THE DBF4-RAD53 INTERACTION

CHARACTERIZATION OF THE INTERACTION BETWEEN DBF4 AND RAD53
DURING REPLICATION STRESS IN BUDDING YEAST

By LINDSAY MATTHEWS, B. Sc. (Hons.)

A Thesis Submitted to the School of Graduate Studies in Partial Fulfillment

of the Requirements for the Degree Doctor of Philosophy

McMaster University

DOCTOR OF PHILOSOPHY (2012)   McMaster University

(Biochemistry and Biomedical Sciences)   Hamilton, Ontario

TITLE: Characterization of the Interaction Between Dbf4 and

                          Rad53 During Replication Stress in Budding Yeast

AUTHOR: Lindsay Matthews, B. Sc. (Hons.) (McMaster University)

SUPERVISOR: Dr. Alba Guarné

NUMBER OF PAGES: ii, 127

Abstract

All living things must replicate their DNA. Despite being essential for life, this process is also inherently dangerous. Replication stress, which induces replication fork stalling, is an unavoidable risk that can trigger potentially harmful changes to the genome. Eukaryotes have a replication checkpoint pathway that stabilizes stalled replication forks to prevent damage. One of the critical protein interactions in this pathway, between Dbf4 and Rad53, pauses the cell cycle in budding yeast. This is important to give the cell time to recover from stress. The molecular details of this interaction were investigated to shed light on how this association is regulated by the cell. The structure of an N-terminal domain from Dbf4 was solved through X-ray crystallography and discovered to have a modified _BR_CA-1 _C-t_erminal (BRCT) fold, which included an additional N-terminal helix. This domain could interact with the _F_ork_H_ead _A_ssociated 1 (FHA1) domain from Rad53 _in vitro_, and the additional helix was necessary for complex formation. Although the FHA1 domain has a well-characterized binding site for phospho-epitopes, a combination of chemical cross-linking and NMR spectroscopy experiments demonstrated that the N-terminal domain from Dbf4 is contacting an alternative surface. However, the full-length Dbf4 protein _in vivo_ may be contacting both this distal site and the phospho-epitope binding pocket. This bipartite interaction between Dbf4 and Rad53 would lend specificity to the complex and also suggests a kinase may be regulating the association. As FHA and BRCT domains are prevalent in eukaryotic nuclear proteins, these findings are instructive for how these domains mediate interactions in other signaling pathways.

Acknowledgements

I would like to thank my supervisor, Dr. Alba Guarné, for her guidance, leadership and support. This project took us in unexpected directions and I truly enjoyed the journey with you. It has been a privilege to train in your lab and I appreciate everything that you invested in me.

I would also like to thank my committee members, Dr. Gerry Wright and Dr. Giuseppe Melacini, for all of their valuable advice. I'd like to extend a special thank you to Dr. Melacini for bravely allowing a crystallographer to train in his lab in NMR spectroscopy. In addition, thank you to Dr. Bernard Duncker for introducing me to Dbf4 and for helpful discussions on the project.

A PhD can be a difficult road, but I've had the good fortune of having excellent company. So I would like to thank all of the past and present members of the Guarné lab for making the challenging times easier and the good times sweeter.

Thank you also to my family for putting up with me running back and forth to the lab. To my mom, Marnie Matthews, thanks for instilling in me a fierce love of reading that has been a source of great strength. To my dad, Michael Matthews, I remember so well that every morning before school when I was young you would tell me "work hard, play hard, have fun". To me, science has always been equal measures of hard work and pure fun. Thank you for this lesson and so much more.

Table of Contents

List of Figures

List of Tables

List of Abbreviations and Symbols

AAA+	ATPases Associated with diverse cellular Activities
APBS	Adaptive Poisson-Boltzmann Solver
APC	Anaphase Promoting Complex
ARS	Autonomously Replicating Sequence
ASK	Activator of S-phase Kinase
ASKL1	Activator of S-phase Kinase Like protein - 1
ATP	Adenosine triphosphate
BNL	Brookhaven National Laboratory
BRCA-1	Breast Cancer - 1
BRCT	BRCA-1 C-terminal
BS^3	Bis[sulfosuccinimidyl]suberate
CDC	Cell Division Cycle
CDK	Cyclin-Dependent Kinase
Cdt1	Cdc10-Dependent Transcript 1
CEN	Centromere
Chk1/2	Checkpoint Kinase 1/2
Chs5	Chitin Synthesis 5
CMG	Cdc45−Mcm2-7−GINS
CMS	Cdc45−Mcm2-7−Sld3
COOT	Crystallographic Object-Oriented Toolkit
CYMAL-7	7-Cyclohexyl-1-Heptyl-β-D-Maltoside
Dbf4	Dumbbell Former 4
DDK	Dbf4-Dependent Kinase
DNA	Deoxyribonucleic Acid
Drf1	Dbf4-Related Factor
dsDNA	Double Stranded Deoxyribonucleic Acid
DTT	Dithiothreitol
EDTA	Ethylenediaminetetraacetic Acid
FHA	Forkhead Associated
G1 phase	Gap 1 phase
G2 phase	Gap 2 phase
GINS	Go-Ichi-Ni-San (Japanese for 5-1-2-3)
H-BRCT	Helix-BRCA-1 C-terminal
HP	High Performance
HPV	Human Papillomavirus
HSQC	Heteronuclear Single Quantum Coherence
HU	Hydroxyurea
IPTG	Isopropyl β-D-1-thiogalactopyranoside
LDAO	Lauryldimethylamine-N-Oxide
M phase	Mitotic phase
MCM	Minichromosome Maintenance

Mec1	Mitosis Entry Checkpoint protein 1
MMS	Methyl Methanesulfonate
Motif C	Motif C-terminus
Motif M	Motif Middle
Motif N	Motif N-terminus
MPD	2-Methyl-1,3-Propanediol
Nbs1	Nijmegan Breakage Syndrome protein 1
NMR	Nuclear Magnetic Resonance
NSLS	National Synchrotron Light Source
OD_{600}	Optical Density (600 nm)
ORC	Origin Recognition Complex
PCR	Polymerase Chain Reaction
PDB	Protein Data Bank
PEE	Pentaerythritol Ethoxylate
PEG	Polyethylene Glycol
PHENIX	Python-based Hierarchical Environment for Integrated Xtallography
Pol ε	Polymerase ε
Pre-RC	Pre-Replication Complex
PSIPRED	Protein Structure Prediction
PXRR	Protein Xtallography Research Resource
Rad53	Radiation 53
RCI	Random Coil Index
RMSD	Root Mean Square Deviation
RPA	Replication Protein A
S phase	Synthetic phase
S.D.	Standard Deviation
SC	Synthetic Complete media
SCD	Serine-Glutamine/Threonine-Glutamine Cluster Domain
SDS	Sodium Dodecyl Sulfate
SeMet	Selenomethionine
SiRNA	Small Interfering RNA
Sld2/3	Synthetically Lethal with Dpb11-1 2/3
Spo4/6	Sporulation protein 4/6
ssDNA	Single Stranded Deoxyribonucleic Acid
SUMO	Small Ubiquitin-like Modifier
TEV	Tobacco Etch Virus
TLS	Translesion Synthesis
tRNA	Transfer Ribonucleic Acid
V_M	Matthews coefficient

Declaration of Academic Achievement

All experiments detailed in this thesis were designed, conducted and analyzed by Dr. Guarné and myself unless stated otherwise in the author's preface of each chapter. This includes budding yeast experiments that were performed by Ajai Prasad, Darryl Jones and Dr. Bernard Duncker (Chapters 2 and 3) as well as constructs that were made by Ajai Prasad and Andrew Duong (Chapter 2). In addition, I conducted the NMR spectroscopy experiments in collaboration with Rajeevan Selvaratnam and Dr. Giuseppe Melacini (Chapter 4).

Chapter 1

Introduction

1.1 Eukaryotic Cell Division

Rudolf Virchow famously wrote "*omnis cellula e cellula*" (every cell from a pre-existing cell) (Virchow, 1858). The ability of life to reproduce itself may be a central tenet in modern biology, but the notion that cells can divide was highly controversial when it was first proposed. Until late in the 19th century, life was thought to arise spontaneously from inanimate matter. Even after Schleiden and Schwann identified cells as the basic building blocks of organisms, the mechanism behind reproduction remained elusive (Schleiden, 1838; Schwann, 1839). Schleiden described the birth of new cells as a spontaneous organization of molecules – a type of crystallization. However, evidence began to accumulate that countered the spontaneous generation of life, which inspired Virchow to write his famous line. Each cell has the instructions necessary to build a complete replica. The mechanism behind this feat was first witnessed by one of Virchow's students, Walther Flemming (Paweletz, 2001). He observed thread-like structures in salamander cells that stained readily (Flemming, 1882). These structures lined up together and then pulled apart to opposite sides of the cell prior to division. Flemming described these threads as being made of chromatin, to reflect their affinity for stain, and named the process mitosis (from *mitos* - the Greek word for thread) in their honor (Flemming, 1882).

These thread-like structures are the chromosomes, which consist of deoxyribonucleic acid (DNA) as well as proteins. Their orderly partitioning into daughter cells during mitosis ensures that each cell inherits a complete copy of the genome. DNA

had first been identified ten years prior to Flemming's experiments. Friedrich Miescher was attempting to characterize the protein content of leucocytes isolated from surgical wrappings. After obtaining the nuclei, he found this organelle contained a strange substance that precipitated in acid and was rich in phosphorous (Miescher, 1871). Puzzling over this discovery, Meischer named his substance "nuclein" and mused that it might be related to heredity (Dahm, 2005; Miescher, 1874). However, DNA was determined to consist of only four different building blocks, the nucleotides adenine (A), guanine (G), cytosine (C), and thymine (T). That a molecule with so little chemical diversity could be responsible for the variation evident in life seemed unlikely. Consequently, DNA was relegated to playing a structural role in the chromosomes, while the associated protein was thought to contain the genetic information. The arguments in favor of this view were so persuasive that even after DNA was shown to be the hereditary molecule in pneumococci (Avery et al., 1944), much of the scientific world remained skeptical (McCarty, 2003). One who was convinced was Erwin Chargaff, who went on to formulate the so-called "Chargaff rule": when quantifying the amount of each base in DNA, A always equals T and G always equals C (Chargaff, 1951). Together with the knowledge that DNA forms a double helix from X-ray diffraction patterns, Watson and Crick formulated the structure of this macromolecule (Watson and Crick, 1953).

Confronted with such an elegant model, the arguments against DNA as the genetic material dissolved. It was clear from the onset that such a double helix would be perfectly suited for replication. Pulling apart the two strands by breaking the hydrogen bonds between complementary bases would reveal single stranded templates that could direct the synthesis of their partner. This semi-conservative method of replication was confirmed by Meselson and Stahl, who showed that newly replicated DNA consists of one parental strand and one new daughter strand (Meselson and Stahl, 1958). DNA had finally taken its rightful place after more than eighty years since its initial discovery. However, investigations into how cells divide are still ongoing. It is clear that a eukaryotic cell spends most of its time outside of mitosis, which is referred to as interphase (Figure 1.1). While this period would appear quite static under Flemming's

microscope since the DNA has not yet condensed, an amazing array of activity is taking place. The cell is busy replicating its macromolecules in preparation for the impending cell division, including its genome. The DNA only condenses during prophase, which is the first stage of mitosis (Figure 1.1). The duplicated chromosomes are aligned in the middle of the cell by the mitotic spindle apparatus during metaphase, and then pulled apart at anaphase. Finally, telophase and cytokinesis occur and the cell is physically divided into two.
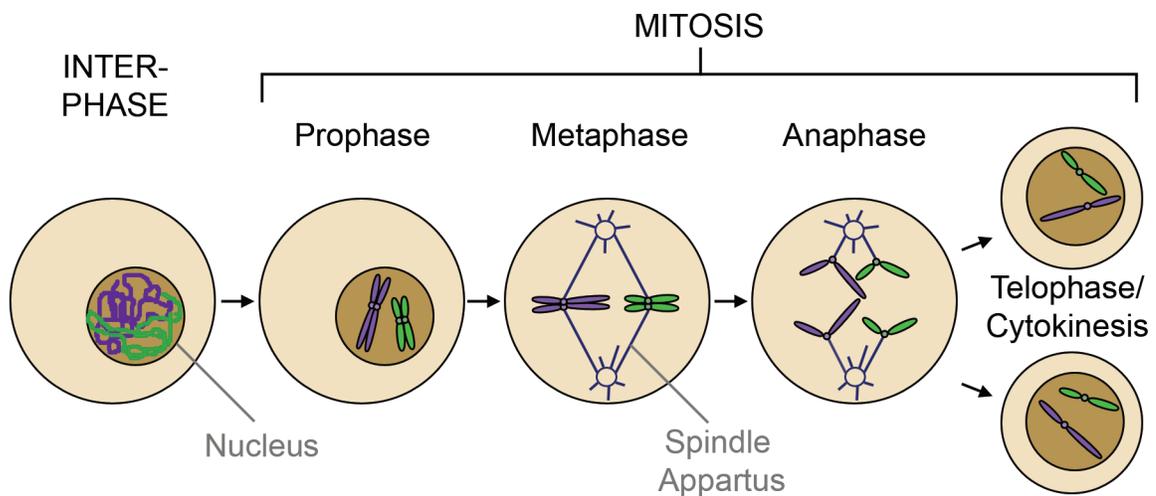
Figure 1.1 Division of a eukaryotic cell. A schematic of a cell that transitions from interphase through mitosis (prophase, metaphase, anaphase, telophase, and cytokinesis). The two chromosomes are colored purple and green. The nucleus and spindle apparatus are labeled in grey. Note that the nuclear membrane breaks down during metaphase and then reforms during telophase. Figure adapted from (Sullivan and Morgan, 2007).

Interphase can be divided into three distinct periods of time (Figure 1.2). The first is a _G_ap phase (G1), during which the cell duplicates all of its molecules except for DNA. The genome is instead exclusively replicated during the subsequent _S_ynthetic (S) phase. This is followed by a second _G_ap phase (G2), which immediately precedes mitosis. New cells begin their cycle in G1 phase and must make a decision to progress further. For a eukaryotic cell, environmental cues such as nutrient levels or the presence of growth

factors contribute to this decision (Sherr, 1994). Past this point, the cell is committed to completing the cell cycle without any further influence from the environment. Instead, kinases whose activity oscillates in tune with the cell cycle phase transitions complete the tasks required for cell division (Morgan, 1997). This machinery is controlled by checkpoint pathways that act as feedback loops to ensure that the cell cycle pauses if necessary steps are incomplete (Weinert, 1998). The majority of these checkpoint pathways are sensitive to the integrity of the DNA, and it is possible to pause the cell cycle before DNA synthesis begins (G1 checkpoint), while DNA synthesis is ongoing (S phase checkpoint) or prior to separating the duplicated chromosomes into daughter cells (G2 phase checkpoint) if the DNA is damaged (Figure 1.2) (Weinert, 1998). Mitosis also has a checkpoint pathway that ensures the mitotic spindle apparatus has attached properly to all of the chromosomes. This makes certain that daughter cells inherit a full complement of the genome. Despite all of these precautions, however, cell division remains a risky enterprise for any organism.

1.2 Replication Forks and Genetic Instability

Duplex DNA is unwound at replication forks, which are associated with two major enzymatic activities: the helicase activity required to separate the parental strands and the polymerase activity that then uses these strands as a template. Many agents can interfere with this process, including free radicals generated by metabolism, UV radiation and environmental toxins (Wogan et al., 2004). The damage these mutagens inflict can trigger the DNA damage checkpoint during S phase to pause the cell cycle so the genome can be repaired before it is used as a template (Labib and De Piccoli, 2011). However, S phase can also be paused by the replication checkpoint. This pathway is triggered due to an endogenous danger: stalled replication forks (Figure 1.2) (Labib and De Piccoli, 2011). Stalled forks occur when the polymerase activity is blocked while the helicase activity continues unabated (Labib and De Piccoli, 2011). As the helicase begins spooling out cytotoxic single-stranded DNA (ssDNA), a number of detrimental events can be triggered. Stalled forks are prone to collapse by dissociation of the replication machinery,

they can lose genetic information through the action of nucleases and may even induce chromosomal rearrangements (Labib and De Piccoli, 2011). Thus, although replication forks appear misleadingly innocuous, once stalled they represent an important source of genetic instability.
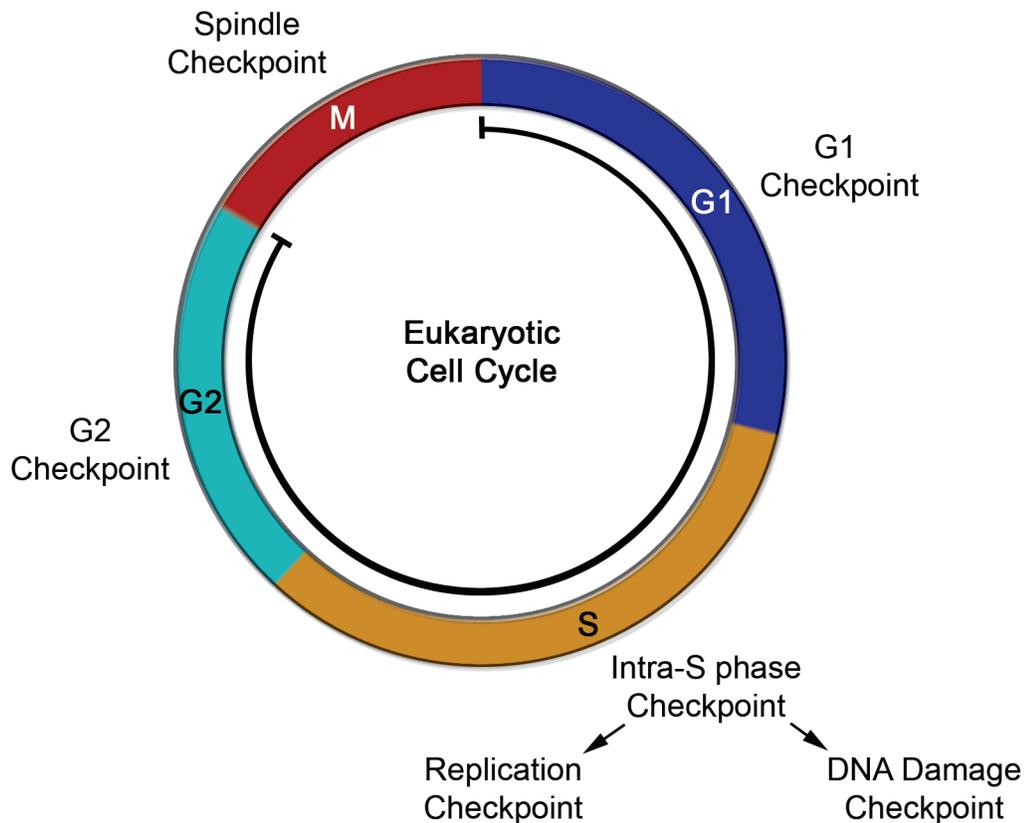
Figure 1.2 Cell cycle checkpoints. The four phases of the eukaryotic cell cycle are shown (Gap 1 (G1), Synthetic (S), Gap 2 (G2), and Mitosis (M)) with their associated checkpoints. The partial circle in the middle of the cycle indicates the stages that constitute interphase (G1, S and G2).

Replication forks are first generated at specific sites within the genome called origins of replication, where the duplex DNA is initially melted creating a ssDNA bubble. The association of a helicase at each fork framing this bubble, followed by recruitment of

the remaining replisome machinery required for elongation, allows the two forks to travel in opposite directions (Waga and Stillman, 1998). The total number of replication forks present in a cell can vary widely. Bacteria such as *E. coli* use only a single origin of replication, which fires to generate two replication forks that move bidirectionally along its single, circular chromosome until reaching the termination sites at the opposite end. Eukaryotes require multiple origins, with budding yeast using approximately 400 (despite having a genome only about 2.5 times larger than *E. coli*) and humans employing tens of thousands (Mechali, 2010). This allows for a high density of replication forks, each of which only has to replicate a relatively short distance before colliding with an oncoming fork originating from a neighboring origin, or reaching the end of the chromosome itself. However, not every origin fires simultaneously as the cell transitions into S phase. A subset of origins will fire early in S phase while others will only be activated late (Mechali, 2010). A third group of origins remain dormant and will not fire under normal circumstances, but rather are triggered in order to rescue replication forks that have stalled in their vicinity (Figure 1.3) (Mechali, 2010).

Eukaryotes control the length of S phase by regulating origin firing. By preventing late origins from being activated, the replication checkpoint is able to pause S phase. While the use of late origins offers an advantage in slowing S phase during times of stress, it is also a major reason why eukaryotic cells spend so much time replicating their genome compared to *E. coli*. In fact, in some circumstances, eukaryotes can alter their replication strategy and fire their origins in quick succession. This is the case for *Xenopus* embryonic cells, which can duplicate their genomes in less than 20 minutes, rivaling *E. coli* for expediency (Marheineke and Hyrien, 2004). However, staggering early and late origin firing also balances how many replication forks are active at any given time. As replication forks deplete the limited resources available in a cell, too many origins firing will lead to stalling and genetic instability (Blow et al., 2011). However, on the other end of the spectrum, too few origins will increase the distance that any individual fork has to travel before colliding with another (Blow et al., 2011). This also increases the chances a fork will encounter an obstacle and stall. Thus, the division

between early and late firing origins not only provides a mechanism for cells to pause S phase by inhibiting late origins, but also maintains an optimal number of active replication forks.
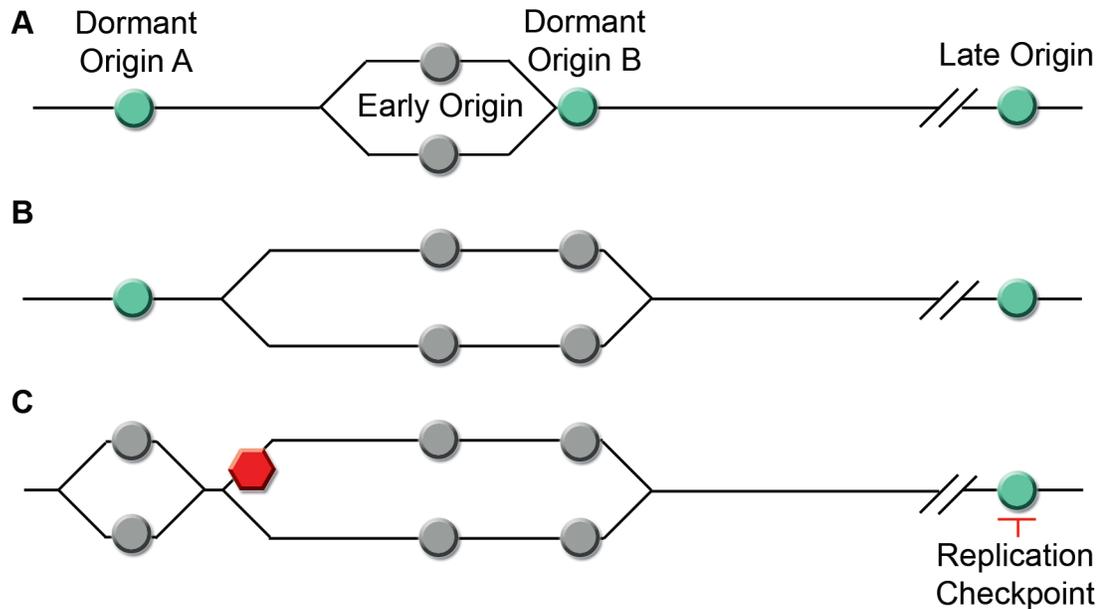
Figure 1.3 Comparison of early, dormant and late origins. Origins capable of initiation are represented as green circles, while newly replicated origin DNA is shown as grey circles. **(A)** Early origins fire creating bidirectional replication forks. **(B)** Dormant origins are passively replicated by neighboring replication forks (Dormant Origin B). **(C)** If forks stall due to encountering an obstacle (red hexagon), then a nearby dormant origin can fire to complete replication of the intervening area (Dormant Origin A). Late origins, in contrast, are suppressed to keep the cell in S phase.

Early, late and dormant origins are functionally defined, but have otherwise remained enigmatic. While the origin in the *E. coli* chromosome (*oriC*) has a precise sequence, eukaryotic origins lack this specificity with the sole exception of budding yeast. This organism has modular origins of replication called ARSs (*A*utonomously *R*eplicating *S*equences) in which one module, the A element, contains a conserved sequence (Mechali, 2010). In other eukaryotes, origins are likely specified by epigenetics instead of particular sequences (Mechali, 2010). It is clear that early origins tend to be located in

topologically open euchromatin, while late origins are found in heterochromatin (Mechali, 2010). However, there are few other clues as to what determines an origin in a eukaryote.

1.3 Replication Stress and Disease

The potential danger posed by replication forks is related to the frequency with which they stall. The obstacles that trigger stalling are pervasive and in many cases inevitable. DNA damage that cannot be accommodated by the polymerase active site will block a replication fork, but even a perfect template is inherently dangerous. For example, regions of the chromosome tightly bound to proteins can stall replication forks (Mirkin and Mirkin, 2007). Additionally, ssDNA sequences exposed by the helicase that are prone to forming secondary structures can block the associated polymerase (Mirkin and Mirkin, 2007). Replication stress, which is stress that induces stalled replication forks, is consequently an unavoidable threat to the integrity of DNA.

Replication forks tend to stall in specific regions of the genome termed common fragile sites. The name comes from their propensity to suffer gaps or breaks during mitosis due to incomplete replication in the preceding S phase (Mirkin and Mirkin, 2007). They are referred to as "common" since these sites are a normal feature of the human genome and are highly conserved amongst metazoa (Durkin and Glover, 2007). Under ordinary circumstances the replication checkpoint prevents them from becoming damaged (Mirkin and Mirkin, 2007). However, if cells are challenged with agents that block the polymerase, gaps and breaks are generated at a high frequency within fragile sites (Casper et al., 2002). The mechanism behind this sensitivity to replication stress is unclear, though these regions are often associated with nucleotide repeats that tend to form secondary structures that can inherently block the polymerase (Durkin and Glover, 2007). Thus, even though drugs that inhibit polymerase activity would cause fork stalling in all areas of the genome, common fragile sites are rendered particularly susceptible to forming gaps and breaks.

Replication stress is implicated in the beginning stages of tumorogenesis as an important source of genetic instability (Arlt et al., 2009). In this case, it is likely that the stress originates from oncogene over-expression that drives a high replication rate in cancer cells (Bartek et al., 2007). The limited cellular resources devoted to serving active replication forks become exhausted and forks begin to stall resulting in potentially pathogenic changes in the genome (Bartek et al., 2007). As expected, these changes are often DNA gaps at common fragile sites (Durkin and Glover, 2007). Replication stress can also trigger cell death. Although this may be viewed as an adaptive response by eliminating mutant cells from the population, it can also play a role in pathogenesis. Alzheimer's disease, which leads to dementia, is associated with the deposition of proteinaceous plaques in the brain as well as neural cell death (Kosik, 1992). The latter occurs when neurons in the diseased brain inappropriately enter the cell cycle and encounter replication stress (Yurov et al., 2011). However, although this stress contributes to the progression of diseases in rare instances, it is important to note that replication stress is common to all organisms whenever DNA is being replicated. It is the price to pay for the elegant mechanism of separating the parental strands of DNA to serve as templates first described by Watson and Crick almost sixty years ago (Watson and Crick, 1953). In fact, the inevitable erosion of the genome due to replication stress has been hypothesized to be the underlying cause of aging (Burhans and Weinberger, 2007).

1.4 DDK Activity Oscillates During the Cell Cycle

The connection between DNA replication and disease inspired Leland H. Hartwell to dissect the cell cycle in budding yeast during the early 1970s by studying temperature sensitive mutants (Hartwell, 1974). *S. cerevisiae* offers a number of advantages when studying cell division, one of which is that the cell morphology changes with the phases of the cell cycle allowing one to track the stage of a cell using only a microscope. With this strategy, the events of the cell cycle were mapped using mutant genes that caused cell cycle arrest (Hartwell, 1974). These genes were named with the acronym '*CDC*' for *C*ell

*D*ivision *C*ycle, and they form the corner stone of eukaryotic cell cycle biology. This led to Dr. Hartwell sharing the 2001 Nobel Prize in physiology or medicine.

One of the genes identified in these seminal studies, *CDC7*, resulted in dumbbell shaped cells when mutants were synchronized in G1 phase and then shifted to the restrictive temperature (Hartwell, 1973). In *S. cerevisiae*, small buds appear on the mother cell in late G1, which grow in size until mitosis, when they are separated and the cell wall repaired with chitin, resulting in a noticeable scar (Hartwell, 1974). Interestingly, the nuclear membrane does not break down during this process, but rather the whole nucleus is stretched through the neck of the bud and cleaved in half as well (Hartwell, 1974). The Cdc7 mutant phenotype showed that these cells were not capable of replicating their DNA, and arrested with buds that continued to grow in size until creating the characteristic dumbbell shape (Hartwell, 1973). However, if the cells were allowed to enter S phase prior to shifting to the restrictive temperature, then they could transition through the cell cycle normally until entering the next cycle, at which point they again arrested as dumbbells (Hartwell, 1973). This suggested that Cdc7 was specifically involved in DNA replication initiation at the G1/S phase transition.

Soon after the discovery of *CDC7*, another group extended these studies with more temperature sensitive *S. cerevisiae* mutants. One of these, *DBF4* (for *D*umb*B*ell *F*ormer 4), resulted in the same dumbbell phenotype as the *CDC7* mutant at the restrictive temperature (Johnston and Thomas, 1982). Furthermore, temperature sensitive *CDC7* mutant cells could be rescued by over-expressing Dbf4, suggesting that these proteins form a complex (Kitada et al., 1992). This was later confirmed using a yeast two-hybrid assay (Jackson et al., 1993). *CDC7* encodes a serine/threonine kinase that is only active when bound to Dbf4 (Jackson et al., 1993) and is required to fire both early and late origins of replication (Bousset and Diffley, 1998; Donaldson et al., 1998). While Cdc7 protein levels are constant in *S. cerevisiae*, the levels of Dbf4 oscillate, peaking at the G1/S transition and disappearing at the end of mitosis (Chapman and Johnston, 1989). This is due to the presence of destruction boxes in its N-terminus, which target it for

degradation by the <u>A</u>naphase <u>P</u>romoting <u>C</u>omplex (APC) (Ferreira et al., 2000; Nougarede et al., 2000; Weinreich and Stillman, 1999). The oscillation of Dbf4 levels ensures that despite stable levels of the Cdc7 kinase, the activity of this enzyme similarly cycles. However, the purpose of this is presently unclear. Mutation of the destruction boxes to maintain Cdc7 activity in G1 phase did not result in any noticeable defects in budding yeast (Ferreira et al., 2000).

The oscillation of Dbf4 levels to control the Cdc7 kinase is reminiscent of another kinase complex: the <u>C</u>yclin-<u>D</u>ependent <u>K</u>inase (CDK). The catalytic subunit (Cdc28 in budding yeast) has constant levels, but its activity changes based on the association of regulatory proteins (the cyclins), which have oscillating levels during the cell cycle. This similarity has resulted in the Cdc7/Dbf4 complex being referred to as DDK for <u>D</u>bf4-<u>D</u>ependent <u>K</u>inase, however there is no homology between these different complexes. CDK drives the progression of the cell cycle (Arellano and Moreno, 1997). In fact, the decision point in G1 phase where a cell chooses to divide is due to the expression of cyclin proteins triggered by the presence of environmental cues, such as mitogens or nutrients (Arellano and Moreno, 1997). After the first CDK complex is formed, it induces a program of cascading cyclin expression patterns that carry a cell through division without any further influence by the environment. The G1 phase CDK complex specifically activates the expression of the cyclin in charge of the G1/S phase transition, which triggers the expression of the S phase cyclin, until finally the mitotic cyclin-CDK complex is activated (Arellano and Moreno, 1997). The APC ultimately degrades the cyclins (Koepp et al., 1999) along with Dbf4 such that daughter cells enter G1 phase with low Cdc7 and Cdc28 kinase activity until they decide to cycle again.

In contrast to cyclins, budding yeast only have a single Dbf4 homolog that is expressed during the cell cycle. However, other eukaryotes have two homologs. In fission yeast, the second Dbf4 is called Spo4 and it regulates a kinase homologous to Cdc7 called Spo6 (Nakamura et al., 2002). This alternative complex is not utilized in the mitotic cell cycle but instead is required for meiosis (Nakamura et al., 2002). Metazoan

cells, in contrast, have two Dbf4 homologs that compete for a single Cdc7 kinase (Montagnoli et al., 2002). The second Dbf4 in this case (called alternatively _D_bf4 _R_elated _F_actor 1 [Drf1] or _A_ctivator of _S_ Phase _K_inase _L_ike 1 [ASKL1]) is predominately expressed in embryonic cells, whereas Dbf4 is conversely the major form in adult cells (Takahashi and Walter, 2005). Why this switch occurs during development is not clear, however it is notable that these two DDK complexes are not functionally equivalent. While both can activate origins of replication, in cancer cells (which express both complexes) only siRNA knockdown of the Dbf4 regulatory subunit is lethal (Yoshizawa-Sugata et al., 2005).

1.5 DDK Activates the Replicative Helicase at Origins

The paradigm for replication initiation is the binding of an initiator protein to origin DNA, which was originally proposed in the early 1960s (Jacob et al., 1963) . In _E. coli_, the initiator is the AAA+ ATPase DnaA. This enzyme has a low intrinsic activity and does not use ATP hydrolysis to power DNA melting (Messer, 2002). Instead, binding to ATP induces a conformational change that allows for DnaA to form a filament on _oriC_ through recognition of specific sites called DnaA boxes (Messer, 2002). This bends the DNA forcing an AT-rich region to melt (Messer, 2002). DnaA is then able to load the helicase, DnaB, onto the ssDNA through a direct interaction (Messer, 2002) (Figure 1.4). DnaB is a hexameric ring of identical subunits, which couples the hydrolysis of ATP to the movement of the helicase along the lagging strand of the DNA template (LeBowitz and McMacken, 1986). Consequently, in order to control initiation, the binding of DnaA to _oriC_ is tightly regulated. This occurs through sequestering _oriC_ from DnaA as well as by stimulating the ATPase activity of DnaA to prevent the conformational change required for initiation (Messer, 2002).

In eukaryotes, the equivalent initiator protein to DnaA is a six subunit complex called the _O_rigin _R_ecognition _C_omplex (ORC), which is bound to origins during the entire cell cycle. The association is not regulated because the ORC itself is not capable of loading the replicative helicase or melting origin DNA. Instead, a seventh subunit called

Cdc6 is required. This is only present in the G1 phase of the cell cycle, and associates with ORC to form a ring like structure (Figure 1.4) (Speck et al., 2005). This activates the ATPase activity of the ORC (Speck et al., 2005). Unlike DnaA, ATP hydrolysis by the ORC is necessary for the loading of the heterohexameric helicase (MCM2-7) (Bowers et al., 2004). This does not occur through a direct interaction, but rather requires a mediator protein called Cdt1 (Figure 1.4) (Chen et al., 2007). Cdt1 is tightly regulated such that helicase loading can only occur in G1 phase. This prevents origins that have fired in S phase from re-associating with a helicase and firing a second time, creating a harmful amplification event (Feng and Kipreos, 2003).

Two MCM2-7 helicases are initially loaded as a double hexamer, which forms due to a head-to-head packing (Evrin et al., 2009; Remus et al., 2009). As shown in Figure 1.4, the N-terminal domains are on the inside mediating double hexamer formation, while the C-terminal domains are on the outside and contain the ATPase activity. At this stage, the assembly found at the origin is called the pre-*R*eplication *C*omplex (pre-RC) and the origin is licensed. However, even after the association of the helicase with the origin DNA, replication still does not precede. The helicase is inactive within the pre-RC and there is no melting of origin DNA, which is still double-stranded when threaded through the central channel of the helicase (Evrin et al., 2009). The activation of MCM2-7 depends on two kinase complexes activated at the G1/S phase transition: S-CDK (CDK bound to the S phase specific cyclin) and DDK (Labib, 2011). Consequently, helicase activation is temporally regulated to only occur in S phase (Figure 1.4) (Labib, 2011).

