DOMAIN-BASED STRUCTURAL STUDIES OF REPLICATION PROTEIN A: ANALYSIS OF AN RPA32N PHOSPHO-MIMIC MUTANT AND THE ROLE OF RPA70N IN BINDING SSDNA

By

Dalyir Imelda Pretto Garcia

Dissertation

Submitted to the Faculty of the

Graduate School of Vanderbilt University

in partial fulfillment of the requirements

for the degree of

DOCTOR OF PHILOSOPHY

in

Biochemistry

August, 2010

Nashville, Tennessee

Approved:

Professor Walter J. Chazin

Professor Richard Armstrong

Professor Ellen Fanning

Professor David Cortez

Professor Brandt Eichman

TABLE OF CONTENTS

Page

AKNOWLEDGEMENTS..... iv

LIST OF TABLES vi

LIST OF FIGURES ……..…… vii

LIST OF ABBREVIATIONS..... ……..…ix

Chapter

I. INTRODUCTION.....….…….…..…..1

DNA Replication .....…….....……..1
Replication Protein A .....…….....…….. 2
RPA ssDNA binding activity.....…….....4
RPA Protein Interactions .....…….....……6
Disposition of RPA70N in the RPA ssDNA binding activity 10
Phosphorylation of RPA . .....…….....…….. 13
Hyperphosphorylation of RPA by Cyclin Dependent Kinases.....17
Hyperphosphorylation of RPA as a response to DNA damage 20
DNA-PK Mediated Phosphorylation of RPA.....23
ATM Mediated phosphorylation of RPA 27
ATR Mediated phosphorylation of RPA 31
Experimental Methods .....……..... 35
Small Angle X-Ray Scattering (SAXS)..... 35
Nuclear Magnetic Resonance (NMR)... 40
Research overview.....43

II. STRUCTURAL DYNAMICS AND SSDNA BINDING ACTIVITY OF THE THREE N-TERMINAL DOMAINS OF THE LARGE SUBUNIT OF REPLICATION PROTEIN A FROM SMALL ANGLE X-RAY SCATTERING 45

Introduction .....…….. 45
Results.....……48
Structural dynamics of RPA70AB from analysis of SAXS..…… 50
ssDNA binding to RPA70AB…… .....……54
Small angle X-ray scattering of RPA70NAB. 56
Effect of ssDNA binding on the structural dynamics of RPA70NAB.....58
Discussion… .....…… 63
Experimental Procedures.....……67
Expression of RPAs 67

Protein purification.....68
Preparation of protein-DNA complexes 68
Size excussion chromatography – multi-angle light scattering 68
Small angle X-ray scattering 69
Computational Modeling … 70

III. TOWARD STRUCTURAL CHARACTERIZATION OF THE HYPERPHOSPHORYLATED FORM OF RPA 72
Introduction 72
Results…..... 76
Production of recombinant RPA32N wt and phosphomimetic peptides. ..76
Wild-type and Phosphomimetic RPA32N peptides are unstructured 78
Isolated RPA70 N, RPA70A and RPA70B domains interact with phosphomimetic RPA32N 80
RPA32N-D8 interaction is specific to the RPA70 subunit.....86
Production and characterization of RPAD9 phospho-mimic mutant... 87
Discussion.....91
Experimental Procedures 93
Cloning.....93
Protein expression.....95
Proteins purification.....96
DNA sequencing.....98
MALDI Mass.....98
NMR chemical shift perturbation assays.....98
NMR resonance assignments.....99
Small angle x-ray scattering.....99

IV. DISCUSSION AND FUTURE DIRECTIONS 101

Influence of RPA70N on RPA ssDNA binding activity.....101
The role of flexibly linked domains in RPA function.....101
RPA structural dynamics 104
Role of protein dynamics in DNA processing 107
RPA32N interactions within RPA 108
The role of hyperphosphorylation in the function of RPA.....109
Hyperphosphorylation and ssDNA binding activity.....110
The effect of hyperphosphorylation on RPA: The compaction model. 111
The compaction model and RPA hyperphosphorylation in replication, recombinatorial repair and checkpoint signaling.....114
Significance.....117

V. REFERENCES 119

AKNOWLEDGEMENTS

I would like to express my gratitude to Dr. Walter J. Chazin for his astute guidance, timely advice, extreme patience, and constant encouragement during my graduate studies at Vanderbilt University. For always caring and believing that I had the potential for reaching the set goals and improving as I went on; for all the time that he dedicated to my mentorship showing support in my ideas and confidence in me; and especially, for always keeping the bar high and not letting me fall below, I would like to thank Dr Chazin.

My thanks and appreciation to the Interdisciplinary Graduate Program and the Department of Biochemistry at Vanderbilt University for the excellent management and organization of the graduate student program from which I have directly benefited. I would particularly like to thank the members of my committee, Dr. Richard Armstrong, Dr. Ellen Fanning, Dr. David Cortez, and Dr. Brandt Eichman, who generously committed time to advice me and evaluate my progress by contributing ideas and critical concerns that brought me to a successful end point. I would like to also thank the National Institute of Health for providing the funding that made this research possible.

I would like to thank Dr. Janos Sumegy and Dr. Eva Uzvolgy whose guidance and friendship during my undergraduate career was invaluable. I always felt at home in their laboratories, and for that I will always be thankful. I am also grateful to Dr. Josiane Eid from whom I learned to love the excitement and anticipation that a single experiment can bring. I will never forget the many fun moments in the laboratory with Dr. Eid. It was the research experience in these laboratories that truly awakened my interest in science and motivated me to pursue graduate studies to obtain a PhD.

What I will cherish the most from my graduate student life are the friends that I made all along the way. I will never forget Anne Karpay, she was the strongest person I have ever met. I thank my friends Yoana Dimitrova, Bonnie Garcia and Christina Williams for all the laughs, silliness and the many moments that they were around to listen and be supportive. I am happy to have met and work with all the members of the Chazin laboratory, including extended members.

For the most happiness and fun that I could have ever dream of having, I thank my family. To my parents, brother and sister, biggest supporters, who have been always so close and loving in spite of the long distance that separate us. I will always appreciate Alberto, who has been there all along the way in spite of almost exhausting his patience, and for being at the same time my pillar and my biggest challenger. He has made me be stronger when I needed to be and given me enough happy memories to last a lifetime. I feel fortunate for having him in my life. To my children: Ivanna, Marcela, and Alessandro who I love more than anything in this world, I thank them for understanding when mom was sometimes busy and for being great loving and caring kids. For all the fun, the laughs, the special moments that never end, for their curiosity that never ceases to amaze me and inspire me, and for their faith in me, I am truly thankful. My ultimate thanks are to God, for blessing me with so many gifts and giving direction to my life.

LIST OF TABLES

Table Page

1.1. RPA domain organization, function and structures available.....4

1.2 Summary of RPA phosphorylation and RPA32N residues involved 14

2.1. SAXS measurements 51

LIST OF FIGURES

Figure **Page**

1.1. Domain Organization of RPA 2

1.2. Cartoon diagram of RPA70A OB-fold domain 3

1.3. RPA is composed of globular domains and flexible disordered linkers 5

1.4. RPA protein-protein interactions that have been characterized 10

1.5. Cyclins/cdk expression levels during cell cycle progression 17

1.6. Representation of SAXS data collection 36

1.7. Analysis of SAXS data 38

1.8. P(r) is sensitive to the shape of the molecule under study 39

1.9. SAXS envelope of RPA70AB bound to ssDNA 40

1.10. 2D ^{15}N-^{1}H HSQC spectrum of ^{15}N-enriched RPA70AB 42

2.1. Domain organization of RPA 46

2.2. TSK-gel G200sw elution profiles for RPA70AB and RPA70NAB prior to SAXS data collection 49

2.3. Superdex 200 SEC elution profiles for RPA70AB and RPA70NAB prior to SAXS data collection 50

2.4. RPA70AB scattering curves 51

2.5 CRYSOL fit to experimental data for RPA70AB models 52

2.6. Plot of $\chi 2$ fit parameter versus radius of gyration for RPA70AB and RPA70NAB conformers generated by BILBOMD 53

2.7. CRYSOL fit to experimental RPA70AB-8mer data 55

2.8. RPA70NAB scattering curves 57

2.9. CRYSOL fit to experimental data for RPA70NAB BILBOMD models.....60

2.10. Models of RPA70NAB in the presence of the ssDNA 14-mer 62

3.1. RPA minitrimer core.....74

3.2. Purification of pBG102 RPA32Nwt $_{1-35}$ and RPA32Nwt$_{1-46}$ peptides.....77

3.3. Purification of pBG102 RPA32N-D8 peptide 78

3.4. ^{15}N ^{1}H HSQC spectra of RPA32N peptides 79

3.5. ^{15}N-^{1}H HSQC spectra RPA70A and RPA70B upon titration of RPA32N-D8 82

3.6. Contact points in RPA70A and RPA70B interacting with RPA32N-D8.....84

3.7. RPA32N-D8 backbone specific resonance assignments 85

3.8. RPA70N, RPA70A and RPA70B electrostatic potential maps 86

3.9. ^{15}N-^{1}H HSQC titrations of RPA32N-D8 with RPA32D/14 and RPA32C 88

3.10. Design of pBG106 RPA constructs 89

3.11. RPA-D9 SAX scattering curves.....90

3.12. RPA-D9 Guinier Analysis 91

4.1. RPA is composed of globular domains combined with disordered regions 102

4.2. Structural dynamics of RPA 106

4.3. Surface representation of phosphorylated RPA32N peptide 108

4.4. Model for compaction upon hyperphosphorylation.....112

4.5. Hyperphosphorylation model of extended RPA 113

LIST OF ABBREVIATIONS

ATP	adenosine triphosphate
BME	β-mercaptoethanol
CD	circular dichroism
Cdc	cell division cycle protein
Cdk	cyclin dependent kinase
CSPA	chemical shift perturbation assay
DTT	dithiothreitol
EDTA	ethylenediaminotetraacetic acid
FancJ	Fanconi anemia complentation group D protein
HSQC	heteronuclear single quantum coherence
IPTG	isopropyl β-D-1-thiogalactopyranoside
ITC	isothermal titration calorimetry
IR	ionizing radiation
Kd	dissociation constant
LC	liquid chromatography
MALDI-TOF	matrix assisted laser desorption ionization – time of flight
MALLS	multi anglelaser light scattering
Mcm	minichromosome maintenance protein
MS	mass spectrometry
NiNTA	nickel-nitrilotriacetic acid
NMR	nuclear magnetic resonance
NOESY	nuclear Overhauser effect spectroscopy

OD	optical density
OBD	origin binding domain
OB	oligonucleotide/oligosaccharide binding
ORC	origin recognition complex
PDB	protein database
pol	DNA polymerase
pol-prim	DNA polymerase α/primase
pre-RC	pre-replication complex
RPA	replication protein A
SAXS	small angle x-ray scattering
SANS	small angle nuclear scattering
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
SEC	size exclusion chromatography
SEC-MALS	SEC using a multi-angle laser light scattering detection system
ssDNA	single-stranded deoxyribonucleic acid
SV40	simian virus 40
Tag	large T antigen
TopBP1	topoisomerase II binding protein 1
topo I	topoisomerase I
UV	ultraviolet

CHAPTER I

INTRODUCTION

DNA replication

DNA replication is fundamental to the survival of cells and preservation of species throughout evolution. Defects in duplication of the genome can lead to detrimental consequences even at the organism level. Mistakes in replication synthesis can introduce unfavorable mutations resulting in genetic instability and translating into the production of disabled or faulty proteins that can no longer perform their unique cellular tasks. Mutations can ultimately result in genetic disorders and progress to fatal diseases such as cancer. It is the accrual of multiple genetic mutations that result in massive cellular differentiation and gives rise to what in advance stages becomes uncontrolled tumor growth.

DNA replication in humans is mediated by complex multi-protein machinery. This machinery is built by timely expressed proteins that assemble strategically during the early stages of cell cycle, recognizing the origin of replication and licensing replication for initiation at the start of the synthesis phase where accurate and efficient elongation and termination of replication take place. At the center of the replication machinery is replication protein A (RPA), the major single strand DNA (ssDNA) binding protein. RPA binds and protects ssDNA from degradation by endonucleases while at the same time preventing the strands form re-annealing and forming secondary structures that may interfere with the smooth transition of replication. Remarkably, when replication encounters damaged DNA the replication process is stopped and depending of the type of damage, specialized repair mechanisms are triggered that respond to the

specific type of damage. One of the signals that accompany recognition of DNA damage is the hyperphosphorylation of RPA, which contributes to stalling of replication. However, it is unknown what changes in RPA are induced by hyperphosphorylation, and how these signal the switch from replication to damage response and repair.

Replication Protein A

RPA is the primary ssDNA binding protein in humans and other eukaryotes. It is modular in composition, formed by three subunits of 70, 32 and 14 KDa, and named on the basis of their respective molecular weights (Figure 1). The RPA70 subunit contains four oligonucleotide-oligosaccharide binding (OB fold) domains spanning residues 1-120 (RPA70N), 181-290 (RPA70A), 300-422 (RPA70B) and 436-616 (RPA70C). The RPA32 subunit contains one OB fold spanning residues 46-171 (RPA32D) flanked by an unstructured N-terminal domain (1-45, RPA32N) and a winged helix-loop-helix C-terminal domain (200-270, RPA32C). RPA14

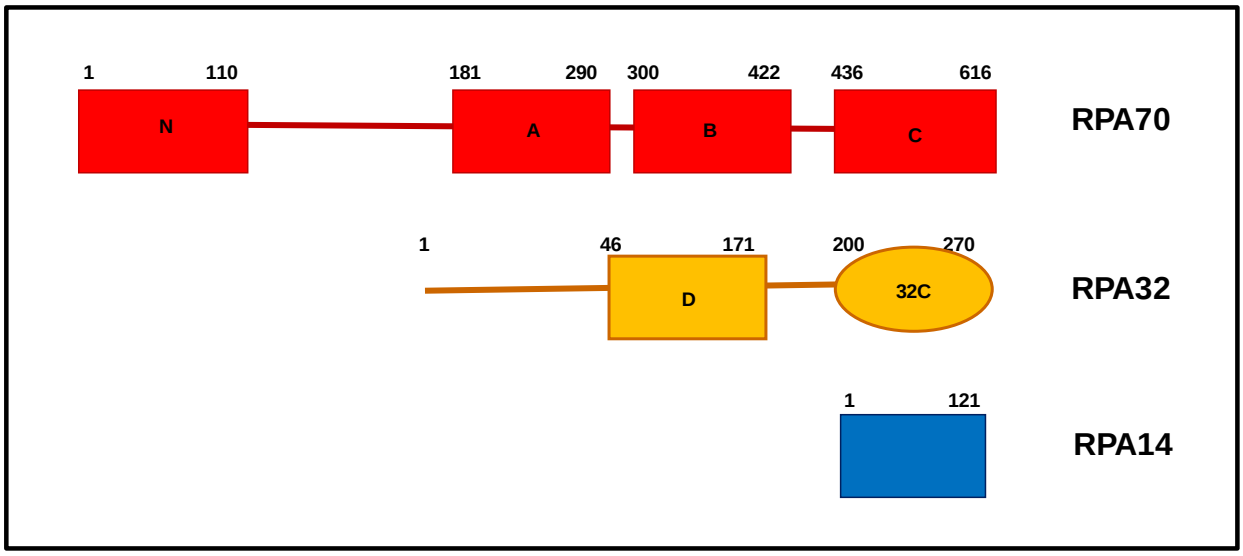


Figure 1.1. Domain organization of RPA. The domain composition of each of RPA subunits, RPA70, RPA32 and RPA14 is different. RPA has six OB-fold domains and one HLH domain.

is a single OB fold domain. Trimerization of RPA occurs through the association of RPA70C, RPA32D and RPA14 inter-domain interactions and it is only this heterotrimeric form of RPA that is active in DNA processing. The ssDNA binding activity of RPA has been mapped to RPA70A, RPA70B, RPA70C and RPA32D domain, whereas protein interaction are mediated by RPA70N, RPA70A, RPA70B, and RPA32C domains (Figure 1).

OB-folds are characterized by 5-stranded antiparallel β sheets, with an α helix between the third and fourth strands [1] (Figure 2). This motif folds into a closed β-barrel, which typically harbors a notable basic patch positioned in between loops 1-2 and 4-5 that is the preferred site of interaction for ssDNA and proteins. The basicity of the basic cleft differs from one OB-fold to another in accord with function. Each ssDNA interacting domain in RPA has a different affinity of interaction with ssDNA and binds to different specific proteins (Table 1.1).

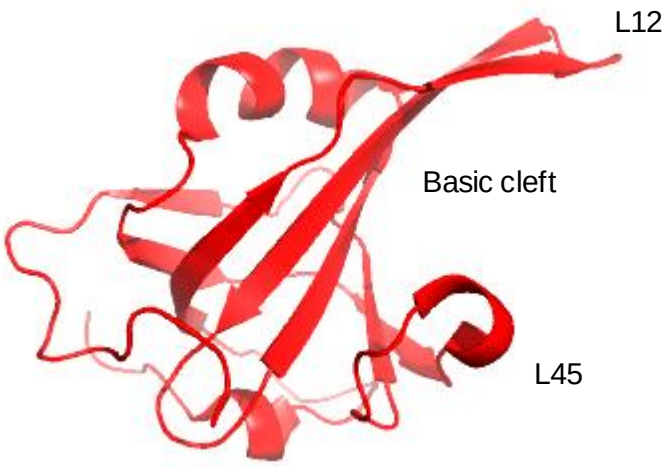


Figure 1.2. Cartoon diagram of RPA70A OB fold domain.

Subunit	Domains	DNA binding	Protein binding	Trimer core	Fold	Structure type
RPA70	N		√		OB	NMR
	A	√	√		OB	Crystal
	B	√	√		OB	Crystal
	C	√		√	OB	Crystal
RPA32	N				Disordered	
	D			√	OB	Crystal
	C		√		HLH	NMR
RPA14	14			√	OB	Crystal

Table 1.1. RPA domain organization, function and structures available.

Flexible linkers joining RPA domains provide great flexibility, which has precluded determination of RPA quaternary structure by NMR and X-ray crystallography. A simpler strategy has been effectively adopted. The tertiary structures of RPA domains in various combinations have been determined (Table 1.1). This approach, combined with functional analysis, has revealed much information about RPA function. Based on this information, several models for RPA quaternary structure have been proposed [2-4]. However, in order to understand RPA dynamics and functional control, the spatial organization of the domains within the intact protein must be characterized.

RPA ssDNA binding activity

As the universal cellular ssDNA binding protein active in DNA processing, RPA binds any ssDNA sequence. RPA uses a sequence-nonspecific ssDNA binding mechanism involving

dynamic remodeling of the binding surface, coupled with RPA domains interacting with other proteins to advance DNA processing [5].

