Characterization of Novel FHA Domain Ligand Binding Motifs

by

Aaron Robertson

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Master of Science
in
Biology

Waterloo, Ontario, Canada, 2016

Author's Declaration

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.
I understand that my thesis may be made electronically available to the public.

Acknowledgements

I would like to thank my supervisor, Dr. Bernard Duncker, for providing me with a chance to work in the lab. I've learned so much during my time here and I am thankful for the opportunity to learn so much. I would also like to thank my committee members, Dr. Christine Dupont and Dr. Brendan McConkey for their valuable feedback on my project, as well as for the time taken to evaluate my thesis work. Special thanks to Dr. McConkey for answering all my incessant questions about bioinformatics.

Thank you to Jeremy Adams for helping me with several bioinformatics projects and for taking the time to always explain things to me. A big thank you to Darryl Jones and Evelyn Suman for helping me immensely these past two years. Special thanks to Darryl for starting the project that I had the pleasure of working on. Thank you to all the former and current Duncker lab members who have helped me during my time at UW. Thank you to all my friends (including labmates and lab neighbours) for an enjoyable grad school experience. I will miss our afternoon gatherings in the office, so thank you for keeping me social. I'd also like to thank the Guarné lab, specifically Ahmad Almawi for the creation of several mutants. Last but not least, I want to thank my family for always being there for me in my times of need and for continually supporting and encouraging me to pursue my interests.

Abstract

Rad53, a key checkpoint protein in *Saccharomyces cerevisiae*, has a central catalytic kinase domain as well as two forkhead-associated (FHA) domains, FHA1 and FHA2. Along with the extensively studied phosphothreonine binding pocket, characteristic of all FHA domains, the Duncker lab in collaboration with Alba Guarné's lab, recently reported that FHA1, but not FHA2, has an additional conserved region on one of its lateral surfaces. Mutation of this novel FHA1 motif abrogated the interaction between Rad53 and Dbf4 (the regulatory subunit of the Dbf4-Cdc7 kinase complex), and rendered cells sensitive to genotoxic stress. When corresponding residues are mutated in Dun1 and Mek1 (other *S. cerevisiae* FHA domains), affinities for the Rad53 FHA1 ligand Dbf4 can be altered, highlighting the importance of the region. Additionally, structural and biochemical data has shown that several previously unidentified residues on Dbf4 are important for the interaction with the lateral surface of FHA1.

I have since investigated whether this non-canonical FHA1 interaction surface mediates binding to other Rad53 ligands, and have found that it has a variable requirement. FHA1 has numerous binding partners in the cell, each with distinct cell cycle progression and checkpoint roles. While the same FHA1 lateral domain mutants with a disrupted Dbf4 interaction retained wild type affinity for Sgs1 and Cdc7, an intermediate effect was observed with Sld3. Taking this data as a whole, it shows that this lateral region of FHA1 is specific for mediating the interaction with Dbf4, and mutations of this region do not disrupt binding to other FHA1 ligands.

Through bioinformatics analysis, several novel conserved regions in numerous other proteins with FHA domains in *S. cerevisiae*, including Dun1 and Mek1, have been identified. Similar to Rad53 FHA1, mutation of conserved residues in the novel Dun FHA motif disrupted its interaction with Sml1, a Dun1 ligand. Furthermore, these mutants showed wild type levels of

interaction when interacting with the known Dun1 ligand Rad53. Spotting plate assays were conducted with two genotoxic agents: hydroxyurea (HU) and methyl methanesulfate (MMS). A moderate increased sensitivity was seen on the HU plates for the Dun1 FHA lateral surface mutants, however cells were not completely sensitive. This work highlights the importance of newly identified FHA domain motifs in mediating non-canonical protein-protein interactions.

Table of Contents

Author's Declaration ii
Abstract iii
Acknowledgements iv
Table of Contents vi
List of Abbreviations/Acronyms viii
List of Figures and Tables. ix

Chapter 1: Introduction 1
1.1 Background 2
1.1.1 Yeast as a Model System 2
1.1.2 Yeast Cell Cycle 4
1.2 DNA Replication and the Intra S-Phase Checkpoint 8
1.2.1 Initiation of DNA Replication 8
1.2.2 Replication Stress and Rad53 11
1.3 FHA Domains 17
1.3.1 Structure and Function of FHA domains 17
1.3.2 FHA Domains and their ligands in *S. cerevisiae* 19
1.4 Research Objectives 21

Chapter 2: Materials and Methods 23
2.1 Yeast Strains 24
2.2 Genomic DNA Preparation 24
2.3 Plasmid Construction 25
2.4 Yeast Transformation 26
2.5 Yeast Two-hybrid Assay 27
2.6 Whole Cell Extract Preparations and Western Blotting. 29
2.7 Spotting Plate Assays 30
2.8 Co-immunipreciptation 31
2.9 Bioinformatics Analysis 32

Chapter 3: Characterization of the importance of the lateral surface of Rad53 FHA1 in ligand specificity 34
3.1 Introduction 35
3.2 Results 37
3.2.1 FHA Interactions with DDK 37
3.2.2 Characterizing the FHA1-Sgs1 interaction 46
3.3.3 Characterizing the FHA1-Sld3 interaction 48
3.3 Discussion 49

Chapter 4: Characterization of the importance of the lateral surface of various additional *S. cerevisiae* FHA domains 52
4.1 Introduction 53
4.2 Results 54
4.2.1 The role of conserved Lateral Surfaces in Ligand Specificity for additional FHA domains 54
4.2.2 Lateral Surfaces on Other FHA Domains 61
4.3 Discussion 69

Chapter 5: General Discussion 74
5.1 Rad53 has a lateral surface patch that has a variable requirement or ligand binding 75
5.2 Lateral surfaces on *S. cerevisiae* FHA domains 77
5.3 Relation to Cancer 78
5.4 Future Directions 80
5.5 Significance 81

References 83

Appendix A: Chapter 3 Supplementary Material 91

Appendix B: Chapter 4 Supplementary Material 93

List of Abbreviations and Acronyms

APC/C:Anaphase promoting complex/cyclosome
ARS: Autonomously replicating sequence
Cdc: Cell division cycle
CDK: Cyclin dependent kinase
CMG: Cdc45, Mcm2-7, GINS
Co-IP: Co-immunoprecipitation
DDK: Dbf4-dependent kinase
DNA: Deoxyribonucleic acid
dNTP: Deoxynucleotide phosphate
DTT: Dithiothreitol
EDTA: Ethylenediaminetetraacetic acid
FACS: Fluorescence activated cell sorting
FL: Full Length
GAL/RAFF: Galactose/Raffinose
GINS: Go-Ichi-Ni-San
HA: Hemagglutinin
HU: Hydroxyurea
Kb: Kilobase
LEU: Leucine
MCM: Minichromosome maintenance
MMS: Methyl methanesulfonate
NLS: Nuclear Localization Signal
ONPG: 2-Nitrophenyl-β-D-galactopyranoside
ORC: Origin recognition complex
PCNA: Proliferating cell nuclear antigen
PCR: Polymerase chain reaction
PMSF: Phenylmethylsulphonyl fluoride
Pre-IC: Pre-initiation complex
Pre-LC: Pre-loading complex
Pre-RC: Pre-replicative complex
RPA: Replication protein A
SC: Synthetic complete
SDS: Sodium dodecyl sulfate
TRP: Tryptophan
URA: Uracil
WCE: Whole cell extract
WT: Wild-type
YPD: Yeast extract, peptone, dextrose

List of Figures and Tables

Figure 1: Cell cycle stages in *Saccharomyces cerevisiae* 8

Figure 2: Proteins involved in the initiation of DNA replication in *S. cerevisiae* 9

Figure 3: Triggering of Intra-S Phase Checkpoint via Rad53 12

Figure 4: Computational surface mapping of Rad53 FHA domains 14

Figure 5: A lateral patch on Rad53 FHA1 mediates the interaction with Dbf4 15

Figure 6: The rad53-NEVF mutant is sensitive to genotoxic stress 16

Figure 7: Checkpoint control of S phase transcription and dNTP pools 20

Figure 8: The yeast two-hybrid assay detects protein interactions 27

Figure 9: Sequence Alignment of Various FHA domains 39

Figure 10: Phylogenetic tree of Various FHA domains 41

Figure 11: FHA1 Lateral surface mutants fail to abrogate the interaction with Cdc7 42

Figure 12: T484 helps mediate the Rad53 FHA1-Cdc7 interaction 43

Figure 13: The FHA1-Dbf4 interaction is mediated by two contact points 45

Figure 14: Yeast two-hybrid analysis of Sgs1 greater helicase region 46

Figure 15: FHA1 lateral surface mutants do not mediate the FHA1-Sgs1 interaction 47

Figure 16: Co-immunopreciptation of Rad53 FHA1 and Sld3 48

Figure 17**:** Structural alignment of Dun1 FHA and Mek1 FHA against Rad53 FHA1 56

Figure 18: Dun1 FHA R124L leads to no change in Dbf4 Binding 58

Figure 19: Mek1 FHA V122E leads to increased association with Dbf4 60

Figure 20: Several FHA domains have conserved lateral patches 61

Figure 21: The Lateral Surface of Dun1 FHA mediates the interaction with Sml1 63

Figure 22: The Lateral Surface of Dun1 FHA is not required for interaction with Rad53 65

Figure 23: Dun1 FHA Mutants have differential sensitivities to DNA damage agents 68

Table 1: Antibodies used for immunoblotting 30

Table 2: Blossum 62 scores for three FHA domains when compared to the FHA1 lateral patch residues 55

Table 3: Blosum 62 scores for Dun1 and Mek1 when compared to the lateral patch residues of Rad53 FHA1 57

Chapter 1: Introduction

Portions of this chapter appear in the following journal article and are reproduced with permission. This research was originally published in The Journal of Biological Chemistry.

Matthews, L., Selvaratnam, R., Jones, D., Akimoto, M., McConkey, B., Melacini, G., Duncker, B.P., Guarné A. (2014). A Novel Non-canonical Forkhead-associated (FHA) Domain-binding Interface Mediates the Interaction between Rad53 and Dbf4 Proteins. *The Journal of Biological Chemistry*, *289*, 2589-2599.

1.1 Background

1.1.1 Yeast as a Model System

There are many model organisms today used today in laboratories, and a particularly prevalent type are microorganisms. Microorganisms have extremely high growth rates, are easy to maintain in the lab, and are easy to genetically manipulate (Mokdad-Gargouri *et al.,* 2012). Initially, common model microorganisms such as bacteria such as *Escherichia coli*, were among the first to have a completely sequenced genome (Blattner *et al.,* 1997). When selecting the model organism for study, one has to consider the applications of one's research. When studying DNA replication, if one wants to generalize results found to higher eukaryotes such as humans, the model organism in question must have similar cell cycle pathways and conserved proteins. This is the main reason why microorganisms like *E. coli* were not considered for the purpose of this study. Prokaryotes lack the ability to process proteins in a similar manner to eukaryotes, and this leads to the build-up of inclusion bodies, with proteins that no longer function properly or have failed to adopt their correct three-dimensional structure during experiments featuring recombinant proteins (Porro *et al.,* 2005).

In this project, the yeast *Saccharomyces cerevisiae* is used as a model system, as it is easy to work with, manipulate and culture. *S. cerevisiae* combines the aforementioned advantages of a microorganism, in addition to possessing the proper machinery to fold and assemble eukaryotic proteins. It requires inexpensive media and an incubator, which is cheap compared to the cost of growing animals in a laboratory setting (Mokdad-Gargouri *et al.,* 2012). In most recombinant protein expression systems, researchers strive to isolate the protein in the foreign system. This usually involves a protocol where the cells need to lysed, and *S. cerevisiae* has no native oncogenic DNA or endotoxins (Pereira *et al.,* 2012). Researchers can be confident that working

with *S. cerevisiae* is safe, a fact the U.S. Food and Drug Administration affirmed when they gave *S. cerevisiae* a rating of GRAS (generally regarded as safe) (Porro *et al.,* 2005).

As a eukaryote, another advantage of *S. cerevisiae* is the fact that most major cellular processes are conserved throughout all eukaryotes. In a 2002 study, it was shown that over 40% of yeast proteins share an amino acid sequence with at least one human protein (Hughes, 2002). Due to the similarities between yeast and human proteins, many human cellular processes are first studied in yeasts, particularly *S. cerevisiae.* An example of such modeling is cell-cycle regulation, which was first studied in yeast, and then applied to the study of human cancer biology (Pereira *et al.,* 2012). Other processes that are better conserved in *S. cerevisiae* than prokaryotes when compared to humans are cell division, protein folding, metabolism and intracellular transport (Fields and Johnston, 2005). *S. cerevisiae* has played an important role in understanding human metabolic conditions such as Mohr-Tranebjaerg syndrome, Friedreich ataxia and Leigh syndrome by allowing the study of analogous proteins in yeasts (TIM8, YFH1, SDHA respectively) (Rea *et al.,* 2010). Although this work consists of basic research, the long-term goal for this project is to use the knowledge gained from the *S. cerevisiae* FHA domains to develop potential therapies for a wide variety of human disorders, such as Li-Fraumani syndrome in addition to lung and colon cancers. By selecting *S. cerevisiae,* an organism that shares numerous conserved pathways and analogous proteins with humans, there will be fewer problems when trying to apply this research to potential therapies, a trait common to common eukaryotic model organisms.

Budding yeast possess a number of qualities that make them extremely adaptable to a number of environments, including low temperature and high salt (reviewed in Walker, 1998; Turk *et al.*, 2011). Metabolically, *S. cerevisiae* is a facultative anaerobe that can meet its cellular ATP

demands by fermentation of glucose to ethanol, allowing cells to survive in the presence or absence of oxygen (Altmann *et al.,* 2007). The fact that two genetic forms of *S. cerevisiae* exist (haploid and diploid) allows *S. cerevisiae* to survive the addition or deletion of gene copies (Altmann *et al.,* 2007). The strong survivability that *S. cerevisiae* displays means that genes of interest, can be very easily inserted, deleted or modified. The distinction between *S. cerevisiae* and other model organisms in this regard is that the added genes can start being expressed as proteins significantly faster than other model organisms (Mokdad-Gargouri *et al.,* 2012). In addition to being extremely easy to manipulate on a molecular level, *S. cerevisiae* became an ideal model organism because it was the first eukaryotic genome to be completely sequenced (Goffeau *et al.,* 1996). Furthermore, there are numerous deletion and over-expression yeast strain collections that have been developed through years of research that exist to study genes, processes and pathways (Mojca *et al.,* 2012). Collections like the Yeast Deletion Project are publically accessible as databases that list all *S. cerevisiae* proteins in addition to the mutant *S. cerevisiae* strains generated by deletion and addition studies (http://www-sequence.stanford.edu/group/yeast_deletion_project/data_sets.html). This massive body of knowledge generated by previous experiments, gives researchers an incredible tool, as thousands of mutants are readily available for laboratory use. With the development of these collections, *S. cerevisiae* has become the eukaryotic model of choice when studying cell cycle regulation.

1.1.2 Yeast Cell Cycle

The cell cycle of budding yeast consists of four main stages that are highly conserved: G1 (Gap1) phase, S (Synthesis of DNA) phase, G2 (Gap2) Phase and M (Mitosis) phase (shown in Figure 1). The cell cycle timings and transitions between phases is regulated by a highly conserved regulatory network that is present from yeast to humans. One of the main checkpoints

is the START checkpoint that occurs during late G1 phase. At this point several internal and external cell signals are evaluated such as cell size, presence of nutrients and the existence of mating signals (Pringle & Hartwell, 1981). Cells have to be at a suitable size for cell division in order for two viable cells to arise from the parent. Nutrient availability is an important environmental indicator of whether it is a suitable time to divide. Lastly, the presence of a mating signals such as alpha-factor would facilitate the start of pairing of yeasts of different mating types. In this case, cells do not continue with the mitotic cycle and instead undergo sexual reproduction with cells of the opposite mating type. It is critical that these factors are evaluated in order to determine whether the conditions are appropriate to divide. At this point cells either commit to entering another round of the cell cycle (mitosis) or choose to undergo an alternative form of development such as mating, sporulation or entering into stationary phase (G0). G0 phase is an extended version of G1 where cells are not getting ready to divide. These cells are quiescent and will re-enter the cell cycle when conditions become favorable for cells to divide again. Progression within the cell cycle is highly dependent on a group of kinases called cyclin dependent kinases (CDKs). The five CDKs that are responsible for regulating the cell cycle in budding yeast are Cdc28, Pho85, Kin28, Ssn3 and Ctk1. Cdc28 is considered the most important CDK in regulating cell cycle timings with Pho85 having a lesser importance (Mendenhall and Hodge, 1998). The three other CDKs play a role in regulating levels of transcription of RNA polymerases. Although Cdc28 has numerous binding partners in the cell, the ligands that are responsible for its core catalytic activity are the G1 cyclins (Cln1-3) and the B type cyclins (Clb1-6). The variable levels of cyclins and their bound CDKs is what regulates progression throughout the cell cycle. Although cyclin expression levels vary throughout the cell cycle, most have a D or KEN box motif which is recognized by the anaphase promoting complex (APC) for

degradation. In late G1 phase, the pre-replicative complex (pre-RC) is then formed. Low levels of active CDKs allow the pre-RC to form with active CDK levels spiking after G1 phase. This increase prevents DNA re-replication or formation of new pre-RCs (Nguyen *et al.,* 2001). CDKs also play a role in the initiation of DNA replication by phosphorylating key replication machinery components.

Once cells have progressed through G1 they enter S phase, where DNA is replicated. The process is called semiconservative replication, as each original DNA strand acts as a template for a newly synthesized DNA strand. This creates two new DNA molecules, each with one parent strand and one daughter strand. The process of DNA replication occurs in the 5' to 3' direction, with DNA being elongated from 3' OH groups. Due to the fact that DNA strands are in an antiparallel conformation, replication cannot be continuous for both strands as one strand is orientated in the 3' to 5' direction with respect to the direction of replication fork movement. The leading strand has continuous DNA synthesis in the 5' to 3' direction whereas the lagging strand has discontinuous DNA synthesis where short fragments of DNA called Okazaki fragments are made. These fragments are later joined by the enzyme DNA ligase. The helicase that is responsible for melting DNA strands and allowing replication forks to progress is the Mcm2-7 helicase. This unwinding creates topological stress in DNA which is relieved by a group of enzymes called topoisomerases, specifically topoisomerase II in yeast. The unwound DNA strands are then stabilized by DNA binding proteins (RPA in budding yeast). Next, the DNA polymerase-α synthesizes short segments of RNA complimentary to the template which prime the DNA elongation reaction. The leading strand is synthesized continuously by the DNA polymerase-ε which moves with the fork in the 5' to 3' direction. The lagging strand uses polymerase-α/primase to synthesize primers, and polymerase-δ to elongate strands, but as this

reaction is moving in the opposite direction to the replication fork, short 150 base stretches are synthesized. An endonuclease called Fen1 is responsible for replacing the RNA primers in conjunction with polymerase-δ and the short fragments are joined together by DNA ligase 1.

The final phase before mitosis is G2. In this phase, cells rapidly grow in order to have a critical size and produce all of the necessary proteins required for entry into mitosis. In G2, centrosomes which are required for mitosis form. Once cells are ready to progress to mitosis they must pass the G2/M checkpoint. This checkpoint is regulated by Cdc28 and its interaction with Clb2, which in complex is called mitosis promoting factor (MPH) and stimulates mitosis (Surana *et al.,* 1993). Mitosis is divided into five different phases that are conserved among eukaryotes: prophase, prometaphase, metaphase, anaphase and telophase. In prophase chromosomes condense, and the mitotic spindle starts to form, with the centrosomes starting to move away from each other. In prometaphase the nuclear envelope starts to fragment and microtubules start to extend into the nuclear area, with some microtubules attaching to kinetochores on chromosomes. By the end of metaphase centrosomes are at opposite poles of the cell with chromosomes aligning. These first two phases of mitosis are slightly different in budding yeast than in most eukaryotes, as in *S. cerevisiae* chromosomes do not condense and the nuclear envelope does not fragment in prometaphase. Due to the lack of condensation, yeast chromosomes do not align on a traditional metaphase plate (Straight, 1997). In anaphase, the shortest phase of mitosis, cohesin proteins are cleaved which allow daughter chromosomes to be pulled apart to the poles of the cell. The main proteins acting at this time are the subunits of the anaphase promoting complex (APC/C) which has a role in marking cyclins for degradation in addition to activating the separase enzyme, which is responsible for degrading cohesins (Manchado *et al.,* 2010). The last phase in mitosis is telophase, where budding yeast also differs

compared to other eukaryotes. Throughout the cell cycle, buds are enlarged until telophase, where proteins are synthesized and localized to the neck region to prepare for cytokinesis. In this stage a contractile actomyosin ring forms at the neck of the bud of the cell (reviewed in Morgan, 2007). This division is asymmetrical as the budding daughter cells are smaller than their mother cells. This process differs from other eukaryotes as most cells undergoing cytokinesis split into two approximately equal size cells.

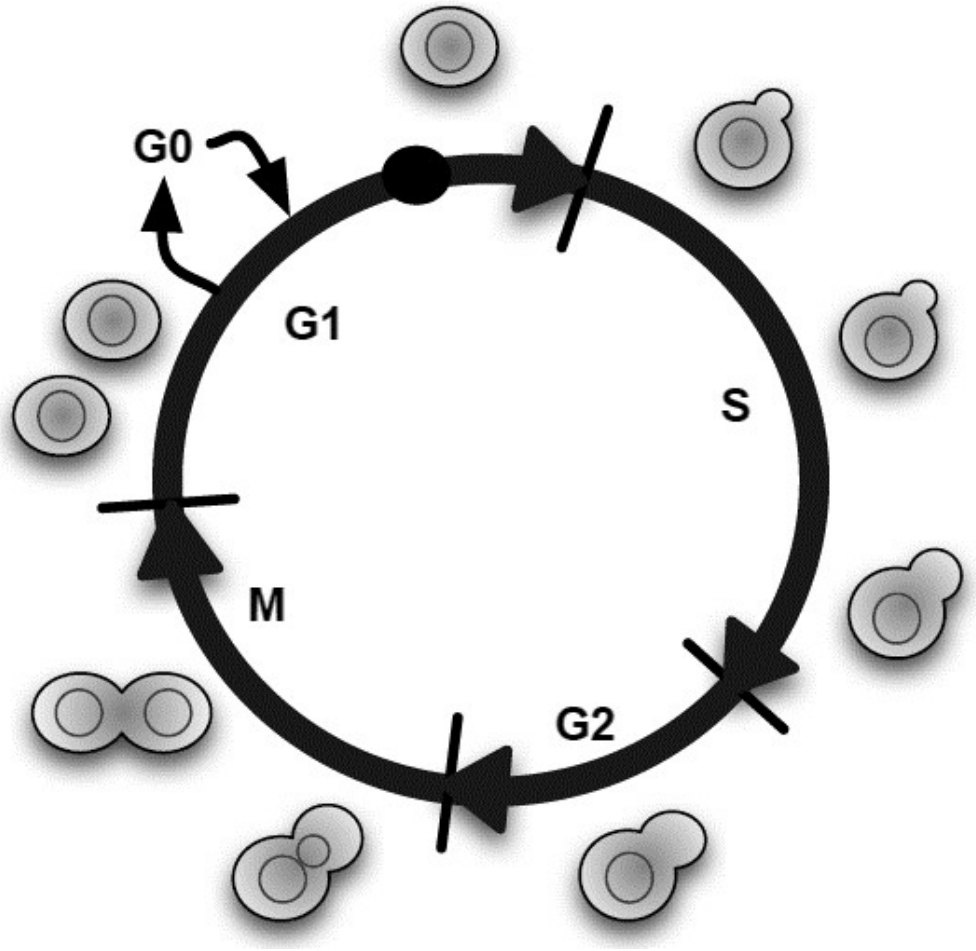

Figure 1: Cell cycle stages in *Saccharomyces cerevisiae.* Life cycle of yeast including G1, S, G2 and M phases, including the budding of a daughter cell. Solid black circle in G1 shows the position where cells can be arrested in G0 via mating factors and other external triggers. Adapted from Herkowitz, 1988.

1.2 DNA Replication and the Intra S-Phase Checkpoint

1.2.1 Initiation of DNA Replication

In budding yeast, origins are specific sites in the genome called Autonomously Replicating Sequences (ARS). These origins fire temporally throughout S-phase and fire no more than once per cell cycle (Nieduszynski *et al.*, 2006; Mantiero *et al.,* 2011). Yeast ARSs are approximately 150 base pairs and there are approximately 400 ARSs on the 16 chromosomes in the haploid yeast genome. The initiation of DNA replication is split into two main phases: DNA licensing which happens in late G1 phase and prepares origins to fire which is followed by the activation of DNA replication (Figure 2).

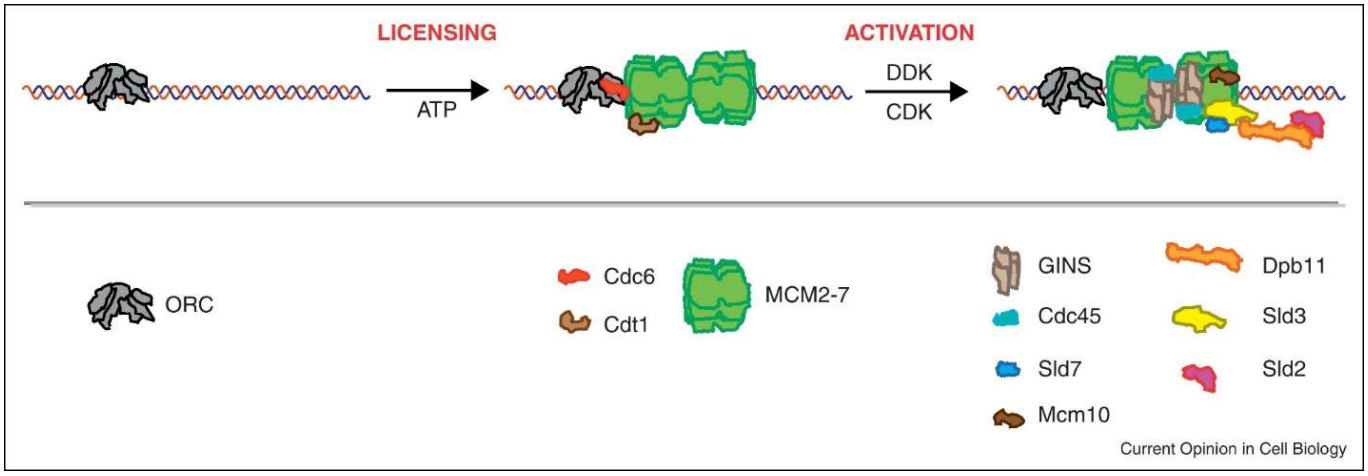

Figure 2: Outline of proteins involved in the initiation of DNA replication in *Saccharomyces cerevisiae*. In the licensing phase, the Mcm2-7 helicase is loaded onto origins with the help of ORC1-6, Cdt1 and Cdc6. The pre-RC is activated when two kinases, DDK and CDK, phosphorylate key proteins in the pre-RC Reproduced with permission from Boos *et al.,* 2012.

Under normal cellular conditions, the first protein complex that binds to origins is the Origin Recognition Complex (ORC), which serves as a base for the binding of other essential proteins. ORC is a complex of six proteins (ORC1-6) and is loaded onto origins during late G1 phase. With the help of cell division cycle protein 6 (Cdc6) and chromatin licensing and DNA

replication factor 1 (Cdt1), minichromosome maintenance protein (Mcm) 2-7 helicase is loaded onto the origin (Randell *et al.,* 2006). MCM helicases are heterohexameric proteins that are responsible for unwinding DNA during replication and are dependent on ATP hydrolysis. After the MCMs are loaded, origins are licensed and ready to be activated, as at this stage they are inactive.

Following this, the Cyclin Dependent Kinase (CDK) and the Dbf4 Dependent Kinase (DDK) activate the origins by phosphorylating key replication proteins (Boos *et al.,* 2012). One important group of proteins that help with the initiation of replication are the Sld proteins (Synthetic lethal with Dpb11). DDK is a complex that consists of cell division cycle 7 (Cdc7) and dumbbell former 4 (Dbf4), which phosphorylates the Mcm2-7 helicase. This phosphorylation event allows the Sld7-Sld3 complex in addition to Cdc45 to associate with the pre-replicative complex (Heller *et al.*, 2011). CDK contributes to initiating DNA replication by phosphorylating two proteins: Sld2 and Sld3. Phosphorylation of Sld2 and Sld3 allows for N-terminal and C-terminal binding of Dpb11 (DNA polymerase B 11). This phosphorylation event recruits several replication proteins to the pre-replicative complex including Sld2, Pol ε, Dpb11 and GINS (reviewed in Li and Araki, 2013). With the addition of the GINS heterotetramer (Psf1-3; Sld5), the CMG helicase (Cdc45, Mcm2-7, GINS complex) forms and activates the helicase (Yabuuchi *et al.*, 2006). In addition to its role in starting the initiation of DNA replication, CDK also plays a role in stopping re-licensing and re-firing of origins, ensuring the DNA is replicated only once per cell cycle. It contributes to this as MCM loading can only occur in G1 phase when CDK levels are low, and origins fire only when CDK levels are significantly higher in S-phase. The protein Mcm10 is important for RPA recruitment to forks, which means that a complex similar to a CMG can be formed without Mcm10, however it will not be functional (Kanke *et al.,*

2012; Yeeles *et al.,* 2015). This lack of functionality is due to the absence of RPA recruitment to forks. With the addition of the appropriate polymerases and topoisomerases, we now know that this list is exhaustive, as initiation of DNA replication with DNA synthesis has be reconstituted with purified proteins (Yeeles *et al.,* 2015). A figure outlining this process is shown in Figure 2.

1.2.2 Replication Stress and Rad53

Knowledge of the molecular mechanisms by which cells safeguard the integrity of their DNA is important to better understand the way they defend themselves against environmental carcinogens. When cells experience conditions of genotoxic stress, they trigger a response that helps maintain both the integrity of the genome and the viability of the cell. Two of the major proteins involved in this are Mec1 and Tel1 (ATR and ATM in humans). Mec1 primarily responds to excess ssDNA in cells, whereas Tel1 I is activated when double stranded breaks (DSBs) are in excess (reviewed in Cimprich & Cortez 2008). These two kinases are at the top of the DNA damage signalling pathway that activates the effector kinase Rad53 via several other proteins in the DNA damage response to DNA damage and replicative stress (reviewed in Carr, 2002).

Under conditions where a DNA damaging agent, such as ultra-violet (UV) light which causes pyrimidine dimers is present, the cell triggers the intra-S phase checkpoint response. Rad53 becomes activated and phosphorylates Dbf4. When this occurs, there is a dissociation of Dbf4 from replication forks and DNA replication is inhibited so that the cell can combat the DNA damage. The initiation of DNA replication in normal conditions and in stress inducing conditions is shown in Figure 3 (adapted from Duncker and Brown, 2003). The checkpoint response stops replication initiation by blocking further origins of DNA replication from firing and by slowing replication fork progression.