Two neighboring subunits in the MCM2-7 hexameric ring (Mcm2 and Mcm5) are only weakly associated (Figure 1.5). The resulting gap in the ring between these subunits is referred to as the MCM gate and is important for helicase function (Vijayraghavan and Schwacha, 2012). This is by providing an entry point for DNA to enter the central channel of the helicase. Two additional factors are required for helicase activity: Cdc45
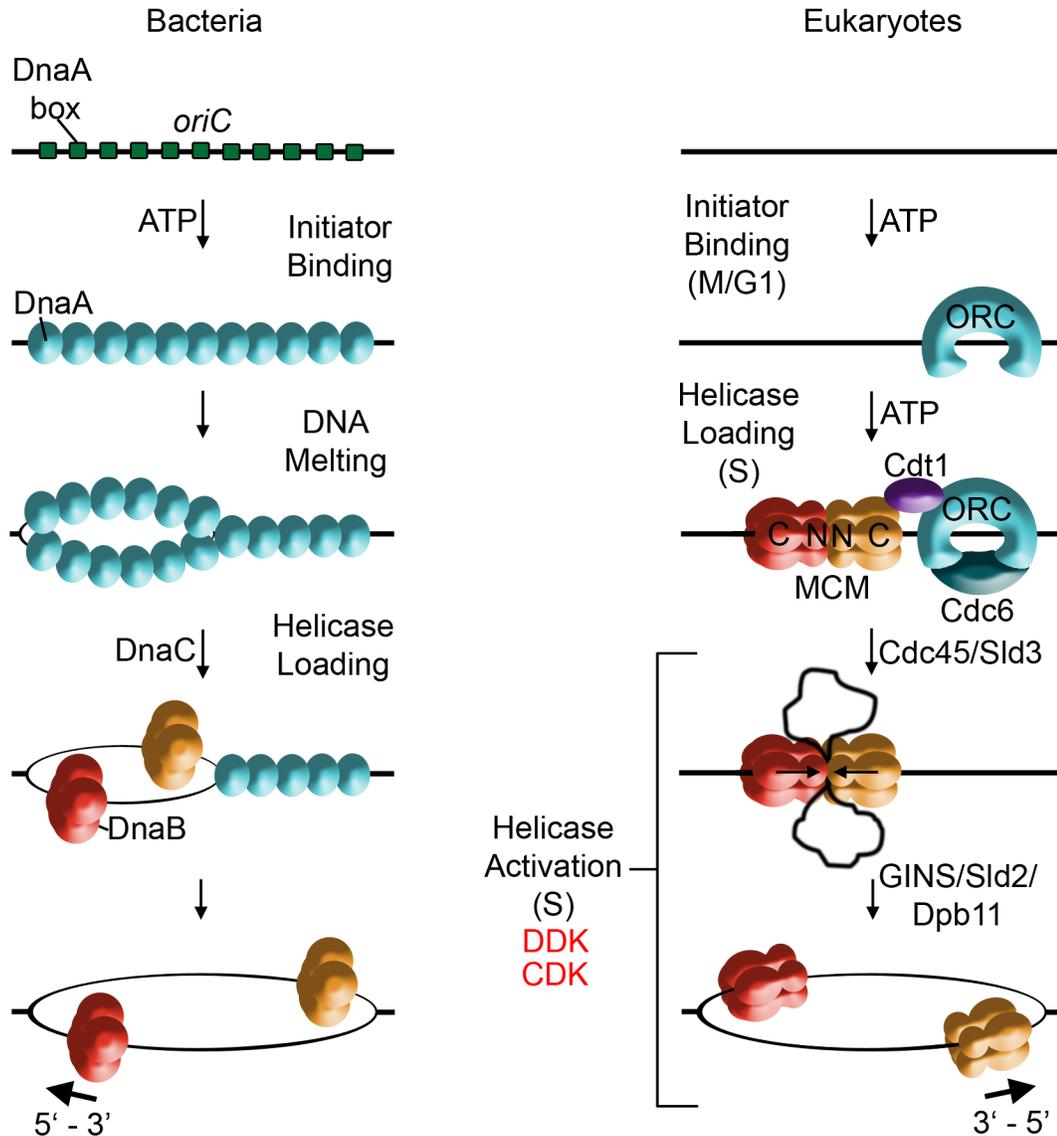
Figure 1.4 Comparison of replication initiation in bacteria and eukaryotes. Cartoon of the steps leading to replication initiation. The initiator proteins (DnaA in bacteria and ORC in eukaryotes) are colored in blue, while the helicases are colored in orange and red. The phase of the cell cycle is indicated for each step in the eukaryotic pathway. Required factors not represented in the cartoon are listed at the arrow of their relevant step. For the eukaryotic helicase, the N- and C-terminal domains are labeled at the helicase loading step, while the directionality of the helicase is indicated in the last step for both pathways. Adapted from (Labib, 2010; Messer, 2002; Vijayraghavan and Schwacha, 2012).

and the tetrameric GINS complex (Vijayraghavan and Schwacha, 2012). Both of these factors bind in close proximity to the MCM gate, with Cdc45 contacting both Mcm2 and Mcm5 while GINS contacts Mcm5 and its other neighbor, Mcm3 (Figure 1.5) (Costa et al., 2011). The association of Cdc45 and GINS stimulates the ATPase activity of the MCM2-7 complex and strengthens its interaction with DNA (Ilves et al., 2010).
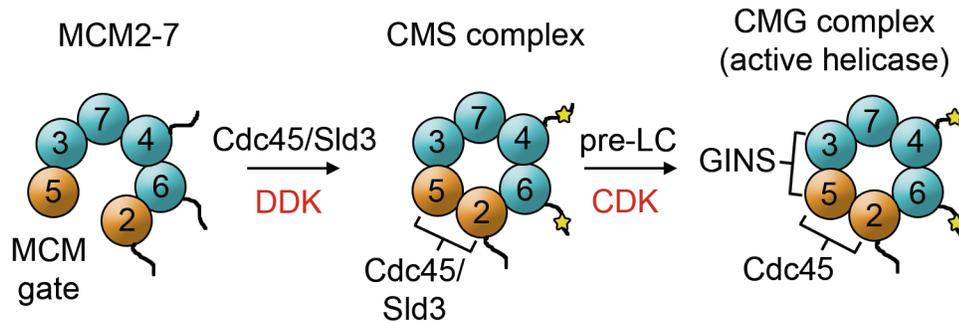
Figure 1.5 Schematic of the MCM2-7 helicase. Cartoon of the MCM2-7 helicase viewed down the helical axis with the subunits numbered. The progression from the MCM2-7 complex to the Cdc45−MCM2-7−Sld3 (CMS) complex and finally the Cdc45−MCM2-7−GINS (CMG) complex is illustrated. The Mcm subunits that make the MCM gate (Mcm2 and Mcm5) are colored in orange. Yellow stars represent phosphorylation of the N-terminal tails of Mcm4 and 6 by DDK. The proteins that associate with MCM2-7 at each step are indicated above the arrows, while the kinase primarily responsible for triggering the interaction is labeled in red below. The interactions by GINS and Cdc45 bridge two Mcm subunits together, as indicated by the bracket. Note that the pre-loading complex (pre-LC) consists of Dpb11, Sld2, GINS, and Pol ε.

Cdc45 is recruited first and relies on DDK activity, which phosphorylates the unstructured N-terminal tails of Mcm4 and Mcm6 (Figure 1.5) (Randell et al., 2010; Sheu and Stillman, 2006, 2010). This accomplishes at least two things. First, it alleviates a negative influence that the Mcm4 N-terminal tail specifically has on MCM2-7 helicase activity (Sheu and Stillman, 2010). Second, it induces a conformational change in the N-terminal domains of Mcm4 and Mcm6 that is transduced through the MCM2-7 hexamer until it reaches Mcm5, which can then bind Cdc45 (Hoang et al., 2007). Importantly, a point mutation in the N-terminal domain of Mcm5 (P83L), which likely mimics this

conformational change, allows for the helicase to be activated in the absence of DDK activity (Fletcher and Chen, 2006; Hardy et al., 1997). Additionally, the deletion of the N-terminal tail from Mcm4 also bypasses the need for DDK (Sheu and Stillman, 2010).

Aside from inducing conformational changes, serine/threonine kinases can also promote protein interactions by generating phospho-epitopes that are recognized by specialized domains. CDK works to recruit the GINS complex using this alternative strategy by phosphorylating the replication proteins Sld2 and Sld3 (Muramatsu et al., 2010; Tanaka et al., 2007b; Zegerman and Diffley, 2007). While Sld3 associates with Cdc45 at origins (Kamimura et al., 2001), Sld2 binds to a soluble complex containing not only GINS but also the polymerase Pol ε (Tanaka et al., 2007a) (Figure 1.5). The protein Dpb11 then acts as an adaptor that connects phosphorylated Sld2 and Sld3 together (Muramatsu et al., 2010; Tanaka et al., 2007b; Zegerman and Diffley, 2007). Dpb11 consists of _BR_CA-1 _C-T_erminal (BRCT) domains, which have a four stranded, parallel β-sheet that is surrounded by three α helices. In certain instances, conserved positively charged residues create a phospho-epitope recognition pocket (Figure 1.6). However, the third residue C-terminal to the phosphorylated site is also required for a robust interaction, and this residue occupies a second pocket created at the interface between the BRCT domain and a second module encoded in tandem (Williams et al., 2004) (Figure 1.6). Dpb11 has tandem BRCT domains at both its N- and C-terminus, with one pair binding to phosphorylated Sld2 and the other pair recognizing Sld3 (Tanaka et al., 2007b; Zegerman and Diffley, 2007).

While subjecting mixtures of purified proteins to size exclusion chromatography, it was discovered that the association of GINS and Sld3 with the Cdc45−MCM2-7 complex was mutually exclusive (Bruck and Kaplan, 2011). This suggests that after DDK stimulates the association of Cdc45 and consequently also Sld3, CDK activity is subsequently required to swap Sld3 with GINS to create a fully functional helicase (Figure 1.5). This also requires the origin DNA to melt, the double MCM2-7 hexamer to separate and the encapsulation of ssDNA in the hexameric cores to form two bidirectional
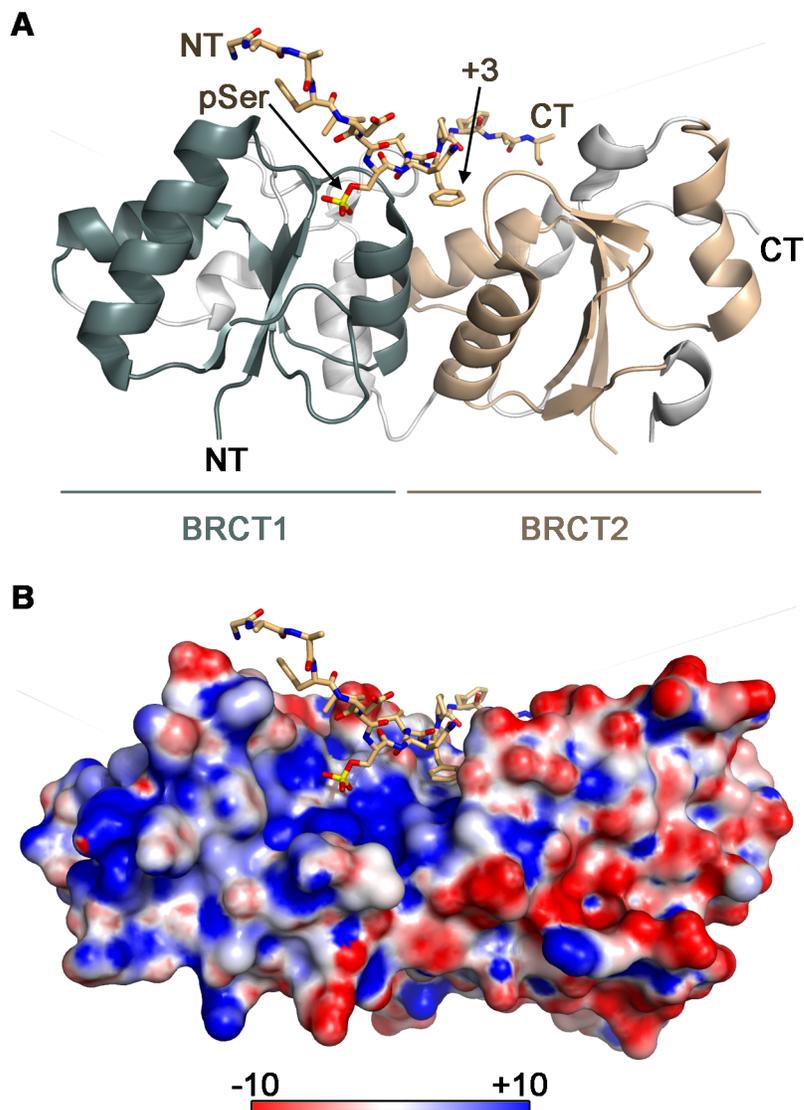
Figure 1.6 Interaction of a phospho-peptide with a tandem BRCT domain. **(A)** Ribbon diagram of the tandem BRCT domain from BRCA-1 (PDB ID: 1T2V) with an associated phospho-peptide shown as sticks. The phospho-serine residue (pSer) and a phenylalanine residue three amino acids downstream (+3) are labeled. The N- and C-termini are indicated for both the tandem BRCT and peptide. **(B)** A surface representation of the structure from panel A 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 charge of an electron. This figure was generated using APBS software (Baker et al., 2001) and visualized in PyMOL.

replication forks. It has been suggested that in the Cdc45−MCM2-7−Sld3 (CMS) complex, each helicase in the double hexameric assembly pumps the dsDNA threaded through its central chamber towards the other (Vijayraghavan and Schwacha, 2012). This strain forces the DNA to melt and emerge from the middle of the double hexamer as two ssDNA loops (Figure 1.4) (Vijayraghavan and Schwacha, 2012). Once Sld3 is replaced by GINS forming the Cdc45−MCM2-7−GINS (CMG) complex, a rearrangement allows for the two hexamers to dissociate from each other and encircle the ssDNA of the leading strand (Vijayraghavan and Schwacha, 2012). This hypothesis therefore associates DDK activity specifically with origin melting. However, the situation is complicated by the fact that Cdc7 is an acidophilic kinase, which requires acidic groups in its target sites in order to efficiently phosphorylate substrates. These groups can be neighboring acidic amino acids, or alternatively they can be serine or threonine residues that have already been phosphorylated by other kinases. These priming kinases not only generate target sites for DDK in the N-terminal tails of Mcm4 and Mcm6, but also can help create binding sites to recruit DDK more effectively (Randell et al., 2010). Thus, additional kinases control the efficiency with which DDK performs its essential function in S phase.

1.6 DDK is Inhibited by the Replication Checkpoint

In order to pause the cell cycle, the replication checkpoint must prevent the pre-RCs loaded at late origins from having their associated MCM2-7 helicase molecules activated. This requires a disruption of two pathways: the DDK pathway that controls Cdc45 association and the CDK pathway that controls the recruitment of the GINS complex. The replication checkpoint pathway begins with the activation of the Mec1 kinase, which occurs due to large stretches of exposed ssDNA associated with stalled replication forks (Friedel et al., 2009). A complex called the 9-1-1 clamp is loaded at the ssDNA/dsDNA junction at the stalled fork and Mec1 is recruited through a subunit that recognizes the ssDNA binding protein RPA (Friedel et al., 2009). The clamp activates Mec1 in budding yeast, while also recruiting the BRCT domain containing protein Dpb11 through a phospho-epitope (Navadgi-Patil and Burgers, 2009). In other eukaryotes, it is

Dpb11, and not the clamp itself, that activates Mec1 (Navadgi-Patil and Burgers, 2009). Once active, Mec1 recruits and phosphorylates the next kinase in the checkpoint pathway, Rad53 (Pellicioli and Foiani, 2005). Rad53 is activated through hyperphosphorylation and is released from the stalled fork in order to execute stress response functions distal to this site. Interestingly, DDK is one of the kinases charged with hyperphosphorylating Rad53 indicating that it plays a role in initiating the replication stress checkpoint (Ogi et al., 2008).

Aside from aiding in checkpoint activation, DDK remains active at the site of the stalled fork as its activity is essential for cell survival (Ogi et al., 2008). This suggests DDK plays a role in recovery of the stalled fork. Mcm2, which also has an N-terminal tail like Mcm4 and 6, is a major target of DDK during replication stress (Stead et al., 2012). This may prevent the helicase from dissociating, which is irreversible since helicase loading requires factors (Cdc6 and Cdt1) only present in G1 phase (Figure 1.4) (Feng and Kipreos, 2003). DDK can also assist in resolving DNA damage that has blocked the replication fork in fission yeast. This is by phosphorylating the 9-1-1 clamp and causing it to dissociate from the stalled fork, which allows room for repair enzymes to access the site (Furuya et al., 2010). However, if the cell cannot remove the obstacle then the stalled fork can be resolved by instead waiting for an impinging replication fork originating from a neighboring origin (Mirkin and Mirkin, 2007). Alternatively, the obstacle can be bypassed by either translesion synthesis (TLS) or a template switching mechanism (Branzei and Foiani, 2007). TLS involves swapping the replicative polymerase for an error prone polymerase instead, and thus contributes to mutagenesis. It has been noted that Cdc7 plays a role in induced mutagenesis in yeast (Hollingsworth et al., 1992; Kilbey, 1986; Njagi and Kilbey, 1982), and this has recently been linked to an active role in TLS (Pessoa-Brandao and Sclafani, 2004; Vaziri and Masai, 2010). Therefore, DDK has numerous functions at the site of a stalled replication fork.

Rad53 also helps stabilize the replisome at stalled forks, but additionally brings about a global response to stress by increasing the synthesis of DNA repair enzymes and

nucleotides as well as inhibiting late origins to pause the cell cycle (Branzei and Foiani, 2009). This latter function is not essential for viability, however it may increase the genetic stability of the cell. To inhibit the assembly of an active CMG helicase at late origins, Rad53 blocks both DDK and CDK. DDK is directly inhibited through phosphorylating Dbf4 (Duch et al., 2011; Kihara et al., 2000; Lopez-Mosqueda et al., 2010; Zegerman and Diffley, 2010). In contrast, CDK is not modified but rather Rad53 targets Sld3, which then prevents the necessary switch with the GINS complex by blocking the interaction between Sld3 and Dpb11 (Figure 1.7) (Lopez-Mosqueda et al., 2010; Zegerman and Diffley, 2010). This difference is likely because CDK activity prevents any origins from firing more than once in a single S phase, and therefore directly inhibiting it would lead to harmful gene amplification events (Labib, 2010). Additionally, Dbf4 and Sld3 are limiting replication initiation factors in addition to Sld2 and Dpb11 (Mantiero et al., 2011). The relatively low levels of these proteins is what ensures only a certain number of origins can fire at any given time and prevents too many replication forks from being active at once, making them strategic targets for inhibiting late origins (Mantiero et al., 2011).

Rad53 has a central kinase domain flanked by two *F*ork*h*ead *A*ssociated (FHA) domains. Each FHA domain is also preceded by a Serine Glutamine/Threonine Glutamine (*S*Q/TQ) *C*luster *D*omain (SCD) (Figure 1.8A) (Pellicioli and Foiani, 2005). Upstream kinases control Rad53 activity by phosphorylating serine and threonine residues within its SCDs (Lee et al., 2003). This creates phospho-epitopes that can trigger protein interactions with substrates. In addition, the FHA domains are essential for Rad53 function as they conversely recognize phospho-epitopes in a variety of binding partners (Schwartz et al., 2003). This not only increases the repertoire of substrates for Rad53, but also helps the kinase localize to specific sites in the cell. For example, the FHA1 domain targets Rad53 to stalled replication forks where it is activated through hyperphosphorylation (Chen and Zhou, 2009).
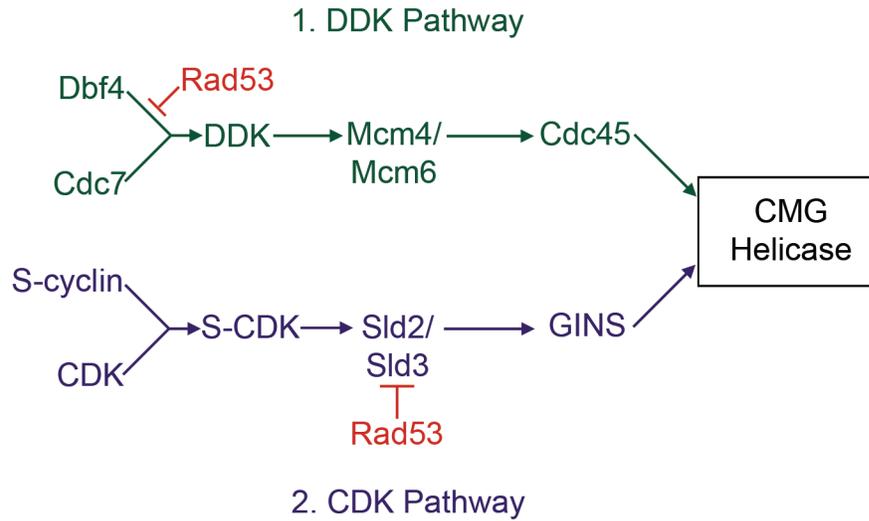
Figure 1.7 Rad53 inhibits CMG helicase assembly to block late origin firing. The DDK (green) and CDK (blue) regulated pathways leading to CMG helicase assembly are outlined. In each case, the first step is the interaction between the regulatory and catalytic subunits of the respective kinases. The next step is the phosphorylation of their targets (Mcm4 and Mcm6 for DDK; Sld2 and Sld3 for CDK). The last step is the recruitment of Cdc45 and GINS for helicase activation. The steps inhibited by Rad53 for each pathway are indicated. For the cyclin and CDK complex, "S" indicates the cyclin is specific to the S phase of the cell cycle.

The homolog of Rad53 in higher eukaryotes is referred to as Chk2 and only has a single FHA domain at its N-terminus, as shown in Figure 1.8B. A second effector kinase, Chk1, also mediates cellular responses to DNA damage. While Rad53/Chk2 and Chk1 have some overlapping functions, they primarily protect against different types of damage. In yeast, Rad53 plays the primary role in replication stress while in higher eukaryotes Chk1 is the main effector kinase that responds to stalled replication forks (Chen and Sanchez, 2004). This is perplexing considering that Chk1 lacks FHA domains (Figure 1.8C), which are critical for Rad53 function. How human Chk1 can play an analogous role to yeast Rad53 without the use of FHA domains is still being investigated.

The power of an FHA domain is its ability to mediate protein interactions. Structurally, this domain consists of a β-sandwich that is capable of binding to phospho-
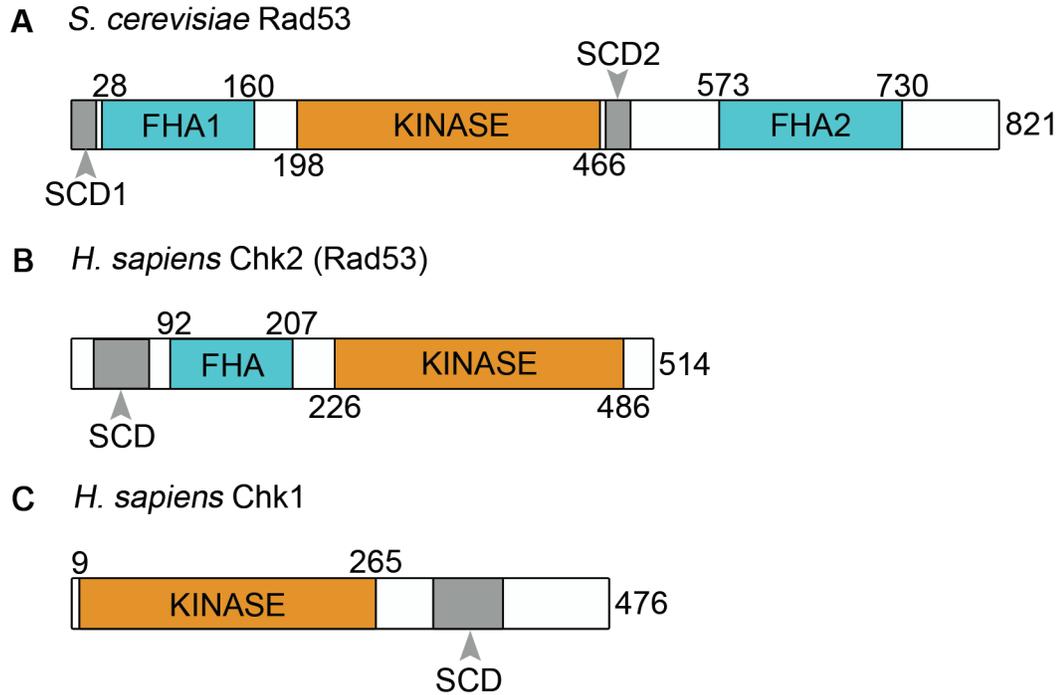
Figure 1.8 Comparison of Rad53, Chk1 and Chk2. A bar diagram illustrating the organization of the domains in checkpoint kinases followed by a number indicated the total amino acid length. The amino acid boundaries of the FHA and kinase domains are labeled. The kinases are **(A)** Rad53 from *S. cerevisiae*, **(B)** the Rad53 homolog from humans (Chk2) and **(C)** Chk1 from humans.

epitopes using the loops connecting the β-strands at one end of the molecule (Figure 1.9A) (Durocher and Jackson, 2002). Like BRCT domains, this involves recognizing the phosphorylated amino acid as well as the third residue C-terminal to this site using two distinct pockets. However, unlike BRCT domains, both binding sites are contained in a single domain module, and FHA domains are never encoded in tandem. Additionally, FHA domains are unique from other phospho-epitope binding units because they are the only known domains to distinguish between phospho-threonine and phospho-serine residues (Durocher et al., 2000). This is due to a small hydrophobic nook within the phospho-threonine binding site in the FHA domain that accommodates the methyl group and stabilizes the correct rotamer for optimal binding (Figure 1.9B) (Pennell et al., 2010).
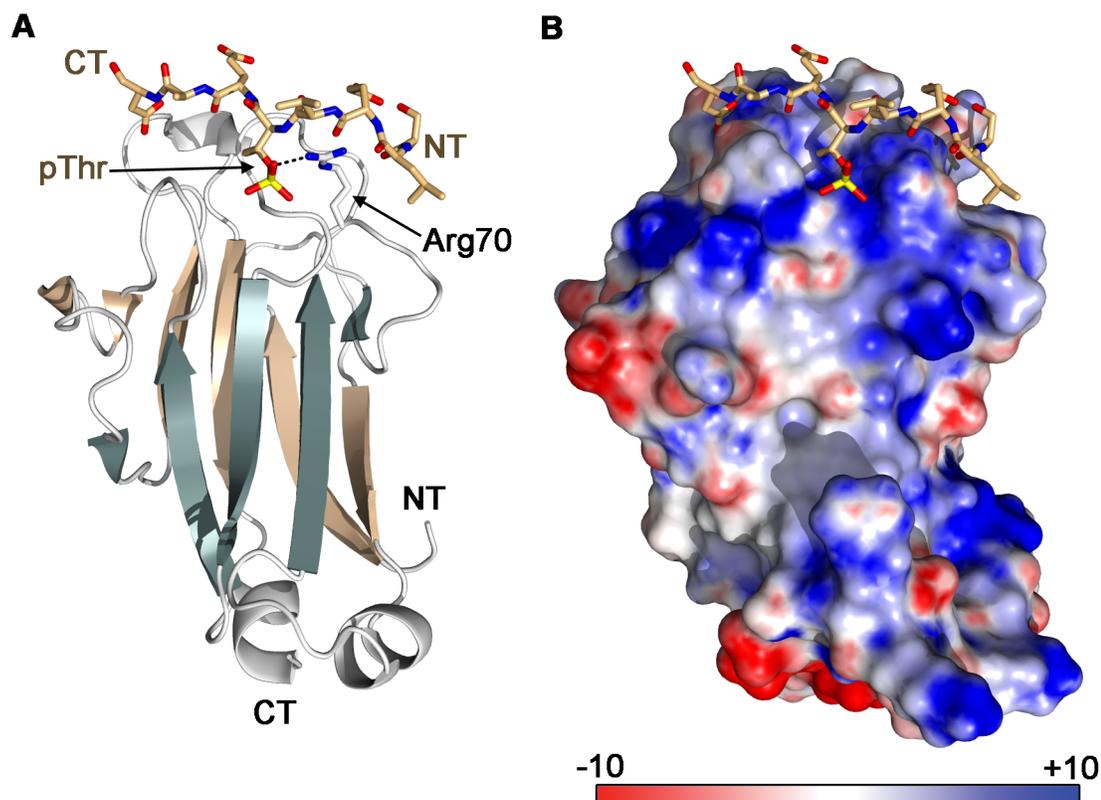
Figure 1.9 Interaction of a phospho-peptide with the FHA1 domain. (A) Ribbon diagram of Rad53^{FHA1} (PDB ID: 1G6G) interacting with a phospho-peptide shown as sticks. The hydrogen bond between Arg70 and the phospho-threonine (pThr) of the peptide is represented by a dotted line. The peptide has nitrogen atoms colored blue, oxygen atoms colored red and the phosphate colored yellow. The N- and C-termini of both Rad53^{FHA1} and the peptide are indicated. **(B)** A surface representation of Rad53^{FHA1} in the same orientation as panel A 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 charge of an electron. This figure was generated using APBS software (Baker et al., 2001) and visualized in PyMOL.

The N-terminal FHA domain in Rad53 (FHA1) is known to mediate more than 30 protein-protein interactions, and is specifically involved in the replication checkpoint. In contrast, the C-terminal FHA domain (FHA2) is involved in checkpoint responses to DNA damage (Pike et al., 2004). One of the binding partners for FHA1 is Dbf4, and this specifically requires a conserved arginine (Arg70) in the FHA1 domain that is critical for

recognizing a phospho-threonine residue in its target (Figure 1.9A) (Duncker et al., 2002). Chk1 in humans also phosphorylates Dbf4 to prevent late origin firing despite lacking an FHA domain (Kim et al., 2003). Purified Chk1 can phosphorylate Dbf4 *in vitro*, and the proteins interact when over-expressed together *in vivo* based on pull down assays (Heffernan et al., 2007). This is important for tumor progression as most cancer cells over-express Dbf4 due to a gene amplification event (Bonte et al., 2008). Although some cancers also over-express Cdc7, this rarely confers a proliferative advantage with the exception of melanoma (Nambiar et al., 2007). In most cases, the over-expression of Dbf4 instead allows newly synthesized protein to replace Chk1-phosphorylated Dbf4 during replication stress. This maintains Cdc7 activity and allows late origins to continue firing, which could contribute to genetic instability.

Chk1 is able to perform many of the same functions as Rad53. However, the Rad53 FHA domains also have unique functions in yeast not shared by Chk1. For example, the FHA1 domain is able to induce a change in yeast morphology (Shi et al., 2007). This helps pathogenic strains of fungi, such as *Candida albicans*, to evade the immune system. *C. albicans* grows normally as ovoid cells that are recognized and engulfed by macrophages. These immune cells bombard the fungus with reactive oxygen species generating DNA damage (Lorenz et al., 2004), which activates the yeast's Rad53 kinase. This kinase relies on its FHA1 domain to induce a dramatic elongation of the yeast cells into hyphae, which bursts the macrophages and allows *C. albicans* to continue the infection cycle. Rad53 in *S. cerevisiae* also triggers an elongation of the cell during replication stress, however this yeast is not pathogenic (Jiang and Kang, 2003).

Fungi are not the only pathogens that take advantage of replication stress. Viruses encode their own helicases free of the regulatory constraints that control the MCM2-7 helicase (Frick and Lam, 2006). The replication checkpoint that prevents helicase activation at late origins is therefore not able to stop viral replication. However, even more deviously, inducing replication stress in the host cell to purposefully suppress its late origins will result in resources that support replication to be shunted entirely to the

virus (Lilley et al., 2007). It will also ensure that cells remain in S phase, which maintains the environment necessary for viral replication (Lilley et al., 2007). Induced replication stress can also have other benefits for a virus. For example, _H_uman _P_apilloma_v_irus (HPV) 16 is a high-risk form of the virus for causing cervical cancer (Lorincz et al., 1992). It encodes an oncoprotein, E7, which drives cells from G1 phase into S phase where the virus can replicate (Spardy et al., 2009). This also inherently triggers replication stress. The stalled replication forks induce changes in the genome, such as chromosomal rearrangements and dsDNA breaks (Spardy et al., 2009). Importantly, the virus may take advantage of these breaks to integrate its genetic material into the chromosomes of the host (Spardy et al., 2009). It is important to emphasize that replication stress is not unique to eukaryotes. Bacteria are thought to have at least one of their two replication forks stall during every cell division (Cox et al., 2000). Replication stress in bacteria also has an impact on pathogenesis, as it can induce biofilm formation (Gotoh et al., 2008). Thus, medication that causes replication stress could possibly trigger perturbations in the microbial population in the body.

1.7 Dbf4 Domain Organization

DNA replication and the response to stalled replication forks are pathways that are inherently intertwined. It is not possible to disturb DNA replication without impacting the natural balance of replication forks. It is therefore essential to understand the varied cellular responses to replication stress and pinpoint common effector proteins. Dbf4 is of special interest as it acts as a scaffold that is able to mediate an interaction with the Cdc7 and Rad53 kinases in order to allow crosstalk between DNA replication and the response to stress. Despite its essential role, Dbf4 only has three small regions of conservation making the identification of homologs difficult (Masai and Arai, 2000). For example, Dbf4 was thought to be completely absent from pathogenic protozoa until it was discovered in a pull down assay using antibodies that recognize Cdc7 from _Giardia duodenalis_ (Lalle et al., 2012). Thus, Dbf4 is likely present in other protozoa as well, but is too divergent to be identified using genome sequences.

The three conserved regions are named motifs N, M and C, which refers to their position in the polypeptide (the _N_-terminus, _M_iddle, and _C_-terminus, respectively) (Masai and Arai, 2000). While this order is always maintained, the architecture of Dbf4 proteins has three major variations depicted in Figure 1.10. Each of the three motifs is responsible for mediating one or more interactions, and they can also influence each other. Aside from directly activating Cdc7, Dbf4 is also essential for contacting a variety of Cdc7 substrates. Interestingly, these targets change with the phase of the cell cycle, much the same as the targets of Cdc28. However, Cdc28 is able to swap cyclin partners during phase transitions and different cyclins have affinities for different targets (Bloom and Cross, 2007). In contrast, Cdc7 is bound to a single Dbf4 protein during the entire cell cycle, except for G1 when Dbf4 is degraded. This implies that the interactions mediated by Dbf4 must be exquisitely controlled to keep in tune with the cellular environment.

1.7.1 The Dbf4-Cdc7 Interaction

The interaction between Dbf4 and Cdc7 has been well characterized due to a recent crystal structure of the complex (Hughes et al., 2012) (Figure 1.11). Although the full-length proteins were not amenable to crystallization, truncations and internal deletions of flexible regions allowed a modified complex to yield diffraction quality crystals. Both motifs M and C of Dbf4 are capable of binding to Cdc7 simultaneously (Hughes et al., 2012; Ogino et al., 2001). Motif C is a zinc finger domain that binds the N-terminal lobe of the Cdc7 kinase domain (Hughes et al., 2012). This orients an α-helix that harbors essential catalytic residues so that it can engage ATP, which is a common strategy for allosteric regulators to activate their kinase binding partners (Hughes et al., 2012). Conversely, motif M binds to the C-terminal lobe of the Cdc7 kinase domain (Hughes et al., 2012). This motif is predominantly random coil, but has β-strands that extend a β-sheet in Cdc7 (Hughes et al., 2012). The distinguishing feature of this motif is four conserved proline residues that induce kinks in the backbone (Hughes et al., 2012; Masai and Arai, 2000). Overall, motif M forms an extended interaction surface, although
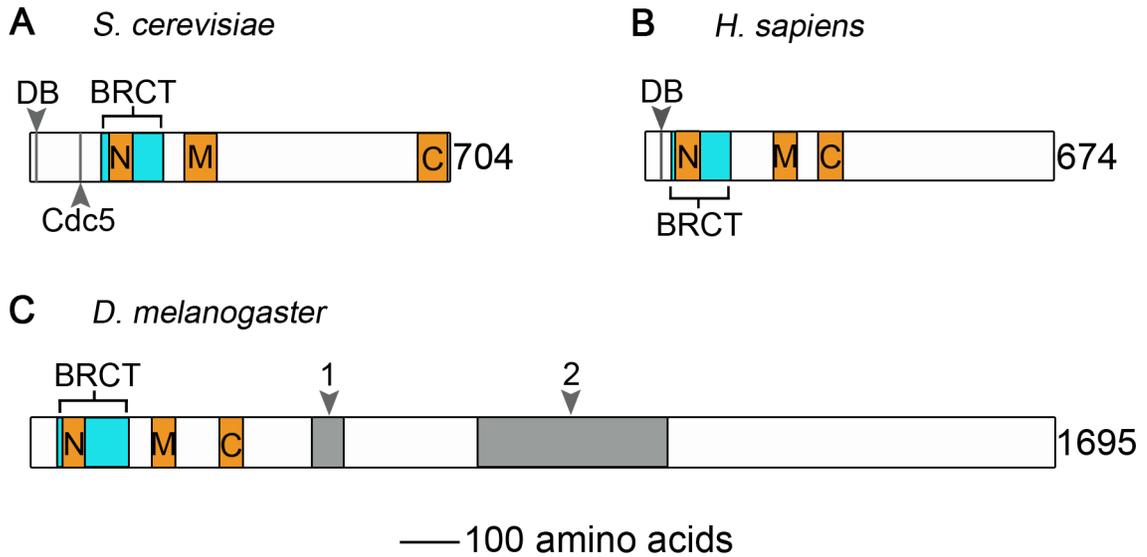
Figure 1.10 Comparison of Dbf4 homologs. Bar diagrams of Dbf4 homologs from **(A)** *S. cerevisiae*, **(B)** *Homo sapiens* (*H. sapiens*) and **(C)** *Drosophila melanogaster* (*D. melanogaster*). A number indicating the total amino acid length follows each diagram. The motifs N (*S. cerevisiae* residues 135-179; *H. sapiens* residues 48-92; *D. melanogaster* residues 45-89), M (*S. cerevisiae* 260-309; *H. sapiens* residues 214-253; *D. melanogaster* residues 200-237) and C (*S. cerevisiae* residues 656-697; *H. sapiens* residues 291-331; *D. melanogaster* residues 309-350) are conserved amongst all homologs and boxed in orange. Motif N is a part of a larger, putative BRCT domain boxed in blue. Grey boxes represent additional motifs that are not common to all Dbf4 homologs. The destruction box (DB), Cdc5 interaction motif (Cdc5) as well as two conserved regions in the C-terminal tail of the homolog from *D. melanogaster* (1 and 2) are labeled. Adapted from (Masai and Arai, 2000).

it only modestly increases the catalytic activity of Cdc7 (Hughes et al., 2012).