RPA binds ssDNA with 5' to 3' polarity as a result of the differences in its domain's affinities for ssDNA [6]. RPA70A has approximately 10-fold higher affinity for ssDNA than RPA70B [7]. RPA70C and RPA32D affinities are weaker still. However, due to the tethering of multiple domains and resulting high local concentrations, the combined ssDNA binding affinity of the RPA heterotrimer is in the low nM range [8]. RPA70A and RPA70B bind ssDNA in the OB-fold basic cleft, wrapping loops 1-2 and 4-5 around the ssDNA ligand [5]. RPA70A and RPA70B are connected by a 10 amino acid flexible linker, which promotes coupled binding of

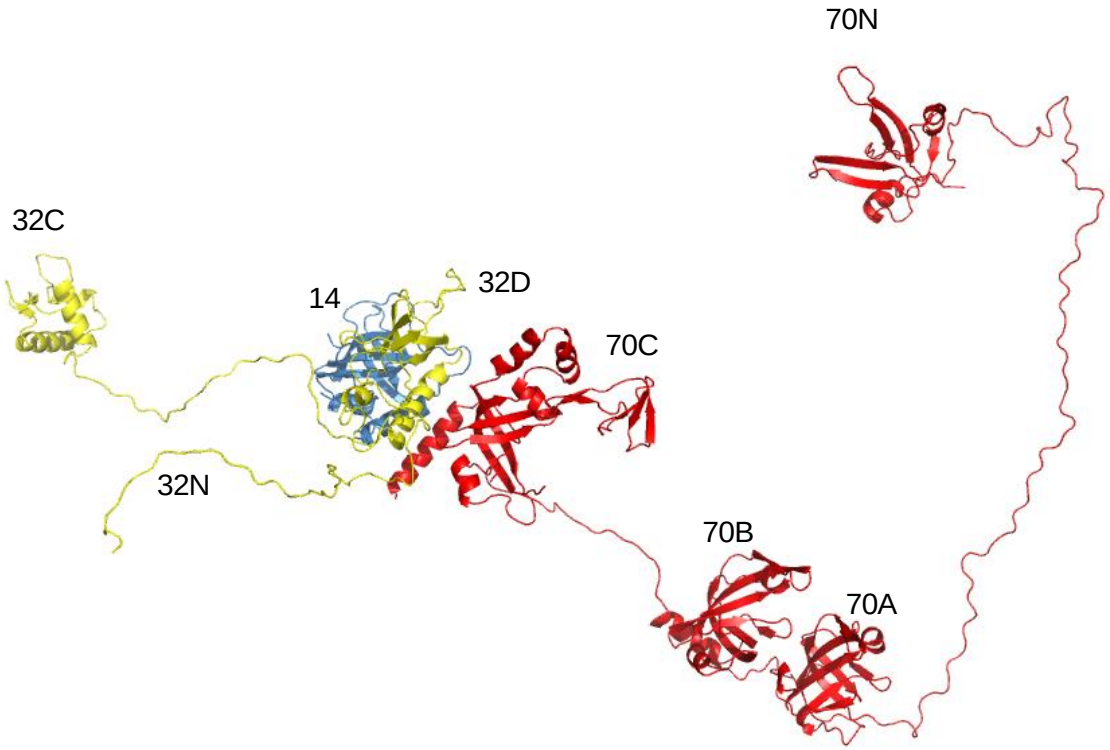


Figure 1.3. RPA is composed of globular domains and flexible disordered linkers. Each subunit is colored differently. RPA70N is connected to RPA70AB and in turn RPA70C/32D/14 is connected to RPA32N and RPA32C. Coordinates for this model have been obtained from Chris Brosey.

RPA70A and RPA70B [7]. Although this step is sequential, due to intrinsic differences in binding affinities, the RPA70A and RPA70B binding events cannot be separated. Hence, RPA DNA binding studies have shown that the first detectable binding mode consists of 8 to 10 nucleotides, followed by second and third modes occluding 13 to 20 nucleotides, and 28 to 30 nucleotides, respectively [8-10]. The X-ray crystal structure of RPA70AB in complex with dC_8 revealed that the first mode corresponds to RPA70AB binding [11]. It is proposed that sequential binding of RPA70C corresponds to the second binding event and the binding of RPA32D corresponds to the third step. Although two X-ray structures of RPA32D/14 and one of RPA70C/32D/14 provide insights into structures [12-14], and mass spectrometry finger printing analysis provided information on the distances spanning the ssDNA bound to RPA [15], exactly how the ssDNA threads onto the RPA domains has not been determined. In all, this tri-modal mechanism establishes RPA's canonical modes of ssDNA binding, which is conserved through the various DNA processing events in the cell.

RPA-protein interactions

RPA was first identified as an essential component for the *in-vitro* SV40 replication system [16, 17]. This system of viral replication has provided a simplified model for the direct analysis of chromosomal DNA replication in eukaryotes. Only one viral protein, the Large T-antigen, participates in the process, which is otherwise entirely dependent on the host replicative apparatus [18]. Thus, the use of this cell-free system allowed the identification and characterization of the proteins involved in the eukaryotic replication activities [17, 19]. RPA was identified as necessary once ssDNA regions were generated at the origin of replication, and

it was determined that RPA facilitates the unwinding of parental DNA strands by the Large T-antigen during the elongation phase of SV40 replication [17]. Since then, characterization of RPA interactions in the SV40 replication system has progressed considerably leading to a better understanding of eukaryotic replication, and the list of proteins that interact with RPA has become very extensive. It is now known that the originally identified "replication" protein A is utilized in a range of DNA processing events. However, although many interactions have been reported, little work has been completed to fully characterize most of them.

In SV40 replication, RPA interacts with the SV40 large T-antigen (Tag) and the human DNA polymerase α-primase (pol-prim). The interaction with Tag helicase is important for the loading of RPA onto ssDNA [20] while interaction with pol-prim is necessary for *de novo* RNA-DNA primer synthesis. Tag origin binding domain (Tag OBD) interacts with both the RPA70AB and RPA32C domain [21, 22]. Pol-prim p58C subunit also interacts with RPA32C (Sivaraja Vaithiyalingham, Erick Warren, Brandt Eichman and Walter J. Chazin, unpublished). Both of these interactions are necessary for initiation of replication to take place.

Other viral systems also utilize RPA during replication. Papilloma viruses depend on host replication machinery for replication of their viral genome. Papilloma virus proteins E1 (functionally equivalent to SV40 T-antigen) and E2 directly bind RPA70 subunit [23]. The uncovering of other similar modes of binding can create a vision for the overall mechanism by which these events are regulated.

In homologous recombination, Rad51, the eukaryotic recombinase active in mitosis and meiosis, interacts with RPA70A through its N-terminus (Rad51N) at the same site as ssDNA, suggesting competition between ssDNA and Rad51N as a mechanism for displacement of RPA [24, 25]. Rad51 forms a nucleoprotein filament on the ssDNA and mediates displacement of

RPA in the initiation stage of genetic recombination [25]. Interestingly, RPA70A has greater affinity for ssDNA than Rad51N, which does not explain the exact mechanism for RPA displacement. An additional RPA-Rad51interaction to RPA32/14 has been observed [24], and preliminary studies have narrowed these interactions to the RPA32C domain. Moreover, an interaction of Rad51 with the RPA70C domain has been observed (M. Stauffer, D.I. Pretto, W.J. Chazin, unpublished results). To our knowledge this is the only putative interaction to RPA70C that has been identified. Rad52, a recombination mediator protein associated with RPA displacement mechanism in recombinatorial repair, also binds the RPA32C domain [26]. Rad52 assists in Rad51 loading and in the formation of the Rad51 nucleoprotein filament on to the ssDNA [27]. BRCA2, another protein involved in regulation of Rad51 localization and DNA binding during recombinatorial double strand break repair, has also been shown to interact with RPA through its N terminus transcriptional domain, and this interaction is disrupted in mutant BRCA2 [28]. Mutations in BRCA proteins predispose women to familial breast cancer.

The same Rad52-interacting RPA32C binding interface is used by XPA (Xeroderma Pigmentosum complementation group A), and by the Nuclear Uracil-DNA Glycosylase UNG2 [26]. These proteins are involved in different DNA repair pathways. XPA is active in nucleotide excision repair (NER) where it recognizes DNA lesions. RPA-XPA interaction enhances XPA's affinity for damaged DNA and is necessary for repair to occur. UNG2 is the major enzyme in base excision repair (BER) of deaminated cytosine (U/G) and possibly initiating BER of misincorporated uracil (U/A) [29]. The Rad52, XPA and UNG2 shared binding site lends support for the molecular hand-off mechanism [26] (see below).

RPA also binds transcription factors GAL4 and Vp16 through RPA70 subunit [30]. Transcription factor and tumor suppressor P53 also interacts with RPA through the RPA70N

domain [31, 32]. P53 is reported to modulate the activity of Werner Syndrome Protein (WRN), a member of the Rec Q 3'to 5' helicase family, which also harbors 3' to 5' exonuclease activity [33]. Defects in WRN activity result in rare premature age diseases and pre-disposition to various cancers. Interestingly, WRN also directly interacts with RPA. This interaction has been mapped to the RPA70N domain [34]. Related to WRN, Foci-Forming Activity 1 protein, (FFA-1), the *Xenopus laevis* functional homologue of human WRN, is also thought to directly interact with RPA [35]. The assembly of RPA into *X. laevis* replication foci requires FFA-1, which is thought to stably associate with replication foci on nuclear chromatin and in so doing generates binding sites for RPA [36]. Yet another RecQ family member, the Bloom Syndrome Protein (BLM), which results in even more serious aging disorders and greater pre-disposition to cancer is also known to bind to RPA through the RPA70 subunit [37]. All this is consistent with our laboratory's hypothesis that helicases assist in the loading of RPA on ssDNA as observed with the T-ag helicase [20]. Moreover, the interaction with RPA also influences the helicase activities. For example, RPA has been shown to stimulate WRN branch migration activity [38].

More recently, RPA interactions with cell cycle checkpoint control proteins have been identified. The ATR interacting protein, ATRIP [39], NBS1, Mre11 [40] and Rad9 [41] proteins interact with RPA through it RPA70N domain. MRE11/Rad50/NBS1 complex as well as BRCA proteins, Rad51, Rad52 are all involved in early stages of double strand break repair. RPA interaction with all these protein complexes suggests RPA may have activity as an orchestrator of the formation of repair foci (Figure 1.4).

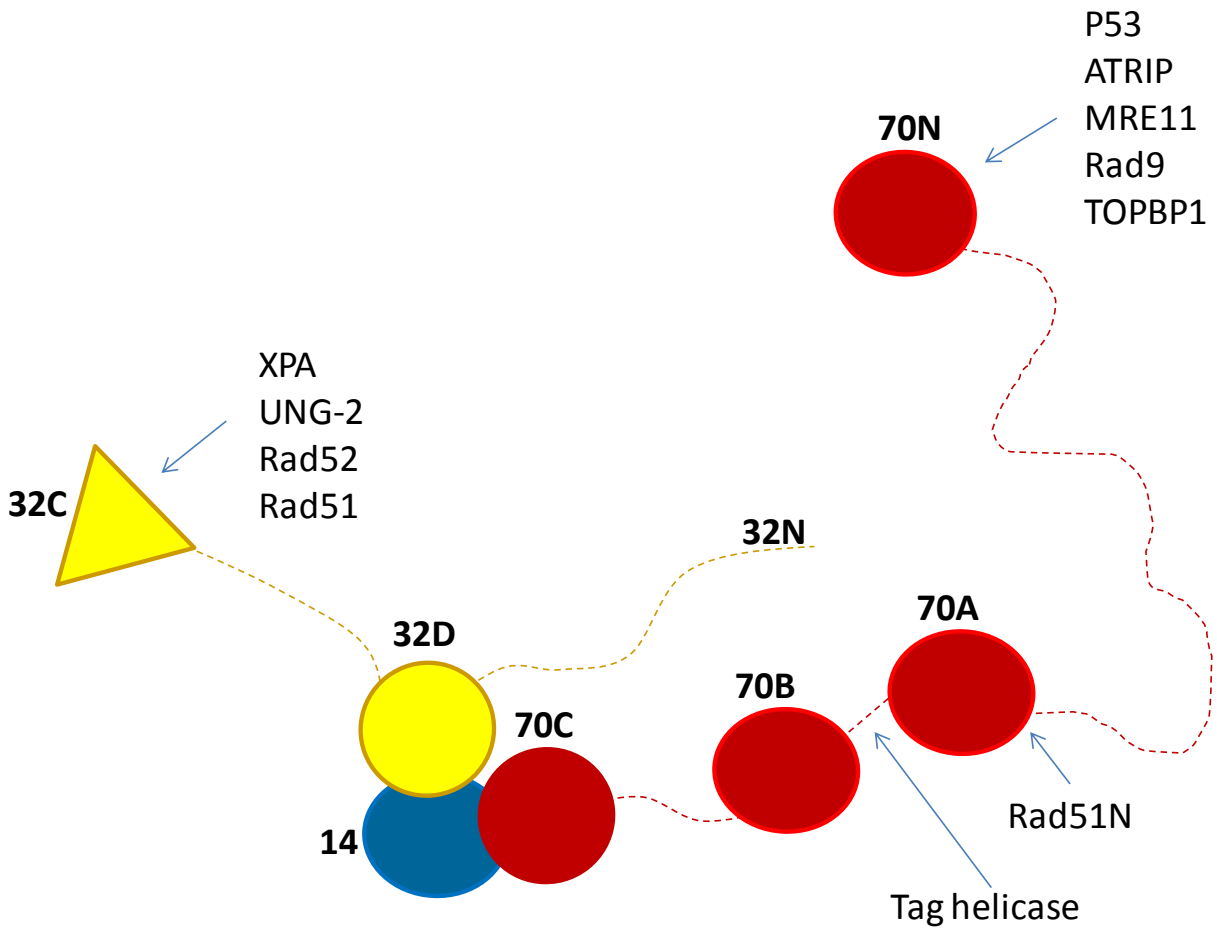


Figure 1.4. RPA protein-protein interactions that have been characterized.

Disposition of RPA70N in the RPA ssDNA binding activity

RPA70N OB fold domain is similar to the other RPA DNA binding OB fold domains (RPA70ABC/32D) given is an OB fold with a basic surface that recognizes complementary acidic surfaces [42]. This property is important for the binding of target proteins. It has also been proposed that RPA70N binds ssDNA [43, 44]. However, there is a critical difference. RPA70N lacks functionally conserved aromatic residues shown to be critical for

RPA70ABC/32D interaction with DNA [1]. In particular, beyond the extensive network of hydrogen bonds, the RPA70A domain contributes two phenylalanine residues, F238 and F269, which form stacking interactions with the DNA bases to form a stable RPA-ssDNA complex. These aromatic residues are conserved in RPA70B, with W361 and F386 taking part in what the authors describe as a domino interaction, involving stacking of the aromatic residues with the DNA bases [45]. The conservation of these key amino acid residues is extended to RPA70C (Phe532 and Tyr581) and RPA32D (Trp107 and Phe135), which participate in the second and third mode of ssDNA binding respectively [14]. The RPA14 subunit is also an OB-fold and like 70N it lacks aromatic residues in the basic cleft. No DNA binding activity has been detected for the human RPA14 subunit.

In spite of the lack of aromatic residues in its basic cleft, a model has been proposed in which RPA70N is suggested to actively participate in the destabilization of the DNA double helix through direct binding to DNA [44]. DNA helix destabilization at origins of replication is necessary for the assembly of the replication machinery. This process is described in two steps known as the nucleation step, which involves presumably binding of the protein and initial unwinding of the double helix creating a small ssDNA bubble; and the melting step, which is the extra dsDNA separation that leads to a double helix destabilization creating more ssDNA.

It is of importance to clarify that this RPA helix destabilization activity is not like the unwinding activities of a DNA helicase. Helicases are proteins dedicated to rapidly and efficiently unpacking large stretches of dsDNA to give way to replication. Massive amounts of ssDNA need to continuously be made available to the replication machinery and the RPA melting would not suffice in creating enough ssDNA. A study aimed at understanding the mechanism by which RPA denatures pseudo-origin substrates (segments of DNA that contain

SV-40 origin of replication sequences) found that only a third of the substrates were denatured indicating that RPA could bind to the partial duplex substrates without causing their complete denaturation [46]. RPA can unwind approximately 30 base pairs of dsDNA by itself (nucleation step), and it was shown that it may load itself onto the ssDNA bubble [47, 48]. However, further studies have shown that rather, RPA is loaded on to the ssDNA by the Tag helicase (Jiang 2006). Overall, RPA is believed to stabilize the formation of ssDNA preventing it from forming any secondary structure or reforming a double helix. This activity promotes partial denaturation of the origin of replication.

Interestingly, helix destabilization studies link RPA70N activities to the hyperphosphorylation of RPA32N domain. In *in vitro* DNA helix destabilization experiments, an RPA70N knock-out mutant is defective in DNA nucleation and DNA strand separation during dsDNA helix destabilization steps that precede DNA replication [3]. The same is true when a pseudo-hyperphosphorylated RPA mutant is used in the reaction. Previous NMR chemical shift perturbation experiments demonstrated a direct, although weak, interaction between ^{15}N-RPA70N and a synthetic pseudo-hyperphosphorylated RPA32N peptide [44]. Together, this lead to a hypothesis that RPA70N intersubunit interaction with hyperphosphorylated RPA32N domain removes RPA70N from its putative DNA bound location, and that this action causes a defect in DNA helix destabilization since RPA70N is no longer present to secure the RPA-DNA interaction that is thought to promote the initial unwinding of the double helix [44]. However, evidence in the literature also link RPA70N and hyperphosphorylation of RPA32N to regulation of RPA70N protein-protein interactions such is the case of RPA70N with the MRN complex [40]. It is more likely that hyperphosphorylation of RPA affects RPA functions through modulation of protein-protein interactions.

The idea that RPA70N directly participates in helix destabilization binding to ssDNA is difficult to reconcile with RPA ssDNA binding mechanism. The recognized RPA ssDNA binding domains, RPA70A, RPA70B, RPA70C and RPA32D, bind ssDNA with 5' → 3' directionality. The ssDNA binding activity is initiated by the high affinity ssDNA binding domains RPA70AB, sequentially engaging RPA70C and RPA32D. This sequential mode of binding is counter to the position of RPA70N in the intact protein. In addition, multiple protein-protein interactions have been identified through RPA70N. I therefore set out to investigate RPA70N behavior in the ssDNA binding activity of the heterotrimer, which forms the basis of the studies described in Chapter II.

Phosphorylation of RPA

Phosphorylation is a common mechanism of regulation of protein activity, and often occurs on more than one distinct site on a given protein. A protein is said to be hyperphosphorylated when multiple sites are phosphorylated, and hyperphosphorylation of proteins during the cell cycle has been previously observed. Cellular proteins including the tumor suppressor Rb (human) involved in checkpoint control and the metaphase checkpoint protein Wee (human) are inactivated as a result of hyperphosphorylation [49]; other proteins appear to be activated by hyperphosphorylation [50]. RPA becomes partially phosphorylated (termed hypophosphorylated) during cell cycle activities and shifts to a hyperphosphorylated state as a response to DNA damage. There is great deal of confusion about which kinases phosphorylate which RPA residues, the order of phosphorylation events and their outcomes. Therefore, as part of my studies, I have gathered all available information and attempted to reconcile the data to help in assessing the current state of the field.