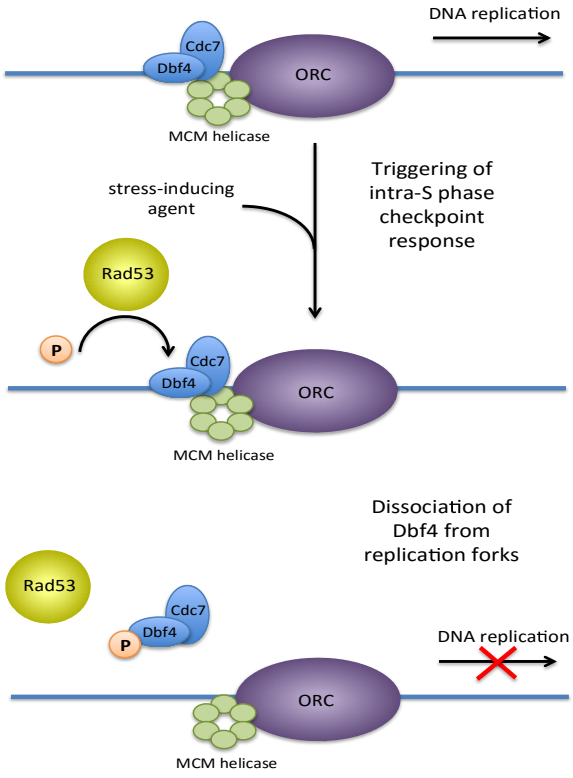

Figure 3: Triggering of Intra-S Phase Checkpoint Response via Rad53 Kinase in *S. cerevisiae*. In the absence of DNA damaging agents, DDK binds to the MCM2-7 helicase promoting DNA replication. When the intra-S phase checkpoint response is triggered, Rad53 phosphorylates Dbf4 which causes a dissociation of Dbf4 from replication forks and stops DNA replication. Adapted from Duncker and Brown, 2003.

Rad53 has a central catalytic kinase domain in addition to two forkhead associated (FHA) domains. FHA domains are protein-protein binding motifs that are well studied for their interactions with phospho-threonines. FHA domains have high ligand specificity. For example, Rad53 FHA1 has been shown to preferentially bind to the DNA replication factor Dbf4, whereas Rad53 FHA2 has been shown to preferentially bind to checkpoint protein Rad9 (Duncker *et al.,* 2002; Schwartz *et al.,* 2002). It is important to note that Rad53 is different from other FHA domain containing proteins as it has two FHA domains whereas all but one other known FHA domain containing protein contain only one FHA domain. In addition to the conserved phospho-threonine binding pocket found on all FHA domains, the Duncker lab along with its collaborators recently reported that FHA1 has a region on one of its surfaces consisting of ten conserved amino acid residues (herein called the lateral patch) that are not present in FHA2 (Matthews *et al.,* 2013). The conserved patch includes residues on strands β1 (Arg35, Ile37, Thr39), β7 (Asn112, Gly113), β10 (Gln126, Asp128, Gln129), and β11 (Val144, Phe146). Figure 4 shows Rad53 FHA1 and Rad53 FHA2 with the highly conserved lateral patch on FHA1 that is absent on FHA2 in red. When these domains are rotated 180 °, it is shown that FHA1 has a second conserved patch (also shown in red) while FHA2 does not. It was hypothesized that this lateral patch could be the explanation for the different ligand specificities seen with FHA1 and FHA2 domains.

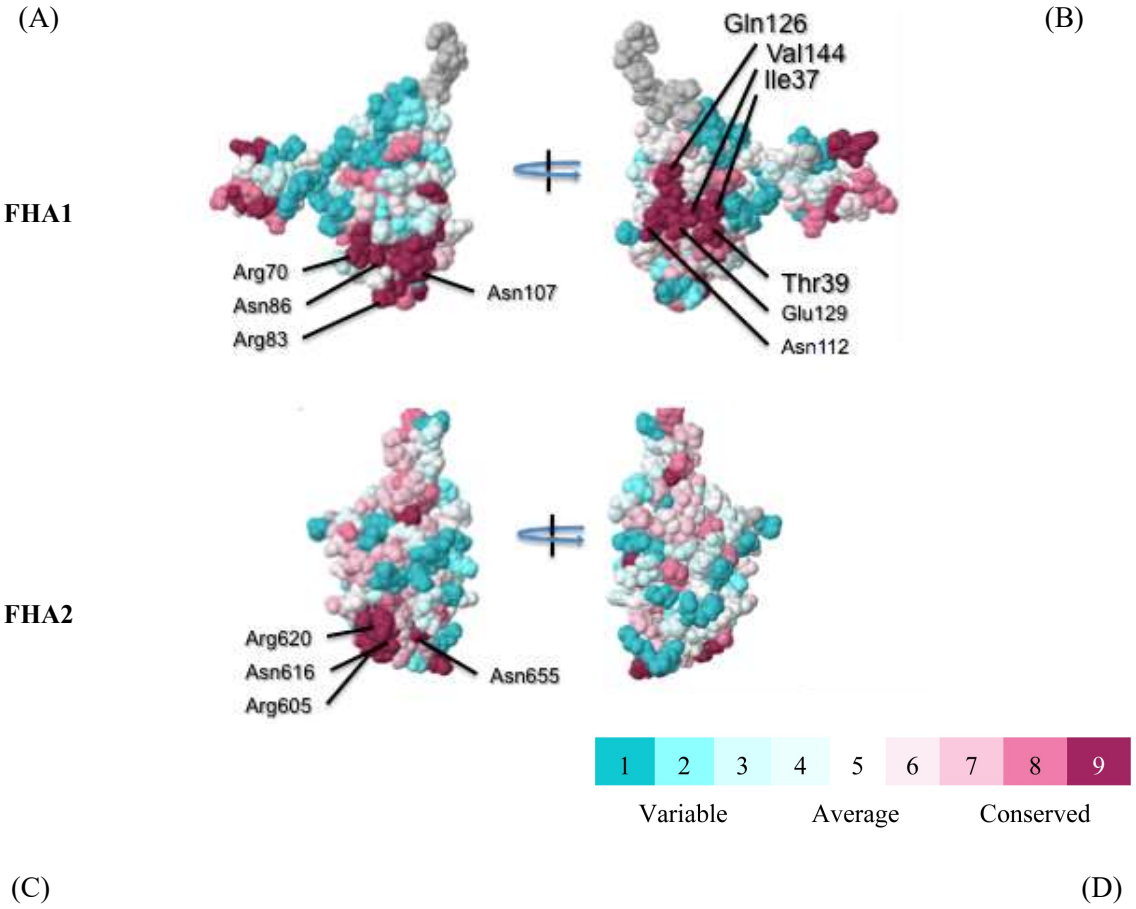

Figure 4: Computational surface mapping of Rad53 FHA domains. Multiple sequence alignment files of Rad53 orthologs were generated and mapped onto PDB structures using Consurf. (A) FHA1: Highly conserved surface patch #1 (phosphopeptide binding site). (B) FHA1: Highly conserved surface patch #2. (C) FHA2: Highly conserved surface patch (phosphopeptide binding site). (D) FHA2: Absence of a second highly conserved surface patch. From Matthews *et al.,* 2014.

This hypothesis was investigated when this region was examined for its importance in interacting with Dbf4. This was of particular interest as the region of Dbf4 necessary for sufficient interaction with Rad53 has previously been determined and crystalized as a conserved BRCT domain with an additional N-terminal alpha helix (Matthews *et al.*, 2012). However, while the domain of Rad53 required for Dbf4 interaction has been identified as FHA1, it had yet to be determined which particular region of FHA1 is responsible for conferring greater specificity than FHA2 for Dbf4 (Matthews *et al.*, 2012). The lateral patch was promising, given

that when all candidate threonines in the H-BRCT domain were mutated, none showed an abrogation of Rad53 binding, suggesting that the Rad53-Dbf4 interaction may be non-canonical. Via two-hybrid analysis, Darryl Jones, a previous PhD student showed that quadruple mutations of the lateral surface completely abrogated the interaction between the two proteins, as shown in Figure 5 (Matthews *et al.,* 2014).

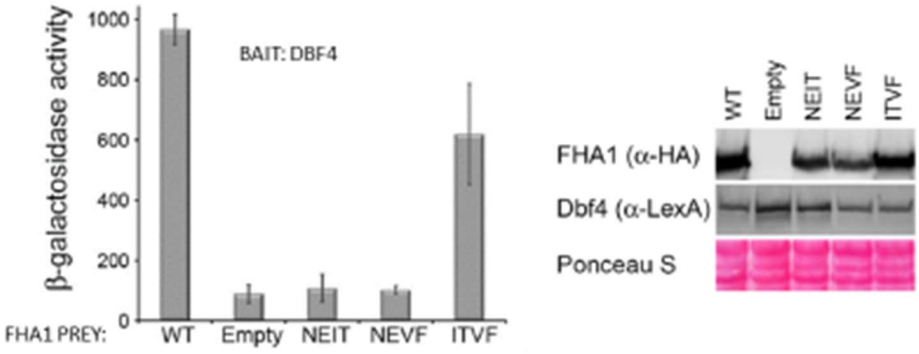

Figure 5: A conserved lateral patch on Rad53 FHA1 mediates the interaction with Dbf4. Yeast-two hybrid analysis (Left), along with corresponding immunoblot data for bait and prey expression (Right) NEIT=N112A/E129A/I37A/T39A, NEVF=N112A/E129A/V144N/F146A ITVF=I37A/T39A /V144N/F146A. Reproduced with permission from Matthews *et al*., 2014.

Three combinations of quadruple mutations were used, and for those involving glutamic acid 129 and asparagine 112, the interaction between Rad53 FHA1 and Dbf4 fell to background levels. Recently, our collaborators at McMaster University solved the co-crystal structure of Rad53 FHA1 and Dbf4 HBRCT motif, confirming that Dbf4 interacts with FHA1 via its lateral patch, and that there are two contact points in this interaction. This double contact point is the reason why the IVTF quad mutant did not significantly alter the interaction as it was only disrupting one of two sites of interaction, whereas NEIT and NEVF disrupted both.

Once it was established that the Rad53 FHA1 lateral patch is important for mediating its interaction with Dbf4, the phenotypes of these mutants were examined under DNA damaging

conditions. This work was done by transforming wild-type (WT) and mutant versions of the DNA sequence encoding FHA1 in a CEN vector (single copy per cell) into a yeast strain background that is *rad*53Δ, *sml*1Δ. This background was chosen as it knocked out endogenous Rad53 expression in the cells. As this Rad53 knockout is not viable on its own, it is combined with *sml*1Δ mutation, as this restores viability by increasing levels of dNTP in cells.

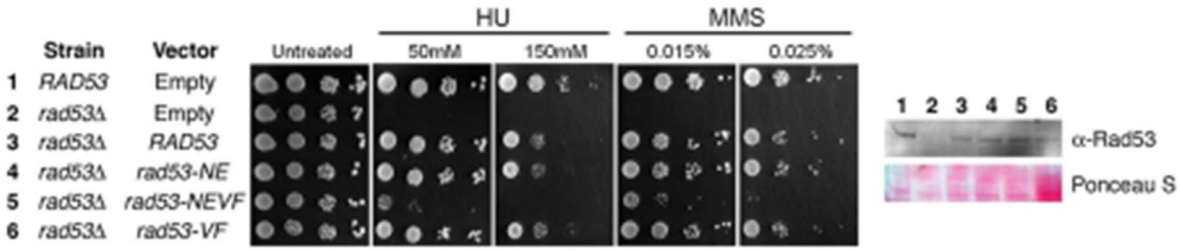

Figure 6: The rad53-NEVF mutant is sensitive to genotoxic stress. *rad*53Δ, *sml*1Δ cells were transformed with vectors expressing the listed Rad53 variants. 10-fold serial dilutions of cells were spotted onto plates with the indicated concentrations of HU or MMS. Reproduced with permission from Matthews *et al.,* 2014.

In Figure 6 we see that the NEVF lateral surface mutation results in a significantly increased sensitivity to genotoxic stress when compared to the WT controls. It is known that the exposure of yeast cells to either hydroxyurea (HU) or methyl methanesulfonate **(**MMS) triggers the intra-S phase response, which the NEVF mutant cannot participate in as efficiently the WT Rad53 as it cannot bind to Dbf4. These DNA damaging agents trigger the intra-S phase response in two different ways: HU restricts dNTP pools by suppressing the enzyme dihydrophosphate reductase and eventually causing double strand breaks whereas MMS alkylates DNA.

1.3 FHA Domains

1.3.1 Structure and Function of FHA Domains

In order to transduce signals throughout the cell, a mechanism has to be in place in order to regulate protein-protein interactions with both upstream and downstream binding partners. An example of such a protein-protein interaction domain involved in cellular processes is the FHA domain. First identified in 1995, FHA domains have been found in over 200 species, from prokaryotes to higher eukaryotes (including humans) (Hofman & Bucher, 1995; Durocher & Jackson, 2002). The shuffling of exons encoding FHA domains around genomes has provided a way for an evolutionarily-rapid formation of new multi-domain proteins with extremely diverse functions, which created complex and selective signal transduction systems and cellular pathways (Durocher & Jackson, 2002). There has been an explosion in the structural knowledge of FHA domains in the recent literature, with FHA domains being crystallized from numerous organisms (Machida & Yuan, 2013; Weng *et al.,* 2015; Wysoczanski *et al.,* 2015; Luo *et al.,* 2015). Although FHA domains are not highly conserved on the sequence level, they all share a similar structure. This structure includes a β-sandwich composed of two β-sheets that are connected by short loops. Each β-sheet contains five to six strands and the entire domain is approximately 100 amino acids in size. In the literature, the six short loops that connect the β-sandwich are primarily responsible for mediating interactions with FHA ligands. These loops are the best studied regions of FHA domains and their role in mediating ligand specificity has been extensively characterized (Durocher *et al.,* 2000; Mahajan *et al.,* 2005). An exception to the typically low consensus sequence when aligned, is the conserved pThr binding pocket. This binding pocket consists of a key conserved arginine with several other amino acids that stabilize

its interaction with the phosphorylated threonine in its ligand. FHA domains almost exclusively recognize phosphorylated threonines due to the recognition of the methyl side chain on threonine as compared to serine. This additional methyl side chain acts as a key that can interact with a small hydrophobic cavity created by a histidine that is conserved among FHA domains while the pThr phosphate binds to a conserved arginine located on the FHA domain (Pennel *et al.,* 2010). Although controversial, there is a proposed rule that gives FHA domains their ligand specificity, the pThr + 3 rule. According to this rule, the amino acids on the FHA domain that interact with the amino acid three positions C-terminal to the pThr on the ligand are responsible for giving FHA domains their ligand specificity. There is some evidence for this heuristic rule as some proteins follow this rule, for instance Rad53 FHA1 prefers ligands with an Asp in the +3 position (Yongkiettrakul *et al.,* 2004). However there are numerous exceptions to the rule. When Dun1 interacts with Rad53 it recognizes two phosphothreonines (Thr 5 and Thr8) instead of one. Additionally the serine at position +3 relative to Thr8 of Rad53 does not form strong interactions with Dun1 FHA (Lee, H. *et al.,* 2008). Furthermore, in specific murine and *Mycobacterium tuberculosis* FHA domains it has been shown that the FHA domain binds ligands three residues N-terminal to the pThr (Alderwick *et al.,* 2006; Berstein *et al.,* 2005). In summary, the pThr +3 rule is a useful tool to predict FHA ligands, however numerous exceptions have already been documented.

There is a small but growing body of literature that documents non-canonical FHA interactions. In humans, the protein Ki67 does recognize a pThr, however its FHA domain binds to a region that is remote from this phosphorylated amino acid (Byeon *et al.,* 2005). In *M. tuberculosis* the FHA pThr binding pocket is responsible for mediating interactions that are both phosphorylation dependent and phosphorylation independent (Nott *et al.,* 2009). In

Schizosaccharomyces pombe a recently identified FHA domain in the protein MDB1 has been crystalized, and interestingly lacks all conserved pThr binding residues (Luo *et al.,* 2015). The MDB1 FHA domain is primarily responsible for mediating interactions with itself to form MDB dimers, and this allows the protein to localize to DNA damage sites and bind ligands involved in the DNA damage response.

1.3.2 FHA Domains and their ligands in *S. cerevisiae*

In total, there are thirteen FHA domains in *Saccharomyces cerevisiae*, all of which play a role in the cell cycle. Two particular proteins of interest are Dun1 (DNA damage uninducible) and Mek1 (meiotic kinase). Dun1 is a homolog of Rad53 and these two proteins are replaced by a single cell cycle protein in higher eukaryotes. Dun1 has a single FHA domain and interacts with Rad53 upstream and has four ligands downstream (Sml1, Dif1, Crt1, Wtm1) in order to regulate dNTP pools in the cell. When phosphorylated, Dun1 will interact with Sml1 to promote the degradation of the catalytic domain of Rnr1 (ribonuclease reductase one) (Zhao *et al.,* 1998). Ribonuclease reductases are a group of enzymes that catalyze the formation of dNTP from ribonucleotides. Sml1 also has a role in regulating dNTP levels in normal S phase, as an extreme sensitivity to DNA damaging agents is seen in cells lacking Mec1 or Rad53, and this sensitivity is drastically reduced in cells that are also lacking Sml1. Dun1 phosphorylation of Dif1 abrogates the nuclear import of Rnr2-4, three other ribonuclease reductases (Lee *et al.,* 2008) and interaction with Wtm1 leads to a disassociation with Rnr2-4 (Lee and Ellege, 2006). Under normal conditions, Wtm1 maintains an interaction with Rnr2-4 in the nucleus, so a decrease in this interaction allows Rnr2-4 to move to the cytoplasm and bind Rnr1 forming and active Rnr

complex. Dun1 can also modulate transcription levels of ribonuclease reductase inhibitors by interaction with Crt1. A model showing these interactions can be seen in Figure 7.

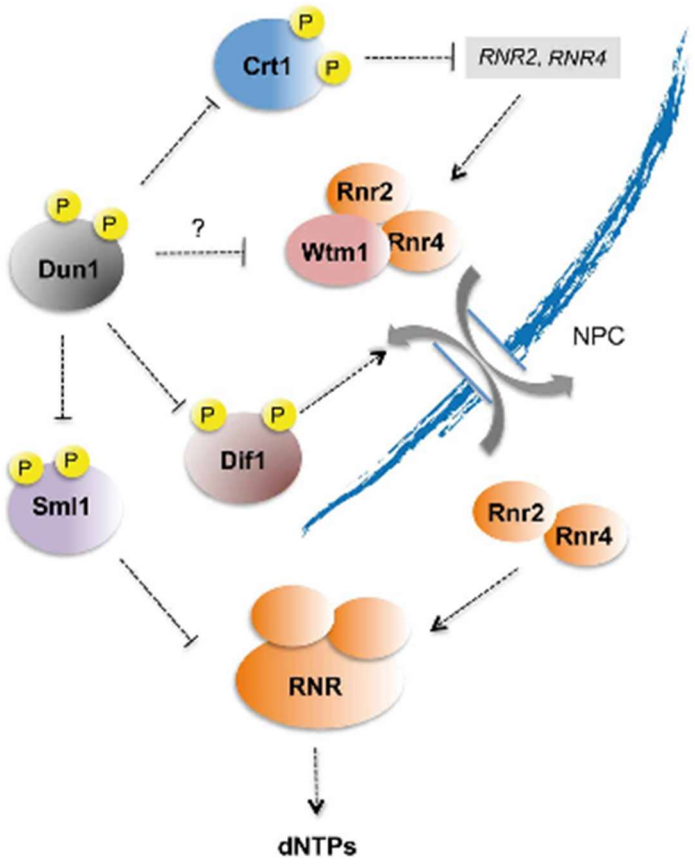

Figure 7: Checkpoint control of S phase transcription and dNTP pools. Dun1 regulates dNTP pools by binding to various partners. Sml1 and Crt1 directly help to increase dNTP levels in cells where Dif1 assists with nuclear import of RNRs. NPC stands for nuclear pore complex. Reproduced with permission from Jossen & Bermejo, 2013.

A major part of my project was looking at the lateral surface of Rad53 FHA1 and the role it plays in ligand binding. Rad53 has two FHA domains each with two completely different ligand affinities as FHA1 has a multitude of binding partners such as Cdc7, Sgs1 and Sld3 whereas FHA2 is only known to bind Rad9 and itself (Smolka *et al.,* 2006).

Small growth repressor 1 (Sgs1) is a member of the RecQ helicase family with clear homologs in humans including BLM and WRN (Hickson, 2003). In addition to binding Top3

Sgs1 also contributes to the intra-S checkpoint phase response via Rad53 binding. When phosphorylated in response to DNA damage, Sgs1 changes its binding preference to have a higher affinity for Rad53. It has been proposed that this binding event helps recruit Rad53 to Mec1-Ddc2, two proteins that are typically associated with stalled replication forks (Hegnauer *et al.,* 2012). This recruitment activates the Mec1-Ddc2 complex, allowing it to then phosphorylate Rad53, which promotes the intra-S phase checkpoint response.

The last ligand that was tested with FHA1 is Sld3, a protein known to interact with Rad53 (Zegerman and Diffley, 2010). Sld3 is a protein that is essential for replication initiation but not elongation, and it is hypothesized that it also plays a role in preventing late origin firing in checkpoint dependent manner.

1.4 Research Objectives

The main goals of my project were to further characterize the role that the lateral binding patch of Rad53 FHA has on ligand binding specificity, as well as identify and characterize additional conserved regions in other FHA domain containing proteins in the budding yeast *Saccharomyces cerevisiae.* My hypothesis was that any conserved FHA lateral patches found would be important for ligand specificity, and that mutations in these region would cause decreased protein-protein interactions between FHA domains and their known ligands. My work focused on determining the role of the lateral surface of FHA1 in binding four ligands: Dbf4, Cdc7, Sgs1 and Sld3. In addition to FHA1, all budding yeast FHA domains were examined using bioinformatics methods to look for conserved lateral surfaces and the Dun1 FHA lateral surface's role in dNTP regulation was investigated. The main applications of this research relates to cell-cycle regulation and particularly cancer. By knowing the exact amino acids required for mediating important checkpoint interactions, our basic knowledge of cell cycle regulation is

furthered in addition to the potential therapies that can be created in order to target specific proteins which would alter the cell cycle and cell growth.

Chapter 2: Methods

Chapter Two: Methods

2.1 Yeast Strains

Two yeast strains were used in this study. DY-1 (*MATa, ade2-1, can1-100, trp1-1, his3-11,-15, ura3-1, leu2-3,-112, pep4:LEU2*) is the strain that was used in two-hybrid assays while DY-294 (*MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 Dun1Δ*) was used in the spotting plate assays performed to examine Dun1 mutants.

2.2 Genomic DNA Preparation

DY-1 cells were grown to mid log-phase in 10 mL of YPD media, washed with 30 mL water, and resuspended in 2 mL screw cap tubes with 200 µL yeast genomic prep mix (2% Triton X-100, 1% SDS, 100 mM NaCl, 10 mM Tris-Cl (pH 8), 1 mM EDTA). Subsequently, 0.3 g of 0.5 mm glass beads was added in addition to 200 µL phenol:chloroform:isoamylalcohol (25:24:1). The solution was vortexed for 4 minutes when 200 µL of TE was added and the tube centrifuged at 13,200 rpm for 5 minutes. The supernatant was then transferred to a fresh tube with 1 mL of 100% ethanol. Tubes were inverted ten times and then centrifuged for 2 minutes at 13,200 rpm. The supernatant was removed via decanting and the pellet was resuspended in 400 µL of TE. Pellets were then resuspended in 10 µL of RNaseA (10 mg/mL; Sigma) incubated at 37ºC for 10 minutes. After this, 10 µL of 4M ammonium acetate and 1 mL of 100% ethanol was added and mixed by inversion. This was centrifuged at 13,200 rpm for 2 minutes. Lastly, the supernatant was discarded and the pellet was allowed to air dry for ten minutes in a sterile environment. The pellet was then resuspended in 50 µL TE.

2.3 Plasmid Construction

In order to assay protein-protein interaction via a yeast two-hybrid analysis, two protein-coding DNA sequences need to be cloned into specific vectors. For this study, FHA domains were cloned into the empty pJG4-6 vector and FHA ligands were cloned into the empty pEG202 vector (Gyuris *et al.,* 1993; Ausubel *et al*., 1994). Prior to the commencement of my project, pJG4-6 Rad53 FHA1, pJG4-6 Rad53 FHA2, pJG4-6 Dun1 and pJG4-6 Mek1 had been cloned in the lab.

Before the current project commenced, pEG202 Dbf4, pEG202 Cdc7, pEG202 Sml1, and pEG202 Rad9 had been constructed. The SGS1 region corresponding to amino acids 245-1179 was cloned using forward and reverse primers with BamHI and XhoI restriction sites, respectively, using a full length pEG202 SGS1 plasmid, which is an unpublished plasmid created by former Duncker lab member Darryl Jones, as template. pEG202 Rad53 was created by PCR amplification of genomic Rad53 using forward and reverse primers corresponding to amino acids 1-821 with NcoI and EcoRI restriction sites, respectively.

pCM190Myc Sld3 was previously created in the Duncker laboratory. pCM190myc is a tetracycline-repressible expression vector that has titratable levels of protein expression correlating with levels of tetracycline. Recombinant proteins expressed on this vector are Myc tagged for detection on immunoblots (Garí *et al.,* 1997). pCM190Myc Dun1 was generated via PCR amplification of genomic Dun1 using forward and reverse primers corresponding to sequence encoding amino acids 1-513 using BamHI and NotI restriction sites, respectively.

For all constructs, once PCR was completed the products were purified with a PCR cleanup kit (Geneaid) and ligated in to the appropriate vector (pJG4-6, pEG202 or pCM190Myc). After ligation, the entire reaction was transformed into DH5α $CalCl_2$ competent

E.coli cells. Prospective positive colonies, as determined by colony PCR, were grown in 10 mL of LB+Amp media overnight and plasmid DNA was isolated using a plasmid prep kit (Geneaid). All plasmids were sent for sequencing to make sure that no new mutations were generated in the PCR. Two sequencing centres were used in this project: The London Regional Genomic Center and TCAG DNA sequencing facility (Sick Kids Hospital).

For FHA mutant constructs, and the Cdc7 T484A mutant specific bases in the gene were mutated using the QuikChange II XL Site-Directed Mutagenesis Kit (Agilent Technologies). Wild-type versions of FHA domains in pJG4-6 and Cdc7 in pEG202 were the templates in mutagenic PCR reactions. Mutagenic primers were designed in such a way as to minimize the energy cost of mismatching cases. PCR cycles amplified the entire plasmid, both the parental and mutagenic versions. Copies of the parental plasmid were then digested with the enzyme Dpn1 and the remaining mutant copies of the plasmid were subsequently transformed into XL-10-Gold Ultracompetent cells using the heat shock method. Cells were plated onto LB-AMP plates, grown overnight at 37° C and plasmid DNA purified from specific transformants was sent off for sequencing using primers flanking the mutagenic region to ensure that the mutagenic PCR was successful. A full protocol is in the QuikChange II XL Site-Directed Mutagenesis instruction manual: http://stanxterm.aecom.yu.edu/wiki/data/Product_manuals_attach/quikchange2xl.pdf.

2.4 Yeast Transformation

Cultures were initially grown to mid-log phase in 50 mL of Synthetic Complete (SC) media with the appropriate amino acids. SC media consists of yeast nitrogen base (0.67%), ammonium sulfate (1%), glucose (2%) and a 1X amino acid mix. Cells were then counted using a hemocytometer under a light microscope to ensure a suitable number of cells (approximately 1 x 10^7) was used in the transformation. Cells were then centrifuged for 5 minutes at 4,000 rpm,

resuspended in sterile 1X TE (Tris-EDTA), spun a second time and resuspended in 2 mL of sterile 100 mM lithium acetate (LiAc)/0.5X TE. The resuspended cultures were incubated for 10 minutes at room temperature. 750 µg of plasmid DNA constructs were added to microcentrifuge tubes with 100 µg of salmon sperm DNA and 100 µL of yeast suspension mix. Subsequently 300 µL of sterile 100 mM LiAc/40% PEG4000 (polyethylene glycol 4000)/ 1XTE solution was added. Contents were mixed gently using a P1000 pipetteman. Transformation tubes were then incubated at 30°C for thirty minutes. Subsequently, 40 µL of sterile DMSO (dimethylsulfoxide) was then added before heat shocking the tubes at 42°C for 7 minutes and placing the tubes on ice for 2 minutes. Cells were spread onto selective media plates to detect for successful transformants and incubated at 30°C for 2-3 days.

2.5 Yeast Two-hybrid Assay

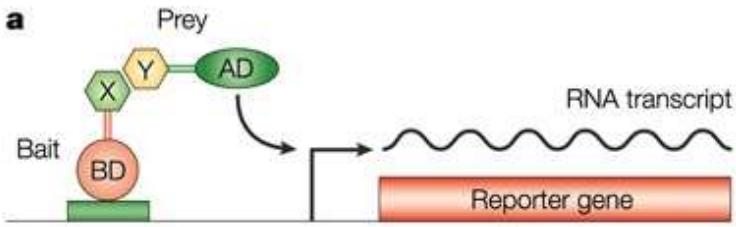

Figure 8: **The yeast two-hybrid assay detects protein interactions**. Each putative interacting protein (X and Y) is fused to one of two functionally distinct domains of the transcription factor. The bait comprises a protein fused to the DNA-binding domain, and the prey comprises a protein fused to the transcriptional activation domain. Physical interaction between bait and prey brings the DNA-binding domain and the activation domain into close proximity, thereby reconstituting a transcriptionally active transcription factor that can bind to the reporter plasmid, starting transcription of the reporter gene. Reproduced with permission from Grünenfelder & Winzeler, 2002.

