IDENTIFICATION OF NOVEL RPA-PROTEIN INTERACTIONS USING THE YEAST TWO-HYBRID ASSAY AND IDENTIFICATION OF REGIONS IMPORTANT FOR INTERACTION BETWEEN RPA AND RAD24

A Thesis
Submitted to the Graduate Faculty
of the
North Dakota State University
of Agriculture and Applied Science

By

Gunjan Piya

In Partial Fulfillment
for the Degree of
MASTER OF SCIENCE

Major Department:
Chemistry and Biochemistry

July 2013

Fargo, North Dakota

North Dakota State University
Graduate School

Title

Identification of novel RPA-protein interactions using the yeast two-hybrid assay and identification of regions important for interaction between RPA and Rad24

By

Gunjan Piya

The Supervisory Committee certifies that this ***disquisition*** complies with North Dakota State University's regulations and meets the accepted standards for the degree of

MASTER OF SCIENCE

SUPERVISORY COMMITTEE:

Stuart J. Haring
Chair

Erika G. Offerdahl

Glenn P. Dorsam

Kendra J. Greenlee

Approved:

7/25/2013 — Date

Gergory R. Cook — Department Chair

ABSTRACT

Replication Protein A (RPA) [Replication Factor A (RFA) in yeast] is an ssDNA binding protein composed of Rpa1, Rpa2, and Rpa3 and involved in numerous DNA processing pathways such as Replication, Recombination, and Repair. It participates in such diverse pathways by its ability to interact with numerous proteins. The goal of my project was to find novel RPA-protein interactions using the yeast two hybrid assay. Using this method, we identified several known and unknown proteins that interact with Rfa1 and showed that these interactions were dependent on the phosphorylation state of Rfa2.

Next, we determine the region important for interaction between Rfa1 and Rad24. Rad24 is a checkpoint protein important for initiation of the DNA damage checkpoint signaling. By using the β-galactosidase assay, we determined the N-terminal region of Rfa1 (DBD-F) and the C-terminal region of Rad24 (460-660 aa) to be necessary for their interaction.

ACKNOWLEDGMENTS

I want to thank my advisor Dr. Stuart Haring, Haring lab members, and my advisory committee for their support and guidance throughout my time at NDSU. Special thanks go to Dr. Susan Gasser (Friedrich Miescher Institute, University of Basel, Switzerland) for providing the Rpa1 plasmids for the beta-galactosidase assay. Finally, I am thankful for all the love and support from my family.

.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AD…….. Activation Domain

APC…… Anaphase Promoting Complex

ARS…… Autonomously Replication Sequence

ATM…… Ataxia Telangiectasia Mutated

ATP…… Adenosine Triphosphate

ATR…… Ataxia Telangiectasia Mutated and Rad3 related

ATRIP…… ATR Interacting Protein

BASC…….. BRCA1 Associated genome Surveillance Complex

BER…… Base Excision Repair

BRCT……...Carboxyl-Terminal domain of BRCA1

CDK……... Cyclin Dependent Kinase

CHO…… Chinese Hamster Ovary

CMG……... CDC45/Mcm2-7/GINS

DBD……... DNA Binding Domain

DDK…… DNA Dependent Kinase

DNA……... Deoxyribose Nucleic Acid

DNA-PK…… DNA Protein Kinase

GG-NER…… Global Genomic NER

HR…….. Homologous Recombination

IR…… Infra Red

LB…… Lysogeny Broth

MAT…… MAting Type

MCM……... Minichromosome Maintenance

MMS…….. Methyl Methanesulfonate

MRN…… Mer11-Rad50-Xrs2

MRX…… Mer11-Rad50-Nbs1

MU……Miller Unit

NER…… Nucleotide Excision Repair

NHEJ…….. Non Homologous End Joining

OB…….. Oligosaccharide/ Oligonucleotide Binding

ONPG……..Ortho-Nitrophenyl-β-galactosidase

OR…….. Origin of Replication

ORC…… Origin Recognition Complex

PCNA……. Proliferating Cell Nuclear Antigen

PCR…… Polymerase Chain Reaction

PD…….. Phosphorylation Domain

PI3K……... Phosphoinositide 3 Kinase

RC……... Replication Complex

RFA……... Replication Factor A

RFC…… Replication Factor C

RNA…… Ribonucleic Acid

RPA…… Replication Protein A

SD……... Synthetic Dextrose

SG……... Synthetic Galactose

SGD…… Saccharomyces Genome Database

UNG…… Uracil DNA Glycosylase

UV…….. Ultra Violate

WHD……... Winged Helix Domain

CHAPTER 1: INTRODUCTION OF RPA AND VARIOUS DNA METABOLIC PATHWAYS

1.1. Introduction

DNA processing is at the core of our survival. Since the revelation of the DNA double helix by Watson and Crick in 1953, it has become an icon of mankind's achievement in understanding the mysteries within our cells. However, this double helix structure must be separated into single strands of DNA to accomplish replication, transcription of genes and repair of DNA damage. Unlike the stable double helix structure, single-stranded DNA (ssDNA) is bound by proteins to protect it from being degraded by nucleases and prevent unwanted annealing with other homologous DNA strands. Replication Protein A (RPA) is the major ssDNA binding protein in eukaryotes. It binds ssDNA with high affinity, but the binding is not sequence-specific. It can also bind to double-stranded DNA but with very low affinity (Wold, 1997). Replication Protein A is a heterotrimeric protein complex composed of the three subunits Rpa1, Rpa2, and Rpa3. It was first discovered to be important for *in vitro* SV40 viral DNA replication. Since then, its involvement in ssDNA binding has been well characterized, and a lot more about its function in cells has been discovered (Wold & Kelly, 1988). As ssDNA is an intermediate structure formed in numerous cellular DNA metabolic processes such as DNA replication, DNA repair, DNA recombination (mitosis and meiosis), telomere maintenance, and DNA damage checkpoint activation, RPA has also been found to be essential in all of these pathways. The evolution of a simpler homotetrameric ssDNA binding proteins from prokaryotes, to a higher ordered heterotrimeric protein in eukaryotes hints at its rising complexity and capabilities to modulate such diverse DNA metabolic processes. All three subunits of RPA are conserved among higher eukaryotes and deletion of any of three of these leads to cell inviability as they are not capable of carrying out DNA replication. Human Rpa1, Rpa2, and Rpa3 subunits are called Rfa1, Rfa2, and Rfa3 in *S.cerevisiae* respectively and share significant structural homology. Hence, working with yeast can provide insight about RPA's function in humans.

How is RPA able do so many different things? It takes part in such multifarious processes by its ability to interact with a number of proteins involved in various pathways through its Rpa1 and Rpa2 subunits, directing the assembly and disassembly onto the DNA. RPA is post translationally modified in a cell cycle-dependent manner and subsequently after DNA damage, which is thought to cause subtle

changes in its three dimensional structure, making it suitable for specific RPA-protein interactions favoring DNA repair instead of DNA replication (Binz et al, 2004). Along with being at the crossroads of so many pathways, it also plays regulatory roles in which process to activate under certain conditions, much like a traffic light directing the flow of traffic through busy intersections. Because my thesis work is mostly based on identifying and characterizing RPA-protein interactions, this chapter will focus mainly on reviewing known RPA-protein interactions and the role of RPA in different facets of DNA metabolism.

1.1.1. Structure of RPA

Replication Protein A is made up of three subunits of approximately 70 kDa, 32 kDa, and 14 kDa, called Rpa1, Rpa2 and Rpa3, respectively. Table 1.1 shows sizes of different RPA subunits in a variety of species. This complex is highly stable, even in 6M urea (Wold, 1997). Due to the large and complex architecture of the RPA, the entire complex has not been crystallized yet; rather individual subunits have been crystallized separately. This leaves us with a fragmented and incomplete picture of the RPA heterotrimer from a structural biology prospective. Replication Protein A has six DNA binding domains (DBDs), DBD-F/A/B/C in Rpa1, DBD-D in Rpa2, and DBD-E in Rpa3. Each DBD is made up of an oligonucleotide/oligosaccharide binding (OB) fold, where five beta sheets run in opposite directions, along with an alpha helix between helix 3 and 4, forming a beta-barrel structure (Murzin, 1993). The OB fold is able to bind ssDNA via base pair stacking and electrostatic interactions. DBD-A and DBD-B binds ssDNA with the highest affinity [$K_a \sim 10^6$-10^8 M^{-1}] while DBD-C, DBD-D, and DBD-F bind DNA with lower affinity. RPA can also bind dsDNA and RNA but with affinities at least three orders of magnitude lower than ssDNA (Brill et al, 1989).

DBD-F is necessary for binding partially duplex DNA and plays an important role in the DNA damage response pathway as it is shown to interact with various repair proteins. Finally, DBD-C (Rpa1), DBD-D (Rpa2), and DBD-E (Rpa3) forms the trimerization core of the RPA complex (Bochkareva et al, 2002) which is assumed to stabilize the protein complex. While the individual subunits are functional in DNA binding, the whole complex is necessary for most DNA metabolic activities mediated by RPA. Recently, DBD-C, D, E has also been shown to be involved in stabilizing and unfolding G-quadruplex providing further evidence that different DBDs of RPA have different binding modes and recognize different structure (Prakash et al, 2011).

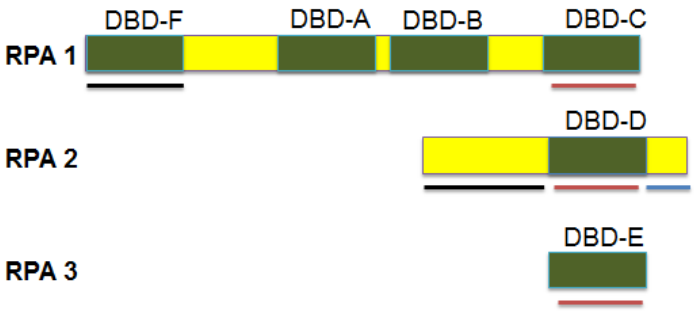


Figure 1.1: Different BDB domains of RPA. DNA Binding Domains are denoted with DBD. Red line indicates regions necessary for complex formation (Cai et al, 2007). Black line indicates regions important for phosphorylation during the cell cycle (Din et al, 1990) and in response to DNA damage (Carty et al, 1994; Liu & Weaver, 1993). Blue line indicates the winged helix domain (Mer et al, 2000)

The OB fold domain of DBD-C contains a Zinc finger, which contributes to strong binding affinity between RPA and ssDNA (Lao et al, 2000). DBD-F also contains a basic cleft region, shown to be important for mediating RPA-protein interaction with RAD9 and regulating ATR signaling in humans, a prominent DNA checkpoint-signaling pathway (Xu et al, 2008). Another Rpa1 mutant *rfa1-t11* (K45E), located in DBD-F has been shown to cause defects in double-strand break (DSB) repair and single-strand annealing (SSA) (Umezu et al, 1998). DBD-F is connected to the rest of the protein via a linker region which is thought to give flexibility to the protein but its precise function is unknown.

The Rpa2 subunit has an N-terminal unstructured region referred to as the phosphorylation domain (PD). This domain is phosphorylated in a cell cycle dependent manner and hyper-phosphorylated following DNA damage by events such as UV radiation (Shao et al, 1999). Following genotoxic stress such as dsDNA break, Rpa2 is phosphorylation on S4 and S8 by DNA-PK. This prevents cells from entering mitosis by blocking homologous recombination demonstrated by reduced Rad51 foci formation at sites of DNA damage (Liaw et al, 2011). This could presumably buy time for DNA repair to occur. This study is complemented by a different study which showed phosphorylation of Rpa2 to have reduced binding affinity for ssDNA (Wold, 1997). Hyper-phosphorylation of Rpa2 may also play a role in modulating cellular pathways by altering DBD-B mediated RPA-DNA and RPA-protein interactions. The authors came to this conclusion after seeing the results demonstrating resistance of Lys-343, Arg-335, and Arg-382 to chemical modification, fluorescence quenching to Trp-461, and reduced proteolysis of DBD-B after chymotrypsin treatment (Liu et al, 2005). This mechanism of phosphorylated Rpa2 N-terminal domain interacting with DBD-B can mimic a negatively charged ssDNA strand and weaken

overall RPA-DNA binding affinity. Preliminary results in our lab also support this model by demonstrating that protein interactions mediated specifically by DBD-B but not DBD-F are affected in the presence of Rpa2-Asp (the hyperphosphorylated form of Rpa2).

Species	Rpa1 (kDa)	Rpa2 (kDa)	Rpa3(kDa)
Homo sapiens	68	29	13.5
Saccharomyces cerevisiae	70.3	29.9	13.8
Schizosaccharomyces pombe	68.3	30.3	11.8
Drosophila melanogaster	70	30	8
Caenorhabditis elegans	73.2	32.1	19
Bos taurus	70	30	11

Table 1.1: Size comparison between different RPA subunits among various species (Wold, 1997)

Despite continued effort to understand the meaning behind this phosphorylation event, the significance of this phenomenon remains largely unclear. The Rpa2 PD has multiple serine and threonine residues that are targeted by various PI3K related Kinases such as CDK, DNA-PK, ATM, and ATR (Wang et al, 2001). Human studies have shown that Cdc2 can phosphorylate Ser-23 and Ser-29 (Dutta & Stillman, 1992) during normal the cell cycle, while numerous possible sites are identified on Rpa2 following DNA damage. Potential phosphorylation sites after UV radiation include Ser-4, Ser-8, Ser-11, Ser-12, Ser-13, Thr-21, Ser-23, Ser-29, and Ser-33 on Rpa2. Additional sites including Ser 569, Thr 580, Ser 585, and Thr 590 in Rpa1 were also shown to be phosphorylated following UV radiation (Nuss et al, 2005).

Studying individual mutations of Rpa2 in humans is extremely difficult because it constitutes knocking down the wild-type Rpa2 using techniques such as siRNA and supplementing the mutant form of Rpa2. Even then, there is always some residual amount of wild-type Rpa2 interfering with the studies. Using yeast as a model system to study this process can make experiments easier and more efficient as we can knock down wild-type genes in yeast easier than we can in mammalian cells and testing the individual mutants are also less time consuming. Hence, in order to understand which specific residues are important for RPAs function in the cell cycle and DNA damage, our lab replaced each potential phosphorylation sites of Rpa2 with Asp and Ala mutants and looked at their effects in Δ*Rpa2* yeast strains. After extensive testing of these individual Rpa2 phospho-mutants, no single residue appeared

critical in the response to DNA damage. These observations suggested that simultaneous phosphorylation of multiple Ser/Thr residues might be necessary for DNA replication and damage response. To further test and validate this, our lab is utilizing a "multi-mutants" approach where clusters of Ser/The residues are mutated to mimic phosphorylated or unphosphorylated forms of RPA and they will be tested to see if any such cluster might be of importance for DNA damage response.

Rpa2 also has a C-terminal winged-helix domain (WHD) which is important in RPA-protein interaction (Bochkarev et al, 1999). It has previously been shown to interact with human uracil-DNA glycosylase ((Nagelhus et al, 1997) and Xeroderma Pigmentosum A (Li et al, 1995) important in Base Excision Repair (BER) and Nucleotide Excision Repair (NER) pathways respectively. An alternative form of Rpa2, called Rpa4, is found to be present in human cells (Keshav et al, 1995). Rpa4 is able to interact with Rpa1 and Rpa3 just like Rpa2, but this alternative RPA complex (aRPA) containing Rpa4 does not support *in vitro* or *in vivo* DNA replication (Haring et al, 2010; Mason et al, 2009). Instead, studies show Rpa4 complex with Rpa1 and Rpa3 may be important specifically during DNA repair (Kemp et al, 2010). Similarly, other organisms also have alternative forms of RPA where subunits can shuffle to form alternative complexes to mediate different processes. Rpa3 is mostly considered to be important in stabilizing the RPA complex with little role in DNA metabolism but recent studies have shown that it is also capable of binding ssDNA (Salas et al, 2009). Co-expressing Rpa2 and Rpa3 can form a stable complex while co-expressing Rpa1 with either Rpa2 or Rpa3 alone fails to do so, indicating that Rpa2 and Rpa3 may form a sub complex which then stabilizes Rpa1 (Henricksen et al, 1994).

1.1.2. DNA Replication Overview

DNA replication is a highly regulated process of making copies of DNA strands, which are ultimately segregated from mother to daughter cells during mitosis. Cells have high fidelity polymerases to minimize errors while copying and repair enzymes to correct any mistakes during and after DNA replication. DNA replication is primed during G1 phase by binding of licensing factors to the origin of replication sequences. The actual replication process is carried out during the S phase by DNA polymerases and chromosome segregation eventually takes place during the S-phase of the cell cycle. We will discuss some major events during this process in further detail.

1.1.3. Replication Origin

DNA replication starts at distinct sites within the genome called Origins of Replication (OR), which in yeast are also called Autonomously Replication Sequences (ARS). These sites are distributed throughout the genome and are recognized by the Origin Recognition Complex (ORC), which is a six-subunit complex composed of ORC (1-6). The subunits Orc1-5 are important for recognizing and binding the ARS, while Orc6 is necessary for stabilizing the pre-Replication Complex (pre-RC) composed of Cdc6p, Cdt1p, and Mcm2-7p. Orc1 also has an ATPase domain which is critical for ORC binding to DNA (Speck et al, 2005). ORC, in the presence of ATP, can summon the necessary proteins at this site, forming a pre-RC complex during the G1 phase (Bell & Stillman, 1992). This event is necessary for licensing the chromosomes before they can start replication in S-phase. *Saccharomyces cerevisiae* origin of replication is comprised of the A element with an 11-17 bp consensus sequence and B1 elements flanking the A element spanning up to 150 bp. A close relative *Schizosaccharomyces pombe* ARSs are relatively larger than *S. cerevisiae* with about 500-1000 bp long AT rich sequences, and are distributed randomly throughout the genome (Segurado et al, 2003).

While *S. cerevisiae* and *S. pombe* possess a somewhat strict consensus sequence for ORC binding, other higher eukaryotic species have a less obvious consensus sequence (Bouton et al, 1987). In humans, the OR consensus sequence and replication firing timing are much more indiscriminate compared to yeast, but still preserve the characteristic AT rich sequence (Karnani et al, 2010). The ORC bound to OR can now assemble Cdc6p, Cdt1p, and Mcm2-7p to form the pre-RC complex, which is triggered by kinases to initiate DNA replication once a cell enters S-phase. Interestingly, one of the proteins we discovered using our two-hybrid assay was Mcm5, which is part of the Mcm2-7 complex. Interaction between RPA and Mcm5 (part of Mcm2-7 helicase) has not been previously characterized and will be discussed in further detail in chapter 2.

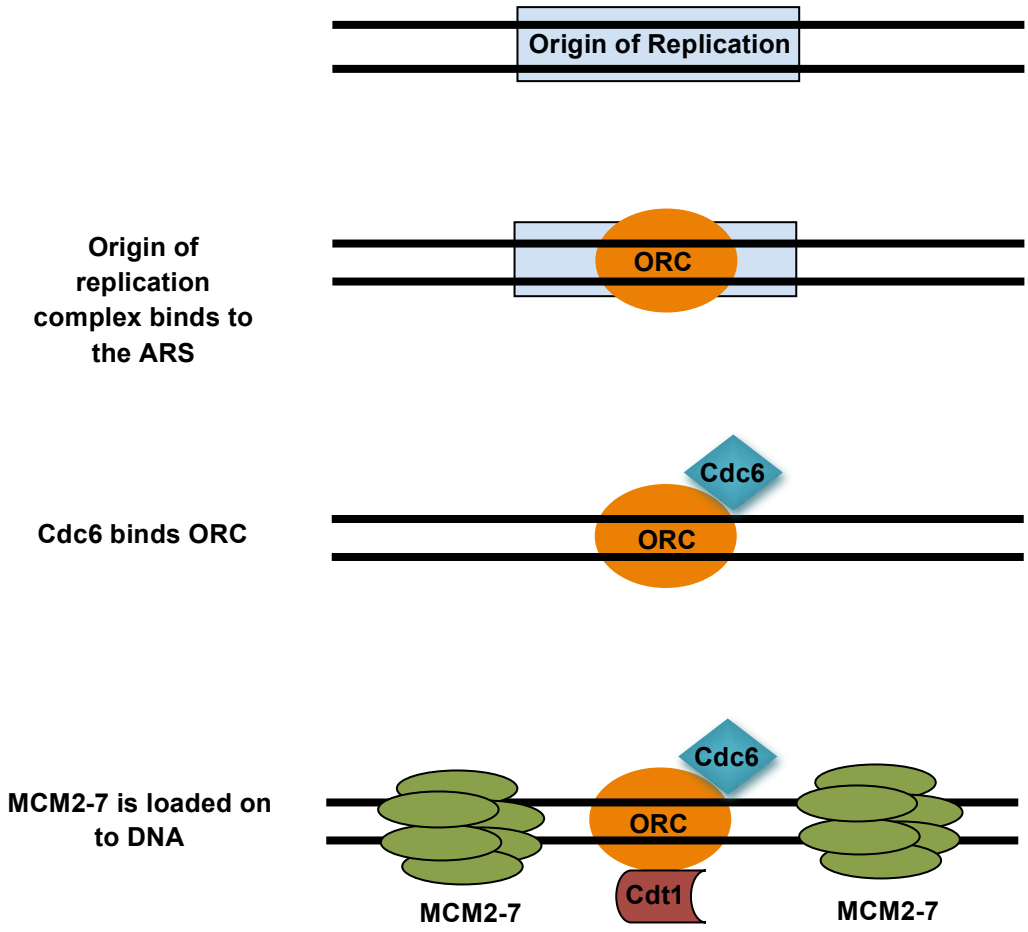


Figure 1.2: Loading of replication factors on to the origin of replication. During the G1 Phase, The pre-RC complex is assembled onto the ARS. First, ORC (1-6) recognizes the OR and then recruits MCM2-7 with the help of Cdc6 and Cdt1. This process is also acts as a licensing factor for DNA replication during the S-phase. The pre-RC complex is activated into an active replicative complex via the action of two kinases S-CDK and Cdc7-Dbf. Upon activation it can start copying the genome.

1.1.4. Initiation of DNA Replication and Fork Progression

The Mcm2-7 complex can recruit additional factors Cdc45 and GINS forming a CMG complex which can activate the helicase domain of Mcm2-7 (Gambus et al, 2006). The CMG complex in yeast is activated by the simultaneous action of Clb5-Cdk1 (S-CDK) and Cdc7-Dbf4 (DDK) to make sure the replication only happens once per cell cycle and is limited to S-phase (Nougarède et al, 2000). As mentioned later in section 3 of this chapter, DDK is a CDK-cyclin complex where Dbf4 presence during late G1/S phase activates Cdc7 and in turn activates the Mcm2-7 complex. Dbf4 is targeted for proteosomal degradation by the Anaphase Promoting Complex (APC) during the G1 phase which keeps Cdc7 from activating the MCM complex (Ferreira et al, 2000). It is to be noted that we were able to isolate

Apc4 (a subunit of the APC complex) with our yeast two hybrid screen as a novel protein interacting partner of Rpa1. I will discuss Apc4 in the next chapter. After the MCM helicase starts to unwind the dsDNA at the OR, RPA binds the newly formed ssDNA and loads additional factors required for replication (Tanaka & Nasmyth, 1998). First DNA Polymerase Alpha (Pol-α) is loaded on to the ssDNA with help of Mcm10 and Ctf4 which stimulates the polymerase activity of Pol-α and is necessary to synthesize an RNA primer that is required for synthesis of short stretches of DNA. Consequently, based on current knowledge, Pol-α falls off the strand and is replaced by DNA Polymerase Epsilon (Pol-ε) and DNA Polymerase Delta (Pol-δ) on the leading and lagging strands. These highly possessive polymerases are loaded on to DNA by a clamp-clamp loader system where the 'clamp loader' Replication Factor C (RFC) loads the 'clamp' Proliferating Cell Nuclear Antigen (PCNA) onto the DNA to create a complete replisome complex (Kunkel & Burgers, 2008). PCNA is a homotrimeric protein which acts as a processivity factor for both Pol-ε and Pol-δ, similar to a molecular tool belt holding all the 'tools' for replication intact onto DNA. Through post-translational modifications (ubiquitylation and SUMOylation), PCNA coordinates repair pathways that can occur by replication errors. A similar 'clamp-clamp loader' system also persists in DNA damage repair pathways, where they are necessary to initiate the checkpoint signal and recruit DNA polymerase required for translesion synthesis.

Two replication forks are created when the replisome start moving in opposite directions from the OR. Each fork has a leading and lagging strand. The leading strand is synthesized continuously because of the innate ability of Pol-α to synthesize DNA only from 5'→3' direction; while the lagging strand is synthesized discontinuously in short Okazaki fragments. Okazaki fragments create nicks in the DNA that need to be joined by phosphodiester bond formation, similar to those generated by DNA damage and repair. The common pathways to repair these nicks include a concerted action of Pol-δ, FEN1 and DNA ligase I. Pol-δ has a special ability to remove primer template via its 5'→3' exonuclease activity. The resulting flap is cleaved off by FEN1 flap endonuclease. The removal of primer creates a gap which is filled by Pol- δ via DNA synthesis. Since, Pol-δ cannot add nucleotides at the 5' end, a nick remains on the lagging strand DNA which is joined together by DNA ligase I to complete the process. Pol32, a subunit of Pol-δ was isolated as an RPA-interacting partner in our yeast-two hybrid assay. Exo1, a flap endonuclease and Dna2, an endonuclease/helicase complex has also been shown to work alongside

FEN1 processing Okazaki fragments via alternative pathways. We isolated fragments of Dna2 in our yeast-two hybrid assay as well, demonstrating further example of RPA's involvement in the various DNA replication process.

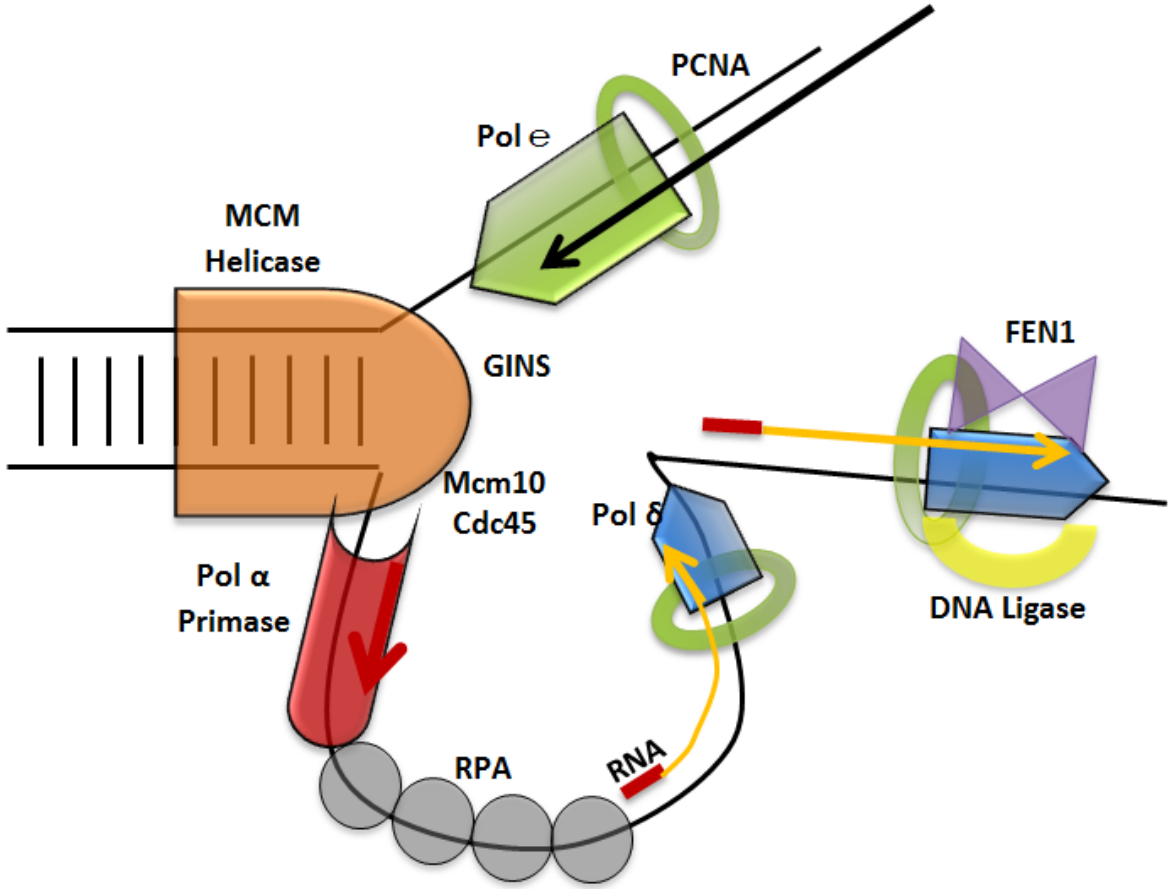


Figure 1.3: DNA replication fork. MCM helicase unwinds the DNA at the replication fork and creates two ssDNA which is immediately coated by RPA to protect it from nucleases and unwanted re-annealing. Pol-α can lay down a RNA primer required for DNA synthesis. It is replaced by Pol-ε and Pol-δ on leading and lagging strand respectively by the action of RFC and PCNA. Okazaki fragments are joined by the collective action of DNA Pol δ, FEN1, and DNA Ligase I.

