Protein-protein interactions that mediate cell cycle events

Characterizing protein-protein interactions that regulate cell cycle processes.

By AHMAD WASSIM ALMAWI, B.M.Sc. (Honours Specialization)

A Thesis Submitted to the School of Graduate Studies in Partial Fulfilment of the Requirement for the Degree Doctor of Philosophy

DOCTOR OF PHILOSOPHY (2018) McMaster University

(Biochemistry and Biomedical Sciences) Hamilton, Ontario

TITLE: Characterizing protein-protein interactions that regulate cell cycle processes.

AUTHOR: Ahmad Wassim Almawi, B.M.Sc. (Western University)

SUPERVISOR: Dr. Alba Guarné

NUMBER OF PAGES: 180

TABLE OF CONTENTS

ABSTRACT

Molecular recognition is at the core of all biological processes whereby protein-protein interactions (PPI) relay messages to drive signaling events. However, many regulatory responses are driven by weak or transient PPI. These interactions are difficult to study using structural biology techniques because they are labile and result in heterogeneous populations. Moreover, interactions reconstituted using peptides are difficult to interpret because they lack context. In this thesis, I characterized key signaling complexes implicated in the replication checkpoint response (Dbf4-Rad53-Cdc7 complex), mitosis (Dbf4-Cdc5 complex), and DNA mismatch repair (clamp-MutL complex). I solved the crystal structures of the Dbf4-Rad53 and clamp-MutL weak complexes by generating fusions of the binding partners. The structures revealed that Dbf4 and MutL undergo subtle conformational movements upon engaging their binding partners, which were sufficient to alter both interfaces. Overall, the structures offer insight as to how Rad53 could inhibit Dbf4-Cdc7 during the replication checkpoint and how the clamp could activate MutL during mismatch repair. Acquiring the Dbf4-Cdc5 co-crystal structure required optimization the Dbf4 peptide. The Dbf4-Cdc5 and Dbf4-Rad53 complexes were relatable because both interactions were phosphorylation-independent even though Rad53 and Cdc5 are known to recognize phosphorylated targets. Dbf4 engaged a binding site on Cdc5 located opposite to the phosphoepitope binding pocket, which is reminiscent to its interaction with Rad53. Collectively, the structures of Dbf4 and its binding partners reveal that Rad53 and Cdc5 function beyond phosphoepitope recognition whereby they utilize additional binding surfaces to engage substrates.

ACKNOWLEDGEMENTS

I would like to thank my supervisor, Dr. Alba Guarné, for her guidance, leadership, and support. The projects often took us by surprise, and I will forever cherish those meetings where we exchanged crazy ideas, which more often than not, propelled the projects forward. It has been a privilege to train in your lab, and I am grateful for acquiring numerous transferable skills from shadowing and learning from you.

I would like to thank my committee members, Dr. Giuseppe Melacini, Dr. Bradley Doble, and Dr. Murray Junop for their invaluable advice. I would like to extend a special thank you to Dr. Melacini and Stephen Boulton for collaborating with us on multiple projects. Also, I am grateful to Dr. Bernard Duncker and Larasati for accommodating me into their laboratory for several weeks to learn yeast two-hybrid technique. I would like to thank past and present members of the Guarné lab for making my Ph.D. studies more manageable and presenting me with invaluable advice that pushed my projects forward. I would like to extend a special thank you to Dr. Lindsay Matthews for her guidance in the Dbf4-Rad53 project as well as Dr. Monica Pillon for her support in the ß clamp-MutL project. Also, I would like to express my gratitude to Yao Shen, Dr. Paula Balisi, and Jeremy Caron for their assistance in the Cdc5 project.

Finally, I would like to thank my family for supporting me throughout the years. Thank you, dad, for introducing me to research and the laboratory environment. A Ph.D. is a long and challenging road, but having the undying will and perseverance to get to your goal and fulfill your vision is what strengthens your character by the end of the journey.

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS AND SYMBOLS

ADP	Adenosine Diphosphate
APC	Anaphase Promoting Complex
APS	Advanced Photon Source
ATP	Adenosine Triphosphate
ß-clamp	Sliding ß-clamp
Bfa1	Mitotic checkpoint protein BFA1
BNL	Brookhaven National Laboratory
BRCA-1	Breast Cancer-1
BRCT	BRCA-1 C-Terminal
Bub	Budding Uninhibited by Benzimidazole
CaM	Calmodulin
CDC	Cell Division Cycle
Cdk	Cyclin-Dependent Kinase
Cdt1	Cdc10-Dependent Transcript 1
Chk	Checkpoint Kinase
Clb	Cyclin protein
Cln	Cyclin protein
CLS	Canadian Light Source
CMG	Cdc45-Mcm2-7-GINS
COOT	Crystallographic Object-Oriented Toolkit

CTD	C-Terminal Domain
Dbf4	Dumbbell Former 4
Ddc2	DNA Damage Checkpoint protein 2
DDK	Dbf4-Dependent Kinase
DNA	Deoxyribonucleic Acid
Dpb11	DNA replication regulator DPB11
dsDNA	Double Stranded Deoxyribonucleic Acid
EDTA	Ethylenediaminetetraacetic Acid
Esp1	Separin
Exo1	Exonuclease I
FHA	Forkhead Associated
ΔG	Gibbs Free Energy
G_1 phase	Gap 1
G_2 phase	Gap 2
GINS	Go-Ichi-Ni-San
HBRCT	Helix BRCA-1 C-Terminal
Ino80	Inositol-requiring protein 80
IPTG	Isopropyl ß-D-1-thiogalactopyranoside
K_D	Dissociation constant
Lte1	Low Temperature Essential protein 1
Mad	Mitotic Arrest Deficient
Map205	Microtubule-Associated Protein 205

M phase	Mitosis phase
MCM	Minichromosome maintenance
Mec1	Mitosis Entry Checkpoint protein 1
Mih1	M-phase Inducer phosphatase 1
MLH	MutL homolog
MMR	Mismatch Repair
Mob1	Maintenance of Ploidy protein 1
Motif C	Motif C-terminus
Motif M	Motif Middle
Motif N	Motif N-terminus
Mtmr	Matrimony protein
Mut	Mutagenic
Net1	Nucleolar protein 1
Ng	Neurogranin
Nm	Neuromodulin
NMR	Nuclear Magnetic Resonance
NSLS	National Synchrotron Light Source
OD_{600}	Optical Density (600 nm)
ORC	Origin Recognition Complex
PCNA	Proliferating Cell Nuclear Antigen
PCR	Polymerase Chain Reaction
PBD	Polo-box domain
PDB	Protein Data Bank

PEG	Polyethylene Glycol
PHENIX	Python-based Hierarchical Environment for Integrated Xtallography
PIP Box	PCNA-Interacting Protein
PISA	Protein Interfaces Structures Assemblies
Plk	Polo-Like Kinase
PMS2	Postmeiotic Segregation Increase 2
Pol	Polymerase
Pre-RC	Pre-Replication Complex
Rad	Radiation
RFC	Replication Factor C
RGD	Regulatory Domain
RMSD	Root Mean Square Deviation
RPA	Replication Protein A
RPS4	Disease resistance protein 4
RRS1	Disease resistance protein 1
S phase	Synthetic phase
SC	Synaptonemal Complex
SCD	Serine-Glutamine/Threonine-Glutamine Cluster Domain
SDS	Sodium Dodecyl Sulfate
SiRNA	Small Interfering Ribonucleic Acid
Sld	Synthetically Lethal with Dpb11-1
Slk19	Synthetic Lethal KAR3 protein 19
Spc72	Spindle Pole Component 72

Spo12	Sporulation-specific Protein 12
ssDNA	Single stranded Deoxyribonucleic Acid
SUMO	Small Ubiquitin-like Modifier
Swe1	Wee1 homolog
TEV	Tobacco Etch Virus
tRNA	Transfer Ribonucleic Acid
XDS	X-ray Detector Software

DECLARATION OF ACADEMIC ACHIEVEMENT

All experiments covered in this thesis were designed, conducted, and analyzed by Dr. Guarné and myself unless stated otherwise in the connecting text for each chapter. This includes the yeast two-hybrid experiments conducted in collaboration with Larasati from the laboratory of Dr. Duncker at the University of Waterloo (Chapter 2) and the Saturation Transfer Difference NMR experiments performed in conjunction with Stephen Boulton from the laboratory of Dr. Melacini at McMaster University (Chapter 3).

CHAPTER ONE

INTRODUCTION

1.1 THE CELL

In 1665, Robert Hooke published his book, Micrographia, whereby he reported on a multitude of tiny pores on a thin slice of cork under his microscope (Gest, 2004). These micro-structures were termed cells. Due to the limitation of microscope magnification at the time, Hooke could not discern the precise function of cells and what lies within them. Shortly after that in 1676, Anton van Leeuwenhoek used a more powerful microscope and discovered that cells are motile objects (Gest, 2004). In his letter to The Royal Society, van Leeuwenhoek argued that cells are living microorganisms since motility is a quality of life. Cells were thought to arise spontaneously; however, the works of van Leeuwenhoek on the process of fertilization between a sperm and egg cell terminated that idea (Mazzarello, 1999). The significant hallmarks of cell biology were achieved in the 1830s from the works of Theodor Schwann, Matthias Jakob Schleiden, and Rudolf Virchow, who postulated the three tenets of cell theory: 1) all living organisms originate from one or more cells, 2) the cell is the most basic unit of life, and 3) cells arise from pre-existing cells (Robinson, 2014).

The ability of life to propagate is a core concept in molecular biology; however, the process of cell division at the time was poorly understood (Gest, 2004; Mazzarello, 1999). The time-span required for a cell to progress through the entire cell cycle was established for various tissues by the 1950s. Nonetheless, it remained unclear as to how the cell spends its time at each stage of the cell cycle, and the stage at which the cell begins cycling (Bertalanffy and Lau, 1962; Leblond and Walker, 1956). It was not until 1963 when Defendi and Manson wrote: “the primum movens (the cause of) of cell division must be searched for, not at mitosis, but earlier in the cycle at the time of initiation of DNA synthesis” (Defendi and Manson, 1963). They were referring to

data acquired from healthy and cancerous mammalian cells showing that most eukaryotic cells require a similar period for DNA synthesis in preparation for mitosis (Defendi and Manson, 1963). This is in contrast to what Gelfant reported around the same period, that cells in adult mammalian epidermis replicate their DNA but remain quiescent indefinitely (Gelfant, 1962). Given these contradictory findings, Defendi and Manson deemed it necessary to organize the stages of the cell cycle, and thus subdivided them into early interphase, prosphase, dichophase (phase of DNA synthesis), and mitosis (Delfant and Manson, 1963).

1.2 THE CELL CYCLE

1.2.1 Stages of the cell cycle

Eukaryotic cells spend the majority of their time outside mitosis in a region known as interphase, which is subdivided into first gap phase (G_1), synthetic phase (S), and second gap phase (G_2) (**Figure 1.1**). During G_1-phase, the cell increases its production of proteins and organelles, and thus grows in size (Cooper, 2000). S-phase initiates once the cell commits to replicating its DNA. Chromosomes are replicated once S-phase is completed, resulting in doubling of the genomic content. The rates of transcription and translation are low during S-phase since most cellular resources are devoted to DNA replication (Cooper, 2000). G_2-phase signals protein synthesis and rapid growth in preparation for the final cell cycle stage, mitosis (**Figure 1.1**).

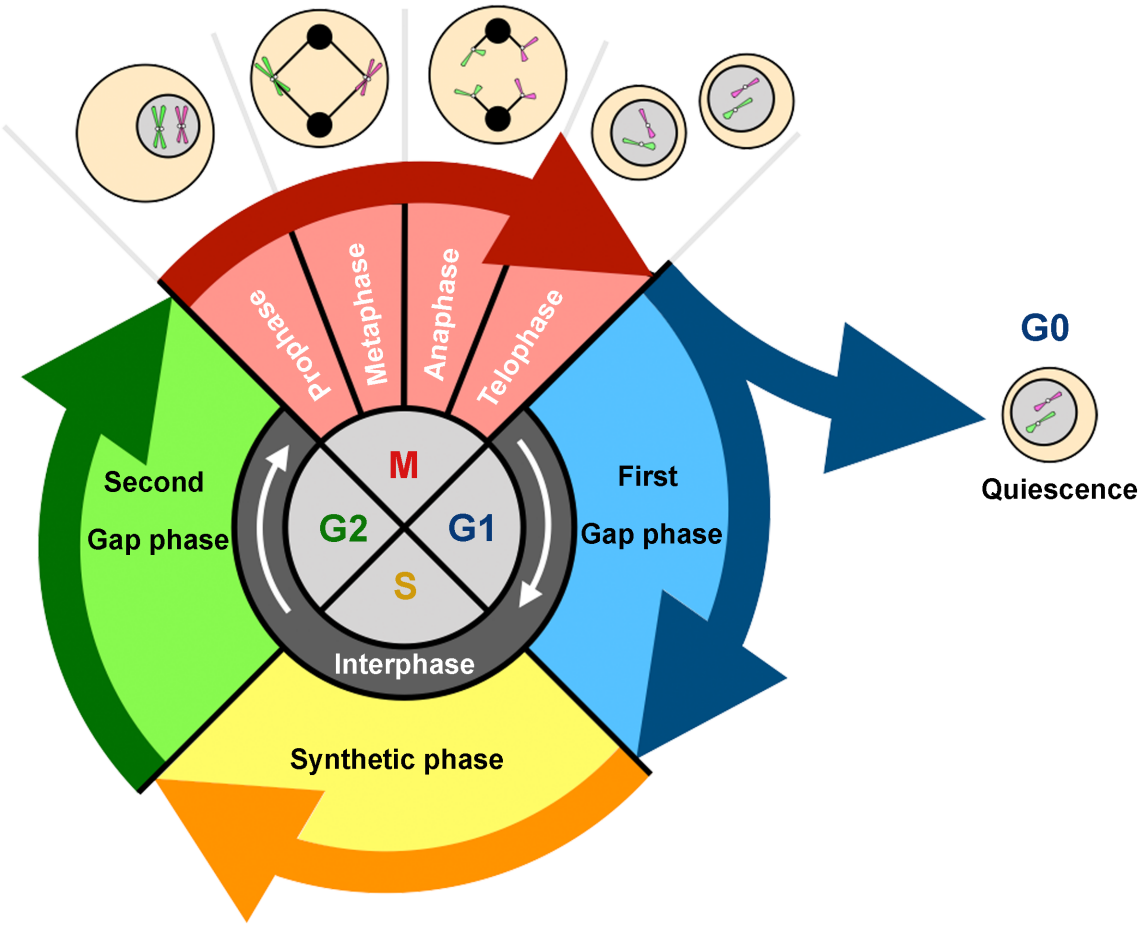

Figure 1.1. Overview of the cell cycle. The stages of the cell cycle: first gap phase (G_1), synthetic phase (S), second gap phase (G_2), mitosis (M), and quiescent state (G_0). The sub-stages in mitosis (prophase, metaphase, anaphase, and telophase) are presented with a pictorial diagram of a cell at each sub-stage. The cell (wheat) has the nucleus (grey) containing two chromosomes (green and pink). The spindle apparatus (black) is shown.

Mitosis is further sub-divided into prophase, metaphase, anaphase, and telophase (**Figure 1.1**). Duplicated DNA condenses only during prophase, which is the first stage of mitosis (**Figure 1.1**). Chromosomes align in the middle of the cell during metaphase, facilitated by the mitotic spindle machinery, and are then pulled apart during anaphase. Finally, the cell divides into two identical daughter cells during cytokinesis (**Figure 1.1**). Following division, cells can continue to grow, re-enter interphase, or arrest in a quiescent state known as G_0 (**Figure 1.1**) (Cooper, 2000).

Non-proliferative or fully-differentiated cells in multicellular organisms generally enter G_0 instead of G_1 and may remain quiescent indefinitely, as seen in neurons.

1.2.2 Protein kinases control the cell cycle

While a better understanding of the stages of the cell cycle was achieved during the mid-late 20th century, it remained unclear as to how the cell controls its different stages. The elegant works of Leland Hartwell, Paul Nurse, and Tim Hunt addressed this conundrum through their discovery of the cyclin proteins coupled to cyclin-dependent kinases (Cdks) (Hartwell, 1974). As their name implies, cyclin-dependent kinases bind cyclins thereby regulating cell cycle events (Enserink, 2011). Hartwell, Nurse, and Hunt shared the Nobel prize in Physiology or Medicine in 2001 because their discovery profoundly impacted our understanding of how cell cycle stages are temporally regulated.

Cdks were initially characterized from *Saccharomyces cerevisiae* (*S. cerevisiae*) and *Schizosaccharomyces pombe*. Using the information discovered from yeast cell cycle studies, significant progress has been made regarding the mammalian cell cycle. It has been determined that yeast and mammalian cell cycles are similar and that Cdks, either directly or indirectly, affect the progression of the cell cycle. *S. cerevisiae* harbors one Cdk (Cdk1), which associates with several cyclins to coordinate cell cycle events (Enserink, 2011) (**Figure 1.2A**). Higher eukaryotes contain at least nine Cdks (Enserink, 2011), four of which (Cdk1-4) are involved in cell cycle regulation (**Figure 1.2B**). Other Cdks, like Cdk7, function indirectly as Cdk-activating enzymes (Morgan, 2007).

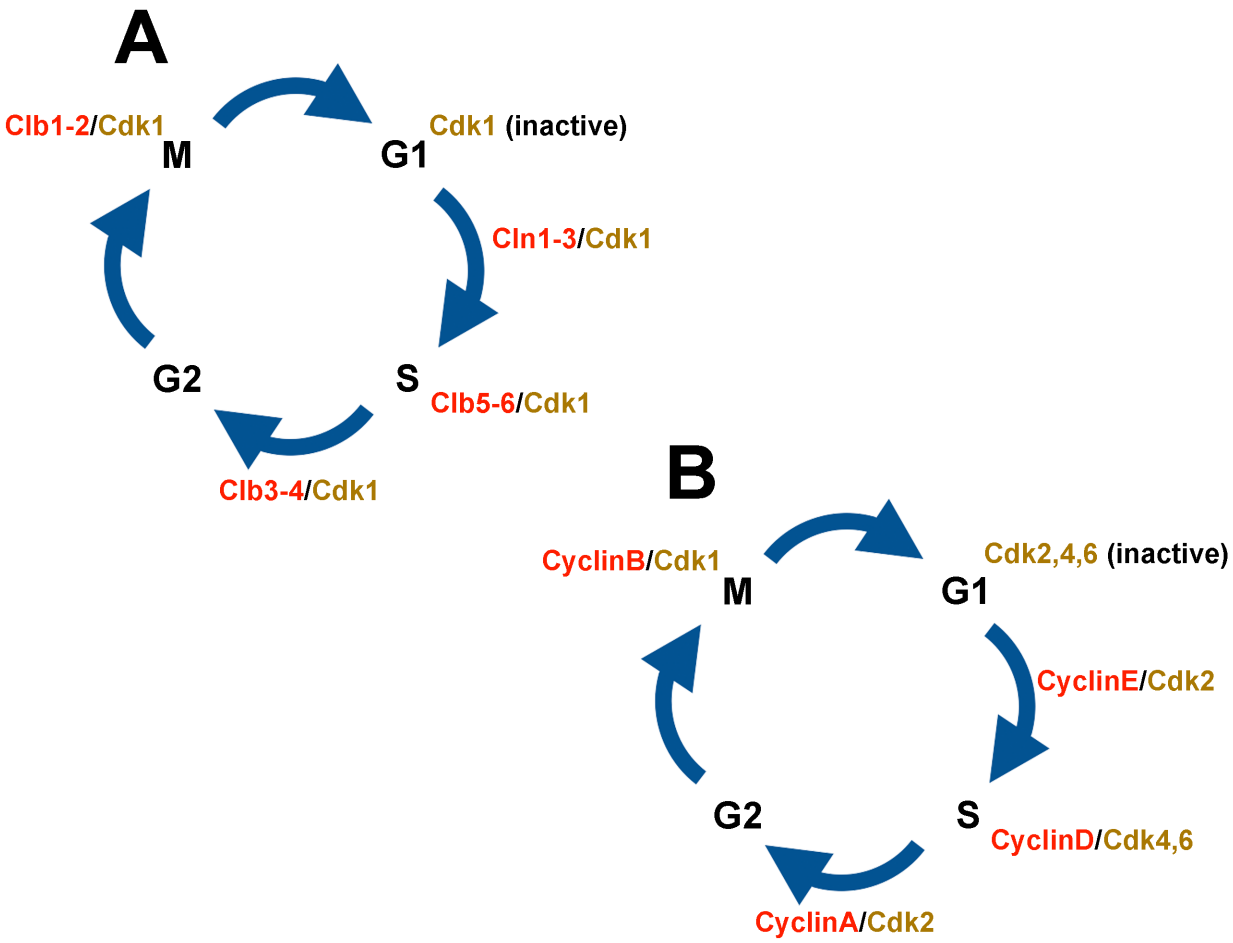

Figure 1.2. Cyclin/Cdks regulate cell cycle progression. (A) Cell cycle control by Cyclin/Cdk proteins in budding yeast and **(B)** in higher eukaryotes. Cyclin (red) and Cdk (brown) complexes are denoted as CyclinX/CdkX; where X is the letter or number associated for each protein.

In humans, four Cdks (Cdk1-4) and five cyclins (Cyclin A-E) associate to coordinate cell cycle events (Enserink, 2011). During G_1, Cyclins C and D interact with Cdk3 and Cdk4, respectively, to prepare the cell for DNA replication (Enserink, 2011). At the G_1-S-phase transition, CyclinE/Cdk2 promote entry into S-phase for initiation of DNA synthesis (Enserink, 2011) (**Figure 1.2B**). CyclinA/Cdk2 coordinate events leading to G_2 to prepare the cell for mitosis (**Figure 1.2B**). Finally, CyclinB/Cdk1 mediate mitotic functions needed to ensure proper cell division (Enserink, 2011) (**Figure 1.2B**). Cdk protein levels remain relatively constant throughout the cell cycle and most of their regulation is mediated through post-translational

modification (Morgan, 1995). In contrast, cyclin levels fluctuate at the various stages of the cell cycle and their regulation is controlled through both synthesis and targeted degradation (Morgan, 1995). In humans, the anaphase-promoting complex (APC) is responsible for mediating cyclin destruction (Morgan, 1995).

Since the discovery of Cdks, other families of protein kinases involved in cell cycle regulation were identified. Protein kinases were shown to not only promote cell cycle progression but also to halt the cell cycle through signaling cascades termed checkpoints (Enserink, 2011). As their name implies, checkpoints "check" that each stage of the cell cycle has faithfully passed before entering the next stage. Throughout interphase, checkpoints monitor cell growth and determine the availability of resources cells require before DNA replication. Once the cycle passes through S-phase, the checkpoint ensures that the duplicated genome is intact before cell division (Enserink, 2011). Checkpoints are also activated in response to exogenous or endogenous stimuli or stress, most notably DNA damage during or following replication (Weinert 1989). Furthermore, mitosis has several checkpoint pathways, which ensure that the mitotic spindle apparatus is appropriately attached to chromosomes (Weinert 1989). Failure of checkpoints to regulate replication- or mitotic-related stress promotes proliferation and genomic instability, which are major hallmarks of carcinogenesis (Hanahan and Weinberg, 2011).

1.3 DNA REPLICATION

Cellular proliferation requires accurate duplication of the genetic content of the parental cell so that the resulting daughter cells each have an identical copy. Cells replicate their genome throughout S-phase, however, spend considerable time during late G_1-phase to prime their genome for replication initiation.

1.3.1 Initiation

Initiation of DNA replication occurs at sites in the genome termed origins of replication. The number of origins varies among different organisms. While budding yeast contains around 400 origins of replication, the human genome includes tens of thousands of such origins (Mechali, 2010). The presence of multiple origins enables eukaryotic cells to promptly copy their entire genome. However, for replication to commence, origins must be modified through a sequence of events involving assembly, licensing, and activation of the pre-replication complex (pre-RC) (**Figure 1.3**).

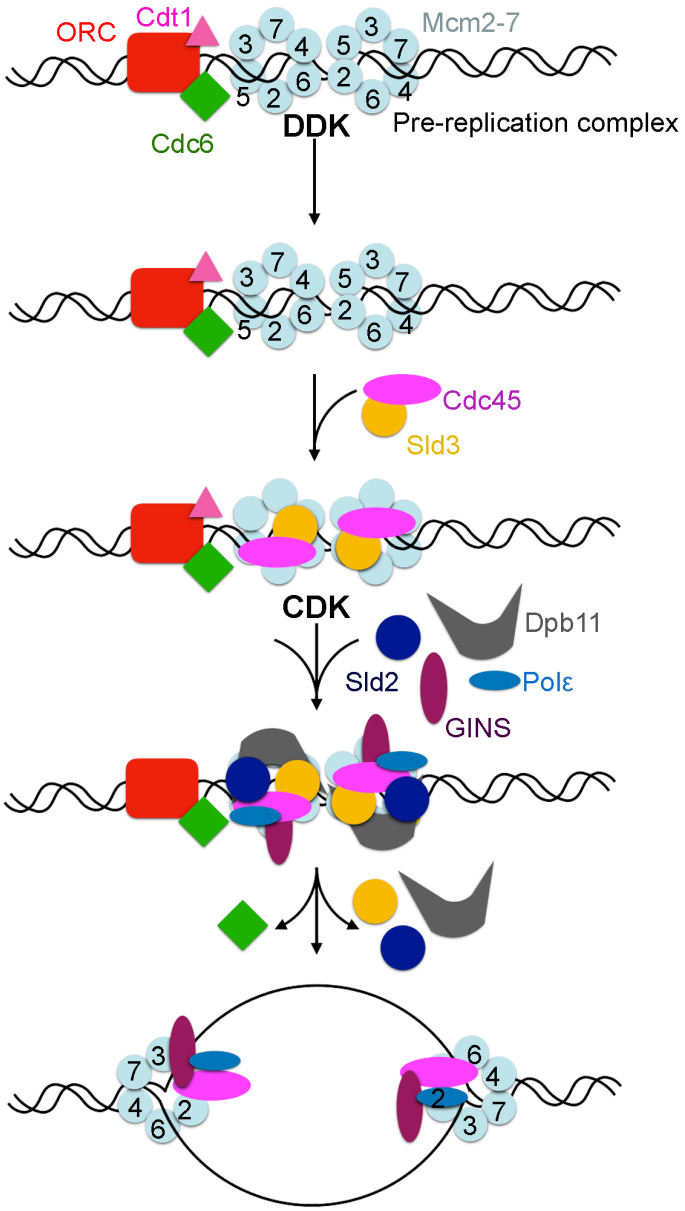

Figure 1.3. Model of DNA replication initiation. DNA replication is initiated by the assembly of the pre-replication complex during G_1-phase. Dbf4-dependent kinase (DDK) and cyclin-dependent kinase (CDK) initiate a series of phosphorylation events to generate the active form of the Cdc45-Mcm2-7-GINS helicase.

Assembly of the pre-RC begins with the binding of the initiator protein, origin recognition complex (ORC), to the origins of replication (**Figure 1.3**). ORC constitutively binds to origins of replication but is inactive for the majority of the cell cycle (Fragkos et al., 2015). ORC activation requires the recruitment of the replication factor, Cdc6 (**Figure 1.3**) (Fragkos et

al., 2015; Speck et al., 2005). Cdc6 expression levels are controlled throughout the cell cycle so that Cdc6 is only available in G_1 (Fragkos et al., 2015; Speck et al., 2005). ORC activation is a prerequisite for the assembly of the Mini-chromosome maintenance (Mcm) proteins, which constitute the core components of DNA helicase (Bowers et al., 2004). Mcm consists of six homologous subunits, denoted Mcm2, 3, 4, 5, 6, and 7, which assemble into a hetero-hexamer (**Figure 1.3**).

Loading of Mcm2-7 onto the pre-RC requires a third replication factor, Cdt1 (**Figure 1.3**) (Chen et al., 2007; Fragkos et al., 2015). Similar to Cdc6, Cdt1 expression levels fluctuate so that the cell controls its assembly with ORC-Cdc6 during G_1 (Feng and Kipreos, 2003). Association of Cdt1 and Cdc6 with ORC at the pre-RC in G_1-phase functions as a safeguard mechanism, aimed at preventing the cell from re-activating replication origins during S-phase (Feng and Kipreos, 2003). Cdt1-mediated loading of Mcm2-7 at the pre-RC involves two Mcm2-7 hexamers assembling in a head-to-head manner as a double hexamer (**Figure 1.3**) (Li et al., 2015; Nugochi et al., 2017; Remus et al., 2009). The association of ORC-Cdc6-Cdt1 with the Mcm2-7 double hexamer represents a licensed but inactive pre-RC (**Figure 1.3**).

Activation of the pre-RC depends on the activity of cyclin-dependent kinase (CDK) and Dbf4-dependent kinase (DDK) (**Figure 1.3**). CDK is a heterodimer composed of the regulatory subunit S-phase cyclin and the catalytic subunit Cdk1, while DDK consists of the regulatory subunit Dbf4 and the catalytic subunit Cdc7 kinase (Labib and De Piccoli, 2011). CDK and DDK recruit accessory factors, Cdc45 and the tetrameric GINS complex, which in turn associate with Mcm2-7 (**Figure 1.3**) (Larasati and Duncker, 2016; Vijayraghavan and Schwacha, 2012). Cdc45

is recruited first in a DDK-dependent manner, while the recruitment of GINS is controlled by CDK (Fragkos et al., 2015; Larasati and Duncker, 2016; Randell et al., 2010; Sheu and Stillman, 2006). DDK phosphorylates Mcm4 and Mcm6 to induce a conformational change throughout the hexamer propagating to Mcm5, which is then primed to interact with Cdc45 (Hoang et al., 2007; Sheu and Stillman, 2010; Larasati and Duncker, 2016).

CDK recruits the GINS complex by phosphorylating the replication factors, Sld2 and Sld3 (Muramatsu et al., 2010; Tanaka et al., 2007; Zegerman and Diffley, 2007). Cdc45 typically associates with Sld3 at the origins of replication whereas Sld2 binds to GINS coupled with the Polε replicative polymerase (**Figure 1.3**) (Kamimura et al., 2001; Tanaka et al., 2007). Recruitment of Sld2 and Sld3 to the replication complex is further dependent on the scaffold protein, Dpb11, which recognizes their phosphorylated forms (**Figure 1.3**) (Fragkos et al., 2015; Muramatsu et al., 2010; Zegerman and Diffley, 2007). Association of Cdc45 and GINS with Mcm2-7 stimulates Mcm2-7 ATPase activity thereby enhancing its interaction with DNA (Ilves et al., 2010; Li et al., 2015). The Cdc45-Mcm2-7-GINS (CMG) complex represents the functional form of DNA helicase capable of unwinding duplex DNA (**Figure 1.3**).

Recent cryo-electron microscopy (EM) structures of *S. cerevisiae* CMG and Mcm2-7 double hexamer bound to DNA reveal an unexpected process for the formation of bidirectional replication forks at origins of replication (Li and O'Donnell, 2018; Noguchi et al., 2017; Sun et al., 2015). Cdc45 and GINS associate together and interact with Mcm2 and Mcm5 respectively, thus sparing the Mcm2-7 central channel for DNA binding (**Figure 1.4A)** (Li et al., 2015). When

the Mcm2-7 double hexamer binds DNA, it threads the duplex through its central channel (**Figure 1.4B**) (Noguchi et al., 2017).

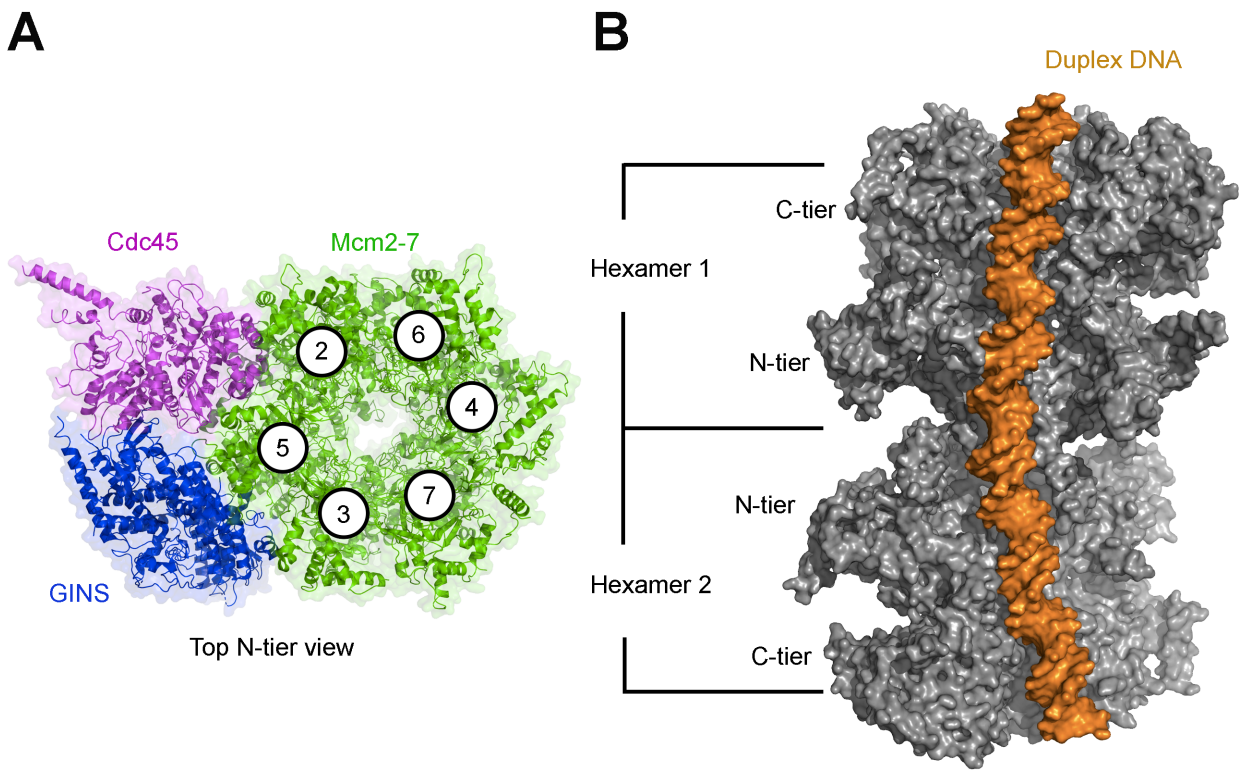

Figure 1.4. EM structures of CMG and Mcm2-7 double hexamer. (A) Surface and ribbon diagrams of CMG (PDB ID: 3JC5), showing Mcm2-7 (green) ring, and the Cdc45 (magenta) and GINS (blue) accessory factors. The circled numbers indicate Mcm subunits. **(B)** Cut-open view of the Mcm2-7 double hexamer (grey) bound to duplex DNA (orange) (PDB ID: 5BK4). Mcm2-7 and DNA are shown as surface diagrams. Each hexamer as well as their N- and C-tiers are labeled. Adapted from (Li and O'Donnell, 2018; Noguchi et al., 2017).

DNA zigzags inside the central channel through extensive interactions with Mcm2-7 (Noguchi et al., 2017). The N-tier ring of each Mcm2-7 hexamer tilts and shifts laterally to convert the double hexamer into the active form found in CMG (Noguchi et al., 2017). Moreover, EM studies show that CMG associates with DNA polymerases and represents the organizing center of the replisome (Langston et al., 2014; Li et al., 2015; Simon et al., 2014; Sun et al., 2015).

Collectively, structural biology studies reveal the molecular details involving the multi-step process of CMG assembly and activation prior to DNA synthesis (Li and O'Donnell, 2018).

1.3.2 Elongation

The replication machinery, termed the replisome, consists of CMG coupled to the replicative polymerases, Polε and Polδ (Li and O'Donnell, 2018, Yao and O'Donnell, 2010). Replication factor C (RFC) is the clamp loading protein responsible for tethering CMG to the replicative polymerases (**Figure 1.5**) (Hedglin et al., 2013).

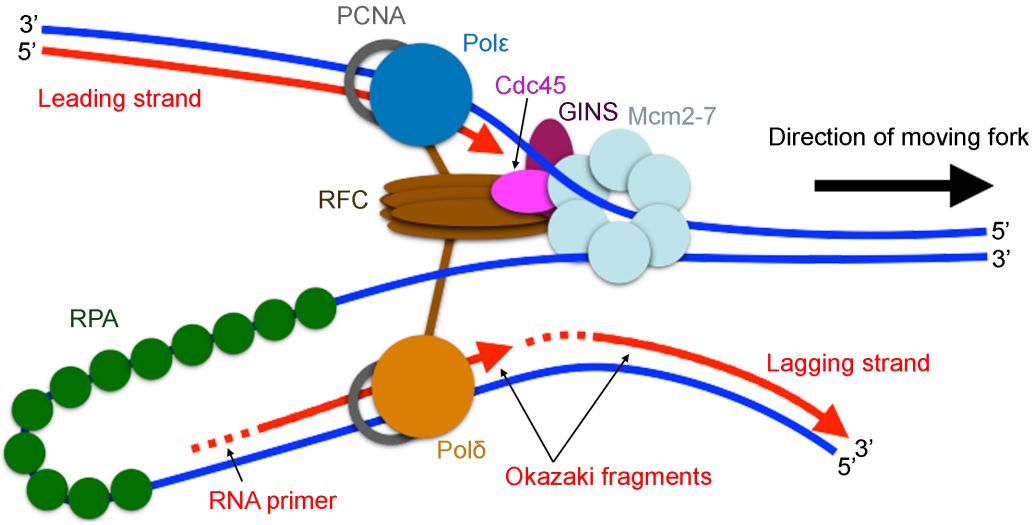

Figure 1.5. Model of a moving replication fork. Duplex DNA is unwound by CMG, and RPA covers ssDNA. RFC binds the moving CMG complex, and bridges the replicative polymerases. PCNA tethers the polymerase to the template DNA to improve polymerase processivity. On the lagging strand, the polymerase extends the RNA primers to synthesize Okazaki fragments.

However, recent structural studies demonstrated that CMG forms a direct complex with Polε, which is referred to as CMGE, independently of RFC (Langston et al., 2014; Sun et al., 2015). CMG can also individually bind the Ctf4 replication factor, which crosslinks CMGE to Polα-primase (Simon et al., 2014; Sun et al., 2015).

The replication fork forms when CMG unwinds duplex DNA, where the hydrogen bonds between the two strands break. DNA unwinding results in the formation of a structure with two

branching prongs comprised of single-stranded (ss) DNA (**Figure 1.5**). As CMG unwinds DNA, stretches of ssDNA become exposed, which can fold back on itself forming secondary structures (Alberts et al., 2002). These structures interfere with the movement of DNA polymerase and hence block it. The ss-binding protein, replication protein A (RPA), coats ssDNA until DNA polymerase completes synthesis of the second strand, thus preventing the formation of secondary structures (**Figure 1.5**) (Alberts et al., 2002). Moreover, Primase uses ssDNA to lay RNA primers, which are needed by DNA polymerases for extending the daughter strands (**Figure 1.5**). The leading strand receives one RNA primer while the lagging strand receives several primers. From each primer, the polymerase continuously extends the leading strand but discontinuously the lagging strand, resulting in the formation of Okazaki fragments (**Figure 1.5**).

Multiple DNA polymerases take on different roles during the elongation process. Polε controls leading strand synthesis while Polδ is responsible for lagging strand synthesis (**Figure 1.5**) (Burgers et al., 2016; Jinks-Robertson and Klein, 2015; Kunkel and Burgers, 2008). This view was recently challenged, and a role for Polδ in leading strand synthesis was suggested (Johnson et al., 2015; Stillman, 2015). Genetic studies demonstrated that Polδ replicates both leading and lagging strands, while Polε's proofreading activity is vital for removing Polδ-generated errors from the leading strand (Burgers et al., 2016; Johnson et al., 2015).

DNA polymerases are processive because they associate with the proliferating cell nuclear antigen (PCNA) protein, often referred to as the sliding clamp (**Figure 1.5**) (Hedglin et al., 2013). RFC loads PCNA onto DNA by recognizing the junction between the DNA template and RNA primers (Choi et al., 2004; Hedglin et al., 2013). Structurally, PCNA is ring-shaped, and thus threads DNA through its central cavity (Gulbis et al., 1996; McNally et al., 2010).

PCNA increases the processivity of DNA polymerase by creating a topological link with DNA thereby preventing the polymerase from dissociating.

Following DNA strand synthesis, RNase removes the RNA primers, and the low processivity DNA polymerase fills the gaps. Once completed, a single nick on the leading strand and several nicks on the lagging strand are generated. DNA ligase fills these nicks, thus marking the completion of the newly replicated DNA molecule.

1.3.3 The replication checkpoint response

DNA at the replication fork is prone to breakages because it is exposed. The potential danger of DNA damage associated with replication forks is correlated with the frequency at which they stall (Labib and De Piccoli, 2011; Wogan et al., 2004). These events often lead to the loss of genetic material and gross chromosomal rearrangements (Hanahan and Weinberg, 2011; Labib and De Piccoli, 2011; Wogan et al., 2004). Replication stress can also originate from oncogene over-expression, which in turn drives a high replication rate in cancer cells (Bartek et al., 2007). Over-stimulation of oncogenes starves the cell of needed resources thereby limiting their ability to repair other stalled forks (Bartek et al., 2007).

Cells have evolved defense mechanisms to respond to damaged replication forks in the form of the intra-S-phase replication checkpoint response. Upon encountering a DNA lesion, the replisome stalls, while CMG may continue unwinding duplex DNA (**Figure 1.6**) (Branzei and Foiani 2010; Mazouzi et al., 2014).

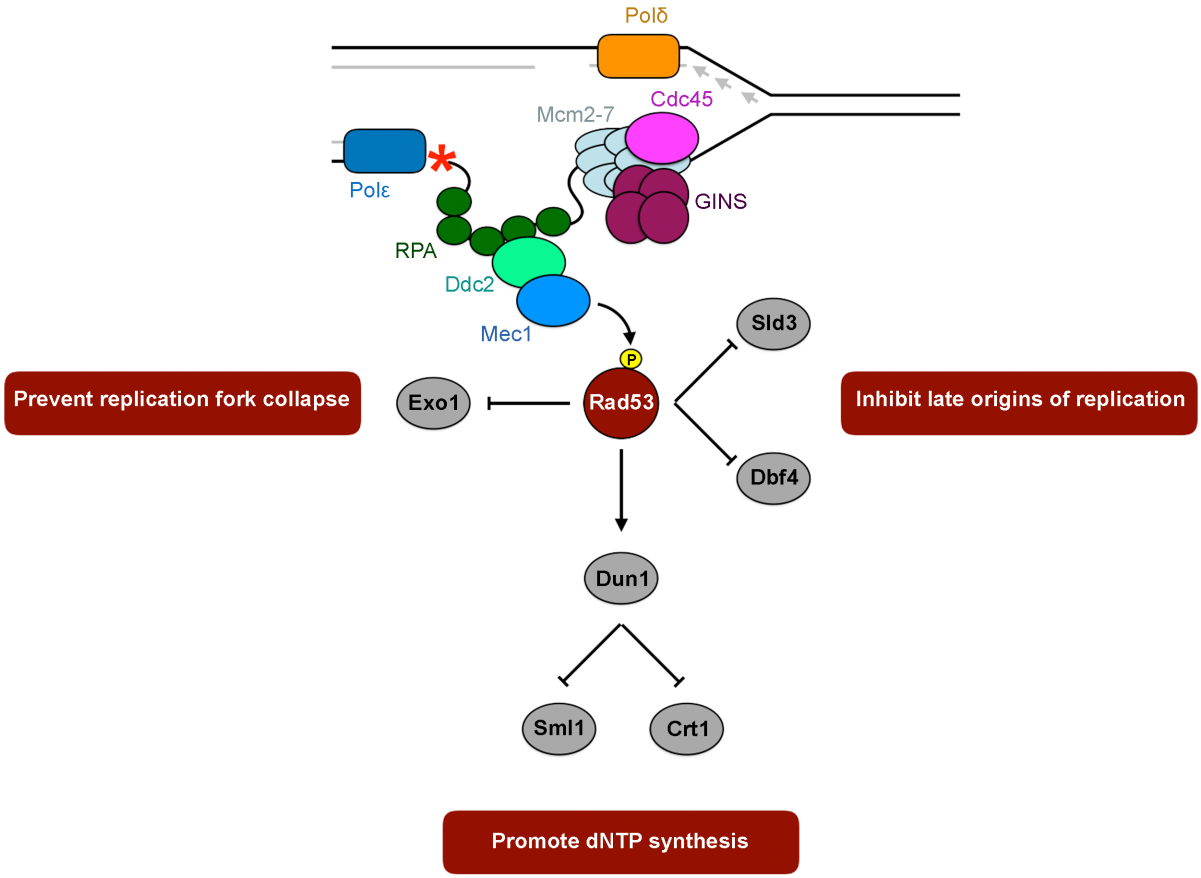

Figure 1.6. Model of the DNA replication checkpoint response. When DNA replication fork stalls at a lesion (red star), the CMG helicase may continue unwinding DNA. RPA coats ssDNA, and recruits the Ddc2-Mec1 complex. Mec1 phosphorylates (yellow P) and activates Rad53 kinase. Rad53 inhibits firing of late origins by phosphorylating Sld3 and Dbf4. Rad53 prevents replication fork collapse by inhibiting Exo1. Rad53 promotes dNTP synthesis by activating Dun1, which phosphorylates and inhibits Sml1 and Crt1 to relieve inhibition of ribonucleotide reductase.

As CMG unwinds DNA, RPA-coated ssDNA accumulates and serves as a signal recognized by the checkpoint proteins, Ddc2-Mec1. (**Figure 1.6**) (Friedel et al., 2009; Mazouzi et al., 2014; Zou and Elledge 2003). Mec1 is the transducer kinase of the replication checkpoint and acts by phosphorylating and activating the effector kinase, Rad53 (**Figure 1.6**) (Branzei and Foiani 2010; Friedel et al., 2009; Mazouzi et al., 2014; Pellicioli and Foiani, 2005). Besides Mec1, DDK phosphorylates and activates Rad53, thus implicating it in the replication checkpoint response (Larasati and Duncker, 2016; Ogi et al., 2008). DDK remains active at stalled forks and

plays a vital role in their recovery (Ogi et al., 2008). For example, DDK phosphorylates Mcm 2 thereby preventing CMG collapse at damaged forks (Bruck et al., 2015; Larasati and Duncker, 2016; Stead et al., 2012).

