork progression." Genome Research **26**(3): 315-330.

- Sheu, Y.-J. and B. Stillman (2006). "Cdc7-Dbf4 Phosphorylates MCM Proteins via a Docking Site-Mediated Mechanism to Promote S Phase Progression." Molecular Cell **24**(1): 101-113.
- Sheu, Y. J. and B. Stillman (2010). "The Dbf4-Cdc7 kinase promotes S phase by alleviating an inhibitory activity in Mcm4." Nature **463**(7277): 113-117.
- Shima, N., A. Alcaraz, I. Liachko, T. R. Buske, C. A. Andrews, R. J. Munroe, S. A. Hartford, B. K. Tye and J. C. Schimenti (2007). "A viable allele of Mcm4 causes chromosome instability and mammary adenocarcinomas in mice." Nature Genetics **39**(1): 93-98.
- Silva, T., R. H. Bradley, Y. Gao and M. Coue (2006). "Xenopus CDC7/DRF1 complex is required for the initiation of DNA replication." Journal of Biological Chemistry **281**(17): 11569-11576.
- Simon, A. C., J. C. Zhou, R. L. Perera, F. van Deursen, C. Evrin, M. E. Ivanova, M. L. Kilkenny, L. Renault, S. Kjaer, D. Matak-Vinkovic, K. Labib, A. Costa and L. Pellegrini (2014). "A Ctf4 trimer couples the CMG helicase to DNA polymerase alpha in the eukaryotic replisome." Nature **510**(7504): 293-297.
- Singleton, M. R., M. S. Dillingham and D. B. Wigley (2007). "Structure and mechanism of helicases and nucleic acid translocases." Annual Review of Biochemistry **76**: 23-50.
- Söding, J., A. Biegert and A. N. Lupas (2005). "The HHpred interactive server for protein homology detection and structure prediction." Nucleic Acids Research **33**(Web Server issue): W244-248.
- Stewart-Morgan, K. R., N. Petryk and A. Groth (2020). "Chromatin replication and epigenetic cell memory." Nature Cell Biology **22**(4): 361-371.
- Stiff, T., M. Alagoz, D. Alcantara, E. Outwin, H. G. Brunner, E. M. Bongers, M. O'Driscoll and P. A. Jeggo (2013). "Deficiency in origin licensing proteins impairs cilia formation: implications for the aetiology of Meier-Gorlin syndrome." PLOS Genetics **9**(3): e1003360.
- Sun, J., C. Evrin, S. A. Samel, A. Fernandez-Cid, A. Riera, H. Kawakami, B. Stillman, C. Speck and H. Li (2013). "Cryo-EM structure of a helicase loading intermediate containing ORC-Cdc6-Cdt1-MCM2-7 bound to DNA." Nature Structural & Molecular Biology **20**(8): 944-951.
- Suski, J. M., N. Ratnayeke, M. Braun, T. Zhang, V. Strmiska, W. Michowski, G. Can, A. Simoneau, K. Snioch, M. Cup, C. M. Sullivan, X. Wu, J. Nowacka, T. B. Branigan, L. R. Pack, J. A. DeCaprio, Y. Geng, L. Zou, S. P. Gygi, J. C. Walter, T. Meyer and P. Sicinski (2022). "CDC7-independent G1/S transition revealed by targeted protein degradation." Nature **605**(7909): 357-365.
- Takahashi, T. S., A. Basu, V. Bermudez, J. Hurwitz and J. C. Walter (2008). "Cdc7-Drf1 kinase links chromosome cohesion to the initiation of DNA replication in *Xenopus* egg extracts." Genes & Development **22**(14): 1894-1905.
- Takahashi, T. S. and J. C. Walter (2005). "Cdc7-Drf1 is a developmentally regulated protein kinase required for the initiation of vertebrate DNA replication." Genes & Development **19**(19): 2295-2300.
- Takeda, D. Y., J. D. Parvin and A. Dutta (2005). "Degradation of Cdt1 during S phase is Skp2-independent and is required for efficient progression of mammalian cells through S phase." Journal of Biological Chemistry **280**(24): 23416-23423.