Such bipartite interaction surfaces have been reported for other protein complexes with varying purposes. One benefit is that bipartite interactions can increase the avidity and allow weak interactions to synergize. However, deleting either motif M or motif C does not prevent the interaction between Dbf4 and 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
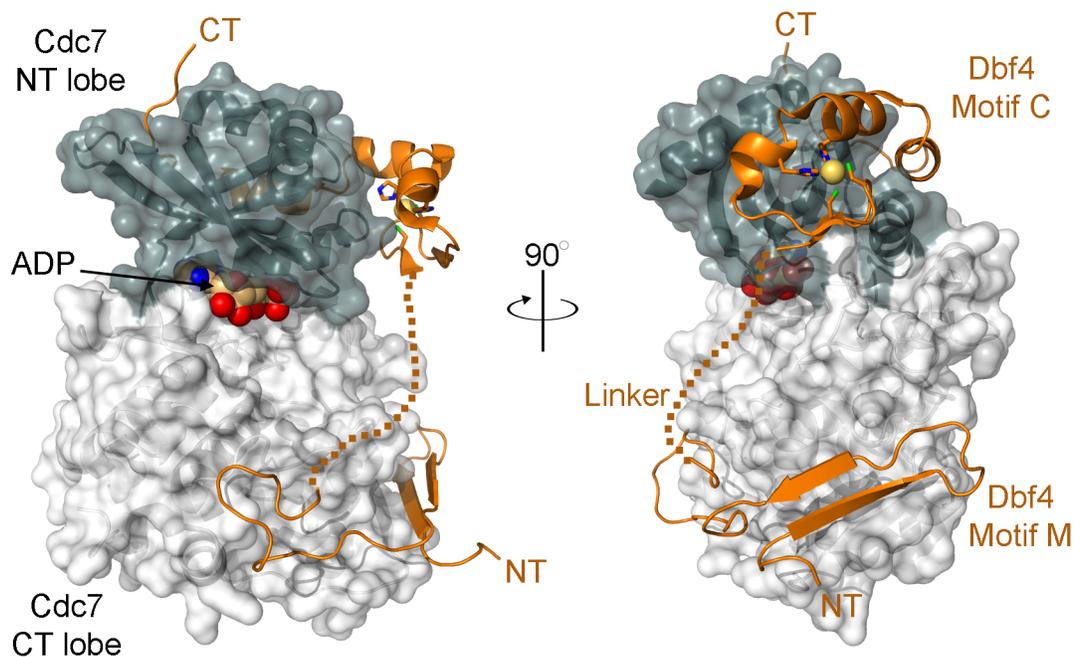
Figure 1.11 Bipartite interaction between Dbf4 and Cdc7. The complex (PDB ID: 4F9A) is shown as a ribbon diagram with a semitransparent surface on Cdc7. The N-terminal (NT) and C-terminal (CT) lobes of the Cdc7 kinase domain are colored green and grey, respectively. A molecule of ADP is represented with spheres to demonstrate the relative position of the kinase active site to the Dbf4 regulatory subunit. Dbf4 is colored orange and has a disordered linker region that is represented by a dotted line. Two cysteine residues (Cys296 and Cys299) and two histidine residues (His309 and His315) in motif C chelate a zinc cation (yellow sphere).

(Hughes et al., 2012). Bipartite interactions across the N- and C-terminal lobes of a kinase domain can also serve to orient the two lobes in an optimal position relative to each other. For example, the scaffold protein Ste5 activates a MAP kinase through this strategy (Bhattacharyya et al., 2006). However, this would place constraints on the linker joining motifs M and C together. The length of the linker is highly variable when comparing different Dbf4 homologs (Figure 1.10). It is even possible to express motifs M and C as separate polypeptides (i.e. without a linker) and achieve moderate Cdc7 activity in fission yeast (Ogino et al., 2001). These observations are not in keeping with the mechanism used by Ste5. Instead, the current hypothesis is that motif M can sustain

an interaction with Cdc7, while motif C engages or disengages to turn the activity of the kinase on or off (Hughes et al., 2012).

Although motifs M and C work together to bind and activate Cdc7, Dbf4 is also able to use additional regions to fine-tune the kinase activity. Specifically, the C-terminal tail in Dbf4 homologs from animals (Figure 1.10) can also bind Cdc7 (Sato et al., 2003). This association decreases Cdc7 activity, which can be alleviated by phosphorylating the C-terminal tail or having it interact with another binding partner (Hughes et al., 2010). The C-terminal tail can also confer unique functions. For example, the C-terminus is so expanded in Dbf4 homologs from insects that the protein is more than twice the size of those from yeast and animals (Figure 1.10). In *D. melanogaster*, the C-terminal tail includes two additional conserved regions that are important for causing repeated origin firing near the chorion genes (Tower, 2004). These are genes that encode protein products required to build the chorion, which is a protective layer around the egg. The amplification of these genes during replication is essential for proper egg development and consequently reduced levels of Dbf4 causes female sterility in fruit flies (Tower, 2004).

1.7.2 The Dbf4-Rad53 Interaction

While motifs M and C bind Cdc7, motif N is left free to interact with the Rad53 kinase. Deletions or point mutations within this region destroy the interaction in a yeast two-hybrid assay (Gabrielse et al., 2006; Varrin et al., 2005). As expected, the mutants were also hypersensitive to replication stress due to an inefficient checkpoint response (Gabrielse et al., 2006; Varrin et al., 2005). Motif N resembles a BRCT domain (Masai and Arai, 2000). These domains are notoriously difficult to identify by primary sequence alone, as there is not one single invariant amino acid (Bork et al., 1997). However, they do have a characteristic pattern of hydrophobic residues that form the core of the fold (Bork et al., 1997). A single BRCT domain is approximately 100 amino acids (Callebaut and Mornon, 1997), and so motif N is a little less than half the expected size. However, this motif was expanded to include additional flanking sequences that also exhibited

conserved patterns of hydrophobic residues (amino acids 117-218) (Gabrielse et al., 2006). Although these new boundaries were consistent with the size of a BRCT domain, the pattern of residues in these additional regions deviated from known BRCT folds in a sequence alignment (Gabrielse et al., 2006). Consequently, at the onset of this thesis the tertiary structure of motif N was unknown.

The molecular mechanism by which motif N recognizes Rad53 was also unknown. Interactions mediated by single BRCT domains are relatively uncharacterized compared to phospho-epitope recognition by tandem pairs (Figure 1.6). Indeed, single modules are best known for their DNA binding properties, rather than protein interactions. However, considering that Arg70 in the FHA1 domain of Rad53 is required for the interaction with Dbf4 (Figure 1.9), it was hypothesized that a phospho-threonine is present in motif N that is recognized by FHA1 (Duncker et al., 2002). In addition to Rad53, Dbf4 also uses its N-terminal domain to bind to the Orc2 subunit found in pre-RCs (Duncker et al., 2002). Although originally suspected of being a mechanism by which the ORC recruits DDK to origins prior to activation of their associated MCM2-7 helicases, this was ruled out when deleting motif N did not result in replication defects (Gabrielse et al., 2006). It is now understood that Dbf4 motif M and C as well as Cdc7 bind to MCM subunits showing that the recruitment of DDK to origins is highly redundant (Jones et al., 2010; Sheu and Stillman, 2006; Varrin et al., 2005).

1.8 Thesis Objectives

During replication stress, Dbf4 must bind to the Cdc7 and Rad53 kinases to suppress late origin firing. While the molecular details of the interaction with Cdc7 have been largely elucidated, relatively little is known of how the interaction between Dbf4 and Rad53 occurs. To characterize this interaction we have pursued the following aims:

(i)   Defining the minimal domain containing motif N by comparing the solubility of different sized fragments from the N-terminus of Dbf4 (Chapter 2, 3).

(ii)   Determining the minimal fragment of Dbf4 that is sufficient for the interaction with Rad53 (Chapter 3).

(iii)   Comparing the mode of interaction between this minimal fragment of Dbf4 and the FHA1 domain of Rad53 with the interaction between full-length proteins *in vivo* (Chapter 4).

Chapter 2

Crystallization and Preliminary X-ray Diffraction Analysis of Motif N from *Saccharomyces cerevisiae* Dbf4

2.1 Author's Preface

This chapter describes the domain boundaries surrounding motif N in Dbf4, which were determined by monitoring different sized fragments of Dbf4 for solubility in an *E. coli* expression host. The best-behaved fragment was then purified and crystallized. A. Duong and A.A. Prasad (from the laboratory of Dr. Duncker, University of Waterloo) performed the cloning of the Dbf4 fragments. I purified the proteins, conducted the crystallography experiments and wrote the manuscript with Dr. Guarné.

2.2 Abstract

The Cdc7-Dbf4 complex plays an instrumental role in the initiation of DNA replication and is a target of replication checkpoint responses in *Saccharomyces cerevisiae*. Cdc7 is a conserved serine/threonine kinase whose activity depends on association with its regulatory subunit, Dbf4. A conserved sequence near the N-terminus of Dbf4 (motif N) is necessary for the interaction of Cdc7-Dbf4 with the checkpoint kinase Rad53. To understand the role of the Cdc7-Dbf4 complex in checkpoint responses, a fragment of *Saccharomyces cerevisiae* Dbf4 encompassing motif N was isolated, overproduced and crystallized. A complete native data set was collected at 100 K from crystals that diffracted X-rays to 2.75 Å resolution and structure determination is currently under way.

2.3 Introduction

The initiation of DNA replication requires the formation of a pre-replicative complex at the origins and the activity of two kinases, which are known as Cdc28 and Cdc7 in the budding yeast *Saccharomyces cerevisiae* (Bell and Dutta, 2002). In this organism, the activity of Cdc28 is regulated by six different cyclins, while that of Cdc7 depends on a single regulatory subunit, the Dbf4 protein (Diffley, 1998; Jackson et al., 1993; Ogino et al., 2001). The Cdc7-Dbf4 complex, also known as DDK (*D*bf4-*d*ependent *k*inase), phosphorylates targets found at licensed origins, triggering the onset of DNA replication, and plays an integral role during cellular checkpoint responses (Masai and Arai, 2000; Sclafani, 2000). During genotoxic stress, the Cdc7-Dbf4 complex is thought to phosphorylate the checkpoint kinase Rad53, leading to its full activation (Ogi et al., 2008). Dbf4 is then phosphorylated in a Rad53-dependent manner, causing its dissociation from chromatin and the inhibition of Cdc7 kinase activity (Kihara et al., 2000; Weinreich and Stillman, 1999), consequently regulating DNA replication by suppressing the activation of unfired origins (Bousset and Diffley, 1998; Donaldson et al., 1998; Pasero et al., 1999).

Dbf4 orthologs are variable in length and sequence but encompass three conserved motifs named N, M and C based on their relative location in the polypeptide chain (Masai and Arai, 2000). In *S. cerevisiae,* motifs M (residues 260-309) and C (residues 659-696) are included within the regions that have been shown to interact with Cdc7 (Dowell et al., 1994; Hardy and Pautz, 1996). Additionally, motif M mediates the association between Dbf4 and the Mcm2 protein, a subunit of the hexameric MCM ring within the pre-replicative complex (Lei et al., 1997; Varrin et al., 2005). Motif N (residues 129-177) mediates interactions with both the Orc2 subunit of the origin recognition complex (ORC) and the checkpoint kinase Rad53 (Varrin et al., 2005). Based on primary sequence analysis, it has been suggested that residues 117-218, encompassing the entire motif N, could resemble a BRCT domain with a divergent C-terminus (Gabrielse et al., 2006). Originally identified at the C-terminus of the BRCA-1 protein, BRCT domains are widely found in DNA repair and replication checkpoint proteins (Bork et al., 1997), reinforcing the idea that Dbf4 could mediate protein-protein interactions during checkpoint responses. However, the specific interactions that mediate the association of Cdc7-Dbf4 with Rad53 and the consequent phosphorylation of Rad53 and Dbf4 remain unclear.

In an effort to understand how Dbf4 mediates the interaction between Cdc7-Dbf4 and Rad53, we overproduced, purified and crystallized Dbf4-N (residues 120-250), a segment of *S. cerevisiae* Dbf4 encompassing motif N. Here, we describe the methods used to improve crystal nucleation and growth, as well as the preliminary X-ray analysis of Dbf4-N crystals diffracting to 2.75 Å resolution.

2.4 Experimental Methods

2.4.1 Cloning and Solubility Assays

Three constructs encompassing motif N of *S. cerevisiae* Dbf4 were generated: residues 120-250 (pAG8206), residues 120-220 (pAG8207) and residues 105-250 (pAG8208) (Fig. 2.1). These DBF4 fragments were subcloned into the pET15b vector

(Novagen) using *Nde*I and *Bam*HI restriction sites and confirmed by DNA sequencing (MOBIX Laboratory, McMaster University). The resulting recombinant proteins included a histidine tag (removable by thrombin digestion) at their N-termini. To test solubility, the proteins were overproduced in *Escherichia coli* BL21 Star (DE3) cells (Invitrogen Life Technologies) containing a plasmid encoding tRNA molecules rarely used in *E. coli* (pRARELysS). The cell cultures were grown to an $OD_{600} \sim 0.7$ at 310 K prior to inducing protein production by the addition of 1 mM β-D-1-thiogalactopyranoside. Cells were harvested by centrifugation after either 3 h at 310 K, 5 h at 298 K or 12 h at 289 K and were lysed using a combination of lysozyme (0.5 mg/mL), salt (0.5 M NaCl) and detergent (0.03% lauryl dimethylamine N-oxide). The solubility of each fragment was assessed on 15% SDS-polyacrylamide gels stained with Coomassie Brilliant Blue.

2.4.2 Protein Expression and Purification

Cells producing Dbf4-N (residues 120-250) were harvested after 5 h at 298 K and washed with phosphate-buffered saline. Cell pellets were resuspended in buffer A (20 mM Tris pH 8.0, 500 mM NaCl, 1.4 mM β-mercaptoethanol, 15 mM imidazole and 5% glycerol) and lysed by sonication. Cell lysates were subsequently clarified by centrifugation (39 000 g at 277 K for 40 min). The supernatent was loaded onto a 5 mL HiTrap HP column (GE Healthcare) charged with Ni^{2+}. The column was then washed with buffer A, followed by buffer A supplemented with 36 mM imidazole. Dbf4-N was eluted with 300 mM imidazole and further purified by ion exchange on a MonoS 10/100 GL column (GE Healthcare) equilibrated with 20 mM Tris pH 8.0, 1 mM EDTA, 5 mM DTT, 100 mM NaCl and 5% glycerol. Pure Dbf4-N was eluted from the column using a linear gradient to 500 mM NaCl, concentrated to 7 mg/mL and stored in 20 mM Tris pH 8.0, 100 mM NaCl, 1 mM DTT and 5% glycerol (Fig. 2.2A). Protein concentration was estimated using a Bradford assay (Bradford, 1976).

2.4.3 Crystallization and Data Collection

Crystallization drops were prepared at 277 K by mixing 1 μL protein solution (7 mg/mL) with 1 μL crystallization solution [27-28% PEG 500, 0.05 M MgCl$_2$, 0.1 M Tris pH 8.5 and 4-5.5% pentaerythritol ethoxylate (PEE)] and subsequently equilibrated against increasing concentrations of ammonium sulfate (0.5-2.5 M) using a modified vapor-diffusion protocol as described previously (Dunlop and Hazes, 2005; Newman, 2005). Crystals of Dbf4-N were flash-cooled in liquid nitrogen. A complete data set was collected on the X29 beamline of the National Synchrotron Light Source (NSLS), Brookhaven National Laboratory (BNL) using 1° oscillations and a crystal-to-detector distance of 300 mm. Diffraction data were indexed, integrated and scaled using the HKL-2000 package (Otwinowski and Minor, 1997).

2.5 Results and Discussion

2.5.1 Domain Boundaries of Motif N

Motif N was originally reported as encompassing residues 129-177 and its resemblance to BRCT domains has been noted (Masai and Arai, 2000). In a subsequent study, the region surrounding motif N (residues 117-218) was predicted to define a BRCT domain with a divergent sequence at its C-terminus that is exclusively found in Dbf4 proteins (Gabrielse et al., 2006). Based on these two reports and secondary-structure predictions of multiple Dbf4 orthologs using PSIPRED (Bryson et al., 2005), three different constructs were generated that differed in the structural elements included at the domain boundaries (Figure 2.1). While the fragment encompassing residues 120-220 included the predicted BRCT motif of Dbf4, it exhibited minimal expression and solubility levels. The inclusion of an additional fragment at the C-terminus (residues 120-250) increased both the expression and the solubility of the fragment, while the addition of residues 105-119 at the N-terminus was without effect. Therefore, subsequent crystallographic analysis was conducted using the fragment encompassing residues 120-250, which is hereafter referred to as Dbf4-N.
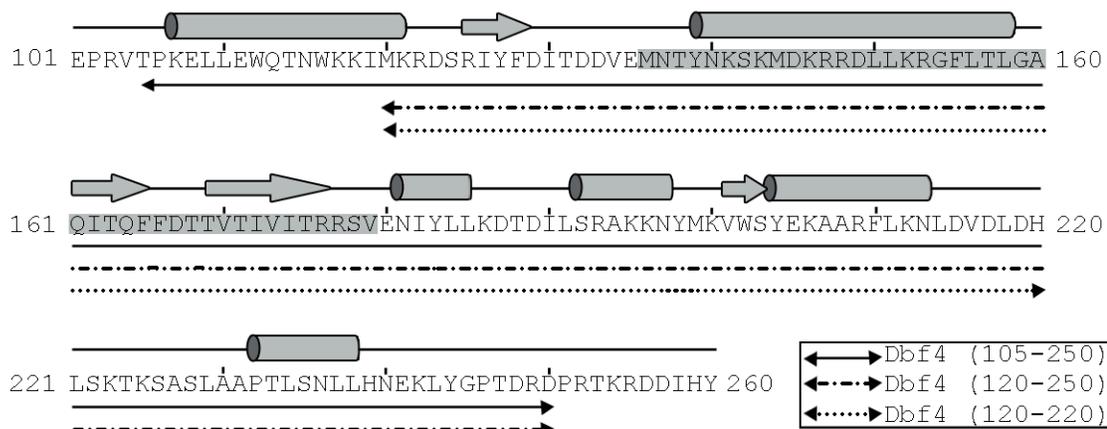
Figure 2.1 Sequence and secondary structure prediction of Dbf4 (residues 101-260). The fragments generated for crystallization are indicated by double-headed arrows and labeled. Residues 129-177, originally identified as motif N (Masai and Arai, 2000), are shaded. Predicted secondary structure elements are shown as cylinders (α-helices) and arrows (β-strands).

2.5.2 Crystallization of Dbf4-N

Two different isoforms of Dbf4-N were isolated using cation exchange chromatography (Figure 2.2). Both isoforms behaved as monomers in solution as judged by size-exclusion chromatography (data not shown). However, mass spectrometric analysis revealed that while the major isoform (17,346 Da) corresponded to unmodified Dbf4-N, the minor isoform was a mixture of two modified species with molecular weights of 17,389 and 17,525 Da. This likely corresponds to carbamylation (43 Da) and gluconoylation (178-179 Da), respectively. Therefore, only the major isoform was used for subsequent crystallization analysis.

Dbf4-N crystals only appeared after stepwise dehydration with 0.5-2.5 M ammonium sulfate for at least two weeks and grew to their maximum size in about two months (Figure 2.3). Optimization of the protein or precipitant concentration, the pH, the ionic strength or the type of precipitant did not change the nucleation of Dbf4-N crystals. We entertained the idea that delayed crystallization was a consequence of protein

degradation, but mass spectrometry of the crystals revealed that they contained intact Dbf4-N (data not shown). It has previously been shown that ammonium sulfate can trigger pH changes owing to ammonia transfer following ammonium-ammonia equilibrium (Mikol et al., 1989), suggesting that the formation of Dbf4-N crystals was partially driven by a pH gradient.
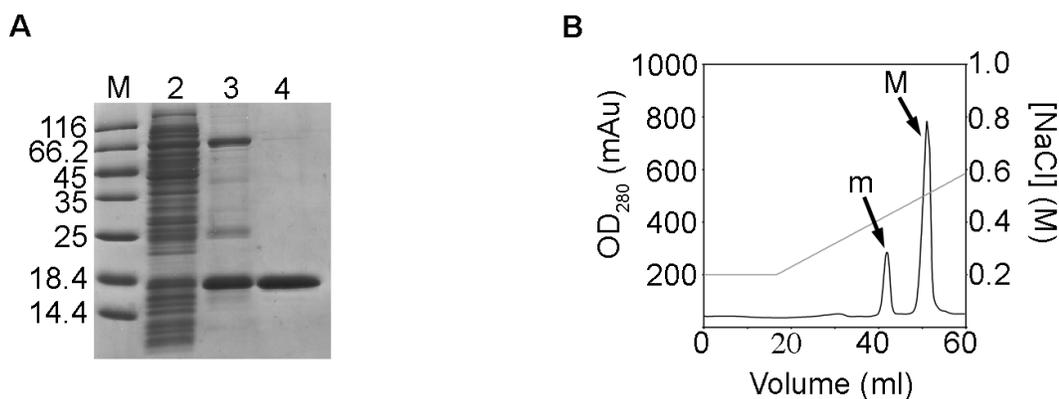
Figure 2.2 Protein expression and purification of Dbf4 fragments encompassing motif N. (A) Coomassie-stained 15% SDS-polyacrylamide gel depicting the purification steps of the major isoform of Dbf4-N (120-250). Lane M, molecular-weight markers (labeled in kDa); lane 2, cell lysate; lane 3, after elution from Ni-affinity column; lane 4, after elution from MonoS column. **(B)** Chromatogram corresponding to the elution of Dbf4-N (120-250) from a MonoS 10/100 GL column (GE Healthcare). The major (M) and minor (m) isoforms are indicated.

In order to grow suitable crystals for X-ray diffraction, we screened a battery of crystallization additives. Pentaerythritol ethyoxylate (PEE) decreased the solubility of Dbf4-N since crystals started appearing in drops equilibrated against 1.5 M rather than 2.5 M ammonium sulfate. Although PEE was initially identified as a promising precipitant that increased the quality of crystals (Gulick et al., 2002), in this instance it presumably acted by enhancing crystal nucleation. However, we found that it could only substitute for PEG 400 in limited amounts because the PEE concentration was positively correlated with the number of crystals obtained (Figure 2.3A). As such, PEE represented

a trade-off between decreasing the time required for crystallization and a reduction in crystal size owing to enhanced nucleation. Optimal Dbf4-N crystals were grown in 27-28% PEG 400, 0.05 M MgCl$_2$, 0.1 M Tris pH 8.5 and 4-5.5% PEE and grew to approximate dimensions of 0.2 x 0.1 x 0.05 mm (Figure 2.3B).
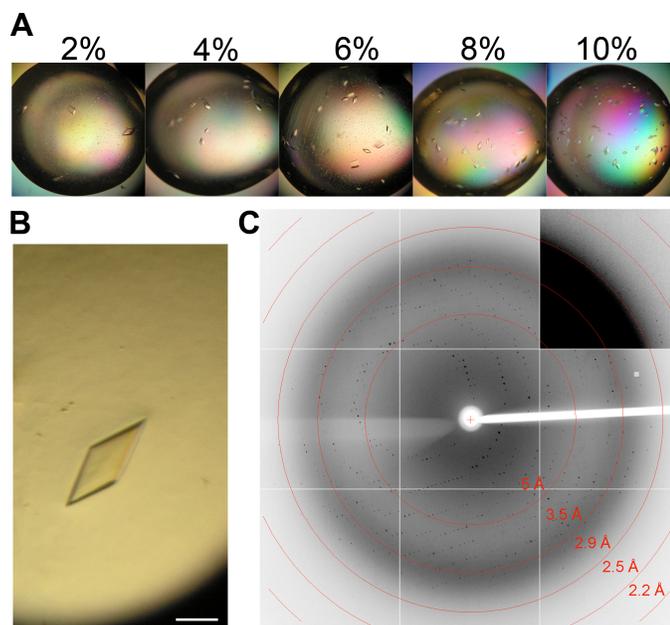
Figure 2.3 Crystallization and diffraction of Dbf4-N (120-250). **(A)** The effect of pentaerythritol ethoxylate (PEE) on the crystallization of Dbf4-N. Increasing concentrations of PEE favor nucleation but decrease crystal size (left to right). **(B)** Representative crystal of Dbf4-N (120-250) grown in 27.6% PEG 400, 0.05 M MgCl$_2$, 0.1 M Tris pH 8.5 and 4 – 5% PEE. The scale bar represents 0.1 mm. **(C)** Representative X-ray diffraction pattern of a Dbf4-N (120-250) crystal collected on beamline X29 (NSLS, BNL). Data were collected at a crystal-to-detector distance of 300 mm and a wavelength of 1.29 Å. The resolution at the detector edge is 2.6 Å. In the top right inset, the darkness of the image has been adjusted to highlight weak reflections.

2.5.3 Diffraction Analysis of Dbf4-N Crystals

The Dbf4-N (120-250) crystals belonged to space group P2$_1$, with unit-cell parameters a = 89.7, b = 79.1, c = 127.3 Å, β = 110.6° (Table 2.1). According to the Matthews coefficient calculation, there could be between eight (V$_M$ = 3.02 Å^3 Da^{-1}) and

twelve (V_M = 2.01 Å^3 Da^{-1}) molecules in the asymmetric unit. Analysis of the self-rotation function calculated with MOLREP revealed the presence of perpendicular twofold axes, as well as two pseudo-fivefold axes (κ = 71.75°) (Collaborative Computational Project Number 4, 1994). The quasi-fivefold symmetry supports the presence of ten molecules (V_M = 2.42 Å^3 Da^{-1}) in the asymmetric unit over the other possible cell contents (Table 2.1). Interestingly, it has previously been shown that oligomers of the Cdc7-Dbf4 kinase exist in *S. cerevisiae* and it has been suggested that oligomerization may play a regulatory role (Shellman et al., 1998). However, the nature of these oligomers is still unclear.

We tried to solve the Dbf4-N structure by molecular replacement using various structures containing single BRCT domains (PDB codes 2JW5, 2EBU, 1L7B and 2EBW). Structures containing tandem BRCT repeats were excluded from our list because the interdomain linker is believed to pose restrictions on the overall BRCT fold (Glover et al., 2004), which would not exist in proteins containing a single BRCT domain such as Dbf4. Dbf4-N only has around 10% sequence identity and 30-35% sequence similarity to these BRCT domains; however, these values are similar to those obtained when comparing *bona fide* BRCT domains using structure-guided sequence alignments.

Unfortunately, none of these models yielded a solution for the structure of Dbf4-N. The lack of successful results using various molecular replacement protocols could have been caused by several factors. Firstly, BRCT domains are notoriously flexible and flexibility is known to jeopardize structure determination by molecular replacement. However, the results obtained using search models generated as ensembles of multiple structures were similar to those obtained using models containing a single BRCT structure. Secondly, the asymmetric unit of the Dbf4-N crystal presumably includes ten Dbf4-N molecules, creating a disproportionate ratio between the volume of the search model and the volume being searched. Thirdly, Dbf4-N is proposed to fold as a noncanonical BRCT domain with its C-terminus adopting a structure that is unique to the Dbf4 family (Gabrielse et al., 2006). Indeed, the N- and C-terminal ends of the BRCT

Table 2.1 Data collection statistics for Dbf4-N (120-250) crystals

Experimental Conditions	
X-ray source	X29 (NSLS, BNL)
Wavelength (Å)	1.29
Temperature (K)	100
Detector	ADSC Q315 CCD
No. of images	100
Exposure time (s)	20
Oscillation angle (°)	1
Data processing	
No. of measured reflections	77,857
No. of unique reflections	41,469
Space group	$P2_1$
Unit-cell parameters (Å,°)	$a = 89.7$, $b = 79.1$, $c = 127.3$, $\beta = 110.6$
Resolution (Å)[a]	25-2.75 (2.85-2.75)
Completeness (%)[a]	95.1 (97.9)
Multiplicity[a]	1.9 (1.8)
Mean $I/\sigma(I)$[a]	14.8 (1.8)
R_{merge}[a]	0.057 (0.462)
No. of molecules per ASU[b]	10
Matthews coefficient V_M (Å^3 Da^{-1})	2.42
Solvent content (%)	49

[a]Values in parentheses are for the highest resolution shell
[b]Asymmetric unit

domains used for molecular replacement are extremely variable, suggesting that the BRCT fold can withstand both flexibility and sequence variability. To overcome these

potential problems, structure determination using multiple anomalous diffraction is currently under way.

2.6 Acknowledgements

Chapter 3

Saccharomyces cerevisiae Dbf4 has a Unique Fold Necessary for the Interaction with the Rad53 Kinase

3.1 Author's Preface

      This chapter refines the domain boundaries identified in Chapter 2 and presents the crystal structure of two variants of the BRCT domain from Dbf4. These are the first structures for any Dbf4 protein and a comparison with canonical BRCT folds revealed unique features that were not anticipated based on Dbf4 sequence alignments. In addition, an extensive search was conducted for any potential phospho-threonine residues that mediate the interaction with the FHA1 domain from Rad53. D.R. Jones, A.A. Prasad and Dr. Duncker performed all of the yeast two-hybrid analyses and survival assays. Molecular cloning was conducted by both D.R. Jones and myself. I also crystallized, solved and refined the structures of Dbf4. The manuscript was written by Dr. Guarné and myself except for the sections detailing the yeast experiments, which were provided by Dr. Duncker.

3.2 Abstract

Dbf4 is a conserved eukaryotic protein that functions as the regulatory subunit of the Dbf4-dependent kinase (DDK) complex. DDK plays essential roles in DNA replication initiation and checkpoint activation. During the replication checkpoint, *Saccharomyces cerevisiae* Dbf4 is phosphorylated in a Rad53-dependent manner, and this, in turn, inhibits initiation of replication at late origins. We have determined the minimal region of Dbf4 required for the interaction with the checkpoint kinase Rad53 and solved its crystal structure. The core of this fragment of Dbf4 folds as a BRCT domain, but it includes an additional N-terminal helix unique to Dbf4. Mutation of the residues that anchor this helix to the domain core abolish the interaction between Dbf4 and Rad53, indicating that this helix is an integral element of the domain. The structure also reveals that previously characterized Dbf4 mutants with checkpoint phenotypes destabilize the domain, indicating that its structural integrity is essential for the interaction with Rad53. Collectively, these results allow us to propose a model for the association between Dbf4 and Rad53.

3.3 Introduction

The Dbf4-dependent kinase (DDK) complex acts as the ultimate trigger of DNA replication by phosphorylating protein targets found at origins. Essential in all eukaryotic cells, DDK is a heterodimer composed of the Cdc7 kinase and its regulatory subunit Dbf4 (Bruck and Kaplan, 2009). Although the levels of Cdc7 are constant throughout the cell cycle, Dbf4 is only synthesized late in G1 phase and is subsequently degraded during mitosis, causing the kinase activity of Cdc7 to cycle accordingly (Cheng et al., 1999; Ferreira et al., 2000; Jackson et al., 1993; Weinreich and Stillman, 1999). Beyond activating origins of replication, the DDK complex also participates in other cellular processes, including meiosis (Marston, 2009; Matos et al., 2008; Patterson et al., 1986), mitotic exit (Miller et al., 2009), and the intra-S phase checkpoint (Pessoa-Brandao and Sclafani, 2004; Zegerman and Diffley, 2010). This latter function is especially important

because it acts to suppress further origin firing as well as to stabilize and ultimately restart stalled replication forks, which otherwise would become sites of genetic instability.

Rad53 is an effector kinase of the replication checkpoint in *Saccharomyces cerevisiae* that delays entry into M phase and suppresses harmful rearrangements of DNA (Lopes et al., 2001; Lucca et al., 2004; Santocanale and Diffley, 1998; Shirahige et al., 1998). Rad53 is activated through hyperphosphorylation mediated both by Rad53 *in trans* and additional kinases (Pellicioli and Foiani, 2005; Pellicioli et al., 1999). DDK phosphorylates Rad53 *in vitro*, and deletion of Cdc7 from yeast cells prevents Rad53 from achieving its hyperphosphorylated state, leading to an increased sensitivity to genotoxic stress (Kihara et al., 2000). Reciprocally, Dbf4 is one of the Rad53 targets during the checkpoint response, causing a significant reduction in Cdc7 kinase activity towards Mcm2 (Kihara et al., 2000). This prevents licensed origins that have not yet fired from initiating replication (Lopez-Mosqueda et al., 2010; Zegerman and Diffley, 2010). Collectively, this suggests that DDK is both an upstream regulator of Rad53 and a downstream target during the replication checkpoint.

Among Dbf4 homologs, only three short sequences are conserved. They are referred to as motifs N, M, and C to denote their location in the polypeptide chain (Masai and Arai, 2000). Motifs M and C (residues 260-309 and 656-697, respectively, in *S. cerevisiae* Dbf4) are required to bind and activate Cdc7 (Ogino et al., 2001), whereas motif N (residues 135-179 in *S. cerevisiae* Dbf4) is necessary for the interaction with Rad53 and the origin recognition complex (Duncker et al., 2002; Varrin et al., 2005). Deletions or point mutations within motif N of Dbf4 manifest as an increased sensitivity to the ribonucleotide reductase inhibitor hydroxyurea and DNA-damaging agents, suggesting that these mutants have an inefficient checkpoint response (Gabrielse et al., 2006; Varrin et al., 2005). An interaction between the human homolog Dbf4 (ASK) and the checkpoint kinase Chk1 has also been described. Chk1 mediates the S phase checkpoint response in higher eukaryotes and phosphorylates ASK *in vitro* (Heffernan et

al., 2007). Although it remains unclear whether the human DDK complex plays the same role in the checkpoint as the yeast DDK complex.

Rad53 contains two Forkhead-associated (FHA) domains. Dbf4 primarily interacts with the FHA1 domain of Rad53, although it also has weak affinity for the FHA2 domain of the protein (Duncker et al., 2002; Varrin et al., 2005). The FHA1 domain of Rad53 specifically recognizes phospho-threonine residues found in unstructured loops (Durocher et al., 1999). Therefore, the interaction between Rad53 and most of its binding partners can be recreated using phospho-threonine containing peptides. Because mutation of the FHA1 phosphate-binding pocket compromises the ability of this domain to recognize Dbf4, it was originally proposed that Rad53 would recognize a phospho-epitope within Dbf4 (Duncker et al., 2002). However, the phospho-threonine in Dbf4 responsible for this interaction has not been identified. Recent studies have unveiled additional phosphorylation-independent modes of interaction by which FHA domains can interact with their binding partners (Nott et al., 2009; Pike et al., 2004). One of these modes of interaction still involves the phosphate-binding pocket of the FHA domain (Nott et al., 2009), indicating that the interaction between Dbf4 and Rad53 could also be phosphorylation-independent.