In a yeast two-hybrid analysis, the interaction of one protein, the "bait" with a second protein, the "prey" results in the expression of a reporter gene, which in this project encodes a protein product called β-galactosidase. An illustration of this technique can be seen in Figure 8 (reviewed in Grünenfelder and Winzeler, 2002). The prey plasmids in this study are pJG4-6 derived plasmids and the bait plasmids are pEG202-derived plasmids. A third reporter plasmid,

pSH 18-34, has a *LacZ* gene which encodes the enzyme β-galactosidase that is transcribed and then converts the substrate ONPG (2-Nitrophenyl-B-D-galatopyranoside) to 2-Nitrophenol and galactose, a reaction that results in a yellow color change. Yeast two-hybrid analysis starts by growing cells to mid log-phase (approximately $1x10^7$ cells/mL) in synthetic complete media that is deficient in uracil, tryptophan and histidine. These selectable markers are omitted to select for cells with the three required plasmids for a two-hybrid assay, as genes encoding proteins that are responsible for their biosynthesis are deleted in the yeast strain used (DY-1). These aforementioned selectable markers are encoded by the three yeast two-hybrid plasmids and can confer cells the ability to grow in specific amino acid or base deficient media. Cells were then washed to remove glucose and recombinant protein expression was induced by growing cells for six hours in 2% galactose 1% raffinose SC media deficient in uracil, tryptophan and histidine. By switching the carbon source from glucose to galactose, recombinant protein production on the pJG4-6 vector is switched from an "off" to an "on" state. Using a hemocytometer, cells were counted to ensure they were mid at log phase and 5 $x10^6$ cells were set aside for yeast two-hybrid analysis. The separated cells were centrifuged at 13,200 x g for 10 minutes. Cells were then suspended in 5 mL Z-buffer (0.06 M Na_2HPO_4, 0.04 M NaH_2PO_4, 0.01 M KCl, 0.001 M $MgSO_4$, 0.05 M β-mercaptoethanol) with 0.1 mL of chloroform and 0.05 mL of 0.1% SDS also being added. Tubes were mixed by vortexing for 10 seconds followed by a five minute incubation at 28°C. A 100 µL solution of ONPG (4 mg/mL) was added to all tubes simultaneously to start the reaction and the reaction was stopped by the addition of 250 µL sodium carbonate, with the total elapsed time recorded. The level of β-galactosidase activity can be measured via spectroscopy at 420 nm and is proportional to the strength of the interaction between the two proteins. The relative β-gal activity can then be calculated through the equation β-gal activity = 1,000 x *A420* /

(t x V x $A600$), where t = time of the reaction (in minutes), and V = volume of culture used in the assay (in millitres). A stronger yellow color upon reaction completion is indicative of a stronger interaction between the two proteins being tested. Yeast transformants with empty prey and/or bait vectors were used as negative controls and combinations of bait and prey vectors expressing proteins that are known to interact were used as positive controls.

2.6 Whole Cell Extract Preparations and Western Blotting

Following removal of culture aliquots for β-galactosidase activity measurements, the remainder of the cultures used in yeast two-hybrid analysis were used to prepare whole cell extracts (WCE). Cells were spun down at 4000 x g for five minutes and resuspended in 250 µL lysis buffer (10 mM Tris-HCl, pH 8; 140 mM NaCl; 1% Triton X-100; 1 mM EDTA; 1 mM PMSF and 1X HALT protease inhibitor cocktail from Thermo-Fisher). Cells in lysis buffer were transferred to 2 mL cryogenic vials and 0.3 g of glass beads were added. A 4°C, cells were lysed using a Biospec bead beater with four alternating cycles of one minute in the bead beater and one minute on ice. Vials were then spun at 13,200 x g for one minute and the protein concentration of the supernatant was determined via a Bradford assay using a BioRad dye. In order to prepare each sample for SDS-page, proteins were heated for fifteen minutes at 65°C after being mixed with loading dye (60% 4X buffer [15% SDS; 40% glycerol, 166 mM tris-base]; 0.26 M DTT; 7% bromophenol blue). 55 µg of prepared proteins were then run on a 10% SDS-page gel and subsequently transferred to nitrocellulose using a wet transfer. The transfer buffer used consisted of 200 mM glycine, 25 mM tris-base, 20% methanol and 0.05% SDS. Transfers were run between 2-12 hours and membranes were then stained with .1% Ponceau-S to determine the equal loading and proper transfer of proteins. Membranes then were destained with 1X-TENT (20 mM Tris-HCl; 1mM EDTA; 0.14 M NaCl; 0.05% Tween 20) and blocked between 2 and 12 hours in a 5% skim milk 1X TENT solution.

Primary antibodies were in a 3% bovine serum albumin 1X TENT solution and incubated with membranes for 1-2 hours. After three washes with 1X TENT for ten minutes, membranes were incubated with the appropriate secondary antibody (in 5% skim milk 1X TENT solution) for 1-2 hours. Following three more ten minute washes with 1X TENT, membranes were imaged using a BioRad Pharos imager and BioRad Quantity One software. A table outlining the antibodies used and their respective dilutions is shown below.

Table 1: Antibodies used for immunoblotting. Antibody, dilution used and manufacturer are listed.

Antibody	Dilution	Source
Anti-HA (Mouse monoclonal)	1:3000	Sigma
Anti-LexA (Rabbit polyclonal)	1:5000	Cedarlane
Alexaflour 488 goat anti-mouse	1:3000	Invitrogen
Alexaflour 647 goat anti-rabbit	1:3000	Invitrogen

2.7 Spotting Plate Assays

Spotting plate assays can compare wild type and mutant yeast strains, assessing the viability when exposed to DNA damaging agents. In spotting plate assays, cells are first grown to saturation (1 x10^8 cells/mL) in the appropriate selective media (SC-U). SC media consists of yeast nitrogen base (0.67%), ammonium sulfate (1%), glucose (2%) and a 1X amino acid mix. With the omission of the amino acid uracil, the media selects for cells containing the pCM190myc plasmid which has the URA3 selectable marker. Saturated cultures (30 µL) were diluted into 270 µL of fresh medium. Three more 10X serial dilutions were performed in order to

create a range of four cell concentrations. Each cell concentration had 5 µL spotted onto solid media containing different concentrations of DNA damaging agents. Cells were spotted on plates with no genotoxic agents (control), 25 and 50 mM hydroxyurea (HU). Cells were also spotted onto plates with 0.1% and 0.2% methyl methanesulfonate. Plates were incubated at 30°C for 72 hours and imaged with the Alphaimager HP (Alpha Innotech).

2.8 Co-immunoprecipitation

DY-1 cells were transformed with pJG4-6 (HA-tag) or pCM190myc derived expression vectors and grown to mid log-phase in SC-UT as described previously (Varrin *et al.,* 2005). Cells were then spun down at 4000 rpm for 3 minutes, washed with sterile water, and resuspended in 2% galactose 1% raffinose media lacking uracil and tryptophan. Cells were induced in the GR-UT media for six hours at thirty degrees. GR-UT is the same as synthetic complete (SC) media, except the sugars are galactose and raffinose instead of glucose. All remaining steps in this protocol were performed at 4°C. Whole cell extracts were then prepared as described above, but with a different buffer lysis buffer (50 mM Hepes, pH 7.5; 140 mM NaCl; 1 mM Na-EDTA; 1% Triton X-100) supplemented with 10 µL/mL Halt Protease Inhibitor Cocktail (Thermo-Fisher). WCEs were all diluted to the same concentration with lysis buffer (10 µg/µL), and 400 µL of each extract was incubated overnight on a rotating wheel with 15 µL Protein A sepharose beads (Sigma) saturated with mouse monoclonal anti-Myc antibody (Sigma). After 18 hours, the supernatant was removed and saved, following a centrifugation at 13,200 X g for thirty seconds. Sepharose beads were washed 3 times with 500 µL of fresh lysis buffer, and resuspended in a final volume of 35 µL lysis buffer.

2.9 Bioinformatics Analysis

Within this thesis two main bioinformatics projects were carried out. The first was looking for conserved regions of FHA domains among yeast and fungal species. This was done by constructing multiple sequence alignment (MSA) files using the UniProt Clusters 90% database with sequences of FHA domains of interest. This was done using EBI-hosted PSI BLAST, a basic local alignment search tool that allows one to run several iterations of regular BLAST, with each subsequent query being based upon the combined results of the previous queries. This was done to find distantly-related homologs, which allows the multiple sequence alignment to be more stringent. UniProt Clusters was used because it contains sequences that are no more than 90% identical, which also aids in finding more distantly-related homologs. MSA files were constructed from ~20 representative sequences recovered after 2 or 3 iterations, the majority of which were from yeast species and some being from closely related fungal species.

3D models of FHA domains were either retrieved from the PDB database (Rad53 FHA1-1J4O, Dun1 FHA – PDB ID: 2JQJ) or constructed with the use of such modeling servers as SwissModel (http://swissmodel.expasy.org/) (Fkh2 FHA) and Robetta (http://robetta.bakerlab.org/) (all other FHA domains). It must be noted that while the structures created with the use of these modeling servers are of acceptable quality, they may not represent the true conformations of the respective FHA domains in vivo.

The Consurf server (http://consurf.tau.ac.il) was used to map the MSA files onto 3D protein structures and produce a comprehensive model of amino-acid evolutionary conservation, where the degree of conservation is reflected in a normalized scale of colour change, from blue (not conserved) to white (somewhat conserved) to red (highly conserved).

For the second bioinformatics project, FHA domains that are conserved in higher eukaryotes were first identified. This was done both by examining the existing literature and probing *S. cerevisiae* FHA domain sequences against *S. pombe, X. laevis,* and *H. sapiens* and evaluating the top matches. If the identified candidate protein could produce the FHA domain in *S. cerevisiae* as a top match in a reciprocal search, the protein was used in this conservation study. The amino acids sequences of these FHA domains were obtained from SGD for *S. cerevisiae* (*Saccharomyces* Genome Database, 2015), PomBase for *S. pombe* (Wood *et al.,* 2002) or NCBI for higher eukaryotes (National Center for Biotechnology, 2015). Sequences were then trimmed in order to include only the core FHA domain fold, to produce a high quality alignment. A multiple sequence alignment was then generated using MUSCLE (European Molecular Biology Laboratory, 2015) and then phylogenetic trees were constructed using PhyML (Guindon *et al*., 2010).

Chapter 3: Characterization of the importance of the lateral surface of Rad53 FHA1 in ligand specificity

Figure 14 and 15 represent experiments carried out in collaboration with with Polina Myrox

Figure 20 represents work carried out by Boris Simovic

3.1 Introduction

Forkhead associated domains (FHA domains) are protein-protein interaction domains that bind phosphoepitopes and are often found in proteins that regulate the cell cycle, including the response to DNA damage. Proteins with FHA domains contribute to the DNA damage response in many ways, as they can pause DNA replication, assist with stalled forks and help regulate dNTP levels in the cell (Labib & DePiccoli, 2011). FHA domain binding to phosphothreonine (pThr) residues has been extensively studied, as it is the phosphoepitope of choice (Durocher *et al.,* 2000). Although FHA domains are best known for mediating pThr interactions, there is a growing body of literature investigating the role of non-canonical interactions (Byeon et al., 2005, Nott et al., 2009, Luo et al., 2015). These non-canonical interactions are an expanding area of research as they contribute to key DNA damage responses. One of the major FHA domain containing proteins that is involved in the DNA damage response in budding yeast is Rad53. Rad53 is an interesting protein as it contains two FHA domains, FHA1 and FHA2, whereas most FHA containing proteins only contain one. Interestingly, the two FHA domains have different affinities for their ligands, for example FHA1 preferentially binds to Dbf4 whereas FHA2 preferentially binds to Rad9 (Sun *et al.,* 1998). When looking at other ligands, FHA1 has numerous binding partners that contribute to various pathways in the cell cycle, whereas FHA2 only has one other known ligand: itself. The interaction between FHA1 and Dbf4 is dependent on an N-terminal region of Dbf4 called the H-BRCT motif, a classic BRCT domain with an extra α-helix (Matthews *et al.,* 2012). An interest of the Duncker lab was determining the surfaces that gave the FHA1 domain its ligand specificities as it was not known what surface of FHA1 interacted with Dbf4. By mapping a multiple sequence alignment onto known Rad53 FHA domain structures it was identified that a novel conserved lateral surface exists on FHA1, and

this region mediates the interaction between Rad53 FHA1 and the Dbf4 H-BRCT motif (Matthews *et al.,* 2014). This region consists of ten amino acid residues on four strands: β1 (Arg35, Ile37, Thr39), β7 (Asn112, Gly113), β10 (Gln126, Asp128, Gln129), and β11 (Val144, Phe146). This lateral surface was not as conserved in FHA2, showing that FHA1 gains ligand specificity by increasing the number of interaction surfaces. In addition to having been shown to mediate the interaction between FHA1 and Dbf4 via two-hybrid analysis, the mutation of this region also resulted in increased genotoxic sensitivity to both HU and MMS, highlighting its importance. This data has been expanded upon, as the Guarné lab at McMaster University recently solved the crystal structure of the H-BRCT region in complex with FHA1 (Alba Guarné , personal communication). This structure validates the role of the lateral surface of FHA1 in mediating the interaction with Dbf4, in addition to identifying several residues that may be important on Dbf4 for interaction with FHA1.

One of the major goals of my thesis research was to elucidate the exact mechanisms of interaction between FHA1 and several known ligands by using mutants that are known to disrupt the lateral surface as well as mutants that disrupt pThr binding site, two distinct regions of FHA1. This analysis was of crucial importance, to determine whether these lateral surface mutants abolish interaction with all FHA1 binding partners or are specific to disrupting Dbf4 binding.

A well-known FHA domain ligand that was a priority to test was Cdc7. Cdc7 is the kinase regulated by Dbf4 and together they comprise the Dbf4-dependent kinase (DDK), a complex that is important for DNA replication and checkpoint responses (Jackson *et al.,* 1993, Kitada e*t al.,* 1992). This interaction was important to test, as it is possible that the FHA1 lateral surface mediates the interaction with Cdc7 in addition to that with Dbf4. However, this seems

somewhat unlikely, as a candidate threonine for the FHA1-Cdc7 interaction has been identified in a high throughput screen, implying that this interaction is mediated through the classical pThr binding pocket and not the lateral surface (Aucher *et al.,* 2010).

In previous work by Bjergbaek *et al.* (2005), it was shown that both FHA1 and FHA2 interact with the central helicase domain (including RQC domain) of Sgs1, however FHA1 has a significantly stronger interaction. Sgs1 is a RecQ helicase which helps to maintain the integrity of the genome both by binding to Rad53, playing a role in the intra-S phase response in addition to binding Top3 to assist in the double strand break pathway. One of the mains goals of my thesis research was to elucidate whether this difference in Sgs1 affinity between FHA1 and FHA2 is due to the presence of the FHA1 lateral patch. The last ligand to be investigated was Sld3. Sld3 is a checkpoint protein that Rad53 FHA1 binds. It is hypothesized that this interaction may occur in order to halt DNA replication during the intra-S phase response. There is no literature on whether this interaction is canonical or not, making it a promising interaction to fully characterize.

For the present study, the role of the FHA1 lateral surface was identified across numerous ligands, showing that its requirement is variable in ligand specificity. We have uncovered that this region is not required for the interaction with Cdc7 and Sgs1, as these ligands are mediated through the pThr binding pocket. This region is required for interaction with Dbf4 and partially mediates the association with Sld3.

3.2 Results

3.2.1 FHA1 interactions with DDK

In order to better expand on the bioinformatics literature on interaction surfaces of FHA domains, FHA domains from a variety of proteins were examined to see if the lateral surface required for Dbf4 binding is conserved. To do this, core FHA domain folds were first identified in *Saccharomyces cerevisiae* (Rad53, Dun1, Mek1, Pml1, Fkh2, Tos4, Dma1, Xrs2) and subsequent homologs were identified in *Schizosaccharomyces pombe* (Dma1, Cds1, Nbs1, Tos4), *Xenopus laevis* (Nbs1), and *Homo sapiens* (Chfr, Chk2). The sequences were trimmed to include only the core FHA domain fold, in order to ensure that the alignment had high accuracy. Although FHA domains share a high structural similarity, they typically align very poorly on the sequence level. By eliminating extraneous sequence, this made the alignment of the core (where the pThr binding pocket and lateral surface are located) better. Sequence alignments performed in MUSCLE are shown in Figure 9.

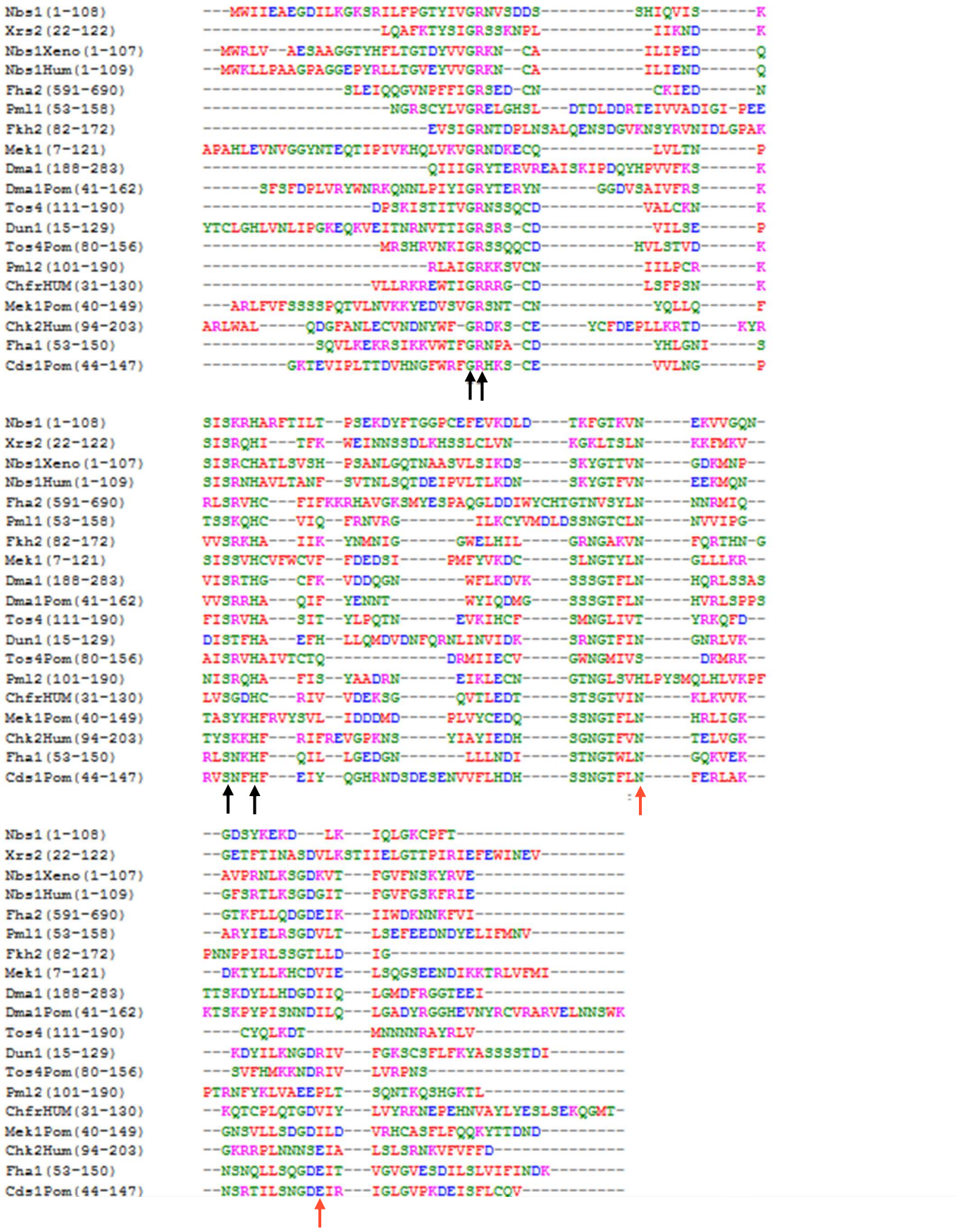

Figure 9: Sequence Alignment of Various FHA domains. Proteins labelled Pom from *S. pombe,* Xeno for *X. laevis*, and Hum for *H. sapiens*. All other sequences are from *S. cerevisiae*. Residues of the pThr binding pocket are denoted by a black arrow, other residues of interest are denoted by an orange arrow.

This initial alignment revealed several interesting conserved features of FHA domains. First, the conserved phosphothreonine binding pocket which consists of a key arginine is present and in an equivalent location for all FHA domains examined. In addition to the key arginine, other important residues for the pThr binding pocket are exclusively conserved across all FHA domains (glycine, serine, and histidine). This is shown by the residues highlighted with black arrows in Figure 9. This is an expected result and gives confidence the alignment is of high quality when examining other areas, particularly other potentially important residues. Two residues that were important for interaction of Rad53 FHA1 with Dbf4 were Asn112 and Glu129, residues that are distinct from the pThr binding pocket and located on β-strands 7 and 10 (Matthews *et al.,* 2014). These two residues are highlighted on the alignment with an orange arrow. Interestingly, these residues are conserved in Cds1 (*S. pombe)* and Chk2 (*H. sapiens*), homologs of Rad53. In the case of Asn112, it is in an eight amino acid region that appears to be important, as the degree of conservation implies that this region may be important for protein stability or function. In Rad53 FHA1, this region mainly consists of a short loop distinct from the pThr binding pocket.

Once the alignments were obtained, phylogenetic trees were constructed to examine the evolutionary relationships between these FHA domains. To do this, alignments were used as inputs for PhyML, a software package that is based on the maximum likelihood principle (Guindon, 2010).

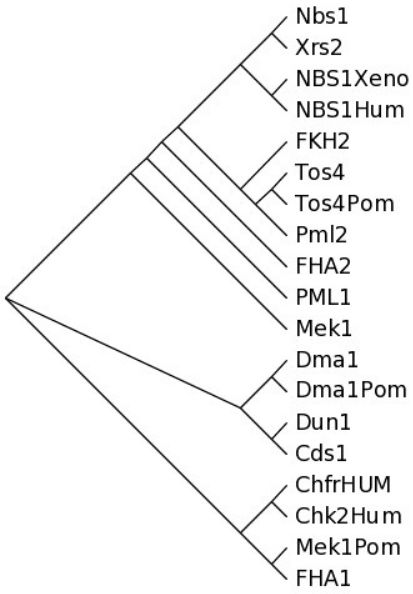

Figure 10: Phylogenetic tree of Various FHA domains. Alignment generated in figure 9 was used with PhyML to generate a phylogenetic tree. Proteins labelled Pom from *S. pombe,* Xeno for *X. laevis*, and Hum for *H. sapiens*. All other sequences are from *S. cerevisiae.* Bootstrapping values for 100 iterations can be found in Appendix A.

When looking at the tree presented in Figure 10 and the bootstrapping values in Appendix A, it gives confidence that the alignment was performed with high accuracy. Almost all homologous proteins are grouped together regardless of the organisms of origin. This is especially true of Nbs1/Xrs2, which make up an entire branch at the very top of the tree, and Chk2/Cds1/FHA1 at the bottom. For both of these cases, FHA domains from humans*, Xenopus*, fission and budding yeast are grouping together. Other proteins, such as Dma1 and Tos4 are conserved among yeast species. Due to this expected grouping of FHA domains in the tree, inferences drawn from the multiple sequence alignment are validated, as the MSA produced a phylogenetic tree that makes evolutionary sense. Taking in this work as a whole, it shows that residues distinct from the pThr binding pocket may be important. Of particular note, Asn112 a residue on the lateral surface of FHA1, is present in most eukaryotes studied, consistent with this

amino acid being important for ligand binding. Although there were no other regions that were completely conserved across all FHA domains aligned, several important regions were identified between orthologs, such as the region which corresponds with Rad53 FHA1 105-112. This region consists of a series of polar amino acids followed by an aromatic amino acid. This region and others should be further characterized through structural and biochemical approaches, similar to the ones used in this study.

Having determined that certain features of the lateral surface of FHA1 that mediates association with Dbf4 are conserved in other FHA domains, further ligands of FHA1 were next examined to see if this region could also be playing a role in their interactions. One of the first ligands that was investigated was Cdc7, the binding partner of Dbf4 in the DDK complex. In the following experiments, yeast two-hybrid analysis was conducted using the same FHA1 lateral surface mutants that abrogated the interaction with Dbf4 in addition to a R70A mutant, a previously well characterized mutation that disrupts the canonical phosphothreonine binding site.

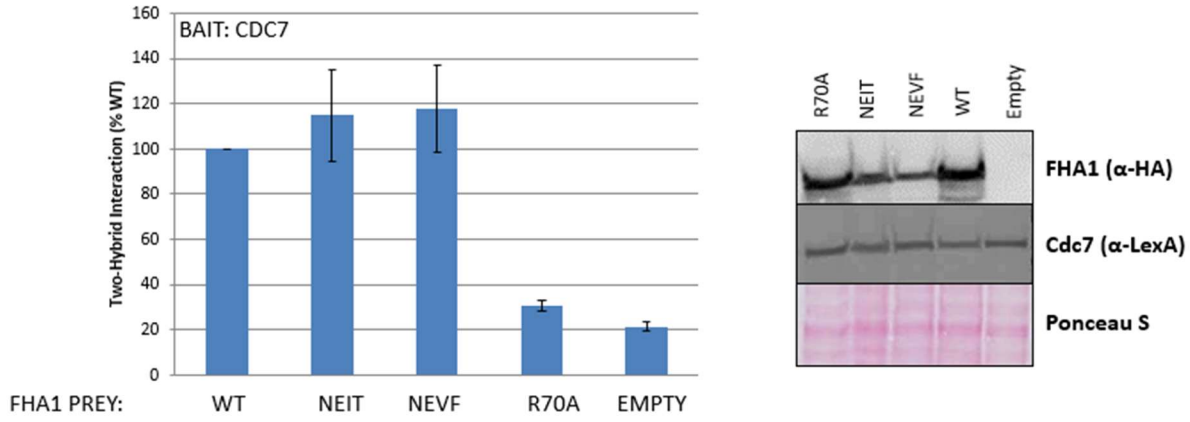

Figure 11: FHA1 Lateral surface mutants fail to abrogate the interaction with Cdc7.Yeast Two-Hybrid analysis shows (left) that the interaction between Rad53 FHA1 and Cdc7 is mediated through the phophothreonine binding pocket and not the lateral surface of FHA1. Immunoblot indicates protein expression levels (right). Experiments were performed in triplicate. Errors bars on graph represent standard error across three trials. NEIT=N112A/E129A/I37A/T39A, NEVF=N112A/E129A/V144N/F146A. Technical assistance was provided by Polina Myrox.

As Figure 11 shows, the interaction between Rad53 FHA1 and Cdc7 is mediated through the classical phosphothreonine binding pocket. The lateral surface mutants (NEIT, NEVF) failed to abrogate the interaction between FHA1 and Cdc7, whereas the R70A mutation did lower the interaction to background empty vector levels. This is consistent with the literature, as a previous high throughput screen identified a candidate threonine (T484) on Cdc7 that may be important for the interaction with FHA1 (Aucher *et al.,* 2010). As the identified key interaction residue on Cdc7 was a threonine, it was not expected that the lateral surface of FHA1 would play a major role in the interaction, as confirmed by our assays. Sequence encoding a mutation of this threonine to alanine (T484A) was generated using the Quickchange XL site directed mutagenesis system, in order to stop any phosphorylation from occurring at that site. Once this mutant two-hybrid bait construct was confirmed by sequencing, it was transformed into yeast in order to test its interaction with Rad53 FHA1 bait, compared to the wild type version of Cdc7. When T484 on Cdc7 is mutated to alanine the interaction between FHA1 and Cdc7 is reduced to 40-50% of WT levels (Figure 12).

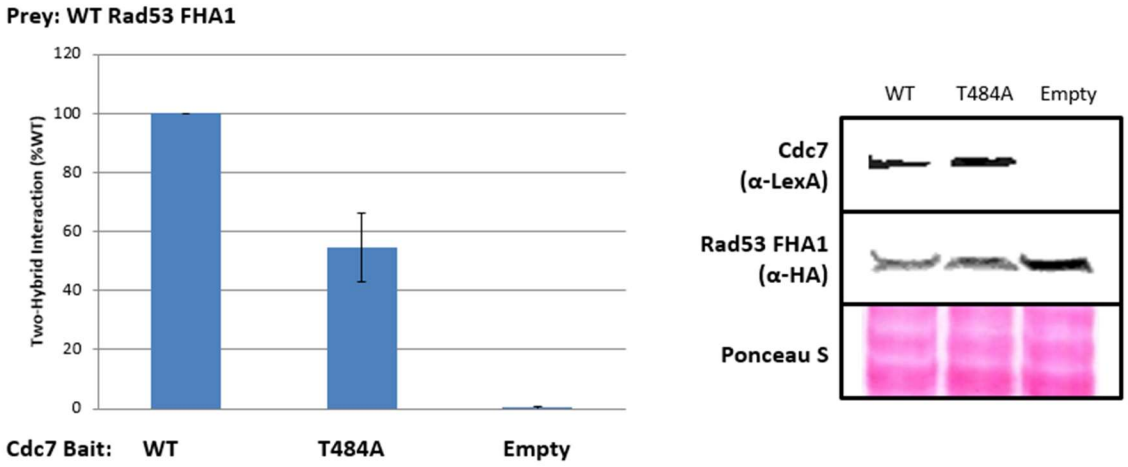

Figure 12: T484 plays a role in mediating the interaction between Rad53 FHA1 and Cdc7.Yeast Two-Hybrid analysis shows (left) that the interaction between Rad53 FHA1 and Cdc7 is partially abrogated by a T484A mutation. Immunoblot confirms equal expression levels (right). Experiments were performed in triplicate. Errors bars on graph represent standard error across three trials. Technical assistance was provided by Polina Myrox.

This alanine mutation changes the amino acid at position 484 to one that cannot be phosphorylated. Interestingly, the T484A mutation does not reduce the interaction between Rad53 FHA1 and Cdc7 to background levels. Although this was an initial puzzle, recent structural data from the Guarné lab has shed light on this result. Her group was recently able to co-crystalize Rad53 FHA1 simultaneously binding to a fragment (H-BRCT domain) of Dbf4 in addition to a short Cdc7 peptide including T484. Although the main amino acid mediating the interaction between FHA1 and Cdc7 is T484, there are other residues that are also responsible for partially mediating the interaction. There is an aspartic acid three amino acids upstream as well as several residues downstream of the threonine that in the crystal structure is shown to be partially assisting with the interaction (Alba Guarné, personal communication). It is possible that an R70A mutation may alter the pThr binding pocket structure in a way that would abrogate secondary interactions, where a T484A would not significantly alter the secondary structure, allowing for secondary interactions to occur and explaining the lack of reduction seen in Figure 12.

This co-crystal structure proved that the lateral surface of FHA1 is the interaction surface for Dbf4. In this structure, it was shown that the interaction surface has two contact points on the lateral surface on FHA1, which explains why certain quadruple mutations were able to disrupt the interaction between Rad53 FHA1 and Dbf4. Previous work mutating single and pairs of amino acids failed to disrupt the interaction because in these particular cases only one point of contact was being disrupted. The mapped interface showed that on Rad53, one contact point consisted of the amino acids N112 and G127 while the second contact point included I37 and F146. The amino acids that are responsible for interaction with FHA1 on Dbf4 at the first point are Y198 and K200, while the second contact point is mediated through W112 and W116.

Previous FHA mutants that reduced the interaction with Dbf4 contained amino acids at both contact points (N112 & F146) while mutants that failed to reduce interaction with Dbf4 maintained one point of contact. Similar to the FHA domain mutants, it was thought that an individual substitution of Dbf4 K200 to alanine would have minimal effects on FHA1-Dbf4 interaction, but if both contact points were disrupted, a complete abrogation would be seen. Two-hybrid analysis was conducted with single Dbf4 mutants in combination with WT FHA1, or a quadruple FHA1 mutant (IVTF) that failed to abrogate the interaction with WT Dbf4 in previous experiments.

