Spring 5-5-2018

Towards the Identification of the Molecular Mechanism Responsible for RPA:RAD52 Complex Formation

Lucas Struble
University of Nebraska Medical Center

Towards the Identification of the Molecular Mechanism Responsible for RPA:RAD52 Complex Formation

By

Lucas Struble

A DISSERTATION

Presented to the Faculty of

the University of Nebraska Graduate College

in Partial Fulfillment of the Requirements

for the Degree of Doctor of Philosophy

Molecular Biology and Biochemistry
Graduate Program

Under the Supervision of Professor Gloria E. O. Borgstahl

University of Nebraska Medical Center
Omaha, Nebraska

April, 2018

Supervisory Committee

Tadayoshi Bessho, Ph.D.

Pi-Wan Cheng, Ph.D.

Amarnath Natarajan, Ph.D.

Towards the Identification of the Molecular Mechanism Responsible for RPA:RAD52 Complex Formation

Lucas Struble, Ph.D.

University of Nebraska, 2018

Supervisor: Gloria E.O. Borgstahl, Ph.D.

Human cells are routinely exposed to DNA-damaging conditions, from both external sources like ionizing radiation and internal sources like normal oxidative metabolism. Damage in the form of double strand breaks (DSBs) is especially problematic. DSBs occurring outside of replication forks can be repaired through two forms of homologous recombination. The first of these is genetic conversion involving either RPA, BRCA1, PALB2, BRCA2, and RAD51, or RPA, RAD52, RAD51, and other unknown factors. The second is single strand annealing involving RPA and RAD52. Familial breast cancers, among numerous others, are characterized by homozygous pathological mutations in the BRCA2 pathway and must therefore rely on the RAD52 pathway for remediation of DSBs. Inhibiting the interaction between RPA and RAD52 should therefore selectively terminate such cancer cells without harming healthy cells. Structural information regarding RAD52 and the phosphorylation state of RPA during active DNA repair must be elucidated to achieve this. Initial structural data for RAD52 was acquired using small angle X-ray scattering (SAXS). Using the available RAD52(1- 212) crystal structure we were able to estimate the orientation of the RAD52(1-303) SAXS structure. The application ITASSER allowed for the modeling of one of the RAD52(213-303) sections which is outside of the RAD52(1-212) crystal structure, and which includes the RPA binding domain. Utilizing available information about known DSB-induced

phosphorylation sites of RPA, paired with data from Phosida and PhosphositePlus, eleven candidate sites were selected for structural and DNA binding studies. Phosphorylation was mimicked by mutation of candidate sites to glutamic acid, and 6 of the combinations tested retained heterotrimer stability. Phosphomimetic mutations to the RPA70 subunit decreased DNA binding affinity. Identification of these stable phosphomimetics with confirmed DNA binding activity provides tools for experiments delving into the activities of RPA functioning in BIRDSB repair, these tools will also be used for structural experiments involving the binding of RPA to DSB repair proteins, including the SAXS compatible RAD52(1-303).

Dedication

I would first like to thank my mentor, Dr. Gloria E.O. Borgstahl. Every in the lab was a day I learned something new. Her willingness to talk about everything from problems with experiments to everyday life events has made this journey interesting. On the day I interviewed to do a rotation in her lab I sat down in her office and explained that I was “deaf, blind, and I just came from dumping a thermos of water over the keyboard of my last interviewer.” She classily replied with “Huh, never heard that as an introduction before, how much do you know about structural biology?” I replied that I knew practically nothing of it. She told me “as long as you can take notes and learn, it will be fine, go and get a notebook from Carol.” From that point on I feel like I was welcomed into the amazing, eccentric, and very often confusing Borgstahl laboratory. Every day spent in this lab has been a treasure. The people I have meet, the things I have done, the things I have learned, all of it was thanks to Gloria. We even developed a friendly rivalry of “who had the most ridiculous family drama” that ended up being almost therapeutic. Gloria, thank you again. All of this was possible because of you, and it is something that I shall always cherish. I consider you a good friend, and the mentor that I needed most.

Next I would like to thank my committee members, Dr. Tadayoshi Bessho, Dr. Pi-Wan Cheng, and Dr. Amarnath Natarajan. They were all great teachers of my classes, as well as amazing sources of advice as I progressed through my degree. As each came from a different scientific background, each challenged me in a way that forced me to move outside my comfort zone and become a better scientist. Thank you all.

I would like to thank all the members of the Borgstahl lab, but a few deserve special recognition. First of these is Carol Kolar. At a time when I was most nervous about my place in the world she helped my feel welcomed and taught as much as she could about

everything she was able to. She was the glue that held all of us students together. Whenever there was a problem of some kind, the solution was usually a quick conversation with Carol followed by a 5-10 minute fix. Carols sense of humor was very close to my own, which is still shocking to me. Pretty much all of my experiments needed Carols help at one time or another, so I am indebted to her for all of the progress made because of her help. I am honored to be one of the last students that will graduate under her sight, and will miss her terribly as she moves across the country. Carol, you have been a true friend, and your friendship means more to me than I can express. Thank you for everything. Jeff Lovelace has been the one in the lab who helped me keep things in perspective, as well as helped keep all of the equipment running (no matter what we students accidently did to it). His help was invaluable in all of my SAXS work, as well as always being quick with a joke whenever I was feeling down. Jeff, thank you. I think the lab may have burned down without you. Concerning former students, I would like to thank Kerry Brader. When I first joined the lab she was the one who showed be the ropes and helped me learn what being a student was about at UNMC. We exchanged audiobooks often to keep awake during some of the more prolonged and dull experiments. Kerry, thank you for all of your help, I know it was a good thing but I was still sad to see you graduate. I would like to also thank Mona Al-Mugotir, as she has been a close friend through my time in this lab. We ended up helping each other out on projects so much that it is difficult to say which projects would have succeeded or failed without the other ones help. Despite her strange insistence that children are a good thing, I still consider her a decent person. Thank you, Mona. Jahaun Azadmanesh joined the lab more recently, but quickly became someone who I looked forward interacting with each day. It was always nice to interact with someone who was so genuine in everything that they did. Thank you, Jahaun, it was nice to finally find someone else who considers pushups a necessary part

of research. Will Lutz has been a member of the lab for about a year as I write this, and in that time he had become a person that I trust and look forward to working with. He has spent more time than anyone but Gloria and Carol in helping me with my projects. Thank you, Will.

Most of all, I wish to thank my friends and family. You have kept me sane these many years, and I know that was no small task. Thank you for putting up with the ever-shifting schedule, the random, long, stressful nights, the multi-day freak outs, and the mood swings that came from this. To my parents, thank you for instilling in me a love of science and exploring the unknown. You made me the person I am today. To my wife, I am sorry that you had to deal with so much more of the worst aspects of graduate school than anyone else in my life. If it hadn't been for your belief that I could do this, I would not have been able to succeed. Your strength let me get through this, and for that I am eternally grateful. Thank you all.

Finally, I would like to thank the department of education for their assistance in my graduate training through their GAANN program as well as NASA. My years of graduate training would not have been possible without your help.

Table of Contents

List of Figures

List of Tables

Abbreviations

A_{280}	Absorbance at 280 nm
AP	Apurinic/apyrimidinic
ATM	Ataxia-telangiectasia-mutated
ATR	ATM and Rad3-related
BER	Base excision repair
BIR	Break induced repair
BLM	Bloom syndrome RecQ like helicase
BME	2-mercaptoethanol
BRCA1	Breast cancer type 1 susceptibility protein
BRCA2	Breast cancer type 2 susceptibility protein
C	Carboxyl
CAK	CDK-activating kinase
Cas9	CRISPR associated protein 9
CDK	Cyclin dependent kinase
CETN2	Centrin 2
CRISPR	Clustered regularly interspaced short palindromic repeats

CSA	WD repeat protein
CSB	Cockayne syndrome protein
CtIP	CtBP-interacting protein (Cofactor of the Mre11-RAD50-NBS1 complex)
CTD	Carboxyl terminal domain
CV	Column volume
DBD	DNA binding domain
DDB1	Damage specific DNA binding protein 1
DDB2	Damage specific DNA binding protein 2
DDR	DNA damage response
DI	Deionized water
DNA	Deoxyribose nucleic acid
DNA2	DNA replication helicase/nuclease 2
DNA-PK	DNA-dependent protein kinase
dRP	Deoxyribophosphate
DSB	Double strand break
DSBR	DSB repair
EDTA	Ethylenediaminetetraacetic acid

ERCC1	Excision repair cross-complementing 1
EXO1	Exonuclease 1
GC	Gene conversion
GG-NER	Global genome NER
HI	HEPES inositol
HR	Homologous recombination
IPTG	Isopropyl β-D-1-thiogalactopyranoside
LIG1	DNA ligase 1
LIG3	DNA ligase 3
LB	Lysogeny broth
MCS1	Multiple cloning site 1
MCS2	Multiple cloning site 2
MLH1	MutL homolog 1
MMR	Mismatch repair
MRE11	Mitotic recombination 11
MRN	MRE11, RAD50, NBS1
MSH2	MutS homolog 2
MSH3	MutS homolog 3

MSH6	MutS homolog 6
MutLα	Heterodimer of MLH1 and PMS2
MutSα	Heterodimer of MSH2 and MSH6
MutSβ	Heterodimer of MSH2 and MSH3
MW	Molecular weight
MWCO	Molecular weight cut-off
NBS1	Nijmegen breakage syndrome 1
NER	Nucleotide excision repair
NHEJ	Non-homologous end joining
NMR	Nuclear magnetic resonance
OB-fold	Oligosaccharide binding fold
OD600	Optical density at 600 nm
PALB2	Partner and localizer of BRCA2
PCNA	Proliferating cell nuclear antigen
PDF	Pairwise distribution function
PIKK	Phospatidylinositol 3-kinase related kinase
PMS2	Post-meiotic segregation protein 1
POLB	DNA polymerase β
POLD	DNA polymerase δ

PVDF	Polyvinylidene fluoride
RAD23B	RAD23 homologue B
RFC	Replication factor C
R_g	Radius of gyration
RNA Pol II	RNA polymerase II
RPA	Replication protein A
SAXS	Small angle X-ray scattering
SDSA	Synthesis dependent strand annealing
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
SEC	Size exclusion chromatography
SEC-MALS	SEC with multi-angle light scattering
shDNA	Short hairpin DNA
SPR	Surface plasmon resonance
SSA	Single strand annealing
TC-NER	Transcription-coupled NER
TFIIH	Transcription initiation factor II H
UNG	Uracil DNA glycosylase

USP7	Ubiquitin-specific-processing protease 7
UTR	Untranslated region
UV-DDB	Ultraviolet radiation-DNA damage-binding protein complex, composed of DDB-1 and DDB2
UVSSA	UV-stimulated scaffold protein A
WD	Tryptophan-aspartic acid
wHLH	Winged helix-loop-helix
XPA	Xeroderma pigmentosum group A-complementing protein
XPB	Xeroderma pigmentosum group B-complementing protein
XPC	Xeroderma pigmentosum group C-complementing protein
XPD	Xeroderma pigmentosum group D-complementing protein
XPF	Xeroderma pigmentosum group F-complementing protein
XPF-ERCC1	The XPF-ERCC1 complex

XPG	Xeroderma pigmentosum group G-complementing protein
XRCC1	X-ray cross complementing 1

Chapter I: Introduction

1.1 DNA Damage and Repair

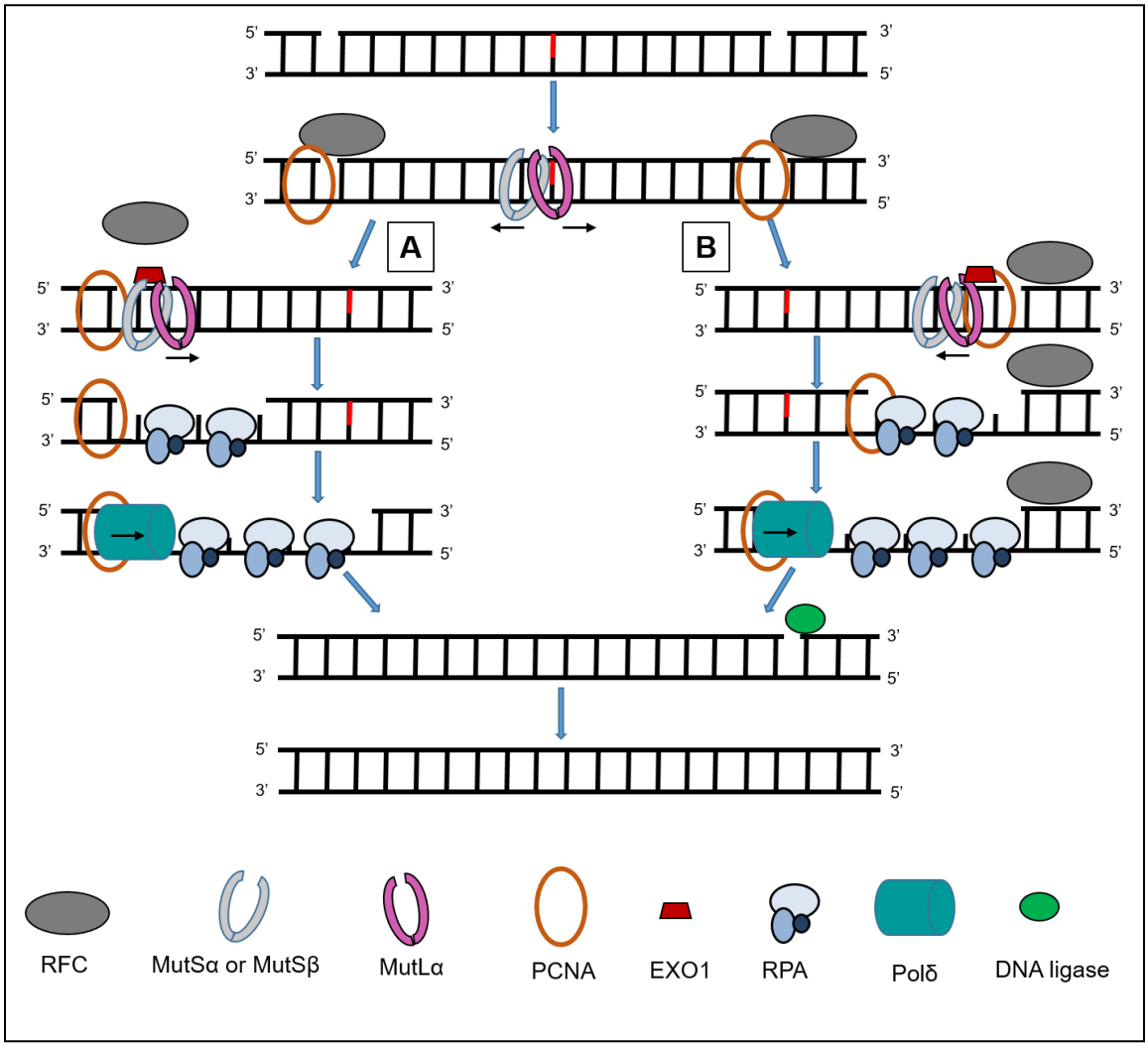
It is impossible for an organism to live and thrive without developing ways to protect and repair its genome from deoxyribose nucleic acid (DNA) damage. Factors that induce DNA damage come from a variety of sources with a multitude of them being unavoidable. Exogenous sources include environmental toxins such as byproducts of tobacco usage, environmental radiation sources such as radon gas, and ionizing radiation such as ultraviolet light and X-rays. If an organism were able to somehow avoid all exogenous factors, endogenous factors can still threaten cells with DNA damage. Many of the factors are intrinsically linked to how cells function. These include replication errors, free radical generation through metabolism, and alkylating agents [1]. All of these factors combined result in a genome that is constantly under attack, and without a way to detect and repair this damage the cells will die or turn cancerous.

This damage is mitigated through the activation of the DNA damage response (DDR) and can result in the activation of multiple DNA repair pathways. The pathway chosen depends on the type of damage and phase of the cell cycle. Damage to a single strand of DNA can be repaired through either the mismatch repair pathway (MMR), the base excision repair pathway (BER), or the nucleotide excision repair pathway (NER).

The MMR pathway is used if an incorrect nucleotide is used during DNA replication or a nucleobase becomes damaged, for instance through deamination (**Figure 1.1**) [2]. The mismatched bases are recognized and bound by the sliding clamp proteins MutSα or MutSβ. These will recruit the sliding clamp protein MutLα. This pair of sliding clamps will shift either in the 5’ or 3’ direction from the mismatch site. If the clamps move 5’ of the site, they will encounter the protein replication factor C (RFC). RFC loads the homotrimeric proliferating cell nuclear antigen (PCNA) onto the 3’ end of a nick or Okazaki

Figure 1.1 MMR pathway overview

The MMR pathway begins when MutSα or MutSβ binds to the mismatch site (red line) and recruits MutLα. These proteins undergo a conformational shift allowing the proteins to act as sliding clamps and move away from the mismatch site. A) Clamps migrating in the 5' direction will encounter RFC bound to the 5' terminus of the break. RFC is displaced and EXO1 is loaded onto the DNA. This endonuclease will then degrade the DNA in a 3' direction with the resulting gap stabilized by RPA. The removal of the mismatch halts the stimulation of EXO1 by MutSα and begins the inhibition of EXO1 by MutLα. POLD binds PCNA at the 5' end of the gap and fills the gap. The remaining nick is repaired by LIG1. B) Clamps migrating in the 3' direction come upon PCNA bound to the 3' terminus of the break. EXO1 is recruited and loaded onto the DNA, possibly multiple times, until the region between the mismatch and the break in degraded. RFC bound at the break will prevent EXO1 from degrading further away from the mismatch in a 3' direction. Degradation of the mismatch results in the inhibition of EXO1 by RPA and MutLα. PCNA will recruit POLD and this complex will fill in the gap. The remaining nick is repaired by LIG1.


5'
3'
A
B
RFC
MutSα or MutSβ
MutLα
PCNA
EXO1
RPA
Polδ
DNA ligase

fragment, becomes displaced, and exonuclease 1 (EXO1) is loaded onto the DNA. EXO1 is stimulated by MutSα and degrades the DNA in a 3’ direction until it degrades the mismatch. This removal of the mismatch ends the stimulation by MutSα and initiates inhibition of EXO1 by MutLα. DNA polymerase δ (POLD) binds to PCNA and fills in the gap. The remaining nick is repaired by DNA ligase 1 (LIG1). Should the sliding clamps shift 3’ of the mismatch site encounter PCNA that was recruited by RCF to the 3’ terminus of a nick, and EXO1 will be loaded onto the DNA. RCF prevents EXO1 from degrading further from the mismatch in a 3’ direction. EXO1 may need to be loaded onto the DNA multiple times to complete the degradation of the DNA to the mismatch site. Inhibition of EXO1 happens as described previously, and PCNA recruits POLD. POLD fills in the gap, and the remaining nick is repaired by a LIG1.

The BER pathway is used to repair a damaged base or apurinic/apyrimidinic (AP) site (**Figure 1.2**) [3]. The pathway begins with a glycosylase recognizing and binding to the damaged base and excising it to create an AP site. An AP nuclease will then make a nick 5’ of the AP site. A polymerase, such as polymerase β (POLB), binds and fills in the gap caused by the nick and missing base. In long patch BER the polymerase will displace bases beyond the AP site, and flap endonuclease 1 (FEN1) then removes the displaced ssDNA flap. The remaining nick is repaired by LIG1. In short patch NER, POLB binds, fills in the single nucleotide gap, and removes the 5’-terminal deoxyribophosphate (dRP). The remaining nick is repaired by DNA ligase 3 (LIG3).

The NER pathway is used when a damaged DNA lesion is distorting the DNA helix, such as a DNA adduct or thymidine dimer (**Figure 1.3**) [4]. The early steps of this pathway can occur in two ways, with the first being global genome nucleotide excision repair (GG-NER). In this form of the pathway the protein complex xeroderma pigmentosum group C-complementing protein (XPC), RAD23 homologue B (RAD23B),

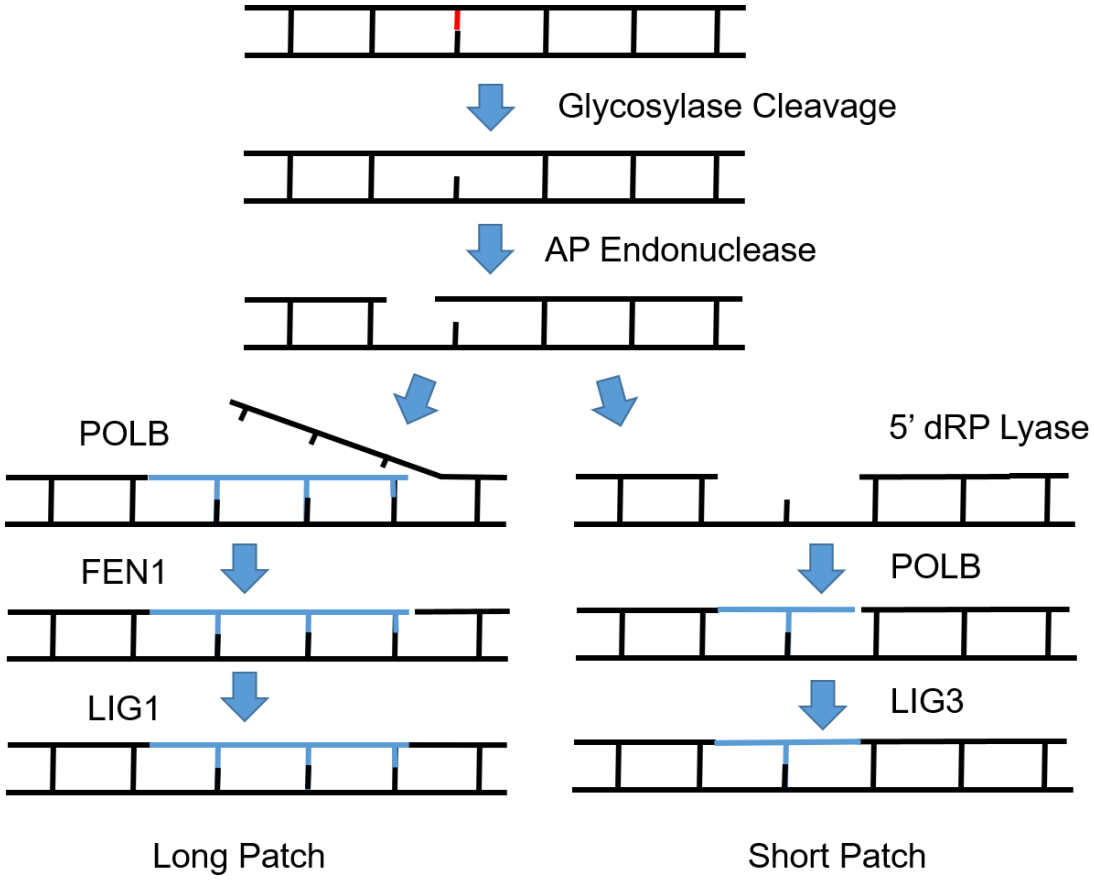
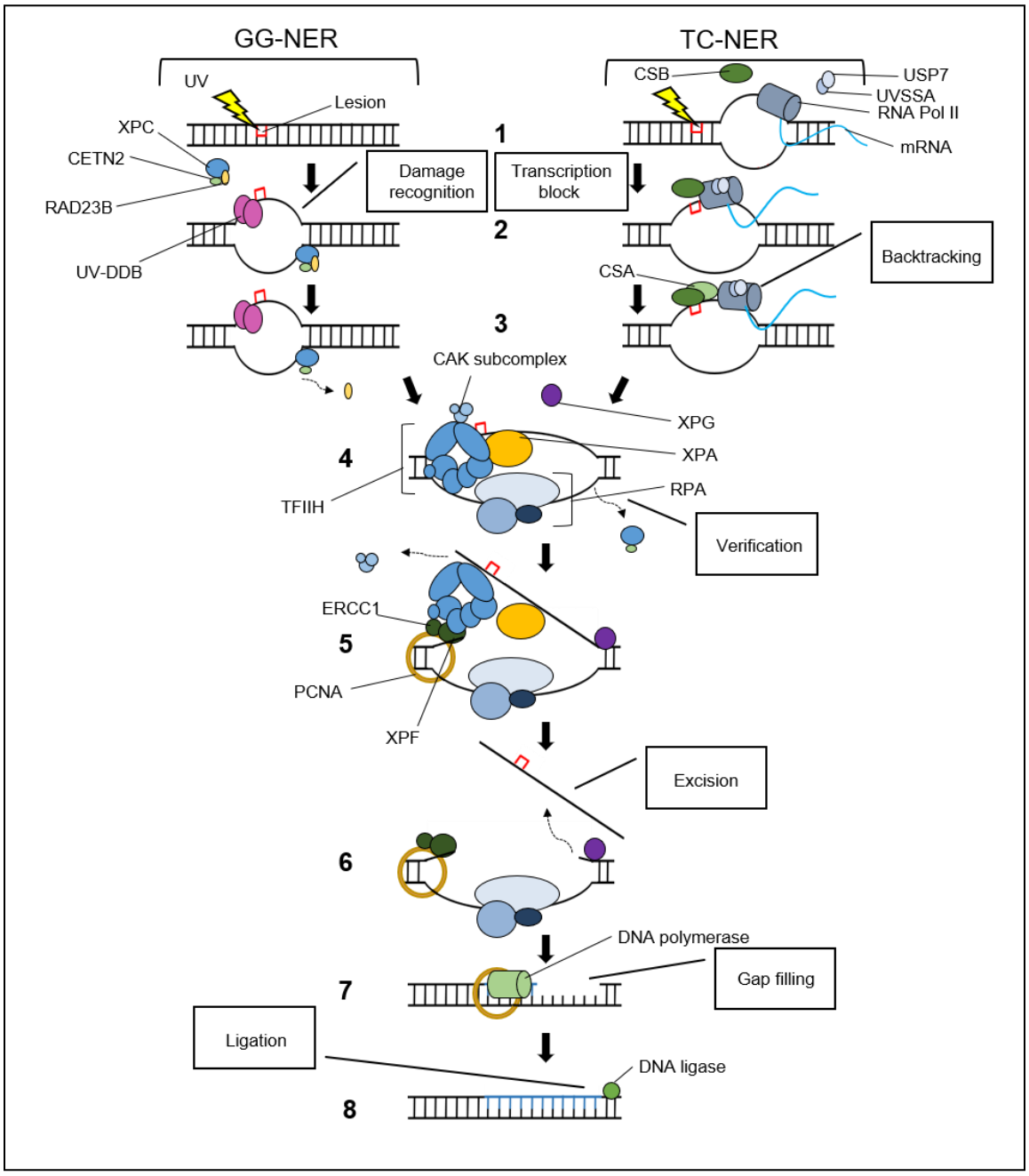


Figure 1.2 BER pathway overview

A damaged base is recognized and acted upon by a glycosylase to create an AP site. The AP endonuclease then makes an incision 5' of the AP site. In long patch BER (left) a DNA polymerase such as POLB fills in the gap and displaces a section of ssDNA longer than a single base. FEN1 cleaves the ssDNA flap, and LIG1 repairs the nick. In short patch BER (right) a polymerase such as POLB will fill in the missing nucleotide and cleave the remaining 5' terminal dRP. The nick is repaired by LIG3.

Figure 1.3 NER pathway overview

1-3) In GG-NER (left) the XPC, RAD52B, and CETN2 complex and UV-DDB complex detect lesions which distort the DNA. When that form of damage is detected RAD23B leaves the complex. In TC-NER (right) the DNA damage is detected by the stalling during elongation of RNA Pol II bound with UVSSA, USP7, and CSB. Once RNA Pol II stalls at the lesion, CSB forms a complex with CSA, possibly inducing reverse translocation of RNA Pol II. 4) TFIIH binds to the DNA lesion along with XPG. TFIIH uses its helicase activity to further unwind the DNA around the lesion. The TFIIH subunits XPD and XPB verify the presence of the lesion. RPA then binds the undamaged strand of DNA and XPA is recruited. 5) CAK dissociates from TFIIH. XPF-ERCC1 is recruited by XPA and positioned by RPA, where it then makes a cut 5' from the lesion. XPF-ERCC1 loads PCNA onto the DNA strand. 6) XPG cuts 3' of the lesion site resulting in the release of the damaged DNA segment. 7) PCNA recruits an appropriate DNA polymerase for the gap filling reaction. 8) The new section of DNA is ligated by either LIG1 or LIG3.


GG-NER
TC-NER
UV
Lesion
XPC
CETN2
RAD23B
UV-DDB
Damage recognition
1
Transcription block
CSB
USP7
UVSSA
RNA Pol II
mRNA
2
Backtracking
CSA
3
CAK subcomplex
XPG
4
XPA
TFIIH
RPA
Verification
ERCC1
5
PCNA
XPF
Excision
6
DNA polymerase
7
Gap filling
Ligation
DNA ligase
8

and centrin 2 (CETN2) as well as the ultraviolet radiation-DNA damage-binding protein (UV-DDB) complex detect lesions which distort the DNA helix. When a damaged site is found, the complexes will bind and RAD23B will dissociate. In the second form of the pathway, known as transcription-coupled NER (TC-NER), damage is recognized by the stalling of RNA polymerase II (RNA Pol II) at a lesion during transcription elongation. The proteins UV-stimulated scaffold protein A (UVSSA), ubiquitin-specific-processing protease 7 (USP7), and Cockayne syndrome protein (CSB) all transiently interact with RNA Pol II. Once RNA Pol II stalls at the DNA damage CSB forms a complex with the WD repeat protein (CSA), possibly inducing reverse translocation of RNA Pol II. Both of the forms of the pathway converge at this point, with transcription initiation factor II H (TFIIH) binding to the DNA damage. The xeroderma pigmentosum group G-complementing protein (XPG) will bind to the repair complex, and the CDK-activating kinase (CAK) will dissociate from TFIIH. TFIIH will use its helicase activity to further unwind the DNA around the lesion. The TFIIH subunits XPD and XPB verify the presence of the lesion. RPA will then bind the undamaged strand of DNA and xeroderma pigmentosum group A-complementing protein (XPA) is recruited. The protein complex of xeroderma pigmentosum group F-complementing protein (XPF) and excision repair cross-complementing (ERCC1) (XPF-ERCC1) is recruited by XPA and positioned by RPA, where it then makes a cut 5’ from the lesion. XPG then cuts 3’ of the lesion site resulting in the release of excision of the damaged DNA segment. PCNA is loaded onto the DNA stand by XPF-ERCC1 and recruits an appropriate DNA polymerase for the gap filling reaction. The new section of DNA is ligated by either LIG1 or LIG3.

If the damage to a single strand of DNA is exacerbated or if stronger DNA damaging conditions occur then the most dangerous form of DNA damage, known as a double strand break (DSB), can ensue [5, 6]. This form of DNA damage impairs both DNA

strands simultaneously, preventing the use of the other strand as a template for repair. Despite being so dangerous, this form of DNA damage is not uncommon, as during S phase a human cell will undergo approximately 50 DSBs [7]. Repairs for DSBs occur using three pathways, the choice of which being dependent on the phase of the cell cycle among other factors [8].

The most prevalent pathway used for DSB repair is non-homologous end joining (NHEJ) (**Figure 1.4**) [9, 10]. In this pathway the ends of the DSB are processed by different complexes of polymerases, nucleases, and ligases. Regions of existing microhomology can be used to guide the repair process, but nucleotides can be added or removed on both sides of the DNA break to create homology as well. The ends of the processed DSB are then ligated together [11-18].