- Tanaka, S. and J. F. Diffley (2002). "Interdependent nuclear accumulation of budding yeast Cdt1 and Mcm2-7 during G1 phase." Nature Cell Biology **4**(3): 198-207.
- Tanaka, S., T. Umemori, K. Hirai, S. Muramatsu, Y. Kamimura and H. Araki (2007). "CDK-dependent phosphorylation of Sld2 and Sld3 initiates DNA replication in budding yeast." Nature **445**(7125): 328-332.
- Taylor, S. S. and A. P. Kornev (2011). "Protein kinases: evolution of dynamic regulatory proteins." Trends in Biochemical Sciences **36**(2): 65-77.

- Terwilliger, T. C., S. J. Ludtke, R. J. Read, P. D. Adams and P. V. Afonine (2020). "Improvement of cryo-EM maps by density modification." Nature Methods **17**(9): 923-927.
- Thomae, A. W., J. Baltin, D. Pich, M. J. Deutsch, M. Ravasz, K. Zeller, M. Gossen, W. Hammerschmidt and A. Schepers (2011). "Different roles of the human Orc6 protein in the replication initiation process." Cellular and Molecular Life Sciences **68**(22): 3741-3756.
- Ticau, S., L. J. Friedman, K. Champasa, I. R. Correa, Jr., J. Gelles and S. P. Bell (2017). "Mechanism and timing of Mcm2-7 ring closure during DNA replication origin licensing." Nature Structural & Molecular Biology **24**(3): 309-315.
- Ticau, S., L. J. Friedman, N. A. Ivica, J. Gelles and S. P. Bell (2015). "Single-molecule studies of origin licensing reveal mechanisms ensuring bidirectional helicase loading." Cell **161**(3): 513-525.
- Tocij, A., K. F. On, Z. Yuan, J. Sun, E. Elkayam, H. Li, B. Stillman and L. Joshua-Tor (2017). "Structure of the active form of human origin recognition complex and its ATPase motor module." eLife **6**.
- Tsuji, T., E. Lau, G. G. Chiang and W. Jiang (2008). "The Role of Dbf4/Drf1-Dependent Kinase Cdc7 in DNA-Damage Checkpoint Control." Molecular Cell **32**(6): 862-869.
- Tudzarova, S., P. Mulholland, A. Dey, K. Stoeber, A. L. Okorokov and G. H. Williams (2016). "p53 controls CDC7 levels to reinforce G1 cell cycle arrest upon genotoxic stress." Cell Cycle **15**(21): 2958-2972.
- van Deursen, F., S. Sengupta, G. De Piccoli, A. Sanchez-Diaz and K. Labib (2012). "Mcm10 associates with the loaded DNA helicase at replication origins and defines a novel step in its activation." EMBO Journal **31**(9): 2195-2206.
- Varadi, M., S. Anyango, M. Deshpande, S. Nair, C. Natassia, G. Yordanova, D. Yuan, O. Stroe, G. Wood, A. Laydon, A. Židek, T. Green, K. Tunyasuvunakool, S. Petersen, J. Jumper, E. Clancy, R. Green, A. Vora, M. Lutfi, M. Figurnov, A. Cowie, N. Hobbs, P. Kohli, G. Kleywegt, E. Birney, D. Hassabis and S. Velankar (2022). "AlphaFold Protein Structure Database: massively expanding the structural coverage of protein-sequence space with high-accuracy models." Nucleic Acids Research **50**(D1): D439-D444.
- Varrin, A. E., A. A. Prasad, R. P. Scholz, M. D. Ramer and B. P. Duncker (2005). "A mutation in Dbf4 motif M impairs interactions with DNA replication factors and confers increased resistance to genotoxic agents." Molecular and Cellular Biology **25**(17): 7494-7504.
- Vashee, S., P. Simancek, M. D. Challberg and T. J. Kelly (2001). "Assembly of the human origin recognition complex." Journal of Biological Chemistry **276**(28): 26666-26673.