It has been proposed that motif N is part of a larger structurally conserved unit that resembles a BRCT domain (Gabrielse et al., 2006; Weinreich and Stillman, 1999). In support of this idea, motif N does not constitute an independent folding unit by itself, but the region of Dbf4 encompassing residues 120-250 can be overproduced on its own, and it is well behaved in solution (Matthews et al., 2009). BRCT domains are commonly found in proteins that respond to DNA damage, and many of them function as tandem repeats, which associate together using conserved hydrophobic residues to create an intervening hydrophobic pocket (Leung and Glover, 2011; Mohammad and Yaffe, 2009). These repeats act as a single unit to recognize binding partners by using both this pocket as well as a phospho-serine binding site contained within one of the BRCT domains (Williams et al., 2001). However, the mechanism by which single BRCT domains, such

as the one presumably present in Dbf4, participate in protein-protein interactions is poorly understood.

In an effort to clarify whether Dbf4 has a *bona fide* BRCT domain and whether this domain mediates the interaction with Rad53, we have characterized the minimal region of Dbf4 necessary for the interaction with Rad53 and determined its crystal structure. We have found that the fragment of Dbf4 encompassing residues 105-221 is sufficient to mediate this interaction. The crystal structures of this sufficient domain of Dbf4 and a shorter fragment encompassing residues 120-221 unveil a structural feature that is unique to Dbf4. Residues 105-119 form an α-helix that is embedded in the core of the domain defining a novel structure that we name Helix-BRCT (H-BRCT). Using a combination of structure-guided site-directed mutagenesis and yeast two-hybrid analysis, we demonstrate that the presence of this N-terminal helix is critical for both folding of the domain and the interaction with Rad53. Collectively, our work demonstrates that the structural integrity of the H-BRCT domain is essential for the ability of Dbf4 to interact with Rad53 and promote cell survival under genotoxic stress.

3.4 Experimental Procedures

3.4.1 Protein Purification and Crystallization

The His-tagged version of the BRCT domain of Dbf4 was obtained by transforming the pAG8206 plasmid encoding residues 120-250 from *S. cerevisiae* Dbf4 (NP_010337) in either BL21(DE3) or B834(DE3) cells carrying the pRARELysS plasmid. SeMet-labeled protein was subsequently overproduced in minimal media supplemented with SeMet as described earlier (Hendrickson et al., 1990). Protein overproduction and purification for both native and SeMet-labeled Dbf4 was performed as described previously (Matthews et al., 2009). The histidine tag was removed with thrombin (Sigma), and the protein was further purified by ionic exchange in a MonoS (10/100) GL column (GE Healthcare). Purified tagged and untagged Dbf4 were concentrated to 5 mg/mL in 20 mM Tris pH 8.0, 1 mM DTT, 100 mM NaCl and 5%

glycerol (storage buffer). Longer fragments of Dbf4 (encoding residues 65-221 and 105-220, pAG8391 and pAG8209, respectively) were produced and purified similarly. Protein concentration was determined using the Beer-Lambert law with an extinction coefficient of 23,950 M^{-1} cm^{-1}.

SeMet crystals of a His_6-tagged Dbf4 fragment encompassing the predicted BRCT domain (residues 120-250) were grown using the hanging drop method in 0.1 M Tris pH 8.5, 0.05 M $MgCl_2$ and 29% PEG 400 (v/v). Native crystals of the untagged BRCT domain were grown in 100 mM sodium/potassium phosphate buffer, pH 7.1, and 22% MPD (v/v) and cryoprotected by increasing the MPD concentration to 40% (v/v). Crystals of the minimal domain of Dbf4 sufficient for the interaction with Rad53 (residues 65-221, Helix-BRCT) were grown in 1.9 M $(NH_4)_2SO_4$, 0.014 M $MgSO_4$, 0.05 M $(CH_3)_2AsO_2Na$ pH 6.5, and 0.03 mM CYMAL-7, improved by streak seeding, and cryoprotected by the addition of 20% glycerol to the mother liquor prior to data collection. All crystals were grown from crystallization drops set at 4°C by mixing 1 μL of protein with 1 μL of reservoir solution.

Complete data sets of SeMet-labeled tagged and untagged BRCT domain crystals were collected using the X29 beamline, whereas data for the crystals of the Helix-BRCT domain were collected using the X25 beamline at the National Synchrotron Light Source (Brookhaven National Laboratory). All data sets were processed with HKL2000 (Otwinowski and Minor, 1997) (Table 3.1).

3.4.2 Structure Determination and Refinement

The structure of the His-tagged BRCT domain was determined by the single anomalous diffraction method. Thirty of the forty selenium sites were found and refined using standard protocols in PHENIX (Adams et al., 2002). The initial model was refined by iterative cycles of manual model building in Coot (Emsley and Cowtan, 2004) and refinement using PHENIX and imposing non-crystallographic symmetry constraints that related the 10 molecules in the asymmetric unit pairwise. This protocol yielded a final

model that includes residues 122-224 and has 93.6% of the residues in the most favored region of the Ramachandran plot and none in the disallowed regions. One BRCT monomer from this structure was subsequently used as the search model to solve the structure of the untagged BRCT domain using molecular replacement with Phaser (McCoy et al., 2007). The structure of the untagged BRCT domain was refined using standard protocols in Coot and PHENIX, and the final model includes residues 119-242 for one of the five monomers in the asymmetric unit and residues 122-220 for the other four, has 94.2% of the residues in the most favored region of the Ramachandran plot, and has none in the disallowed regions.

The structure of the sufficient domain (H-BRCT) was solved by molecular replacement using the BRCT domain as the search model. The final model includes residues 98-220 and has 94% residues in the most favored regions of the Ramachandran plot and none in the disallowed regions. The refinement statistics for the three structures are presented in Table 3.1. Figures depicting molecular structures were generated using PyMOL (DeLano, 2002).

3.4.3 Protein Solubility Assays

Cells harboring plasmids encoding variants of the Dbf4 fragment encompassing residues 105-220 [H-BRCT-T171A (pAG8494), H-BRCT-GA159LL (pAG8393), H-BRCT-W202A (pAG8492), and H-BRCT-Δ132-144 (pAG8496)] were transformed in BL21(DE3) cells carrying the pRARE plasmid were grown at 37°C to $OD_{600} \sim 0.7$ in the presence of ampicillin (50 μg/mL) and chloramphenicol (25 μg/mL). At that point, protein expression was induced by addition of 1 mM IPTG. Cell cultures were incubated for 4 hours at 30°C and harvested by centrifugation. Cell pellets were resuspended in lysis buffer (20 mM Tris pH 8.0, 1.4 mM β-mercaptoethanol, 1 mg/mL lysozyme) and allowed to lyse for 30 minutes. Cell lysates were subsequently supplemented with 0.5 M NaCl and 0.05% (v/v) LDAO (lauryldimethylamine-oxide), mixed by inverting and incubated for 15 minutes. Soluble proteins were separated from insoluble proteins and cell debris by centrifugation at 13,000 g for 20 minutes. Samples taken before induction

with IPTG (-), after 4 hours incubation (+), and from the soluble (S) and insoluble (I) fractions after cell lysis were resolved in 15% SDS-polyacrylamide gels to assess the level of over-expression, as well as, the solubility of each Dbf4 variant.

3.4.4 Size Exclusion Chromatography

Dbf4 variants were purified as described earlier (Matthews et al., 2009) and stored in 20 mM Tris pH 8.0, 100 mM NaCl, 5% glycerol, and 1 mM DTT. Samples were spun down at 13,000 g for 15 minutes to remove any particulates and loaded onto a Superdex-75 column (GE Healthcare) equilibrated with 20 mM Tris pH 8.0, 100 mM NaCl, 5% glycerol, and 1-5 mM DTT.

3.4.5 Two-Hybrid Analysis

Two-hybrid analysis was carried out as described previously (Jones et al., 2010; Varrin et al., 2005). pEG-Dbf4-FL was used to express full-length wild-type Dbf4 bait, whereas pJG-Rad53 was used to express Rad53 prey (Duncker et al., 2002; Varrin et al., 2005). Additional two-hybrid bait constructs expressing various regions of Dbf4 were cloned by PCR using pPD32 as template (Dohrmann et al., 1999) and forward and reverse primers including 5' EcoRI and XhoI sites, respectively, with sequences at the 3'-end complementary to the appropriate regions of DBF4. PCR products were digested with EcoRI and XhoI and ligated into pEG-202 (Ausubel et al., 1995) digested with the same enzymes. Dbf4 variants encompassing point mutations were generated using QuikChange (Invitrogen) and verified by DNA sequencing (MOBIX, McMaster University; Robarts Research Institute).

3.4.6 Plasmid Shuffle Growth Spotting Assays

Diploid budding yeast strain DY-13 (MATa/α, his3Δ1/his3Δ1, leu2Δ0/leu2Δ0, met15Δ0/+, lys2Δ0/+, ura3Δ0/ura3Δ0, kanR::DBF4/+) was transformed with pCM190-Dbf4-FL (Duncker et al., 2002). Following sporulation, a haploid pCM190-Dbf4-

FL/ΔDBF4 strain was established. Plasmid shuffle with this strain was then carried out using CEN vector pPD32 expressing DBF4 under its natural promoter (Dohrmann et al., 1999) or one of the mutant variants described below, following standard procedures (Cormack and Castano, 2002). Bsu36I fragments from pEG-Dbf4-T171R and pEG-Dbf4ΔN (Varrin et al., 2005) were used to replace the equivalent DBF4 fragment of pPD32, to generate pPD32-T171R and pPD32-Δ135-179, respectively. Growth spotting assays were performed as described previously (Varrin et al., 2005).

3.5 Results

3.5.1 Motif N of Dbf4 is Part of a BRCT Domain

We had previously shown that motif N (residues 135-179) does not constitute an independent folding unit. Conversely, the fragment of Dbf4 encompassing residues 120-250 includes the BRCT domain and is well behaved in solution (Gabrielse et al., 2006; Matthews et al., 2009). The structure of this fragment of Dbf4, including a removable hexahistidine tag at the N-terminus of the protein, was determined by single anomalous diffraction using selenomethionine-labeled crystals, and 10 molecules were readily identified in the experimental electron density maps. About 20 residues at the C-terminus of each monomer (Ala^{230}-Asp^{250}) were disordered and could not be included in the final model. Residues Arg^{122}-Thr^{224} define a *bona fide* BRCT domain, consisting of a central four-stranded parallel β-sheet surrounded by α-helices (Figures 3.1).

In comparison with the structures of other BRCT domains determined previously, this domain of Dbf4 has a significantly longer helix α1 that projects an additional β-strand (β1') out to the solvent (Figure 3.1A and B). This additional strand interacts with the β2-strand of a neighboring molecule, defining a mixed β-sheet with four parallel strands (β4-β3-β1-β2) from one molecule and one anti-parallel (β1') strand from the adjacent one. As this intermolecular β1'-β2 interaction propagates through the crystal, it defines a right-handed helical filament (Figure 3.2). Formation of this filament was not due to the presence of the exogenous tag, because this supramolecular organization was also found
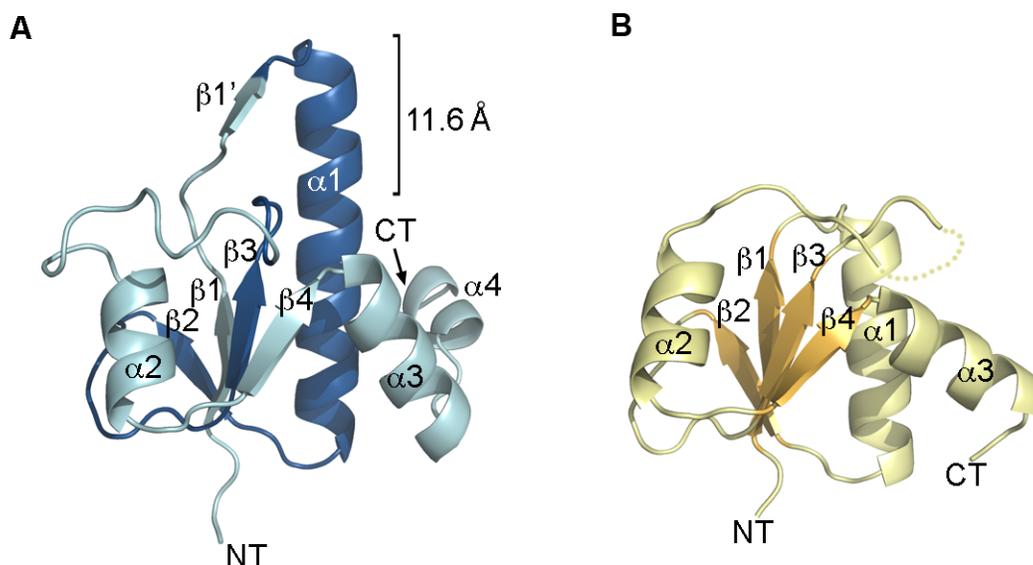
Figure 3.1. The fragment of Dbf4 including residues 120-250 folds as a BRCT domain. (A) Ribbon diagram of the structure of Dbf4 (residues 120-250) colored in blue, with motif N (residues 135-179) highlighted in a darker shade of blue. The N- and C-termini, as well as secondary structure elements are labeled. The section of helix α1 protruding from the BRCT core is indicated. **(B)** Ribbon diagram of the first BRCT from BRCA-1 (Protein Data Bank entry 1JNX, residues 1649-1737), shown as a pale yellow ribbon in the same orientation as in panel A. Secondary structure elements are shown in two shades of yellow and labeled, with the disordered β3-α2 loop depicted as a dotted line.

Table 3.1 Data collection and refinement

Dbf4 Fragment	Residues 120-250	Residues 120-250 (without His tag)	Residues 65-220 (without His tag)
Data Collection			
Space group	$P2_1$	$P2_12_12$	$P6_422$
Unit Cell			
a, b, c (Å)	90.2, 79.4, 127.4	83.7, 99.7, 127.0	83.7, 83.7, 103.7
α, β, γ (°)	90, 110.7, 90	90, 90, 90	90, 90, 120
V_M (Å^3 Da^{-1})	2.44	3.44	2.78
Solvent content (%)	49.6	64.2	55.9
Wavelength (Å)	0.9796	1.081	1.1

Continued on Page 55...

Table 3.1 continued from page 54...

Dbf4 Fragment	Residues 120-250	Residues 120-250 (without His tag)	Residues 65-220 (without His tag)
Monomers in ASU[b]	10	5	1
Resolution (Å)[a]	50-2.70 (2.75-2.70)	50-2.40 (2.44-2.40)	50 – 2.69 (2.74-2.69)
Completeness (%)[a]	99.3 (99.1)	99.8 (100)	99.7 (99.7)
R_{merge} (%)[a]	0.10 (0.79)	0.192 (0.74)	0.06 (0.42)
$I/\sigma(I)$[a]	16.9 (2.8)	8.9 (2.6)	31.0 (3.5)
Redundancy[a]	7.0 (7.2)	6.2 (5.9)	6.9 (6.2)
Reflections (total)	323,273	262,516	43,786
Reflections (unique)	46,033	42,318	6,280
Refinement			
Resolution (Å)	35.5-2.7	46.4-2.4	38.8-2.69
Reflections (R_{work}/R_{free})	44,042/2,233	42,245/2,146	6,061/300
R_{work}/R_{free} (%)	20.5/24.4	20.3/22.9	22.4/27.3
No. of atoms			
Protein	8,276	4,362	1,058
Solvent	99	230	10
RMSDs			
Bonds (Å)	0.009	0.007	0.008
Angles (°)	1.161	0.959	1.034
Ramachandran Plot			
Most favored (%)	93.6	94.2	94.0
Additionally allowed (%)	6.2	5.8	6.0

[a] Data in the highest resolution shell are shown in parentheses.
[b] Asymmetric unit

in crystals grown after removal of the hexahistidine tag (Table 3.1 and Figure 3.2B and C). However, size exclusion chromatography and analytical centrifugation experiments indicated that this fragment of Dbf4 exists predominantly as a monomer in solution
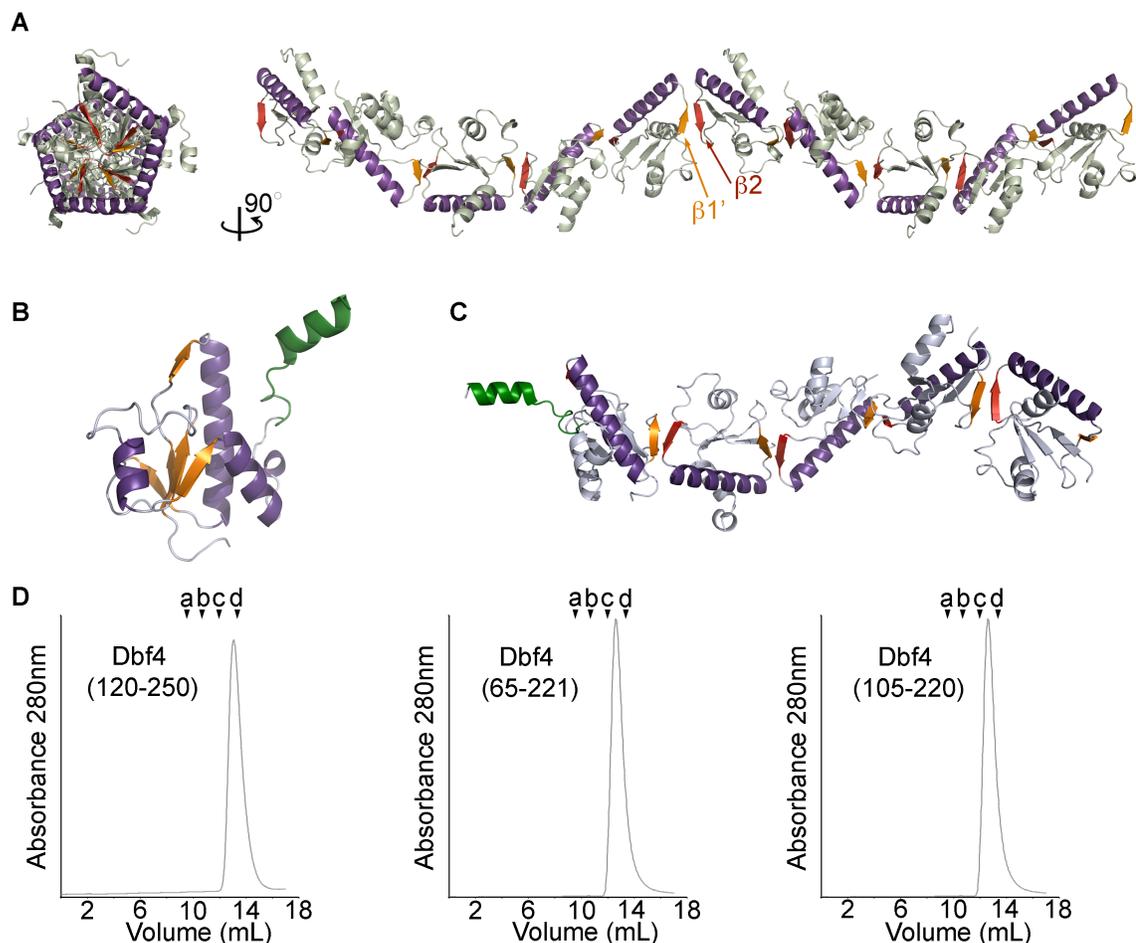
Figure 3.2. Oligomerization of the BRCT domain of Dbf4. (A) Orthogonal views of the 5-fold helical filament defined by the 10 molecules found in the asymmetric unit of the His6-BRCT crystals. Strands β1' and β2 are labeled and colored orange and red, respectively. **(B)** Ribbon diagram of the BRCT domain crystallized in the absence of the His6 tag with α-helices shown in purple, β-strands in orange, and loops in light blue. The C-terminal region of the domain (residues 225-242, colored in green) was ordered in one of the molecules of the asymmetric unit due to crystal contacts with a symmetry mate. **(C)** Organization of the asymmetric unit of the BRCT crystals. The five molecules form a 5-fold helical filament very similar to the one seen in the structure with the His6 tag. **(D)** Size exclusion chromatography profiles of three variants of Dbf4 over a Superdex-75 column (GE Healthcare). From left to right: Dbf4 (residues 120-250), Dbf4 (residues 65-221) and the sufficient domain of Dbf4 for the interaction with Rad53 (residues 105-220). The three samples elute as apparent monomers. The molecular markers used to calibrate the column are (a) albumin (67 kDa); (b) ovoalbumin (43 kDa); (c) chymotypsinogen A (25 kDa); and (d) ribonuclease A (13.7 kDa).

56

(Figure 3.2D). Therefore, although the protein-protein interactions seen in the crystals do not reflect the oligomeric state of Dbf4, they could indicate how Dbf4 interacts with other cellular partners.

3.5.2 The BRCT Domain of Dbf4 is Not Sufficient to Mediate an Interaction with Rad53

Using yeast two-hybrid analysis, we found that the fragment of Dbf4 encompassing residues 120-221 could not sustain the interaction with Rad53 (Figure 3.3A). Because it had been previously shown that a fragment of Dbf4 encompassing residues 1-296 binds Rad53 similarly to wild-type Dbf4 (Duncker et al., 2002), we generated fragments of Dbf4 encompassing additional residues either N- or C-terminal to the BRCT domain. Extending the C-terminal end to include residues 120-250 did not restore the interaction with Rad53. Conversely, extending into the N-terminal region (residues 65-220) did restore this interaction (Figure 3.3A). Fragments of Dbf4 starting at residue Leu65 where part or the entire BRCT domain (residues 65-179 or 65-109) was deleted did not support binding to Rad53 either, indicating that the BRCT domain is necessary for the interaction (Figure 3.3A).

We entertained the possibility that the N-terminal extension (residues 65-119) was merely functioning as a linker between the BRCT domain and the bulky N-terminal LexA tag used in these experiments. Therefore, we generated a fragment of Dbf4 encompassing residues 1-221 but lacking the intervening Gln66-Iln119 region (an internal deletion that changes the sequence connecting LexA to the BRCT domain while maintaining a similar spacing between the two proteins). Notably, this fragment of Dbf4 (1-221Δ66-119) did not support binding to Rad53 (Figure 3.3A), indicating that the integrity of the intervening region (residues 66-119) is also necessary for the Dbf4-Rad53 interaction. This demonstrates that both the region encompassing residues Leu65-Ile119 and the intact BRCT domain (Met120-Leu221) are necessary for the Dbf4-Rad53 interaction, whereas neither feature alone can support binding (Figure 3.3A). In turn, this
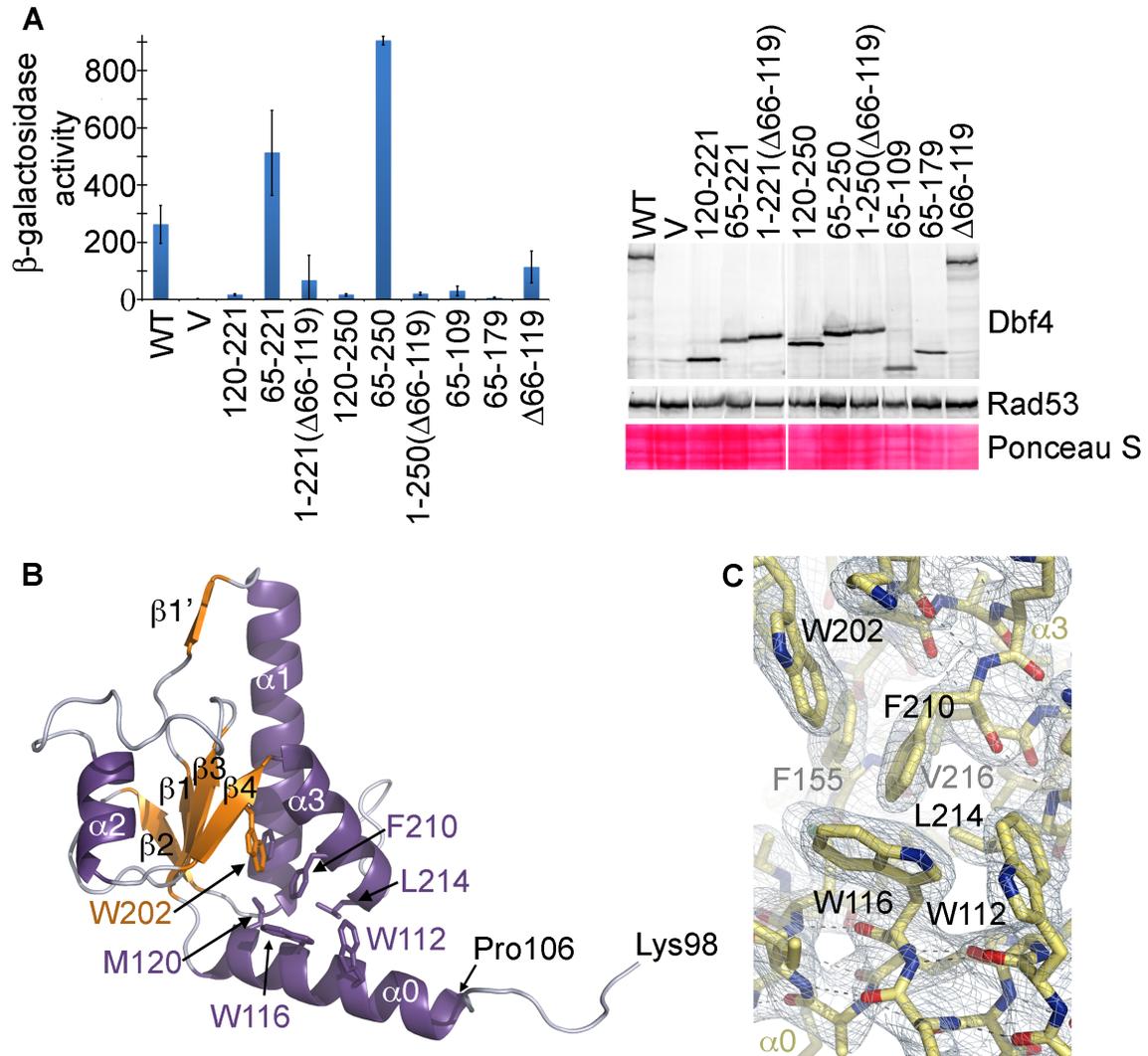
Figure 3.3 The BRCT domain and an N-terminal helix are both necessary for the interaction with Rad53. **(A)** Two-hybrid analysis using Rad53 as prey and either full-length Dbf4 (WT), empty vector (V), or regions of Dbf4 as indicated (numbering refers to amino acid positions within full-length Dbf4) as baits. The interaction is shown as β-galactosidase activity units and in each case represents the average of three independent measurements. Error bars, S.D. As a control, whole cell extracts were prepared from transformants following prey induction and analyzed by Western blot using rabbit anti-LexA antibody (Cedarlane Laboratories) to detect bait protein and a mouse monoclonal anti-HA antibody (Sigma) to detect prey protein, along with Alexa Fluor 647 goat anti-rabbit and 488 goat anti-mouse secondary antibodies (Invitrogen), respectively. Prior to *Continued on page 59…*

Figure 3.3 continued from page 58…
detection, the membrane was stained with Ponceau S to assess relative protein loading. **(B)** Ribbon diagram of the H-BRCT domain of Dbf4 colored as in Figure 3.2B. The N- and C-termini as well as secondary structure elements are labeled. The additional N-terminal helix is labeled as α0, and the residues defining the hydrophobic core are shown as sticks and labeled. **(C)** Detail of the $2F_o$-F_c electron density map around Trp116 contoured at 1σ. The H-BRCT model is shown as color-coded sticks with the residues latching the N-terminal helix (α0) to the hydrophobic core of the H-BRCT domain (β4, β1, and α3) labeled.

suggests that the fragment of Dbf4 encompassing residues Leu65-Leu221 may define a unique structure responsible for mediating the interaction with Rad53.

3.5.3 Structure of Sufficient Domain of Dbf4 for Interaction with Rad53

To elucidate whether this N-terminal extension (residues 65-119) altered the BRCT fold, we determined the crystal structure of the sufficient interaction region of Dbf4 (residues 65-221). Crystals of this fragment diffracted X-rays to 2.6 Å resolution (Table 3.1) and contained only one molecule in the asymmetric unit, confirming that the filament seen in the previous structures was indeed a crystallographic artifact.

Residues 98-221 were clearly defined in the electron density map; however, the region encompassing residues 65-97 was not. We suspected that this fragment of Dbf4 was susceptible to degradation, and indeed, mass spectrometry confirmed that the crystals contained a fragment of Dbf4 encompassing residues 84-221. As expected, residues 122-221 adopt the BRCT fold seen in the previous structures of Dbf4 (Figures 3.1 and 3.2), whereas residues 107-121 form a well defined amphipathic α-helix (α0) that is packed against the hydrophobic surface defined by the C-terminal end of α1, β1, and α3 (Figure 3.3B and C). Despite the low sequence conservation, the presence of a helix N-terminal to the BRCT domain is predicted in most Dbf4 homologs from lower eukaryotes (Figure 3.4) (data not shown). The residues upstream of this helix (residues 98-106) adopt an extended and relatively flexible conformation as judged by the quality of the electron density.
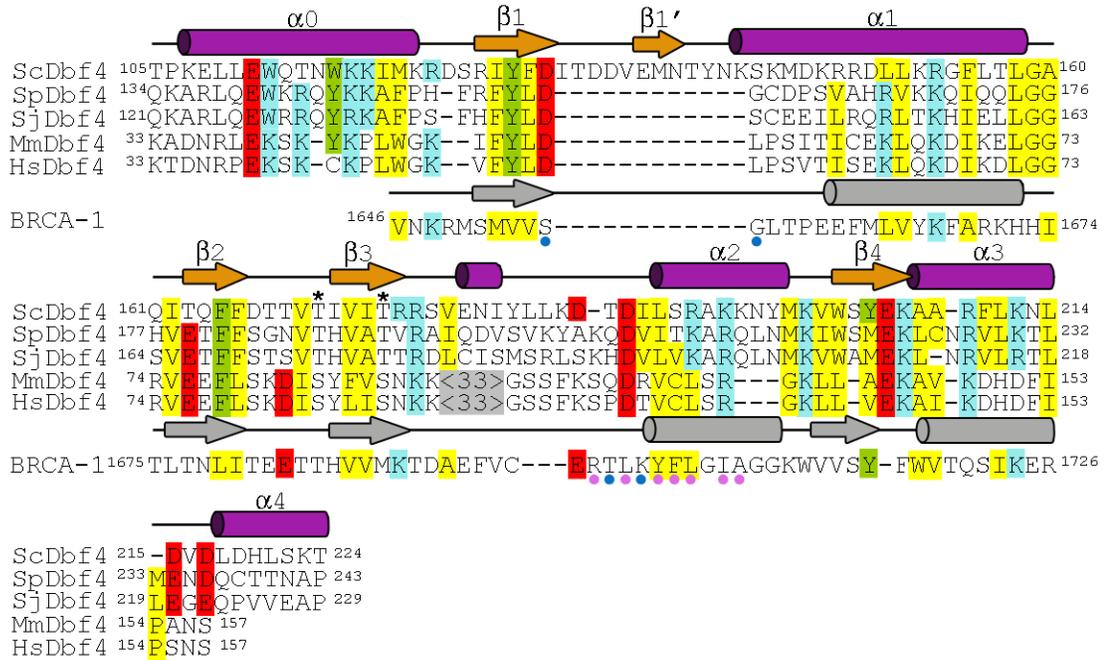
Figure 3.4 Sequence alignment of the region of Dbf4 that mediates the interaction with Rad53. Structure-based sequence alignment of the H-BRCT domains found in Dbf4 homologs from *S. cerevisiae* (Sc), *Schizosaccharomyces pombe* (Sp), *Schizosaccharomyces japonicus* (Sj), *Mus musculus* (Mm), and *Homo sapiens* (Hs) as well as the first BRCT domain found in human BRCA-1 (Williams et al., 2001). Secondary structure elements of ScDbf4 and BRCA-1 are shown as arrows (strands) and cylinders (helices) and colored as in Figure 3.2B. Conserved hydrophobic (yellow), positively charged (blue) and negatively charged (red) residues are highlighted. The positions of the conserved Thr171 and Thr175 are marked with an asterisk. Residues of BRCA-1 necessary for the recognition of a phospho-serine residue are marked with a blue dot, and those involved in stabilizing the BRCT repeat are marked with a purple dot.

To test whether helix α0 was the only relevant feature for binding Rad53 within the region encompassing residues 65-119, we repeated the yeast two-hybrid analysis with shorter fragments of Dbf4. Constructs consisting of residues 94-221 or just 105-221 were capable of binding Rad53 (Figure 3.5; note the difference in scale compared with Figure 3.3A). Therefore, we conclude that the region of Dbf4 encompassing residues 105-221 is the minimal structural unit of Dbf4 required for binding to Rad53. Interestingly, the fragment encompassing residues 105-221 interacts with Rad53 more efficiently than Dbf4

itself, suggesting that the full-length protein has additional mechanisms that down-regulate the interaction between the two proteins. This inhibitory effect may arise from other regions of the protein occluding the interaction interface or the competition between Rad53 and other cellular Dbf4-binding proteins. Indeed, it has been shown that Cdc5 interacts with Dbf4 through residues 83-93, and hence, it is conceivable that binding to endogenous Cdc5 restricts binding to Rad53 in our yeast two-hybrid assay (Chen and Weinreich, 2010).
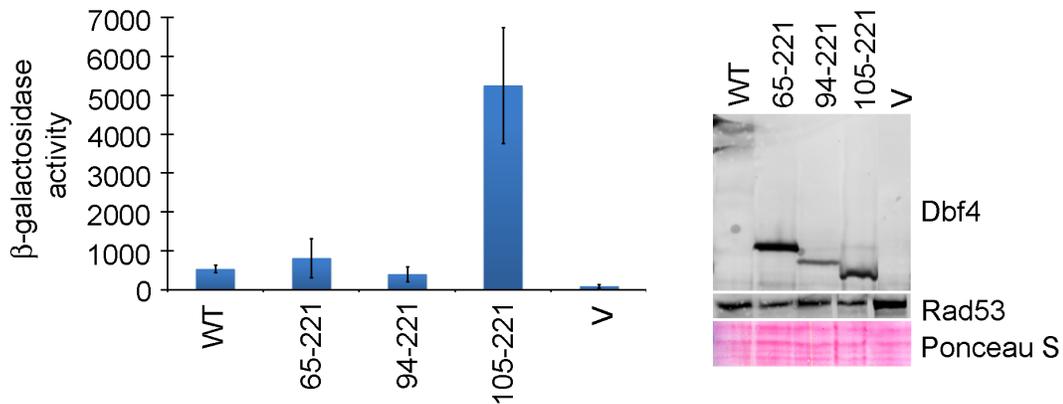
Figure 3.5 The fragment of Dbf4 including residues 105-221 is the minimal domain of Dbf4 that mediates the interaction with Rad53. Two-hybrid analysis using Rad53 as prey and either full-length Dbf4 (WT) or fragments of Dbf4 encompassing residues 65-221, 94-221, and 105-221 as baits. Data is presented as in Figure 3.3A. See also Table 3.2.

3.5.4 The H-BRCT Fold is Necessary for the Dbf4-Rad53 Interaction

Arguably, helix α0 could be an independent motif adjacent to the BRCT domain whose relative orientation had been determined by crystal packing. However, the side chains of Trp116 and Met120 and, to a lesser extent, Trp112 have extensive van der Waals and aromatic interactions with the hydrophobic core defined by residues Phe155, Trp202, Phe210, and Leu214 (Figure 3.3B and C). Trp116 and Met120 are not strictly conserved in other Dbf4 homologs, but aromatic or bulky hydrophobic residues are found

at these positions (Figure 3.4), suggesting a common mechanism to occlude the hydrophobic surface defined by Phe155, Trp202, Phe210, and Leu214 and hence supporting the role of this helix in maintaining the structural integrity of the domain. Mutation of Trp116 and Met120 abolishes the interaction between Dbf4 and Rad53 (Figure 3.6), indicating that the integrity of the hydrophobic core of the domain is critical for the interaction. These results confirm that the helix N-terminal to the BRCT domain is an inherent part of the domain of Dbf4 that we herein refer to as H-BRCT. Recent work has shown that at least another BRCT domain also requires additional structural elements to perform its function (Kobayashi et al., 2009). However, the H-BRCT domain of Dbf4 constitutes the first example of a BRCT domain where this additional structural motif is an integral part of the folding unit.

Interestingly, the hydrophobic and polar faces of the α0 helix contribute differently to the interaction with Rad53. Mutation of Trp112 and Leu109, which contribute only minimally to the hydrophobic core of the H-BRCT domain, also disrupts the Dbf4-Rad53 interaction. Conversely, mutation of the polar face of the helix (K121D/R122E or E111R/K118D) did not affect the interaction (Figure 3.6). Collectively, these results indicate that the concave surface defined by α0, β4 and α3 is important for the interaction with Rad53.