The three primary functions of Rad53 are to prevent replication fork collapse, promote free nucleotide (dNTP) expression to repair forks, and to inhibit S-phase (**Figure 1.6**). Rad53 stabilizes the replication fork by phosphorylating Exonuclease 1 (Exo1) (Hong et al., 2014; Mazouzi et al., 2014; Segurado and Diffley, 2008). While the details of the Rad53-Exo1 interaction are poorly understood, it was shown that phosphorylated Exo1 fails to cleave exposed DNA (Segurado and Diffley, 2008). Rad53 functions in repairing damaged forks by promoting the transcription of proteins which regulate nucleotide metabolism (Hong et al., 2014). For example, Rad53 phosphorylates Dun1, which in turn inactivates Sml1, an inhibitor of ribonucleotide reductase (**Figure 1.6**) (Zhao and Rothstein, 2002). Dun1 also inhibits Crt1, a transcriptional repressor of genes encoding ribonucleotide reductase (**Figure 1.6**) (Branzei and Foiani, 2009; Jaehnig et al., 2009; Zhao and Rothstein, 2002.)

Rad53 inhibits S-phase by preventing the activation of late licensed origins of replication (Mazouzi et al., 2014; Zegerman and Diffley, 2010). Both DDK and CDK are targeted by Rad53 during the replication checkpoint response (Maas et al. 2010; Mazouzi et al., 2014; Zegerman and Diffley, 2010). Rad53 inhibits firing of late origins by phosphorylating Sld3 and Dbf4 **(Figure 1.6)** (Duch et al., 2011; Kihara et al., 2000; Larasati and Duncker, 2016; Lopez-Mosqueda et al., 2010; Zegerman and Diffley, 2010). Phosphorylated Sld3 cannot bind Dpb11, and thus, the GINS complex is no longer recruited to Mcm2-7 at late licensed origins (Lopez-Mosqueda et al., 2010; Zegerman and Diffley, 2010). Similar to Cdc6 and Cdt1, Dbf4 and Sld3

protein levels fluctuate throughout the cell cycle (Mantiero et al., 2011). Dbf4 and Sld3 are expressed during G_1-phase, but their levels decline throughout S-phase. This ensures that only a limited number of late origins fire at any time, which in turn prevents several replication forks from activating at once. The temporal regulation of Dbf4 and Sld3 protein levels makes them strategic targets for Rad53 in inhibiting late licensed origins.

Rad53 has a central kinase domain flanked by two Forkhead-associated (FHA) domains, each preceded by a serine-glutamine/threonine-glutamine cluster domain (SCD) (Pellicioli and Foiani, 2005). Rad53 activity is controlled by upstream kinases, which act by phosphorylating serine and threonine residues within the SCD (Lee et al., 2003). However, the Rad53 function extends beyond the replication checkpoint response, as its FHA2 domain is implicated in the DNA damage checkpoint response (Pike et al., 2004). The primary roles of its FHA domains are to mediate protein-protein interactions by recognizing phosphorylated epitopes in binding substrates (Schwartz et al., 2003). FHA-mediated interactions enable Rad53 to localize to specific sites within the cell. For example, the FHA1 domain targets Rad53 to stalled replication forks (Chen and Zhou, 2009). The Rad53 homolog in higher eukaryotes is the effector kinase, Chk2, which harbors a single FHA domain (Hong et al., 2014). A second effector kinase, Chk1, also responsive to DNA damage and lacks an FHA domain (Chen and Sanchez, 2004; Hong et al., 2014). Collectively, this highlights that Rad53 and its homologs function similarly within the cell, but through distinct interaction mechanisms.

The FHA domain folds as a ß-sandwich defining a conserved phosphoepitope binding pocket (**Figure 1.7A**) (Durocher et al., 2000; Durocher and Jackson, 2002). The pocket is

composed mainly of positively-charged residues capable of binding to phosphorylated targets (**Figure 1.7B**).

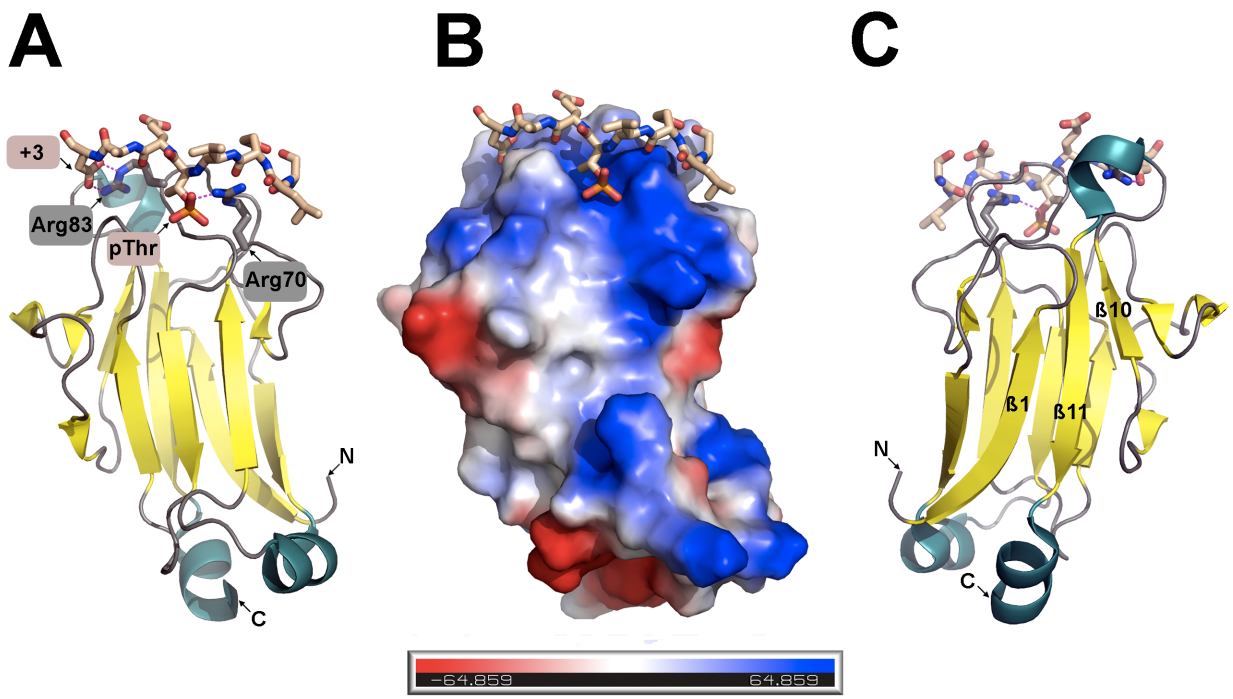

Figure 1.7. The FHA1 domain of Rad53. (A) Ribbon diagram of the FHA1 domain of Rad53 (PDB ID: 1G6G). Rad53 is colored according to secondary structure: α-helices (teal), ß-strands (yellow), and random coil (grey). The phosphorylated peptide (wheat) is shown as colour-coded sticks, with the phosphothreonine (pThr) and the pThr +3 position (+3) labeled. Arg70 and Arg83 (shown as sticks) interact with the pThr and the pThr +3 respectively. **(B)** Surface diagram of FHA1 presented in the same orientation as panel A, and colored according to electrostatic potential. The scale is given in units of k_bT/e_c, where k_b is the Boltzmann's constant, T is the temperature in Kelvins and e_c is the electron charge. **(C)** FHA1 presented similarly as in panel A but rotated 180 ˚ along the y-axis.

The FHA1 domain of Rad53 recognizes the consensus sequence, pThr-X-X-(Asp/Glu), in binding substrates. Conserved arginine residues, located on loops at the phosphoepitope binding pocket, mediate the interactions with the phosphorylated target. In FHA1, Arg70 stabilizes the phosphothreonine while Arg83 recognizes the third residue C-terminal to phosphothreonine (**Figure 1.7A**) (Durocher et al., 2000). Dbf4 is a binding partner of the Rad53 FHA1 domain,

and this interaction depends on the conserved Arg70 residue in FHA1 (Duncker et al., 2002). However, the phosphoepitope recognized by FHA1 in DDK lies within Cdc7 kinase (Aucher et al., 2010). It should be noted that the interaction between Dbf4 and FHA1 is phosphorylation-independent and is mediated through a ß-sheet surface (includes ß1, ß10, and ß11) opposite to the conserved phosphoepitope binding pocket (**Figure 1.7C**) (Matthews et al., 2014).

1.4 CELL DIVISION

The cell is ready to divide after it replicates its DNA throughout S-phase. The chromosomes condense in prophase to form chromatids, which are then aligned at the center of the spindle to facilitate genetic exchange during metaphase. The sister chromatids are then physically pulled apart into the daughter cells during the last stage of mitosis, anaphase. Mitosis is followed by cytokinesis whereby the cytoplasm, organelles, and cell membrane divide into two new cells containing roughly equal shares of cellular components. Transitioning between each stage in mitosis leading up to cytokinesis depends on the interplay between Cdks and ubiquitin-mediated proteolysis of protein targets (Kraft, 2003). The APC protein is the ubiquitin ligase that regulates mitotic progression by targeting factors for proteasomal degradation (Peters, 1998).

1.4.1 Mitotic entry

Mitotic entry commences when the cell transitions from G_2-phase to prophase in mitosis. The G_2-M phase transition is regulated by a counteracting kinase/phosphatase mechanism (Kraft, 2003). Mitotic entry ultimately requires the activation of the Cdk1-Clb2 complex which activates substrates required for chromosome condensation during prophase (Kraft, 2003). In late G_2-

phase, the Swe1 kinase inhibits Cdk1 thereby preventing it from interacting with Clb2 (**Figure 1.8**) (Booher et al. 1993; Hu et al. 2008).

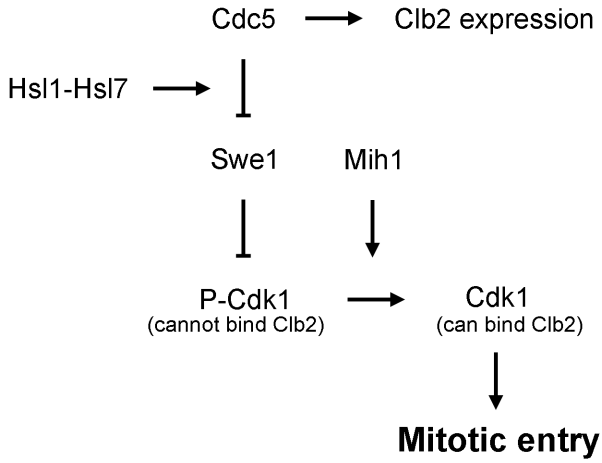

Figure 1.8. Mitotic entry depends on the activation of Cdk1-Clb2. Protein-protein interactions involved in mitotic entry.

The Hsl1-Hsl7 kinase complex recruits Cdc5 kinase during late G_2-phase to counteract Swe1 (Asano et al., 2005; Sakchaisri et al., 2004). Hsl1-Hsl7 and Cdc5 phosphorylate Swe1 thereby targeting it for APC-mediated proteasomal degradation (Asano et al., 2005). The Mih1 phosphatase dephosphorylates Cdk1, which can then interact with Clb2 to drive mitosis (**Figure 1.8**) (Russell et al., 1989; Sia et al., 1996). Nuclear exclusion of Cdc5 is linked to mitotic delay, implying that Cdc5 acts within the nucleus during mitotic entry (Botchkarev et al., 2014; Nakashima et al., 2008). Upon nuclear import, Cdc5 phosphorylates transcription factors that promote expression of Clb2 (**Figure 1.8**) (Darieva et al., 2006). Once mitotic entry is complete, the Cdk1-Clb2 complex activates substrates required for spindle formation in preparation for metaphase (Kraft, 2003).

1.4.2 The spindle assembly checkpoint response

The cell is at risk as it enters metaphase because errors related to spindle pole alignment frequently occur (Mantikou et al., 2012). Anaphase lag occurs when the movement of one chromatid is impeded due to its failure to attach to the mitotic spindle during metaphase. This process gives rise to aneuploid cells that have too few or too many of one or more chromosomes, a condition strongly associated with cancer (Draviam et al., 2004; Santaguida et al., 2015). Incorrect attachment of chromatids onto the spindle is attributed to the inability of microtubules to bind to the kinetochore, which is a proteinaceous structure assembled at the centromere (Cheeseman, 2014). The spindle assembly checkpoint (SAC) is responsible for inhibiting mitosis when kinetochores fail to properly attach to the spindle during metaphase (**Figure 1.9**) (Lara-Gonzalez et al., 2012).

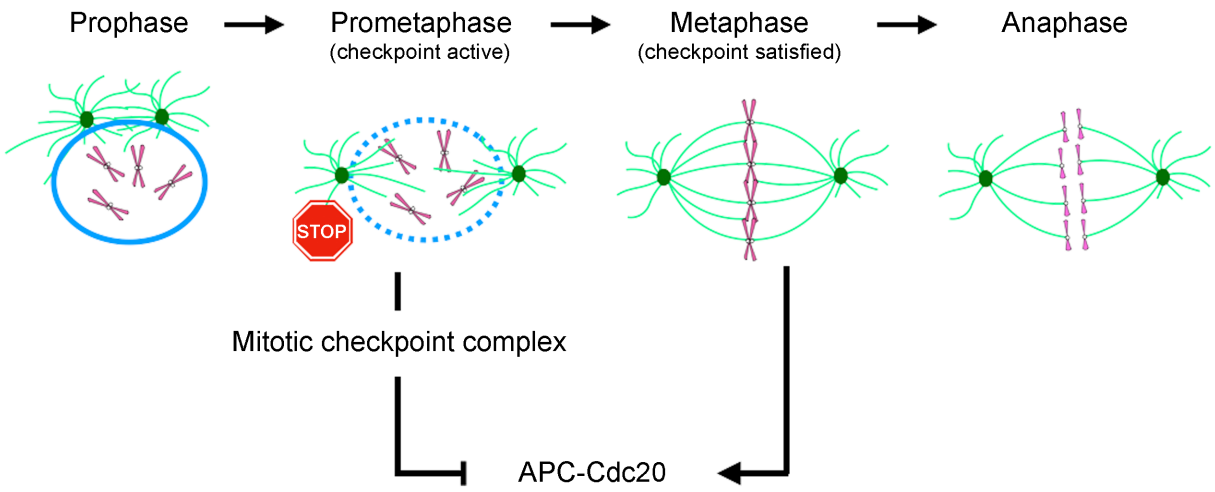

Figure 1.9. Role of the spindle assembly checkpoint. The nuclear envelope (blue circle) breaks down upon mitotic entry. The spindle assembly checkpoint (SAC) is activated during prometaphase when the kinetochore of sister chromatids (pink) fails to attach to the spindle (green lines). SAC proteins are recruited to the unattached kinetochores to form the mitotic checkpoint complex. This complex binds stably to the anaphase-promoting complex (APC)-Cdc20 and inhibits it. SAC is turned off during metaphase when all kinetochores have attached

and are under tension. The mitotic checkpoint complex then disassembles, thus relieving inhibition of APC-Cdc20. The active APC-Cdc20 complex targets its substrates for degradation, resulting in sister chromatid separation and mitotic exit during anaphase.

While it remains unclear as to how SAC discriminates properly from improperly attached chromatids, previous studies showed that the kinetochore generates the signal for SAC (Howel et al., 2004; Maresca and Salmon, 2010; Rieder et al., 1995; Shah et al., 2004). Chromatid attachment to the spindle generates tension, resulting in kinetochore stretching which prevents binding of SAC proteins (Rieder et al., 1995).

Genetic screens in budding yeast identified the components of SAC including the mitotic-arrest deficient (Mad) proteins (Mad1, Mad2, and Mad3) and the budding uninhibited by benzimidazole (Bub) proteins (Bub1 and Bub3) (Lara-Gonzalez et al., 2012). Mad2 and Mad3-Bub3 bind to Cdc20, the APC mitotic co-activator, thus forming the mitotic checkpoint complex (Lischetti and Nilsson, 2015). Mad3 is then activated through phosphorylation, mediated by Cdc5 and to a lesser extent by Cdk1 (Rancati et al., 2005). By sequestering Cdc20, the mitotic checkpoint complex inhibits APC thereby preventing it from degrading mitotic substrates (**Figure 1.9**) (Howel et al., 2004; Shah et al., 2004). The mitotic checkpoint complex can also directly bind to APC and inhibit its E3 ubiquitin ligase activity (Herzog et al., 2009; Sudakin et al., 2001). Once kinetochores correctly attach to microtubules, the mitotic checkpoint complex disassembles. The APC-Cdc20 complex is then reactivated to promote anaphase entry whereby mitotic exit ensues (**Figure 1.9**).

1.4.3 Mitotic exit

Mitotic Exit is the transition point in late anaphase that signifies the end of mitosis and the onset of a new G_1-phase for the cell. Mitotic exit is an irreversible process where the cell relies on specific control mechanisms to ensure that once it exits mitosis, it never returns until it has gone through G_1, S, and G_2 phases of the next cell cycle. Protein factors including cyclins, Cdks, APC, and inhibitors of Cdks regulate mitotic exit to ensure that cell cycle events occur correctly. Two sequential pathways coordinate mitotic exit termed the Cdc-fourteen early anaphase release (FEAR) and the mitotic exit network (MEN) (**Figure 1.10**). This results in the nucleolar release of Cdc14 phosphatase, which then drives mitotic exit through inactivation of Cdk1 and de-phosphorylation of Cdk1 substrates (Machin et al., 2016; Stegmeier and Amon, 2004).

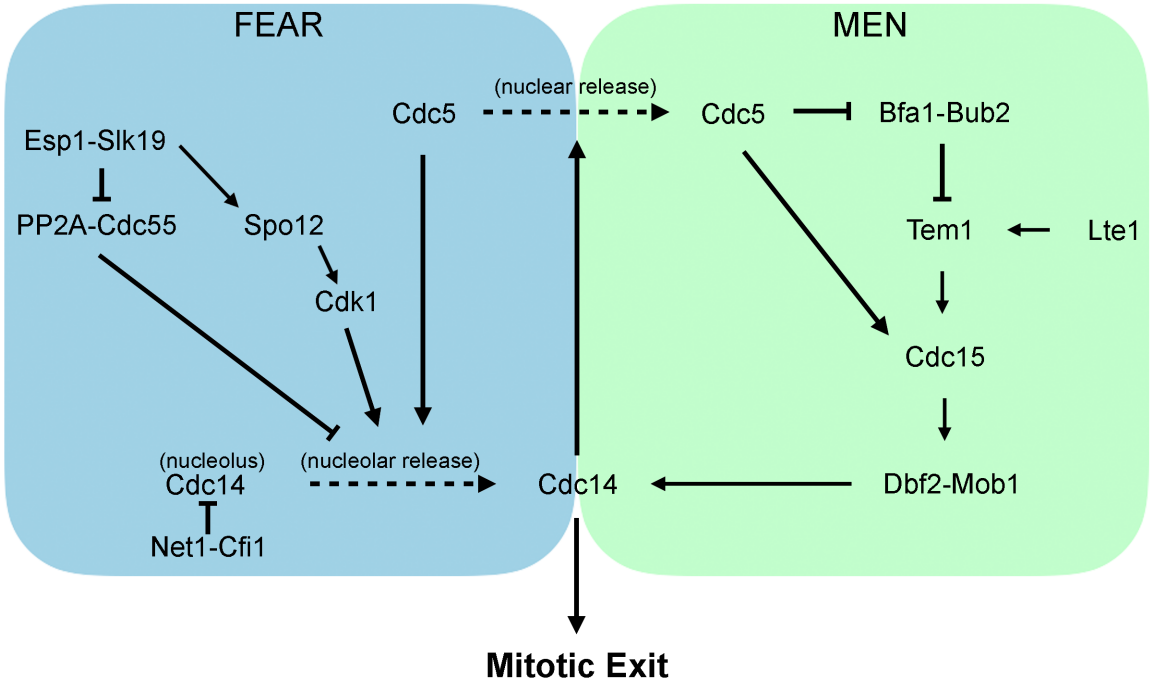

Figure 1.10. Nucleolar release of Cdc14 drives mitotic exit. Protein-protein interactions in the Cdc Fourteen-Early Anaphase Release (FEAR) and the Mitotic Exit Network (MEN) pathways. The dashed-lines represent a translocation from one cellular compartment to another.

Cdc14 is sequestered in the nucleolus by the Net1-Cfi1 adaptor proteins from G_1-phase to metaphase (**Figure 1.10**) (Shou et al., 1999). As the cell enters anaphase, the FEAR components, Esp1-Slk19 and Spo12, promote Cdk1-mediated phosphorylation of Net1-Cfi1 and Cdc5 (**Figure 1.10**) (D'Amours and Amon, 2004; Stegmeier et al., 2002). Phosphorylation of Cdc5 is necessary for its nuclear export to gain access to its cytoplasmic substrates (Botchkarev and Haber, 2018). Meanwhile, the phosphorylated form of Net1-Cfi1 binds weakly to Cdc14, resulting in the gradual release of Cdc14 from the nucleolus (**Figure 1.10**) (Rossio et al., 2010). As Cdc14 exits the nucleolus, it gains access to its nuclear and cytoplasmic substrates to initiate the second mitotic exit cascade, the mitotic exit network.

MEN functions to activate the Tem1 GTPase (**Figure 1.10**) (Rock and Amon, 2011). The Bfa1-Bub2 GTPase-activating protein complex inactivates Tem1 before anaphase commences (Geymonat et al., 2002). As the spindle poles correctly orient in the daughter cells during late anaphase, Cdc5 phosphorylates and inactivates Bfa1 on each spindle-pole body (Geymonat et al. 2002). Afterward, the Lte1 guanine exchange factor can activate Tem1 in each daughter cell (**Figure 1.10**) (Bardin et al., 2000; Pereira et al., 2000). Tem1 activation triggers the MEN through engaging its substrate, Cdc15 kinase (**Figure 1.10**) (Bardin et al., 2000; Lee et al., 2001; Visintin and Amon, 2001). Cdc5 can also directly phosphorylate Cdc15 independently of Tem1 GTPase activity (Rock and Amon 2011). As the cell exists anaphase prior to entering cytokinesis, the spindle breaks down and microtubules shorten.

1.4.4 Polo-like kinases are master regulators of mitosis

Similar to Cdk1, Cdc5 also functions at multiple stages of mitosis and hence is regarded as a master regulator of mitotic events. Cdc5 belongs to a conserved family of kinases called polo-like kinases (Plks). Plks were identified over thirty years ago from genetic screens in *Drosophila melanogaster* (*D. melanogaster*) and were deemed necessary for cell division because polo mutants were defective in both meiotic and mitotic processes (Llamazares et al., 1991; Sunkel and Glover, 1988). Human Plk1 participates in numerous mitotic functions including centrosome maturation, Golgi fragmentation, spindle assembly, kinetochore maintenance, and cytokinesis (Archambault et al., 2015; Golsteyn et al., 1995; Lane and Nigg, 1996; Sutterlin et al., 2001).

Plks are composed of a serine/threonine kinase domain followed by a polo-box domain that mediates interactions with substrates (Elia et al., 2003; Lee et al., 1998). Similar to the FHA domain, the polo-box domain contains a conserved phosphoepitope binding pocket recognizing the consensus motif, X-S-(pS/T)-P, in phosphorylated substrates (Cheng et al., 2003; Elia et al., 2003). The polo-box domain is made up of two consecutive polo boxes and a polo cap that wraps behind the second polo box (**Figure 1.11A**).

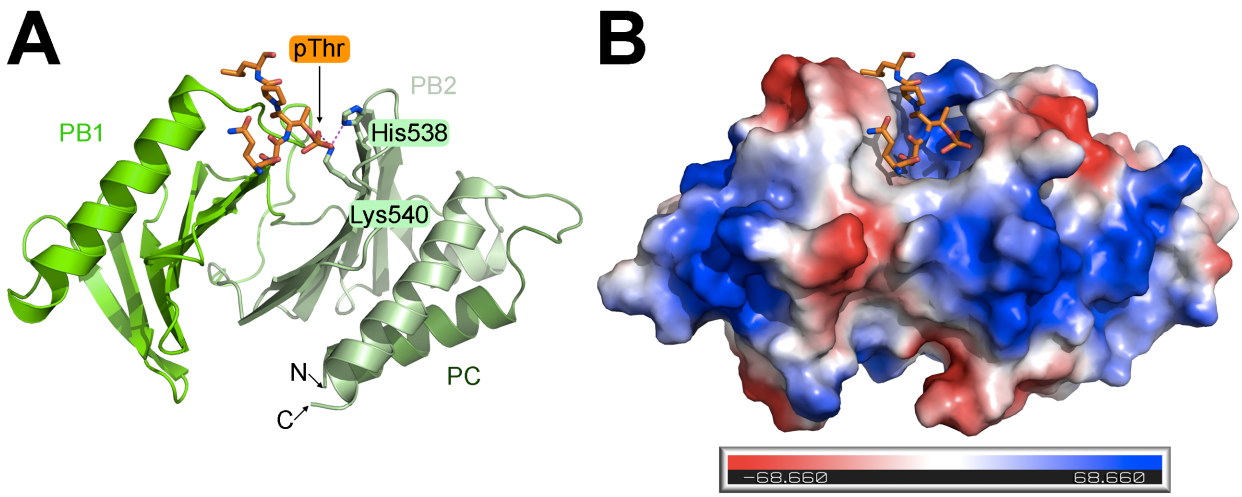

Figure 1.11. The interaction between a phosphorylated peptide and the polo-box domain of Plk1. (A) Ribbon diagram of the polo-box domain of Plk1 (green) (PDB ID: 1Q4K) in complex with a phosphorylated peptide (orange). The polo-box domain includes the polo cap (PC; dark green), polo box 1 (PB1; green), and polo box 2 (PB2; pale green). Plk1 residues (His538 and Lys540) engaging the phosphothreonine (pThr) are shown as sticks. The N- and C- termini of the polo-box domain are labeled. **(B)** Surface diagram of the polo-box domain presented in the same orientation as in panel A, and colored according to electrostatic potential. The scale is given in units of k_bT/e_c, where k_b is the Boltzmann's constant, T is the temperature in Kelvins and e_c is the electron charge.

The phosphoepitope binding pocket contains several positively-charged residues, which enable binding to phosphorylated targets (**Figure 1.11B**). His538 and Lys540, located in polo box 2, are the residues in Plk1 that recognize the phosphothreonine or phosphoserine **(Figure 1.11A)** (Cheng et al., 2003). Phospho-priming of Plk substrates is attributed to Cdk1 since it is a proline-directed kinase (Lowery et al., 2005). Cdk1-mediated phospho-priming was demonstrated for several Plk1 targets, including Mad3 at kinetochores and Swe1 during mitotic entry (Watanabe et al., 2005). However, the proline residue is not an absolute requirement for polo-box domain phosphorylated binding since Plk1 itself can phosphorylate polo-box domain-binding substrates, especially when Cdk1 activity is low during the cell cycle (Neef et al., 2007).

Apart from mitosis, Plks coordinate cellular adaptation, which is a response that allows the cell to bypass checkpoints when it encounters irreparable DNA damage (Serrano and

D'Amours, 2014). Cdc5 activity is elevated in checkpoint-arrested cells and promotes adaptation by phosphorylating specific effectors of cell cycle re-entry (Hu et al. 2001; Liang and Wang 2007). Under normal checkpoint responses, Rad53 targets Cdc5 and antagonizes its functions in promoting mitotic progression (Zhang et al., 2009). The mechanism behind how Cdc5 is re-activated to trigger the adaptation response is poorly understood. If Rad53 alone controls Cdc5 activity, then it is expected that the inhibitory effect of Rad53 towards Cdc5 will need to be lifted to promote adaptation. Cdc5 has been proposed to promote adaptation by targeting the kinase domain of Rad53 (Vidanes et al., 2010). Similar to DDK, Cdc5 also phosphorylates Rad53, but unlike DDK, Cdc5-mediated phosphorylation of Rad53 inhibits its kinase activity (Vidanes et al., 2010). The interplay between Cdc5-Rad53 dual phosphorylation remains poorly understood. It is also unclear why Cdc5 avoids targeting Rad53 in the early stages of the checkpoint response but commits later to inhibiting Rad53 in promoting adaptation.

1.5 DBF4 ENABLES DDK TO INTERACT WITH RAD53 AND CDC5

Beyond its functions in S-phase, DDK is implicated in mitosis through the interaction between Dbf4 and Cdc5 (Argunhan et al., 2017; Chen and Weinreich, 2010). Dbf4 acts as a scaffold to interact with Rad53 and Cdc5 thereby promoting crosstalk between replication, replication checkpoint, and cell division.

1.5.1 Dbf4 domain organization

Despite its essential role, Dbf4 has only three short conserved regions, which makes it challenging to identify homologs (Masai and Arai, 2000; Matthews and Guarné, 2013). The Dbf4

conserved regions, referred to as motifs, include N, M, and C denoted relative to their position in the polypeptide (N: N-terminus, M: Middle, and C: C-terminus) (**Figure 1.12**) (Masai and Arai, 2000; Matthews and Guarné, 2013).

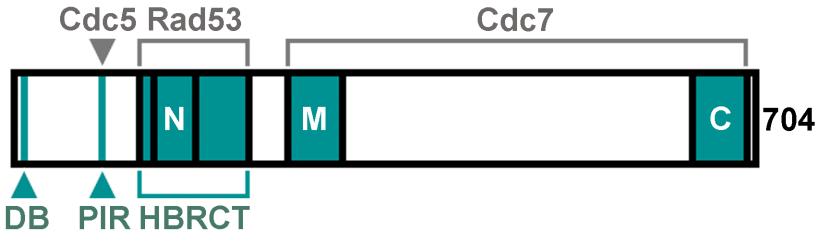

Figure 1.12. Dbf4 protein architecture. Dbf4 is a 704 amino acid protein that consists of N (residues 135-179), M (residues 260-309), and C (residues 656-697) motifs (motifs are colored teal). Motif N is located within the Helix-BRCT (HBRCT; residues 98-220) domain, which interacts with Rad53, while motifs M and C interact with Cdc7. The polo-interacting region (PIR: residues 83-88) is located N-terminal to HBRCT and interacts with Cdc5. The destruction box (DB) is found at the extreme N-terminus.

The motifs comprise a small portion of the polypeptide and mediate almost all of Dbf4's interactions with the cellular kinases (**Figure 1.12**) (Matthews and Guarné, 2013). Dbf4 is unique, as it can provide for different modes of protein-protein interactions with its substrates. For example, Dbf4 is constitutively bound to Cdc7 throughout the cell cycle, in contrast to Rad53, which interacts weakly with Dbf4 and only during the replication checkpoint response (Matthews et al., 2012; Matthews et al., 2014). Meanwhile, the Dbf4-Cdc5 interaction can be recapitulated using a peptide representing the Cdc5-interacting region in Dbf4 (Chen and Weinreich, 2010).

1.5.2 The Dbf4-Cdc7 interaction

Dbf4 interacts with Cdc7 through its M (residues 260-309) and C (residues 656-697) motifs (**Figure 1.12**). The Dbf4-Cdc7 interaction was well characterized because of the crystal structure of the complex (**Figure 1.13**) (Hughes et al., 2012).

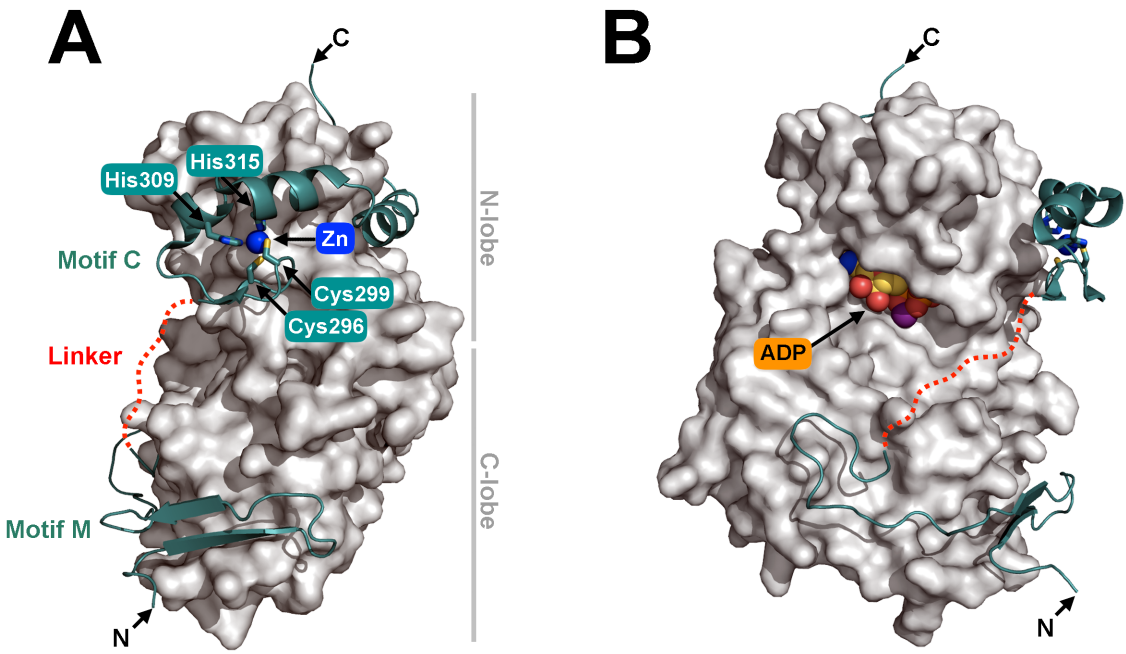

Figure 1.13. Crystal structure of the Dbf4-Cdc7 complex. (A) The crystal structure of Dbf4 (teal) and Cdc7 (white) complex from human (PDB ID: 4F9A). Cdc7 is shown as a surface diagram {(N-lobe: N-terminal lobe) and (C-lobe: C-terminal lobe)}. The linker (red) between motifs M and C represents the disordered region. The Zn ion (blue sphere) at motif C is held by residues Cys296, Cys299, His308, and His315 (shown as sticks). The N- and C-termini of Dbf4 are labeled. **(B)** 90 ˚ rotation along the y-axis of the structure presented in panel A. The ADP molecule is shown as color-coded spheres. The N- and C-termini as well as the disordered region of Dbf4 are shown.

Motif C folds as a zinc finger domain whereby histidine and cysteine residues chelate a zinc ion (**Figure 1.13A**) (Hughes et al., 2012). Binding of motif C orients the conserved catalytic residues of the Cdc7 kinase domain so that they are primed to engage ATP (**Figure 1.13B**) (Hughes et al., 2012). Motif M is not as structured as motif C but folds as a ß-sheet (**Figure 1.13A**). The Dbf4-Cdc7 complex reveals a bipartite interaction whereby motifs M and C engage the C- and N-

terminal lobes of Cdc7 respectively (**Figure 1.13A**). A bipartite interaction could serve to increase the avidity between Dbf4 and Cdc7 thereby synergizing their interaction. However, deleting either motif M or motif C does not affect complex formation, indicating that each motif can independently interact with Cdc7 (Harkins et al., 2009; Hughes et al., 2012; Ogino et al., 2001). While this shows that either motif can sustain the interaction alone, the kinase activity is not fully activated unless both motifs are present (Hughes et al., 2012).

It was shown in *Schizosaccharomyces pombe* that Dbf4 motifs M and C can be expressed separately and achieve moderate Cdc7 activity (Ogino et al., 2001). Deleting motif M modestly reduces the kinase activity of Cdc7, implying that motif C plays a critical role in regulating the kinase activity (Hughes et al., 2012). The current consensus is that motif M is responsible for Dbf4-binding while motif C regulates Cdc7 kinase activity (Hughes et al., 2012). While motifs M and C work together to bind and activate Cdc7, Dbf4 also uses additional regions to fine-tune its kinase activity. The extreme C-terminal tail in Dbf4 homologs from higher eukaryotes binds to and inhibits Cdc7 (Sato et al., 2003). This inhibition can then be alleviated by phosphorylating the C-terminal tail or by having it interact with another binding partner (Hughes et al., 2010).

1.5.3 The Dbf4-Rad53 interaction

While Dbf4 motifs M and C recognize Cdc7, motif N interacts with the FHA1 domain of Rad53 (Matthews et al., 2012; Matthews et al., 2014; Varrin et al., 2005). Yeast two-hybrid studies demonstrated that disrupting motif N abrogates the Rad53 interaction (Gabrielse et al., 2006; Varrin et al., 2005). As expected, these mutants were hypersensitive to replication stress due to an inefficient checkpoint response (Gabrielse et al., 2006; Varrin et al., 2005). Motif N

was predicted to be part of a BRCA1 C-terminal (BRCT) domain (Masai and Arai, 2000). The crystal structure of a fragment including motif N (residues 98-220) indicated that motif N does not fold as the BRCT domain, but instead is part of one (**Figure 1.14A**) (Matthews et al., 2012). Interactions mediated by single BRCT domains are less characterized compared to tandem pairs that recognize phosphoepitopes in substrates (Matthews and Guarné, 2013). Moreover, single BRCT domains are best known for their DNA binding properties rather than mediating protein - protein interactions.

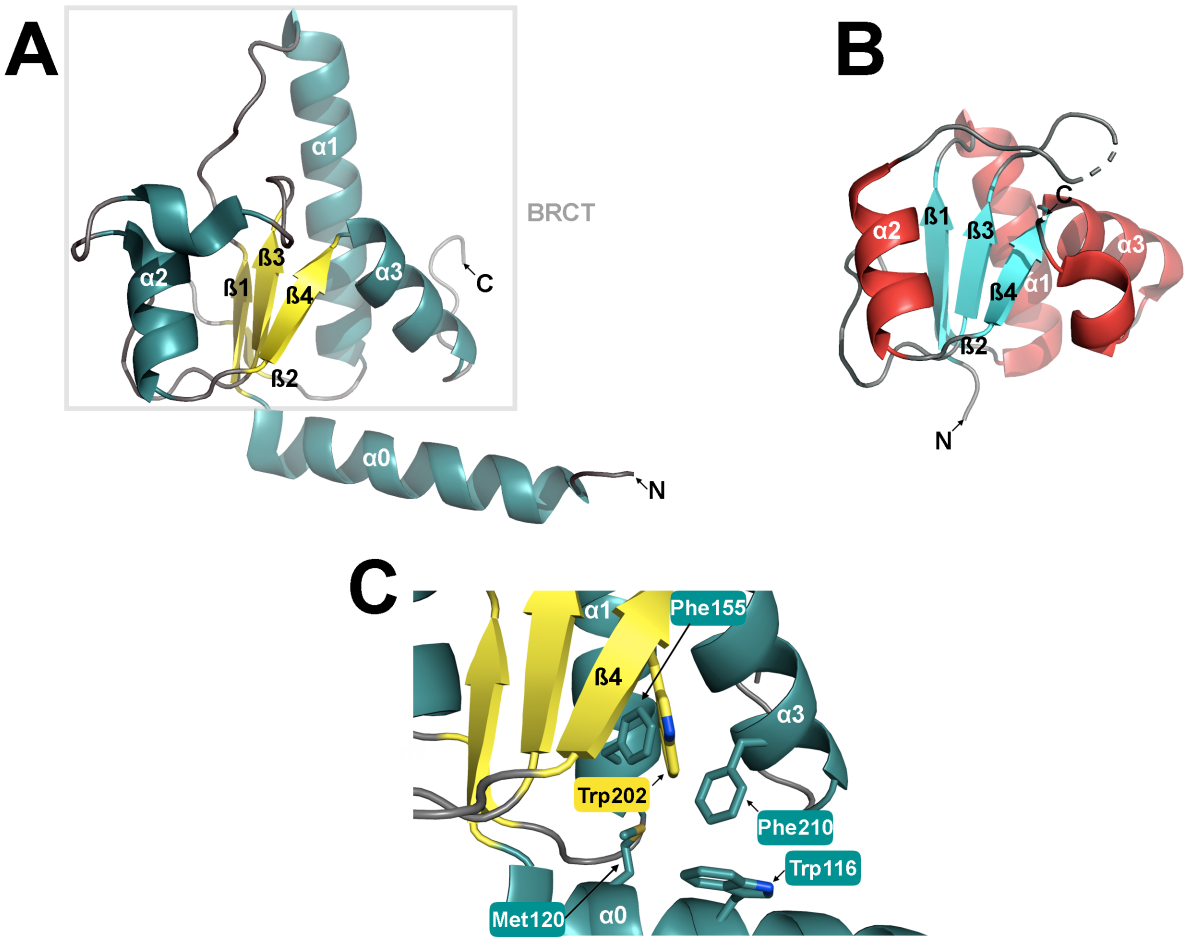

Figure 1.14. The Helix-BRCT domain of Dbf4. (A) Ribbon diagram of the Helix-BRCT (HBRCT) domain of Dbf4 (PDB ID: 3QBZ) colored according to secondary structure: α-helices (teal), ß-strands (yellow), and random coil (grey). The BRCT is separately boxed. **(B)** The first BRCT domain of human BRCA1 (PDB ID: 1JNX, residues 1649–1737) oriented similarly as in panel A, and colored according to secondary structure: α-helices (red), ß-strands (cyan), and

random coil (grey). The dashed-lines between α2-ß3 represent missing residues. **(C)** Hydrophobic contacts that anchor α0 to the BRCT domain. The secondary structure elements pertaining to the hydrophobic residues are labeled.

The structure revealed that the BRCT domain of Dbf4 contains an extra element, the α0 helix, which is necessary for the interaction with Rad53 (**Figure 1.14A**) (Matthews et al., 2012; Matthews et al., 2014; Matthews and Guarné, 2013). Collectively, the fragment was termed the Helix-BRCT (HBRCT) domain of Dbf4 (Matthews et al., 2012). The BRCT domain of Dbf4 resembles other BRCT domains, except that it has a more extended α1 helix (**Figure 1.14A-B**) (Matthews et al., 2012). Moreover, the α0 helix is dispensable for the folding of the BRCT domain (Matthews et al., 2012). In this regard, the crystal structure of a Dbf4 fragment lacking α0 showed that it was identical to the BRCT domain in HBRCT (Matthews et al., 2012).

Neither α0 nor BRCT interacted with FHA1 *in vitro*, implying that Dbf4 requires the entire HBRCT domain to bind FHA1 (Matthews et al., 2012). Mutations of hydrophobic residues on α0 abrogated the interaction with FHA1 in a yeast two-hybrid assay (Matthews et al., 2012). These residues are packed with additional hydrophobic residues located in the BRCT domain, and contribute to the structural integrity of the HBRCT domain (**Figure 1.14C**). The HBRCT domain exhibits a low-affinity interaction with FHA1, and this complex was challenging to capture *in vitro* (Matthews et al., 2014). At the onset of this thesis, the molecular details of the HBRCT-FHA1 interaction were not fully understood.

Apart from Rad53, Dbf4 uses motif N to engage ORC at pre-RCs (Duncker et al., 2002). It was initially anticipated that Dbf4 binding to ORC serves to recruit DDK to origins of replication (Duncker et al., 2002). However, this idea was overruled since deleting motif N did

not result in replication defects (Gabrielse et al., 2006; Jones et al., 2010; Sheu and Stillman, 2006).

1.5.4 The Dbf4-Cdc5 interaction

Although Dbf4 utilizes its motifs to interact with Cdc7 and Rad53, its interaction with Cdc5 is mediated through an unstructured region located N-terminal to the HBRCT domain, termed the polo-interacting region (**Figure 1.12**). The Dbf4 polo-interacting region (^{83}RSIEGA^{88}) interacts with the Cdc5 polo-box domain even though it does not resemble the consensus phosphorylated motif recognized by polo-box domains (Chen and Weinreich, 2010). In this regard, Dbf4 recognizes an unknown surface on the Cdc5 polo-box domain and binds to it in a phosphorylation-independent manner (Chen and Weinreich, 2010). At the onset of this thesis, the molecular details of the Dbf4-Cdc5 interaction were unknown.

DDK is implicated in the mitotic checkpoint response by inhibiting Cdc5 when spindle poles mis-position in daughter cells during anaphase (Miller et al., 2009). Dbf4 binding to Cdc5 enables Cdc7 to phosphorylate the Cdc5 polo-box domain *in vitro* (Miller et al., 2009). It is presumed that phosphorylated Cdc5 is inactive, given its failure to recognize its substrate, Cdc15, in the MEN pathway (Miller et al., 2009). The Drosophila matrimony (Mtrm) protein represents another polo-box domain-interacting partner that inhibits Polo function (Xiang et al., 2007). Mtrm functions acts as a negative regulator of Polo thereby arresting cells in G_2-phase (Xiang et al., 2007). However, unlike Dbf4, Mtrm binds the polo-box domain of Polo in a phosphorylation-dependent manner.

DDK also functions with Cdc5 during meiosis, which is a process that drives gamete production in sexually-reproducing organisms. Sister chromatid exchange involves sister chromatid organization around a proteinaceous axis to form a meiosis-specific chromosomal structure termed the synaptonemal complex (SC) (Cahoon and Hawley, 2016). Sister chromatid exchange marks an irreversible commitment to meiosis, with SC disassembly during metaphase I (Tsuchiya et al., 2014). Cdc5 and DDK coordinate the destruction of the SC (Argunhan et al., 2017; Sourirajan & Lichten, 2008). Dbf4 serves as a regulator by interacting with and being phosphorylated by Cdc5. DDK and Cdc5 promote SC disassembly by phosphorylating the Red1 and Zip1 SC proteins (Argunhan et al., 2017). This coordinated action leads to the reactivation of Rad51, which promotes the repair of any persisting double-stranded breaks (DSBs) before chromosomes are separated during anaphase I. By facilitating the removal of SC and triggering Rad51-dependent DSB repair, DDK and Cdc5 collaborate to ensure faithful inheritance of the genome.

1.7 THESIS OBJECTIVES

Unlike Cdc7, which is always bound to Dbf4, Rad53 and Cdc5 interact with Dbf4 at specific points during the cell cycle. The crystal structure of the Cdc7-Dbf4 complex revealed the molecular details of how motifs M and C engage and regulate Cdc7. At the onset of this thesis, it was unclear how the HBRCT domain interacts weakly with the FHA1 domain of Rad53, or how the polo-interacting region recognizes the polo-box domain of Cdc5. To characterize these Dbf4-mediated interactions, we pursued the following aims:

(i) Determining the Dbf4-Rad53 complex at atomic resolution. (Chapter 2)

(ii) Identifying the surface on the polo-box domain of Cdc5 that binds to Dbf4. (Chapter 3)

We also characterized another low-affinity interaction involving the ß clamp and MutL, which is implicated in DNA mismatch repair. Like Dbf4 and Rad53, the structures of the ß clamp and MutL were solved (Burnouf et al., 2014; Guarné et al., 2004; Kong et al., 1992; Pillon et al., 2010) and their interaction was extensively characterized *in vitro* (Pillon et al., 2015). We thus pursued the following aims:

(iii) Stabilizing the ß clamp-MutL low-affinity interaction for structural studies. (Chapters 4)

(iv) Understanding how to stabilize low-affinity protein-protein interactions. (Chapter 5)

CHAPTER TWO

‘AND’ logic gates at work: Crystal structure of Rad53 bound to Dbf4 and Cdc7.