If the cell is in S or G2 phase it can use the homologous recombination (HR) pathway to repair DSBs (**Figure 1.5**) [19-21]. HR uses a section of homologous sequence, usually on a sister chromatid or a nearby repeated sequence, to fill in the gap left by the DSB [22]. There are three forms to this pathway: gene conversion (GC), single strand annealing (SSA), and break induced repair (BIR). GC has two forms called the double strand break repair pathway (DSBR) and the synthesis dependent strand annealing pathway (SDSA) [23]. These pathways differ in their resolution, but begin the same. On each side of a DSB the 5' strand is resected to leave a 3' overhang using the MRN complex (composed of the proteins mitotic recombination 11 (MRE11), RAD50, and nijmegen breakage syndrome 1 (NBS1)), as well as the proteins CtBP-interacting protein (CtIP), bloom syndrome RecQ like helicase (BLM), exonuclease 1 (EXO1), and DNA replication helicase /nuclease DNA2 [24, 25]. RAD51 is loaded onto the 3' strand with the help of multiple mediator proteins [23], and performs strand invasion on a homologous sequence, possibly on a sister chromatid [22, 26]. After D-loop formation, this sequence

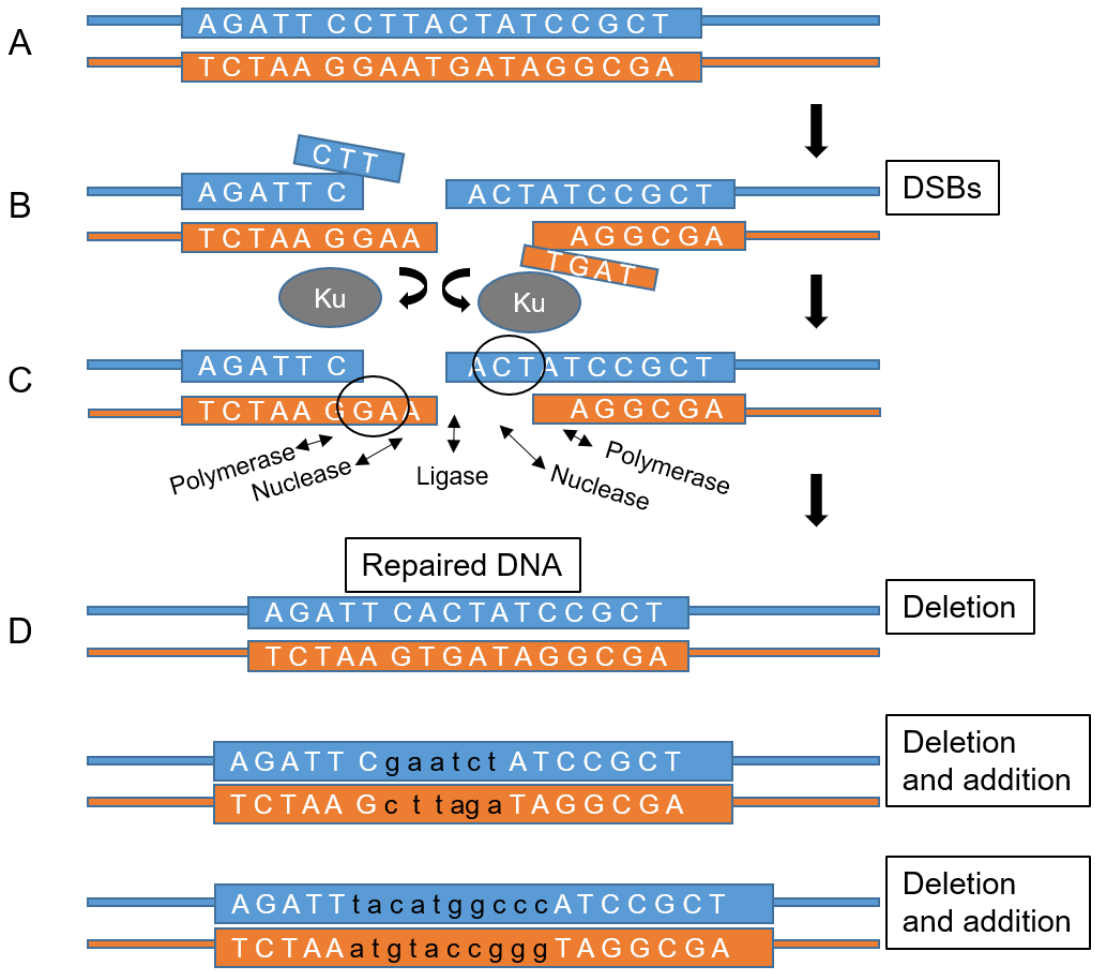
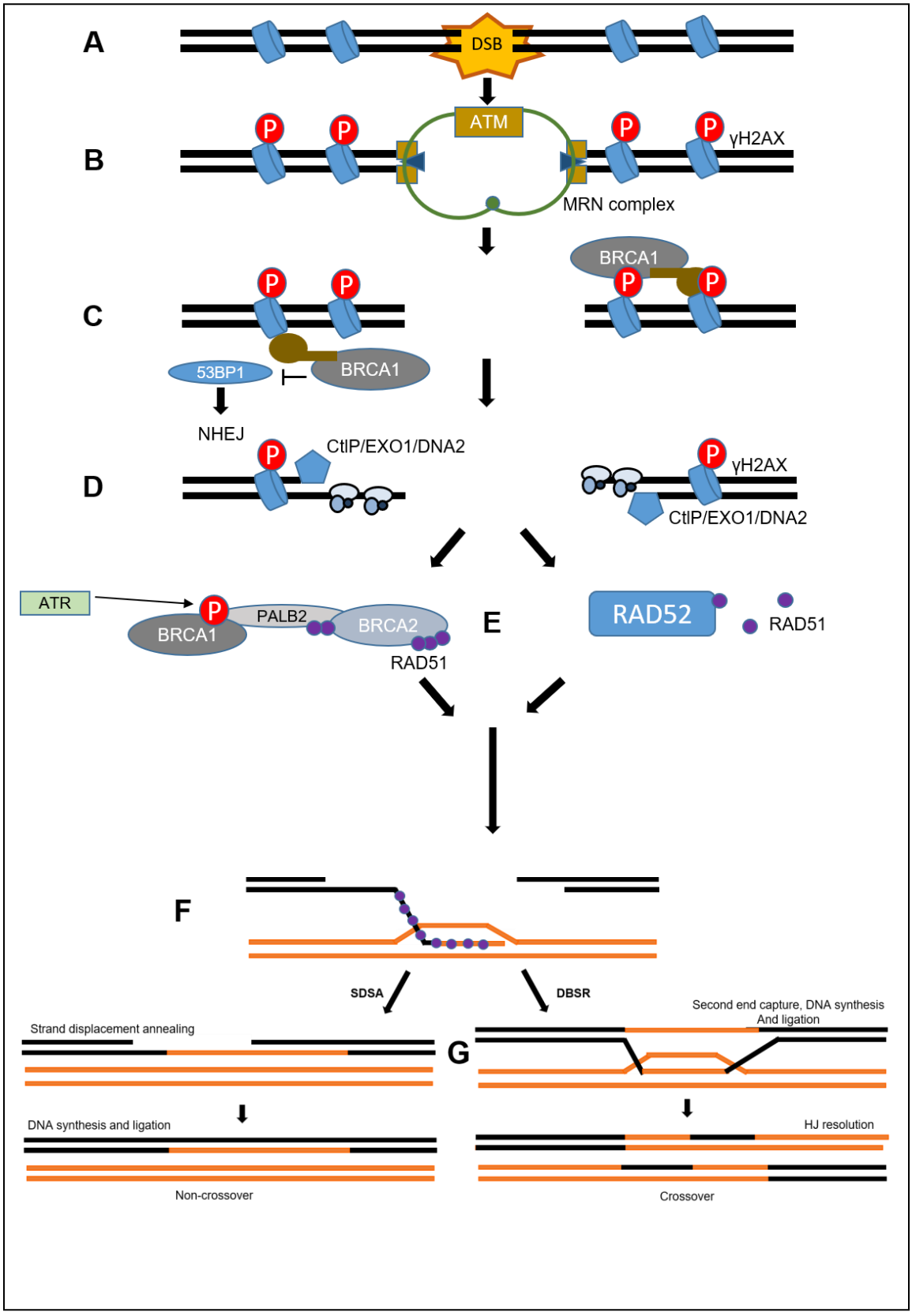


Figure 1.4 NHEJ pathway overview

An intact DNA strand (A) that undergoes a DSB (B). The KU70-KU80 heterodimer binds to the DSB ends and assists the binding of NHEJ polymerase, nuclease, and ligase complexes (C). Regions of microhomology (circled) can be used to guide the repair process (D, top). Bases can be removed or added to create sequence homology as well (D, middle and bottom, black text indicates added bases).

Figure 1.5 HR GC pathway overview

A) A DSB is formed. B) The ends of the DSB are bound by the MRN complex. ATM is recruited, which leads to the phosphorylation of γH2AX. C) BRCA1 interacts with the damage site and inhibits 53BP1. This inhibits activation of the NHEJ pathway. D) CtIP, EXO1, and DNA2 resect the 5' ends of the DSB leaving a 3' overhanging ssDNA which quickly becomes bound by RPA. E) RPA then assists in activating ATR, which in turn phosphorylates PALB2 and RPA. This phosphorylation promotes the interaction between BRCA1 and PALB2, leading to the binding and activation of BRCA2. BRCA2 replaces RPA on the ssDNA and recruits RAD51 (left). There is an alternate pathway involving the protein RAD52 replacing RPA on the ssDNA and recruiting RAD51 (right). Less is known about this pathway, and key components are still awaiting discovery. F) RAD51 forms a nucleofilament and performs strand invasion on a homologous chromosome. G) If the SDSA pathway is used, then crossover products are not produced (left). If the DBSR pathway is used, then crossover products can be formed (right).


A
DSB
B
ATM
P
γH2AX
MRN complex
C
BRCA1
53BP1
NHEJ
D
CtIP/EXO1/DNA2
ATR
PALB2
BRCA2
RAD51
E
RAD52
F
SDSA
DBSR
Strand displacement annealing
Second end capture, DNA synthesis
And ligation
G
DNA synthesis and ligation
HJ resolution
Non-crossover
Crossover

is used to fill in the missing sections of the damaged strands. There are two sets of mediators that recruit and activate the RAD51 recombinase. The primary pathway involves the interaction of replication protein A (RPA), breast cancer type 1 susceptibility protein (BRCA1), partner and localizer of BRCA2 (PALB1), and breast cancer type 2 susceptibility protein (BRCA2) to recruit RAD51, while the alternate pathway involves RPA, RAD52, and some unknown mediators to recruit RAD51.

The second HR pathway, single strand annealing (SSA), does not require alternate DNA strands to function. In this pathway the DNA is again resected on both sides of the break by exonucleases, and the 3' overhanging strands are annealed together at a section containing a repeated sequence by RAD52. The overhanging flaps are cut off, and the DNA strands are ligated [27, 28].

The final pathway is BIR, which is used to repair one sided DSBs that occur at replication forks and to assist in the maintenance of telomere length (**Figure 1.6**) [29-33]. It follows the same pattern of resection of the 5' stand to create a 3' overhang that the other two pathways use, but due to the nature of the break at the replication fork there is only a single end that seeks out a homologous sequence.

This research focuses on the interaction between RPA and RAD52, specifically as it applies to the regulation of RPA:RAD52 binding during DSBs.

1.2 RPA Activities

RPA is an essential player in nearly all forms of the DDR. The MMR pathway sees RPA binding to the nicked DNA and recruiting the MMR initiation complex, which then later promotes mismatch-promoted excision [34]. During BER, RPA will stimulate the completion of the final steps of long-patch BER and will bind with the protein uracil DNA

glycosylase (UNG), a potentially important element of the pathway [35, 36]. In the NER pathway, RPA will interact with XPF-ERCC1 and possibly XPA, both of which are

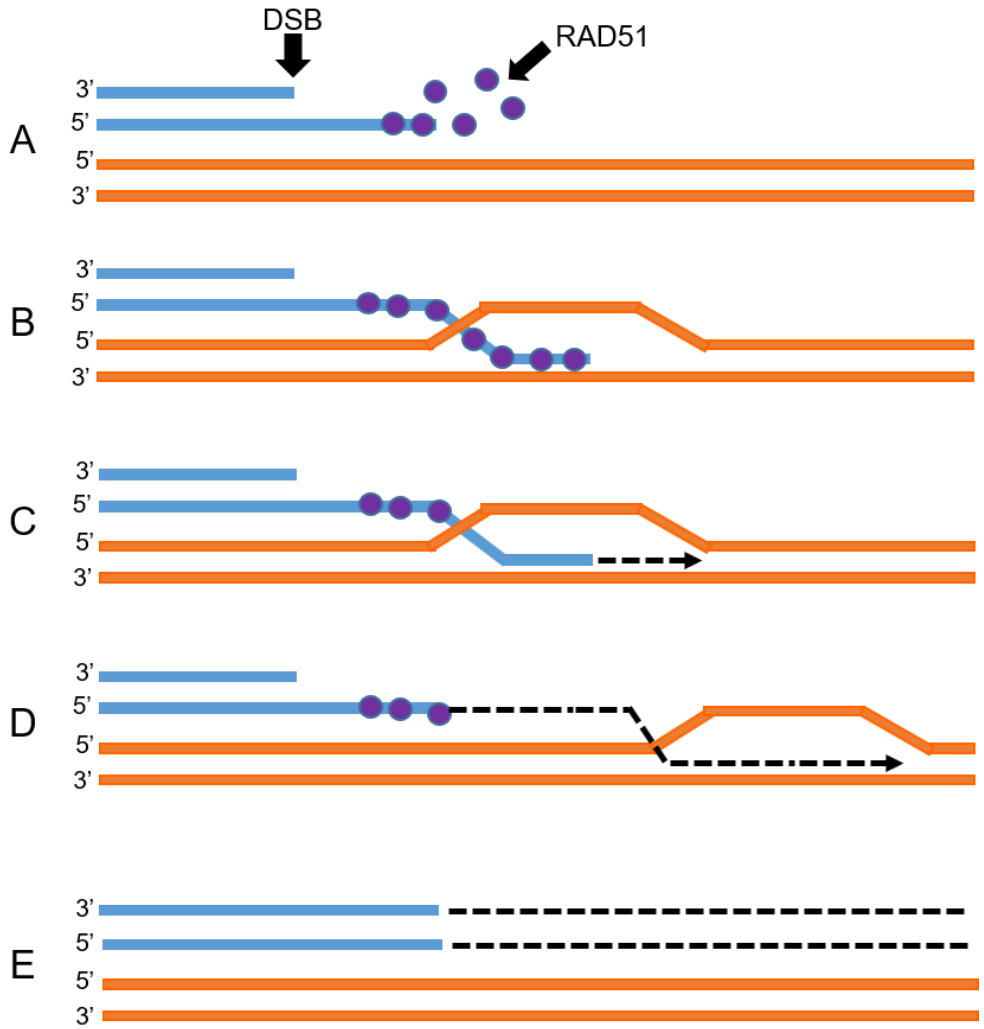


Figure 1.6 BIR pathway overview

A) The 5' end of the broken strand (blue) is resected, generating a length of 3' ssDNA. B) RAD51 forms a nucleofilament and performs strand invasion onto a homologous chromosome (orange), forming a D-loop. C) New DNA (black) is synthesized using the homologous chromosome as a template. D) The BIR replication bubble follows DNA synthesis, resulting in an extended ssDNA tail behind the bubble. E) DNA is synthesized along the extended ssDNA.

required for the repair process [37-40]. RPA will also bind and stabilize the exposed length of single stranded DNA (ssDNA) that occurs during this pathway, stimulate the rate of the NHEJ pathway, and co-localize at KU80 foci [37, 41, 42]. In HR pathways, RPA is responsible for binding and stabilizing the 3' overhang that occurs in each pathway, recruits a host of downstream mediator and DNA repair proteins, and interacts with multiple checkpoint proteins [43-49]. During HR RPA co-localizes with various HR factors at DNA damage sites, and will directly associate with RAD51 [50-52]. Of particular interest is the interaction between RPA, BRCA2, and RAD52.

1.3 RPA Structure

The unique architecture of RPA involves the presence of six oligosaccharide binding folds (OB-folds) divided amongst the heterotrimeric protein subunits (**Figure 1.7**). These subunits are named for their respective molecular weight (MW) and are therefore labeled RPA70, RPA32, and RPA14 ordered largest to smallest, respectively. RPA70 contains four OB-folds making up the DNA binding domains (DBDs) labeled A, B, C, and F domains of the protein. DBD-A and B have the highest ssDNA affinity and are thought to initiate the ssDNA binding process. DBD-F has low ssDNA affinity and is responsible for the interaction with multiple other proteins. DBD-C binds ssDNA with higher affinity than DBD-F but with lower affinity than DBD-A and B as well as interacts with the other RPA subunits to form the trimer core. There is evidence that a region of DBD-A (amino acids 169-326) can bind with the protein RAD52 [47]. RPA32 is comprised of the DBD-D, the winged helix-loop-helix (wHLH) domain responsible for further interaction with other proteins, and an N-terminal unstructured region known to become hyper-phosphorylated in circumstances of DNA damage [53, 54]. The final subunit, RPA14, is composed of a single

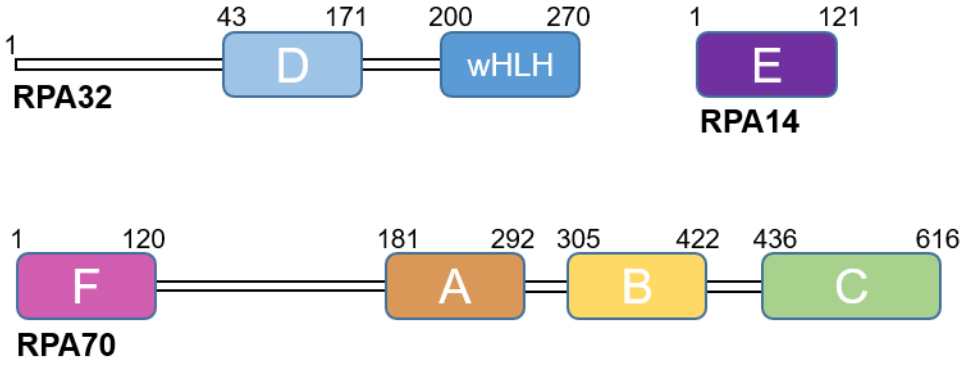


Figure 1.7 RPA domain map

The original domain map published by Dr. Marc Wold (1997). RPA is a heterotrimeric protein composed of the RPA70, RPA32, and RPA14 subunits, named for their respective MWs. Thicker regions represent structured domains, while thin white bands represent unstructured linker regions. The six OB fold domains are DBD-A (orange), DBD-B (yellow), DBD-C (green), DBD-D (blue), DBD-E (purple), and DBD-F (pink). DBDs C, D, and E interact with each other to form the trimer core. Interactions with other proteins occur through DBD-F, DBD-A, and the wHLH domain.

OB-fold domain named DBD-E and does not participate in DNA binding. This subunit functions to stabilize the heterotrimer [43, 55].

1.4 RPA Post -Translational Modifications

A form of post-translational modification of RPA is phosphorylation. This phosphorylation occurs both during the cell cycle and in response to DNA damage. Phosphorylation occurs at Ser23 and Ser29 during the S/G1 phase transition and during M phase [53, 56-60]. RPA undergoes different patterns of phosphorylation depending on the type of DNA damage [61, 62]. The hyper-phosphorylation of RPA that occurs in response to DNA damage has been well documented but the specifics of its effect on activity is poorly characterized [61-63]. Hyper-phosphorylation has been linked to the breaking up of the heterotrimer as well as steps leading to apoptosis [64, 65]. DNA damage-induced hyper-phosphorylation of RPA is currently thought to be carried out by the PIKK family proteins DNA-PK (for NHEJ), ATM, and ATR (for HR) [66-68]. In the appropriate cell cycle for HR where cells have undergone DSBs, capillary isoelectric focusing data has shown that there can be up to 14 phosphorylations on RPA [69]. Even in cells without DSBs the majority of RPA has between 2 and 4 phosphorylations. The phosphorylation state of RPA has been shown to regulate its interaction with other proteins, such as RAD52 [47, 70, 71]. Evidence has shown that there is a possibility that hyperphosphorylation of RPA can shift its role from DNA replication to DNA repair [72-74].

1.5 RAD52 Activities

RAD52 was originally identified in yeast during a screen of mutations responsible for sensitivity to radiation, and is the primary mediator of recombination in yeast [75-77].

This protein is essential in yeast, as knockout mutations show severe defects in all HR pathways, SSA, mating type switching, meiosis, integration of homologous DNA into the genome, and spore viability [78-80]. In mammals however, the protein is non-essential. A knockout of the protein in mouse embryonic stem cells showed no increase in sensitivity to radiation, but there was a decrease in levels of HR [81]. $RAD52^{-/-}$ mice showed no changes in their viability and fertility [81]. This drastic shift from the changes which occur in yeast in the same conditions was repeated in chicken DT40 cells [82]. These cells showed no change in the formation of RAD51 foci or levels of DSB repair. Because of these studies interest in RAD52 waned for a period. Research eventually surfaced, however, that showed that RAD52 did in fact have a vital role in HR. In cells with deficiencies with BRCA1, PALB2, and BRCA2, knockdown of RAD52 proved lethal [83, 84]. RAD52 is part of an alternate HR pathway, where it takes up the mediator role of BRCA2. In this alternate HR pathway RAD52 will bind to RPA, assist in displacing it from the ssDNA, and recruit RAD51 [77, 85-87]. Outside of HR, RAD52 has been shown to carry out multiple functions including the repair of stalled replication forks, strand exchange between short lengths of ssDNA and dsDNA, and reverse strand exchange between RNA and DNA [88-92]. It is also possible that RAD52 promotes BIR independent of RAD51 [29].

1.6 RAD52 Structure

RAD52 is composed of an ordered N-terminal region and a disordered C-terminal region (**Figure 1.8**). In the N-terminal region amino acids 39-80 form the DBD and amino acids 85-159 form the self-binding domain [90, 93, 94]. The unstructured carboxyl (C) -terminal domain (CTD) contains another RPA binding region mapped to the amino acids 221-280, as well as the RAD51 binding domain at 290-330 [46, 47, 95].

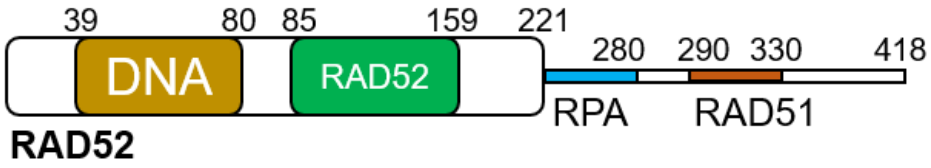


Figure 1.8 RAD52 domain map

The original domain map published by Park *et. al* (1996). The N-terminal half of RAD52 contains the DNA binding domain (orange) and the RAD52 self-binding domain (green). The C-terminal half contains the RPA binding domain (blue) and the RAD51 binding domain (brown). The thick N-terminus indicates the structured region of this protein, while the thin C-terminus indicates the unstructured region.

A nuclear magnetic resonance (NMR) structure of the RPA32 wHLH domain shows the amino acid sequence 257-274 of the RAD52 RPA binding domain forming a α-helical structure and interacting with the RPA32 wHLH domain through amino acids Arg260, Gln261, and Lys262 [54]. Electron micrographs of the full length recombinant protein show it assembled into a heptametrical (7-mer) ring [96]. The unstructured CTD has prevented crystallization of the full length protein, and the size of the protein has prevented analysis by NMR, but crystal structures exist for the ordered N-terminal regions of RAD52(1-209) and RAD52(1-212) [97, 98]. In these crystals the protein forms an undecamer (11-mer) ring with the DNA binding grooves of each subunit open on the same face of the oligomer. Amino acids 218-418 have been shown to be responsible for the formation of even larger aggregates of RAD52 [99]. With the RPA binding domain outside of the structured region of the protein, and the only structural information on the interaction between RAD52 and RPA currently being a 18 amino acid peptide sequence (RAD52(257-274)) bound to a single RPA domain, more research into the full mechanism of RPA and RAD52 interaction is needed.

1.7 RAD52 and RPA Interactions

RPA and RAD52 are confirmed to interact in multiple situations during the DDR. DSBs incurred outside of S or G2 phase can use SSA to repair the damage. The exposed ends of the DSB are bound by the MRN complex and CtIP [24, 25]. The helicase BLM is then recruited along with the endonucleases EXO1 and DNA2 [24]. These proceed to resect the 5’ strand, creating long 3’ overhanging ends on both sides of the DSB. RPA binds these exposed 3’ strands and recruits RAD52. RAD52 then displaces RPA on the ssDNA and proceeds to anneal the 3’ strands together at a region with a repeated sequence on both sides of the break. The remaining overhanging 3’ flaps are cleaved by

the XPF-ERCC1 complex, and the nicks ligated to form an intact DNA strand, albeit one with sections adjacent to the break cut away resulting in a loss of genetic information [100]. If DNA damage occurs during S or G2 phase, then RPA and RAD52 can also participate in an alternate HR pathway. After initiation, HR begins with strand resection as described previously, with the MRN complex binding the broken DSB end resecting the 5' ends around the DSB to provide exposed 3' ssDNA. RPA binds to this ssDNA and begins the recruitment of BRCA1. BRCA1 in turn recruits PALB2, which recruits/activates BRCA2. BRCA2 will bind to RPA and remove it from the DNA strand while recruiting, binding, and stimulating RAD51 recombinase. RAD51 will assemble into a helical nucleoprotein filament and perform strand invasion into the homologous sequence of a sister chromatid, forming a Holliday junction. If the DSBR pathway is used then the 3' ssDNA which was not involved in strand invasion will form a second Holliday junction with the invaded chromosome. A nicking endonuclease is then utilized to resolve the double Holliday junctions [23]. Depending on how the Holliday junctions are cut by the nuclease, a crossover product can be formed. If the SDSA pathway is used, only a single Holliday junction is formed, and the extended 3' invading stand is released in a process called branch migration. It then rebinds with the other 3' resected strand and polymerases fill in the missing nucleotides[23]. The SDSA pathway does not produce crossover products [101]. GC pathways result in no loss of genetic information as any damaged or missing components are copied over from the homologous sequence. There is an alternate form of this pathway that is less explored. In this alternate pathway following RPA binding to the newly resected ssDNA strands, RAD52 is recruited. It then assists in displacing RPA from the ssDNA, and recruits RAD51. The pathway then proceeds much as the BRCA2 pathway. This alternate pathway is of great interest because cancers which have deficiencies in key members of the BRCA2 pathway are dependent on the RAD52

alternate pathway to accurately repair their DSBs. If the molecular mechanism behind the interaction of RPA and RAD52 could be discovered, then a focus on the disruption of that mechanism could lead to new potential targets for cancer drugs that would target cells with this so-called RAD52 addiction while leaving healthy cells intact.

Chapter II: Deduction of an Experimental Model of RAD52(1-303)

2.1 Introduction

The inactivation of the alternate HR pathway involving the interaction between RPA and RAD52 has been found to be synthetically lethal with deficiencies in the BRCA2 mediated HR pathway [83, 84]. This means that cells which have either functional BRCA2 HR or RAD52 HR will continue to divide and thrive. If both pathways are made deficient simultaneously the cells are unable to survive. This is of great interest as members of the BRCA2 pathway are known to be mutated in various cancers including breast, pancreatic, and ovarian cancers [102-106]. In cancers featuring a defective BRCA2 pathway, inhibition of RAD52 activity would provide a way to selectively terminate the cancer cells while leaving healthy cells unharmed. To this end finding a way to inhibit the interaction between RPA and RAD52 is a great potential drug target.

This potential is held back by the fact that structural data on the full individual proteins as well as of the complex itself is not available. RPA and RAD52 both contain significant levels of unstructured regions that prevent the crystallization of these proteins as well as inhibiting the effectiveness of NMR techniques with them. It was possible to both crystalize and collect NMR data on specific ordered domains for both proteins [54, 97, 98, 107-110]. This leads us to an alternate way of getting the structure of these proteins as they interact.

Small angle X-ray scattering (SAXS) does not suffer the crystallization requirements of X-ray crystallography nor does it have the size and mobile region restrictions of NMR, however these bonuses come at the cost of resolution. This drawback can be partially overcome by docking the existing domain structures for the proteins with their respective SAXS envelopes. This allows us to compare the shifts in the SAXS envelopes of the individual proteins to identify where the structured domains are oriented,

giving us the binding surface. SAXS data can be interpreted in the following ways. The scattering data can be plotted in three ways, the Guinier plot, Kratky plot, and the pair distance distribution function (P(r)) (**Figures 1.9 and 1.10**). The Guinier plot will give information about the radius of gyration (R_g) as well as serving as a data quality control. The Guinier plot should be linear, as linear means that the sample was monodisperse. If the plot smiles or frowns then there was a problem with aggregation or repulsion effects and further processing cannot be done. The SAXS unit can be calibrated with a MW standard so that the intensity at zero concentration will give an estimate of MW. The second plot that can be used is the Kratky plot. This plot gives information on how globular or well-folded the protein is. A folded protein shows a peak followed by a return to the baseline while unfolded proteins show more of a horizontal asymptote depending on the degree to which they are unfolded. The third plot is the P(r) plot. This plot shows the probability that a random pair of atoms will have a specific distance from each other. Envelopes are generated from the P(r) function. The r intercept of the P(r) function is the farthest distance between a pair of atoms in the molecule or D_{max}. The $P(r)_{max}$ shows the most common distance between pairs of atoms.

The objective for this portion of the project is to acquire initial structural data of the RPA binding domain on the RAD52 unstructured C-terminal domain using SAXS. SAXS data was gathered on RAD52(1-212) and on RAD52(1-303) and *ab initio* protein envelopes generated. RAD52(1-303) was chosen for these experiments because it contains the RPA binding domain and can be purified to a polydispersity compatible with this technique. RAD52(1-212) was selected because docking its existing crystal structure with the RAD52(1-212) SAXS envelope can give us the alignment of the structure. The RAD52(1-212) SAXS envelope was then aligned with the RAD52(1-303) SAXS envelope, conferring the alignment of the crystal structure and allowing us to infer the placement of

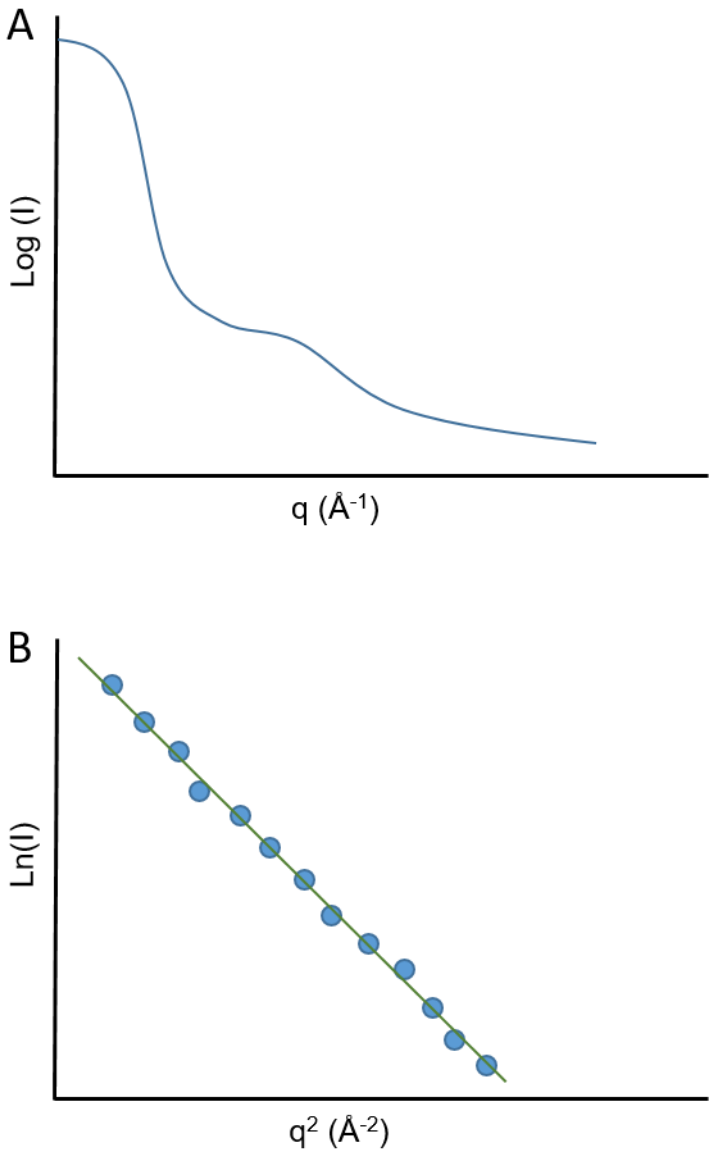


Figure 1.9 Interpreting the SAXS scatter plot and Guinier plot

A) A SAXS scatter plot. Y axis is a log of the intensity value, X axis is the inverse distance from the center of the scattering plot. Low resolution data is closer to the Y axis. The Guinier and Kratky plot are made by regraphing this data.

B) Corresponding Guinier Plot. This plot is made from the low q range data. The slope of the best fit line gives the R_g of the molecule. The Y-intercept value gives the intensity at 0 concentration (I(0)), which when paired with MW standards can give the MW of your molecule. Should this line curve up (or “smile”) it would indicate that aggregation has occurred. If the line should curve down (or “frown”) it would indicate that intraparticle repulsion is occuring. Both situations require the data collected to be disgarded.

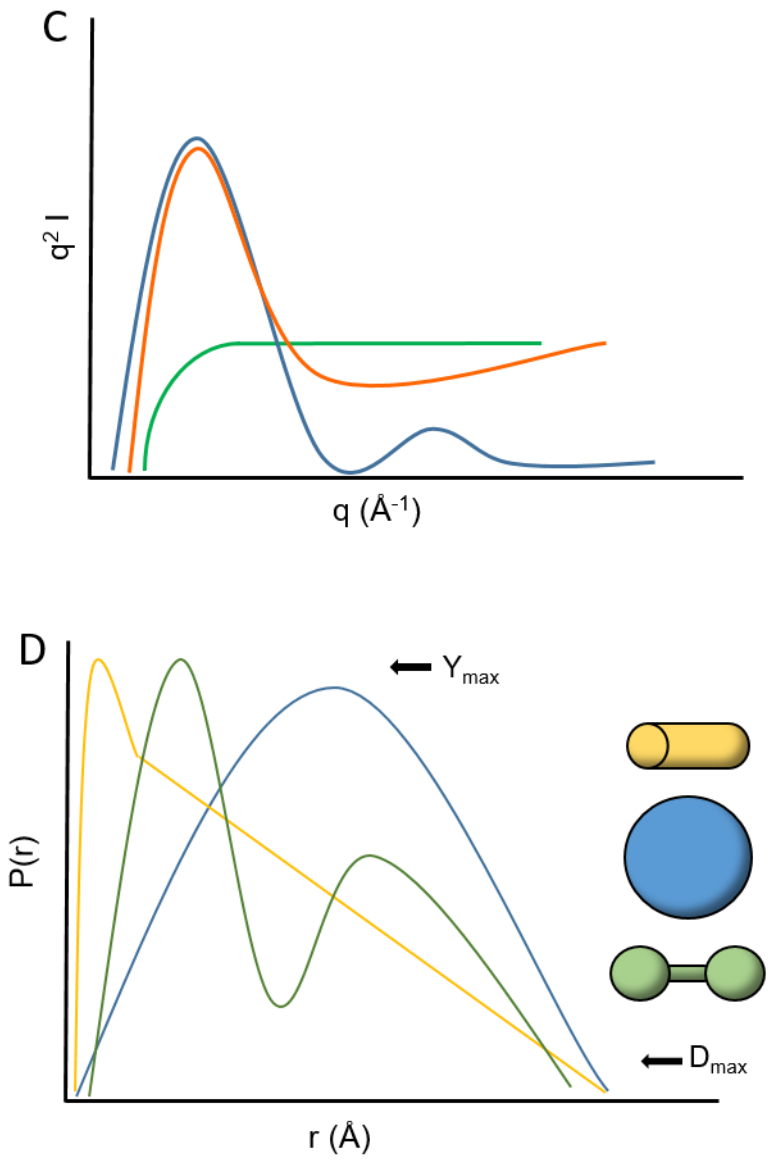


Figure 1.10 Interpreting the Kratky plot and Pair Distribution Function

C) A cartoon of the Kratky plot. Globular (well-folded) proteins will form a bell shaped curve that returns to the baseline (blue). Extended or partially unfolded molecules will not have this return to the baseline, and will often have a rise in higher q values (orange). Completely unfolded proteins rise into a horizontal asymptotic curve and show little or no drop back to the baseline (green).