3.5.5 Point Mutations that Destabilize the H-BRCT Fold Disrupt the Dbf4-Rad53 Interaction

Internal deletions disrupting the H-BRCT fold disrupt the interaction with Rad53 (Figure 3.3 and Table 3.2), demonstrating that the tertiary structure of the H-BRCT domain is important for this function. Therefore, we subsequently investigated whether previously described mutations impairing Dbf4 function could also have folding defects. Three highly conserved residues within the H-BRCT domain (Gly159, Gly/Ala160, and Trp202) are important for Dbf4 function. Mutation of any of these residues causes
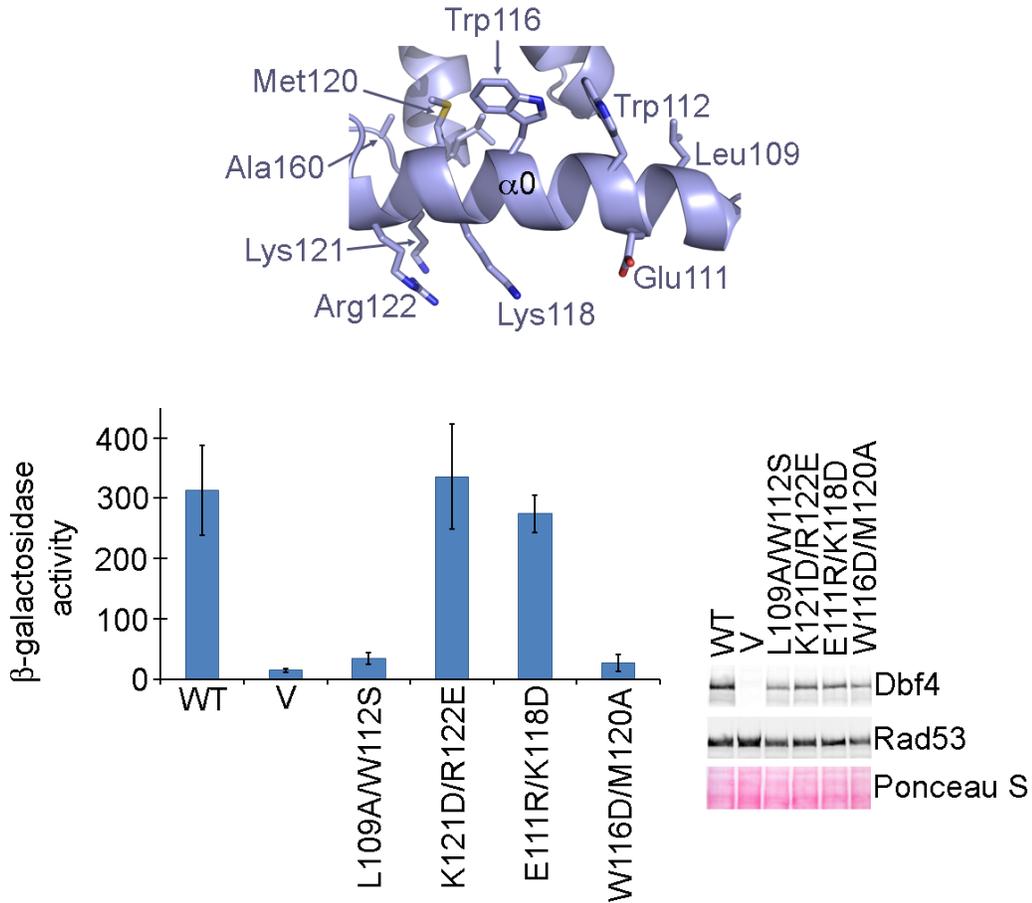
Figure 3.6 Mutation of the residues anchoring helix α0 to the core of the BRCT domain abolishes the interaction between Dbf4 and Rad53. Two-hybrid analysis carried out using full-length Rad53 as prey and either wild-type Dbf4 (WT), empty vector (V), Dbf4-L109A/W112S (L109A/W112S), Dbf4-K112D/R122E (K112D/R122E), Dbf4-E111R/K118D (E111R/K118D), or Dbf4-W116D/M120A (W116D/M120A) as baits. A detail of helix α0 with the side chains of the residues assayed shown as sticks is shown at the top for reference. Data are presented as in Figure 3.3A. See also Table 3.2. Error bars, S.D.

increased sensitivity to both DNA-damaging agents and hydroxyurea (Gabrielse et al., 2006) and mutation of the Leu158-Gly159 doublet destroys the interaction between Dbf4 and origin DNA in a yeast one-hybrid assay (Ogino et al., 2001). We found that the

Dbf4-W202E and Dbf4-G159L/A160L variants also lost their ability to interact with Rad53 (Figure 3.7A and Table 3.2). The side chain of Trp202 lies at the hydrophobic pocket defined by the central β-sheet and the surrounding α0, α1 and α3 helices.

Table 3.2 Effect of point mutations on the Dbf4-Rad53 interaction in a yeast two-hybrid

Mutation	Residual interaction (%)	S.D.
Wild type	100	--
Δ132-144	28	10
Δ168-178	4	3
W116D/M120A	7	1.5
L109A/W112S	11	4
E111R/K118D	90	15
K121D/R122E	110	28
GA159LL	1	2
W202E	1	2
T157A	80	9
T157R	66	11
T163A	52	9
T163R	78	14
T171A	52	14
T171R	47	10
T171E	56	12
T175A	111	18
T224A	73	16
T224R	63	18
T95R	98	24
T105R	123	16
T114A	97	50
T171A/T175A	1	1
T168A/T169A[a]	93	7

[a] In contrast to all other Dbf4 variants that were assayed for their interaction with full-length Rad53, Dbf4-T168A/T169A was assayed against the FHA1 domain of Rad53 [Rad53(4-165)].
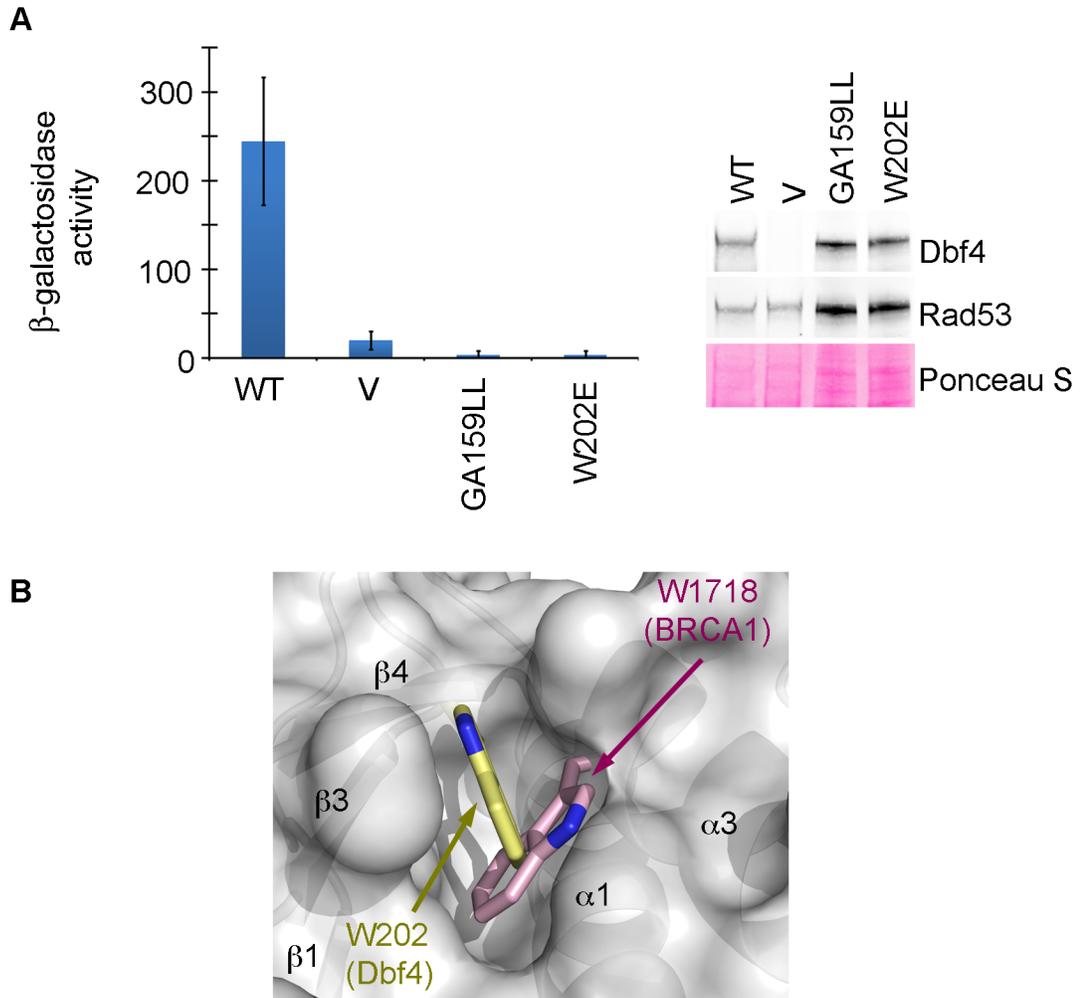
Figure 3.7 Dbf4 mutants that are sensitive to genotoxic agents fail to interact with Rad53. **(A)** Two-hybrid analysis was carried out using Rad53 as prey and wild-type Dbf4 (WT), Dbf4-G159L/A160L (GA159LL), or Dbf4-W202E (W202E) as baits with data presented as in Figure 3.3A. **(B)** Superimposition of Trp1718 from BRCA-1 (pink color-coded sticks) onto the structure of Dbf4-N. Dbf4-N is shown as a semitransparent white surface with the secondary structure elements shown as a ribbon diagram. Trp202 from Dbf4 is shown as yellow color-coded sticks. See also Table 3.2. Error bars, S.D.

Although this residue is not conserved in other BRCT domains, a structurally equivalent tryptophan is found at the N- termini of helix α3 in most BRCT domains (Figure 3.7B), reinforcing the idea that a hydrophobic, bulky residue at this position is necessary to stabilize the hydrophobic core of the H-BRCT domain. Therefore, the defects associated

with mutating Trp202 are probably due to the structural destabilization of the H-BRCT domain.

The Gly159-Ala160 doublet is located in the short loop connecting α1 and β2 that is conspicuously exposed to solvent. This loop reverses the orientation of the polypeptide chain, and hence, the substitution of these two small and flexible residues to bulky and hydrophobic amino acids probably reduces the plasticity of the α1-β2 loop, potentially destabilizing the central β-sheet (Figure 3.1A). Indeed the ϕ/ψ angles adopted by Gly159 to make this turn would not be favored for any residue other than glycine, in turn explaining the strict conservation of these two residues among BRCT domains. The H-BRCT-G159L/A160L variant was also insoluble when over-expressed in bacteria using conditions identical to those used to produce the H-BRCT domain (Figure 3.8), reinforcing the idea that these point mutations destabilize the H-BRCT fold. Collectively, these results strongly suggest that the integrity of the H-BRCT domain is critical for the interaction between Dbf4 and Rad53.
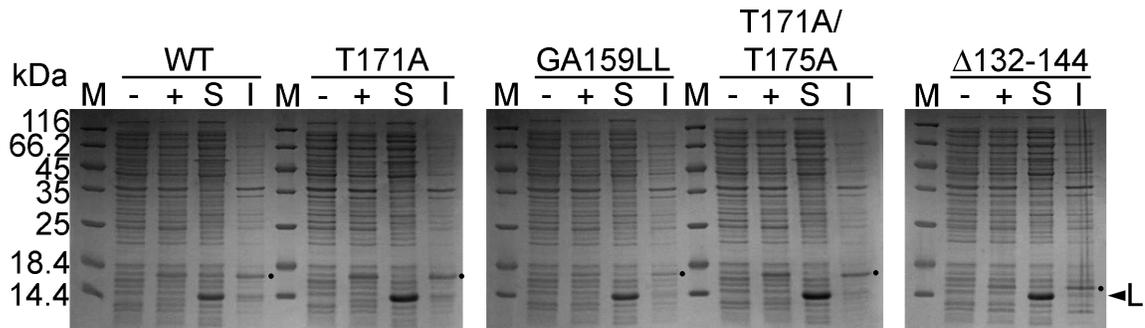
Figure 3.8. Solubility assays of Dbf4 variants encompassing H-BRCT destabilizing point mutations. 15% SDS-polyacrylamide gels showing the levels of expression of H-BRCT (residues 105-220) variants as indicated. From left to right, molecular weight markers (M), uninduced cell lysate (-), induced cell lysate (+), soluble fraction of the cell lysate (S), and insoluble fraction of the cell lysate (I) is shown for every Dbf4 variant. The band corresponding to Dbf4 is marked with a black dot and that corresponding to lysozyme is marked with an arrowhead and labeled "L".

3.5.6 Thr171 Aids the Interaction Between Dbf4 and Rad53

Mutation of Arg70 within the FHA1 domain of Rad53, a residue important for the recognition of phospho-threonine (pThr) targets, weakens the interaction with Dbf4, and hence, it had been previously suggested that Rad53 would recognize a phospho-epitope in Dbf4 (Duncker et al., 2002). We confirmed the importance of Arg70 for this interaction using full-length protein (data not shown). Therefore, we subsequently scanned the sequence of the H-BRCT domains from various Dbf4 homologs to find threonines that could potentially mediate the interaction with Rad53. Only two threonine residues, Thr171 and Thr175, are conserved in the H-BRCT domain of Dbf4 (marked with asterisks in Figure 3.4). Thr175 is buried in the hydrophobic core of the H-BRCT domain and is involved in a hydrogen bond network that stabilizes the β3-α2 loop (Figure 3.9A). This loop is highly variable among BRCT domains, suggesting that Thr175 may have a conserved structural role in Dbf4 by stabilizing the β3-α2 loop (Figure 3.9B). Due to its location in the domain, we did not expect Thr175 to be the phosphorylated site recognized by Rad53. Accordingly, mutation of Thr175 to alanine did not affect the interaction between Dbf4 and Rad53 (Table 3.2).

Conversely, Thr171 is relatively exposed (Figure 3.9A), and the equivalent serine in higher eukaryotes is part of a Chk1 kinase phosphorylation site (Kim et al., 2003). Mutation of Thr171 to either alanine or arginine significantly reduced binding to Rad53 (Figure 3.9C and Table 3.2). Importantly, the Dbf4-T171R variant also showed increased sensitivity to genotoxic agents (Figure 3.10), reinforcing the idea that this residue is important for the Rad53-dependent checkpoint response. However, the H-BRCT-T171A variant was also insoluble when overproduced in bacteria using experimental conditions identical to those used to produce the H-BRCT domain (Figure 3.8), suggesting that this mutation also destabilizes the H-BRCT fold. The side chain of Thr171 is hydrogen-bonded to the main chain of residue Arg125, and this interaction indirectly orients the α0 helix. Therefore, it is not surprising that mutation of Thr171 destabilizes the H-BRCT fold and weakens the interaction between Dbf4 and Rad53.
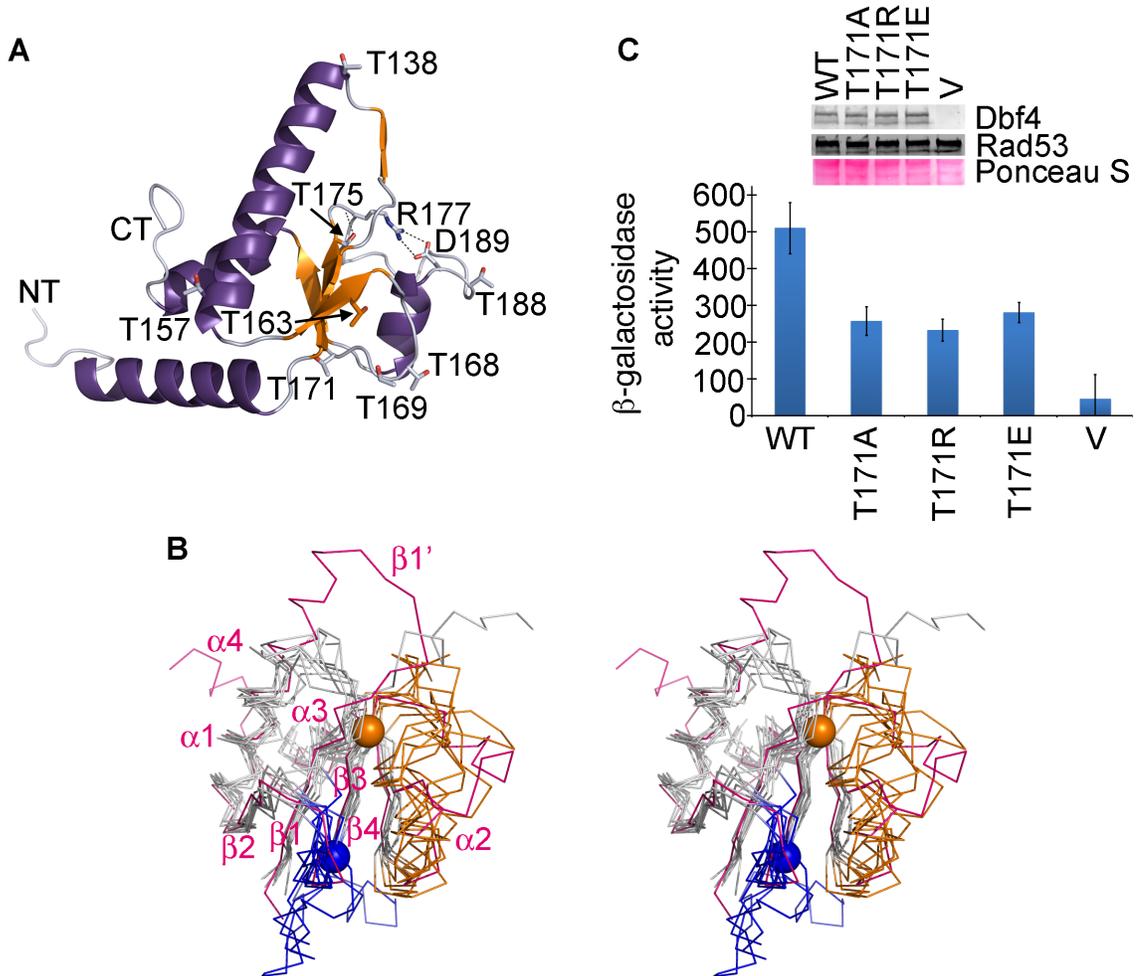
Figure 3.9 Role of Thr171 and Thr175 in mediating the Dbf4-Rad53 interaction. (A) Ribbon diagram of the H-BRCT domain colored as in Figure 3.2B. The side chains of the threonine residues found in the domain are shown as sticks. The hydrogen bond network restraining the β3-α2 loop is shown as black dashed lines. **(B)** Stereo view of the BRCT domain of Dbf4 shown as a pink Cα-tracing with the Cα atoms of Thr171 and Thr175 shown as spheres (blue and orange, respectively) and secondary structure motifs labeled. BRCT domains from proteins containing single or multiple BRCT repeats (PDB IDs: 1JNX, 1CDZ, 1GZH, 2NTE, 3L3E, and 3L46) reveals that the β2-β3 (shown in blue) and the β3-α2 (shown in orange) loops are the most variable regions of the domain. **(C)** Two-hybrid analysis carried out using full-length Rad53 as prey and either full-length Dbf4 or Dbf4 variants carrying point mutations at Thr171 as baits. Data are presented as in Figure 3.3A. Error bars, S.D.

Should Thr171 be the pThr recognized by Rad53, mutation of Thr171 to a phosphomimetic residue would maintain or even strengthen the interaction between the two proteins. However, the Dbf4-T171E variant also showed reduced binding affinity for Rad53 (Figure 3.9C and Table 3.2). This result reinforces the idea that Thr171 has a structural role rather than being the phospho-epitope recognized by Rad53. We cannot exclude the possibility that the T171E mutation is not a good mimic for a pThr. Indeed, in studies using pThr-containing peptides that recognize FHA1 domains in the absence of any structural context, aspartate cannot efficiently mimic a phosphorylated threonine (Durocher et al., 1999). However, we have demonstrated that the structural integrity of the H-BRCT domain of Dbf4, and hence, the structural context of the putative phospho-epitope, must be important. Therefore, we favor the idea that Thr171 also has a structural role, because the double mutation of Thr171 and Thr175 to alanine completely disrupts
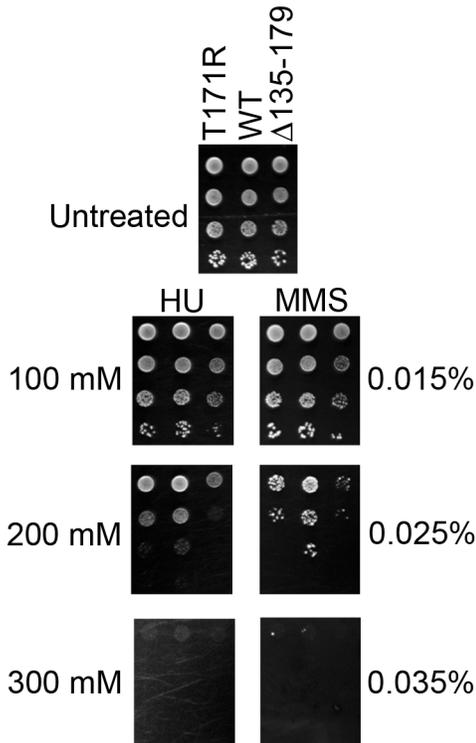
Figure 3.10. The Dbf4-T171R Mutant is susceptible to HU and MMS. ΔDBF4 cells transformed with CEN vectors (one copy per cell) expressing either wild-type Dbf4 (WT), Dbf4-T171R, or Dbf4 Δ135-179 were spotted in 10-fold dilution series onto SC-Leu (synthetic complete medium, lacking leucine in order to select for plasmid maintenance) plates containing the indicated concentrations of hydroxyurea (HU) or the alkylating agent methyl methanesulfonate (MMS). Each plate was incubated at 30°C for 3 days.

the interaction between Dbf4 and Rad53 (Table 3.2), indicating the synergy between these two destabilizing mutations and restating the importance of the structural integrity of the H-BRCT fold.

Other Thr residues within this domain of Dbf4 have been proposed as damage-induced phosphorylation sites (Duch et al., 2011; Lopez-Mosqueda et al., 2010; Zegerman and Diffley, 2010); however, all non-conserved threonine residues within the H-BRCT domain except Thr163 could be replaced without affecting the Dbf4-Rad53 interaction. The Dbf4-T163A variant has reduced affinity for Rad53 (Table 3.2). Interestingly, Thr163 resides on the same surface of the H-BRCT domain as Thr171, and it could thus contribute to delineating the interaction interface (Figure 3.11).
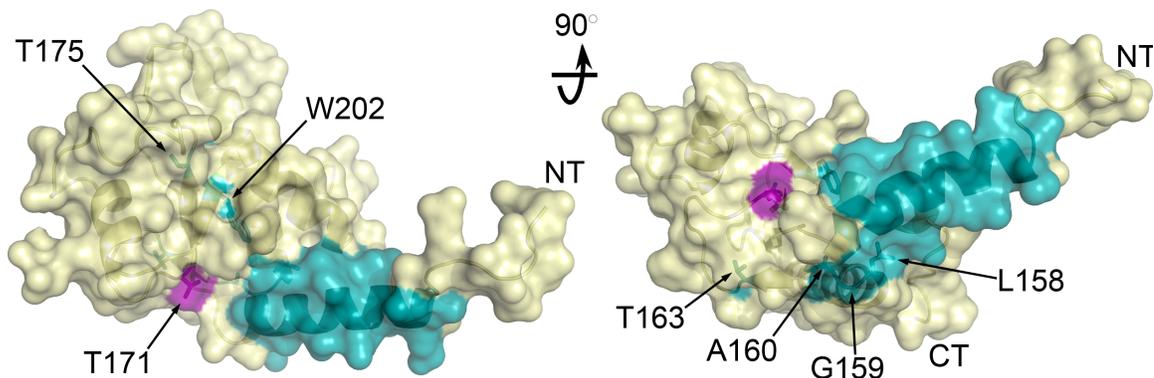
Figure 3.11 Surface mapping of the Dbf4 residues important for the interaction with Rad53. Orthogonal views of the H-BRCT domain shown as a yellow ribbon inside a semitransparent surface. Residues that are important for the interaction with Rad53 are colored in cyan except for Thr171, which is shown in purple for clarity.

3.6 Discussion

We have shown that motif N of Dbf4 is part of a unique structure built upon a BRCT domain that includes an additional α-helix strictly required for the function of the

domain. We refer to this domain as H-BRCT to signify the location and the nature of the additional structural motif. The H-BRCT domain is necessary and sufficient to mediate the interaction between Dbf4 and Rad53. Interestingly, other proteins containing single BRCT domains also include additional structural elements surrounding the domain that aid their function, suggesting that the BRCT fold is a core whereupon complex structures with different specificities can be built. Tandem BRCT repeats, known to function as a single unit, could be seen as one of the types of structures that can be built upon the BRCT fold. Likewise, single BRCT domains should be considered as part of bigger functional units where the additional BRCT domain found in tandem repeats has been replaced by either a different domain, as in the case of Nbs1 or Chs5 (Lloyd et al., 2009; Martin-Garcia et al., 2011; Williams et al., 2009), or additional structural elements that complement the BRCT fold. Collectively, our results and those from others suggest that other single BRCT domains may also require additional structural motifs, thereby explaining why single BRCT domains do not share a common mechanism of interaction with their binding partners. In turn, this would also explain why BRCT domains, which are commonly found in DNA replication and repair proteins, are not particularly stable and do not show high sequence conservation (Williams et al., 2001). The most variable regions within the BRCT domain are the loops connecting β2-β3 and β3-α2, which provide the means to gain specificity for different binding partners (Figure 3.9B). Here, we propose that additional secondary structural elements at the N- or C-terminus of the domain can also increase target specificity and present the first example of a BRCT where the additional structural elements have become an integral part of the domain.

BRCT and FHA domains are protein folds with important functions in DNA repair and checkpoint activation. Their mechanisms of pThr recognition have been studied in depth (Campbell et al., 2010; Durocher et al., 2000; Glover et al., 2004). However, recent studies have identified novel mechanisms of interaction between FHA and BRCT domains (Lloyd et al., 2009; Williams et al., 2009). Therefore, the FHA1 domain recognizing a single pThr in the absence of any structural context may no longer be the paradigm that best describes FHA interactions.

Our results reveal that the structural integrity of the H-BRCT fold is more relevant than the presence of a phospho-epitope for the interaction between Dbf4 and Rad53. This finding was somewhat unexpected because it had been proposed that Rad53 would recognize a phospho-epitope in Dbf4 (Duncker et al., 2002). Furthermore, previous structural work had shown that Rad53 recognizes phospho-epitopes in the absence of a structural context (Durocher et al., 2000). However, an elegant study from the Smerdon laboratory has shown that the FHA domain of *Mycobacterium tuberculosis* Rv1827 recognizes several binding partners in a phosphorylation-independent manner but involving residues integral to the canonical pThr-binding surface (Nott et al., 2009). Therefore, it is conceivable that the FHA1 domain of Rad53 does not recognize a phospho-epitope in the H-BRCT domain of Dbf4, although mutation of the conserved Arg70 in the FHA1 domain disrupts the interaction with Dbf4 (Duncker et al., 2002).

The structurally unstable Dbf4 Δ135-179, Dbf4-G159L/A160L, Dbf4-W202A, and Dbf4-T171R variants have both a weakened interaction with Rad53 and an increased sensitivity to DNA-damaging agents and hydroxyurea (this work) (Gabrielse et al., 2006). Interestingly, Gly159, Ala160, Thr163, Thr171, and Trp202 are all in close proximity to helix α0, potentially demarcating the surface of Dbf4 that interacts with Rad53.

The BRCT domain of Dbf4 could be replaced by the BRCT domain found in Rev1 without loss of viability during genotoxic stress, although several other BRCT domains were unable to do so (Harkins et al., 2009). It was suggested that Rev1 and Dbf4 BRCT domains might share a common interaction partner. However, the ability of this Rev1-Dbf4 chimera protein to interact with Rad53 and suppress late origin firing was not examined in that study, and therefore, the need for an intact H-BRCT domain did not become apparent.

Collectively, our work unveils a phosphorylation-independent interaction with the FHA1 domain of Rad53 that relies on the unique H-BRCT domain found at the N-terminus of Dbf4. In a broader context, the identification of the H-BRCT domain suggests that other single BRCT domains may also require additional motifs for their

function.   The variable nature of these additional motifs would explain why the mechanisms of interaction of single BRCT domains have remained elusive, whereas those of tandem BRCT domains are well characterized.

3.7 Acknowledgments

Chapter 4

In vitro Analysis of the Interaction Between the Rad53 FHA1 and Dbf4 H-BRCT Domains Reveals an Alternative FHA Binding Surface

4.1 Author's Preface

This chapter presents the first *in vitro* analysis of the interaction between *S. cerevisiae* Dbf4 and Rad53, with the goal of investigating the necessary components of both proteins for molecular recognition. The role of the α0 helix from Dbf4$^{\text{H-BRCT}}$ and the residue Arg70 from Rad53$^{\text{FHA1}}$, both necessary for the interaction *in vivo*, were assessed in the context of the isolated domains. The primary focus is on the FHA1 domain of Rad53 to complement the work already accomplished on the characterization of the H-BRCT domain of Dbf4 (chapter 3). The NMR experiments were performed in collaboration with Rajeevan Selvaratnam and Dr. Giuseppe Melacini (McMaster University). I conducted the experiments, interpreted the data, prepared the figures, and wrote the chapter.

4.2 Abstract

FHA (*Fork*H*ead *A*ssociated) and BRCT (*BR*CA-1 *C*-*T*erminal) domains are highly prevalent in proteins that respond to stress and DNA damage, and are best understood for their ability to bind phospho-epitopes. The interaction between Dbf4 and Rad53 during the replication checkpoint in budding yeast involves both types of domains, making this an exceptional case study for probing the molecular basis of recognition between FHA and BRCT domains. Here we report the first *in vitro* analysis of the association between the isolated FHA1 domain from Rad53 and H-BRCT domain from Dbf4 using NMR spectroscopy and chemical cross-linking. The interaction requires a noncanonical interaction surface on both modules. The Dbf4 H-BRCT domain has a pre-formed surface created partly by a unique N-terminal helix, which binds to the FHA1 domain of Rad53 directly and independently of phosphorylation. We also show that the Arg70 residue in FHA1 critical for phosphate recognition is not required for binding to the H-BRCT domain of Dbf4. Combining these results with previous yeast two-hybrid assays, we propose a dual binding site mechanism for the interaction between full-length Dbf4 and Rad53.

4.3 Introduction

Dbf4 is an essential protein that is conserved amongst all eukaryotes. It is the regulatory subunit of the Cdc7 kinase, which is required in order to initiate DNA replication by activating origins of replication (Bousset and Diffley, 1998; Donaldson et al., 1998). Dbf4-Cdc7 also plays a role in the replication checkpoint, and in budding yeast this is partly through an N-terminal BRCT domain in Dbf4 that mediates an interaction with the FHA1 domain from the checkpoint kinase Rad53 (Duncker et al., 2002; Gabrielse et al., 2006; Matthews et al., 2012; Varrin et al., 2005). The replication checkpoint preserves the genetic integrity of the cell by stabilizing stalled forks, as well as boosting DNA repair enzyme levels and pausing the cell cycle (Labib and De Piccoli, 2011). Rad53 plays integral roles in all of these facets of the checkpoint (Branzei and

Foiani, 2006), and utilizes its N-terminal FHA domain (FHA1) to mediate a diversity of interactions including with Dbf4 (Duncker et al., 2002). The association with Dbf4 is specifically linked to pausing the cell cycle, as this allows the kinase domain of Rad53 to phosphorylate Dbf4 and inhibit the activity of its associated Cdc7 kinase subunit preventing late origins from firing (Duch et al., 2011; Kihara et al., 2000; Lopez-Mosqueda et al., 2010; Zegerman and Diffley, 2010). We previously showed that the interaction between the Rad53 FHA1 and Dbf4 H-BRCT domains was independent of phosphorylation (Matthews et al., 2012). This makes the Dbf4-Rad53 complex an interesting case study as phosphorylation-independent interactions mediated by FHA or BRCT domains are poorly understood.

The BRCT domain is found in proteins from all kingdoms of life and has a profound influence on human health. Proteins involved in responding to DNA damage are enriched in BRCT domains, and mutations within these modules can lead to cancer (Callebaut and Mornon, 1997). It was originally discovered as a tandem pair in the C-terminus of the protein _Br_east _C_ancer _A_ssociated-1 (BRCA-1), from which the domain gets its name. Mutations in the tandem BRCT domains in BRCA-1 can cause hereditary breast and ovarian cancer syndrome (Glover, 2006). This carries up to an 80% - 90% risk of developing breast cancer and 40% - 65% increase for ovarian cancer (Ford et al., 1994). These mutations often cause folding defects that interfere with the function of BRCA-1 in the homologous recombination pathway used by eukaryotic cells to deal with double stranded DNA breaks that can be generated behind replication forks (Powell and Kachnic, 2003).

In order to better understand the inner workings of a BRCT domain, structures from a variety of proteins, including Dbf4 (Matthews et al., 2012) and BRCA-1 (Williams et al., 2001), have been solved. However, it has been difficult to extrapolate results from one protein to another due to the surprisingly low sequence identity between BRCT domains. There is not one single invariant residue, and structural studies still turn up BRCT domains unexpectedly (Rappas et al., 2011). Despite poor sequence conservation,

the tertiary structure of a BRCT domain is well defined with a hydrophobic core consisting of a four-stranded parallel β-sheet with three helices packed against the sheet to protect it from the solvent.

The BRCT domain is found in prokaryotic NAD+-dependent DNA ligases where it functions in binding to DNA (Bork et al., 1997; Sheng et al., 2011). In eukaryotes, tandem BRCT domains arose through a duplication event, which recognize phospho-epitopes on protein interaction partners rather than the phosphate backbone of DNA (Sheng et al., 2011). Tandem BRCT domains coalesce using their peripheral helices: α2 of the N-terminal domain forms a three helical bundle with α1 and α3 from the C-terminal domain. Importantly, this three helical bundle provides a deep groove important for recognizing the third residue C-terminal to the phosphorylation site (see Figure 1.6 in Chapter 1). Thus, the N-terminal BRCT domain contains a phosphate binding pocket, while the three helical bundle interface between the two modules provides specificity (Williams et al., 2004). The function of tandem BRCT domains is well characterized, owing in part to their importance in DNA damage checkpoints and human health. Despite the experimental bias towards tandem domains, single BRCT domains, such as in Dbf4, are more prevalent in nature. Unlike the single domain present in prokaryotic ligases that bind to DNA, single domains in eukaryotes can engage in a myriad of molecular interactions with both DNA and proteins.

While the BRCT domain from Dbf4 resembles the canonical BRCT fold, it has an additional N-terminal helix (α0) that is an integral part of the domain in the crystal structure and is absolutely required for the interaction with Rad53 *in vivo* (Matthews et al., 2012). As Dbf4 proteins in higher eukaryotes have BRCT domains that lack this feature, we have termed the fungal domain H-BRCT (Helix-BRCT) to distinguish it. This structure expanded the knowledge on how additional α-helices impact BRCT domain function. For instance, Rfc1 has an N-terminal α-helix connected to the BRCT domain by a flexible linker that is required for robust DNA binding (Kobayashi et al., 2009). Additionally, the Rtt107 tandem BRCT domain also has an additional N-terminal

helix that is required for structural stability and forms a three helical bundle, not unlike the one found at the interface between the two BRCT modules (Li et al., 2012).

The H-BRCT domain in Dbf4 is one of many partners known to bind to the FHA1 domain of Rad53. This domain is an eleven-stranded β-sandwich that can recognize phospho-threonine peptides using the loops on one of the apical surfaces of the domain (Durocher and Jackson, 2002). Due to the reliance on phosphorylation, most FHA domains have been studied in complex with synthesized peptides. However, interactions with full-length partners likely involve much larger or even alternative surfaces than are utilized by peptides (Li et al., 2004; Tong et al., 2010). For example, the tandem BRCT domain from BRCA-1 interacts with the FHA domain from the checkpoint kinase Chk2, which is the human homolog of Rad53 (Li et al., 2002). This interaction not only involves the phospho-threonine binding site, but also a hydrophobic patch on the face of one of the FHA β-sheets. Mutation of the β-sheet prevents the interaction and may lead to Li Fraumeni syndrome, which predisposes patients to multi-organ cancers (Li et al., 2002). However, this hydrophobic patch is not conserved in FHA1 from Rad53 making extrapolation of these results to the Rad53-Dbf4 interaction difficult.