The possibility that RPA is post-translational modified was first investigated following the discovery that the unwinding of the origin of replication by T-antigen and cellular proteins was dependent on the cell cycle stage from which the cell extracts were isolated [51]. As a result, it was investigated if variations in the amount of RPA could account for this cell cycle specific unwinding activity; however, this was not the case. The abundance of RPA remained unchanged throughout the cell cycle, but a cell cycle dependent phosphorylation of RPA was detected in human and yeast cells [52] and later in Xenopus oocyte extracts [53]. It was observed that phosphorylation was limited to the S and G2 phases of the cell cycle, but not the G1 phase with de-phosphorylation occurring late in mitosis, thereby resetting the phosphorylation cycle [52]. The association of RPA32 with the other two subunits, RPA70 and RPA14, was shown to be essential for phosphorylation to occur [52]. Interestingly, it has also been reported that phosphorylation dissociates the heterotrimer and its been speculated that phosphorylation and changes in subunit interaction are required for the proposed role of RPA during the polymerase switch at replication forks [54]. Table 1.2 provides a summary of the sites on RPA that are phosphorylated.

Although RPA phosphorylation was detected primarily on the RPA32N domain, phosphorylation has been observed in other RPA domains [55]. Up to five phosphorylation sites in the RPA70C subunit (Ser569, Thr580, Ser585, Thr590, and a single methionine), and one additional site in the RPA32D subunit (Thr98) were observed using a cocktail of kinases *in vitro.* However, these results were not reproduced *in vivo* when HeLa cells were arrested in S and G2/M transitions using aphydicolin followed by treatment with either hydroxyurea to stall replication forks or using UV light to cause ssDNA damage. Instead, only phosphorylation in the RPA32N domain and a possible single phosphorylation site in the RPA70N-A linker region

Event	Hypo-phosphorylation	Hyper-phosphorylation
Occurrence	During normal cell cycle activity	In response to ssDNA damage
RPA32N phosphorylated residues	Thr21, Ser23	Ser4, 8, 11, 12, 13, 23, 29, 33 Thr21
Responsible Kinases	Cyclin Dependent Kinases	PI3K family of Kinases

Table 1.2. Summary of RPA phosphorylation and RPA32N residues involved.

(between residues 112 and 157) were identified. A specific functional role for the RPA70N-A linker phosphorylation has yet to be assigned [50]. It has been suggested that phosphorylation sites observed *in vitro* correlate with a normal process in cell cycle progression, however, previous studies have not identified modification in the RPA70 subunit during progression of the cell cycle [52]. Thus, much attention has been focus in understanding the phosphorylation of the RPA32N domain instead.

RPA32N is an unstructured region that harbors nine potential phosphorylation sites: Eight serines and one threonine (Ser 4, Ser 8, Ser 11, Ser12, Ser13, Thr21, Ser23, Ser29, Ser33). These residues are targets for phosphorylation by cell cycle checkpoint kinases and are presumably modified in a cell cycle dependent manner and as a response to DNA damage. The specific role of these phosphorylation events has begun to be understood. The timing of phosphorylation suggests a regulatory role during DNA synthesis, but questions still remain as to how many and which sites are phosphorylated in the hyperphosphorylated state.

At least four distinct species of phosphorylated RPA have been identified based on their migration pattern on SDS-PAGE gels [56]. These have been called RPA32 Forms 2, 3, 4, and 5.

The non-phosphorylated RPA is referred to as Form 1. Forms with increased level of phosphorylation travel slower in the polyacrylamide gels and are designated a higher number based on their slow speed of migration. Form 5 therefore is the hyperphosphorylated RPA32. Form 2 and Form 3 are termed hypophosphorylated forms [57, 58].

To characterize the various RPA32N phosphorylation forms, SDS-PAGE gel patterns of purified and trypsinized RPA from non-irradiated and UV irradiated cells were compared. These phosphopeptide maps show that Form 2 and Form 3 remain essentially the same pre or post UV irradiation, which suggests these hypophosphorylated RPA forms exist during normal cell cycle. Form 5 was most frequently observed post UV irradiation, and it was shown this hyperphosphorylated form contains the same phosphorylation sites from Forms 2 and 3 in addition to other sites. A comparison of predicted phosphopeptide maps to the authentic maps was used as a tool to identify the possible residues involved in the various phosphorylated forms of RPA, identifying Thr21, Ser23, Ser29, Ser 33 and either Ser11, Ser12 or Ser 13 and probably Ser4 and Ser8 as sites of phosphorylation in the hyperphosphorylated form of RPA produced by UV irradiation. For form 2, three phosphorylation events are proposed and at least five phosphorylation events for form 3 and seven for form 5 [57]. This is consistent with the notion that hypophosphorylated forms of RPA act as a precursor for hyperphosphorylated forms.

Evidence suggests that RPA hypophosphorylation is a normal cell cycle dependent process in regulation of replication. Hypophosphorylation occurs at the replication initiation complex [59], beginning at the G1 to S transition phase with dephosphorylation occurring late in mitosis [52]. It has been shown that hypophosphorylated RPA has decreased interaction with DNA pol α, ATM and DNA-PK, and that it may have an effect in RPA binding to dsDNA at both damage and undamaged sites probably affecting RPA unwinding activity or, alternatively,

protein-protein interactions [60]. However, the primary activity of RPA is to bind ssDNA rather than dsDNA. RPA binding to dsDNA was shown to occur via denaturation of thermally unstable cisplatin damaged DNA while binding to undamaged dsDNA was not significant [61]. Thus, during the normal progression of DNA synthesis RPA does not bind and unwind dsDNA. This putative RPA hypophosphorylation role in the regulation of RPA function in replication has not been investigated any further and more emphasis has been given to understanding the effect of hyperphosphorylation.

Hypophosphorylation of RPA by Cyclin Dependent Kinases

RPA becomes hypophosphorylated in a cell cycle specific manner under normal conditions by cyclin/cdk complexes. The hypophosphorylation is thought to regulate RPA activity in DNA replication [52]. Cyclin dependent kinases (cdk) are serine/threonine kinases involved in the regulation of cell cycle, transcription and mRNA processing. Activation of CDKs occurs through binding of cyclins, which are cell cycle proteins expressed at specific stages of the cell replication cycle (Figure 1.5). Cyclin-CDK active complexes then phosphorylate their targets controlling the progression of cells through the cell cycle.

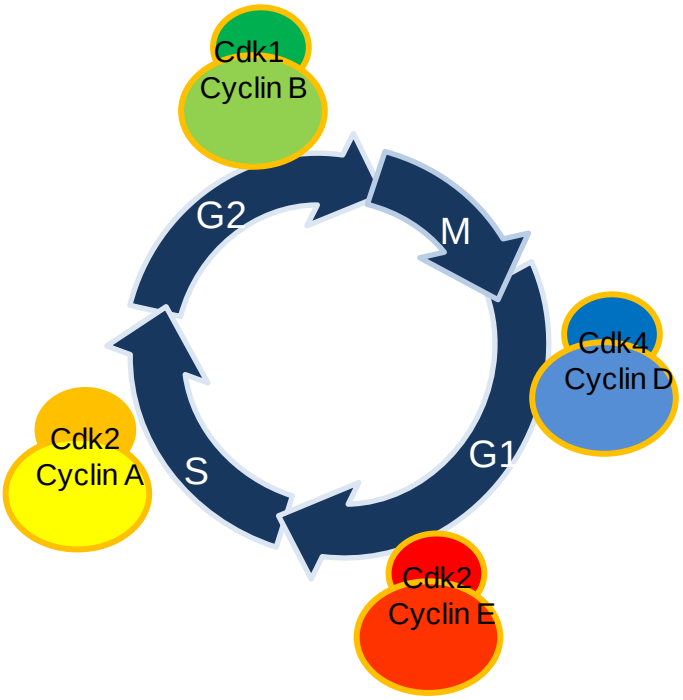


Figure 1.5. Cyclins/cdk expression levels during cell cycle progression.

Figure 1.5. Cyclins/cdk expression levels during cell cycle progression.

Analysis of the RPA32 primary structure implicated residues Ser23 and Ser29 as candidates for cyclin/cdk phosphorylation based on the established cyclin/cdk preferred substrate recognition sites (Ser or Thr followed by Pro) [62].

Cell cycle kinases are responsible for various RPA phosphorylation events during cell cycle progression. Using cell fractionation techniques, cytoplasmic fractions from HeLa cells were purified that contained cyclin A and cyclin B related kinase activities. These fractions were shown to be active in phosphorylating RPA32. In addition, comparison of the levels of RPA32 phosphorylation in mutant (Ser23→Ala and Ser29→Ala) and WT stable cell lines demonstrated that these residues are necessary for phosphorylation of RPA by the cdc2 kinase related activity in NH3T3 cells [63]. However, it is difficult to know with certainty which combination of kinases actively phosphorylate RPA *in-vivo* because cdc2 and cdk2 can both associate with cyclin A and cyclin B [64, 65].

Cyclin B/cdc2, cyclin A/cdc2 and cyclin A/cdk2 complexes were all able to phosphorylate RPA32 *in-vitro* [63, 65]). Phosphopeptide mapping of an *in-vitro* phosphorylated and chymotrypsin digested RPA32 synthetic peptide containing residues 2-42 showed that the cyclin A/cdc2 complex can phosphorylate the same RPA32 peptides as the cyclin B/cdc2 complex, and that these phosphorylation events were only a subset of all phosphorylation events identified in 293 cells. These results revealed that other kinase activities could be involved in the phosphorylation of RPA, as was also observed by others [59]. Sequencing of the peptides confirmed that Ser23 and Ser29 were the targets for these cyclin dependent kinases and judging from the apparent levels of phosphorylation observed by SDS-PAGE gels, it can be inferred that cyclin B/cdc2 phosphorylation activity was more efficient than cyclin A related phosphorylation activity [63]. Similarly, *in-vitro* phosphorylation assays using purified cyclin/kinase complexes

demonstrated that the cyclin B/cdc2 complex was 5-times more efficient at phosphorylating RPA32 than the cyclin A/cdk2 complex [65]. Further, the *in-vitro* reconstitution of cdc2 and cdk2 kinase activities by addition of purified either cyclin A or B1 demonstrated that phosphorylation was only observed in the presence of cdc2 and not of cdk2 with either cyclin A of B1, and that phosphorylation in the presence of cyclin B1/cdc2 was again 4 to 5 times more efficient than phosphorylation by cyclin A/cdc2 [65]. These results lead to the conclusion that cyclin B/cdc2, and most likely cyclin A/cdk2, are the preferred combinations of cyclin/kinase complexes that phosphorylate RPA32 in the cell. Interestingly, phosphorylation products resulting from addition of a fraction enriched with cyclin A/cdk activity (5S glycerol gradient fraction which also contained cdc2 as determined by immunobloting) or addition of commercially purified cyclin B/cdc2 were of similar sizes, indicating that the number of sites phosphorylated were similar [56]

Another study addresses the role of cdc2 and cdk2 in RPA phosphorylation during the cell cycle by examining the course of phosphorylation in *Xenopus laevis* egg extracts. These cells are in a constant cycle between S and M phase, therefore lacking the G1 and G2 phases. It was determined that cell cycle phosphorylation of RPA occurred during replication initiation because treatment with aphidicolin, a polymerase inhibitor, did not abolish phosphorylation of RPA in the nucleus despite arresting elongation. Interestingly, RPA in the cytoplasm was not phosphorylated, thus it was concluded that the kinase phosphorylating RPA or a precursor for recognition by the kinase must reside in the nucleus [53]. Using immuno-depletion and rescue experiments it was determined that cdc2 phosphorylates free RPA during mitosis but not during S phase and that cdk2 is essential, but not sufficient, for phosphorylation of DNA-bound RPA in an S-phase dependent manner. This indicated the existence of another kinase activity during S

phase and suggested that cdk2-mediated RPA32 phosphorylation does not occur in a direct manner. These findings are consistent with previous evidence that cdc2 kinase phosphorylates RPA32 [63], and also that a kinase unrelated to cdc2 or cdk2 is involved in the phosphorylation of RPA32 [59, 63].

Cdc2 and cdk2 are cyclin dependent kinases with redundant roles in regulating G1/S transition. Cdk2 is constitutively expressed in *Xenopus* egg extracts, however it was demonstrated that cdk2 cannot compensate for the loss of cdc2 in mitosis and its depletion from interphase extracts can block DNA replication during the initiation steps [66]. Consequently, it is generally established that cdk2 pushes cells through G2 and cdc2 in complex with cyclin B regulates the entry and exit to mitosis [64]. The kinetics of these kinases and the literature reviewed are therefore consistent with a cyclin A/cdk2-mediated RPA phosphorylation occurring during S phase and a cdc2-mediated RPA phosphorylation event occurring during G2-M phases of the cell cycle, as part of the normal cell cycle progression.

Hyperphosphorylation of RPA as a response to DNA damage

RPA hyperphosphorylation occurs when DNA damage is inflicted by UV, IR or chemical agents [57, 58, 67]. Moreover, transient HeLa cell transfections with wild type myc-RPA32 following knockdown of endogenous RPA32 by siRNA exhibit RPA foci formation at sites of DNA damage post UV irradiation as shown by colocalization with the DNA damage marker γH2AX. The same phenotype is observed when cells are transfected with a pseudo-phosphorylated RPA32 mutant (Ser→Asp), but not with a phospho-deficient mutant (Ser→Ala) that, instead, shows deficiency in foci formation [68]. This suggests that when RPA becomes hyperphosphorylated it remains localized with damaged DNA but the functional effect of

hyperphosphorylation on RPA function is not very well understood. RPA-protein interactions seem to be altered when RPA is hyperphosphorylated, favoring RPA binding to proteins involved in DNA repair activities such as Rad51 and Rad52 [69], contrary to what is observed with the hypophosphorylated RPA and DNA pol α, reflecting a more avid binding of DNA repair proteins to hyperphosphorylated RPA. These observations led to proposals that RPA hyperphosphorylation serves as a recognition signal for eliciting DNA repair responses.

The regulation of the phosphorylation state of RPA is also an active area of research. Recently, it has been shown that in cells recovering from hydroxyurea (HU) induced damage, the serine/threonine protein phosphatase 2A (PP2A) can dephosphorylate RPA, and although dephosphorylation does not appear to be important for normal checkpoint activation and re-entry into the cell cycle, persistent hyperphosphorylation of RPA results in defective DNA repair [70]. Consistent with this, PP2AC has been used throughout the literature to dephosphorylate RPA in order to confirm that the bands in SDS-PAGE gels correspond to phosphorylated species. Similarly, protein phosphatase 4 complex (PP4) can also dephosphorylate RPA, and when PP4 is absent, elevated levels of hyperphosphorylation impede homologous recombination repair of double strand breaks by preventing loading of Rad51, thus increasing the cells susceptibility to DNA damaging agents [71]. Information regarding exactly which kinases phosphorylate RPA is much more complex, as there are a great number of studies and multiple kinases are involved in the phosphorylation of RPA. RPA32N is hypophosphorylated by cyclin dependent kinases, and hyperphosphorylated by kinases active in the cellular response to DNA damage. Evidence for the kinases implicated in these two separate RPA phosphorylation states is discussed in the next sections.

Phosphatidylinositol-3 kinase-like (PI3-K) kinases are responsible for hyperphosphorylation of RPA. It is well established that this family of Ser/Thr kinases is involved in the response of DNA damage. DNA-dependent protein kinase (DNA-PK), ataxia telangiectasia mutated (ATM) kinase, and ATM- and Rad3-related (ATR) kinase are members of this family [71]. DNA-PK and ATM recognize damage to dsDNA and share similar sequence recognition patterns. In cell studies, dsDNA damage is inflicted by used of ionizing radiation (IR) or bleomycin (a radiomimetic agent) or camptothesin (CPT), which trap topoisomerase 1 cleavage complexes by obstructing their DNA interface creating a collision of DNA replication forks [72].

ATR responds to a variety of damage signals that include intrastrand crosslinks, oxidative damage and polymerase toxins [73], thus mostly recognizing ssDNA damage. This type of damage can be inflicted by the use of UV (typically 10, 30, or 60 J/m^2) or hydroxyurea (HU), which inhibits DNA replication by inactivating ribonuclease reductase. ATR works in conjunction with the ATR-interacting protein (ATRIP), which is essential for all ATR activities [73]. Importantly, ATRIP is required for ATR's localization to damage-induced foci [39]. Large-scale proteomic analysis demonstrated that these kinases phosphorylate a large number of proteins and identified more than 900 regulated phosphorylation sites encompassing over 700 proteins [74]. RPA is a target of DNA-PK, ATM and ATR/ATRIP, resulting in the hyperphosphorylation of RPA.

The work presented in Chapter III is concerned with the hyperphosphorylation of RPA32N. Our initial interest in RPA32N was motivated by observation that the NMR spectrum of the full length wild type RPA32/14 heterodimer in the presence of ssDNA showed the appearance of glycine resonances that belonged to RPA32N [26]. A similar result was observed

in NMR studies of full-length RPA, but only in the presence of ssDNA [4]. This data suggested that remodeling of RPA in the presence of ssDNA may result in the displacement of RPA32N, presumably allowing greater access to kinases. Evidence of intersubunit interactions involving a synthetic RPA32N peptide that mimics hyperphosphorylation also indicated a potential for RPA remodeling involving RPA32N [44]. In addition, the fact that this negatively charged peptide could interact with the basic cleft of RPA70N suggested that RPA32N interaction may also be mediated through other RPA domains. Together, these observations suggested RPA32N intersubunit interactions may be involved in some type of regulatory mechanism for RPA function. I therefore set out to investigate intersubunit interactions involving RPA32N, as described in Chapter III.

DNA-PK mediated phosphorylation of RPA

DNA Protein Kinase (DNA-PK) is a nuclear serine/threonine kinase composed of a catalytic subunit, DNA-$PK_{Cs,}$ (470 kDa), and a DNA targeting factor, Ku70/Ku80 heterodimer or Ku antigen, (150 kDa). Ku antigen binds at dsDNA break sites and recruits and activates DNA-PKcs. This positions DNA-PK near target proteins at DNA break sites, such as RPA.

DNA PK was first found to phosphorylate purified RPA *in vitro* following reports that cyclin dependent kinases can phosphorylate RPA [56]. Incubation of DNA PK with purified RPA in the presence or absence of ssDNA resulted in three slow migrating forms of RPA distinctly visualized on SDS-PAGE gels. Comparing migration patterns to two slow migrating RPA forms from cyclin A/cdk or cyclin B/cdc2 mediated RPA phosphorylation, these experiments showed that DNA-PK phosphorylates RPA at different sites than the cyclin dependent kinases. When purified RPA was co-incubated with DNA PK and either cyclin

A/cdk or cyclinB/cdc2, five RPA32 slow migrating forms similar to those observed when phosphorylation via DNA-PK incubation with cyclin A activated G1 extracts were observed. This suggests that the slowest RPA32 migrating forms result from the combined actions of cyclin dependent kinases and DNA-PK. The same pattern was produced when DNA-PK is present in the G1 extracts prior to the addition of cyclin A or commercially purified cyclin B/cdc2. Since the faster migrating forms of phosphorylated RPA32 (attributed to the action of cyclin A or cyclin B/cdc2 complexes) accumulate earlier in the reaction than the slowest RPA32 migrating forms resulting from the combined action of DNA-PK and cyclin dependent kinases, it was proposed that there is a requirement for cyclin A/cdk activity for the phosphorylation of RPA32 by DNA-PK. However, in the absence of cyclin A, or when p21 (Cdk-interacting protein 1 or Cip1), a specific inhibitor for cyclin dependent kinases [75], was added to the *in-vitro* phosphorylation reactions, hyperphosphorylation of RPA was nearly abolished, but not hypophosphorylation [56]. Thus, it was proposed that cyclin A/cdk and not cdc2 activity may act as a precursor to DNA-PK-mediated RPA32 phosphorylation, perhaps by activating Ku antigen promoting DNA-PK activity. It is also possible that cdc2 activity may be responsible for RPA32 hypophosphorylation that persists upon p21 treatment.