D) A schematic drawing of the of the P(r) plot. The line gives the pair-wise distance between atoms in the molecule. The maximum r value (D_{max}) gives the longest dimension of the molecule. The maximum Y value (Y_{max}) gives the most common distance between atoms in the molecule. The curvature of the plot contains information about the shape of the molecule. Pictured here are representative curves for rod (yellow), sphere (blue), and barbell (green) shapes.

the unstructured C-terminal amino acids. With this in mind a model of the C-terminal 91 amino acids was generated and added to the crystal structure. Put together, these actions provided the orientation of the RAD52(1-303) molecule as well as some insight into where the C-terminal amino acids, including the RPA binding domain, may be located.

2.2 Methods

2.2.1 Recombinant Protein Expression

The RAD52(1-212) pET28a expression plasmid was created by Gilson Baia in Dr. Gloria E.O. Borgstahl's Lab. The pET28a expression plasmids for RAD52 and RAD52(1-303) were a gift from Dr. Min Park. Rosetta2 *E. coli* cells were transformed using the manufacturers' protocols (Novagen). Selection of transformed cells was done on agarose plates containing 30 µg/ml kanamycin and 34 µg/ml chloramphenicol. Colonies that survived the selection process were grown in lysogeny broth (LB) media incubated at 37 °C with 170 rpm shaking until they reached an optical density at 600 nm (OD600) between 0.7 and 0.9. At this point the cells were induced with 1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) and further incubated for an additional 4 hours. Cells were harvested by centrifugation at 14,000 xg and stored at -20 °C.

2.2.2 Protein Purification

Cell pellets were thawed and suspended in HisTrap running buffer (50 mM Tris-HCl, pH 7.8, 300 mM KCl, 10 mM imidazole, 10% glycerol, 2 mM 2-mercaptoethanol (BME)). A protease inhibitor cocktail (PIC) (SIGMA) compatible with histidine-tagged proteins was added at a level of 250 µl per gram of cell pellet. Lysis of the bacteria was achieved by three passes through an Emulsiflex-C3 at 15,000 psi. Centrifugation for 30 minutes at 40,000 xg separated the cellular debris from the protein-containing supernatant. The supernatant passed through a 0.45 µm HV Durapore^{R} membrane filter (MILLIPORE). Purification of the protein was achieved using the process described in Kagawa W. *et al.*, (2002), and Ranatunga W., *et al.*, (2001), with the following changes. All column chromatography purification was done using an ÄKTApure (GE Lifesciences). The protein was first purified on a HisTrap^{tm} HP chromatography column (GE

Lifesciences). The column was equilibrated with 5 CV of HisTrap running buffer, then the protein lysate was loaded. A gradient to 1 M imidazole was used to elute the protein. Fractions containing the protein of interest were pooled and dialyzed against the heparin running buffer (50 mM Tris-HCl, pH 7.5, 200 mM KCl, 10% glycerol, 0.5mM ethylenediaminetetraacetic acid (EDTA)). The second step in purification utilized a HiTrap^{tm} Heparin HP column (GE Lifesciences). The column was equilibrated with the heparin running buffer and loaded with the protein as described previously. Elution was achieved with a gradient to 1 M KCl. The protein peak was dialyzed overnight against RAD52 size exclusion chromatography (SEC) running buffer (20 mM Bis-Tris, pH 6.0, 10% glycerol, 400 mM NaCl, 100 mM KCl, 1 mM EDTA). The final step in purification used a HiLoad 16/600 Superdex 200 pg column (GE Lifesciences) equilibrated with 1.3 CV of RAD52 SEC running buffer collecting 1 ml fractions. Protein concentrations for wtRAD52, RAD52(1-212), and RAD52(1-303) were acquired using absorbance at 280 nm (A_{280}) with extinction coefficients (ε) of 40300 $M^{-1}cm^{-1}$, 20400 $M^{-1}cm^{-1}$, and 20400 $M^{-1}cm^{-1}$ and MWs of 48.09 kDa, 25.22 kDa, and 34.6 kDa, respectively.

2.2.3 Dynamic Light Scattering

Individual protein fractions were concentrated in the RAD52 SEC running buffer using a 10 kDa molecular weight cut-off (MWCO) spin concentrator (GE Lifesciences). Multiple concentrated fractions were checked for monodispersity using a DynaPro MS/X with a 12 µl quartz cuvette (Wyatt Technology Corporation). Data was collected and processed using Dynamics 6.7.7 software (Wyatt Technology Corporation). A minimum concentration of 1 mg/ml was required for accurate measurement of monodispersity. With this system, fractions with a polydispersity of less than 20% are considered to be monodisperse. The washing procedure for the cuvette involved one wash with 1% Liquinox followed by 5 rinses with deionized water (DI). This wash procedure was

repeated three times. Dynamic light scattering of DI water and buffer was checked with the instrument to ensure the cuvette was clean before use and that the only readings were from the protein being tested. All samples were centrifuged at 13,000xg for 5 minutes and then filtered using a 0.45 µM polyvinylidene fluoride (PVDF) membrane syringe filter (MicroSOLV) before loading 17 µl of analyte into the cuvette in a manner that prevented bubble formation. The equipment performed 10 scans of 10 seconds each to acquire the polydispersity value. Monodisperse fractions that were adjacent to each other in the elution order were combined, prepared again as they were originally, and then re-analyzed.

2.2.4 Size Exclusion Chromatography with Multi-Angle Light Scattering

Protein samples were prepared by centrifugation at 13,000xg for 5 minutes followed by filtration using a 0.45 µM PVDF membrane syringe filter. Size exclusion chromatography with multi-angle light scattering (SEC-MALS) was performed using an Agilent 1260 chromatography system paired with an ultraviolet detector, miniDAWN TREOS, Optilab T-rEX, and a WTC SEC column with WTC SEC guard column (Wyatt Technology Corporation). Analysis of molecular weights was done using ASTRA 6 software (Wyatt Technology Corporation). Samples were run at 0.5 ml/min using RAD52 SEC running buffer.

2.2.5 Small Angle X-ray Scattering

SAXS analysis was performed using a BioSAXS1000 (Rigaku) with attached auto-sampler and a FR-E rotating anode X-ray generator (λ = 1.54 Å). Concentrations were obtained using A_{280} as described previously. Images were collected for 90 minutes with subframes taken every 10 minutes to check for radiation damage, of which none was observed. Data analysis was done using the Automated Analysis Pipeline (AAP) in the SAXLab software. SAXSLab is a GUI for running the ATSAS package [111] running scripts to control the ATSAS programs, after which the results are returned to the user. Buffer subtraction, Guinier plot, R_g, MW, D_{max}, volume, P(r), Kratky plot, the infinite dilution scattering curve, and *ab initio* bead models are all calculated through this software. Docking of the models with the SAXS envelopes was done with the Situs package using the Colores program [112, 113]. Figures were made using PyMOL (Schrödinger, LLC).

2.3 Results

2.3.1. RAD52(1-212)

2.3.1.1 Expression and Purification

RAD52(1-212) expression was checked through visual comparison of uninduced and induced cells using SDS-PAGE. RAD52(1-212) elutes at 150 mM and 300 mM imidazole from the HisTrap^{tm} HP column. Both peaks contained RAD52(1-212) but the peak at 300 mM imidazole was cleaner when checked by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE). This purification was simplified by replacing the concentration gradient with 3 CV steps at 50 mM, 150 mM, 300 mM, and 1000 mM imidazole. Fractions collected from the 300 mM imidazole elution were more monodisperse and were better candidates for SAXS. RAD52(1-212) eluted from the HiTrap^{tm} Heparin HP column at 450 mM KCl. After this step the protein was a single band when investigated with SDS-PAGE and could be used for assays that did not require extreme levels of purity. Polished protein eluted from the HiLoad 16/600 Superdex 200 pg column at approximately 62 ml after injection.

2.3.1.2 Dynamic Light Scattering

Despite its purity, RAD52(1-212) purified in these conditions often would only have a few fractions that were monodisperse. Fractions from SEC purification had to be concentrated to a minimum of 1 mg/ml to reduce polydispersity, and all fractions were checked. DLS of monodisperse fractions revealed a polydispersity of 12.2% (**Figure 2.1**). Adjacent monodisperse fractions could be combined without any gain in polydispersity.

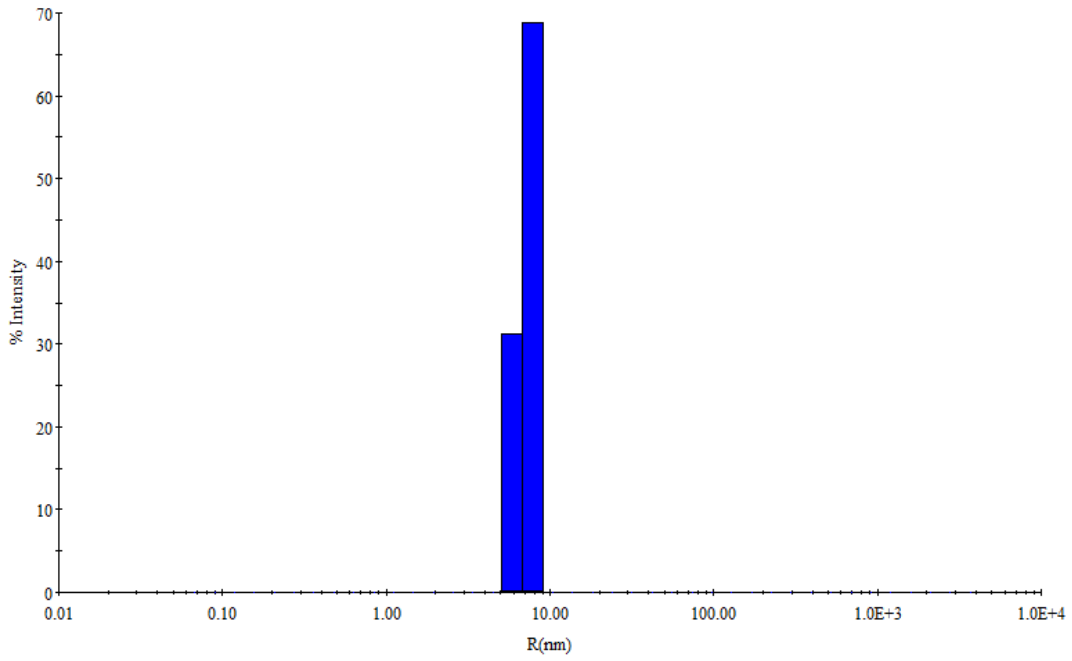


Figure 2.1 DLS of RAD52(1-212)

An example of DLS from fractions of RAD52(1-212) which were used for SAXS. The polydispersity of this polydispersity was 12.2%.

2.3.1.3 Size Exclusion Chromatography with Multi-Angle Light Scattering

SEC-MALS analysis also showed the protein to be monodisperse with a MW of 229 kDa across the peak (**Figure 2.2**). This suggests a protein consisting of a 9.8 subunit ring. Using this information a model 10-mer of the RAD52(1-212) ring was created (by G. Borgstahl) using the program CNS-SOLVE [114, 115] (**Figure 2.3**). To this end six pairs of amino acid contacts were defined and the subunits were placed in a simulated excited state and allowed to reform, combining into a modified but familiar ring shape.

2.3.1.4 Small Angle X-ray Scattering

The 7.48 mg/ml concentration by itself provided the best scattering data and SAXS envelope, and so was not processed with accompanying concentrations (**Figure 2.4**).The Guinier plot for this concentration was linear indicating a monodisperse sample with no aggregation or concentration dependent effects (**Figure 2.5**). The Kratky plot shows a well folded protein (**Figure 2.6**). The P(r) plot showed a D_{max} of 117.7 Å and the most common dimension to be 54 Å (**Figure 2.7**). Instead of the expected and familiar ring form we observed a partially hollow half sphere with a diameter of 120 Å using *ab initio* bead modeling (**Figure 2.8**). The 10-mer model created with CNS SOLVE was docked with the SAXS envelope and the best orientation was found to have the DNA binding region of the ring residing inside the curved portion of the half-sphere (**Figure 2.9**).

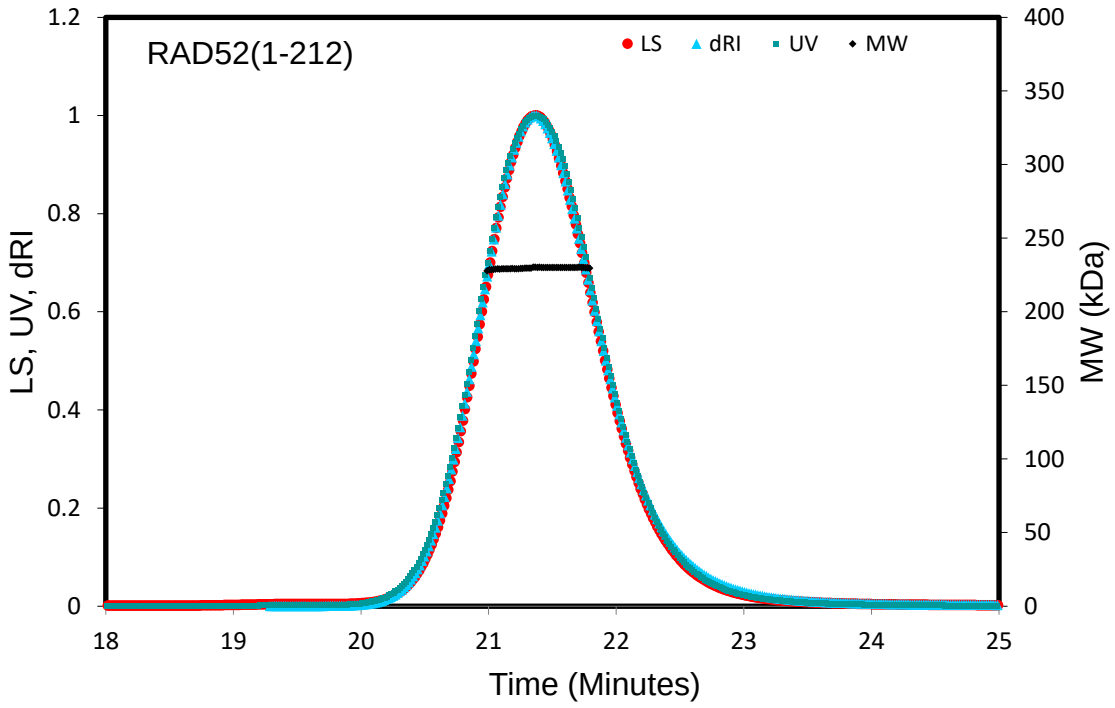


Figure 2.2 SEC-MALS of RAD52(1-212)

Light scattering is shown in red, differential refractive index in blue, A_{280} in teal, and MW in black. The level MW across the peak indicates monodispersity. The MW for this molecule was 229 kDa, which corresponds to 9.8 subunits.

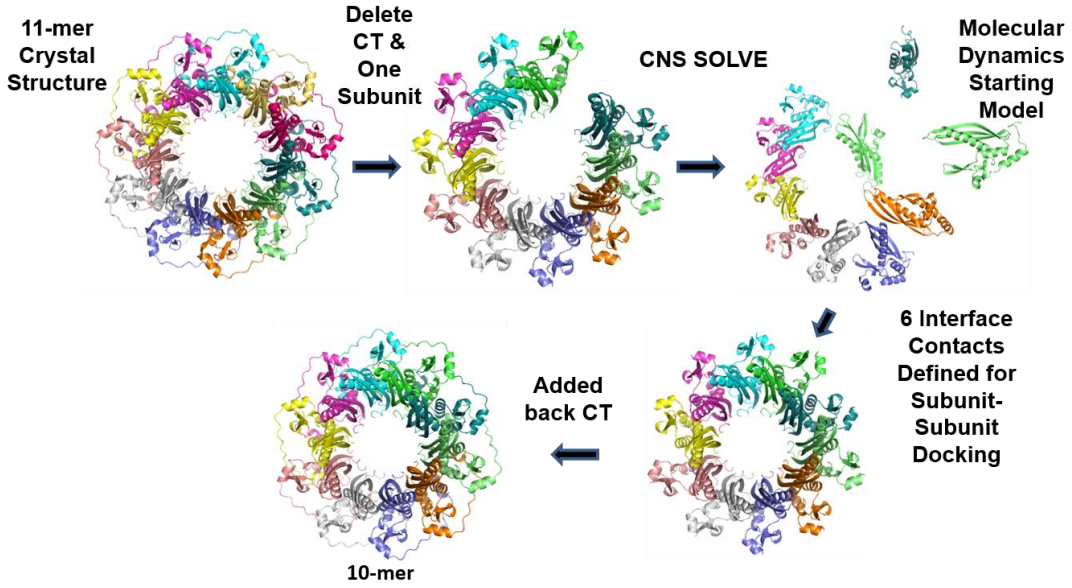


Figure 2.3 Creation of the RAD52(1-212) 10-mer model

The 10-mer model of RAD52(1-212) was created by removing a 1mer from the 11mer crystal structure. CNS SOLVE was used to perform molecular dynamics with six interface contacts acting as restraints. This resulted in 10mer rings.

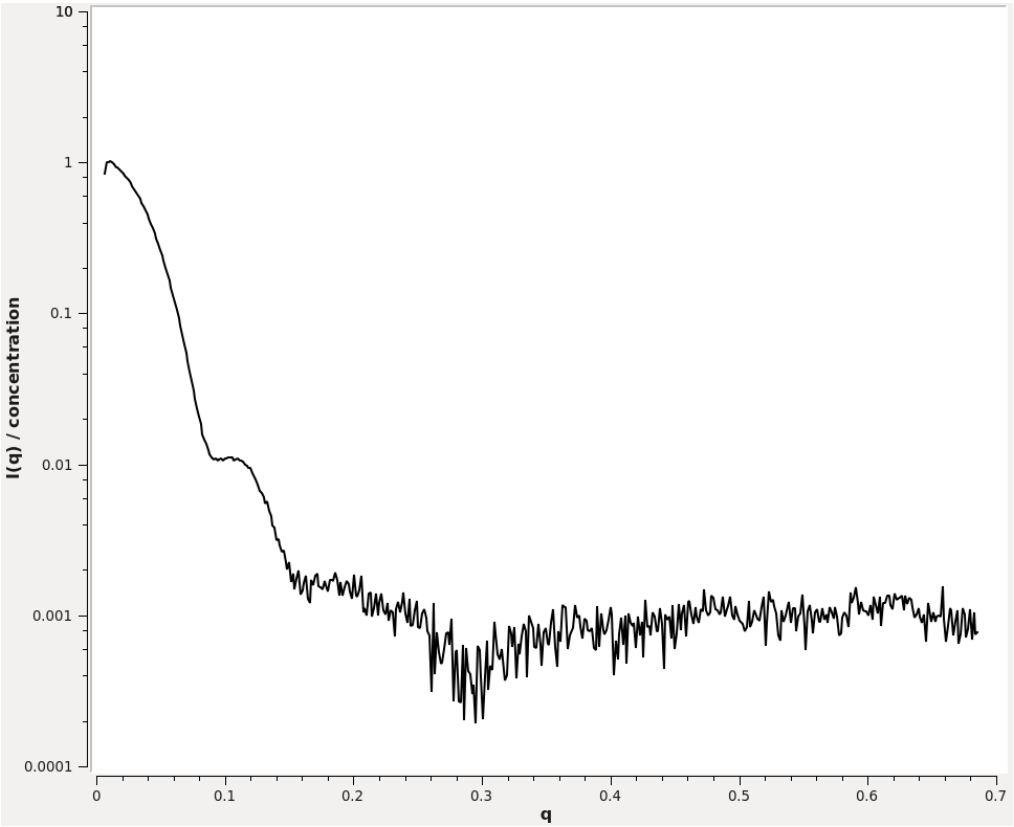


Figure 2.4 SAXS scatter plot for RAD52(1-212)

The scattering pattern for the RAD52(1-212) protein. The concentration for this data is 7.48 mg/ml.

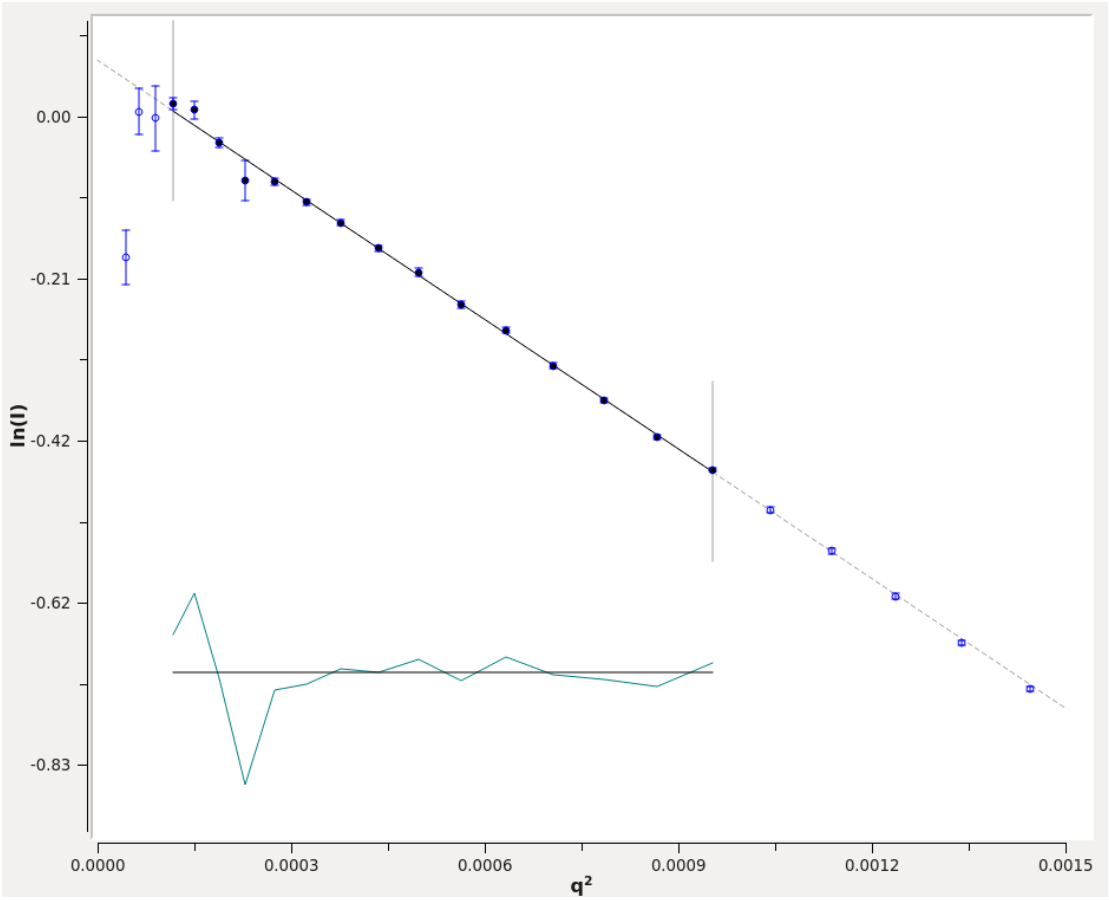


Figure 2.5 Guinier plot of RAD52(1-212)

The Guinier Plot of the low resolution RAD52(1-212) data is linear and therefore does not indicate the presence of aggregation or repulsion effects.

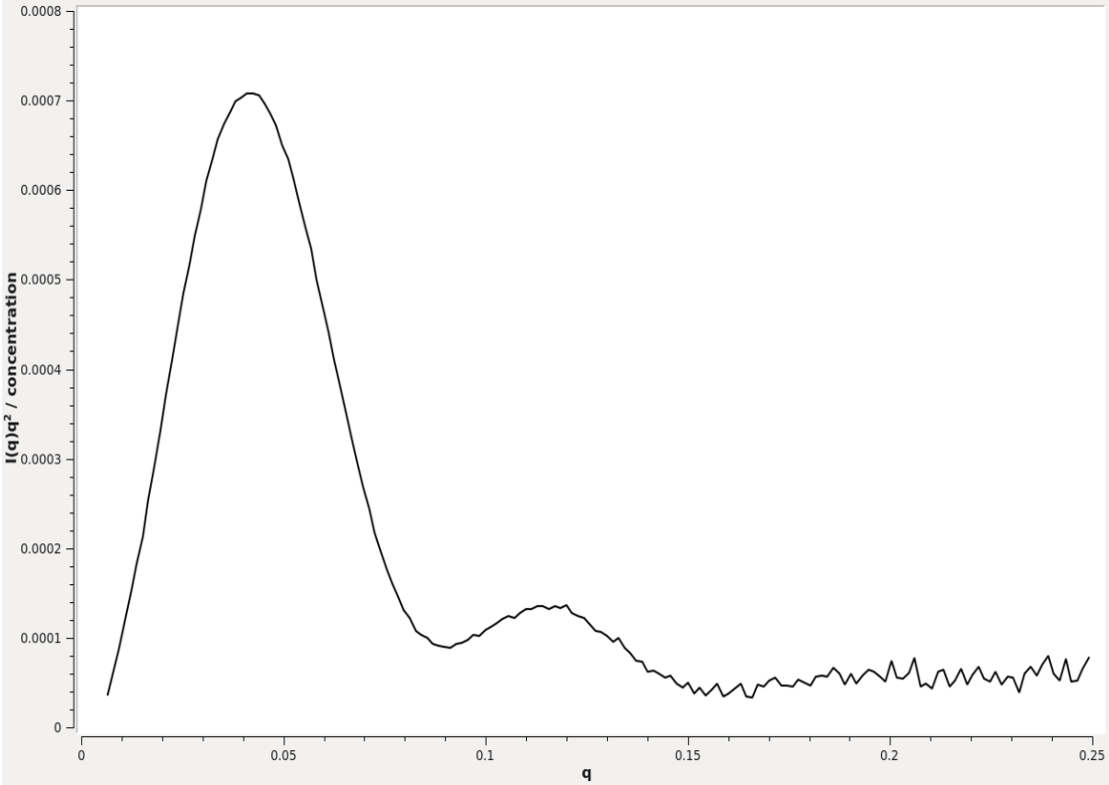


Figure 2.6 Kratky plot of RAD52(1-212)

The Kratky plot for RAD52(1-212) shows a globular, well-folded protein with the peak rising and falling back close to the baseline at the higher q values.

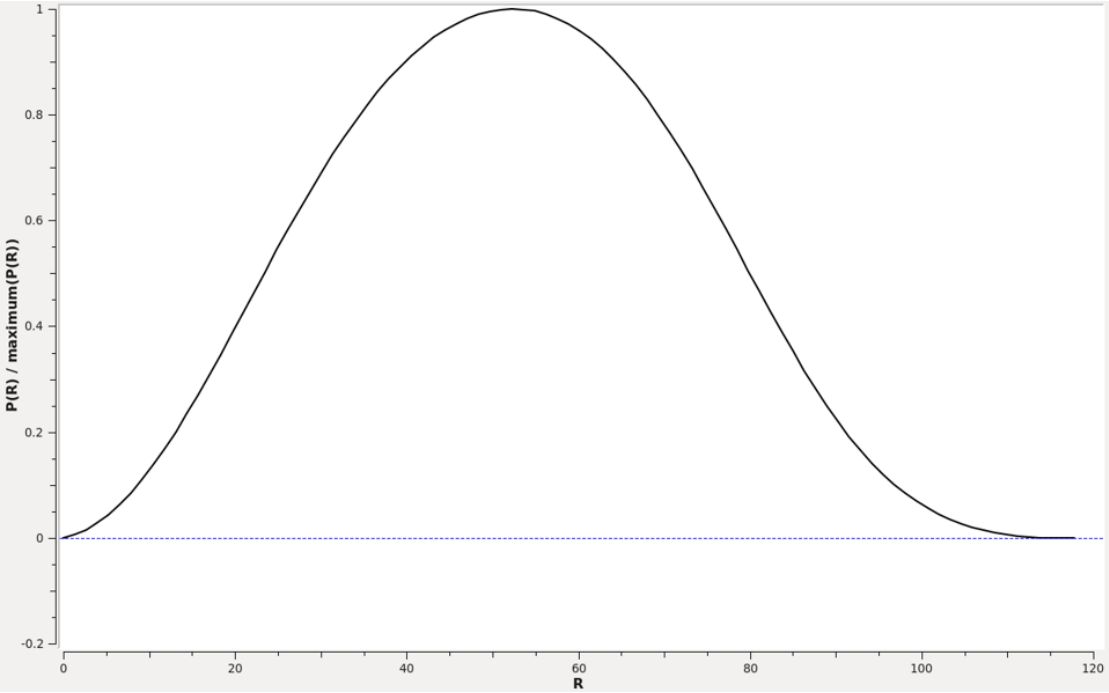


Figure 2.7 P(r) plot of RAD52(1-212)

The P(r) of RAD52(1-212) shows the D_{max} to be 117.7 Å.

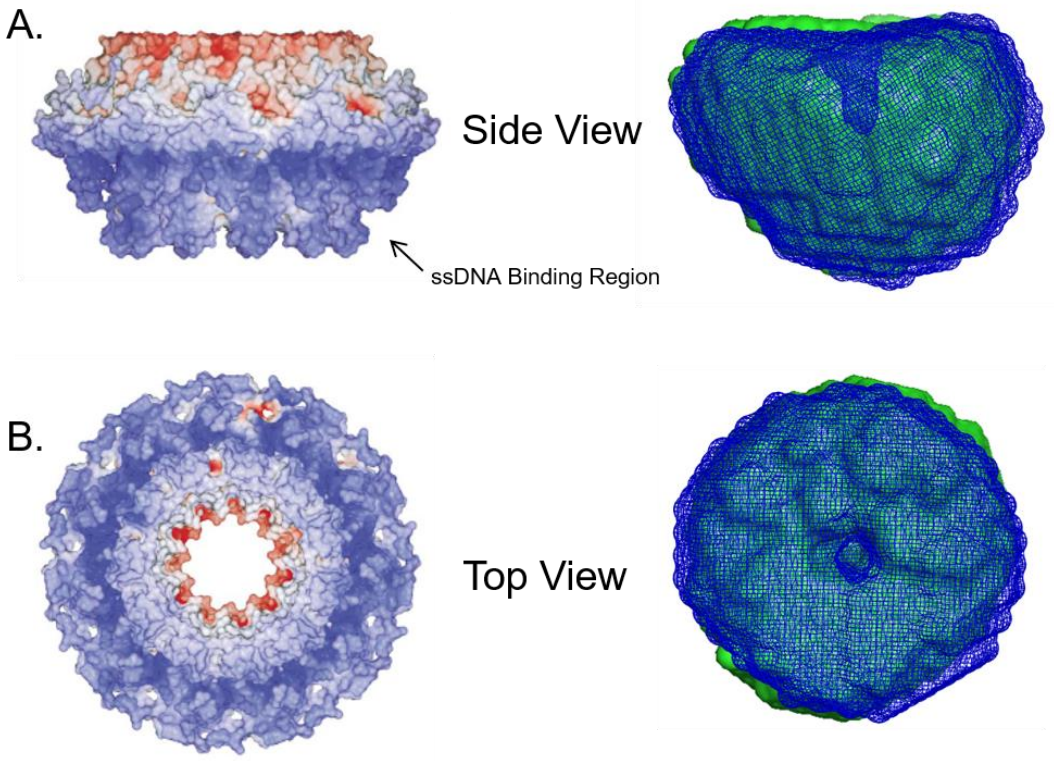


Figure 2.8 Comparison of the electrostatic surface map of the crystal structure and the SAXS *ab initio* model of RAD52(1-212)

RAD52(1-212) viewed from the A) side and B) top. The electrostatic surface (left) and the SAXS *ab initio* bead model of RAD52(1-212) (right, green) are shown. A blue wireframe model has been superimposed over the bead model to highlight the central cavity.

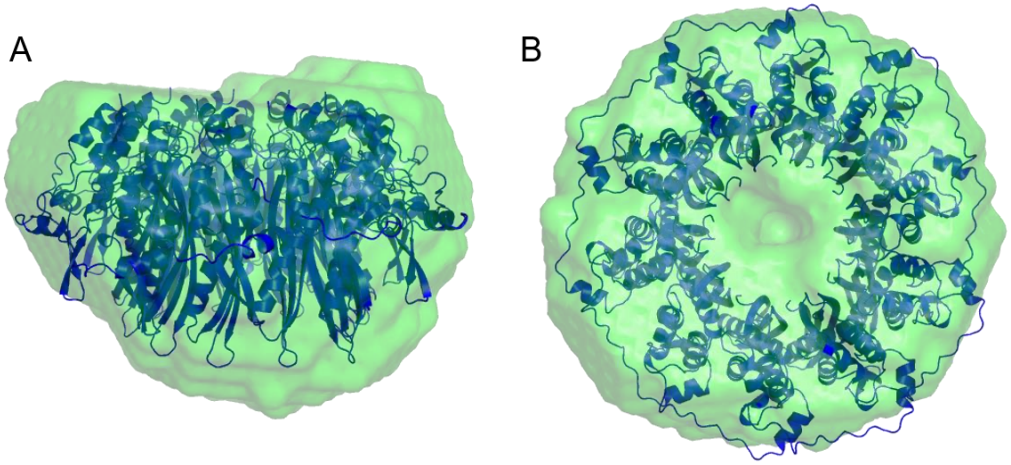


Figure 2.9 Docking the RAD52(1-212) 10-mer with the RAD52(1-212) SAXS envelope

The 10-mer RAD52(1-212) model docked with the RAD52(1-212) SAXS envelope. The best orientation was found to have the DNA binding domains of the ring situated inside the curved face of the SAXS envelope. A) side view. B) top view.