Previous experiments have demonstrated that the phospho-threonine binding site in FHA1 is required for the Dbf4 interaction as mutating a conserved arginine that hydrogen bonds to the phosphate (Arg70Ala) abrogates this interaction in a yeast two-hybrid assay (Duncker et al., 2002). However, an exhaustive search for a possible phospho-threonine in the H-BRCT domain of Dbf4 resulted in only one candidate, Thr171, which is the only surface exposed conserved threonine and is also in close proximity to α0 (Matthews et al., 2012). It is notable that the phosphate binding pocket and surrounding loops in other FHA domains have been known (or suspected) to accommodate α-helices, suggesting that α0 could be required because it directly mediates the interaction in addition to a phospho-threonine (Lee et al., 2004; Li et al., 2004). However, Thr171 was not identified as a phosphorylated residue in two independent

studies that mapped phospho-epitopes in Dbf4 after replication stress through mass spectrometry (Duch et al., 2011; Lopez-Mosqueda et al., 2010).

An additional layer of regulation has recently emerged for FHA domains, which involves nonphosphorylated binding partners competing with phospho-epitopes for the same interface. Such a case has been described in *Mycobacterium tuberculosis*, where a FHA domain in the protein Rv1827 can engage a phospho-epitope intramolecularly (Nott et al., 2009). Nonphosphorylated citric acid cycle enzymes can then compete for binding to the FHA domain, although the molecular details of this remain elusive. The coiled-coil region in human kinesin KIF1A is also suspected of interacting intramolecularly with the FHA domain contained in this protein independently of phosphorylation (Lee et al., 2004).

Here we analyze the Dbf4-Rad53 interaction using the purified H-BRCT and FHA1 domains to determine the required molecular components. The α0 helix from the H-BRCT domain is necessary for the interaction and forms a stable interaction surface, while point mutations within this helix known to destroy the interaction are found to act by unlatching this element from the rest of the BRCT domain. To assess the role of Arg70 in the FHA1 domain, the interaction between the Rad53 FHA1 and Dbf4 H-BRCT domains was monitored *in vitro* using NMR and chemical cross-linking. A phospho-peptide cannot compete with the H-BRCT domain for Rad53, and mutating Arg70 does not disrupt the interaction. This suggests that Arg70 is required only for binding full-length Dbf4 and likely contacts a region outside of the H-BRCT domain, which independently binds a noncanonical surface on the FHA1 domain. Thus, a dual binding site model is presented similar to the case of the Chk2−BRCA-1 complex and the functional implications of this for the replication checkpoint are discussed.

4.4 Experimental Procedures

4.4.1 Cloning, Expression and Purification of SUMO Tagged Dbf4$^{\text{H-BRCT}}$ Constructs

A gene fragment encoding a HIS$_6$ tag fused to SUMO followed by a $\underline{T}$obacco $\underline{E}$tch $\underline{V}$irus (TEV) protease cleavage site was amplified using a PCR with pTriEX serving as the template, which was a gift from the Melacini lab. The PCR primers contained an *Nco*I and *Nde*I site at the 5' ends and the resulting PCR product was subcloned into pJET1.2 (Fermentas) and confirmed through DNA sequencing (MOBIX, McMaster University). After digesting with *Nco*I and *Nde*I, the fragment was ligated into a pET15b vector (Novagen) digested by the same enzymes to create the pET-SUMO vector. A gene encoding the H-BRCT domain of Dbf4 (amino acids 105-220 – herein referred to as Dbf4$^{\text{H-BRCT}}$) was codon optimized to improve expression levels and synthesized with flanking *Nde*I and *Bam*HI restriction sites by GeneArt® (Invitrogen). The Trp116Asp/Met120Ala, Leu109Ala/Trp112Ser and Thr171Ala Dbf4$^{\text{H-BRCT}}$ variants were generated as described previously (Matthews et al., 2012). The PCR products were ligated into pJET1.2 (Fermentas) and sequenced (MOBIX, McMaster University). The codon optimized construct and the three mutants were then all digested with *Nde*I and *Bam*HI and ligated into the pET-SUMO vector to generate plasmids pAG8586 (WT), pAG8587 (Trp116Asp/Met120Ala), pAG8639 (Leu109Ala/Trp112Ser) and pAG8611 (Thr171Ala). These HIS$_6$-SUMO tagged Dbf4$^{\text{H-BRCT}}$ variants were overproduced in BL21(DE3) cells that also contained a plasmid encoding tRNAs that recognize rare codons (pRARE). Cultures were grown until reaching an OD$_{600}$ ~ 0.7, then chilled on ice and induced by adding 1 mM IPTG prior to incubating overnight at 16°C. Cells were harvested and washed once with 1X phosphate buffered saline before storing the cell pellets at -80°C.

The proteins were purified by first resuspending the cell pellets in 20 mL of buffer A (20 mM Tris pH 8.0, 500 mM NaCl, 1.4 mM β-mercaptoethanol, and 5% glycerol)

supplemented with 15 mM imidazole and lysed using sonication. The cell lysate was centrifuged at 39,000 g and the supernatant was then loaded onto a 1 mL HiTRAP chelating HP column (GE Healthcare) charged with Ni^{2+}. After washing with buffer A containing 15 mM imidazole, a linear gradient was set from 15 mM to 300 mM imidazole in 48 minutes at 0.25 mL/min. Fractions containing the HIS_6-SUMO-tagged protein were pooled and diluted 1/5 in buffer A lacking NaCl. The HIS_6-SUMO tag was then digested using TEV protease for 2 hours at room temperature. The digested protein was diluted in half with buffer A, filtered and reloaded onto the Ni^{2+}-charged 5 mL HiTRAP chelating HP column. After loading, the column was washed with 3 column volumes of buffer A supplemented with 24 mM imidazole. The unbound protein collected during loading as well as the 24 mM imidazole wash were combined and concentrated prior to storing at -80°C.

4.4.2 Cloning, Expression and Purification of Rad53^{FHA1}

For the NMR experiments, the FHA1 domain from Rad53 (amino acids 14-164 – herein referred to as Rad53^{FHA1}) was amplified using PCR primers with 5' ends containing *Kas*I or *Eco*RI sites. For cross-linking experiments, which required a size difference between Rad53^{FHA1} and Dbf4^{H-BRCT} that could be readily resolved on a SDS-polyacrylamide gel, a longer fragment (amino acids 4-165) was cloned in the same manner. Both PCR products were subcloned into pJET1.2 (Fermentas), confirmed through DNA sequencing (MOBIX, McMaster University), and then digested using *Kas*I and *Eco*RI and ligated into the pPROEX HTa (Invitrogen) multi-cloning site digested with the same enzymes.

Rad53^{FHA1} (either 14-164 or 4-165) was overproduced in BL21(DE3) pRARE cells. After growing the culture to an $OD_{600} \sim 0.7$, expression was induced using 1 mM IPTG and the cultures were chilled in an ice bath prior to incubating at 30°C for 4 hours with orbital agitation. The cells were harvested through centrifugation, washed with 1X phosphate buffered saline, and stored at -80°C. To purify Rad53^{FHA1}, the cell pellet was resuspended, lysed and cleared by centrifugation as described previously for the H-BRCT

variants (see section 4.4.1). The cell lysate was then loaded onto a 5 mL HiTRAP chelating HP column (GE Healthcare) charged with Ni^{2+} and washed with buffer A supplemented with 15 mM imidazole. The protein was eluted using a linear gradient of imidazole from 15 mM to 300 mM in 60 minutes at 1 mL/min. Eluted fractions containing Rad53^{FHA1} were pooled prior to diluting, digesting the HIS$_6$ tag with TEV protease, and purifying the untagged protein as described for the H-BRCT variants (see section 4.4.1).

4.4.3 Trypsin Proteolysis

Wild-type, Trp116Asp/Met120Ala, Leu109Ala/Trp112Ser, and Thr171Ala Dbf4^{H-BRCT} variants were loaded onto a Superdex-75 column (GE Healthcare) equilibrated with buffer B (20 mM Tris pH 8.0, 100 mM NaCl, 1 mM EDTA, 5 mM DTT, and 5% glycerol). Fractions containing the desired protein were pooled together and concentrated. The final concentration was determined using the Beer Lambert Law with the extinction coefficient 23,950 M^{-1} cm^{-1} (for wild-type and Thr171Ala) or 18,450 M^{-1} cm^{-1} (for Trp116Asp/Met120Ala and Leu109Ala/Trp112Ser). The samples were then diluted to 20 μM in buffer B.

Each reaction consisted of 9 μL of a 20 μM Dbf4^{H-BRCT} variant mixed with 1 μL of 50 mM MgCl$_2$ and 1 μL of trypsin protease ranging from 0.195 μg/mL to 12.5 μg/mL. After incubating for 1 hour at room temperature, 10 μL loading dye (56 mM Tris pH 6.8, 1.8% SDS, 0.01% bromophenol blue, 23% glycerol, and 1 M β-mercaptoethanol) was added and a 15% SDS-polyacrylamide gel was run at 150 V for 75 minutes.

4.4.4 Chemical Cross-linking

Wild-type and Trp116Asp/Met120Ala Dbf4^{H-BRCT} variants as well as Rad53^{FHA1} (amino acids 4-165) were loaded onto a Superdex-75 column (GE Healthcare) equilibrated with buffer C (20 mM HEPES pH 7.7, 100 mM NaCl, 1 mM EDTA, 5 mM DTT, and 5% glycerol). Fractions containing the desired protein were pooled and

concentrated. Protein concentration was determined using the extinction coefficients of the Dbf4[H-BRCT] variants described in section 4.4.3, and 12,950 M^{-1} cm^{-1} for Rad53[4-165]. The samples were then adjusted to 45 μM using buffer C.

The cross-linking reactions consisted of 4 μL of each protein at 45 μM, while reactions with only a single protein were compensated with 4 μL of buffer C. The phosphorylated peptide PEP1-pThr [DGES(pThr)DEDDVVS] and a control peptide lacking phosphorylation (PEP1) were synthesized by Biomatik. These peptides represent Cdc7 residues 480-492 as described previously (Aucher et al., 2010). The lyophilized peptides were dissolved in buffer C to a final concentration of 1.44 mM, 0.72 mM or 0.36 mM. 1 μL of the peptide was then added to the reaction (or 1 μL of buffer C in the control reaction). The cross-linker BS[3] (Sigma) was prepared by making a 0.75 mM stock in 20 mM HEPES pH 7.7, and 1 μL was added to each reaction. After incubating for 30 minutes at room temperature, 1 μL of 0.5 M Tris was added to quench the reactions for a further 15 minutes. Then 10 μL of loading dye was added and an 18% SDS-polyacrylamide gel was run at 150 V for 90 minutes.

4.4.5 NMR Spectroscopy

Dbf4[H-BRCT] (pAG8209 encoding amino acids 105-220) was uniformly labeled by growing in M9 media containing $^{15}NH_4Cl$ and ^{13}C-D-glucose using the same expression and purification protocols described previously (Matthews et al., 2012). The protein was loaded onto a Superdex-S75 column (GE Healthcare) equilibrated with buffer B and concentrated. The concentration was measured by using the Beer Lambert Law with an extinction coefficient of 23,950 M^{-1} cm^{-1} and adjusted to 0.5 mM. NMR experiments were carried out using a Bruker AV 700 spectrometer equipped with a TCI cryoprobe. Heteronuclear Single-Quantum Coherence (HSQC) spectra and triple resonance experiments (HNCA, HNCO, HNCOCA, and CBCA(CO)NH) were collected as previously described (Selvaratnam et al., 2012). The data was processed with NMRPipe (Delaglio et al., 1995) for resonance assignment using SPARKY (Goddard and Kneller).

Order parameters were calculated using the Random Coil Index (RCI) (Berjanskii and Wishart, 2005).

Rad53^{FHA1} (amino acids 14-164) was uniformly labeled by growing in M9 media containing $^{15}NH_4Cl$. It was expressed and purified as described in section 4.4.2 followed by loading onto a Superdex-S75 column (GE Healthcare) equilibrated with buffer D (10 mM sodium phosphate buffer pH 6.5, 1 mM EDTA, and 5 mM DTT). Unlabelled Dbf4^{H-BRCT} was purified as described previously and also exchanged with buffer D using size exclusion chromatography (Matthews et al., 2012). The sample of ^{15}N-Rad53^{FHA1} was mixed with an equal molarity of unlabeled Dbf4^{H-BRCT} (or buffer D for a control lacking Dbf4) to a final concentration of 0.15 mM – 0.3 mM. The NMR experiments were conducted as described previously (Yongkiettrakul et al., 2004) with a Bruker 700 MHz spectrometer (see above). The HSQC spectrum obtained matched the previously published data such that we were able to utilize available resonance assignments, which accounted for ~ 50% of the Rad53^{FHA1} residues (Yongkiettrakul et al., 2004).

To test for competitive binding between Dbf4 and a phospho-peptide, PEP1-pThr and PEP1 (see section 4.4.4) were resuspended in 10 mM sodium phosphate buffer pH 6.5 and titrated into a sample of ^{15}N-Rad53^{FHA1} with an equal molarity of unlabelled Dbf4^{H-BRCT} (0.12 mM). An HSQC spectrum was collected after each addition of peptide and the total volume added was < 20% of the total protein volume.

4.4.6 Statistical Analysis

The intensity loss for peaks in the ^{15}N-Rad53^{FHA1} HSQC spectrum after addition of unlabeled Dbf4^{H-BRCT} was analyzed based on a method published previously (Diaz-Moreno et al., 2005). The peak fit height was measured at four different contour levels with SPARKY using a Lorentzian fit to integrate the peaks (Goddard and Kneller). The mean peak fit height was then calculated and normalized by dividing by the most intense peak, which was the C-terminal residue Arg164, prior to being compared with a similarly analyzed data set from a sample lacking Dbf4. The peak height fraction of the two

samples was expressed as a percentage and the mean was taken excluding Arg164, which was used for normalization and would therefore bias the average to a higher value. In addition, Ile145 was also excluded as its value exceeded the average beyond four standard deviations. Perturbed residues were defined as having a peak height ratio below one standard deviation from the mean.

4.5 Results

4.5.1 α0 is Anchored to the BRCT Domain in Solution

Although α0 is anchored to the BRCT domain in the crystal structure of Dbf4[H-BRCT], the N-terminal tail preceding the helix is engaged in crystal contacts (Matthews et al., 2012). It was unclear, therefore, whether the BRCT domain and α0 form a single unit or whether the association is dynamic and only one state had been captured in the crystal structure. To address this further, backbone chemical shifts were assigned based on 3D NMR experiments with ^{13}C,^{15}N-Dbf4[H-BRCT] (Figure 4.1). In total, 85% of the residues could be assigned.
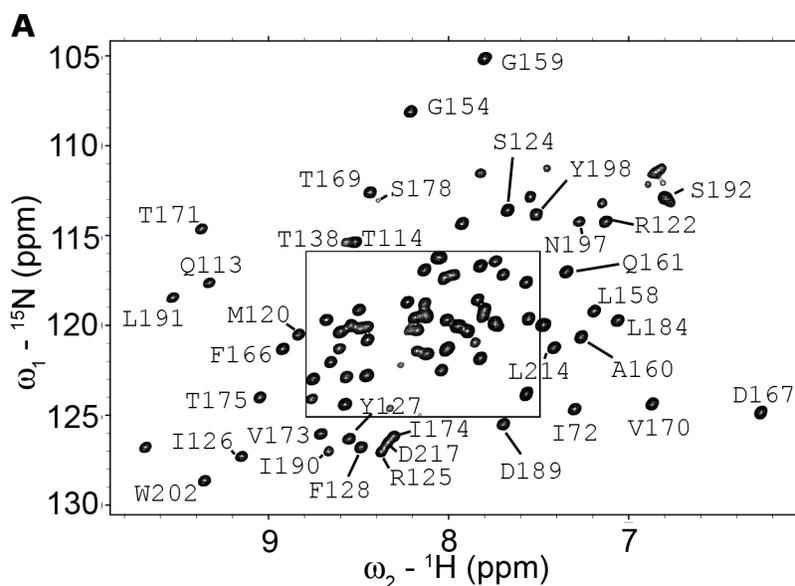
Figure 4.1 continued on page 87...

Figure 4.1 continued from page 86...
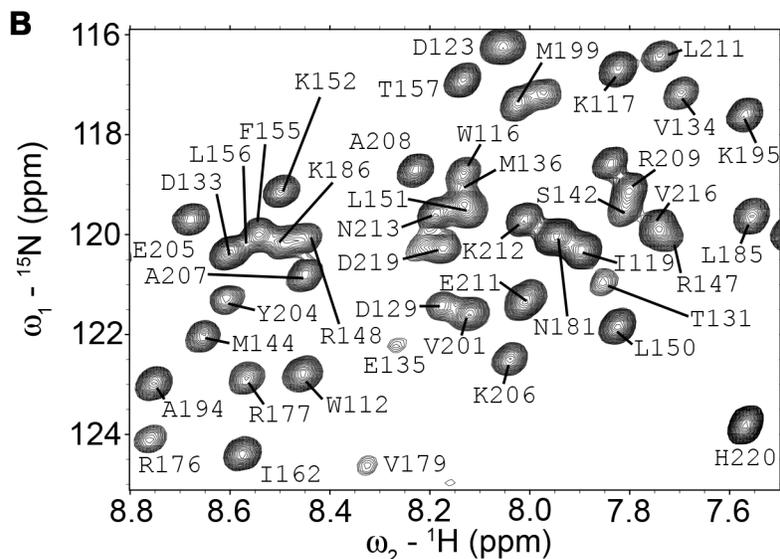
Figure 4.1. ^{15}N-Dbf4$^{H\text{-BRCT}}$ **HSQC spectrum. (A)** HSQC spectrum with ^{15}N-H resonances assigned to residues in Dbf4$^{H\text{-BRCT}}$. **(B)** Enlarged view of boxed region from panel A with assignments.

Using the backbone chemical shifts, the flexibility of Dbf4$^{H\text{-BRCT}}$ can be addressed by calculating the order parameter (S^2) for each residue, which ranges from 1 (rigid) to 0 (completely flexible) (Figure 4.2). The residues within the α0 helix have order parameters that average to 0.90, which is on par with the residues from the other helices in the domain (S^2 of 0.89 for α1, 0.83 for α2 and 0.89 for α3) and indicates that this is a stable region.

4.5.2 Point Mutations Unlatch α0 from the BRCT Domain

The stability of the α0 helix was also investigated using limited proteolysis. The linker between α0 and the BRCT domain core is rich in positively charged residues, but the susceptibility of this region to trypsin will vary depending on whether α0 is anchored or not. Wild-type Dbf4$^{H\text{-BRCT}}$ degraded as a single unit as no degradation product
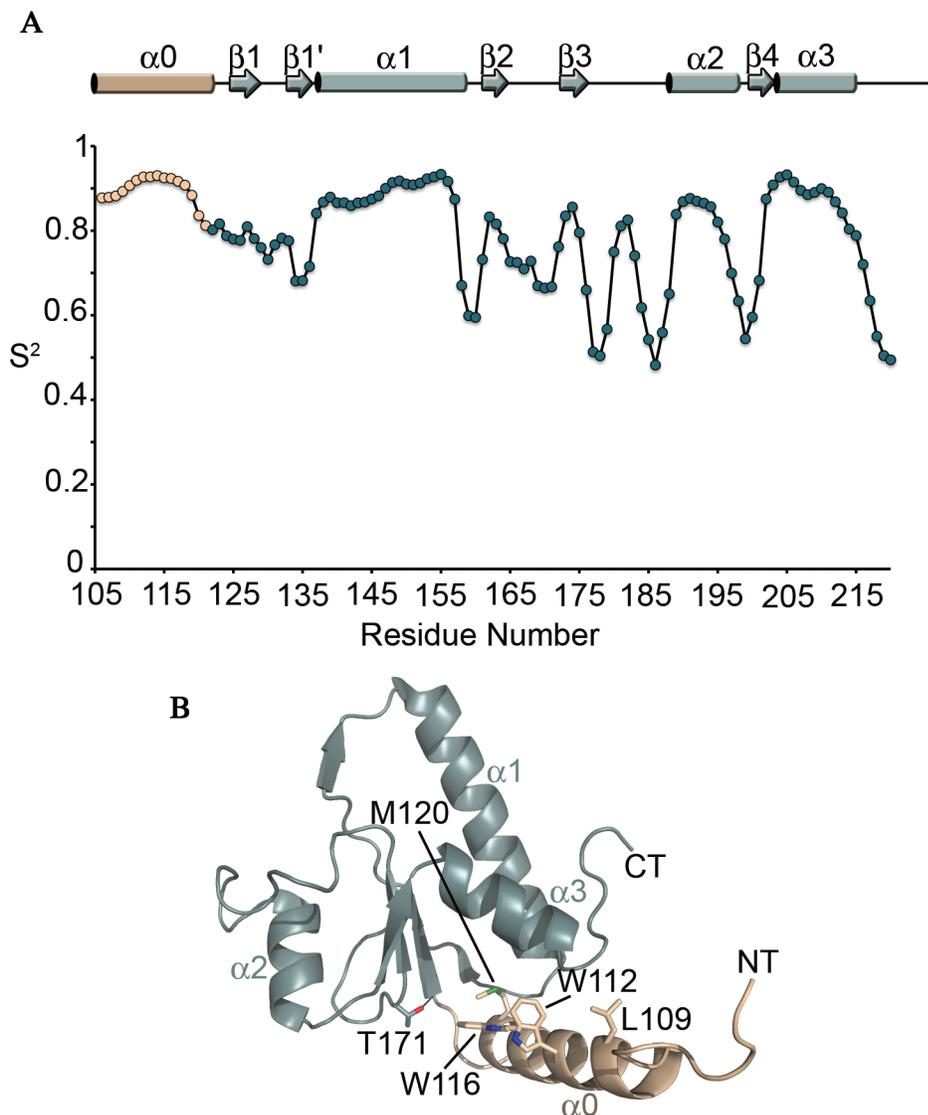
Figure 4.2. Flexibility of Dbf4[H-BRCT] α0 residues in solution. (A) Plot of the order parameter squared (S^2) versus residue number for Dbf4[H-BRCT] based on backbone chemical shift assignments. As a reference, the secondary structure from the crystal structure is provided above the graph, and the data points are color-coded to match (BRCT domain in green and α0 helix residues in tan). **(B)** Ribbon diagram of H-BRCT crystal structure (PDB ID: 3QBZ) with the BRCT (green) and α0-helix (tan) engaged through hydrophobic interactions. The side chains making up the hydrophobic face of α0 (L109, W112, W116, and M120) are shown as sticks as is T171. The α helices and N- and C-termini are labeled.

accumulated more than the others (Figure 4.3A). In contrast, the Thr171Ala mutant, which has a reduced affinity for Rad53 (Matthews et al., 2012), accumulated a fragment consistent with the BRCT domain (Figure 4.3B). Double-point mutations on the hydrophobic face of α0 (Trp116Asp/Met120Ala and Leu109Ala/Trp112Ser) that completely destroy the interaction with Rad53 (Matthews et al., 2012) also generated this pattern (Figure 4.3C and D; see also Figure 4.2B).
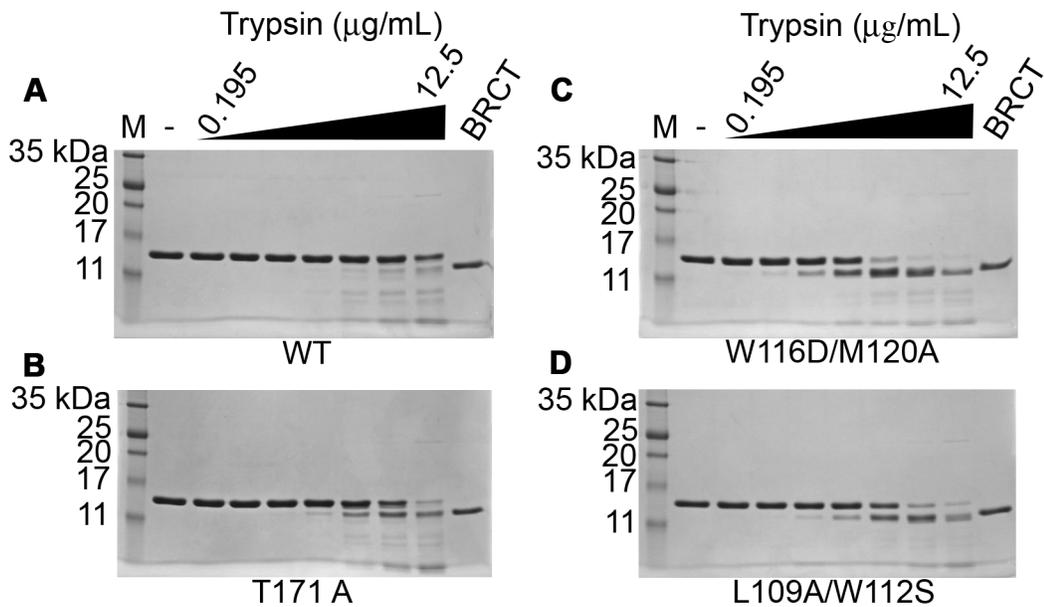
Figure 4.3. Limited proteolysis of Dbf4^{H-BRCT} mutants. 15% SDS-polyacrylamide gels showing Dbf4^{H-BRCT} variants that were digested with increasing concentrations of trypsin protease (0.195 μg/mL – 12.5 μg/mL). The second lane of each gel is a control that lacks trypsin (-) and the last lane has the purified Dbf4 BRCT domain as a molecular weight standard. The Dbf4^{H-BRCT} construct used is indicated beneath each gel.

4.5.3 Dbf4^{H-BRCT} and Rad53^{FHA1} Interact Directly *In Vitro*

In order to identify the molecular features required for the interaction between Dbf4^{H-BRCT} and Rad53^{FHA1}, we first needed to establish an experimental system that could detect this weak complex. An HSQC spectrum of uniformly labeled ^{15}N-Rad53^{FHA1} was collected with and without an equal molar ratio of Dbf4^{H-BRCT}. As Dbf4^{H-BRCT} is

unlabeled, it cannot generate any peaks. However, it can induce changes in Rad53^{FHA1} that affect the spectrum. While no chemical shifts were observed, eight new peaks (labeled a-h in Figure 4.4A) clustered in the same region of the HSQC spectrum. To demonstrate that this was due to a specific interaction, the same experiment was repeated with the non-functional BRCT domain (i.e. lacking α0) where no additional peaks formed (Figure 4.4B).
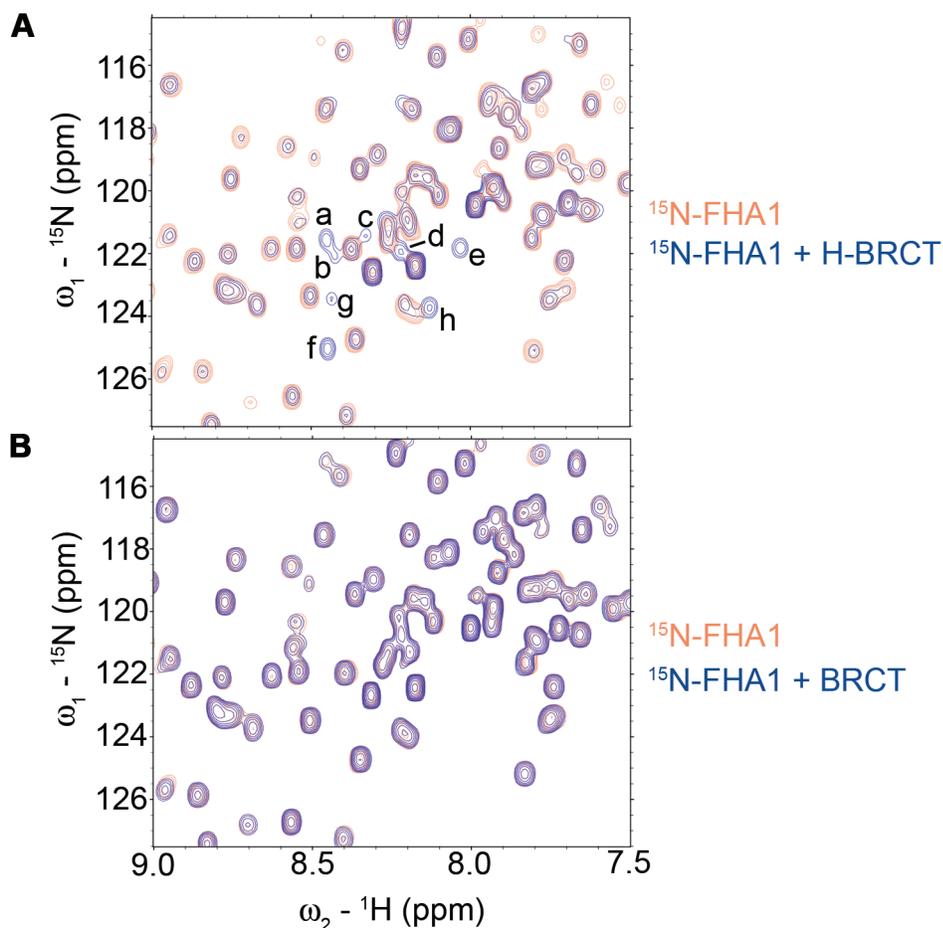
Figure 4.4. Changes induced in ^{15}N-Rad53^{FHA1} HSQC spectrum by Dbf4^{H-BRCT}. HSQC spectra of ^{15}N-Rad53^{FHA1} alone (orange) are overlaid with the same sample mixed with unlabeled Dbf4 (blue). The H-BRCT domain of Dbf4 is used in **(A)**, whereas the nonfunctional BRCT domain is used as a control in **(B)**. Peaks that appear only in the presence of Dbf4^{H-BRCT} are labeled a-h.

Another change in the HSQC spectrum of ^{15}N-Rad53^{FHA1} after the addition of Dbf4^{H-BRCT} was peak broadening resulting in a lower signal. This can occur when the molecular mass of the species being studied increases (due to forming a complex), and is a reliable indicator that an interaction is taking place. In this case, the Rad53^{FHA1}-Dbf4^{H-BRCT} complex (31.1 kDa) is almost twice the size of Rad53^{FHA1} alone (17.1 kDa).

Peaks that represent residues involved in mediating molecular interactions can also undergo line broadening. Thus, peaks that broaden beyond what is expected for an increase in molecular mass alone can be used to map the interaction interface. To determine whether the interaction surface on Rad53^{FHA1} could be identified, the signal intensity after adding Dbf4^{H-BRCT} was measured by integrating the peaks with previously published assignments and measuring their height using SPARKY as described in the experimental procedures (Goddard and Kneller). These values were then compared to the same peaks of a sample of ^{15}N-Rad53^{FHA1} alone to reveal specific changes induced by Dbf4. On average the peaks were reduced to 45% of their original height after the addition of the H-BRCT domain (Figure 4.5), whereas they maintained 95% of their height after adding the non-functional BRCT domain (data not shown), again demonstrating that the peak broadening is specific to cases where the interaction is taking place. The perturbed residues were defined as those whose peak heights were reduced beyond one standard deviation from the mean (Figure 4.5). However, no peaks were reduced beyond two standard deviations.

Interestingly, the perturbed residues in Rad53^{FHA1} predominantly all fall within the same β-sheet of the sandwich (colored tan in Figure 4.5). Although only assigned residues were included in this analysis, these were distributed throughout the entire domain. Thus, this observation cannot be attributed to a sampling bias. We therefore mapped these residues onto the Rad53^{FHA1} crystal structure and determined that many of them clustered to define a surface on the opposite side of the domain from Arg70 (Figure 4.6).
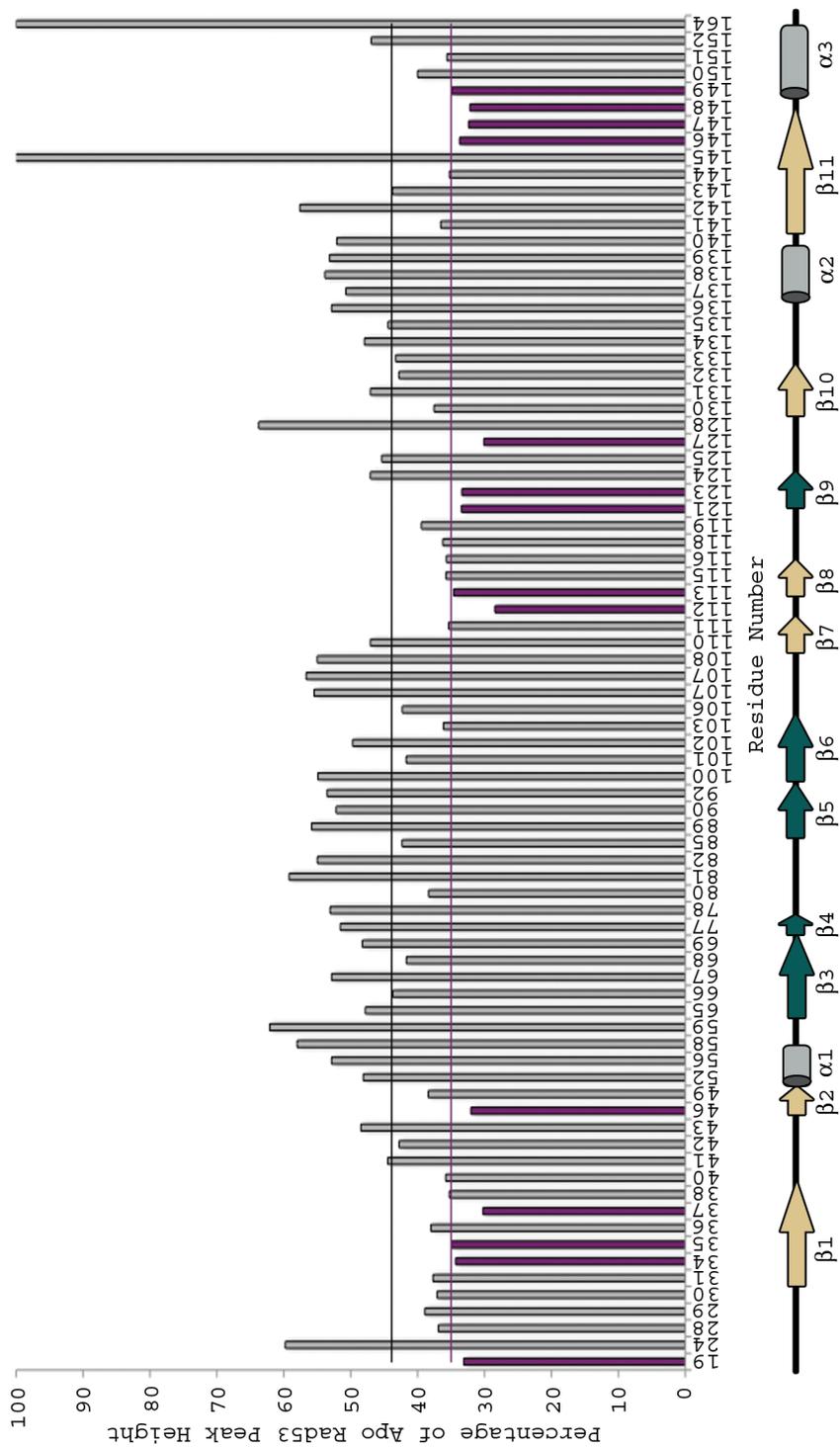
Figure 4.5. Identification of Rad53^{FHA1} Residues Perturbed by Dbf4^{H-BRCT}. The peak height from the HSQC spectrum of ¹⁵N-Rad53^{FHA1} with Dbf4^{H-BRCT} (Figure 4.4A) was measured as a percentage of the peak height when Rad53 was alone (Apo Rad53). The black line marks the mean percentage while the purple line marks one standard deviation below this value. Bars that represent perturbed residues are coloured purple. The secondary structure from the crystal structure of Rad53^{FHA1} is shown below the x-axis as a reference with the β-strands that make up the two sides of the β-sandwich colored tan and green, respectively.
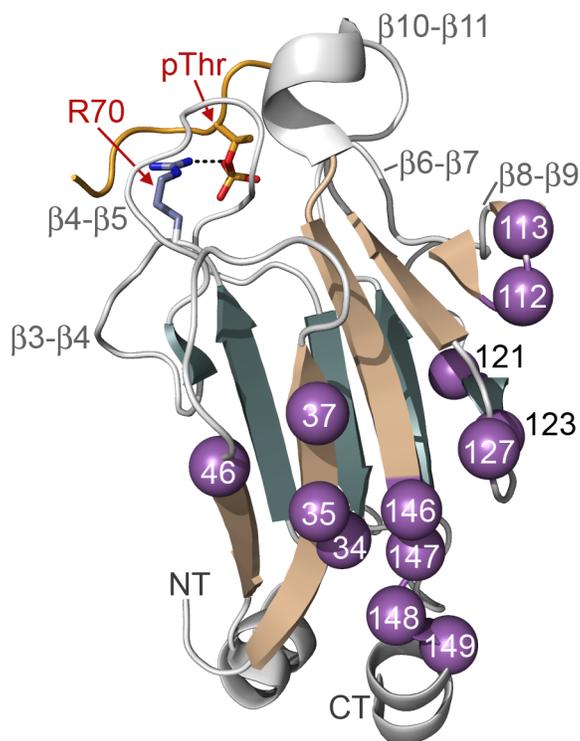
Figure 4.6. Mapping Rad53[FHA1] Residues Perturbed by Dbf4[H-BRCT]. Ribbon diagram of the Rad53[FHA1] crystal structure (PDB ID: 1G6G) with the Cα atoms of perturbed residues identified in Figure 4.5 shown as purple spheres and labeled by residue number. The side chain of Arg70 and an engaged phospho-threonine (pThr) from a peptide are also shown as a reference. The loops that surround the peptide binding site are labeled in grey according to the β-strands that they connect. The N- and C-termini are labeled.