Given that RPA phosphorylation is dependent on cell cycle dependent kinases, in order to understand its role in regulation of DNA replication, the SV40 *in vitro* replication system was used to examine RPA phosphorylation and its effect in DNA replication in HeLa cells under genotoxic stress [58]. UV-induced DNA damage generated by a 10 J/m^2 dose resulted in DNA replication arrest in HeLa cells for approximately 8 hours and complete recovery by 24 hours. However cell cycle progression in cells irradiated with 30 J/m^2 doses did not recover after 24 hours. The replication arrest was consistent with the appearance of primarily

hypophosphorylated forms of RPA, which decreased in concentration between 0 and 8 hours, appearing again at 12 hours and increasing in concentration up to 24 hours when 10 J/m^2 dose was used, suggesting that these forms are important for recovery of replication activity. Hyperphosphorylated forms were observed as late as 18 hours after irradiation and remained constant until 24 hours when 10 J/m^2 dose was used. In contrast when 30 J/m^2 dose was used, hypophosphorylated RPA decreased during the first 2 hours and did not reach high concentration by 24 hours. The hyperphosphorylated forms predominated as early as 4-8 hours and persisted until 24 hours. This suggested that hyperphosphorylated RPA is not active in replication. Notably, addition of purified RPA restored the replication activity, and SDS-PAGE analysis of the *in vitro* assays shows that the added RPA did not substantially increase the levels of hyperphosphorylated RPA form, but did increase the levels of hypophosphorylated forms. Thus, exposure of HeLa cells to UV irradiation slows DNA synthesis and alters the phosphorylation pattern of RPA. However, while the hyperphosphorylated RPA is not active in replication, SV-40 *in-vitro* replication assays showed that DNA replication activity was restored whether or not the hyperphosphorylated form of RPA is present, suggesting that its activity is not necessarily inhibitory [58].

Further confirmation of DNA-PK mediated RPA phosphorylation came from a study showing that DNA-PK immunodepleted extracts were unable to phosphorylate RPA32. However, immunodepletion of DNA PK had no effect on the extent of SV-40 replication *in vitro* [76]. This suggested that DNA PK is not involved in the hypophosphorylation of RPA. Later studies, comparing two dimensional SDS-PAGE migration patterns showed that, *in-vitro,* DNA-PK can phosphorylate the same set of RPA32 phosphopeptides that are phosphorylated in HeLa cells after UV irradiation. Comparison to predicted phosphopeptide maps suggested that the

phosphorylated residues in the observed peptides may include a combination of Thr 21 and Ser 23, or Ser 29 and Ser 33, and one site on Ser 11, 12, or 13 [57]. Notably, a reduction of 70% in RPA phosphorylation was observed in the DNA PK *in-vitro* assays, which may be explained by the absence of cyclin dependent kinases in the *in-vitro* reaction, consistent with the cyclin kinases contribution to the overall phosphorylation of RPA. Consistent with the DNA PK predictions, chemical sequencing and mass spectrometry analysis identified Thr21 and Ser 33 RPA32N residues as the specific targets of DNA-PK phosphorylation [77]. These sites are consistent with the consensus preferred phosphorylation sequence for DNA-PK, a serine or threonine residue followed or preceded by a glutamine residue. Additionally, this study also demonstrated that Ser29 is the predominant, although presumably not the sole cdc2 phosphorylation site. This is consistent with previous evidence suggesting that Ser29 was a target for cyclin dependent kinases as discussed above.

DNA-PK phosphorylation of RPA has also been studied in DNA-PK cell lines. SV-40 *in-vitro* replication assays using cytoplasmic extracts from HeLa, M059J (which lack DNA PK activity and have low expression of ATM), and AT cells (which lack ATM activity) treated with IR or CPT showed a reduction in replication activity [78]. The inhibition was only reversed when the extracts were treated with wortmannin, a non specific kinase inhibitor, either prior or post treatment and recombinant wild type RPA was added. However, it was concluded that DNA PK mediated phosphorylation of Tag and not of RPA was the cause of the replication arrest, since reactions with phosphorylated RPA from HeLa cell extracts did not inhibit replication while reaction with phosphorylated Tag did. Moreover, replication assays with a 32N knock out mutant had no effect on replication activity [78]. On the contrary, another study using the same *in vitro* assay using cytoplasmic extracts from M059K cells (which have normal levels

of DNA PKcs and ATM) and M059J cells showed a role for DNA PK in replication arrest through modulation of RPA activity [79]. Extracts derived from M059K cells 4 hours post UV irradiation exhibited reduced replication capacity. However, they recovered replication capacity starting at 8 hours and full replication activity was observed at 24 hours post UV. In contrast, although no decrease in replication was observed in M059J cells 4 hours post UV, a gradual and irreversible decrease in DNA replication activity was observed as early as 8 hours post UV and replication activity was minimal at 24 hours. The same gradual decline in replication effect was observed in M059K extracts of cells subjected to UV irradiation when they were incubated with 20 uM wortmannin, a PI-3 Kinase inhibitor, preceding the replication assay. Additionally, the effect is only attributed to DNA-PK and not other PI-3K kinases since addition of DNA-PK to DNA-PK-immunodepleted extracts resulted in replication arrest. This lead to the conclusion that DNA-PK activity is involved in UV-induced replication arrest through modulation of RPA and may be is also important for restoring the replication activity. The timing of these phosphorylation events in the replication reactions may explain why the previous study had opposing conclusions. Moreover, the type of genotoxic stress used in these studies was different and the response mechanisms may be different as well.

ATM-mediated phosphorylation of RPA

ATM kinase is necessary for the immediate response to DNA damage incurred in all phases of the cell cycle [80]. To determine its role in RPA activity several studies have used AT cell lines (defective in ATM) to probe the genotoxic stress induced hyperphosphorylation of RPA. There are a number of different AT cell lines that code for inactive mutant ATM including AT3ABR, AT5ABR and AT5BI, which have known mutations, and AT3Be and AT2SF, which

have unknown mutations. These have been used to assess the necessity of ATM in RPA phosphorylation in various studies. GM08436, GM01526 and GM03189 cells are homozygous for AT mutations and don't express ATM [67]. AT3ABR, AT5ABR, and AT2SF do not express detectable levels of ATM and low levels are expressed in AT5BI [81]. AT5BIVA and AT3BISV cells are also defective in ATM [82].

Rapid RPA hyperphosphorylation was observed in response to ionizing radiation when Raji cells exposed to 10, 50 or 100 Gy of IR, pulsed with ^{32}P and immunoprecipitated with RPA32 or RPA70 monoclonal antibodies. [67] Phosphopeptide mapping of S/G2 hyperphosphorylated RPA in comparison to the species resulting from γ irradiated cells showed these contained the same peptides, and sequencing revealed that only serine residues were phosphorylated. Hyperphosphorylation occurred mainly in G1 phase and diminished through S phase, and it was estimated that cdc2 phosphorylation at Ser23 and Ser29 is required for γ induced phosphorylation mediated by ATM. The phosphorylation of RPA in response to low γ doses was seen after 2 hours in AT cells (GM08436, GM01526, and GM03189). This is very fast compared to normal cells, in which hyperphosphorylation was not noticeable until after 45 minutes. Similarly, using 5kJ/m2 of UV, hyperphosphorylation was observed as early as 2 hours [67].

ATM immunoprecipitated from HeLa cells irradiated with 10 Gy ionizing radiation (IR) supplemented with DNA was shown to phosphorylate RPA32 *in-vitro*. This was further corroborated by the inability of AT2SF immunoprecipitates to phosphorylate RPA. Additionally, no phosphorylation of RPA32 was observed in the absence of the added mixture of ssDNA and dsDNA [81]. Stimulation of this phosphorylation reaction by DNA was corroborated by others showing that M13 single stranded closed circular DNA, in the presence of

sheared DNA, greatly stimulated ATM-mediated *in-vitro* RPA32 phosphorylation in assays containing ATM isolated from human placenta in the presence of $MnCl_2$ [83]. This is consistent with the notion that ATM binds dsDNA prior to phosphorylating its targets. Two dimensional phosphoamino acid analysis showed that ATM phosphorylates threonine and serine residues of RPA [81].

The timeline for RPA hyperphosphorylation observed in two human skin-derived normal fibroblasts cell lines, LM217 and GM00637, irradiated with 10 J/m^2 UV [82] was similar to that observed in HeLa cells [58]. While hypophosphorylation of RPA remained throughout the cell cycle up to 32 hours after irradiation, hyperphosphorylation occurred after 4 hours, peaked at 12 hours, and remained up to 32 hours post irradiation. However, UV induced RPA hyperphosphorylation was not observed in AT5BIVA and AT3BISV cells up to 24 hours after UV irradiation with 30 J/m^2, but inducible expression of ATM restored the hyperphosphorylation of RPA. SDS-PAGE analysis shows RPA hyperphosphorylation in AT cells exposed to 10 J/m^2 UVC is less robust after 8 hours post exposure. Cell lysates from asynchronous AT3BISV cells show that the hyperphosphorylated RPA (form 5) exist both pre- and post-UVC irradiation, reaching maxima approximately 12 hours post irradiation. However, only very faint bands are observed and it appears that RPA form 5 is fainter after UV induction than before. In contrast, nuclear extracts from the same cell line are free of the hyperphosphorylated RPA at 8 hours, hence their claim that RPA hyperphosphorylation is not detected. It should be noted these gels are not shown at later time points as in complete cell lysates.

ATM deficient AT1ABR cells engineered into ATM inducible cells undergo hyperphosphorylation 8 hours post a 30J/m2 dose of UV irradiation. Accordingly, they failed to observe RPA32 hyperphosphorylation when ATM expression is reduced by at least 50% in cells

that normally express ATM. Moreover, normal lymphoblast cell lines transfected with an inducible antisense ATM cDNA expression vector, C3ABR cells, were defective in RPA hyperphosphorylation. Hypophosphorylated forms of RPA did not seem to depend on ATM kinase activity. All this indicated that ATM kinase targets RPA [82]. In addition to this, MH59J cells not only lack DNA PK but also have low ATM activity, which indicated that the low hyperphosphorylation activity observed could be a consequence of either DNA PK or ATM activity. It was also shown that hyperphosphorylation of RPA 32 is dependent on ongoing replication, since inhibition of DNA replication by addition of aphidicolin prevented induction of RPA hyperphosphorylation. In addition to this purified ATM was shown to phosphorylate RPA *in vitro* at many of the same sites of phosphorylation observed in vivo in response to IR. Two dimensional peptide maps showed that ATM phosphorylates RPA at multiple sites and although some appear to be the same as DNA PK, there are some differences [82]. ATM substrate specificity in *in-vitro* phosphorylation reaction was compared to that of DNA-PK using two dimensional peptide maps of hyperphosphorylated RPA. The differences observed between the ATM and DNA-PK RPA32 phosphorylation are obvious. However a pattern obtained from both kinases together combines the peptides observed from maps of each individual kinase suggesting that both are active in the hyperphosphorylation of RPA. Interestingly, the levels of DNA-PK in ATM mutated cell lines is comparable to that found in wild-type cells, however the contrary is not true. Endogenous levels of ATM are reduced in the M059J DNA-PKcs deficient cells but not in the M059K DNA-PKcs expressing cells. Thus, ATM could also contribute to the reduced RPA32 hyperphosphorylation observed in DNA-PKcs- M059J cells. However, whether ATM driven RPA hyperphosphorylation is a direct or indirect effect was not clarified.

The type of DNA damaging agent used to elicit RPA hyperphosphorylation can result in a different response. IR and Camptothesin (CPT) treatments result primarily in dsDNA breaks whereas UV causes mutation and ssDNA damage, therefore the kinases recruited in each event may be different. CPT induces replication dependent DNA lesions by trapping topoisomerase I cleavage complexes and resulting in the arrest of cells in the S and G2 phases of the cell cycle [84]. CPT treated AT-fibroblast and ATM normal fibroblasts have the same level of RPA hyperphosphorylation, suggesting that CPT induced RPA hyperphosphorylation is not ATM dependent [84]. Dominant negative ATM cells, expressing a non functional ATM protein, have a 5-fold increase in DNA-PK activity post CPT treatment and DNA-PK kinetics coincide with an increase in RPA32 phosphorylation suggesting that DNA-PK hyperphosphorylates RPA32 in response to CPT treatment. In addition, aphidicolin, a DNA polymerase inhibitor blocked CPT- and IR (18 Gy)-induced RPA32 hyperphosphorylation. Therefore, when replication is halted at the elongation step, replication dependent DNA lesions caused by CPT or IR will not result in RPA hyperphosphorylation. Thus, for the kinases to phosphorylate RPA in response to IR and CPT induced damage, replication must be ongoing. This is consistent with DNA-PK-mediated phosphorylation occurring at S phase of the cell cycle. A common denominator in replication defects is the generation of ssDNA that becomes immediately coated by RPA. When double strand breaks occur, in order to repair the damage, the DNA is resected and ssDNA is generated. This intermediate is targeted by other kinases such as ATR and ATRIP complex. This prompted studies of the possibility that these kinases are involved in RPA phosphorylation in response to DNA damage.

ATR-mediated phosphorylation of RPA

The yeast ATR homologue, MEC1, has been shown to be important for phosphorylation of RPA during normal cell cycle and also in response to UV and ionizing radiation (IR) [85]. In that study, yeast RPA phosphorylation was monitored in cells deficient for MEC1, Mec1-1 cells. While MEC1 expressing cells exhibited RPA phosphorylation 30 minutes after release from G1 phase, which corresponds to the entrance to S phase of cell cycle, RPA failed to become phosphorylated in Mec1-1 cells. Curiously, in spite of this difference both cell lines progressed through the various phases of cell cycle with identical kinetics as shown by flow cytometry. These results indicate that ATR may be involved in the initial steps of RPA phosphorylation during the S-phase of cell cycle progression. Moreover, in the presence of HU, IR, or 60 J/m^2 UV irradiation, the MEC1-1 cells exhibited a deficient level of RPA phosphorylation compared to MEC1 cells. The authors estimate that the mobility of this phosphorylated form of RPA in SDS-PAGE gels is different from other forms observed on normal cells. The low resolution of the gels makes it difficult to estimate if this band is a hypo or hyper phosphorylated form of RPA or a mix. In HU treated cells, TEL1 (the ATM homologue) was shown to only partially restore RPA phosphorylation in the absence of MEC1 while a MEC1-TEL1 double deletion resulted in no RPA phosphorylation. However, addition of MEC1 restored RPA phosphorylation in the double mutant cell line. This seems to indicate that ATR phosphorylation of RPA is important for RPA hyperphosphorylation in response to DNA damage and that this kinase activity is independent from the ATM kinase but may have a functional relationship. ATR kinase is necessary to achieve normal levels of phosphorylation regardless of ATM activity, while ATM activity does not seem to be necessary for a full ATR response after HU treatment. Interestingly, the same group reported later that phosphorylation of RPA 70 subunit is also observed in

response to HU but unlike RPA32 phosphorylation, RPA70 phosphorylation is not normally present during the cell cycle [86].

ATR participation in RPA32 hyperphosphorylation has also been studied in human cells [87]. Using an SV40 transformed fibroblast cell line (GM847) expressing a doxycyclin inducible kinase-inactive allele of ATR, it was demonstrated that, in response to UV or HU, RPA32 hyperphosphorylation was reduced upon over-expression of the kinase-inactive allele. In addition, SDS-PAGE analysis of RPA32 detected with phospho-specific antibodies against Ser4, Ser8, and Ser33 demonstrated that phosphorylation levels on these residues were reduced. Interestingly, Thr21 and Ser33 are ATR consensus sites but Ser4 and Ser8 are not. Although this demonstrates that absence of ATR affects phosphorylation of RPA, it is important to mention that over-expression of the kinase inactive allele of ATR also resulted in loss of Chk1 phosphorylation by ATR, which is necessary for Chk1 activation. Ser→Ala mutations on Ser 4, Ser8, Ser11, Ser12, and Ser13 do not abolish ATR phosphorylation of RPA32, but Ser→Ala mutations of Ser33 or Thr21 do. This suggests that ATR directly phosphorylates RPA32 Ser33 and Thr21. The contribution of other PI3K kinases was also investigated in this study by using retroviral infection of siRNAs in U2OS cells against DNA-PK and ATM in parallel with ATR and also Chk1. RPA32 hyperphosphorylation was monitored after UV and HU treatments demonstrating that depletion of DNA-PK, ATM or Chk1 does not affect the overall levels of UV- or HU-induced RPA32 hyperphosphorylation, but more specifically phosphorylation of Ser4, Ser8 and Ser33. Interestingly, RPA32 phosphorylation was reportedly increased when Chk1 expression was knocked down. The authors point out the consistency with reports that Chk1 inhibition results in an increase of ATR activity [87, 88]. A similar observation was made in Chk1 depleted human colon cancer HCT116 cells arrested in S-phase. Inhibition of

replication by addition of thymidine to these cells resulted in accumulation of RPA foci, primarily in the hyperphosphorylated form [89].

It has been demonstrated that RPA-coated ssDNA serves as an intermediate to stimulate ATRIP binding to ssDNA. RPA70 expression is necessary for ATR/ATRIP localization to sites of DNA damage induced by IR [90]. The same study also showed that in-vitro RPA-coated ssDNA is an important intermediate for the activation of ATR substrates such as Rad17, which is phosphorylated at Ser635, and Chk1 kinase, which is phosphorylated at Ser345. In this study, chromatin fractions of RPA70 siRNA HeLa cells exhibited a loss of phosphorylation of the Chk1 kinase post UV irradiation. Interestingly, a more recent study in SW80.SN3 cells (human colorectal carcinoma cell line) showed no effect in soluble fractions from RPA70 siRNA on Chk1 phosphorylation after DNA damage induced by HU [91]. It is possible that the type of DNA damage inflicted in these two events caused a different ATR activation response or that the response is localized to the nucleus and the effect was overlooked in soluble fractions. Thus, although ATR activation does not always occur via ssDNA-RPA recruitment of ATRIP, there is no evidence that suggests that the ATR-mediated phosphorylation of RPA32 post DNA damage is dependent on the activation pathway for ATR. ATR directly phosphorylates Ser33 and Thr21. It also has an indirect effect on phosphorylation of Ser 4, Ser8, Ser11, Ser12, and Ser13. Whether or not Chk1 is involved in the phosphorylation of these residues is not known. However, observations that RPA32 hyperphosphorylation is increased in the absence of Chk1 suggest that this is not the case.