1.1.5. DNA Damage Repair

Selective mutation is necessary for evolution and adaptation of species. On the other hand, unwanted mutations can cause diseases such as cancer. Our genome is always under attack from environmental stresses, such as ultraviolet (UV) radiation, chemical carcinogens, certain viruses, reactive oxygen species, and internal DNA processing errors. To maintain the integrity of our genome and

safeguard the information encoded within, it is crucial that our cells have efficient replication machinery, as well as an efficient and accurate repair system. DNA damage can take place in many ways. Damage to the nucleotide or sugar base is common and repaired by Nucleotide Excision Repair (NER) or Base Excision Repair (BER) respectively. Double-Strand Breaks (DSBs) are much more deleterious to the cells and are repaired primarily via either non-homologous end joining or homologous recombination (Chapman et al, 2012). DSB are induced by the cell itself to carry out certain events such as chromosomal crossing-over during meiosis, V(D)J recombination in immunoglobulin (Ig) genes, and mating-type (MAT) locus switching in yeast. Double-strand breaks can also be caused by collapsed replication fork, UV radiation, and carcinogens, and if left unrepaired, this can and does lead to chromosomal rearrangements and cell death (apoptosis). Furthermore, certain loci in our genome are inherently more error prone than others. For example, the probability of mutation in NF1 gene (Neurofibromatosis Type I) is higher in the general population than it is in HTT gene (Huntington's disease) just by random mutation (Crowe, F. W., et al. 1956). In cancer cells, kinases including EGFR, ABL, MET, FLT3, and KIT also show mutation at a higher frequency especially in their activation loop compared to normal cells (Dixit et al, 2009). Similarly, tumor suppressor gene p53 also contains hotspots for UV-induced DNA lesions, in which resulting mutations lead to the formation of various types of skin cancer (Dumaz et al, 1997). To understand how cells detect DNA damage and subsequently activate the repair process, we need to understand more about how the cell cycle progresses and also about the various checkpoints.

1.1.6. The Cell Cycle and Checkpoints

The cell cycle is primarily divided into four phases, Gap 1 (G1), Synthesis (S), Gap2 (G2), and Mitosis (M), and is a highly regulated and synchronized process. Cells are required to successfully replicate their entire genome and then divide one copy of that information into each daughter cell. This task is completed in different phases of the cell cycle, with DNA synthesis occurring in S-phase and chromosome segregation occurring during M-phase. The G1 and G2 phases separate the S and M phase, allowing the cell to grow to a critical mass and prepare for S and M phase. In fact, three scientists, Leland Hartwell, Paul Nurse, and Tim Hunt were awarded the 2001 Nobel Prize in Physiology or Medicine for discovering key regulators of the cell cycle. Hartwell was able to identify *Cdc28* in yeast as the gene

required for progression through G1-phase and hence also called ‘START’. This finding was later expanded by Nurse when he discovered *cdc2* gene in *S. pombe* (CDK1 in humans), which was identical to the *Cdc28* gene that Hartwell found earlier, but was necessary for transition from G2 to M-phase. These findings showed that CDC genes control all aspects of the cell cycle progression in yeast. It should be noted that unlike Cdc28/Cdc2 in yeast, humans have numerous CDK proteins that control various aspects of the cell cycle. About a decade later, Hunt discovered cyclins, which are created and degraded periodically in the cell; and binds CDKs to regulate their function via targeted phosphorylation.

These phases of the cell cycle are ‘guarded’ by checkpoints to ensure the proper order and timing of the cell cycle transition events and to make sure the cells have completed the requirements for moving on to the next phase. The G1/S checkpoint takes place towards the end of the G1 phase and can delay the onset of S phase depending on lack of cell size and nutritional requirements. If cell division is deemed unfavorable, a cell can enter the G0 phase (resting phase). G1/S is a critical checkpoint since the cell is committed to completing the cell cycle after this point onwards, until the next G1/S checkpoint is reached (Hartwell, Culotti et al. 1974). Activation of intra-S phase checkpoint upon DNA damage can stall ongoing replication, stabilize the replisome at stalled forks, and initiate late firing of ORs. Since DNA replication occurs during S-phase, replication stress can generate fragile, unstable DNA structures (*e.g.*, fork-stalling, DSBs), triggering the activation of the S-phase checkpoint. The damage is sensed by sensor proteins, which in turn relay the signal to adaptor proteins and ultimately activate effecter proteins to generate the necessary response. Rad24 (Rad17 in humans and *S. Pombe*) is one of the damage sensor proteins involved in checkpoint signaling in *S. cerevisiae*. We isolated Rad24 in our two-hybrid assay as an interacting partner of RPA, and the details of this interaction and its implication in checkpoint signaling will be discussed in chapter 2. Activation of the G2/M checkpoint can arrest cells at G2/M transition and delay the onset of mitosis (Metaphase to Anaphase) thereby preventing chromosome segregation. After cells have attained the desired size, the Anaphase Promoting Complex (APC) can now degrade securin via cyclin B and promote the separation of sister chromatids. The mother cell can now divide into two daughter cells and start new round of the cell cycle.

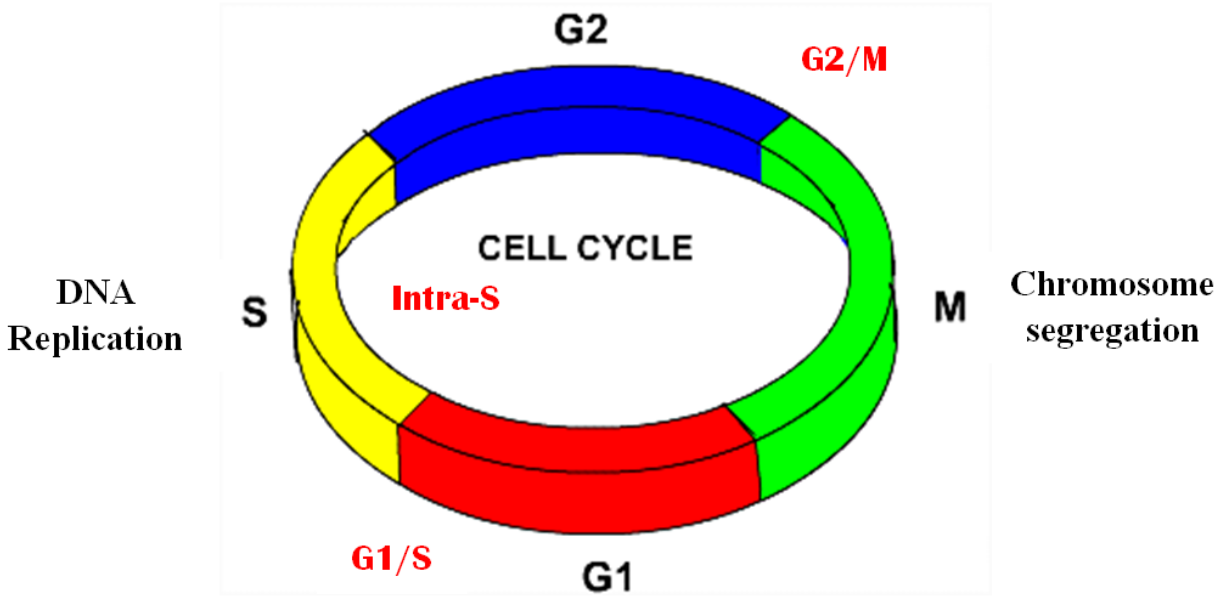


Figure 1.4: The mitotic cell cycle. The pre replication complex forms in the G1 phase cell has to pass through the G1/S checkpoint to commit for the cycle. DNA is replicated during the S phase and intra-S phase checkpoint signals for repair of any damaged DNA. G2/M checkpoint ensures the replication was completed accurately and the cell has reached the critical mass. During the M phase, cells undergo cell division producing two identical daughter cells from one mother cell.

1.1.7. Nucleotide Excision Repair

Nucleotide Excision Repair is useful in correcting DNA damage caused by UV radiation. Ultraviolet-induced damage can produce bulky DNA adducts such as cyclobutane pyrimidine dimers (CPDs) and 6,4-photoproducts (6,4PPs), which create a distorted helical structure and can stall ongoing replication by posing barrier to DNA polymerase. Defects in genes involved in NER pathway are linked to diseases, such as Xeroderma Pigmentosum and Cockayne syndrome and also linked to increased susceptibility for skin cancer (Chu & Mayne, 1996). Nucleotide Excision Repair is carried out via two different pathways: the Global Genome NER (GG-NER) and Transcription-Coupled NER (TC-NER). These pathways recognize DNA damage by separate mechanisms but ultimately converge to fix DNA.

GG-NER uses DNA damage binding protein DDB1, DDB2, and XPC-Rad23B complex to scan through the entire genome to find distorted structures while TC-NER uses stalled RNA Polymerase II as a damage recognition signal at actively transcribed genes (Nouspikel, 2009). Xeroderma Pigmentosum (XPA) is required to verify this DNA damage in the open conformation and recruit the rest of the NER proteins at the site. Transcription factor TFIIH can separate the DNA strands at the damage site XBP and

XPD helicase. The newly formed ssDNA is bound by RPA, which can then recruit XPG and XPF-ERCC1 endonucleases required for removing the DNA strand around the lesion. Once the gap is about 24-32 nucleotides long, DNA Polymerase fills in the proper nucleotides, and the remaining nick is sealed by DNA ligase (Hess et al, 1997). XPF-ERCC1 mutant that is deficient in interacting with RPA has been shown to be defective in NER in Chinese Hamster Ovary (CHO) cells, because of its inability to properly localize in the nucleus (Fisher et al, 2011).

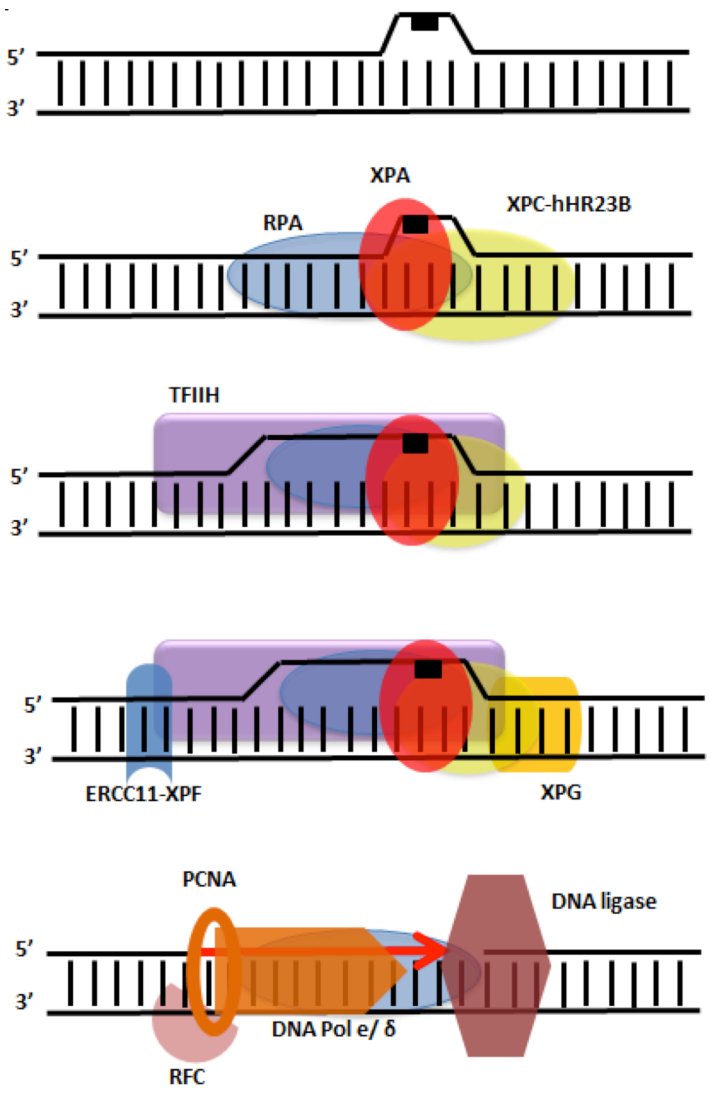


Figure 1.5: GG-NER pathway to resolve photo adducts. DNA damage is sensed by XPC-Rad23B along with DDB1 and DDB2. XPA can verify the damage and recruit the TFIIH and RPA to the site. Once the DNA around the lesion is opened, ERCC1-XPF and XPG can cleave the ends of the damaged strand. This gap can now be filled by the normal replication machinery and sealed by DNA ligase.

1.1.8. Base Excision Repair

Base Excision Repair is an important DNA damage repair pathway to remove small non-helix distorting single base lesions. DNA bases can be altered via oxidation, alkylation, and deamination which can affect the ability to base pair with the proper nucleotide. This can result in events like nucleotide switching, uracil incorporation into the DNA, or formation of apurinic/apyrimidinic (AP) sites that can all lead to DNA mutation. Base Excision Repair (BER) is initiated by an enzyme called N-glycosylase which is necessary for recognition of the damaged base and cleavage of the N-glycosidic bond at the 3' deoxyribose site. Once the base is removed, AP endonuclease can cleave both 3' and 5' phosphodiester bond at the AP site (Haukanes et al, 1988). After this step, the single-strand break can be repaired via either the long patch BER or short patch BER pathway. RPA is shown to be involved in long patch BER pathway, where the damaged base is excised and three additional enzymes FEN1, DNA polymerase, and DNA ligase come in to perform strand incision, gap filling, and ligation respectively (Robertson et al, 2009). The preference between these two pathways is based on factors such as the cell cycle stage and type of DNA lesion (Fortini & Dogliotti, 2007). Similarly, uracil is detected and removed by Uracil DNA Glycosylase (UNG), which has been also been shown to interact with RPA via its Rpa2 C-terminal WHD domain (Nagelhus et al, 1997).

1.1.9. Non-Homologous End Joining

Higher eukaryotes use Non Homologous End Joining (NHEJ) as the predominant pathway to repair dsDNA breaks. In contrast, *S. cerevisiae* utilizes Homologous Recombination (HR) to repair DSB. Following a break, the broken DNA ends are recognized by a heterodimeric protein KU 70/80 (Walker et al, 2001). Next, DNA-Dependent Protein Kinase catalytic subunit (DNA-PKcs) is recruited to the site of damage which is then auto-phosphorylated and recruits DNA ligase IV-XRCC4-XLF complex to the damage site (Ding et al, 2003). This complex can now ligate the broken ends of DNA in a sequence independent manner (Ahnesorg et al, 2006). It may be necessary to process the DNA ends creating a short 1-6 base pair (bp) region of micro-homology, before DNA ligase IV can act on it (de Jager et al, 2001). Although NHEJ has an inherent tendency to be mutagenic, it is necessary in G1 phase because there is no homologous DNA sequence present for HR to occur.

1.1.10. Homologous Recombination

Homologous Recombination utilizes sequence similarity between two DNA molecules to exchange and repair broken DNA. It is also important in meiosis for carrying out crossing-over and in V(D)J recombination during immunoglobin formation. Homologous Recombination is the predominant pathway to repair DSBs in *S. cerevisiae*, and it also has other repair functions in higher eukaryotes. Homologous Recombination (HR) is carried out by a series of steps involving end resection, homology scanning, D-loop formation, strand synthesis, and Holliday junction (HJ) resolution controlled by the Rad52 epistasis group of proteins. These steps are carried out by multiple proteins via a hand-off mechanism. The DSB is first detected by the Mre11-Rad50-Xrs2 (MRX) complex, which recruits the Tel1 kinase. This prevents NHEJ by displacing Ku complex at the DSB site. Sae2 undergoes Mec1 and Tel1 dependent phosphorylation (Baroni et al, 2004) which prompts it to initiate the 5' end resection, removing a 50-100 bp region by its exonuclease activity. Tel1 kinase can phosphorylate Mre11 and Xrs2 subunits of the MRX complex, which can then dissociate from the DNA, and additional exonucleases come in to further resect the 3' ssDNA tail as explained in Figure 8. The newly formed ssDNA is then bound by RPA. This ssDNA bound RPA can recruit Rad17/Mec3/Ddc1 complex via Rad24-RFC to initiate the checkpoint signal, which can activate Rad52 and Chk1 kinase ultimately leading to halt the cell cycle (Majka et al, 2006). Rad52 assists Rad51 displacement of RPA to initiate homology search and subsequent strand invasion (Song & Sung, 2000). RPA-depleted cells fail to form Rad52 foci after DNA damage as it needs ssDNA bound RPA as a substrate to load Rad51. Furthermore, RPA has been shown to interact with both Rad51 and Rad52, making it a central player in HR pathway. The Rfa1-t11 mutant also showed recombination deficiency by its inability to be replaced by Rad51 from ssDNA (Kantake et al, 2003).This creates a D-loop structure where the 3' ends can now be extended by DNA polymerase using the template strand, while RecQ helicase unwinds the template DNA helix. Once the DNA synthesis is complete, the 3' end re-anneals to the original strand, resulting a double HJ. These HJ structures are resolved to restore the broken segment of DNA (Lisby et al, 2004).

S. Cerevisiae	*H. Sapiens*	Cellular Function
	Sensor Proteins	
Mec1-Ddc2	ATR-ATRIP	Checkpoint Kinase
Mre11-Rad50-Xrs2	Mre11-Rad50-Nbs1	Recognizes DSB
Rad24-Rfa2-5	Rad17-Rfc2-5	Clamp loader
Ddc1-Rad17-Mec3 (1-17-3)	Rad9-Rad1-Hus1 (9-1-1)	Checkpoint clamp
Dpb11	TopBP1	Binds and stabilizes Topoisomerase
	Double Strand Break Processing Proteins	
Sae2	CtIP	Endonuclease
Exo1	EXO1	5'→3' exonuclease and flap-endonuclease
Sgs1	BLM	RecQ DNA helicase
Dna2	DNA2	DNA helicase and nuclease
	Adaptor/Mediator Proteins	
Rad9	53BP1, BRCA1, MDC1	Activates Chk1 and Rad53
Mrc1	Claspin	Stabilizes Pol2p at replication fork
	Effector Proteins	
Rad53	Chk2	Cell cycle arrest
Chk1	Chk1	Cell cycle arrest by phosphorylation of Psd1
	Homologous Recombination Proteins	
Rad52	Rad52	Facilitates Rad51 binding to ssDNA
Rad51	Rad51	Promotes strand exchange

Table 1.2: Homologs of proteins involved in DSB recognition and checkpoint signal transduction in yeast and humans.

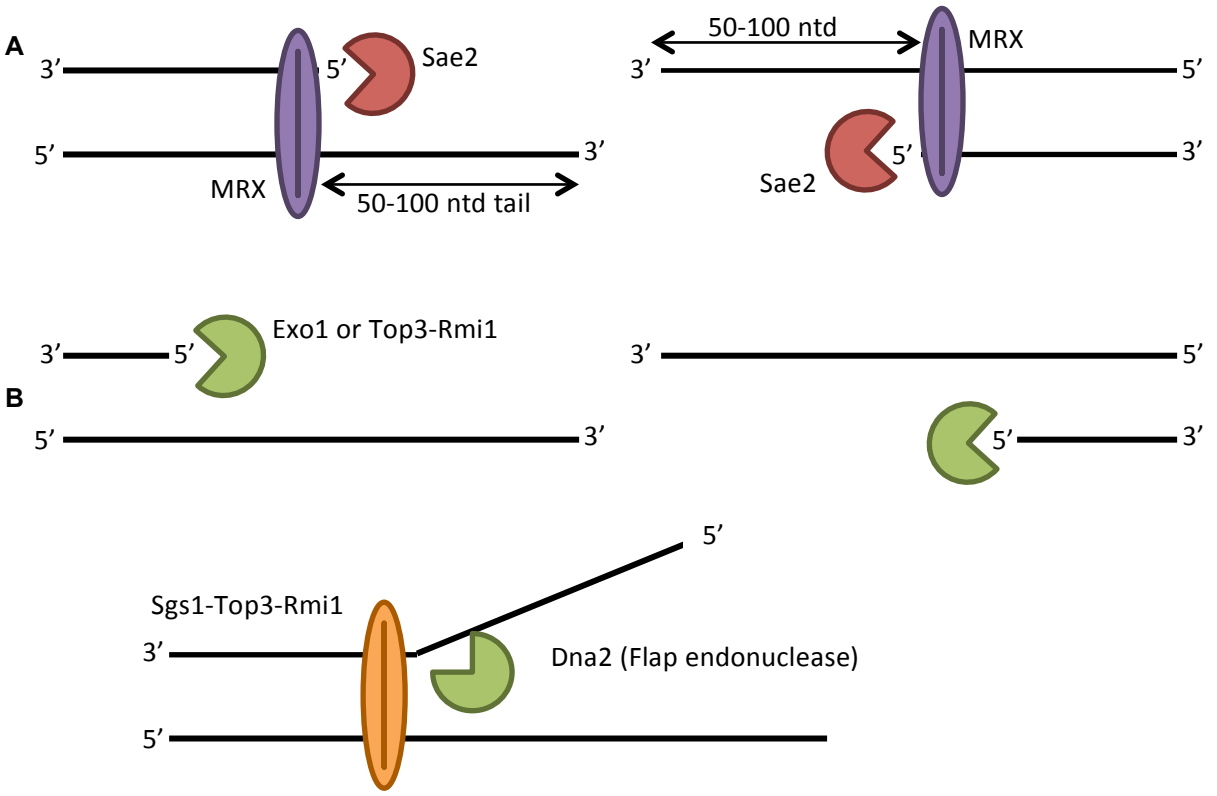


Figure 1.6: End resection after DSB formation. A) MRX complex binds at the break spot and Sae2 resects a small section from the 5'end. B) End resection is carried out by either Exo1/Top3-Rmi1 or Sgs1-Top3-Rmi1 helicase coupled with Dna2 (Finn et al, 2012).

1.1.11. Known RPA-Protein Interactions

Replication Protein A takes part in many aspects of DNA metabolism by its ability to interact with different proteins. It interacts with proteins like ATM (Plug et al, 1997), ATR/ATRIP (Zou & Elledge, 2003), Rad9-Rad1-Hus1 (Wu et al, 2005), p53 (Abramova et al, 1997), and Rad17 (Zou et al, 2003) to mediate the checkpoint response. It also interacts with DNA Pol-alpha (Dornreiter et al, 1992), PCNA (Bylund & Burgers, 2005), RFC (Kim & Brill, 2001), and STAT3 (Kim et al, 2000) to mediate DNA replication. Interaction between RPA and XPA (Vasquez et al, 2002), XPF/ERCC1 (Fisher et al, 2011), UNG (Nagelhus et al, 1997), and MRN (Oakley et al, 2009) has been shown to play important roles in various DNA repair pathways. Finally, interaction between RPA and Rad51 (Stauffer & Chazin, 2004), Rad52 (Park et al, 1996), and DMC1(Golub et al, 1998) play important roles in homologous recombination during

mitosis and meiosis. Phosphorylation of RPA has been assumed to be a major factor in determining its interaction with all these proteins hence dictating its role in so many different processes. Learning about these interactions is important because they can help us identity the mechanism by which these processes take place and also potentially help us find drug targets for controlling and attacking diseases like cancer which hijack various aspects of these processes for proliferation and evading cell death.

Replication Protein A is involved with some major players in the checkpoint pathway. In humans, ATM and ATR-ATRIP are among the first kinases that localize at the site of DNA damage mediated via the MRN complex. ATM and ATR after being recruited to the DNA damage sites can phosphorylate various downstream substrates to activate response elements such as Chk1 and Rad53. Another protein complex containing Rad9-Rad1-Hus1 and Rad17/RFC2-5 also localizes at the DNA damage site to initiate checkpoint signaling and this is mediated by the presence of RPA bound to ssDNA (Majka et al). The 9-1-1 complex can translocate along the damaged DNA and also leads to downstream activation of downstream kinases like chk1 to initiate the checkpoint. Studies in human cells show that ATR forms foci with RPA after DNA damage and stalled replication forks. Furthermore, in the same study RNAi mediated knock down of Rpa70 lead to corresponding decrease of ATR nuclear foci in cells treated with aphidicolin (Dart et al, 2004). This is evidence of RPA as a necessary factor for ATR recruitment to site of DNA damage. Another study demonstrated that ATR is recruited by RPA to site of DNA damage by its interaction with another protein ATRIP and this event was important for ATR mediated phosphorylation of Rad17 and activation of Chk1 activation (Zou et al, 2003). In yeasts a mutation in Rpa1 (L45E) known to be defective in checkpoint activation is unable to recruit Ddc2 (ATRIP in humans) on to ssDNA suggesting that RPA-Ddc2 interaction is mediated via the DBD-F of Rpa70. Very recently, interesting studies have shown that damage dependent hyper-phosphorylation of Rpa32 is important in RPA-p53 interaction. Simultaneous phosphorylation of Rpa32 and p53 on S37 and S46 by DNA-PK and ATM/ATR respectively after DNA damage allows the RPA-p53 complex to dissociate freeing both proteins. This also presents preliminary data on cooperative interplay between HR and NHEJ pathways mediated by regulation of specific protein-protein interaction by action of various Kinases (Serrano et al).

DNA replication is a complex process involving many proteins. Two distinct regions on RPA have been shown to interact with DNA Pol-alpha. The N-terminal region (1-170 aa) of Rpa70 is shown to be

important for stimulating polymers activity of DNA Pol-aplha, while region (170-327) increases its processivity (Braun et al). Loading of PCNA onto primed DNA is done by the RFC complex in an ATP dependent manner and is mediated via interaction between RFC and RPA. Using Surface Plasmon Resonance, binding of PCNA and RFC is mapped to amino acid 36-55 and 196-215 region of PCNA (Zhang et al). RFC binds to 3' end of DNA and opens the PCNA ring using ATP. The ring is closed following ATP hydrolysis and PCNA can firmly clamp to DNA. In chapter 3, we will discuss interaction between Rad24-RFC and RPA which is a complex similar to RFC but important in DNA damage. These findings show the importance of RPA in choreographing various proteins by RPA in order to carry out DNA replication.