- Wagner, T., F. Merino, M. Stabrin, T. Moriya, C. Antoni, A. Apelbaum, P. Hagel, O. Sitsel, T. Raisch, D. Prumbaum, D. Quentin, D. Roderer, S. Tacke, B. Siebolds, E. Schubert, T. R. Shaikh, P. Lill, C. Gatsogiannis and S. Raunser (2019). "SPHIRE-crYOLO is a fast and accurate fully automated particle picker for cryo-EM." Communications Biology **2**: 218.
- Watase, G., H. Takisawa and M. T. Kanemaki (2012). "Mcm10 plays a role in functioning of the eukaryotic replicative DNA helicase, Cdc45-Mcm-GINS." Current Biology **22**(4): 343-349.
- Wei, Z., C. Liu, X. Wu, N. Xu, B. Zhou, C. Liang and G. Zhu (2010). "Characterization and structure determination of the Cdt1 binding domain of human minichromosome maintenance (Mcm) 6." Journal of Biological Chemistry **285**(17): 12469-12473.

- Weinreich, M. and B. Stillman (1999). "Cdc7p-Dbf4p kinase binds to chromatin during S phase and is regulated by both the APC and the RAD53 checkpoint pathway." EMBO Journal **18**: 5334-5346.
- Willhoft, O. and A. Costa (2021). "A structural framework for DNA replication and transcription through chromatin." Current Opinion in Structural Biology **71**: 51-58.
- Wilmes, G. M., V. Archambault, R. J. Austin, M. D. Jacobson, S. P. Bell and F. R. Cross (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 & Development **18**(9): 981-991.
- Wilmes, G. M. and S. P. Bell (2002). "The B2 element of the *Saccharomyces cerevisiae* ARS1 origin of replication requires specific sequences to facilitate pre-RC formation." Proceedings of the National Academy of Sciences of the United States of America **99**(1): 101-106.
- Wohlschlegel, J. A., B. T. Dwyer, S. K. Dhar, C. Cvetcic, J. C. Walter and A. Dutta (2000). "Inhibition of eukaryotic DNA replication by geminin binding to Cdt1." Science **290**(5500): 2309-2312.
- Wybenga-Groot, L. E., C. S. Ho, F. D. Sweeney, D. F. Ceccarelli, C. J. McGlade, D. Durocher and F. Sicheri (2014). "Structural basis of Rad53 kinase activation by dimerization and activation segment exchange." Cellular Signalling **26**(9): 1825-1836.
- Xu, N., Q. Lin, H. Tian, C. Liu, P. Wang, C. M. Suen, H. Yang, Y. Xiang and G. Zhu (2022). "Cryo-EM structure of human hexameric MCM2-7 complex." iScience **25**(9): 104976.
- Yamashita, N., J. M. Kim, O. Koiwai, K. Arai and H. Masai (2005). "Functional analyses of mouse ASK, an activation subunit for Cdc7 kinase, using conditional ASK knockout ES cells." Genes to Cells **10**(6): 551-563.
- Yang, J. and Y. Zhang (2015). "I-TASSER server: new development for protein structure and function predictions." Nucleic Acids Research **43**(W1): W174-181.
- Yao, L., J. Chen, X. Wu, S. Jia and A. Meng (2017). "Zebrafish cdc6 hypomorphic mutation causes Meier-Gorlin syndrome-like phenotype." Human Molecular Genetics **26**(21): 4168-4180.
- Yeeles, J. T. P., T. D. Deegan, A. Janska, A. Early and J. F. X. Diffley (2015). "Regulated eukaryotic DNA replication origin firing with purified proteins." Nature **519**(7544): 431-435.
- Yeeles, J. T. P., A. Janska, A. Early and J. F. X. Diffley (2017). "How the Eukaryotic Replisome Achieves Rapid and Efficient DNA Replication." Molecular Cell **65**(1): 105-116.