This research was originally published in Scientific Reports. Almawi AW, Matthews LA, Larasati, Myrox P, Boulton S, Lai C, Moraes T, Melacini G, Ghirlando R, Duncker BP, and Guarné A. (2016). ’AND’logic gates at work: Crystal structure of Rad53 bound to Dbf4 and Cdc7. Sci Rep 6:34327 . 10.1038/srep34237.

2.1 CONNECTING TEXT

Chapter 2 analyzes the crystal structure of the Dbf4-Rad53-Cdc7 complex. The Dbf4-Rad53 interaction was captured by creating a fusion of the two proteins. The C-terminus of the HBRCT domain of Dbf4 was fused to the N-terminus of the FHA1 domain of Rad53 to create the fusion. Multiple fusions with different linker lengths were generated, but only one of them yielded diffraction-quality crystals. Larasati and P. Myrox (from the laboratory of Dr. Duncker, University of Waterloo) performed the *in vivo* experiments. Dr. R. Ghirlando (National Institutes of Health) conducted, interpreted, and analyzed the ultracentrifugation data. Dr. L. Matthews (from the laboratory of Dr. Guarné, McMaster University) performed the NMR experiments and provided technical and intellectual input to the crystallographic work. S. Boulton (from the laboratory of Dr. Melacini, McMaster University) analyzed the NMR data. C. Lai (from the laboratory of Dr. Moraes, University of Toronto) provided technical and intellectual input to analyze the affinities of the Dbf4-Rad53 and Dbf4-Rad53-Cdc7 complexes. With the guidance of my supervisor, Dr. Alba Guarné, I conducted the experiments, interpreted the data, prepared several figures, and assisted with the writing of the manuscript.

2.2 ABSTRACT

Forkhead-associated (FHA) domains are phosphopeptide recognition modules found in many signaling proteins. The Saccharomyces cerevisiae protein kinase Rad53 is a key regulator of the DNA damage checkpoint and uses its two FHA domains to interact with multiple binding partners during the checkpoint response. One of these binding partners is the Dbf4-dependent kinase (DDK), a heterodimer composed of the Cdc7 kinase and its regulatory subunit Dbf4. Binding of Rad53 to DDK, through its N-terminal FHA (FHA1) domain, ultimately inhibits DDK kinase activity, thereby preventing firing of late origins. We have previously found that the FHA1 domain of Rad53 binds simultaneously to Dbf4 and a phosphoepitope derived from Cdc7 kinase, suggesting that this domain functions as an 'AND' logic gate. Here, we present the crystal structures of the FHA1 domain of Rad53 bound to Dbf4, in the presence and absence of a Cdc7 phosphorylated peptide. Our results reveal how the FHA1 uses a canonical binding interface to recognize the Cdc7 phosphopeptide and a non-canonical interface to bind Dbf4. Based on these data we propose a mechanism to explain how Rad53 enhances the specificity of FHA1-mediated transient interactions.

2.3 INTRODUCTION

Stress generated during DNA replication is one of the biggest hurdles proliferating cells face to preserve genome integrity. Therefore, eukaryotic cells have conserved surveillance mechanisms, known as cell cycle checkpoints, to detect and repair damage generated during DNA replication (Branzei and Foiani, 2005; Putnam et al., 2009; Segurado and Tercero, 2009; Zegerman and Diffley, 2009). Rad53, and its mammalian ortholog the checkpoint kinase 2 (Chk2), are key effector kinases of the DNA replication checkpoint (Bartek and Lukas, 2003;

Rouse and Jackson, 2002). Loss-of-function mutations in RAD53 cause loss of viability due to an essential function in maintaining dNTP levels during DNA replication, while hypomorphic RAD53 mutations result in DNA damage sensitivity and deficits in checkpoint responses (Allen et al., 1994; Moore, 1978; Paulovich and Hartwell, 1995; Fay et al., 1997). Similarly, loss-of-function mutations in Chk2 lead to a defective checkpoint response (Matsuoka et al., 1998; Hirao et al., 2000).

Rad53 contains two forkhead-associated (FHA) domains, as well as two SQ/TQ cluster domains (SCD), flanking its kinase domain. FHA domains are commonly found in DNA damage response proteins and mediate protein-protein interactions by recognizing phosphorylated epitopes on their binding partners (Durocher et al., 2000). During the replication checkpoint, phosphorylation-dependent interactions mediated by the FHA domains of Rad53 trigger hyperphosphorylation of the N-terminal SCD domain and lead to the full activation of Rad53 (Pellicioli and Foiani, 2005). It was generally believed that FHA domains recognize unstructured sequences containing a phosphorylated amino acid –often a threonine. Recent studies, however, have shown that FHA domains can also use alternate surfaces for protein oligomerization and to mediate protein-protein interactions (Luo et al., 2015; Matthews et al., 2014; Nott et al., 2009; Raasch et al., 2014; Weng et al., 2015). The Dbf4-dependent kinase (DDK) and Rad9 are two binding partners of Rad53 during the replication checkpoint response. Dbf4 preferentially interacts with the N-terminal FHA domain (FHA1) of Rad53 (Duncker et al., 2002), whereas Rad9 binds the C-terminal FHA domain (FHA2) (Sun et al., 1998), reinforcing the idea that the two FHA domains recognize distinct features on their binding partners. DDK, a heterodimer composed of the Ser/Thr kinase Cdc7 and its regulatory subunit Dbf4, is one of the kinases

known to hyperphosphorylate Rad53 (Ogi et al., 2008). Reciprocally, Rad53 phosphorylates DDK to inhibit its activity, thereby preventing the firing of late replication origins (Zegerman and Diffley, 2010). This is crucial as it inhibits S-phase progression and allows cells to recover from replication stress.

The interaction between Rad53 and DDK is of special interest because it involves multiple interfaces of the FHA1 domain. The phosphoepitope-binding site recognizes an epitope present in DDK (Aucher et al., 2010), whereas one of the lateral surfaces of FHA1 interacts with the modified BRCT domain of Dbf4 (Matthews et al., 2014), herein referred to as HBRCT domain due to the presence of an additional α-helix at the N-terminus of the domain. However, like many other relevant signaling interactions, the Rad53 and Dbf4 association is weak and presumably transient. The latter are especially difficult to study for effector proteins like Rad53, because they often interact with multiple partners using a common interface. To understand how Rad53 manages its multiple interactions during the steps leading to Cdc7 inhibition, we stabilized the Rad53:Dbf4 complex using glycine-rich linkers. We generated chimeras expressing the HBRCT (Dbf4) and FHA1 (Rad53) domains in tandem and solved the crystal structures of these chimeras in the absence and presence of a phosphorylated epitope derived from Cdc7. The structure of the ternary complex (Dbf4-Rad53-Cdc7) revealed that Dbf4 and Cdc7 simultaneously engage different surfaces on the FHA1 domain of Rad53. These are the first structures of an FHA domain bound to a binding partner through a non-canonical interface and they reveal a unique bipartite interface between Rad53 and Dbf4 that provides exquisite specificity despite the minimal interaction surface.

2.4 METHODS

2.4.1 Cloning and expression

Dbf4-Rad53 chimeras were created by subcloning a codon-optimized fragment of Dbf4 encompassing amino acids 105–220 followed by the FHA1 domain of Rad53 (amino acids 22–162) into a modified pET15b vector including His6-SUMO tag with a TEV protease cleavage site (pAG8586). The two protein fragments were connected directly (Dbf4(0)Rad53) or separated by a five-residue linker (Dbf4-SGASG-Rad53, herein referred to as Dbf4(5)Rad53). Clones were confirmed by DNA sequencing (MOBIX, McMaster University). Plasmids encoding the Dbf4(0)Rad53 (pAG8801) and Dbf4(5)Rad53 (pAG8805) chimeras were transformed in BL21(DE3) cells containing a plasmid encoding rare tRNAs. Cultures were grown to A600 = 0.7, induced by addition of 1 mM isopropyl β-D-1-thiogalactopyranoside, and incubated overnight at 16 °C with orbital agitation.

2.4.2 Protein purification

Cell pellets were resuspended in buffer A (20 mM TRIS-HCl pH 8.0, 500 mM NaCl, 1.4 mM 2-mercaptoethanol, 5% glycerol) and lysed by sonication. Lysates were cleared by centrifugation at 39,000 g, and the supernatants were loaded onto a HiTrap nickel-chelating HP column (GE Healthcare) equilibrated with buffer A. The His6-SUMO-tagged chimeras were eluted with a linear gradient to 300 mM imidazole. The fractions containing the chimera were pooled and injected onto a HiPrep 26/10 desalting column (GE Healthcare) equilibrated with buffer B (20 mM TRIS-HCl pH 8.0, 150 mM NaCl, 1.4 2-mercaptoethanol, 5% glycerol). The

His6-SUMO tag was removed with tobacco etch virus (TEV) protease, and the tagless chimeras further purified by affinity (HiTrap nickel-chelating HP column, GE Healthcare) and size-exclusion chromatography (Superdex 75 (10/300) GL column, GE Healthcare). The purified proteins were concentrated to 9–12 mg/mL and stored in buffer B. Protein concentrations were determined using the Beer-Lambert equation with an extinction coefficient of 36,440 M−1 cm−1.

2.4.3 Analytical ultracentrifugation

Sedimentation velocity experiments were conducted at 50,000 rpm and 20 °C on a Beckman Coulter ProteomeLab XLI analytical ultracentrifuge following standard protocols (Zhao et al., 2013). Samples of the Dbf4(5)Rad53 chimera were studied at various loading concentrations ranging from 2 to 310 μM in 0.2 M NaCl, 0.02 M TRIS-HCl pH 8.0, 1.4 mM 2-mercaptoethanol and 5% v/v glycerol. Samples were loaded in 2-channel centerpiece cells and data were collected using both the absorbance (280 nm) and Rayleigh interference (655 nm) optical detection systems when possible. Standard 12 mm centerpieces were used, whereas shorter 3 mm centerpieces were used for the higher concentration protein samples (>70 μM). Sedimentation data were time-corrected (Ghirlando et al., 2013) and analyzed in SEDFIT 15.01b (Schuck, 2000) in terms of a continuous c(s) distribution of sedimenting species with a resolution of 0.05 S and a maximum entropy regularization confidence level of 0.68. The solution density, solution viscosity and protein partial specific volume were calculated in SEDNTERP (Cole et al., 2008, http://sedn- terp.unh.edu/), and sedimentation coefficients s were corrected to standard conditions s20,w. Weighted-average sedimentation coefficients obtained by integration of the c(s) distributions were used to create an isotherm that was analyzed in SEDPHAT 13.0a in terms of a reversible monomer-dimer equilibrium. The protein extinction coefficient at 280 nm and the

interference signal increment were calculated based on the amino acid composition in SEDFIT 15.01b (Zhao et al., 2011).

2.4.4 Crystallization, structure determination, and refinement

Crystals of the Dbf4(5)Rad53 grew in 50 mM sodium cacodylate pH 6.5, 12% PEG 4000 (v/v), and 250 mM MgCl2 and cryo-protected by addition of 10% ethylene glycol. A complete data set was collected at the X29 beamline of NSLS-I (Brookhaven National Laboratory). Data was processed and scaled in HKL2000 (**see Table 2.1**) (Otwinowski and Minor, 1997). A phosphorylated peptide (pPEP) derived from Cdc7 (^{480}DGESpTDEDDVVS^{491}) was purchased from GenScript and resuspended in buffer B. The Dbf4(0)Rad53 chimera was mixed with the phosphorylated peptide at a 10-fold molar excess and incubated at 4 °C for one hour prior to crystallization trials. Crystals of the Dbf4(0)Rad53-pPEP complex were grown in 100 mM TRIS pH 8.5, and 12.5% PEG 3350 (v/v) and cryo-protected by addition of 15% glycerol. A complete data set was collected at the O8B1-1 beamline of the Canadian Light Source and processed using XDS (**see Table 2.1**) (Kabsch, 2010).

Table 2.1. Data collection and refinement statistics.

	Dbf4(5)Rad53	Dbf4(0)Rad53 + pPEP
Data Collection		
Beamline	X29 (NSLS)	08B1-1 (CLS)
Wavelength (Å)	1.1	0.979
Space group	P 1	P 2_1
Cell dimensions		
a, b, c	57.7, 66.6, 86.6	64.5, 87.3, 66.1
α, β, γ	109.5, 90.1, 90.1	90, 94, 90
Resolution	35–2.3 (2.34–2.3)	44.6–2.25 (2.31–2.25)
Reflections (total/unique)	887,843 / 55,038	101,589 / 36,580
Completeness (%)	87.2 (57.3)	98.3 (97.4)
CC1/2 (%)	97.1 (92.5)	99.3 (31.8)
I/σ((I)	13.6 (1.4)	8.15 (1)
Redundancy	1.6 (1.4)	2.8 (2.8)
Refinement		
Resolution (Å)	35–2.66	44.6–2.4
Completeness (%)	91.1	98.3
R_{work} / R_{free} (%)	20.6 / 23.7	20.7 / 22.9
Atoms refined	15,338	8,312
Solvent Atoms	175	192
rmsd in bonds (Å)	0.004	0.003
rmsd in angles (°)	0.834	0.733
Mean B values (Å^2)	45.6	51.5

Both structures were determined by molecular replacement using the FHA1 domain of Rad53 (PDB 1G6G) and the HBRCT domain of Dbf4 (PDB 3QBZ) as search models. The initial

models were refined by iterative cycles of manual model building in Coot and refinement in PHENIX (Adams et al., 2010). The refined models have 98% (Dbf4(5)Rad53) and 98.4% (Dbf4(0)Rad53-pPEP) of the residues in the most favored regions of the Ramachandran plot and none in the disallowed regions. Quantitative analysis of the Dbf4(L)Rad53 (±pPEP) interfaces was done using the online Protein Interfaces Structures and Assemblies (PISA) server (Krissinel and Henrick, 2007). Figures showing molecular structures were generated using PyMOL (DeLano, 2002).

2.4.5 Analysis of the NMR data

Gradient and sensitivity enhanced [^{1}H-^{15}N] heteronuclear single quantum coherence (HSQC) spectra were acquired at 306 K using a Bruker AV-700 MHz spectrometer equipped with a 5 mm TCI cryoprobe. Samples were prepared as described by ref. 16 with an equimolar concentraion of FHA1 and HBRCT in either the absence or presence of 200 μM phosphorylated Cdc7 peptide (pPEP, ^{480}DGESpTDEDDVVS^{491}). Spectra were processed using NMRPipe (Delaglio et al., 1995) and analyzed in Sparky.

2.4.6 Yeast two-hybrid experiments

Point mutations within FHA1 (F146A and N112A/F146A) were generated by site-directed mutagenesis using pJG4-6 FHA1 (including residues 1–165 of Rad53) as the template. Single point mutations on Dbf4 (Y198A and K200A) were generated from the pEG202-Dbf4 full-length template. All constructs were verified by DNA sequencing (MOBIX, McMaster University). Two-hybrid analysis was carried out as described previously using pEG202-Dbf4-FL as the bait and the indicated pJG4-6 FHA1 variants as the prey36. The lacZ reporter pSH18-34, pEG-202-Dbf4-FL-derived bait and pJG4-6-Rad53-FHA1-derived prey plasmids

were transformed into DY-1 yeast strain. Cells were grown to a concentration of 2 × 107 cells/ mL in Synthetic Defined (SD) medium lacking uracil, histidine, and tryptophan. Prey expression was induced for 5–6 h in 20 mL of 2% galactose-1% raffinose medium (BD Bioscience). Using a total of 5 × 106 cells, the interactions between fusion proteins were detected by the quantitative β-Galactosidase (β-Gal) assay. The β-Gal activity was calculated using the following equation: β-Gal activity = 1,000 × A420/(t × V × A600), where t = time of reaction (in min) and V = volume of culture used in the assay (in mL). The remaining culture was centrifuged at 2,000 g for 3 min; the cell pellets were resuspended in lysis buffer (10 mM Tris-HCl pH 8.0, 140 mM NaCl, 1 mM EDTA, 1% Triton X-100, with protease inhibitors) and lysed with a bead beater (BioSpec Products, Inc.). Cell lysates were clarified by centrifugation and the extracted supernatant represented the whole-cell extract. Protein concentration was quantified using a Bradford assay, and protein expression was analyzed by Western blotting. The LexA-tagged Dbf4 bait was detected using a rabbit polyclonal anti-LexA antibody (Invitrogen), and the HA-tagged FHA1 prey was detected using a mouse monoclonal anti-HA antibody (Sigma). Alexa Fluor 647 goat anti-rabbit and Alexa Fluor 488 goat anti-mouse polyclonal secondary antibodies (Molecular Probes) were used.

2.5 RESULTS

2.5.1 The Dbf4(L)Rad53 chimeras have a weak self-association

We have previously shown that the HBRCT domain of Dbf4, consisting of a BRCT fold immediately preceded by a helix, is necessary and sufficient for the interaction with the FHA1 domain of Rad53 (Matthews et al., 2012). The interaction with this domain of Dbf4 is mediated by one of the lateral surfaces of the FHA1 domain rather than its phosphopeptide-binding pocket

(Matthews et al., 2014). However, the instability of the HBRCT domain at high concentrations prevented the characterization of the reciprocal surface in Dbf4. Based on our biochemical, genetic and structural data, we generated a preliminary model of the Dbf4-Rad53 complex using the Rosetta software. In this model, the lateral surface of the FHA1 domain interacted with the concave surface of the HBRCT domain of Dbf4 leaving the termini of both domains in close proximity. Therefore, we anticipated that we could stabilize the interaction by producing the two domains as a single polypeptide chain. We fused the FHA1 domain of Rad53 at the C-terminus of the HBRCT domain of Dbf4 directly, or using a five-residue glycine/serine-rich linkers (Dbf4(0)Rad53 and Dbf4(5)Rad53; **Figure 2.1A**).

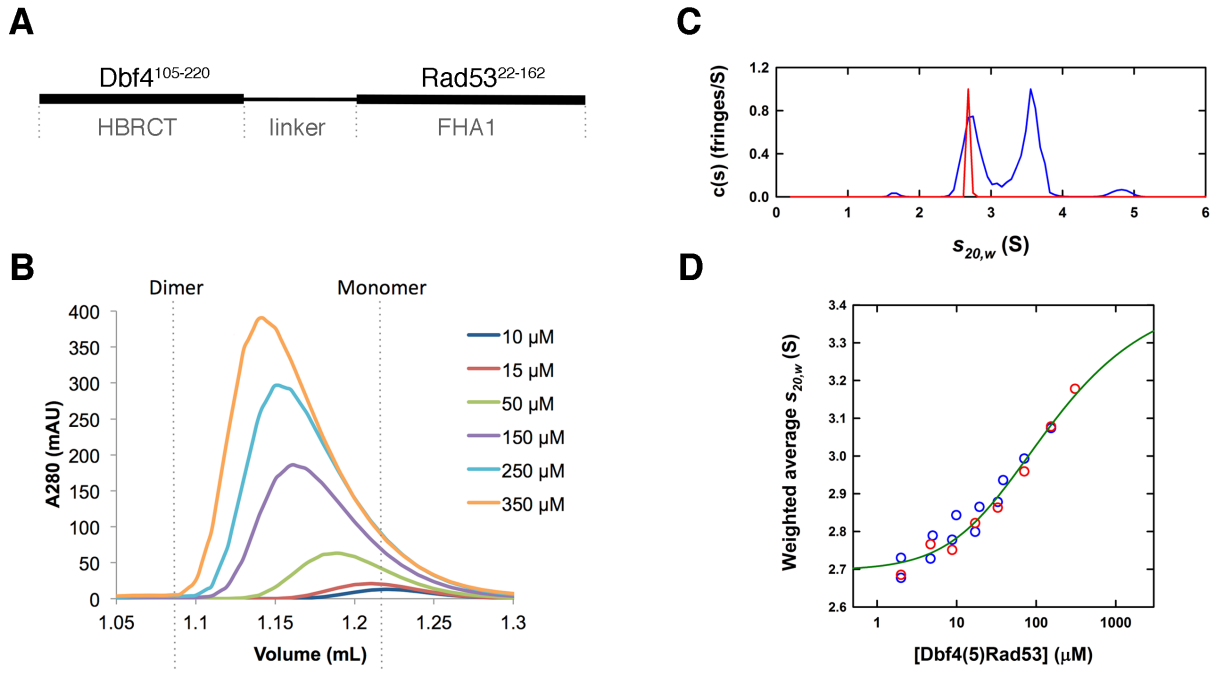

Figure 2.1. The Rad53(5)Dbf4 chimera exists in a monomer-dimer equilibrium. (A) Cartoon depicting how the Rad53(L)Dbf4 chimeras were generated. **(B)** Size exclusion chromatography profiles of Dbf4(5)Rad53 at increasing protein concentrations. Elution volumes for a ideal Dbf4(5)Rad53 monomer and dimer are indicated as dashed lines. **(C)** Normalized interference c(s) profiles for Dbf4(5)Rad53 at 2 µM (red) and 310 µM (blue) supporting a reversible monomer-dimer equilibrium. **(D)** Dependence of the weighted-average s20,w on the loading concentration for absorbance (blue) and interference (red) sedimentation velocity data, along with the global best-fit monomer-dimer isotherm (green).

The resulting chimeras were monodisperse and behaved as monomers in solution as judged by dynamic light scattering and size exclusion chromatography. Despite being predominantly monomeric, the elution times from an analytical size exclusion chromatography varied in a concentration-dependent manner suggesting a weak intermolecular association (**Figure 2.1B and Figure 2.2**).

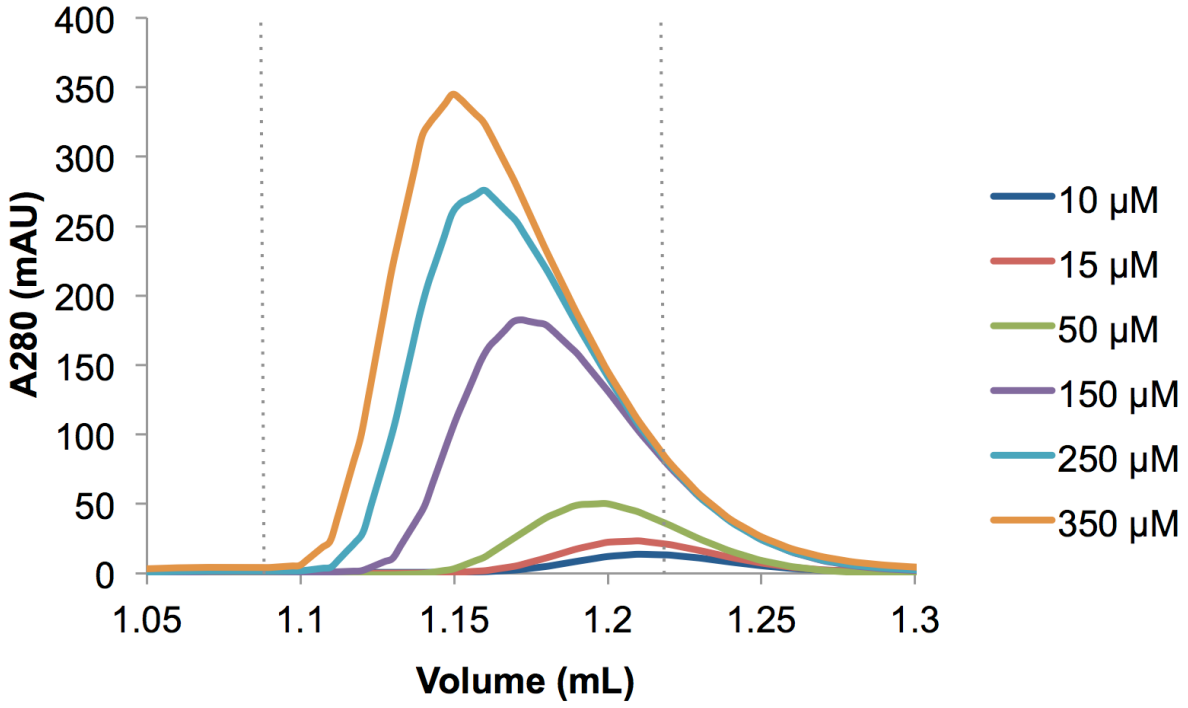

Figure 2.2. The Dbf4(0)Rad53 chimera exist in a monomer:dimer equilibrium. Size exclusion chromatography profiles of Dbf4(L0)Rad53 at increasing concentrations.

We carried out a series of sedimentation velocity experiments on the Dbf4(5)Rad53 chimera at increasing loading concentrations. The sedimentation experiments demonstrated the absence of very large species and yielded c(s) profiles that supported a reversible monomer-dimer equilibrium (**Figure 2.1C**). Dimerization of chimeric proteins is not uncommon and it indicates that the two components of the chimera associate intermolecularly (Foss et al., 2005; Wang et al., 2008; Williams et al., 2014). A weighted-average sedimentation coefficient isotherm

was constructed and analyzed in terms of reversible monomer-dimer equilibrium (**Figure 2.1D**), to obtain a dissociation constant Kd of 130 μM. As the isotherm does not adequately cover the high concentration region there may be significant errors in this value. Based on the reduced chi-squared, using the method of F-statistics (Johnson, 1992), we estimate 68% and 95% confidence limits of the Kd to be 70–260 μM and 50–400 μM. These values indicate the order of magnitude of the interaction and confirm that the HBRCT domain of Dbf4 and the FHA1 domain of Rad53 associate weakly.

2.5.2 The Dbf4(L)Rad53 chimeras recreate the Dbf4:Rad53 interaction

To avoid constraints imposed by the presence of the linker joining the two proteins, we set crystallization trials of two chimeras: Dbf4(0)Rad53 and Dbf4(5)Rad53. The chimera containing a five-residue linker readily yielded diffraction-quality crystals (**Table 2.1**). The asymmetric unit contained four copies of the Dbf4(5)Rad53 chimera arranged to form four Dbf4:Rad53 complexes. The C-terminal end of Dbf4 (residues 216–220) and the N-terminal end of Rad53 (residues 22–29), as well as the five amino acid linker, were disordered in the structure (**Figure 2.3**).

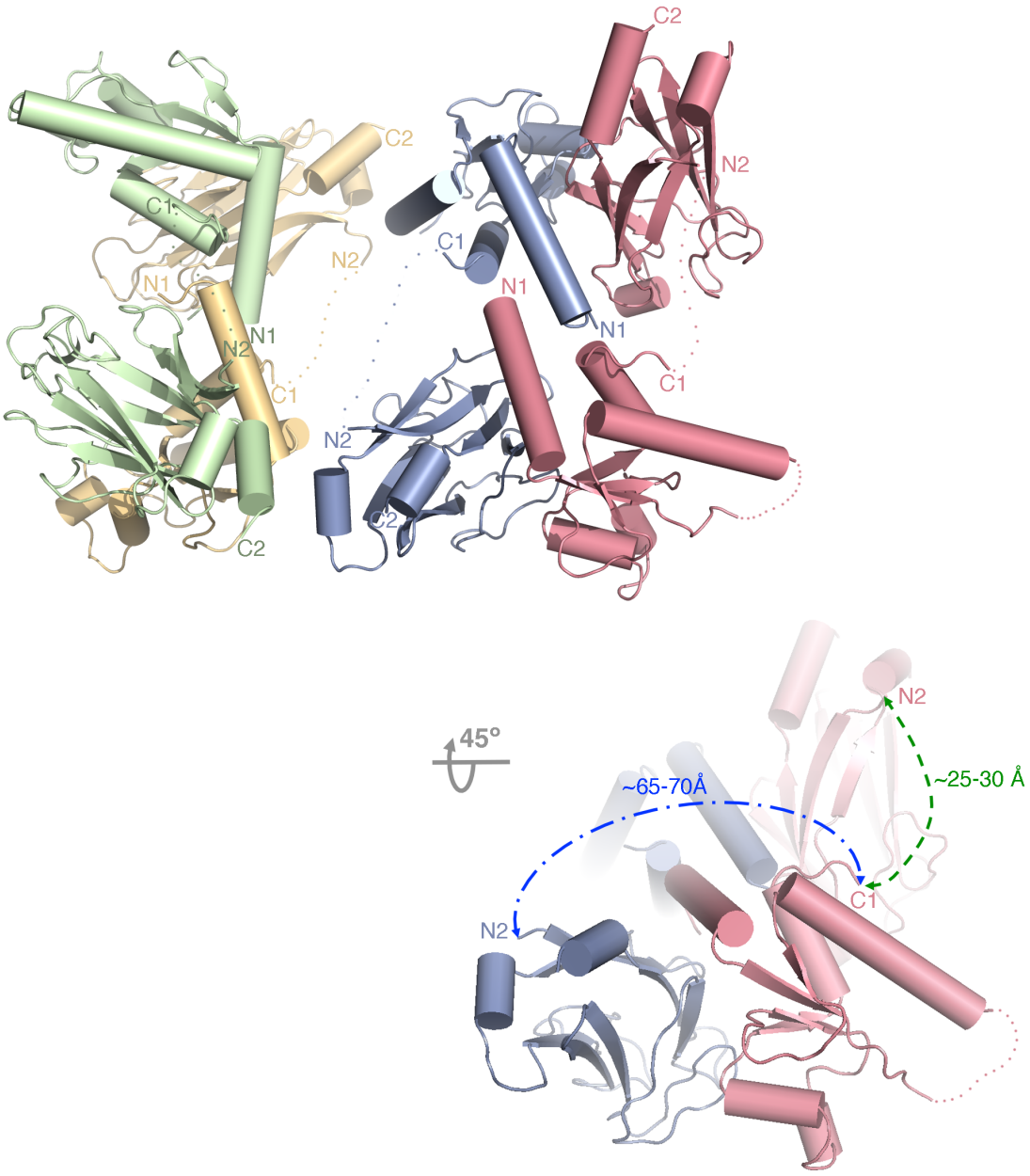

Figure 2.3. Crystal packing of the binary complex. The four molecules in the asymmetric unit are shown color-coded with the N- and C-termini of the HBRCT and FHA1 domains labeled. The bottom panel shows the approximate distances between the C-terminus of the HBRCT and the N-terminus of its closest FHA1 neighbors.

This results in almost twenty amino acids missing in each polypeptide chain. The distance between the last ordered residue of the HBRCT domain of Dbf4 and the first ordered residue in the closest FHA1 neighbors, the crystal packing contacts and the behavior in solution of the

chimeras, confirms that the Dbf4:Rad53 complex forms inter-molecularly. Importantly, the four complexes in the asymmetric unit had identical interfaces, indicating that the linkers did not constrain complex formation.

The FHA1 and HBRCT domains have identical architectures in the complex as in their unbound structures (**Figure 2.4A**).

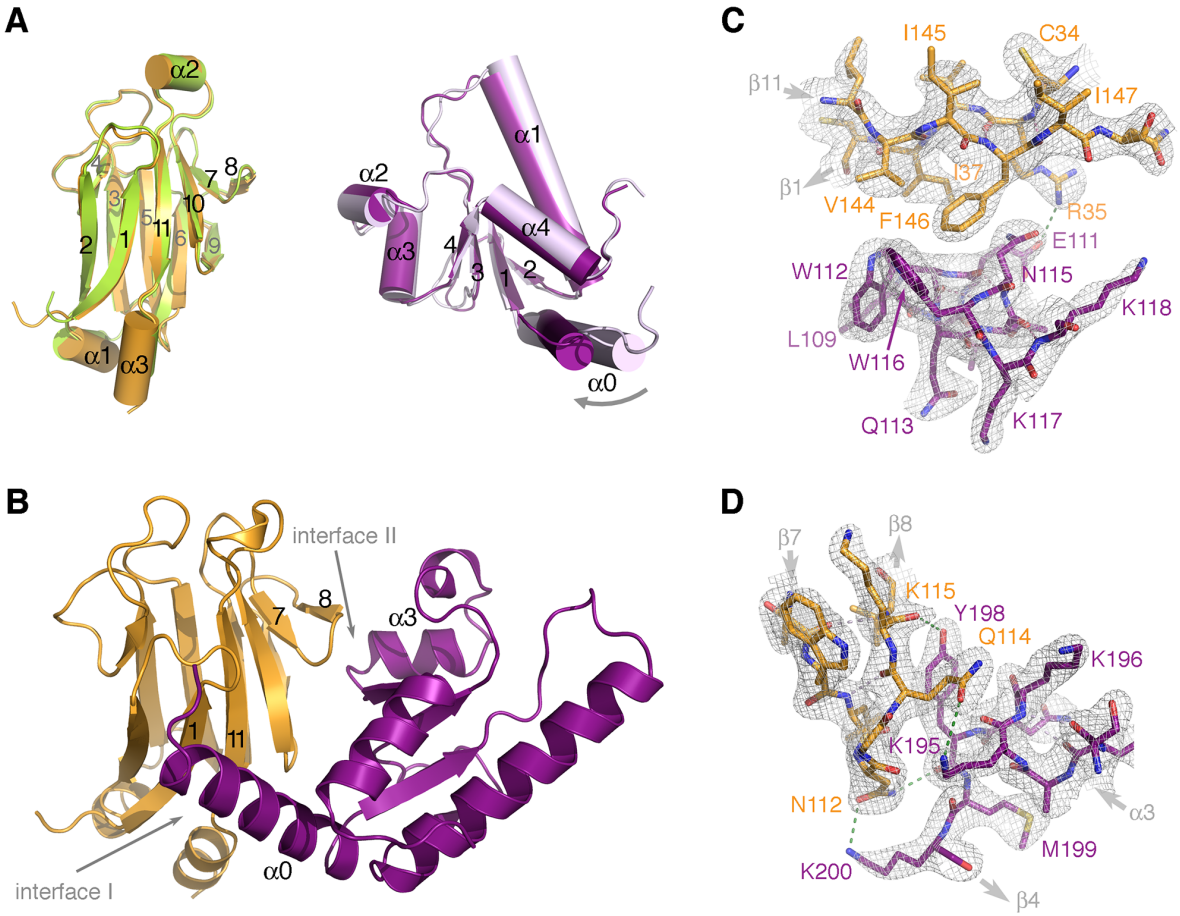

Figure 2.4. Structure of the Rad53(5)Dbf4 chimera. (A) Comparison of the structures of the FHA1 domain of Rad53 and the HBRCT domain of Dbf4 when crystallized on their own (PDB ID: 1G6G and 3QBZ) or forming a complex. The FHA1 domain is shown in green (1G6G) or gold (complex) and the HBRCT domain is shown in lilac (3QBZ) or purple (complex) with secondary structure elements labeled for clarity. **(B)** Ribbon diagram of the crystal structure of the Rad53:Dbf4 complex with Rad53 colored gold and Dbf4 colored purple. The interfaces mediating the complex, as well as the pThr-binding groove, are labeled. Detailed views of the interactions defining interface I **(C)** and interface II **(D)**. Rad53 and Dbf4 residues are shown as sticks colored as in **(B)** and labeled. Refined 2Fo-Fc electron density maps are shown as a grey mesh contoured at σ = 1.2. Hydrogen bonds are shown as dashed lines.

However, the helix α0 of the HBRCT domain swivels about twenty degrees upon complex formation (**Figure 2.4A**). In good agreement with our previous results showing that the pThr-binding pocket of the FHA1 domain does not mediate the interaction with Dbf4 (Matthews et al., 2014), the complex forms through the lateral surface of the FHA1 domain defined by the β2-β1-β11-β10-β7-β8 strands and the concave surface of the HBRCT domain (**Figure 2.4B**). This interface, however, is quite limited because Dbf4 only contacts two small regions on each side of the lateral surface of Rad53. On one side of the interface, the side chains of residues Arg35 (β1), Ile37 (β1), Val144 (β11) and Phe146 (β11) of Rad53 are cradled by the α0 helix of the HBRCT domain, specifically by residues Glu111, Trp112, Asn115 and Trp116, defining interface I (**Figure 2.4C**). On the other side, the loop containing residues Tyr198 and Lys200 of Dbf4 wraps around the β7/β8 loop of the FHA1 domain enabling the interaction between the amine group of Lys200 and Asn112 (β7) in Rad53 defining interface II (**Figure 2.4D**). Globally the two interfaces bury a mere 10% of the total accessible surface area of the FHA1 (755 out of 6,664 Å2) and the HBRCT (801 out of 7,799 Å2) domains, a value that is below the cutoff for specific interactions as judged using the PISA server (Krissinel and Henrick, 2007). This is not surprising in light of the dissociation constant estimated from the sedimentation velocity and NMR analysis (**Figure 2.1** and Matthews et al., 2014).

2.5.3 Rad53 and Dbf4 contribute asymmetrically to the interface of the complex

The residues of the FHA1 domain mediating the interaction with Dbf4 in the crystal structure of the Dbf4(5)Rad53 chimera are the same as those previously identified using NMR (Matthews et al., 2014). Our previous work, however, could not explain why multiple point

mutations on the FHA1 surface were required to abrogate complex formation (Matthews et al., 2014). These results were intriguing because Rad53 and Dbf4 interact with low affinity and, hence, we had not anticipated the need of multiple mutations to abrogate the interaction. Since the point mutations in Rad53 were designed based on sequence conservation, we had a better sampling of interface I than interface II (**Figure 2.4**). Therefore, we decided to dissect the contributions of both interfaces to the complex formation.

We generated single point mutations in the FHA1 domain affecting either interface I (Phe146Ala) or interface II (Asn112Ala), as well as a double point mutation affecting both interfaces (Asn112Ala/Phe146Ala). We then measured the ability of these variants to interact with full-length Dbf4 using a yeast two-hybrid assay. As we expected from our previous work, the FHA1-Asn112Ala had a mild, yet significant, binding defect (**Figure 2.5A** and **Figure 2.6A**).

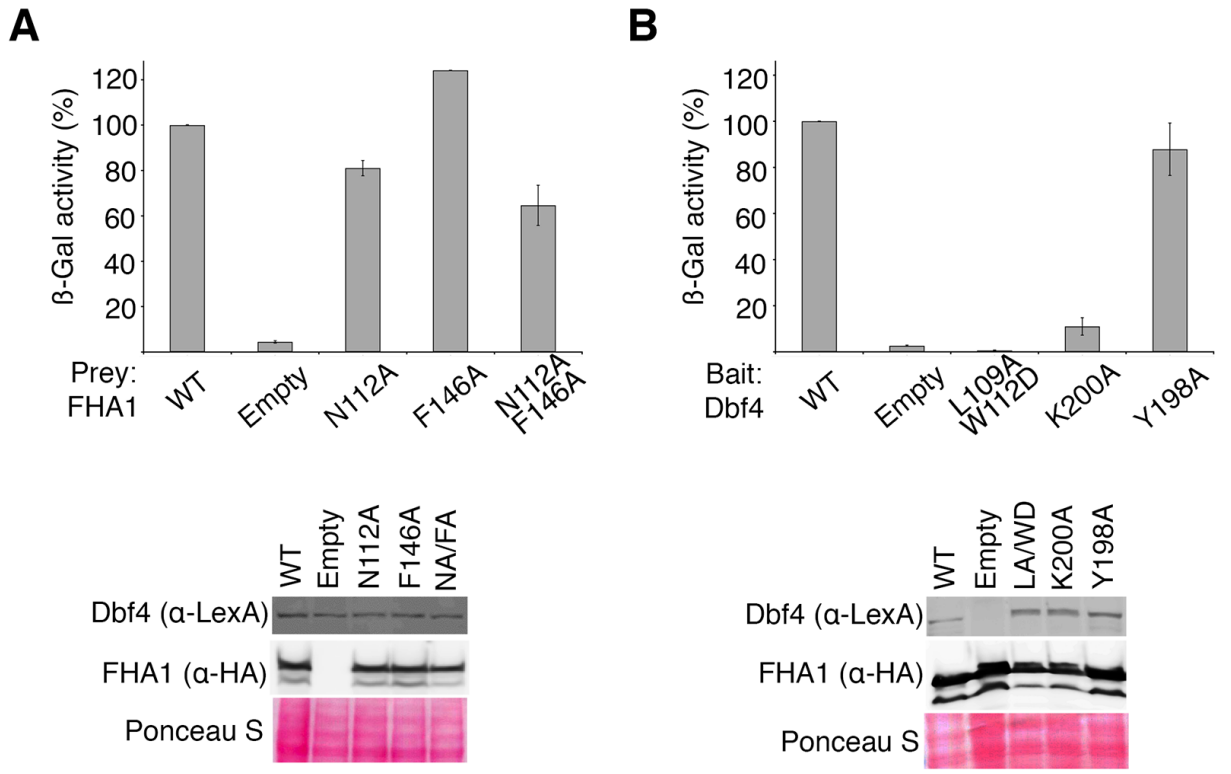

Figure 2.5. Two discrete interfaces contribute to the Rad53:Dbf4 interaction. (A) Yeast two-hybrid analysis using wild type Dbf4 as the bait and variants of the FHA1 domain of Rad53 as

the preys. **(B)** Yeast two-hybrid analysis using variants of Dbf4 as the baits and the wild type FHA1 domain of Rad53 as the prey. In each case, the interaction is shown as a percentage of β-galactosidase activity for the interaction between wild-type proteins and represents the mean of three independent measurements (error bars represent S.D). Bait and prey expression levels were analyzed by Western blotting and relative protein loading assessed by Ponceau S staining. See **Figure 2.6** for original gels/blots.

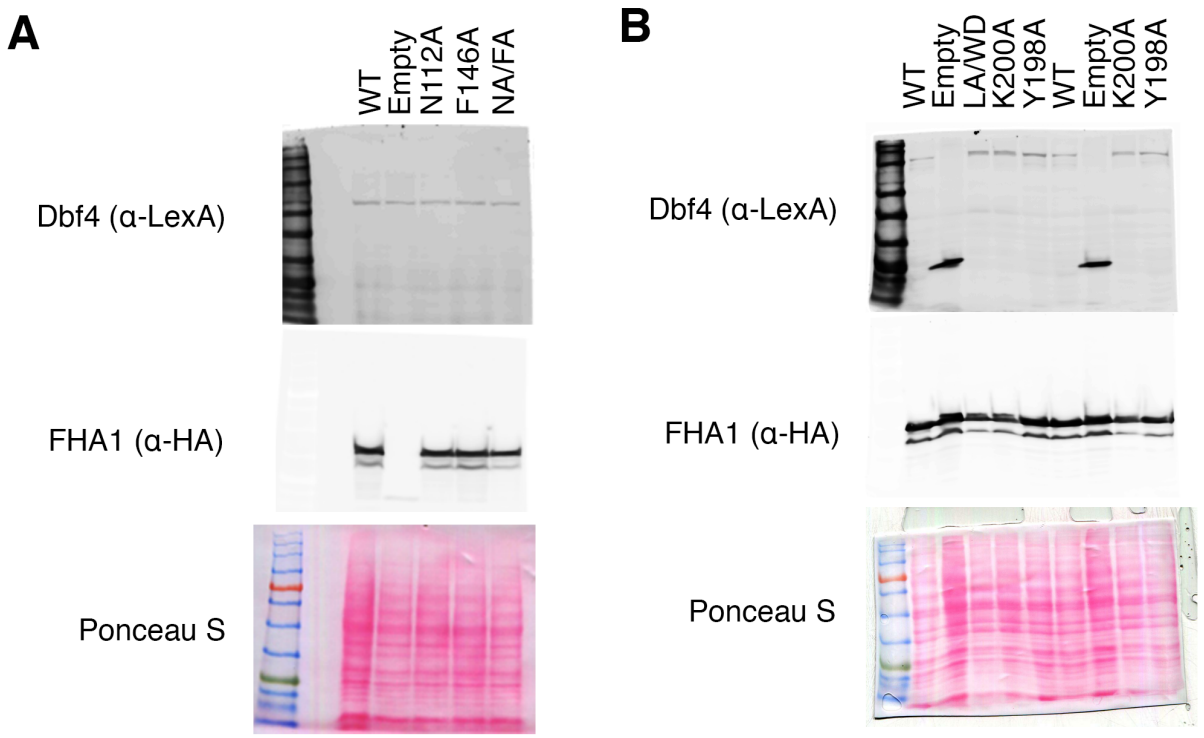

Figure 2.6. Original gels and blots for Figure 2.5. (A) To control for the two-hybrid bait and prey expression levels in **Figure 2.5**, whole cell extracts were prepared from transformants following prey induction and analyzed by Western blotting using rabbit anti-LexA antibody (bait) and mouse anti-HA monoclonal antibody (prey), along with Alexa Fluor 647-conjugated goat anti-rabbit and Alexa Fluor 488-conjugated goat anti-mouse secondary antibodies, respectively. Prior to detection, the membrane was stained with Ponceau S to assess relative protein loading. **(B)** Yeast two-hybrid analysis using variants of Dbf4 as the baits tested using the wild type FHA1 domain of Rad53 (left lanes) or the FHA1-F146A variant of Rad53 (right lanes) as the prey.

Conversely, the FHA1-Phe146Ala variant interacted with Dbf4 better than wild type FHA1 suggesting that a smaller side chain at this position may help accommodate helix α0 of the HBRCT. The combination of both changes had a stronger defect than the FHA1-Asn112Ala variant, but retained about half of the residual binding to Dbf4 (**Figure 2.5A**). When we

conducted the reciprocal experiment, the results were more drastic. The Dbf4-Leu109Ala/ Trp112Asp variant (affecting interface I) completely abrogated the interaction with the FHA1 domain, whereas variants affecting interface II had wide-ranging effects (**Figure 2.5B**). On our structure, the loop containing residues Tyr198 and Lys200 of Dbf4 wraps around the β7/β8 loop of the FHA1 domain enabling the interaction with Asn112 (β7) in Rad53 (**Figure 2.4D**). Mutation of Tyr198Ala did not affect the interaction with the FHA1 domain, whereas mutation of Lys200Ala disrupted binding to the FHA1 domain (**Figure 2.5B** and **Figure 2.6B**). Collectively these data suggest that hydrophobic contacts and the relative rigid body movement of helix α0 drive the interaction at interface I, whereas polar interactions determine interface II. Furthermore, they confirm that Dbf4 and Rad53 do not contribute equally to each interface, but both contact points are important for complex formation.

2.5.4 Rad53 interacts simultaneously with Dbf4 and a phosphorylated peptide

The combination of hydrophobic and polar interactions segregated in two different contact areas could provide the means to regulate complex formation upon binding of additional partners. Since a phosphorylated binding epitope is necessary for the interaction of Rad53 with DDK, we sought to determine the structure of the Dbf4(L)Rad53 chimera bound to a phosphorylated peptide (**Table 2.1**).

The fragment of Cdc7 encompassing residues 294–493 interacts reproducibly with the FHA1 domain of Rad53. This region only contains one TXXD motif (^{484}TDED^{487}) that is conserved and has high phosphorylation probability (Aucher et al., 2010). In good agreement, a Cdc7 variant encompassing a Thr484Ala point mutation reduces the interaction of Cdc7 with the FHA1 domain of Rad53 to 50% of wild type (**Figure 2.7A** and **Figure 2.8A**).