2.3.2 Rad52(1-303)

2.3.2.1 Expression and Purification

Purification of RAD52(1-303) was identical to that of RAD52(1-212) except for the following changes. HisTrap^{tm} HP purification could be done either with a gradient that would result in RAD52(1-303) elution at 270 mM, or through steps at 150 mM, 300 mM, and 1000 mM imidazole. When purified with steps the 300 mM imidazole elution contained the highest concentration of protein that would be monodisperse in later purification steps. RAD52(1-303) further purified on a HiTrap^{tm} Heparin HP column eluted at 430 mM KCl. At this point the protein was clean when checked by SDS-PAGE. Dialysis against the RAD52 SEC running buffer for at least 12 hours followed by SEC was required for protein monodispersity. During SEC RAD52(1-303) would elute from the Superdex 200 column at approximately 53 ml.

2.3.2.2 Dynamic Light Scattering

RAD52(1-303) purified in these conditions had only a few fractions that were monodisperse, requiring all factions from SEC to be tested separately. Adjacent monodisperse fractions would be combined for later experiments. DLS of these fractions typically resulted in a polydispersity of around 14% (**Figure 2.10**).

2.3.2.3 Size Exclusion Chromatography with Multi-Angle Light Scattering

SEC-MALS analysis also showed the protein to be close to, but not monodisperse with an average MW of 328.7 kDa across the peak, indicating a ring composed of 9.4 subunits (**Figure 2.11**). The DLS information and SEC-MALS graph indicate that there are

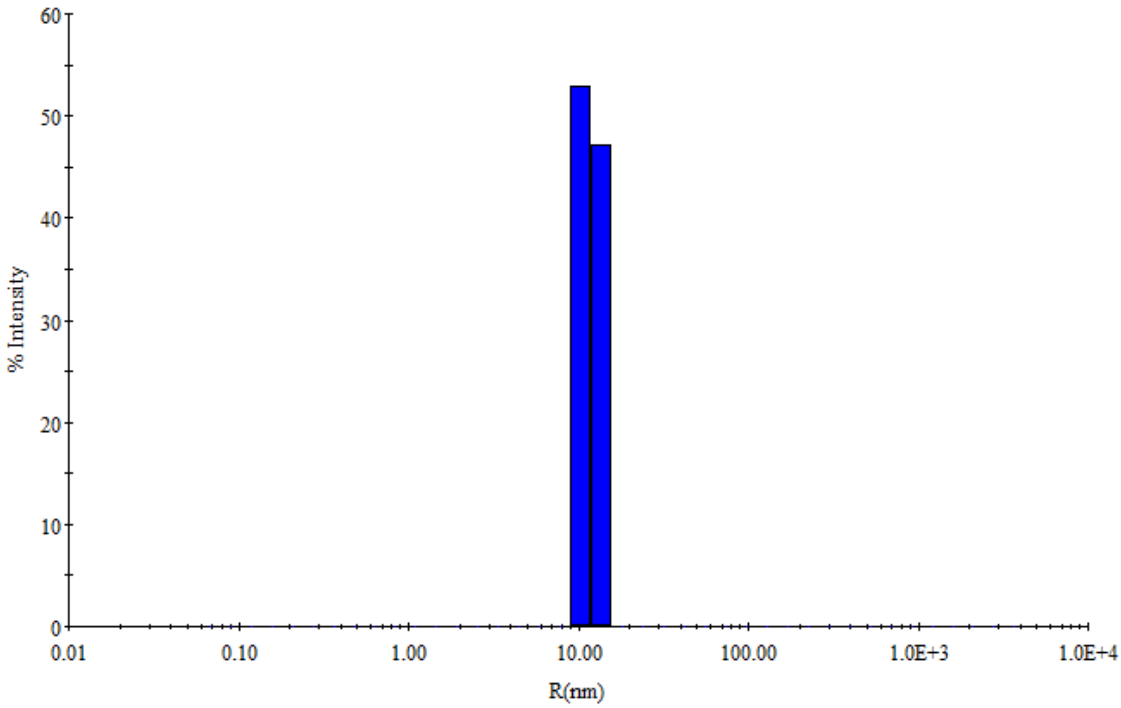


Figure 2.10 DLS of RAD52(1-303)

The DLS of a SEC fraction of RAD52(1-303). The polydispersity of this faction was 14%.

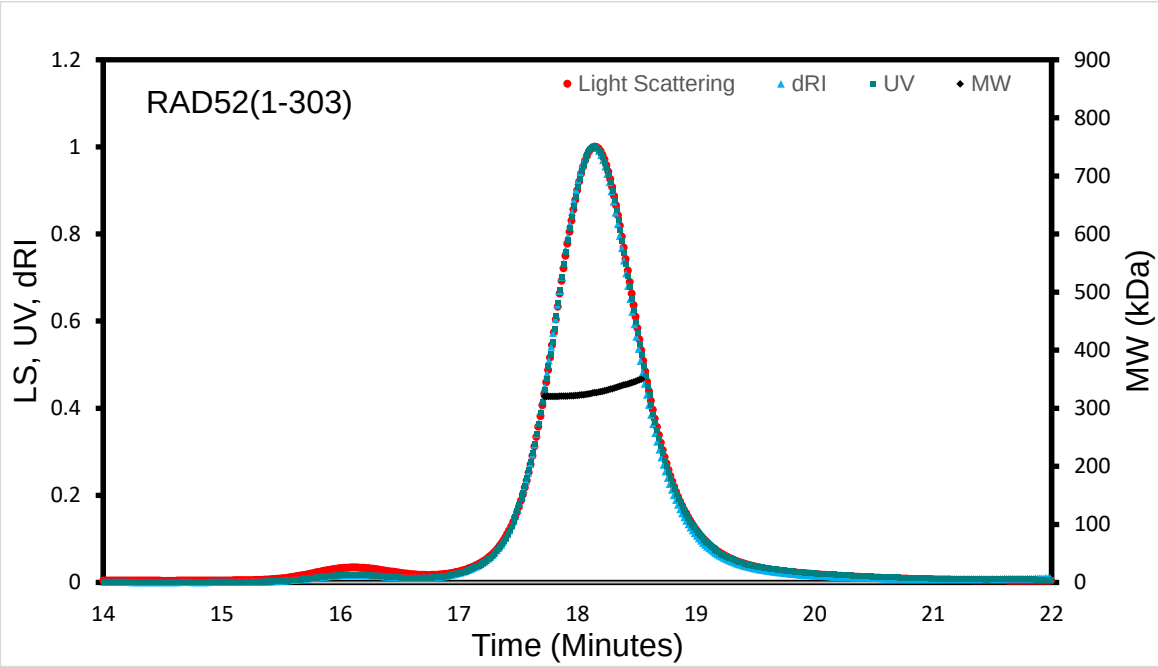


Figure 2.11 SEC-MALS of RAD52(1-303)

Light scattering is shown in red, differential refractive index in blue, A_{280} in teal, and MW in black. The tilt in MW across the peak indicates that although the peak is nearly Gaussian, the protein is not truly monodisperse and has slightly higher MW on the right side of the peak. The calculated MW across the peak is 328.7 kDa, indicating 9.4 subunits.

two monodisperse peaks that blend together, yet look like a single peak. A MW of 328.7 kDa indicates a protein consisting of a mix of nine and ten subunit rings with a higher concentration of nine membered rings. Using the same process as described before, a 9 subunit model of the RAD52(1-212) ring was assembled.

2.3.2.4 Small Angle X-ray Scattering

Scattering for 1.56 mg/ml and 2.66 mg/ml were collected (**Figure 2.12**). The Guinier plots were linear for all concentrations, indicating that there were no aggregation or concentration dependent effects of the protein under these conditions (**Figure 2.13, 2.14, 2.15**). Kratky plot analysis indicated that the protein was well-folded (**Figure 2.16**). The P(r) showed a D_{max} of 174 Å and a Y_{max} of 59.6 Å (**Figure 2.17**). The SAXS envelope for this protein was pear-shaped with the wider region having a diameter of approximately 120 Å and an end-to-end length of 174 Å (**Figure 2.18**).

2.3.2.5 Docking RAD52(1-303)/RAD52(1-212)

Using the curved dome present in both SAXS envelopes for alignment, the RAD52(1-212) and RAD52(1-303) SAXS envelopes were docked (**Figure 2.19**). This positioning based on the common structural feature of the two models shows a clear, common orientation of the two. From this we can interpret that the rounded portion of the wider end of the envelope will likely have the DNA binding surface of the RAD52 ring oriented facing this curve.

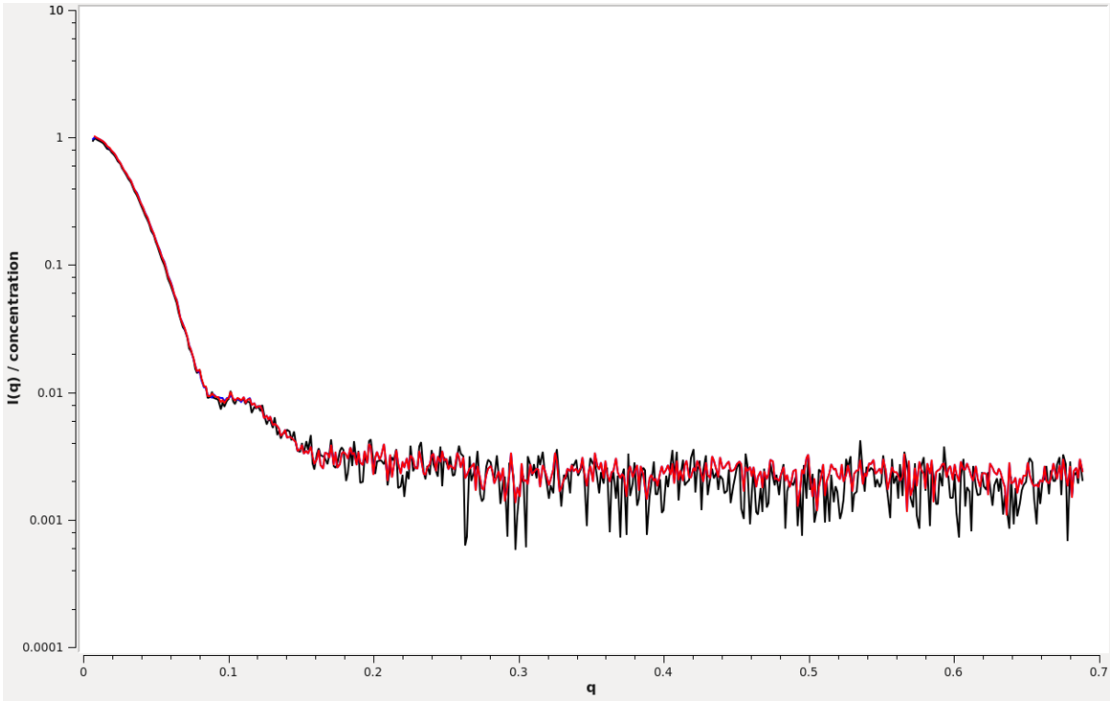


Figure 2.12 SAXS scatter plot of RAD52(1-303)

Scattering curve for 1.56 mg/ml shown in black. Scattering curve for 2.66 mg/ml shown in blue. Extrapolated (0 mg/ml) scattering curve shown in red.

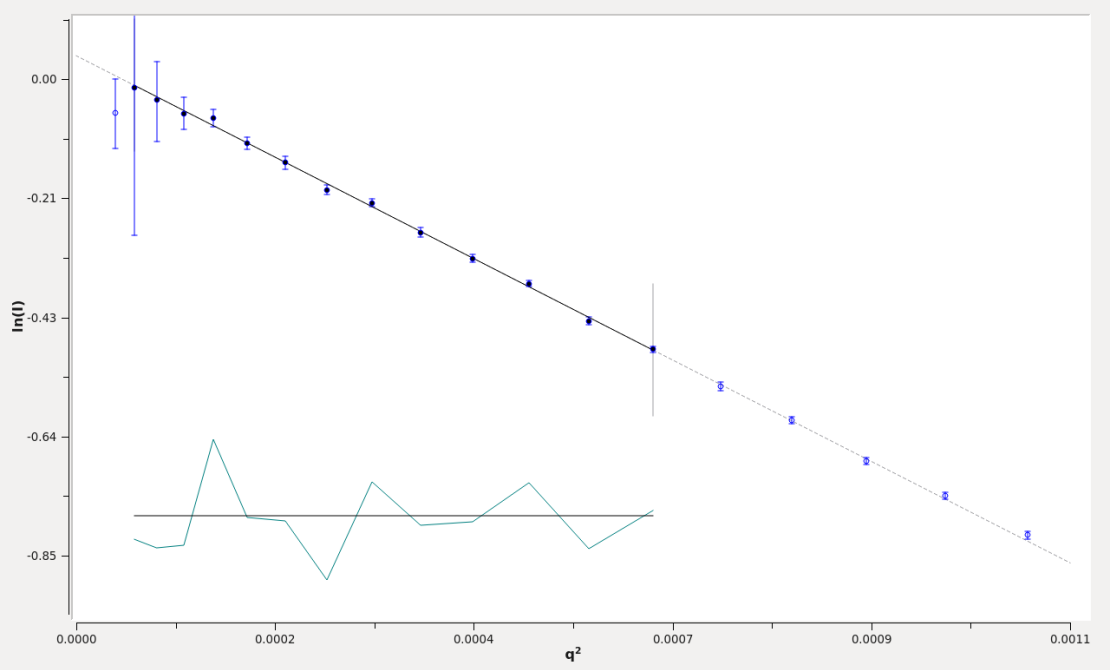


Figure 2.13 Guinier plot of RAD52(1-303) 1.56 mg/ml

The Guinier Plot of the low resolution data is linear and does not show the presence of aggregation or repulsion dependent effects.

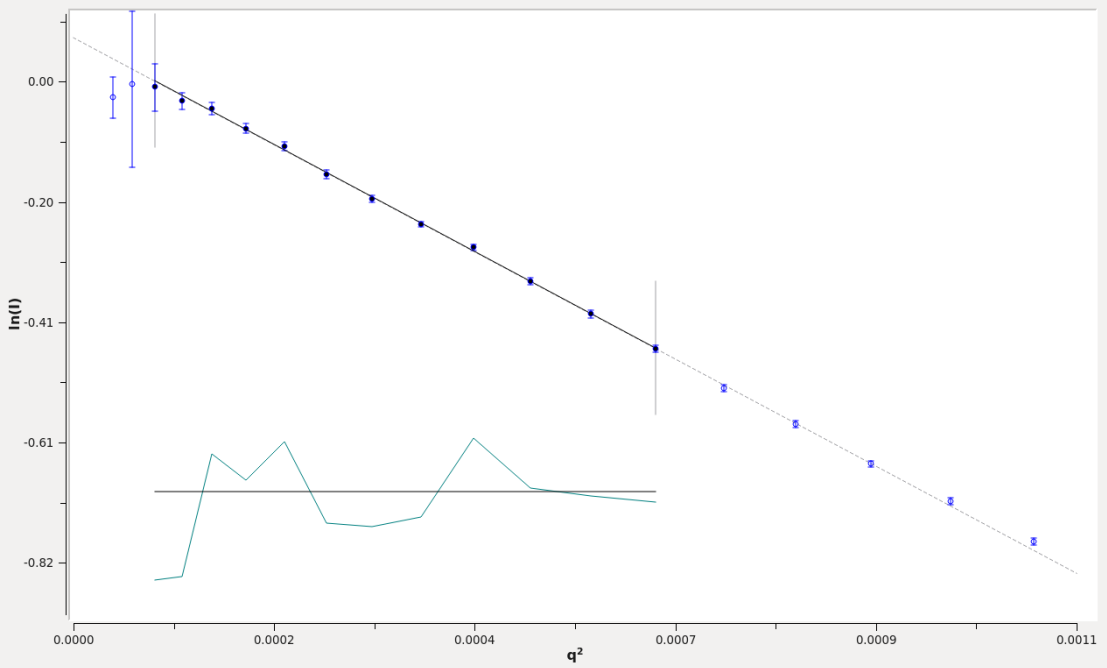


Figure 2.14 Guinier plot of RAD52(1-303) 2.66 mg/ml

The Guinier Plot of the low resolution data is linear and does not show the presence of aggregation or repulsion dependent effects.

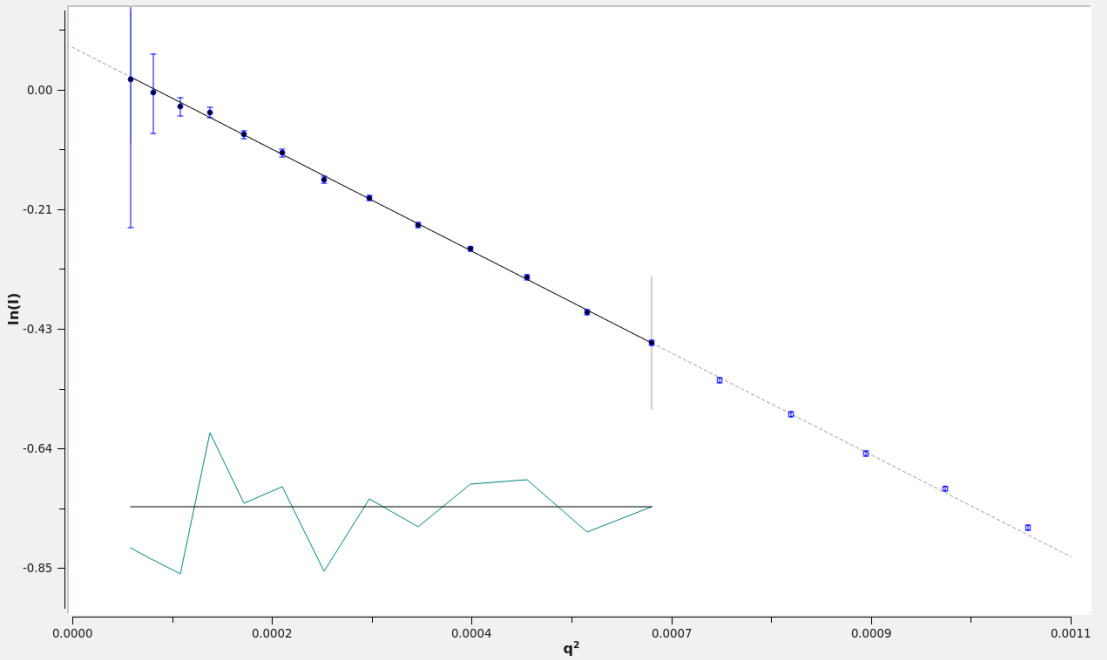


Figure 2.15 Guinier plot of RAD52(1-303) extrapolated data (0 mg/ml)

The Guinier Plot of the low resolution data is linear and does not show the presence of aggregation or repulsion dependent effects.

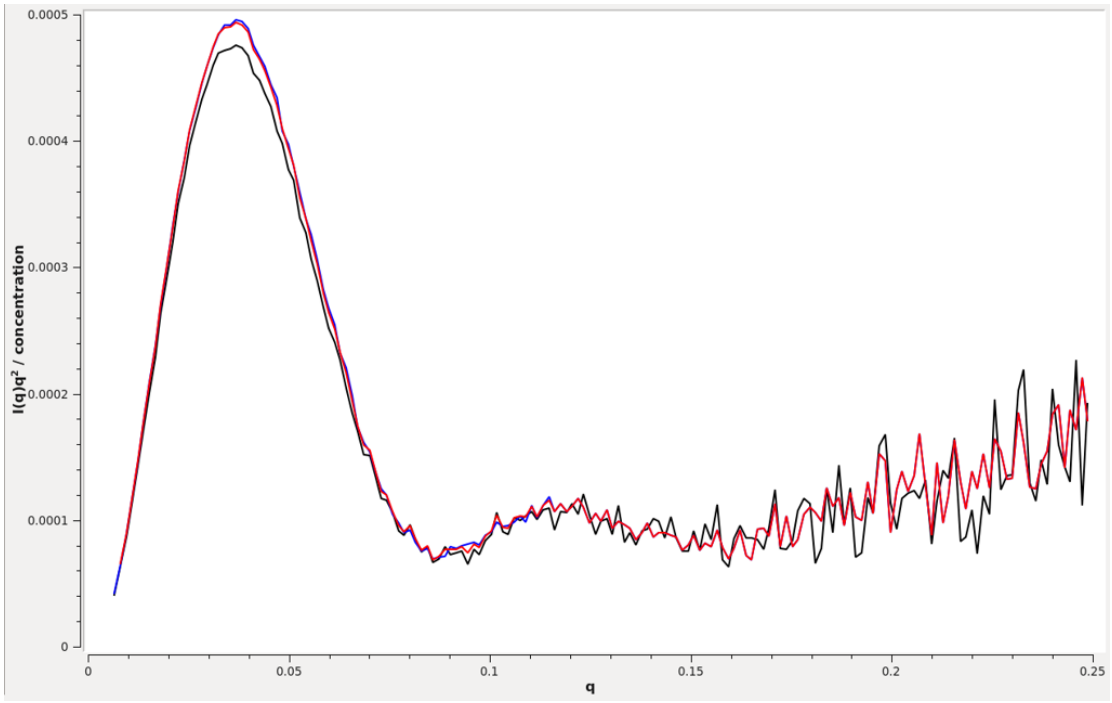


Figure 2.16 Kratky plot of RAD52(1-303)

Data for 1.56 mg/ml shown in black. Data for 2.66 mg/ml shown in blue. Extrapolated (0 mg/ml) Data shown in red. The Kratky plot for RAD52(1-303) shows a globular, well folded protein with the peak rising and falling back close to the baseline.

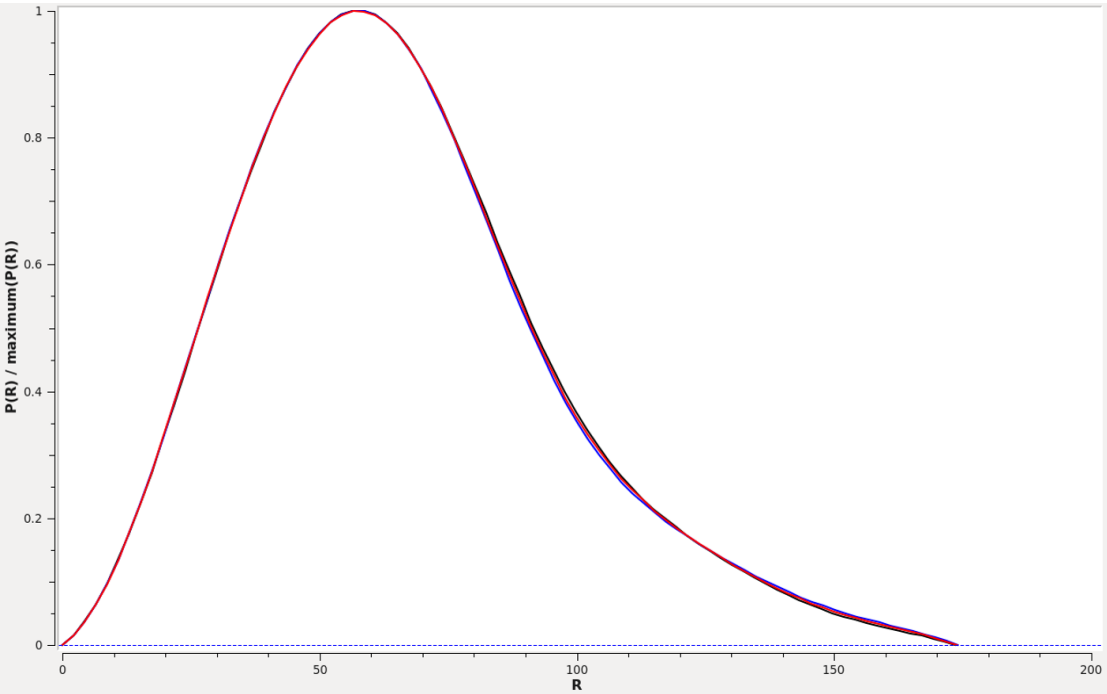


Figure 2.17 PDF plot of RAD52(1-303)

Data for 1.56 mg/ml shown in black. Data for 2.66 mg/ml shown in blue. Extrapolated (0 mg/ml) Data shown in red. The PDF plot of RAD52(1-303) shows the D_{max} to be 174 Å, and the shape is likely a combination of a sphere and rod.

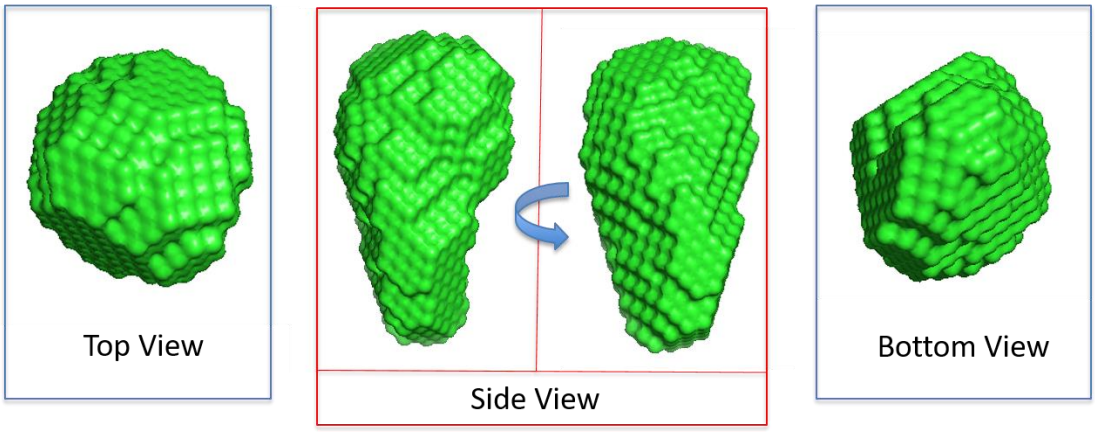


Figure 2.18 RAD52(1-303) SAXS envelope

A view of the molecule from the larger diameter region down (left), from the side (center), and from the narrower diameter region looking up (right). The molecule is 174 Å from top to bottom (side view), with the larger spherical head close to 120 Å in diameter (top view).

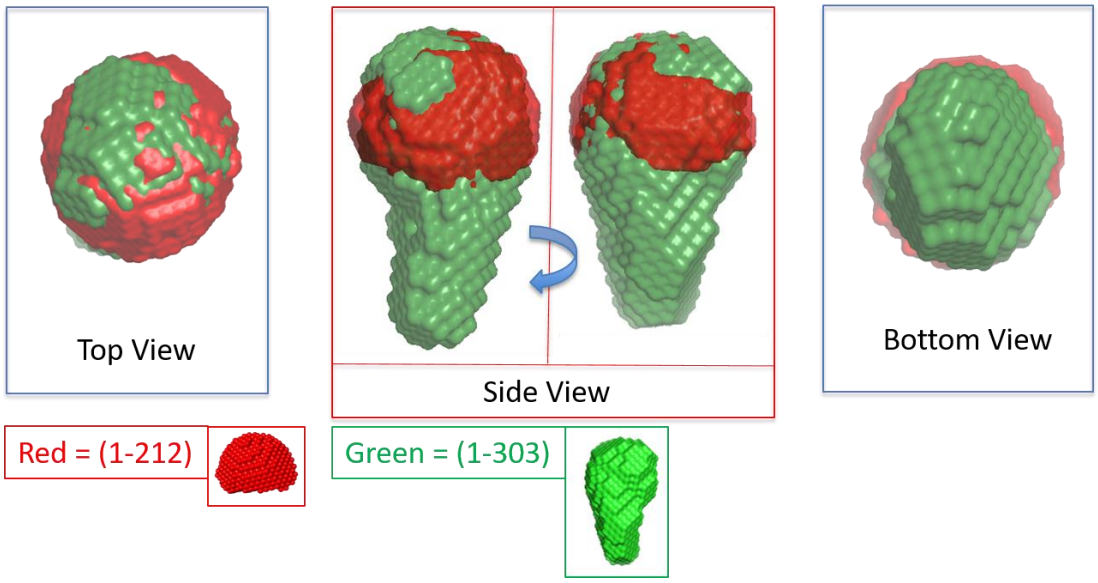


Figure 2.19 Docking of the SAXS envelopes for RAD52(1-212) and RAD52(1-303)

The curved face of the RAD52(1-212) SAXS envelope (red) fits very well into the larger domed region of RAD52(1-303) (green), showing the alignment of the RAD52(1-303) molecule. With this it can be extrapolated that the DNA binding domain of RAD52(1-212) orients toward the top of this dome, with the C-terminal region oriented into the narrow lower portion of the envelope.

2.4 Discussion

Published methods for purification for RAD52(1-212) produce pure and monodisperse RAD52(1-212). To transition this method to SAXS analysis, the end buffer needed to be made more accommodating to the restrictions imposed by this technique, so a modified buffer was employed. The lower pH of the RAD52 SEC buffer (compared to the other RAD52 buffers described earlier) moved the charge of RAD52 further from its isoelectric point, while the increased salt levels assisted with both the stability of RAD52 and protected against prolonged X-ray exposure-induced aggregation effects in SAXS. DLS and SEC-MALS both showed a monodisperse protein, but SEC-MALS also showed a shift from the crystal structure of RAD52(1-212) to a 10-mer configuration. To reflect this, a 10-mer model was constructed. The Kratky plot showed the protein to be well-folded with little to no unfolded regions. The Guinier plot showed no concentration dependent effects as well as little to no radiation damage of the protein over the time spent in the path of the X-ray beam. The PDF function indicated that the D_{max} of this molecule was 117.7 Å with the most common inter-atom distance being 54 Å. The *ab initio* bead model gave a SAXS envelope that looks like a half sphere with a hollowed region in the center as opposed to the expected ring shape. This half sphere model was confirmed through multiple attempts and varying concentrations. The 11-mer and 10-mer models were docked with the SAXS envelope for RAD52(1-212). The 11-mer model would not fit inside of the SAXS envelope without a significant portion of the model extending outside. It was found that the DNA binding surface of the RAD52(1-212) 10-mer would fit well towards the curved region of the RAD52(1-212) SAXS envelope.

Purification of the RAD52(1-303) protein was comparable to the purification of RAD52(1-212), aside from a few details. RAD52(1-303) eluted from the HisTrap^{tm} HP by

a gradient does not appear as a single concentrated peak. The majority of the protein was a single peak, but a significant amount would continue to elute as the imidazole concentration increased, eventually tapering off. This was most easily observed on SDS-PAGE gels. Our current explanation for this is that RAD52 is known to form rings and those rings have been shown to associate with each other. It is possible that we are seeing multiple forms of the RAD52(1-303) oligomer. Taking RAD52(1-303) from a HisTrap^{tm} HP to HiTrap^{tm} Heparin HP purification required dialysis of the protein into the heparin running buffer. Simple dilution of the protein resulted in an absence of protein binding to the column. RAD52(1-303) elution from the heparin column was very pure when examined by SDS-PAGE, but it was not monodisperse. To achieve monodispersity the protein had to be dialyzed overnight into RAD52 SEC running buffer, and purified on a Superdex 200 column. The fractions from SEC needed to be concentrated to greater than 1 mg/ml before investigation by DLS. Not all fractions from SEC were monodisperse, likely because SEC separated RAD52 rings with different numbers of subunits. Adjacent monodisperse fractions could be combined and used for further experiments. SEC-MALS of RAD52(1-303) gave a MW of 328.7 kDa indicating 9.4 subunits. This most likely means that under these conditions we had a mixture of a 9-mers and 10-mers, the majority being 9-mers. The Guinier plot of RAD52(1-303) showed no curvature, indicating that for all concentrations tested the protein was free from effects due to aggregation, concentration, and X-ray damage. The Kratky plot revealed the protein to be well folded, which was a little surprising given the unstructured nature of the C-terminal region. The P(r) function shows a maximum dimension of 174 Å and a most common atom to atom distance of 59.6 Å. Averaging by the ATSAS program suite shows the SAXS envelop to be an oblong ovoid 174 Å from end to end at its longest axis, and with a shape almost like a large sphere with an attached narrower cylinder. The top curved region of the larger spherical hemisphere

of this envelope has almost identical dimensions to the curved face of the RAD52(1-212) SAXS envelope. The 9-mer model was docked with the RAD52(1-303) SAXS envelope in the same orientation as the RAD52(1-212) envelope, with the DNA binding region toward the larger curved end. There are still 91 amino acids in RAD52(1-303) that are unaccounted for by the RAD52(1-212) crystal structure. The missing amino acid sequence was put into I-TASSER (by G. Borgstahl) to predict the most likely folding of this unstructured region [116-118]. This program's methodology combines composite homology, where the target sequence is aligned with sequences of known structure, and *ab initio* loop modeling to generate likely structures for unfamiliar sequences. The most likely conformation had a Z score of 1.14, where scores above 1 indicate a good alignment, and a C score of -2.77, where scores greater than -1.5 indicate a correctly predicted global topology. It was modeled attached to a single subunit of the 9-mer model. The folded 91 amino acids fit within the SAXS envelope (**Figure 2.20**). Additional experiments will be needed to determine the formal structure of this unstructured CTD. Performing SAXS on RAD52(1-303) in the presence of a strong magnetic field could help organize the multiple unstructured coils increasing the confidence of the SAXS envelope, as well as using a molecule to bind to the C-terminal amino acids and comparing the changes to the unbound SAXS envelope.