- Yoshizawa-Sugata, N., A. Ishii, C. Taniyama, E. Matsui, K. Arai and H. Masai (2005). "A second human Dbf4/ASK-related protein, Drf1/ASKL1, is required for efficient progression of S and M phases." Journal of Biological Chemistry **280**(13): 13062-13070.
- Yuan, Z., L. Bai, J. Sun, R. Georgescu, J. Liu, M. E. O'Donnell and H. Li (2016). "Structure of the eukaryotic replicative CMG helicase suggests a pumpjack motion for translocation." Nature Structural & Molecular Biology **23**(3): 217-224.
- Yuan, Z., R. Georgescu, L. Bai, D. Zhang, H. Li and M. E. O'Donnell (2020a). "DNA unwinding mechanism of a eukaryotic replicative CMG helicase." Nature Communications **11**(1): 688.
- Yuan, Z., A. Riera, L. Bai, J. Sun, S. Nandi, C. Spanos, Z. A. Chen, M. Barbon, J. Rappsilber, B. Stillman, C. Speck and H. Li (2017). "Structural basis of Mcm2-7 replicative helicase loading by ORC-Cdc6 and Cdt1." Nature Structural & Molecular Biology **24**(3): 316-324.
- Yuan, Z., S. Schneider, T. Dodd, A. Riera, L. Bai, C. Yan, I. Magdalou, I. Ivanov, B. Stillman, H. Li and C. Speck (2020b). "Structural mechanism of helicase loading

- onto replication origin DNA by ORC-Cdc6." Proceedings of the National Academy of Sciences of the United States of America **117**(30): 17747-17756.
- Zachariae, W., M. Schwab, K. Nasmyth and W. Seufert (1998). "Control of cyclin ubiquitination by CDK-regulated binding of Hct1 to the anaphase promoting complex." Science **282**(5394): 1721-1724.
- Zegerman, P. and J. F. Diffley (2007). "Phosphorylation of Sld2 and Sld3 by cyclin-dependent kinases promotes DNA replication in budding yeast." Nature **445**(7125): 281-285.
- Zegerman, P. and J. F. Diffley (2010). "Checkpoint-dependent inhibition of DNA replication initiation by Sld3 and Dbf4 phosphorylation." Nature **467**(7314): 474-478.
- Zhai, Y., E. Cheng, H. Wu, N. Li, P. Y. K. Yung, N. Gao and B. K. Tye (2017). "Opening structure of the Cdt1-Mcm2-7 complex as a precursor of the MCM double hexamer." Nature Structural & Molecular Biology **24**(3): 300-308.
- Zhang, H. (2021) "Regulation of DNA Replication Licensing and Re-Replication by Cdt1." International Journal of Molecular Sciences **22** DOI: 10.3390/ijms22105195.
- Zhang, K. (2016). "Gctf: Real-time CTF determination and correction." Journal of Structural Biology **193**(1): 1-12.
- Zhao, X., E. G. Muller and R. Rothstein (1998). "A suppressor of two essential checkpoint genes identifies a novel protein that negatively affects dNTP pools." Molecular Cell **2**(3): 329-340.
- Zheng, S. Q., E. Palovcak, J. P. Armache, K. A. Verba, Y. Cheng and D. A. Agard (2017). "MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy." Nature Methods **14**(4): 331-332.
- Zivanov, J., T. Nakane, B. O. Forsberg, D. Kimanius, W. J. H. Hagen, E. Lindahl and S. H. W. Scheres (2018). "New tools for automated high-resolution cryo-EM structure determination in RELION-3." eLife **7**: e42166.
- Zivanov, J., T. Nakane and S. H. W. Scheres (2019). "A Bayesian approach to beam-induced motion correction in cryo-EM single-particle analysis." International Union of Crystallography Journal **6**(1): 5-17.
- Zivanov, J., T. Nakane and S. H. W. Scheres (2020). "Estimation of high-order aberrations and anisotropic magnification from cryo-EM data sets in RELION-3.1." International Union of Crystallography Journal **7**(2): 253-267.