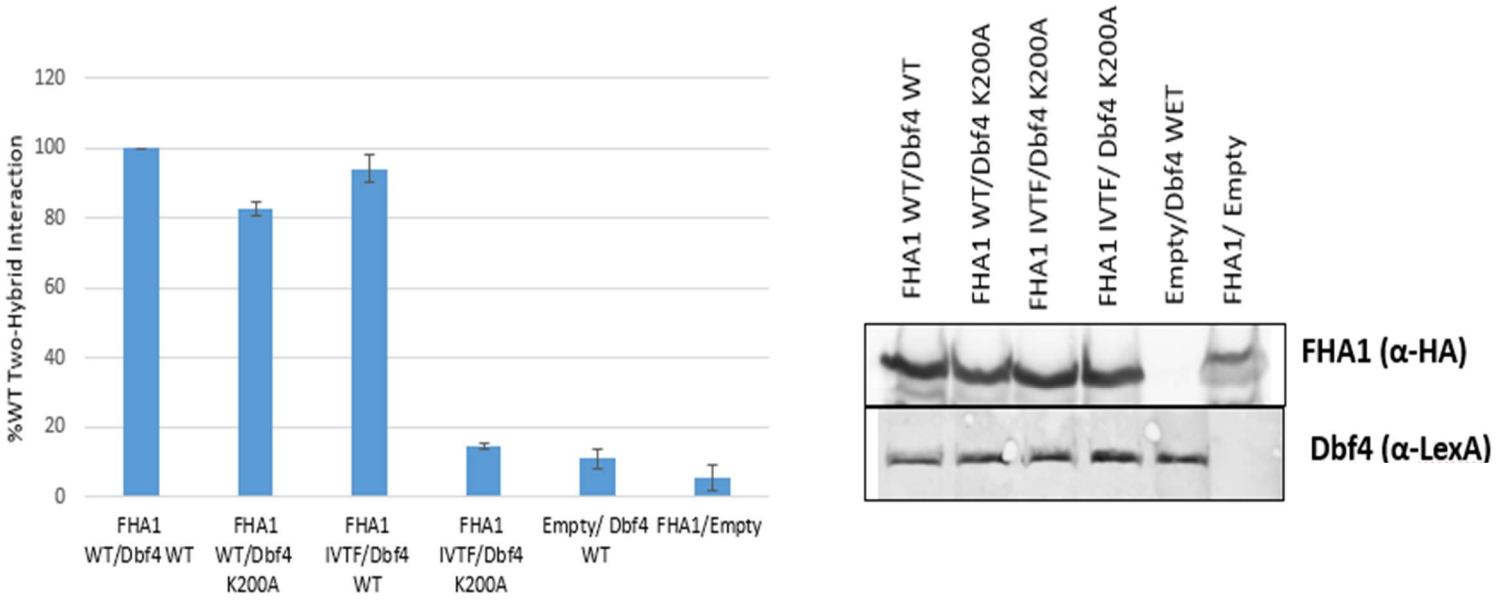

Figure 13: The FHA1-Dbf4 interaction is mediated by two contact points. Yeast Two-Hybrid analysis shows (left) that the interaction between Rad53 FHA1 and Dbf4 is abrogated when contacts point on both FHA1 and Dbf4 are disrupted. Immunoblot indicates protein expression levels (right). Experiments were performed in triplicate. Errors bars on graph represent standard error across three trials. IVTF= I37A/T39A/V144N/F146A.

Looking at the data in Figure 13, it supports the double contact point hypothesis. Both the Dbf4 K200A and FHA1 IVTF single mutants fail to abrogate the interaction on their own with their WT counterparts, showing approximately 80-95% of the double wild type interaction. Interestingly, when the single mutants were combined, a knockdown to empty vector levels was

observed. These mutants are at the two different sites of contact on FHA1 and Dbf4. This abrogation of Dbf4 interaction is similar to what was observed with the Rad53 FHA1 NEVF mutant, which disrupted the two contact points both on the FHA1 domain. This shows that if the two contact points are disrupted from the Dbf4 and/or FHA1 side, the interaction will be abrogated.

This data taken as a whole solidifies our knowledge of the Rad53 FHA1-Dbf4 interaction. FHA1 is responsible for mediating the interaction with Cdc7 via the phosphothreonine binding site and with Dbf4 via the newly identified lateral surface. The threonine on Cdc7 that is responsible for mediating its interaction with Rad53 FHA1 is T484, while the interaction between FHA1 and Dbf4 is mediated through the H-BRCT domain of Dbf4, with two contact points being present.

3.2.2 Characterizing the Rad53 FHA1-Sgs1 Interaction

One of the major questions coming out of this initial result was the degree to which the Rad53 FHA1 lateral surface is involved in binding yet other ligands. As there was a variable requirement for this region in binding the members of the DDK complex, it was of interest to investigate whether this region could play a role in binding other checkpoint ligands.

Another previously identified ligand of Rad53 FHA1 is the greater helicase region of Sgs1 (Bjergbaek *et al.,* 2005). I confirmed this through two-hybrid analysis using the FHA1, FHA2 and helicase domains of Rad53 (Figure 14). It was interesting to see that FHA1 had a significantly higher interaction when compared to FHA2. An initial hypothesis was that the

lateral surface may play a role in the FHA1-Sgs1 interaction, as it is present in FHA1 and not found in FHA2.

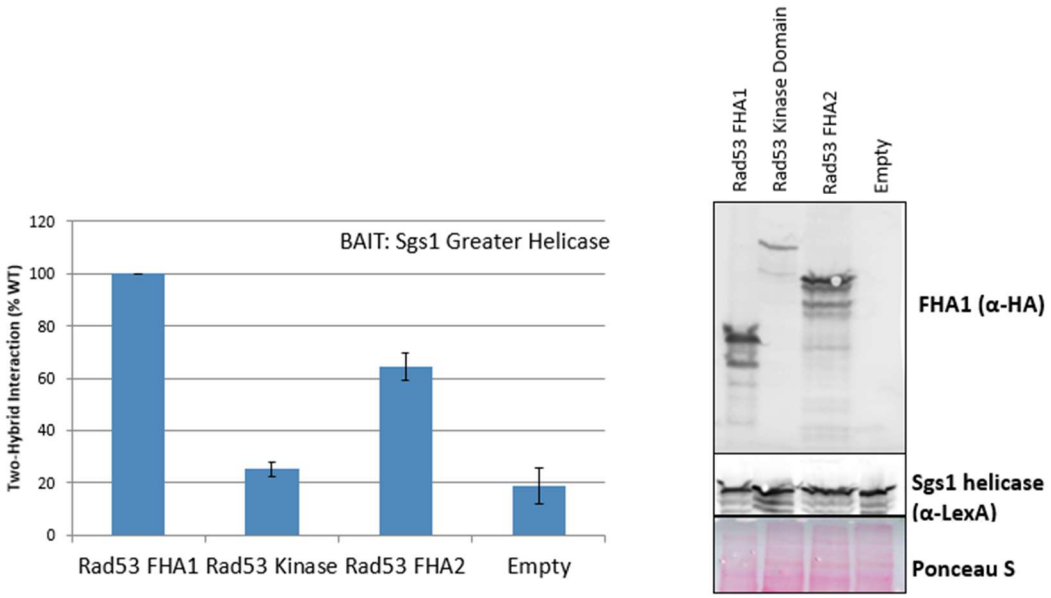

Figure 14: Yeast two-hybrid analysis of Sgs1 greater helicase region (AA 291-1129) Both FHA1 and FHA2 interact with the Sgs1 greater helicase region, with FHA1 showing significantly stronger association (left). Immunoblot indicates bait and prey expression levels (right). Experiments were performed in triplicate. Errors bars on graph represent standard error across three trials.

The next step was to perform two hybrid analysis with the same FHA1 mutants that abrogated the interaction between Rad53-Dbf4. In addition to the FHA1 lateral surface mutants that previously abrogated the interaction with Dbf4, the FHA1 R70A mutant was included.

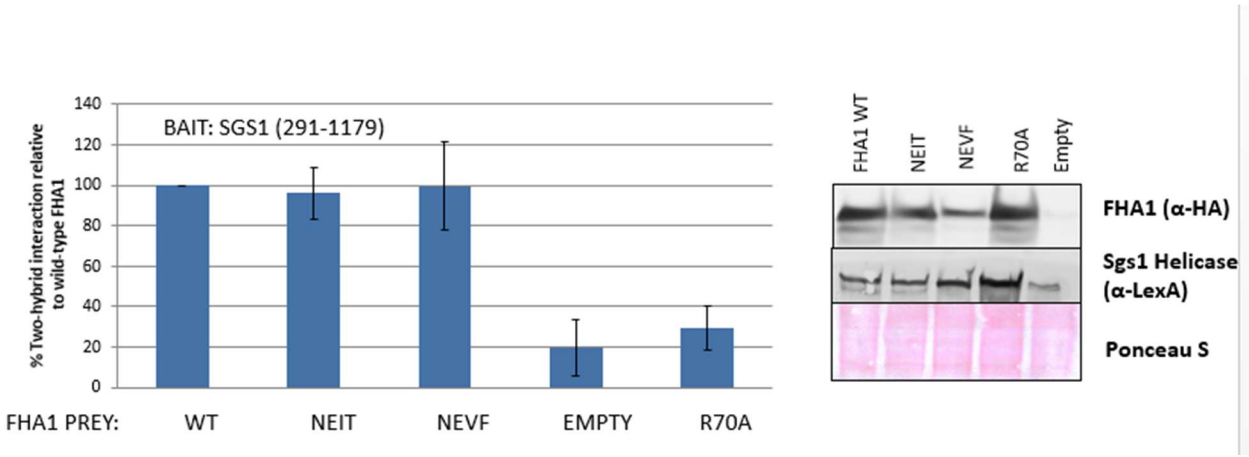

Figure 15: FHA1 lateral surface mutants do not mediate interaction with the Sgs1 greater helicase region (AA 291-1129). FHA lateral surface mutants do not show a reduction in interaction when compared to the WT. FHA1 R70A shows a reduction in interaction to empty vector levels (left). Immunoblot indicates bait and prey expression levels (right). Experiments were performed in triplicate. Errors bars on graph represent standard error across three trials. NEIT=N112A/E129A/I37A/T39A, NEVF=N112A/E129A/V144N/F146A.

Consistent with the literature, it appears that the interaction between Sgs1 and Rad53 FHA1 is mediated through the pThr binding pocket (Hegnauer *et al.,* 2012). When looking at the FHA1 lateral surface mutants, both mutants tested (NEIT, NEVF) interacted with Sgs1 at wildtype levels. Due to the interaction levels of the R70A mutant being near empty vector levels, strong evidence is provided that this interaction is mediated through the pThr binding pocket.

3.2.3 Characterizing the Rad53 FHA1-Sld3 Interaction

The last ligand of FHA1 that was examined for the importance of the lateral patch was Sld3. Sld3 interacts with Rad53 in the intra-S phase checkpoint response, which can stop DNA replication by inhibiting the CDK pathway (Zegerman & Diffley, 2010). Initially, lateral surface mutants were examined via two-hybrid analysis to determine their role in Sld3 binding. Unfortunately, a positive signal for Sld3 and wild-type FHA1 could not be obtained in this assay. It was hypothesized that the N-terminal tag (LexA DNA binding domain) added to Sld3 for the two-hybrid analysis may have interfered with it binding to FHA1. To get around this, co-immunoprecitation (Co-IP) assays were used to detect protein-protein interactions. This solved

the perceived problem as the Co-IP vector uses a C-terminal Myc tag for immunoprecipitation. A Co-IP was performed with Sld3-Myc and three HA-tagged versions of FHA1 (WT, R70A, and NEVF), as described in Materials and Methods.

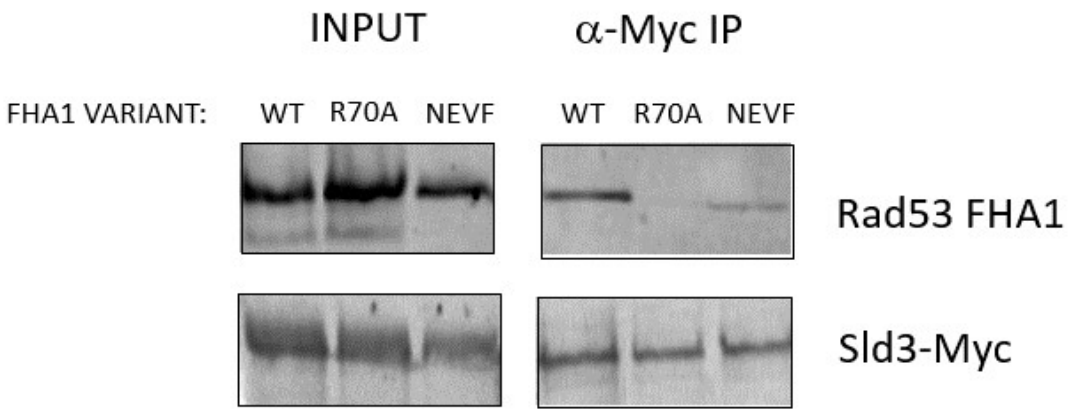

Figure 16: Co-immunopreciptation of Rad53 FHA1 and Sld3. (Left) Input. (Right) Immunoprecipited samples as detected by Western blot. NEVF=N112A/E129A/V144N/F146A. Experiment performed by Boris Simovic.

The Co-IP data showed that there is a partial requirement for the FHA1 lateral patch in interacting with Sld3. When looking at the relative amounts of pulldown, the WT version of FHA1 has the strongest pull down levels (darkest band) as expected. This pulldown is almost completely abrogated in the R70A mutant, showing that the pThr binding domain is required in order to bind Sld3. Lastly, the NEVF mutant showed an intermediate level of interaction compared to WT and R70A mutant. This data is consistent with the lateral surface of FHA1 playing a role in interacting with Sld3, but not at the primary interaction surface.

3.3 Discussion

Overall, this data suggests that the lateral surface of FHA1 is variably required for ligand binding and specificity. This surface is absolutely required for interactions with Dbf4, and interestingly, is conserved in higher eukaryotes. This level of conservation of the FHA1 lateral

surface implies that such a non-canonical interaction interface may be more widespread than previously thought. As we have shown in this work, the interaction between Rad53 FHA1 and Dbf4 depends on two contact points. Furthermore, a complete abrogation of the interaction can be seen by mutating both contact points, either on Dbf4 or FHA1. This interaction interface shares some similarities with Chk2 and BRCA-1, which interact in humans. In this association, the FHA domain in Chk2 has two binding surfaces: the pThr binding pocket and a conserved hydrophobic lateral region (Li *et al.,* 2002). Interestingly, mutation at either contact point stops the interaction, and mutation of the lateral surface on Chk2 has been linked to Li-Fraumeni syndrome, a precursor to cancer in multiple organs (Bell *et al.,* 1999). This lateral patch on Chk2 is not in a similar position to that of Rad53 FHA1, showing that these regions are extremely plastic in determining ligand specificity. Due the extreme importance of Rad53 and Chk2 interaction with Dbf4 and BRCA-1 in halting late origin firing and double strand repair, respectively, it makes intuitive sense that cells have developed strict selectivity mechanism to mediate these associations.

When looking at other ligands it was seen that the conserved lateral region in FHA1 is not required for binding to Cdc7 or Sgs1, as these interactions are mediated through the pThr binding patch. This was because lateral surface mutants failed to show a reduced interaction, whereas an R70A mutation completely abrogated the interaction. The most likely Thr residue on Cdc7 is T484, as mutation of this residue to an alanine reduces the interaction to approximately 40% of the wild-type level. This is consistent with the literature, as Cdc7 T484 had previously been identified in a high throughput screen (Aucher *et al.,* 2010). For Cdc7, this was a critical experiment to conduct. This result confirms that the interaction abrogation seen in FHA1-Dbf4 two hybrids was not due to endogenous Cdc7 mediating an indirect interaction. As FHA1 binds

Cdc7 via its pThr binding site while Dbf4 binds via the FHA1 lateral surface, it is hypothesized that the FHA1-DDK interaction is bipartite in nature. For Sgs1, the results I obtained were consistent with previous work showing a loss of interaction with the R70A mutant (Hegnauer et al., 2012) but also raise more questions. Using bioinformatics methods, the only difference in conserved regions between FHA1 and FHA2 was the lateral binding surface on FHA1. As this region is not involved in Sgs1 binding, another surface must give FHA1 its high ligand specificity with Sgs1 when compared to FHA2.

For Sld3, the lateral surface is partially required for binding FHA1, consistent with this not being the primary interaction surface. Via co-immunoprecipitation, it was seen that a lateral surface mutant had a pull-down level in between WT FHA1 and FHA1 R70A. Although beyond the scope of this project, future research could be directed in order to find the minimal region of Sld3 necessary to interact with FHA1. While budding yeast Sld3 differs considerably in size and sequence when compared to its orthologs, one region is conserved across all proteins in the Sld3 family, the C-terminal treslin domain, so a future experiment could look at whether it is required for interaction with FHA1 (Itou *et al.,* 2014). With this data, it may be possible to narrow down the possible threonines that may be mediating the interaction.

Future work can be done using a similar approach to the one outlined here to identify novel interaction surfaces on other FHA domains that can determine ligand specificity. As there are thirteen FHA domains in *S. cerevisiae* there are numerous possibilities, and more information about non-canonical FHA domain interactions is yet to be gathered. This work shows that FHA domains can have more than one interaction surface, and that these interaction surfaces are important in conferring ligand specificities.

Chapter 4: Characterization of the importance of the lateral surface of various additional *S. cerevisiae* FHA domains

4.1 Introduction

The fork-head associated domain (FHA) is a protein-protein interaction domain whose most characterized feature has been its interaction with phosphorylated threonines via a pThr binding pocket. These domains have been identified from prokaryotes to eukaryotes in over 2000 proteins to date (Hofmann and Bucher, 1995). These domains are involved in numerous processes, at every stage of the cell cycle, including cell signalling, the DNA damage response and cell growth. Recent work in the Duncker laboratory has focussed on characterizing the FHA1 domain of Rad53, and how it interacts with its binding partners. This protein is responsible for mediating the intra-S phase checkpoint response, which helps to stop DNA replication during times of cellular stress. In addition to its pThr binding pocket, FHA1 has a second conserved lateral surface that is important for ligand specificity, and is required to varying degrees for different ligands. In addition to Rad53 FHA1, there are 12 other FHA domains that have been characterized in budding yeast. Interestingly, the majority of these domains have functional equivalents higher eukaryotes.

As part of the effort to fully characterize the role of the Rad53 FHA1 lateral surface in ligand specificity, initial experiments were conducted to determine the importance of this region in binding Dbf4, a known FHA1 ligand (Duncker *et al.,* 2002). Given that this region was shown to be important for Dbf4 binding, it was hypothesized that mutations in the corresponding regions in other structurally similar FHA domains may also change ligand affinity. Previous unpublished research by the Duncker laboratory and Brendan McConkey (University of Waterloo) has shown that there are other FHA containing proteins that have second conserved lateral patches similar to Rad53 FHA1. These lateral patches were present on most FHA domains studied, and the relative position of these surfaces varied in relation to the classical pThr binding domain. Three proteins of particular interest that have this lateral patch are DNA-damage

Uninducible1 (Dun1), Meiotic Kinase 1 (Mek1) and Plasmid Maintenance 2 Plm2). Dun1 is a serine-threonine kinase that is involved in the DNA damage response by phosphorylating both Rad55 and Sml1, which can arrest the cell cycle (Tsaponina & Chabes 2013). It has a lateral surface that is conserved on the surface opposite of the pThr binding pocket, similar to Rad53 FHA1. In addition to its interaction with Sml1, Dun1 further plays a role in regulating dNTP levels in cells by binding to other ligands. Two particular proteins involved in this regulation are Dif1 and Crt1, both of which are involved in regulating ribonucleotide reductases (RNRs). RNRs are enzymes that catalyze the formation of dNTPs from ribonucleotides. Dif1 is a protein that is involved in the import of RNRs into the nucleus, whereas Crt1 is responsible for the transcriptional regulation of RNRs. Mek1 is meiotic serine/threonine kinase that acts as a meiotic checkpoint factor, suppressing double strand breaks during homologous recombination (Wan *et al.,* 2003). The conserved lateral surface on Mek1 is on the same protein face as the pThr binding pocket, differently positioned compared to those of both Rad53 and Dun1. Plm2 is a FHA containing protein that arose from a whole genome duplication. This protein is a target of the SBF transcription factor and is activated in the DNA damage response in addition to its role in extending telomeres (Huang & Elledge, 2000; Nautiyal *et al.,* 2002). Plm2 FHA's conserved lateral region is on the opposite face compared to the position of the pThr binding pocket.

4.2 Results

4.2.1 The role of conserved Lateral Surfaces in Ligand Specificity for additional FHA domains

A multiple sequence alignment of the domains of interest (Rad53 FHA1, Dun1 FHA, Mek1 FHA, Plm2 FHA) was performed in MUSCLE (Edgar, 2004). A custom program was then created in Python in collaboration with Jeremy Adams from Andrew Doxey's lab (University of

Waterloo) to classify these domains based on their similarity to the lateral patch of FHA1. Blossum62 scoring matrix was the scoring system used and the results of this analysis can be found in Table 2. The Blossum62 matrix assigns scores based on the probability of each possible amino acid substitution. This allows for more accurate conservation scores compared to simple match/mismatch scores when looking at the particular conserved residues. Based on the amino acids one is looking at, a maximum score is established. An alignment with a perfect score would mean that all residues being examined are completely conserved, with any other scores indicative of differing degrees of conservation. Any score above zero is indicative of some level of conservation, as the residues in questions appear to be similar.

Table 2: Blossum 62 scores for three FHA containing proteins when compared to the lateral patch residues of Rad53 FHA1.

FHA-Containing Protein	Lateral Patch Conservation Score (7 Amino Acids)
Dun1	16/32
Mek1	11/32
Plm2	14/32

When looking at these results, one can see that Dun1 FHA is more similar to the lateral patch of Rad53 FHA1 than Mek1 FHA, with Plm2 having an intermediate score. It is important to note that this analysis was done for seven out of ten lateral patch amino acids as three (Arg35, Ile37, Thr39) do not have clear alignments in Dun1 and Mek1 to Rad53 FHA1. These three amino acids on Rad53 FHA1 align to gap regions. To better understand how well the lateral patch is conserved in these proteins, a structural alignment was performed in USCF Chimera and pseudo Blossum62 scores were given for the Dun1 and Mek1 gap regions based on the closest

residues to the conserved patch residues. At this point, the decision was made not to perform this structural analysis with Plm2, as previous two-hybrid experiments conducted by a previous undergraduate thesis student showing Dun1 FHA and Mek1 FHA have different affinities for Dbf4 (E. Hazra, B.J. McConkey and B.P. Duncker, unpublished). Due to these differential WT affinities, Plm2 was dropped as the role of the FHA1 lateral patch on Dbf4 binding could be determined by using Dun1 and Mek1 FHA. The structural alignment was done in collaboration with Jeremy Adams, as he was involved in writing the program that determined computationally the closest residues to Rad53 FHA1 in the Dun1 and Mek1 gap regions. Images of the structural alignments can be found in Figure 17.

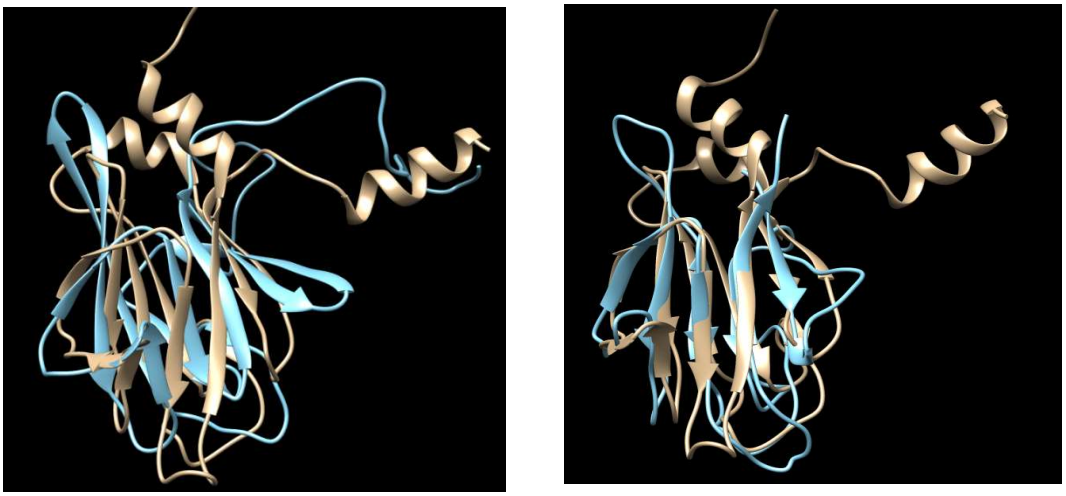

Figure 17: Structural alignment of Dun1 FHA and Mek1 FHA against Rad53 FHA1. (Left) Structural alignment of Dun1 FHA (Blue) against the template Rad53 FHA1 (Brown). **(Right)** Structural alignment of Mek1 FHA (Blue) against the template Rad53 FHA1 (Brown).

The root-mean-square deviation (RMSD) is a measure of how well superimposed proteins align, as it measures the average distance between specific atoms. The RMSD for the Dun1 alignment as 5.8 and the RMSD for Mek1 was 2.6. This indicates that Mek1 has a very good structural alignment with Rad53 and Dun1 has a moderately good structural alignment with Rad53. The quality of this structural alignment is important as it allows one to make inferences about the closest residues structurally to gap regions in our alignment. After assigning pseudo

Blossum62 scores to gap regions in both Dun1 and Mek1, the previous table can be updated to reflect the all 10 lateral patch residues (Table 3). With the addition of these three residues, the maximum Blossum62 score goes from 36 to 54.

Table 3: Blosum 62 scores for Dun1 and Mek1 when compared to the lateral patch residues of Rad53 FHA1.

FHA-Containing Protein	Surface Patch #2 Conservation Score (10 Amino Acids)
Dun1	15/54
Mek1	14/54

It is important to note that although they have similar overall scores, the residues that are conserved for these two proteins, relative to Rad53 FHA1, differ. This information was important for choosing residues to be altered by site directed mutagenesis in order to test which of them in the FHA1 lateral patch are most important for ligand specificity. Previous research (Matthews et al., 2014) has shown that Glutamic acid 129 is an important residue for FHA1-Dbf4 ligand specificity, making the corresponding amino acid in other FHA domains an ideal candidate for mutation. The Dun1 FHA domain was mutated such that a polar arginine residue was changed to a nonpolar leucine residue, while for the Mek1 FHA domain, mutation from a non-polar valine to a glutamic acid was carried out. It was hypothesized that this would decrease Dun1 FHA binding to Dbf4 (as it would be “less” FHA1 like) and increase Mek1 FHA binding to Dbf4 (as it is would be “more” FHA1 like). Looking at the preliminary Dun1 data in Figure 18, there is a slight decrease in Dbf4 binding, however judging by the error bars this decrease does not appear significant. This is not the result we expected, and more research needs to be done to investigate why Dbf4 binding was not knocked down further. An initial hypothesis is

that the presence of other conserved lateral surface residues is mediating the interaction with Dbf4 in the absence of Glu129. Interestingly, the mutation did knock down Rad9 binding for all trials performed. Future experiments can focus on the role of the lateral surface in interaction with Rad9, as this is a previously uncharacterized interaction in the lab.

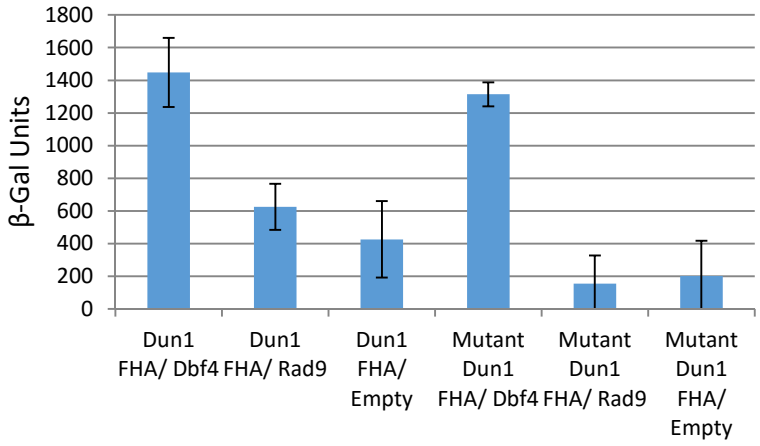

Figure 18: Mutation of Dun1 FHA domain (Arginine 124 to Leucine) leads to no change in binding to Dbf4. Yeast two-hybrid analysis of WT and mutant Dun1 FHA domains with Dbf4 and Rad9 ligands. Experiments were performed in triplicate, and error bars represent standard deviation. Preliminary data, immunoblotting not shown.

When looking at the Mek1 data (Figure 19), we see a result consistent with our hypothesis. There is a significant increase in Mek1-Dbf4 interaction when valine is mutated to glutamic acid at position 129. The mutant Mek1-Dbf4 interaction is approximately six times stronger when compared to the WT Mek1-Dbf4 interaction (Figure 19, A). This increase in interaction shows that the corresponding to Rad53 FHA1 Glutamic acid 129 in Mek1 FHA is important as it strengthens the Mek1 FHA/Dbf4 interaction *in vivo*. An important note is that although there were larger than expected error bars for both sets of two-hybrid data, all trends stayed proportional and consistent over three trials (i.e. mutant Mek1 always had stronger Dbf4 interaction when compared to WT, Mutant Mek1 always had stronger Dbf4 interaction when

compared to WT). This work as a whole highlights the importance of the lateral surface of FHA domains in conferring Dbf4 affinity, as these properties can be altered in Mek1 FHA.