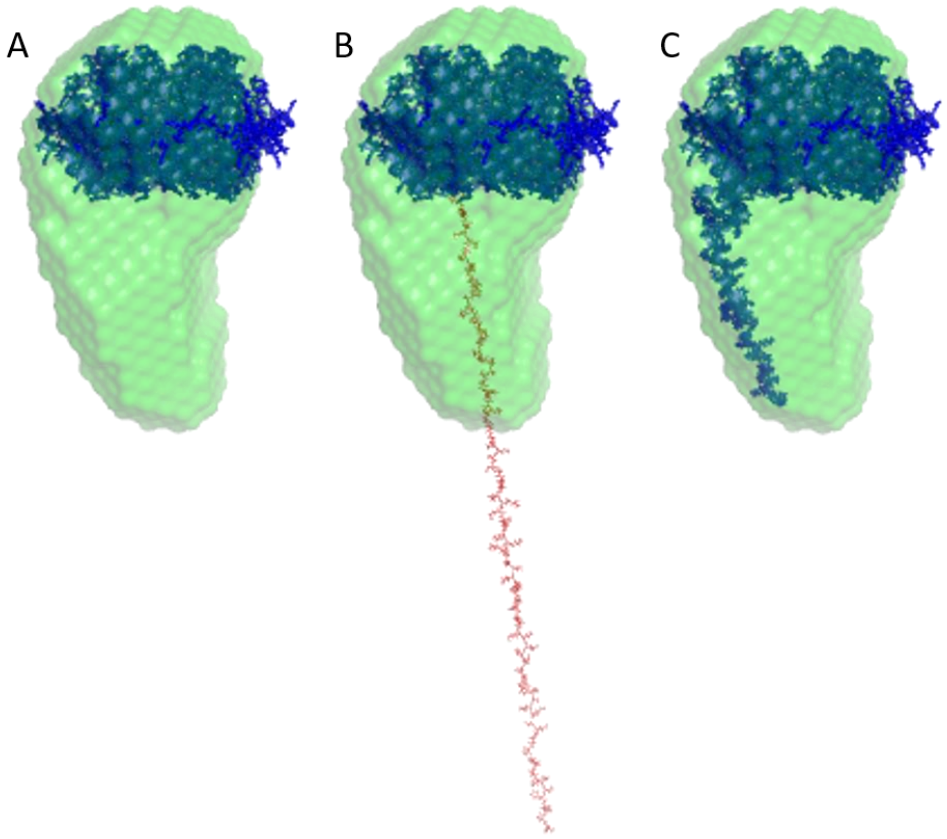


Figure 2.20 RAD52(1-303) with attached CTD

A) The 9-mer model of RAD52(1-212) docked with the RAD52(1-303) SAXS envelope. The DNA binding domain of the 9-mer is oriented toward the top of the envelope.

B) Amino acids (213-303) shown attached but fully extended to a single subunit, demonstrating the level of folding that must occur for this sequence to fit inside the envelope.

C) The most likely model of amino acids (213-303) as predicted by ITASSER attached to a single subunit. The model of these C-terminal amino acids fits well inside the SAXS envelope.

Chapter III: Creating DNA Damage Response Relevant RPA Phosphomimetics Compatible with Small Angle X-ray Scattering

3.1 Introduction

RPA is an important and well-studied ssDNA binding protein in humans. The RPA heterotrimer (composed of subunits of 70-, 32-, and 14-kDa labeled RPA70, RPA32, and RPA14, respectively) is regulated through extensive phosphorylation. RPA is involved in the initiation of almost every form of DNA repair [43, 61, 119-122]. RPA becomes hyperphosphorylated in response to DDR activation, and the N-terminus of RPA32 contains the majority of the phosphorylation sites in unstructured regions of the protein [53, 63, 72, 123, 124]. After activation of the DDR pathway three kinases belonging to the phosphatidylinositol 3-kinase-related kinase (PIKK) family, known as ataxia-telangiectasia-mutated (ATM), ataxia-telangiectasia and Rad3-related (ATR), and DNA-dependent protein kinase (DNA-PK), become active and are known to phosphorylate RPA [66-68]. Historically the RPA32 subunit was known to be phosphorylated due to its shift on SDS-PAGE so an 8D construct (amino acids Ser8, Ser11, Ser12, Ser13, Thr21, Ser23, Ser29, and Ser33 on RPA32, mutated to Asp residues) was made in the laboratory of Dr. Marc Wold to mimic this hyperphosphorylated form of RPA32 [125]. Since then much more RPA phosphorylation data has been produced. Databases have hundreds of phosphorylation sites involving all three subunits. It is time to create phosphomimetics based on the new data. Even with the extensive research and reviews available, as well as the fact that RPA plays such a key regulatory role in the DDR, information on which phosphorylation sites are important to the DDR let alone their individual contributions to the regulation of RPA are not known. Using information gathered in previous experiments

by our lab we attempted to make phosphomimetic variants of RPA that has undergone full hyperphosphorylation due to DDR activation and versions with fewer sites.

Previously capillary isoelectric focusing immunoassay data showed that RPA in UM-SCC-38 cells synchronized in G2 phase contained up to 7 phosphorylations with 1, 2 and 4 phosphorylations being most frequent (**Figure 3.1**, blue line). Once DSBs were introduced RPA had up to 14 phosphorylations present, again with 1, 2 and 4 phosphorylations being the most populated (**Figure 3.1,** pink line). RPA phosphorylation-specific antibodies showed that Ser4/8, Ser12, Thr21, Ser23, and Ser33 on RPA32 all become phosphorylated in cells with DSBs, while Ser23 or Ser23 and Ser33 are phosphorylated in control cells due to cell cycle regulation (**Figure 3.2**). Using the website PhosphoSitePlus we acquired a list of all RPA phosphorylations that have been observed at least once, and selected from that list ones that are likely phosphorylated by the PIKK family of kinases (ATM and ATR), as those kinases are known to phosphorylate RPA after DSBs have been formed. This was done by checking if the phosphorylation site matched the ATM and ATR phosphorylation consensus sequence of S/TQ (Ser or Thr residue followed by a Gln). Including sites observed through western blotting in response to DNA damage, this left us with: Ser38, Thr180, Ser207, and T483 on RPA70, and Ser4/8, Ser11/12/13, Thr21, Ser23, Ser33, Ser52, Ser72, and Ser174 on RPA32. We further pared down the list of candidate sites by removing Ser38 and Ter483 from the RPA70 group. Residue Ser38 in an uncharacterized CDK site in the RPA70-F domain and is unlikely to contribute to complex formation with RAD52 or alter ssDNA interaction. Amino acid Thr483 is in the zinc-finger region and its mutation would destabilize the heterotrimer. This resulted in our final list of phosphorylation sites to investigate: Thr180 and Ser207 on RPA70 and Ser4/8, Ser11/12/13, Thr21, Ser23, Ser33, Ser52, Ser72, and Ser174 on

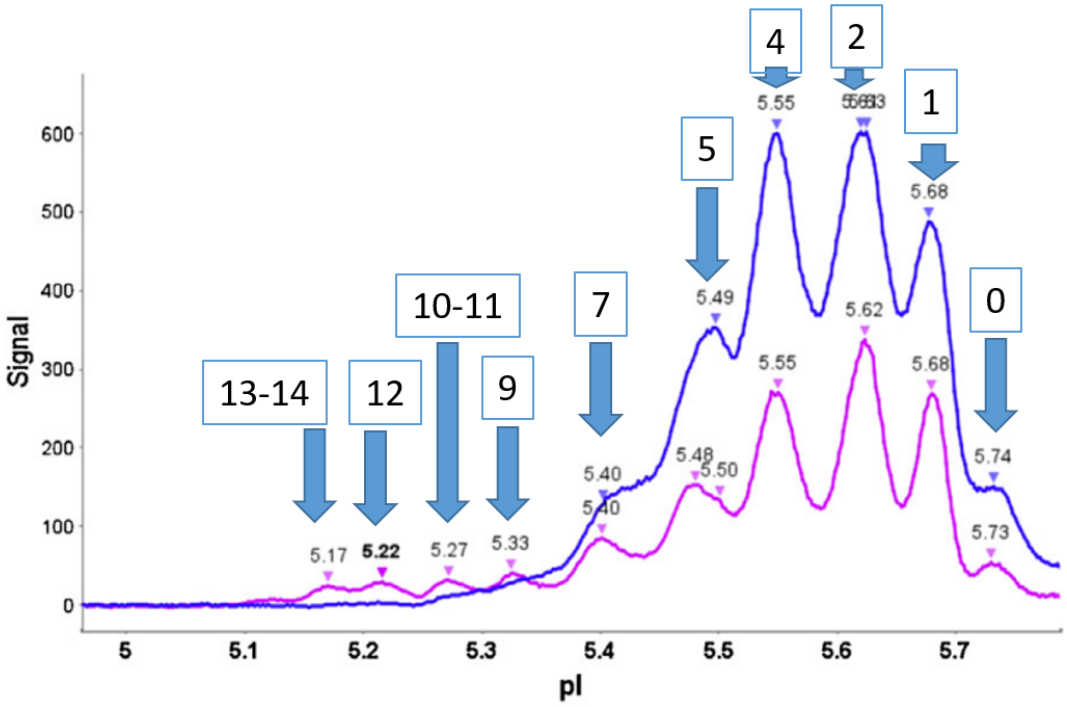


Figure 3.1 Capillary isoelectric focusing of RPA from whole cell lysates

Whole cell lysates from cells without DSBs (blue) and with DSBs (pink) were run through capillary isoelectric focusing before being cross-linked to the capillary wall and probed with α-RPA70. The shifts in pI indicate changes in phosphorylation state. Number of phosphorylations corresponding to the shift in pI are listed above each peak. (Modified from G. Borgstahl *et. al, 2014)*

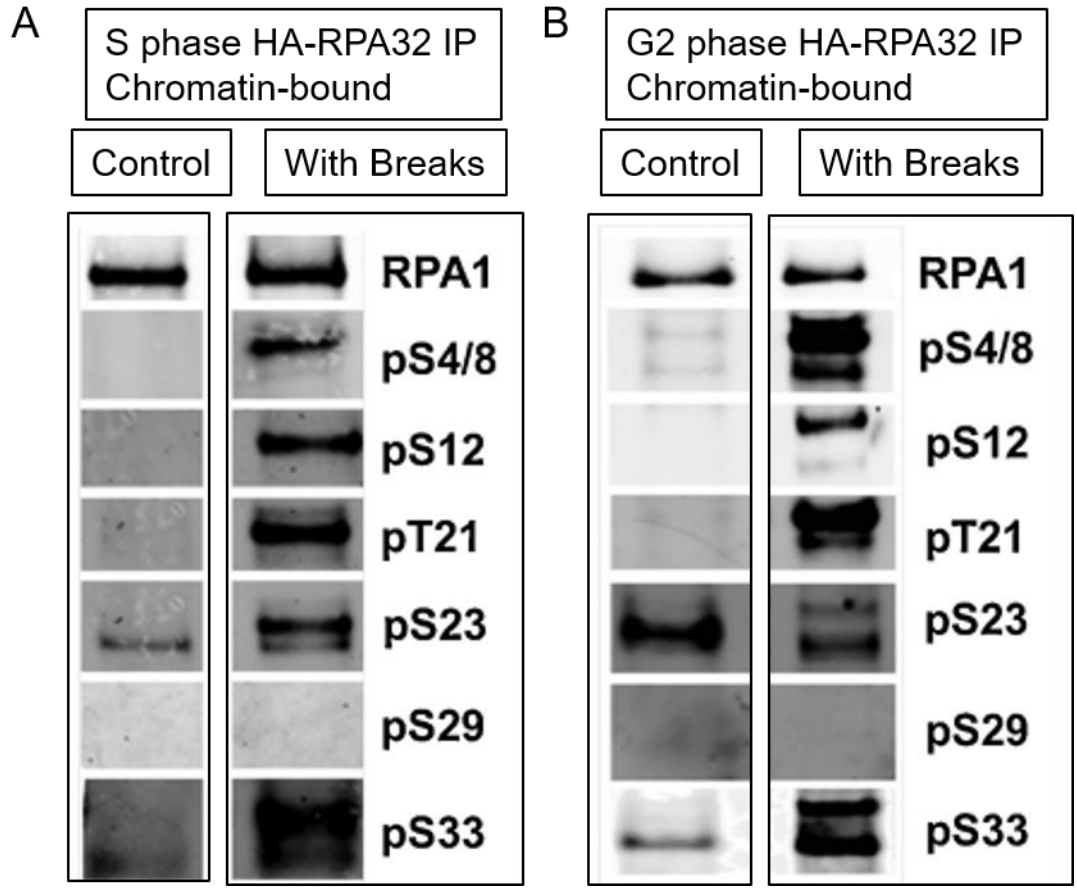


Figure 3.2 Western blot of DSB induced phosphorylation

UM-SCC-38 cells without DSBs (Control) and with induced DSBs (With Breaks). Cells were synchronized in either A) S phase or B) G2 phase. In S phase the cells showed phosphorylation of the CDK site S23, even in the absence of DNA damage. In G2 phase the CDK sites of S23 as well as S33 also were phosphorylated in the absence of DNA damage. (Modified from G. Borgstahl *et. al, 2014)*

RPA32 (**Table 3.1, Figure 3.3**). Residue Ser207 is in DBD-A on RPA70 close to the actual DNA binding interface, and phosphorylation of this residue could affect DNA binding (**Figure 3.4A**). Amino acid Ser52 is located in DBD-D on RPA32 in the interface between RPA32 and RPA14, and phosphorylation could affect heterotrimer stability (**Figure 3.4B**). Site Ser72 is also in DBD-D on RPA32, but is located on the same face of the subunit as both Ser174, which is just outside the structured region, and the RPA32 N-terminus, which contains the majority of the candidate phosphorylation sites. This face of RPA32 could contain a high amount of negative charges (**Figure 3.4C**).

Recombinant proteins with different combinations of candidate sites mutated to Glu residues were expressed and purified. This generally followed two paths, with the first path creating phosphomimetic combinations relevant to cells with DSB-induced phosphorylation, and the second path creating phosphomimetics found in control cells. Combinations which could form heterotrimers were repurified using the full four step purification and had their DNA affinity tested using surface plasmon resonance (SPR).

Protein	Residue	Notes
RPA70	Thr 180	unstructured
	Ser 207	ssDNA binding site A
RPA32	Ser 4,8	unstructured
	Ser 11,12,13	unstructured
	Thr 21	unstructured
	Ser 23	unstructured, G2 cyclin dependent kinase site
	Ser 33	unstructured
	Ser 52	D domain, RPA14/32 protein-protein interface
	Ser 72	D domain
	Ser 174	unstructured

Table 3.1 Candidate amino acids for mutation to Glu on RPA70 and RPA32

Residues are listed as either part of RPA70 (top) or RPA32 (bottom), as well as their location on the subunit.

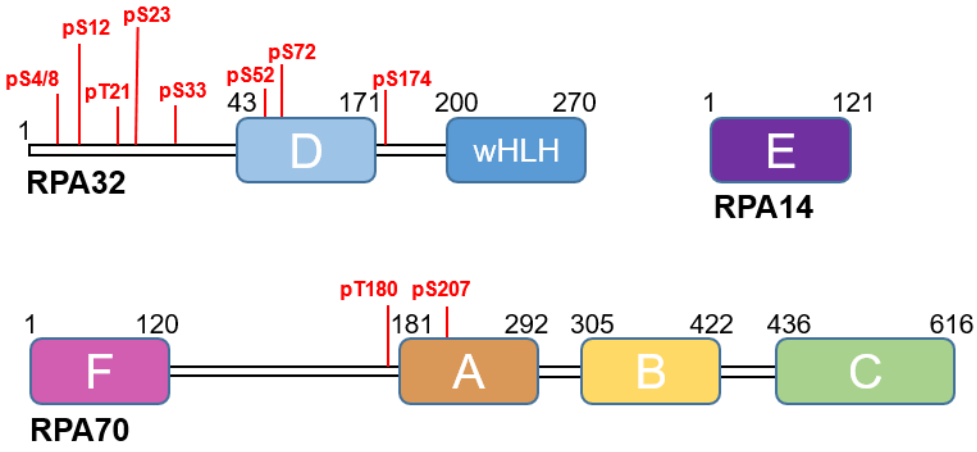


Figure 3.3 RPA domain map with candidate phosphorylation sites

While most of the candidate sites are located in unstructured linker regions, with the majority of them located on the RPA32 C-terminus, some are located in ordered domains.

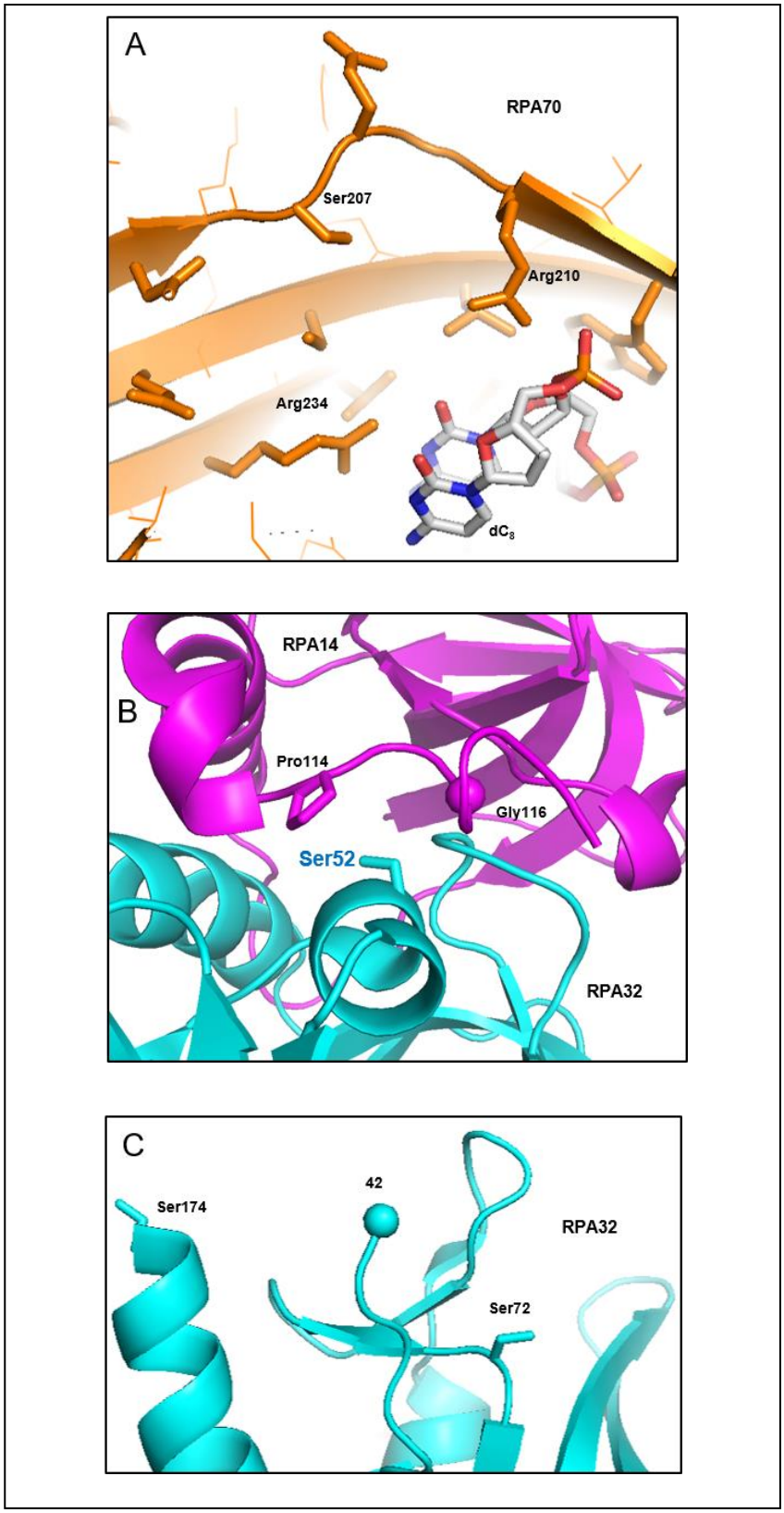
Figure 3.4 Location of candidate phosphorylation sites in available crystal structures

A) Ser207 is located on RPA70 DBD-A in a position to interfere with DNA binding.

B) Ser52 is located on RPA32 DBD-D in binding interface between RPA32 and RPA14.

C) Ser72 is located on RPA32 DBD-D, on the same face as Ser174 and the RPA32 N-terminal unstructured region containing the majority of the phosphorylation candidate sites.

Figure prepared by G. Borgstahl using Pymol.


A
RPA70
Ser207
Arg210
Arg234
dC8
B
RPA14
Pro114
Gly116
Ser52
RPA32
C
Ser174
42
RPA32
Ser72

3.2 Methods

3.2.1 RPA Phosphomimetics

Plasmids containing recombinant proteins were ordered from GenScript. Full length human RPA14 was cloned into the MCS1 (Ncol/Notl) restriction site and full length human RPA32 was cloned into the MCS2 (Ndel/Xhol) restriction site of the pACYCDuet-1 vector. Full length human RPA70 was cloned into the MCS2 (Ndel/Xhol) restriction site of a pCOLADuet-1 vector (**Table 3.2**). Codons for these RPA subunits were optimized for expression in *E. coli.* The combinations of candidate sites chosen for each plasmid variant were based on the appropriate PIKK and CDK sites identified in previous studies [69].

3.2.2 Recombinant Protein Expression

BL21(DE3) *E. coli* cells were transformed with pACYCDuet-1 and pCOLADuet-1 plasmids concurrently following methods provided by the supplier (Novagen) except for the incubation time in the shaking incubator following heat shock. The cells had to be incubated for a minimum of 30 minutes longer than the provided protocol recommended. Dual selection of cells occurred through growth on agarose plates containing 30 µg/ml kanamycin and 34 µg/ml chloramphenicol. If the cells were not given the additional time in the shaking incubator stated earlier then no colonies would form. Surviving colonies were selected and grown in LB media incubated at 37 °C with shaking at 170 RPM. Cells were grown until they reached an OD600 between 0.7 and 0.9, after which they were induced with 1 mM IPTG and incubated for 4 hours at 37 °C. Cells were pelleted by centrifugation at 14000 xg and stored at -20 °C.

RPA70 in pCOLADuet-1	RPA70 mutations
wtRPA70	none
2ERPA70	T180E, S207E
1EaRPA70	S207E
1EbRPA70	T180E

RPA32/14 in pACYCDuet-1	RPA32 mutations
wtRPA32/14	none
9ERPA32/14	S4E;S8E;S12E;T21E;S23E;S33E;S52E;S72E;S174E
8ERPA32/14	S4E;S8E;S12E;T21E;S23E;S33E; S72E;S174E
7EaRPA32/14	S4E;S8E;S12E;T21E;S23E;S33E; S72E
7EbRPA32/14	S4E;S8E;S12E;T21E;S23E;S33E; S174E
7EcRPA32/14	S4E;S8E;S12E;T21E;S23E;S33E; S52E
6ERPA32/14	S4E;S8E;S12E;T21E;S23E;S33E
2ERPA32/14	S23E;S33E
1ERPA32/14	S23E

Table 3.2 Plasmid combinations for expressing phosphomimetics of RPA

Plasmids contained glutamic acid mutations for RPA70 or RPA32 candidate sites as listed above. If a plasmid contained the same number of phosphomimetic mutations as another plasmid, a letter was added to the name as a distinction.

3.2.3 Purification

Protein purification was performed as described in the literature with the following changes [126]. The lysis buffer had 500 mM of sodium thiocyanate (NaSCN) and 0.1% (v/v) PIC (SIGMA) added to the HEPES Inositol (HI) buffer (30 mM HEPES, pH 7.8, 0.5% inositol, 0.25 mM EDTA and 2 mM 2-mercaptoethanol) and was used for resuspension of cells stored in pellet form (5 ml/g cell pellet). Cell lysis was achieved with three passes through an Emulsiflex-C3. Cell lysate was centrifuged at 40,000 xg and the supernatant was filtered using a 0.45 µm filter (EMD Millipore) to remove cellular debris. This clarified lysate was dialyzed against HI-200 mM KCl. An ÄKTApure was used for all column chromatography purification steps. The first step in purification for all RPA proteins was a HiTrap^{tm} Blue HP column (GE Lifesciences). This column was equilibrated with 5 CV of HI-0 buffer after which the clarified lysate was loaded using injections of up to 15 ml with 7 CV of HI-0 buffer as a wash after each injection. After the column was loaded with the full lysate amount (25-50 ml) it was rinsed with 5 CV of HI-800 mM KCl buffer, followed by 5 CV of HI-500 mM NaSCN. RPA heterotrimer elution takes place with a 5 CV step of HI-750 mM NaSCN, which was collected and dialyzed into HI-200 mM KCl. The second step in purification used a column packed with ceramic hydroxyapatite (HAP) (BIO-RAD) using HI-0 as a running buffer and HI-160 mM sodium phosphate ($NaPO_4$) to create an elution gradient. Protein solution was loaded onto the column as described previously, and a 20 CV gradient from 0 to 100% HI-160 mM $NaPO_4$ eluted the protein. To clean the column 1 CV of HI-0 followed by 5 CV of HI-500 mM KPO_4 was used, and the column was stored in 0.5 M NaOH. The third purification step is a HiTrap^{tm} Q FF column. The protein solution from the HAP purification step was diluted with an additional 4 volumes of HI-0 buffer before being loaded onto the HiTrap^{tm} Q FF column. HI-0 was used as the running buffer, and an elution gradient was formed to HI-1M NaCl. The final step in purification was a

HiLoad 16/600 Superdex 200 pg column equilibrated with 1.3 CV of HI-300 mM KCl. Elution of RPA phosphomimetics occurred between 70 ml and 80 ml after injection, with 1 ml fractions collected.

3.2.4 Size Exclusion Chromatography with Multi-Angle Light Scattering

Samples of phosphomimetic RPA were prepared by centrifugation at 13,000 xg for 5 minutes, after which the sample was filtered using a 0.45 µM PVDF membrane syringe filter. If the sample volume was less than 100 µl an insert was used to assist in sample containment. SEC-MALS was performed as described in section 2.2.5. The running buffer for these experiments was HI-300 mM KCl.

3.2.5 Surface Plasmon Resonance

Samples prepared for SPR underwent the full 4 step purification process to ensure that no DNA contamination was present. This was confirmed through $A_{260/280}$, for which all samples had values less than 0.7. SPR data was collected using an OpenSPR (Nicoya Lifesciences). For experiments involving the binding affinity between RPA and DNA the running buffer was phosphate buffered saline (PBS) (Fisher Bioreagents) supplemented with 1 M KCl, making the final concentration of the buffer 137 mM NaCl, 10 mM Na_2HPO_4, 1.8 mM KH_2PO_4, 1002.7 mM KCl, pH 7.4. The regeneration buffer was 10 mM HCl, pH 2. Streptavidin sensor chips (Nicoya Lifesciences) with a 3’ Biotinylated 25-mer ssDNA ligand (sequence 5’-CCACCCCCCCCCCCCCCCCCCCCC-3’) were utilized. Ligand was bound to the sensor chip at 20 µl/min while analyte samples and regeneration buffer flowed at 60 µl/min. After ligand was bound an average of seven injections of regeneration buffer was required to condition the sensor and provide a stable baseline. Injections of the

RPA analyte were performed using two hundred and fifty microliters of 250 nM, 125 nM, 62.5 nM, 31.25 nM, and 15.625 nM concentrations. Kinetic analysis was performed using the TRACEDRAWER software package.

3.3. Results

3.3.1. Expression and Purification

Seventeen combinations of phosphomimetic candidate sites were tested (**Figure 3.5**). After transformation all combinations were checked for expression of all three subunits by SDS-PAGE. Using a HiTrap^{tm} Blue HP column followed by SEC with a Superdex 200 column it was found that 12 combinations would not form stable heterotrimers. Five combinations formed a stable heterotrimer (**Figure 3.6**). These five were put through the full four step purification protocol before being tested for DNA binding affinity using SPR (**Figure 3.7**).

3.3.2. Size Exclusion Chromatography with Multi-Angle Light Scattering

SEC-MALS showed that the combinations wtRPA70+9ERPA32/14, 1EbRPA70+9ERPA32/14, 1EaRPA70+wtRPA32/14, 1EbRPA70+wtRPA32/14, and wtRPA70+2ERPA32/14 had a MW close to the expected 110 kDa (**Table 3.3**).

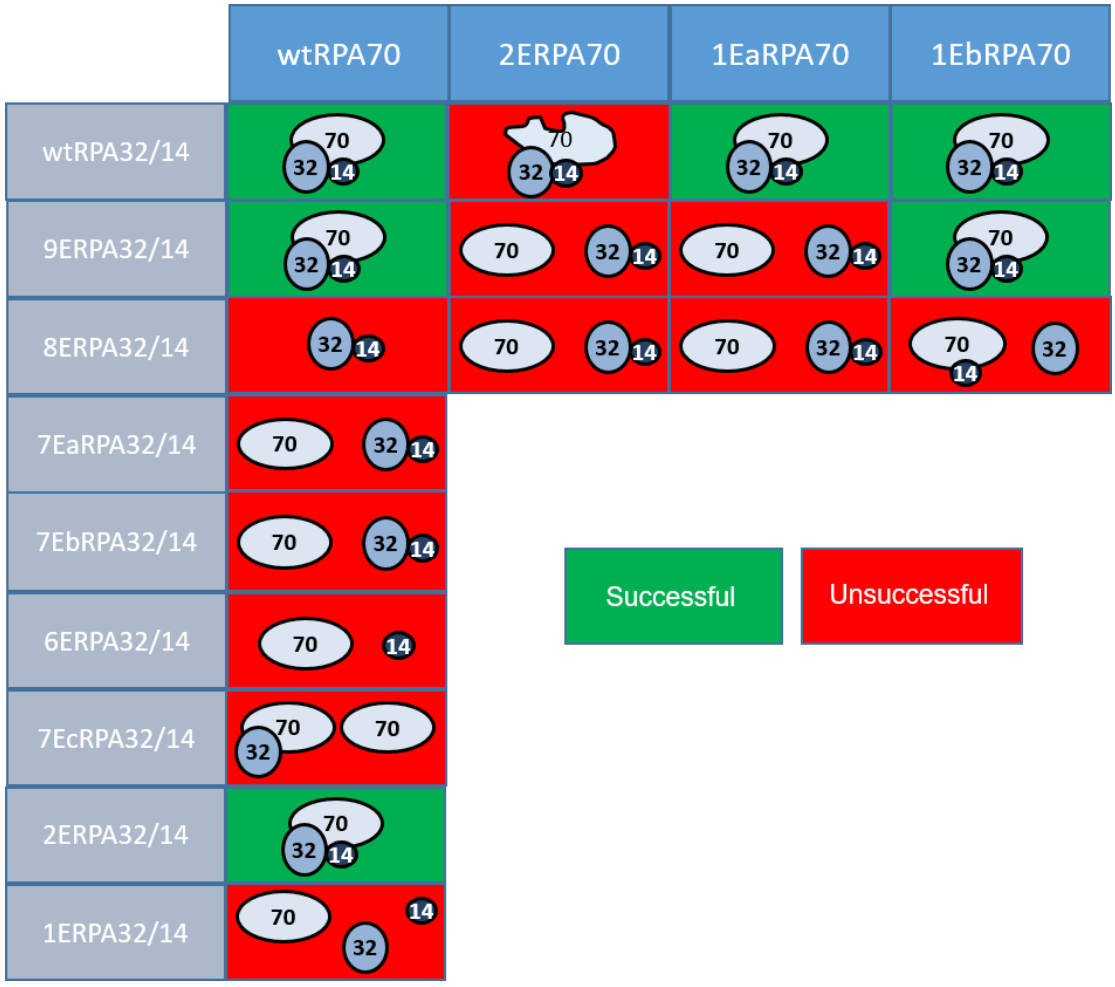


Figure 3.5 RPA phosphomimetic combinations and purification results

RPA70 candidate site combinations are listed in columns (dark blue). RPA32/14 candidate site combinations are listed in in rows (grey). Combinations resulting in purifiable heterotrimers are shown with a green background, with unsuccessful combinations shown with a red background. Cartoons of the RPA70, 32, and 14 subunits indicate if the subunit was present at the end of purification, with overlapping cartoon subunits indicating if they purified in the same fraction.

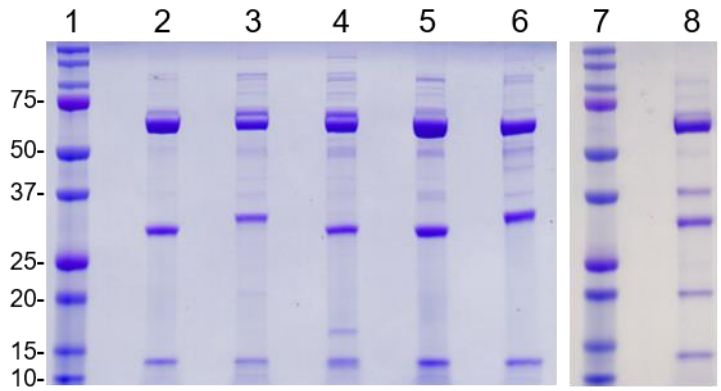


Figure 3.6 Purified RPA phosphomimetics

Proteins were purified using HiTrap^{tm} Blue affinity chromatography and SEC before being run on SDS-PAGE. Gel layout: 1) MW ladder, 2) wtRPA70+wtRPA32/14, 3) wtRPA70+9ERPA32/14, 4) 1EaRPA70+wtRPA32/14, 5)1EbRPA70+wtRPA32/14, 6) 1EbRPA70+9ERPA32/14, 7) MW ladder, and 8) wtRPA70+2ERPA32/14.

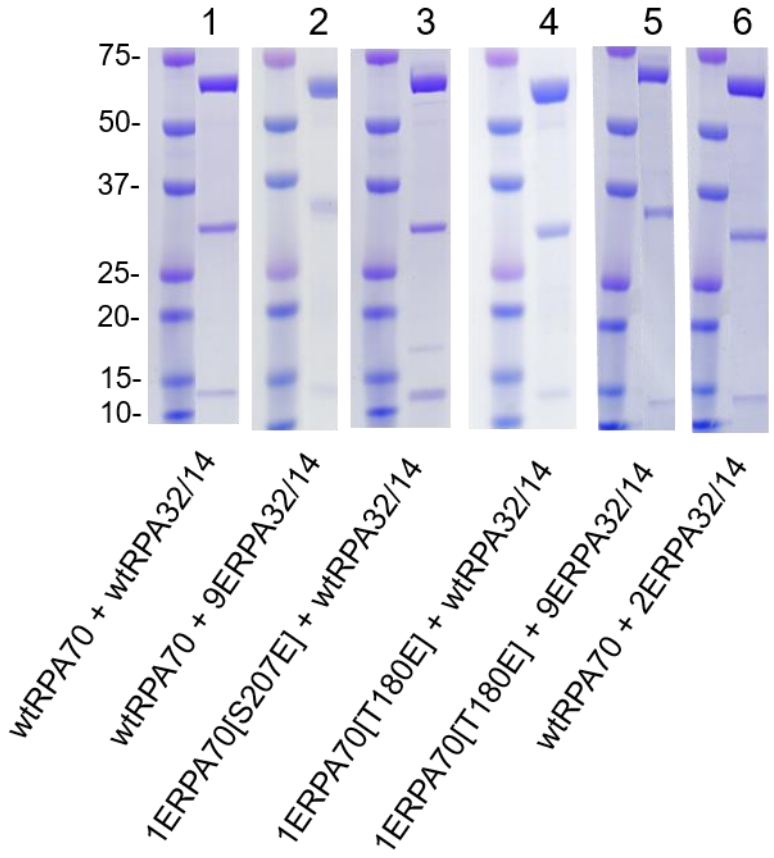


Figure 3.7 Four-step purified RPA phosphomimetics

wtRPA and RPA phosphomimetics that could form stable heterotrimers purified using HiTrap^{tm} Blue affinity, Hitrap^{tm} Q FF affinity, HAP affinity, and SEC.