4.5.4 A Phospho-peptide Does Not Outcompete Dbf4[H-BRCT] for Binding Rad53[FHA1]

To directly test whether the H-BRCT domain from Dbf4 engages Arg70, a phospho-peptide known to contact this residue was utilized to see whether it could compete with Dbf4. The Dbf4[H-BRCT]-Rad53[FHA1] complex was monitored through chemical cross-linking. A covalently cross-linked product consistent with a heterodimer was generated, which was also confirmed through mass spectrometry (Figure 4.7A). This represents a specific interaction as the efficiency of cross-linking was significantly reduced for the heterodimer when using the nonfunctional Trp116Asp/Met120Ala H-BRCT mutant (Figure 4.7A). Note that the residues mutated in the control do not directly contribute to cross-linking, which instead targets primary amines.
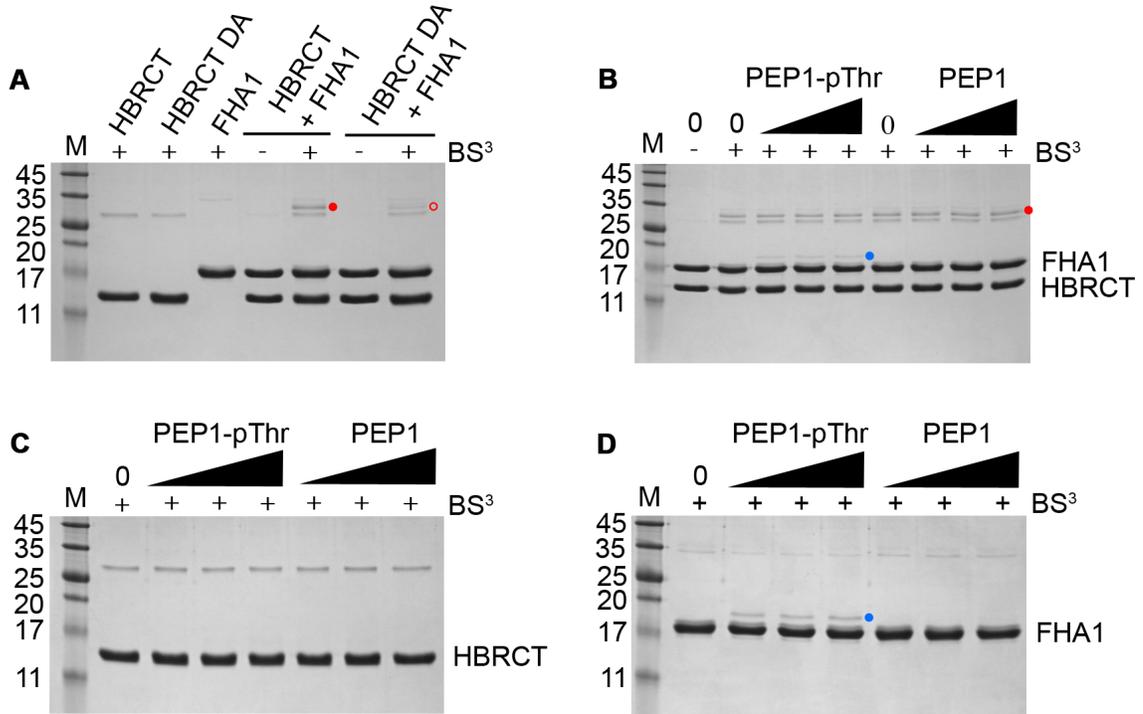
Figure 4.7. Chemical cross-linking of Dbf4[H-BRCT] and Rad53[FHA1] in the presence of a phospho-peptide. (A) 18% SDS-polyacrylamide gel demonstrating chemical cross-linking of the Dbf4[H-BRCT]-Rad53[FHA1] heterodimer (marked with a red dot). Dbf4[H-BRCT], either wild-type (HBRCT) or Trp116Asp/Met120Ala (HBRCT DA), and Rad53[FHA1] (FHA1) were incubated with (+) or without (-) BS[3]. **(B)** Wild-type Dbf4[H-BRCT] (HBRCT) and Rad53[FHA1] (FHA1) were incubated with increasing concentrations of a peptide containing a phospho-threonine moiety (PEP1-pThr) or unmodified (PEP1) prior to cross-linking using the same conditions as panel A. A band unique to reactions containing PEP1-pThr is marked with a blue dot. This experiment was repeated with wild-type Dbf4[H-BRCT] (HBRCT) **(C)** and Rad53[FHA1] (FHA1) **(D)** separately to determine which protein was creating the cross-linked band in the presence of PEP1-pThr marked with a blue dot. The marker (M) is labeled in kDa.

For the titration, a phospho-peptide based on Cdc7 (PEP1-pThr) was used for two reasons. Firstly, it has already been established that it binds to Rad53[FHA1] in a phosphorylation-dependent manner making it suitable for our assay (Aucher et al., 2010). Secondly, as Cdc7, Dbf4 and Rad53 form a complex together, the phospho-epitope in

Cdc7 is a prime candidate in the cell for influencing the Dbf4-Rad53 interaction based on proximity. To ensure that any changes in cross-linking efficiency were due solely to the binding of PEP1-pThr to Rad53[FHA1], a non-phosphorylated version of the same peptide (PEP1) was used as a control. Note that PEP1-pThr and PEP1 have the same number of primary amines, and differ only in the presence or absence of a phosphate group. The peptides were added in excess (ranging from 2- to 8-fold molar excess over Rad53[FHA1]) to ensure that the FHA1 domain would be saturated. The presence of PEP1-pThr did not prevent the Dbf4[H-BRCT]-Rad53[FHA1] heterodimer from being cross-linked (Figure 4.7B). However, a new band appeared in the presence of PEP1-Thr, but not unmodified PEP1, which consisted of only Rad53[FHA1] (Figure 4.7C and D). This is likely a consequence of Rad53[FHA1] binding specifically to PEP1-pThr.

The phospho-peptide titration was also monitored by NMR. A sample of ^{15}N-Rad53[FHA1] and unlabeled Dbf4[H-BRCT] was incubated with increasing amounts of the modified peptide (PEP1-pThr) or control peptide (PEP1) while collecting HSQC spectra. The peaks that signify the interaction between Dbf4 and Rad53 were monitored to see if saturating Rad53[FHA1] with PEP1-pThr would disrupt the interaction with Dbf4 and cause them to disappear. Although peak 'g' was difficult to monitor since a close neighbor (V116) shifted on top of it during the titration, the other peaks clearly remained (Figure 4.8). This is the case even when PEP1-pThr was present in a 2-fold molar excess (data not shown). A total of 28 assigned peaks shifted in the presence of PEP1-pThr that predominantly mapped to residues in the loops surrounding Arg70. These loops are the known binding site of phospho-peptides and were also perturbed with a titration published previously using a peptide similar to PEP1-pThr (Liao et al., 2000). In addition, the control peptide (PEP1) did not result in any chemical shifts demonstrating the observed perturbations were specific to an interaction (Figure 4.8).
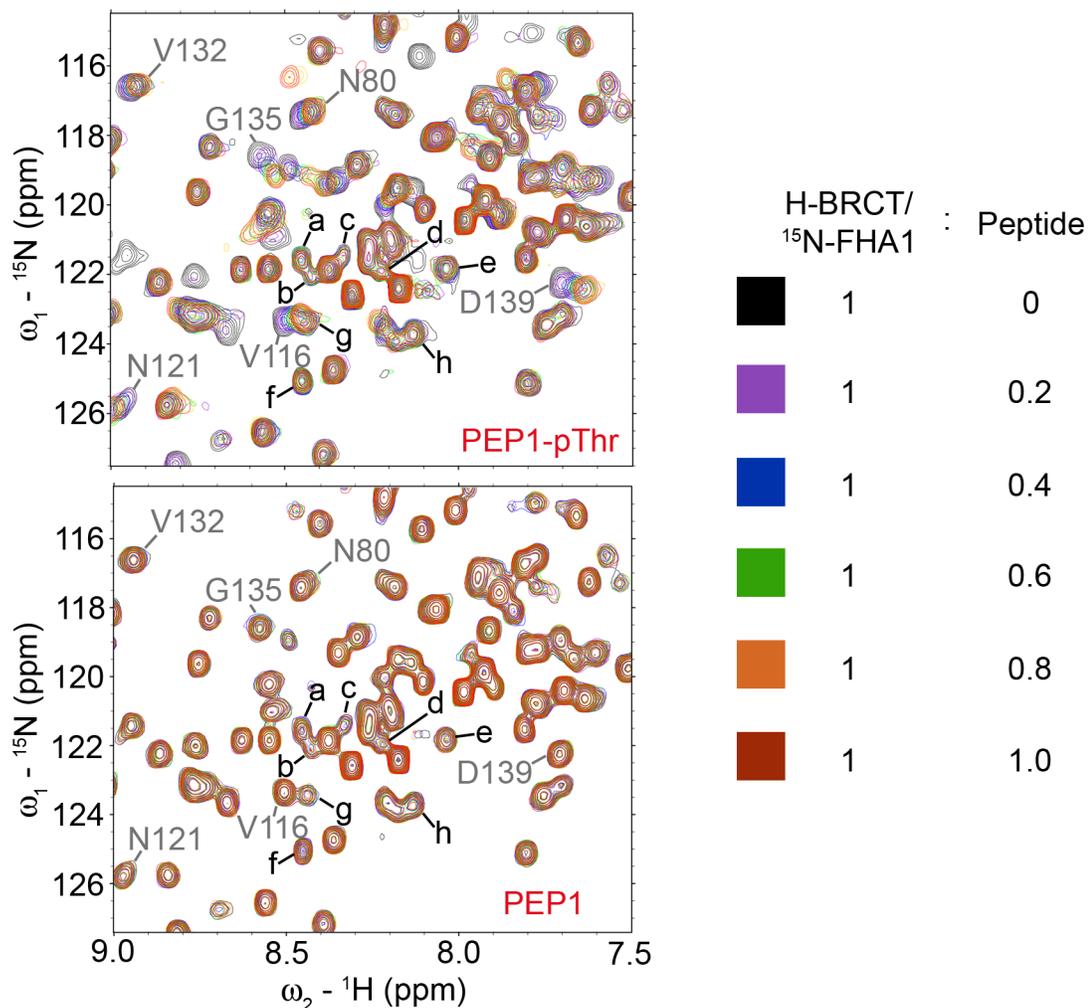
Figure 4.8. Phospho-peptide titration into sample of ^{15}N-Rad53^{FHA1}/Dbf4^{H-BRCT}. An HSQC spectrum of a 1:1 molar ratio solution of ^{15}N-Rad53^{FHA1} and unlabeled Dbf4^{H-BRCT} (black) is overlaid with spectra collected after titrating in PEP1-pThr (top) or PEP1 (bottom). Shifted peaks with known assignments are labeled in grey and peaks induced by the interaction with Dbf4^{H-BRCT} are labeled in black (a-h). The legend on the right hand side indicates the molar ratio of ^{15}N-Rad53 to peptide.

4.5.5 Rad53^{FHA1} Arg70Ala Mutant Interacts with Dbf4^{H-BRCT}

We next assessed whether Arg70 in Rad53 plays an indirect role in the interaction with Dbf4 by stabilizing the structure of the FHA domain. To the best of our knowledge, this is the first analysis of the structural consequences of this mutation. An HSQC

spectrum of the ^{15}N-labelled Arg70Ala Rad53 mutant was compared to a spectrum of the wild-type protein. In total, 23 assigned peaks shifted (Figure 4.9). However, only one residue (Leu141) was close to the surface suspected of binding Dbf4 (Figure 4.6). The rest of the perturbed residues localized to the loops that create the phospho-peptide binding site.
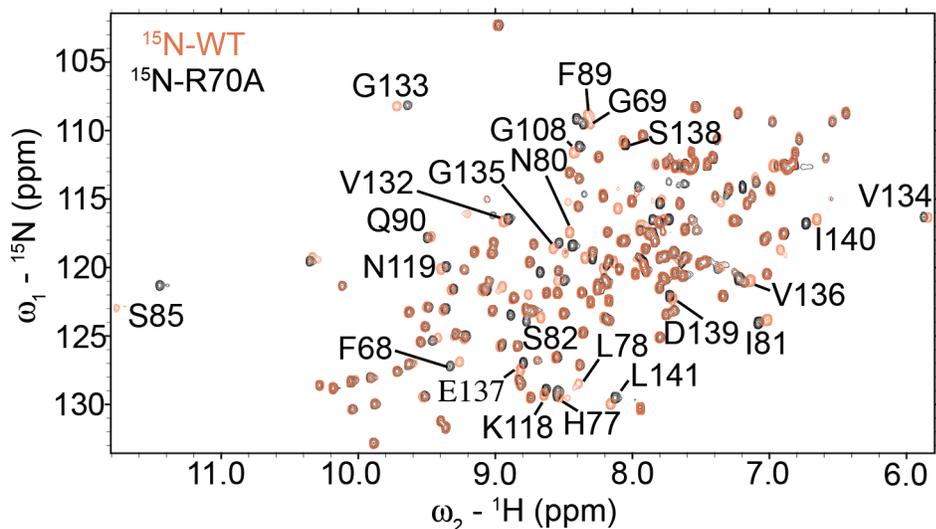
Figure 4.9 Analysis of Rad53^{FHA1} Arg70Ala Mutant. Overlay of HSQC spectra of ^{15}N-Rad53^{FHA1} wild-type (orange) with ^{15}N-Rad53^{FHA1} Arg70Ala (black). Assigned peaks that shifted are labeled.

A comparison of peaks that shifted during the PEP1-pThr titration (Figure 4.8) with those that shifted for the Arg70Ala mutant (Figure 4.9) showed a significant overlap (Table 4.1). This indicates that Arg70 not only plays a key role in binding to phospho-peptides by hydrogen bonding to the phosphate group, but also stabilizes the peptide binding surface in the apo protein. However, these results also indicate that Arg70 does not play a role in binding to the H-BRCT domain of Dbf4. To test this directly, the NMR experiment shown in Figure 4.4 was repeated with the Arg70Ala mutant of Rad53. As shown in Figure 4.10, Dbf4^{H-BRCT} induced the appearance of the eight peaks indicative of an interaction in the ^{15}N-Rad53^{FHA1} Arg70Ala HSQC spectrum.

Table 4.1 Comparison of perturbed Rad53^{FHA1} residues

Rad53^{FHA1} Structural Feature[a]	Residues Perturbed by PEP1-pThr	Residues Perturbed by Arg70Ala
β3-β4 loop	68, 69	68, 69
β4-β5 loop	77, 78, 80, 81, 82, 85	77, 78, 80, 81, 82, 85
β5-Strand	89	89, 90
β6-β7 loop	102, 106, 107, 108	108
β8-β9	116, 118, 119, 121, 123	118, 119
β10-β11	132, 133, 134, 135, 136, 137, 138, 139, 140	132, 133, 134, 135, 136, 137, 138, 139, 140
β11-Strand	143	141

[a]See Figure 4.6 for a reference to the structure of Rad53^{FHA1}
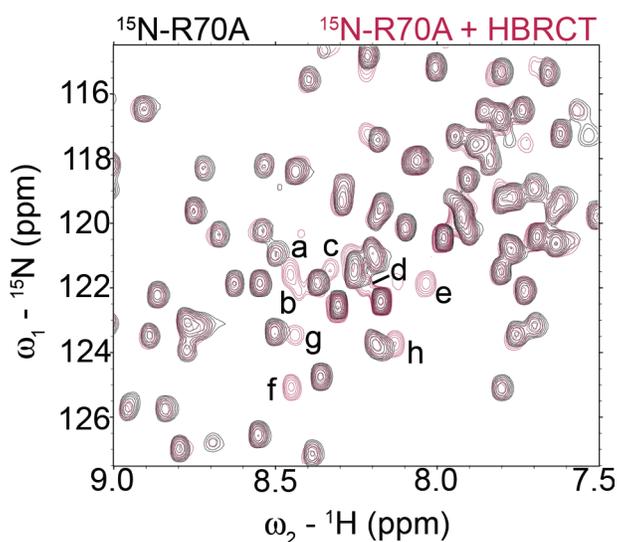
Figure 4.10 The interaction between Rad53^{FHA1} Arg70Ala and Dbf4^{H-BRCT}. Overlay of HSQC spectra of ^{15}N-labelled Rad53^{FHA1} Arg70Ala mutant (R70A) with (red) and without (black) unlabeled Dbf4^{H-BRCT} (HBRCT) in a 1:1 ratio. Peaks that represent an interaction are labeled a-h as in Figure 4.4.

4.6 Discussion

4.6.1 The H-BRCT Domain is Stable in Solution

Our previous crystallographic studies demonstrated the presence of an N-terminal amphipathic helix (α0) that was anchored to the core of the Dbf4 BRCT domain through its hydrophobic face (Matthews et al., 2012). Here we verified that this H-BRCT fold

also exists in solution through an analysis of order parameters calculated from chemical shifts (Figure 4.2) and trypsin digestions (Figure 4.3). Importantly, point mutations that destroy the interaction with Rad53 were found to unlatch the α0 helix from the rest of the domain (Figure 4.3). This includes the Thr171Ala mutation, which we had previously suspected of being structurally destabilizing (Matthews et al., 2012). Although the side chain of Thr171 is hydrogen bonded to the loop connecting α0 to the BRCT domain (Figure 4.11A), this residue may not be directly involved in latching the helix. Many BRCT domains have a conserved threonine (or serine) in this same position that lack an N-terminal helix. For domains with available structures, this residue uses its side chain as a hydrogen bond donor to β1 similarly to Thr171. This may increase the stability of the central β-sheet in the core of the domain. In the N-terminal BRCT domain from BRCA-1, mutation of the equivalent residue (Thr1685) destroys the function of this protein and is causally linked to hereditary breast cancer (Figure 4.11B) (Easton et al., 2007). Thus, Thr171 is playing a conserved structural role in the BRCT domain of Dbf4, and mutating it may indirectly unlatch the α0 helix by causing a general instability.

4.6.2 Potential Interaction Surface on Rad53[FHA1]

In order to detect and study the Dbf4-Rad53 interaction we have used NMR spectroscopy. This revealed a patch of residues on one of the β-sheets that were perturbed in the presence of Dbf4[H-BRCT]. Other FHA domains have been known to use their β-sheets to associate with binding partners (Li et al., 2004; Li et al., 2002; Tong et al., 2010). By using site-directed mutagenesis combined with yeast two-hybrid assays, the role of this surface in the interaction between Dbf4 and Rad53 can be directly tested in the future. In addition, it will be instructive to determine whether the Arg70Ala mutant can interact with the isolated H-BRCT domain from Dbf4 in a yeast two-hybrid assay. We have shown that Arg70 does not play a role in the interaction *in vitro*, which is conflicting with previous yeast two-hybrid data (Duncker et al., 2002). However, it is important to note that the yeast experiments were done with full-length proteins and thus it is likely that another region outside of the H-BRCT domain of Dbf4 contacts Arg70.
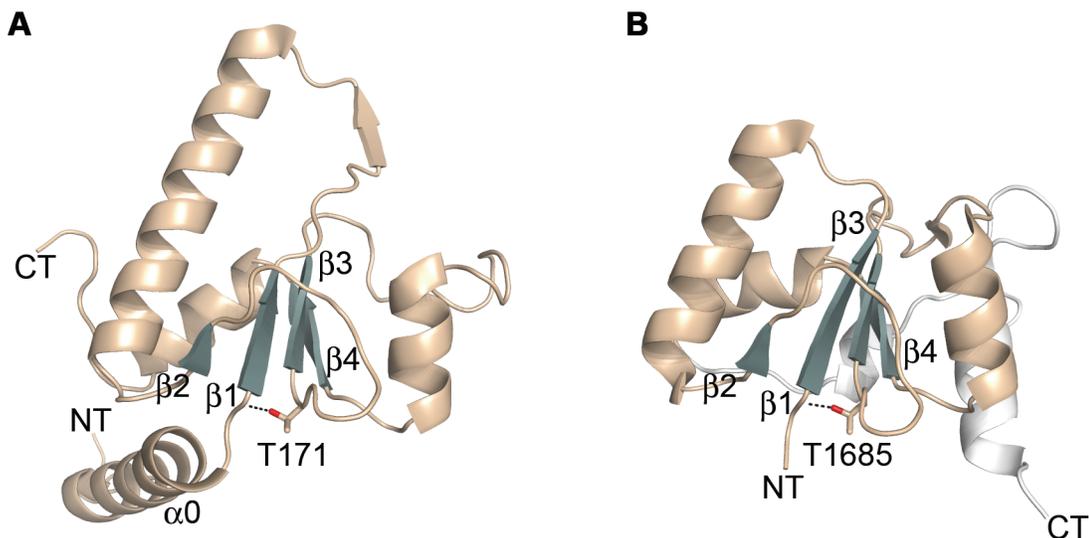
Figure 4.11 Comparison of Thr171 from Dbf4 and Thr1685 from BRCA-1. (A) Ribbon diagram of Dbf4$^{\text{H-BRCT}}$ (PDB 3QBZ) colored in tan with the central β-sheet colored green and labeled. The side chain of Thr171 (T171) is shown as sticks with the oxygen atom in red. A dotted line represents the hydrogen bond between Thr171 and the loop between β1 and α0. **(B)** The N-terminal BRCT domain (residues 1649 – 1757) from the tandem pair in BRCA-1 (PDB 1T2V) represented the same as panel A. The N- and C-termini are labeled.

4.6.3 Binding Mode Between Rad53 and Dbf4

A thought-provoking view of FHA domains as logic gates was published after the discovery of an FHA module that requires two phospho-threonines in order to recognize a peptide (Zhang and Durocher, 2008). Logic gates, which accept inputs and produce a specific output, are an instructive way of viewing the flow of information between proteins in signaling pathways. When the output is considered to be peptide binding, the FHA1 domain from Rad53 acts as an 'OR' logic gate, because it can only bind to one phospho-threonine within a multiply phosphorylated peptide (Zhang and Durocher, 2008). In contrast, the FHA domain from Dun1 was considered an 'AND' logic gate by requiring two phospho-threonines that bind simultaneously (Lee et al., 2008; Zhang and Durocher, 2008).

Here we have shown that if the output is instead considered to be binding to full-length Dbf4, Rad53[FHA1] can also behave as an 'AND' logic gate, requiring the H-BRCT domain as well as an undetermined second "input" that engages Arg70. This would explain why disabling either contact point in the yeast two-hybrid assay (by mutating Arg70 of FHA1 or mutating the H-BRCT domain of Dbf4) destroys the interaction. This is reminiscent of the FHA domain from the human homolog of Rad53 (Chk2), which also requires the tandem BRCT domain from BRCA-1 to contact two distal surfaces simultaneously (Li et al., 2002). These examples underline the experimental difficulty in detecting dual input FHA domains, as the complete abolishment of the interaction through mutation of only one of the required surfaces masks the presence of a second site.

The view of FHA domains behaving as 'AND' logic gates was suggested to indicate that the linked "outputs" are costly and thus require complex input signals (Zhang and Durocher, 2008). This proposal also fits with the Rad53-Dbf4 and Chk2-BRCA-1 interactions, which can also be viewed as leading to costly outputs (inhibition of late origin firing and homologous recombination, respectively). In the case of Rad53-Dbf4, the second input (that engages Arg70) is still unknown. However, it is likely a phospho-epitope that may reside in Dbf4 outside of the H-BRCT domain. As the yeast two-hybrid assay takes place in the presence of endogenous proteins, it is also possible that the phospho-threonine is actually in cellular Cdc7 that is bound to our bait protein, as it has already been shown to contain such an epitope that binds to Rad53[FHA1] (Aucher et al., 2010). This would be similar to the kinesin KIF13, which binds to two different molecules simultaneously through a surface equivalent to the one containing Arg70 in Rad53 (though in this case the interaction is phosphorylation-independent as the residues required for phosphate recognition are not conserved) as well as a second surface on the face of one of its β-sheets (Tong et al., 2010). Lastly, it is possible that binding to any phospho-epitope causes a change in Rad53[FHA1] that increases the affinity for Dbf4 (cooperative binding).

Another complication of the endogenous proteins present in the yeast two-hybrid assay is in changing the relative contribution of the dual inputs in the Rad53-Dbf4 interaction. This has been noted in the past, and seems to be due to the fact that other protein-interaction modules flank the H-BRCT domain. On the N-terminal side, a small motif is present that binds to Cdc5 (Chen and Weinreich, 2010). Motif M is on the C-terminal side, which binds to a component of the replicative helicase (Mcm2) as well as Cdc7 (Varrin et al., 2005). The removal of the N-terminal region containing the Cdc5 interaction site reciprocally strengthens the interaction of the mutant protein with Rad53 in a yeast two-hybrid (Matthews et al., 2012). This suggests there may be a competition between our prey Rad53 protein and endogenous targets for neighboring motifs in the Dbf4 bait. In fact, the isolated H-BRCT domain has a higher affinity for Rad53 relative to full-length Dbf4 in a yeast two-hybrid analysis, even though it is missing the second input (presumably a phospho-epitope) that binds Arg70 based on the work we presented here (Matthews et al., 2012). It will therefore be very important in the future to continue searching for the second input in the context of the full-length proteins as various truncations may alter the basal interaction strength of the H-BRCT domain itself for Rad53, greatly complicating (at best) or even masking (at worst) the contribution of the other regions of Dbf4.

The ligand that engages Arg70 when the full-length proteins interact during the replication checkpoint is likely a phospho-threonine, suggesting that a threonine mutagenesis approach would be appropriate. This also has implications for how the replication checkpoint functions. If a phospho-epitope can be identified, then the kinase responsible for generating this site will reveal how the cell regulates the Dbf4-Rad53 interaction. FHA domains employing dual binding sites appears to be a growing trend among this highly prevalent and versatile module, and indicates how an FHA domain can balance its high diversity of binding partners with maintaining the control and specificity required of proteins responding to DNA damage and stress.

4.7 Acknowledgments

CONCLUSIONS

5.1 BRCT Domains Act as Structural Scaffolds

The first two objectives of this research project were to define the boundaries of the putative BRCT domain from Dbf4 and solve its crystal structure. We discovered that Dbf4 contains a *bona fide* BRCT domain with a unique surface created by the α0 helix that is recognized by Rad53 (amino acids 105-220). However, α0 was initially difficult to detect. The BRCT domain without the α0 helix (amino acids 120-250) still encoded a soluble protein fragment that could be over-expressed in an *E. coli* expression host (Chapter 2). It was also stable enough to grow diffraction quality crystals. This is surprising given that monitoring solubility is normally a reliable indicator that a complete domain has been isolated (Dyson, 2010). Truncated domains usually do not fold properly, and are poorly behaved in solution. The importance of the α0 helix was also obscured because its sequence is highly variable between species. This is because α0 is only present in fungal Dbf4 homologs. Consequently, sequence alignments that included higher eukaryotes masked the conservation of this region. However, we have shown that α0 is an integral component of the domain that shares the hydrophobic core of the BRCT fold (Chapter 3 and 4).

The boundaries of a functional BRCT domain can be difficult to define since they behave like structural scaffolds. While these scaffolds contain binding sites, they are usually not sufficient for molecular interactions. For instance, the N-terminal BRCT in a tandem pair has a pocket that binds to a phosphorylated residue, but is still not able to bind a phospho-peptide. Instead, the interaction also requires another pocket created by

the association of a second BRCT tandem partner (Manke et al., 2003). In addition, the BRCT domain from Dbf4 is required for the interaction with Rad53, but is not functional without the α0 helix (Chapter 3). The Rfc1 BRCT domain is unable to bind DNA despite having a positively charged pocket that recognizes the phosphate backbone. Instead, it also needs an N-terminal helix that binds in the major groove (Kobayashi et al., 2009). It is the combination of these additional elements with the BRCT scaffold that makes the whole unit functional. The term "element" in this sense can refer to single helices such as the H-BRCT domain from Dbf4, or whole domains as with a tandem BRCT pair. As demonstrated for Dbf4, this can make it difficult to decipher domain boundaries empirically since these elements may not be necessary for structural stability. Additionally, they may not even pack against the BRCT domain so it is not always possible to detect them through limited proteolysis. However, in all cases it is true that they are required for proper function. Therefore, it is highly important when examining a BRCT domain that the function of the fragment being studied is rigorously analyzed. Otherwise, important features could easily be truncated leaving what may be a structurally sound BRCT scaffold, but little to inform biological function.

5.2 Detecting Weak Protein-Protein Interactions

The third objective was to identify the molecular features necessary for the interaction between Rad53^{FHA1} and Dbf4^{H-BRCT}. This was a technical challenge, as the association between these domains is weak. Many physiologically relevant interactions have weak affinities, particularly between molecules whose association must be dynamic. Although X-ray crystallography was entertained for studying the complex, this could be problematic. Crystal formation requires a homogenous sample, and a weak affinity would lead to a mixture of predominately monomeric Rad53^{FHA1} and Dbf4^{H-BRCT} in addition to the complex. One way to overcome this bottleneck is using high protein concentrations to displace the equilibrium towards the complex and favor its crystallization. However, high protein concentrations can also trigger precipitation. Ultimately, we decided to study the interaction in solution in order to avoid these issues.

Chemical cross-linking is well suited for detecting weak interactions because the complex is captured using covalent bonds. This also did not require any modification of the proteins themselves, as they were both naturally rich in primary amines and therefore good substrates for the BS^3 cross-linker. One major limitation of cross-linking is that it can link proteins that associate nonspecifically. Therefore, we combined our cross-linking analysis with NMR spectroscopy, which is also well suited for studying weak interactions (Vaynberg and Qin, 2006). We were able to develop a method for monitoring the interaction by identifying eight peaks in an HSQC spectrum that only appeared when the proteins formed a complex (Chapter 4). While the assignments of these peaks are unknown, they nonetheless can be used to reliably identify samples competent for the interaction. In addition to this, NMR allows for monitoring the complex at the level of the individual residues. However, chemical shift perturbations accompanying weak interactions tend to be small (Vinogradova and Qin, 2012) and in fact none were detected for Rad53^{FHA1} in the presence of Dbf4^{H-BRCT}. One of the β-sheet faces was pinpointed as being a possible interaction surface based instead on changes in peak height (Chapter 4). This is fitting with a growing trend in the literature that the β-sheets of FHA domains may contribute to interactions.

5.3 Final Thoughts and Future Directions

One of the major contributions of this work is in revealing the presence of two distinct binding surfaces on the FHA1 domain of Rad53. We were able to identify the presence of both surfaces by studying the interaction *in vivo* with full-length proteins, and comparing the results to *in vitro* experiments with the isolated domains. Such a bipartite interaction may serve to increase the avidity of the complex in the cell. By engaging one of the two sites, the remaining surface is forced into close proximity to its target. This is a highly prevalent strategy used throughout nature (Oehler and Muller-Hill, 2010). One of these sites is likely the face of the β-sheet in Rad53^{FHA1}, and requires the α0 helix in Dbf4^{H-BRCT}. The second interaction surface includes Arg70 in Rad53^{FHA1}. As this residue is well known for recognizing phospho-threonines, the interaction may be

regulated by a kinase that phosphorylates Dbf4 somewhere outside of the H-BRCT domain. The confirmation of the β-sheet face interaction surface by a yeast two-hybrid assay with full-length Dbf4 will be an important future direction in the project. A long-term goal will then be to pinpoint the location of the putative phospho-threonine.

This research project characterized the interaction between Dbf4 and Rad53. However, the revelations of the dual binding sites on the FHA1 domain as well as the unique features of the H-BRCT fold also expanded the interaction modes for FHA and BRCT domains in general. This has increased our understanding of how these domains function in the cell. It also provides critical insight into why these domains are so prevalent in proteins charged with safeguarding the eukaryotic genome, which must have the versatility to interact with numerous partners.

References

Adams, P.D., Grosse-Kunstleve, R.W., Hung, L.W., Ioerger, T.R., McCoy, A.J., Moriarty, N.W., Read, R.J., Sacchettini, J.C., Sauter, N.K., and Terwilliger, T.C. (2002). PHENIX: building new software for automated crystallographic structure determination. Acta Crystallogr D Biol Crystallogr *58*:1948-1954.

Arellano, M., and Moreno, S. (1997). Regulation of CDK/cyclin complexes during the cell cycle. Int J Biochem Cell Biol *29*:559-573.

Arlt, M.F., Mulle, J.G., Schaibley, V.M., Ragland, R.L., Durkin, S.G., Warren, S.T., and Glover, T.W. (2009). Replication stress induces genome-wide copy number changes in human cells that resemble polymorphic and pathogenic variants. Am J Hum Genet *84*:339-350.

Aucher, W., Becker, E., Ma, E., Miron, S., Martel, A., Ochsenbein, F., Marsolier-Kergoat, M.C., 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*:2745-2759.

Ausubel, F., Brent, R., Kingston, R.E., Moore, D.D., Seidman, J.G., Smith, J.A., and Struhl, K. (1995). Short Protocols in Molecular Biology, 3 edn (New York, N.Y.: Wiley).

Avery, O.T., Macleod, C.M., and McCarty, M. (1944). Studies on the Chemical Nature of the Substance Inducing Transformation of Pneumococcal Types : Induction of Transformation by a Desoxyribonucleic Acid Fraction Isolated from Pneumococcus Type Iii. J Exp Med *79*:137-158.

Baker, N.A., Sept, D., Joseph, S., Holst, M.J., and McCammon, J.A. (2001). Electrostatics of nanosystems: application to microtubules and the ribosome. Proc Natl Acad Sci U S A *98*:10037-10041.

Bartek, J., Bartkova, J., and Lukas, J. (2007). DNA damage signalling guards against activated oncogenes and tumour progression. Oncogene *26*:7773-7779.

Bell, S.P., and Dutta, A. (2002). DNA replication in eukaryotic cells. Annu Rev Biochem *71*:333-374.

Berjanskii, M.V., and Wishart, D.S. (2005). A simple method to predict protein flexibility using secondary chemical shifts. J Am Chem Soc *127*:14970-14971.

Bhattacharyya, R.P., Remenyi, A., Good, M.C., Bashor, C.J., Falick, A.M., and Lim, W.A. (2006). The Ste5 scaffold allosterically modulates signaling output of the yeast mating pathway. Science *311*:822-826.

Bloom, J., and Cross, F.R. (2007). Multiple levels of cyclin specificity in cell-cycle control. Nat Rev Mol Cell Biol *8*:149-160.

Blow, J.J., Ge, X.Q., and Jackson, D.A. (2011). How dormant origins promote complete genome replication. Trends Biochem Sci *36*:405-414.

Bonte, D., Lindvall, C., Liu, H., Dykema, K., Furge, K., and Weinreich, M. (2008). Cdc7-Dbf4 kinase overexpression in multiple cancers and tumor cell lines is correlated with p53 inactivation. Neoplasia *10*:920-931.

Bork, P., Hofmann, K., Bucher, P., Neuwald, A.F., Altschul, S.F., and Koonin, E.V. (1997). A superfamily of conserved domains in DNA damage-responsive cell cycle checkpoint proteins. FASEB J *11*:68-76.

Bousset, K., and Diffley, J.F. (1998). The Cdc7 protein kinase is required for origin firing during S phase. Genes and Development. *12*:480-490.

Bowers, J.L., Randell, J.C., Chen, S., and Bell, S.P. (2004). ATP hydrolysis by ORC catalyzes reiterative Mcm2-7 assembly at a defined origin of replication. Mol Cell *16*:967-978.

Bradford, M.M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal Biochem *72*:248-254.