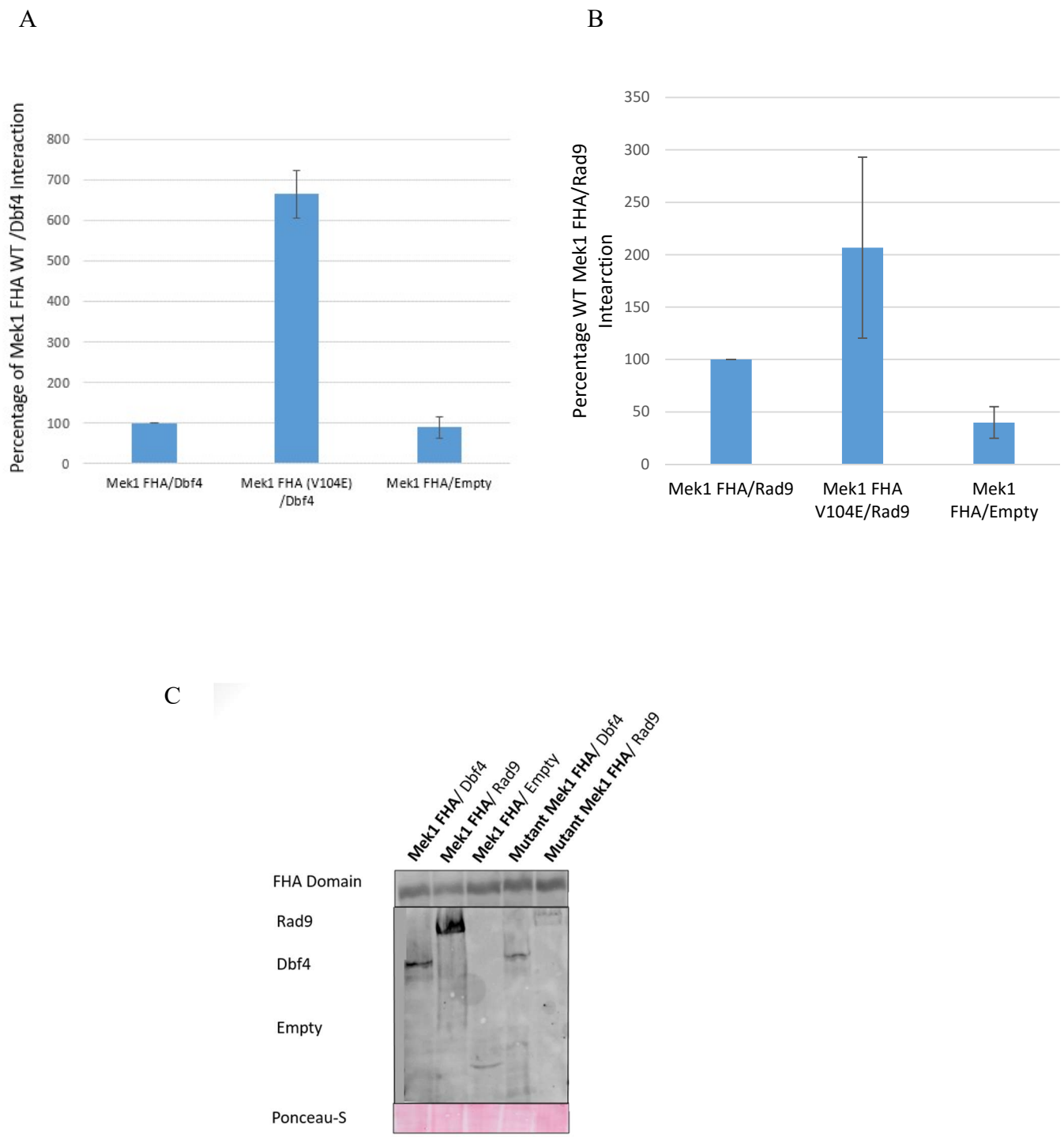

Figure 19: Mutation of Mek1 FHA domain (Valine 122 to Glutamic Acid) leads to increased association with Dbf4. Yeast two-hybrid analysis of WT and mutant Mek1 FHA domains with Dbf4 and Rad9 ligands. Experiments were performed in triplicate. Errors bars on graph represent standard deviation Anti-HA immunoblot analysis of both WT and mutant Mek1 FHA domain expression (Top). Anti-lexA immunoblot analysis of Dbf4 (FHA1 ligand), Rad9 (FHA2 ligand) or empty vector (just encodes HA-tag) expression. Ponceau-S staining was performed to assess protein loading (Bottom).

4.2.2 Lateral surfaces on other FHA Domains

As previously shown, Rad53 FHA1 has lateral surface residues that can mediate interaction with Dbf4. These residues can be mutated to abrogate the interaction and conversely, FHA1 like properties can be attained when these conserved residues are mutated in other FHA domains. As a very interesting avenue of research, a similar approach to what was conducted with FHA1 was carried out with other FHA domains. Using PSI-BLAST, an MSA file was constructed by probing the NCBI database for closely related yeast and fungal orthologs to known *S. cerevisiae* FHA domains. Once this MSA file was complete, the conservation of residues was either mapped onto known crystal structures from the PDB or structures generated using modelling software. FHA domains in budding yeast underwent this protocol, with three representative domains shown in Figure 20 with an additional figures in Appendix B.

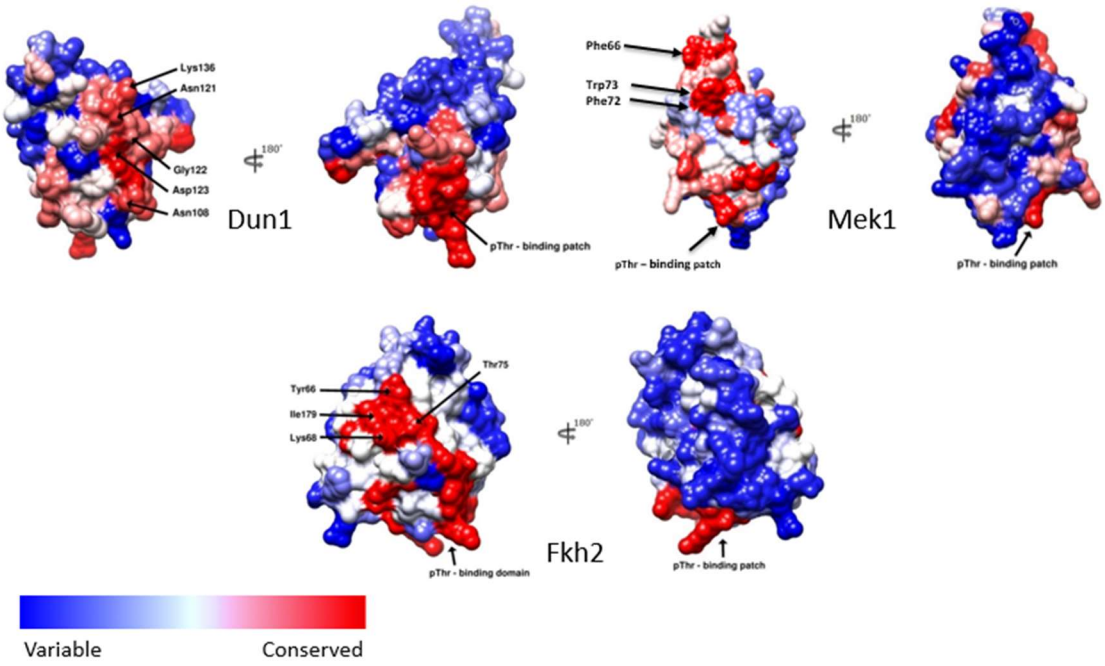

Figure 20: Several FHA domains have conserved lateral patches. (Top Left) MSA file of Dun1 FHA orthologs mapped into PDB structure 2JQJ. **(Top Right)** MSA file of Mek1 FHA orthologs mapped into Rosetta model. **(Bottom)** MSA file of Fkh2 FHA orthologs mapped into Rosetta model.

Looking at Figure 20, it is clear that our models were created with high accuracy. First, although FHA domains typically do not have well conserved sequences, FHA domains are similar on a structural level. This structure is a beta-sandwich connected by short loops. As Figure 20 shows space filling models, it is impossible to see these structural features, however in all cases proteins are orientated such that the majority of the short helical loops are facing towards the bottom. In addition to the similar structure, in all three FHA domains shown the pThr binding pocket is conserved, as denoted by a group of red residues at the bottom of each structure. In this diagram the more conserved a residue is, the more red it appears and the less conserved a residue is the more blue it appears. Clearly evident on each protein is a second well conserved domain that is just as conserved as the pThr binding site. Although present on all three proteins, there appears to be three variations of the location of these conserved regions in relation to the pThr binding pocket. On Dun1, the conserved lateral surface is directly opposite to its pThr recognition domain, similar to Rad53 FHA1. This makes sense as Dun1 and Rad53 are two evolutionarily closely related proteins, whose function is replaced by one protein in higher eukaryotes (Zhou and Elledge, 2000). On Mek1, the conserved surface appears to be on the same side as the pThr binding pocket. Lastly, Fkh2 has a conserved lateral surface located 90° to the pThr binding site, in a structurally close but distinct region of the protein. When looking at the FHA domains with conserved lateral surfaces, they fell into one of the above categories in terms of the location of the lateral surface. One representative image can be found in Figure 20 and Appendix B of FHA domain pairs that have arisen in *S. cervisiae* due to gene duplication events (Far10/Vps64, Fkh1/Fkh2).

Once these additional conserved lateral surfaces were identified on FHA domains in budding yeast, it was decided that Dun1 FHA would be the first examined in detail. Two reasons

Dun1 FHA was a promising domain of study are its evolutionary relationship with Rad53 and the fact that the crystal structure of Dun1 has been solved, showing that it takes the typical shape of FHA domains, forming a β-sandwich with the pThr binding site on a short loop (Lee *et al.,* 2008). Of the amino acids identified as part of the lateral patch, five residues appeared to be extremely important, Asn108, Gly122, Asp123, Asn121 and Lys136. Yeast two-hybrid vector pJG4-6 Dun1 FHA WT had already been created in the lab by a previous student, Ema Hazra. Mutants were generated using the Quickchange XL system, initially with a triple mutant being generated (Asn108A, Gly122A and Asp123A). The triple mutant was created first due to only three important residues being identified in an initial bioinformatic screen. Due to a lack of interaction abrogation (results not shown), further mutations to Dun1 FHA were made. Quadruple (Asn108A, Gly122A, Asp123A and Asn121) and penta mutations (Asn108A, Gly122A, Asp123A, Asn121 and Lys136A) were made using the same protocol. In addition to the lateral surface mutants, a pThr binding site mutant was also created. By aligning Dun1 FHA against Rad53 FHA1, it was clear that the analogous residue to FHA1 R70 is Dun1 FHA R60 (data not shown). This mutant was created using pJG4-6 Dun1 FHA WT as a template and included in the two-hybrid analysis. This was done in order to determine what the primary interaction surface on Dun1 FHA was during interactions with its ligands. Mutations for all residues tested were to alanine, as it is a non-bulky, chemically inert amino acid that does not significantly alter secondary structures. These Dun1 FHA mutants and WT plasmids were tested for their interactions with Sml1, a known Dun1 ligand (Zhao and Rothstein, 2002). This ligand was chosen due to the similarities between Dbf4 and Sml1, as both FHA domain ligands have alpha-helices. These helices were identified using IUPRED and I-TASSER two programs that

determine how structured proteins regions are and create a structural model (Dosztányi *et al.,* 2005; Yang *et al.,* 2015).

The two hybrid results in Figure 21 show that the lateral surface residues of Dun1 play a role in mediating the interaction with Sml1.

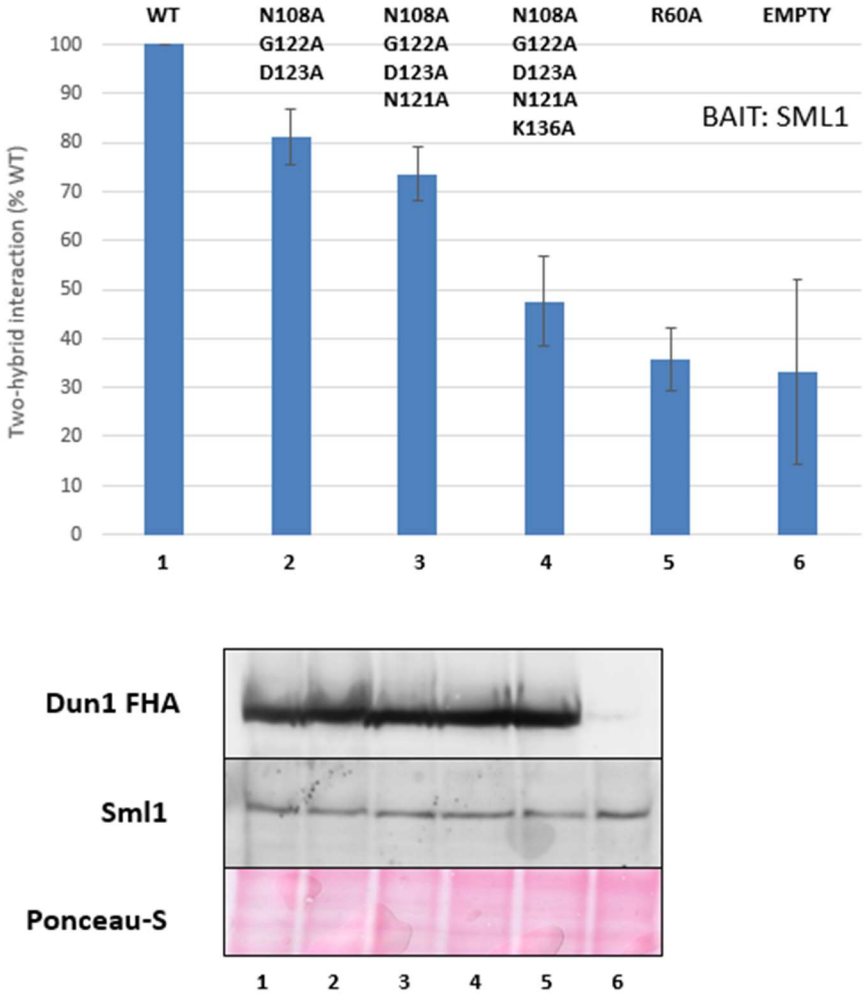

Figure 21: The Lateral Surface of Dun1 FHA mediates the interaction with Sml1.Yeast Two-Hybrid analysis shows (top) that the interaction between Dun1 FHA and Sml1 is abrogated when lateral surface residues are disrupted. Immunoblot confirms equal expression levels for both bait and prey proteins (bottom). Experiments were performed in triplicate. Errors bars on graph represent standard deviation.

When looking at the triple Dun1 FHA mutant (lane 2) very minimal knockdown is observed (~20%). This reduction is slightly increased when a quadruple Dun1 FHA mutant is

used and is reduced to empty levels with the penta Dun1 FHA mutation. Importantly, the penta mutation of the lateral surface of Dun1 FHA is at similar levels in the two hybrid assay when compared to R60A, the pThr binding pocket mutation. Looking at the immunoblot, expression levels are equal for all constructs, showing that any reduction of binding in the assay is due to decreased affinities. This result implies that the Dun1 lateral surface plays a significant role in mediating the interaction with Sml1. Taking lessons from Rad53 FHA1, the reasons such a high number of mutations need to be made before a decrease in binding affinity is seen may be due to the presence of multiple contact points between the two proteins. With the correct amino acids mutated to disrupt all contact points, it may be possible to see reduced interaction levels with fewer mutations. Another trend seen both in Dun1 FHA and Rad53 is that mutation of either the lateral surfaces or the pThr binding pocket appear to knockdown interaction levels.

In order to fully characterize the role of the Dun1 FHA lateral patch in ligand specificity, other binding partners needed to be examined. One particular binding partner of interest was Rad53. It has been previously shown that Dun1 and Rad53 interact, with Dun1 FHA binding through its pThr binding site (including R60) to two threonines in the N-terminal SQ/TQ cluster domain of Rad53 (Lee, H. *et al.,* 2008). By examining the role of the Dun1 FHA lateral patch in Rad53 binding it can be determined whether mutation of this surface also disrupts this interaction. Rad53 WT was cloned from the existing pJG4-6 Rad53 WT two-hybrid prey plasmid and ligated into the pEG202 two-hybrid bait plasmid. This Rad53 bait construct was tested in combination with pJG4-6 Dun1 WT, penta mutant and R60A FHA prey constructs.

Looking at the immunoblot results in Figure 22, all prey and bait protein expression levels are equal, which allows inferences to be made about binding affinities.

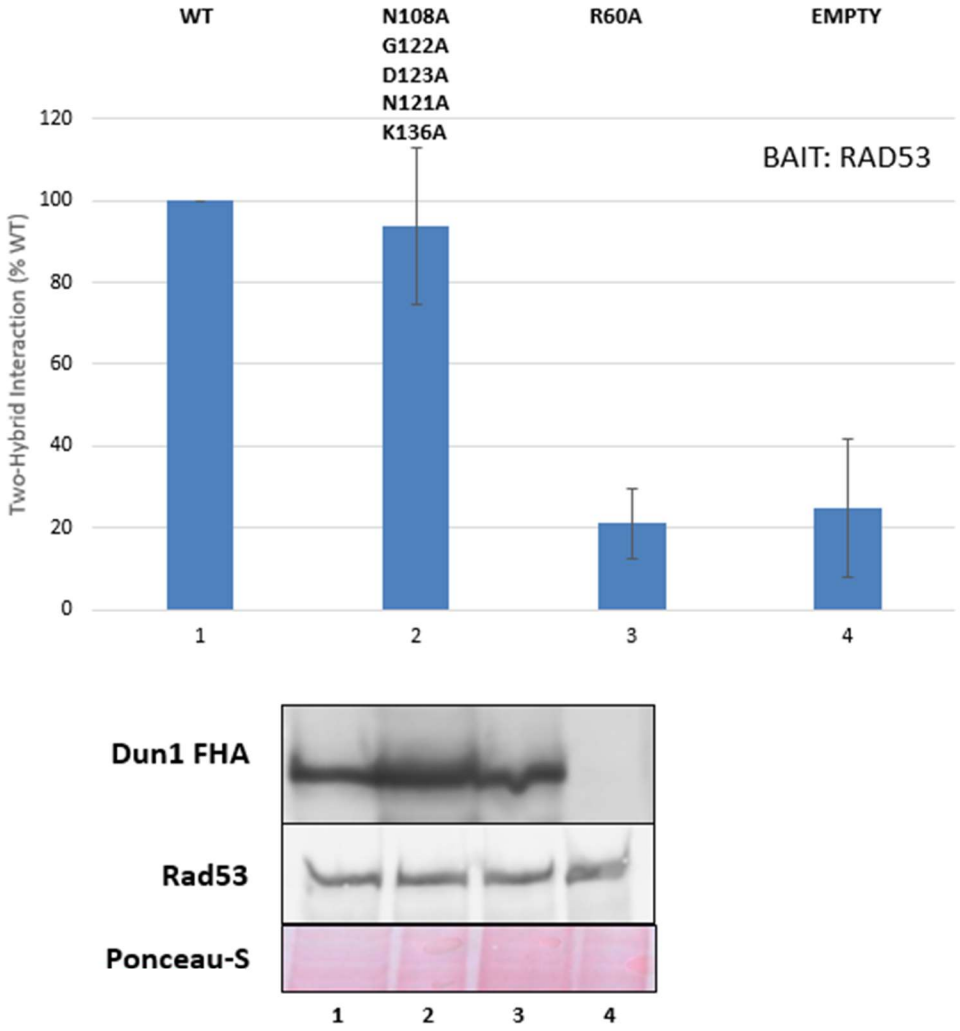

Figure 22: The Lateral Surface of Dun1 FHA is not required for interaction with Rad53.Yeast Two-Hybrid analysis shows (left) that the interaction between Dun1 FHA and Rad53 is not abrogated when lateral surface residues are disrupted. Immunoblot confirms equal expression levels (bottom). Experiments were performed in triplicate. Errors bars on graph represent standard error.

The Dun1 lateral surface mutations that reduced the interaction with Sml1 failed to reduce the interaction with Rad53, showing that this region is not important for Rad53 binding and that the structure of Dun1 FHA is not significantly altered with mutations to this region. Consistent with the literature, the Dun1 R60A mutation showed interaction levels at the same level as for an empty control, showing that this interaction is mediated primarily through the Dun1 pThr binding pocket (Lee, H. *et al.,* 2008).

My work has shown that Dun1 has a second conserved region that mediates a specific interaction with Sml1. In order to highlight the importance of this lateral region, spotting plate assays were conducted in order to detect a possible phenotype. It would be expected that the Dun1 FHA lateral surface mutants that exhibited a disrupted interaction with Sml1 would have increased sensitivity to genotoxic agents when compared to the WT Dun1. WT, penta mutation, R60A mutants and empty vector were transformed into a ΔDun1 strain (MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 Dun1Δ). Full length versions of Dun1 WT and mutants were used. The plasmid used was pCM190, which has a *URA3* (encodes a protein involved in uracil synthesis) selectable marker and whose recombinant protein expression levels can be controlled by tetracycline. As proteins levels can be controlled by the addition or removal of tetracycline from the experimental medium, recombinant Dun1 protein levels can be controlled to prevent overexpression. These yeast transformants were grown to saturation, serially diluted and plated onto media in the presence or absence of genotoxic agents (MMS and HU). It was expected that Dun1 mutant cells (R60A + Penta mutation) would be more sensitive to HU as compared to MMS. Although MMS provokes a checkpoint response by alkylating DNA, the majority of Dun1 ligands play a role in mediating dNTP synthesis. As HU restricts dNTP pools in cells it is expected that this reagent would trigger the most severe phenotype in Dun1 mutant yeast cells.

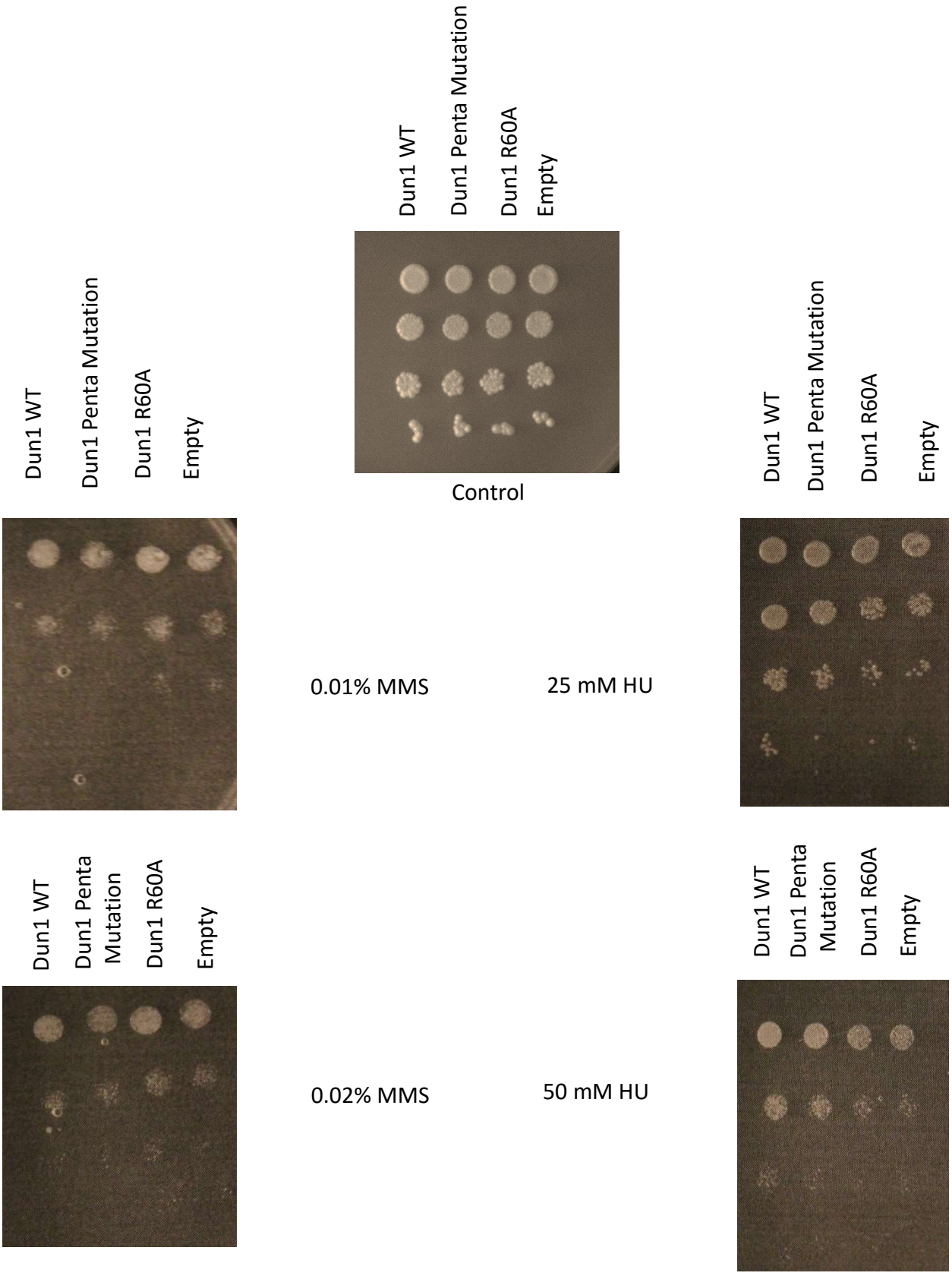

Figure 23: Dun1 FHA mutants have differential sensitivities to DNA damage agents. Full length versions of Dun1 WT, Dun1 Penta Mutation, Dun1 R60A mutation in addition to an empty vector control were exposed to MMS (Left), HU (Right) and no DNA damaging agent (Top).

Looking at the phenotypes in Figure 23, it appears that the conserved Dun1 lateral surface may be playing a role in DNA damage responses. As expected, there appears to be no significant difference between yeast transformants when exposed to MMS. This makes sense, as Dun1 has a main role in regulating dNTP levels as opposed to other aspects of the intra-S phase checkpoint response (Zhao & Rothstein 2002). Surprisingly, none of the yeast transformants were extremely sensitive to HU, however there were differential sensitivities. This was surprising as the lateral surface mutants were expected to have a higher sensitivity to HU than seen, due to two-hybrid experiments presented above. Additionally, it was expected that the empty vector control should have been very sensitive to HU at 50 mM as it has been previously shown (Chen *et al.,* 2007). Both the Dun1 R60A and empty pCM190 transformants had the highest sensitivity when exposed to HU, with the WT Dun1 transformant having the lowest sensitivity. Additionally, the penta Dun1 FHA1 mutant had an intermediate sensitivity when comparing the two other forms of Dun1. From this data, it can be interpreted that the lateral surface on Dun1 does play a role in mediating dNTP levels in response to HU, however there may be other redundant pathways also in play. This intuitively makes sense, if the lateral patch mutation is specific to the interaction with Sml1 and the R60A mutation is important for all Dun1 interactions, explaining the increased sensitivity with the latter. Future work can be done to test this hypothesis, to see whether the Dun1 lateral surface is important for any other interactions.

4.3 Discussion

It is clear that numerous FHA domains have second highly conserved regions that are distinct from their pThr binding site. These surfaces play a major role in conferring ligand specificity, as they can be modified in order to decrease or increase affinities. This was seen

when lateral surfaces residues of Rad53 FHA1 were mutated on the corresponding faces of Dun1 and Mek1, to see if ligand specificity could be altered with respect to an FHA1 ligand, Dbf4. As previously shown, a residue important for Dbf4 binding to FHA1 is E129. The corresponding residue is conserved on Dun1 FHA, whereas it is not conserved in Mek1 FHA. Additionally, all other lateral surface residues in addition to E129 were evaluated for conservation using both sequence and structural alignment methods. Although Dun1, Mek1, and Plm2 had relatively close conservation scores both from the sequence and structural alignments, these scores were compiled with all lateral surface residues having the same weighting. Recent experiments have shown that in interactions with Dbf4 this may not be the case, as some residues may contribute to binding more than others (Glu129, Asn112) (Matthews *et al.,* 2014). When the corresponding residues to FHA1 Glu129 had their properties altered (Dun1 became non-polar, Mek1 became polar) Dun1 retained its specificity for Dbf4, whereas Mek1 significantly increased its affinity. This increase in affinity proves that this residue is important for binding Dbf4, as it confers a new specificity when this alteration was made to Mek1FHA In Dun1, the lack of disruption of its binding to Dbf4 could be due to other residues in the pThr binding pocket that maintain the positive charge and allow it to bind Dbf4 at WT levels. It is important to note that Dbf4 is not thought to be an *in vivo* ligand for either Dun1 or Mek1. This work further supports the existing literature that Glu129 is a key residue on the primary FHA1 interaction surface with Dbf4.

After these initial experiments, other FHA domains were examined using bioinformatic methods. Interestingly, of the FHA domains examined all had a conserved region that was distinct but as conserved as the known pThr binding pocket. These conserved surfaces typically consist of a series of charged residues that occur in one of three orientations relative to the pThr binding site: on opposite, adjacent or the same face. Evolutionarily, this tells a very interesting

story, as the presence of a potential secondary binding site is common, however the relative position of this site is poorly conserved. With further characterization of the importance of these sites it may become clear why the location of these sites appears to be variable. One initial hypothesis is that the location of the lateral FHA domain binding site may be determined by the structure of the ligand it is trying to bind. These lateral binding sites may have evolved and changed position in order to reduce steric hindrance and increase the strength of their respective interactions. The existing data supports this hypothesis, as both Rad53 FHA1 and Dun1 have lateral surfaces directly opposite their pThr binding pockets, which makes sense due to their close evolutionary history. Interestingly, the ligands that they bind (Dbf4 and Sml1, respectively) both have core structural alpha-helices, as shown in crystal structure and structural prediction data (Matthews *et al.,* 2012).

When looking at the role of the Dun1 FHA lateral surface in binding Sml1, it was seen that a mutation of five residues in this region caused a reduction in interaction similar to empty vector levels in two-hybrid analysis. Interestingly, this reduction in interaction was at similar levels to a Dun1 FHA R60A mutation, a mutant that is known to disrupt the pThr binding site. This result was similar to the Rad53 FHA1-Dbf4 interaction, as both the lateral surface and pThr binding pocket appear to play a role in conferring ligand specificity. Due to the fairly large number of amino acids mutated, it is hypothesized that the interaction may be similar to FHA1-Dbf4, and multiple contact points may exist in the Dun1 FHA-Sml1 interaction. This would explain why smaller sets of mutations appear to have minimal effect on the interaction. The Dun1 FHA- Rad53 interaction has been previously been characterized, so it was a good control to test the specificity of our Dun1 FHA lateral surface mutants (Lee, H. *et al.,* 2008). In our two-hybrid analysis it was shown that the Dun1 FHA lateral surface mutants displayed WT levels of

interaction with Rad53, and that the R60A mutant showed empty vector interaction levels. This shows that the interaction knockdown seen with Dun1 FHA lateral surface mutants and Sml1 is ligand specific, and that alterations in this region do not disrupt Dun1 FHA structure in a way that compromises all ligand binding. When these same mutants were expressed using a titratable expression vector (pCM190) in a strain lacking endogenous Dun1 and exposed to genotoxic agents, it was seen that cells were sensitive to HU but not as sensitive to MMS. This is an expected result as the Dun1-Sml1 interaction is primarily involved in dNTP synthesis and not as involved with other facets of the DNA checkpoint response. Although more sensitive than WT, cells with the lateral surface Dun1 penta mutant were not as sensitive as expected. This may be due to this mutant being able to maintain enough of a residual interaction with Sml1 to mediate the levels of dNTPs in the cell. Looking at the initial two-hybrid result, this hypothesis has some credence, as the levels of interaction of the lateral surface mutant were slightly elevated when compared to the R60A mutant. Additionally, cells overall in our spotting assays were not as sensitive as expected. Future experiments should refine the amount of tetracycline used, as an overexpression of proteins may have caused a reduced sensitivity. Another possible explanation for the lack of sensitivity could be a decreased potency of the HU used in my experiments. In order to test for this, newly ordered HU should be used in future experiments, in addition to transformants that have previously been tested in the laboratory and showed a high level of sensitivity.

This work clearly shows that there are more interaction surfaces of FHA domains than previously thought. These lateral surfaces are distinct regions from the pThr binding pocket and can be found on multiple FHA domains in *S. cerevisiae*. This work provides a framework to

investigate novel secondary protein binding patches in this critical protein-protein interaction domain.