DNA damage repair after both ssDNA and dsDNA breaks is mediated via protein interactions between RPA and other repair proteins. ssDNA breaks are repaired via three distinct mechanism involving NER, BER, and MMR. In NER, RPA is important in the gap-filling reaction by working alongside PCNA and DNA Pol-delta to repair damaged nucleotides (Aboussekhra et al). Junction cutting activity of nucleases XPF-ERCC1 and XPG involved in NER is also stimulated by their interaction with RPA (Matsunaga et al, 1996). Xeroderma Pigmentosum group A (XPA) which is a core part of the incision complex of the NER system also interacts with RPA. Studies have shown that RPA via Rpa70 and Rpa32 interacts with XPA on residues 153-176 and 1-58 respectively and this interaction is necessary for XPA to bind damaged DNA with greater affinity (Li et al). In BER, RPA is important for coordinating the long-patch repair pathway by interacting with Uracil-DNA glycosylase (DeMott et al). This discovery was made after the identification of their interaction using the yeast two-hybrid assay where the N-terminal region of human UNG2 was shown to physically interact with the C-terminal region of Rpa32 (Nagelhus et al). In MMR, the template strand is protected by RPA from nuclease degradation and it also promotes MMR complex assembly and excision of mismatch DNA template (Ramilo et al). Unphosphorylated RPA was shown to be important for the excision phase of MMR while phosphorylation of RPA seems to facilitate DNA gap filling by DNA-Pol-delta (Guo et al). This is a good example of how RPA phosphorylation regulates its role in different DNA metabolic processes.

Homologous Recombination (HR) is important for dsDNA break repair in mitosis as well as recombination in meiosis. Two homologs of *E.coli* recombinase RecA, Rad51 and Dmc1 mediate

homologous recombination in mitosis and meiosis respectively. Interestingly, RPA has been shown to interact with both Rad51 and Dmc1 using co-immunoprecipitation. Specifically Rpa70 residue 169-326 is important for interacting with N-terminal domain of Rad51 (Golub et al). Using mutant Rad51 defective in binding RPA, it was shown that RPA is not properly displaced from ssDNA (Stauffer & Chazin, 2004).This demonstrates the necessity for physical interaction between RPA and Rad51 to displace RPA from ssDNA by Rad51. Rpa1-t11 mutants fail to carry out recombination because of their inability to interact with Rad52. Since Rad52 is important for recruiting Rad51 on to ssDNA coated with RPA this defect can be explained (Sugiyama & Kantake, 2009). These experiments all point to the importance of RPA-protein interaction in mediating homologous recombination.

1.2. Thesis Objectives

Replication Protein A has been known to take part in numerous DNA metabolic pathways. It does so by its ability to physically interact with DNA and with proteins necessary for these pathways. The Rpa2 undergoes post-translational modification by various kinases during different stages of the cell cycle, which make it suitable for protein interactions or signaling to carry out its different functions. Identification of proteins (novel or known) that interact with RPA will potentially identify more pathways that RPA plays an active role in and/or allow for further understanding of RPA function in different pathways through further characterization of currently known RPA-protein interactions. Keeping this goal in mind, for the first phase of my thesis work we used the yeast two-hybrid assay to identify proteins that could interact with a subunit of RPA. Using this method, we identified several known proteins, as well as several novel proteins that interact with RPA. Further characterization of these interactions will allow one to understand how RPA participates in a specific function in the cell. Therefore, the second phase of my thesis work focused on characterizing the interaction between RPA and Rad24 to understand the significance of this interaction in DNA damage checkpoint. We identified the regions of Rpa1 and Rad24 that are important for their interaction, and these regions suggest a hand-off mechanism for RPA recruitment of Rad24-Rfc2-5 to sites of DNA damage.

1.3. References

Abramova NA, Russell J, Botchan M, Li R (1997) Interaction between replication protein A and p53 is disrupted after UV damage in a DNA repair-dependent manner. *Proc Natl Acad Sci U S A* **94:** 7186-7191

Baroni E, Viscardi V, Cartagena-Lirola H, Lucchini G, Longhese MP (2004) The functions of budding yeast Sae2 in the DNA damage response require Mec1- and Tel1-dependent phosphorylation. *Mol Cell Biol* **24:** 4151-4165

Bell SP, Stillman B (1992) ATP-dependent recognition of eukaryotic origins of DNA replication by a multiprotein complex. *Nature* **357:** 128-134

Binz SK, Sheehan AM, Wold MS (2004) Replication protein A phosphorylation and the cellular response to DNA damage. *DNA Repair (Amst)* **3:** 1015-1024

Bochkarev A, Bochkareva E, Frappier L, Edwards AM (1999) The crystal structure of the complex of replication protein A subunits RPA32 and RPA14 reveals a mechanism for single-stranded DNA binding. *EMBO J* **18:** 4498-4504

Bochkareva E, Korolev S, Lees-Miller SP, Bochkarev A (2002) Structure of the RPA trimerization core and its role in the multistep DNA-binding mechanism of RPA. *EMBO J* **21:** 1855-1863

Bouton AH, Stirling VB, Smith MM (1987) Analysis of DNA sequences homologous with the ARS core consensus in Saccharomyces cerevisiae. *Yeast* **3:** 107-115

Bylund GO, Burgers PM (2005) Replication protein A-directed unloading of PCNA by the Ctf18 cohesion establishment complex. *Mol Cell Biol* **25:** 5445-5455

Cai L, Roginskaya M, Qu Y, Yang Z, Xu Y, Zou Y (2007) Structural characterization of human RPA sequential binding to single-stranded DNA using ssDNA as a molecular ruler. *Biochemistry* **46:** 8226-8233

Carty MP, Zernik-Kobak M, McGrath S, Dixon K (1994) UV light-induced DNA synthesis arrest in HeLa cells is associated with changes in phosphorylation of human single-stranded DNA-binding protein. *EMBO J* **13:** 2114-2123

Chapman JR, Taylor MR, Boulton SJ (2012) Playing the end game: DNA double-strand break repair pathway choice. *Mol Cell* **47:** 497-510

Chu G, Mayne L (1996) Xeroderma pigmentosum, Cockayne syndrome and trichothiodystrophy: do the genes explain the diseases? *Trends Genet* **12:** 187-192

de Jager M, van Noort J, van Gent DC, Dekker C, Kanaar R, Wyman C (2001) Human Rad50/Mre11 is a flexible complex that can tether DNA ends. *Mol Cell* **8:** 1129-1135

Din S, Brill SJ, Fairman MP, Stillman B (1990) Cell-cycle-regulated phosphorylation of DNA replication factor A from human and yeast cells. *Genes Dev* **4:** 968-977

Ding Q, Reddy YV, Wang W, Woods T, Douglas P, Ramsden DA, Lees-Miller SP, Meek K (2003) Autophosphorylation of the catalytic subunit of the DNA-dependent protein kinase is required for efficient end processing during DNA double-strand break repair. *Mol Cell Biol* **23:** 5836-5848

Dixit A, Yi L, Gowthaman R, Torkamani A, Schork NJ, Verkhivker GM (2009) Sequence and structure signatures of cancer mutation hotspots in protein kinases. *PLoS One* **4:** e7485

Dornreiter I, Erdile LF, Gilbert IU, von Winkler D, Kelly TJ, Fanning E (1992) Interaction of DNA polymerase alpha-primase with cellular replication protein A and SV40 T antigen. *EMBO J* **11:** 769-776

Dumaz N, Duthu A, Ehrhart JC, Drougard C, Appella E, Anderson CW, May P, Sarasin A, Daya-Grosjean L (1997) Prolonged p53 protein accumulation in trichothiodystrophy fibroblasts dependent on unrepaired pyrimidine dimers on the transcribed strands of cellular genes. *Mol Carcinog* **20:** 340-347

Dutta A, Stillman B (1992) cdc2 family kinases phosphorylate a human cell DNA replication factor, RPA, and activate DNA replication. *EMBO J* **11:** 2189-2199

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

Fisher LA, Bessho M, Wakasugi M, Matsunaga T, Bessho T (2011) Role of interaction of XPF with RPA in nucleotide excision repair. *J Mol Biol* **413:** 337-346

Fortini P, Dogliotti E (2007) Base damage and single-strand break repair: mechanisms and functional significance of short- and long-patch repair subpathways. *DNA Repair (Amst)* **6:** 398-409

Gambus A, Jones RC, Sanchez-Diaz A, Kanemaki M, van Deursen F, Edmondson RD, Labib K (2006) GINS maintains association of Cdc45 with MCM in replisome progression complexes at eukaryotic DNA replication forks. *Nat Cell Biol* **8:** 358-366

Golub EI, Gupta RC, Haaf T, Wold MS, Radding CM (1998) Interaction of human rad51 recombination protein with single-stranded DNA binding protein, RPA. *Nucleic Acids Res* **26:** 5388-5393

Haring SJ, Humphreys TD, Wold MS (2010) A naturally occurring human RPA subunit homolog does not support DNA replication or cell-cycle progression. *Nucleic Acids Res* **38:** 846-858

Haukanes BI, Helland DE, Kleppe K (1988) Analysis of cleavage products of DNA repair enzymes and other nucleases. Characterization of an apurinic/apyrimidinic specific endonuclease from mouse plasmacytoma cells. *Nucleic Acids Res* **16:** 6871-6882

Henricksen LA, Umbricht CB, Wold MS (1994) Recombinant replication protein A: expression, complex formation, and functional characterization. *J Biol Chem* **269:** 11121-11132

Hess MT, Schwitter U, Petretta M, Giese B, Naegeli H (1997) Bipartite substrate discrimination by human nucleotide excision repair. *Proc Natl Acad Sci U S A* **94:** 6664-6669

Kantake N, Sugiyama T, Kolodner RD, Kowalczykowski SC (2003) The recombination-deficient mutant RPA (rfa1-t11) is displaced slowly from single-stranded DNA by Rad51 protein. *J Biol Chem* **278:** 23410-23417

Karnani N, Taylor CM, Malhotra A, Dutta A (2010) Genomic study of replication initiation in human chromosomes reveals the influence of transcription regulation and chromatin structure on origin selection. *Mol Biol Cell* **21:** 393-404

Kemp MG, Mason AC, Carreira A, Reardon JT, Haring SJ, Borgstahl GE, Kowalczykowski SC, Sancar A, Wold MS (2010) An alternative form of replication protein a expressed in normal human tissues supports DNA repair. *J Biol Chem* **285:** 4788-4797

Keshav KF, Chen C, Dutta A (1995) Rpa4, a homolog of the 34-kilodalton subunit of the replication protein A complex. *Mol Cell Biol* **15:** 3119-3128

Kim HS, Brill SJ (2001) Rfc4 interacts with Rpa1 and is required for both DNA replication and DNA damage checkpoints in Saccharomyces cerevisiae. *Mol Cell Biol* **21:** 3725-3737

Kim J, Kim D, Chung J (2000) Replication protein a 32 kDa subunit (RPA p32) binds the SH2 domain of STAT3 and regulates its transcriptional activity. *Cell Biol Int* **24:** 467-473

Kunkel TA, Burgers PM (2008) Dividing the workload at a eukaryotic replication fork. *Trends Cell Biol* **18:** 521-527

Lao Y, Gomes XV, Ren Y, Taylor JS, Wold MS (2000) Replication protein A interactions with DNA. III. Molecular basis of recognition of damaged DNA. *Biochemistry* **39:** 850-859

Li L, Lu X, Peterson CA, Legerski RJ (1995) An interaction between the DNA repair factor XPA and replication protein A appears essential for nucleotide excision repair. *Mol Cell Biol* **15:** 5396-5402

Liaw H, Lee D, Myung K (2011) DNA-PK-dependent RPA2 hyperphosphorylation facilitates DNA repair and suppresses sister chromatid exchange. *PLoS One* **6:** e21424

Lisby M, Barlow JH, Burgess RC, Rothstein R (2004) Choreography of the DNA damage response: spatiotemporal relationships among checkpoint and repair proteins. *Cell* **118:** 699-713

Liu VF, Weaver DT (1993) The ionizing radiation-induced replication protein A phosphorylation response differs between ataxia telangiectasia and normal human cells. *Mol Cell Biol* **13:** 7222-7231

Liu Y, Kvaratskhelia M, Hess S, Qu Y, Zou Y (2005) Modulation of replication protein A function by its hyperphosphorylation-induced conformational change involving DNA binding domain B. *J Biol Chem* **280:** 32775-32783

Majka J, Binz SK, Wold MS, Burgers PM (2006) Replication protein A directs loading of the DNA damage checkpoint clamp to 5'-DNA junctions. *J Biol Chem* **281:** 27855-27861

Mason AC, Haring SJ, Pryor JM, Staloch CA, Gan TF, Wold MS (2009) An alternative form of replication protein a prevents viral replication in vitro. *J Biol Chem* **284:** 5324-5331

Mer G, Bochkarev A, Gupta R, Bochkareva E, Frappier L, Ingles CJ, Edwards AM, Chazin WJ (2000) Structural basis for the recognition of DNA repair proteins UNG2, XPA, and RAD52 by replication factor RPA. *Cell* **103:** 449-456

Murzin AG (1993) OB(oligonucleotide/oligosaccharide binding)-fold: common structural and functional solution for non-homologous sequences. *EMBO J* **12:** 861-867

Nagelhus TA, Haug T, Singh KK, Keshav KF, Skorpen F, Otterlei M, Bharati S, Lindmo T, Benichou S, Benarous R, Krokan HE (1997) A sequence in the N-terminal region of human uracil-DNA glycosylase with homology to XPA interacts with the C-terminal part of the 34-kDa subunit of replication protein A. *J Biol Chem* **272:** 6561-6566

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

Nouspikel T (2009) DNA repair in mammalian cells : Nucleotide excision repair: variations on versatility. *Cell Mol Life Sci* **66:** 994-1009

Nuss JE, Patrick SM, Oakley GG, Alter GM, Robison JG, Dixon K, Turchi JJ (2005) DNA damage induced hyperphosphorylation of replication protein A. 1. Identification of novel sites of phosphorylation in response to DNA damage. *Biochemistry* **44:** 8428-8437

Oakley GG, Tillison K, Opiyo SA, Glanzer JG, Horn JM, Patrick SM (2009) Physical interaction between replication protein A (RPA) and MRN: involvement of RPA2 phosphorylation and the N-terminus of RPA1. *Biochemistry* **48:** 7473-7481

Park MS, Ludwig DL, Stigger E, Lee SH (1996) Physical interaction between human RAD52 and RPA is required for homologous recombination in mammalian cells. *J Biol Chem* **271:** 18996-19000

Plug AW, Peters AH, Xu Y, Keegan KS, Hoekstra MF, Baltimore D, de Boer P, Ashley T (1997) ATM and RPA in meiotic chromosome synapsis and recombination. *Nat Genet* **17:** 457-461

Prakash A, Natarajan A, Marky LA, Ouellette MM, Borgstahl GE (2011) Identification of the DNA-Binding Domains of Human Replication Protein A That Recognize G-Quadruplex DNA. *J Nucleic Acids* **2011:** 896947

Robertson AB, Klungland A, Rognes T, Leiros I (2009) DNA repair in mammalian cells: Base excision repair: the long and short of it. *Cell Mol Life Sci* **66:** 981-993

Salas TR, Petruseva I, Lavrik O, Saintomé C (2009) Evidence for direct contact between the RPA3 subunit of the human replication protein A and single-stranded DNA. *Nucleic Acids Res* **37:** 38-46

Segurado M, de Luis A, Antequera F (2003) Genome-wide distribution of DNA replication origins at A+T-rich islands in Schizosaccharomyces pombe. *EMBO Rep* **4:** 1048-1053

Shao RG, Cao CX, Zhang H, Kohn KW, Wold MS, Pommier Y (1999) Replication-mediated DNA damage by camptothecin induces phosphorylation of RPA by DNA-dependent protein kinase and dissociates RPA:DNA-PK complexes. *EMBO J* **18:** 1397-1406

Song B, Sung P (2000) Functional interactions among yeast Rad51 recombinase, Rad52 mediator, and replication protein A in DNA strand exchange. *J Biol Chem* **275:** 15895-15904

Speck C, Chen Z, Li H, Stillman B (2005) ATPase-dependent cooperative binding of ORC and Cdc6 to origin DNA. *Nat Struct Mol Biol* **12:** 965-971

Stauffer ME, Chazin WJ (2004) Physical interaction between replication protein A and Rad51 promotes exchange on single-stranded DNA. *J Biol Chem* **279:** 25638-25645

Tanaka T, Nasmyth K (1998) Association of RPA with chromosomal replication origins requires an Mcm protein, and is regulated by Rad53, and cyclin- and Dbf4-dependent kinases. *EMBO J* **17:** 5182-5191

Umezu K, Sugawara N, Chen C, Haber JE, Kolodner RD (1998) Genetic analysis of yeast RPA1 reveals its multiple functions in DNA metabolism. *Genetics* **148:** 989-1005

Vasquez KM, Christensen J, Li L, Finch RA, Glazer PM (2002) Human XPA and RPA DNA repair proteins participate in specific recognition of triplex-induced helical distortions. *Proc Natl Acad Sci U S A* **99:** 5848-5853

Walker JR, Corpina RA, Goldberg J (2001) Structure of the Ku heterodimer bound to DNA and its implications for double-strand break repair. *Nature* **412:** 607-614

Wang H, Guan J, Perrault AR, Wang Y, Iliakis G (2001) Replication protein A2 phosphorylation after DNA damage by the coordinated action of ataxia telangiectasia-mutated and DNA-dependent protein kinase. *Cancer Res* **61:** 8554-8563

Wold MS (1997) Replication protein A: a heterotrimeric, single-stranded DNA-binding protein required for eukaryotic DNA metabolism. *Annu Rev Biochem* **66:** 61-92

Wold MS, Kelly T (1988) Purification and characterization of replication protein A, a cellular protein required for in vitro replication of simian virus 40 DNA. *Proc Natl Acad Sci U S A* **85:** 2523-2527

Wu X, Shell SM, Zou Y (2005) Interaction and colocalization of Rad9/Rad1/Hus1 checkpoint complex with replication protein A in human cells. *Oncogene* **24:** 4728-4735

Xu X, Vaithiyalingam S, Glick GG, Mordes DA, Chazin WJ, Cortez D (2008) The basic cleft of RPA70N binds multiple checkpoint proteins, including RAD9, to regulate ATR signaling. *Mol Cell Biol* **28:** 7345-7353

Zou L, Elledge SJ (2003) Sensing DNA damage through ATRIP recognition of RPA-ssDNA complexes. *Science* **300:** 1542-1548

Zou L, Liu D, Elledge SJ (2003) Replication protein A-mediated recruitment and activation of Rad17 complexes. *Proc Natl Acad Sci U S A* **100:** 13827-13832

CHAPTER 2: IDENTIFICATION OF NOVEL PROTEINS INTERACTING WITH RFA1-FAB USING THE YEAST TWO HYBRID ASSAY AND THE EFFECT OF RFA2 PHOSPHORYLATION STATE ON THESE INTERACTIONS

2.1. Abstract

Replication Protein A is a heterotrimeric protein important in DNA replication, repair, and recombination. It is the primary ssDNA binding protein in eukaryotes being expressed in abundance throughout the cell cycle. The formation of ssDNA is an intermediate in multiple DNA metabolic pathways, summoning RPA to these sites. ssDNA bound RPA can then interact with various proteins to carry out specific pathways that it is involved in. Learning about these interactions and their mechanism will help us learn a great deal about how different DNA metabolic pathways are initiated and if these processes are regulated by phosphorylation state of RFA. We performed a global assay to identify and understand novel RFA-protein interactions and their importance in specific pathways. Using the yeast two-hybrid assay, we were able to isolate and identify 15 proteins: 3 known, 2 known but uncharacterized, and 10 novel proteins, most of which either have implications in DNA metabolism or at least are normally localized in the nucleus. Furthermore, we determined that the phosphorylation state of Rfa2 affects the interaction between Rfa1-FAB and a subset of these proteins. Lastly, we identified another subset of these interactions that are disrupted when treated with DNA damaging agent MMS, indicating that post-translational modification on either Rfa1-FAB and/or the target protein can abrogate their interaction. Since these proteins are involved in a broad range of DNA processing pathways as mentioned in chapter 1, further work will be necessary to elucidate the significance of each of these interactions, which can give us insights about RFAs role in the whole DNA metabolic network.

2.2. Introduction

Protein-protein interactions virtually control every single process in the cell including DNA metabolism, replication, repair, cell migration, signal transduction, scaffolding, enzymatic actions, cell cycle regulation and much more. While some enzymes function individually, most proteins usually function as part of a larger complex or are part of a network of proteins working together to accomplish a cellular task. Such inter-dependence among proteins in a network means that mutation in one protein has the ability to disrupt the entire pathway or the process it is involved in. While this might not always be the

case, there are numerous examples of proteins that are central to a metabolic process and mutation of which leads to cellular dysfunction ultimately leading to disease. For example, in humans BRCA1 is part of a larger protein complex called BRCA1-Associated Genome Surveillance (BASC) complex and mutation in this gene is associated with patients having predisposition to breast cancer (Wang et al, 2000). BRCA1 is expressed in breast tissue and helps repair DNA damage by interacting with DNA damage sensors, tumor suppressors, and checkpoint signaling molecules (Caestecker & Van de Walle, 2013). Similar to BRCA1 many proteins function in parallel ways, helping us understand the importance a single gene and protein-protein interaction in a larger cellular network.

The first goal of my thesis was to identify novel proteins that could interact with RFA *in vivo* and investigate if those interactions were dependent on the phosphorylation state of Rfa2. Indeed, a prior study showed that the N-terminal region of Rfa1 is important for RFA-MRN interaction and phosphomimetic form of RFA2 weakens RFA-MRN interaction implying the importance of the phosphorylation state of Rfa2 in regulating RFA-protein interactions (Oakley et al, 2009). The two-hybrid screen can help us identify similar protein interactions with RFA and their dependency on the phosphorylation state of Rfa2. It will help us understand RFA's role in pathways that are unknown and also characterize interactions that are known but the significance of which are not yet clear. We wanted to capture these protein interactions *in vivo* and isolate the gene (or gene fragment) encoding for the interacting protein. The isolated gene (or gene fragment) could then be sequenced and analyzed. We chose to perform a global screen using the yeast two hybrid assay to accomplish this task. The yeast two-hybrid assay is a molecular biology tool used to identify *in vivo* protein-protein interaction (Figure 2.1). It is especially useful to identify protein-protein interactions in their native environment where other cellular factors are present as compared to some of the other techniques such as co-immunoprecipitation, ChIP, and Tag Affinity Purification. Hence we can observe the physiological relevance of such protein interactions using this method.

Introduced by Stanley Fields in 1989, the yeast two-hybrid assay exploited the fact that the yeast transcription factor GAL4 has two distinct domains, the DNA Binding Domain (DBD) and the Transcription Activation Domain (AD). As neither of these domains can activate transcription on their own, they can be separated. The BD can be fused to protein X (often referred to as 'bait'), and the AD can be fused to

protein Y (often referred to as 'prey'). When proteins X and Y interact with each other *in vivo*, the DBD and AD of the Gal4 protein will be brought into proximity once again, restoring Gal4's function as a transcriptional activator. The Gal4 DBD interacts with a specific sequence, called UAS_{GAL}, located upstream of galactose-inducible genes. If one inserts multiple copies of the UAS_{GAL} sequences upstream of a reporter gene, this gene would only be expressed if protein X and Y can interact with each other.

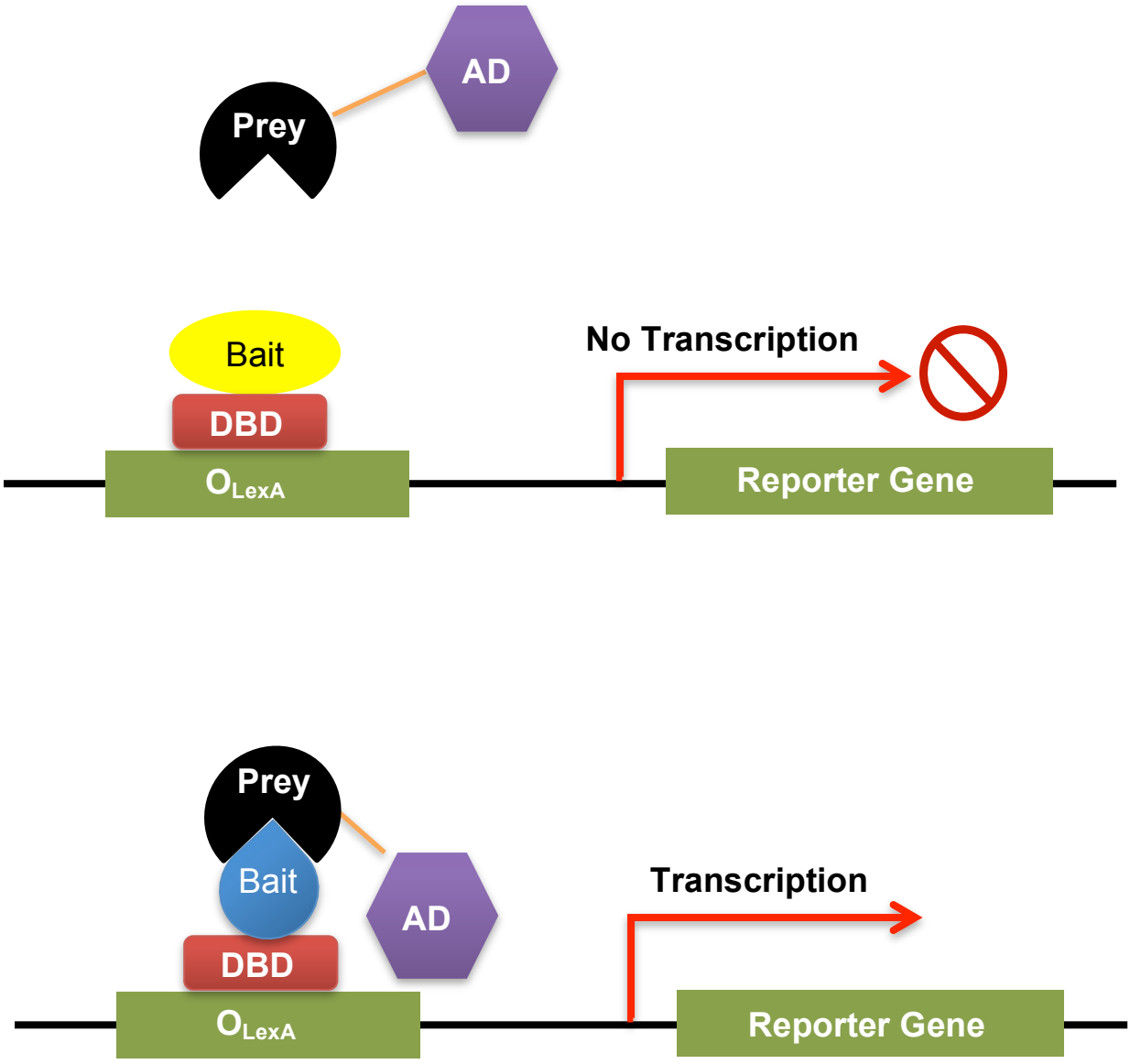


Figure 2.1: Yeast two hybrid assay. If the bait and prey proteins cannot interact with each other, then the DBD and AD are not in close proximity to initiate transcription of the reporter gene (picture above). On the other hand, if the bait and prey protein are capable of interaction with each other than the DBD and AD are brought in close proximity and the reporter gene can is transcribed. This event can be detected by growing yeast cells on selective media which only allows cells expressing the reporter gene to grow. The prey plasmid can be isolated from these yeast colonies and sequenced to identify the gene fragment cloned in it. Hence, we can determine two proteins interacting *in vivo*.