RPA Phosphomimetic	Purified as a Trimer	$A_{260/280}$	Molecular Weight (kDa)	Polydispersity	mg yield/g of cells
wtRPA70+wtRPA32/14	Yes	.59	105.5	1.000	.27
wtRPA70+9ERPA32	Yes	1.52	105.3	1.003	.16
wtRPA70+2ERPA32/14	Yes	.69	NA	NA	.45
1EaRPA70+wtRPA32/14	Yes	.45	108.6	1.000	.29
1EbRPA70+wtRPA32/14	Yes	.60	108.4	1.000	.14
1EbRPA70+9ERPA32/14	Yes	.82	106.3	1.000	.45

Table 3.3 Purification and polydispersity of RPA phosphomimetics

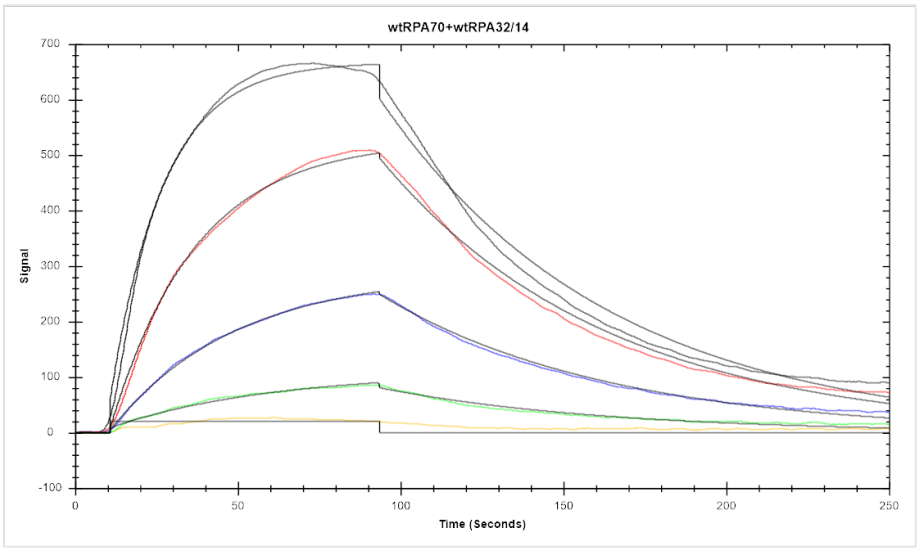
3.3.3. Surface Plasmon Resonance

The affinity of wtRPA for ssDNA is known to be very high, with a K_D 10^{-9} M [126]. Initial SPR experiments with wtRPA confirmed this (data not shown). To promote equilibrium binding 1M KCl was added to the buffer so that differences between each phosphomimetic could be seen, as 1 M KCl partially disrupts the binding between ssDNA and RPA. DNA binding affinity for the combinations of wtRPA70+wtRPA32/14, wtRPA70+9ERPA32/14, and wtRPA70+2ERPA32/14 were similar with K_D values of $6.7x10^{-8}$ M, $7.4x10^{-8}$ M, and $8.8x10^{-8}$ M respectively. In contrast the combinations of 1EaRPA70+wtRPA32/14, 1EbRPA70+wtRPA32/14, and 1EbRPA70+9ERPA32/14 had significantly diminished affinity for ssDNA with values of $15.4x10^{-8}$ M, $13.2x10^{-8}$ M, and $35.5x10^{-8}$ M. The 1EbRPA70+9ERPA32/14 combination required fitting with a two state reaction model (**Figures 3.8, 3.9, 3.10, 3.11 and Table 3.4**).

3.4 Discussion

3.4.1 Outcomes of Purification

The standard RPA purification methods, A stepwise process involving Affigel Blue affinity, HAP affinity, strong anion exchange, and finally SEC, will produce very pure RPA. This process was lengthy and inefficient for the initial testing of the phosphomimetics. The process was streamlined in two ways. First, HAP affinity was removed from the procedures as it was not a robust step. Various combinations of the RPA phosphomimetics would, for reasons unknown, simply flow through the resin instead of binding. Second, QFF was also dropped from the shortened



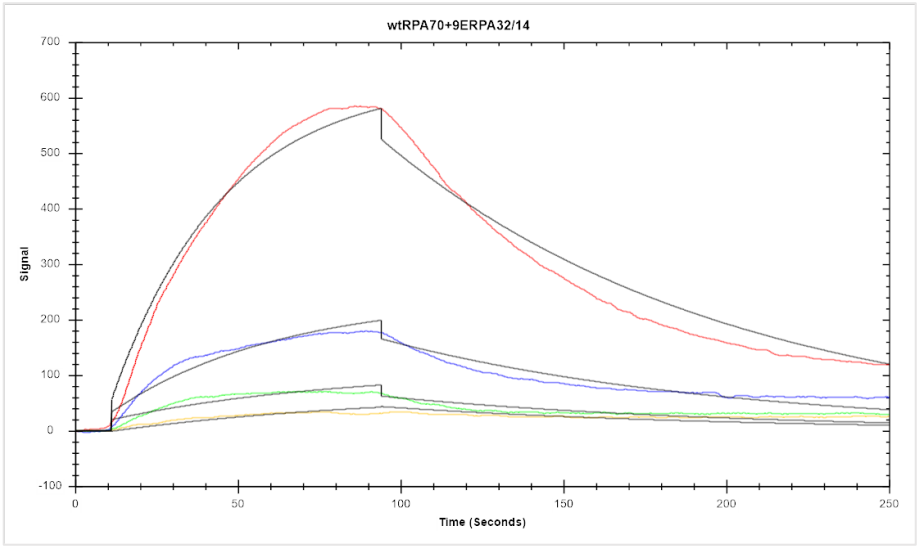
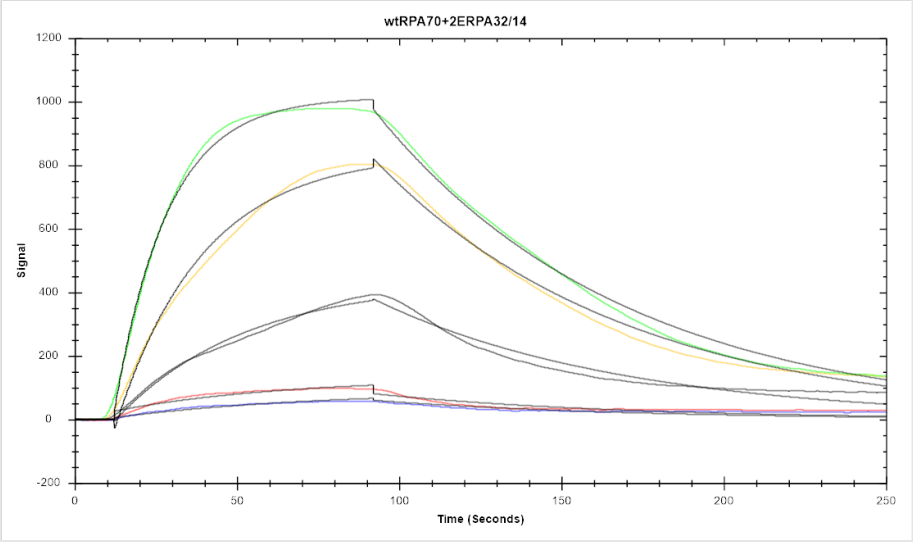


Figure 3.8 SPR curves for wtRPA70+wtRPA32/14 and wtRPA70+9ERPA32/14

Concentrations for each curve are 250 nM (black), 125 nM, (red), 62.5 nM (blue), 31.25 nM (green), and 15.63 nM (yellow).



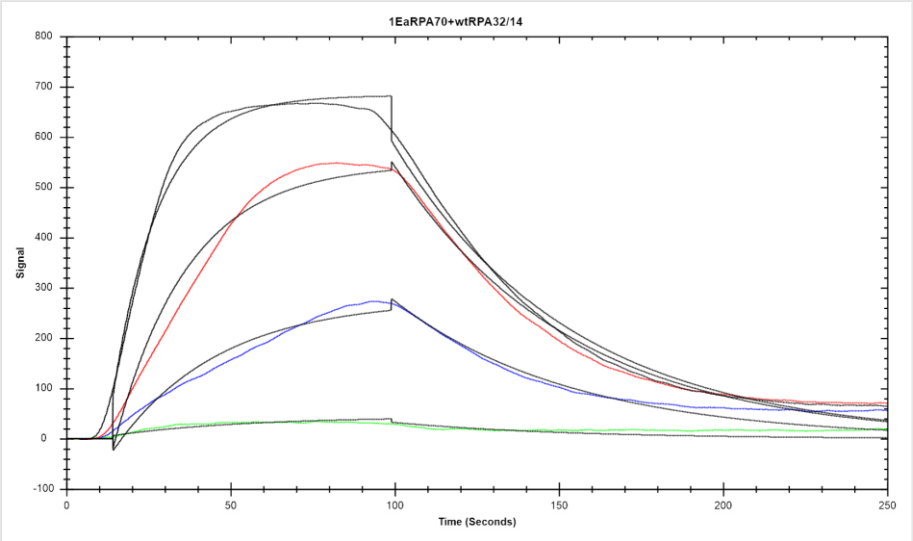
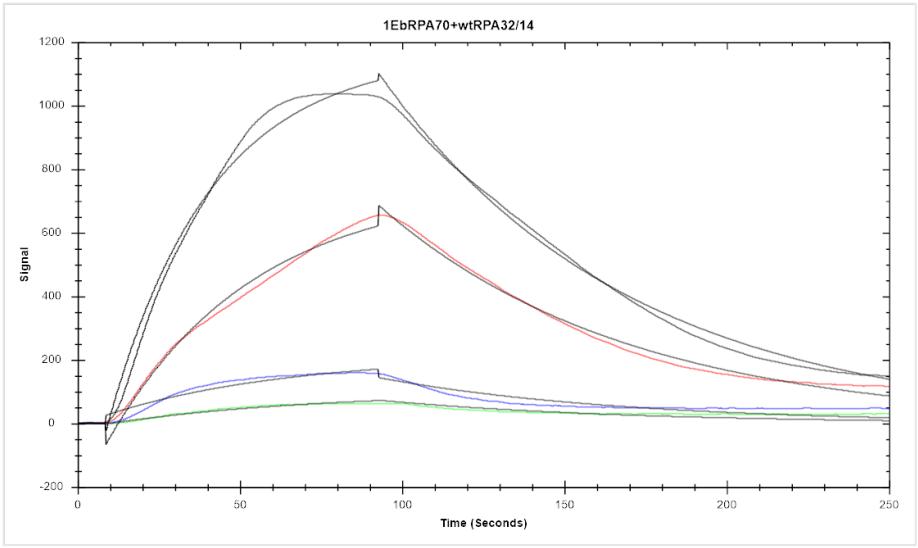


Figure 3.9 SPR curves for wtRPA70+2ERPA32/14 and 1EaRPA70+wtRPA32/14

Concentrations for each curve are top: 250 nM (green), 125 nM, (yellow), 62.5 nM (black), 31.25 nM (red), 15.63 nM (blue), and bottom: 250 nM (black), 125 nM, (red), 62.5 nM (blue), 31.25 nM (green), and 15.63 nM (yellow).



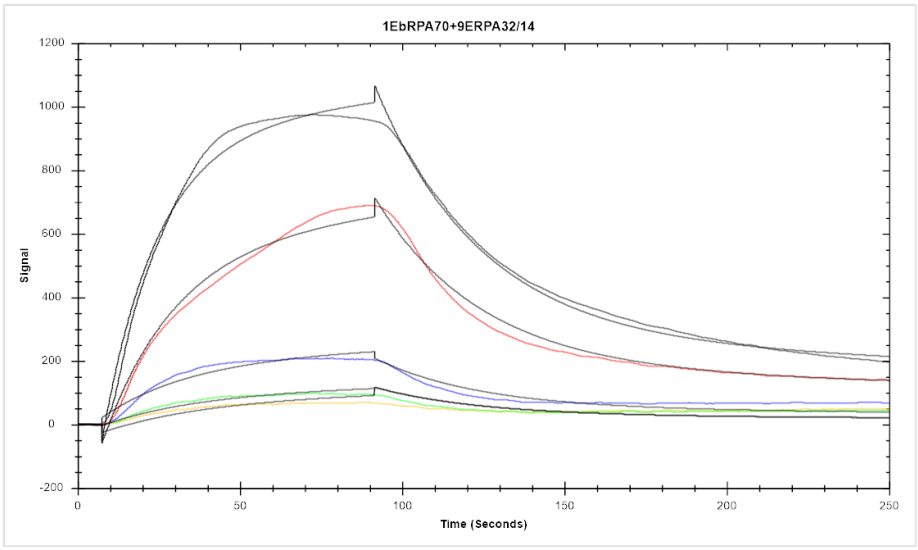


Figure 3.10 SPR curves for 1EbRPA70+wtRPA32/14 and 1EbRPA70+9ERPA32/14

Concentrations for each curve are 250 nM (black), 125 nM, (red), 62.5 nM (blue), 31.25 nM (green), and 15.63 nM (yellow). 1EbRPA70+9ERPA32/14 required fitting with a two-state binding reaction model.

Candidate Site Combinations	K_a (1/(M*s))	K_d (1/s)	K_D (M)
wtRPA70 + wtRPA32/14	$2.5x10^5$ ($4.2x10^4$)	$1.5x10^{-2}$ ($5.5x10^{-4}$)	$6.7x10^{-8}$ ($.65x10^{-8}$)
wtRPA70 + 9ERPA32/14	$1.5x10^5$ ($1.4x10^4$)	$1.1x10^{-2}$ ($3.5x10^{-4}$)	$7.4x10^{-8}$ ($.50x10^{-8}$)
wtRPA70 + 2ERPA32/14	$1.6x10^5$ ($7.6x10^3$)	$1.4x10^{-2}$ ($3.2x10^{-4}$)	$8.8x10^{-8}$ ($.22x10^{-8}$)
1EaRPA70 + wtRPA32/14	$1.1x10^5$ ($6.7x10^3$)	$1.7x10^{-2}$ ($1.5x10^{-4}$)	$15.4x10^{-8}$ ($.86x10^{-8}$)
1EbRPA70 + wtRPA32/14	$9.9x10^4$ ($5.4x10^3$)	$1.3x10^{-2}$ ($3.5x10^{-4}$)	$13.2x10^{-8}$ ($.72x10^{-8}$)

Two-State Binding					
Candidate Site Combinations	K_{a1} (1/(M*s))	K_{d1} (1/s)	K_{a2} (1/s)	K_{d2} (1/s)	K_D (M)
1EbRPA70 + 9ERPA32/14	$4.6x10^4$ ($4.5x10^3$)	$3.7x10^{-2}$ ($3.3x10^{-3}$)	$4.1x10^{-3}$ ($1.5x10^{-4}$)	$1.8x10^{-3}$ ($3.1x10^{-5}$)	$35.5x10^{-8}$ ($5.8x10^{-8}$)

Table 3.4 Phosphomimetic RPA DNA binding affinity

Average values from three experiments are given with standard deviation in parenthesis. Experiments were done in triplicate. Only the 1EbRPA70+9ERPA32/14 required fitting with a two-state binding reaction model.

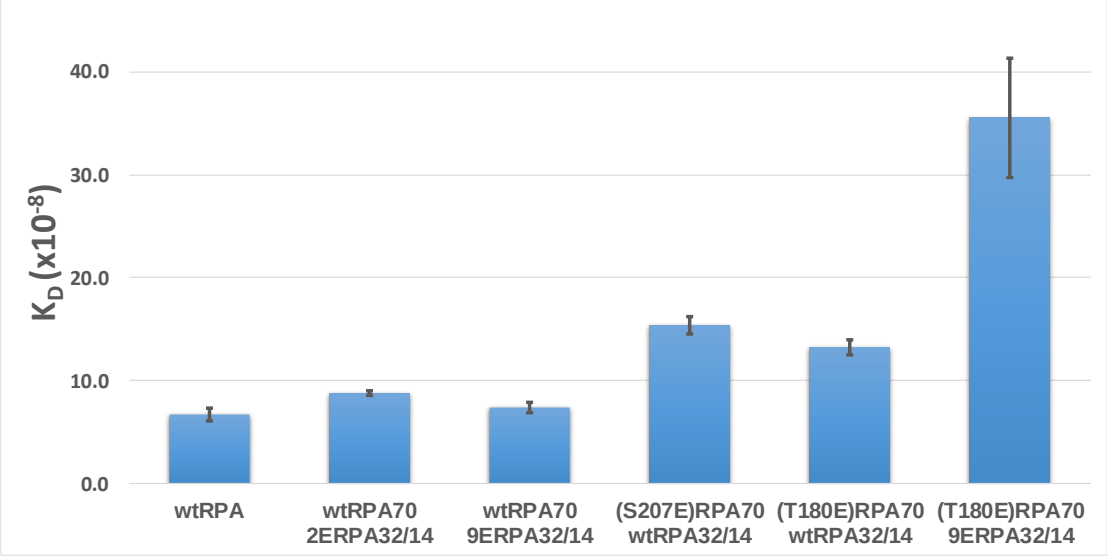


Figure 3.11 Respective K_D values of phosphomimetic RPA combinations

wtRPA70+wtRPA32/14 had a K_D of 6.7x10^{-8} M.

wtRPA70+2ERPA32/14 had a K_D of 8.8x10^{-8} M.

wtRPA70+9ERPA32/14 had a K_D of 7.4x10^{-8} M.

1Ea(S207E)RPA70+wtRPA32/14 had a K_D of 15.4x10^{-8} M.

1Eb(T180E)RPA70+wtRPA32/14 had a K_D of 13.2x10^{-8} M.

1Eb(T180E)RPA70+9ERPA32/14 had a K_D of 35.5x10^{-8} M.

protocols as it would purify the phosphomimetics regardless of whether they formed a heterotrimer or not. QFF fractions containing RPA also contained enough impurities to obscure the RPA subunits when examined by SDS-PAGE, making it difficult to select fractions to use if further purification steps. For these reasons HiTrap^{tm} Blue affinity was kept as the first step, but the second and final step for these experiments was SEC. HiTrap^{tm} Blue affinity allowed for not only the initial robust selection of the RPA phosphomimetics, but it also provided the step where DNA contamination was removed. SEC proved to be the perfect second step as it was a gentle yet thorough procedure to ensure that the recovered RPA was bound together in its heterotrimer form. RPA phosphomimetics that would not form stable heterotrimers had their subunits purify into separate fractions as monomers or heterodimers.

We tested 17 combinations of RPA70 and RPA32/14 candidate sites, of which 12 proved unstable and could not be purified into a heterotrimer by the method described above (**Figure 3.5**). Failure of the RPA phosphomimetics to form a stable heterotrimer was most often correlated with a distinct overexpression of the RPA32 subunit compared to total soluble protein when checked by SDS-PAGE. Of the 12 combinations that were unable to form a heterotrimer only four did not overexpress RPA32, namely, 2ERPA70+wtRPA32/wtRPA14 wtRPA70+6ERPA32/14, wtRPA70+7EcRPA32/14, and wtRPA70+1ERPA32/14. It is predicted that phosphomimetics that were unable to form a stable heterotrimer but were still able to pass the Affigel Blue affinity step were able to find a binding partner to stabilize their interaction with the Affigel Blue resin or were stabilized by the resin itself. There were 7 ways that these combinations failed to form heterotrimers. (1) The combinations of 2ERPA70+9ERPA32/14, 2ERPA70+8ERPA32/14, 1EaRPA70+9ERPA32/14, 1EaRPA70+8ERPA32/14, wtRPA70+7EaRPA32/14, and wtRPA70+7EbRPA32/14 all produce RPA70 in a separate purified fraction and a stable

RPA32/14 heterodimer. (2) The wtRPA70+8ERPA32/14 combination produced only a stable RPA32/14 heterodimer with no RPA70 detectable in any fractions. (3) The 1EbRPA70+8ERPA32/14 combination showed the RPA70/14 subunits in a single fraction with RPA32 alone in a separate fraction. (4) 2ERPA70+wtRPA32/14 assembled into a trimer but the RPA70 subunit showed significant degradation. (5) wtRPA70+6ERPA32/14 generated RPA70 and RPA14 in separate fractions with no RPA32 in any fractions. (6) wtRPA70+7Ec RPA32/14 exhibited RPA70 and RPA32 together in an early fraction, and then RPA70 in multiple later fractions without RPA32 or RPA14 present. (7) Finally, wtRPA70+1ERPA32/14 yielded RPA70, RPA32, and RPA14 all in separate fractions. As RPA32/14 have been found to be soluble on their own but RPA70 is not this makes circumstances 1,3, and 5-7 stand out, as they involve either RPA70 alone or with an incomplete RPA32/14 [126, 127]. The successful purification of the control combination wtRPA70 and wtRPA32/14 indicates that the above failures were not due to the pET Duet system used for expression. Indeed, 5 of the 17 RPA70 and RPA32/14 phosphomimetic combinations were able to be purified as intact heterotrimers and were as follows: wtRPA70+9ERPA32/14, 1EbRPA70+9ERPA32/14, 1EaRPA70+wtRPA32/14, 1EbRPA70+wtRPA32/14, and wtRPA70+2ERPA32/14. When examined by SDS-PAGE, mutated subunits run noticeably higher, which is typical. Only the wtRPA70+9ERPA32/14 candidate combination purified with DNA contamination as indicated by its $A_{260/A280}$ of 1.52 (**Table 3.3**). This implies that this combination has retained ssDNA through the purification process and means that it may have different ssDNA binding properties. When this combination was taken through all four steps of purification the DNA contamination was removed, allowing it to still be tested for DNA affinity by SPR in later experiments. Examination of these proteins by SEC-MALS showed that all tested combinations were

monodisperse except for the wtRPA70+9ERPA32/14 combination, which is likely a result of impurities or ssDNA contamination.

3.4.2 RPA Phosphoisoforms in G2 phase in control cells

The majority of RPA in cells exists in a phosphorylated state. Currently, the majority of cell-free experiments do not reflect this in their methodology, choosing to instead use wtRPA for protein:protein interactions. Experiments with physiologically relevant forms of RPA require the use of stable phosphomimetics. The phosphomimetic combinations of wtRPA70+1ERPA32 and wtRPA+2ERPA32/14 were tried. These combinations were selected due to the constant phosphorylation of S23 or S23 and S33 due to the position in the cell cycle (**Figure 3.2**). We were surprised to find the combination of wtRPA70+1ERPA32 would not purify, as western blotting showed its presence in control cells, and capillary isoelectric focusing has confirmed that species of RPA with a single phosphorylation can be found in the cell (**Figure 3.1**). The combination of wtRPA+2ERPA32/14 was purifiable. When this combination was tested by SPR it was found to have ssDNA affinity lower than but still similar to that of wtRPA. The K_D values were $8.8x10^{-8}$ M and $6.7x10^{-8}$ M respectively (**Table 3.4**). It is possible that this combination reflects one of the phosphorylation states observed in control cells (**Figure 3.1**, blue line).

3.4.3 RPA Phosphoisoforms in G2 After Induction of Double Strand Breaks

There are up to 14 phosphorylation sites on nine isoforms of RPA in cells with DSBs. (**Figure 3.1,** pink line). Out of the sixteen DSB phosphomimetic combinations tested, we found four isoforms that would purify as stable heterotrimers:

wtRPA70+9ERPA32/14, 1EaRPA70+wtRPA32/14, 1EbRPA70+9ERPA32/14, and 1EbRPA70+wtRPA32/14. When DNA binding affinity was examined by SPR we discovered some very interesting results. Phosphomimetic combinations containing mutations solely on the RPA32 subunit had K_D values similar to the $6.7x10^{-8}$ M of wtRPA, as wtRPA70+9ERPA32/14 and wtRPA+2ERPA32/14 had K_D values of 7.4x10-8 and 8.8x10-8 respectively. Combinations with Glu mutations on RPA70 significantly decrease the ssDNA binding affinity when compared to wtRPA (**Table 3.4, Figure 3.11**). This in itself is not terribly surprising, as the majority of the ssDNA binding affinity for RPA is the result of DBD-A and –B on the RPA70 subunit, so modifications here would likely have a stronger result. The combinations of 1EaRPA70+wtPA32/14 and 1EbRPA70+wtRPA32/14 had values of $15.4x10^{-8}$ M and $13.2x10^{-8}$ M respectively. These values are almost double that of wtRPA. There was surprising data with the fact that 1EbRPA70+9ERPA32/14, the phosphomimetic RPA with mutations on both the RPA70 and the RPA32 subunit, had such a drastic drop in ssDNA affinity compared to all the others. The 1EbRPA70+9ERPA32/14 combination had a K_D $35.5x10^{-8}$ M, significantly higher than even 1EaRPA70+wtPA32/14 and 1EbRPA70+wtRPA32/14 (**Figure 3.4, Figure 3.11**). A second unexpected detail about this combination is that it required fitting with a two state reaction model. This indicates this combination is acting as if the initial binding of ssDNA is enhancing the binding of a second ssDNA binding site, a phenomenon not observed in the other tested phosphomimetic combinations. From interpreting all of this data together we find an interesting possibility. All of the RPA phosphomimetics that contained RPA70 mutations had significantly reduced ssDNA binding affinity. It is therefore possible that the phosphorylation state that allows for RPA to pass ssDNA to RAD52 in the HR pathways is based on RPA70 phosphorylation. As the DBDs with the highest ssDNA affinity on RPA are located on the RPA70 subunit, and phosphomimetic

mutations to this subunit result in the most drastic changes in K_D, it is possible that these mutations are somehow changing the way that RPA interacts with DNA through a conformational shift. This conformational shift may be altering RPA:ssDNA binding in a fashion that is more conducive to other proteins binding or interacting with the ssDNA and replacing RPA on the ssDNA.

Chapter IV: Conclusion and Future Directions

4.1 Conclusion

RAD52(1-303) has been purified to a quality that renders it monodisperse and of a quality compatible with SAXS analysis. Using the orientation we found for the RAD52(1-212) molecule we were able to extend that alignment to the RAD52(1-303) SAXS envelope. This places the DNA binding region of RAD52 facing upwards from the larger spherical region while the unstructured C-terminal strands occupy the lower, smaller diameter space. SAXS structures are low resolution however, so further experiments need to be performed to learn more about how the RAD52 molecule behaves, specifically how it interacts with phosphorylated and unphosphorylated RPA.

RPA is usually phosphorylated in normal cells, with between one and four phosphorylations being the most common isoforms. When DSBs are present up to 14 phosphorylations can be detected. However, cell free experiments using RPA do not reflect this, and often use RPA with no post-translational modifications. If the interactions of RPA are to be accurately investigated, then information on the effects of RPA's phosphorylation state need to be discovered. Using western blot data [69] from cells without DSBs identifying RPA with a single phosphorylation (S23) in S phase and two phosphorylations (S23 and S33) in G2 phase we created Glu phosphomimetic mutants of those combinations. Surprisingly, the S23 mutant was unstable and could not be purified. The S23/S33 mutant could be purified and showed a DNA binding affinity akin to that of wtRPA. As both of these are CDK phosphorylation sites and S23 phosphorylation has been identified without accompanying S33 phosphorylation, it is surprising that a mutation to a single site induces instability while mutation of both of them results in a stable protein. As this phosphomimetic may represent RPA with a basal level of phosphorylation, further characterization experiments should be performed. After narrowing down a list of DSB

induced phosphorylation sites to eleven candidate sites we tested 15 different combinations of those sites. The four combinations of wtRPA70+9ERPA32/14, 1EaRPA70+wtRPA32/14, 1EbRPA70+9ERPA32/14, and 1EbRPA70+wtRPA32/14 were found to form stable heterotrimers. ssDNA affinity of these combinations was tested using SPR, and it was found that 1EaRPA70+wtRPA32/14,1EbRPA70+wtRPA32/14, and 1EbRPA70+9ERPA32/14 all had significantly lower ssDNA binding affinity than wtRPA. In addition, 1EbRPA70+9ERPA32/14 had significantly reduced ssDNA affinity when compared to both 1EaRPA70+wtRPA32/14 and 1EbRPA70+wtRPA32/14, as well as required fitting with a two state reaction model. All other tested combinations did not require this. These results are surprising because they show that the significant decreases seen in these mutants all involved a mutation to the RPA70 subunit. It is possible that these RPA70 mutations are inducing some form of conformational change that could make the ssDNA more accessible to other DNA binding proteins, or at least makes it easier for other DNA binding proteins to remove RPA from the ssDNA strand. If that is the case, then it is possible that one of the RPA70 phosphomimetic mutations tested in these experiments is the site responsible for allowing the transfer of ssDNA from RPA to RAD52 in HR. More experiments are required to characterize these phosphomimetic mutants to find the correct DSB relevant form of RPA.

4.2 Future Directions

Structural analysis of RAD52(1-303) can be improved through a number of methods. The first of these would be by performing the X-ray scattering of either the whole RAD52(1-303) molecule or just the RAD52(212-303) domain in the presence of a strong magnetic field. This would restrict the motion of the CTDs allowing for potential fiber diffraction (personal communication with Scott Barton of SAXSLAB and [128]).

Experiments to break up the RAD52 ring also show potential to produce more SAXS structures to be used for further identification of the various regions of RAD52(1-212) and RAD52(1-303). The small molecule 6-hydroxy-dopa has been previously identified to break the undecameric ring into dimers [129]. This drastic change in oligomer state, if it could be purified to a quality compatible with SAXS analysis, would give a very different view of the RAD52(1-303) molecule and its C-terminal region. If the results could be made monodisperse, it may even be possible to use this technique to acquire SAXS data on full length RAD52.

The addition of RPA(172-270) (the wHLH domain) to RAD52(1-303) would also be of interest. The wHLH domain could bind to the previously identified CTD sequence, and this "decoration" of the CTD could restrict movement and alter the SAXS envelope, highlighting the bound region. This kind of experiment may also have an unexpected outcome, as recent experiments have shown that the K_D between RAD52(1-303) and the RPA32(172-270) is approximately $1.2x10^{-9}$ M using SPR (Mona Al-Mugotir, unpublished data). This is a stark contrast to the information presented in the Mer *et al.* (2000) which said that the binding between RPA(172-270) and RAD52(257-274) had a $K_D \approx 10^{-6}$ M. A binding simulation (G. Borgstahl, unpublished data) between RPA(172-270) and RAD52(1-303) did not show RPA(172-270) binding the RAD52(257-274) region in the CTD, but instead showed it binding in the region where the RAD52 CTD connects to the RAD52(1-212) crystal structure. This information is worth following up on, as depending on the actual properties of this binding the wHLH domain could be what breaks up the RAD52 ring. If this is true, that means it could be possible to form crystals of RAD52(1-303) and RPA(172-270), and discover the true interaction between these two molecules. Experiments to use SEC-MALS to investigate the interaction between RAD52(1-303) and RPA(172-270) were attempted, but not completed. It was found that salt concentrations

low enough to promote binding between the two molecules also promoted binding between RAD52(1-303) and the SEC column (unpublished data). These experiments were attempted with a WTC-050S5 column but the protein bound to the column. The experiments should be continued with a different SEC column, for instance one that is Superose based such as the one used in Deng *et al.* (2009).

Concerning RPA phosphomimetics, first and foremost the table of the phosphomimetic combinations presented in this document is incomplete (**Figure 3.5**). There is a chance that other combinations which can form stable trimers have not been identified. To this end, the rest of the table should be filled in by a future student to identify any other stable isoforms of RPA phosphomimetics.

We are currently generating preliminary data on phosphomimetic RPA binding with RAD52 using SPR. For these experiments RAD52 is bound to a NTA sensor chip and the RPA phosphomimetics are tested for interaction. If an RPA phosphomimetic is found to have different or improved binding with RAD52 then it can be used to plan follow-up experiments on the ssDNA:RPA:RAD52 complex as employed by Deng & coworkers [130]. The RPA phosphomimetic that affects the complex in a similar way to truly phosphorylated RPA could be identified. Deng noted a handoff of ssDNA from RPA to RAD52 when RPA was phosphorylated; our data suggests this could be caused by RPA70 phosphorylation.

Dr. Alexander Mazin at Drexel University will be testing the ability of the RPA phosphomimetics to promote inverse strand exchange with DNA and RNA using the assay described in his paper [92], further assisting with the identification and validation of the appropriate RPA phosphomimetic.

The identification of the relevant RPA phosphomimetic will allow for the SAXS analysis of the ssDNA:RPA:RAD52 complex, giving insight into the mechanism behind

how this complex forms, including the characteristics of the actual binding surfaces between these proteins.

Finally, the activity of the phosphomimetics could be investigated in human cells and tested for biological phenotype. RPA could be knocked out, and rescue experiments performed to identify how each of the previously validated phosphomimetics interacts with the various DNA repair pathways. This would be very difficult to do, as RPA is an essential protein and interference in its activities is extremely harmful to cells. These experiments are possible though, and clustered regularly interspaced short palindromic repeats (CRISPR)/ CRISPR associated protein 9 (Cas9) (CRISPR/Cas9) gene editing combined with short hairpin DNA (shDNA) knockdown targeting the 5’ untranslated region (UTR) of the wtRPA subunits show promise for conducting these experiments.