Chapter 5: General Discussion

5.1 Rad53 FHA1 has a lateral surface patch that is variably required for ligand binding

In my experiments, I have extensively characterized the role of the lateral surface of FHA1 in interacting with Dbf4, and the variable requirement of this region in binding other FHA1 ligands using a structural and biochemical approach. Previously it was shown that the lateral surface of FHA1 is responsible for mediating the interaction with the Dbf4 HBRCT domain (Matthews *et al.,* 2014). Although this lateral region was shown to be important, little was known about the interaction interface. Here, I report a more fully characterized interaction where the lateral surface of FHA1 interacts with the HBRCT domain via two contact points. This hypothesis stemmed from the recent co-crystallization data of Rad53 FHA1-Dbf4 HBRCT obtained from the Guarné lab in addition to several biochemical yeast two-hybrid analyses. Disruption of a single contact point is not enough to significantly decrease protein binding affinity, whereas disruption of both contact points (whether on FHA1 or Dbf4) is enough to decrease interaction levels to empty background levels.

Due the novelty of this lateral surface on Rad53 FHA1, it was important to test the specificity of this region in mediating interaction with Dbf4. To do this, several ligands were evaluated using different protein-protein interaction assays. Cdc7, the catalytic subunit of DDK, was an initial ligand tested via two-hybrid analysis. In this assay the same lateral surface FHA1 mutants that decreased the interaction to empty levels with Dbf4, failed to affect the interaction with Cdc7. This initial result was important as proved two things: that the lateral surface mutants are structurally stable as they retained WT Cdc7 binding activity in addition to the fact that this region is specific to Dbf4 as it is not important for binding Cdc7. This interaction was further shown in my research to be regulated through the pThr binding pocket when a previously identified threonine was mutated to alanine and a 60% reduction of interaction was seen in a

two-hybrid experiment. As Cdc7 has been shown to bind the pThr binding pocket of FHA1 whereas Dbf4 binds the lateral surface, we propose a bipartite interaction model where Cdc7 binds first and slightly changes the conformation of FHA1 to allow Dbf4 binding. This model is supported by recent structural data, as the conformation of FHA1 has been shown to slightly shift depending on the ligand bound (Alba Guarné, personal communication). The second ligand investigated was Sgs1, a helicase involved both in the DNA unwinding and damage response. It was shown that the same mutants for the lateral surface of FHA1 that could abrogate interaction with Dbf4 did not have an effect on interaction with Sgs1. Consistent with the literature, it appeared that this interaction is mediated through the pThr binding pocket, as mutation of this region disrupted interaction (Hegnauer *et al.*, 2012). The last ligand investigated with lateral FHA1 domain mutants was Sld3. Via co-immunoprecipitation, it was shown that the lateral surface of FHA1 is only partially required for binding Sld3, as opposed to the absolute requirement for the pThr binding pocket. This intermediate requirement of the FHA1 lateral surface in binding Sld3 shares similarities with the FHA1-Dbf4 interaction, as it appears that both the pThr binding pocket and lateral surface play a role. The variable requirement of this region of FHA1 is interesting, as the FHA1 lateral surface has now been shown to influence both the DDK (via Dbf4) and CDK (via Sld3) branches of the DNA damage response. This suggests a pair of evolutionarily redundant pathways to halt DNA replication in times when checkpoints are triggered.

An interesting aspect of this research is applying lessons learned from Rad53 FHA1 interactions to other FHA domains. This is of particular interest as second conserved patches appear to be more common than previously thought. The presence of these additional conserved regions may indicate that FHA domains may be able to bind ligands in a bipartite manner. As

these lateral surfaces are conserved, it is expected that FHA domains with similar lateral surfaces would bind ligands that are structurally similar. When looking at phosphorylation independent interactions, there appears to be great variability in the active site used. In Rv1827, a FHA domain from *Mycobacterium tuberculosis*, the pThr binding pocket is used to mediate both phosphorylation dependent and phosphorylation independent interaction (Nott *et al.,* 2009). In both Rad53 and its human counterpart Chk2, it has been shown that interaction with binding partners is dependent on both the pThr binding pocket in addition to another lateral surface (Li *et al.,* 2002. Matthews *et al.,* 2014). Looking at the above examples it is clear that FHA domains can mediate phosphorylation independent interactions, and this work helps characterize some of the novel surface types that may be involved.

5.2 Lateral Surfaces in *S. cerevisiae* FHA domains

In my work I have highlighted the importance of lateral surfaces in conferring ligand specificity in multiple FHA domains. In my first experiment, I effectively showed the importance of the lateral surface of FHA1 in mediating a direct interaction with Dbf4, the regulatory subunit of DDK. I did this by mutating the corresponding lateral surface regions on Dun1 FHA and Mek1 FHA in order to disrupt or enhance Dbf4 binding. These two protein domains were chosen due to their structural similarities with Rad53 FHA1. It was seen when the key conserved Dun1 FHA residue was mutated to a non-polar amino acid, it had little reduction on Dbf4 affinity, whereas mutating the non-conserved Mek1 FHA lateral surface to a polar amino acid led to a six-fold increase in Dbf4 affinity when compared to its wild type counterparts in a two-hybrid experiment. Although FHA domains are divergent on the sequence level, they are structurally conserved, explaining how an FHA domain can be engineered to have an increased affinity for a non-native ligand.

Due to the extreme physiological importance of the Rad53 FHA1-Dbf4 interaction, other FHA domains in *S. cerevisiae* were examined and many had a second conserved patch that was distinct from the pThr binding pocket, yet conserved roughly to the same extent. The conserved surfaces appear in a number of different orientations suggesting that structure of the FHA ligand in question may determine the orientation of the lateral surface that it may bind to. An interesting example of this can be seen in Pml1, a protein that contains an FHA domain and is involved in the pre-mRNA splicing complex. In this protein, the pThr binding site is not conserved, and its interaction with its ligand Snu17 is mediated through a region of the protein completely distinct from the pThr binding site (Wysoczanski *et al.,* 2015). Furthermore, this region is the minimal sufficient domain required to bind Snu17, highlighting the importance in secondary surfaces on FHA domains in mediating a diverse set of cell-cycle interactions. Conserved lateral surfaces on FHA domains have been identified on Xrs2 (involved in double strand break repair), Dun1 (involved in dNTP regulation) and Fkh2 (involved in expression of genes that regulate the cell cycle) showing that these FHA lateral surfaces are not confined to one biological pathway. Future experiments can focus on whether these conserved lateral surfaces play similar roles in DSB repair and gene expression compared to what they do in mediating checkpoint responses and dNTP levels (Rad53 FHA1 and Dun1 FHA).

5.3 Relation to Cancer

As many of the proteins and pathways studied in this research contribute to genome integrity and maintenance, it is a logical assumption that a disruption of these interactions can lead to cancer phenotypes. Although this work is performed in a budding yeast system that is not cancerous, the degree of conservation of DNA replication and repair processes allows findings to be generalized to higher eukaryotes. One of the most studied protein complexes is DDK,

comprised of Dbf4 and Cdc7. My work has helped to fully characterize the interaction interface between Rad53 FHA1 and DDK, with this interaction being important for maintenance of the genome. When looking at the literature it is clear that increased expression levels of DDK is present in both tumor cells lines and cancerous tissues (Bonte *et al.*, 2008, Kulkarni *et al.*, 2009, Cheng *et al.,* 2013). A potential therapy would be using the FHA1 lateral surface region to titrate Dbf4 when it is overexpressed in cancerous cell lines. Based on my data, it would be expected that this titration would primarily affect Dbf4, as other FHA1 ligands appear to have a minimal dependence on this lateral surface region. The additional insight I have contributed to the mechanism of Rad53 interaction with Dbf4 is an important step in gaining a better understanding of how the former keeps DDK activity in check following DNA damage.

FHA domains have been implicated in several cancers when mutated. When Chk2, the human equivalent of Rad53 is mutated (I157T, R145W) it causes a variant of Li-Fraumani syndrome in addition to colon cancer (Bell *et al.,* 1999). Additionally, when this I157T mutation is combined with a mutation in the Chk2 kinase domain an increased susceptibility to breast cancer is seen. These naturally occurring mutations predispose individuals to cancer as Chk2 has a tumor suppressor role. In addition to Chk2, FHA domains in several other human proteins have been implicated in having causal roles in tumor formation when disrupted, including that of CHFR (Scolnick & Halazonetis, 2000) and Nbs1 (Tauchi *et al.,* 2001). Interestingly, CHFR is a promising candidate for use as a biomarker as levels are significantly reduced in lung, colon and gastic cancers (Koga *et al.,* 2011; Tanaka *et al.,* 2011; Satoh *et al.,* 2003). It is also efficient to screen for CHFR levels in humans, as existing protocols currently exist, making it an effective potential biomarker for the future (Herman *et al.,* 1996).

Although cancer biology has not been the focus of this project, the major proteins and

pathways are conserved in eukaryotes. The more that is understood about the mechanisms of FHA domain interactions will allow us to learn more about the cell cycle, and in turn, cell cycle dysregulation and cancer.

5.4 Future Directions

This work has opened up several new avenues for potential research, as little is known about secondary interaction surfaces on FHA domains. Although the conservation of FHA1 lateral surface residues was evaluated in this project, this was done at the bioinformatics level. In order to determine whether the interaction interface seen in this study between Rad53-Dbf4 is conserved in higher eukaryotes, further biochemical and structural experiments must be performed. I identified that a key amino acid (E129) in the Rad53-Dbf4 interaction is conserved in both Cds1 (*S. pombe*) and Chk2 (*H. sapiens*), leading me to believe that this amino acid may also be mediating interactions on FHA domains in both yeast and human species. By doing these experiments in a variety of organisms, it will shed light on how these interaction surfaces evolved to be structurally distinct from the pThr binding site.

In this study I showed that Dun1 FHA maintains an interaction with Sml1 via a lateral surface region. Mutations to the lateral surface of Dun1 FHA appear to reduce the affinity of interaction between Dun1 FHA-Sml1 to empty vector levels. Due to the close evolutionary history between Rad53 and Dun1, I propose that a similar binding interface is occurring with Dun1-Sml1 compared to Rad53 FHA1-Dbf4. Future structural experiments can be conducted to fully map the Dun1 FHA-Sml1 interface, to determine what residues on both proteins form the primary surface of interaction.

In order to fully characterize the role of the lateral surface of Dun1 FHA in dNTP regulation, other ligands must be examined. Two proteins of particular interest are Dif1 and Crt1. Dif1 is a protein that controls ribonuclease reductase import into the nucleus during normal and checkpoint conditions, while Crt1 is responsible for the transcription of ribonuclease reductases when it is phosphorylated by Dun1. By better understanding the role of the lateral surface of Dun1 in binding its partners, lessons will be learned about how dNTP levels are regulated in cells in addition to how secondary interaction surfaces on FHA domains play roles in mediating interactions.

In addition to Rad53 FHA1 and Dun1 FHA, several other FHA domains were identified in this study as having conserved lateral surfaces that may be mediating interactions. By examining these regions using similar approaches as outlined in this study, it will be possible to expand the literature surrounding non-canonical FHA domain interaction surfaces, extending our knowledge surrounding cell cycle and checkpoint regulation.

5.5 Significance

This work has helped better characterize FHA domain structure and function by identifying multiple protein-protein interactions that occur via non-canonical surfaces. This work is significant as it joins a growing body of literature documenting interaction surfaces distinct from pThr binding sites of FHA domains that play a role in a number of diverse physiological roles. These surfaces have been implicated both the DDK and CDK arms of the DNA damage response (via Dbf4 and Sld3 respectively). This study has also shown that a non-canonical interaction surface on Dun1 FHA may also be playing a role in regulating dNTP levels in the cell during exposure to DNA damage. The presence of a second conserved patch in multiple other

FHA domains may indicate an interaction ability with ligands in a bipartite manner and with phosphorylation independent interactions could be more common in FHA domains than previously thought. This has been shown to be true in the case of Rad53 FHA1 and DDK, showing for the first time an FHA domain simultaneously binding two ligands via distinct interaction surfaces. This work has shown a direct relationship between the interaction surfaces, as the binding of Cdc7 changes the conformation of FHA1 in order to better bind Dbf4. Using Rad53 FHA1 as a model, other FHA domains (such as Dun1 FHA) may be mediating bipartite interactions in a similar manner. By better understanding the basic mechanisms of how FHA domains contribute to cell cycle checkpoints, lessons can be learned surrounding how these interactions are disrupted in formation and progression of cancers.

References

Alderwick, L., Molle, V., Kremer, L., Cozzone, A., Dafforn, T., Besra, G., & Futterer, K. (2006). Molecular structure of EmbR, a response element of Ser/Thr kinase signaling in Mycobacterium tuberculosis. *Proceedings of the National Academy of Sciences of the United States of America, 103*, 2558-2563.

Altmann, K., Dürr, M., & Westermann, B. (2007). *Saccharomyces cerevisiae* as a model organism to study mitochondrial biology: general considerations and basic procedures. *Methods in Molecular Biology*, *372*, 81-90.

Aucher, W., Becker, E., Ma, E., Miron, S., Martel, A., Ochsenbein, F., Marsolier-Kergoat., M.C., & Guerois, R. (2010). A Strategy for Interaction Site Prediction between Phospho-binding Modules and their Partners Identified from Proteomic Data. *Molecular & Cellular Proteomics, 9*, 2745-2759.

Ausubel, F., Brent, R., Kingston, R., Moore, D., Seidman, J.,Smith, J., & Struhl, K. (1994). Current Protocols in Molecular Biology. John Wiley & Sons, New York.

Bell, D., Varley, J., Szydlo, T., Kang, D., Wahrer, D.,Shannon K., Lubratovich, M., Verselis S., Isselbacher K., Fraumeni, J., & Birch, J. M (1999). Heterozygous germ line hCHK2mutations in Li-Fraumeni syndrome. *Science 286*, 2528-253.

Bernstein, N., Williams, R., Rakovszky, M., Cui, D., Green, R., Karimi-Busheri, F & Glover, J. (2005). The Molecular Architecture of the Mammalian DNA Repair Enzyme, Polynucleotide Kinase. *Molecular Cell, 17*, 657-670.

Bjergbaek, L, Cobb, J., Tsai-Pflugfelder, M., & Gasser, S. (2005). Mechanistically distinct roles for Sgs1p in checkpoint activation and replication fork maintenance. *EMBO Journal***,** *24,* 405-417.

Blattner, F., Plunkett, G., Bloch, C., Perna, N., Burland, V., Riley, M., Collado-Vides, J., Glasner, J., Rode, C., Mayhew, G., Gregor, J., Davis, N., Kirkpatrick, H., Goeden, M., Rose, D., Mau, B., & Shao, Y. (1997). The Complete Genome Sequence of Escherichia coli K-12. *Science, 277*, 1453-1462.

Bonte, D., Lindvall, C., Liu, H., Dykema, K., Furge, K., & Weinrich, M. (2008) Cdc7-Dbf4 kinase overexpression in multiple cancers and tumour cell lines in correlated with p53 inactivation. *Neoplasia, 10*, 920-31.

Boos, D., Frigolaa, J., & Diffley, J. (2012). Activation of the replicative DNA helicase: breaking up is hard to do. *Current Opinion in Cell Biology*, *24*, 423-430.

Byeon, I., Li, H., Song, H., Gronenborn, A., & Tsai, M. (2005). Sequential phosphorylation and multisite interactions characterize specific target recognition by the FHA domain of Ki67. *Nature Structural & Molecular Biology, 12*, 987-993.

Carr, A.M. (2002). DNA structure dependent checkpoints as regulators of DNA repair. *DNA Repair*, *1*, pp.983–994.

Chen, S., Smolka, M., and H. Zhou. (2007). Mechanism of Dun1 activation by Rad53 phosphorylation in *Saccharomyces cerevisiae*. *Journal of Biological Chemistry***,** *282*, 986-95.

Cheng, A. N., Jiang, S. S., Fan, C.-C., Lo, Y.-K., Kuo, C.-Y., & Chen, C.-H. (2013). Increased Cdc7 expression is a marker of oral squamous cell carcinoma and overexpression of Cdc7 contributes to the resistance to DNA-damaging agents. *Cancer Letters, 337*, 218–25.

Cimprich, K. & Cortez, D. (2008). ATR: an essential regulator of genome integrity. *Nature Reviews Molecular Cell Biology*, *9*, 616–27.

Dosztányi, Z., Csizmók, V., Tompa, P., & Simon, I. (2005). IUPred: web server for the prediction of intrinsically unstructured regions of proteins based on estimated energy content. *Bioinformatics 21*, 3433-3434.

Duncker, B. P., and Brown, G. (2003). Cdc7 kinases (DDKs) and checkpoint responses: lessons from two yeasts. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis*, *532*, 21–27.

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. *Proceedings of the National Academy of Sciences of the United States of America, 99*, 16087–16092.

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

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

Edgar, Robert C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput, *Nucleic Acids Research, 32,* 1792-1797.

EMBI. "Multiple Sequence Comparison by Log- Expectation", European Molecular Biology Laboratory. http://www.ebi.ac.uk/Tools/msa/muscle. (September 1st 2015)

Fields, S., & Johnston, M. (2005). Whither model organism research? *Science, 307,* 1885–1886.

Garí, E., Piedrafita, L., Aldea, M., & Herrero, E. (1997). A Set of Vectors with a Tetracycline-Regulatable Promoter System for Modulated Gene Expression in *Saccharomyces cerevisiae*. *Yeast, 13*, 837-848.

Goffeau, A., Barrell, B., Bussey, H., Davis, R., Dujon, B., Feldmann, H., Galibert, F., Hoheisel, J., Jacq, C., Johnston, M., Louis, E., Mewes, H., Murakami, Y., Philippsen, P., Tettelin, H., Oliver, S. (1996). Life with 6000 Genes. *Science, 274*, 546-567.

Grünenfelder, B., & Winzeler, E. (2002). Treasures and traps in genome-wide data sets: case examples from yeast. *Nature Reviews Genetics*, *3*, 653-661.

Guindon, S., Dufayard, J., Lefort, V., Anisimova, M., Hordijk, W., & Gascuel, O. (2010). New Algorithms and Methods to Estimate Maximum-Likelihood Phylogenies: Assessing the Performance of PhyML 3.0. *Systematic Biology, 59*, 307-321.

Gyuris, J., Golemis, E., Chertkov, H., & Brent, R. (1993). Cdi1, a human G1 and S phase protein phosphatase that associates with Cdk2. *Cell, 75*, 791-803.

Hegnauer, A., Thomä, N., Hustedt, N., Gasser, S., Van Attikum, H., Guénolé, A., Van Leeuwen, F., Rubin, S., Amsler, P., Vogel, M., Pike, B., & Shimada, S. (2012) An N-terminal acidic region of Sgs1 interacts with Rpa70 and recruits Rad53 kinase to stalled forks. *The EMBO Journal, 31,* 3768-3783.

Heller, R. C., Kang, S., Lam, W. M., Chen, S., Chan, C. S., & Bell, S. P. (2011). Eukaryotic Origin-Dependent DNA Replication in vitro Reveals Sequential Action of DDK and S-CDK Kinases. *Cell, 146*, 80–91.

Herman, J. G., Graff, J. R., Myohanen, S., Nelkin, B. D., & Baylin, S.B. (1996). Methylation-specific PCR: a novel PCR assay for methylation status of CpG islands. *Proceedings of the National Academy of Sciences of the United States of America, 93*, 9821–9826.

Herskowitz, I. (1988). Life cycle of the budding yeast *Saccharomyces cerevisiae*. *Microbiology Reviews, 52*, 536-553.

Hickson, I. (2003). RecQ helicases: caretakers of the genome. *Nature Reviews Cancer*, *3*, 169-178.

Hofmann, K. & Bucher, P. (1995). The FHA domain: A putative nuclear signaling domain found in protein kinases and transcription factors. *Trends in Biochemical Science*, 20, 347-349.

Huang, M., & Elledge, S. (2000). The FHA domain, a phosphoamino acid binding domain involved in the DNA damage response pathway. *Cold Spring Harbor Symposia on Quantitative Biology, 65*, 413-21

Hughes, T. (2002). Yeast and drug discovery. *Functional & Integrative Genomics, 2,* 199–211.

Itou, H., Muramatsu, S., Shirakihara, Y., & Araki, H. (2014). Crystal Structure of the Homology Domain of the Eukaryotic DNA Replication Proteins Sld3/Treslin. *Structure, 22*, 1341-1347.

Jackson A. L., Pahl P. M., Harrison K., Rosamond J., & Sclafani R. A. (1993). Cell cycle regulation of the yeast Cdc7 protein kinase by association with the Dbf4 protein. *Molecular and Cellular Biology*, *13*, 2899–2908.

Jossen, R., & Bermejo, R. (2013). The DNA damage checkpoint response to replication stress: A Game of Forks. *Frontiers in Genetics, 4*, 1-13.

Kanke, M., Kodama, Y., Takahashi, T. S., Nakagawa, T., & Masukata, H. (2012). Mcm10 plays an essential role in origin DNA unwinding after loading of the CMG components. *The EMBO Journal*, *31*, 2182–94.

Kitada K., Johnston L. H., Sugino T., & Sugino A. (1992). Temperature-sensitive *cdc7* mutations of *Saccharomyces cerevisiae* are suppressed by the *DBF4* gene, which is required for the G_1/S cell cycle transition. *Genetics 131*, 21–29.

Koga, T., Takeshita, M., Yano, T., Maehara, Y., & Sueishi, K. (2011). CHFR hypermethylation and EGFR mutation are mutually exclusive and exhibit contrastive clinical backgrounds and outcomes in non-small cell lung cancer. *International Journal of Cancer, 128*, 1009–1017.

Kulkarni, A., Kingsbury, S. R., Tudzarova, S., Hong, H.-K., Loddo, M., & Rashid, M. (2009). Cdc7 kinase is a predictor of survival and a novel therapeutic target in epithelial ovarian carcinoma. *Clinical Cancer Research, 15*, 2417–25.

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

Labib, K. & De Piccoli, G. (2011). Surviving chromosome replication: the many roles of the S-phase checkpoint pathway. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, *366*, 3554–61.

Lee, Y. D., and Elledge, S. J. (2006). Control of ribonucleotide reductase localization through an anchoring mechanism involving Wtm1. *Genes & Development*, *20,* 334–344.

Lee, Y. D., Wang, J., Stubbe, J., & Elledge, S. J. (2008). Dif1 is a DNA-damage-regulated facilitator of nuclear import for ribonucleotide reductase. *Molecular Cell, 32,* 70–80.

Lee, H., Yuan, C., Hammet, A., Mahajan, A., Chen, E., Wu, M., Su M., Heierhorst J., & Tsai, M. (2008). Diphosphothreonine-Specific Interaction between an SQ/TQ Cluster and an FHA Domain in the Rad53-Dun1 Kinase Cascade. *Molecular Cell, 30*, 767-778.

Li, Y., and Araki, H. (2013). Loading and activation of DNA replicative helicases: the key step of initiation of DNA replication. *Genes to cells: devoted to molecular & cellular mechanisms*, *18*, 266–77.

Li, J., Williams, B., Haire, F., Goldberg M, Wilker E, Durocher D, Yaffe MB, Jackson SP & Smerdon SJ. (2002). Structural and functional versatility of the FHA domain in DNA-damage signaling by the tumor suppressor kinase Chk2. *Molecular Cell*, *9*, 1045–54.

Luo, S., Xin, X., Du, L., Ye, K., & Wei, Y. (2015). Dimerization Mediated by a Divergent Forkhead-associated Domain Is Essential for the DNA Damage and Spindle Functions of Fission Yeast Mdb1. *Journal of Biological Chemistry, 290*, 21054-21066.

Machida, S., & Yuan, Y. (2013). Crystal structure of Arabidopsis DDL FHA domain. *Molecular Plant, 6*, 1290-1300.

Mahajan, C. Yuan, B. L. Pike, J. Heierhorst, C. F. Chang, & M. D. Tsai. (2005). FHA domain-ligand interactions: Importance of integrating chemical and biological approaches. *Journal of the American Chemical Society. 127,* 14572–14573.

Manchado, E., Eguren, M., & Malumbres, M. (2010). The anaphase-promoting complex/cyclosome (APC/C): cell-cycle-dependent and -independent functions. *Biochemical Society Transactions*, *38*, 65–71.

Mantiero, D., Mackenzie, A., Donaldson, A., & Zegerman, P. (2011). Limiting replication initiation factors execute the temporal programme of origin firing in budding yeast. *The EMBO Journal*, *30*, 4805–14.

Matthews, L., Jones, D., Prasad, A., Duncker, B., & Guarne, A. (2012). *Saccharomyces cerevisiae* Dbf4 Has Unique Fold Necessary for Interaction with Rad53 Kinase. *Journal of Biological Chemistry, 287*, 2378-2387.

Matthews, L., Selvaratnam, R., Jones, D., Akimoto, M., McConkey, B., Melacini, G., Duncker, B.P., & Guarné A. (2014). A Novel Non-canonical Forkhead-associated (FHA) Domain-binding Interface Mediates the Interaction between Rad53 and Dbf4 Proteins. *The Journal of Biological Chemistry*, *289*, 2589-2599.

Mendenhall MD & Hodge AE. (1998). Regulation of Cdc28 cyclin-dependent protein kinase activity during the cell cycle of the yeast *Saccharomyces cerevisiae*. *Microbiology and Molecular Biology Reviews. 62,* 1191–1243.

Mokdad-Gargouri, R., Abelmoula-Soussi, S., Hadiji-Abbes, N., Amor, I. Y., Borchani-Chabchoub, I., & Gargouri, A. (2012). Yeasts as a Tool for Heterologous Gene Expression. *Methods in Molecular Biology*, *824*, 359-370.

Montagnoli, A., Valsasina, B., Croci, V., Menichincheri, M., Rainoldi, S., & Marchesi, V. (2008). A Cdc7 kinase inhibitor restricts initiation of DNA replication and has antitumor activity. *Nature Chemical Biology*, *4*, 357–65.

Mojca, M., Petrovic, U., and Krizaj, I. (2012). Yeast as a model eukaryote in toxicology: A functional genomics approach to studying the molecular basis of action of pharmacologically active molecules. *Toxicon*, *60,* 558-571.

Morgan, D.O. (2007). The Cell Cycle: Principles of Control. New Science Press Ltd. U.S.A. pp. 12-84, 158-172, 196-205.

Natoni, A., Coyne, M. R. E., Jacobsen, A., Rainey, M. D., O'Brien, G., & Healy, S. (2013). Characterization of a dual CDC7/CDK9 inhibitor in multiple myeloma cellular models. *Cancers, 5*, 901–18.

Nautiyal, S., Derisi, J., & Blackburn, E. (2002). The genome-wide expression response to telomerase deletion in *Saccharomyces cerevisiae*. *Proceedings of the National Academy of Sciences, 99*, 9316-9321.

NCBI. "National Center for Biotechnology" National Center for Biotechnology- Proteins. http://www.ncbi.nlm.nih.gov/guide/proteins (September 1st, 2015)

Nguyen, V., Co, C., & Li, J. (2001). Cyclin-dependent kinases prevent DNA re-replication through multiple mechanisms. *Nature*, *411*, 1068–1073.

Nieduszynski, C. A., Knox, Y. & Donaldson, A. D. (2006). Genome-wide identification of replication origins in yeast by comparative genomics. *Genes Development, 20*, 1874-1879.

Nott, T., Kelly, G., Stach, L., Li, J., Westcott, S., Patel, D., Hunt, D., Howell, S., Buxton, RS. O'Hare H & Smerdon, S. (2009). An Intramolecular Switch Regulates Phosphoindependent FHA Domain Interactions in Mycobacterium tuberculosis. *Science Signaling, 2,* ra12.

Pennell, S., Westcott, S., Ortiz-Lombardía, M., Patel, D., Li, J., Nott, T & Smerdon, S. (2010). Structural and Functional Analysis of Phosphothreonine-Dependent FHA Domain Interactions. *Structure, 18*, 1587-1595.

Pereira, C., Coutinho, I., Soares, J., Bessa, C., Leao, M., & Saraiva, L. (2012). New insights into cancer-related proteins provided by the yeast model. *The FEBS Journal, 279,* 697-712.

Porro, D., Sauer, M., Branduadi, P., & Mattanovich, D. M. (2005). Recombinant Protein Production in Yeasts. *Molecular Biotechnology*, *31*, 245-259.

Pringle, J., & Hartwell, L. (1981). The *Saccharomyces cerevisiae* cell cycle. In *The Molecular biology of the yeast Saccharomyces, life cycle and inheritance*. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory.

Randell J.C., Bowers, J.L., Rodriguez, H.K., & Bell, S.P. (2006) Sequential ATP hydrolysis by Cdc6 and ORC directs loading of Mcm2-7 helicase. *Molecular Cell, 21*, 29-39.

Rea, S., Graham, B., Nakamaru-Ogiso, E., Kar, A., & Falk, M. (2010). Bacteria, yeast, worms, and flies: exploiting simple model organisms to investigate human mitochondrial diseases. *Developmental Disabilities Research Reviews*, *16*, 200-218.

Satoh, A., Toyota, M., Itoh, F., Sasaki, Y., Suzuki, H., & Ogi, K. (2003). Epigenetic inactivation of CHFR and sensitivity to microtubule inhibitors in gastric cancer. *Cancer Research, 63*, 8606–8613.

Schwartz, M.F., Duong, J.K., Sun, Z., Morrow, J.S., Pradhan, D., & Stern, D.F. (2002). Rad9 phosphorylation sites couple Rad53 to the *Saccharomyces cerevisiae* DNA damage checkpoint. *Molecular Cell 9*, 1055-1065.

Scolnick, D., & Halazonetis, T. (2000). Chfr defines a mitotic stress checkpoint that delays entry into metaphase. *Nature, 406*, 430-435.

SGD. "*Saccharomyces* Genome database", *Saccharomyces cerevisiae* Genome Snapshot/Overview. http://www.yeastgenome.org. (September 1st, 2015)

Smolka, M., Chen, S., Maddox, P., Enserink, J., Albuquerque, C., Wei, X., Desai, A., Kolodner, R.D., & Zhou, H. (2006). An FHA domain-mediated protein interaction network of Rad53 reveals its role in polarized cell growth. *The Journal of Cell Biology, 175*, 743-753.

Straight, A. F. (1997). Mitosis in living budding yeast: Anaphase A but no metaphase plate. *Science*, *277*, 574–578.

Sun, Z., Hsaio, J.,Fay, D., & Stern, D. (1998). Rad53 FHA domain associated with phosphorylated Rad9 in the DNA damage checkpoint. *Science, 281*, 272–274.

Surana, U., Amon, A, Dowzer, C., McGrew, J., Byers, B., and Nasmyth, K. (1993). Destruction of the CDC28/CLB mitotic kinase is not required for the metaphase to anaphase transition in budding yeast. *The EMBO Journal*, *12*, 1969–78.