In summary, RPA32 residues Ser4, Ser8, Ser11, Ser12, Ser13, Thr21, Ser23, Ser29, and Ser33 are targets for phosphorylation by cellular kinases. Cyclin dependent kinases phosphorylate Ser23 and Ser29 during cell cycle progression, with Ser29 is the preferred site for

cyclin B/cdc2. DNA-PK, ATM and ATR all have the capacity to phosphorylate Thr21 and Ser33 in response to DNA damage. Phosphorylation of Ser4 and Ser8 may also depend upon ATR and DNA-PK. *In-vitro*, DNA-PK is able to phosphorylate all of these residues. Cyclin A/cdk is a DNA-PK precursor for RPA32 phosphorylation. It seems more and more apparent that the type of genotoxic stress inflicted dictates which kinase will respond to phosphorylate RPA. The literature points to PI3-K kinases DNA-PK, ATM and ATR being involved in DNA damage hyperphosphorylation of RPA32. IR causes double strand breaks and the repair of this type of damage requires the resection of the strands to generate single stranded DNA that is then coated with RPA. ATM is a kinase responsive to dsDNA breaks and IR-induced hyperphosphorylation of RPA is dependent on ATM. However, upon resection of the strands, other kinases such as ATR may also be involved. In contrast, UV- and HU-induced hyperphosphorylation of RPA32 is largely ATR dependent. Both cause ssDNA formation, although we cannot disregard that DNA-PK and ATM may also be involved in RPA phosphorylation with these types of damage. Blocking DNA replication elongation by treatment with aphidicolin resulted in lack of DNA-PK-mediated RPA hyperphosphorylation and persistence of hypophosphorylated forms of RPA32. In addition, aphidicolin also blocks CPT and IR induced RPA hyperphosphorylation. Together, these data indicate that hyperphosphorylation occurs during the S phase of cell cycle while RPA is loaded on the ssDNA and replication is ongoing. Hypophosphorylated forms of RPA are a precursor for the hyperphosphorylated RPA generated in response to the type of damage inflicted. It is possible that the work of these kinases in phosphorylating RPA overlaps to ensure that a response mechanism is triggered in any event in which DNA damage is produced, initiating a cascade of events through RPA interactions leading to repair of the damage in actively replicating cells.

Experimental Methods

Small Angle X-Ray Scattering (SAXS)

SAXS is a powerful structural technique for the study of macromolecular systems, especially those which represent a challenge to more established techniques such as NMR and X-ray crystallography. Moreover, SAXS has proven extremely valuable when used in combination with these complementary structural techniques, allowing the description of structures and conformational changes associated with different functional states of macromolecules and molecular complexes [92, 93].

SAXS experiments can be performed on a broad range of samples, included proteins from ~10 to 300 kDa in size [94]. During the SAXS experiment, a 2-10 mg/ml protein solution (concentration depends on the protein size) is exposed to x-rays and the scattering intensity (I) is

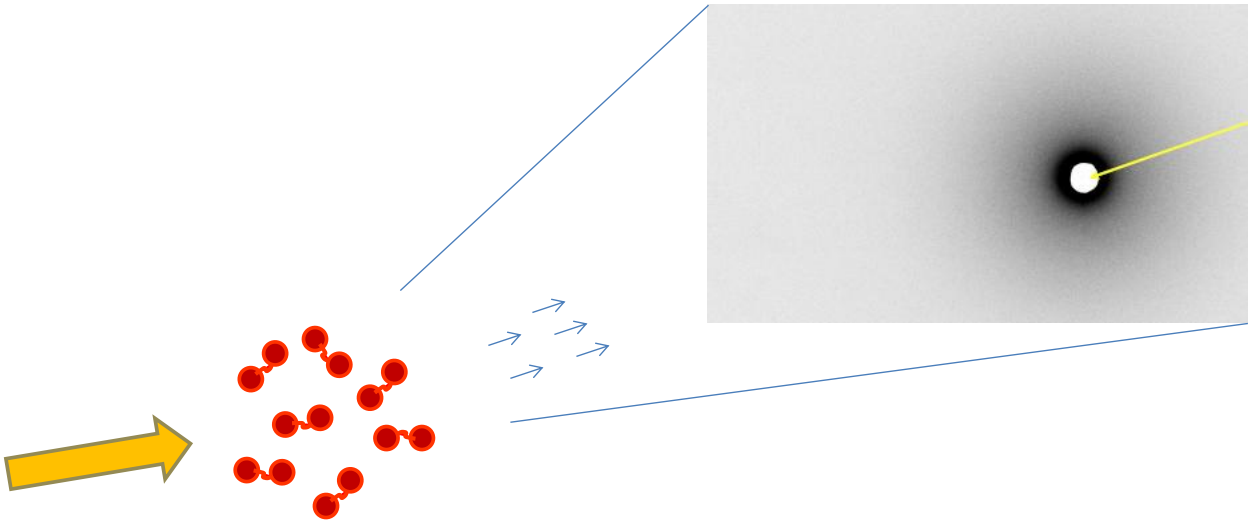


Figure 1.6. Representation of SAXS data collection. An intense beam scatters electrons from the protein solution and scattering intensities are recorded by an image detector. Raw scattering data for RPA70AB is shown. The yellow arrow indicates the beam stop. Scattering intensities at low angle are surrounding the beam stop. Scattering intensities at higher angles are more distant from the beam stop.

recorded as a function of the scattering angle (q) (Figure 1.6). Because of the random positions and orientations of each protein particle, the intensity is isotropic and proportional to the scattering from a single particle averaged over all orientations [94].

Once data is collected the contribution from the solvent is removed by subtracting scattering intensities from the solvent alone from scattering intensities of the sample, leaving only the scattering intensities of the protein. These are then spherically averaged to obtain the scattering profile of the protein particle. This procedure highlights the importance of performing these experiments with pure monodispese samples, as it would be impossible to separate the scattering from impurities, aggregates or polydisperse samples. Moreover, scattering intensities will be dominated by the solvent, thus appropriate concentrations are necessary to obtain good intensity signals for the solute.

A log plot of the protein scattering intensities, I(q), as a function of the scattering angle, q ($Å^{-1}$), gives us the SAXS profile for the protein in reciprocal space (Figure 1.7a). Scattering data at higher angles (larger q values) correspond to scattering that is far from the beam stop and contains detailed structural information about the protein shape [92]. Scattering data at the lowest angles (small q values) corresponds to data closest to the beam stop, and is used to check against aggregation that may arise during data collection. Guinier analysis for globular proteins (Figure 1.7b), or Debye Formula for elongated proteins, is used to check for aggregation. Guinier analysis also allows estimating the radius of gyration (Rg) of the protein directly from the scattering data. Rg is the radial space occupied by the protein based on its center of mass. Other types of analysis are also possible that estimate the Rg and hydrated volume of globular proteins, such as the Porod invariant (Figure 1.7c). Because globular proteins have uniform

electron densities, their molecular weight is directly related to their volume [92]. Porod volume in nm^3 corresponds to approximately twice the molecular weight of the protein in KDa.

Kratky analysis, a plot of $q^2*I(q)$ as a function of q, is another useful tool derived directly from the scattering data (Figure 1.7d). For well folded samples the plot has a parabolic appearance, $q^2*I(q)$ values converge back to zero at large q. For unfolded samples it reaches a plateau with high $q^2*I(q)$ values at large q and partially unfolded proteins fall in between [92]. Kratky profiles for flexible proteins, such are RPA70NAB and RPA70AB, show an inherent degree of randomness with $q^2*I(q)$ higher than zero, but far from high $q^2*I(q)$ values [95]. Therefore, this type of analysis can be used assess the degree of flexibility, the presence of disorder, and conformational change [93].

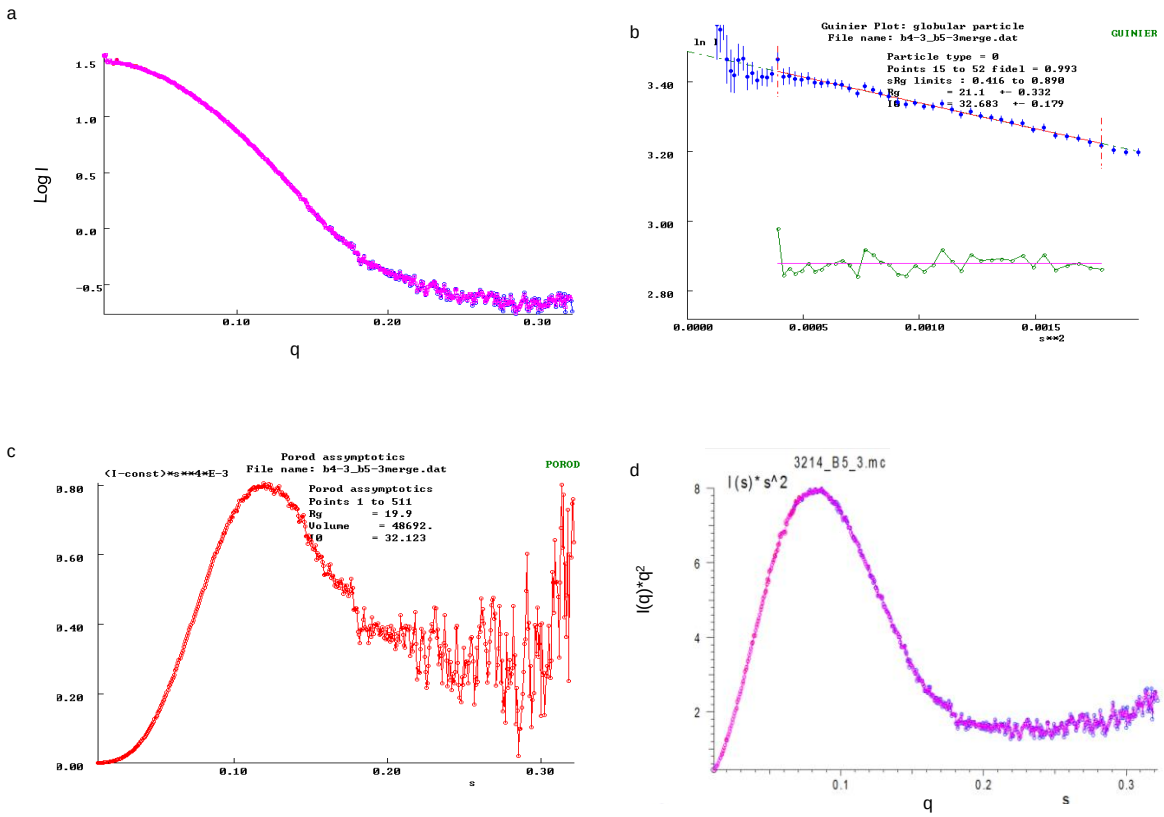


Figure 1.7. Analysis of SAXS data: (a) Scattering plot; (b) Guinier analysis; (c) Porod analysis; (d) Kratky analysis.

Further transformation of the scattering data produces the P(r)-distribution function, a plot that describes the distribution of inter-atomic distances, P(r), in real space, r (Å). P(r) is sensitive to the shape of the molecule under study (Figure 1.8). The maximum dimension (*dmax*) of the particle is one of the key characteristics that can be obtained from the P(r) function, and like Rg, is a useful description of the molecule under study. Low resolution three-dimensional *ab initio* models can also be generated either directly from the scattering data or the P(r) function. To construct the three dimensional model, GASBOR uses the scattering data to fit a chain of dummy residues, much like a chain of amino acids, within a spatial arrangement that satisfies the SAXS data [94]. Similarly, DAMMIN uses an ensemble of densely packed beads and simulated annealing to find a configuration that fits the SAXS data [96]. Due to the inherent imprecision, 10 GASBOR or DAMMIN models are typically averaged (i.e. using DAMAVER) (Figure 1.9). The fit of these conformational envelopes onto known crystal or NMR structures can be used to assess consistency. In our research, we found that these calculations cannot be applied to highly flexible systems, such as RPA70AB (unbound) and RPA70NAB, because intrinsic interdomain flexibility is not accounted for in these models. Rigid body molecular dynamics is a common approach used to generate conformational ensembles for proteins with substantial inter-domain flexibility. BILBO-MD is one such

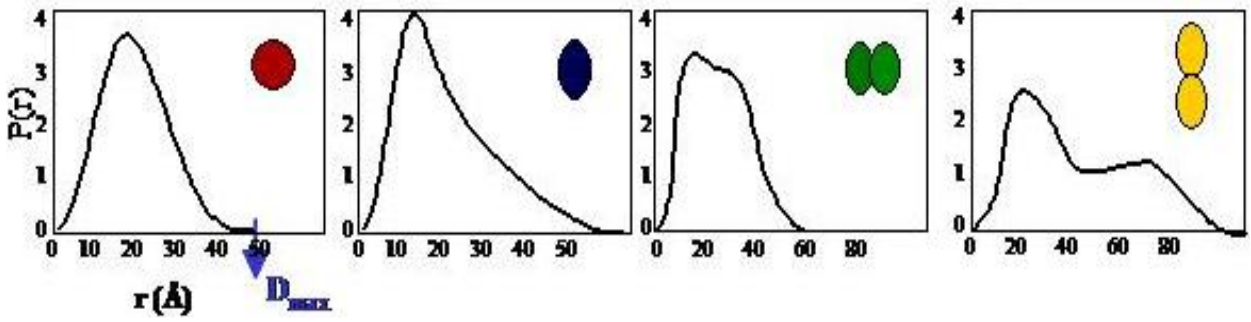


Figure 1.8. P(r) is sensitive to the shape of the molecule under study

program, and was used in our studies. It uses CRYSOL to backcalculate a scattering profile for each conformer. The fit between the calculated and experimental scattering curves are then used to evaluate how well each conformer matches the experimental data, as reflected in a χ^2 fitting parameter. The Minimal Ensemble Search (MES) program can be applied to assess whether multiple conformers provide a better fit to the data than single structures. However, as will be discussed in Chapter II, it is not possible to distinguish the relative merit of one conformational subset over another [95, 97].

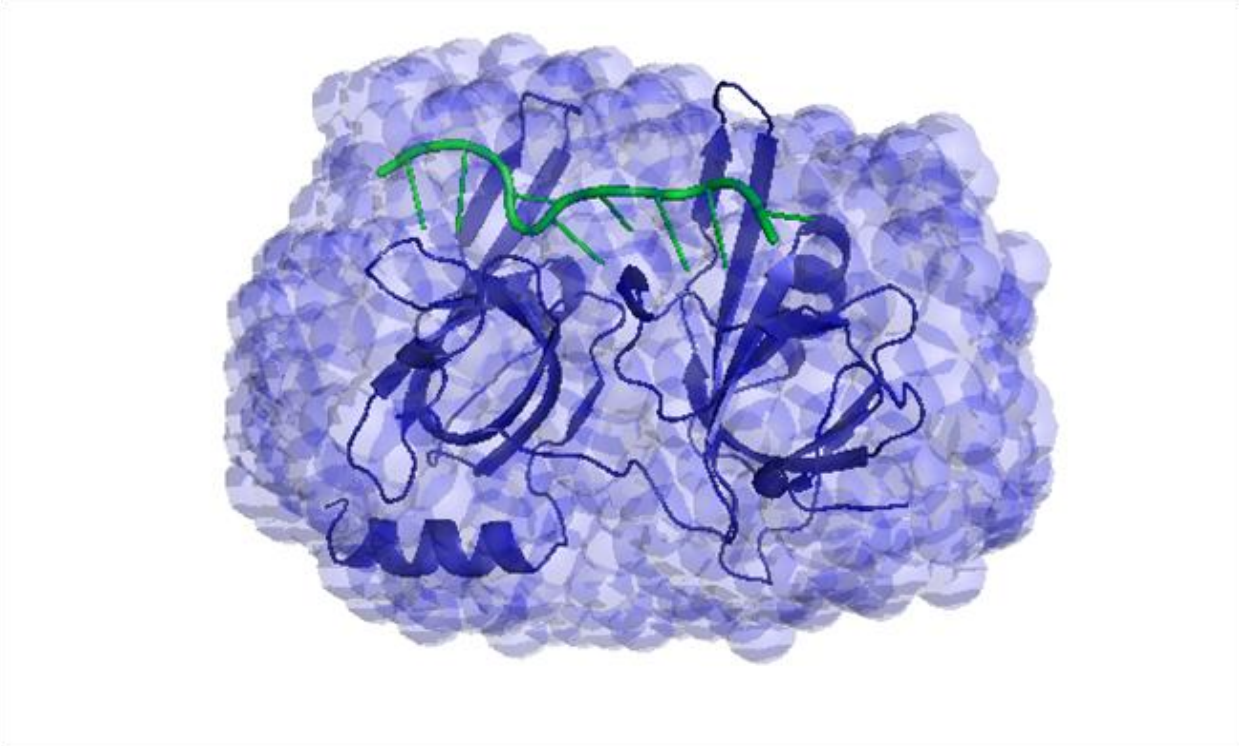


Figure 1.9. SAXS envelope of RPA70AB bound to ssDNA.

Nuclear Magnetic Resonance (NMR) spectroscopy

Nuclear Magnetic Resonance (NMR) spectroscopy is a powerful means to characterize the structure and dynamics of molecules [98-103]. One particularly useful application of NMR is for the study of weakly interacting systems, including protein-protein interactions. Chemical

shift perturbation is the most widely used method to map protein interfaces through the identification of the residues involved in an interaction. In this experiment, a 2D ^{15}N-^{1}H HSQC (Heteronuclear Single Quantum Coherence) spectrum of an ^{15}N-enriched protein (Figure 1.10) is monitored when the unlabeled interacting partner is titrated in. Signals correlating the ^{1}H chemical shift to the ^{15}N chemical shift appear in the 2D spectrum for every backbone and side chain amide in the protein. The perturbations of the chemical shift resulting from changes in the environment surrounding a residue, as a consequence of the interaction with the titrant, are recorded for every titration point. Perturbations of the signal arise in the spectrum in the form of broadening of a peak, disappearance of a peak, or shift in peak position. The kinetics of the interaction determines the change in chemical shifts during the titration. Peaks are broadened as a result of intermediate exchange, whereas in slow exchange one set of peaks disappears from the spectrum as a second set appears at a different position. Fast exchange, which is often observed for weaker binding interactions, results in constant shifting of peak positions. This refers to a fast dissociation of the complex and thus a single set of resonances is observed at all times.