Appendix 1

Introduction

RPA contains two domains which are traditionally responsible for its interaction with other proteins. The first of these is the RPA70(1-120) region, also known as domain F, and RPA32(200-270), known to contain the wHLH domain. Structural evidence exists of interaction between the wHLH domain and peptides of the UNG, XPA, and RAD52 proteins [54]. This protein's affinity for RAD52 makes it a point of interest for determining the full structural interaction between RPA and RAD52. To this end, attempts to crystallize RPA70(172-270) and examine the stoichiometry of RAD52(1-303)/RPA32(172-270) interaction were attempted.

Expression of RPA32(204-270)

The sequence for RPA32(204-270) was cloned into pET28a vector by Dr. Marc Wold (GenScript) and was a gift. BL21(DE3) *E. coli* cells were transformed using the manufacturers protocol (Novagen). Selection of cells occurred through growth on agarose plates containing 30 µg/ml kanamycin. Surviving colonies were selected and grown in LB media incubated at 37 °C and shaken at 170 RPM. Cells were grown to an OD600 between 0.7 and 0.9, after which they were induced with 1 mM IPTG and incubated for 4 hours at 37 °C. Cells were pelleted by centrifugation at 14000 xg and stored at -20° C.

Purification of RPA32(204-270)

RPA32(204-270) was purified as in Mer G. *et. al* (2000) with the following changes. A HisTrap^{tm} HP column was equilibrated with 5 CV of 10 mM $NaPO_4$, pH 7.5, 150 mM NaCl, 10 mM imidazole, 2 mM BME, running buffer before the protein was loaded. Elution occurred using a gradient to 1M imidazole. The protein peak was collected and dialyzed

into 20mM NaPO4, 50mM NaCl, 2mM BME, pH 7.5 running buffer for anion exchange. A HiTrap^{tm} Q FF column was used for the anion exchange step. The running buffer was the same as the dialysis buffer listed above, and elution was accomplished by a gradient to 1M NaCl. Elution of the protein occurred at 335 mM NaCl. At this point the protein was a single band when examined by SDS-PAGE. Proteins to be used for crystallography were additionally purified by SEC using HiLoad 16/60 Superdex 75 pg (GE Lifesciences). The running buffer for SEC was 10mM Tris-HCl, 170mM NaCl, 2mM BME, 0.5mM EDTA. Due to a lack of Trp amino acids in this protein and the fact that the Bradford as `say was not found to be reliable, it was necessary to use a Microplate BCA^{tm} Protein Assay Kit – Reducing agent Compatible (ThermoScientific) to determine the concentration of the protein. This assay was performed as per the manufacturer's instructions, and was read on a EL$_x$808 plate reader.

Dynamic Light Scattering of RPA32(204-270)

DLS was performed as described in section 2.2.3. The protein was easily made monodisperse using this purification protocol, and would typically have a polydispersity of 12.5% (**Figure A.1**).

SEC-MALS of RPA32(204-270)

SEC-MALS was performed as described in section 2.2.4. Experiments involving the binding of RAD52(1-303) and RPA32(204-270) were done with multiple running buffers. It was determined that a salt concentration of 175 mM KCl was too high to allow interaction between the proteins. The salt type and concentration was adjusted over multiple experiments until reaching a low salt value of 50 mM NaCl. During this process the proteins switched from not interacting with each other, to having RAD52 bind strongly

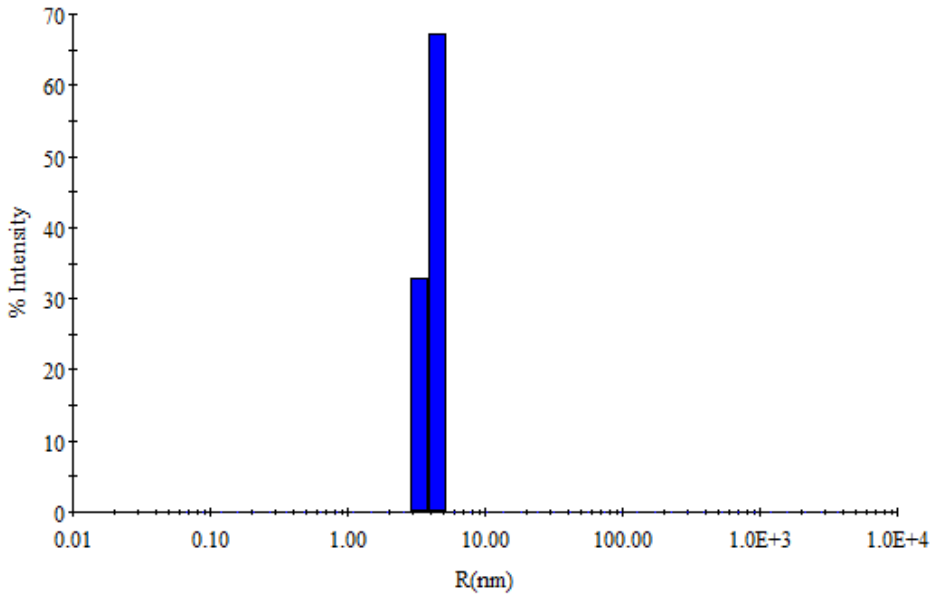


Figure A.1 RPA32(204-270) DLS

The average polydispersity of monodisperse wHLH fractions was 12.5%

with the column, and at no middle value could the proteins be detected as interacting. A column using different packing material should be used for future experiments.

Crystal Trays of RPA32(204-270)

RPA wHLH that retained its monodispersity was concentrated to 20 mg/ml and used for setting up crystal trays. The majority of these trays showed that RPA wHLH was more soluble than expected, and did not precipitate very often at the concentrations used. The following commercial crystallization condition screens were attempted: NeXtal Anions, NeXtal MPD, NeXtal pH Clear 2, NeXtal The Classics, NeXtal The Classics Lite, NeXtal AmSO4, NeXtal Cations, NeXtal PEGS, NeXtal PEGSII, NeXtal Cryos, NeXtal Cryos 1 & 2, Hampton Index, Hampton Index II, Hampton HT, Hampton JCSG Core I, Emerald Biosystems Cryo 1 & 2, and Molecular Dimensions Midas.

Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis Results of Tested Phosphomimetic Combinations

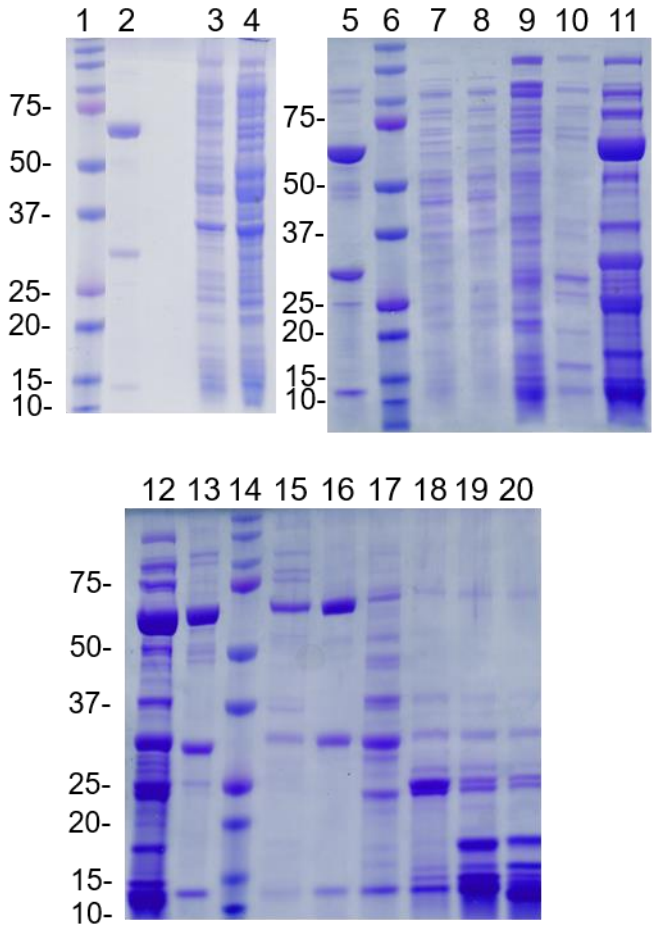


Figure A.2 SDS-PAGE of wtRPA70+2ERPA32/14

1) Ladder, 2) wtRPA, 3) pre-induction cells, 4) post-induction cells, 5) wtRPA, 6) ladder, 7) cell lysate, 8) Affigel Blue flow-through, 9) 800 mM KCl fraction, 10) 500 mM NaSCN fraction, 11) 1500 mM NaSCN fraction, 12) SEC input sample, 13) wtRPA, 14) ladder, 15-20) SEC fractions

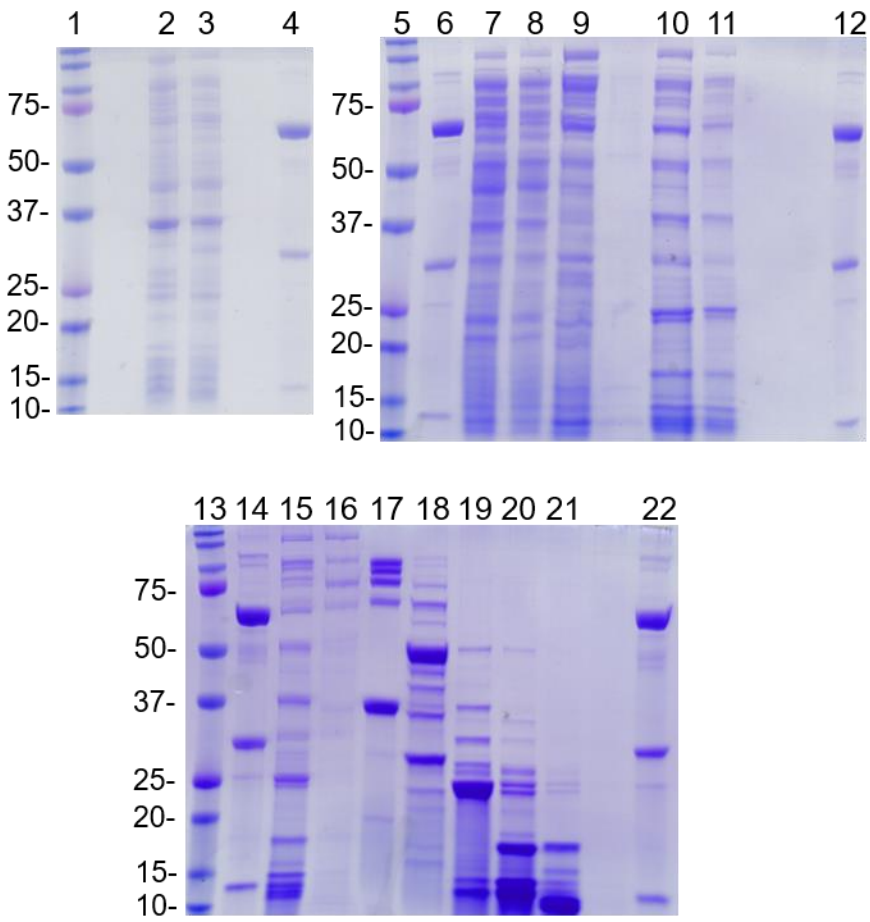


Figure A.3 SDS-PAGE of wtRPA70+1ERPA32/14

1) Ladder, 2) pre-induction cells, 3) post-induction cells, 4) wtRPA, 5) ladder, 6)wtRPA, 7) cell lysate, 8) Affigel Blue flow-through, 9) 800 mM KCl fraction, 10) 500 mM NaSCN fraction, 11) 1500 mM NaSCN fraction, 12) wtRPA, 13) ladder, 14) wtRPA, 15-21) SEC fractions, 22) wtRPA

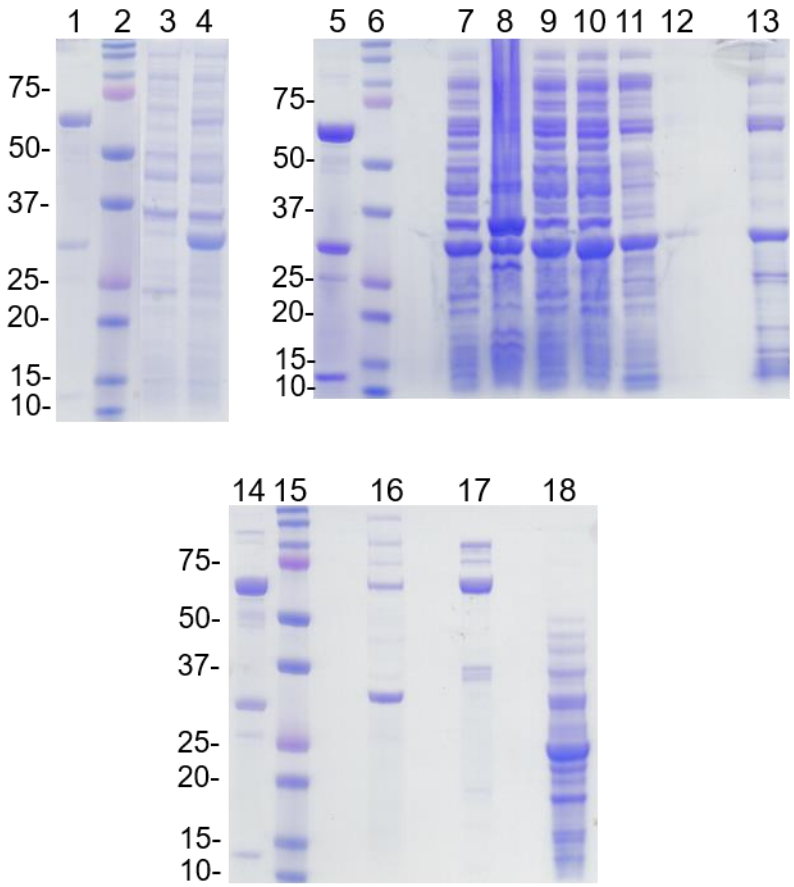


Figure A.4 SDS-PAGE of wtRPA70+7EcRPA32/14

1) wtRPA, 2) ladder, 3) pre-induction cells, 4) post-induction cells, 5) wtRPA, 6) ladder, 7) cell lysate, 8) cell pellet, 9) clarified supernatant, 10) Affigel Blue flow-through, 11) 800 mM KCl fraction, 12) 500 mM NaSCN fraction, 13) 1500 mM NaSCN fraction, 14) wtRPA, 15) ladder, 16-18) SEC fractions

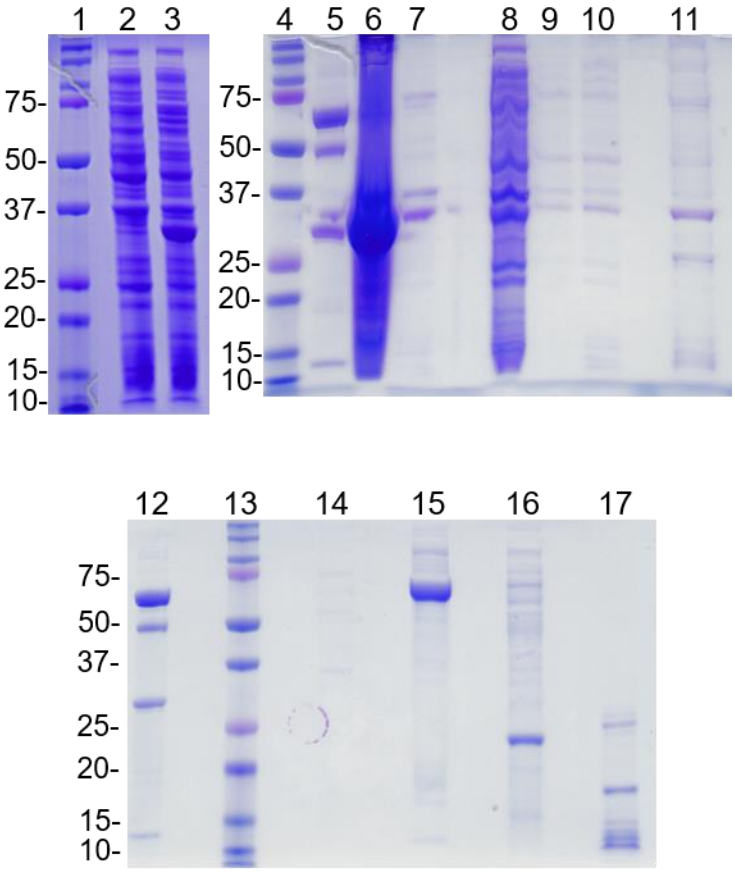


Figure A.5 SDS-PAGE of wtRPA70+6ERPA32/14

1) Ladder, 2) pre-induction cells, 3) post-induction cells, 4) ladder, 5) wtRPA, 6) cell pellet, 7) cell lysate, 8) Affigel Blue flow-through, 9) 800 mM KCl fraction, 10) 500 mM NaSCN fraction, 11) 1500 mM NaSCN fraction, 12) wtRPA, 13) ladder, 14-17) SEC fractions

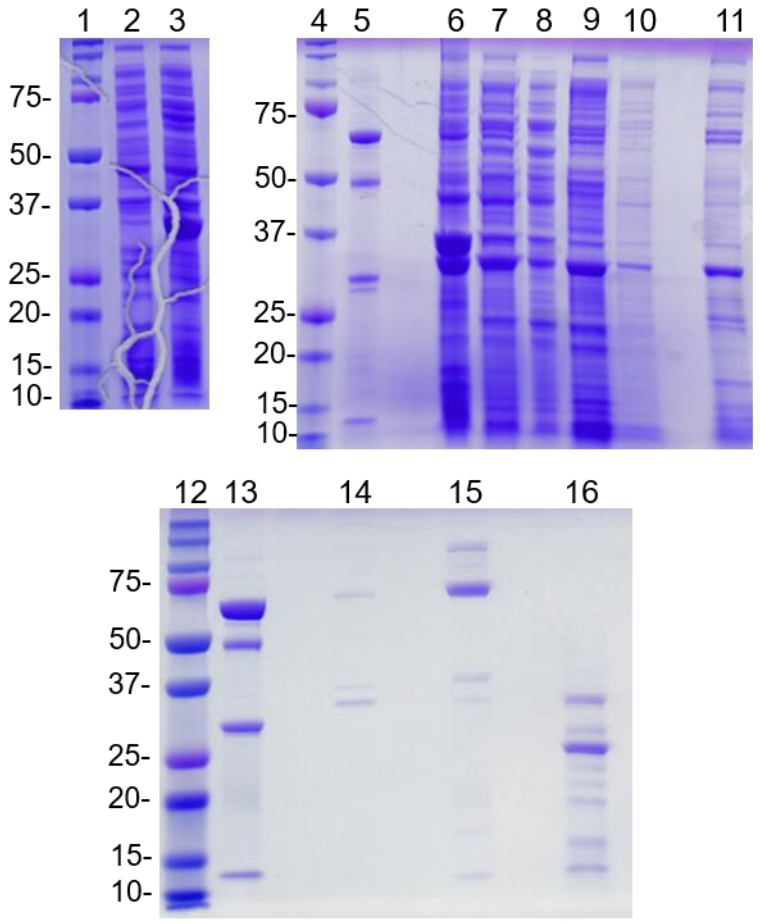


Figure A.6 SDS-PAGE of wtRPA70+1EbRPA32/14

1) Ladder, 2) pre-induction cells, 3) post-induction cells, 4) ladder, 5) wtRPA, 6) cell pellet, 7) cell lysate, 8) Affigel Blue flow-through, 9) 800 mM KCl fraction, 10) 500 mM NaSCN fraction, 11) 1500 mM NaSCN fraction, 12) ladder, 13) wtRPA, 14-16) SEC fractions.

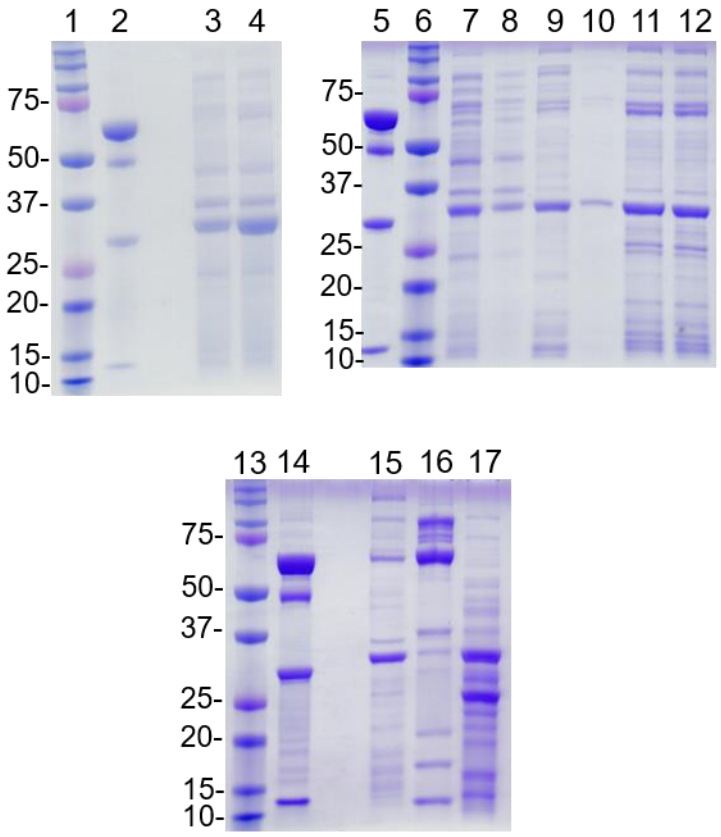


Figure A.7 SDS-PAGE of wtRPA70+7EaRPA32/14

1) Ladder, 2) wtRPA, 3) pre-induction cells, 4) post-induction cells, 5) wtRPA, 6) ladder, 7) clarified cell lysate, 8) Affigel Blue flow-through, 9) 800 mM KCl fraction, 10) 500 mM NaSCN fraction, 11) 1500 mM NaSCN fraction, 12) SEC input sample, 13) ladder, 14) wtRPA, 15-17) SEC fractions

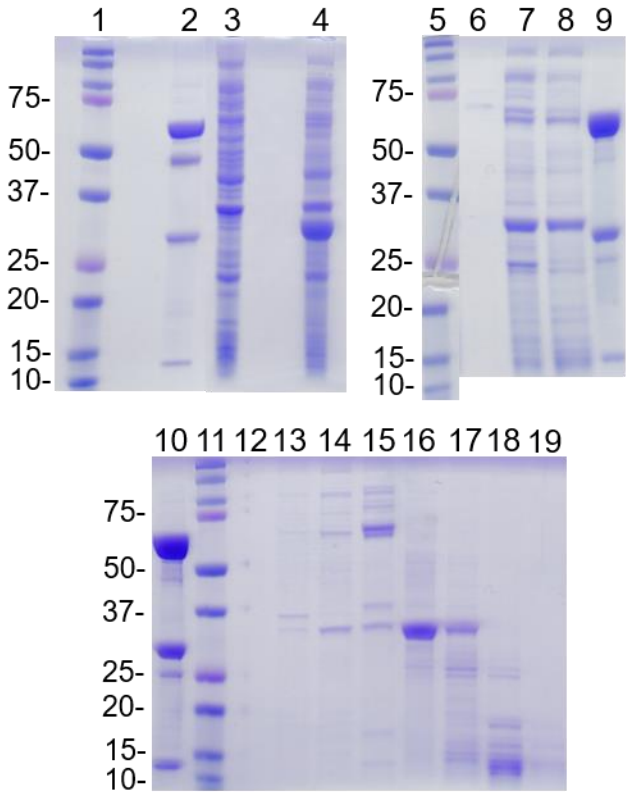


Figure A.8 SDS-PAGE of 1EbRPA70+8ERPA32/14

1) Ladder, 2) wtRPA, 3) pre-induction cells, 4) post-induction cells, 5) ladder, 6) 800 mM KCl fraction, 7) 500 mM NaSCN fraction, 8) 1500 mM NaSCN fraction, 9-10) wtRPA, 11) ladder, 12-19) SEC fractions

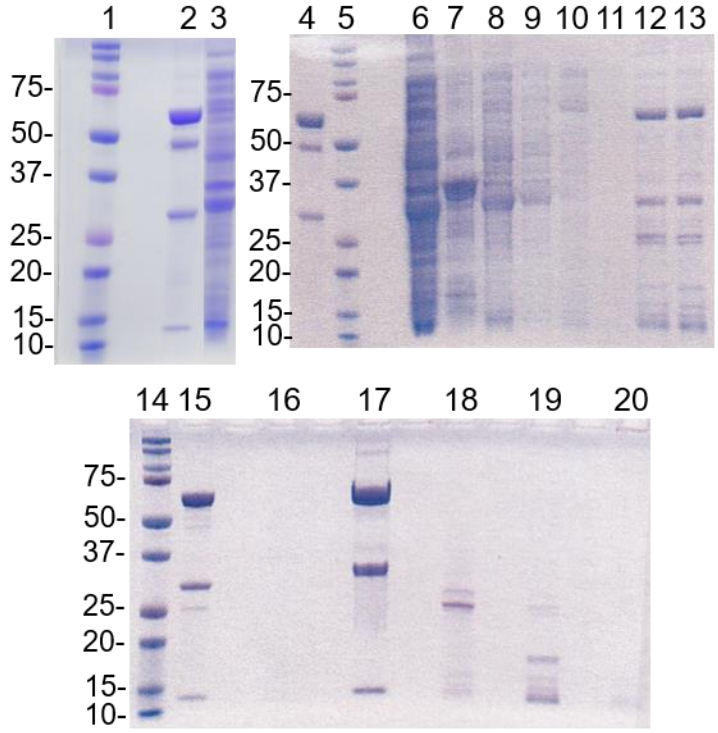


Figure A.9 SDS-PAGE of 1EbRPA70+9ERPA32/14

1) Ladder, 2) wtRPA, 3) post-induction cells, 4) wtRPA, 5) ladder, 6) cell lysate, 7) cell pellet, 8) Lysate after dialysis, 9) Affigel Blue flow-through, 10) 800 mM KCl fraction, 11) 500 mM NaSCN fraction, 12) 1500 mM NaSCN fraction, 13) SEC input sample, 14) ladder, 15) wtRPA, 16-20) SEC fractions

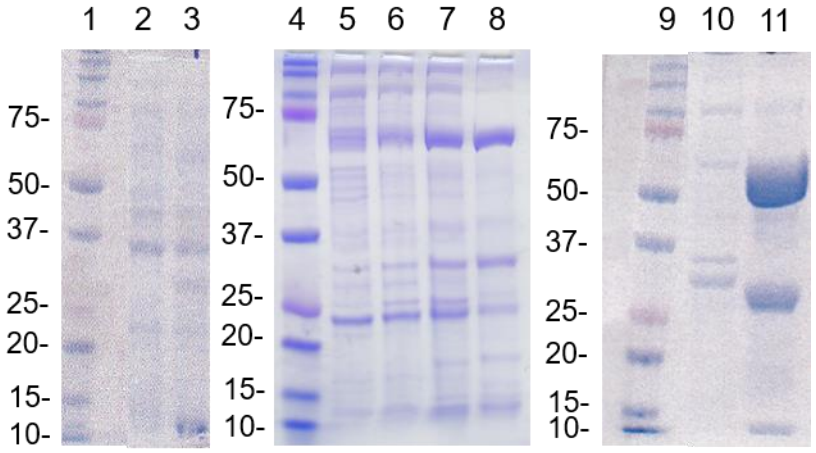


Figure A.10 SDS-PAGE of 1EbRPA70+wtRPA32/14

1) Ladder, 2) pre-induction cells, 3) post-induction cells, 4) ladder, 5) Affigel Blue flow-through, 6) 500 mM NaSCN fraction, 7) 1500 mM NaSCN fraction, 8) SEC input sample, 9) ladder, 10-11) SEC fractions

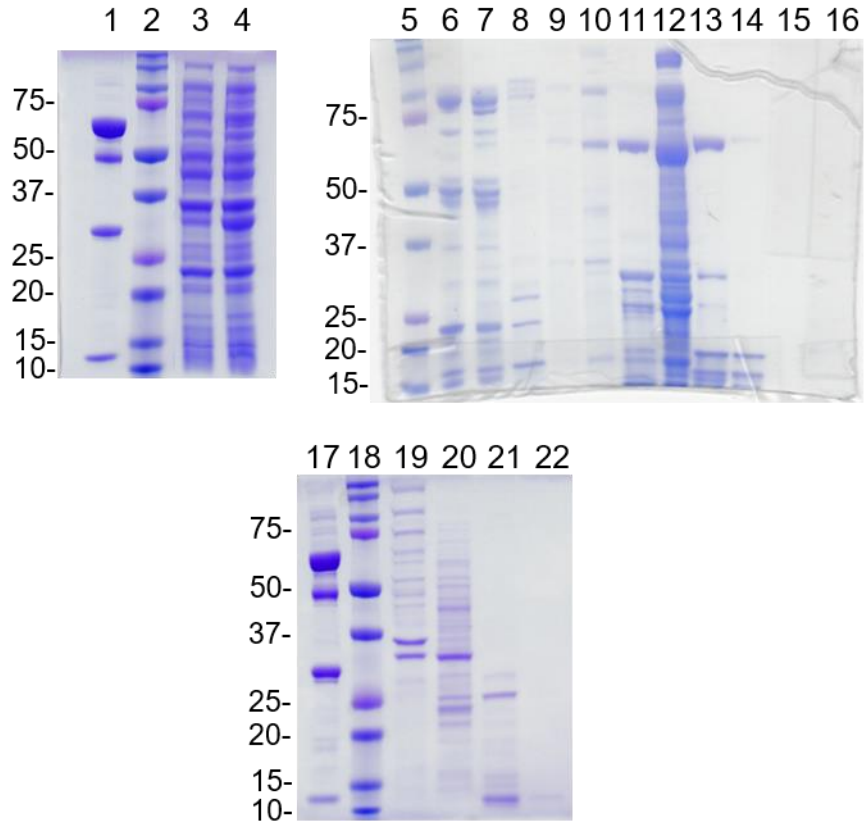


Figure A.11 SDS-PAGE of wtRPA70+8ERPA32/14

1) wtRPA, 2) ladder, 3) pre-induction cells, 4) post-induction cells, 5) ladder, 6-7) 800 mM KCl fractions, 8) 500 mM NaSCN fraction, 9-16) 1500 mM NaSCN fractions, 17) wtRPA, 18) ladder, 19-22) SEC fractions

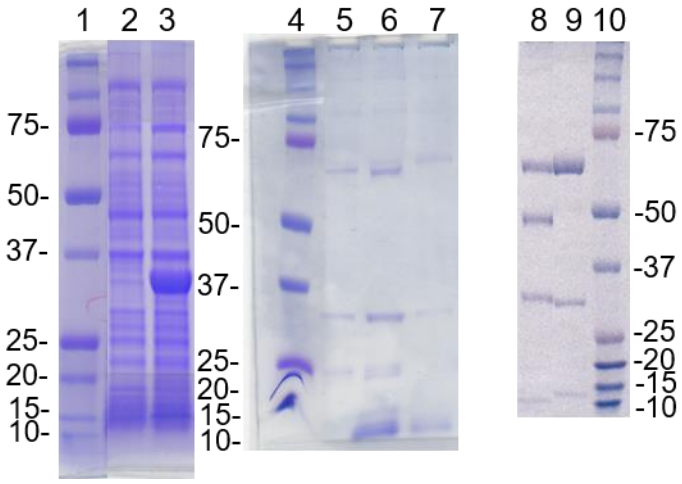


Figure A.12 SDS-PAGE of wtRPA70+9ERPA32/14

1) Ladder, 2) pre-induction cells, 3) post-induction cells, 4) ladder, 5-7) 1500 mM NaSCN fractions, 8) SEC fraction, 9) wtRPA, 10) ladder

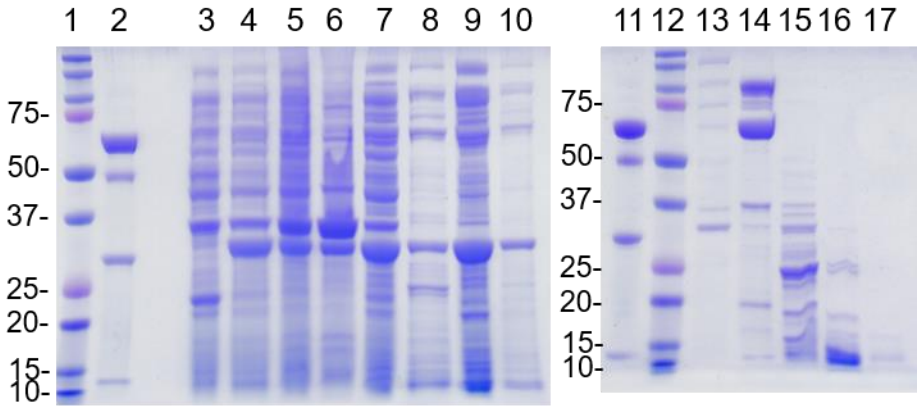


Figure A.13 SDS-PAGE of 1EaRPA70+8ERPA32/14

1) Ladder, 2) wtRPA, 3) pre-induction cells, 4) post-induction cells, 5) cell pellet, 6) cell lysate, 7) Affigel Blue flow-through, 8) 500 mM NaSCN fraction, 9) 1500 mM NaSCN fraction, 10) SEC input sample, 11) wtRPA, 12) ladder, 13-17) SEC fractions

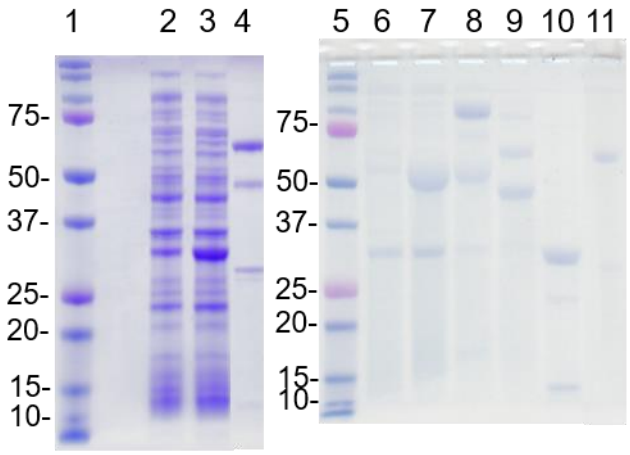


Figure A.14 SDS-PAGE of 1EaRPA70+9ERPA32/14

1) Ladder, 2) pre-induction cells, 3) post-induction cells, 4) wtRPA, 5) ladder, 6-10) SEC fractions, 11) wtRPA