Tanaka, M., Chang, P., Li, Y., Li, D., Overman, M., Maru, D. M., Sethi, S., Phillips, J., Bland, G., Abbruzzese, J., & Eng, C. (2011). Association of CHFR Promoter Methylation with Disease Recurrence in Locally Advanced Colon Cancer. *Clinical Cancer Research, 17*, 4531-4540.

Tauchi, H., Kobayashi, J., Morishima, K., Matsuura, S., Nakamura, A., Shiraishi, T & Komatsu, K. (2001). The Forkhead-associated Domain of NBS1 Is Essential for Nuclear Foci Formation after Irradiation but Not Essential for hRAD50 hMRE11 NBS1 Complex DNA Repair Activity. *Journal of Biological Chemistry, 276*, 12-15.

Tsaponina, O., & Chabes, A. (2013). Pre-activation of the genome integrity checkpoint increases DNA damage tolerance. *Nucleic Acids Research*, *41*, 10371-10378.

Turk, M., Plemenitaš, A., & Gunde-Cimerman, N. (2011). Extremophilic yeasts: plasma-membrane fluidity as determinant of stress tolerance. *Fungal Biology*, *115*, 950–8.

Varrin, A.E., Prasad, A., Sclolz, R., Ramer, M., & Duncker BP. (2005). A mutation in Dbf4 motif M impairs interactions with DNA replication factors and confers increased resistance to genotoxic agents. *Molecular and cellular biology*, *25,* 7494–504.

Walker, G.M. (1998). Yeast physiology and biotechnology. Wiley Press, New York, NY. pp. 100-135.

Wan, L., los Santos, T. d., Zhang, C., Shokat, K., & Hollingsworth, N. (2003). Mek1 Kinase Activity Functions Downstream of RED1 in the Regulation of Meiotic Double Strand Break Repair in Budding Yeast. *Molecular Biology of the Cell*, *15*, 11-23.

Weng, J., Hsieh, Y., Huang, C., Wei, T., Lim, L., Chen, Y. & Tsai, M. (2015). Uncovering the Mechanism of Forkhead-Associated Domain-Mediated TIFA Oligomerization That Plays a Central Role in Immune Responses. *Biochemistry, 54*, 6219-6229.

Wood V, Gwilliam R, Rajandream MA, Lyne M, Lyne R, Stewart A, Sgouros J, Peat N, Hayles J, & Baker S. (2002). The genome sequence of *Schizosaccharomyces pombe*. *Nature*, *21*, 871-80.

Wysoczanski, P., Becker, S., & Zweckstetter, M. (2015). Structures of intermediates during RES complex assembly. *Scientific Reports, 5*.

Yabuuchi, H., Yamada, Y., Uchida, T., Sunathvanichkul, T., Nakagawa, T., & Masukata, H. (2006). Ordered assembly of Sld3, GINS and Cdc45 is distinctly regulated by DDK and CDK for activation of replication origins. *The EMBO journal*, *25*, 4663-74.

Yang, J, Yan, R, Roy, A Xu, D, Poisson, J & Y Zhang. (2015) The I-TASSER Suite: Protein structure and function prediction. *Nature Methods, 12*, 7-8.

Yeeles, J., Deegan, T., Janska, A., Early, A., & Diffley, J. (2015). Regulated eukaryotic DNA replication origin firing with purified proteins. *Nature, 519*, 431-435.

Yongkiettrakul S, Byeon IJL, Tsai MD. (2004). The ligand specificity of yeast Rad53 FHA domains at the +3 position is determined by nonconserved residues. *Biochemistry 43*, 3862–3869.

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

Zhao, X., Muller, E. G., & Rothstein, R. (1998). A suppressor of two essential checkpoint genes identifies a novel protein that negatively affects dNTP pools. *Molecular Cell.* 2, 329–340.

Zhao, X., & Rothstein, R. (2002). The Dun1 checkpoint kinase phosphorylates and regulates the ribonucleotide reductase inhibitor Sml1. *Proceedings of the National Academy of Sciences of the United States of America, 99*, 3746-3751.

Zhou, B., & Elledge, S. (2000). The DNA damage response: Putting checkpoints in perspective. *Nature, 408*, 433-439.

Appendix A: Chapter 3 Supplementary Material

Tree Node	Bootstrapping Value
Nbs1/Xrs2	89
Nbs1Xeno/Nbs1Hum	36
Fkh2/Tos4/Tos4Pom/Plm2	84
Tos4/Tos4Pom/Plm2	56
Dma1/Dma1Pom/Dun1/Cds1	44
Dma1/Dma1Pom	31
Dun1/Cds1	53
ChfrHum/Chk2Hum/Mek1Pom/FHA1	62
ChfrHum/Chk2Hum	82
Mek1Pom/FHA1	61

Figure 1. Bootstrapping values for Figure 10 (Page 40). Iterations run 100 times, values are for nodes specified on tree.

Appendix B: Chapter 4 Supplementary Material

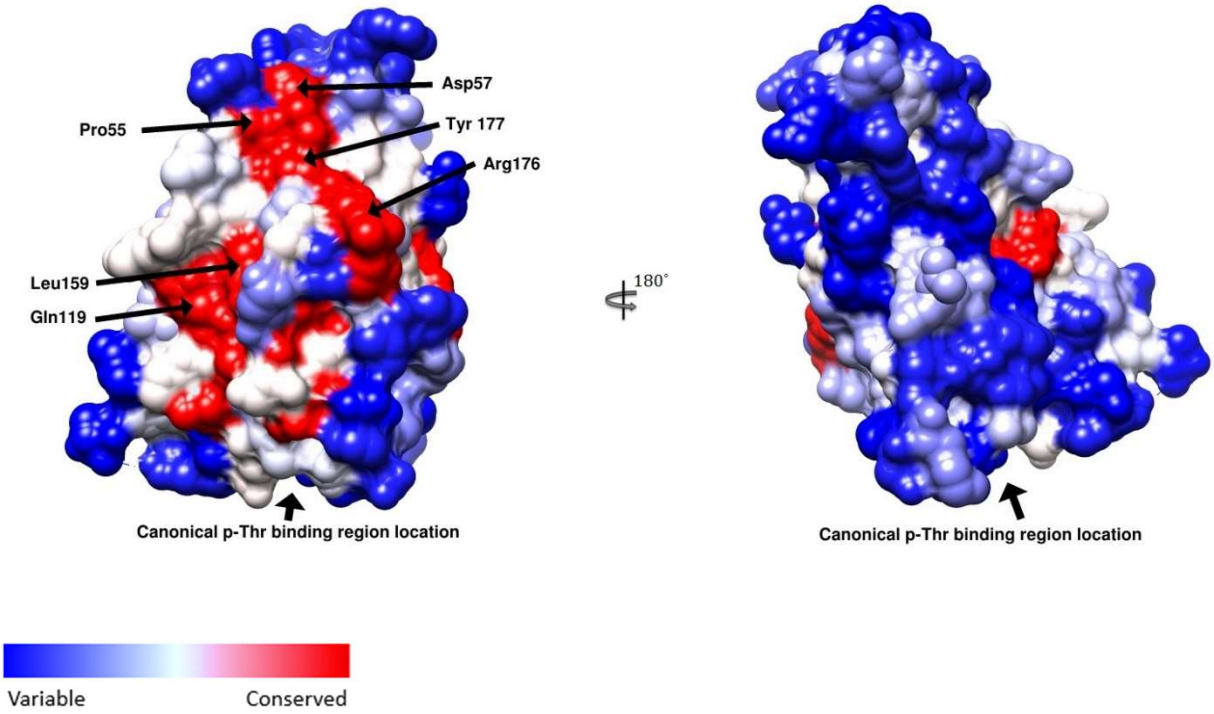

Figure 1. MSA file of Pml1 orthologs mapped into PDB structure 2JKD**.**

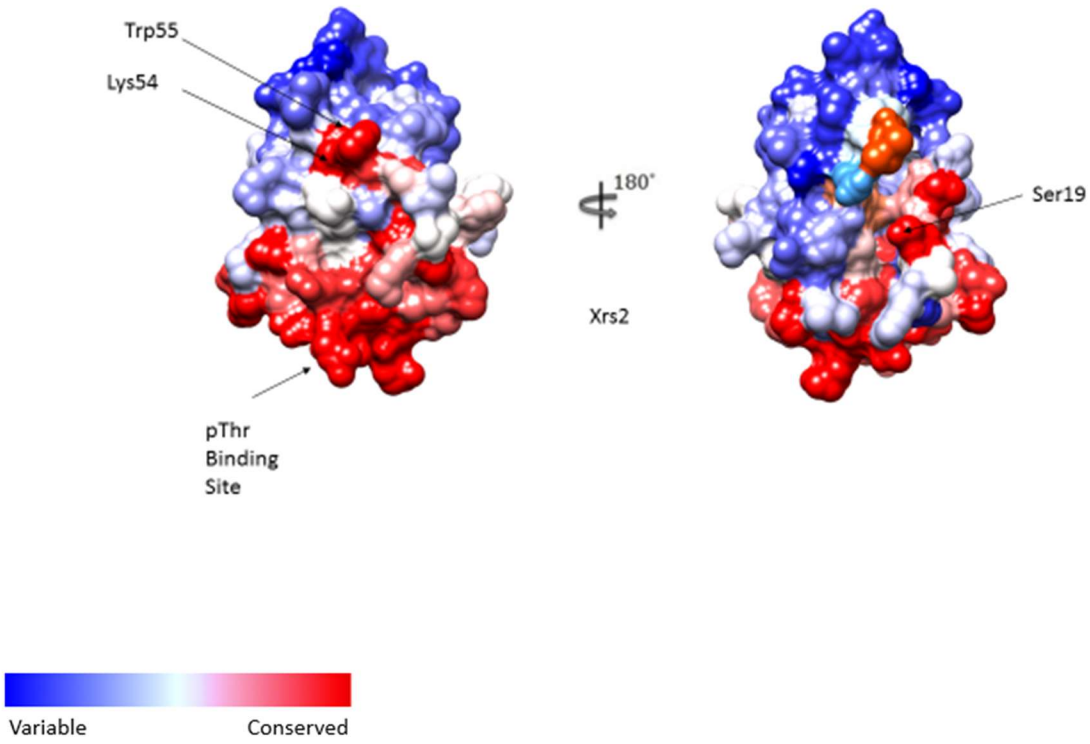

Figure 2. MSA file of Xrs2 orthologs mapped onto a Robetta Xrs2 model.

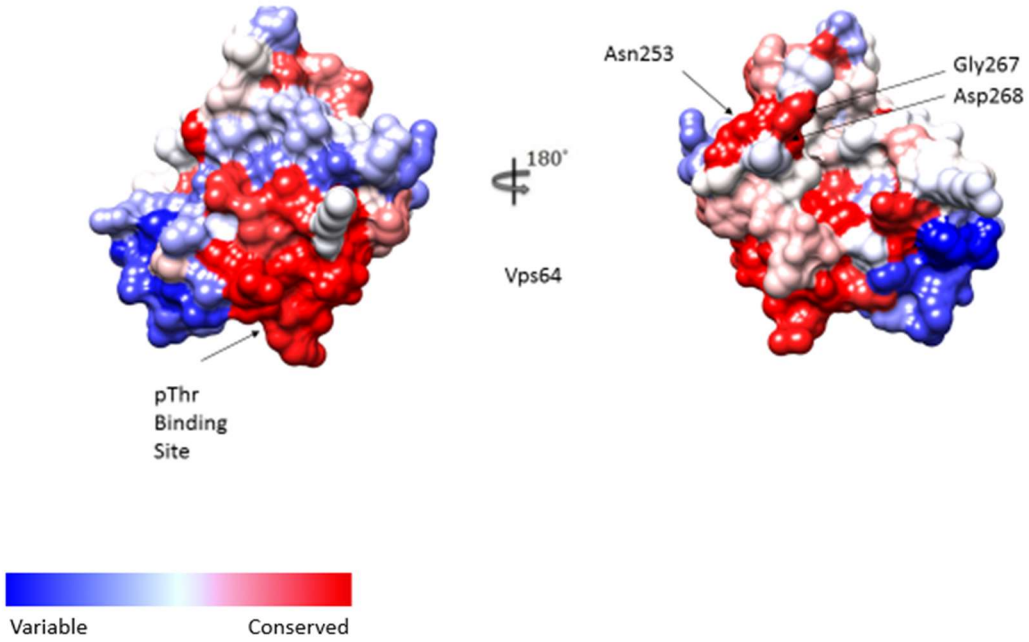

Figure 3. MSA file of Vps64 orthologs mapped onto a Robetta Vps64 model.

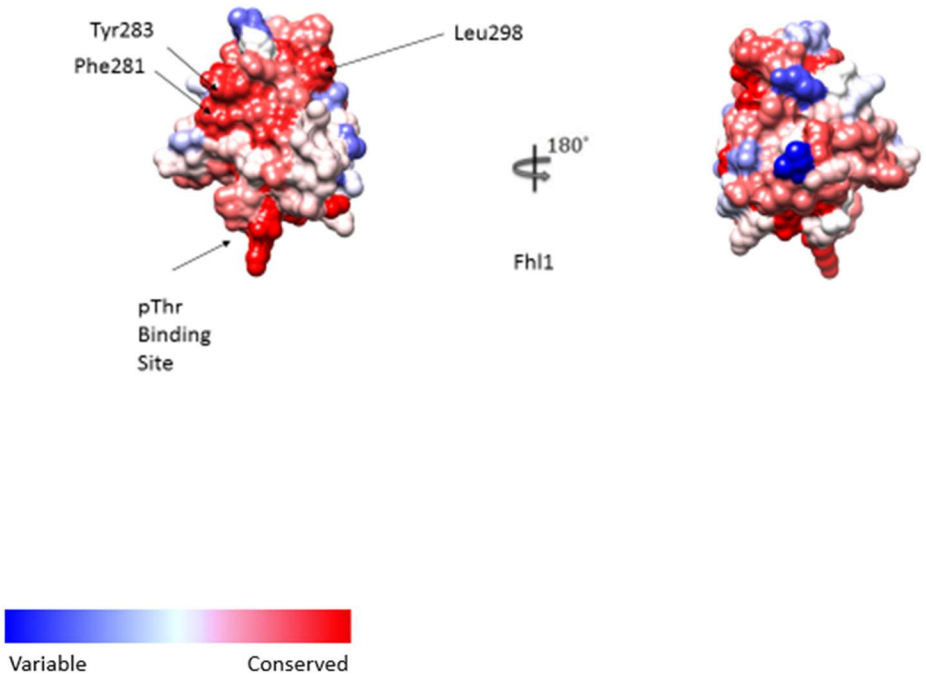

Figure 4. MSA file of Fhl1 orthologs mapped onto a Robetta Fhl1 model.

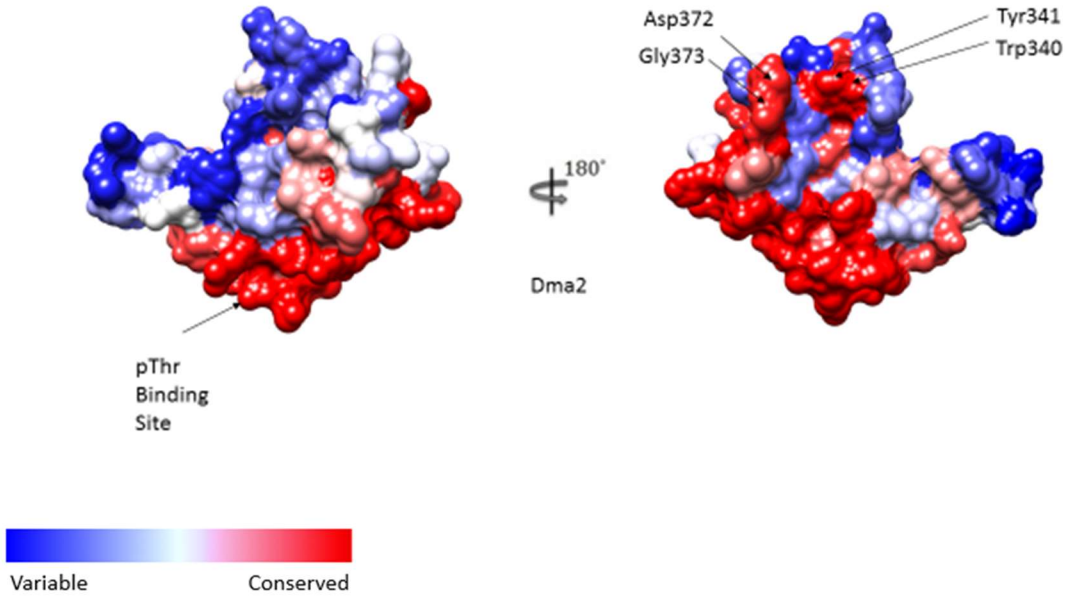

Figure 5. MSA file of Dma2 orthologs mapped onto a Robetta Dma2 model.

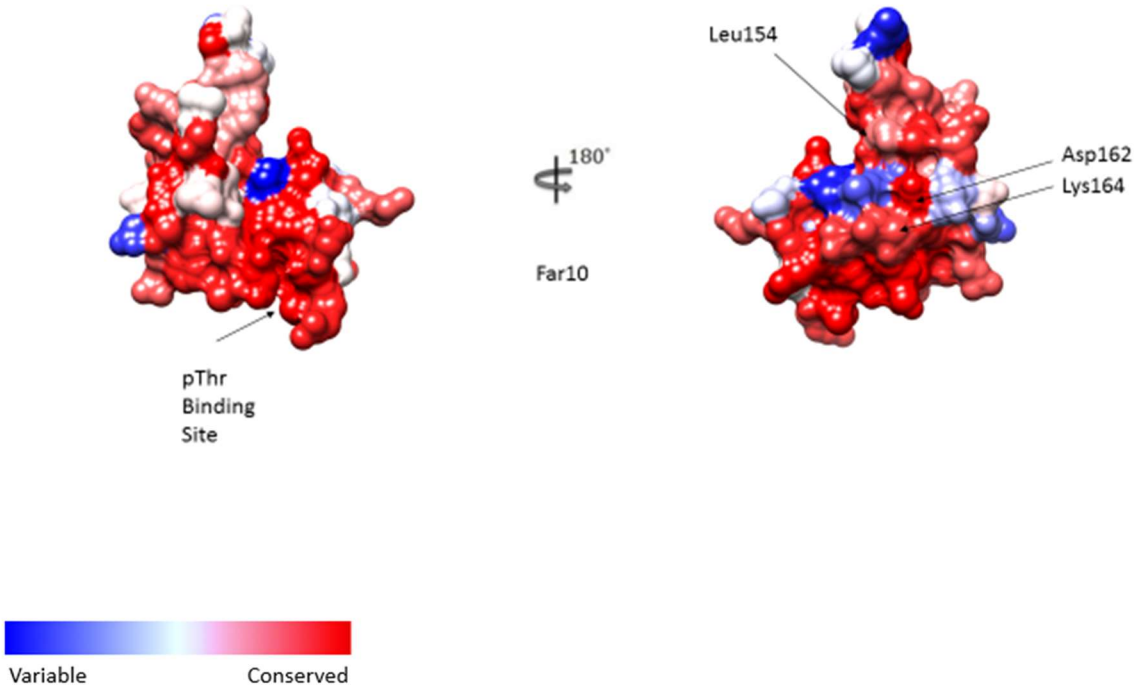

Figure 6. MSA file of Far10 orthologs mapped onto a Robetta Far10 model.

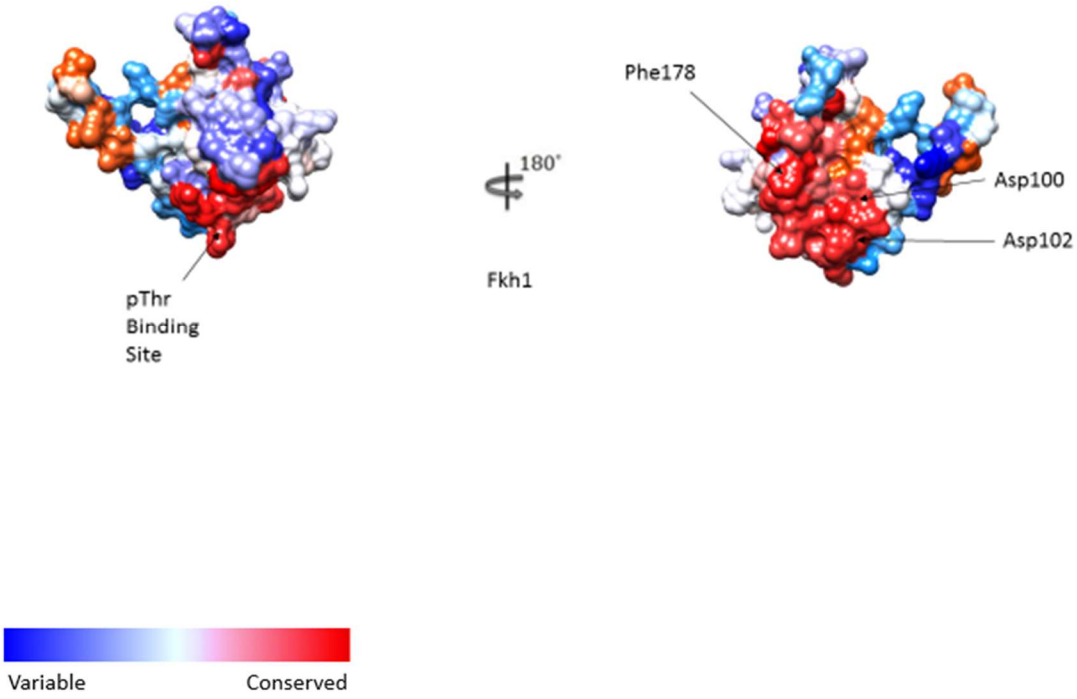

Figure 7. MSA file of Fkh1 orthologs mapped onto a Robetta Fkh1 model. Orange and cyan residues indicate insufficient data.

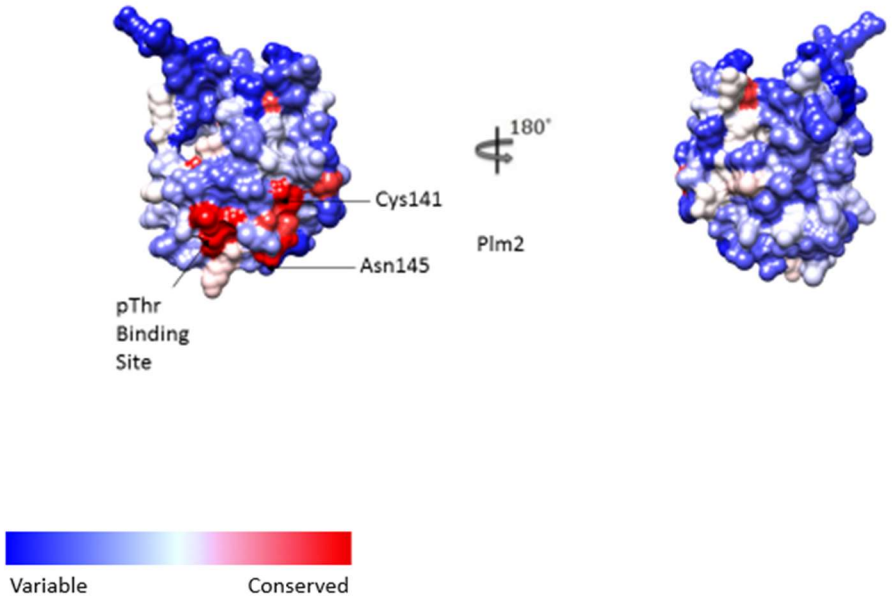

Figure 8. MSA file of Plm2 orthologs mapped onto a Robetta Plm2 model.

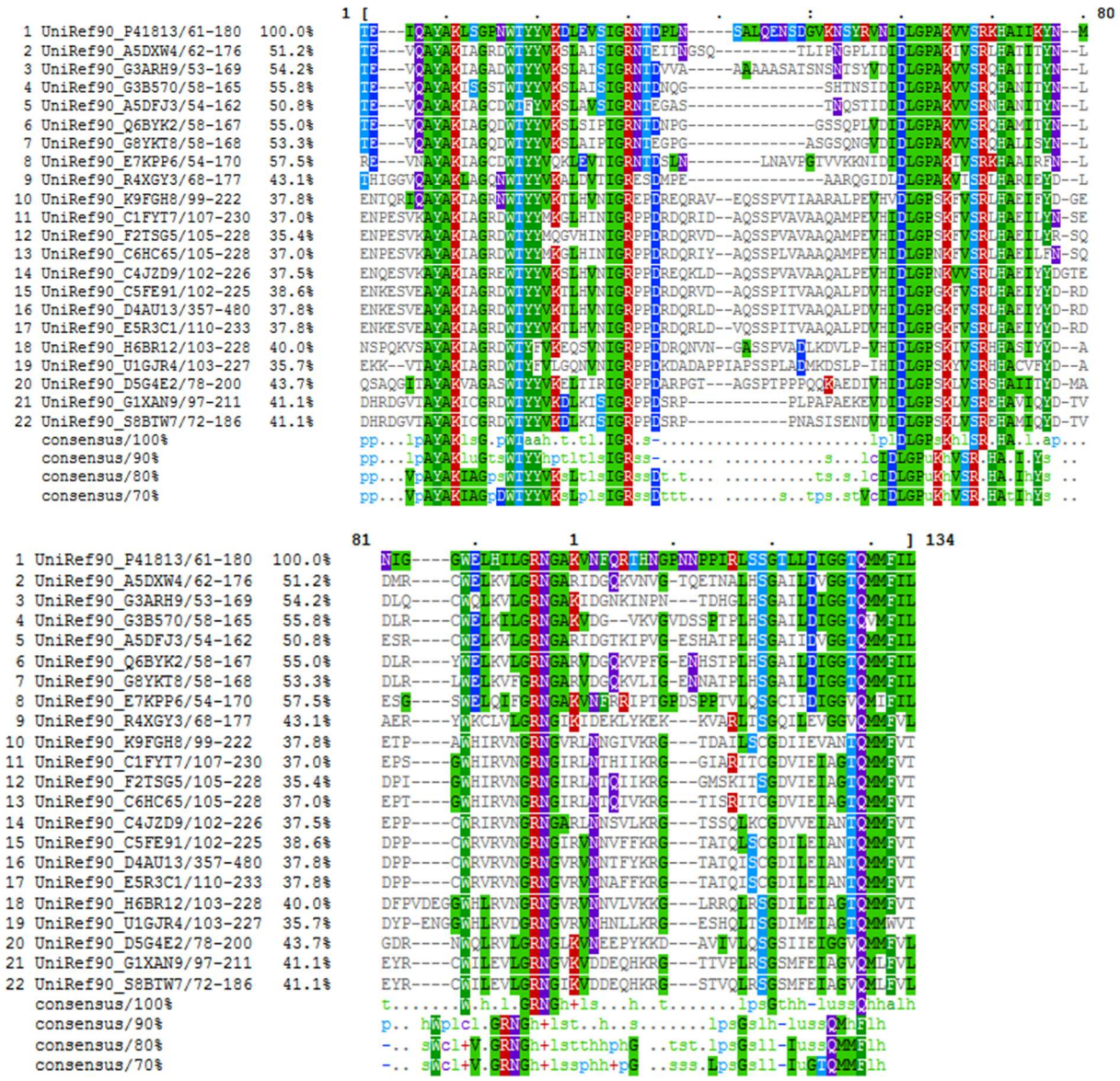

Figure 9. MSA file of Fkh2 orthologs aligned to *S. cerevisiae* Fkh2 (UniRef90_P41815). Sequences obtained as noted from the Uniprot database, with numbers following UniRef90_ referring to the specific protein used.

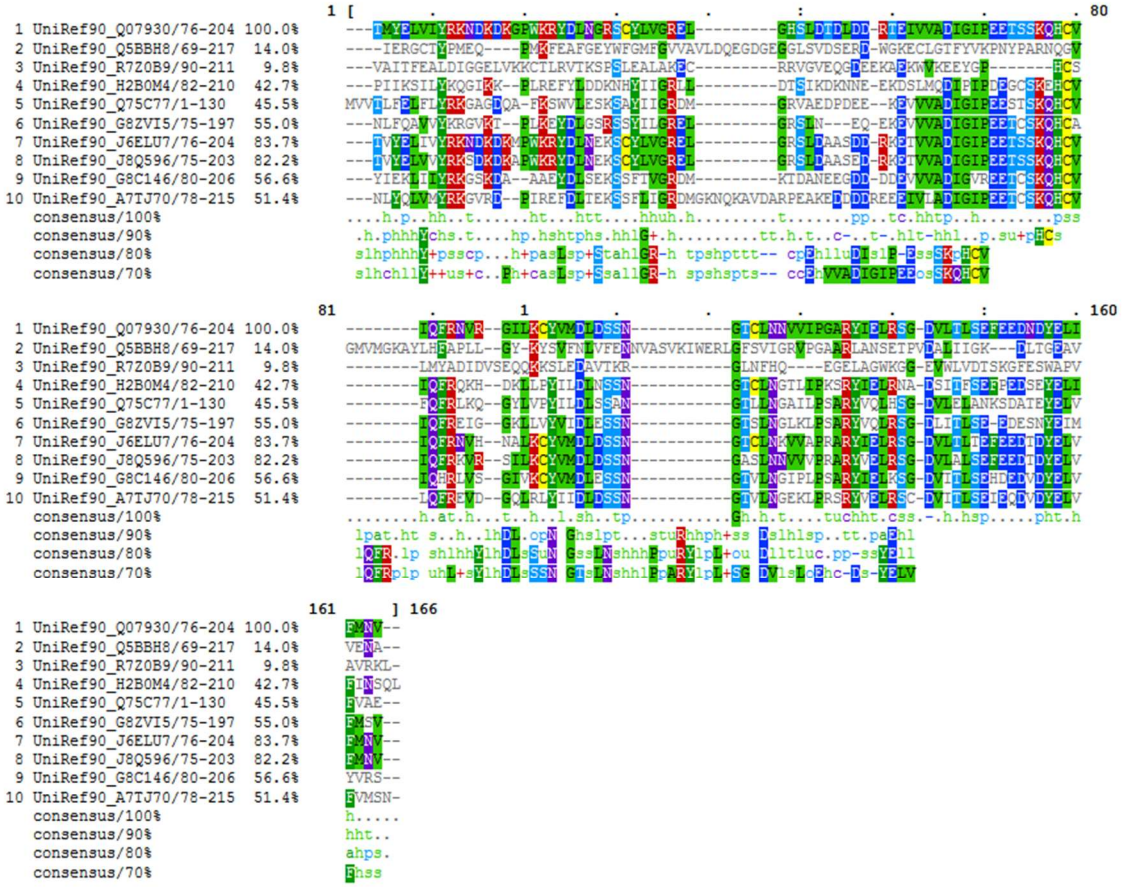

Figure 10. MSA file of Pml1 orthologs aligned to *S. cerevisiae* Pml1 (UniRef90_Q07930). Sequences obtained as noted from the Uniprot database, with numbers following UniRef90_ referring to the specific protein used.