Finally, one can isolate yeast colonies that support the interaction by growing them in a selective media, isolating the plasmids from the yeast, and sequencing them to identify the cloned gene fragment (Fields & Sternglanz, 1994). The biggest drawback of this assay is the possibility of getting false-positive results (*i.e.*, growth on selective media in the absence of interaction between protein X and Y). This can occur because protein X fused with the DBD might itself contain a transcriptional activation domain, resulting in its recruitment upstream of the reporter gene, thereby eliminating the need to interact with another protein containing an AD. Another reason for getting false-positive results is due to the fact that the interactions are occurring *in vivo* (in the yeast cell with potentially thousands of other proteins). It is possible that cellular proteins (not cloned into the prey plasmid) could interact with protein X bait, resulting in transcription of the reporter gene without specific interaction with the intended prey. Hence, before doing this screen, it is crucial to test the bait plasmid for any auto-activation activity by doing the assay with empty prey plasmid to monitor for unwanted cell growth on selective media. Another thing to be careful while performing the yeast two hybrid assay is the potential for false negatives. There might be legitimate protein-protein interaction between the bait and prey which may not result in growth on the selective media. In this case we will never discover these protein interactions to begin with. Hence yeast two hybrid is a powerful tool but also has its fair share of drawbacks. Since its inception in 1989, different variations of this system has been created allowing researchers to detect *in vivo* protein interactions with higher precision while minimizing false-positive interactions. Since GAL4 is a eukaryotic transcriptional activator native to yeast it could produce large number of false-positive interactions, making it very difficult to screen real candidates. To minimize pitfalls of the traditional GAL4 system, we used the DupLEX-A yeast two-hybrid system (OriGene), which uses prokaryotic DBD of LexA protein (recognizes the O_{lexA} sequence) and transcription activation domain from prokaryotic B42 protein. Hence we can detect targeted expression of proteins based on their fusion with either LexA-DBD or B42-AD proteins. The B42-AD domain was cloned under the control of a constitutive GAL1 promoter which makes sure the gene is only expressed when galactose is present as a primary carbon source. This is crucial to our experiment as we would only be able to see interaction in the presence of galactose media which provides us yet another checkpoint to eliminate false-positive results.

Finding the suitable bait plasmid turned out to be a rather challenging task for this assay. It has long been known that Rfa2 is hyper-phosphorylated after DNA damage. Our hypothesis was that Rfa2 hyper-phosphorylation after DNA damage causes the RFA complex to interact or dissociate with different sets of proteins than it would during the normal cell cycle. This hyper-phosphorylated Rfa2 could act as a signal to recruit proteins necessary for signaling cell cycle arrest and for initiating the checkpoint. So to test this hypothesis, we wanted to use Rfa2-WT along with the Rfa2 phospho-mimetic mutants: Rfa2-Asp, Rfa2-Ala, and Rfa2-ΔN plasmid (referred to as Rfa2-X plasmid) as our bait. To make the Rfa2-X mutant plasmids all the Serine and Threonine in the N-terminal of Rfa2 were substituted with either Aspartic acid, Alanine, or the domain was deleted entirely using site directed mutagenesis as shown in Figure 2.2. Rfa2-WT would have been the ideal candidate for this study because it would help us identify RFA-protein interactions that were directly mediated via the Rfa2 subunit. Furthermore, we could see if these interactions were directly affected by the phosphorylation state of Rfa2 by testing the interactions with Rfa2-X mutant plasmids. Unfortunately, all of these Rfa2-X plasmids caused auto-activation, leaving us to find alternative options for the bait plasmid. In the process of creating Rfa1-WT as our bait plasmid; we inadvertently ended up cloning Rfa1-FAB fragment in our bait due to a nonsense mutation during the PRC reaction which we ended up using for our two-hybrid screen.

Along with isolating proteins that interact with Rfa1 *in vivo*, we were also interested in understanding if these interactions were dependent on the phosphorylation state of Rfa2. So once we identified our potential genes from the two-hybrid screen using RFA1-FAB as bait, we retested those interactions in a mutant EGY48 strains (referred to as EGY48-Rfa2-X strains), which had Rfa2-WT mutated to Rfa2-Asp, Rfa2-Ala, and Rfa2-ΔN in their genome. We utilized the two-step gene replacement method to remove the WT copy of Rfa2 and inserted the mutant form of Rfa2 into its location in the genome (Figure 2.3). This would make sure only Rfa2 mutants were present in the cells at any given time and help us understand how this setup would affect the Rfa1-FAB interaction with its interacting partners. We were successful in identifying a subset of interactions between Rfa1-FAB and the 15 HITS isolated from our initial screen that are dependent on the phosphorylation state of Rfa2. It was a surprising finding since Rfa1-FAB cannot form complex with Rfa2/Rfa3 as it is missing DBD-C necessary for forming the trimerization core of RFA complex but nevertheless its interaction with proteins was affected by the

phosphorylation state of Rfa2. These observations will be discussed further in the results section of this chapter.

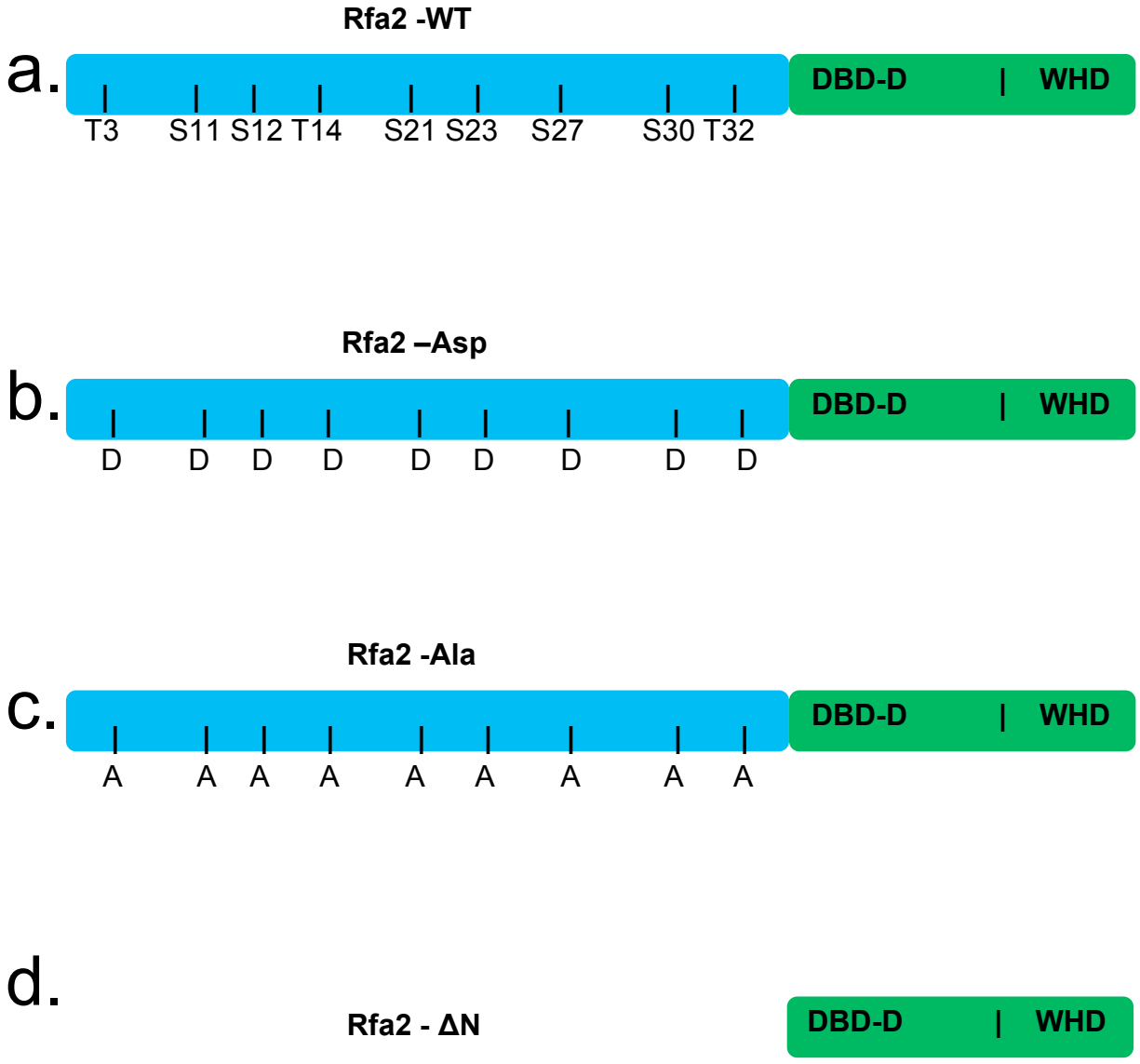


Figure 2.2: Schematic representation of Rfa2-X mutants: a) Rfa2-WT protein is shown with all the Serine and Threonine in the N-terminal Phosphorylation Domain (PD). b) All the Serine and Threonine residues were changed to Aspartic acid to mimic phosphorylation. c) All the Serine and Threonine residues were changed to Alanine to mimic e unphosphorylated state. d) The N-terminal of Rfa2 (1-38 aa) was deleted completely.

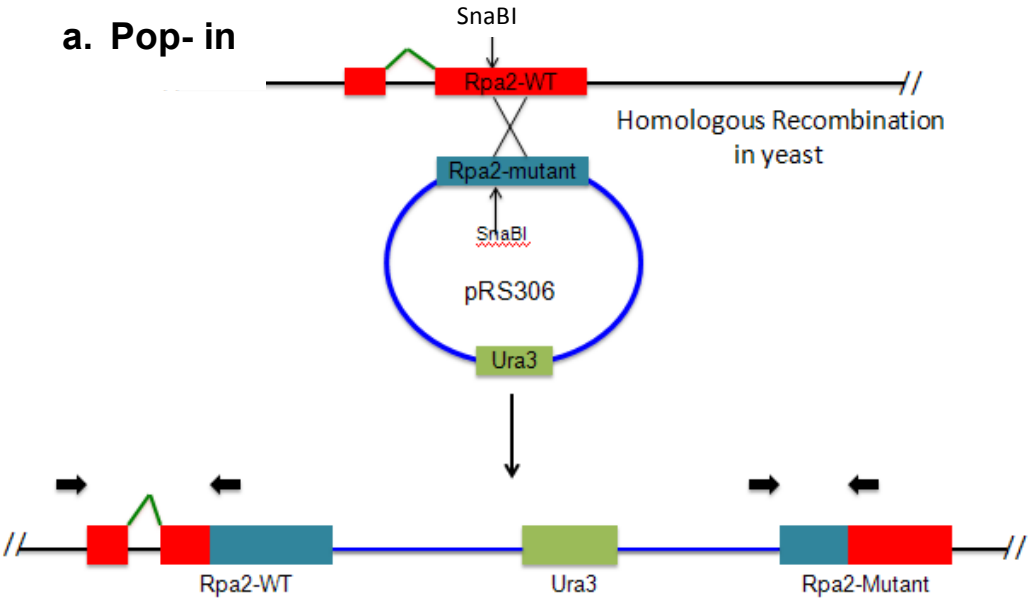


Figure 2.3: Two-step gene replacement in yeast using homologous recombination. pRS306 was digested with SnaBI and transformed into EGY48. A faction of the yeast cells will incorporate linearized pRS306 into their genome using homologous recombination as they have sequence homology. The transformed cells were selected on plates lacking URA and random 'pop-in' events were discovered by using PCR directly to the isolated genomic DNA. Presence of two bands about 100bp different in length indicated successful pop-in candidates. The reason for two separate bands produced by Rpa2 is because the WT copy has an extra 100 bp intron compared to the plasmid copy.

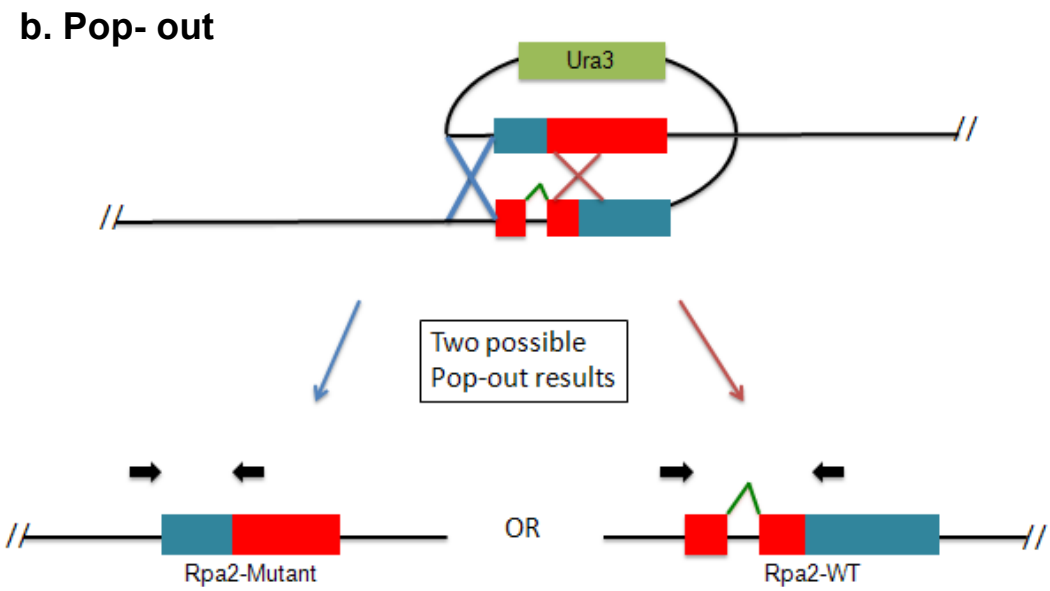


Figure 2.3: Two-step gene replacement in yeast using homologous recombination (continued). The successful 'pop-in' candidates were then plated then on 5-FOA to select for candidates excluded the pRS306 plasmid. Since the recombination event can happen on either side of the Rpa2 gene before the plasmid was excluded, we rescreened them using PCR similarly as done earlier. This time we were looking for candidates that only had the gene from pRS306 indicated by the presence of only one band running lower than the WT copy.

Along with isolating proteins that interact with Rfa1 *in vivo*, we were also interested in understanding if these interactions were dependent on the phosphorylation state of Rfa2. So once we identified our potential genes from the two-hybrid screen using RFA1-FAB as bait, we retested those interactions in a mutant EGY48 strains (referred to as EGY48-Rfa2-X strains), which had Rfa2-WT mutated to Rfa2-Asp, Rfa2-Ala, and Rfa2-ΔN in their genome. We utilized the two-step gene replacement method to remove the WT copy of Rfa2 and inserted the mutant form of Rfa2 into its location in the genome (Figure 2.3). This would make sure only Rfa2 mutants were present in the cells at any given time and help us understand how this setup would affect the Rfa1-FAB interaction with its interacting partners. We were successful in identifying a subset of interactions between Rfa1-FAB and the 15 HITS isolated from our initial screen that are dependent on the phosphorylation state of Rfa2. It was a surprising finding since Rfa1-FAB cannot form complex with Rfa2/Rfa3 as it is missing DBD-C necessary for forming the trimerization core of RFA complex but nevertheless its interaction with proteins was affected by the phosphorylation state of Rfa2. These observations will be discussed further in the results section of this chapter.

2.3. Methods and Materials

2.3.1. Plasmids and Yeast Strains

The pEG202 (bait), pJG4-5 (prey), and pSH18-34 (reporter) vectors were provided with the DupLEX-A Yeast Two-Hybrid System (OriGene) (Table 2.1). Yeast genomic library cloned into pJG4-5 was also purchased from OriGene. To clone the Rfa2-X into bait plasmid, PCR was first used to amplify Rfa2-WT, Rfa2-Asp, Rfa2-Ala, and Rfa2-ΔN with common reverse primer O-271 and four individual forward primers O-270, O-310, O-311, and O-312 respectively (Table 2.2). These PCR products were cloned into linearized pEG202 (with NcoI) and transformed into EGY48 yeast strains using *in vivo* cloning by homologous recombination. Transformed colonies containing potential pEG202-Rfa2-X clones were selected by growing them on SD-His plates at 30°C for 2-3 days. These colonies were further grown in SD-His media overnight in 3 mL culture tubes and used for isolating Yeast Genomic DNA. This genomic DNA was transformed into DH10B bacterial cells using electroporation. Plasmid from bacteria was isolated using miniprep and digested with appropriate restriction enzyme to identify successful clones and yeast colonies.

All yeast strains used in my experiments were derived from EGY48 strains. EGY48-Asp, EGY48-Ala, and EGY48-ΔN mutants were generated using two-step gene replacement method to replace the WT copy of Rfa2 with Rfa2-Asp, Rfa2-Ala, or Rfa2-ΔN. Table 2.3 lists all yeast strains with their genotypes.

Plasmid	Description
pEG202	HIS3, 2 μ, AmpR, (constitutive ADH promoter expresses LexA and is followed by a polylinker for making the bait fusion protein)
pJG4-5	TRP1, 2 μ, AmpR, (inducible GAL1 promoter expresses B42-HA tag and is followed by a polylinker for making target fusion protein expression libraries from cDNA);
pSH18-34	URA3, 2 μ, AmpR, 8 ops.-LacZ (high sensitivity)
pRS306	URA3, 2 μ, AmpR

Table 2.1: Plasmids used in the yeast two hybrid assay and their properties

Primer ID	Primer Name	Primer Sequence (5'→3')
O-270	pEG202-Rfa2-FOR	TATTCGCAACGGCGACTGGCTGGAATTCCCGGG GATCCTTATGGCAACCTATCAACCATA
O-271	pEG202-Rfa2-REV	AAATTCGCCCGGAATTAGCTTGGCTGCAGGTCG ACTCGAGTCATAGGGCAAAGAAGTTAT
O-310	pEG202-Rfa2-Asp-FOR	TATTCGCAACGGCGACTGGCTGGAATTCCCGGG GATCCTTATGGCAGACTATCAACCATA
O-311	pEG202-Rfa2-Ala-FOR	TATTCGCAACGGCGACTGGCTGGAATTCCCGGG GATCCTTATGGCAGCTTATCAACCATA
O-312	pEG202-Rfa2-N38-FOR	TATTCGCAACGGCGACTGGCTGGAATTCCCGGG GATCCTTATGGCACCTGTGACGATCAA

Table 2.2: Primer ID, name, and sequence used to make Rfa2-X bait plasmids

Strain	Genotype/Description
EGY48	MATα *trp1 his3 ura3 leu2*::6 LexAop-*LEU2* (high sensitivity)
EGY48-Rpa2-Asp	MATα *trp1 his3 ura3 leu2*::6 LexAop-*LEU2* (high sensitivity), (Rpa2-WT →Rpa2-Asp)
EGY48-Rpa2-Ala	MATα *trp1 his3 ura3 leu2*::6 LexAop-*LEU2* (high sensitivity), (Rpa2-WT →Rpa2-Ala)
EGY48-Rpa2-ΔN	MATα *trp1 his3 ura3 leu2*::6 LexAop-*LEU2* (high sensitivity), (Rpa2-WT → Rpa2-ΔN)
EGY188	MATa trp1 his3 ura3 leu2::2 LexAop-LEU2 (low sensitivity)
EGY194	MATa trp1 his3 ura3 leu2::4 LexAop-LEU2 (medium sensitivity)

Table 2.3: Yeast strains used in the two hybrid and their genotypes

2.3.2. Yeast and Bacterial Media

Yeast media was prepared in 500 mL batches by adding 10 gm of Dextrose or Galactose (referred as SD or SG respectively), 2.5 gm of Ammonium Sulfate, 0.85 gm of Yeast N2 Base w/o aa, and appropriate drop out amino acid in 500 mL distilled water. The solution was autoclaved like this for liquid media and autoclaved by adding 10 gm Agar to make solid media (plates). Similarly, MMS containing plates were made by adding 120 uL of MMS solution after the media was cooled to room temperature after autoclaving.

Bacterial Luria Broth (LB) media was prepared in 500 mL batches by adding 10 gm Tryptone, 2.5 gm Yeast Extract, 5 gm NaCl in 500 mL distilled water. The solution was autoclaved like this for liquid media and autoclaved by adding 10 gm Agar to make solid media (plates). In either case, appropriate antibiotic was added to final concentration of 100 microgram/ mL after the media was autoclaved to make the desired selective media (e.g. LB+AMP, LB+KAN).

2.3.3. Polymerase Chain Reaction Protocol

The PCR reaction was set up by adding 50 – 500 ng of plasmid to 5 uL 10 X PCR Buffer, 25 mM MgCl2, 5 uM Primer-FOR, 5 uM, 10 mM dNTPs, 5 U/uL Taq DNA Polymerase, and bringing the total volume to 50 uL by adding ddH_2O. After all the components were properly added, the reaction tube was placed in a thermo-cycler and PCR was allowed to proceed under following reaction condition:

# Cycles	1	35			1	
Temp	94°C	94°C	55°C	72°C	72°C	4°C
Time	3 min	1 min	1 min	1 min/kbp	2. min	∞

2.3.4. *In vivo* Cloning by Homologous Recombination

100 uL EGY48 competent cells was transformed with 325 uL 50% PEG, 40 uL 1M LiAc, 25 uL 2 mg/mL whale sperm (carrier) DNA, ~1 ug of linearized plasmid and 25 uL of PCR products. The mixture was incubated at 30°C for 30 min and additional 42°C for 15 min. 250 uL was plated onto selective media and grown for 2-3 days at 30°C. Yeast cells are capable of integrating the linearized plasmid with our PCR fragments (with flanking regions homologous to the cloning site on the linearized plasmid) by their ability to fix dsDNA breaks using homologous recombination. Once the transformation was completed,

yeast colonies that successfully retained functional copies of these plasmids could grow on the selective media. These plasmids were isolated from the yeast colonies using yeast genomic DNA isolation protocol and transformed into bacteria (DH10B) cells using electroporation. Once the plasmid was duplicated by bacteria in abundance, it was harvested using the miniprep protocol and digested with appropriate restriction enzymes to verify the insertion of the correct PCR product.

2.3.5. Isolation of Yeast Genomic DNA

Yeast cells were grown in appropriate media at 30°C overnight while shaking at 220 RPM. Cells were spun at 4,000 RPM for 10 min and resuspended in 0.2 mL yeast lysis buffer [2% Triton X-100, 1% SDS, 100 nM NaCl, TE Buffer (pH 8.0)]. The resuspended cells were added to 2 mL centrifuge tubes containing 0.2 mL (425-600 um) acid-washed glass beads. Then 0.3 mL Phenol: Chloroform: Isoamyl alcohol (P:C:I) was added to the cell/ bead mixture making sure the tubes were closed tightly to prevent leaks in the subsequent steps. The tube was then vigorously mixed using a vortex at max speed for 10-15 minutes. After this step was completed, we centrifuged the tube at 14,000 RPM for 10 minutes and the aqueous (top) layer containing DNA/RNA was carefully removed and collected in a new 1.5 mL microfuge tube. We added 1 mL cold 95% EtOH to the tubes and vortex the solution for 10 second before placing it in -20°C for 20 minutes to allow for DNA precipitation. Once the while cloudy precipitate was visible, we centrifuged the tube at 14,000 RPM for 15 minutes at 4°C and removed the supernatant using a vacuum aspirator. The pellet was washed again by adding 1 mL 70% EtOH and centrifuged at 14,000 RPM for 15 min at 4°C. The supernatant was removed using vacuum aspirator once again and the pellet was air dried using a vacuum centrifuge for 30-60 minutes making sure to leave the microfuge cap open. Finally the DNA pellet was suspended in 100 uL ddH_20 and stored in 4°C for future use.

2.3.6. Bacterial Transformation and Plasmid Isolation

Bacterial transformation was done using electroporation. Frozen DH10B electro-competent cells stored at -75°C were thawed on ice for 20 minutes along with electroporation cuvettes. 1 uL of desired plasmid was then mixed with 40 uL of thawed DH10B cells and thoroughly mixed by pipetting 8-10 times before being transferred to the chilled cuvettes. The mixture was chilled on ice for additional 5 minutes allowing the contents to settle at the bottom. Then the cuvette was placed in the BTX ECM 600 electroporation system and adjusted to the following setting: 2.5 kV, 200 Ohms, and 25 uF. Once the

system was ready, the capacitors were charged using the pulse button which last about 1 second. When the capacitors discharge, a beep is heard indicating the completion of the process (In certain occasions, a blue spark may be observed at this step in which case the procedure needs to be repeated in a separate cuvette). Next, the cuvette was filled with 1 mL LB and gently mixed by pipetting up and down and transferred to a new 1.5 ml microfuge tube. This tube was incubated at 37°C for 1 hour and 250 uL of this solution was then plated on to LB+AMP plates and incubated at 37°C for 16 - 24 hours. Once the colonies were formed on the plates, they were further grown overnight in LB+AMP liquid culture at 37°C and used to isolate plasmid DNA using the Omega Bio-tek's HiBind E.Z.N.A. Plasmid Miniprep kit as per the manufactures guidelines.

2.3.7. Restriction Digest and Gel Electrophoresis

Once the mini-prep was completed, plasmids were digested for verification of the clones. 7 uL of plasmid was added to 1 uL 10 X BSA, 1 uL NEB 3 Buffer, 0.5 uL BamHI, and 0.5 uL of XhoI bringing the total volume to 10 uL. The mixture was incubated at 37°C for 2 hours and loaded on to medium sized 1% agrose (1 gm in 100 mL 1 X TAE solution). Gel electrophoresis was done at 85 V for 2 hours followed by staining with 10 ug/mL EtBr solution with gentle rocking for 20 minutes and distaining with water and visualized under UV light.

2.3.8. Two-Step Gene Replacement

pRS306 plasmid with Rfa2-Asp, Rfa2-Ala, and Rfa2-ΔN were linearized using *Sna*BI and transformed intoEGY48 cells. Transformants were grown on SD-Ura plates to select for colonies with the particular Rfa2-X mutant integrated into their genome at the endogenous Rfa2 locus. Colonies were also replica plated onto YPG to test for possible mitochondrial defects. Yeast genomic DNA was isolated from these candidates to verify the integration of Rfa2 mutant genes using PCR with primers RFA2-UP-NEW [5'-TAGCAATTCCTTTGGCCTCGATGAGCTTCC-3'] and RFA2-DOWN-NEW [5'-GATAAAACCCTGGTCAG TCAAGGTCGTAC-3']. Candidates with newly integrated Rfa2 mutant 'pop-ins' will have two Rfa2 genes that are roughly 100 bp different in length due to the presence of a 100 bp intron in the genomic Rfa2. Hence using primers to amplify the Rfa2 gene we can identify candidates with 'pop-ins' as they will have two separate bands on a agrose gel that differ by 100 bp. Pop-in candidates were grown in 3ml YPD overnight at 30°C to allow for 'pop-out' of the extra copy of Rfa2 and extra vector

sequence. These cells were then plated onto SD plates containing X mg/mL 5-fluororotic acid (5-FOA). Colonies from the 5-FOA plates were grown in YPD and yeast genomic DNA was isolated to test for desired pop-out candidates with PCR as done earlier.