In all NMR spectroscopy experiments, it is important that the protein samples are of high purity and high concentration starting from at least 100 uM to 500 uM. Since ~0.5 mL of solution is required, there is a need for large amounts of protein, which could present a problem for some proteins. The pH of the protein solution must be maintained during the experiment to maintain a stable chemical environment and avoid chemical shift changes for reasons that are not a by-product of intermolecular interaction. Using the same buffer for both proteins eliminates this problem. This method is suitable for a range of proteins. Although it can be manipulated

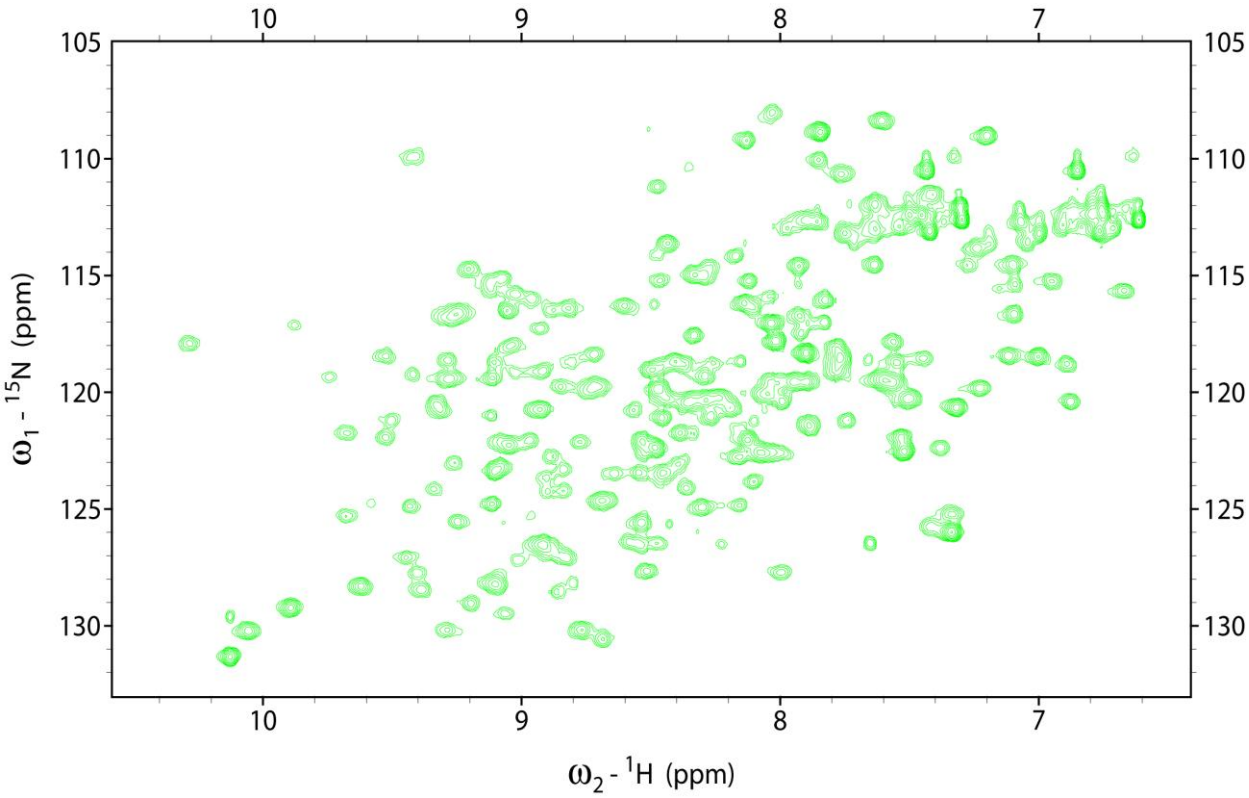


Figure 1.10. 2D ^{15}N-^{1}H HSQC spectrum of ^{15}N-enriched RPA70AB.

to identify resonances of proteins larger than 30 KDa size, its sensitivity is most optimal for proteins of lower molecular weight. Throughout my research, chemical shift perturbation experiments have been used to investigate RPA32N interaction with other RPA domains.

In addition to mapping the interface, the titrations allow the estimation of affinities, stoichiometry and specificity of binding. The dissociation constant can be calculated by following the resonances to the bound position and fitting the fractional shift against a quadratic equation depending on the total protein and partner concentrations. The equation is derived from the formula $K_D = ([P_{tot}] - [PA])([A_{tot}] - [PA]) / [PA]$, where P is the protein, A is the interacting partner and PA is the bound complex.

Resonance assignments are an essential first step for any detailed study by NMR. Assignments are required in order to map chemical shift perturbations on the structure of the protein under study. NMR chemical shift assignments can be obtained using a variety of

strategies and these have been amply reviewed [98]. In my studies, a series of 3D ^{15}N-^{13}C-^{1}H-triple resonance NMR experiments including HNCA, HNCO, HN(CO)CA, HNCACB, and CBCA(CO)NH were collected to complete backbone assignments for RPA32N. Once the resonances have been assigned, the chemical shift perturbations can be related to the protein structure and the binding interface delineated. This in turn allows a highly efficient, direct approach to designing mutations to assess the functional significance of the interaction under investigation.

Research Overview

This dissertation focuses first on the influence of RPA70N on the structural dynamics of the high affinity ssDNA binding domains RPA70AB and effect of hyperphosphorylated RPA32N on the overall architecture of RPA. These studies are aimed at understanding how RPA domains participate in RPA functional activities. In Chapter II we address an ongoing controversy about the role of RPA70N in the ssDNA binding activity of RPA. I have approached this question using SAXS and molecular dynamics to demonstrate the disposition of RPA70N within the context of RPA70NAB in the absence and presence of ssDNA. In RPA70NAB, the linker between RPA70N and RPA70A is 60 residues, and the linker between RPA70A and RPA70B is 10 residues. Protein purification and stable buffer conditions were optimized for the RPA70AB control and RPA70NAB. SEC-MALS was used to assess protein monodispersity, an essential prelude to SAXS analysis. The ssDNA binding properties of these two RPA constructs were analyzed. All SAXS data were collected at the SYBILS beam line located at Berkley National Laboratory in collaboration with Dr. John Tainer's research group, specifically with Dr. Susan Tsutakawa. An RPA70NAB model was constructed based upon

crystal structures of RPA70AB + ssDNA and RPA70N. Molecular dynamics simulations using SAXS data to guide the conformational search demonstrated the inherent flexibility of the linkers and dynamic properties of RPA70NAB. These studies showed RPA70N does not to interact with ssDNA even in the presence of high local concentrations of ssDNA accessible for binding. This work also illustrated the difficulties of SAXS analysis in flexible systems.

Studies described in Chapter III identify RPA32N inter-domain interactions in the hyperphosphorylated state of RPA. This work is aimed at understanding what intersubunit interactions are relevant for the analysis of RPA's actions in response to DNA damage. An RPA32N phosphomimetic mutant peptide (RPA32N-D8) was constructed and characterized. Recombinant pseudo-phosphorylated and wild type RPA32N peptides were produced and potential intersubunit interactions were systematically tested using NMR chemical shift perturbation assays. RPA32N-D8 was found to interact with RPA70N, RPA70A and RPA70B domains, but not with RPA32C, RPA32D, and RPA14 domains. An interaction with RPA70C could not be confirmed. Residues involved in the intersubunit interactions were compared against residues involved in ssDNA and RPA-protein interactions. I also extended the analysis to preliminary experiments on a full-length RPA phosphomimic protein. A full-length RPA hyperphosphorylation mutant was constructed in collaboration with SBDR at Berkley National Lab. Preliminary SAXS data was collected for the hyperphosphorylated mutant. These data suggested a model in which hyperphosphorylation of RPA causes a compaction of the protein.

Finally, the data presented in this thesis are discussed in Chapter IV. Overall, my research advances our understanding of how the properties of RPA domains influence RPA function, and provide a foundation for analysis of dynamic RPA actions in DNA repair.

CHAPTER II

STRUCTURAL DYNAMICS AND SINGLE-STRANDED DNA BINDING ACTIVITY OF THE THREE N-TERMINAL DOMAINS OF THE LARGE SUBUNIT OF REPLICATION PROTEIN A FROM SMALL ANGLE X-RAY SCATTERING*

Introduction

RPA is the primary eukaryotic ssDNA binding protein utilized for diverse DNA transactions in the replication and maintenance of the genome [reviewed by Fanning and co-workers [104]]. RPA functions by binding and protecting ssDNA from degradation by endonucleases, inhibiting formation of ssDNA secondary structure, and providing a scaffold for DNA processing machinery by interacting with numerous DNA processing proteins. RPA biochemical functions and biological activities have been intensively investigated and the structures of its domains determined [11-14, 26, 42, 45, 105]. Despite this detailed information, the mechanisms for RPA function remain poorly understood, largely because of the inherent difficulties of characterizing proteins with modular organization and the fact that RPA function is integrated within complex multiprotein machinery.

RPA is a modular 116 kDa heterotrimer composed of seven structured globular domains and one disordered domain (Figure 2.1). The trimer subunits are named on the basis of their approximate molecular masses: RPA70, RPA32, and RPA14. The RPA70 subunit contains four oligonucleotide-oligosaccharide binding (OB-fold) domains: RPA70N, RPA70A, RPA70B, and RPA70C. RPA70N is linked to RPA70A by a 70 residue linker, which in turn is connected to

*This chapter has been published: Pretto D.I., Tsutakawa S., Brosey C.A., Castillo A., Chagot M.E., Smith J.A., Tainer J.A., Chazin W.J. *Structural dynamics and ssDNA binding activity of the three N-terminal domains of the large subunit of Replication Protein A from small angle X-ray scattering.* Biochemistry. 2010 Apr 6;49(13):2880-9

RPA70B by a 10-residue linker. RPA70B is connected to RPA70C by a 15-residue linker. The RPA32 subunit contains a 45-residue unstructured N-terminal domain (RPA32N), along with a central OB-fold domain (RPA32D) and a C-terminal winged helix domain (RPA32C), which are separated by a 23-residue linker. The RPA14 subunit consists of a single OB-fold domain. High-resolution structures have been determined by X-ray crystallography or NMR spectroscopy for all of the globular domains [11-14, 26, 42, 45, 105]. However, knowledge of the spatial organization of the domains in the intact protein is lacking [2]. Such information is important because remodeling of RPA "architecture" constitutes an essential element of its function in DNA processing machinery.

RPA ssDNA binding activity is associated with the A−D domains of RPA70 and RPA32. The binding of ssDNA by RPA occurs with a 5′ → 3′ molecular polarity in the order A, B, C, and D [6, 106]. The directional binding to ssDNA is the result of differences in the binding affinity of the four ssDNA binding domains. The RPA70A and RPA70B domains have the highest ssDNA affinity and serve as the anchor for all ssDNA binding activities. Two X-ray

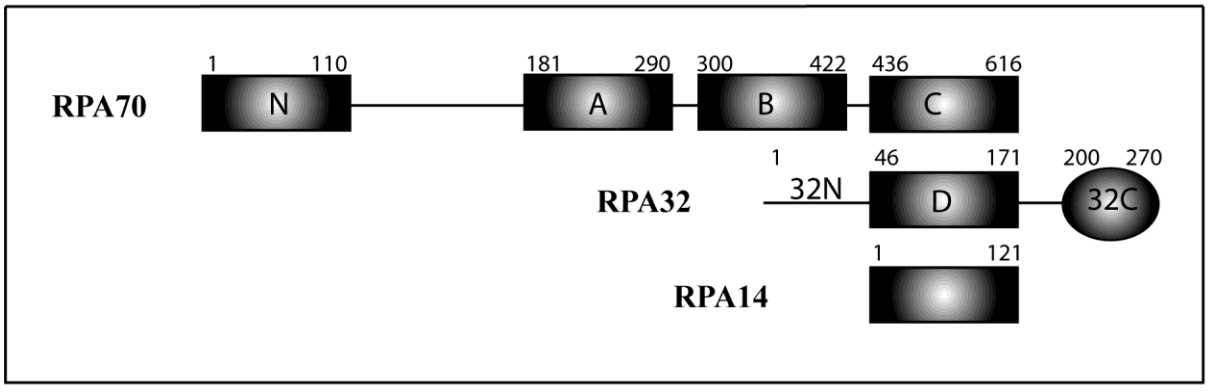


Figure 2.1. Domain organization of RPA. Rectangles are OB-fold domains, and the ellipsoid is a winged helix domain. RPA32 has one unstructured N-terminal domain (RPA32N). Trimerization occurs via non-covalent interactions between RPA70C, RPA32D and RPA14 domains. ssDNA binds to RPA70A, RPA70B, RPA70C and RPA32D.

crystal structures of RPA70AB have been determined, one without ssDNA [11] and another in the presence of d-CCCCCCCC [45]. In the complex with ssDNA, RPA70A and RPA70B are aligned with the two binding loops wrapped around and nearly encircling the DNA. In contrast, two very different orientations of RPA70A with respect to RPA70B were observed in the structure of the free protein. Moreover, NMR analysis indicated the two domains of RPA70AB are structurally independent and implied that the two domains are attached by a flexible tether [7]. However, no direct information about interdomain orientations in solution and the effect of ssDNA on RPA70AB and the relationship with the rest of the protein has been obtained.

RPA70N is suggested to have weak ssDNA binding activity that is important for the DNA unwinding activity of RPA [3, 44]. However, the ssDNA binding affinity of this domain is more than 1000-fold weaker than that of RPA70AB, and RPA70N is generally accepted to be a protein interaction module targeting DNA replication, damage response and repair proteins such as p53 and ATRIP [32, 39, 41]. The only RPA70N structure is of the isolated domain [32, 42], so there is no available information about its disposition with respect to the rest of RPA and hence its availability to influence the DNA binding properties of the protein. NMR spectroscopy of a construct containing RPA70N, RPA70A, and a portion of RPA70B suggested that RPA70N does not interact with the RPA70A domain [43]. An NMR study of full-length RPA and larger multidomain constructs suggests the motion of RPA70N is independent of the remaining DNA-binding domains, both in the absence and in the presence of ssDNA [4].

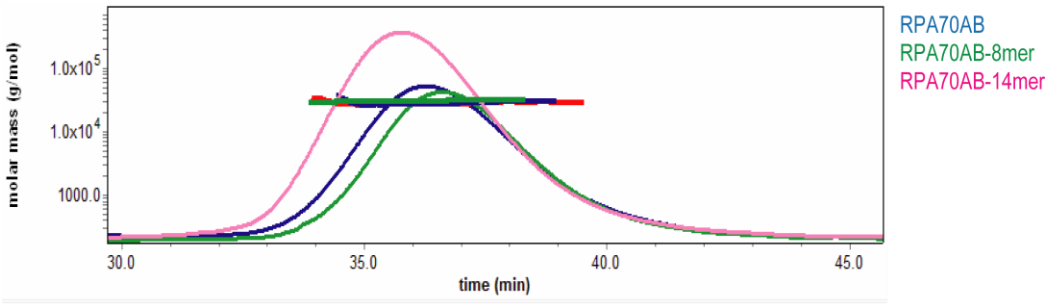
To expand these initial observations and obtain direct information about interdomain orientation, we have turned to small angle X-ray scattering (SAXS), which is a powerful approach to studying proteins under native solution conditions and extracting low-resolution

spatial information for dynamic systems such as RPA. To experimentally address the structural dynamics of the RPA70A, RPA70B, and RPA70N domains, the impact of ssDNA binding on interdomain flexibility, and the effect of RPA70N on the ssDNA binding activity of the tandem high-affinity RPA70AB domains, we purified RPA70AB and RPA70NAB and examined them with and without ssDNA in solution by small angle X-ray scattering (SAXS). Analysis of SAXS data for these dynamic systems was facilitated by the generation of large ensembles of structures with different interdomain orientations using rigid body molecular dynamics simulations. The results show RPA70N is flexibly linked to RPA70AB and has no influence on the binding of ssDNA and have general implications for RPA dynamic architecture and functions.

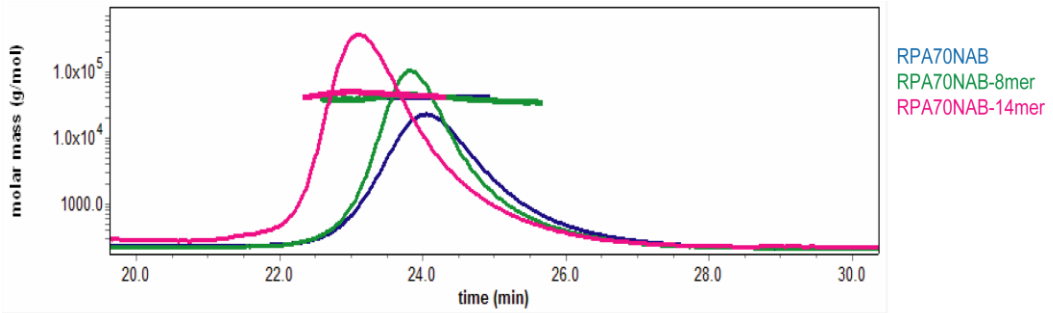
Results

To characterize the effects on the structural dynamics of RPA70A and RPA70B as they bind ssDNA and investigate the influence of RPA70N, small angle X-ray scattering (SAXS) experiments were performed on RPA70AB and RPA70NAB in the absence and presence of ssDNA. SAXS measures the electron pair distribution and is well-suited for characterization of the architecture of molecules in solution [107]. Since scattering data are distorted by scattering from small amounts of aggregation, the monodispersity of the samples was carefully monitored by multi-angle light scattering (MALS) of the peaks eluted during size exclusion chromatography (SEC) (Figures 2.2 and 2.3) and by verification of the data in the Guinier analysis. To ensure monodispersity and remove any free DNA for protein−DNA complexes, each sample was treated by SEC just prior to the collection of data.

A.



B.



C.

Molecular Weight (KDa)	AB	AB-8mer	AB-14mer	NAB	NAB-8mer	NAB-14mer
Expected	27.2	29.6	31.4	46.1	48.5	50.3
Calculated	27.9	27.5	30.0	40.9	41.3	45.6

Figure 2.2. TSK-gel G200sw elution profiles of (a) RPA70AB SEC-MALS experiments prior to SAXS data collection and (b) RPA70NAB SEC-MALS experiments prior to SAXS data collection, and (c) the corresponding calculated and experimentally determined molecular weights.

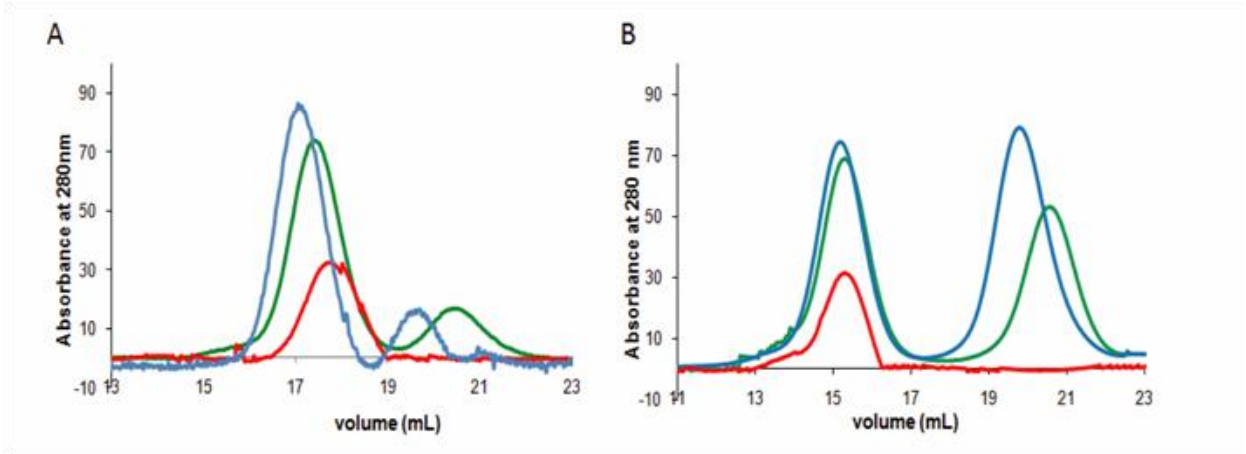


Figure 2.3. Superdex 200 SEC elution profiles for (a) RPA70AB and (b) RPA70NAB prior to SAXS data collection. All protein samples are monomers and the protein:ssDNA complexes are 1:1 ratios. Red= NAB or AB alone; Green = NAB or AB with 8-mer; Blue = NAB or AB with 14-mer. Excess ssDNA peaks are observed centered around 20 mL.