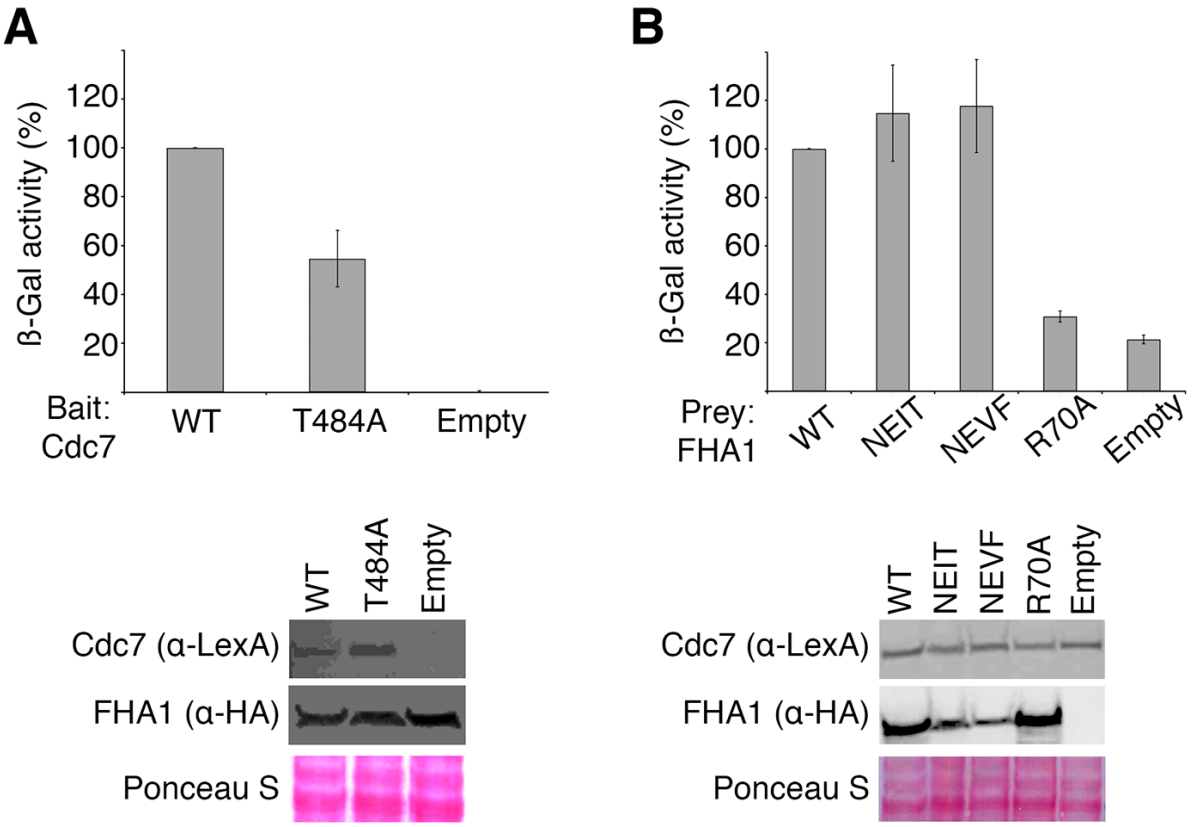

Figure 2.7. Rad53 recognizes a phosphorylated epitope in the Cdc7 subunit of the DDK complex. (A) Yeast two-hybrid analysis using either wild type or a T484A variant of Cdc7 as the baits and wild type FHA1 domain of Rad53 as the prey. **(B)** Yeast two-hybrid analysis using wild type Cdc7 as the bait and variants of the FHA1 domain of Rad53 as the prey. Bait and prey expression levels were analyzed by Western blotting and relative protein loading assessed by Ponceau S staining. See **Figure 2.7** for original gels/blots.

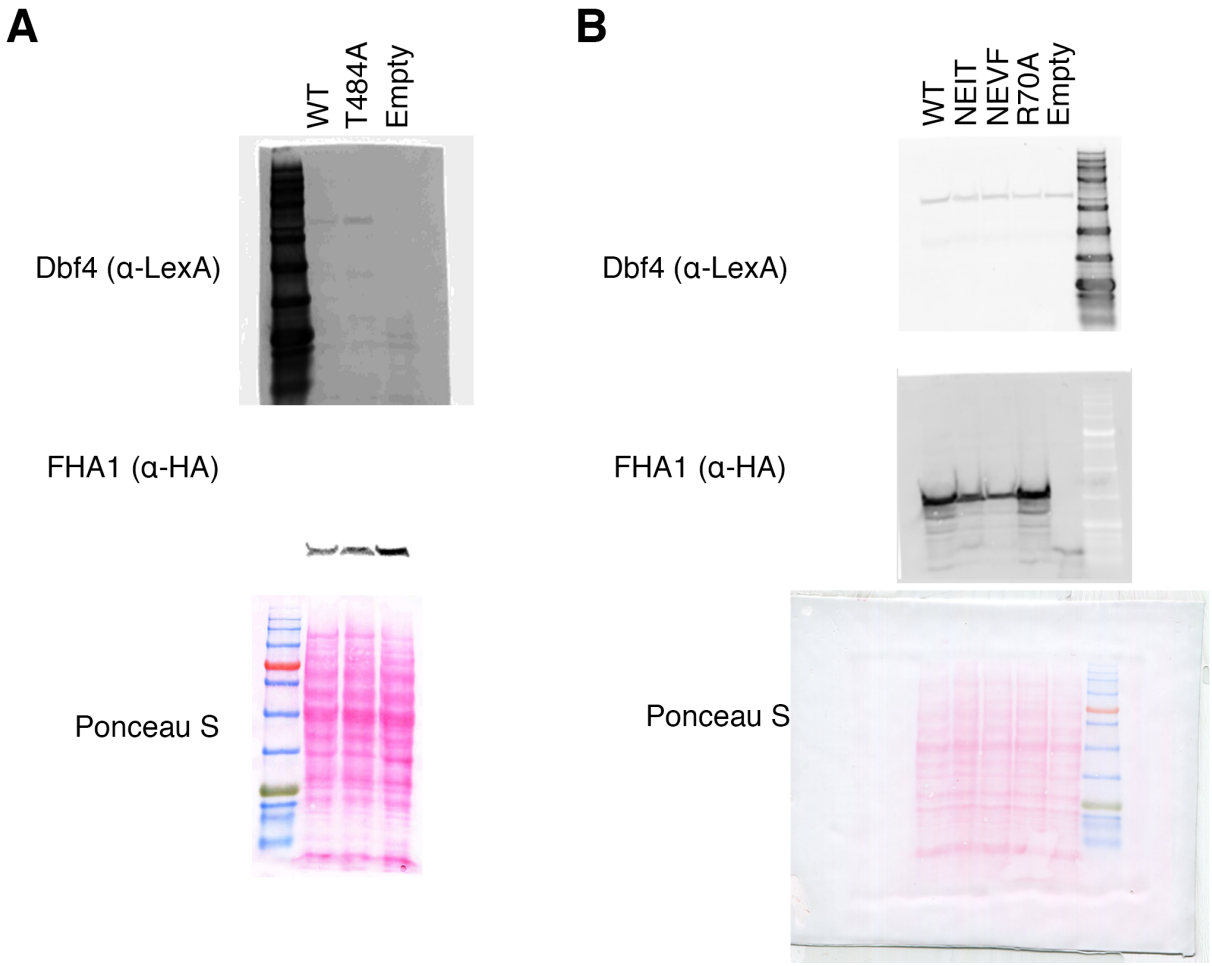

Figure 2.8. Original gels and blots for Figure 2.7. To control for the two-hybrid bait and prey expression levels in **Figure 2.7**, whole cell extracts were prepared from transformants following prey induction and analyzed by Western blotting using rabbit anti-LexA antibody (bait) and mouse anti-HA monoclonal antibody (prey), along with Alexa Fluor 647-conjugated goat anti-rabbit and Alexa Fluor 488-conjugated goat anti-mouse secondary antibodies, respectively. Prior to detection, the membrane was stained with Ponceau S to assess relative protein loading.

Reciprocally, a variant of Rad53-FHA1 unable to bind phosphorylated targets (FHA1-Arg70Ala) abrogates the interaction with Cdc7 (**Figure 2.7B** and **Figure 2.8B**). The differences between the Cdc7-Thr484Ala and Rad53-Arg70Ala variants suggest that Rad53 may be able to bind other epitopes in Cdc7 in the absence of Thr484.

Conversely, variants disrupting the Rad53:Dbf4 interface do not affect binding to Cdc7 (**Figure 2.7B** and Matthews et al., 2014). Since we have previously shown that a peptide derived

from this motif of Cdc7 (pPEP, ^{480}DGESpTDEDDVVS^{491}) binds to the FHA1 domain of Rad53 in a phosphorylation-dependent manner in vitro16, we used this peptide for subsequent crystallographic studies. Crystals of the ternary Rad53-Dbf4-Cdc7 complex grew in the P21 space group and diffracted X-rays beyond 2.3 Å resolution. We determined the structure by molecular replacement using the structures of the individual FHA1 and HBRCT domains as search models. The molecular replacement solution showed well-defined electron density for the two domains, as well as the main chain and most side chains of the phosphorylated peptide (**Figure 2.9A-B**).

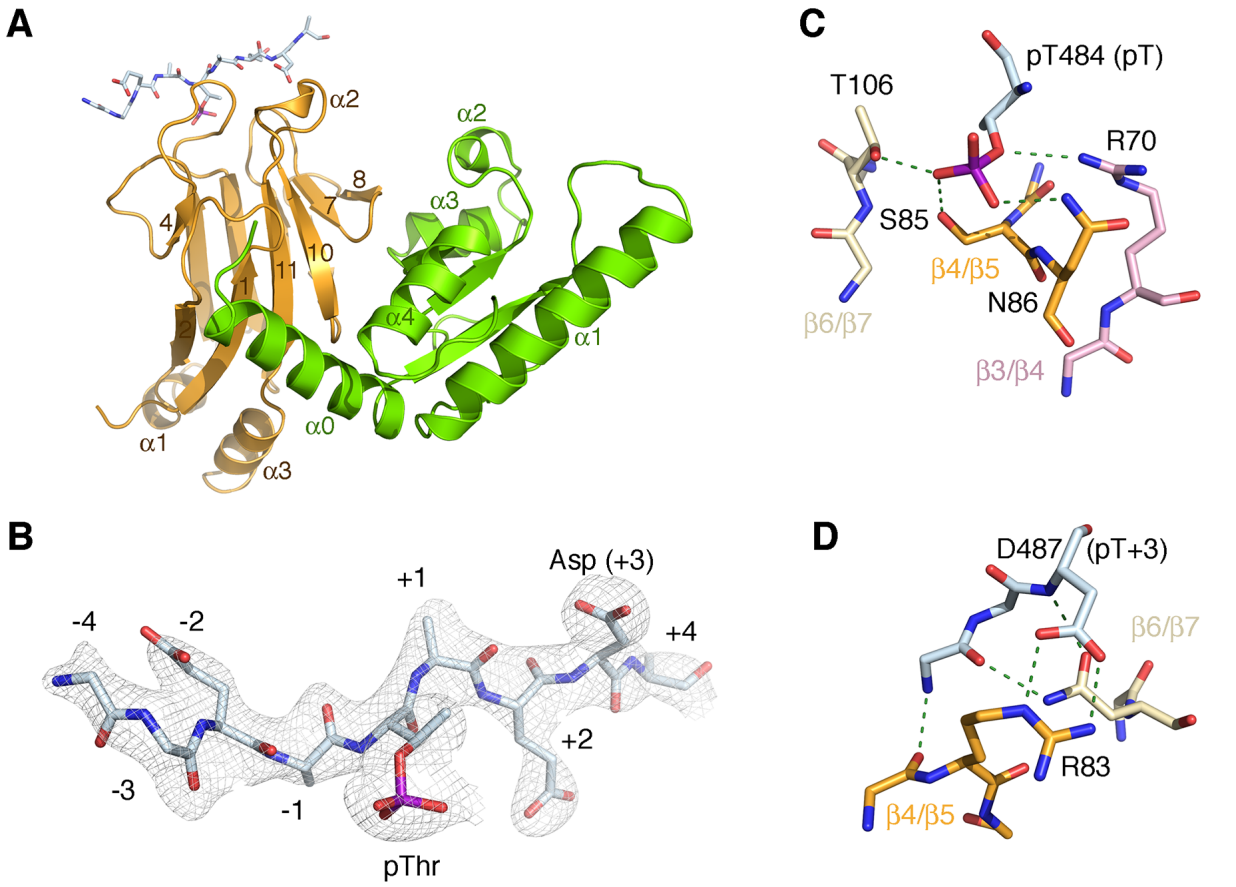

Figure 2.9. Structure of the Rad53:Dbf4:Cdc7 ternary complex. (A) Ribbon representation of the ternary complex on a similar view as in **Figure 2.4**. Rad53 (orange) and Dbf4 (green) are shown as ribbons. The Cdc7- derived peptide (pPEP) is shown as colored coded sticks. **(B)** Detail of the electron density map around the phosphorylated peptide shown as a grey mesh contoured at σ = 1.0. **(C)** Detail of the hydrogen-bond network stabilizing pThr484. **(D)** Detail of

the hydrogen bond interactions defining the specificity at the pT + 3 position of the peptide, as well as additional hydrogen bonds stabilizing the main chain of the peptide. Hydrogen bonds are shown as green dashed lines.

Similar to other structures of FHA1 domains bound to phosphorylated peptides, pPEP is bound at one end of the FHA1 domain and interacts with residues in the β3/β4, β4/β5 and β6/β7 loops (Durocher et al., 2000). The phosphate moiety of pThr484 is held in place through hydrogen bonds with Arg70, Ser85, Asn86 and Thr106 (**Figure 2.9C**); the pT + 3 aspartate residue (Asp487) is anchored through a salt bridge with Arg83; and the main chain of the intervening residues is further stabilized through hydrogen-bonds with the main chain carbonyl of Ser82 and the side chain of Asn107 (**Figure 2.9D**).

The Dbf4:Rad53 interface is similar, but not identical, to the binary complex. Superimposition of the FHA1 domains in the binary and ternary complexes revealed that the β1 strand and the β1/β2 loop were virtually invariant (r.m.s.d. <0.1 Å2). Therefore, we used this region of the FHA1 to superimpose and compare the two complexes. As expected, the loops defining the pThr-binding groove of the FHA1 domain had larger deviations (0.34 < r.m.s.d. <0.63), caused by the binding of the phosphopeptide. Conversely, the residues in FHA1 mediating the interaction with Dbf4 were barely affected by phosphopeptide binding (0.15 < r.m.s.d. <0.35).

Binding of the phosphopeptide, however, induces a small rigid body movement of Dbf4 around the two interfaces holding the complex (**Figure 2.10A**).

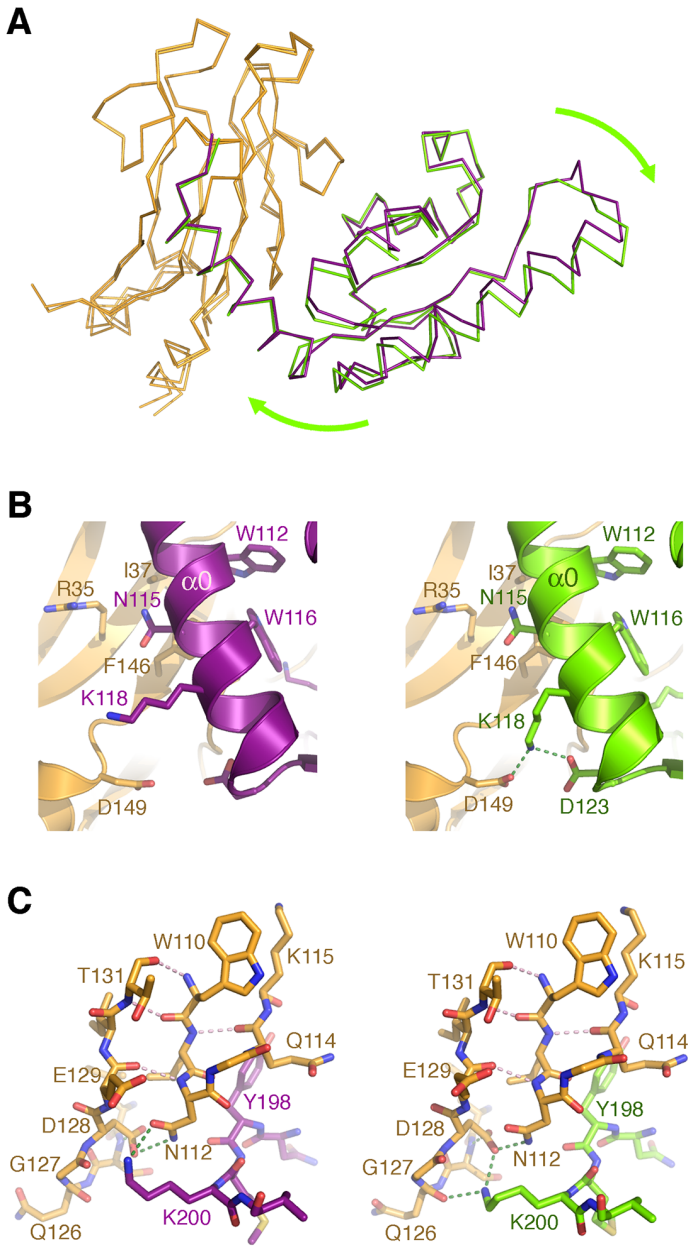

Figure 2.10. Peptide-binding induces a rigid body movement of Dbf4. (A) Opposite views of the Rad53:Dbf4:Cdc7 complex (gold-green) superimposed onto the Rad53:Dbf4 complex (gold-purple). **(B)** Detail of the different interactions of Lys118 (Dbf4) in in the binary (left) and ternary (right) complexes. **(C)** Detail of the conformational change imposed onto the side chain of Lys200 (Dbf4) upon binding of the phosphorylated peptide. In the binary complex (left-side panel) Lys200 interacts with Asn112 (Rad53), whereas in the ternary complex (right-side panel) interacts with Gly127 and Asp128 (Rad53).

The HBRCT domain seesaws pushing helix α1 away from the FHA1 domain while pulling the α0/β1 loop towards the FHA1 domain. This rotation is identical for both complexes in the

asymmetric unit and, though subtle, the movement is enough to reorganize some of the residues at both interfaces. Upon binding to pPEP, the side-chain of Lys118 (interface I in Dbf4) comes close to the side-chains of Asp123 (Dbf4) and Asp149 (Rad53) stabilizing the interaction of the C-terminus of helix α0 in Dbf4 with Rad53 (**Figure 2.10B**). On the ternary complex, Lys200 (interface II in Dbf4) is not hydrogen-bonded to Asn112. Binding to pPEP pushes the α3/β4 loop of Dbf4 closer to the β10 strand where the new conformation of Lys200 is stabilized through hydrogen bonds with Gln126 and Asp128 (**Figure 2.10C**).

2.5.5 Phosphopeptide binding modulates the Rad53:Dbf4 interaction

Given the subtle movement of the HBRCT domain, the analysis of the two interfaces did not show significant differences in the extension of the interface or solvation energy (**Table 2.2**).

Table 2.2. PISA analysis of the Dbf4:Rad53 and Dbf4:Rad53:Cdc7 complexes.

			Solvation Energy contribution			
Complex	Interface (Å^2)	ΔG solvation	Dbf4		Rad53	
		(kcal/mol)	Structure	Average gain	Structure	Average gain
Rad53:Dbf4	684-707	-3.5 – -5.0	-96.3	-3.75	-115	-1.5
Rad53:Dbf4:Cdc7	684	-3.8	-96.7	-2.2	-115.1	-1.3

Dbf4 had a minimal gain in solvation energy suggesting the HBRCT domain has more surface exposed residues in the ternary than the binary complexes (**Table 2.2**). The differences between the Dbf4:Rad53 interface in the binary and ternary complexes although subtle, could indicate that binding of the phos- phopeptide allosterically regulates the interaction. However, these differences could also be due to crystal packing environment or the different linker length of the fusions.

The asymmetric unit of the binary complex included four Dbf4(5)Rad53 molecules defining four Dbf4:Rad53 interfaces, whereas that of the ternary complex included two Dbf4(0)Rad53 molecules defining two Dbf4:Rad53 interfaces. Superimposition of each ternary complex onto any of the binary complexes revealed that the peptide moiety could only be accommodated in half of the complexes, explaining why crystals of the ternary complex grew in different conditions. However, no other crystal contacts mediated by the FHA1 or HBRCT domains enhanced or prevented the movement in Dbf4. Therefore, phosphopeptide binding rather than crystal packing is the likely driving force of the movement.

We have previously shown that losses of cross-peak intensity in the HSQC spectrum of ^{15}N-labeled Rad53 upon binding Dbf4 serve as sensitive reporters to map the interface of the complex (Matthews et al., 2014). If phosphopeptide binding to the FHA1 domain weakened the interaction, we would expect an enhancement of cross-peak intensities for the Rad53 residues at the interface with Dbf4. Despite maintaining similar conformations in the binary and ternary crystal structures, two residues on interface I (Ile37 and Phe146) showed increased cross-peak intensities in the ternary complex (**Figure 2.11**).

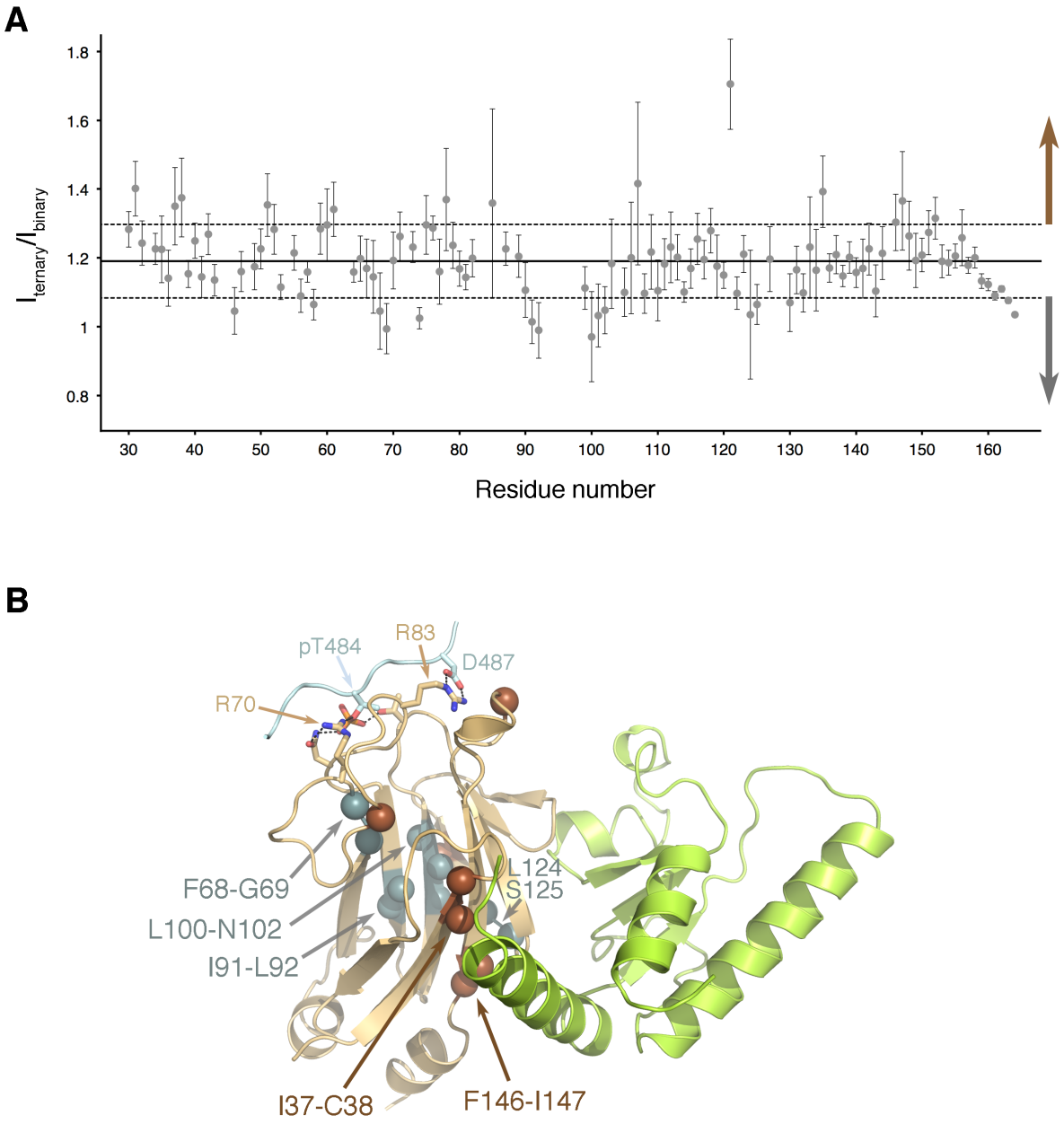

Figure 2.11. ^{15}N-HSQC intensity changes confirm weakening of the Rad53:Dbf4 complex caused by binding of the phosphopeptide to Rad53. (A) Plot of residue specific ratios between cross-peak intensities in the ^{15}N-HSQC spectra of the binary (Rad53:Dbf4) and ternary (Rad53:Dbf4:Cdc7) complexes. The solid and dashed horizontal lines represent the average intensity ratio +/- one standard deviation. **(A)** Residues with values either greater or lower than the average +/- one standard deviation were plotted onto the X-ray structure of the ternary complex as brown or blue spheres, respectively **(B)**.

These observations are consistent with the idea that phosphopeptide binding to the FHA1 domain weakens the Rad53:Dbf4 interaction. However, some surface residues beyond the complex

interface also show variations of cross-peak intensities between the two complexes (**Figure 2.11**). Interestingly, a number of residues within the hydrophobic core of Rad53 display decreased cross-peak intensities upon phosphopeptide binding (**Figure 2.11**). These residues form a continuous net- work from the β3/β4 loop (Phe68-Gly69) to the β9/β10 loop (Leu124-Ser125) that propagates across the β-sheet defined by strands β4-β3-β5-β6-β9. Such intensity losses typically reflect changes in internal dynamics and could possibly reveal an allosteric network to report the presence of the phosphorylated peptide to the Rad53:Dbf4 interface. While these changes could explain how the two inputs of the logic gate sense each other to elicit a single output, the idea awaits further validation.

2.6 DISCUSSION

Yeast genetics has delineated the factors and hierarchy of interactions involved in the DNA damage response, but the molecular detail has remained elusive because most of the interactions driving the checkpoint response are transient. This problem is aggravated for 'AND' logic gates because they recognize two or more inputs to produce a single signal (Hasty et al., 2003; Lee et al., 2008; Zhang and Durocher, 2008), but disruption of any of the inputs disrupts the output leading to technically biased interpretations. We have found that the FHA1 domain of Rad53 functions as a 'AND' logic gate for its interaction with DDK, thereby explaining more than a decade of partly conflicting results (Duncker et al., 2002; Chen et al., 2013; Matthews et al., 2012; Matthews et al., 2014; Matthews and Guarné, 2013). The crystal structures of the FHA1 domain of Rad53 bound to one (HBRCT) or both (HBRCT and phosphoepitope) partners in the DDK complex presented here unveil how this logic gate simultaneously recognizes two

inputs and provide the first image of an FHA domain recognizing a binding partner through a non-canonical interface.

The interaction of Rad53 with the DDK complex is reminiscent of the interaction between Chk2 (the human ortholog of Rad53) and BRCA1, where the tandem BRCT repeat of BRCA1 simultaneously recognizes two distal surfaces in the FHA domain of Chk2 (Li et al., 2002). In the BRCA1:Chk2 complex, the interaction involves the pThr-binding site and a conserved hydrophobic patch on one of the lateral surfaces of the FHA domain. Disruption of either contact point prevents the interaction, and mutation of the hydrophobic patch has been linked to Li-Fraumeni syndrome (Li et al., 2002). However, Dbf4 and BRCA1 do not interact with the same lateral surface of the FHA domains of Rad53 and Chk2, exposing the extreme plasticity of FHA domains to enhance binding specificity.

Both Rad53 and Chk2 dimerize in solution and this is important to promote kinase activation by trans-autophosphorylation (Cai et al., 2009; Wybenga-Groot et al., 2014). Dimerization is triggered by damage-induced phosphorylation of a threonine within the SCD of the kinase. The dimers associate in a face-to-face configuration that promotes the swap of the activation loops for phosphorylation in trans (Cai et al., 2009; Schwarz et al., 2003). In the crystal structure of Chk2, one of the lateral surfaces of the FHA domain also contributes to the dimerization interface and, in fact, the requirement of a phosphorylated threonine residue is bypassed by protein overexpression indicating that the kinase and FHA mediated interactions suffice to stabilize the dimer (Cai et al., 2009). The surface of the FHA domain involved in the dimerization interface is the same as that of Rad53 mediating the interaction with Dbf4 (**Figure 2.12**).

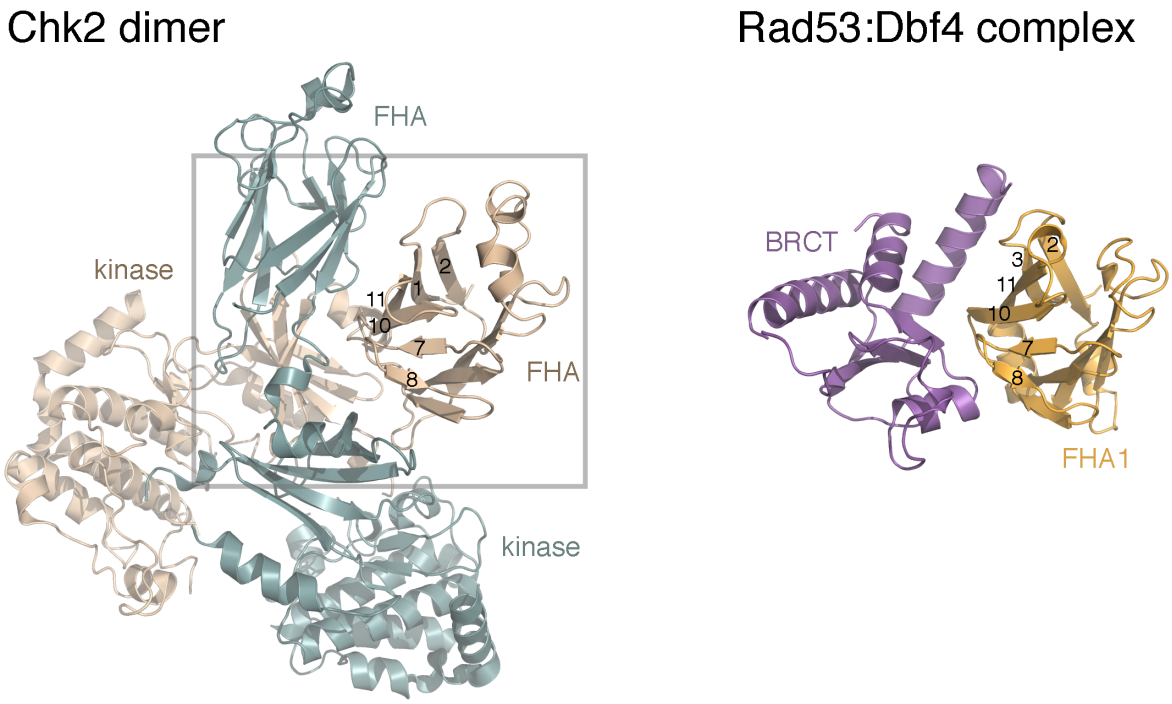

Figure 2.12. Interaction surfaces mediating dimerization of Chk2 and the Rad53:Dbf4 complex. Ribbon diagram of the Chk2 crystal structure (left) showing with the two protomers of the dimer shown in steel blue and tan. The lateral surface of the FHA domain interacts with the FHA domain and the N-terminal lobe of the kinase domain on the second protomer. Comparison with the crystal structure of the Rad53:Dbf4 complex shows that the FHA1 domain of Rad53 (light orange) uses the equivalent lateral surface to interact with the BRCT domain of Dbf4 (purple).

However, the two kinases transition to monomers to phosphorylate downstream targets (Ahn et al., 2002; Ahn and Prives, 2002; Cai et al., 2009) implying that dimer formation and subsequent dissociation may determine the hierarchy of checkpoint events.

In contrast to the Chk2 dimer, where the entire lateral face of the FHA domain contributes to dimer formation, Dbf4 only contacts two points on the lateral face of the FHA1 domain. Interestingly, mutations on the two surfaces are not reciprocal indicating that each partner contributes asymmetrically to the two small interfaces mediating the interaction with Dbf4, thereby suggesting a sophisticated way to gain binding-specificity without strengthening the interaction. To our knowledge, this is the first crystal structure of an FHA domain bound to a

binding partner through a non-canonical interface and lays the foundation to study how FHA domains can exploit canonical and non-canonical interactions to increase binding specificity of low-affinity interactions and, in turn, extend the functional repertoire of this phosphoepitope binding module.

2.7 ACKNOWLEDGEMENTS

We are grateful to Dr. Pawel Grochulski and Michel Fodje at the Canadian Light Source for assistance on data collection and processing of the ternary complex; as well as Aaron G. Robertson for assistance generating the Cdc7-T484A mutant. This work was funded by the Canadian Institutes of Health Research (MOP-67189 to AG); the Natural Sciences and Engineering Research Council (RGPIN 238392 to BPD); of and the Intramural Research Program of the NIH, the National Institute of Diabetes and Digestive and Kidney Diseases (to RG). A.W.A. was funded in part by an Ontario Graduate Scholarship.

CHAPTER THREE

THE CRYSTAL STRUCTURE OF THE CDC5-DBF4 COMPLEX PROVIDES INSIGHT ON POLO-BOX DOMAIN SUBSTRATE RECOGNITION

Manuscript in preparation

Authors: Ahmad W. Almawi, Stephen Boulton, Laurence Langlois-Lemay, Damien D'Amours, Giuseppe Melacini, and Alba Guarné.

3.1 CONNECTING TEXT

Chapter 3 presents the first structural work regarding the polo-box domain of Cdc5 with the goal of identifying the Dbf4-binding surface. In addition to solving the structure of the Dbf4-Cdc5 complex, we also characterized the interaction between Cdc5 and a phosphorylated substrate. Collectively, these structures provide insight on how the polo-box domain utilizes different surfaces to engage its substrates. Stephen Boulton, a graduate student from Dr. Giuseppe Melacini's laboratory at McMaster University, performed the NMR experiments. Laurence Langlois-Lemay, a graduate student from Dr. D'Amours's laboratory at the University of Ottawa, conducted the *in vivo* characterization of Cdc5 variants. This work is in progress and will be included in the final version of the manuscript. With the guidance of my supervisor, Dr. Alba Guarné, I conducted the experiments, interpreted the data, prepared the figures, and wrote the chapter.

3.2 ABSTRACT

Cdc5 is the sole Polo-like kinase (Plk) in budding yeast, and functions in mitotic entry, chromosomal segregation, mitotic exit, and cytokinesis. Similar to other Plks, Cdc5 contains a kinase domain, followed by a polo-box domain which recognizes phosphorylated epitopes in binding partners. The polo-box domain enables Cdc5 to target its substrates in a timely and sequential manner, to ensure correct mitotic progression. The polo-box domain of Cdc5 is also capable of mediating phosphorylation-independent interactions through an ill-defined process. One example is the interaction with the regulatory subunit of the Dbf4-dependent kinase (DDK) complex. Binding of Dbf4 to Cdc5 inhibits it from recognizing its substrates as seen during the anaphase spindle checkpoint. Alternatively, Dbf4 cooperates with Cdc5 to target substrates involved in DNA joint molecule resolution. As such, it remains unclear as to how Dbf4 binding could affect Cdc5 substrate recognition. In this study, we present the crystal structures of the polo-box domain of Cdc5, and in complex with peptides derived from Dbf4 and a canonical phosphorylated substrate, Spc72^{P}. The deduced structures revealed that the two peptides bind opposite surfaces on the polo-box domain. Also, both Cdc5 substrates interact simultaneously and non-competitively with Cdc5. Comparing obtained structures with that of zebrafish Plk1 suggests that the Dbf4-binding surface overlaps with the interface of the polo-box and kinase domains. As such, we propose a mechanism to explain how Dbf4 binding could regulate Cdc5 kinase activity.

3.3 INTRODUCTION

Cells strictly regulate mitotic processes to ensure faithful propagation of their newly replicated genetic content. Mitotic defects, including sister chromatid mis-alignment or spindle

pole damage, lead to improper chromosome segregation, which if left unchecked, promotes polyploidy and aneuploidy (Kops et al., 2005; Pierera and Schiebel, 2005; Seong et al.,). Cells respond to these insults by activating signaling cascades known as checkpoints to inhibit mitosis (Hartwell and Weinert, 1989). Cells with deregulated mitotic checkpoints gain a proliferative advantage, a major hallmark of carcinogenesis (Hanahan and Weinberg, 2000; Kops et al., 2005).

Mitotic events are controlled by a conserved family of kinases known as Polo-like kinases (Plks) (Golsteyn et al., 1995). Human Plk1 functions in mitotic entry, spindle pole dynamics, and cytokinesis (Archambault and Glover, 2009; Golsteyn et al., 1995; Seong et al., 2002). Due to its proliferative capacity, Plk1 is up-regulated in a variety of human tumors, making it an attractive anti-cancer target (Lee et al., 2005; Strebhardt, 2010.). However, inhibition of Plk1 activity is difficult to achieve, because Plk1 shares a high degree of structural similarity with other cellular kinases. In fact, some of the most selective drugs against Plk1 also inhibit Plk2 and Plk3 with comparable potencies (Rudolph et al., 2009; Steegmaier et al., 2007;). Since Plk2 and Plk3 function to protect the cell against DNA damage, they should not be inhibited. A more specific and promising approach to target Plk1 is through interference with cellular interactions that regulate its kinase activity.

Most of our understanding of Plk1 function stems from Cdc5, the sole Plk in budding yeast (Botchkarev and Haber, 2018; Charles et al., 1998; Hartwell et al., 1973; Snead et al., 2007). Cdc5 regulates mitotic exit by promoting nucleolar release of Cdc14 phosphatase, leading to downstream activation of the anaphase-promoting complex, cyclin destruction, and mitotic spindle disassembly (Archambault and Glover, 2009; D'Amours et al., 2004; Hu et al, 2001;

Pereira and Schiebel, 2004; Rahal and Amon, 2008; Stegmeier and Amon, 2004). Cdc5 is inhibited during spindle position checkpoint (SPOC), which operates in response to mispositioned or damaged spindle poles (Miller et al., 2009). Dbf4-dependent kinase (DDK) complex is one of the components of SPOC (Miller et al., 2009). DDK is a heterodimer of Cdc7 kinase and its regulatory subunit, Dbf4, and functions at multiple steps of the cell cycle (Dowell et al., 1994; Jackson et al., 1993; Matos et al., 2008; Miller et al., 2009). Dbf4 binds to Cdc5 and promotes Cdc7-mediated phosphorylation of Cdc5 during the SPOC response (Miller et al., 2009). In turn, phosphorylated Cdc5 fails to recognize its substrates in the mitotic exit pathway (Chen and Weinreich, 2010; Miller et al., 2009). DDK is also involved with Cdc5 in other cellular processes (Argunhan et al., 2017; Princz et al., 2017). For example, Cdc5 collaborates with DDK and CDK1 to phosphorylate the structure-specific nuclease Mus81-Mms4, which in turn activates the nuclease to resolve joint DNA molecules at the onset of mitosis (Princz et al., 2017). Full activation of Mus81-Mms4 depends on the Cdc5-Dbf4 interaction, indicating that DDK and Cdc5 work cooperatively during this process (Princz et al., 2017).

Similar to other Plks, Cdc5 contains a serine/threonine kinase domain, followed by a polo-box domain which mediates substrate recognition and sub-cellular localization (Elia et al., 2003; Lee et al., 2005; Smerdon and Yaffe, 2003). Two polo boxes comprise the polo-box domain and define a conserved phosphoserine/threonine (pS/T)-binding pocket. The pS/T-binding pocket recognize substrates harboring an X-S-(pS/T)-(P/X) consensus phosphorylated site (Elia et al., 2003; Cheng et al., 2003) but can also mediate phosphorylation-independent interactions (Elia et al., 2003; Cheng et al., 2003; Xu et al., 2013). For example, Drosophila microtubule-associated protein 205 (Map205) interacts with the polo-box domain of Polo

(Drosophila homolog of human Plk1) at the pS/T-binding pocket (Archambault et al., 2008; Xu et al., 2013), despite lacking the consensus phosphorylated motif.

Similar to Map205, Dbf4 lacks the phosphorylated consensus sequence and utilizes a unique polo-interacting region (^{83}RSIEGA^{88}) to recognize the polo-box domain of Cdc5 (Chen and Weinreich, 2010). Mutation of the critical residues at the pS/T-binding pocket does not affect the Dbf4-Cdc5 interaction, indicating that Dbf4 binds a different surface of the polo-box domain (Miller et al., 2009; Cheng et al., 2003). However, the interface of the Dbf4-Cdc5 complex is unknown. To better understand how Dbf4 binds and modulates the function of Cdc5, we solved the crystal structure of the polo-box domain of Cdc5 and in complex with peptides derived from the polo-interacting region of Dbf4, and the phosphorylated motif from a bonafide phosphorylated substrate (the spindle pole body protein Spc72). We showed that Dbf4 and Spc72^{P} recognize opposite surfaces of the polo-box domain, and bind simultaneously but non-competitively to Cdc5. Based on structural, biochemical and genetic data, we propose a model to explain the role of Dbf4 in modulating Cdc5 activity.

3.4 METHODS

3.4.1 Cloning and expression

The polo-box domain of Cdc5 (residues 418-705) was sub-cloned into the pPROEX HTa vector (pAG 8531) (Invitrogen Life Technologies). The polo-box domain A567W variant was generated by site-directed mutagenesis using the Q5 kit (New England Biolabs), and clones were confirmed by DNA sequencing (Nanuq, McGill University, Génome Québec Innovation Centre). Plasmids were transformed in BL21 (DE3) cells containing a plasmid encoding for rare tRNAs.

Cultures were grown in Luria-Bertani media to OD600 ~ 0.6, induced by addition of 2 mM isopropyl ß-D-1-thiogalactopyranoside, and incubated overnight at 16 ˚C with orbital agitation.

3.4.2 Protein Purification

Cell pellets were resuspended in 50 mM Tris-HCl pH 7.0, 500 mM NaCl, 1.4 mM 2-mercaptoethanol, 5 % glycerol and lysed by sonication. Lysates were cleared by centrifugation at 39,000 g and the supernatants loaded onto a HiTrap nickel-chelating HP column (GE Healthcare). Histidine-tagged Cdc5 was eluted at 240 mM imidazole. The histidine tag was cleaved by tobacco etch virus (TEV) protease, and tagless Cdc5 was further purified by affinity and size-exclusion (Superdex 75 (10/300) GL, GE Healthcare) chromatography. Purified Cdc5 was kept in storage buffer (50 mM Tris-HCl pH 7.0, 200 mM NaCl, 1.4 mM 2-mercaptoethanol, 5 % glycerol). Protein concentration was determined using the Beer-Lambert equation with an extinction coefficient of 38,850 M-1cm-1 for Cdc5, and 44,350 M-1cm-1 for Cdc5-A567W. Purified proteins were analyzed using dynamic light scattering to assess sample homogeneity.

3.4.3 Crystallization, structure determination, and refinement

Crystals of Cdc5 grew in 100 mM Bis-Tris Propane pH 8.5, 20 % PEG 3350 (v/v), 200 mM sodium formate, and cryoprotected with 8 % glycerol. A complete data set was collected at the 08B1 beamline at the Canadian Light Source, and data were processed and scaled using XDS (Kabsch, 2010.) (Table 1). The Cdc5 structure was determined by molecular replacement using a fragment of the polo-box domain of human Plk1 (PDB ID: 1Q4O) as the search model. The initial model building was done using auto-build in PHENIX (Adams et al., 2010.). The initial model was improved by iterative cycles of model building in Coot and refinement in PHENIX (Adams et al., 2010; Emsley et al., 2010) (**Table 3.1**).

Table 3.1. Data collection and refinement statistics.

	Cdc5	Cdc5-Spc72	Cdc5-Dbf4
Data Collection			
Beamline	08B1; CLS	17-ID; APS	17-ID; APS
Wavelength (Å)	1.2834	0.9795	0.9762
Space group	P 21 21 21	P 1 21 1	P 65
Cell dimensions a, b, c α, β, γ	 51.1, 65.9, 96.1 90, 90, 90	 51.5, 68.2, 86.2 90, 102.59, 90	 135.1, 135.1, 75.2 90, 90, 120
Resolution (Å)*	48.1-1.8 (1.86-1.8)	42.1-2.7 (2.78-2.7)	67.5-3.4 (3.49-3.4)
Reflections (total/unique)	235,329/58,352	93,716/16,158	205,926/10,923
Wilson B factor	22.7	46.8	129.7
Completeness (%)	100 (99.9)	99.9 (99.4)	100 (99.9)
CC1/2 (%)	99.7 (35.8)	99 (31.7)	99.8 (31.5)
I/σ(I)	12.58 (0.96)	9.1 (1.56)	10.92 (1.74)
Redundancy	4 (3.8)	5.8 (4.9)	18.9 (18.8)
Refinement			
Resolution (Å)	48.1-1.8	42.1-2.7	67.5-3.4
Completeness (%)	100	99.9	100
R_{work}/R_{free} (%)	18.06/19.69	21.36/25.73	24.33/28.3
Ramachandran plot (%) Favoured Outliers	 98.3 0	 93.6 0.2	 93 0.4
rmsd in bonds (Å)	0.004	0.004	0.004
rmsd in angles (°)	0.885	0.805	0.711
Mean B values (Å^2)	31.4	62.9	193.4

***** *Data in the highest resolution shell is shown in parentheses.*

Peptides derived from Spc72 (^{227}SLA**QSpSP**AGSQ^{237}) and the polo-interacting region of Dbf4 (^{76}RARIERA**RSIEGA**VQVSKGTG^{96}) were purchased from GenScript and resuspended in storage buffer. The polo-box domain of Cdc5 (5 mg/mL) was mixed with Spc72^{P} at a 1:15 (protein:peptide) molar ratio and incubated overnight at 4 ˚C before crystallization trials. Crystals of Cdc5-Spc72^{P} grew in 100 mM Bis-Tris pH 6.5, 25 % PEG3350 (v/v), 200 mM $MgCl_2$, 100 mM CsCl, and were cryoprotected with 10 % glycerol. Cdc5 (5 mg/mL) was mixed with Dbf4 at a 1:10 (protein:peptide) molar ratio and incubated overnight at 4 ˚C before crystallization trials. Crystals of Cdc5-Dbf4 grew in 20 % PEG 3350, 200 mM sodium potassium tartrate, 10 mM trimethylamine hydrochloride, and were cryoprotected with 14 % glycerol. Complete data sets of the Cdc5-Spc72^{P} and Cdc5-Dbf4 complexes were collected at the 17-ID beamline at Advanced Photon Source (APS) (Argon National Laboratory) and processed using MOSFLM (Battye et al., 2011) (**Table 3.1**). The structures were determined by molecular replacement using the polo-box domain of Cdc5 as the search model. Initial models of the Cdc5-Spc72^{P} and Cdc5-Dbf4 structures were further improved by iterative cycles of manual model building in Coot and refinement in PHENIX (Adams et al., 2010; Emsley et al., 2010.) (**Table 3.1**). Figures showing molecular structures were generated using PyMOL (DeLano, 2002).