2.3.9. Yeast Two-Hybrid Assay

The two-hybrid assay was performed by following the DupLEX-A yeast-two hybrid application guide provided by OriGene. Rfa1-FAB, Rfa2, Rfa3, Rfa2-Asp, Rfa2-Ala, and Rfa2-ΔN cloned into pEG202 were tested for auto-activation by plating the transformed EGY48 cells onto SG-His-Leu plates. To perform the large scale library transformation, PEG202-Rfa1-FAB plasmid was transformed into EGY48 strain along with the yeast genomic library (cloned into pJG4-5) and the *lacZ* reporter plasmid (pSH18-34) and plated onto 88 SD-His-Trp-Ura plates (in two separate batches of Group 1 and 2 with 40 and 44 plates respectively) followed by incubation at 37°C for 3 days. Yeast colonies from all plates from each group were harvested by adding 3 mL of ddH_2O to each plate and scraping off cells with a sterile spatula and collected in a 50 mL conical tube. The cells were then centrifuged at 1,500 x g for 5 minutes and the cell pellet was washed with 10 mL sterile ddH_2O twice. Lastly, the pellet was mixed with 10 mL ddH_2O and half a volume of sterile 50% glycerol before being stocked frozen in 1 mL aliquots at -75°C. Titering of yeast transformation was performed by making 8 serial dilutions of the stock solution by factor of 10 with the last dilution factor being 1 X 10^8 and 100 uL from each dilution was plated onto SD-His-Trp-Ura plates. These plates were grown at 37°C for 2-3 days and the titer was calculated by using the following formula:

Titer = Number of Colony Forming Units (CFU) / Vol. Plated * Total Dilution Factor

The titers for two Group 1 and Group 2 were around 7.73 x 10^5 and 1.21 x 10^6 cells / uL respectively. Then 100 mL stock solution from each group was diluted 10X with SG-His-Leu-Trp-Ura and 150 uL and 40 uL from Group 1 and Group 2 were added to 3.2 mL SG-His-Leu-Trp-Ura in a 50 mL conical tube. The solution was grown at 30°C for 4 hours and 200 uL of this solution was plated into 17 individual SG-His-Leu-Trp-Ura plates to yield approximately 1.68 x 10^7 CFU in order to cover all potential protein interactions. Plates were grown at 30°C for 4 days before picking master plates. Starting on day four, 9 biggest colonies from each plate were picked onto SD-His-Trp-Ura master plate. More master plates were picked using similar strategy until day seven.

The master plates from SD-His-Trp-Ura were then replica plated onto both SD-His-Leu-Trp-Ura and SG-His-Leu-Trp-Ura and grown at 30°C for 3 days. Colonies that grew on galactose plates but not on glucose plates were considered potential positive transformants and were re-picked onto SD-His-Trp-Ura and maintained for further characterization. Prey plasmid from the first 100 colonies of our SD-His-Trp-Ura master plate was identified and isolated using restriction digest (with EcoRI and XhoI) and miniprep. Since all three plasmids had different lengths and digest pattern with EcoRI and XhoI a backbone of 10.2 kb and 6.4 kb with a smaller fragment cloned in it signified the bait and Prey plasmids respectively. A digest pattern with four bands was characteristic of the pSH18-34 (reporter) plasmid. Figure 2.4 illustrates an example of this is selection process. All prey plasmids isolated with approach were sent to Eton Bioscience Inc for sequencing using sequencing primer pJG4-5-UP-Sequence [5'-GATCCAGCCTGACTGGCTGAAATCGAATGG-3'].

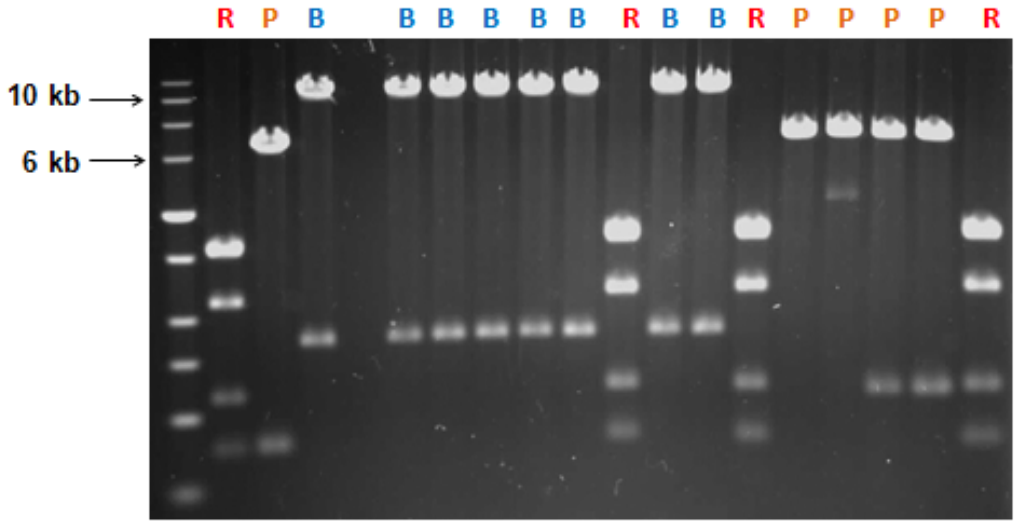


Figure 2.4: Example of banding pattern for pJG4-5, pEG202, and pSH18-34. Plasmids [(pJG4-5 (Prey, P), pEG202 (Bait, B), and pSH18-34 (Reporter, R)] isolated from yeast colonies were digested with EcoRI and XhoI. The prey, bait, and reporter plasmid all produced unique banding pattern. pJG4-5 has a vector backbone of about 6 kb with a random size fragment cloned in it. pEG202 has a vector backbone of about 10 kb with Rfa1-FAB cloned in it which is consistent throughout the gel. pRS18-34 digest is gives 4 separate bands. Since each plasmid has a unique digest pattern, we can identify them and eventually

2.3.10. Survival Assay on MMS

Master plates containing all positive colonies from the initial screen that grew on SG-His-Trp-Ura media were replica plated onto SG-His-Leu-Trp-Ura-MMS plates and grown at 30°C for 2-3 days. Prey

plasmids form colonies that failed to grow on SG-His-Leu-Trp-Ura-MMS were also isolated and sent for sequencing similarly to the first 100 candidates.

2.4. Results

2.4.1. Generating and Testing Auto-Activation of Different RFA Subunit 'Bait' Constructs

Since we ultimately wanted to test RFA-protein interaction dependence on the phosphorylation state of Rfa2, we wanted to use Rfa2 WT and Rfa2 phospho-mimetic mutants (Rfa2-X) as our bait proteins. Our original goal was to isolate proteins from a yeast genomic library using Rfa2, and then test the different phospho-mutants of Rfa2 to determine if the interaction was phosphorylation specific. This would have given us a very clear idea of how RFA interactions would be affected by phosphorylation of Rfa2. Rfa2 cloned into pEG202 (bait vector) were transformed in EGY48 along with empty pJG4-5 (prey vector) and tested for auto-activation by growing them on SD-Leu plates. Unfortunately, we found that Rfa2-WT and every one of Rfa2 phospho-mimetic mutants (Rfa2-Asp, Rfa2-Ala, and Rfa2-ΔN) resulted in growth on plates lacking Leucine in the absence of any interacting protein partner, indicating that these constructs caused auto-activation (Figure 2.5). Therefore, they were unsuitable as bait for performing the two-hybrid screen. In the process of making Rfa2 bait plasmids, I also constructed bait plasmids for Rfa1 (which turned out to be Rfa1-FAB later after sequencing) and Rfa3 subunits. Rfa3 demonstrated auto-activation, whereas Rfa1-FAB did not (see below). Hence, by default only the Rfa1-FAB subunit could be used in the two-hybrid screen as bait.

2.4.2. Prey Candidates Identified by Two-Hybrid Screen

Using the Rfa1-FAB domain in our yeast two-hybrid assay as bait, we isolated numerous gene fragments that could encode meaningful interaction partners of RFA *in vivo*. We originally picked 100 colonies appearing earliest on the screening plates (after 2-3 days) which meant they showed the strongest interactions in our assay and then isolated pJG4-5 (prey plasmids) from them to determine the gene sequence cloned in them. Since each colony contained all three plasmids: pEG202, pJG4-5, and pSH18-34, we isolated the DNA and performed restriction digest to identify only the pJG4-5 prey plasmids before we sent them out for sequencing. Once we received our sequencing results back, we used the Washington University BLAST WU-BLAST2 (http://www.yeastgenome.org/cgi-bin/blast-sgd.pl) software to identify sequence similarity with known yeast genes in the Saccharomyces Genome

Database (SGD) database. Our screen revealed numerous gene fragments corresponding to functional yeast genes, which are listed in Table 2.4.

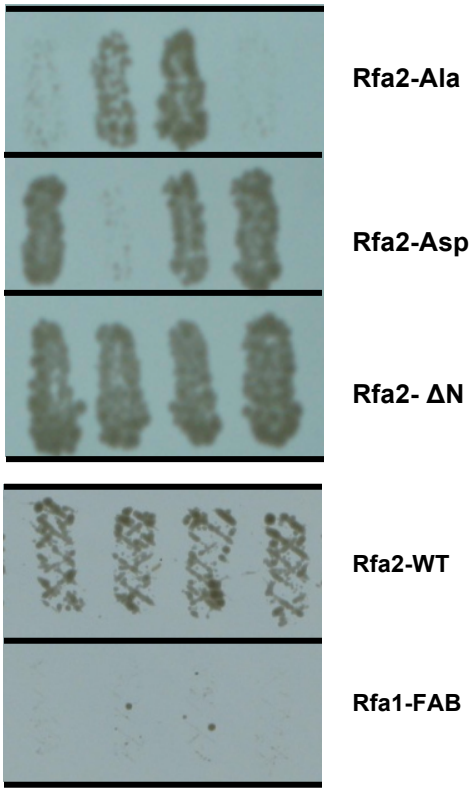


Figure 2.5: Autoactivation test on SD His-Leu plates. Rfa1-FAB does not autoactivate while Rfa2-WT, Rfa2-Asp, Rfa2-Ala, and Rfa2- ΔN all autoactivate. Cells were transformed with linearized pEG202 and PCR amplified Rfa2-WT, Rfa2-Asp, Rfa2-Ala, Rfa2-ΔN, and Rfa1-FAB into EGY48 cells and plated on SD-His plates. Colonies form the plates were picked directly on to SD-His-Leu plates to check for autoactivation. The variation in growth of Rfa2-Ala and Rfa2-Asp colonies is because of the incomplete in vivo cloning. We transformed linearized vector and PCR products in yeast but not all PCR products were incorporated into the pEG202. Some plasmids ligated back together without incorporating a PCR fragment resulting in a functional but empty pEG202 plasmid. To see the fraction of cells that incorporated the PCR product, we just transformed the linearized vector and plated cells as negative control to estimate the re-ligation frequency. Cells with this functional pEG202 can grow just fine on SD-HIS plates but they fail to grow on SD His-Leu media as it is unable to produce Rfa2-X causing auto activation essentially acting as negative control. The colonies that grew were later tested for presence of Rad2-X genes by PCR amplification (results not shown).

Gene	ORF	Frequency	Protein length	Sequence Immediately following EcoRI site (5'→3')
AIM7	YDR063W	1	149	CGGCTAGACGAAATTGAAGATTTGAGCG AATTAGCCGAGATCCTACCTGATAACTC ACCTAGATTTGTAC
DNA2	YHR164C	21	1522	CCCCCCCTTATAGCGTACCAAGGACGCT CCTGCTTCTGAAGCCTAAACTTCGACAG TCATCATTCGCTGC
AXL2	YIL140W	5	823	CCTCTCATTTGCTTCCTAATATTCTGGAG ACGCAGAAGGGAAAATCCAGACGATGA AAACTTACCGCATG
GPM3	YOL056W	1	303	CGCCCCGAAGAGAAGGCTAATGAGCGT CTGCCAGAAAGTGAGTCTCTTTGTGAAG TGGTCGTCAGATTG
SGS1	YMR190C	5	1447	CGCTACATGACTACTAGAGATGAAGAAA AAGAAGAAAACGAATTACTAAATCAAAG CGATTTTGATTTTGT
RFA1	YAR007C	1	621	CAAGTTCCAGTCAATGGAACTACAAAGG GGTGATATCATTCGCGTGATAATTGCAG AACCTGCTATTGTC
EHT1	YBR177C	1	451	CCCCAGATCAACCATACTCAATCGTGGA AAAGAATCCTCGTATTTTGTATTGTAGAA CCGATTTAGGTGG
LDB19	YOR322C	1	818	CGCCCACAAATCCGGGTTGCAGTAAACT CTTTAGAAAATATGCCGTCGCAAAGGCT TCCAGGCGAGCCC
APC4	YDR118W	1	652	CGCCTTGCTATTTACAGAGTCTCGGATC ACGCTAGGCTTGCAGTAATTCCAATAAG AAATATAAACTTGGT
PTC1	YDL006W	3	281	CGCCTGATAATGAAAAGTCGTGTAAATG GTATGCTGGCAGTGACGAGATCGTTAG GGGATAAATTTTTTG
ADH1	YOL086C	1	348	GGCCGACTCGAGAAGCTTTGGACTTCTT CGCCAGAGGTTTGGTCAAGTCTCCAATC AAGGTTGTCGGCT
FAS1	YKL182W	1	2051	CCCGGTGGTGGTTACTTTTCCGCAGCAG GTATGACCGCCGCTATTGATTCTGTGGT TTCTCAGATAGAA
SRP10	YKL154W	1	244	CGCCATGTCAAGTTGCGTTATAAACTCT CAGATTATTTGAAAACAAGAGCCAAATTT GTTAAAGGGTTGA
MAF1	YDR005C	11	395	CGCCCTATAATTTAAATTTGGGTTCTCTT GGTCCCCTAAACCAACCGCCAAGTTGAT CCTAATTTGGTTA

Table 2.4: List of all unique prey candidates isolated from the first 100 colonies

Gene	ORF	Frequency	Protein length	Sequence Immediately following EcoRI site (5'→3')
RAD24	YER173W	18	659	CCTGAAGATTGGTTAAATGTTAGTCTTTA TAAGTACAACGCGGTACATTCTTTCAGG AATATAACTCTAGA
MCM5	YLR274W	1	775	CCTCAATCTGCTAATAACAATGACAAAG ATCCAGAAAATACTAGTATGGATACTGAT TCTCTCTTATTGAA
PAA1	YDR071C	1	191	CGGCTATGTTCAGGTCTTTTCATCAGAG AGATCGAGGGCAAAGAAGTTAAAAAGGA GACACTAATTGGTC
VPS29	YHR012W	1	282	CGGCCAATTGTCTTTGATGTTGAGGATA GCGATGAAGCTGTGACATCAGAGGTGG ACAAGCCGACTAAG
	YDL180W	1	547	CGGCCAGTTGTTAGGTAATACGACGTTA TATTATGGTAGTTTAAATGGGGATGATG ATGATATGACCAAC
MOG1	YJR074W	1	218	CCTTTACAAGAAGACACTGTTCAGCAAG GAACCAAGTTCACCGGACTCGTTATGGA AGTAGCAAATAAGT
MPH1	YIR002C	2	993	CGCTGTACTGTTGACAAGAATAACATGG TATTGAGTTTGGACGATTGGAATTTTTTC AGAAACCGCTATAT
IML1	YJR138W	2	1584	CGCCCGCTTTACAACTTCATTAATGAGC AACAAACATCTTTAGAAAGCTCGGCTAT AAATTTTAAAGATTC
MIF2	YKL089W	2	549	CCTATTTCTGGACAAAAGAAATTAAGCAA CTCATTTAGGACATATATTACGTTCCACG TGATACAGGGAAT
ECM21	YBL101C	2	1117	CGGCTGGAAATTATGCTACGAATCAGTA AACCTGATCCTGAGTGTCCGTCCAAATT AAGACATTATGAGG
SAC6	YDR129C	1	642	CGGCCTAAGAAGGGTAAGGAATTGAATA ATTTTCAGGCTAGTGAAAATGCTAATATT GTTATTAATTCTGC
VMA21	YGR105W	1	77	CCTGCTGCAATGGCCAATGTTGTTCTAA TCGTTTACATTGTTGTAGCGTTCCGCGA GGATACTGAAGATC
	YNR071C	1	342	CCTACCATTGCGCCACTTGGTGCAACTT TGGTAGACCTGAAGGTAAACGGCCAATC AGTCGTTCAAGGTT
UBP1	YDL122W	1	809	CGCTTAAATTTACCGAACGAGAATATTG GTTCCACTTTAAAATTATCTCAGTTATTA AGCGACTGGAGTAA

Table 2.4: List of all unique prey candidates isolated from the first 100 colonies (continued).

Gene	ORF	Frequency	Protein length	Sequence Immediately following EcoRI site (5'→3')
ALR1	YOL130W	1	859	CGGCCTTCCAGACTAGCCCACTCCATGC CACATCAAAGGCAGCTTTATGTGGAAAG TAATATACACACA
GAP1	YKR039W	1	602	CGGCCACAAAGCCAAGATGGTATAGAAT CTGGAATTTCTGGTGTTAATGAATCGAA CAACGACAAAAAAA
PKC1	YBL105C	1	1151	CGCTACGTTACTATAAAAATCGATGATAC GATCAAAGCCAGAACGAAGCCTTCTAGA AATGACAGGTGG

Table 2.4: List of all unique prey candidates isolated from the first 100 colonies (continued). The frequency of occurrence for each unique gene in the screen, total length of the protein product, and sequence immediately downstream of the EcoRI cloning site in the pJG4-5 are listed

2.4.3. RFA Potentially Interacts With 12 Novel Proteins *in vivo* via its Rfa1-FAB Domain

We performed further analysis using SGD and shortlisted genes from our candidate pool that were known to be either important in some DNA metabolic process, were localized in the nucleus, or did not have any previously characterized function. This process helped us narrow the list of proteins that were most likely capable of interacting with RFA. Novel proteins among the shortlisted candidates that showed direct interaction with Rfa1-FAB were Apc4, Maf1, Mif2, Ptc1, Paa1, Pkc1, YNR071C, Ubp1, Pol32, and Yuh1. Proteins that had previously showed indirect association with RFA in other screens and were uncharacterized include Mcm5 and Rad24. Proteins isolated from our screen that were known and fairly well characterized include Dna2 (Bae et al, 2003), Sgs1 (Hegnauer et al, 2012) and Mph1 (Banerjee et al, 2008). A brief summary of these genes is also listed in Table 2.5.

Candidate Gene	Null Mutant Phenotype	Protein Function	Interaction with RPA
Dna2	**Not viable**	DNA nuclease and helicase, required for Okazaki fragment processing, involved in DNA repair	Known on both proteins
Sgs1	Chromosomal instability; viable	RecQ family DNA helicase, plays role in genomic stability	Known on both proteins

Table 2.5: Shortlisted genes isolated from the yeast two hybrid screen using Rfa1-FAB as bait.

Candidate Gene	Null Mutant Phenotype	Protein Function	Interaction with RPA
Mph1	Viable	Helicase involved in 3'→5' error free bypass of DNA lesion	Known on Mph1
Apc4	**Not viable**	E3 ubiquitin ligase, promotes metaphase to anaphase transition by degrading anaphase inhibitors and mitotic cyclins	**NOVEL**
Maf1	Viable	Negative regulator of RNA Pol III, localized in the nucleus during stress conditions	**NOVEL**
Mcm5	**Not viable**	Component of MCM complex, ATP dependent Helicase important for priming of ORC during G1	Co-localization determined; Direct interaction is **NOVEL**
Mif2	**Not viable**	Required for spindle elongation, interacts with H2A, H2B, and H4	**NOVEL**
Ptc1	Viable	Protein phosphotase	**NOVEL**
Paa1	Viable	Acetylates polyamines, may be involved in Replication/ Repair	**NOVEL**
Pkc1	**Not viable**	Protein serine/threonine kinase, important for cell wall remodeling	**NOVEL**
Rad24	Viable	Checkpoint protein, activates checkpoint signaling after damage by loading Rad17-Mec3-Ddc1 onto damage DNA	Indirect interaction determined; Direct interaction is **NOVEL**
YNR071C	Viable	Unknown	**NOVEL**
Ubp1	Viable	Ubiquitin ligase, cleaves polyubiquitin chains at C-terminal	**NOVEL**
Pol32	Viable	Subunit of DNA Pol delta, essential in DNA replication	**NOVEL**
Yuh1	Viable	Ubiquitin C-term hydrolase	**NOVEL**

Table 2.5: Shortlisted genes isolated from the yeast two hybrid screen using Rfa1-FAB as bait (Continued). The null mutant phenotype shows the viability of the gene in knock out yeast. Protein functions of each candidate are listed and their interaction with Rfa1 to date is summarized. Novel and uncharacterized RFA-protein interactions in vivo are Apc4, Maf1, Mcm5, Mif2, Ptc1, Pka1, Paa1, Rad24, YNR071c, Ubp1, Pol32, Yuh1. Previously characterized protein interactions include Sgs1, Dna1, and Mph1.

2.4.4. A Subset of Interactions Between Rfa1-FAB and HITS are Abrogated by Treatment with MMS

Methylmethane sulfonate (MMS) is a DNA alkylating agent that is fatal to cells that are compromised in homologous recombination pathway. It has been long-speculated that MMS causes dsDNA breaks, but recent studies have shown that it might actually cause replication forks to stall by creating base adducts (Groth et al, 2010; Wold, 1997). Nevertheless, MMS treatment causes hyper-phosphorylation of human Rpa2 in the cell (Wold, 1997). To determine if the interaction between Rfa1-FAB and the HITS were disrupted by MMS treatment, we replica plated these yeast colonies on to SG-His-Leu-Trp-Ura plates and incubated at 30°C for 2-3 days. Subsets of these interactions were indeed disrupted by MMS treatment and are summarized in the Table 2.7. Since MMS treatment can potentially phosphorylate a myriad of proteins in the cell, the disruption of interaction between Rfa1-FAB and our HITS could be either due to phosphorylation of Rfa2 or the HITS themselves. Alternatively, but not mutually exclusive, modification of RFA-interacting proteins, or other proteins that make up a complex with these interacting proteins (HITS) might also cause these interactions to be disrupted. Thus, although it is clear that interactions are disrupted or weakened, this assay does not provide a direct explanation of why the disruption occurs. However, as we found multiple candidates that showed interaction-sensitivity when cells were exposed to MMS beyond the first 100 candidates, we isolated prey plasmids from 35 additional candidates and sent them for sequencing analysis. Out of these additional 35 candidates, we isolated 12 unique Rad24 candidates. Since Rad24 is an important checkpoint protein and shows distinct interaction defects with Rfa1 in presence of MMS, we focused on this particular interaction for the next part of my thesis (chapter 3). Our goal was to further map which region on each protein is important for Rfa1-Rad24 interaction.

2.4.5. Testing Reciprocal Interaction Using Prey Candidates as Bait

As the Rfa2 phospho-mimetic plasmids were all causing auto-activation and unusable for the two-hybrid assay, our strategy was to use the prey candidates (HITS) themselves as the bait plasmid and test their interaction with the different Rfa2 phospho-mimetic constructs cloned as prey plasmids. To do this, we amplified all 15 prey candidates with 4-5into202-FOR [5'-AATTCGCCCGGAATTAGCTTGGCTG CAG GTCGACTCGAGAACCTCTGGCGAAGAAGTCC-3'] and 4-5into202-REV [5'-GTTCTCACATCACAT CC GAACATAAACAAAAATGGGTAAGGAAAAGACTCACGTTTCGAG-3'] primers using PCR and

Gene (HITS)	Location on Plate	Growth after treatment with MMS
Dna2	4	No
Sgs1	8	No
Mph1	80	No
Apc4	19	No
Maf1	33	Little Growth
Mcm5	31	No
Mif2	54	Yes
Ptc1	21	Yes
Paa1	35	No
Pkc1	100	Yes
Rad24	4-30	No
YNR071C	72	No
Ubp1	78	No
Pol32	101	No
Yuh1	87	No

Table 2.6: Growth of yeast containing Rpa1-FAB and HITS on SG-His-Leu-Trp-Ura- MMS plates. Previously identified interactions were retested by replica plating them on SG-His-Leu-Trp-Ura-MMS plates to see if cell growth was affected. Interaction between Dna2, Sgs1, Mph1, Apc4, Maf1, Mcm5, Paa1, Rad24, YNR071C, Ubp1, Pol32, and Yuh1 were disrupted in presence of MMS while interaction with Mif2, Ptc1, and Pkc1 was unaffected

transformed it with linearized pEG202 (with NcoI) in EGY48 strains to clone them into pEG202 using *in vivo* cloning. After these clones were made we had to make sure that unlike the Rfa2 phospho-mimetic plasmids; these pEG202-HITS would not cause auto-activation be testing their growth on SD-His-Leu plates. Unfortunately, 12 out of the 15 candidates showed high level of auto-activation (Figure 2.6). Candidates that showed no auto-activation when cloned as bait were Mcm5, Mif2, and Paa1. Since we were unable to use most of the HITS as our bait in our two-hybrid screen, we decided to try an alternative approach by creating Rfa2-X mutant yeast strains. This would allow us to test interactions between Rfa1-

FAB and HITS (1-15) as we did in our initial screen, but in yeast strains with specific Rfa2-X mutations. This approach would make sure that only Rfa2 with specific mutations would be available in the cell at any given time thereby allowing us to see if these different phosphorylation states of Rfa2 had any effects on our RFA-Protein interactions.

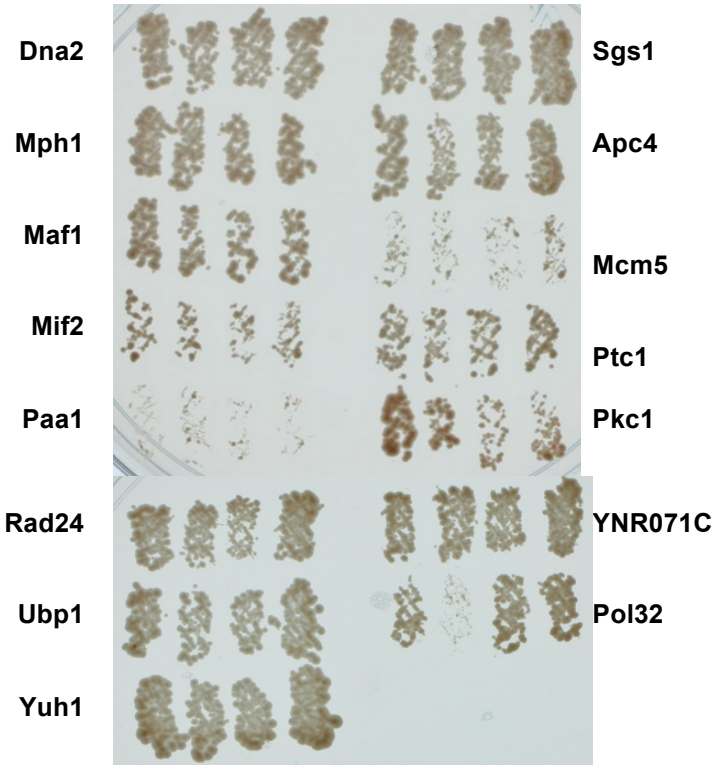


Figure 2.6: Testing autoactivation of prey candidates cloned in pEG202. Prey candidates were cloned into PEG202 to test for autoactivation as bait plasmids. Plasmids were transformed into EGY48 strains and plated on SD-His plates. Colonies were then picked onto SD-His master plates and replica plated onto SD-His-Leu plates. Growth on SD-His-Leu demonstrates autoactivation since the reporter gene (Leu2) is activated without its interaction with a prey plasmid while no growth shows lack of autoactivation. Dna2, Sgs2, Mph1, Apc4, Maf1, Pck1, Rad24, YNR071C, Ubp1, Pol32, and Yuh1 showed growth on SD-His-Leu plates showing autoactivation while Mcm5, Mif2, and Paa1 did not autoactivate. Most of these plasmids were not suitable to be used as bait plasmids.