Structural Dynamics of RPA70AB from Analysis of SAXS

The RPA70AB scattering profile (Figure 2.4a) reveals the high quality of the SAXS data obtained after optimization of the sample and acquisition parameters. A Kratky analysis of the data is consistent with the two globular well-folded domains connected by a 10-residue linker (Figure2.4b). The use of the Kratky plot of SAXS data to detect flexibility in proteins is well-established (see Figure 24 of ref [92]). A Guinier analysis provided a radius of gyration (R_g) of RPA70AB in solution of 25.6 Å. The data were also analyzed with GNOM to derive the probability distribution function $P(r)$, which reflects the distribution of interatomic distances in the molecule (Figure 2.4c). The main peak in the $P(r)$ function corresponds to scattering between atoms within the globular RPA70A and RPA70B domains, which are similar in shape and size. The shoulder at longer distances corresponds to scattering between atoms in one domain and atoms in the other. The maximum distance (D_{max}) in the $P(r)$ function is 100 Å, and the R_g value

derived from $P(r)$ is 25.7 Å, consistent with the reciprocal space R_g derived directly from the scattering data (Table 2.1).

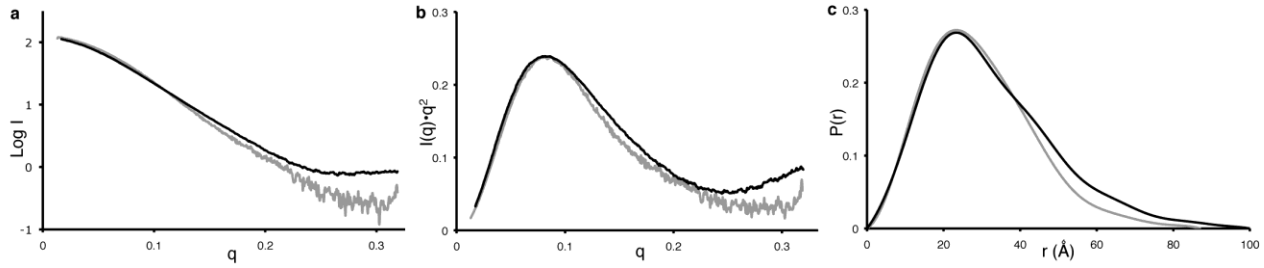


Figure 2.4. Scattering curves (a), kratky analysis (b) and P(r) functions (c) for RPA 70AB in absence (black) and presence (gray) of d-(CCACCCCC).

Proteins	**AB**	**AB-8mer**	**AB-14mer**	**NAB**	**NAB-8mer**	**NAB-14mer**
Rg (Å), Guinier Analysis	25.6 ± 0.2	23.4 ± 0.3	24.4 ± 0.3	39.5 ± 0.4	37.4 ± 0.3	37.9 ± 0.2
Rg (Å), P(r) analysis	25.7 ± 0.1	23.2 ± 0.1	24.7 ± 0.1	43.7 ± 0.2	42.1 ± 0.2	41.6 ± 0.2
Dmax (Å), P(r) analysis	100	87	100	165	165	165

Table 2.1. SAXS measurements.

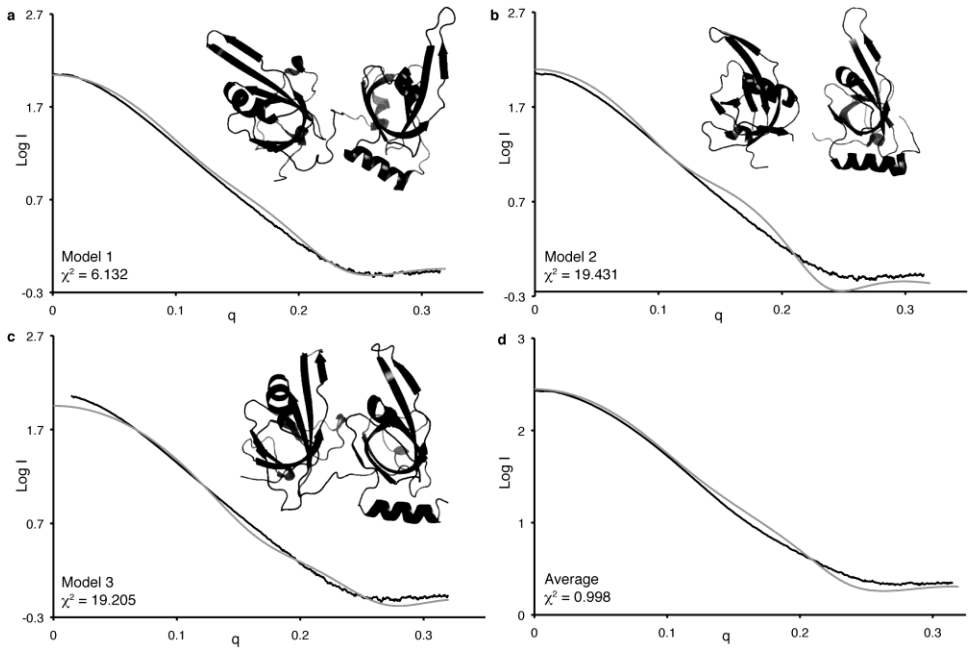


Figure 2.5. CRYSOL fit to experimental data for each of the crystal structure models and their average. Comparison of RPA70AB experimental SAXS scattering curves (black) and back-calculated scattering curves (gray) for models 1 (a), 2 (b), 3 (c), and the average (d). Models 1 and 2 correspond to the two molecules in the unit cell of RPA70AB (1FGU.pdb). Model 3 is the protein molecule after ssDNA atoms were extracted from RPA70AB/d-C8 (1JMC.pdb).

To examine the implications of the SAXS data for the solution structure, we back-calculated scattering curves from three crystallographic models with CRYSOL and superimposed them on the experimental data (Figure 2.5 a-c). Models 1 and 2 correspond to the two different molecules in the asymmetric unit of the RPA70AB X-ray crystal structure, which have very different orientations of the two domains. Model 3 was obtained from the X-ray crystal structure of RPA70AB bound to d-CCCCCCCC, from which the ssDNA coordinates were removed. The χ^2parameter reflecting the fit of the calculated curve to the experimental data is also included for each model. Although χ^2 is dependent in part on the experimental noise, a lower χ^2 value generally corresponds to a better fit of the model(s) to the experimental data. This analysis shows that although model 1 provides a better fit than the two other models, none of the

models fit especially well to the data. Similarly, each of the crystallographic models poorly fit the ab initio envelope predicted from the scattering data, with model 1 being most similar. These results suggest two possibilities. (i) The structure is not accurately represented by any of these three specific models, or (ii) the two domains occupy multiple interdomain orientations.

To distinguish these two possibilities, we employed rigid body molecular dynamics simulations with BILBO-MD [97] to generate 6400 conformers with a wide range of interdomain orientations. For these calculations, the RPA70A and RPA70B domains were treated as rigid bodies and the linker between them was allowed to fully sample conformational space (Figure 2.6a). The radius of gyration was allowed to vary between 20 and 40 Å. The shape of the distribution of data in this plot is a reflection of the RPA70AB structure. Despite there being small gaps in the plot due to incomplete sampling of the conformational space accessible to RPA70AB, the sampling is sufficient to draw conclusions about the fit to the experimental data. The observation of a wide χ^2 minimum around 4 implies the data cannot be properly represented by a single structure.

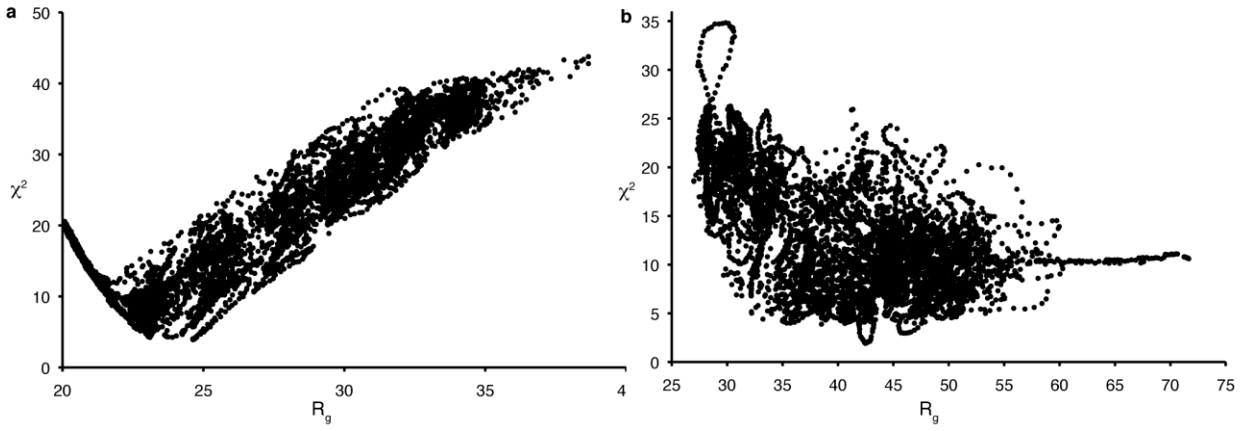


Figure 2.6. Plot of χ2 fit parameter versus radius of gyration for RPA70AB (a) and RPA70NAB (b) conformers generated by bilbomd. The calculated scattering curve for each conformer was generated by CRYSOL.

To analyze the ensemble of structures, we employed a genetic algorithm to determine if select groups of conformers can fit the data better than single conformers [97]. The best fits obtained taking two and three conformers provide χ^2 values of 2.1 and 1.6, respectively, indicating that multiple conformers represent the data far better than any single conformer. Figure 2.5d shows the improved fit to the experimental RPA70AB scattering obtained for the combination of three RPA70AB conformers comprised of models 1, 2, and 3 described above. Overall, analyses of the SAXS data show directly that the two domains in RPA70AB are not fixed in space but rather occupy a range of interdomain orientations.

ssDNA Binding to RPA70AB

To determine the effect of binding ssDNA on the structural dynamics of RPA70AB, the scattering measurements were repeated in the presence of d-CCACCCCC. Previous studies have established that RPA70AB binds oligomers of 8−10 nucleotides with an affinity in the high nanomolar range [7]. The tight binding affinity for ssDNA implies that a monodisperse protein−DNA complex can be prepared and characterized, and the monodispersity was confirmed by SEC-MALS (Figures 2.2 and 2.3).

When the scattering data analyzed in the same manner as for free RPA70AB are compared to those of the free protein, it is evident that binding of ssDNA significantly reduces the protein's structural dynamics (Figure 2.4a). The R_g value derived directly from the Guinier analysis was 23.4 Å, 2.2 Å shorter than the value determined for the free protein (Table 2.1). The range of interatomic distances reflected in the $P(r)$ function is significantly decreased relative to that of the free protein (Figure 2.4c). In addition, the D_{max} value is 87 Å, a reduction of ~13 Å relative to that determined for the free protein. Notably, the D_{max} is significantly larger than

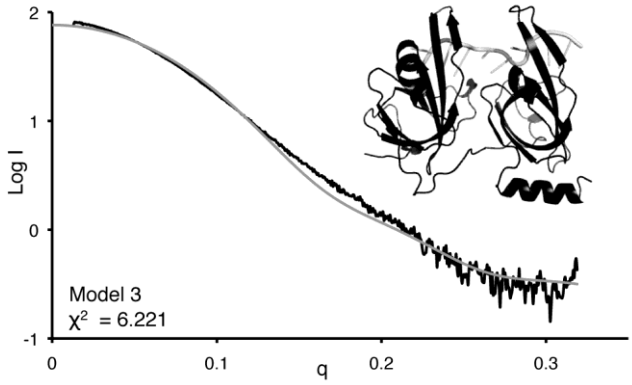


Figure 2.7. CRYSOL fit to experimental RPA70AB-8mer data (black) for the crystal structure model of RPA70AB bound to dC8 (gray). Comparison of experimental scattering curve for the RPA70AB-8mer complex (black), versus the back-calculated scattering curve for the crystal structure (1JMC).

the D_{max} of 67.4 Å measured directly from the crystal structure (PDB entry 1JMC), even when the missing residues at the N- and C-termini are taken into account. The lower value of R_g and the narrowing of the curve and lower D_{max} in the $P(r)$ function all indicate there is an overall compaction of RPA70AB upon binding ssDNA [5]. The reduction in R_g in particular directly reflects the fact that the two domains are on average closer to each other when ssDNA is bound. This interpretation is consistent with the ssDNA serving to further tether the two domains together [7]. The SAXS data show directly that interdomain dynamics is quenched relative to free RPA70AB.

We next asked if the dynamic architecture revealed by the solution scattering data was accurately represented by the X-ray crystal structure of RPA70AB in complex with d-CCCCCCCC. Gasbor calculations were first performed to determine the conformational envelope generated from the data. While the fits to the crystal structure were reasonable in this case, back-calculation provides a more direct and quantifiable assessment of the fits to atomic-resolution models. We therefore turned to back-calculating the scattering curve from the

coordinates of the crystal structure using CRYSOL, and the results were plotted and compared to the experimental data (Figure 2.7).

Overall, there is agreement between the experimental scatter and the crystal structure, which is consistent with the value of 20.0 Å for R_g calculated from the crystal structure. However, the χ^2 fitting parameter is 6.22, which indicates inconsistencies between the X-ray crystal structure and the SAXS data. This result suggests either the complex has a different structure in solution or, as for free RPA70AB, the complex cannot be adequately represented by a single structure. The latter explanation is supported by the observation in the Kratky analysis that the curve does not completely return to the baseline at higher scattering angles, which indicates that some disorder or flexibility is still present even when DNA is bound (Figure 2.4b). The interdomain flexibility probably arises from the flexible linker between domains and from torsional interdomain motions around the bound ssDNA and suggests the RPA70AB complex retains conformational flexibility that is not evident in the crystal structure. However, because of the low resolution of SAXS data, we cannot completely rule out the possibility that the complex can be represented by a single structure that is different from the crystal structure.

Small Angle X-ray Scattering of RPA70NAB

To examine the structural dynamics of RPA70AB and investigate the influence of RPA70N, SAXS data were acquired for the RPA70NAB construct (Figure 2.8). An R_g value of 39.5 Å was derived directly from the data based on the Guinier analysis, which is >14 Å greater than the R_g for RPA70AB. The $P(r)$ function shows the same primary peak centered at ~24 Å as observed for RPA70AB, reflecting scattering within the globular OB-fold domains (Figure 2.8c).

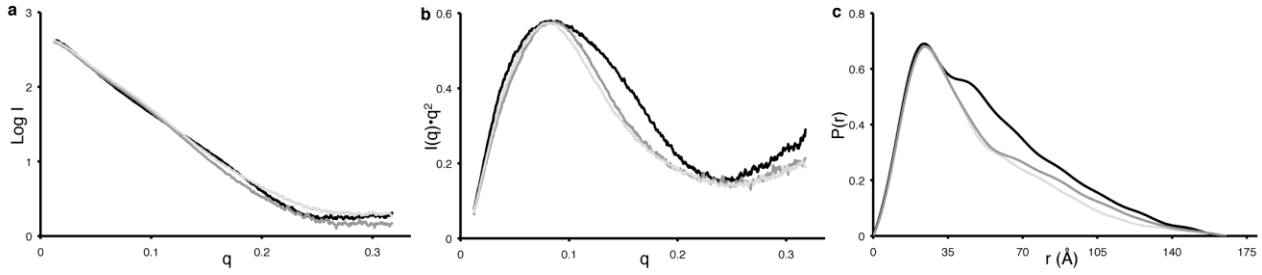


Figure 2.8. Comparison of Scattering curves (a), kratky analysis (b), and P(r) function (c) for RPA70NAB in absence (black) and presence (dark and light gray, respectively) of d-(CCACCCCC) and d-(AAAAAACCACCCCC).

The scattering curves differ substantially at longer distances. For example, D_{max} extends out to 165 Å. Calculation of *ab initio* molecular envelopes directly from the scattering data using GASBOR did not converge. This observation suggests that envelope representations are problematic for highly flexible systems in which domain orientations may vary considerably between conformers. Together, the results from the SAXS analysis indicate that the conformational space sampled by RPA70NAB is substantially larger than that observed for RPA70AB.

To test the implications of the SAXS data with respect to the structural dynamics of RPA70NAB, BILBO-MD was used to generate 6000 conformers of RPA70NAB. For these calculations, each globular domain was treated as an independent rigid body, but the N−A and A−B linkers were allowed to sample conformational space freely (Figure 2.6 b). Notably, χ^2 does not reach a specific minimum but rather reaches a plateau through a wide range of R_g values, even more so than what was observed for RPA70AB. This observation of a broad plateau for the χ^2 minimum indicates the data are not properly represented by a single structure and suggests that RPA70NAB has substantial interdomain flexibility.

To obtain further insight into the solution structure of RPA70NAB, we examined representative conformers from the ensemble generated by BILBO-MD, including fully extended and closely packed arrangements of the three domains. CRYSOL was used to back-calculate scattering curves for each conformer selected (Figure 2.9). The calculated R_g values are ~45 Å for the extended conformation and ~29 Å for the closely packed conformation. The R_g value for both of these conformers is far from the experimentally observed value, and the scattering curves do not match the experimental data. Notably, poor fits to the scattering curve were also obtained even for specific conformers that closely match the experimentally observed R_g value. Thus, the SAXS data indicate that RPA70NAB has extensive interdomain flexibility, which is substantially larger than RPA70AB as a result of the long flexible linker between the N and A domains.

Effect of ssDNA Binding on the Structural Dynamics of RPA70NAB

To investigate the effects of binding ssDNA on the structural dynamics of RPA70NAB, we performed experiments with an 8mer and a 14mer ssDNA oligonucleotide. The 8mer corresponds to the excluded site size of RPA70AB and is designed as a control to characterize the effect of ssDNA binding to these high-affinity domains. The 14mer was designed to determine if RPA70N is able to modulate the ssDNA binding activity by providing six extra nucleotides to the 5′ side of the high-affinity RPA70AB domains. The SAXS experiment is ideally suited to detect ssDNA binding by RPA70N in the context of RPA70NAB because any appreciable interaction would produce a pronounced compaction of the molecule as a result of the alignment of RPA70N with RPA70AB. In particular, since RPA70AB is already strongly bound to eight nucleotides of ssDNA, the binding of RPA70N to the remaining overhang would result in a substantial reduction in the R_g of RPA70NAB to a value in the vicinity of 30 Å.

Two important considerations guided the design of the 14mer oligonucleotide used for these experiments. (i) The number of nucleotides had to allow binding of an additional OB-fold domain without enabling binding of a second molecule, and (ii) the position of RPA70AB on the ssDNA needed to be biased to the 3′ end of the oligo to maximize the availability of free ssDNA for binding to RPA70N. The X-ray crystal structure of RPA70AB bound to d-CCCCCCCC shows each OB-fold domain makes contact with three nucleotides, and two nucleotides bridge the RPA70A and RPA70B domains [45]. Thus, the length of the oligonucleotide needed to be fewer than 16 nucleotides to preclude binding of two protein molecules on the DNA. A ssDNA 14mer was therefore selected because it is too short for binding two molecules yet provides six extra nucleotides for RPA70N to bind. The d-CCACCCCC sequence was used at the 3′ end of the 14mer oligonucleotide on the basis of previous analysis of the sequence preferences of RPA70AB (E. Bochkareva, A. I. Arunkumar, W. J. Chazin, and A. Bochkarev, unpublished results). These studies showed that RPA70AB binds more strongly to cytosine rich sequences than adenine rich sequences and that placement of a single adenine at position 3 in d-CCCCCCCC further enhances binding affinity. To bias RPA70AB to the 3′ end of the oligonucleotide, the 14mer was constructed by adding six adenines to the 5′ side of the high-affinity sequence, resulting in d-AAAAAACCACCCCC. Although RPA binds polyadenine more weakly than polypyrimidines, the affinity is nonetheless in the nanomolar range [8]. Moreover, six free adenines is adequate for binding because the tethering to RPA70AB means RPA70N is present in a higher local concentration relative to that for free diffusion. All studies with 8mer oligonucleotides used the sequence d-CCACCCCC to ensure the accuracy of the comparisons between binding of RPA70NAB to 8mer and 14mer ssDNA.