3.4.4 Isothermal titration calorimetry (ITC) experiments

The Cdc5-Spc72^{P} and Cdc5-Dbf4 interactions were analyzed using the Nano ITC instrument (TA Instruments), while the interaction between Cdc5-A567W and Dbf4 was analyzed using the MicroCaliTC200 instrument (Malvern Instruments Inc.). Peptides derived from Dbf4 (^{73}EKKRARIERA**RSIEGA**VQVSKGTG^{96}), Spc72

(^{222}DKEEFLSLA**<u>QSpSP</u>**AGSQ^{237}), and a non-phosphorylated version of Spc72 were purchased from GenScript and resuspended in storage buffer supplemented with 20 mM EDTA pH 8.0. The experiments run on the Nano ITC instrument were designed by filling the chamber cell with 30 μM protein (in buffer C) and the injection syringe with 280 μM peptide. The ternary complex (Dbf4-Cdc5-Spc72^{P}) was analyzed by filling the chamber cell with either Cdc5-Dbf4 or Cdc5-Spc72^{P} (at a 1:4 molar ratio of protein:peptide), and the injection syringe with 280 μM peptide. Data were processed using the NanoAnalyze program (TA Instruments). The interaction between Cdc5-A567W and Dbf4 was analyzed by filling the chamber cell with 45 μM protein and the injection syringe with 1.2 mM peptide. Data were processed using the MicroCaliTC200 program (Malvern Instruments Inc.).

3.4.5 NMR Saturation Transfer Difference (STD) analysis

All samples were either dissolved or buffer exchanged into 20 mM NaPO4 pH 7.0, 200 mM NaCl, 20 mM EDTA pH 8.0, 1.4 mM 2-mercaptoethanol, >99% D2O. The concentration of Cdc5 was kept at 20 μM while Dbf4 (^{73}EKKRARIERA**<u>RSIEGA</u>**VQVSKGTG^{96}) and Spc72^{P} (^{222}DKEEFLSLA**<u>QSpSP</u>**AGSQ^{237}) were at 500 μM unless otherwise specified. STD experiments were acquired at 298 K on a Bruker Avance 850 MHz spectrometer with TXI probe. The spectra were acquired with 32k points, a spectral width of 16 ppm and a carrier frequency of 4.7 ppm. STD spectra were acquired with 2,048 scans, while saturation transfer reference (STR) spectra were acquired with 256 scans. The saturation frequencies for the selective excitation of Cdc5 were 6.604 and 0.462 ppm. Spectra were processed in Topspin 3.5 using an exponential multiplication window function with 3 Hz line broadening. STD/STR ratios were then compiled for the Spc72^{P} peptide in the absence and presence of Cdc5 ± Dbf4.

3.4.6 Circular Dichroism (CD) analysis

Cdc5 and Cdc5-A567W proteins were buffer exchanged in 50 mM Tris-HCl pH 7.0, 150 mM NaF, 1.4 mM 2-mercaptoethanol, 5 % glycerol to a final concentration of 7 μM. CD runs were carried out at room temperature using the Chirascan instrument (Applied Photophysics). Twenty iterations were done for each run and measurements were recorded between 190-250 nm. Data were analyzed using the Chirascan program (Applied Photophysics).

3.5 RESULTS

3.5.1 The polo-box domain of yeast Cdc5 recognizes phosphorylated substrates similarly to human Plk1

The C-terminal half of Cdc5 (residues 418-705) contains the conserved polo-box domain found in Plks (**Figure 3.1A**). The polo cap encompasses αA, the first polo box includes β1-β2-β3-β4-β5-β6-αB, and the second polo box comprises αC-β7-β8-β9-β10-β11-β12-αD (**Figure 3.1A-B**). The polo-box domains of Cdc5 and human Plk1 are virtually identical, but superimposing them yields a root mean square deviation (rmsd) value of 1.36 Å over 840 atoms. The αA helix in the polo-box domain is mainly responsible for this deviation as it is considerably longer in Cdc5 than in human Plk1 (**Figure 3.1C**).

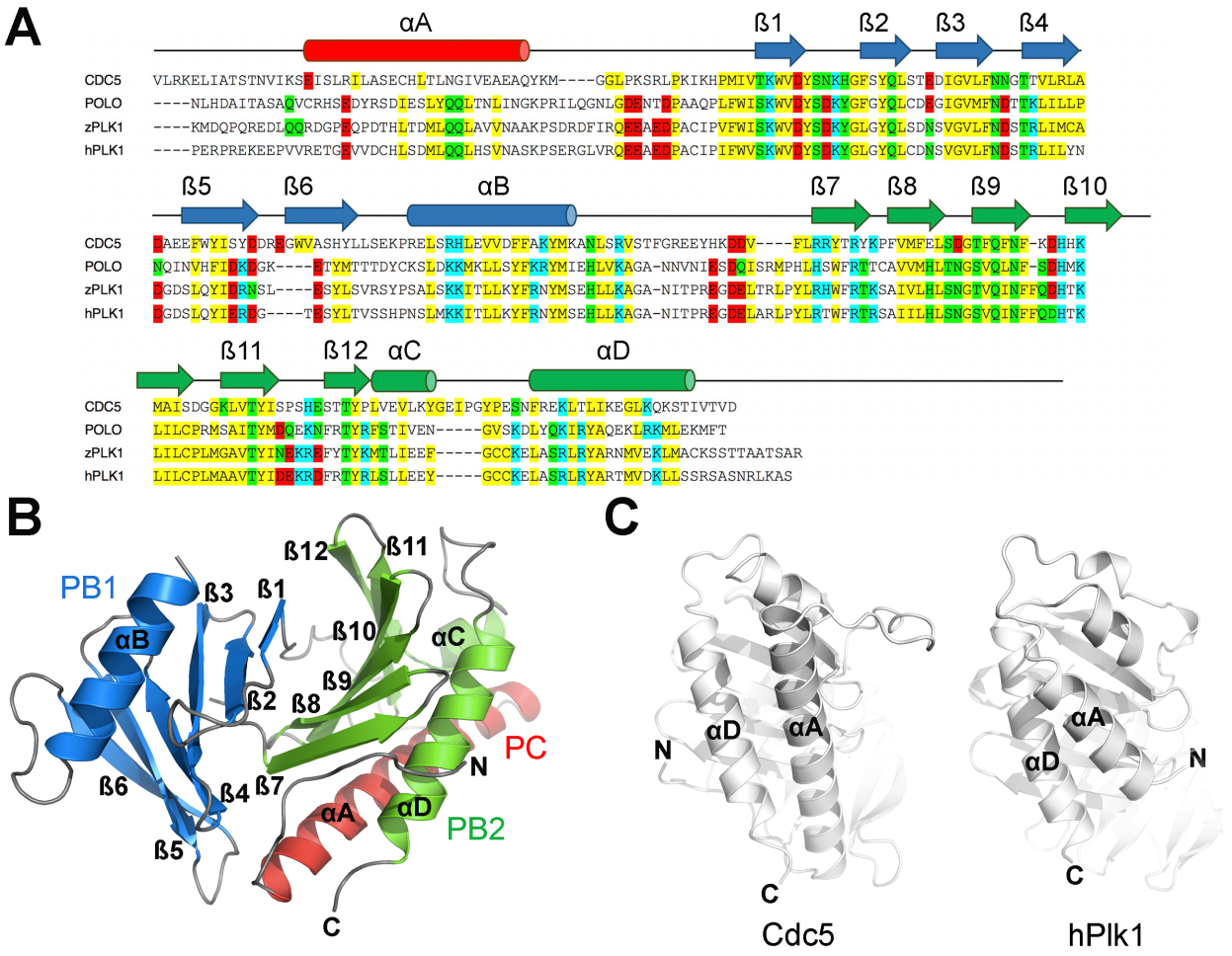

Figure 3.1. The polo-box domain of Cdc5. **(A)**. Sequence alignment of Cdc5, human Plk1 (hPlk1), zebrafish Plk1 (zPlk1), and Drosophila Polo. Conserved non-polar (yellow), positively-charged (cyan), negatively-charged (red), and polar (green) residues are highlighted. The secondary structure for Cdc5 (residues Leu460-Asp705) is presented above its sequence. The polo cap includes αA (red), polo-box1 encompasses β1-β2-β3-β4-β5-β6-αB (blue), and polo-box 2 (PB2) comprises αC-β7-β8-β9-β10-β11-β12-αD (green). **(B)** Ribbon diagram of the polo-box domain of Cdc5. The color scheme is identical as in panel A. The N- and C-termini are labeled. **(C)** Ribbon diagram of the polo-box domains of Cdc5 (white) and human Plk1 (white) rotated 90 ˚ relative to panel B. The N- and C-termini, as well as αA and αD, are labeled.

Next, the interaction between the polo-box domain of Cdc5 and a phosphorylated substrate was structurally investigated and evaluated whether the binding mode is similar to that of human Plk1. The structure of Cdc5 bound to a phosphorylated peptide derived from the spindle-pole body protein Spc72 (^{227}SLAQSpSPAGSQ^{237}) was thus solved. Spc72^{P} is an established Cdc5 binding partner containing the conserved X-S-(pS/T)-(P/X) motif recognized by polo-box

domains (Maekawa et al., 2007). The crystal structure of the Cdc5-Spc72^{P} complex showed clear electron density for phosphopeptide (**Figure 3.2A**).

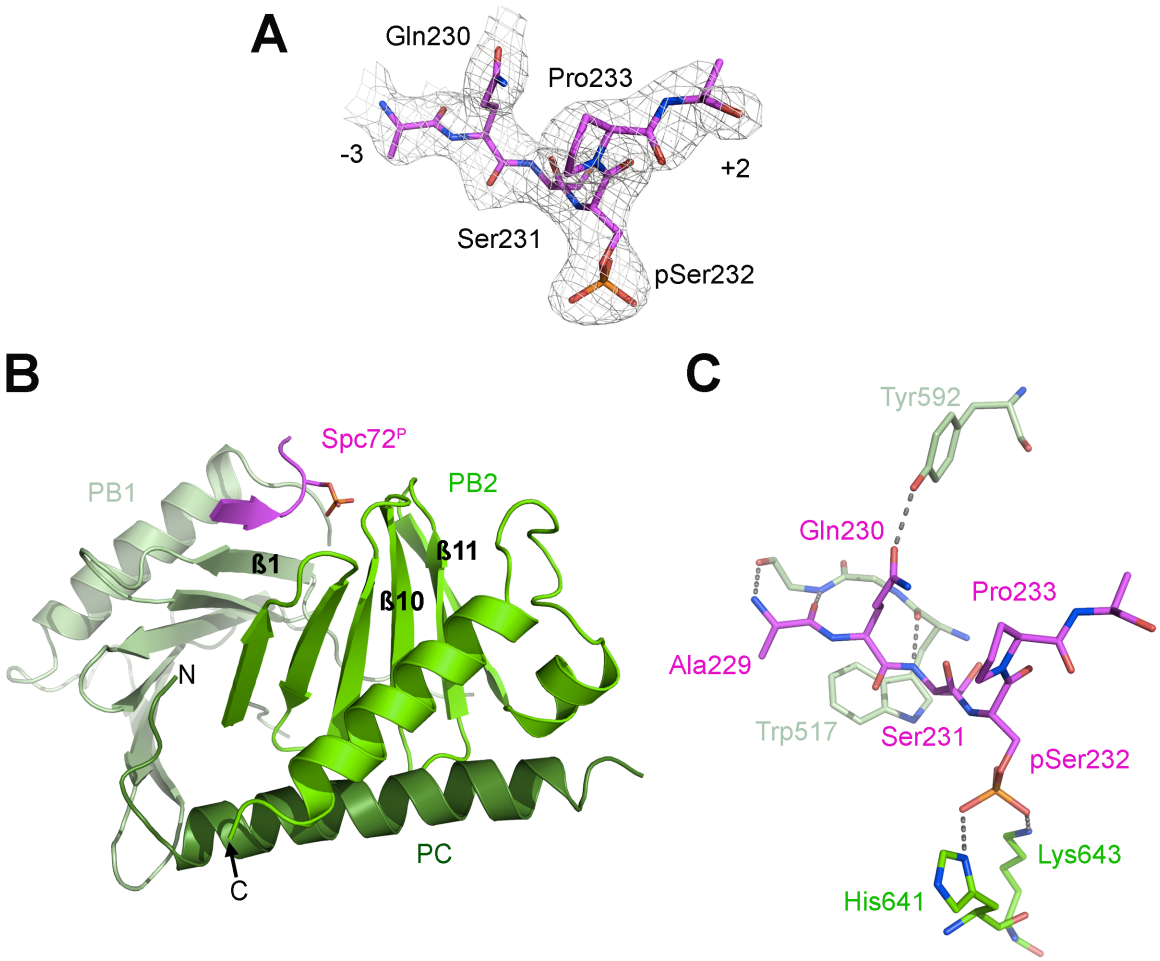

Figure 3.2. Spc72^{P} recognizes the pS/T-binding pocket of Cdc5. (A) The 2Fo-Fc electron density map around the Spc72^{P} peptide contoured at 1σ. **(B)** Ribbon diagram of the polo-box domain of Cdc5 bound to Spc72^{P} (magenta). The side chain of the phosphoserine (pSer232) is shown as a stick. The polo cap (CP; dark green), polo-box 1 (PB1; green), and polo-box 2 (PB2; pale green) are highlighted. Spc72^{P} binds between PB1 and PB2 making a short antiparallel ß-sheet with ß1 of PB1. Secondary structure elements that contribute to Spc72^{P} binding are labeled. **(C)** Detail of the interactions between Spc72^{P} and Cdc5. The color scheme is identical as in panel B.

The N-terminal region of Spc72^{P} adopts a strand conformation, which runs anti-parallel to β1 of polo-box 1 (**Figure 3.2B**). The phosphoserine (pSer232) was stabilized through hydrogen bonding with the side chains of His641 and Lys643, which are the pincer residues at the pS/T-

binding pocket (**Figure 3.2C**). This interaction was further supported by hydrogen bonding between Gln230 from Spc72^{P} and Tyr592 from Cdc5 (**Figure 3.2C**). Overall, the Cdc5-Spc72^{P} interactions were identical to the binding between human Plk1 and a canonical phosphorylated substrate.

3.5.2 The Dbf4-binding surface is opposite to the pS/T-binding pocket on the polo-box domain of Cdc5

Dbf4 recognizes the polo-box domain of Cdc5 through its polo-interacting region (^{83}RSIEGA^{88}) (Chen and Weinreich, 2010; Miller *et al.*, 2009). Mutation of the pincer residues on the pS/T-binding pocket of the polo-box domain did not abolish the *in vitro* interaction with Dbf4 (Chen and Weinreich, 2010), indicating that Dbf4 recognizes a different surface of Cdc5. The crystal structure of the polo-box domain of Cdc5 bound to the polo-interacting region of Dbf4 (^{76}RARIERA**RSIEGA**VQVSKGTG^{96}) was solved to characterize their interaction (**Table 3.1**). Despite its moderate resolution, the structure unequivocally showed that Dbf4 bound Cdc5 at a surface opposite to the pS/T-binding pocket (**Figure 3.3A**).

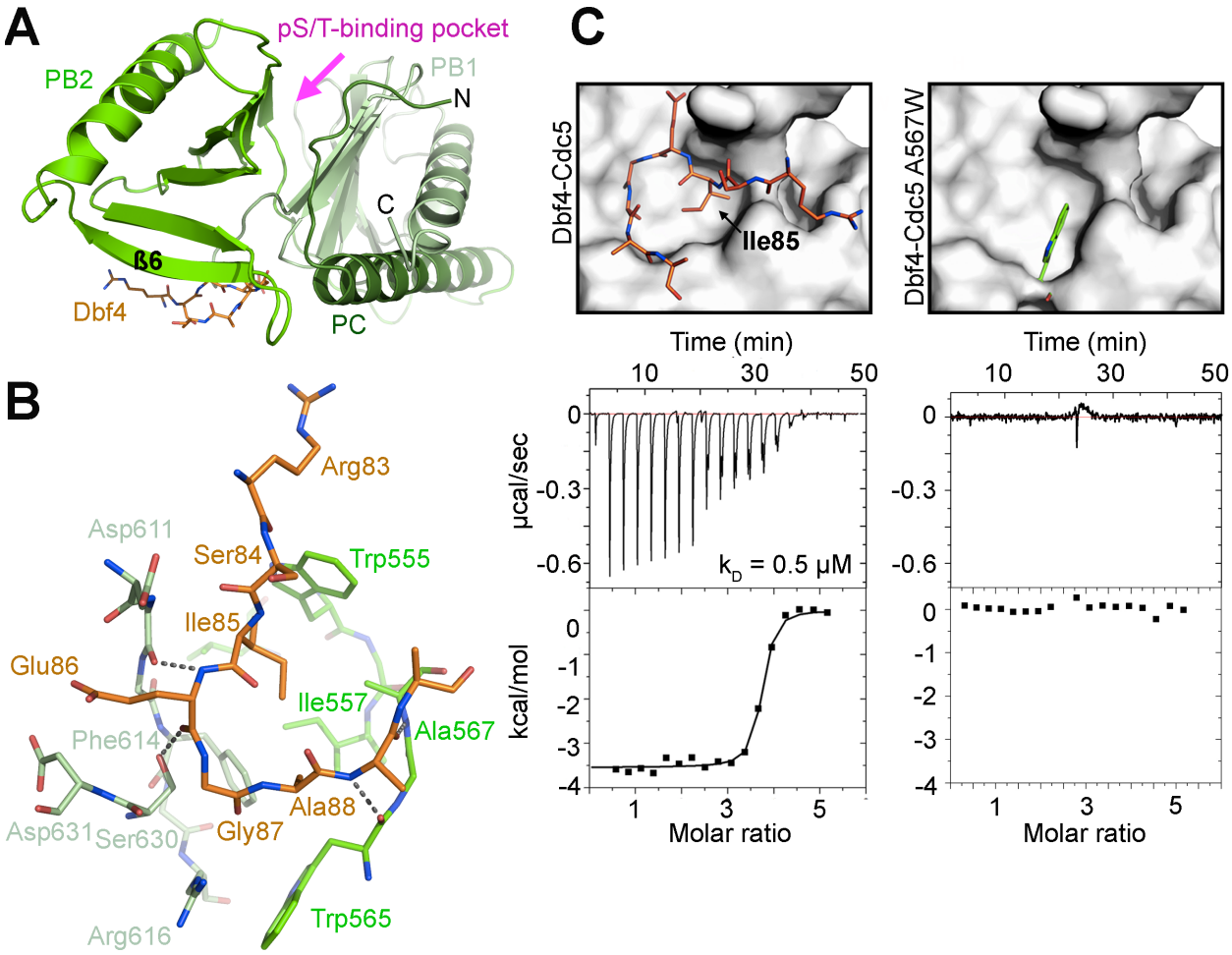

Figure 3.3. The Dbf4-binding surface on the polo-box domain of Cdc5. (A) The crystal structure of Dbf4 (orange sticks) bound to the polo-box domain of Cdc5 (cartoon). The polo-box domain of Cdc5 is colored and labeled identically as in figure 3.2B. The pS/T-binding pocket (magenta) is labeled. **(B)** Detail of the interactions between Dbf4 and Cdc5. Dbf4 and Cdc5 are colored identically as in panel A. **(C)** ITC analysis of the Cdc5-A567W mutant. Cdc5 (white) is shown as a surface and Dbf4 (orange) is presented as a stick (top left). The Cdc5 residue Ala567 is mutated to tryptophan (green stick) (top right). The ITC curves describe the interactions between Dbf4-Cdc5 (bottom left) and Dbf4-Cdc5-A567W (bottom right).

Of the 21 amino acids representing the Dbf4 peptide, only eight residues within the electron density map could be modeled. These amino acids likely represent ^{83}RSIEGA^{88} in Dbf4, since it is the minimal sequence of the polo-interacting region required for the interaction with Cdc5 (Chen and Weinreich, 2010).

The ordered part of the peptide adopted a U-turn conformation where the two arms of the U are stabilized by interactions with residues Asp611-Val613 preceding the β7 strand and residues Trp565-Ala567 of the β6 strand (**Figure 3.3B**). Accommodating the two extended arms forced a tight interaction between the residues at the U, and the surface defined by residues Trp565 (β6), Phe614-Arg616 (β7) and Ser630 (β5-β6 loop) of Cdc5 (**Figure 3.3B**). Accordingly, the turn was designated as Gly87-Ala88, given that these were the only residues that did not promote steric clashes with Cdc5 (**Figure 3.3B**). Two additional features further confirmed the register of the peptide. Ile85 occupies a conserved, hydrophobic pocket defined by residues Leu546, Trp555, Ile557, Ala567, and Phe614 (**Figures 3.3B**). The area of this pocket is roughly 110 $Å^2$, which is larger than the area of the pocket present in human Plk1 (72 $Å^2$). It was previously shown that a Dbf4-E86K mutant peptide bound Cdc5 with higher affinity than wild-type Dbf4 peptide (Chen and Weinreich, 2010). The Cdc5-Dbf4 structure demonstrated that the side chain of Glu86 is sandwiched between the side chains of Asp611 and Asp631, but does not interact with either (**Figure 3.3B**). The mutation of Glu86 to a lysine would favor the interaction with Asp611 thereby justifying the higher affinity of the mutant peptide to Cdc5.

Mutants of Cdc5 that disrupt Dbf4 binding were designed to confirm the interface of the Cdc5-Dbf4 complex. Apart from Ile85, the remainder of the contacts with Cdc5 were mediated by main chain interactions, and thus, could not be disrupted by point mutations. Therefore, Ala567 (ß6) of Cdc5 was mutated to tryptophan to disrupt Ile85 from packing at the hydrophobic pocket (**Figure 3.3C**). The resultant Cdc5-A567W variant was homogenous and folded (**Figure 3.4**). The binding affinities of the Dbf4 peptide to Cdc5 was then compared to the Cdc5-A567W mutant. In agreement with previously published work (Chen and Weinreich, 2010), Dbf4 bound

Cdc5 with a dissociation contast (k_D) ~ 0.5 μM (**Figure 3.3C**). However, Dbf4 failed to interact with the Cdc5-A567W mutant (**Figure 3.3C**), validating the Dbf4-binding surface of the polo-box domain identified in the crystal structure.

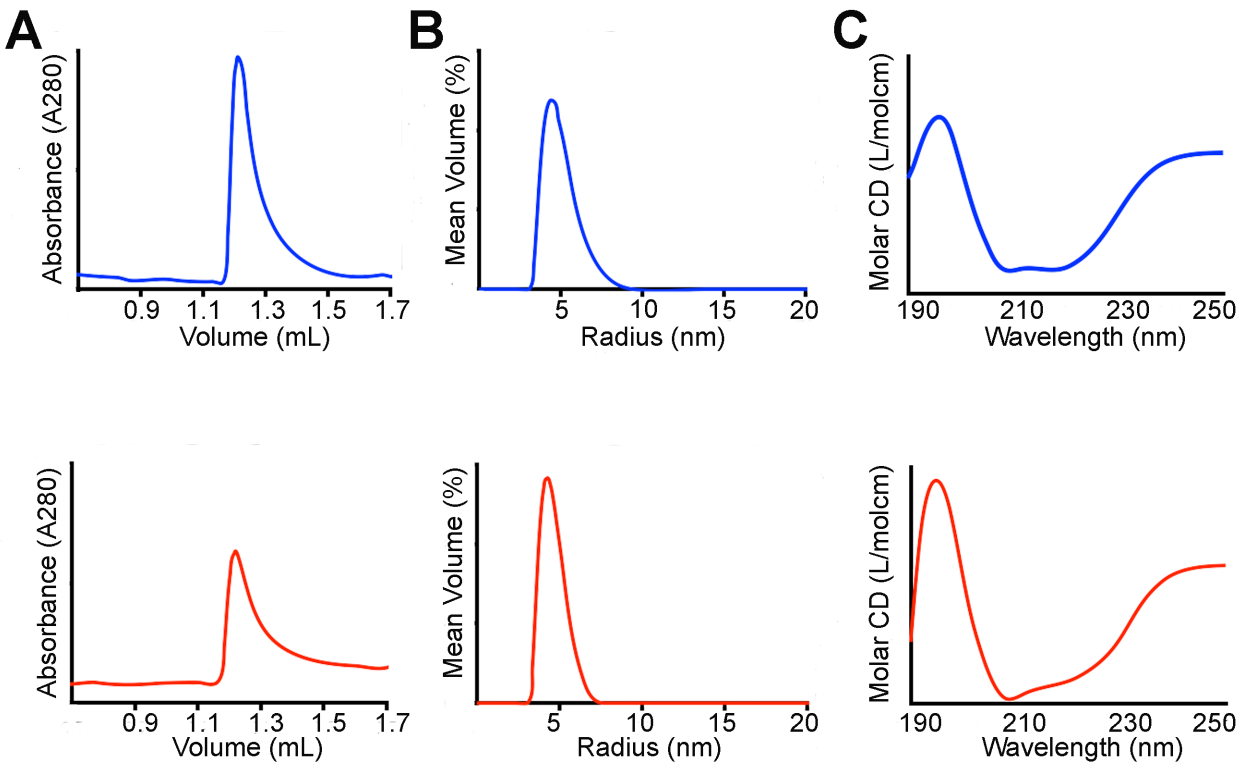

Figure 3.4. Biochemical characterization of the Cdc5-A567W mutant. (A) Size exclusion chromatograms, **(B)** dynamic light scattering mean volume plots, and **(C)** CD profiles for Cdc5 (blue) and Cdc5-A567W (red).

3.5.3 Dbf4 and Spc72^P bind simultaneously and non-competitively to Cdc5

We assessed whether Spc72^P and Dbf4 could simultaneously bind Cdc5 because each substrate recognized an opposite face of the polo-box domain. Using ITC, the interaction between Cdc5 and each substrate was confirmed, and that the binding of Spc72 depended on the phosphorylation of Ser232 (**Figure 3.5A-C**).

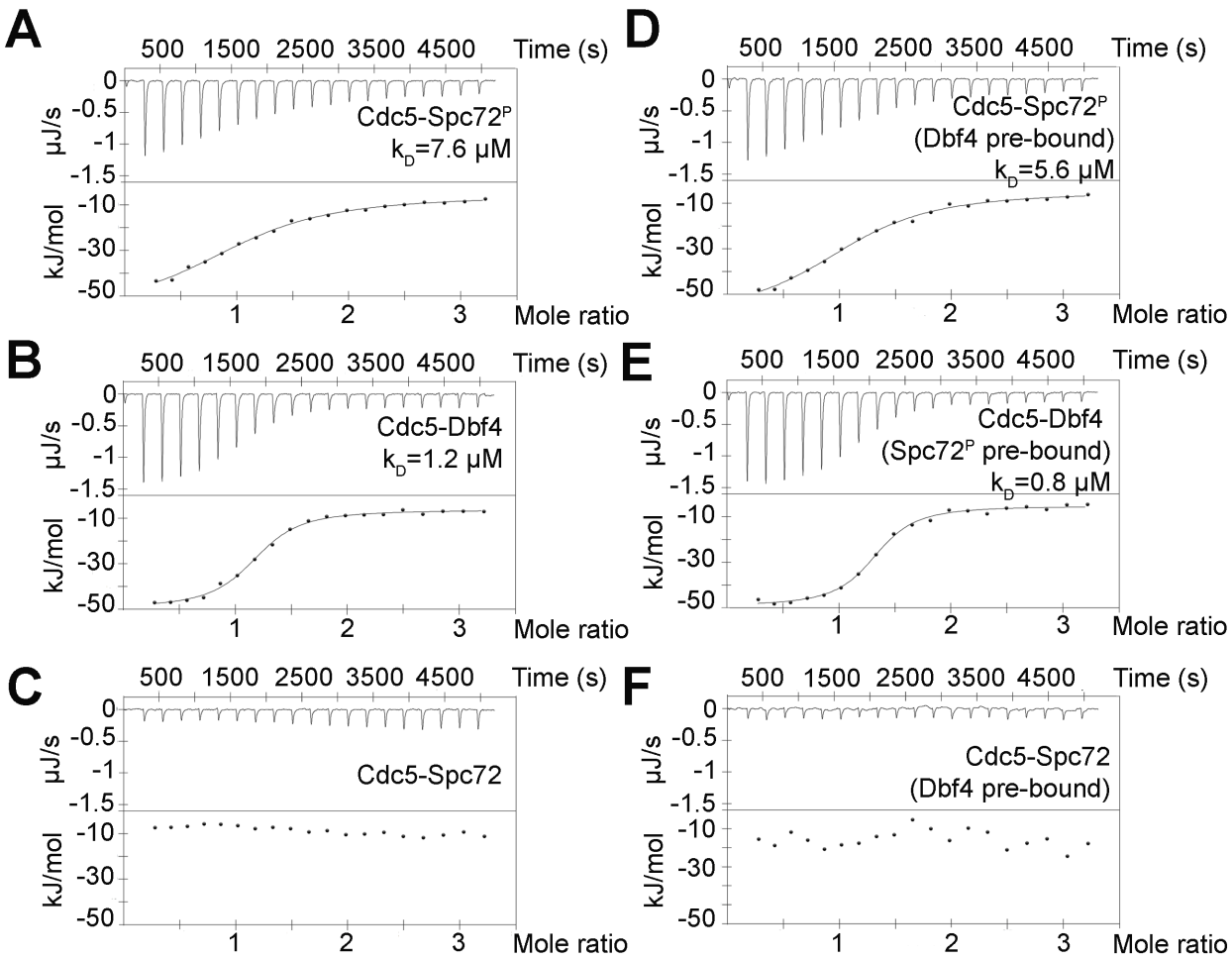

Figure 3.5. Spc72^{P} and Dbf4 interacted with Cdc5 simultaneously and non-competitively. **(A-C)** ITC curves for the interaction between Cdc5 and Spc72^{P} **(A)**, Dbf4 **(B)**, and a non-phosphorylated Spc72 peptide **(C)**. **(D-F)** ITC curves for the interaction with Spc72^{P} when Cdc5 was pre-incubated with Dbf4 (1:4 molar ratio of protein:peptide) **(D)**, the Cdc5-Dbf4 interaction when Cdc5 was pre-incubated with Spc72^{P} (same molar ratio as in D) **(E)**, and the Cdc5-Spc72 interaction when Cdc5 was pre-incubated with Dbf4 (same molar ratio as in D) **(F)**.

The binding curves showed that Cdc5-Dbf4 and Cdc5-Spc72^{P} complexes were saturated at around 1:3.5 molar ratios of Cdc5:peptide (**Figure 5A-B**). To test whether Spc72^{P} could interact with Cdc5 when pre-bound to Dbf4, Cdc5 was pre-incubated with Dbf4 at 1:4 molar ratio before titrating Spc72^{P}. The reciprocal experiment was also conducted whereby Cdc5 was pre-incubated with Spc72^{P} at 1:4 ratio, followed by titrating the Dbf4 peptide. Binding was observed in both cases (**Figure 3.5D-E**), indicating that Cdc5 could engage both substrates simultaneously,

irrespective of the order of peptide binding. As expected, the non-phosphorylated Spc72 peptide did not recognize Cdc5 (**Figure 3.5F**). Measurements at different excess molar ratios yielded similar results (data not shown).

The simultaneous binding of both peptides was further confirmed using NMR saturation-transfer difference (STD). One-dimensional proton NMR spectra were acquired for Cdc5, Dbf4 and Spc72^{P} to select for the frequency of Cdc5 (**Figure 3.6A**).

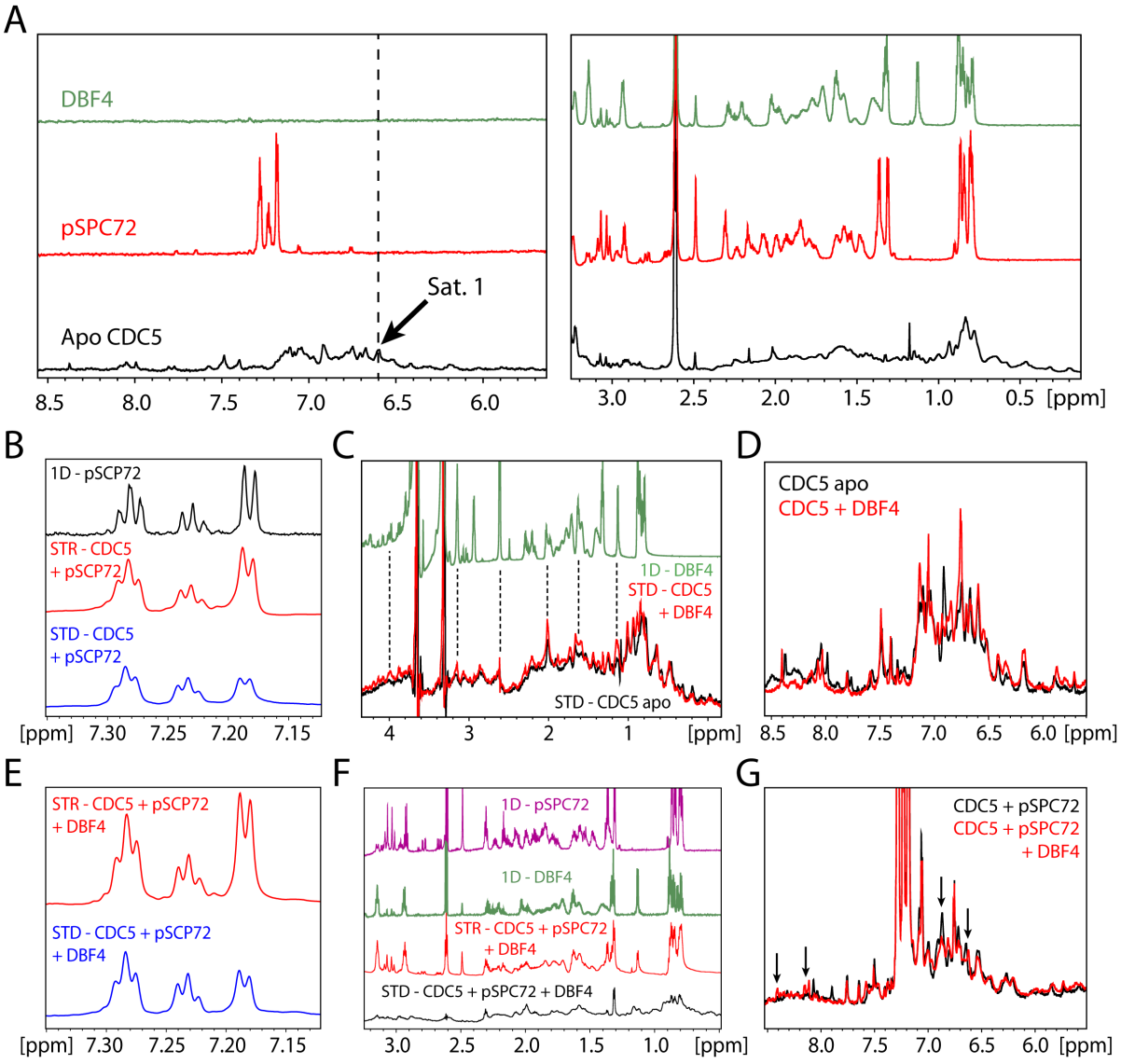

Figure 3.6. STD/STR profiles describing the Cdc5-substrate interactions. (A) One dimensional proton spectra for Cdc5 (black), Scp72^{P} (red), and Dbf4 (green). The saturation

frequency used in the STD experiments is indicated in the spectra. **(B)** STD spectra of Cdc5-Scp72^{P} complex focused on the aromatic resonances of Scp72^{P}. The 1D proton spectra and STD reference spectra (STR) are shown for comparison. **(C)** Overlay of Cdc5 (black) and Cdc5-Dbf4 STD (red) spectra. Dashed lines highlight Cdc5-Dbf4 peaks with increased intensities that correlate with free Dbf4 (green) chemical shifts. **(D)** Overlay of the aromatic region of 1D proton spectra for Cdc5 (black) and Cdc5-Dbf4 (red). **(E)** STD spectra of the Dbf4-Cdc5-Scp72^{P} complex focused on the aromatic resonances of Scp72^{P}. **(F)** Aliphatic region of the Dbf4-Cdc5-Scp72^{P} STD spectra (black). 1D proton spectra for each peptide and the STR spectra of the Dbf4-Cdc5-Scp72^{P} complex are shown for reference. **(G)** Overlay of the aromatic region of 1D proton spectra for Cdc5-Scp72^{P} (black) and Dbf4-Cdc5-Scp72^{P} (red) complexes. Arrows indicate chemical shift changes caused by Dbf4 binding.

A saturation frequency was chosen in the aromatic region of Cdc5 (6.604 ppm) since there was significant spectral overlap in the aliphatic region with Dbf4 and Spc72^{P}. By lacking aromatic residues, Dbf4 avoids the risk of non-specific saturation in the aromatic region. Accordingly, STD experiments for Dbf4 alone revealed no visible signal (**Figure 3.7A**).

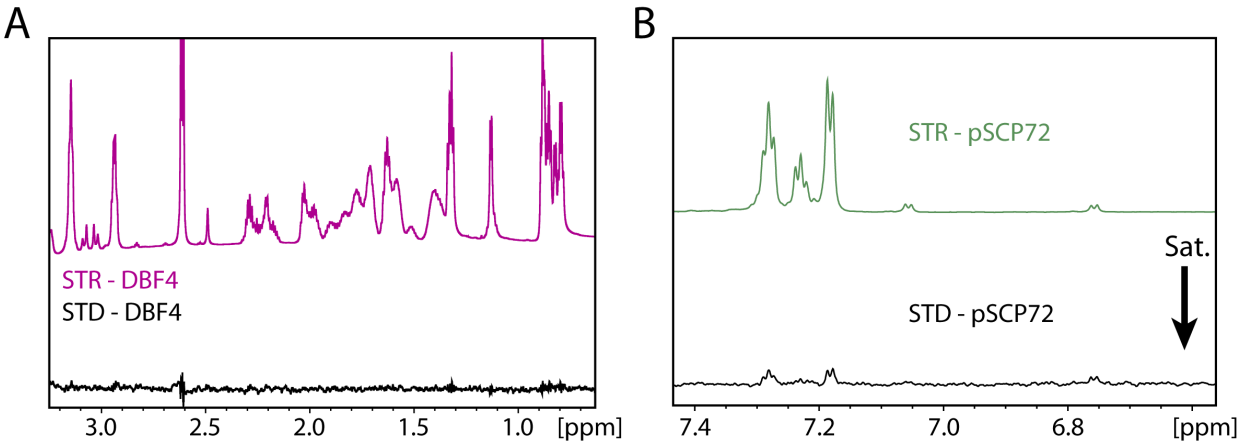

Figure 3.7. STD spectra of the individual substrates. STD spectra of Dbf4 **(A)** and Scp72^{P} **(B)**.

Meanwhile, STD spectra for Spc72^{P} showed a minor peak at around 6.7 ppm (**Figure 3.7B**). The STDs of Spc72^{P} were then compared to those of the Cdc5-Spc72^{P} complex to account for this saturation. Acquiring STD spectra for the Cdc5-Dbf4 and Cdc5-Spc72^{P} complexes demonstrated that the saturation transfer for Spc72^{P} resulted in a significantly higher signal when compared to Spc72^{P} alone (**Figure 3.6B-C, Table 3.2**).

Table 3.2. STD/STR ratios of the Spc72 peptides.

Sample	STD/STR ($x10^{-3}$)		
	1	**2**	**3**
Spc72^{P}	9±1.0	7±2	7±1.0
Spc72	9±3.0	N/D	8±2.0
Cdc5-Spc72^{P}	97.0±0.5	125±1	50.4±0.4
Cdc5-Spc72	70±2.0	107±3	38±1.0
Dbf4-Cdc5-Spc72^{P}	89.8±0.5	114±1	50.5±0.4

*The STD/STR ratios were calculated for the three peaks in Figure 3.6B, left to right.

The saturation transfer for Dbf4 was difficult to interpret because of signal broadening and spectral overlap with Cdc5 (**Figure 3.6C**). The STD spectra of Cdc5 and Cdc5-Dbf4 had several peaks with increased intensities consistent with the positions of Dbf4 resonances in the absence of Cdc5. There was also evidence of a chemical shift in the aromatic resonances of the Cdc5 spectra upon the addition of Dbf4 (**Figure 3.6D**). The spectral differences indicated that the chemical shift changes observed by Cdc5 must be due to Dbf4 binding since Dbf4 had no resonances within this region.

The normalized STD intensities (STD/STR ratios) for Spc72^{P} in the Dbf4-Cdc5-Spc72^{P} ternary complex (**Figure 3.6E**) indicated that its binding affinity was comparable to that of the binary Cdc5-Spc72^{P} complex (**Table 3.2**). On the other hand, impaired STD signal for Dbf4 was observed when analyzing the ternary complex, which could be due to spectral crowding and weak intensities similar to the Cdc5-Dbf4 complex (**Figure 3.6F**). However, chemical shift changes between the Cdc5-Spc72^{P} and Dbf4-Cdc5-Spc72^{P} complexes confirmed that Dbf4 interacted with Cdc5 in the presence of Spc72^{P} (**Figure 3.6G**). Overall, the ITC and NMR

experiments showed that Spc72^{P} and Dbf4 interacted simultaneously with Cdc5, and that binding to one peptide did not influence binding of the other.

3.5.4 The Dbf4-binding surface overlaps with the polo-box-kinase interface of zebrafish Plk1

The crystal structure of zebrafish Plk1 included both kinase and polo-box domains (Xu et al., 2013) (**Figure 3.8A**).

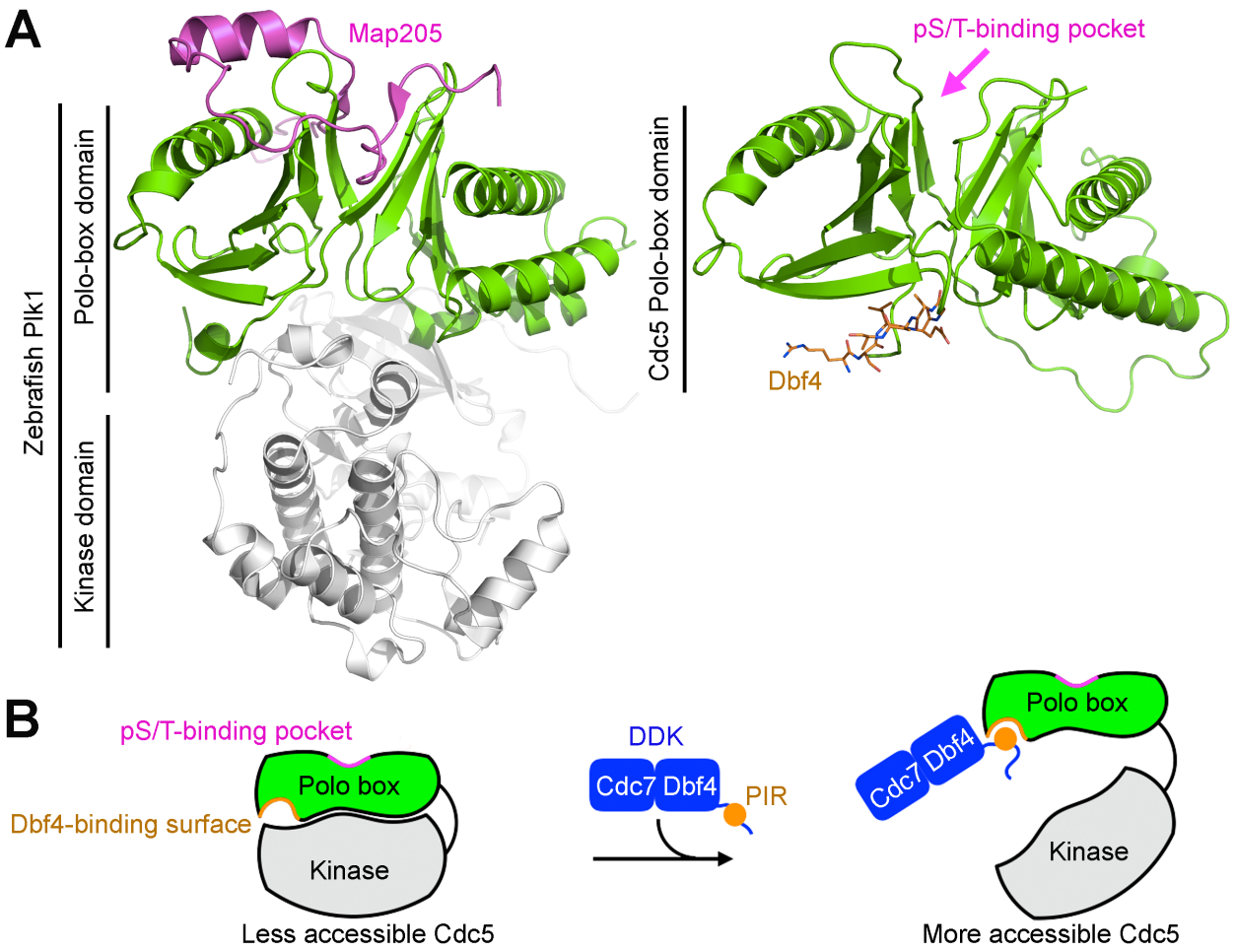

Figure 3.8. Comparison of yeast Cdc5 and zebrafish Plk1. (A) Ribbon diagrams of the zebrafish Plk1 in complex with Map205 (left) (PDB ID:4J7B) and Dbf4 bound to the polo-box domain of Cdc5 (right). The kinase domain (white) and the polo-box domain (green) are labeled. Map205 (magenta) is shown as a cartoon. The pS/T-binding pocket (magenta) on the polo-box domain of Cdc5 is labeled. **(B)** A model explaining how Dbf4-binding regulates Cdc5 activity. The inter-domain linker (black curved line) joins the kinase domain (white) and the polo-box domain (green) of Cdc5. The pS/T- (magenta) and Dbf4-binding surfaces (orange) on the polo-

box domain are labeled and highlighted. The Dbf4 polo-interacting region (PIR; orange ball) in DDK (blue) binds to Cdc5 and detaches the polo-box and kinase domains, thus rendering Cdc5 more accessible.