2.4.6. Sequencing of Rfa1 Subunit Reveals Truncation Which Prevents Complementation

While making the Rfa1 bait plasmid, PCR randomly generated a CAG to UAG nonsense mutation at position 446 creating Rfa1 that constituted only the first 148 aa. I called this truncated protein Rfa1-FAB, since it lacked DBD-C completely, but contained all of the other DBDs (F, A, and B). The Rfa1-FAB fragment cloned into the pEG202 bait plasmid was determined not to auto-activate *LEU2* reporter expression. This construct turned out to be a good alternative to Rfa1-WT for two reasons: 1) it still

preserved regions of Rfa1 that had been previously identified to be important for almost all of RFA-protein interactions, and 2) this truncated protein lacked DBD-C which necessary for Rfa1 to form a complex with Rfa2 and Rfa3, and therefore any interactions identified would most likely be direct interactions with Rfa1 and not an indirect interaction through Rfa2 or Rfa3 in the complex. As expected, Rfa1-FAB was not able to complement Rfa1 in *Rfa1*∆ cells, presumably because of its inability to form the trimerization core and rendering it incapable of being part of a RFA complex.

2.4.7. Interaction Between Rfa1-FAB and HITS is Altered by Different Phosphorylation State of Rfa2

Once we had our potential candidates, we wanted to test if the phosphorylation state of Rfa2 would have any effect on these interactions between Rfa1-FAB and our HITS isolated through the initial two-hybrid screen. To test this, we constructed three mutant strains of EGY48 strains with mutant forms of Rfa2 EGY48-Asp, EGY48-Ala, and EGY48-∆N. These three strains were used to perform the yeast-two hybrid assays for all 15 HITS in parallel to test for any differential interactions between them and Rfa1-FAB. Interestingly, while Rfa1-FAB is not capable of making a hetero-trimeric complex with Rfa2 and Rfa3, the phosphorylation state of Rfa2 still had an effect on interaction between Rfa1-FAB and a subset of our HITS. This hints at the fact that phosphorylation state of Rfa2 in one complex can affect interaction of proteins with Rfa1 in a different RFA complex. In summary, Dna2, Sgs1, and Mph1 showed no difference in interaction with Rfa1-FAB regardless of the phosphorylation state of Rfa2. Interaction between Rfa1-FAB and Apc4, Maf1, Pkc1, Mcm5, Mif2, Ptc1, and Ubp1 were disrupted in the EGY48-Asp (mimicking hyper-phosphorylated Rfa2) strains but were unaffected in EGY48-Ala (unphosphorylatable form) and EGY48-∆N strains. Finally, Interaction between Rfa1-FAB and Paa1, Pol32, YNR071C, and Yuh1 were disrupted in EGY48-Asp strains while they also showed weak interaction in EGY48-Ala and EGY48-∆N strains. The summary of these interactions in various Rfa2-X mutant strains is listed in Table 2.7 and Figure 2.7.

	HITS	EGY48-WT	EGy48-Asp	EGY48-Ala	EGY48-ΔN
1	Dna2	+	+	+	+
2	Sgs1	+	+	+	+
3	Mph1	+	+	+	+
4	Apc4*	+	-	+	+
5	Maf1*	+	+/-	+	+
6	Mcm5*	+	-	+	+
7	Mif2*	+	-	+	+
8	Ptc1*	+	+/--	+	+
9	Paa1	+	-	+/-	+/-
10	Pkc1*	+	+/-	+	+
11	Rad24	+	+	+	+
12	YNR071C*	+	-	+/-	+/-
13	Ubp1*	+	+/--	+	+
14	Pol32*	+	-	+/-	+/-
15	Yuh1*	+	-	+/-	+/-

Table 2.7: Growth on SG-His-Leu-Trp-Ura with testing done in various Rfa2 mutant strains. The shortlisted candidates were re-transformed in EGY48 cells with different Rfa2 mutations. Hence each interaction between Rfa1-FAB and HITS was tested simultaneously in four different yeast strains EGY48-Rpa2-WT, EGY48-Rfa2-Asp, EGY48-Rfa2-Ala, EGY48-Rfa2-ΔN. Cell growth indicates the presence of interaction between Rpa1-FAB and HITS in different EGY48 Rfa2-X strains after 3 day incubation on SG-His-Leu-Trp-Ura plates at 30°C. No growth indicates lack of interaction between Rfa1-FAB and HITS.

\+ = Full growth

+/- = Partial growth

+/-- = Very little growth

\- = No growth

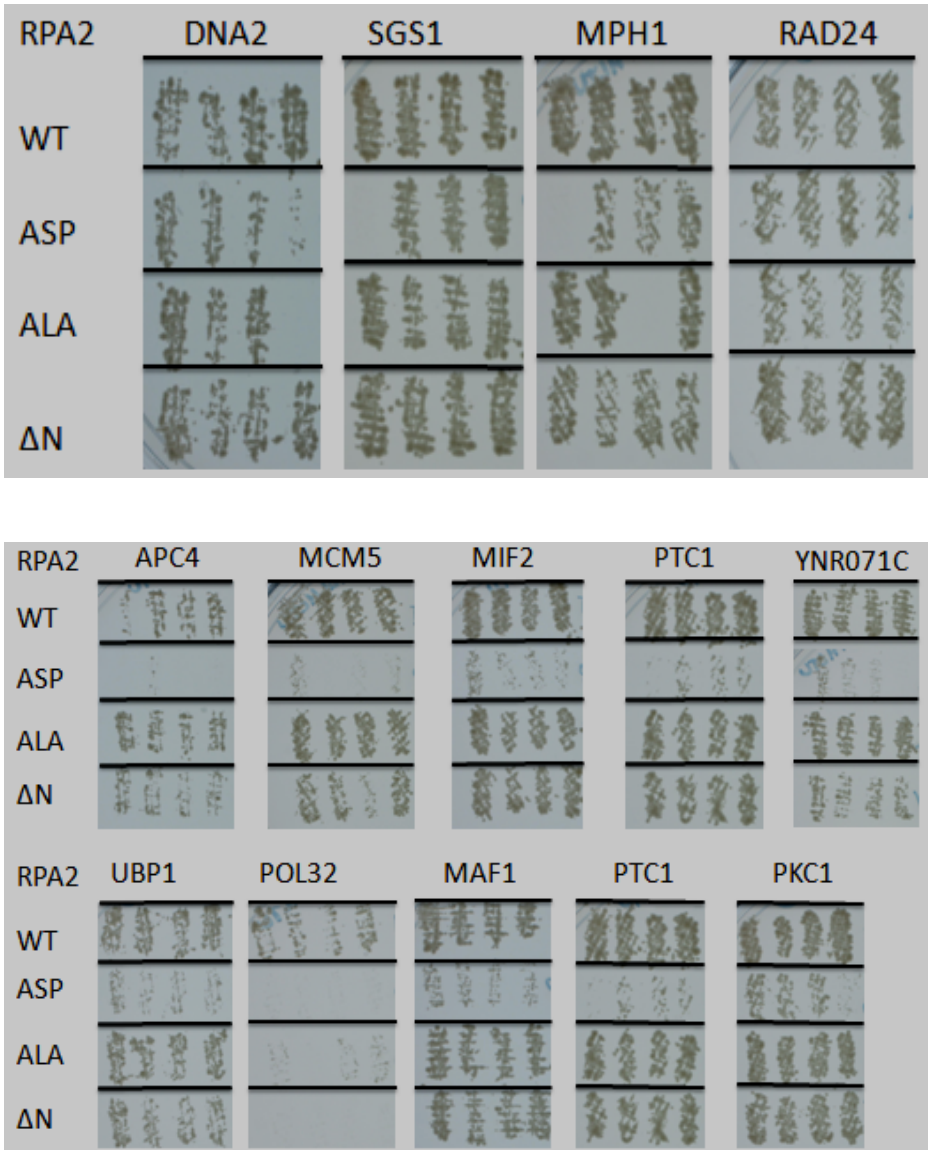


Figure 2.7: Interaction between Rfa1-FAB (bait) and HITS (prey) in Rfa2 mutant strains visualized by growth on SG-His-Leu-Trp-Ura. Dna2, Sgs1, Mph1, and Rad24 interaction with Rfa1-FAB is not affected by phosphorylation state of Rfa2. Apc4, Mcm5, Mif2, Ptc1, YNR071C, Ubp1, Pol32, Maf1, Ptc1, and Pkc1 interaction with Rfa1-FAB is abrogated by presence of Rfa2-Asp mutant demonstrate by lack of cell growth in EGY48-Rpa2-Asp. Additionally, interaction between Rfa1-FAB and Pol32 is also disrupted by Rfa2-Ala and Rfa2-ΔN. These results are also summarized in Table 2.7.

2.5. Conclusion and Discussion

Using the Yeast two-hybrid assay, we were able to isolate 12 novel and 3 known proteins interacting with Rfa1-FAB *in vivo*. Isolation of known interacting partners was important for the validation of our screen and we were excited to isolate a few such known interactions through our screen. In fact, a recent paper demonstrating interaction between Rfa1 and Sgs1 by Susan Gasser's lab was published

shortly after we did our screen identifying this particular interaction (Hegnauer et al, 2012). They identified the same region of Sgs1 that was important for its interaction with Rfa1 as we isolated in our screen. Not only did this validate our screen, but we were able to narrow down the region on Sgs1 inherently, whereas Hegnauear *et al*. had to perform a deletion to series achieve the same thing. All genes characterized as potential candidates (HITS) are involved in various DNA metabolic pathways at different stages of the cell cycle as mentioned throughout chapter 1. Our prey candidates also fall under various functional categories such as helicases and nucleases (Dna2, Sgs1, Mph1), ubiquitin ligase (Apc4, Ubp1, Yuh1), Kinases and phosphotases (Pkc1, Ptc1), polymerase (Pol32, Maf1), and checkpoint proteins (Rad24) all of which are central in processing DNA during various stages of the cell cycle. Given that RFA is an ssDNA binding protein and all these proteins have to encounter ssDNA for carrying out their function at one point or other in the cell cycle, it is perfectly suitable to have these categories of protein in our interaction pool.

The most interesting observation we encountered was that Rfa2 phosphorylation mutant can affect Rfa1-FAB interaction with other proteins in the cell. It is commonly thought that Rpa2 phosphorylation causes three dimensional conformation changes to the overall RFA structure which likely makes it favorable for certain proteins to interact with it during different stages of the cell cycle. Another hypothesis is that Rpa2 hyper phosphorylation causes it to act like a short stretch of ssDNA which can then compete to bind DBD-F and DBD-B on Rpa1 to alter its interaction capability with both other proteins and DNA (Liu et al, 2005). Both models assume that Rpa2 phosphorylation is causing RPA complex to interact with different proteins with its immediate effect on the RPA complex but our observation draws light to yet another possibility. Since multiple RPA molecules can bind ssDNA depending on the length of ssDNA it is possible that adjacent RPA molecules act in concert to modulate RPA-protein and RPA-DNA interaction in the cell. This might lead to mobilization of numerous RPA molecules in a short period of time and scale the effects of phosphorylation as hyperphosphorylated Rpa2 that is not bound to ssDNA can still modulate RPA on ssDNA to interact with specific proteins necessary at various stages of DNA metabolism. This hypothesis is further supported by the fact that Rfa1-FAB was able to interact with all our HITS without it being in a complex with Rfa2/ Rfa3. While Rfa1-FAB cannot substitute Rfa1-WT in the cell, we have shown that it can interact with a subset of proteins that are important in DNA metabolic

pathways. This leads us to believe that Rfa1 does not have to be exclusively in a complex to interact with some of its partners and Rfa2 does not have to be in a complex either to effect protein interactions between Rfa1 and its partners. This might also explain why we isolated numerous helicases and nucleases (e.g. Dna2 and Sgs1), and a regulator of polymerases (Maf1). Since Rfa1-FAB cannot be part of the RFA complex, it is probably only capable of interacting with such group of proteins strongly with its Rfa1-FAB domain.

The disruption of interaction between Rfa1-FAB and HITS due to MMS treatment or different phosphorylation state of Rfa2 demonstrates that these interactions have a functional relationship in cellular metabolism and also gives clue to their grouping in specific pathways. Since MMS causes DNA damage it cannot be confirmed if disruption of interaction between Rfa1-FAB and their protein is caused solely by hyperphosphorylation of Rfa2 as it may potentially phosphorylate numerous other proteins in the cell altering interaction dynamics at all fronts. Understanding the precise mechanism and importance of each interaction in cell will require studying each individual interaction at a molecular level and may reveal interesting details on how RFA mediates complex processes at different phases of the cell cycle and after DNA damage. We decided to further analyze interaction between RFA and Rad24 mostly because of its recurrence as an MMS sensitive candidate in our analysis and due to its involvement in the early stages of DNA damage signaling pathway. The next chapter will feature detailed analysis of Rfa1 and Rad24 interaction and identity regions on both proteins important for this interaction.

2.6. References

Bae KH, Kim HS, Bae SH, Kang HY, Brill S, Seo YS (2003) Bimodal interaction between replication-protein A and Dna2 is critical for Dna2 function both in vivo and in vitro. *Nucleic Acids Res* **31:** 3006-3015

Banerjee S, Smith S, Oum JH, Liaw HJ, Hwang JY, Sikdar N, Motegi A, Lee SE, Myung K (2008) Mph1p promotes gross chromosomal rearrangement through partial inhibition of homologous recombination. *J Cell Biol* **181:** 1083-1093

Caestecker KW, Van de Walle GR (2013) The role of BRCA1 in DNA double-strand repair: past and present. *Exp Cell Res* **319:** 575-587

Fields S, Sternglanz R (1994) The two-hybrid system: an assay for protein-protein interactions. *Trends Genet* **10:** 286-292

Groth P, Ausländer S, Majumder MM, Schultz N, Johansson F, Petermann E, Helleday T (2010) Methylated DNA causes a physical block to replication forks independently of damage signalling, O(6)-methylguanine or DNA single-strand breaks and results in DNA damage. *J Mol Biol* **402:** 70-82

Hegnauer AM, Hustedt N, Shimada K, Pike BL, Vogel M, Amsler P, Rubin SM, van Leeuwen F, Guénolé A, van Attikum H, Thomä NH, Gasser SM (2012) An N-terminal acidic region of Sgs1 interacts with Rpa70 and recruits Rad53 kinase to stalled forks. *EMBO J* **31:** 3768-3783

Liu Y, Kvaratskhelia M, Hess S, Qu Y, Zou Y (2005) Modulation of replication protein A function by its hyperphosphorylation-induced conformational change involving DNA binding domain B. *J Biol Chem* **280:** 32775-32783

Oakley GG, Tillison K, Opiyo SA, Glanzer JG, Horn JM, Patrick SM (2009) Physical interaction between replication protein A (RPA) and MRN: involvement of RPA2 phosphorylation and the N-terminus of RPA1. *Biochemistry* **48:** 7473-7481

Wang Y, Cortez D, Yazdi P, Neff N, Elledge SJ, Qin J (2000) BASC, a super complex of BRCA1-associated proteins involved in the recognition and repair of aberrant DNA structures. *Genes Dev* **14:** 927-939

Wold MS (1997) Replication protein A: a heterotrimeric, single-stranded DNA-binding protein required for eukaryotic DNA metabolism. *Annu Rev Biochem* **66:** 61-92

CHAPTER 3: THE N-TERMINAL REGION OF RPA1 IS RESPONSIBLE FOR INTERACTION WITH THE C-TERMINAL REGION OF RAD24 IN VIVO

3.1. Abstract

Rad24 (Rad17 in humans and *Schizosaccharomyces pombe*) is a checkpoint protein in *Saccharomyces cerevisiae*. It is involved in loading the Rad17-Mec3-Ddc1 complex (Rad9-Rad1-Hus1 complex in humans) onto damaged DNA. Although physical interaction between Rad24 and RFA has not been shown, the recruitment of Rad24 requires RFA to be present on ssDNA (Majka & Burgers, 2003). While the involvement of RFA and Rad24 in loading the 17-3-1 complex was known, the exact region of interaction between these proteins has not been characterized. In our yeast two-hybrid assay, we identified Rad24 as one of the protein partners of Rfa1-FAB via multiple independent interactions. Upon further testing, we revealed that this interaction was alleviated by treatment of cells with DNA damaging agent MMS. Since MMS causes DNA damage and presumably phosphorylation of various proteins, the exact mechanism of how MMS affects RPA-Rad24 binding is not understood. Based on this observation, and the given importance of this interaction in DNA damage signaling, we decided to further investigate and determine which regions on each protein were important for facilitating their interaction. By performing the beta-galactosidase (β-gal) assay on truncated versions of both Rfa1 and Rad24 genes, we determined that the N-terminal region of Rfa1 (DBD-F) and the C-terminal region of Rad24 (527-660 aa) were crucial for their physical interaction. Further work will be required to determine the exact residue (or group of residues) that are important in this interaction at the molecular level. Since this interaction is important in cell-cycle checkpoint and repair, understanding it will greatly increase our knowledge about how the early part of the signal transduction mechanism works during DNA damage response.

3.2. Introduction

Cell cycle checkpoints were initially described as events that governed the transition between different phases of the cell cycle in order to maintain their genetic integrity. This checkpoint pathway is composed of DNA damage sensors, signal transduction proteins, and DNA repair pathways. DNA damage sensing is the initial step in recognizing DNA structures resulting from numerous DNA damaging events such as UV induced cyclobutane dimers, stalled replication forks, and double-strand breaks (DSBs). Especially DSBs are most harmful to the cells as it can cause chromosome translocation and

other unwanted rearrangements in the genome. Furthermore, unrepaired DSB can lead to chromosome loss during mitosis. The Mre11-Rad50-Xrs2 (MRX) complex (de Jager et al, 2001) and the Ku complex (Ku70 and Ku8) (Walker et al, 2001) are among the first known protein complexes to recognize and bind DSBs. The Ku complex is shown to compete with MRX at the site of DNA breaks and this competition is favored during G1 phase while MRX complex can remove Ku from DNA ends in G2 phase of the cell cycle (Clerici et al, 2008). While Ku binding promotes NHEJ, Mre11 of the MRX complex exhibits nuclease activity that is necessary for processing of DNA breaks. The resection of DNA by Mre11 creates ssDNA which is then bound by RPA (Lewis et al, 2004). This RPA bound ssDNA is thought to be a major signaling complex indicating the presence of DNA damage. Consequently, RPA can direct loading of the 17-3-1 complex on to DNA which is mediated via Rad24-Rfc2-5. This process is similar to the RFC 'clamp loader' and PCNA 'clamp' system present in normal DNA replication which uses RFC to load PCNA on to DNA which then acts as a processivity factor for DNA Pol III and other proteins necessary for successful DNA replication (Warbrick, 2000). Similarly, Rad24-Rfc2-5 'clamp-loader' loads Rad17-3-1 'clamp' on to partial duplex DNA in an ATP dependent manner and is necessary for activating downstream effectors of the DNA damage induced signal transduction pathways which ultimately leads to cell cycle arrest and initiates the repair process, or in some cases induces apoptosis (Treuner et al, 1999). Unlike PCNA which can only be loaded on to the 3' end of the DNA strand by RFC, 17-3-1 can be loaded on both 3' and 5' ends of DNA by Rad24-Rfc2-5 indicating the diversity of structures it can recognize (Majka & Burgers, 2003). But RFA confers directionality to this loading by only allowing loading of 17-3-1 complex at 5' DNA junctions. The later stages of DNA damage response are fairly well characterized but molecular details regarding the initial steps of this pathway are not clearly understood. In humans, a PI3K related kinase ATM (Tel1 in yeast) also localizes at the site of dsDNA breaks independent of Rad9-Rad1-Hus1 (17-3-1) complex suggesting the presence of two independent but co-operative pathways for triggering the checkpoint response. ATM is recruited to DNA break sites by its physical interaction with the C-terminal region of Nbs1 (part of MRN complex in humans) and it is necessary to phosphorylate Ser139 on H2AX (Rogakou et al, 1998). This modified histone can recruit other proteins via their BRTC domain and can also ultimately induce chromatin remodeling, cell cycle arrest and apoptosis. Some important proteins recruited in this manner include damage repair pathway protein MDC1 and chromatin remolding complex

p400 (Swr1 in yeast) (Jungmichel & Stucki, 2010). These proteins are then phosphorylated by ATM to recruit other downstream effectors for signal amplification.

Rad24 interacts with four other RFC subunits (Rfc2, Rfc3, Rfc4, and Rfc5) forming the Rad24-Rfc2-5 complex (Green et al, 2000). All four RFC subunits are structurally similar and function with different proteins to regulate various aspects of DNA metabolism. Rfc1-Rfc2-5 is involved in DNA synthesis in unstressed cells, Ctf18-Rfc2-5 in sister chromatic cohesion (Naiki et al, 2001), and Elg1-Rfc2-5 in genomic stability (Ben-Aroya et al, 2003) and unloading PCNA from DNA (Kubota et al, 2013). Similarly, Rad24-Rfc2-5 is necessary for loading another PCNA-like complex, the 17-3-1 complex, onto DNA after DNA damage. In yeast, this is one of two distinct pathways of initiating the signal transduction pathway following DNA damage. Another pathway involves the Rad9 epistasis group; however, both pathways eventually converge at Rad53 phosphorylation (de la Torre-Ruiz et al, 1998). Rad53 (or Chk2 in humans) will eventually activate its downstream targets to halt the cell cycle and initiate either DNA repair or apoptosis (if the damage is deemed irreparable) (Zegerman & Diffley, 2010). Rad24 has been shown to be important in both mitotic checkpoint activation following various types of DNA lesions and meiotic checkpoint arising from recombination events (Lydall et al, 1996; Weinert et al, 1994). It has an ATPase domain which allows it to hydrolyze ATP while the whole complex loads the 17-3-1 complex onto DNA. (Majka et al, 2004). Binding of RPA to ssDNA after DNA damage is a pre-requisite for Rad24-Rfc2-5 complex to localize onto these sites (Zou et al, 2003). While RPA and Rad24 have been shown to work together in this pathway, their direct interaction has not been demonstrated, nor has the exact region of interaction been identified. The identification of these regions involved in their physical interaction will assist in uncovering the mechanism by which Rad24 is recruited to RPA-coated DNA and how is gets activated to load the 17-3-1 complex.

Since our focus was on the understanding the interaction between Rfa1 and Rad24, we decided to perform the β-gal assay with truncated forms of both proteins to identify specific regions that was necessary for RFA to interact with Rad24. To find the region on Rfa1 necessary for interacting with Rad24, we used truncated form of Rfa1 consisting of Rfa1-DBD-F, Rfa1-L, Rfa1-DBD-F-L, Rfa1-DBD-A, Rfa1-DBD-B, Rfa1-DBD-C, and Rfa1-WT as bait and performed the β-gal assays with Rad24-C as prey (Figure 3.1). We also performed the reciprocal analysis by testing fragments of Rad24 including Rad24-

WT, Rad24-ΔN and Rad24-ΔC as prey and Rpa1-FAB as bait to find the region of Rad24 that was necessary for its interaction with Rfa1 (Figure 3.2) . Based on our results from the β-gal assay we have narrowed down regions on both proteins to the first120 aa of Rfa1 and a last 134 aa of Rad24 to be important in their physical interaction. In summary, we have shown that DBD-F of Rfa1 interacts with the C-terminus of Rad24. This finding could be suggestive of a hand-off mechanism where ssDNA bound RFA interacts with the C-terminus of Rad24 via its DBD-F domain and the N-terminus of Rad24 interacts with Rfc2-5 forming a functional clamp loader complex tethered to damaged DNA which can then effectively load the 17-3-1 complex.

3.3. Methods and Materials

3.3.1. Plasmids Used for Beta-Galactosidase Assay

All Rfa1 bait constructs used in our experiment were kindly provided by Dr. Susan Gasser from Friedrich Miescher Institute in Switzerland. Rfa1 bait constructs from the Gasser lab included Rfa1-WT (1504), Rfa1 DBD-C (1512), Rfa1 DBD-B (1510), Rfa1 DBD-A (1508), Rfa1 DBD-FL (1506), Rfa1-L linker region (2068), and Rfa1 DBD-F (2069). We also received the empty bait vector, pGAL-LexA (965), and a full-length Rfa2 bait vector (2065). These bait plasmids were identical to the pEG202 plasmid we used in our original screen except for the LexA-DBD was also cloned downstream of a constitutive GAL1 promoter similar to our pGJ4-5 'prey' plasmid.

Prey constructs used in our experiments were Rad24-WT, Rad24-ΔN, and Rad24-ΔC. Rad24-WT prey construct was made by amplifying Rad24 from the yeast genomic DNA. For this PCR reaction, I used 10 uL 5 x Phusion HF buffer, 1.5 uL 50 mM $MgCl_2$, 2 uL 10mM dNTP, 5 uL 5 uM pJG4-5-Rad24-FOR primer [5'-CTAC CCTTATGATGTGCCAGATTATGCCTCTCCCGAATTCATGGATAGTACGAATTT GAA-3'], 5uL 5uM pJG4-5-Rad24-REV primer [5'TTGACCAAACCTCTGGCGAAGAAGTCCAAAGCTTCT CGAGT TAGAGTA TTTCCAGATCTG-3'], 0.5 uL 2 U/uL Phusion HF DNA Polymerase, 1 uL yeast genomic DNA, and brought the total volume to 50 uL by adding ddH_2O. The tube was placed in a thermocycler under similar condition as shown in chapter 2 to complete the reaction. 25 uL of the final PCR product was transformed into 100 mL EGY48 cells with 10 uL of linearized pJG4-5 plasmids with EcoRI done as explained in chapter 2. 250 uL of transformed cells were plated on SD-Trp plates to select

for candidates with cloned plasmids. Yeast genomic DNA isolation, bacterial transformation, miniprep, and restriction digest to identity proper clones were also performed as explained in chapter 2.

Rad24-ΔC (461-660) prey construct was made by deleting the last 200 aa from the C-terminal of Rad24 from previously made pGJ4-5-Rad24-WT plasmid using *in vitro* site directed mutagenesis. To do this, I added 10 uL 5 x Phusion HF buffer, 1.5 uL 50 mM $MgCl_2$, 2 uL 10 mM dNTP, 5 uL 5uM Pjg4-5-Rad24-Delta(461-660) primer [5'-ATTCATTTAAAGTTCAAGCTTAACTCG-3'], 1 uL pJG4-5-Rad24-WT plasmid, 0.5 uL 2 U/uL Phusion HF DNA Polymerase, and 24 uL ddH_2O bringing the final solution volume of 50 uL. PCR reaction was done under similar condition as mentioned in chapter 2. Once the mutagenesis reaction was completed, 1 uL of DpnI enzyme was added to the tube and incubated at 37˚C for 2 hours. 1 uL of the final product was transformed in bacteria using electroporation and plated on LB+AMP plates. Colonies grown overnight were then inoculated in 3 mL LB+AMP solution and grown overnight 37˚C at 220 RPM. Plasmids were isolated using miniprep and digested with EcoRI and XhoI to identify proper clones. All results were later verified by DNA sequencing (Eton Bioscience Inc.) using pGJ4-5-UP-SEQ primer [5'-GATCCAGCCTGACTGGCTGAAATCGAATGG-3']. A smaller deletion series of the Rad24 C-terminus was generated in pJG4-5 by Erica Mueller by further dividing this region into three equal parts using *in vitro* site directed mutagenesis. These mutants are referred to as Rad24-C1, Rad24-C2, and Rad24-C3 constituting deletions of 461-526 aa, 527-592 aa, and 593-660 aa respectively.