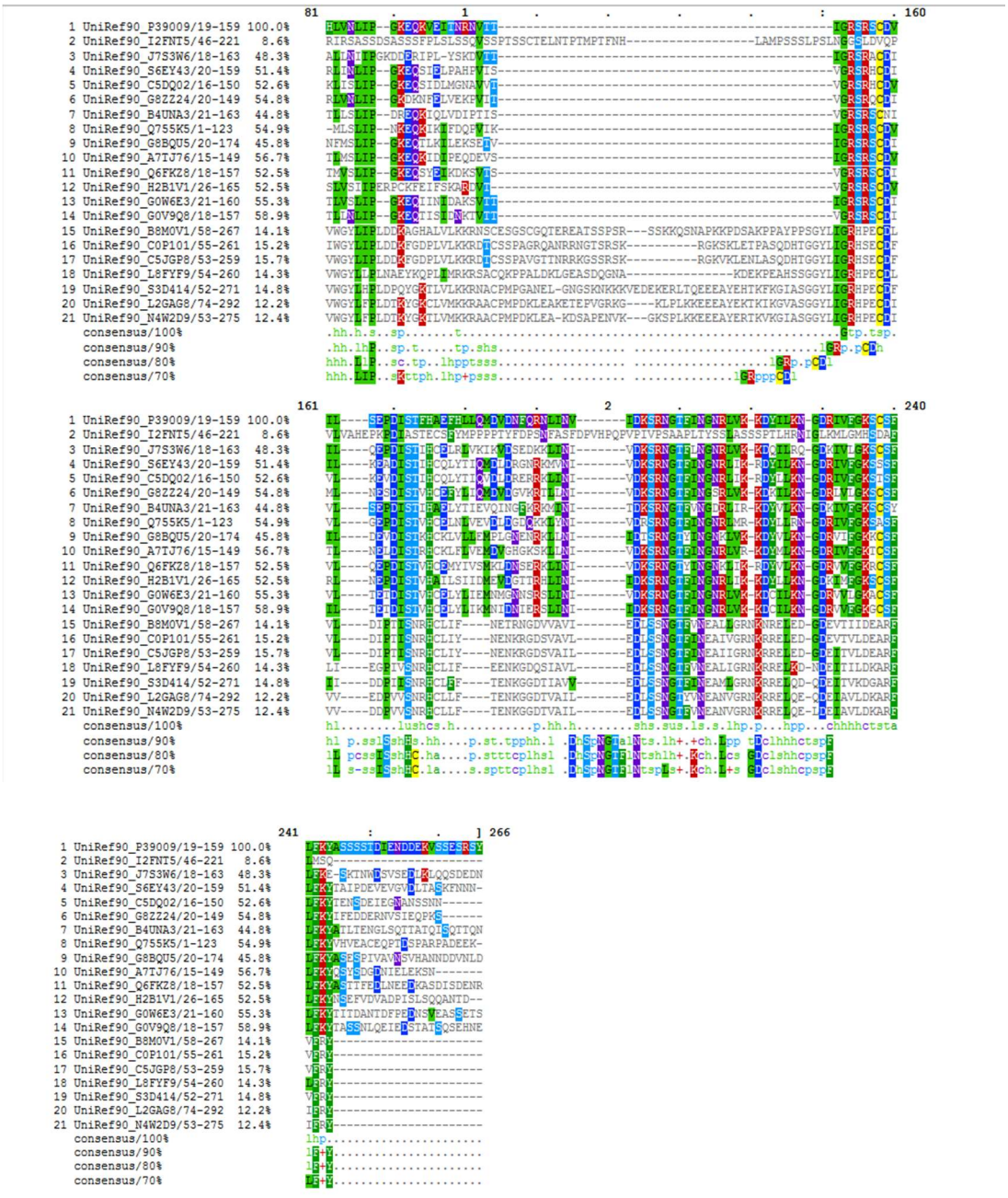

Figure 11. MSA file of Dun1 orthologs aligned to *S. cerevisiae* Dun1 (UniRef_90 P39009). Sequences obtained as noted from the Uniprot database, with numbers following UniRef90_ referring to the specific protein used.

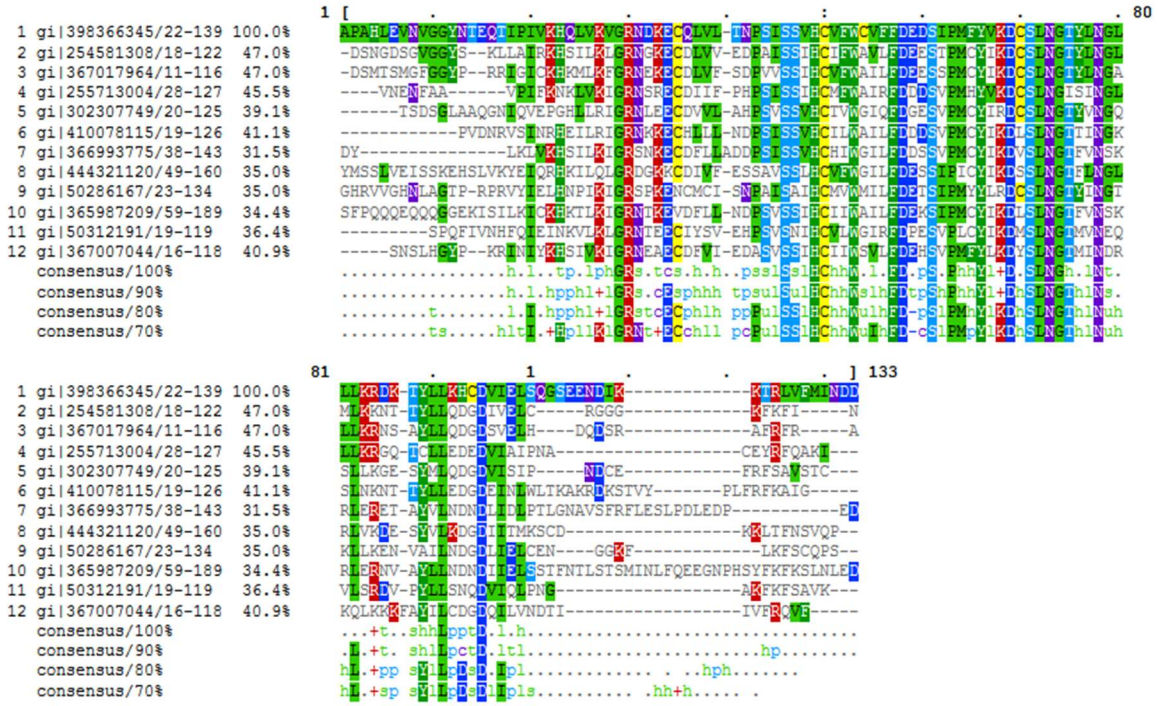

Figure 12. MSA file of Mek1 orthologs aligned to *S. cerevisiae* Mek1 (gi | 398366345). Sequences obtained as noted from the NCBI database, with numbers following gi| referring to the sequence identifier in the NCBI database.

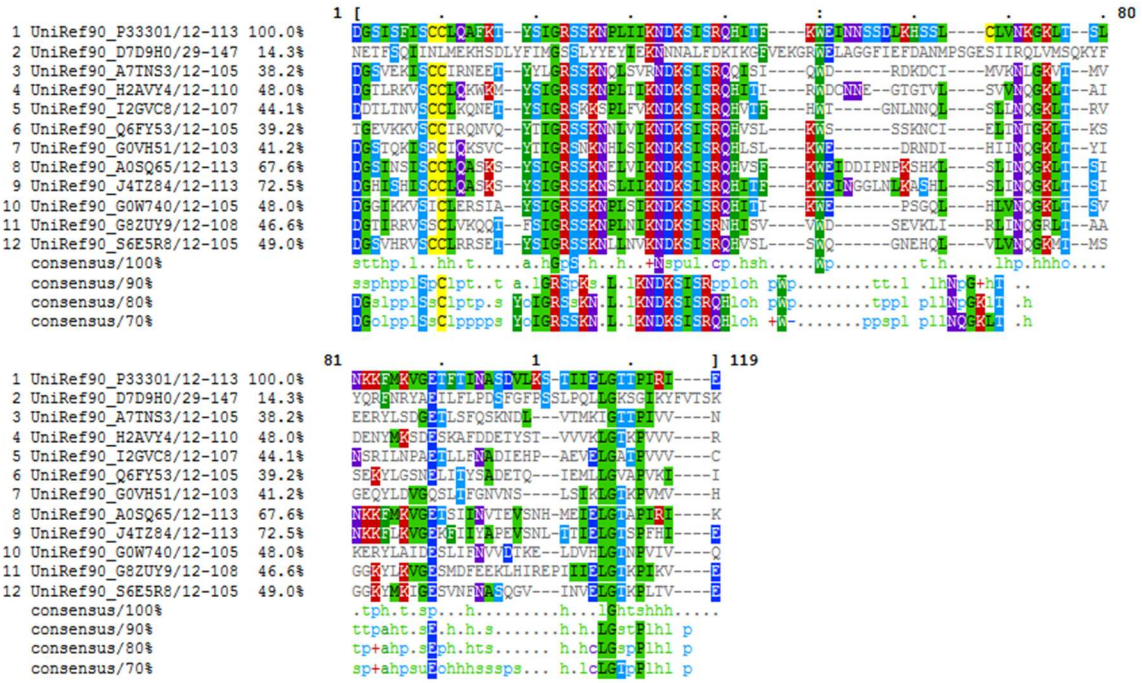

Figure 13. MSA file of Xrs2 orthologs aligned to *S. cerevisiae* Xrs2 (UniRef90_P33301). Sequences obtained as noted from the Uniprot database, with numbers following UniRef90_ referring to the specific protein used.

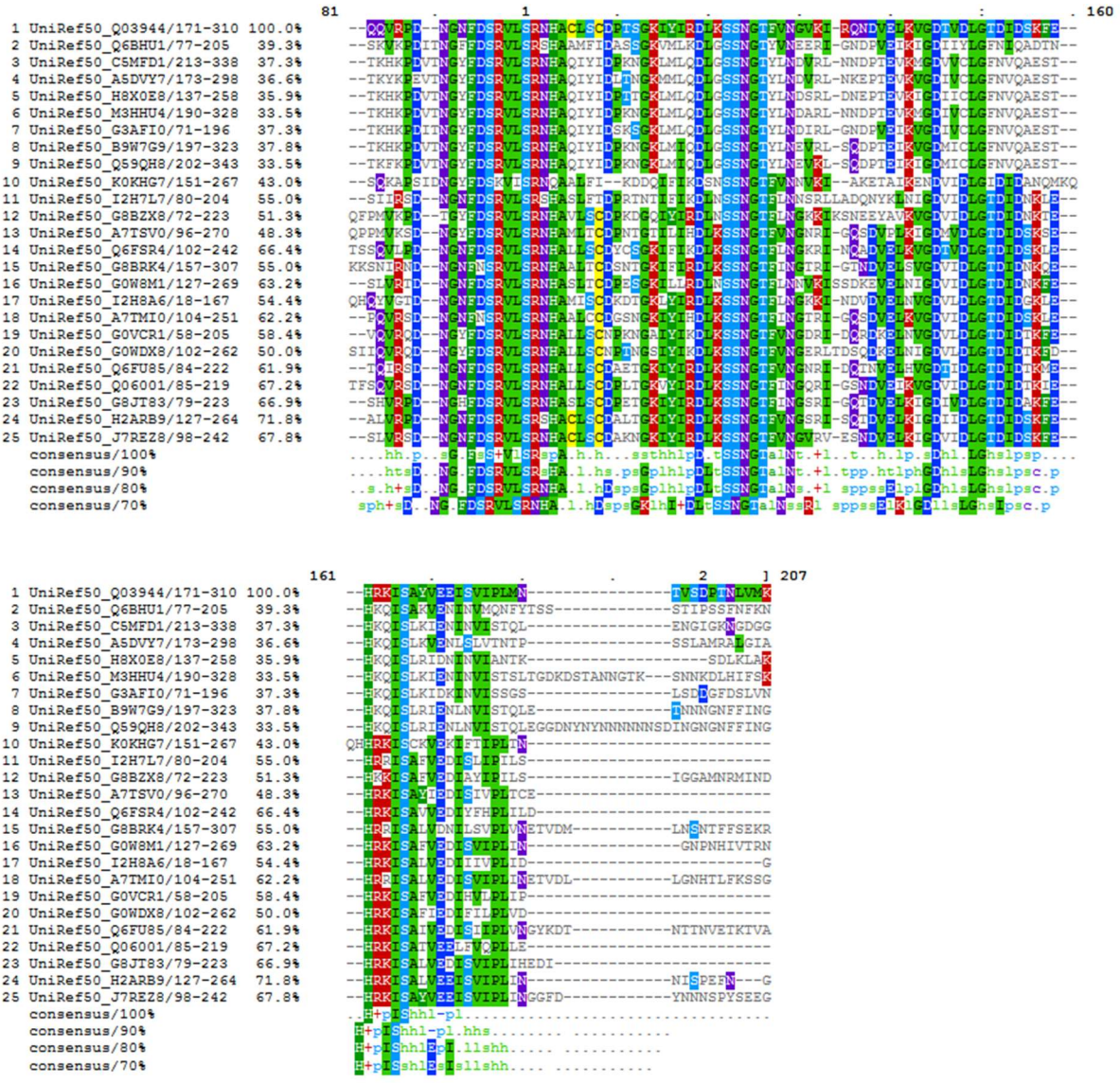

Figure 14. MSA file of Vps64 orthologs aligned to *S. cerevisiae* Vps64 (UniRef50_Q03944). Sequences obtained as noted from the Uniprot database, with numbers following UniRef90_ referring to the specific protein used.

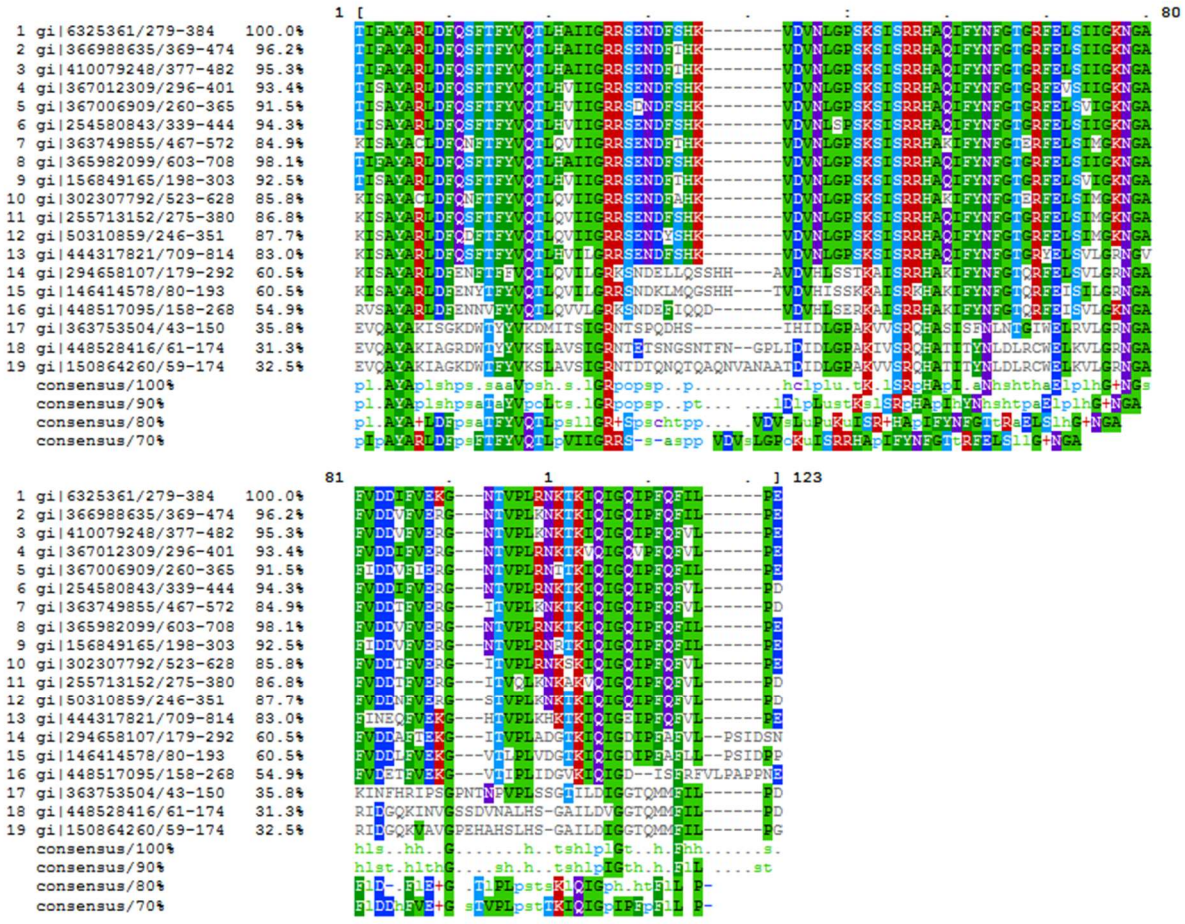

Figure 15. MSA file of Fhl1 orthologs aligned to *S. cerevisiae* Fhl1 (gi | 6325361). Sequences obtained as noted from the NCBI database, with numbers following gi| referring to the sequence identifier in the NCBI database.

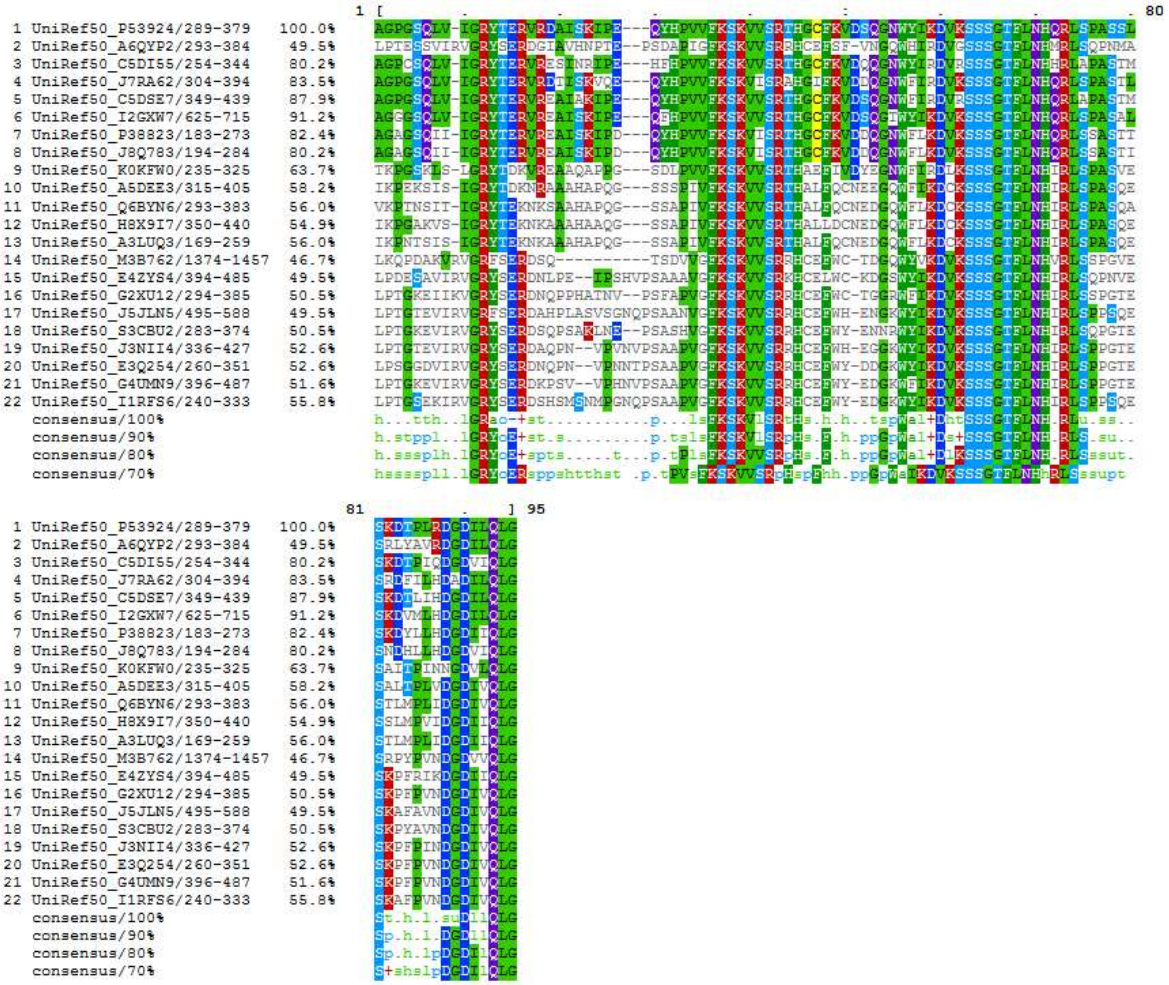

Figure 16. MSA file of Dma2 orthologs aligned to *S. cerevisiae* Dma2 (UniRef50_P53924). Sequences obtained as noted from the Uniprot database, with numbers following UniRef50_ referring to the specific protein used.

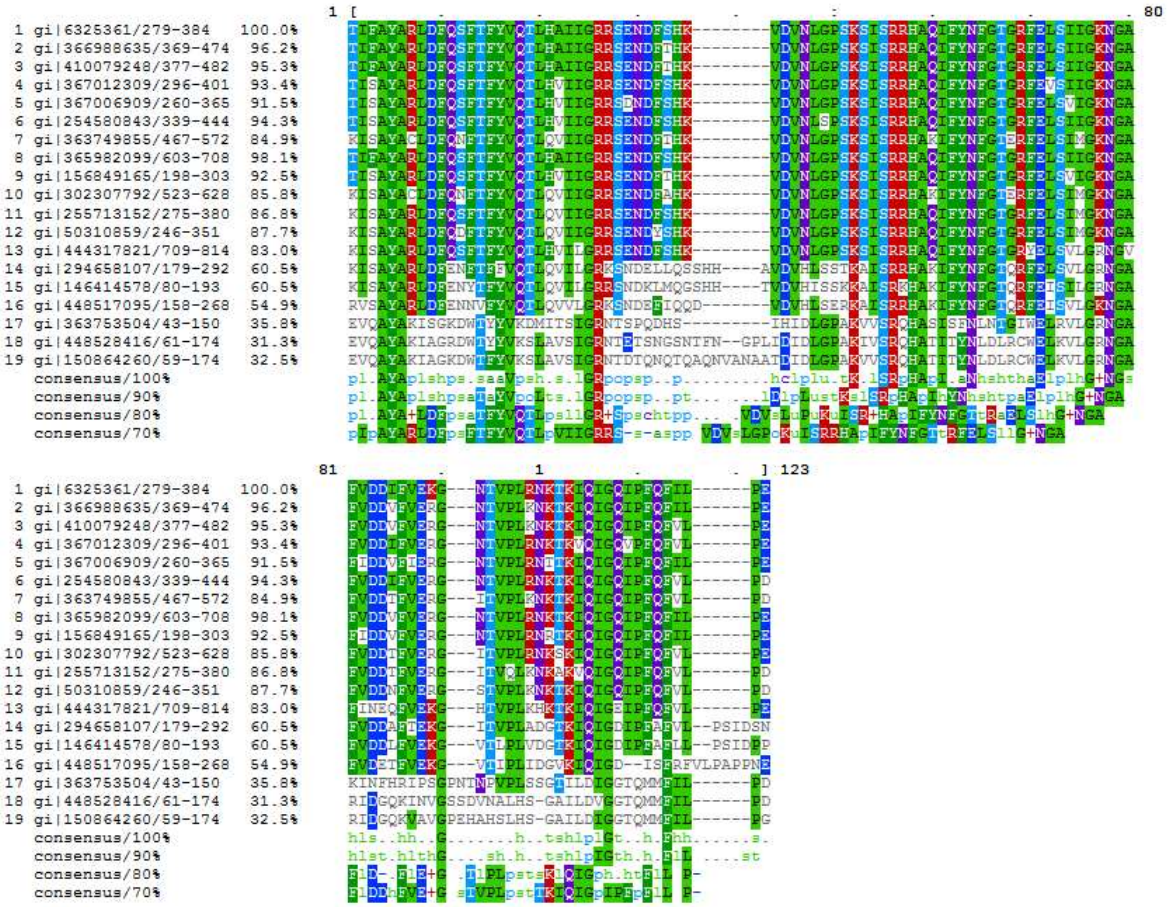

Figure 17. MSA file of Far10 orthologs aligned to *S. cerevisiae* Far10 (gi | 6325361). Sequences obtained as noted from the NCBI database, with numbers following gi| referring to the sequence identifier in the NCBI database.

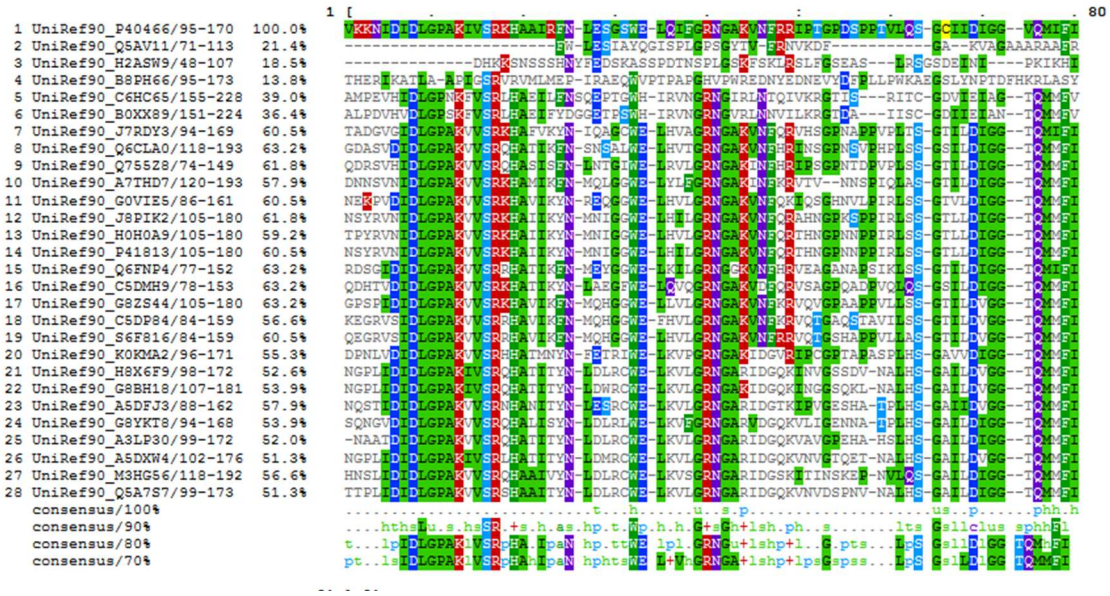

Figure 18. MSA file of Fkh1 orthologs aligned to *S. cerevisiae* Fkh1 (UniRef90_P40466). Sequences obtained as noted from the Uniprot database, with numbers following UniRef90_ referring to the specific protein used.

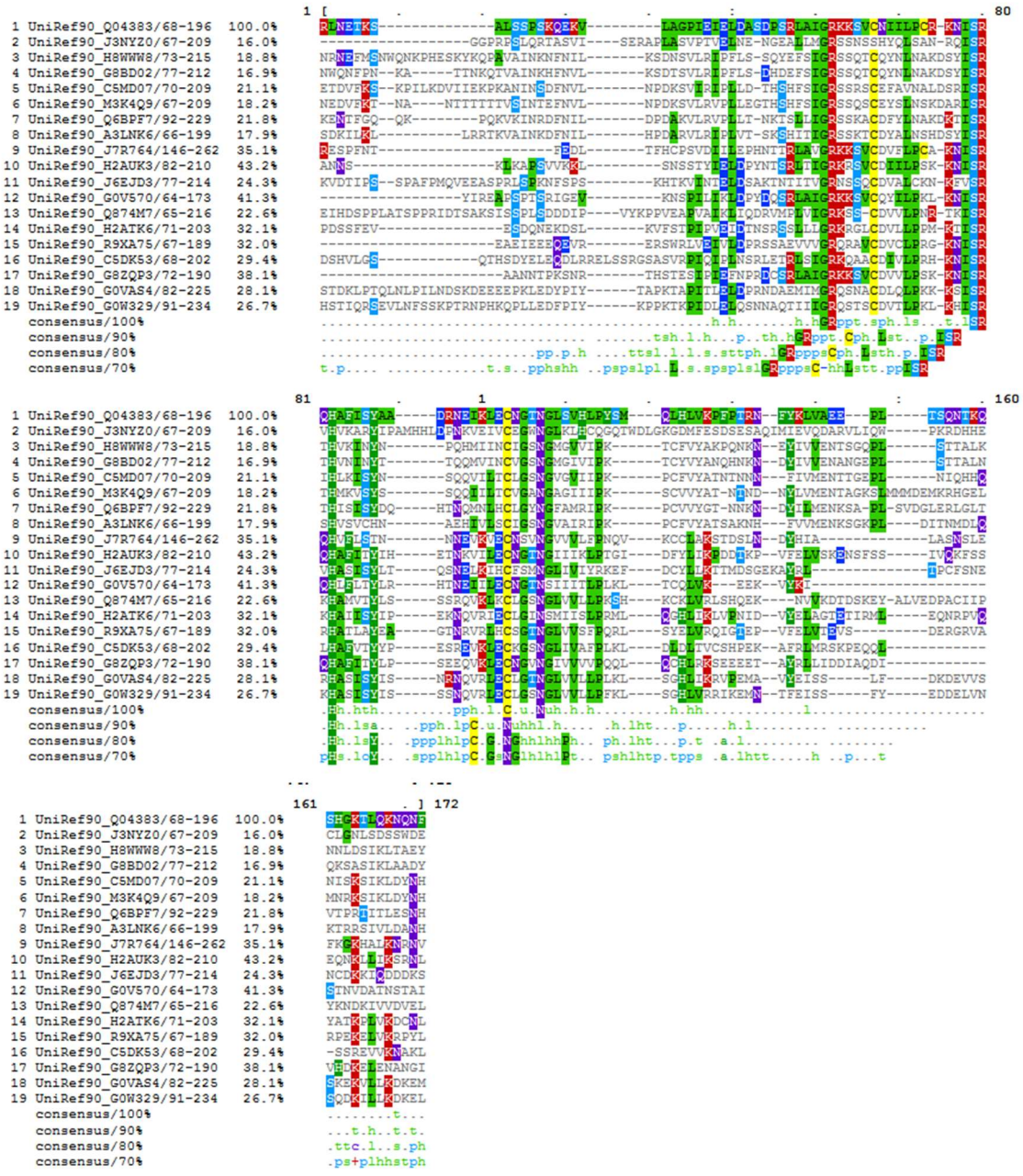

Figure 19. MSA file of Plm2 orthologs aligned to *S. cerevisiae* Plm2 (UniRef90_Q04383). Sequences obtained as noted from the Uniprot database, with numbers following UniRef90_ referring to the specific protein used.