The structure showed that Map205 bound the pS/T-binding pocket of the polo-box domain, and induced a conformational change that enhanced the interaction between the polo-box and kinase domains (**Figure 3.8A**). The polo-box domains of Cdc5 and zebrafish Plk1 are virtually identical (rmsd = 0.68 Å for 761 atoms superimposed). The polo-box and kinase domains of zebrafish Plk1 formed a compact structure covering and interface area of 1,140 $Å^2$ (Krissinel and Henrick, 2007). Superimposing the polo-box domain of Cdc5 onto the polo-box domain of zebrafish Plk1 revealed that the pS/T-binding pocket is solvent-exposed (**Figure 3.8A**). On the other hand, the Dbf4-binding surface overlapped with the interface defined by the polo-box and kinase domains of zebrafish Plk1 (**Figure 3.8A**). Based on this, a model was proposed to explain how Dbf4 could regulate Cdc5 activity (**Figure 3.8B**). Dbf4 binding could detach the polo-box from the kinase domain thereby increasing the accessibility for each domain (**Figure 3.8B**).

The association of the kinase and polo-box domains resulted in auto-inhibition of their functions in zebrafish and human Plk1 (Elia et al., 2003; Xu et al., 2013). *In vitro* kinase assays demonstrated that the kinase domain is more active than full-length Plk1 and that the addition of the polo-box domain inhibited Plk1 kinase activity (Elia *et al.*, 2003; Jang *et al.*, 2002; Mundt *et al.*, 1997). Additionally, the polo-box domain binds phosphorylated peptides more effectively than full-length Plk1, suggesting that the kinase domain inhibits the polo-box domain (Elia et al., 2003). While auto-inhibition has not yet been demonstrated for Cdc5, the interface mediating the interaction between the polo-box and kinase domains is conserved among Plk1 homologs.

Accordingly, the compact form visualized in the zebrafish Plk1 structure depicts the auto-inhibited state for all Plks.

The Dbf4-Cdc5 interaction is required for full activation of the Mms4-Mus81 nuclease (Princz *et al.*, 2017). Accordingly, if Dbf4-binding releases inhibition off Cdc5 (**Figure 3.8B**), then it would explain how the interaction promotes Cdc7- and Cdc5-mediated phosphorylation of Mms4-Mus81. On the other hand, Dbf4 binding to Cdc5 during the SPOC response is consistent with Cdc7-mediated phosphorylation of the polo-box domain of Cdc5, which presumably prevents Cdc5 from recognizing its phosphorylated binding substrates during mitotic exit (Miller *et al.*, 2009). While the Cdc5-Dbf4 structure does not recapitulate the spatial organization of the DDK-Cdc5 complex, it is plausible that DDK partially blocks Cdc5 access to its phosphorylated substrates.

3.6 DISCUSSION

Genetic studies described the hierarchy of interactions involving the polo-box domain of Plks. The molecular details of how a polo-box domain mediates its interaction with phosphorylated targets through its conserved pS/T-binding pocket were also revealed (Cheng et al., 2003; Elia et al., 2003). However, phosphorylation-independent interactions involving polo-box domains received little attention, despite growing number of studies reporting such interactions (Chen and Weinreich, 2010; Miller et al., 2009; Xiang et al., 2007; Xu et al., 2013). We demonstrated how the polo-box domain of Cdc5 utilized a distinct surface in engaging its substrate in a phosphorylation-independent manner. We also showed that binding of a Dbf4 peptide does not prevent Cdc5 from interacting with its canonical phosphorylated substrate, Spc72^{P} (**Figure 3.5**). While this contradicts previously published work describing that Dbf4

binding to Cdc5 prevents the polo-box domain from recognizing its phosphorylated targets (Miller et al., 2009), the possibility that other regions of DDK beyond the polo-interacting region blocking Cdc5 access to its substrates cannot be ruled out.

Both Dbf4 and Map205 interact with polo-box domains in a phosphorylation-independent fashion (Chen and Weinreich, 2010; Xu et al., 2013). Map205 partially blocks the phosphopeptide binding site of zebrafish Plk1, whereas Dbf4 binds a distal surface of the polo-box domain of Cdc5. Binding of Map205 induces a conformational change which tightens the interaction between the polo-box and kinase domains, and as a result auto-inhibits the kinase activity of zebrafish Plk1 (Xu *et al.*, 2013). Conversely, Dbf4 binding likely detaches the polo-box and kinase domains of Cdc5, presumably increasing its kinase activity (**Figure 3.8**). The binding mode of Map205 received significant attention from a therapeutic standpoint, as it provided for means to indirectly inhibit the kinase activity of Plks (Archambault et al., 2015; Xu et al., 2013).

Interfering with the polo-box domain of human Plk1 using phosphomimetic peptide analogs caused mitotic defects, cytokinesis failure, and cell death (Lara and Berkard, 2012; Seong *et al.*, 2002). However, peptide analogs targeting the pS/T-binding pocket of Plks have shown little potency due to their inability to cross the plasma membrane effectively (Qian *et al.*, 2013). Mutations in Plk1 implicated in human pathology cluster at the interface between the kinase and polo-box domains (Burkard et al., 2009; Jang et al., 2002b; Koida et al., 2008; Macurek et al., 2008; Yuan et al., 2002). Two mutations (R337A and L340A) in the kinase domain interfere with ubiquitination and subsequent proteasomal degradation (Lindon and Pines, 2004; Bassermann et al., 2008), and in turn enhances Plk1 activity. As these two residues are

neighbouring the Dbf4-binding interface, it is tempting to suggest that Dbf4 may also modulate ubiquitination.

Our study unveils a new surface of the polo-box domain used in modulating Cdc5 activity. To our knowledge, this is the first crystal structure of a polo-box domain bound to a substrate through a distinct interface and lays the foundation to study how a polo-box domain utilizes surfaces beyond the conserved pS/T-binding pocket to mediate protein-protein interactions and regulate kinase activity.

3.7 ACKNOWLEDGEMENTS

We would like to thank Dr. Junop, Dr. Brown and Ms. Shen for assistance with data collection, Dr. José Carlos Bozelli Jr. from Dr. Richard Epand's laboratory (McMaster University) for assistance with the isothermal calorimetry experiments, and Dr. Paula Balisi for assistance with protein purification at the early stages of the project. This work was funded by the Canadian Institutes of Health Research (MOP-67189 to A.G.). A.W.A. was funded in part by an Ontario Graduate Scholarship.

CHAPTER FOUR

BINDING OF MUTL TO THE SLIDING ß-CLAMP IS SPECIES SPECIFIC.

Manuscript submitted.

Authors: Ahmad W. Almawi, Michelle K. Scotland, Justin R. Randall, Linda Liu, Heather K. Martin, Monica C. Pillon, Lyle A. Simmons, Mark D. Sutton, and Alba Guarné.

4.1 CONNECTING TEXT

The Dbf4-Rad53 complex was captured and crystallized because a chimera of the two proteins stabilized their interaction. Like the Dbf4-Rad53 interaction, the sliding ß-clamp weakly binds to the MutL endonuclease during DNA mismatch repair. The clamp-MutL interaction serves to activate MutL in cleaving the newly-synthesized DNA strand that harbors the mismatch. The clamp-MutL complex is a promising target to capture because the structure of each binding partner is known (Burnouf et al., 2014; Guarné et al., 2004; Kong et al., 1992; Pillon et al., 2010) and their interaction was extensively characterized *in vitro* (Pillon et al., 2015). Chapter 4 describes the crystal structures of the *Escherichia coli* (*E. coli)* and *Bacillus subtilis* (*B. subtilis)* clamp-MutL complexes with the aim of understanding how their binding differs between organisms. Dr. Michelle K. Scotland from the laboratory of Dr. Mark D. Sutton (University of Buffalo) analyzed the mutation frequency of *E. coli* MutL variants. Dr. Justin R. Randall and Heather K. Martin from the laboratory of Dr. Lyle A. Simmons (University of Michigan) analyzed the mutation frequency of *B. subtilis* MutL variants. Linda Liu from the laboratory of Dr. Alba Guarné (McMaster University) contributed in crystallizing the *B. subtilis* clamp-MutL chimera. Dr. Monica C. Pillon (National Institutes of Health) provided technical and intellectual input to the project. With the guidance of my supervisor, Dr. Alba Guarné, I conducted experiments, analyzed data, and assisted in preparing the manuscript.

4.2 ABSTRACT

The sliding ß-clamp is a protein hub central to DNA replication and fork management. Proteins interacting with the ß-clamp typically harbor a conserved clamp-binding motif that is often found in extended regions. Therefore, clamp interactions have –almost exclusively– been studied using short peptides recapitulating the binding motif. This approach has revealed the molecular determinants that mediate the binding, but peptides cannot describe how proteins with clamp-binding motifs embedded in structured domains are recognized. The mismatch repair protein MutL has an internal clamp-binding motif, but its interaction with the ß-clamp has different roles depending on the organism. In *Bacillus subtilis*, the interaction stimulates the endonuclease activity of MutL and it is critical for DNA mismatch repair function. Conversely, the interaction between *Escherichia coli* MutL and the ß-clamp only plays a minimal role in DNA mismatch repair. We have determined the structures of *E. coli* and *B. subtilis* MutL bound to their respective ß-clamps. These structures reveal different binding modes and confirm that binding to the ß-clamp is a two-step process. Comparison of the two structures also allows us to propose a model for the activation of the endonuclease activity of *B. subtilis* MutL that is likely conserved in other organisms lacking methyl-directed mismatch repair.

4.3 INTRODUCTION

The sliding ß-clamp, and its eukaryotic counterpart PCNA, are ring-shaped structures that encircles DNA and tethers their binding partners to DNA. The main role of these clamps is to increase the processivity of DNA polymerase, but they also have critical roles at regulating DNA replication, polymerase switching, and DNA mismatch repair (Heltzel et al., 2009; Kath et al., 2016; Lopez de Saro and O'Donnell, 2001; Moldovan et al., 2007). Initiation of DNA mismatch

repair depends on the coordinated actions of two proteins –MutS and MutL– and both interact with the ß-clamp (Lopez de Saro et al., 2006; Pillon et al., 2011; Simmons et al., 2008). MutS recognizes mismatches and small insertion/deletion loops that have escaped DNA polymerase proofreading, and recruits MutL to mark the newly synthesized strand for repair either directly (methyl-independent repair) or indirectly (methyl-directed repair). *Escherichia coli* (*E. coli*) has a devoted nuclease that recognizes transiently hemimethylated DNA and nicks the unmethylated strand, effectively discriminating the new from the template strand. However, most prokaryotes – and all eukaryotes– lack this gene. Instead, MutL homologs from these organisms harbor their own nuclease activity that is stimulated by the interaction with the ß-clamp or its eukaryotic counterpart PCNA (Genschel et al., 2017; Kadyrov et al., 2006; Kosinski et al., 2008; Pillon et al., 2015; Pluciennik et al., 2010). Accordingly, disruption of the clamp-binding motif in *B. subtilis* MutL causes a strong mutator phenotype, whereas disruption of this motif in *E. coli* MutL –which does not have nuclease activity– only causes a mild mutator phenotype (Pillon et al., 2011; Pillon et al., 2015). Sliding clamps bind DNA with a defined orientation. Therefore, proteins interacting with the ß-clamp also bind DNA in a specific orientation. In the case of MutL, this asymmetry determines the strand of the DNA duplex that gets cut (Pluciennik et al., 2010).

All clamp-binding partners contain a conserved linear motif known as PCNA-interacting protein (PIP) box or clamp-binding motif (CBM) that binds a hydrophobic groove on the sliding clamp (**Figure 4.1A**).

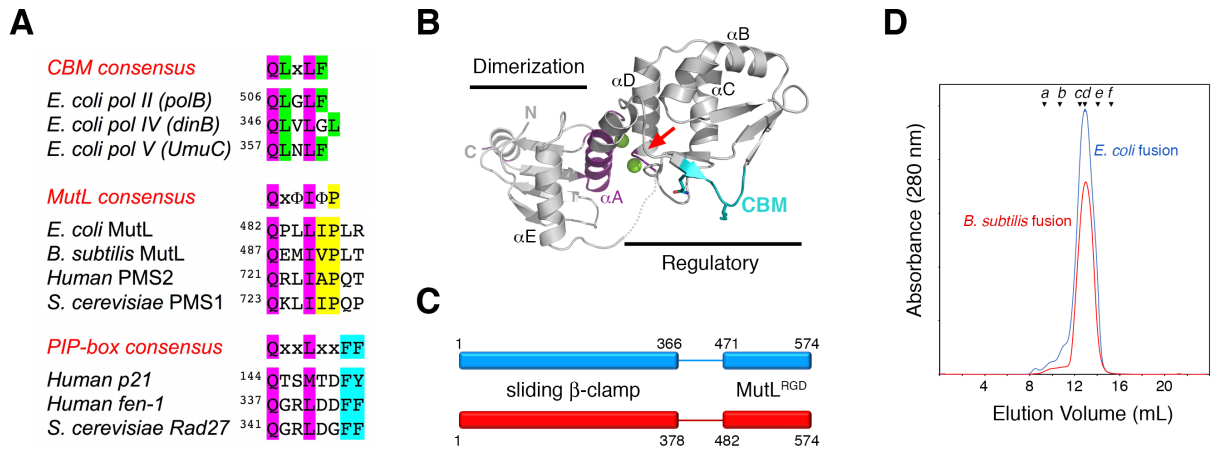

Figure 4.1. Stabilization of the clamp-MutL complex. (A) Sequence conservation of the canonical clamp-binding motif (top), MutL clamp- and PCNA-binding motif (center), and PIP-box (bottom). Universally conserved residues are shown in purple and conserved residues unique to each motif highlighted in green, yellow, and cyan. **(B)** Ribbon representation of the endonuclease domain of *B. subtilis* MutL (433-627). The beginning of the regulatory domain (RGD) is mark with a red arrow. The endonuclease motif is colored purple and the zinc metal ions are shown as green spheres. The clamp- binding motif (CBM) is colored in cyan with the side chains of the conserved Gln and Leu residues shown as sticks. **(C)** Construction of the *E. coli* and *B. subtilis* clamp-MutL^{RGD} fusions shown in blue and red, respectively. **(D)** Size exclusion chromatography column profiles of the *E. coli* and *B. subtilis* clamp-MutL^{RGD} fusions. Elution volumes of reference proteins are: a– thyroglobulin (669 kDa), b– ferritin (440 kDa), c– catalase (232 kDa), d– aldolase (158 kDa), e– albumin (67 kDa), and f– ovoalbumin (43 kDa).

PIP boxes have a strict QxxLxxFF consensus sequence, where the conserved leucine and phenylalanine residues form a 3_{10} helix that defines a "three-forked plug" interaction with PCNA (Boehm and Washington, 2016). Only one PIP-box lacking the two aromatic residues has been described to date (Armstrong et al., 2012), underscoring the importance of these two residues. Clamp-binding motifs are shorter and have a general QL(D/S)LF consensus sequence. This consensus sequence tolerates greater variability than the PIP-box consensus, a feature that often challenges the identification of CBMs (Dalrymple et al., 2001). PIP boxes and CBMs are normally found in flexible terminal regions and, therefore, most structural information comes from complexes of PCNA or the ß-clamp bound to short peptides derived from binding partners

(Bunting et al., 2003; Georgescu et al., 2008; Patoli et al., 2013). These studies have shown that all clamp-binding partners share a characteristic bidentate interaction, where the conserved glutamine and leucine residues of the motif are bound to adjacent pockets at the C-terminus of the ß-clamp. The loops surrounding these two pockets are flexible and close in to define the walls of the groove upon binding of the peptide. Studies using peptides, however, fail to explain how the ß-clamp balances multiple binding partners or how additional surfaces of the ß-clamp regulate partner switching (Heltzel et al., 2009; Kath et al., 2016).

The clamp-binding motif of MutL is located within a structured region of the endonuclease domain (Pillon et al., 2010). Internal clamp-binding motifs are less common than terminal CBMs, but this may be due to the challenge of identifying internal motifs in the absence of a strong consensus sequence (Dalrymple et al., 2001). DNA polymerase V (UmuC) has an internal CBM located between its little finger and C-terminal domains, but since it is found within the linker connecting the two domains its interaction with the ß-clamp was also studied structurally using a short peptide (Patoli et al., 2013). The a catalytic and e proofreading subunits of DNA polymerase III also contain internal CBMs (Dohrmann and McHenry, 2005; Jergic et al., 2013; Toste Rego et al., 2013). Interaction of Polα with the ß-clamp was also studied structurally using a short peptide (Georgescu et al., 2008). Therefore, as in previous analyses, these structures do not provide information as to how the surrounding domains affect binding to the ß-clamp.

Although the clamp-binding motif of MutL resides in a surface exposed loop of the endonuclease domain, binding to the ß-clamp must impose significant rearrangements to avoid steric hindrance. The interaction between MutL and the sliding ß-clamp is conserved from bacterial to human (Genschel et al., 2017; Pillon et al., 2010), therefore the same binding motif

mediates the interaction with either the ß-clamp and PCNA (**Figure 4.1A**). To understand how MutL homologs interact with the ß-clamp, we have determined the crystal structures of *B. subtilis* and *E. coli* ß-clamp bound to their respective MutL homologs. We find that *B. subtilis* MutL interacts with the ß-clamp through the characteristic bidentate interaction, whereas *E. coli* MutL forms a monodentate interaction. Comparison of the two structures reveals a trade-off between domain flexibility and the formation of a bidentate interaction with the ß-clamp. We propose a model describing how the increased flexibility helps *B. subtilis* MutL, and likely human MutLα, align their nuclease active site with the central cavity of the ß-clamp.

4.4 METHODS

4.4.1 Design of the clamp-MutL fusion proteins

The fragment of *E. coli* MutL encoding residues 471-574 (regulatory domain) was subcloned in a modified pET15b expression vector containing an N-terminal His_6-tag removable with tobacco etch virus (TEV) protease site using the BamHI/BlpI restriction sites (pAG8902). Full-length *E. coli* ß-clamp was subsequently subcloned using the NdeI/BamHI restriction sites to create the *E. coli* clamp-MutL fusion (pAG8903). A SalI restriction site was engineered between the two protein fragments to introduce a glycine/serine-rich linker (SGASG). Oligonucleotides encoding the linker flanked by SalI/BamHI sites were purchased from Integrated DNA Technologies and ligated into to generate the final clamp-SGASG- MutL fusion expression plasmid (pAG8918). The *B. subtilis* fusion expression plasmid (pAG9021) was generated analogously and included full-length ß-clamp and residues 482-574 from *B. subtilis*,

corresponding to its regulatory domain joined by a three-amino acid linker (VDS). The identity of all plasmids was confirmed by DNA sequencing (MOBIX, McMaster University).

4.4.2 Protein expression and purification

The *E. coli* and *B. subtilis* clamp-MutL fusions were produced in BL21 (DE3) cells supplemented with a plasmid encoding rare tRNAs. Cells were grown in Luria-Bertani media to an OD_{600} ~ 0.7 and protein expression was induced by addition of 1.5 mM isopropyl β-D-1-thiogalactopyranoside (IPTG). Cells were harvested by centrifugation (1,000 x *g* for 15 min) after overnight incubation at 16 °C. Cell pellets were re-suspended in buffer A (20 mM TRIS-HCl pH 8, 500 mM NaCl, 1.4 mM b-mercaptoethanol, 5 % glycerol) and lysed by sonication. Lysates were clarified by centrifugation at 39,000 x *g* for 40 minutes. The supernatant was loaded onto a Nickel-chelating affinity column (GE Healthcare) pre-equilibrated with buffer A. The column was washed with 63 mM imidazole and the His-tag fusions were eluted with 195 mM imidazole. The eluted protein was then pooled, and the salt concentration was diluted in half prior to injecting into a Q-Sepharose ion exchange column (GE Healthcare) equilibrated with buffer B (20 mM TRIS-HCl pH 7.5, 150 mM NaCl, 1.4 mM b-mercaptoethanol, 5% glycerol). The protein was eluted off the column using a linear gradient to 500 mM NaCl (fusions elute at ~350 mM NaCl). The protein was further purified through size exclusion chromatography using a Superdex200 (S200) 10/300 GL size-exclusion column (GE Healthcare) equilibrated with buffer B. The eluted samples were concentrated to 20 mg/mL in buffer B and protein concentration was calculated using the Beer-Lambert equation with an extinction coefficient of

29,450 $M^{-1}cm^{-1}$ and 24,710 $M^{-1}cm^{-1}$ for the *E. coli* and *B. subtilis* clamp-MutL fusions, respectively.

4.4.3 Crystallization and structure determination

Crystals of the *E. coli* clamp-MutL fusion grew in 100 mM BIS-TRIS pH 5.5, and 2 M ammonium sulfate and were cryo-protected by addition of 8 % glycerol to the mother liquor. Crystals of the *B. subtilis* clamp-MutL fusion grew in 100 mM HEPES pH 7.5, 25 % PEG 3350 (v/v), and 0.2 M ammonium sulfate, were cryo-protected by addition of 12 % ethylene glycol. Complete data sets were collected at the O8ID-1 and O8B1-1 beam lines of the Canadian Light Source and were processed using XDS (**Table 4.1**) (Kabsch, 2010).

Table 4.1. Data collection and refinement statistics.

	E. coli **β-MutL fusion**	***B. subtilis*** **β-MutL fusion**
Data Collection		
Beamline	08ID-1	08B1-1
Wavelength	0.97936	0.98010
Space Group	$P2_12_12_1$	P1
Cell dimensions a, b, c (Å) α, β, γ (°)	78.4, 103.2, 141.6 90, 90, 90	59.0, 83.9, 128.3 80.3, 83.6, 90
Resolution (Å)[#]	46.0-2.07 (2.12-2.07)	48.2-2.34 (2.42-2.34)
Completeness (%)[#]	97.5 (99.8)	97.1 (97.1)
CC½ (%)	99.8 (35.6)	99.6 (32.7)
I/s(I)	11.5 (1.2)	13.4 (0.95)
Redundancy	4.4 (4.5)	2.2 (1.7)
Refinement		
Resolution (Å)	46 – 2.07	48.2 – 2.34
No. Reflections	53,370	98,638
R_{work} / R_{free} (%)	19.0 / 22.5	21.6 / 25.0
Atoms refined (no H)	7,224	13,850
Solvent atoms	207	195
Rmsd in bonds (Å)	0.004	0.004
Rmsd in angles (°)	0.96	0.73
Mean B values (Å)	54	94.5
Ramachandran Plot (%) Favored Outliers	 97.4 0.1	 95.5 0.4

Data in the highest resolution shell is shown in parentheses.

The structures were determined by molecular replacement using the structures of the individual components as search models (PDB ID: 4K3M and 1X9Z for the *E. coli* fusion and PDB ID: 4TR6 and 3KDK for the *B. subtilis* fusion). The initial models were refined by iterative cycles of manual model building in Coot and refinement in PHENIX (Afonine et al., 2005; Emsley and Cowtan, 2004). Quantitative analysis of the interfaces of both structures was done using the

online PISA server (Krissinel and Henrick, 2007). Figures showing molecular structures were generated using PyMOL (DeLano, 2002).

4.4.4 Analysis of the mutation frequency of *E. coli* MutL variants

E. coli strains MG1655 and JW4128-1 were obtained from the *E. coli* Genetic Stock Center. The Δ*mutL720*::*kan* allele from strain JW4128-1 was transduced into MG1655 using P1*vir* (Sutton, 2004). Cultures of strain MKS108 bearing either pET15b or a pET15b derivative expressing the indicated MutL protein were grown at 37 ˚ for 12 hours with aeration in LB medium (10 g/l Difco tryptone, 5 g/l Difco yeast extract, 10 g/l NaCl) supplemented with ampicillin (150 μg/ml) and kanamycin (40 μg/ml). Saturated cultures were serially diluted in 0.8% saline and 100 μl of the 10^{-6} dilution (n>15) was spread onto LB agar plates to determine culture titers. Three-hundred μl of the same undiluted cultures (n=20) were spread onto LB agar plates supplemented with rifampicin (50 μg/ml) to identify spontaneous mutations. Plates were incubated at 37 ˚ for 12 h prior to counting colonies and the spontaneous mutation frequency of each strain was determined by dividing the number of colony forming units on plates containing rifampicin by the average number of colony forming units on LB lacking rifampicin (Dixon and Massey, 1969, Pillon et al., 2015).

4.4.5 Analysis of the mutation frequency of *B. subtilis* MutL variants

Each strain was created by integration of each *mutL* mutant into the *amyE* locus of the Δ*mutL* strain of *B. subtilis* for ectopic expression with IPTG. A stock for each strain was frozen and stored at -80°C. The wild type PY79 and isogenic *ΔmutL* strains were struck out on LB agar;

while the *mutL*$^{+}$, *mutL*^{HK}, *mutL*^{CMB} were struck on LB agar supplemented with spectinomycin (100 μg/ml) and IPTG (1 mM) for overnight growth. Both plates were incubated overnight at 30°C. The following day, individual colonies were picked and cultured in 2 ml LB and 200 μM IPTG rotating at 37°C for 2-3 hours to OD_{600} of 1- 1.2. At this point, 1 ml from each tube was aliquoted and pelleted at 11,000 x g and saved on ice for Western blot analysis. The remaining 1 ml was also pelleted at 11,000 x g and the supernatant was removed. This pellet was resuspended in 100 μl of 0.85% saline solution and 100 μl was plated on LB agar supplemented with rifampin (100 μg/ml). The cell suspension was then diluted 10^{-6} and 100 μl of each strain was plated on LB agar to determine the total number of viable cells. Plates were then incubated at 30°C and each plate was scored for colony forming units. Mutation rate analysis was conducted using the Ma-Sandri-Sarkar maximum likelihood estimator (MSS-MLE) method described in (Foster, 2006; Hall et al., 2009). In total, 12-15 independent cultures were used per strain.

4.4.6 Cysteine-crosslinking complex formation

Single cysteine variants of ß-clamp (pAG8807) and the endonuclease domain of MutL (pAG8803) were purified as described earlier (Pillon et al., 2015). The single-Cys variant of the regulatory domain of MutL (pAG9148) was generated by site-directed mutagenesis and produced as described elsewhere (Pillon et al., 2011). To form the complexes between the ß-clamp and the endonuclease and regulatory domains of *B. subtilis* MutL, the ß-clamp was incubated with either the endonuclease or regulatory domains at a 1:1 ratio to a final concentration of 20 μM. The samples (1–2 ml) were dialyzed in buffer A (20 mM Tris pH 7.6, 150 mM KCl, 10 mM DTT, 10% glycerol) for 2 hours at 4°C. The mixture was then transferred

into dialysis buffer B (same as A but with 5 mM DTT) for 1 h, followed by 1 h in dialysis buffer C (without DTT). The sample was then left in dialysis buffer C. Complex formation was monitored by resolving samples collected at different time points on denaturing polyacrylamide gradient gels stained with Coomassie Brilliant Blue.

4.5 RESULTS

4.5.1 Stabilization of the MutL-clamp complex

E. coli and *B. subtilis* MutL form weak, yet specific, interactions with their respective sliding ß-clamps (Pillon et al., 2011; Pillon et al., 2015). The dimerization domain of MutL is organized into two independently folded regions connected by an α-helix that, in the case of *B. subtilis* MutL harbors the endonuclease motif (Pillon et al., 2010). The N- and C-terminal ends of the domain of MutL define the dimerization interface of the protein and the intervening region defines an independently folded subdomain (**Figure 4.1B**). This subdomain is often referred to as the regulatory subdomain because it mediates the interaction with the sliding ß-clamp. The clamp-binding motif of MutL is located at the N-terminus of the regulatory subdomain (**Figure 4.1B-C**). Since the sliding ß-clamp interacts with its binding partners through a conserved groove located at its C- terminus, we stabilized the interaction by connecting the two polypeptide chains with a short linker – an approach that has been successfully used to stabilize other weak protein-protein interactions for crystallographic studies (Almawi et al., 2016; Antczak et al., 2006; Kingston et al., 2004; Williams et al., 2014). The *E. coli* and *B. subtilis* fusions could be purified to homogeneity and eluted from a size exclusion chromatography column at retention volumes consistent with the expected mass for each complex (**Figure 4.1C-D**).

4.5.2 *E. coli* MutL forms a monodentate interaction with the sliding ß-clamp

Crystals of the *E. coli* ß-clamp fused to the regulatory domain of MutL (residues 471-574) diffracted to 2.1 Å (**Table 4.1**). The asymmetric unit contained one dimer of the fusion (**Figure 4.2A**), where both the clamp dimer and the two MutL regulatory subdomains had similar overall structures to the individual proteins.

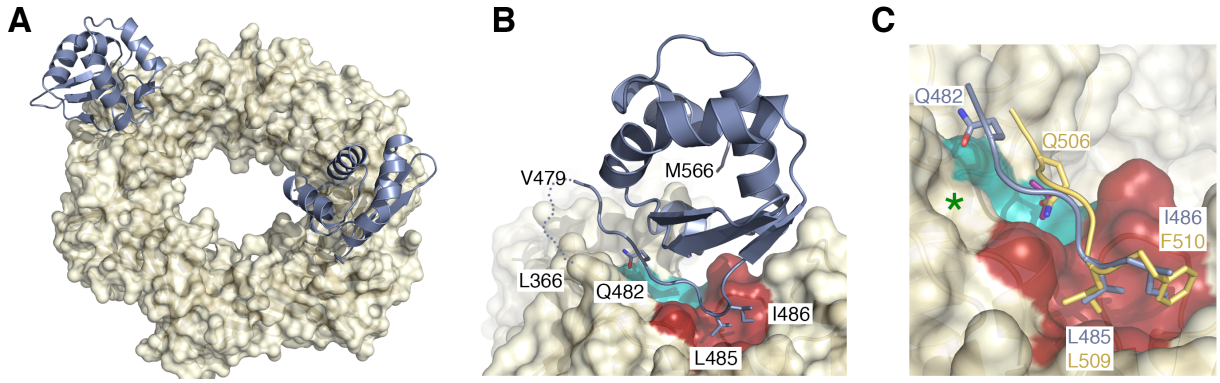

Figure 4.2. Interaction between *E. coli* MutL and the ß-clamp. (A) Crystal structure of the *E. coli* clamp-MutL fusion with the ß-clamp ring shown as a semi-transparent surface and the two regulatory domains of MutL bound to the ring shown as grey ribbons. **(B)** Detail of the interaction between the b- binding motif of *E. coli* MutL and the conserved subsites 1 (red) and 2 (cyan) on the surface of the ß-clamp. The conserved residues of the b-binding motif are shown as color-coded sticks and labeled. **(C)** Comparison of the interaction of *E. coli* MutL and *E. coli* pol II with the ß-clamp. The structure of the ß-clamp bound to pol II (PDB ID 3D1E) is superimposed onto the structure of the clamp-MutL fusion and the pol II peptide is shown in yellow with the conserved Gln, Leu and Phe residues shown as color-coded sticks. The sulfate ion occupying subsite 2 is shown in purple and a green asterisk marks the shallow pocket occupied by the conserved leucine found at position 2 in canonical CMBs.

Superimposition of MutL onto the regulatory subdomain of MutL (PDB ID 1X9Z) domain of the fusion structure results in a root mean square deviation (rmsd) of 0.495 Å (518 atoms), whereas superimposition of the ß-clamp monomer onto a structure of the ß-clamp bound to a canonical b-binding peptide yields an rmsd of 0.628 Å (2,211 atoms). For both protomers of the ring, the clamp- binding motif of MutL (^{482}QPLLIP^{487}) sits atop the binding site of the ß-clamp and

adopts identical orientations. The linker joining both halves of the fusion is disordered in both protomers of the dimer, indicating that the association between the two proteins is held together by the specific interactions between clamp-binding motif of MutL and the ß-clamp rather than constrains imposed by the linker.

The conserved Leu485 of the clamp-binding motif occupies the conserved hydrophobic pocket on the clamp known as subsite 1 (**Figure 4.2B**). Ile486 is also inside this hydrophobic pocket (**Figure 4.2C**). However, the presence of an isoleucine rather than the canonical phenylalanine found at this position imposes a small rotation of the motif around Leu485 that forces the other end of the motif to seesaw towards the rim of the clamp (**Figure 4.2B-C**). This movement pushes the conserved glutamine residue (Gln482) out of the second binding pocket (**Figure 4.2B-C**). Canonical CBMs have a conserved leucine following the glutamine (**Figure 4.1A**), which occupies a shallow pocket opposite to Gln-binding pocket and stabilizes the bidentate interaction between the ß-clamp and its binding partners (**Figure 4.2C**). The second position of the motif is not conserved in the CBM of MutL (**Figure 4.1A**). *E. coli* MutL harbors a proline at this position, and the smaller size of this residue may enable the monodentate interaction with the ß-clamp. In good agreement, point mutations on the conserved Gln482 do not affect mismatch repair function *in vivo* (Pillon et al., 2011).

A sulfate ion from the crystallization solution occupies subsite 2 (**Figure 4.3A**). This sulfate ion is part of a hydrogen-bond network involving the side chains of Asn320 from the ß-clamp and Arg531 and Gln532 from MutL, as well as water-mediated hydrogen bonds with His175, Asn320 and Met362 (**Figure 4.3A**).

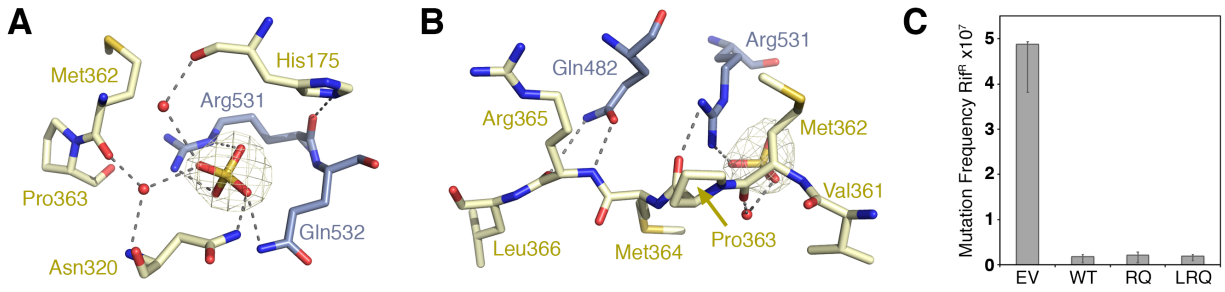

Figure 4.3. Stabilization of subsite 1. (A) Hydrogen-bond network stabilizing the interaction of the sulfate ion found in subsite 2. Residues from MutL and the ß-clamp are, respectively, shown in grey and pale-yellow, and hydrogen-bonds are shown as dashed lines. Water molecules are shown as red spheres. The 2Fo-Fc electron density map for the sulfate ion is shown as a gold mesh contoured at 1.2 s. **(B)** Detail of the interaction of the conserved Gln482 of MutL and the C-terminal strand of the ß-clamp. Additional interactions stabilizing subsite 2 are shown. **(C)** Mutation frequency associated with the MutL-RQ531SA (RQ) and MutL-L528A/RQ531SA (LRQ) variants compared to a *mutL*-deficient *E. coli* strain complemented with an empty vector (EV) or MutL wild-type (WT).

Additionally, the side chain of Leu528 from MutL is sandwiched between the side chains of Phe278 and Met364 from the ß-clamp and stabilizes the other end subsite 2. When the conserved glutamine of the clamp-binding motif occupies subsite 2, as it does in the case of Pol II (**Figure 4.2C**), its side chain forms a hydrogen bond with the carbonyl group of Met362, effectively restricting the width of the groove. Here the sulfate ion forms a water-mediated hydrogen bond with the same group, and the guanidinium group of Arg531 further constrains the pocket by forming a hydrogen-bond with the carbonyl group of Pro363 (**Figure 4.3B**). Additionally, the side chain of Gln482 forms a bidentate hydrogen bond with the main chain of Arg365 from the ß-clamp (**Figure 4.3B**). Therefore, while the interactions in subsite 2 are not conserved, the overall width of the binding groove is similar to other structures of the ß-clamp bound to clamp-binding motifs.

4.5.3 MutL residues interacting with subsite 2 are dispensable for the interaction with the ß-clamp

Given the roles of Arg531, Gln532 and Leu528 at stabilizing the conformation of subsite 2 in the structure, we checked whether point mutations on these residues affected mismatch repair activity *in vivo*. We generated a double-mutant of MutL (MutL-RQ531SA) unable to mediate the electrostatic interactions with the sulfate ion, as well as a triple variant (MutL-L528A/RQ531SA) in which both, electrostatic and van der Waals interactions, had been abrogated. We then measured the frequency of spontaneous mutation of *rpoB* to rifampicin resistance (Rif^R) for the wild-type and the variant *mutL* strains. Both variants displayed similar mutation frequencies to wild-type *mutL* (**Figure 4.3C**), indicating that the interactions mediated by these three residues with subsite 2 are not necessary for mismatch repair activity *in vivo*. We have previously shown that disruption of the MutL-clamp interaction causes a mild- mutator phenotype in *E. coli* (Pillon et al., 2011), therefore we concluded that the Arg531, Gln532 and Leu528 residues of MutL are not necessary for the functional interaction with the ß-clamp during mismatch repair.

4.5.4 *B. subtilis* MutL forms a bidentate interaction with the sliding ß-clamp

The interaction between *B. subtilis* MutL and the ß-clamp is essential for mismatch repair activity *in vivo* (Pillon et al., 2010). Therefore, we sought to determine whether *B. subtilis* MutL also forms a monodentate interaction with the ß-clamp. Crystals of the *B. subtilis* ß-clamp fused to the regulatory domain of MutL (residues 482-574) contained two ß-clamp dimers in the asymmetric unit. The two protomers of the ß-clamp dimer were virtually identical to each other

(rmsd of 0.438 Å over 2,165 atoms) and to the *B. subtilis* ß-clamp on its own (PDB ID 4TR6). The two regulatory domains of MutL ($MutL^{RGD}$) on each ß-clamp dimer interacted with the ß-clamp through the canonical bidentate interaction observed for other clamp-interacting partners (**Figure 4.4A-B**).

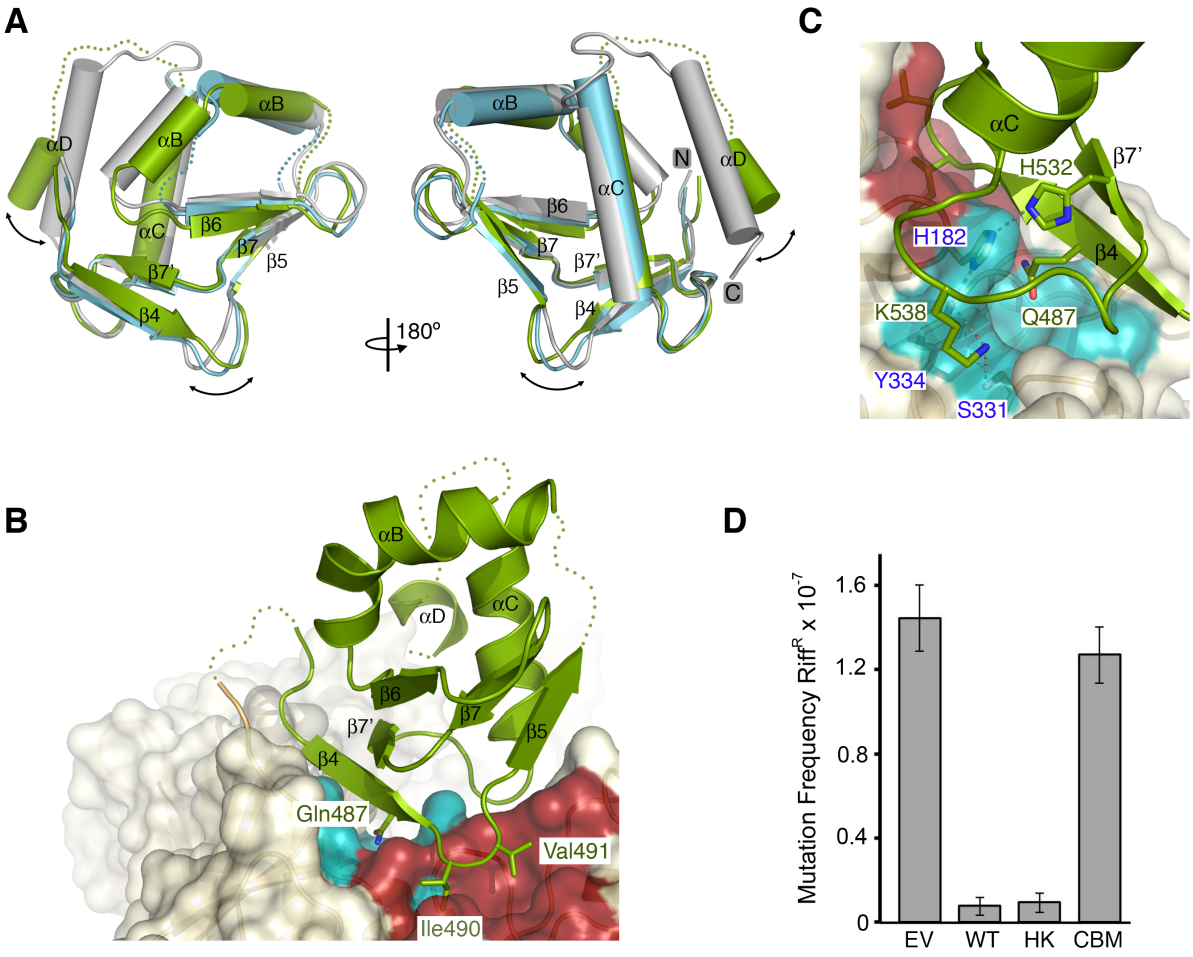

Figure 4.4. Interaction between *B. subtilis* MutL and the ß-clamp. (A) Opposite views of the regulatory domains of MutL from the structure of the *B. subtilis* clamp-MuL fusion (green and blue) superimposed onto the structure of $MutL^{RGD}$ (grey, PDB ID 3KDK). **(B)** Detail of the crystal structure of the *B. subtilis* clamp-MutL fusion with the ß-clamp ring shown as a semi-transparent surface and the regulatory domain of MutL shown as a green ribbon. The conserved residues of the b-binding motif are shown as color- coded sticks and labeled. The conserved ß-clamp binding pockets are colored in red (subsite 1) and blue (subsite 2). **(C)** Detail of the residues of the b7-aC loop contributing to the stabilization of subsite 2. **(D)** Mutation frequency associated with the MutL-H532A/K538A (HK) and MutL-CBM (CBM) variants compared to a *mutL*-deficient strain complemented with an empty vector (EV) or MutL (WT).

The side chains of the conserved isoleucine and glutamine residues within the motif ([487]QEMIVP[492]) occupy subsites 1 and 2 of the ß-clamp. The presence of the linker does not determine the interaction because the linker connecting the *B. subtilis* ß-clamp and MutL halves of the fusion is only visible in one of the protomers of the dimer, but both copies share the same binding mode (**Figure 4.5**).

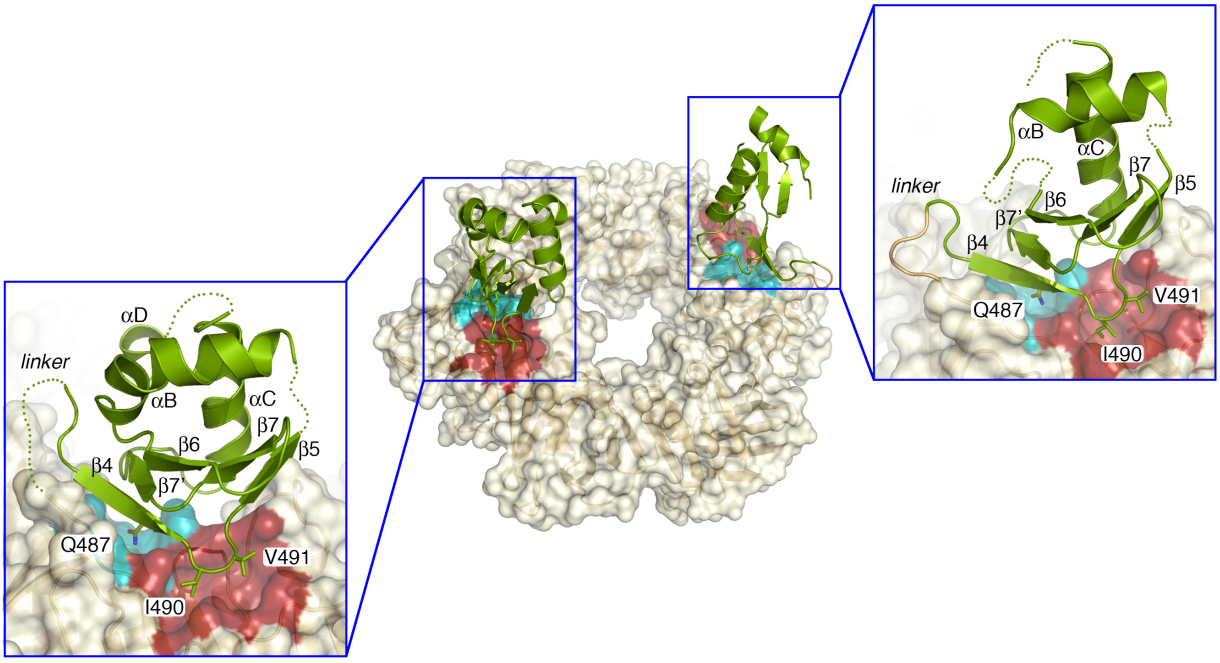

Figure 4.5. Interaction between *B. subtilis* MutL and the ß-clamp. Crystal structure of the *B. subtilis* clamp-MutL fusion with the ß-clamp ring shown as a semi-transparent surface and the two regulatory domains of MutL shown as green ribbons (centre). Detail of the interaction between the b-binding motif on each protomer of *B. subtilis* MutL and the conserved subsites 1 (red) and 2 (light blue) on the ß-clamp.

However, binding to the ß-clamp imposes significant rearrangements onto the regulatory domain of *B. subtilis* MutL. Superimposition of the domain onto the original structure of the endonuclease domain of MutL (PDB ID 3KDK) results in rmsd >1.5 Å, primarily caused by changes on the relative orientation of the clamp-binding motif and helices αB-αD (**Figure 4.4A**). The ß5-αB loop, as well as most of helix αD, are disordered in both protomers. In the protomer

with the ordered linker, the last turn of helix αB, as well as the loop connecting this helix to ß6, are also disordered (**Figure 4.4A**). The last two turns of helix αD, however, are visible in the protomer with the disordered linker (**Figure 4.4B**). This is the first structure of *B. subtilis* ß-clamp bound to one of its binding partners, but kinetic studies have shown that *B. subtilis* ß-clamp binds a model clamp-binding peptide with similar affinity to *E. coli* ß-clamp (28). This suggests that bidentate binding is also the canonical binding mode for *B. subtilis* clamp- binding partners. Our structure also reveals that formation of this bidentate interaction imposes important conformational rearrangements when the clamp-binding motif is embedded in a structured domain.