Branzei, D., and Foiani, M. (2006). The Rad53 signal transduction pathway: Replication fork stabilization, DNA repair, and adaptation. Exp Cell Res *312*:2654-2659.

Branzei, D., and Foiani, M. (2007). Interplay of replication checkpoints and repair proteins at stalled replication forks. DNA Repair (Amst) *6*:994-1003.

Branzei, D., and Foiani, M. (2009). The checkpoint response to replication stress. DNA Repair (Amst) *8*:1038-1046.

Bruck, I., and Kaplan, D. (2009). Dbf4-Cdc7 phosphorylation of Mcm2 is required for cell growth. J Biol Chem *284*:28823-28831.

Bruck, I., and Kaplan, D.L. (2011). GINS and Sld3 compete with one another for Mcm2-7 and Cdc45 binding. J Biol Chem *286*:14157-14167.

Bryson, K., McGuffin, L.J., Marsden, R.L., Ward, J.J., Sodhi, J.S., and Jones, D.T. (2005). Protein structure prediction servers at University College London. Nucleic Acids Res *33*:W36-38.

Burhans, W.C., and Weinberger, M. (2007). DNA replication stress, genome instability and aging. Nucleic Acids Res *35*:7545-7556.

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

Campbell, S.J., Edwards, R.A., and Glover, J.N. (2010). Comparison of the structures and peptide binding specificities of the BRCT domains of MDC1 and BRCA1. Structure *18*:167-176.

Casper, A.M., Nghiem, P., Arlt, M.F., and Glover, T.W. (2002). ATR regulates fragile site stability. Cell *111*:779-789.

Chapman, J.W., and Johnston, L.H. (1989). The yeast gene, DBF4, essential for entry into S phase is cell cycle regulated. Exp Cell Res *180*:419-428.

Chargaff, E. (1951). Structure and function of nucleic acids as cell constituents. Fed Proc *10*:654-659.

Chen, S., de Vries, M.A., and Bell, S.P. (2007). Orc6 is required for dynamic recruitment of Cdt1 during repeated Mcm2-7 loading. Genes Dev *21*:2897-2907.

Chen, S.H., and Zhou, H. (2009). Reconstitution of Rad53 activation by Mec1 through adaptor protein Mrc1. J Biol Chem *284*:18593-18604.

Chen, Y., and Sanchez, Y. (2004). Chk1 in the DNA damage response: conserved roles from yeasts to mammals. DNA repair *3*:1025-1032.

Chen, Y.C., and Weinreich, M. (2010). Dbf4 regulates the Cdc5 Polo-like kinase through a distinct non-canonical binding interaction. J Biol Chem *285*:41244-41254.

Cheng, L., Collyer, T., and Hardy, C.F. (1999). Cell cycle regulation of DNA replication initiation factor Dbf4p. Mol Cell Biol *19*:4270-4278.

Collaborative Computational Project Number 4 (1994). The CCP4 suite: programs for protein crystallography. Acta Crystallogr D Biol Crystallogr *50*:760-763.

Cormack, B., and Castano, I. (2002). Introduction of Point Mutations into Cloned Genes. Methods Enzymol *350*:199-218.

Costa, A., Ilves, I., Tamberg, N., Petojevic, T., Nogales, E., Botchan, M.R., and Berger, J.M. (2011). The structural basis for MCM2-7 helicase activation by GINS and Cdc45. Nat Struct Mol Biol *18*:471-477.

Cox, M.M., Goodman, M.F., Kreuzer, K.N., Sherratt, D.J., Sandler, S.J., and Marians, K.J. (2000). The importance of repairing stalled replication forks. Nature *404*:37-41.

Dahm, R. (2005). Friedrich Miescher and the discovery of DNA. Developmental Biology *278*:274-288.

Delaglio, F., Grzesiek, S., Vuister, G.W., Zhu, G., Pfeifer, J., and Bax, A. (1995). NMRPipe: a multidimensional spectral processing system based on UNIX pipes. J Biomol NMR *6*:277-293.

DeLano, W.L. (2002). The PyMOL Molecular Graphic Systems (Palo Alto, CA, USA).

Diaz-Moreno, I., Diaz-Quintana, A., Molina-Heredia, F.P., Nieto, P.M., Hansson, O., De la Rosa, M.A., and Karlsson, B.G. (2005). NMR analysis of the transient complex

between membrane photosystem I and soluble cytochrome c6. J Biol Chem *280*:7925-7931.

Diffley, J.F. (1998). Replication control: choreographing replication origins. Curr Biol *8*:R771-773.

Dohrmann, P.R., Oshiro, G., Tecklenburg, M., and Sclafani, R.A. (1999). RAD53 regulates DBF4 independently of checkpoint function in Saccharomyces cerevisiae. Genetics *151*:965-977.

Donaldson, A.D., Fangman, W.L., and Brewer, B.J. (1998). Cdc7 is required throughout the yeast S phase to activate replication origins. Genes Dev *12*:491-501.

Dowell, S.J., Romanowski, P., and Diffley, J.F. (1994). Interaction of Dbf4, the Cdc7 protein kinase regulatory subunit, with yeast replication origins in vivo. Science *265*:1243-1246.

Duch, A., Palou, G., Jonsson, Z.O., Palou, R., Calvo, E., Wohlschlegel, J., and Quintana, D.G. (2011). A Dbf4 mutant contributes to bypassing the Rad53-mediated block of origins of replication in response to genotoxic stress. J Biol Chem *286*:2486-2491.

Duncker, B.P., Shimada, K., Tsai-Pflugfelder, M., Pasero, P., and Gasser, S.M. (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*:16087-16092.

Dunlop, K.V., and Hazes, B. (2005). A modified vapor-diffusion crystallization protocol that uses a common dehydrating agent. Acta Crystallogr D Biol Crystallogr *61*:1041-1048.

Durkin, S.G., and Glover, J.N. (2007). Chromosome fragile sites. Annu Rev Genet *41*:169-192.

Durocher, D., Henckel, J., Fersht, A.R., and Jackson, S.P. (1999). The FHA domain is a modular phosphopeptide recognition motif. Mol Cell *4*:387-394.

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

Durocher, D., Taylor, I.A., Sarbassova, D., Haire, L.F., Westcott, S.L., Jackson, S.P., Smerdon, S.J., and Yaffe, M.B. (2000). The molecular basis of FHA domain:phosphopeptide binding specificity and implications for phospho-dependent signaling mechanisms. Mol Cell *6*:1169-1182.

Dyson, M.R. (2010). Selection of soluble protein expression constructs: the experimental determination of protein domain boundaries. Biochem Soc Trans *38*:908-913.

Easton, D.F., Deffenbaugh, A.M., Pruss, D., Frye, C., Wenstrup, R.J., Allen-Brady, K., Tavtigian, S.V., Monteiro, A.N., Iversen, E.S., Couch, F.J., and Goldgar, D.E. (2007). A systematic genetic assessment of 1,433 sequence variants of unknown clinical significance in the BRCA1 and BRCA2 breast cancer-predisposition genes. Am J Hum Genet *81*:873-883.

Emsley, P., and Cowtan, K. (2004). Coot: model-building tools for molecular graphics. Acta Crystallogr D Biol Crystallogr *60*:2126-2132.

Evrin, C., Clarke, P., Zech, J., Lurz, R., Sun, J., Uhle, S., Li, H., Stillman, B., and Speck, C. (2009). A double-hexameric MCM2-7 complex is loaded onto origin DNA during licensing of eukaryotic DNA replication. Proc Natl Acad Sci *106*:20240-20245.

Feng, H., and Kipreos, E.T. (2003). Preventing DNA re-replication--divergent safeguards in yeast and metazoa. Cell Cycle *2*:431-434.

Ferreira, M.F., Santocanale, C., Drury, L.S., and Diffley, J.F. (2000). Dbf4p, an essential S phase-promoting factor, is targeted for degradation by the anaphase-promoting complex. Mol Cell Biol *20*:242-248.

Flemming, W. (1882). Zellsubstanz, kern und zelltheilung (Leipzig: FCW Vogel).

Fletcher, R.J., and Chen, X.S. (2006). Biochemical activities of the BOB1 mutant in Methanobacterium thermoautotrophicum MCM. Biochemistry *45*:462-467.

Ford, D., Easton, D.F., Bishop, D.T., Narod, S.A., and Goldgar, D.E. (1994). Risks of cancer in BRCA1-mutation carriers. Breast Cancer Linkage Consortium. Lancet *343*:692-695.

Frick, D.N., and Lam, A.M. (2006). Understanding helicases as a means of virus control. Curr Pharm Des *12*:1315-1338.

Friedel, A.M., Pike, B.L., and Gasser, S.M. (2009). ATR/Mec1: coordinating fork stability and repair. Curr Opin Cell Biol *21*:237-244.

Furuya, K., Miyabe, I., Tsutsui, Y., Paderi, F., Kakusho, N., Masai, H., Niki, H., and Carr, A.M. (2010). DDK phosphorylates checkpoint clamp component Rad9 and promotes its release from damaged chromatin. Mol Cell *40*:606-618.

Gabrielse, C., Miller, C.T., McConnell, K.H., DeWard, A., Fox, C.A., and Weinreich, M. (2006). A Dbf4p BRCA1 C-terminal-like domain required for the response to replication fork arrest in budding yeast. Genetics *173*:541-555.

Glover, J.N. (2006). Insights into the molecular basis of human hereditary breast cancer from studies of the BRCA1 BRCT domain. Fam Cancer *5*:89-93.

Glover, J.N., Williams, R.S., and Lee, M.S. (2004). Interactions between BRCT repeats and phosphoproteins: tangled up in two. Trends Biochem Sci *29*:579-585.

Goddard, T.D., and Kneller, D.G. SPARKY 3. Unversity of California, San Francisco.

Gotoh, H., Zhang, Y., Dallo, S.F., Hong, S., Kasaraneni, N., and Weitao, T. (2008). Pseudomonas aeruginosa, under DNA replication inhibition, tends to form biofilms via Arr. Res Microbiol *159*:294-302.

Gulick, A.M., Horswill, A.R., Thoden, J.B., Escalante-Semerena, J.C., and Rayment, I. (2002). Pentaerythritol propoxylate: a new crystallization agent and cryoprotectant induces crystal growth of 2-methylcitrate dehydratase. Acta Crystallogr D Biol Crystallogr *58*:306-309.

Hardy, C.F., and Pautz, A. (1996). A novel role for Cdc5p in DNA replication. Mol Cell Biol *16*:6775-6782.

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, L.H. (1973). Three additional genes required for deoxyribonucleic acid synthesis in Saccharomyces cerevisiae. J Bacteriol *115*:966-974.

Hartwell, L.H. (1974). Saccharomyces cerevisiae cell cycle. Bacteriol Rev *38*:164-198.

Heffernan, T.P., Unsal-Kacmaz, K., Heinloth, A.N., Simpson, D.A., Paules, R.S., Sancar, A., Cordeiro-Stone, M., and Kaufmann, W.K. (2007). Cdc7-Dbf4 and the human S checkpoint response to UVC. J Biol Chem *282*:9458-9468.

Hendrickson, W.A., Horton, J.R., and LeMaster, D.M. (1990). Selenomethionyl proteins produced for analysis by multiwavelength anomalous diffraction (MAD): a vehicle for direct determination of three-dimensional structure. EMBO J *9*:1665-1672.

Hoang, M.L., Leon, R.P., Pessoa-Brandao, L., Hunt, S., Raghuraman, M.K., Fangman, W.L., Brewer, B.J., and Sclafani, R.A. (2007). Structural changes in Mcm5 protein bypass Cdc7-Dbf4 function and reduce replication origin efficiency in Saccharomyces cerevisiae. Mol Cell Biol *27*:7594-7602.

Hollingsworth, R.E., Jr., Ostroff, R.M., Klein, M.B., Niswander, L.A., and Sclafani, R.A. (1992). Molecular genetic studies of the Cdc7 protein kinase and induced mutagenesis in yeast. Genetics *132*:53-62.

Hughes, S., Elustondo, F., Di Fonzo, A., Leroux, F.G., Wong, A.C., Snijders, A.P., Matthews, S.J., and Cherepanov, P. (2012). Crystal structure of human CDC7 kinase in complex with its activator DBF4. Nat Struct Mol Biol *19*:1101-1107.

Hughes, S., Jenkins, V., Dar, M.J., Engelman, A., and Cherepanov, P. (2010). Transcriptional co-activator LEDGF interacts with Cdc7-activator of S-phase kinase (ASK) and stimulates its enzymatic activity. J Biol Chem *285*:541-554.

Ilves, I., Petojevic, T., Pesavento, J.J., and Botchan, M.R. (2010). Activation of the MCM2-7 helicase by association with Cdc45 and GINS proteins. Mol Cell *37*:247-258.

Jackson, A.L., Pahl, P.M., Harrison, K., Rosamond, J., and Sclafani, R.A. (1993). Cell cycle regulation of the yeast Cdc7 protein kinase by association with the Dbf4 protein. Mol Cell Biol *13*:2899-2908.

Jacob, F., Brenner, S., and Cuzin, F. (1963). On the regulation of DNA replication in bacteria. Cold Spring Harbor Symp. Quant. Biol. *28*:329-348.

Jiang, Y.W., and Kang, C.M. (2003). Induction of S. cerevisiae filamentous differentiation by slowed DNA synthesis involves Mec1, Rad53 and Swe1 checkpoint proteins. Mol Biol Cell *14*:5116-5124.

Johnston, L.H., and Thomas, A.P. (1982). A further two mutants defective in initiation of the S phase in the yeast Saccharomyces cerevisiae. Mol Gen Genet *186*:445-448.

Jones, D.R., Prasad, A.A., Chan, P.K., and Duncker, B.P. (2010). The Dbf4 motif C zinc finger promotes DNA replication and mediates resistance to genotoxic stress. Cell Cycle *9*:2018-2026.

Kamimura, Y., Tak, Y.S., Sugino, A., and Araki, H. (2001). Sld3, which interacts with Cdc45 (Sld4), functions for chromosomal DNA replication in Saccharomyces cerevisiae. EMBO J *20*:2097-2107.

Kihara, M., Nakai, W., Asano, S., Suzuki, A., Kitada, K., Kawasaki, Y., Johnston, L.H., 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*:35051-35062.

Kilbey, B.J. (1986). cdc7 alleles and the control of induced mutagenesis in yeast. Mutagenesis *1*:29-31.

Kim, J.M., Yamada, M., and Masai, H. (2003). Functions of mammalian Cdc7 kinase in initiation/monitoring of DNA replication and development. Mutat Res *532*:29-40.

Kitada, K., Johnston, L.H., Sugino, T., and Sugino, A. (1992). Temperature-sensitive cdc7 mutations of Saccharomyces cerevisiae are suppressed by the DBF4 gene, which is required for the G1/S cell cycle transition. Genetics *131*:21-29.

Kobayashi, M., AB, E., Bonvin, A.M.J.J., and Siegal, G. (2009). Structure of the DNA-bound BRCA1 C-terminal Region from Human Replication Factor C p140 and Model of the Protein-DNA Complex. J Biol Chem *285*:10087-10097.

Koepp, D.M., Harper, J.W., and Elledge, S.J. (1999). How the cyclin became a cyclin: regulated proteolysis in the cell cycle. Cell *97*:431-434.

Kosik, K.S. (1992). Alzheimer's disease: a cell biological perspective. Science *256*:780-783.

Labib, K. (2010). How do Cdc7 and cyclin-dependent kinases trigger the initiation of chromosome replication in eukaryotic cells? Genes Dev *24*:1208-1219.

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*:3554-3561.

Lalle, M., Camerini, S., Cecchetti, S., Sayadi, A., Crescenzi, M., and Pozio, E. (2012). Interaction network of the 14-3-3 protein in the ancient protozoan parasite Giardia duodenalis. J Proteome Res *11*:2666-2683.

LeBowitz, J.H., and McMacken, R. (1986). The Escherichia coli dnaB replication protein is a DNA helicase. J Biol Chem *261*:4738-4748.

Lee, H., Yuan, C., Hammet, A., Mahajan, A., Chen, E.S.W., Wu, M.R., Su, M.I., Heierhorst, J., and Tsai, M.D. (2008). Diphosphothreonine-specific interaction between an SQ/TQ cluster and an FHA domain in the Rad53-Dun1 kinase cascade. Mol Cell *30*:767-778.

Lee, J.R., Shin, H., Choi, J., Ko, J., Kim, S., Lee, H.W., Kim, K., Rho, S.H., Lee, J.H., Song, H.E., *et al.* (2004). An intramolecular interaction between the FHA domain and a coiled coil negatively regulates the kinesin motor KIF1A. EMBO J *23*:1506-1515.

Lee, S.J., Schwartz, M.F., Duong, J.K., and Stern, D.F. (2003). Rad53 phosphorylation site clusters are important for Rad53 regulation and signaling. Mol Cell Biol *23*:6300-6314.

Lei, M., Kawasaki, Y., Young, M.R., Kihara, M., Sugino, A., and Tye, B.K. (1997). Mcm2 is a target of regulation by Cdc7-Dbf4 during the initiation of DNA synthesis. Genes Dev *11*:3365-3374.

Leung, C.C., and Glover, J.N. (2011). BRCT domains: easy as one, two, three. Cell Cycle *10*:2461-2470.

Li, H., Byeon, I.J., Ju, Y., and Tsai, M.D. (2004). Structure of human Ki67 FHA domain and its binding to a phosphoprotein fragment from hNIFK reveal unique recognition sites and new views to the structural basis of FHA domain functions. J Mol Biol *335*:371-381.

Li, J., Williams, B.L., Haire, L.F., Goldberg, M., Wilker, E., Durocher, D., Yaffe, M.B., Jackson, S.P., and Smerdon, S.J. (2002). Structural and functional versatility of the FHA domain in DNA-damage signaling by the tumor suppressor kinase Chk2. Mol Cell *9*:1045-1054.

Li, X., Liu, K., Li, F., Wang, J., Huang, H., Wu, J., and Shi, Y. (2012). Structure of C-terminal tandem BRCT repeats of Rtt107 protein reveals critical role in interaction with phosphorylated histone H2A during DNA damage repair. J Biol Chem *287*:9137-9146.

Liao, H., Yuan, C., Su, M.I., Yongkiettrakul, S., Qin, D., Li, H., Byeon, I.J., Pei, D., and Tsai, M.D. (2000). Structure of the FHA1 domain of yeast Rad53 and identification of binding sites for both FHA1 and its target protein Rad9. J Mol Biol *304*:941-951.

Lilley, C.E., Schwartz, R.A., and Weitzman, M.D. (2007). Using or abusing: viruses and the cellular DNA damage response. Trends Microbiol *15*:119-126.

Lloyd, J., Chapman, J.R., Clapperton, J.A., Haire, L.F., Hartsuiker, E., Li, J., Carr, A.M., Jackson, S.P., and Smerdon, S.J. (2009). A supramodular FHA/BRCT-repeat architecture mediates Nbs1 adaptor function in response to DNA damage. Cell *139*:100-111.

Lopes, M., Cotta-Ramusino, C., Pellicioli, A., Liberi, G., Plevani, P., Muzi-Falconi, M., Newlon, C.S., and Foiani, M. (2001). The DNA replication checkpoint response stabilizes stalled replication forks. Nature *412*:557-561.

Lopez-Mosqueda, J., Maas, N.L., Jonsson, Z.O., Defazio-Eli, L.G., Wohlschlegel, J., and Toczyski, D.P. (2010). Damage-induced phosphorylation of Sld3 is important to block late origin firing. Nature *467*:479-483.

Lorenz, M.C., Bender, J.A., and Fink, G.R. (2004). Transcriptional response of Candida albicans upon internalization by macrophages. Eukaryot Cell *3*:1076-1087.

Lorincz, A.T., Reid, R., Jenson, A.B., Greenberg, M.D., Lancaster, W., and Kurman, R.J. (1992). Human papillomavirus infection of the cervix: relative risk associations of 15 common anogenital types. Obstet Gynecol *79*:328-337.

Lucca, C., Vanoli, F., Cotta-Ramusino, C., Pellicioli, A., Liberi, G., Haber, J., and Foiani, M. (2004). Checkpoint-mediated control of replisome-fork association and signalling in response to replication pausing. Oncogene *23*:1206-1213.

Manke, I.A., Lowery, D.M., Nguyen, A., and Yaffe, M.B. (2003). BRCT repeats as phosphopeptide-binding modules involved in protein targeting. Science *302*:636-639.

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*:4805-4814.

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*:28071-28081.

Marston, A.L. (2009). Meiosis: DDK is not just for replication. Curr Biol *19*:R74-76.

Martin-Garcia, R., de Leon, N., Sharifmoghadam, M.R., Curto, M.A., Hoya, M., Bustos-Sanmamed, P., and Valdivieso, M.H. (2011). The FN3 and BRCT motifs in the exomer componenet Chs5p define a conserved module that is necessary and sufficient for its function. Cell Mol Life Sci *68*:2907-2917.

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*:228-232.

Matos, J., Lipp, J.J., 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*:662-678.

Matthews, L.A., Duong, A., Prasad, A.A., Duncker, B.P., and Guarne, A. (2009). Crystallization and preliminary X-ray diffraction analysis of motif N from Saccharomyces cerevisiae Dbf4. Acta Crystallogr Sect F Struct Biol Cryst Commun *65*:890-894.

Matthews, L.A., Jones, D.R., Prasad, A.A., Duncker, B.P., and Guarne, A. (2012). Saccharomyces cerevisiae Dbf4 has unique fold necessary for interaction with Rad53 kinase. J Biol Chem *287*:2378-2387.

McCarty, M. (2003). Discovering genes are made of DNA. Nature *421*:406.

McCoy, A.J., Grosse-Kunstleve, R.W., Adams, P.D., Winn, M.D., Storoni, L.C., and Read, R.J. (2007). Phaser crystallographic software. J Appl Crystallogr *40*:658-674.

Mechali, M. (2010). Eukaryotic DNA replication origins: many choices for appropriate answers. Nat Rev Mol Cell Biol *11*:728-738.

Meselson, M., and Stahl, F.W. (1958). The replication of DNA. Cold Spring Harb Symp Quant Biol *23*:9-12.

Messer, W. (2002). The bacterial replication initiator DnaA. DnaA and oriC, the bacterial mode to initiate DNA replication. FEMS Microbiol Rev *26*:355-374.

Miescher, F. (1871). Ueber die chemische Zusammensetzung der Eiterzellen. Med.-Chem. Unters *4*:441-460.

Miescher, F. (1874). Die Spermatozoen einiger Wirbelthiere. Ein Beitrag zur Histochemie. Verh. Nat. forsch. Ges. Basel *6*:138-208.

Mikol, V., Rodeau, J.L., and Giege, R. (1989). Changes in pH during biomolecular crystallization by vapor diffusion using ammonium sulfate as the precipitant. Journal of Applied Crystallography *22*:155-161.

Miller, C.T., Gabrielse, C., Chen, Y.C., and Weinreich, M. (2009). Cdc7p-dbf4p regulates mitotic exit by inhibiting Polo kinase. PLoS Genet. *5*:e1000498.

Mirkin, E.V., and Mirkin, S.M. (2007). Replication fork stalling at natural impediments. Microbiol Mol Biol Rev *71*:13-35.

Mohammad, D.H., and Yaffe, M.B. (2009). 14-3-3 proteins, FHA domains and BRCT domains in the DNA damage response. DNA repair *8*:1009-1017.

Montagnoli, A., Bosotti, R., Villa, F., Rialland, M., Brotherton, D., Mercurio, C., Berthelsen, J., and Santocanale, C. (2002). Drf1, a novel regulatory subunit for human Cdc7 kinase. EMBO J *21*:3171-3181.

Morgan, D.O. (1997). Cyclin-dependent kinases: engines, clocks, and microprocessors. Annu Rev Cell Dev Biol *13*:261-291.

Muramatsu, S., Hirai, K., Tak, Y.S., Kamimura, Y., and Araki, H. (2010). CDK-dependent complex formation between replication proteins Dpb11, Sld2, Pol epsilon, and GINS in budding yeast. Genes Dev *24*:602-612.

Nakamura, T., Nakamura-Kubo, M., and Shimoda, C. (2002). Novel fission yeast Cdc7-Dbf4-like kinase complex required for the initiation and progression of meiotic second division. Mol Cell Biol *22*:309-320.

Nambiar, S., Mirmohammadsadegh, A., Hassan, M., Mota, R., Marini, A., Alaoui, A., Tannapfel, A., Hegemann, J.H., and Hengge, U.R. (2007). Identification and functional characterization of ASK/Dbf4, a novel cell survival gene in cutaneous melanoma with prognostic relevance. Carcinogenesis *28*:2501-2510.

Navadgi-Patil, V.M., and Burgers, P.M. (2009). A tale of two tails: activation of DNA damage checkpoint kinase Mec1/ATR by the 9-1-1 clamp and by Dpb11/TopBP1. DNA Repair (Amst) *8*:996-1003.

Newman, J. (2005). Expanding screening space through the use of alternative reservoirs in vapor-diffusion experiments. Acta Crystallogr D Biol Crystallogr *61*:490-493.

Njagi, G.D., and Kilbey, B.J. (1982). cdc7-1 a temperature sensitive cell-cycle mutant which interferes with induced mutagenesis in Saccharomyces cerevisiae. Mol Gen Genet *186*:478-481.

Nott, T.J., Kelly, G., Stach, L., Li, J., Westcott, S., Patel, D., Hunt, D.M., Howell, S., Buxton, R.S., O'Hare, H.M., and Smerdon, S.J. (2009). An intramolecular switch regulates phosphoindependent FHA domain interactions in Mycobacterium tuberculosis. Sci Signal *2*:ra12.

Nougarede, R., Della Seta, F., Zarzov, P., and Schwob, E. (2000). Hierarchy of S-phase-promoting factors: yeast Dbf4-Cdc7 kinase requires prior S-phase cyclin-dependent kinase activation. Mol Cell Biol *20*:3795-3806.

Oehler, S., and Muller-Hill, B. (2010). High local concentration: a fundamental strategy of life. J Mol Biol *395*:242-253.

Ogi, H., Wang, C.Z., 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*:31376-31387.

Otwinowski, Z., and Minor, W. (1997). Processing of X-ray Diffraction Data Collected in Oscillation Mode. Methods Enzymol *276*:307-326.

Pasero, P., Duncker, B.P., Schwob, E., and Gasser, S.M. (1999). A role for the Cdc7 kinase regulatory subunit Dbf4p in the formation of initiation-competent origins of replication. Genes Dev *13*:2159-2176.

Patterson, M., Sclafani, R.A., Fangman, W.L., and Rosamond, J. (1986). Molecular characterization of cell cycle gene CDC7 from Saccharomyces cerevisiae. Mol Cell Biol *6*:1590-1598.

Paweletz, N. (2001). Walther Flemming: pioneer of mitosis research. Nat Rev Mol Cell Biol *2*:72-75.

Pellicioli, A., Lucca, C., Liberi, G., Marini, F., Lopes, M., Plevani, P., Romano, A., Di Fiore, P.P., and Foiani, M. (1999). Activiation of Rad53 kinase in response to DNA damage and its effect in modulating phosphorylation of the lagging strand DNA polymerase. EMBO J *18*:6561-6572.

Pellicioli, A., and Foiani, M. (2005). Signal transduction: how rad53 kinase is activated. Curr Biol *15*:R769-771.

Pennell, S., Westcott, S., Ortiz-Lombardia, M., Patel, D., Li, J., Nott, T.J., Mohammed, D., Buxton, R.S., Yaffe, M.B., Verma, C., and Smerdon, S.J. (2010). Structural and functional analysis of phosphothreonine-dependent FHA domain interactions. Structure *18*:1587-1595.

Pessoa-Brandao, L., and Sclafani, R.A. (2004). CDC7/DBF4 functions in the translesion synthesis branch of the RAD6 epistasis group in Saccharomyces cerevisiae. Genetics *167*:1597-1610.

Pike, B.L., 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*:39636-39644.

Powell, S.N., and Kachnic, L.A. (2003). Roles of BRCA1 and BRCA2 in homologous recombination, DNA replication fidelity and the cellular response to ionizing radiation. Oncogene *22*:5784-5791.

Randell, J.C., Fan, A., Chan, C., Francis, L.I., Heller, R.C., Galani, K., and Bell, S.P. (2010). Mec1 is one of multiple kinases that prime the Mcm2-7 helicase for phosphorylation by Cdc7. Mol Cell *40*:353-363.

Rappas, M., Oliver, A.W., and Pearl, L.H. (2011). Structure and function of the Rad9-binding region of the DNA-damage checkpoint adaptor TopBP1. Nucleic Acids Res *39*:313-324.

Remus, D., Beuron, F., Tolun, G., Griffith, J.D., Morris, E.P., and Diffley, J.F. (2009). Concerted loading of Mcm2-7 double hexamers around DNA during DNA replication origin licensing. Cell *139*:719-730.

Santocanale, C., and Diffley, J.F. (1998). A Mec1- and Rad53-dependent checkpoint controls late-firing origins of DNA replication. Nature *395*:615-618.

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*:451-463.

Schleiden, M. (1838). J. Arch. Anat. Physiol. Wiss. Med. *13*:137-176.

Schwann, T. (1839). mikroskopische untersuchungen über die übereinstimmung in der struktur und dem wachstum der tiere and pflanzen (Berlin: Sander'shen Buchhandlung).

Schwartz, M.F., Lee, S.J., Duong, J.K., Eminaga, S., and Stern, D.F. (2003). FHA domain-mediated DNA checkpoint regulation of Rad53. Cell Cycle *2*:384-396.

Sclafani, R.A. (2000). Cdc7p-Dbf4p becomes famous in the cell cycle. J Cell Sci *113 (Pt 12)*:2111-2117.

Selvaratnam, R., Mazhab-Jafari, M.T., Das, R., and Melacini, G. (2012). The Auto-Inhibitory Role of the EPAC Hinge Helix as Mapped by NMR. PLoS One *7*:e48707.

Shellman, Y.G., Schauer, I.E., Oshiro, G., Dohrmann, P., and Sclafani, R.A. (1998). Oligomers of the Cdc7/Dbf4 protein kinase exist in the yeast cell. Mol Gen Genet *259*:429-436.

Sheng, Z.Z., Zhao, Y.Q., and Huang, J.F. (2011). Functional Evolution of BRCT Domains from Binding DNA to Protein. Evol Bioinform Online *7*:87-97.

Sherr, C.J. (1994). G1 phase progression: cycling on cue. Cell *79*:551-555.

Sheu, Y.J., and Stillman, B. (2006). Cdc7-Dbf4 phosphorylates MCM proteins via a docking site-mediated mechanism to promote S phase progression. Mol Cell *24*:101-113.

Sheu, Y.J., and Stillman, B. (2010). The Dbf4-Cdc7 kinase promotes S phase by alleviating an inhibitory activity in Mcm4. Nature *463*:113-117.

Shi, Q.M., Wang, Y.M., Zheng, X.D., Lee, R.T., and Wang, Y. (2007). Critical role of DNA checkpoints in mediating genotoxic-stress-induced filamentous growth in Candida albicans. Mol Biol Cell *18*:815-826.

Shirahige, K., Hori, Y., Shiraishi, K., Yamashita, M., Takahashi, K., Obuse, C., Tsurimoto, T., and Yoshikawa, H. (1998). Regulation of DNA-replication origins during cell-cycle progression. Nature *395*:618-621.

Spardy, N., Covella, K., Cha, E., Hoskins, E.E., Wells, S.I., Duensing, A., and Duensing, S. (2009). Human papillomavirus 16 E7 oncoprotein attenuates DNA damage checkpoint control by increasing the proteolytic turnover of claspin. Cancer Res *69*:7022-7029.

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*:965-971.

Stead, B.E., Brandl, C.J., Sandre, M.K., and Davey, M.J. (2012). Mcm2 phosphorylation and the response to replicative stress. BMC Genet *13*:36.

Sullivan, M., and Morgan, D.O. (2007). Finishing mitosis, one step at a time. Nat Rev Mol Cell Biol *8*:894-903.

Takahashi, T.S., and Walter, J.C. (2005). Cdc7-Drf1 is a developmentally regulated protein kinase required for the initiation of vertebrate DNA replication. Genes Dev *19*:2295-2300.

Tanaka, S., Tak, Y.S., and Araki, H. (2007a). The role of CDK in the initiation step of DNA replication in eukaryotes. Cell Div *2*:16.

Tanaka, S., Umemori, T., Hirai, K., Muramatsu, S., Kamimura, Y., and Araki, H. (2007b). CDK-dependent phosphorylation of Sld2 and Sld3 initiates DNA replication in budding yeast. Nature *445*:328-332.

Tong, Y., Tempel, W., Wang, H., Yamada, K., Shen, L., Senisterra, G.A., MacKenzie, F., Chishti, A.H., and Park, H.W. (2010). Phosphorylation-independent dual-site binding of the FHA domain of KIF13 mediates phosphoinositide transport via centaurin alpha1. Proc Natl Acad Sci USA *107*:20346-20351.

Tower, J. (2004). Developmental gene amplification and origin regulation. Annu Rev Genet *38*:273-304.

Varrin, A.E., Prasad, A.A., Scholz, R.P., Ramer, M.D., and Duncker, B.P. (2005). A mutation in Dbf4 motif M impairs interactions with DNA replication factors and confers increased resistance to genotoxic agents. Mol Cell Biol *25*:7494-7504.

Vaynberg, J., and Qin, J. (2006). Weak protein-protein interactions as probed by NMR spectroscopy. Trends Biotechnol *24*:22-27.

Vaziri, C., and Masai, H. (2010). Integrating DNA replication with trans-lesion synthesis via Cdc7. Cell Cycle *9*:4818-4823.

Vijayraghavan, S., and Schwacha, A. (2012). The Eukaryotic Mcm2-7 Replicative Helicase. In The Eukaryotic Replisome: a Guide to Protein Structure and Function, S. MacNeill, ed. (Dordrecht: Springer), pp. 113-134.

Vinogradova, O., and Qin, J. (2012). NMR as a unique tool in assessment and complex determination of weak protein-protein interactions. Top Curr Chem *326*:35-45.

Virchow, R. (1858). Die Cellularpathologie in ihre begrundung auf physiologische and pathologische Gewebelehre (Berlin: Hirschwald).

Waga, S., and Stillman, B. (1998). The DNA replication fork in eukaryotic cells. Annu Rev Biochem *67*:721-751.

Watson, J.D., and Crick, F.H. (1953). Genetical implications of the structure of deoxyribonucleic acid. Nature *171*:964-967.

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

Weinreich, M., and Stillman, B. (1999). Cdc7p-Dbf4p kinase binds to chromatin during S phase and is regulated by both the APC and the RAD53 checkpoint pathway. EMBO J *18:*5334-5346.

Williams, R.S., Dodson, G.E., Limbo, O., Yamada, Y., Williams, J.S., Guenther, G., Classen, S., Glover, J.N., Iwasaki, H., Russell, P., and Tainer, J.A. (2009). Nbs1 flexibly tethers Ctp1 and Mre11-Rad50 to coordinate DNA double-strand break processing and repair. Cell *139*:87-99.

Williams, R.S., Green, R., and Glover, J.N. (2001). Crystal structure of the BRCT repeat region from the breast cancer-associated protein BRCA1. Nature Structural Biology *8*:838-842.

Williams, R.S., Lee, M.S., Hau, D.D., and Glover, J.N. (2004). Structural basis of phosphopeptide recognition by the BRCT domain of BRCA1. Nat Struct Mol Biol *11*:519-525.

Wogan, G.N., Hecht, S.S., Felton, J.S., Conney, A.H., and Loeb, L.A. (2004). Environmental and chemical carcinogenesis. Semin Cancer Biol *14*:473-486.

Yongkiettrakul, S., Byeon, I.J., and Tsai, M.D. (2004). The ligand specificity of yeast Rad53 FHA domains at the +3 position is determined by nonconserved residues. Biochemistry *43*:3862-3869.

Yoshizawa-Sugata, N., Ishii, A., Taniyama, C., Matsui, E., Arai, K., and Masai, H. (2005). A second human Dbf4/ASK-related protein, Drf1/ASKL1, is required for efficient progression of S and M phases. J Biol Chem *280*:13062-13070.

Yurov, Y.B., Vorsanova, S.G., and Iourov, I.Y. (2011). The DNA replication stress hypothesis of Alzheimer's disease. ScientificWorldJournal *11*:2602-2612.

Zegerman, P., and Diffley, J.F. (2007). Phosphorylation of Sld2 and Sld3 by cyclin-dependent kinases promotes DNA replication in budding yeast. Nature *445*:281-285.

Zegerman, P., and Diffley, J.F. (2010). Checkpoint-dependent inhibition of DNA replication initiation by Sld3 and Dbf4 phosphorylation. Nature *467*:474-478.

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