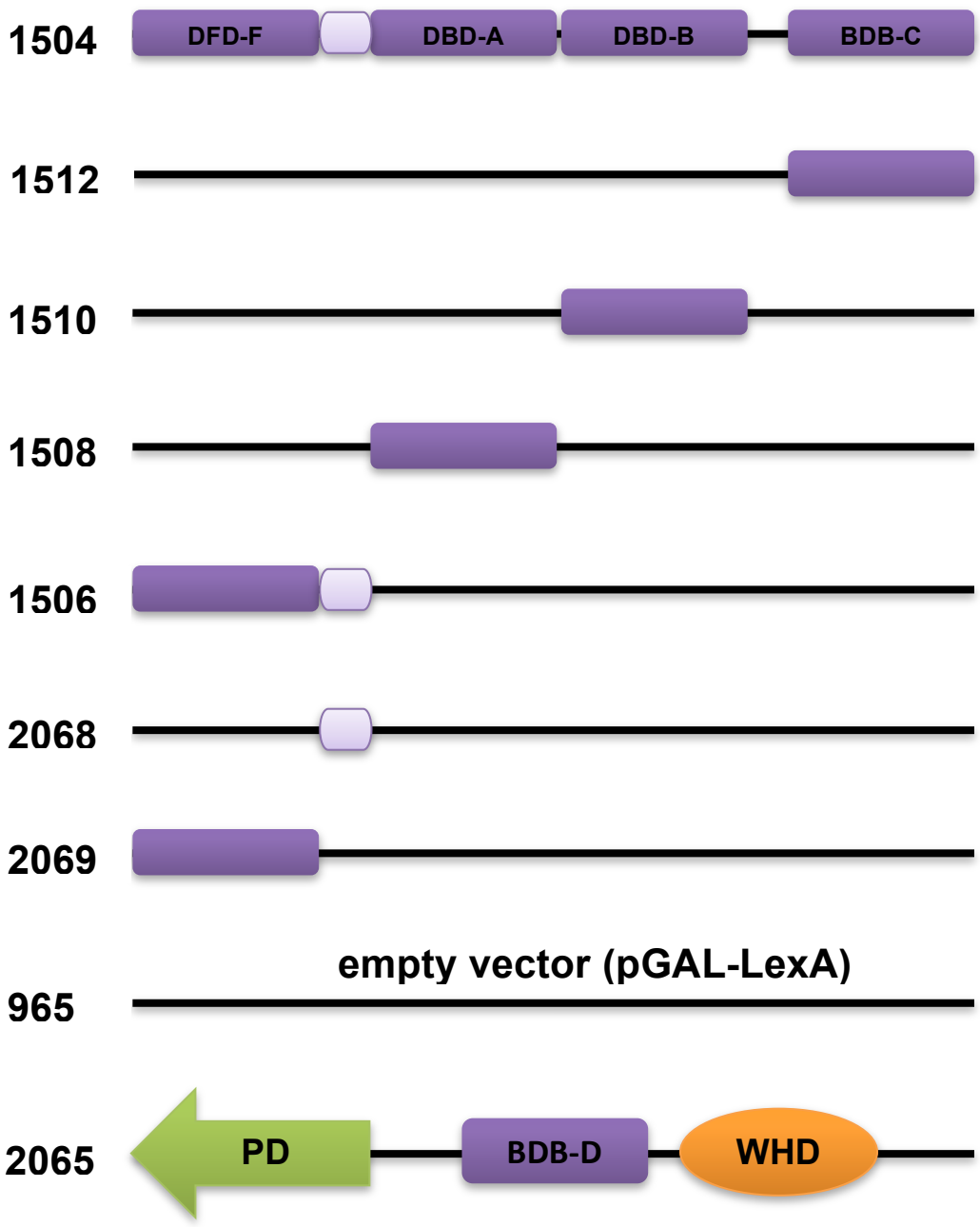


Figure 3.1: Rfa1 fragments and Rfa2 cloned into pGAL-LexA. Schematic representation of Rfa1 and its various DBD domains (purple) used as bait in the beta galactosidase assay. All Rfa1 domains and Rfa2 were fused downstream of LexA DNA binding domain under the Gal1 promoter. 1504 (Rfa1-WT), 1512 (Rfa1-C), 1510 (Rfa1-B), 1508 (Rfa1-A), 1506 (Rfa1-FL), 2068 (Rfa1-L), 2069 (Rfa1-F), 965 (pGAL-LexA), 2065 (Rfa2) were obtained from Dr. Susan Gasser (FMI, Switzerland). These 'prey' plasmids were transformed along with Rad24ΔN 'bait' plasmid and pSH18-34 'reporter' plasmid to test for beta-galactosidase activity using ONPG.

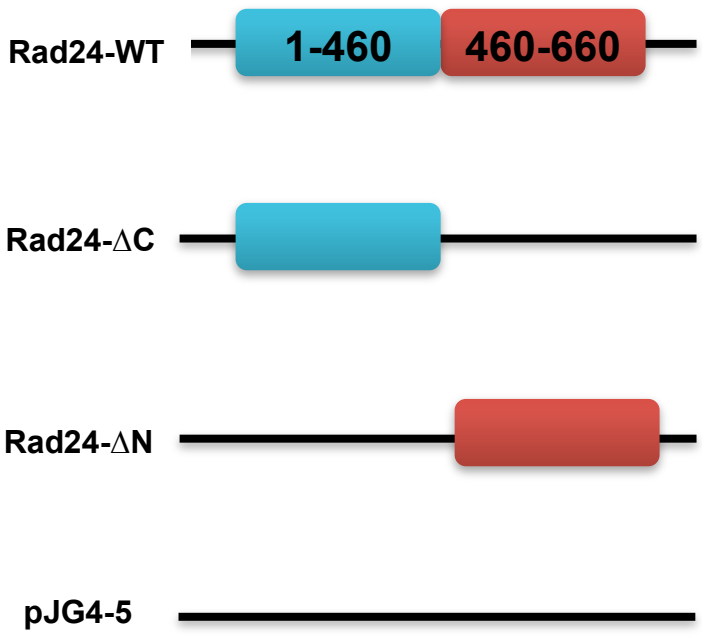
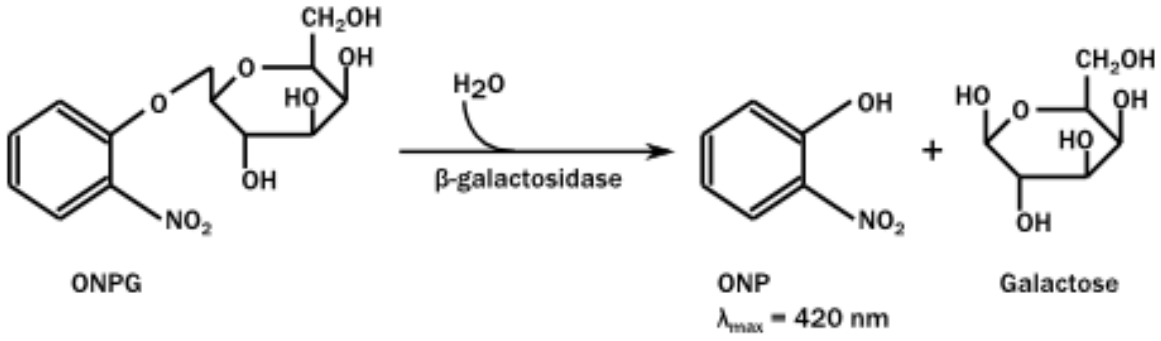


Figure 3.2: Rad24 fragments cloned in pJG4-5. PCR amplified Rad24-WT and Rad24-ΔC were cloned into pJG4-5 vector downstream of a Gal1 promoter. Rad24-ΔN was made my deleting the last 200 aa from Rad24-WT using the site directed mutagenesis. The WT and truncated copies of Rad24 prey plasmids were transformed along with Rfa1-FAB 'bait, plasmid and pSH18-34 'reporter' plasmid to test for beta-galactosidase activity by using ONPG.

3.3.2. Liquid Beta-Galactosidase Assay Using ONPG

While selecting for growth on plates lacking Leu provides us a method for identifying protein interactions, it is not a quantitative assay and only tells us if the interaction is detected or not. To determine the strength of each protein interaction, we utilized a derivative of the Leu selection method called the liquid beta-galactosidase assay. This will let us quantify the interaction between two proteins by measuring the enzymatic activity of the beta-galactosidase protein (product of the *lacZ* reporter gene). When the bait and prey proteins interact strongly, the reporter *lacZ* gene is expressed, which ultimately leads to production of the beta-galactosidase protein. The beta-galactosidase protein normally cleaves lactose into glucose and galactose in the cell, but *ortho*-Nitrophenyl-β-D-galactosidase (ONPG), a synthetic compound similar to lactose is also cleaved by it to produce galactose and o-nitrophenol. The o-nitrophenol compound is bright yellow in color and its intensity can be quantified by colorimetric assay done at 420 nm to calculate the enzymatic activity of β-gal and ultimately infer *lacZ* expression. Hence higher expression of the *lacZ* gene results in increased production of β-gal protein and results in higher enzymatic activity which can be detected and quantified.



To perform the β-gal assay, EGY188 cells transformed with the desired bait, prey, and reporter plasmid were inoculated in 3 mL SD-His-Trp-Ura media and grown overnight at 30˚C and 220 RPM. Next day, 300 uL of the culture was inoculated in new 3 mL SD-His-Trp-Ura media and grown for 6 hours at 30˚C and 220 RPM. Then 2.5 mL of the new culture was inoculated in 25 mL SA-His-Trp-Ura media and grown overnight under similar conditions as described above. Following day, 300 uL 20% galactose was added to each tube and incubated for additional 5 hours at 30˚C at 220 RPM. To perform the actual enzymatic assay, 1 mL of culture was collected in a 2 mL microfuge tube and cells were pelleted by centrifugation at 2500 RCF for 5 minutes. The supernatant was discarded and cell pellet was resuspended in 1 mL Z buffer [4.27 gm Na_2HPO_4, 2.75 gm $NaH_2PO_4.H_2O$, 0.375 gm KCl, 0.125 gm $MgSO_4.7H_2O$, adjust pH to 7.0 and bring volume to 500 mL with ddH_2O]. 200 uL of this new solution was loaded on a 96 well plate to measure cell concentration by taking the absorbance at OD_{600}. The remaining 800 uL solution was mixed with 120 uL chloroform and 80 uL 0.1% SDS and mixed vigorously on a vortex using max speed for 15 seconds to break open the cells and release beta-galactosidase into the solution. The tubes were then incubated at 30˚C for 15 minutes along with 4 mg/mL ONPG solution. Finally, the enzymatic reaction was started by adding 0.2 mL of 4 mg/mL ONPG solution to each sample and mixed for 3 seconds noting the precise time of ONPG addition. As soon as the solution turned pale yellow the reaction was stopped by adding 1M Na_2CO_3 to the solution noting the exact time of this as well (Figure 3.3). Once the reaction was stopped, cells were pelleted by centrifugation for 10 minutes at 21,000 RCF. 200 uL of the supernatant was again added on to a 96 well plate and absorbance at OD_{420} was measured

to determine the concentrations of o-nitrophenol (Figure 3.4). The final beta-galactosidase activity in Miller Units (MU) was calculated using the following equation:

Miller units = OD_{420} / [OD_{600} of assayed culture (mL) * volume assayed (mL)* time (min)]

To ensure the reproducibility of this experiment, MU for each experiment was measured for four individual colonies with any given interaction and the whole set was repeated twice on different days.

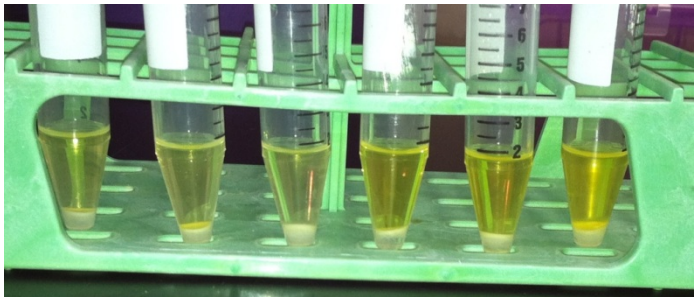


Figure 3.3: Example of beta-galactosidase assay. *Ortho*-Nitrophenyl-β-D-galactosidase (ONPG) is cleaved by beta-galactosidase it to produce galactose and o-nitrophenol. The o-nitrophenol is bright yellow in color which maximum absorbance at 420 nm which can be measure to quantify the LacZ expression and inter protein interaction.

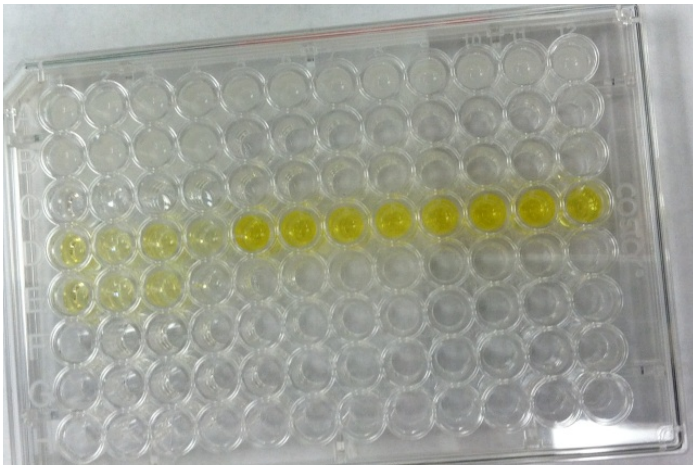
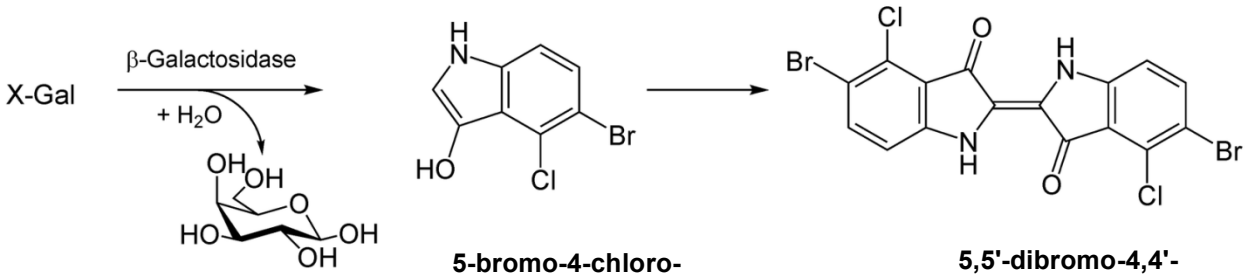


Figure 3.4: Example of 96 well plate loaded with 200 uL samples for spectroscopic reading. Absorbance at OD_{600} and OD_{420} were taken to calculate cell and o-nitrophenol concentrations respectively.

3.3.3. Testing Beta-Galactosidase Expression on X-gal Plates

EGY188 cells transformed with the desired bait, prey, and reporter plasmids were grown on SD-His-Trp-Ura plates. These cells were repicked onto another SD-His-Trp-Ura master plate with 9 individual

colonies from each interaction and grown for 2-3 days at 30˚C. This master plate was replica-pated on to SD-(X-gal)-His-Trp-Ura plates and grown at 30˚C for additional 2-3 days. X-gal (5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside) similar to ONPG is an analog of lactose and also hydrolyzed by beta-galactosidase producing galactose and 5-bromo-4-chloro-3-hydroxyindole. The product 5-bromo-4-chloro-3-hydroxyindole can spontaneously dimerize and get oxidized to 5, 5'-dibromo-4, 4'-dichloro-indigo which is deep blue in color. Hence interaction between bait and prey proteins will cause the expression of the *lacZ* gene which in turn will make the yeast cells blue in presence of X-gal making it easy to identify them.



5-bromo-4-chloro-3-hydroxyindole

5,5'-dibromo-4,4'-dichloro-indigo (Blue)

3.4. Results

3.4.1. The N-terminus DNA Binding Domain of Rfa1 is Necessary for Interaction Between Rfa1 and Rad24

Rad24-Rfc2-5 loading the 17-3-1 complex on to DNA in the presence of RFA is part of the initial steps of the DNA damage signal transduction pathway. However, the exact mechanism of how this process works is not understood. Through our β-gal assay we have determined that Rfa1 DBD-F (1504) is necessary for interaction between Rfa1 and Rad24. Miller units from 1504 (Rfa1-FABC) and Rad24ΔN were set arbitrarily to 1 and all other interactions was normalized against it for analyzing relative interaction strength of each region. Based on our analysis, we find that Rfa1-FL shows the strongest relative interaction with Rad24-WT with 2.5 X stronger interaction than with Rfa1-WT. Rfa1-F also showed 1.5 X stronger interaction with Rad24-WT than Rfa1-WT. Rfa1-C, Rfa1-B, and Rfa1-A fragments showed interactions as weak as the empty vector pGAL-LexA and pJG4-5, which demonstrates that these regions are not involved in interacting with Rad24-WT (Figure 3.5). Rfa2-WT also showed 1.5 X stronger

interaction with Rad24 than Rfa1-WT but this data is surprising considering our Rfa2-WT plasmid auto activates. Nevertheless, their (Susan Gasser) plasmid pGAL-LexA is different from our pEG202 as they have a Gal1 promoter which is not present in the pEG202, may account for such discrepancy. However of this interaction is in fact true, then Rfa1 and Rfa2 both might contribute in interacting with Ras24 either via direct or indirect interaction These results were further supported by presence of blue color in the colonies transformed with Rfa1 DBD-F, Rfa1-F, and Rfa1-FL on the X-gal plates (Figure 3.6). This study along with a recent paper showing interaction between the same region of Rfa1 and another protein, Sgs1, demonstrated that Rpa1 DBD-F is an important region in RFA for mediating various protein interactions. In our results we see that the DBD-FL (1506) and DBD-F (2069) fragments bind Rad24 C term better than Rfa1-WT (1504). These might be a few reasons why this is observed. First, it is possible that the fragments have less steric hindrance associated with its fold since it is not part of the whole RFA complex. This may provide a better binding surface for Rad24 and hence makes for a stronger interaction. Second, the Rfa1-WT (1504) has to compete with the native Rfa1 native to the cell and hence even if the interaction between Rfa1-WT (1504) and Rad24 C term is just as strong as other fragments, the competition will affect the binding strength as less of Rad24 is available for binding to start with. Hence, to solidify our results we will need to perform an *in vitro* binding assay such as co-immunoprecipitation and correlate the results with our β-galactosidase assay.

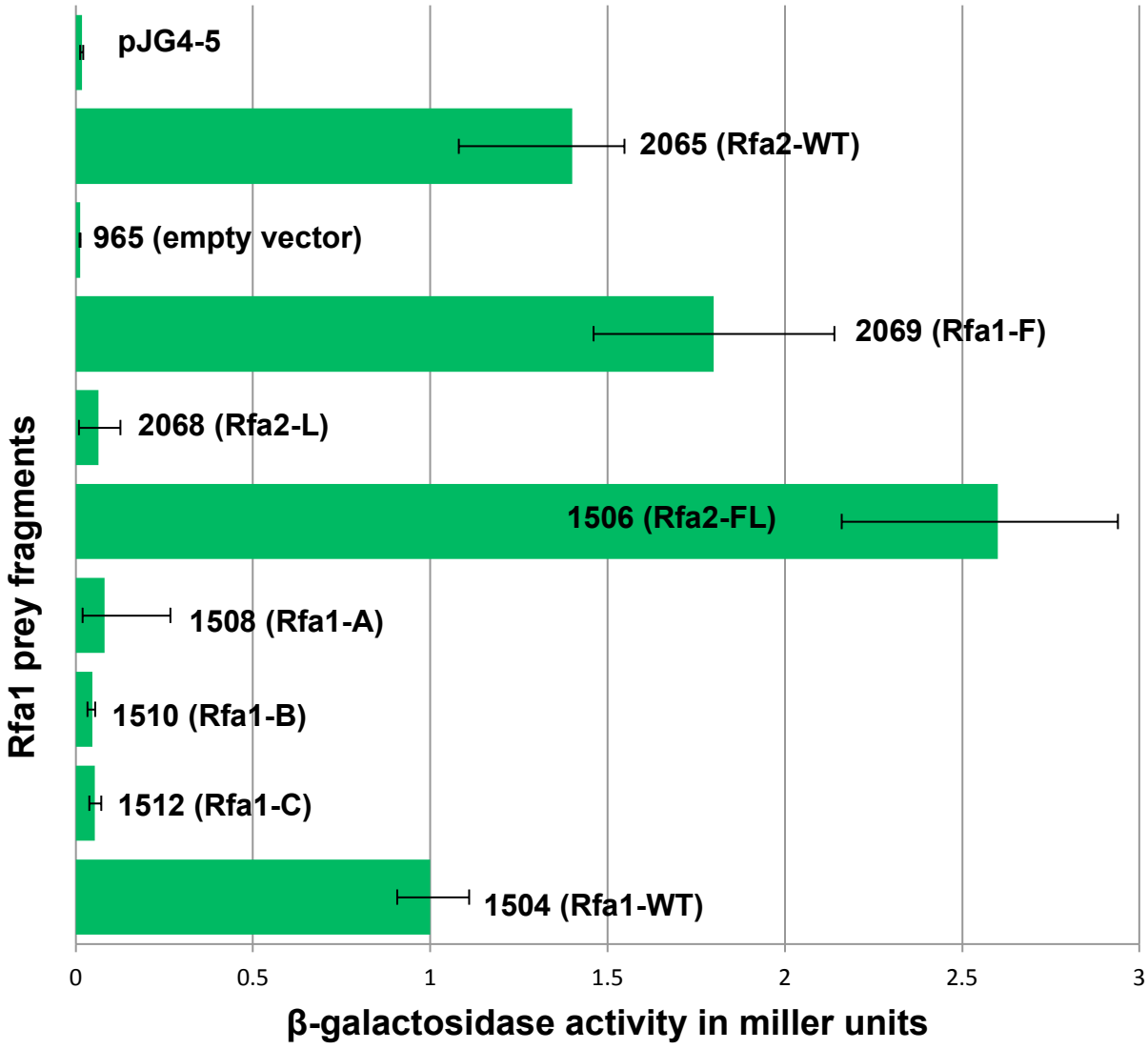


Figure 3.5: Relative MU showing interaction between for various Rfa1 fragments and Rad24ΔN. MU values for Rfa1-WT and Rad24ΔN was arbitrarily set to 1 and all other MU from other Rfa1 fragments were normalized against to measure relative interaction strength. Rfa1-FL and Rfa1-F showed twice as strong interaction with Rad24 than Rfa1-WT indicating that interaction between these two Rad24 and Rfa1 is mediated via this domain. Rfa2 also showed slightly stronger interaction with Rad24ΔN that Rfa1-WT. Hence both Rfa1 and Rfa2 might be involved in RFA complex interaction with Rad24. All other fragments showed interaction as weak as the empty pGAL-LexA.

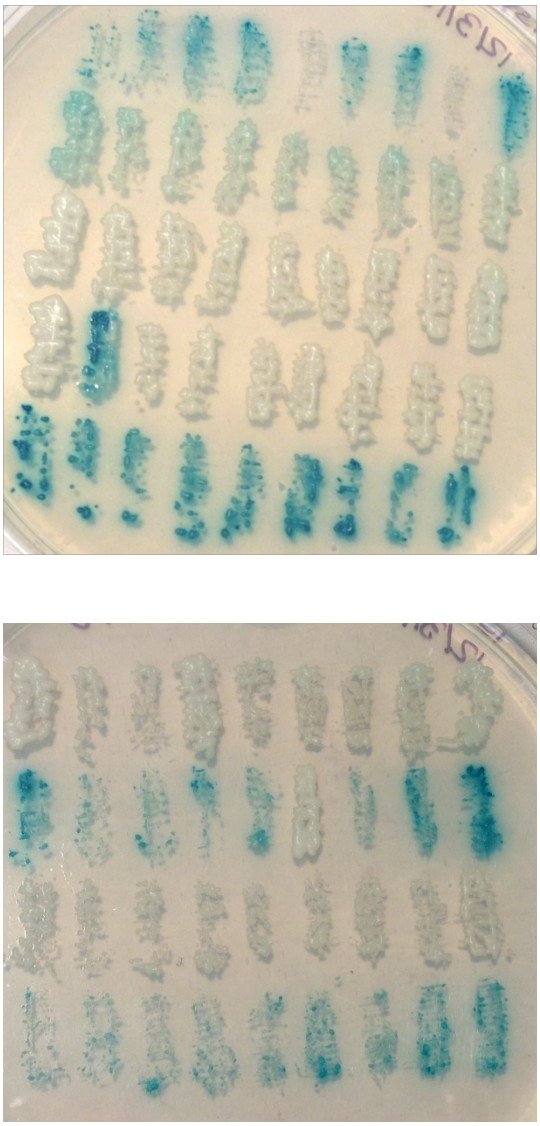


Figure 3.6: X-gal plates with yeast colonies showing interaction between Rfa1 fragments and Rad24. The blue colonies represent the expression of lacZ in the yeast cells. If the bait and prey plasmids interact with each other, then LacZ gene on the reporter plasmid can be transcribed and the blue colors can be visualized on X-gal plates. Data shown here supports the results from in the liquid beta-galactosidase assay where Rfa1-FABC, Rfa1-FL, Rfa1-F, and Rfa2-WT shows interaction with Rad24ΔN. Colonies transformed with Rfa1-C, Rfa1-B, Rfa1-A, Rfa1-L show no blue color like the empty pGAL-LexA plasmid indicating lack of protein interaction.

3.4.2. The C-terminal End of Rad24 is Necessary for Interaction Between Rfa1 and Rad24

Once we determined that the region of Rfa1 necessary for interaction with Rad24, we wanted to now identify the region(s) on Rad24 that was necessary for its interaction with RPA. Since we isolated Rad24ΔN as an interacting partner with Rpa1-FAB from our two-hybrid screen, we decided to clone the full-length Rad24 and only the N-terminus (Rad24ΔC; deletion of amino acids 460-660) into pJG4-5. We

then tested these fragments with full-length Rfa1 for interaction to identify which region was important. Miller units from Rfa1-FAB and Rad24ΔN were set arbitrarily to 1 and all other interactions were normalized against it for analyzing relative interaction strength of each region. As expected, Rad24ΔN showed the strongest relative interaction which was 4 times stronger than Rad24-WT. However, Rad24-ΔC showed interaction signal as weak as the empty pJG4-5 vector, suggesting this region is not necessary for Rfa1-Rad24 interaction (Figure 3.7). To further narrow the region on the Rad24 C-terminus (460-600 aa) we fragmented it into three smaller regions and tested their interaction with Rfa1. Based on the X-gal experiments done with these sub clones, we narrowed down the region to the last 132 aa as being important for its interaction with Rfa1. The Rad24-C2 and Rad24-C3 showed no activity on the X-gal plates whereas Rad24-C1 had plenty of blue colonies indicating strong interaction between this region and Rfa1-FAB (Figure 3.8). To further understand how this region mediates protein interaction we preformed bioinformatics analysis of the DNA sequence in this region. Since there was no known 3D structure of this region published, we used protein modeling software to predict its putative structure. Structure prediction algorithms used in COILS software suggests presence of a putative coiled-coil domain in the C-terminal region of Rad24 and this might be important for its physical interaction with Rfa1 (Figure 3.9). Further analysis will be required to validate this hypothesis. As mentioned previously, the Rad24 C term fragment show a stronger interaction with Rfa1-WT than Rad24-WT which can be attributed to a potential fold within the C terminal of Rad24 which poses a lesser steric hindrance and a better binding site for Rfa1-WT. Furthermore, the Rad24-WT has to compete with the native Rad24 present in the cell which might reduce the number of Rfa1-WT present for binding in the first place.