4.5.5 Residues beyond the clamp-binding motif are not required for the interaction of *B. subtilis* MutL and ß-clamp

The induced flexibility of the *B. subtilis* MutL regulatory domain allows the ^{487}QEMIVP^{492} motif to reach into the hydrophobic groove of the binding ß-clamp. This, in turn, brings the surrounding loops of the regulatory subdomain close to the surface of the ß-clamp. The loop connecting the ß7' strand to the αC helix wraps around subsite 2 defining an intricate hydrogen-bond network involving the side chains of His532 and Lys538 from MutL and the side chains of His182, Ser331 and Tyr334 from the ß-clamp, as well as the main chain carbonyl of Trp535 (MutL) and the amino group of Ala287 (ß-clamp) (**Figure 4.4C**). To test the relevance of this hydrogen-bond network, we generated a strain of *mutL* (*mutL^{HK}*) where His532 and Lys538 had been mutated to alanine residues. The D*mutL*/*mutL^{HK}* strain had similar mutation frequency to a D*mutL* strain complemented with wild-type MutL (D*mutL*/*mutL*$^{+}$), in contrast to the strain

with a mutated clamp-binding motif (D*mutL*/*mutL*^{CBM}) that had equivalent mutation frequency to the D*mutL* strain bearing the empty vector (**Figure 4.4D**). These results indicate that the increased flexibility of the domain enhances the shape and charge complementarity with the ß-clamp, but the functional interaction of MutL with the ß-clamp remains exclusively mediated by the clamp- binding domain.

4.5.6 Domain flexibility aligns the endonuclease motif of MutL with the central ß-clamp channel

The increased flexibility of the regulatory subdomain of *B*. *subtilis* MutL allows not only the clamp-binding motif to reach both binding subsites on the surface of the ß-clamp, but it also aligns the endonuclease site of MutL with the central cavity of the ß-clamp. Superimposition of the endonuclease domain of MutL (PDB ID 3KDK) using the clamp-binding motif as reference resulted in significant clashes between the second protomer of the MutL dimer and the ß-clamp (**Figure 4.6A**).

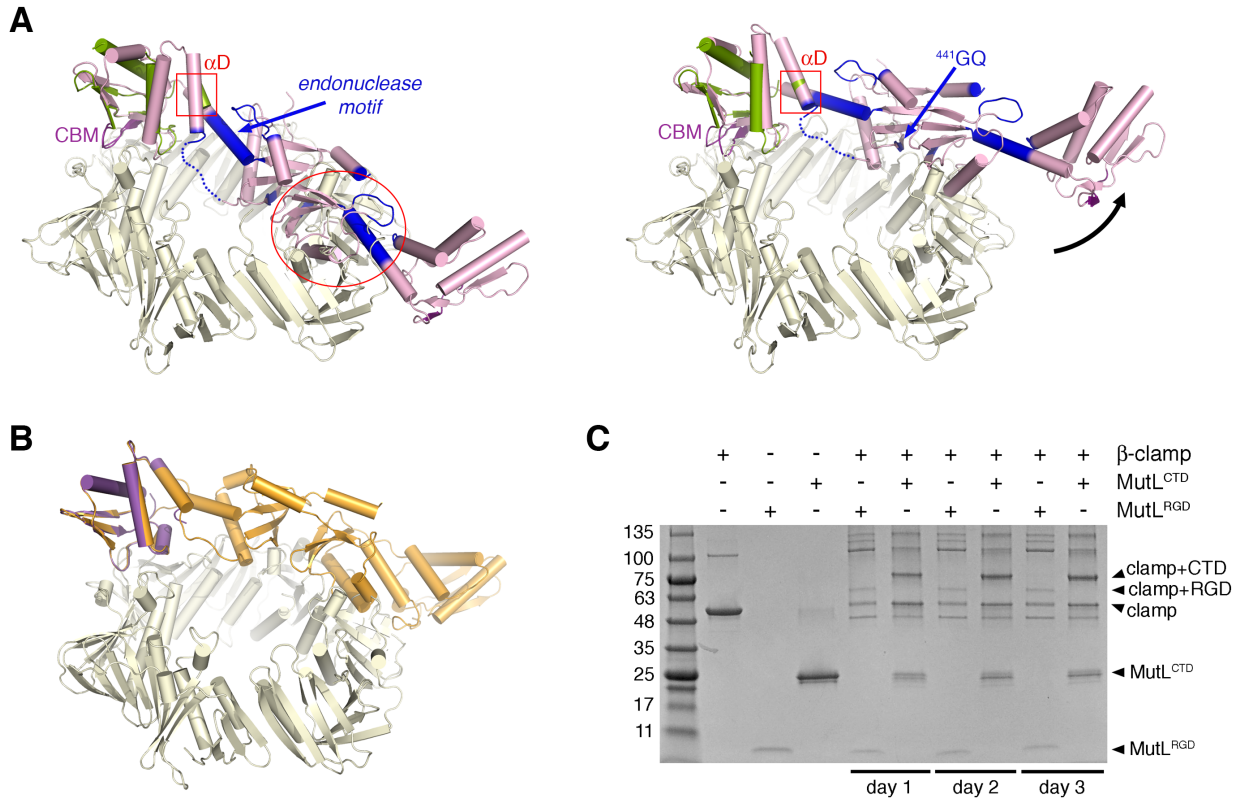

Figure 4.6. Flexibility of the MutL^{RGD} is necessary to align the endonuclease motif of MutL. (A) Ribbon diagram of the endonuclease domain dimer of *B. subtilis* MutL (pink) superimposed onto the *B. subtilis* clamp-MutL fusion (pale yellow and green) using either the clamp-binding motif (left) or helix aD (right) as reference. Re-orientation of helix aD aligns the endonuclease motif of MutL with the central cavity of the ß-clamp and prevents clashes with the second protomer of the dimer (red circle). **(B)** Ribbon diagram of the dimerization domain of *E. coli* MutL (gold) superimposed onto the *E. coli* clamp-MutL fusion (pale yellow and purple). **(C)** Variants of ß-clamp, and the endonuclease (MutL^{CTD}) and regulatory (MutL^{RGD}) including a single cysteine residue at the C-terminus (ß-clamp) or preceding the clamp binding motif (MutL-E485C) were purified and equimolar mixtures of either clamp-MutL^{CTD} or clamp- MutL^{RGD} were incubated in the absence of reducing agents. Samples withdrawn from the reaction at the indicated time points were resolved on denaturing gels in the absence of b-mercaptoethanol.

However, superimposition of the dimer taking into account the re-orientation of helix αD resolved most clashes and aligned the helix containing the nuclease motif of MutL with the central cavity of the ß-clamp (**Figure 4.6A**), reinforcing the idea that the clamp enhances the endonuclease activity of MutL by threading DNA onto its nuclease activity site (Pluciennik et al., 2010). This is also in good agreement with data showing that *B. subtilis* MutL only requires the

endonuclease site from the protomer bound to the ß-clamp for nuclease activity (Liu et al, unpublished). The dimerization domain of *E. coli* MutL could be directly superimposed onto the structure of the fusion without major clashes (**Figure 4.6B**), indicating that the conformational changes of the regulatory subdomain in *B. subtilis* MutL may be necessary the engage the second binding subsite.

MutL homologs harboring the endonuclease motif also harbor four additional motifs (Kosinski et al., 2008). The structure of the endonuclease domain of *B. subtilis* MutL revealed that three of these motifs (^{572}SCK, ^{604}CPHGRP, ^{623}FKR), together with the endonuclease motif (^{462}DQHAAQERIKYE), define the endonuclease site and coordinate the two zinc metal ions necessary for MutL activity (Pillon et al., 2010). The fourth motif (^{443}GQ) is located at the dimerization interface of MutL and could not be assigned a specific function from the structure of the endonuclease domain of BsMutL. The interaction between MutL and the ß-clamp places this motif in close proximity of the α1-ß2 (Ala22-Thr33), ß3-ß4 (Ser49-Asp52) and α4-ß11 (Ser156-Gly166) loops of the ß-clamp and, therefore, the conserved small and polar residues at these positions may prevent steric hindrance. Based on this model, we predicted that additional surfaces beyond the regulatory domain of *B. subtilis* MutL may stabilize the interaction with the ß-clamp. Using single-cysteine variants of ß-clamp, as well as the endonuclease and regulatory domains of *B. subtilis* MutL (Pillon et al., 2015), we tested whether the endonuclease dimer interacted more readily with the ß-clamp than the regulatory domain. Incubation of the two variants with the ß-clamp in the absence of reducing agents resulted in the formation of crosslinked species at 63 kDa (for the regulatory domain) and 75 kDa (for the endonuclease domain), consistent with the interaction of both domains with the ß-clamp (**Figure 4.6C**).

However, the 75 kDa species was formed more efficiently and accumulated to a larger extent than the 63 kDa species, supporting the idea that the dimerization region of the domain helps stabilize the interaction of MutL with the ß-clamp.

4.6 DISCUSSION

The structures of the ß-clamp bound to MutL are the first describing how the ß-clamp recognizes internal clamp-binding motifs embedded within structured domains. *B. subtilis* MutL has the characteristic bidentate interaction with its ß-clamp, whereas *E. coli* MutL forms a monodentate interaction with its ß-clamp. Previous studies suggested that linear clamp-binding motifs bind to the *E. coli* ß-clamp in a two-step process that engages the hydrophobic subsite first, and this is followed by the orientation and binding of flanking residues to subsite 2 (Yin et al., 2013). The structure of the *E. coli* clamp-MutL fusion visualizes the first binding step. The interaction between *E. coli* MutL and the ß-clamp has only a minimal role in DNA mismatch repair *in vivo* (Pillon et al., 2011), and we presume that *E. coli* MutL does not need to engage the second subsite. Conversely, *B. subtilis* MutL depends on the interaction with the ß-clamp to enhance its endonuclease activity. Upon binding through subsite 1, conformational re-arrangements of the regulatory domain engage subsite 2. The increased flexibility of the domain also affords steric and electrostatic complementarity of the surface of MutL surrounding the CBM and the alignment of the endonuclease motif with the central cavity of the ß-clamp. Therefore, these structures recapitulate the relevance of the interaction in each species.

PCNA-interacting protein (PIP) motifs, MLH1-interacting protein (MIP) motifs and Rev1- interacting region (RIR) motifs all feature adjacent aromatic residues challenging the traditional paradigm where the two aromatic residues determine binding specificity to PCNA

(Boehm and Washington, 2016). This, in turn, suggests that MutL homologs retain the minimum motif able to confer binding specificity for both the ß-clamp or PCNA. Additionally, the motif includes a conserved proline residue – a feature that is commonly found in internal clamp binding motifs. While this proline helps expose the clamp binding motif to the solvent, flexibility is required to form a canonical bidentate interaction between *B. subtilis* MutL and the ß-clamp. We predict that domain flexibility may be a general feature in all MutL homologs harboring an endonuclease activity, as well as other proteins containing internal clamp-binding motifs. Docking of the nuclease domain of *Saccharomyces cerevisiae* MutLα (PDB ID 4E4W) onto the structure of PCNA does not result in steric clashes between MLH1 of the dimer, but the endonuclease motif of PMS1 is not aligned with the central cavity of PCNA (**Figure 4.7**).

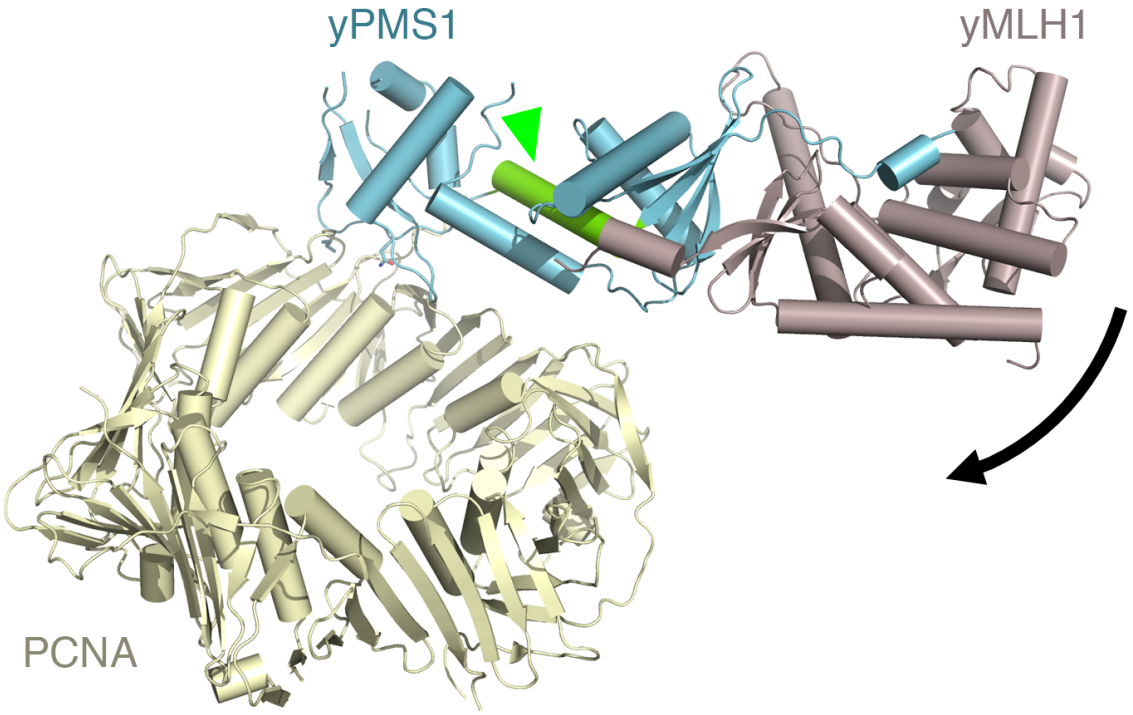

Figure 4.7. Model of the interaction between PCNA and yeast MutLα. Ribbon diagram of the endonuclease domain dimer of *S. cerevisiae* MutLα (yPMS1 (blue), yMLH1 (taupe)) superimposed onto human PCNA (pale yellow) using the PIP box of yPMS1 to guide the superimposition. The helix harboring the nuclease motif of yPMS1 is colored green.

PCNA enhances the endonuclease activity of human and yeast MutLα and provides directionality to the nicking reaction (Kadyrov et al., 2006, Kadyrov et al., 2007; Puciennik et al., 2010), reinforcing the idea that the endonuclease motif of PMS1 must be aligned with the central cavity of PCNA. Therefore, flexibility of the regulatory domain of yeast PMS1 and human PMS2 may also be necessary to align the endonuclease motif with the central cavity of PCNA.

The ß-clamp has been identified as a potential antimicrobial target (Kling et al., 2015). However, despite the structural conservation of the binding sites across species, different organisms recognize clamp-binding motifs with different affinities (Wolff et al., 2014). The structures of the *E. coli* and *B. subtilis* clamp-MutL fusions reveal two different interacting modes for the same motif, suggesting that despite the conservation of the motif and the ß-clamp subsites, interactions beyond the binding motif may modulate the interaction. Work with short peptides cannot recapitulate these additional interactions and, therefore, structural characterization of larger clamp-bound complexes remains a priority to explore the true potential of the ß-clamp as a drug target.

4.7 ACKNOWLEDGEMENTS

We thank former and current members of the Guarné laboratory for stimulating discussions. This work was funded by grants from the Canadian Institutes of Health Research (MOP-67189) and the Natural Sciences and Engineering Research Council of Canada (288295) to AG, NIH R01 GM066094 to MDS and NSF grants MCB1050948 and MCB1714539 to LAS. JRR was supported in part by NIH Cellular Biotechnology Training Grant (T32 GM008353) and a pre-doctoral fellowship from the Rackham Graduate School at the University of Michigan.

CHAPTER FIVE

UNDERSTANDING HOW CHIMERAS STABILIZE LOW-AFFINITY PROTEIN-PROTEIN INTERACTIONS.

Authors: Ahmad W. Almawi, Rodolfo Ghirlando, and Alba Guarné.

5.1 CONNECTING TEXT

Chapter 5 investigates how chimeras stabilize low-affinity protein-protein interactions either intermolecularly (as seen for the Dbf4-Rad53 complex) or intramolecularly (as seen for the clamp-MutL complex). We examined several structures of intermolecularly- or intramolecularly-linked complexes from the protein data bank and found differences in their linker sequence, length, and more importantly, the position of the linker relative to the weak interface. Dr. Rodolfo Ghirlando (National Institutes of Health) conducted, analyzed, and interpreted the ultracentrifugation experiments. I conducted the experiments, analyzed the data, and wrote the chapter.

5.2 DISCUSSION

Protein-protein interactions drive virtually all signaling events that regulate downstream cellular processes. Protein interactions are diverse and can be categorized as constitutively-bound, high-affinity, low-affinity, and transient (Chichili et al., 2013). Low-affinity or transient interactions coordinate numerous cellular processes, especially when the binding partners associate dynamically. These types of interactions are implicated in the replication checkpoint response, DNA mismatch repair, calcium metabolism, phage infection, chaperone maintenance, viral polymerase function, and regulation of neuronal receptors (Antczak et al., 2006; Kingston et al., 2004; Kumar et al., 2013; Matthews et al., 2014; Pillon et al., 2015; Williams et al., 2014; Ye et al., 2005).

Through structural biology techniques, such as X-ray crystallography and electron microscopy, protein complexes can be characterized at the molecular level thereby increasing our understanding of how proteins interact. However, the use of structural biology techniques requires a strictly homogenous sample. Achieving sample homogeneity for low-affinity or transient complexes is challenging because these complexes yield a mixture of predominately monomeric binders together with the actual complex (Qin et al., 2014). In X-ray crystallography, the protein concentrations can be increased to displace the equilibrium towards the complex thereby favoring its crystallization. However, high protein concentrations often trigger aggregation and precipitation. Moreover, increasing protein concentration is sometimes unsuitable for electron microscopy. Given the difficulty in stabilizing low-affinity or transient interactions, limited structural studies describing such complexes are reported in the protein data bank (PDB) (Qin et al., 2014).

Low-affinity or transient interactions can be studied in vitro by chemical crosslinking (Back et al., 2003; Matthews et al., 2014). This involves the formation of covalent bonds between the protein binding partners using bifunctional chemical crosslinkers that react with amino acid functional groups, such as primary amines and sulfhydryl groups (Sinz, 2006). However, the use of crosslinkers depends on the proximity of amino acids, including aspartate, cysteine, glutamate, and lysine, to the interaction site (Pierce, 2006). Optimization of crosslinking reactions is also necessary to reduce the formation of higher-order oligomers (Trakselis et al., 2005). In this regard, crosslinkers often generate a heterogeneous population of crosslinked products thereby rendering the sample unfitting for structural biology analysis (Chichili et al., 2013; Pierce, 2006; Trakselis et al., 2005). Chemical crosslinking is not limited to using bifunctional crosslinkers since several protein complexes were successfully characterized by creating disulfide bonds between their binding partners (Pillon et al., 2015). Moreover, an all-hydrocarbon cross-linking system was utilized to enhance the helicity and hence, metabolic stability of various peptides (Schafmeister et al., 2000).

An alternative approach in stabilizing low-affinity or transient complexes is by covalently joining the binding partners using a small peptide linker thereby forming a single polypeptide chain termed a chimera (Chichili et al., 2013). The linker acts as a spacer that bridges the binding partners (Chichili et al., 2013). Moreover, the linker is optimized for each complex to avoid constraining the conformation of the linked partners (Kumar et al., 2013; Williams et al., 2014). Chimeras stabilize interactions by forcing the binding partners to remain close to each other thereby displacing the equilibrium towards the complex (Matsushima et al., 2008). Using this approach, numerous complexes were successfully characterized through structural biology

techniques (Almawi et al., 2016; Antczak et al., 2006; Foss et al., 2005; Kingston et al., 2004; Wang et al., 2008; Williams et al., 2014; Ye et al., 2006). By keeping binding partners close to each other, chimeras can also stabilize intrinsically disordered proteins which gain a folded structure upon binding to their partner (Chichili et al., 2013). Chimeras are also widely used in industrial research owing to their ability to join independent domains to create a ligand-binding site or recognition sequence for a target protein (Chichili et al., 2013).

Chimeras can promote weak interactions either intermolecularly (as seen for the Dbf4-Rad53 complex; Chapter 2) or intramolecularly (as seen for the clamp-MutL complex; Chapter 4). However, the number of reported structures of intramolecularly-linked complexes is less than intermolecularly-stabilized complexes (**Table 5.1**). Upon examining these structures, we found differences in their linker properties including sequence and length (**Table 5.1**).

Table 5.1. Linker properties of chimeras.

Captured complex	Chimera	Linker sequence	Linker length
Intermolecular	Dbf4-L[#]-Rad53	SGASG	5
	Phosphoprotein P-L-Nucleocapsid N	GSGSGSGS	8
	RRS1-L-RPS4	GSGGS	5
	CaM-L-NgIQ	GGGGG	5
	CaM-L-Calcineurin	GGGGG	5
	G3P-L-TolA	GGGSEGGGSEGGG SEGGGSG	20
	Transthyretin-L-Transthyretin	GSGGGTGGGSG	11
Intramolecular	Ec^ Clamp-L-MutL	SGASG	5
	Bs* Clamp-L-MutL	VDS	3
	Asf1-L-H3α3	AGAATAA	7
	PXR-L-SRC1	SGGSGG	6

#. L: linker ^ Ec: *E. coli* * Bs: *B. subtilis*

The majority of intermolecularly- and intramolecularly-linked complexes harbor a glycine/serine-rich linker (**Table 5.1**) because these amino acids are known to act as independent units and do not affect the function of the individual proteins to which they are attached to (Chichili et al., 2013; Kumar et al., 2013). The structures presented in **Table 5.1** show that intermolecularly-linked complexes possess a more flexible sequence than intramolecularly-linked complexes. These chimeras must adopt oligomers (often dimers) to stabilize intermolecularly-linked complexes, which is achieved by the presence of flexible linkers.

Apart from their linking properties, glycine-rich linkers exist in many proteins in biological systems and sometimes instill a functional role for the protein (Xu et al., 1999). The human PAX6 transcription factor consists of two domains joined by a glycine-rich linker (Mishra et al., 2002). The crystal structure of the PAX6-DNA complex shows that the extended linker makes minor groove contacts with the DNA (Wilson et al., 2005). Various compositions of linkers were previously used for structural studies (Chichili et al., 2013). These comprised mostly glycine and other small side-chained amino acids. However, other amino acid compositions are being tested for possible use in the design of new artificial linkers (Matsushima et al., 2008).

In addition to having flexible linkers, the intermolecularly-linked complexes posses longer linkers (5-20 amino acids) than the intermolecularly-linked complexes (3-7 amino acids) (**Table 5.1**). The shorter linker forces the binding partners to be near each other, thus favoring intramolecular complex formation. On the other hand, intermolecularly-linked complexes can tolerate a broader range of linker lengths as it provides chimeras with the conformational

flexibility to form dimers. The structures of the dimers show that the interfaces of the weak complexes are shared, and thus form a continuous surface. Analyzing the interface properties of the dimers reveals a synergistic increase in the buried surface area, the interaction energy, and the number of contacts in comparison to the weak complex (**Table 5.2**).

Table 5.2. Interface properties of intermolecularly-linked complexes.

Intermolecularly-linked complex	Low-affinity complex				Dimer			
	IA# (A^2)	ΔG (kcal/mol)	HB^	SB*	IA (A^2)	ΔG (kcal/mol)	HB	SB
Dbf4-Rad53	790	-7	6	5	2,121	-20	12	5
Phosphoprotein P-Nucleocapsid N	634	-15	4	0	1,623	-29	8	2
RRS1-RPS4	709	-4	13	2	1,786	-8	26	6
CaM-NgIQ	763	-10	6	4	2,055	-25	13	10
CaM-Calcineurin	664	-8	5	10	1,689	-24	17	10

#. IA: interface area. ^. HB: number of hydrogen bonds. *. SB: number of salt bridges.

The interface area of the dimer is more than twice the interface area of the weak complex (**Table 5.2**). The increase in the interface area is entropically favorable and results in lowering of the energy of interaction of the dimer when compared to the weak complex (**Table 5.2**). Overall, dimerization enables the two weak complexes to share their interfaces, and thus create a continuous surface that synergistically improves the thermodynamic parameters (**Table 5.2**). Based on these observations, we anticipate that a longer linker will prevent the formation of a extended interface, thereby destabilizing the complex. To test this, we evaluated the weak self-association of a Dbf4-Rad53 chimera harboring a 30 amino acid linker (GSGSGGASGSGSGSGGSGSGGSGSGGSGSG), in a manner analogous to the Dbf4(5)Rad53 chimera (Almawi et al., 2016; Chapter 2). The Dbf4(30)Rad53 chimera was injected into the

Superdex 75 (S75) PC3.2/30 size-exclusion column at increasing protein concentrations (**Figure 5.1A**).

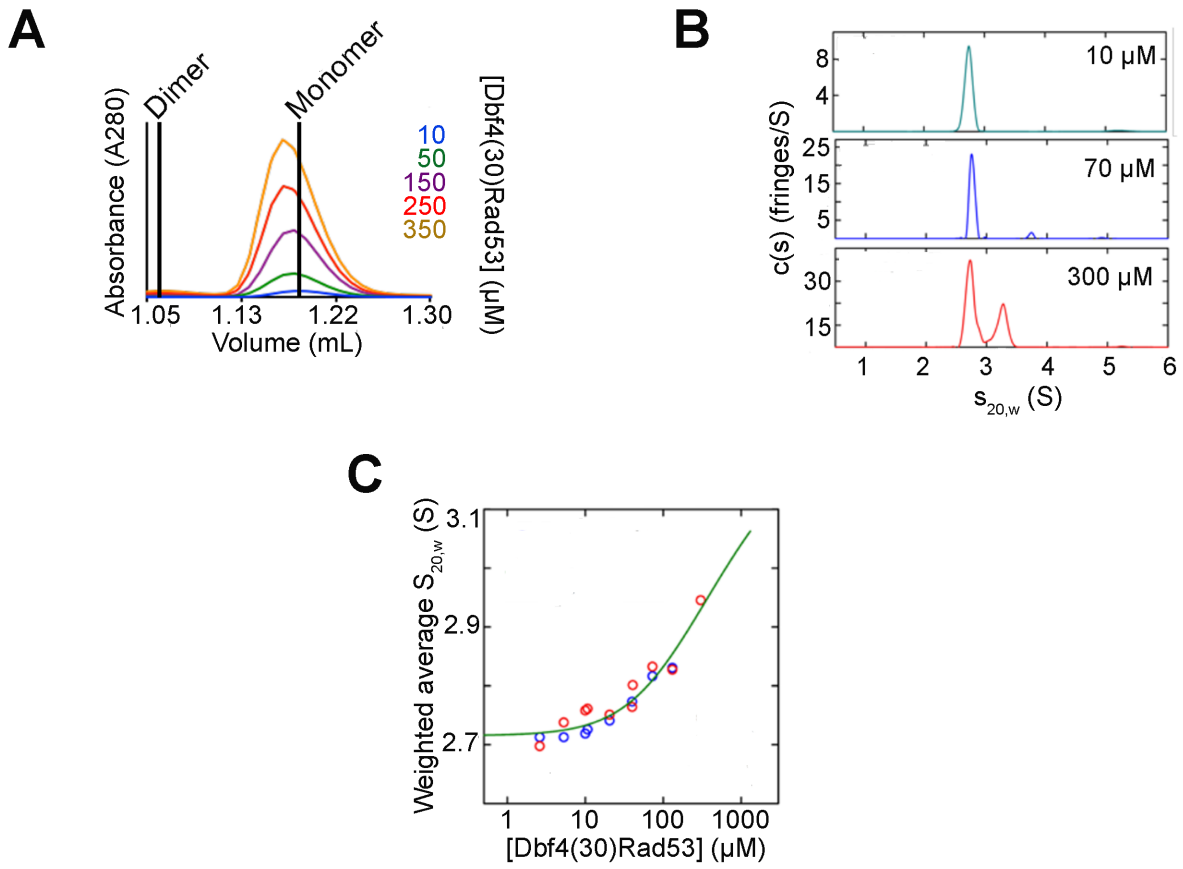

Figure 5.1. The Dbf4(30)Rad53 chimera has weaker self-association than Dbf4(5)Rad53. **(A)** Size exclusion chromatography profile of Dbf4(30)Rad53 at increasing protein concentrations. Elution volumes for a ideal Dbf4(30)Rad53 monomer and dimer are indicated as lines. **(B)** Normalized interference c(s) profiles for Dbf4(30)Rad53 at increasing protein concentrations. **(C)** Dependence of the weighted-average $s_{20,w}$ on the loading concentration for absorbance (blue) and interference (red) sedimentation velocity data, along with the global best-fit monomer-dimer isotherm (green) for Dbf4(30)Rad53.

The size-exclusion profile of Dbf4(30)Rad53 revealed that the chimera was predominantly monomeric (**Figure 5.1A**). However, the elution times of Dbf4(30)Rad53 varied in a concentration-dependent manner to a lesser extent than the Dbf4(5)Rad53 chimera (Chapter 2; **Figure 2.1B**), suggesting weaker self-association. We then performed a series of

sedimentation velocity experiments for the Dbf4(30)Rad53 chimera with increasing loading concentrations (**Figure 5.1B**). The sedimentation experiments confirmed the absence of large species at lower protein concentrations (**Figure 5.1B**), but the c(s) profiles showed a monomer-dimer equilibrium for Dbf4(30)Rad53 at higher concentration (300 μM) when compared to Dbf4(5)Rad53 (**Figure 5.1B** and Chapter 2; **Figure 2.1C**).

A weighted-average sedimentation coefficient isotherm was then constructed and analyzed regarding reversible monomer-dimer equilibrium for Dbf4(30)Rad53 yielding an estimated dissociation constant (Kd) of 590 μM (**Figure 5.1C**). Based on reduced chi-squared and F-statistics (Johnson, 1992), the estimated 68% confidence limits of the Kd value varied between 170 - 3,100 μM. These estimates were significantly higher than those calculated for Dbf4(5)Rad53 (Chapter 2; **Figure 2.1D**). Size-exclusion chromatography and ultracentrifugation analyses suggest that larger species were more prominent for shorter linkers. Collectively, this indicates weaker self-association for the Dbf4(30)Rad53 chimera when compared to Dbf4(5)Rad53 (**Figure 5.1A-C**; Chapter 2; **Figure 2.1B-D**). While this data explains the behavior of the Dbf4-Rad53 chimera, we believe that our findings may also apply to the other chimeras, although this awaits further validation.

Beyond stabilizing weak interactions, altering the linker length also affects the stability, solubility, oligomeric state, and proteolytic resistance of several chimeras (Chichili et al., 2013; Nagi et al., 1997). A chimera of Arc repressor protein, which naturally exists as a homodimer, was tested for its stability, protein folding kinetics, and biological activity by varying the length of the linker. While longer linkers provided maximum conformational freedom, the authors failed to ensure optimal stability for the Arc chimera (Robinson and Sauer, 1998).

Variations in linker sequence and length do not explain why the *E. coli* clamp-MutL complex was stabilized differently from the Dbf4-Rad53 complex (**Table 5.1**). Although the linker sequence and length is identical (SGASG) between both chimeras, the position of the linker relative to the interface of the weak complex is different (**Figure 5.2**).

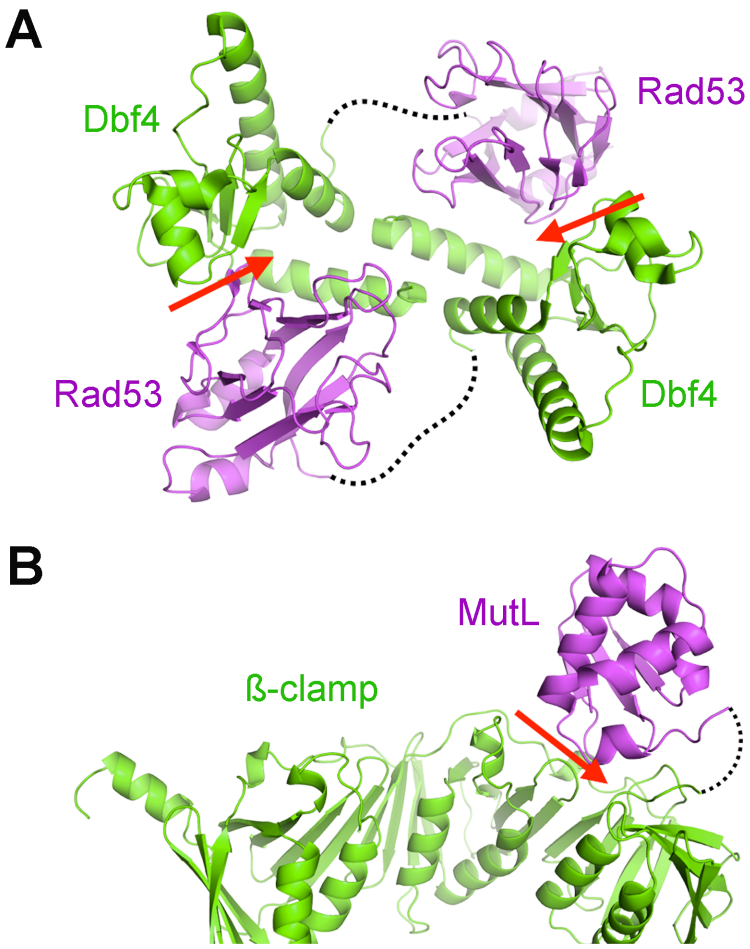

Figure 5.2. Linker position relative to the interface of the weak complex. The linker (dashed line) is positioned further away from the interface (red arrow) of the Dbf4 (green)-Rad53 (magenta) chimera **(A)** in comparison to the clamp (green)-MutL (magenta) chimera **(B)**. Structures are shown as ribbon diagrams.

The linker is positioned further away from the interface of the Db4-Rad53 complex (**Figure 5.2A**) when compared to the interface of the clamp-MutL complex (**Figure 5.2B**). This feature was observed for other chimeras, whereby the average distance between the linker and the

interface of the weak complex was 26 Å for intermolecularly-linked complexes (**Table 5.3**), which was roughly double the distance (13 Å) of intramolecularly-linked complexes (**Table 5.3**).

Table 5.3. Distance between the linker and the interface of the weak complex.

Captured complex	Chimera	Distance between linker and interface of weak complex (Å)	Average distance (Å)
Intermolecular	Dbf4-L-Rad53	26	26
	Phosphoprotein P-L-Nucleocapsid N	20	
	RRS1-L-RPS4	28	
	CaM-L-NgIQ	19	
	CaM-L-Calcineurin	24	
	G3P-L-TolA	36	
	Transthyretin-L-Transthyretin	31	
Intramolecular	Ec Clamp-L-MutL	14	13
	Bs Clamp-L-MutL	18	
	Asf1-L-H3α3	11	
	PXR-L-SRC1	10	

When the linker is close to the interface of the weak complex, the binding partners are ideally positioned to interact intramolecularly. On the other hand, positioning the linker further away from the interface displaces the binding partners thereby favoring dimerization of the chimera. Overall, analysis of linker properties reveals that the sequence, length, and the position of the linker relative to the weak interface are determining factors in stabilizing weak complexes either intermolecularly or intramolecularly.

When designing chimeras, the linker length and amino acid composition are considered, but the linker position is often ignored. The structures analyzed from the PDB show that the linker position relative to the weak interface influences how a weak complex is captured (**Table**

5.3). The conclusions derived from studying the Dbf4-Rad53 and clamp-MutL complexes is that linker position should be coupled with linker length and sequence when designing chimeras. Also, having information about the structure of the individual binding partners as well as their interface will be beneficial in deciding on the position of the linker. Ultimately, this may determine whether the complex is captured either intermolecularly or intramolecularly.

5.7 ACKNOWLEDGMENTS

We would like to thank Dr. Murray Junop and members of the Guarné lab for insightful discussions. This work was funded by the Canadian Institutes of Health Research (MOP-67189 to A.G.). A.W.A. was funded in part by an Ontario Graduate Scholarship.

CHAPTER SIX

CONCLUSIONS

6.1 The importance of analyzing weak protein-protein interactions

6.1.1 The Dbf4-Rad53 interface is asymmetric

The structures of the Dbf4-Rad53 and clamp-MutL complexes revealed that their interfaces represent more than just minimal contacts between binding partners. The Dbf4-Rad53 complex exhibited an asymmetric interface whereby Dbf4 was the major contributor to the formation of the complex. Asymmetric interfaces were previously reported for a variety of biological systems (Dey et al., 2010) and were not restricted to protein-protein interactions since they have been seen in protein-DNA complexes (Narayanan et al., 2014). Studying asymmetric interfaces also offers therapeutic potential, as exemplified with the HER2-HER3 complex in breast cancer (Claus et al., 2018).

6.1.2 The ß clamp-MutL interaction is different between *E. coli* and *B. subtilis*

Examining the *E. coli* and *B. subtilis* clamp-MutL complexes revealed differences in their binding modes even though the interacting partners were structurally similar. The structures of the weak complexes allowed us to explain how previous motifs of unknown function on *B. subtilis* MutL (Kosinski et al. 2008, Pillon et al., 2015) may contribute to the binding to the sliding clamp. Given the affinity of a weak complex, it is often discouraging to analyze it structurally. However, the structures of the Dbf4-Rad53 and clamp-MutL complexes revealed critical details that explained how each binding partner regulates the weak complex. These features would have been overlooked, and hence their biological significance missed had we not rigorously examined weak protein-protein interactions.

6.2 FHA and polo-box domains function beyond phosphoepitope recognition

FHA and polo-box domains were traditionally characterized as phosphoepitope binding domains. The Dbf4-Rad53 and Dbf4-Cdc5 structures revealed how these domains mediate phosphorylation-independent interactions at molecular detail. Identifying an additional binding surface on the polo-box domain of Cdc5, as well as the unique features that the FHA1 domain of Rad53 exploits in mediating canonical and non-canonical interactions has expanded our understanding of the binding modes involving these domains. Collectively, FHA and polo-box domains appear to be more plastic than initially anticipated. This idea was discussed for FHA domains in a review that we published recently in collaboration with Dr. Lindsay Matthews (Almawi et al., 2017). By demonstrating that the Cdc5 polo-box domain and that the Rad53 FHA1 domain utilize multiple surfaces to simultaneously engage substrates, we provide insight into why these domains are prevalent in proteins tasked with regulating cell cycle events, which must have the versatility to interact with numerous partners.

6.3 Future directions

One of the significant contributions of our work is identifying a novel and conserved binding surface on the Cdc5 polo-box domain. This raises the question regarding the function of this surface beyond its interaction with Dbf4. Whether this surface plays a role in regulating the Cdc5 kinase domain is the next avenue to explore for this project. We also found innovative ways to stabilize weak protein-protein interactions, and as such can use the information derived from linkers to design chimeras to study more complicated systems. In the case of mismatch repair, the lingering question in the field revolves around understanding how MutL cleaves DNA even though it binds weakly to its substrate. Using what we learned from stabilizing the clamp-

MutL interaction, we can derive intuitive strategies to stabilize MutL onto DNA for structural biology characterization.

REFERENCES

Adams PD, Afonine PV, Bunkoczi G, Chen VB, Davis IW, Echols N, Headd JJ, Hung LW, Kapral GJ, Grosse-Kunstleve RW, McCoy AJ, Moriarty NW, Oeffner R, Read RJ, Richardson DC, Richardson JS, Terwilliger TC, and Zwart PH. (2010). PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Cryst D 66:213-21.

Afonine PV, Grosse-Kunstleve RW, and Adams PD. (2005). phenix.refine. CCP4 Newsletter 42:contribution 8.

Ahn J and Prives C. (2002). Checkpoint kinase 2 (Chk2) monomers or dimers phosphorylate Cdc25C after DNA damage regardless of threonine 68 phosphorylation. J Biol Chem 277:48418-26.

Ahn JY, Li X, Davis HL, and Canman CE. (2002). Phosphorylation of threonine 68 promotes oligomerization and autophosphorylation of the Chk2 protein kinase via the forkhead-associated domain. J Biol Chem 277:19389-95.

Alberts B, Johnson A, Lewis J, Raff M, Roberts K, and Walter P. (2002). Molecular Biology of the Cell. Garland Science. ISBN 0-8153-3218-1. Chapter 5: DNA Replication Mechanisms.

Allen JB, Zhou Z, Siede W, Friedberg EC, and Elledge SJ. (1994). The SAD1/RAD53 protein kinase controls multiple checkpoints and DNA damage-induced transcription in yeast. Genes Dev 8:2401-15.

Almawi AW, Matthews LA, and Guarné A. (2017). FHA domains: phosphopeptide binding and beyond. Prog Biophys Mol Biol. 127:105-10.

Almawi AW, Matthews LA, Larasati, Myrox P, Boulton S, Lai C, Moraes T, Melacini G, Ghirlando R, Duncker BP, and Guarné A. (2016). ‘AND’ logic gates at work: Crystal structure of Rad53 bound to Dbf4 and Cdc7. Sci Rep 6:34237.

Antczak AJ, Tsubota T, Kaufman PD, Berger JM. (2006). Structure of the yeast histone H3-ASF1 interaction: implications for chaperone mechanism, species-specific interactions, and epigenetics. BMC Struct Biol 6:26-32.

Archambault V, and Glover DM (2009). Polo-like kinases: conservation and divergence in their functions and regulation. Nat Rev Mol Cell Biol 10:265–75.

Archambault V, D'Avino PP, Deery MJ, Lilley KS, and Glover DM. (2008). Sequestration of Polo kinase to microtubules by phophopriming-independent binding to Map205 is relieved by phosphorylation at a CDK site in mitosis. Genes Dev 22:2707-20.

Archambault V, Lépine G, and Kachaner D. (2015). Understanding the Polo Kinase machine. Oncogene 34:4799-807.

Argunhan B, Leung WK, Afshar N, Terentyev Y, Subramanian VV, Murayama Y, Hochwagen A, Iwasaka H, Tsubouchi T, and Tsubouchi H. (2017). Fundamental cell cycle kinases collaborate to ensure timely destruction of the synaptonemal complex during meiosis. EMBO J 36:2488-509.

Armstrong AA, Mohideen F, and Lima CD. (2012). Recognition of SUMO-modified PCNA requires tandem receptor motifs in Srs2. Nature 483:59-63.

Aucher W, Becker E, Ma E, Miron S, Martel A, Ochsenbein F, Marsolier-Kergoat MC, and Guerois R. (2010). A strategy for interaction site prediction between phospho-binding modules and their partners identified from proteomic data. Mol Cell Proteomics 9:2745-59.

Bahassi M. (2011). Polo-like kinases and DNA damage checkpoint: beyond the traditional mitotic functions. Exp Biol Med 236:648–657.

Bardin AJ, Visintin R, Amon A. (2000). A mechanism for coupling exit from mitosis to partitioning of the nucleus. Cell 102:21-31.

Bartek J, Bartkova J, and Lukas J. (2007). DNA damage signalling guards against activated oncogenes and tumour progression. Oncogene 26:7773-9.

Bartek J and Lukas J. (2003). Chk1 and Chk2 kinases in checkpoint control and cancer. Cancer Cell 3:421-9.

Bassermann F, Frescas D, Guardavaccaro D, Busino L, Peschiaroli A, and Pagano M. (2008). The Cdc14B-Cdh1-Plk1 axis controls the G2 DNA-damage-response checkpoint. Cell 134:256-67.

Battye TGG, Kontogiannis L, Johnson O, Powell HR, and Leslie AGW. (2011). IMosflm: a new graphical interface for diffraction-image processing with MOSFLM. Acta Cryst 67:271-81.

Bertalanffy FD and Lau C. (1962). Rates of cell division of transplantable malignant rat tumors. Cancer Res 22:627-31.

Blow JJ, Ge XQ, and Jackson DA. (2011). How dormant origins promote complete genome replication. Trends Biochem Sci 36:405-14.

Boehm EM and Washington MT. (2016). R.I.P. to the PIP: PCNA-binding motif no longer considered specific: PIP motifs and other related sequences are not distinct entities and can bind multiple proteins involved in genome maintenance. Bioessays 38:1117-22.

Booher RN, Deshaies RJ, and Kirschner MW. (1993). Properties of Saccharomyces cerevisiae wee1 and its differential regulation of p34CDC28 in response to G1 and G2 cyclins. EMBO J 12:3417-26.

Botchkarev VV Jr and Haber J. (2018). Functions and regulation of the Polo-like kinase Cdc5 in the absence and presence of DNA damage. Curr Genet 64:87-96.

Bowers JL, Randell JC, Chen S, and Bell SP. (2004). ATP hydrolysis by ORC catalyzes reiterative Mcm2-7 assembly at a defined origin of replication. Mol Cell 16:967-78.

Boxem M. (2006). Cyclin-dependent kinases in C. elegans. Cell Div 12:1-6.

Branzei D and Foiani M. (2005). The DNA damage response during DNA replication. Curr Opin Cell Biol 17:568-75.

Branzei D and Foiani M. (2006). The Rad53 signal transduction pathway: Replication fork stabilization, DNA repair, and adaptation. Exp Cell Res 312:2654-9.

Bruinsma W, Raaijmakers JA, and Medema RH. (2013). Switching Polo-like kinase-1 on and off in time and space. Trends Biochem Sci 37:534-42.

Bunting KA, Roe SM, and Pearl LH. (2003). Structural basis for recruitment of translesion DNA polymerase Pol IV/DinB to the beta-clamp. EMBO J 22:5883-92.

Burgers PMJ, Gordenin D, and Kunkel TA. (2016). Who is leading the replication fork, Pol epsilon or Pol delta? Mol Cell 61:492-3.

Burkard ME, Maciejowski J, Rodriguez-Bravo V, Repka M, Lowery DM, Clauser KR, Zhang C, Shokat KM, Carr SA, Yaffe MB, and Jallepalli PV. (2009). Plk1 self-organization and priming phosphorylation of HsCYK-4 at the spindle midzone regulate the onset of division in human cells. PLoS Biol 7:E1000111-E1000111.

Burnouf DY et al. (2004). Structural and biochemical analysis of sliding clamp/ligand interactions suggest a competition between replicative and translesion DNA polymerases. J Mol Biol 335:1187-97.

Cai Z, Chehab NH, and Pavletich NP. (2009). Structure and activation mechanism of the CHK2 DNA damage checkpoint kinase. Mol Cell 35:818-29.

Callebaut I and Mornon JP. (1997). From BRCA1 to RAP1: a widespread BRCT module closely associated with DNA repair. FEBS Lett 400:25-30.