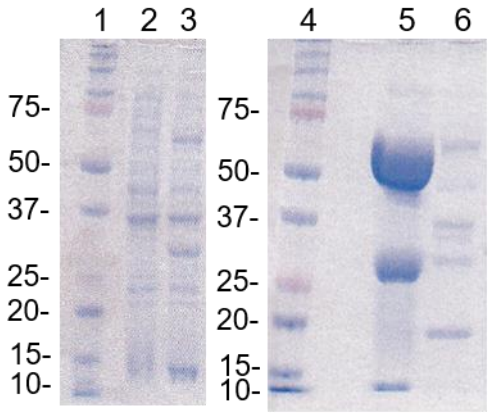


Figure A.15 SDS-PAGE of 1EaRPA70+wtRPA32/14

1) Ladder, 2) pre-induction cells, 3) post-induction cells, 4) ladder, 5-6) SEC fractions

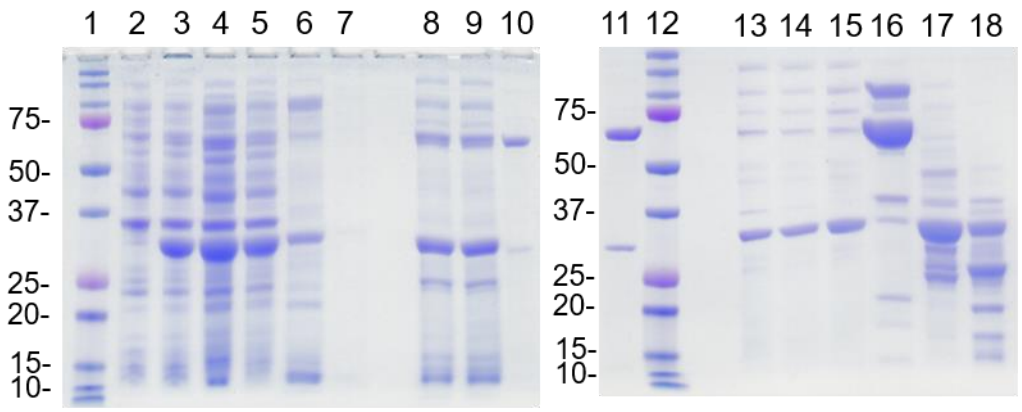


Figure A.16 SDS-PAGE of 2ERPA70+8ERPA32/14

1) Ladder, 2) pre-induction cells, 3) post-induction cells, 4) post dialysis cell lysate, 5) Affigel Blue flow-through 6) 800 mM KCl fraction, 7) 500 mM NaSCN fraction, 8) 1500 mM NaSCN fraction, 9) SEC input sample, 10-11) wtRPA, 12) ladder, 13-18) SEC fractions

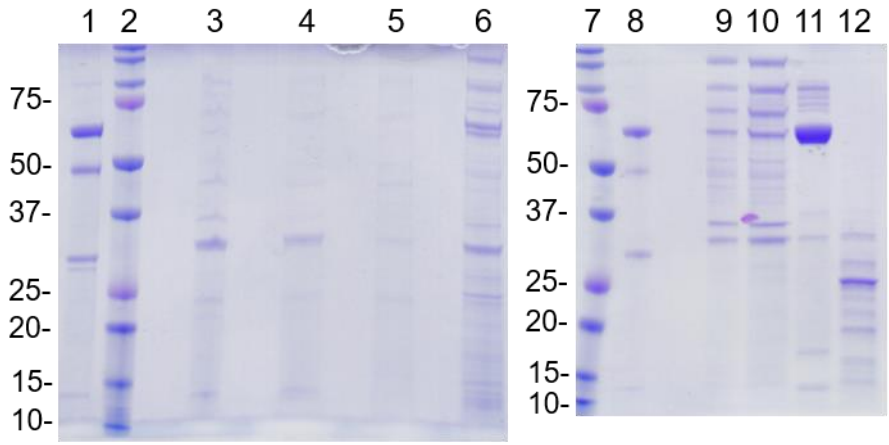


Figure A.17 SDS-PAGE of 2ERPA70+9ERPA32/14

1) wtRPA, 2) ladder, 3) Affigel Blue flow-through, 4) 800 mM KCl fraction, 5) 500 mM NaSCN fraction, 6) 1500 mM NaSCN fraction, 7) ladder, 8) wtRPA, 9-12) SEC fractions

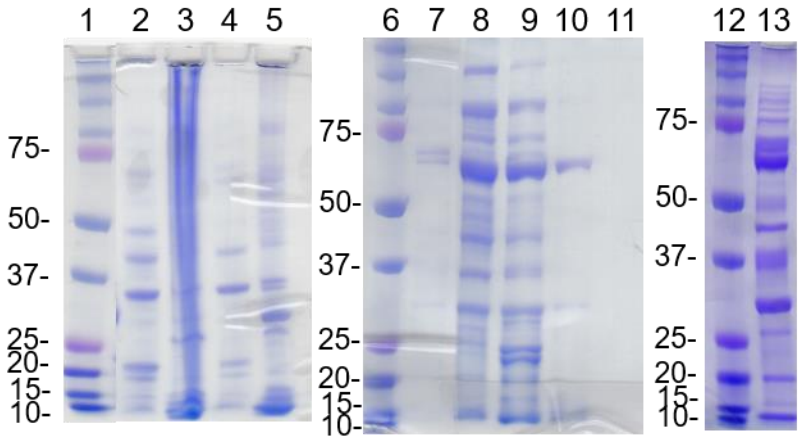


Figure A.18 SDS-PAGE of 2ERPA70+wtRPA32/14

1) Ladder, 2) supernatant from uninduced cells, 3) cell pellet from uninduced cells, 4) supernatant from induced cells, 5) pellet from induced cells, 6) ladder, 7) 500 mM NaSCN fraction, 8-11) 1500 mM NaSCN fractions, 12) ladder, 13) SEC fraction

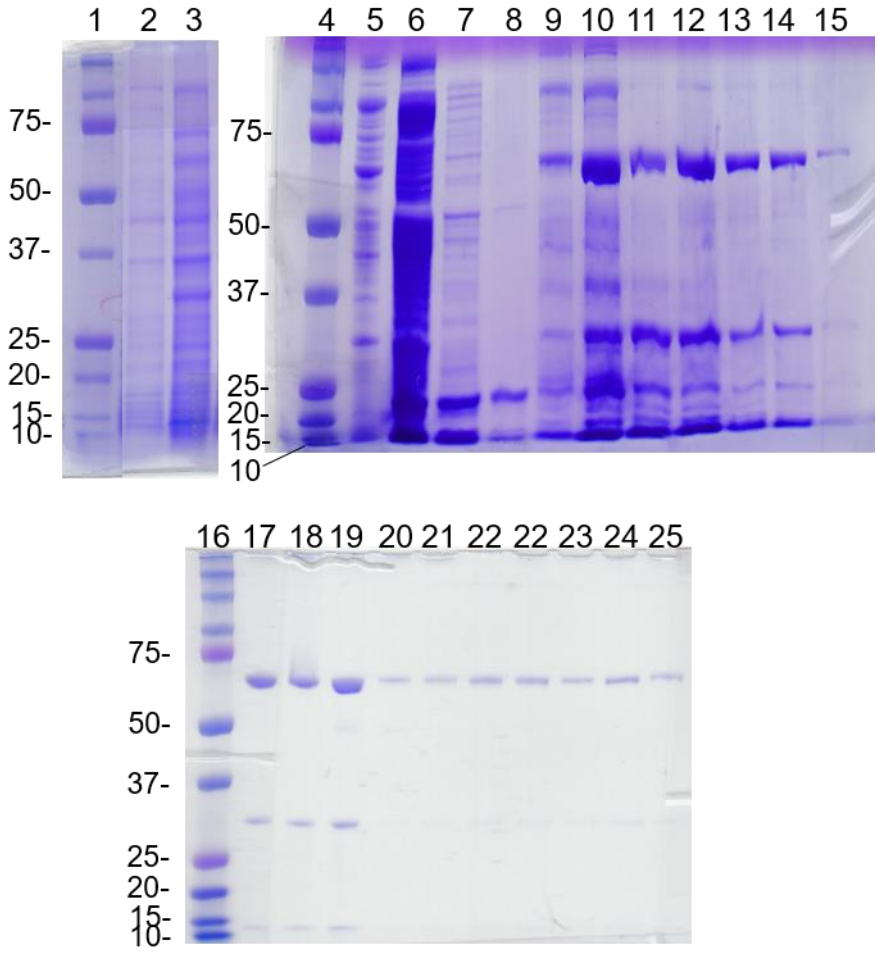


Figure A.19 SDS-PAGE of wtRPA

1) Ladder, 2) pre-induction cells, 3) post-induction cells, 4) ladder, 5-8) 800 mM KCl fractions, 9-15) 1500 mM NaSCN fractions, 16) ladder, 17-25) SEC fractions

References

1. De Bont, R. and N. van Larebeke, *Endogenous DNA damage in humans: a review of quantitative data.* Mutagenesis, 2004. **19**(3): p. 169-85.
2. Jiricny, J., *The multifaceted mismatch-repair system.* Nat Rev Mol Cell Biol, 2006. **7**(5): p. 335-46.
3. Robertson, A.B., et al., *DNA repair in mammalian cells: Base excision repair: the long and short of it.* Cell Mol Life Sci, 2009. **66**(6): p. 981-93.
4. Marteijn, J.A., et al., *Understanding nucleotide excision repair and its roles in cancer and ageing.* Nat Rev Mol Cell Biol, 2014. **15**(7): p. 465-81.
5. Vilenchik, M.M. and A.G. Knudson, *Endogenous DNA double-strand breaks: production, fidelity of repair, and induction of cancer.* Proc Natl Acad Sci U S A, 2003. **100**(22): p. 12871-6.
6. Rothkamm, K., et al., *Pathways of DNA double-strand break repair during the mammalian cell cycle.* Mol Cell Biol, 2003. **23**(16): p. 5706-15.
7. Kolodner, R.D., C.D. Putnam, and K. Myung, *Maintenance of genome stability in Saccharomyces cerevisiae.* Science, 2002. **297**(5581): p. 552-7.
8. Symington, L.S. and J. Gautier, *Double-strand break end resection and repair pathway choice.* Annu Rev Genet, 2011. **45**: p. 247-71.
9. Guirouilh-Barbat, J., et al., *Impact of the KU80 pathway on NHEJ-induced genome rearrangements in mammalian cells.* Mol Cell, 2004. **14**(5): p. 611-23.
10. Chang, H.H.Y., et al., *Non-homologous DNA end joining and alternative pathways to double-strand break repair.* Nat Rev Mol Cell Biol, 2017. **18**(8): p. 495-506.
11. Zha, S., C. Boboila, and F.W. Alt, *Mre11: roles in DNA repair beyond homologous recombination.* Nat Struct Mol Biol, 2009. **16**(8): p. 798-800.
12. DeFazio, L.G., et al., *Synapsis of DNA ends by DNA-dependent protein kinase.* EMBO J, 2002. **21**(12): p. 3192-200.
13. Palmbos, P.L., et al., *Recruitment of Saccharomyces cerevisiae Dnl4-Lif1 complex to a double-strand break requires interactions with Yku80 and the Xrs2 FHA domain.* Genetics, 2008. **180**(4): p. 1809-19.
14. Yano, K., et al., *Ku recruits XLF to DNA double-strand breaks.* EMBO Rep, 2008. **9**(1): p. 91-6.
15. Boulton, S.J. and S.P. Jackson, *Saccharomyces cerevisiae Ku70 potentiates illegitimate DNA double-strand break repair and serves as a barrier to error-prone DNA repair pathways.* EMBO J, 1996. **15**(18): p. 5093-103.
16. Daley, J.M., et al., *DNA joint dependence of pol X family polymerase action in nonhomologous end joining.* J Biol Chem, 2005. **280**(32): p. 29030-7.
17. Wilson, T.E., U. Grawunder, and M.R. Lieber, *Yeast DNA ligase IV mediates non-homologous DNA end joining.* Nature, 1997. **388**(6641): p. 495-8.
18. Ahnesorg, P., P. Smith, and S.P. Jackson, *XLF interacts with the XRCC4-DNA ligase IV complex to promote DNA nonhomologous end-joining.* Cell, 2006. **124**(2): p. 301-13.
19. Mao, Z., et al., *DNA repair by nonhomologous end joining and homologous recombination during cell cycle in human cells.* Cell Cycle, 2008. **7**(18): p. 2902-6.
20. Nepomuceno, T.C., et al., *The Role of PALB2 in the DNA Damage Response and Cancer Predisposition.* Int J Mol Sci, 2017. **18**(9).

21. Hanamshet, K., O.M. Mazina, and A.V. Mazin, *Reappearance from Obscurity: Mammalian Rad52 in Homologous Recombination.* Genes (Basel), 2016. **7**(9).
22. Collins, I. and C.S. Newlon, *Meiosis-specific formation of joint DNA molecules containing sequences from homologous chromosomes.* Cell, 1994. **76**(1): p. 65-75.
23. Sung, P. and H. Klein, *Mechanism of homologous recombination: mediators and helicases take on regulatory functions.* Nat Rev Mol Cell Biol, 2006. **7**(10): p. 739-50.
24. Mimitou, E.P. and L.S. Symington, *Nucleases and helicases take center stage in homologous recombination.* Trends Biochem Sci, 2009. **34**(5): p. 264-72.
25. Makharashvili, N. and T.T. Paull, *CtIP: A DNA damage response protein at the intersection of DNA metabolism.* DNA Repair (Amst), 2015. **32**: p. 75-81.
26. Schwacha, A. and N. Kleckner, *Identification of joint molecules that form frequently between homologs but rarely between sister chromatids during yeast meiosis.* Cell, 1994. **76**(1): p. 51-63.
27. Van Dyck, E., et al., *Visualization of recombination intermediates produced by RAD52-mediated single-strand annealing.* EMBO Rep, 2001. **2**(10): p. 905-9.
28. Ivanov, E.L., et al., *Genetic requirements for the single-strand annealing pathway of double-strand break repair in Saccharomyces cerevisiae.* Genetics, 1996. **142**(3): p. 693-704.
29. McEachern, M.J. and J.E. Haber, *Break-induced replication and recombinational telomere elongation in yeast.* Annu Rev Biochem, 2006. **75**: p. 111-35.
30. Dilley, R.L. and R.A. Greenberg, *ALTernative Telomere Maintenance and Cancer.* Trends Cancer, 2015. **1**(2): p. 145-156.
31. Costantino, L., et al., *Break-induced replication repair of damaged forks induces genomic duplications in human cells.* Science, 2014. **343**(6166): p. 88-91.
32. Minocherhomji, S., et al., *Replication stress activates DNA repair synthesis in mitosis.* Nature, 2015. **528**(7581): p. 286-90.
33. Sakofsky, C.J. and A. Malkova, *Break induced replication in eukaryotes: mechanisms, functions, and consequences.* Crit Rev Biochem Mol Biol, 2017. **52**(4): p. 395-413.
34. Guo, S., et al., *Regulation of replication protein A functions in DNA mismatch repair by phosphorylation.* J Biol Chem, 2006. **281**(31): p. 21607-16.
35. DeMott, M.S., S. Zigman, and R.A. Bambara, *Replication protein A stimulates long patch DNA base excision repair.* J Biol Chem, 1998. **273**(42): p. 27492-8.
36. Nagelhus, T.A., et al., *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, 1997. **272**(10): p. 6561-6.
37. He, Z., et al., *RPA involvement in the damage-recognition and incision steps of nucleotide excision repair.* Nature, 1995. **374**(6522): p. 566-9.
38. Matsuda, T., et al., *DNA repair protein XPA binds replication protein A (RPA).* J Biol Chem, 1995. **270**(8): p. 4152-7.
39. Li, L., et al., *An interaction between the DNA repair factor XPA and replication protein A appears essential for nucleotide excision repair.* Mol Cell Biol, 1995. **15**(10): p. 5396-402.
40. Matsunaga, T., et al., *Replication protein A confers structure-specific endonuclease activities to the XPF-ERCC1 and XPG subunits of human DNA repair excision nuclease.* J Biol Chem, 1996. **271**(19): p. 11047-50.
41. Perrault, R., et al., *RPA facilitates rejoining of DNA double-strand breaks in an in vitro assay utilizing genomic DNA as substrate.* Int J Radiat Biol, 2001. **77**(5): p. 593-607.

42. Grandi, P., et al., *DNA double-strand breaks induce formation of RP-A/Ku foci on in vitro reconstituted Xenopus sperm nuclei.* J Cell Sci, 2001. **114**(Pt 18): p. 3345-57.
43. Wold, M.S., *Replication protein A: a heterotrimeric, single-stranded DNA-binding protein required for eukaryotic DNA metabolism.* Annu Rev Biochem, 1997. **66**: p. 61-92.
44. Rajagopalan, S., et al., *Mapping the physical and functional interactions between the tumor suppressors p53 and BRCA2.* Proc Natl Acad Sci U S A, 2010. **107**(19): p. 8587-92.
45. Wong, J.M., D. Ionescu, and C.J. Ingles, *Interaction between BRCA2 and replication protein A is compromised by a cancer-predisposing mutation in BRCA2.* Oncogene, 2003. **22**(1): p. 28-33.
46. Park, M.S., et al., *Physical interaction between human RAD52 and RPA is required for homologous recombination in mammalian cells.* J Biol Chem, 1996. **271**(31): p. 18996-9000.
47. Jackson, D., et al., *Analysis of the human replication protein A:Rad52 complex: evidence for crosstalk between RPA32, RPA70, Rad52 and DNA.* J Mol Biol, 2002. **321**(1): p. 133-48.
48. Nakaya, R., et al., *Identification of proteins that may directly interact with human RPA.* J Biochem, 2010. **148**(5): p. 539-47.
49. Henricksen, L.A. and M.S. Wold, *Replication protein A mutants lacking phosphorylation sites for p34cdc2 kinase support DNA replication.* J Biol Chem, 1994. **269**(39): p. 24203-8.
50. Gasior, S.L., et al., *Rad52 associates with RPA and functions with rad55 and rad57 to assemble meiotic recombination complexes.* Genes Dev, 1998. **12**(14): p. 2208-21.
51. Tan, T.L., et al., *Mouse Rad54 affects DNA conformation and DNA-damage-induced Rad51 foci formation.* Curr Biol, 1999. **9**(6): p. 325-8.
52. Golub, E.I., et al., *Interaction of human rad51 recombination protein with single-stranded DNA binding protein, RPA.* Nucleic Acids Res, 1998. **26**(23): p. 5388-93.
53. Zernik-Kobak, M., et al., *Sites of UV-induced phosphorylation of the p34 subunit of replication protein A from HeLa cells.* J Biol Chem, 1997. **272**(38): p. 23896-904.
54. Mer, G., et al., *Structural basis for the recognition of DNA repair proteins UNG2, XPA, and RAD52 by replication factor RPA.* Cell, 2000. **103**(3): p. 449-56.
55. Iftode, C., Y. Daniely, and J.A. Borowiec, *Replication protein A (RPA): the eukaryotic SSB.* Crit Rev Biochem Mol Biol, 1999. **34**(3): p. 141-80.
56. Din, S., et al., *Cell-cycle-regulated phosphorylation of DNA replication factor A from human and yeast cells.* Genes Dev, 1990. **4**(6): p. 968-77.
57. Dutta, A. and B. Stillman, *cdc2 family kinases phosphorylate a human cell DNA replication factor, RPA, and activate DNA replication.* EMBO J, 1992. **11**(6): p. 2189-99.
58. Fang, F. and J.W. Newport, *Distinct roles of cdk2 and cdc2 in RP-A phosphorylation during the cell cycle.* J Cell Sci, 1993. **106 (Pt 3)**: p. 983-94.
59. Niu, H., et al., *Mapping of amino acid residues in the p34 subunit of human single-stranded DNA-binding protein phosphorylated by DNA-dependent protein kinase and Cdc2 kinase in vitro.* J Biol Chem, 1997. **272**(19): p. 12634-41.
60. Pan, Z.Q., et al., *Phosphorylation of the p34 subunit of human single-stranded-DNA-binding protein in cyclin A-activated G1 extracts is catalyzed by cdk-cyclin A complex and DNA-dependent protein kinase.* Proc Natl Acad Sci U S A, 1994. **91**(18): p. 8343-7.
61. Oakley, G.G. and S.M. Patrick, *Replication protein A: directing traffic at the intersection of replication and repair.* Front Biosci (Landmark Ed), 2010. **15**: p. 883-900.
62. Binz, S.K., A.M. Sheehan, and M.S. Wold, *Replication protein A phosphorylation and the cellular response to DNA damage.* DNA Repair (Amst), 2004. **3**(8-9): p. 1015-24.

63. Liu, V.F. and D.T. Weaver, *The ionizing radiation-induced replication protein A phosphorylation response differs between ataxia telangiectasia and normal human cells.* Mol Cell Biol, 1993. **13**(12): p. 7222-31.
64. Zuazua-Villar, P., et al., *Extensive RPA2 hyperphosphorylation promotes apoptosis in response to DNA replication stress in CHK1 inhibited cells.* Nucleic Acids Res, 2015. **43**(20): p. 9776-87.
65. Treuner, K., et al., *Phosphorylation of replication protein A middle subunit (RPA32) leads to a disassembly of the RPA heterotrimer.* J Biol Chem, 1999. **274**(22): p. 15556-61.
66. Derheimer, F.A. and M.B. Kastan, *Multiple roles of ATM in monitoring and maintaining DNA integrity.* FEBS Lett, 2010. **584**(17): p. 3675-81.
67. Nam, E.A. and D. Cortez, *ATR signalling: more than meeting at the fork.* Biochem J, 2011. **436**(3): p. 527-36.
68. Martin, M., et al., *ATM and DNA-PKcs make a complementary couple in DNA double strand break repair.* Mutat Res, 2012.
69. Borgstahl, G.E., et al., *Interplay of DNA damage and cell cycle signaling at the level of human replication protein A.* DNA Repair (Amst), 2014. **21**: p. 12-23.
70. Oakley, G.G., et al., *RPA phosphorylation in mitosis alters DNA binding and protein-protein interactions.* Biochemistry, 2003. **42**(11): p. 3255-64.
71. Binz, S.K., et al., *The phosphorylation domain of the 32-kDa subunit of replication protein A (RPA) modulates RPA-DNA interactions. Evidence for an intersubunit interaction.* J Biol Chem, 2003. **278**(37): p. 35584-91.
72. Carty, M.P., et al., *UV light-induced DNA synthesis arrest in HeLa cells is associated with changes in phosphorylation of human single-stranded DNA-binding protein.* EMBO J, 1994. **13**(9): p. 2114-23.
73. Liu, J.S., et al., *Adozelesin triggers DNA damage response pathways and arrests SV40 DNA replication through replication protein A inactivation.* J Biol Chem, 2000. **275**(2): p. 1391-7.
74. Wang, Y., et al., *Roles of replication protein A and DNA-dependent protein kinase in the regulation of DNA replication following DNA damage.* J Biol Chem, 1999. **274**(31): p. 22060-4.
75. Game, J.C. and R.K. Mortimer, *A genetic study of x-ray sensitive mutants in yeast.* Mutat Res, 1974. **24**(3): p. 281-92.
76. Shinohara, A. and T. Ogawa, *Stimulation by Rad52 of yeast Rad51-mediated recombination.* Nature, 1998. **391**(6665): p. 404-7.
77. New, J.H., et al., *Rad52 protein stimulates DNA strand exchange by Rad51 and replication protein A.* Nature, 1998. **391**(6665): p. 407-10.
78. San Filippo, J., P. Sung, and H. Klein, *Mechanism of eukaryotic homologous recombination.* Annu Rev Biochem, 2008. **77**: p. 229-57.
79. Malone, R.E. and R.E. Esposito, *The RAD52 gene is required for homothallic interconversion of mating types and spontaneous mitotic recombination in yeast.* Proc Natl Acad Sci U S A, 1980. **77**(1): p. 503-7.
80. Schiestl, R.H., M. Dominska, and T.D. Petes, *Transformation of Saccharomyces cerevisiae with nonhomologous DNA: illegitimate integration of transforming DNA into yeast chromosomes and in vivo ligation of transforming DNA to mitochondrial DNA sequences.* Mol Cell Biol, 1993. **13**(5): p. 2697-705.

81. Rijkers, T., et al., *Targeted inactivation of mouse RAD52 reduces homologous recombination but not resistance to ionizing radiation.* Mol Cell Biol, 1998. **18**(11): p. 6423-9.
82. Yamaguchi-Iwai, Y., et al., *Homologous recombination, but not DNA repair, is reduced in vertebrate cells deficient in RAD52.* Mol Cell Biol, 1998. **18**(11): p. 6430-5.
83. Lok, B.H., et al., *RAD52 inactivation is synthetically lethal with deficiencies in BRCA1 and PALB2 in addition to BRCA2 through RAD51-mediated homologous recombination.* Oncogene, 2013. **32**(30): p. 3552-8.
84. Feng, Z., et al., *Rad52 inactivation is synthetically lethal with BRCA2 deficiency.* Proc Natl Acad Sci U S A, 2011. **108**(2): p. 686-91.
85. Benson, F.E., P. Baumann, and S.C. West, *Synergistic actions of Rad51 and Rad52 in recombination and DNA repair.* Nature, 1998. **391**(6665): p. 401-4.
86. Sugiyama, T. and S.C. Kowalczykowski, *Rad52 protein associates with replication protein A (RPA)-single-stranded DNA to accelerate Rad51-mediated displacement of RPA and presynaptic complex formation.* J Biol Chem, 2002. **277**(35): p. 31663-72.
87. New, J.H. and S.C. Kowalczykowski, *Rad52 protein has a second stimulatory role in DNA strand exchange that complements replication protein-A function.* J Biol Chem, 2002. **277**(29): p. 26171-6.
88. Kagawa, W., et al., *Identification of a second DNA binding site in the human Rad52 protein.* J Biol Chem, 2008. **283**(35): p. 24264-73.
89. Bi, B., et al., *Human and yeast Rad52 proteins promote DNA strand exchange.* Proc Natl Acad Sci U S A, 2004. **101**(26): p. 9568-72.
90. Kagawa, W., et al., *Homologous pairing promoted by the human Rad52 protein.* J Biol Chem, 2001. **276**(37): p. 35201-8.
91. Sotiriou, S.K., et al., *Mammalian RAD52 Functions in Break-Induced Replication Repair of Collapsed DNA Replication Forks.* Mol Cell, 2016. **64**(6): p. 1127-1134.
92. Mazina, O.M., et al., *Rad52 Inverse Strand Exchange Drives RNA-Templated DNA Double-Strand Break Repair.* Mol Cell, 2017. **67**(1): p. 19-29 e3.
93. Lloyd, J.A., D.A. McGrew, and K.L. Knight, *Identification of residues important for DNA binding in the full-length human Rad52 protein.* J Mol Biol, 2005. **345**(2): p. 239-49.
94. Lloyd, J.A., A.L. Forget, and K.L. Knight, *Correlation of biochemical properties with the oligomeric state of human rad52 protein.* J Biol Chem, 2002. **277**(48): p. 46172-8.
95. Shen, Z., et al., *Specific interactions between the human RAD51 and RAD52 proteins.* J Biol Chem, 1996. **271**(1): p. 148-52.
96. Stasiak, A.Z., et al., *The human Rad52 protein exists as a heptameric ring.* Curr Biol, 2000. **10**(6): p. 337-40.
97. Singleton, M.R., et al., *Structure of the single-strand annealing domain of human RAD52 protein.* Proc Natl Acad Sci U S A, 2002. **99**(21): p. 13492-7.
98. Kagawa, W., et al., *Crystal structure of the homologous-pairing domain from the human Rad52 recombinase in the undecameric form.* Mol Cell, 2002. **10**(2): p. 359-71.
99. Ranatunga, W., et al., *Human RAD52 exhibits two modes of self-association.* J Biol Chem, 2001. **276**(19): p. 15876-80.
100. Bhargava, R., D.O. Onyango, and J.M. Stark, *Regulation of Single-Strand Annealing and its Role in Genome Maintenance.* Trends Genet, 2016. **32**(9): p. 566-575.
101. Helleday, T., et al., *DNA double-strand break repair: from mechanistic understanding to cancer treatment.* DNA Repair (Amst), 2007. **6**(7): p. 923-35.

102. Narod, S.A. and A.A. Rodriguez, *[Genetic predisposition for breast cancer: BRCA1 and BRCA2 genes].* Salud Publica Mex, 2011. **53**(5): p. 420-9.
103. Narod, S.A. and L. Salmena, *BRCA1 and BRCA2 mutations and breast cancer.* Discov Med, 2011. **12**(66): p. 445-53.
104. Koumpis, C., et al., *Prevalence of BRCA1 and BRCA2 mutations in unselected breast cancer patients from Greece.* Hered Cancer Clin Pract, 2011. **9**: p. 10.
105. Rodriguez, A.O., et al., *BRCA1 and BRCA2 mutations among ovarian cancer patients from Colombia.* Gynecol Oncol, 2012. **124**(2): p. 236-43.
106. Greer, J.B. and D.C. Whitcomb, *Role of BRCA1 and BRCA2 mutations in pancreatic cancer.* Gut, 2007. **56**(5): p. 601-5.
107. Bochkareva, E., et al., *Structure of the RPA trimerization core and its role in the multistep DNA-binding mechanism of RPA.* EMBO J, 2002. **21**(7): p. 1855-63.
108. Bochkareva, E., et al., *Structure of the major single-stranded DNA-binding domain of replication protein A suggests a dynamic mechanism for DNA binding.* EMBO J, 2001. **20**(3): p. 612-8.
109. Bochkarev, A., et al., *Structure of the single-stranded-DNA-binding domain of replication protein A bound to DNA.* Nature, 1997. **385**(6612): p. 176-81.
110. Jacobs, D.M., et al., *Human replication protein A: global fold of the N-terminal RPA-70 domain reveals a basic cleft and flexible C-terminal linker.* J Biomol NMR, 1999. **14**(4): p. 321-31.
111. Petoukhov, M.V., et al., *New developments in the ATSAS program package for small-angle scattering data analysis.* J Appl Crystallogr, 2012. **45**(Pt 2): p. 342-350.
112. Chacon, P. and W. Wriggers, *Multi-resolution contour-based fitting of macromolecular structures.* J Mol Biol, 2002. **317**(3): p. 375-84.
113. Wriggers, W., *Using Situs for the integration of multi-resolution structures.* Biophys Rev, 2010. **2**(1): p. 21-27.
114. Brunger, A.T., et al., *Crystallography & NMR system: A new software suite for macromolecular structure determination.* Acta Crystallogr D Biol Crystallogr, 1998. **54**(Pt 5): p. 905-21.
115. Brunger, A.T., *Version 1.2 of the Crystallography and NMR system.* Nat Protoc, 2007. **2**(11): p. 2728-33.
116. Yang, J., et al., *The I-TASSER Suite: protein structure and function prediction.* Nat Methods, 2015. **12**(1): p. 7-8.
117. Roy, A., A. Kucukural, and Y. Zhang, *I-TASSER: a unified platform for automated protein structure and function prediction.* Nat Protoc, 2010. **5**(4): p. 725-38.
118. Zhang, Y., *I-TASSER server for protein 3D structure prediction.* BMC Bioinformatics, 2008. **9**: p. 40.
119. Fanning, E., V. Klimovich, and A.R. Nager, *A dynamic model for replication protein A (RPA) function in DNA processing pathways.* Nucleic Acids Res, 2006. **34**(15): p. 4126-37.
120. Li, G.M., *Mechanisms and functions of DNA mismatch repair.* Cell Res, 2008. **18**(1): p. 85-98.
121. Krejci, L., et al., *Homologous recombination and its regulation.* Nucleic Acids Res, 2012. **40**(13): p. 5795-818.
122. Iyama, T. and D.M. Wilson, 3rd, *DNA repair mechanisms in dividing and non-dividing cells.* DNA Repair (Amst), 2013. **12**(8): p. 620-36.
123. Lee, S.H. and D.K. Kim, *The role of the 34-kDa subunit of human replication protein A in simian virus 40 DNA replication in vitro.* J Biol Chem, 1995. **270**(21): p. 12801-7.

124. Gomes, X.V., L.A. Henricksen, and M.S. Wold, *Proteolytic mapping of human replication protein A: evidence for multiple structural domains and a conformational change upon interaction with single-stranded DNA.* Biochemistry, 1996. **35**(17): p. 5586-95.
125. Vassin, V.M., M.S. Wold, and J.A. Borowiec, *Replication protein A (RPA) phosphorylation prevents RPA association with replication centers.* Mol Cell Biol, 2004. **24**(5): p. 1930-43.
126. Henricksen, L.A., C.B. Umbricht, and M.S. Wold, *Recombinant replication protein A: expression, complex formation, and functional characterization.* J Biol Chem, 1994. **269**(15): p. 11121-32.
127. Deng, X., et al., *Structure of the full-length human RPA14/32 complex gives insights into the mechanism of DNA binding and complex formation.* J Mol Biol, 2007. **374**(4): p. 865-76.
128. Bras, W., et al., *The susceptibility of pure tubulin to high magnetic fields: a magnetic birefringence and x-ray fiber diffraction study.* Biophys J, 1998. **74**(3): p. 1509-21.
129. Chandramouly, G., et al., *Small-Molecule Disruption of RAD52 Rings as a Mechanism for Precision Medicine in BRCA-Deficient Cancers.* Chem Biol, 2015. **22**(11): p. 1491-1504.
130. Deng, X., et al., *Human replication protein A-Rad52-single-stranded DNA complex: stoichiometry and evidence for strand transfer regulation by phosphorylation.* Biochemistry, 2009. **48**(28): p. 6633-43.