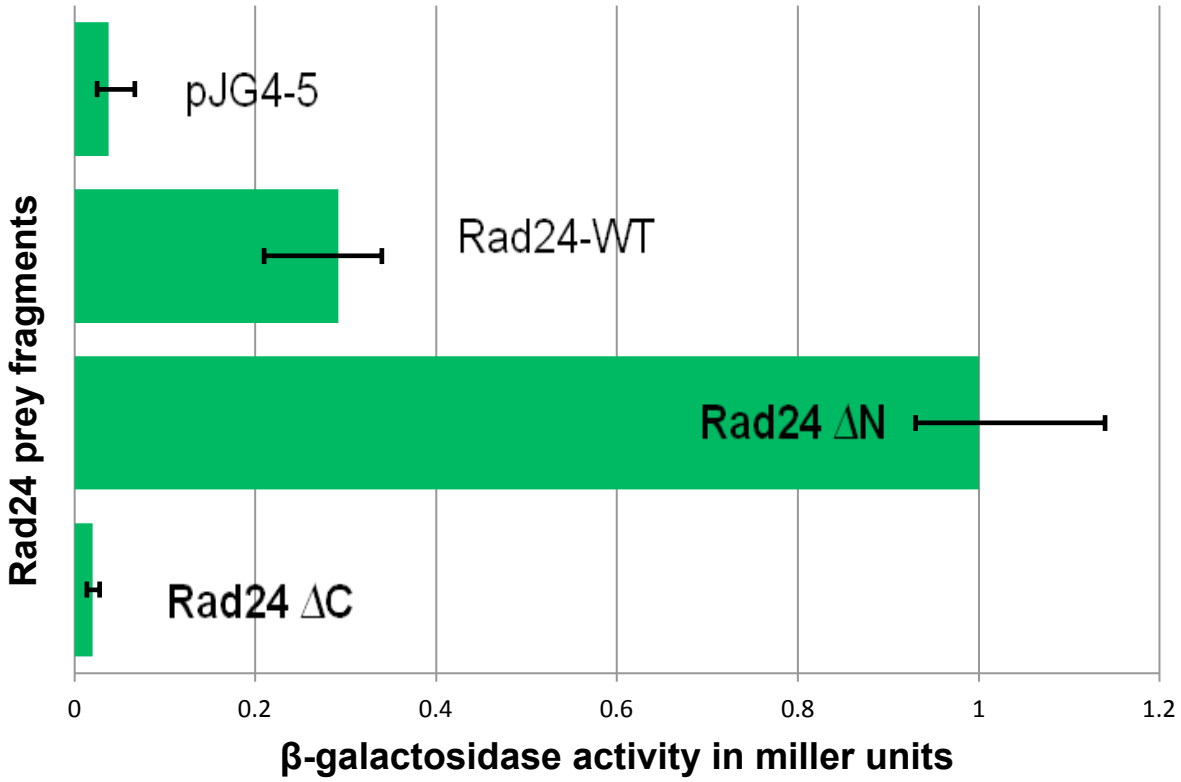


Figure 3.7: Relative MU showing interaction between Rad24 fragments and Rfa1-FAB. MU values for Rfa1-WT and Rad24ΔN was arbitrarily set to 1 and MU from all other fragment were normalized against it measuring relative interaction strength. Rad24ΔN showed strongest interaction with Rfa1-FAB as much as four times higher that Rad24-WT. Rad24ΔC showed interaction as weak as the empty pGP4-5 vector. In summary, interaction between Rfa1-FAB and Rad24 is mediated by the C-term (460-660 aa) of Rad24 which contains a putative coiled-coil domain which might important for their interaction.

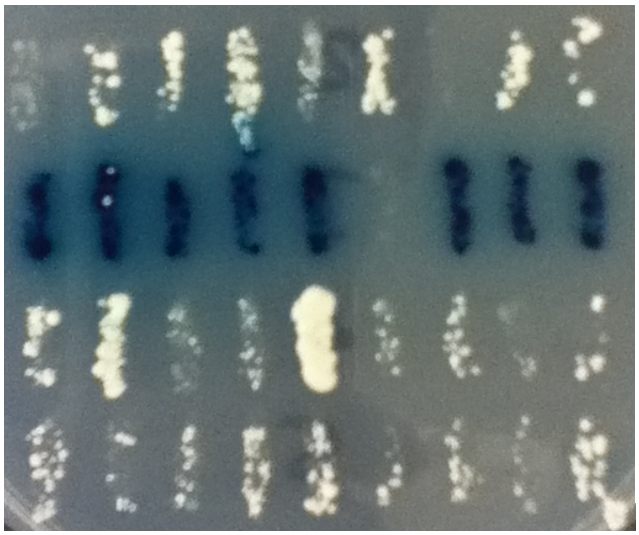


Figure 3.8: X-gal plates showing interaction between Rad24 C-terminal fragments and Rfa1-FAB. Rad24-C1, Rad24-C2, and Rad24-C3 corresponds to deletions of 461-526 aa, 527-592 aa, and 593-660 aa respectively from the Rad24ΔN strain. The blue colonies represent the expression of lacZ in the yeast cells indicating the presence of protein interaction. Based on the preliminary results from this assay we can see that Rad24-C1 shows interaction between Rad24 and Rfa1-FAB whereas Rad24-C2 and Rad24-C3 shows no blue color indication lack of protein interaction.

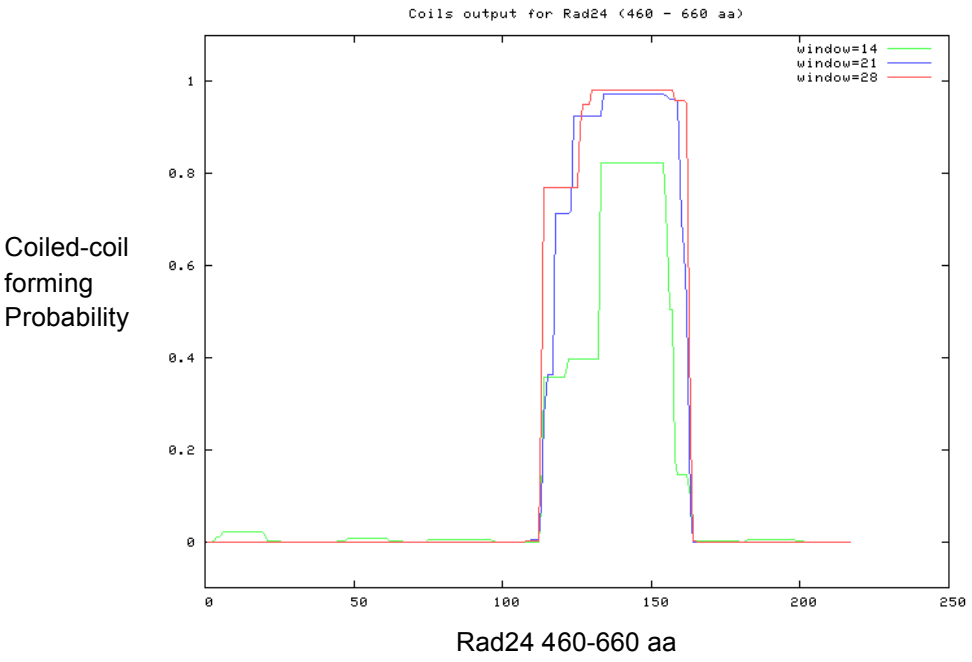


Figure 3.9: Putative coiled coil domain in the C-terminal region of Rad24. COILS, a domain mapping software predicts the C-terminal region of Rad24 to constitute a coiled-coil domain which is a common motif present in proteins regulating gene expression such as transcription factors (e.g. c-Fos and C-jun). Green, blue, and red lines correspond to a scanning window of 12,21, and 28 residues respectively to predict the coiled-coil domain.

3.5. Conclusion and Discussion

Using the β-gal assay we were able to demonstrate that Rfa1 and Rad24 physically interact *in vivo* and the N-terminal DBD-F and C-terminal (460-660 aa) regions on the respective proteins were important for this interaction. Furthermore using deletion mutants of the C-terminal Rad24 region, our preliminary data has mapped the last 132 aa to be crucial for their interaction. This region also contains a putative coiled-coil domain that might facilitate its interaction with Rfa1-DBD-F. It is clear that RFA mediates different processes in the cell by its ability to interact with different proteins at various stages of the cell cycle. After sensing DNA damage, RFA helps in loading of the 17-3-1 complex onto damaged DNA by its ability to interact with Rad24-Rfc2-5 complex. While their overall involvement in this pathway is known the mechanistic details by which these interactions are carried out remains uncharacterized. Another complex Mec1/Ddc2 is recruited on RFA coated ssDNA independent of 17-3-1 and Rad24-Rfc2-5 complex, but their localization on DNA is crucial for initiating and enhancing Mec1 dependent checkpoint signaling in both G1 and G1 phase of the cell cycle (Finn et al, 2012). Activated Mec1 phosphorylates numerous protein including Ddc1, Mec3, Rfa1, Rfa2, Rad24, Rad53, Chk1, and Rad9 all with implications in orchestrating the damage repair process. Activated Rad53 can further phosphorylate and activate Dun1 and Cdc5 for preventing mitotic exit and initiating transcription of repair genes. Activation of Chk1 can activate Pds1 which inhibits anaphase entry by preventing the separation of sister chromatids (Figure 3.10). Since this intricate DNA damage dependent signaling cascade relies on proper loading of 17-3-1 complex, understanding the interaction between RFA and Rad24-Rfc2-5 is vital for understanding early steps in the process.

In this study we have made advancements in our understanding of how RFA modulates Rad24-Rfc2-5 and 17-3-1 loading kinetics on to DNA damage sites by showing their ability to physically interact with each other. This interaction probably helps loading and unloading of Rad24-Rfc2-5 complex on DNA damage sites. K115E mutation in the NTP binding motif of Rad24 has been shown to dissociate interaction between Rad24 and Rfc2-5 proteins (Naiki et al, 2000) suggesting that the N-terminal region of Rad24 is important for it to form a complex with Rfc2-5(2-5).

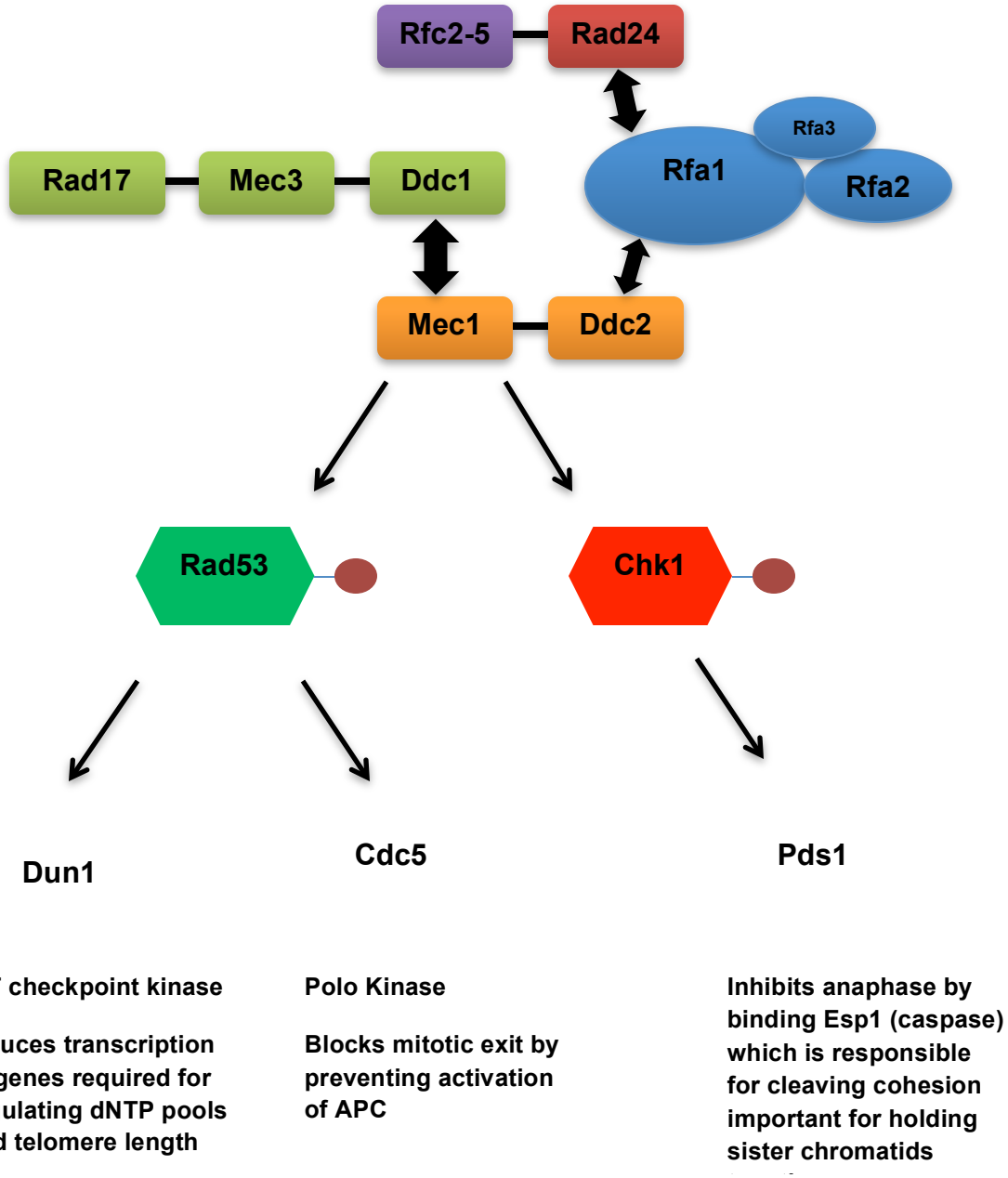


Figure 3.10: Putative model of steps involved in the DNA damage induced signal transduction pathway. SRFA binds ssDNA and recruits Rad24-Rfc2-5 via Rfa1-DBD-F. Rad24-Rfa2-5 can then load 17-3-1 complex to damage site on the 5' junctions. Mec1/Ddc2 is recruited independently at these sites by 17-3-1 greatly stimulates the kinase activity of Mec1 which can then phosphorylate numerous downstream effectors in this signal transduction pathway to initiate DNA repair.

Putting these two results together we find two distinct binding regions on Rad24 that is important for to bind to Rfc2-5 and RPA. Specifically, Rad24 interacts with Rfc2-5 and RFA via its N-terminal and C-terminal regions respectively. In this study we have made advancements in our understanding of how RFA modulates Rad24-Rfc2-5 and 17-3-1 loading kinetics on to DNA damage sites by showing their

ability to physically interact with each other. This interaction probably helps loading and unloading of Rad24-Rfc2-5 complex on DNA damage sites. K115E mutation in the NTP binding motif of Rad24 has been shown to dissociate interaction between Rad24 and Rfc2-5 proteins (Naiki et al, 2000) suggesting that the N-terminal region of Rad24 is important for it to form a complex with Rfc2-5(2-5). Putting these two results together we find two distinct binding regions on Rad24 that is important for to bind to Rfc2-5 and RPA. Specifically, Rad24 interacts with Rfc2-5 and RFA via its N-terminal and C-terminal regions respectively. Upon Rad24-Rfc2-5 binding to RFA, it is able to load the 17/3/1 complex specifically on the 5' ssDNA/dsDNA junction, which can then slide across DNA using ATP hydrolysis and signal downstream effectors to halt the cell cycle buying time for the repair process. Majka et al also looked at Rfa1-DBD-FΔ mutant and Rfa1-t11 mutant interaction with Rad24-Rfc2-5 and found that these mutants were unable to bind Rad24-Rfc2-5 complex which mirrors our results from this study. Furthermore, dissociation of Rad24-Rfc2-5 complex from DNA coated with these mutant forms of RFA was greatly reduced in comparison to wild type RFA. Since we found Rad24 interaction with Rfa1-FAB domain, we can assume that Rad24 and Rfa1 interaction is independent of the Rfa2/Rfa3 subunits of RFA. This suggests the presence of proteins that can interact with specific RFC subunits without the presence of a properly formed RFA complex.

Since this study is done in vivo where the proteins are interacting in their environment, there is possibility that the interaction between Rfa1 and Rad24 might be indirectly mediated via yet another unknown protein. We believe that there is a direct interaction between these two proteins but we can only be sure of it after verifying their interaction in an in vitro assay with purified proteins. Also, since Rfa1-Rad24 interaction was not affected by the phosphorylation state of Rfa2 but rather disrupted by MMS treatment, it presents yet more evidence of this interaction being modulated by actions of other proteins than Rfa2. In fact there might be two separate classes of proteins whose interaction with Rfa1 is modulated in Rfa2 phosphorylation dependent and independent pathways.

3.6. References

Ben-Aroya S, Koren A, Liefshitz B, Steinlauf R, Kupiec M (2003) ELG1, a yeast gene required for genome stability, forms a complex related to replication factor C. *Proc Natl Acad Sci U S A* **100:** 9906-9911

Clerici M, Mantiero D, Guerini I, Lucchini G, Longhese MP (2008) The Yku70-Yku80 complex contributes to regulate double-strand break processing and checkpoint activation during the cell cycle. *EMBO Rep* **9:** 810-818

de Jager M, van Noort J, van Gent DC, Dekker C, Kanaar R, Wyman C (2001) Human Rad50/Mre11 is a flexible complex that can tether DNA ends. *Mol Cell* **8:** 1129-1135

de la Torre-Ruiz MA, Green CM, Lowndes NF (1998) RAD9 and RAD24 define two additive, interacting branches of the DNA damage checkpoint pathway in budding yeast normally required for Rad53 modification and activation. *EMBO J* **17:** 2687-2698

Finn K, Lowndes NF, Grenon M (2012) Eukaryotic DNA damage checkpoint activation in response to double-strand breaks. *Cell Mol Life Sci* **69:** 1447-1473

Green CM, Erdjument-Bromage H, Tempst P, Lowndes NF (2000) A novel Rad24 checkpoint protein complex closely related to replication factor C. *Curr Biol* **10:** 39-42

Jungmichel S, Stucki M (2010) MDC1: The art of keeping things in focus. *Chromosoma* **119:** 337-349

Kubota T, Nishimura K, Kanemaki MT, Donaldson AD (2013) The Elg1 Replication Factor C-like Complex Functions in PCNA Unloading during DNA Replication. *Mol Cell* **50:** 273-280

Lewis LK, Storici F, Van Komen S, Calero S, Sung P, Resnick MA (2004) Role of the nuclease activity of Saccharomyces cerevisiae Mre11 in repair of DNA double-strand breaks in mitotic cells. *Genetics* **166:** 1701-1713

Lydall D, Nikolsky Y, Bishop DK, Weinert T (1996) A meiotic recombination checkpoint controlled by mitotic checkpoint genes. *Nature* **383:** 840-843

Majka J, Burgers PM (2003) Yeast Rad17/Mec3/Ddc1: a sliding clamp for the DNA damage checkpoint. *Proc Natl Acad Sci U S A* **100:** 2249-2254

Majka J, Chung BY, Burgers PM (2004) Requirement for ATP by the DNA damage checkpoint clamp loader. *J Biol Chem* **279:** 20921-20926

Naiki T, Kondo T, Nakada D, Matsumoto K, Sugimoto K (2001) Chl12 (Ctf18) forms a novel replication factor C-related complex and functions redundantly with Rad24 in the DNA replication checkpoint pathway. *Mol Cell Biol* **21:** 5838-5845

Naiki T, Shimomura T, Kondo T, Matsumoto K, Sugimoto K (2000) Rfc5, in cooperation with rad24, controls DNA damage checkpoints throughout the cell cycle in Saccharomyces cerevisiae. *Mol Cell Biol* **20:** 5888-5896

Rogakou EP, Pilch DR, Orr AH, Ivanova VS, Bonner WM (1998) DNA double-stranded breaks induce histone H2AX phosphorylation on serine 139. *J Biol Chem* **273:** 5858-5868

Treuner K, Okuyama A, Knippers R, Fackelmayer FO (1999) Hyperphosphorylation of replication protein A middle subunit (RPA32) in apoptosis. *Nucleic Acids Res* **27:** 1499-1504

Walker JR, Corpina RA, Goldberg J (2001) Structure of the Ku heterodimer bound to DNA and its implications for double-strand break repair. *Nature* **412:** 607-614

Warbrick E (2000) The puzzle of PCNA's many partners. *Bioessays* **22:** 997-1006

Weinert TA, Kiser GL, Hartwell LH (1994) Mitotic checkpoint genes in budding yeast and the dependence of mitosis on DNA replication and repair. *Genes Dev* **8:** 652-665

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

Zou L, Liu D, Elledge SJ (2003) Replication protein A-mediated recruitment and activation of Rad17 complexes. *Proc Natl Acad Sci U S A* **100:** 13827-13832

CHAPTER 4: DISCUSSION AND FUTURE DIRECTION

DNA metabolism is a central process in the life cycle of all organisms. While DNA is neatly organized in the iconic double helix structure, much of its metabolism requires it to unwind and dissociate into single strands. As opposed to the dsDNA structure, ssDNA is much more susceptible to nuclease damage and hairpin formations. Protection of DNA in its single stranded form and mobilization of proteins processing it for effective replication, repair and recombination is coordinated by Replication Protein A. It is able to take part in such diverse pathways by not only interacting with DNA but also because of its ability to interact with various other proteins in the cell. These specific RPA-protein interactions provide RPA the flexibility to take part in diverse pathways. To this date over 175 unique RPA-protein interactions have been identified mostly by global screening methods such as Affinity capture coupled with mass spectroscopy or RNA. Such diversity in protein interaction by any single protein is a rare occurrence in cells. While the three subunits of RPA may offer independent sites for different protein interactions, RPA's regulation by various post-translational modifications also plays a crucial role in guiding RPA-protein interactions. Studies have shown hyper phosphorylation on RPA2 N-terminus can modulate protein interactions between DBD-B of RPA1 subunits (Liu et al, 2005). There are many hypotheses to how post-translational modification of RFA causes it to modulate protein interactions. While we may not have figure out the underlying mechanism we know that it is an important event that dictates RPAs function in the cell.

Keeping this mind, we set out to discover novel RPA-protein interactions in cell using the yeast two-hybrid assay. It is a valuable tool to identify *in vivo* protein interactions and also provides a nice platform for isolating specific regions on proteins that are important for interaction. Taking the experimental paradigm one step further, we wanted to learn if these interactions were dependent on the phosphorylation state of RPA2. Studies show that other types of post-translational modification can also affect RPA function in the cell. For example, SUMOylation of Rpa70 by SUMO2/3 which facilitates recruitment of Rad51 to damaged DNA foci and modulates its role in Homologous Recombination (Dou et al, 2010). Nevertheless, phosphorylation of RPA2 has been studied most extensively to date and is also best characterized and hence we decided to test this particular modification and its ramification on RPA-protein interactions for our experiment. We were excited to identify numerous proteins through our two-

hybrid assay some that were previously characterized and other that were novel. We also found a subset of these protein interactions were dependent of phosphorylation state of Rfa2 subunit. Generally speaking Rfa2-Asp mutant seemed to disrupt protein interactions with RFA mediated via the N-terminal of Rfa1.

Since we used Rfa1-FAB as bait in our two-hybrid screen these protein interactions and effects of phosphorylation on them are assumed to be occurring independent of the whole RPA complex. This observation is novel to our knowledge and highlights the ability of RFA to function at the individual subunit level. It would be interesting to see if proteins that interact with RFA via the Rfa2 subunit could also do so without the formation of a proper RFA complex. Our results also suggest that Rfa2 phosphorylation in one RFA complex can affect the interaction between Rfa1 and proteins outside of that particular complex. We may be looking at a cooperative model where multiple RFA complexes can work together to modulate protein interactions. Another hypothesis for why the presence of Rfa2-Asp disrupts interaction between Rfa1-FAB and HITS is that Rfa2-Asp might have a higher binding affinity for these HITS than Rfa1-FAB itself. In this case all the free proteins will effectively be bound to Rfa2-Asp and 'disrupts' their interaction with Rfa1-FAB by essentially depleting any free proteins. Furthermore we looked into individual interactions that we isolated to get a greater understanding of how these interactions are regulated and their importance in the cell. Among the first two interactions that we looked at, Rfa1-FAB interaction with Rad24 was not affected by the Rfa2-Asp mutant while Rfa1-FAB interaction with YNR071C and APC4 was affected by phospho-mimetic mutation of Rfa2 (data not shown). Interestingly, in sharp contrast to this observation we also mapped those same two interactions to completely different regions of Rfa1. Rad24 binding was mediated by DBD-F on Rfa1 while APC4 and YNR071C binding was mediated by DBD-B. While these results are preliminary and needs much work, we can start to picture a model in which Rfa2 phosphorylation specifically affects RPA-proteins interactions that are mediated specifically by DBD-B while proteins binding to Rfa1 via its N-terminal DBD-F might not be affected by it. Since we have a decent pool of interactions isolated from the two-hybrid screen that are affected by the Rfa2-Asp mutant, we can test each one of them using similar methods as before to help us determine if this model holds true in all cases. If this model is in fact validated, then we have a novel mechanistic insight on how the phosphorylation of Rfa2 affects its interactions with other proteins, especially those important in DNA damage signaling and repair.

As mentioned earlier our hypothesis for why presence of Rfa2-Asp disrupts interaction between Rfa1-FAB and HITS is that Rfa2-Asp can bind to these HITS with a higher binding affinity than Rfa1-FAB, leaving no free HITS in the cell to bind Rfa1-FAB. To test this hypothesis, we can retest these interactions with just Rfa1-WT and HITS as bait and prey in our two hybrid assay. Since Rfa1-FAB is not in a complex with Rfa2-Asp it is hard to tell with certainty if Rfa2-Asp truly disrupts proteins interaction. But if we used Rfa1-WT as our bait, we can be sure that the Rfa2-Asp is in the RFA complex and if we see a direct interaction between Rfa1-WT and these HITS with this set up, then we can be sure that Rfa2-Asp is indeed binding these HITS with higher affinity that Rfa1-FAB alone. And if we see similar results with Rfa1-WT as we did with Rfa1-FAB then we can be sure that Rfa2-Asp is indeed disrupting protein interaction between Rfa1-FAB and these HITS. Since all these interactions are happening inside the cell, various cellular factors can affect these interactions. We can look at cell growth and assume that our bait and prey interacts with each other, but in fact these interactions might be mediated by other unknown proteins in the cell causing an indirect interaction. In that case to fully understand the mechanism underlying these interactions we will need to identify these unknown proteins as well. These interactions can be tested in vitro with purified proteins to back up our two hybrid data and show direct interaction. And if we don't see these interactions in tin vitro experiments, then we can go back and strategically knock out genes in yeast that might be mediating these interactions and look perform the two hybrid screen on those strains. If the interactions are alleviated in these KO strains, then we know that it was important for the protein interaction between Rfa1 and HITS. These interactions can again be constituted in an in vitro experiment and validated to back out the two hybrid data.

Although protein interactions with RFA have been discovered steadily over the years, very little is known about how all these interaction are modulated during different damage conditions, pathways, and cell cycle phases. So it is of critical important to understand how these protein interactions are coordinated at a systems level and understand the importance of each one for DNA metabolism. Although the two-hybrid assay has helped us narrow down proteins that interact with RFA, we need more sophisticated tools to understand their mechanistic details. The next step for this project should be to express of these proteins *in vitro* and verifying direct interaction between them with co-immunoprecipitation. Also, creating yeast deletion strains of these genes either made individually or

couple together and testing their effects on cell survival, cell cycle progression, and damage response will also greatly enhance the role of these individual interactions. The complexity of the RFA structure, its promiscuous nature in the cell and a highly sophisticated DNA and protein binding models have perplexed scientists for over two decades now. As more interactions between RFA and proteins are continued to be discovered and characterized, it will inevitably resolve some of the questions behind its intricate nature but since it sits at the epicenter of a complicated DNA metabolic network, we can only speculate when the complete picture will be revealed.

4.1. References

Dou H, Huang C, Singh M, Carpenter PB, Yeh ET (2010) Regulation of DNA repair through deSUMOylation and SUMOylation of replication protein A complex. *Mol Cell* **39:** 333-345

Liu Y, Kvaratskhelia M, Hess S, Qu Y, Zou Y (2005) Modulation of replication protein A function by its hyperphosphorylation-induced conformational change involving DNA binding domain B. *J Biol Chem* **280:** 32775-32783