Casper AM, Nghiem P, Arlt MF, and Glover TW. (2002). ATR regulates fragile site stability. Cell 111:779-789.

Charles JF, Jaspersen SL, Tinker-Kulberg RL, Hwang L, Szidon A, and Morgan DO. (1998). The Polo-related kinase Cdc5 activates and is destroyed by the mitotic cyclin destruction machin- ery in S. cerevisiae. Curr Biol 8:497-507.

Cheng KY, Lowe ED, Sinclair J, Nigg EA, and Johnson LN. (2003). The crystal structure of the human polo-like kinase-1 polo box domain and its phopho-peptide complex. EMBO J 22:5757-68.

Cheng L, Hunke L, and Hardy CF. (1998). Cell cycle regulation of the Saccharomyces cerevisiae Polo-like kinase cdc5p. Mol Cell Biol 18:7360-70.

Chen S, de Vries MA, and Bell SP. (2007). Orc6 is required for dynamic recruitment of Cdt1 during repeated Mcm2-7 loading. Genes Dev 21:2897-907.s

Chen YC and Weinreich M. (2010). Dbf4 regulates the Cdc5 Polo-like kinase through a distinct non-canonical binding interaction. J Biol Chem 285:41244-54.

Chen YC et al. (2013). DNA replication checkpoint signaling depends on a Rad53-Dbf4 N-terminal interaction in Saccharomyces cerevisiae. Genetics 194:389-401.

Chichili VPR, Kumar V, and Sivaraman J. (2013). A method to trap transient and weak interacting protein complexes for structural studies. Intrinsically Disord Proteins 1:e25464.

Chichili VPR, Kumar V, and Sivaraman J. (2013). Linkers in the structural biology of protein-protein interactions. Protein Sci 22:153-67.

Claus J et al. (2018). Inhibitor-induced HER2-HER3 heterodimerisation promotes proliferation through a novel dimer interface eLIFE 1:e32271.

Cole JL, Lary JW, Moody TP, and Laue TM. (2008). Analytical ultracentrifugation: Sedimentation velocity and sedimentation equilibrium. Biophysical Tools for Biologists: Vol 1 in vitro Techniques 84:143-79.

Cooper GM. (2000).The cell: a molecular approach (2nd ed.). Washington, D.C: ASM Press. ISBN 0-87893-106-6.

Costa A, Ilves I, Tamberg N, Petojevic T, Nogales E, Botchan MR, and Berger JM. (2011). The structural basis for MCM2-7 helicase activation by GINS and Cdc45. Nat Struct Mol Biol 18:471-7.

Craig SN, Wyatt MD, McInnes C. (2014). Current assessment of polo-like kinases as anti-tumor drug targets. Expert Opin Drug Discov 9:773-89.

Dalrymple BP, Kongsuwan K, Wijffels G, Dixon NE, and Jennings PA. (2001). A universal protein- protein interaction motif in the eubacterial DNA replication and repair systems. Proc Natl Acad Sci U S A 98:11627-32.

Defendi V and Manson LA. (1963). Analysis of the life-cycle in mammalian cells. Nature 198:359-61.

Delaglio F et al. (1995). NMRPipe: a multidimensional spectral processing system based on UNIX pipes. J Biomol NMR 6:277-93.

DeLano WL. (2002). The PyMOL Molecular Graphic Systems. DeLano Scientific, Palo Alto, CA, USA.

Dey S, Pal A, Chakrabarti P, and Janin J. (2010). The subunit interfaces of weakly associated homodimeric proteins. J Mol Biol 398:146-60.

Diffley JF. (2011). The Many Faces of Redundancy in DNA Replication Control. Cold Spring Harb Symp Quant Biol 75:135-42.

Dixon WJ and Massey FJ. (1969). Introduction to Statistical Analysis (McGraw-Hill, New York).

Dohrmann PR and McHenry CS. (2005). A bipartite polymerase-processivity factor interaction: only the internal beta binding site of the alpha subunit is required for processive replication by the DNA polymerase III holoenzyme. J Mol Biol 350:228-39.

Dohrmann PR and Sclafani RA. (2006). Novel Role for Checkpoint Rad53 Protein Kinase in the Initiation of Chromosomal DNA Replication in Saccharomyces cerevisiae. Genetics 174:87-99.

Dowell SJ, Romanowski P, and Diffley JF. (1994). Interaction of Dbf4, the Cdc7 protein kinase regulatory subunit, with yeast replication origins in vivo. Science 265:1243-6.

Duch A, Palou G, Jonsson ZO, Palou R, Calvo E, Wohlschlegel J, and Quintana DG. (2011). A Dbf4 mutant contributes to bypassing the Rad53-mediated block of origins of replication in response to genotoxic stress. J Biol Chem 286:2486-91.

Duncker BP, Shimada K, Tsai-Pflugfelder M, Pasero P, and Gasser SM. (2002). An N-terminal domain of Dbf4p mediates interaction with both origin recognition complex (ORC) and Rad53p and can deregulate late origin firing. Proc Natl Acad Sci U S A 99:16087-92.

Durkin SG and Glover JN. (2007). Chromosome fragile sites. Annu Rev Genet 41:169-192.

Durocher, D., and Jackson, S.P. (2002). The FHA domain. FEBS Lett 513:58-66.

Durocher D, Taylor IA, Sarbassova D, Haire LF, Westcott SL, Jackson SP, Smerdon SJ, and Yaffe MB. (2000). The molecular basis of FHA domain:phosphopeptide binding specificity and implications for phospho-dependent signaling mechanisms. Mol Cell 6:1169-82.

D'Amours D, Stegmeier F, and Amon A. (2004). Cdc14 and condensin control the dissolution of cohesin-independent linkages at repeated DNA. Cell 117:455-69.

D'Amours D and Amon A. (2004). At the interface between signaling and executing anaphase—Cdc14 and the FEAR network. Genes Dev 18:2581-95.

Elia, AE, Rellos P, Haire LF, Chao, JW, Ivins FJ, Hoepker K, Mohammad D, Cantley LC, Smerdon SJ, and Yaffe MB. (2003). The molecular basis for phosphodependent substrate targeting and regulation of Plks by the Polo-box domain. Cell 115:83-95.

Elia AE, Cantley LC, and Yaffe, MB. (2003). Proteomic screen finds pSer/pThr-binding domain localizing Plk1 to mitotic substrates. Science 299:1228–31.

Emsley P, Lohkamp B, Scott WG, and Cowtan K. (2010). Features and development of Coot. Acta Crystallogr D Biol Crystallogr 66:486-501.

Enserink JM. (2011). Cell cycle regulation of DNA replication in S. cerevisiae. DNA Rep Curr Adv 4:391-408.

Fay DS, Sun Z, and Stern DF. (1997). Mutations in SPK1/RAD53 that specifically abolish checkpoint but not growth-related functions. Curr Genet 31:97-105.

Feng H and Kipreos ET. (2003). Preventing DNA re-replication--divergent safeguards in yeast and metazoa. Cell Cycle 2:431-4.

Foss TR, Kelker MS, Wiseman RL, Wilson IA, Kelly JW. (2005). Kinetic stabilization of the native state by protein engineering: implications for inhibition of transthyretin amyloidogenesis. J Mol Biol:841-54.

Foster PL. (2006). Methods for determining spontaneous mutation rates. Methods Enzymol 409:195-213.

Fragkos M, Ganier O, Coulombe P, and Méchali M. (2015). DNA replication origin activation in space and time. Nat Rev Mol Cell Biol 16:360-74.

Friedel AM, Pike BL, and Gasser SM. (2009). ATR/Mec1: coordinating fork stability and repair. Curr Opin Cell Biol 21:237-44.

Galgoczy DJ and Toczyski DP. (2001). Checkpoint adaptation precedes spontaneous and damage-induced genomic instability in yeast. Mol Cell Biol 21:1710-8.

Gelfant S. (1962). Initiation of mitosis in relation to the cell division cycle. Exp Cell Res 26:395-403.

Genschel J et al. (2017). Interaction of proliferating cell nuclear antigen with PMS2 is required for MutLα activation and function in mismatch repair. Proc Natl Acad Sci U S A 114:4930-5.

Georgescu RE, Yurieva O, Kim SS, Kuriyan J, Kong XP, and O'Donnell M. (2008). Structure of a small-molecule inhibitor of a DNA polymerase sliding clamp. Proc Natl Acad Sci U S A 105:11116-21.

Georgescu RE et al. (2008). Structure of a small-molecule inhibitor of a DNA polymerase sliding clamp. Proc Natl Acad Sci U S A 105:11116-21.

Gest H. (2004). The discovery of microorganisms by Robert Hooke and Antoni Van Leeuwenhoek, fellows of the Royal Society. Notes and Records of the Royal Society of London 58:187-201.

Geymonat M, Spanos A, Walker PA, Johnston LH, and Sedgwick SG. (2003). In vitro regulation of budding yeast Bfa1/Bub2 GAP activity by Cdc5. J Biol Chem 278:14591-4.

Ghirlando R. et al. (2013). Improving the thermal, radial, and temporal accuracy of the analytical ultracentrifuge through external references. Anal Biochem 440:81-95.

Gjertsen BT and Schoffski P. (2015). Discovery and development of the Polo-like kinase inhibitor volasertib in cancer therapy. Leukemia 29:11-9.

Golsteyn RM, Mundt KE, Fry AM, and Nigg EA. (1995). Cell cycle regulation of the activity and subcellular localization of Plk1, a human protein kinase implicated in mitotic spindle function. J Cell Biol 129:1617-28.

Guarne A, Ramon-Maiques S, Wolff EM, Ghirlando R, Hu X, Miller JH, and Yang W. (2004). Structure of the MutL C-terminal domain: a model of intact MutL and its roles in mismatch repair. EMBO J 23:4134-45.

Hall BM, Ma CX, Liang P, and Singh KK. (2009). Fluctuation analysis CalculatOR: a web tool for the determination of mutation rate using Luria-Delbruck fluctuation analysis. Bioinformatics 25:1564-5.

Hanahan D and Weinberg RA. (2000). The hallmarks of cancer. Cell 100:57-70.

Hanahan D and Weinberg RA. (2011). Hallmarks of cancer: the next generation. Cell 144:646-74.

Harkins V, Gabrielse C, Haste L, and Weinreich M. (2009). Budding Yeast Dbf4 Sequences Required for Cdc7 Kinase Activation and Identification of a Functional Relationship Between the Dbf4 and Rev1 BRCT Domains. Genetics.

Hartwell, LH, and Weinert TA. (1989). Checkpoints: controls that ensure the order of cell cycle events. Science 246:629-34.

Hartwell LH, Mortimer RK, Culotti J, and Culotti M. (1973). Genetic control of the cell division cycle in yeast: V. Genetic analysis of cdc mutants. Genetics 74:267-86.

Hartwell LH. (1974). Saccharomyces cerevisiae cell cycle. Bacteriol Rev 38:164-198.

Hasty J, McMillen D, and Collins JJ. (2002). Engineered gene circuits. Nature 420:224-30.

Hedglin M, Kumar R, and Benkovic SJ. (2013). Replication Clamps and Clamp Loaders. Cold Spring Hard Prospect Biol 5:a010165.

Heltzel JM, Maul RW, Scouten Ponticelli SK, and Sutton MD. (2009). A model for DNA polymerase switching involving a single cleft and the rim of the sliding clamp. Proc Natl Acad Sci U S A 106:12664-9.

Hirao A et al. (2000). DNA damage-induced activation of p53 by the checkpoint kinase Chk2. Science 287:1824-7.

Hoang ML, Leon RP, Pessoa-Brandao L, Hunt S, Raghuraman MK, Fangman WL, Brewer BJ, and Sclafani RA. (2007). Structural changes in Mcm5 protein bypass Cdc7-Dbf4 function and reduce replication origin efficiency in Saccharomyces cerevisiae. Mol Cell Biol 27:7594-602.

Hong WUL, Yang L. and DaoChun K. (2014). Mechanism of chromosomal DNA replication initiation and replication fork stabilization in eukaryotes. Sci China Life Sci 57:482-7.

Hu F, Wang Y, Liu D, Li Y, Qin J, and Elledge SJ. (2001). Regulation of the Bub2/Bfa1 GAP complex by Cdc5 and cell cycle checkpoints. Cell 107:655-65.

Hughes S, Elustondo F, Di Fonzo A, Leroux FG, Wong AC, Snijders AP, Matthews SJ, and Cherepanov P. (2012). Crystal structure of human CDC7 kinase in complex with its activator DBF4. Nat Struct Mol Biol 19:1101-7.

Ilves I, Petojevic T, Pesavento JJ, and Botchan MR. (2010). Activation of the MCM2-7 helicase by association with Cdc45 and GINS proteins. Mol Cell 37:247-58.

Jackson AL, Pahl PM, Harrison K, Rosamond J, and Sclafani RA. (1993). Cell cycle regulation of the yeast Cdc7 protein kinase by association with the Dbf4 protein. Mol Cell Biol 13:2899-908.

Jang YJ, Lin CY, Ma S, and Erikson RL. (2002). Functional studies on the role of the C-terminal domain of mammalian polo-like kinase. Proc Natl Acad Sci USA 99:1984-9.

Jang YJ, Ma S, Terada Y, and Erikson RL. (2002b). Phosphorylation of threonine 210 and the role of serine 137 in the regulation of mammalian polo-like kinase. J Biol Chem 277:44115-20.

Jaspersen SL and Winey M. (2004). The budding yeast spindle pole body: structure, duplication, and function. Annu Rev Cell Dev Biol 20:1-28.

Jergic S et al. (2013). A direct proofreader-clamp interaction stabilizes the Pol III replicase in the polymerization mode. EMBO J 32:1322-33.

Johnson, M. L. (1992). Why, When, and How Biochemists Should Use Least-Squares. Analytical Biochemistry 206:215-25.

Johnson RE, Klassen R, Prakash L, and Prakash S. (2015). A Major Role of DNA Polymerase d in Replication of Both the Leading and Lagging DNA Strands. Mol Cell 59:163-75.

Jones DR, Prasad AA, Chan PK, and Duncker BP. (2010). The Dbf4 motif C zinc finger promotes DNA replication and mediates resistance to genotoxic stress. Cell Cycle 9:2018-26.

Kabsch W. (2010). Xds. Acta Crystallogr D Biol Crystallogr 66:125-32.

Kadyrov FA, Dzantiev L, Constantin N, and Modrich P. (2006). Endonucleolytic function of MutLα in human mismatch repair. Cell 126:297-308.

Kadyrov FA et al. (2007). Saccharomyces cerevisiae MutLalpha is a mismatch repair endonuclease. J Biol Chem.

Kamimura Y, Tak YS, Sugino A, and Araki H. (2001). Sld3, which interacts with Cdc45 (Sld4), functions for chromosomal DNA replication in Saccharomyces cerevisiae. EMBO J 20:2097-107.

Kath JE et al. (2016). Exchange between Escherichia coli polymerases II and III on a processivity clamp. Nucleic Acids Res 44:1681-90.

Kihara M, Nakai W, Asano S, Suzuki A, Kitada K, Kawasaki Y, Johnston LH, and Sugino A. (2000). Characterization of the yeast Cdc7p/Dbf4p complex purified from insect cells. Its protein kinase activity is regulated by Rad53p. J Biol Chem 275:35051-62.

Kingston RL, Hamel DJ, Gay LS, Dahlquist FW, and Matthews BW. (2004). Structural basis for the attachment of a paramyxoviral polymerase to its template. Proc Natl Acad Sci USA 101:8301–8306.

Kling A et al. (2015). Antibiotics. Targeting DnaN for tuberculosis therapy using novel griselimycins. Science 348(6239):1106-12.

Koida N, Ozaki T, Yamamoto H, Ono S, Koda T, Ando K, Okoshi R, Kamijo T, Omura K, and Nakagawara A. (2008). Inhibitory role of Plk1 in the regulation of p73-dependent apoptosis through physical interaction and phosphorylation. J Biol Chem 283:8555-63.

Kops GJ, Weaver BA, and Cleveland DW. (2005). On the road to cancer: aneuploidy and the mitotic checkpoint. Nat Rev Cancer 5:773-85.

Kosinski J, Plotz G, Guarné A, Bujnicki JM, and Friedhoff P. (2008). The PMS2 subunit of human MutLalpha contains a metal ion binding domain of the iron-dependent repressor protein family. J Mol Biol 382:610-27.

Krissinel E and Henrick KJ. (2007). Inference of macromolecular assemblies from crystalline state. Mol Biol 372:774-97.

Kumar V, Chichili VP, Zhong L, Tang X, Velazquez-Campoy A, Sheu FS, Seetharaman J, Gerges NZ, and Sivaraman J. (2013). Structural basis for the interaction of unstructured neuron specific substrates neuromodulin and neurogranin with Calmodulin. Sci Rep 3:1392 DOI:10.1038.

Kunkel TA and Burgers PM. (2008). Dividing the workload at a eukaryotic fork. Trends Cell Biol 18:521-7.

Kunkel TA and Burgers PM. (2014). Delivering nonidentical twins. Nat Struct Mol Biol 21:649-51.

Labib K, and De Piccoli G (2011). Surviving chromosome replication: the many roles of the S-phase checkpoint pathway. Philos Trans R Soc Lond B Biol Sci 366:3554-61.

Lane HA and Nigg EA. (1996). Antibody microinjection reveals an essential role for human polo-like kinase 1 (Plk1) in the functional maturation of mitotic centrosomes. J Cell Biol 135: 1701-13.

Langston LD, Zhang D, Yurieva O, Georgescu RE, Finkelstein J, Yao NY, Indiani C, and O'Donnell ME. (2014). CMG helicase and DNA polymerase ε form a functional 15-subunit holoenzyme for eukaryotic leading-strand DNA replication. Proc Natl Acad Sci USA 111:15390-5.

Lara-Gonzalez P, Westhorpe FG, and Taylor SS. (2012). The spindle assembly checkpoint. Currently Biol 22:R966-80.

Larasati and Duncker BP. (2016). Mechanisms governing DDK regulation of the initiation of DNA replication. Genes (Basel) 8:pii: E3. doi: 10.3390.

Leblond CP and Walker BE. (1956). Renewal of cell populations. Physiol Rev 36:255-76.

LeBowitz JH and McMacken R. (1986). The Escherichia coli dnaB replication protein is a DNA helicase. J Biol Chem 261:4738-48.

Lee H et al. (2008). Diphosphothreonine-specific interaction between an SQ/TQ cluster and an FHA domain in the Rad53-Dun1 kinase cascade. Mol Cell 30:767-78.

Lee KS, Grenfell TZ, Yarm FR, and Erikson RL. (1998). Mutation of the polo-box disrupts localization and mitotic functions of the mammalian polo kinase Plk. Proc Natl
Acad Sci USA 95:9301-6.

Lee KS, Park JE, Asano S, and Park CJ. (2005). Yeast polo-like kinases: functionally conserved multitask mitotic regulators. Oncogene 24:217–29.

Lenart P, Petronczki M, Steegmaier M, Di Fiore B, Lipp JJ, and Hoffmann M. (2007). The small-molecule inhibitor BI 2536 reveals novel insights into mitotic roles of polo-like kinase 1. Curr Biol 17:304-15.

Lera RF and Burkard ME. (2012). High mitotic activity of Polo-like kinase 1 is required for chromosome segregation and genomic integrity in human epithelial cells. J Biol Chem 287:42812-25.

Liang F and Wang Y. (2007). DNA damage checkpoints inhibit mitotic exit by two different mechanisms. Mol Cell Biol 27:5067-78.

Li H and O'Donnell ME. (2018). The Eukaryotic CMG Helicase at the Replication Fork: Emerging Architecture Reveals an Unexpected Mechanism. Bioessays 40:doi10.1002.

Li J et al. (2002). Structural and functional versatility of the FHA domain in DNA-damage signaling by the tumor suppressor kinase Chk2. Mol Cell 9:1045-54.

Lindon C and Pines J. (2004). Ordered proteolysis in anaphase inactivates Plk1 to contribute to proper mitotic exit in human cells. J Cell Biol 164:233-41.

Lischetti T and Nilsson J. (2015). Regulation of mitotic progression by the spindle assembly checkpoint. Mol Cell Oncology 2:e970484.

Llamazares S, Moreira A, Tavares A, Girdham C, Spruce BA, and Gonzalez C. (1991). Polo encodes a protein kinase homolog required for mitosis in Drosophila. Genes Dev 5:2153-65.

Lopez-Mosqueda J and Maas NL. (2010). Damage-induced phosphorylation of Sld3 is important to block late origin firing. Nature 467:479-83.

Lopez de Saro FJ, Marinus MG, Modrich P, and O'Donnell M. (2006). The beta sliding clamp binds to multiple sites within MutL and MutS. J Biol Chem 281:14340-9.

Lopez de Saro FJ and O'Donnell M. (2001). Interaction of the beta sliding clamp with MutS, ligase, and DNA polymerase I. Proc Natl Acad Sci U S A 98:8376-80.

Lowery DM, Lim D, and Yaffe MB. (2005). Structure and function of Polo-like kinases. Oncogene 24:248-59.

Luo S, Xin X, Du LL, Ye K, and Wei Y. (2015). Dimerization Mediated by a Divergent Forkhead-associated Domain Is Essential for the DNA Damage and Spindle Functions of Fission Yeast Mdb1. J Biol Chem 290:21054-66.

Macurek L., Lindqvist A., Lim D., Lampson M.A., Klompmaker R., Freire R., Clouin C., Taylor S.S., Yaffe M.B., Medema R.H. (2008). Polo-like kinase-1 is activated by aurora A to promote checkpoint recovery. Nature 455:119-23.

Maekawa H, Priest C, Lechner J, Pereira G, and Schiebel E. (2007). The yeast centrosome translates the positional information of the anaphase spindle into a cell cycle signal. J Cell Biol 179:423-36.

Mantiero D, Mackenzie A, Donaldson A, and Zegerman P. (2011). Limiting replication initiation factors execute the temporal programme of origin firing in budding yeast. EMBO J 30:4805-14.

Marheineke K and Hyrien O. (2004). Control of replication origin density and firing time in Xenopus egg extracts: role of a caffeine-sensitive, ATR-dependent checkpoint. J Biol Chem 279:28071-81.

Masai, H and Arai K. (2000). Dbf4 motifs: conserved motifs in activation subunits for Cdc7 kinases essential for S-phase. Biochem Biophys Res Commun 275:228-32.

Matos J, Lipp JJ, Bogdanova A, Guillot S, Okaz E, Junqueira M, Shevchenko A, and Zachariae W. (2008). Dbf4-dependent CDC7 kinase links DNA replication to the segregation of homologous chromosomes in meiosis I. Cell 135:662-78.

Matsuoka S, Huang M, and Elledge SJ. (1998). Linkage of ATM to cell cycle regulation by the Chk2 protein kinase. Science 282:1893-7.

Matsushima N, Yoshida H, Kumaki Y, Kamiya M, Tanaka T, Izumi Y, and Kretsinger RH. (2008). Flexible structures and ligand interactions of tandem repeats consisting of proline, glycine, asparagine, serine, and/or threonine rich oligopeptides in proteins. Curr Protein Pept Sci 9:591-610.

Matthew EM, Yen TJ, Dicker DT, Dorsey JF, Yang W, and Navaraj A. (2007). Replication stress, defective S-phase checkpoint and increased death in Plk2-deficient human cancer cells. Cell Cycle 6:2571-8.

Matthews LA and Guarné A. (2013). Dbf4: the whole is greater than the sum of its parts. Cell Cycle 12:1180-8.

Matthews LA, Jones DR, Prasad AA, Duncker BP, and Guarné A. (2012). Saccharomyces cerevisiae Dbf4 has unique fold necessary for interaction with Rad53 kinase. JBC 287:2378-87.

Matthews LA, Selvaratnam R, Jones DR, Akimoto M, McConkey BJ, Melacini G, Duncker BP, Guarné A. (2014). A novel non-canonical forkhead-associated (FHA) domain-binding interface mediates the interaction between Rad53 and Dbf4 proteins. J Biol Chem 289:2589-99.

Mazouzi A, Velimezi G, and Loizou JI. (2014). DNA replication stress: causes, resolution and disease. Exp Cell Res 329:85-93.

Mazzarello P. (1999). A unifying concept: the history of cell theory. Nature Cell Biology 1:E13-5.

McInnes C, Estes K, Baxter M, Yang Z, Farag DB, and Johnston P. (2012). Targeting subcellular localization through the polo-box domain: non-ATP competitive inhibitors recapitulate a PLK1 phenotype. Mol Cancer Ther 11:1683-92.

McNally R, Bowman GD, Goedken ER, O'Donnell M, and Kuriyan J. (2010). Analysis of the role of PCNA-DNA contacts during clamp loading. BMC Struct Biol 10:doi10.1186.

Mechali M. (2010). Eukaryotic DNA replication origins: many choices for appropriate answers. Nat Rev Mol Cell Biol 11:728-38.

Messer W. (2002). The bacterial replication initiator DnaA. DnaA and oriC, the bacterial mode to initiate DNA replication. FEMS Microbiol Rev 26:355-74.

Miller CT, Gabrielse, C, Chen, Y C, and Weinreich, M. (2009). Cdc7p-Dbf4p regulates mitotic exit by inhibiting Polo kinase.. PLoS Genet 5:e1000498.

Mirkin EV and Mirkin SM. (2007). Replication fork stalling at natural impediments. Microbiol Mol Biol Rev 71:13-35.

Mishra R, Gorlov IP, Chao LY, Singh S, and Saunders GF. (2002). PAX6, paired domain influences sequence recognition by the homeodomain. J Biol Chem 277: 49488-94.

Moldovan GL, Pfander B, and Jentsch S. (2007). PCNA, the maestro of the replication fork. Cell 129:665-79.

Moore CW. (1978). Responses of radiation-sensitive mutants of Saccharomyces cerevisiae to lethal effects of bleomycin. Mutat Res 51:165-80.

Morgan DO. (2007). The Cell Cycle: Principles of Control (1st ed.). London: New Science Press. ISBN 978-0-87893-508-6.

Morgan DO. (1995). Principles of Cdk regulation. Nature 374:131-4.

Mundt KE, Golsteyn RM, Lane HA, and Nigg EA. (1997). On the regulation and function of human polo-like kinase 1 (PLK1): effects of overexpression on cell cycle progression. Biochem Biophys Res Commun 239: 377-85.

Muramatsu S, Hirai K, Tak YS, Kamimura Y, and Araki H. (2010). CDK- dependent complex formation between replication proteins Dpb11, Sld2, Pol epsilon, and GINS in budding yeast. Genes Dev 24:602-12.

Myer DL, Bahassi el M, and Stambrook PJ. (2005). The Plk3-Cdc25 circuit. Oncogene 24:299-305.

Nagi AD and Regan L. (1997). An inverse correlation between loop length and stability in a four-helix-bundle protein. Fold Des 2:67-75.

Narayanan A, Kumar S, Evrard AN, Paul LN, and Yernool DA. (2014). An asymmetric heterodomain interface stabilizes a response regulator-DNA complex. Nat Commun 5:doi: 10.1038.

Neef R, Gruneberg U, Kopajtich R, Li X, Nigg EA, and Sillje H. (2007). Choice of Plk1 docking partners during mitosis and cytokinesis is controlled by the activation state of Cdk1. Nat Cell Biol 9:436-44.

Noguchi Y, Yuan Z, Bai L, Schneider S, Zhao G, Stillman B, Speck C, and Li H. (2017). Cryo-EM structure of Mcm2-7 double hexameter on DNA suggest a lagging-strand DNA extrusion model. PNAS 114:E9529-38.

Nott TJ et al. (2009). An intramolecular switch regulates phosphoindependent FHA domain interactions in Mycobacterium tuberculosis. Sci Signal 2:ra12.

Ogi H, Wang CZ, Nakai W, Kawasaki Y, and Masumoto H. (2008). The role of the Saccharomyces cerevisiae Cdc7-Dbf4 complex in the replication checkpoint. Gene 414:32-40.

Ogino K, Takeda T, Matsui E, Iiyama H, Taniyama C, Arai K, and Masai H. (2001). Bipartite binding of a kinase activator activates Cdc7-related kinase essential for S phase. J Biol Chem 276:31376-87.

Otwinowski Z and Minor W. (1997). Processing of X-ray diffraction data collected in oscillation mode. Macromolecular Crystallography, Pt A 276:307-26.

Papamichos-Chronakis M, Krebs JE, Peterson CL (2006) Interplay between Ino80 and Swr1 chromatin remodeling enzymes regulates cell cycle checkpoint adaptation in response to DNA damage. Genes Dev 20:2437-49.

Park JE, Soung NK, Johmura Y, Kang YH, Liao C, and Lee KH. (2010). Polo-box domain: a versatile mediator of polo-like kinase function. Cell Mol Life Sci 67:1957-70.

Patoli AA, Winter JA, and Bunting KA. (2013). The UmuC subunit of the E. coli DNA polymerase V shows a unique interaction with the beta-clamp processivity factor. BMC Struct Biol 13:12.

Paulovich AG and Hartwell LH. (1995). A checkpoint regulates the rate of progression through S phase in S. cerevisiae in response to DNA damage. Cell 82:841-7.

Pellicioli A, Lucca C, Liberi G, Marini F, Lopes M, Plevani P, Romano A, Di Fiore PP, and Foiani M. (1999). Activation of Rad53 kinase in response to DNA damage and its effect in modulating phosphorylation of the lagging strand DNA polymerase. EMBO J 18:6561-72.

Pellicioli A and Foiani M. (2005). Signal transduction: how rad53 kinase is activated. Curr Biol 15:R769-71.

Pereira G, and Schiebel E. (2004). Cdc14 phosphatase resolves the rDNA segregation delay. Nat Cell Biol 6:473-5.

Pereira G, and Schiebel E. (2005). Kin4 kinase delays mitotic exit in response to spindle alignment defects. Mol Cell 19:209–21.

Pereira G, Hofken T, Grindlay J, Manson C, Schiebel E. (2000). The Bub2p spindle checkpoint links nuclear migration with mitotic exit. Mol Cell 6:1-10.

Pike BL, Tenis N, and Heierhorst J. (2004). Rad53 kinase activation-independent replication checkpoint function of the N-terminal forkhead-associated (FHA1) domain. J Biol Chem 279:39636-44.

Pillon MC, Babu VM, Randall JR, Cai J, Simmons LA, Sutton MD, and Guarné A. (2015). The sliding clamp tethers the endonuclease domain of MutL to DNA. Nucleic Acids Res 43:10746-59.

Pillon MC, Miller JH, and Guarne A. (2011). The endonuclease domain of MutL interacts with the beta sliding clamp. DNA Repair (Amst) 10:87-93.

Pillon MC et al. (2010). Structure of the endonuclease domain of MutL: unlicensed to cut. Mol Cell 39:145-51.

Pluciennik A et al. (2010). PCNA function in the activation and strand direction of MutLalpha endonuclease in mismatch repair. Proc Natl Acad Sci U S A 107:16066-71.

Princz LN, Wild P, Bittmann J, Aguado FJ, Blanco MG, Matos J, and Pfander B. (2017). Dbf4-dependent kinase and the Rtt107 scaffold promote Mus81-Mms4 resolvase activation during mitosis. EMBO J 36:665-78.

Putnam CD, Jaehnig JE, and Kolodner RD. (2009). Perspectives on the DNA damage and replication checkpoint responses in Saccharomyces cerevisiae. DNA Repair (Amst).

Qian WJ, Park JE, Lim D, Park SY, Lee KW, Yaffe MB. (2013). Peptide-based inhibitors of Plk1 polo-box domain containing mono-anionic phosphothreonine esters and their pivaloyloxymethyl prodrugs. Chem Biol 20:1255-64.

Qin J, Gronenborn AM. (2014). Weak protein complexes: challenging to study but essential for life. FEBS J 281:1948-9.

Raasch, K. et al. (2014). Interaction of 2-oxoglutarate dehydrogenase OdhA with its inhibitor OdhI in Corynebacterium glutamicum:Mutants and a model. J Biotechnol 191:99-105.

Rahal R and Amon A. (2008). The Polo-like kinase Cdc5 interacts with FEAR network components and Cdc14. Cell Cycle 7:3262-72.

Rancati G, Crispo V, Lucchini G, and Piatti S. (2005). Mad3/BubR1 phosphorylation during spindle checkpoint activation depends on both Polo and Aurora kinases in budding yeast. Cell Cycle 4:972-80.

Randell JC, Fan A, Chan C, Francis LI, Heller RC, Galani K, and Bell SP. (2010). Mec1 is one of multiple kinases that prime the Mcm2-7 helicase for phosphorylation by Cdc7. Mol Cell 40:353-63.

Ratsima H, Ladouceur AM, Pascariu M, Sauvé V, Salloum Z, Maddox PS, and D'Amours D. (2011). Independent modulation of the kinase and polo-box activities of Cdc5 protein unravels unique roles in the maintenance of genome stability. Proc Natl Acad Sci U S A 108:E914-23.

Reddy Chichili VP and Kumar V, Sivaraman J. (2013). Linkers in the structural biology of protein-protein interactions. Protein Sci 22:153-67.

Reindl W, Yuan J, Kramer A, Strebhardt K, and Berg T. (2008). Inhibition of polo-like kinase 1 by blocking polo-box domain-dependent protein-protein interactions. Chem Biol 15:459-66.

Remus D, Beuron F, Tolun G, Griffith JD, Morris EP, and Diffley JF. (2009). Concerted loading of Mcm2-7 double hexamers around DNA during DNA replication origin licensing. Cell 139:719-30.

Robinson CR and Sauer RT. (1998). Optimizing the stability of single-chain proteins by linker length and composition mutagenesis. Proc Natl Acad Sci USA 95:5929-34.

Robinson R. (2014) History of Biology: Cell Theory and Cell Structure. Advameg, Inc.

Rock JM and Amon A. (2011). Cdc15 integrates Tem1 GTPase-mediated spatial signals with Polo kinase-mediated temporal cues to activate mitotic exit. Genes Dev 25:1943-54.

Rossio V. (2010). The RSC chromatin-remodeling complex in uences mitotic exit and adaptation to the spindle assembly checkpoint by controlling the Cdc14 phosphatase. J Cell Biol 191:981-97.

Rouse J and Jackson SP. (2002). Interfaces between the detection, signaling, and repair of DNA damage. Science 297:547-51.

Rudolph D, Steegmaier M, Hoffmann M, Grauert M, Baum A, and Quant J. (2009). BI 6727, a Polo-like kinase inhibitor with improved pharmacokinetic profile and broad antitumor activity. Clin Cancer Res 15:3094-102.

Sato N, Sato M, Nakayama M, Saitoh R, Arai K, and Masai H. (2003). Cell cycle regulation of chromatin binding and nuclear localization of human Cdc7-ASK kinase complex. Genes Cells 8:451-63.

Schafmeister CE, Po J, and Verdine GL. (2000). An all-hydrocarbon cross-linking system for enhancing the helicity and metabolic stability of peptides. J Am Chem Soc 122:5891-2.

Schuck, P. (2000). Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and lamm equation modeling. Biophys J 78:1606-19.

Schwartz MF, Lee SJ, Duong JK, Eminaga S, and Stern DF. (2003). FHA domain-mediated DNA checkpoint regulation of Rad53. Cell Cycle 2:384-96.

Schwarz JK, Lovly CM, and Piwnica-Worms H. (2003). Regulation of the Chk2 protein kinase by oligomerization-mediated cis- and trans-phosphorylation. Mol Cancer Res 1:598-609.

Segurado M and Tercero JA. (2009). The S-phase checkpoint: targeting the replication fork. Biol Cell 101:617-27.

Seong YS, Kamijo K, Lee JS, Fernandez E, Kuriyama R, Miki T, and Lee KS. (2002). A spindle checkpoint arrest and a cytokinesis failure by the dominant-negative polo-box domain of Plk1 in U-2 OS cells. J Biol Chem 277:32282-93.

Serano D and D'Amours D. (2014). When genome integrity and cell cycle decisions collide: roles of polo kinases in cellular adaptation to DNA damage. Syst Synth Biol 8:195-203.

Sheu YJ and Stillman B. (2006). Cdc7-Dbf4 phosphorylates MCM proteins via a docking site-mediated mechanism to promote S phase progression. Mol Cell 24:101-13.

Shou W. (1999). Exit from mitosis is triggered by Tem1-dependent release of the protein phosphatase Cdc14 from nucleolar RENT complex. Cell 97:233-44.

Simmons LA, Davies BW, Grossman AD, and Walker GC. (2008). Beta clamp directs localization of mismatch repair in Bacillus subtilis. Mol Cell 29:291-301.

Simon AC, Zhou JC, Perera RL, van Deursen F, Evrin C, Ivanova ME, Kilkenny ML, Renault L, Kjaer S, Matak-Vinkovic D, Labib K, Costa A, and Pellegrini L. (2014). A Ctf4 trimer couples the CMG helicase to DNA polymerase α in the eukaryotic replisome. Nature 510:293-7.

Snead JL, Sullivan M, Lowery DM, Cohen MS, Zhang C, Randle DH, Taunton J, Yaffe MB, Morgan DO, and Shokat KM. A coupled chemical-genetic and bioinformatic approach to Polo-like kinase pathway exploration. Chem Biol 14:1261-72.

Speck C, Chen Z, Li H, and Stillman B. (2005). ATPase-dependent cooperative binding of ORC and Cdc6 to origin DNA. Nat Struct Mol Biol 12:965-71.

St-Pierre J. (2009) Polo kinase regulates mitotic chromosome condensation by hyperactivation of condensin DNA supercoiling activity. Mol Cell 34:416-26.

Stead BE, Brandl CJ, Sandre MK, and Davey MJ. (2012). Mcm2 phosphorylation and the response to replicative stress. BMC Genet 13:36.

Steegmaier M, Hoffmann M, Baum A, Lenart P, Petronczki M, and Krssak M. (2007). BI 2536, a potent and selective inhibitor of polo-like kinase 1, inhibits tumor growth in vivo. Curr Biol 17: 316-22.

Stegmeier F and Amon A. (2004). Closing mitosis: the functions of the Cdc14 phosphatase and its regulation. Annu Rev Genet 38:203-32.

Stillman B. (2015). "Reconsidering DNA Polymerases at the Replication Fork in Eukaryotes". Mol Cell 59:139-41.

Strebhardt K. (2010). Multifaceted polo-like kinases: drug targets and anti-targets for cancer therapy. Nat Rev Drug Discov 8:643-60.

Sun J, Shi Y, Georgescu RE, Yuan Z, Chait BT, Li H, and O'Donnell ME. (2015). The architechture of a eukaryotic replisome. Nat Struct Mol Biol 22:976-92.

Sunkel CE and Glover DM. (1988). Polo, a mitotic mutant of Drosophila displaying abnormal spindle poles. J Cell Sci 89:25-38.

Sun Z, Hsiao J, Fay DS, and Stern DF. (1998). Rad53 FHA domain associated with phosphorylated Rad9 in the DNA damage checkpoint. Science 281:272-44.

Sutterlin C, Lin CY, Feng Y, Ferris DK, Erikson RL, and Malhotra V. (2001). Polo-like kinase is required for the fragmentation of pericentriolar Golgi stacks during mitosis. Proc Natl Acad Sci USA 98:9128-32.

Sutton MD. (2004). The Escherichia coli dnaN159 mutant displays altered DNA polymerase usage and chronic SOS induction. J Bacteriol 186:6738-48.

Swift LH and Golsteyn RM. (2014). Genotoxic anti-cancer agents and their relationship to DNA damage, mitosis, and checkpoint adaptation in proliferating cancer cells. Int J Mol Sci 15:3403-31.

Tanaka S, Umemori T, Hirai K, Muramatsu S, Kamimura Y, and Araki H. (2007). CDK-dependent phosphorylation of Sld2 and Sld3 initiates DNA replication in budding yeast. Nature 445:328-32.

Toste Rego A, Holding AN, Kent H, and Lamers MH. (2013). Architecture of the Pol III-clamp-exonuclease complex reveals key roles of the exonuclease subunit in processive DNA synthesis and repair. EMBO J 32:1334-43.

Tria G, Mertens HDT, Kachala M, and Svergun DI. (2015). Advanced ensemble modelling of flexible macromolecules using X-ray solution scattering. IUCrJ 2:207-17.

Valerio-Santiago M, de Los Santos-Velazquez AI, Monje-Casas F. (2013). Inhibition of the mitotic exit network in response to damaged telomeres. PLoS Genet 9:e1003859.

Varrin AE, Prasad AA, Scholz RP, Ramer MD, and Duncker BP. (2005). A mutation in Dbf4 motif M impairs interactions with DNA replication factors and confers increased resistance to genotoxic agents. Mol Cell Biol 25:7494-504.

Vidanes GM, Sweeney FD, Galicia S, Cheung S, Doyle JP, Durocher D, Toczyski DP. (2010). CDC5 inhibits the hyperphosphorylation of the checkpoint kinase Rad53, leading to checkpoint adaptation. PLoS Biol 8:e1000286.

Vijayraghavan S and Schwacha A. (2012). The Eukaryotic Mcm2-7 Replicative Helicase. In The Eukaryotic Replisome: a Guide to Protein Structure and Function, Dordrecht: Springer 2:113-34.

Waga S, and Stillman B. (1998). The DNA replication fork in eukaryotic cells. Annu Rev Biochem 67:721-51.

Wang W, Prosise WW, Chen J, Taremi SS, Le HV, Madison V, Cui X, Thomas A, Cheng KC, and Lesburg CA. (2008). Construction and characterization of a fully active PXR/SRC-1 tethered protein with increased stability. Protein Eng Des Sel 21:425-33.

Watanabe N, Arai H, Iwasaki J, Shiina M, Ogata K, and Hunter T. (2005). Cyclin-dependent kinase (CDK) phosphorylation destabilizes somatic Wee1 via multiple pathways. Proc Natl Acad Sci USA 102:11663-8.

Weinert T. (1998). DNA damage checkpoints update: getting molecular. Curr Opin Genet Dev 8:185-193.

Weng JH et al. (2015). Uncovering the Mechanism of Forkhead-Associated Domain-Mediated TIFA Oligomerization That Plays a Central Role in Immune Responses. Biochemistry 54:6219-29.

Williams RS, Green R, and Glover JN. (2001). Crystal structure of the BRCT repeat region from the breast cancer-associated protein BRCA1. Nat Struct Biol 8:838-42.

Williams SJ, Sohn KH, Wan L, Bernoux M, Sarris PF, Segonzac C, Ve T, Ma Y, Saucet SB, Ericsson DJ, Casey LW, Lonhienne T, Winzor DJ, Zhang X, Coerdt A, Parker JE, Dodds PN, Kobe B, and Jones JD. (2014). Structural basis for assembly and function of a heterodimeric plant immune receptor. Science 344:299-303.

Wilson KA, Bar S, Maerz AL, Alizon M, and Poumbourios P. (2005). The conserved glycine-rich segment linking the N-terminal fusion peptide to the coiled coil of human T- cell leukemia virus type 1 transmembrane glycoprotein gp21 is a determinant of membrane fusion function. J Virol 79:4533-9.

Wogan GN, Hecht SS, Felton JS, Conney AH, and Loeb LA. (2004). Environmental and chemical carcinogenesis. Semin Cancer Biol 14:473-486.

Wolff P et al. (2014). Differential modes of peptide binding onto replicative sliding clamps from various bacterial origins. J Med Chem 57:7565-76.

Wybenga-Groot LE et al. (2014). Structural basis of Rad53 kinase activation by dimerization and activation segment exchange. Cell Signal 26:1825-36.

Xiang Y, Takeo S, Florens L, Hughes SE, Huo LJ, and Gilliland WD. (2007). The inhibition of Polo Kinase by matrimony maintains G2 arrest in the meiotic cell cycle. PLoS Biol 5:e323.

Xu HE, Rould MA, Xu W, Epstein JA, Maas RL, and Pabo CO. (1999). Crystal structure of the human Pax6 paired domain-DNA complex reveals specific roles for the linker region and carboxy-terminal subdomain in DNA binding. Genes Dev 13:1263–1275.

Xu J, Shen C, Wang T, and Quan J. (2013). Structural basis for the inhibition of Polo-like kinase 1. Nat Struct Mol Biol 20:1047-53.

Ye Q, Li X, Wong A, Wei Q, and Jia Z. (2006). Structure of calmodulin bound to a calcineurin peptide: a new way of making an old binding mode. Biochemistry 45:738-45.

Yin Z, Kelso MJ, Beck JL, and Oakley AJ. (2013). Structural and thermodynamic dissection of linear motif recognition by the E. coli sliding clamp. J Med Chem 56:8665-73.

Yin Z, Song Y, and Rehse PH. (2013). Thymoquinone blocks pSer/pThr recognition by Plk1 Polo-box domain as a phosphate mimic. ACS Chem Biol 8:303-8.

Yuan J, Eckerdt F, Bereiter-Hahn J, Kurunci-Csacsko E, Kaufmann M, and Strebhardt K. (2002). Cooperative phosphorylation including the activity of polo-like kinase 1 regulates the subcellular localization of cyclin B1. Oncogene 21:8282-92.

Yuan J, Sanhaji M, Kramer A, Reindl W, Hofmann M, and Kreis NN. (2011). Polo-box domain inhibitor poloxin activates the spindle assembly checkpoint and inhibits tumor growth in vivo. Am J Pathol 179:2091-9.

Zegerman P and Diffley JF. (2007). Phosphorylation of Sld2 and Sld3 by cyclin-dependent kinases promotes DNA replication in budding yeast. Nature 445:281-5.

Zegerman P and Diffley JF. (2009). DNA replication as a target of the DNA damage checkpoint. DNA Repair (Amst) 8:1077-88.

Zegerman P and Diffley JF. (2010). Checkpoint-dependent inhibition of DNA replication initiation by Sld3 and Dbf4 phosphorylation. Nature 467:474-8.

Zhang T, Nirantar S, Lim HH, Sinha I, Surana U (2009) DNA damage checkpoint maintains CDH1 in an active state to inhibit anaphase progression. Dev Cell 17:541-51.

Zhang W and Durocher D. (2008). Dun1 counts on rad53 to be turned on. Mol Cell 31:1-2.

Zhao, H., Brown, P. H. & Schuck, P. On the distribution of protein refractive index increments. Biophys J 100, 2309–2317 (2011).

Zhao H, Brautigam CA, Ghirlando R, and Schuck P. (2013). Overview of current methods in sedimentation velocity and sedimentation equilibrium analytical ultracentrifugation. Curr Protoc Protein Sci Chapter 20, Unit20 12.

Zhao X and R Rothstein. (2002). The Dun1 checkpoint kinase phosphorylates and regulates the ribonucleotide reductase inhibitor Sml1. Proc Natl Acad Sci U S A 99:3746-51.

Zou L and SJ Elledge (2003). Sensing DNA damage through ATRIP recognition of RPA- ssDNA complexes. Science 300:1542-8.