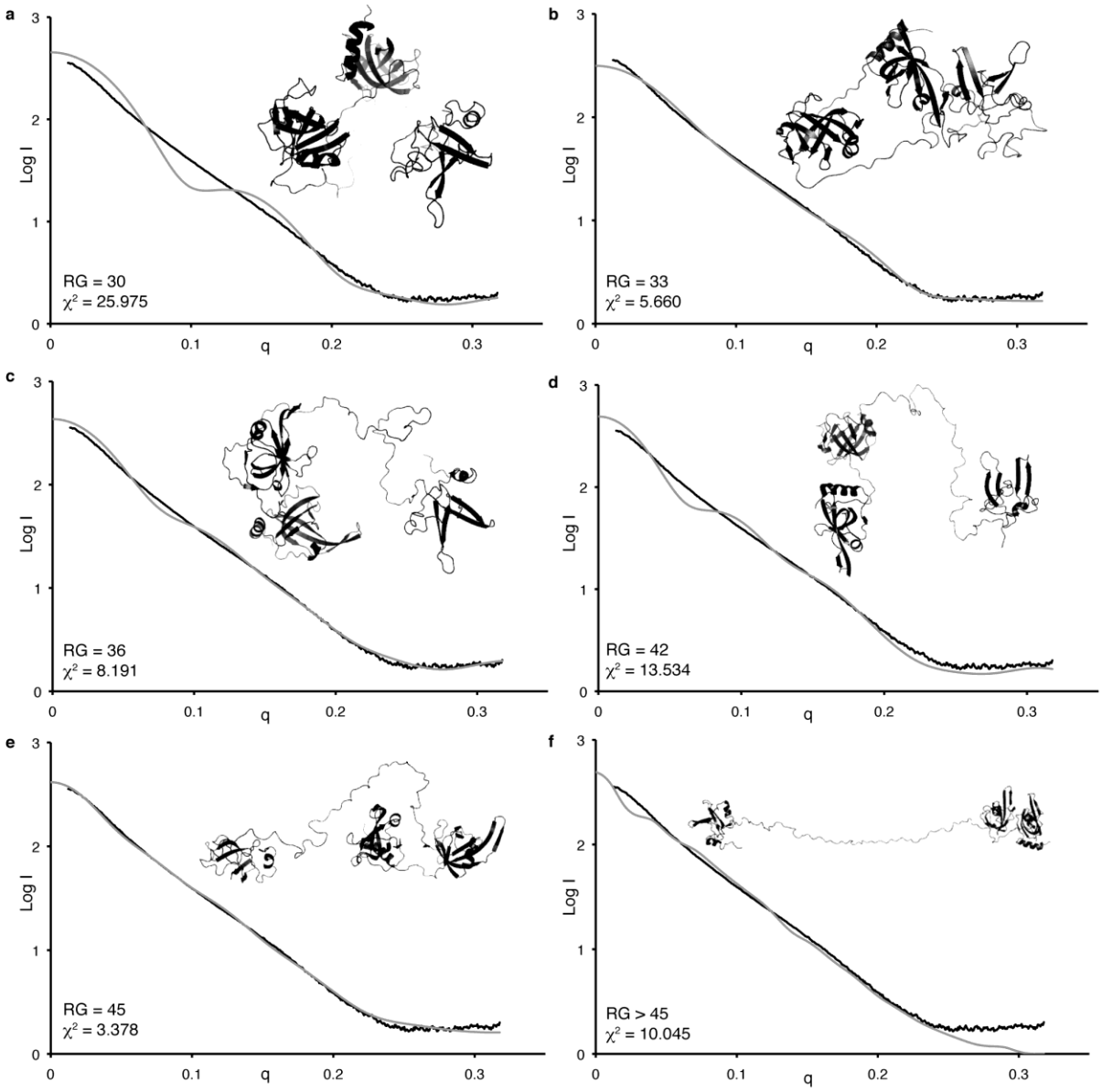


Figure 2.9. Comparison of experimental scattering curves for RPA70NAB (black) against back-calculated scattering curves of selected BILBOMD generated models (gray). Cartoon representations are shown for each of the models, along with the corresponding χ2 fit parameter and Rg values.

Scattering measurements for RPA70NAB bound to d-CCACCCCC (8mer) or d-AAAAAACCACCCCC (14mer) were taken and analyzed following the strategy for the free

protein and the RPA70AB−ssDNA complex. Comparison of the data with free RPA70NAB reveals that binding of ssDNA alters the structural dynamics of the protein, although in a relative sense, the effect on the scattering curve is not as great as for RPA70AB (Figure 2.8a). A Kratky analysis indicated little change in the relative amount of unordered polypeptide in RPA70NAB when either ssDNA oligomer was bound (Figure 2.8b). The R_g values from the Guinier analysis for the 8mer and 14mer are 37.4 and 37.8 Å, respectively, reflecting a reduction of 2.1 and 1.7 Å, respectively, relative to that of the free protein. As was observed for RPA70AB, the peak range of interatomic distances reflected in the $P(r)$ function decreases significantly (Figure 2.8c). These observations indicate reductions in R_g and D_{max} upon binding of ssDNA closely parallel what is observed for RPA70AB. Thus, the data indicate a compaction of the RPA70AB domains as they bind the ssDNA within the complex, which is correlated with quenching of the interdomain dynamics of these domains. However, there is no indication of further compaction of the protein, which would result from the RPA70N domain also engaging the ssDNA. On the contrary, the SAXS data show that RPA70N remains as flexible in the DNA-bound state as in the free protein, confirming the hypothesis proposed on the basis of indirect NMR evidence in our study of intact RPA [4].

To further analyze the data, large ensembles of conformations for both RPA70NAB−ssDNA complexes were generated using BILBO-MD, and then the scattering was back-calculated using CRYSOL and compared to the experimental data. The initial model was built using the coordinates from the X-ray crystal structures of RPA70N and RPA70AB bound to d-CCCCCCCC, along with a purely modeled N−A linker. For these calculations, the entire RPA70AB module was treated as a single rigid body, RPA70N was treated as a second rigid body, and the N−A linker was allowed to sample conformational space freely. For a select

number of conformers, 14mer ssDNA was added back and CRYSOL was used to back-calculate scattering profiles. An examination of a range of different conformers, including highly extended and closely packed arrangements of the three domains, shows that, in fact, scattering is dominated by the relative position of the three globular domains (Figure 2.10).

The key to our interpretation is the fact that all of the observed R_g values for RPA70NAB experiments (Table 2.1) are consistent only with RPA70N populating interdomain orientations

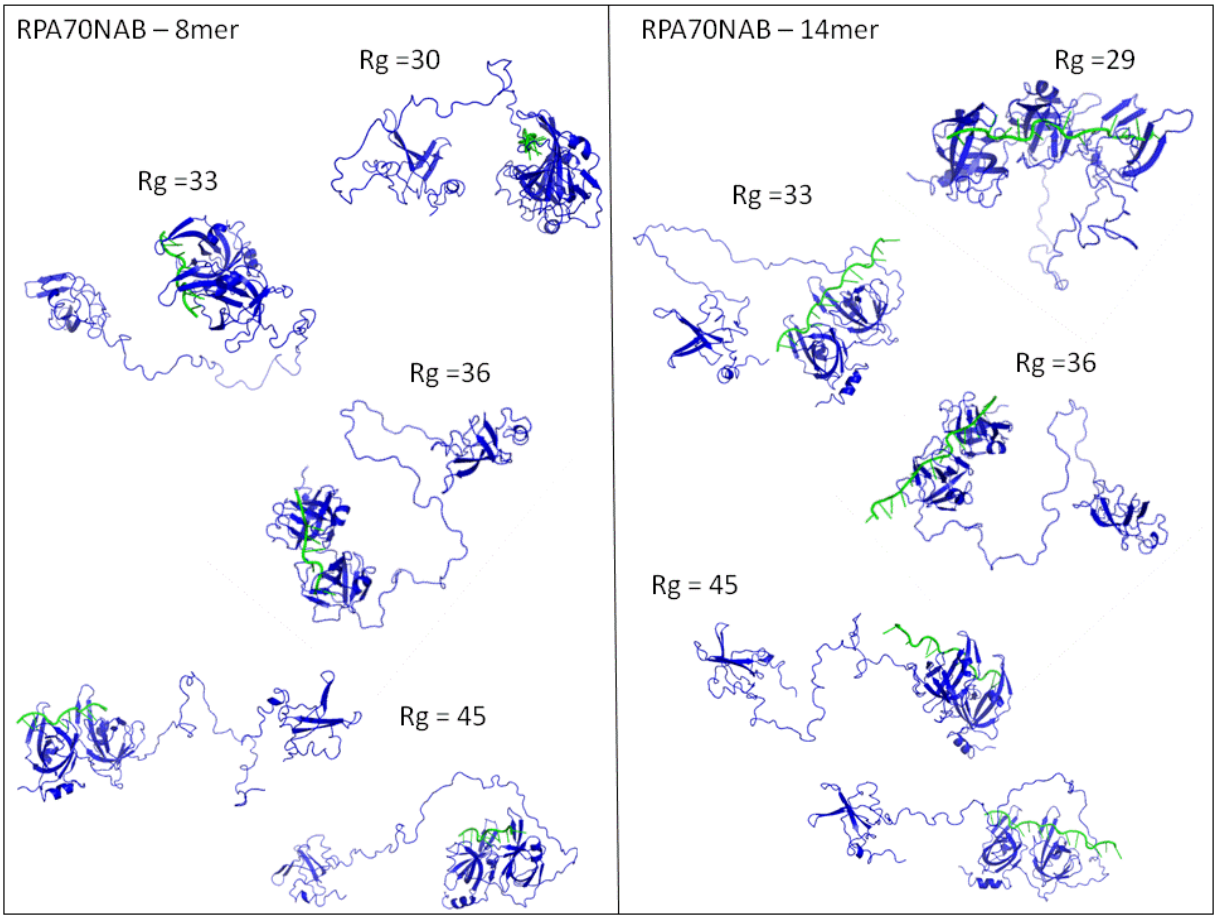


Figure 2.10. Models of RPA70NAB in the presence of the ssDNA 14-mer. The complex with with Rg = 29 was generated by manually positioning the RPA70N domain in a position that mimics ssDNA binding. All other conformers were generated using BILBO-MD as indicated in the Methods section.

where it is distant from RPA70AB. In particular, in experiments with the 14mer, if RPA70N were interacting with the ssDNA it would be closely packed to RPA70AB and the overall shape of the RPA70NAB molecule would be substantially more compact and globular (Figure 2.9). This would result in R_g values significantly lower than those observed in the control experiments with the free protein and the 8mer ssDNA. The models of the complex of RPA70NAB with the 14mer in which RPA70N is packed near the DNA have R_g values on the order of 29 Å, which is far from the experimentally determined value of 41 Å. Thus, our analysis shows that RPA70N does not become ordered even when ssDNA is bound to RPA70NAB.

Discussion

RPA ssDNA binding occurs with 5′ to 3′ directionality, initiated by high-affinity ssDNA binding domain RPA70A [6, 106, 108]. The existence of the short 10-residue A−B linker increases the effective concentration of RPA70B, which promotes its binding to ssDNA [7]. Binding of DNA to RPA70AB is followed in turn by rearrangement of RPA70C and RPA32D. The trajectory of binding is therefore in opposition to the orientation of RPA70N toward the 5′ end of ssDNA sequences.

It has been proposed that RPA70N plays a direct role in binding ssDNA [3, 44]. The ssDNA binding affinity for the isolated RPA70N domain is extremely weak, which on the basis of the evidence in the literature [32, 43, 44] has a lower limit for the dissociation constant (K_d) in the millimolar range. Our study was designed to determine the effect of RPA70N in the context of ssDNA binding to the adjacent RPA70 A and B domains, which more directly addresses the hypothesis put forth in the literature. The effective local concentration of ssDNA in the vicinity of RPA70N is maximized in the experiment with the ssDNA 14mer, therefore providing every

opportunity for the RPA70N domain to engage the ssDNA. If RPA70N had any role to play at all in binding ssDNA, this would have been reflected in a change in the distribution of conformational states occupied by RPA70NAB relative to the distributions observed in the control experiments on free RPA70NAB and the complex with the ssDNA 8mer. The fact that there is no indication of interaction of RPA70N with ssDNA in the context of RPA70NAB is convincing evidence against the proposal that RPA70N plays a direct role in the ssDNA binding activity of RPA.

The 70-residue linker between the N and A domains suggests there would be little correlated movement of the two domains, as suggested by NMR relaxation analysis of a construct containing RPA70N, RPA70A, and a portion of RPA70B [43]. However, it is difficult to draw firm conclusions from that study because the RPA70A domain in the absence of the RPA70B domain has only very weak affinity for ssDNA. Our studies show directly, and in the physiologically relevant context of high-affinity binding of ssDNA, that the dynamic and flexible N−A linker enables a wide range of RPA70N orientations relative to RPA70AB. Thus, the long linker provides the large degree of freedom to RPA70N that is critical to this domain's participation in the recruitment of partner proteins. RPA70N binds multiple proteins involved in DNA replication, damage response, and repair mechanisms, including p53, ATRIP, MRE11, and NBS1[32, 39-41]. Hence, our results support models of RPA function in which RPA70N acts as a general protein recruitment module.

SAXS is an emerging technique in structural biology that measures electron scattering intensities to yield interatomic distances for molecules in solution [92, 107]. For globular proteins, which have fixed rather than variable conformations, it is possible to convert this distance distribution into molecular envelopes that reflect the average shape of a protein or

protein complex in solution [92]. Such coarse-grained structural envelopes are invaluable complements to atomic-resolution information. For systems with high degrees of interdomain flexibility, such as RPA70NAB, a single, "averaged" conformation fails to provide an adequate description of an intrinsically time-varying architecture. The lack of statistically significant correlation reflected by high χ^2 values between experimental scattering profiles from RPA70NAB and those calculated for individual models readily illustrates this point and defines this as a flexible region. Proper interpretation of scattering data from these flexible systems requires an ensemble approach, both to describing the population of feasible conformations and to characterizing their relative frequency within the ensemble at a given moment in time. BILBO-MD [97] allowed us to search a broad range of accessible RPA70NAB conformational space and examine in detail a subset of models (including some that represent the extreme conformations) and their capacity to recapitulate the scattering data. Notably, while averaging theoretical scattering profiles from multiple conformations improves the goodness of fit to the experimental data, the challenge of distinguishing the relative merit of one conformational subset over another remains.

Methods and programs are available to select combinations of conformers that provide better fits to the data, including the minimum ensemble (MES) approach that is part of BILBO-MD [97]. However, while it is possible to define combinations of structures that give improved fits to the data, these combinations are not necessarily unique but represent a minimal identified ensemble that fits the data. In the case of highly flexible proteins, there are many combinations of conformers that fit the data equally well. Consequently, the main conclusion that can be drawn from this type of analysis of a complex highly flexible system such as RPA70NAB is that the protein contains substantial degrees of conformational heterogeneity, a point that is best made

directly from data. Nevertheless, ensemble fitting does provide valuable insight into the dynamics of the native, solution architectures of proteins and macromolecular complexes.

The key role of dynamics in facilitating the organization and progression of large multiprotein machines is increasingly recognized particularly for DNA replication and repair machinery that requires precise coordination to efficiently preserve genome integrity. Our results suggest SAXS offers a robust approach to characterizing protein structural dynamics in solution without the complications of isotopic enrichment or spin labeling required by spectroscopic methods. RPA70N is found to be structurally independent of RPA70AB in the DNA-bound state and therefore able to act flexibly as a protein recruitment module. Notably, this flexible attachment of the RPA DNA and protein binding domains, elucidated by SAXS, enables the interactions of RPA with diverse DNA substrates and protein partners required for effective orchestration of DNA replication and repair. Similar flexible attachments joining protein and DNA binding domains were recently discovered for DNA-PK and polynucleotide kinase [109, 110]. In fact, such RPA domain structural flexibility as experimentally defined here is essential to allow efficient protein handoffs and interface exchanges, as proposed for FEN1-PCNA [111] and BRCA2-Rad51[112]. RPA binding protein partners, such as the Mre11−Rad50−Nbs1 complex, have similar ordered and flexible domains, as shown by Nbs1 SAXS and crystal structures [113]. Such dynamic character may be a hallmark for scaffold proteins such as RPA.

The role of dynamics in facilitating recruitment, organization, and exchange of DNA processing factors has been characterized in several model systems, most notably in a recent study of homotetrameric *E. coli* SSB diffusion dynamics along ssDNA [114]. In that study, SSB diffusion was shown to be critical for resolving DNA secondary structures to enable RecA

filament formation. Unlike the modular, multidomain RPA, the homotetrameric single-domain SSB does not utilize preexisting structural dynamics to facilitate organizing strands of ssDNA for DNA processing. Instead, the compact, globular SSB homotetramer is encircled by the ssDNA and is thought to "roll" along the template via a consecutive unwrapping and wrapping of ssDNA. Thus, while the structural organization of these two SSB systems remains fundamentally different, dynamic motion would appear to be integral aspects of both. Specifically, the nature of the structural dynamics of linked, ordered, and flexible RPA domains as identified here appears to be critical to the accommodation of the large-scale complex conformational changes proposed to regulate RPA-related functions, while preserving the integrity of DNA and protein partner interactions for maintaining genetic fidelity.

Experimental Procedures

Expression of RPAs

Fragments of human RPA were expressed from a pSV281 RPA70AB plasmid containing a TEV cleavable six-His tag at the N-terminus and a pBG100 RPA70NAB plasmid containing an H3C six-His tag also at the N-terminus. TEV and H3C proteases are produced in-house. ssDNA oligomers d(CCACCCCC) and d(AAAAAACCACCCCC) purchased from Integrated DNA Technologies were desalted, lyophilized, and resuspended into autoclaved distilled water.

Recombinant RPA70AB ($RPA70_{181–422}$) and RPA70NAB ($RPA70_{1–422}$) constructs were prepared as described previously [4, 7]. Proteins were expressed in *Escherichia coli* host Rosetta (DE3) cells (Novagen, Madison, WI). Cells were grown in LB medium containing kanamycin at 37 °C, induced with 0.1 M IPTG when the OD reached 0.6, and harvested after 3 h using a JLA 8.1 Beckman rotor at 7500 rpm and 4 °C. Pellets were stored at −20 °C.

Protein Purification

RPA70AB samples were purified using nickel-nitrilotriacetic acid resin (NiNTA) by nickel affinity chromatography in 10 mM Hepes (pH 7.5), 500 mM NaCl, 5 mM BME, and 10% glycerol using an elution gradient from 20 to 300 mM imidazole. Cleavage of the His tag with TEV protease was performed through overnight dialysis in a buffer containing 10 mM Hepes (pH 7.5), 200 mM NaCl, 5 mM BME, 200 mM L-arginine, and 10% glycerol. A second NiNTA purification step was used to remove the His tag. Size exclusion chromatography (SEC) using a Superdex S75 column equilibrated with the dialysis buffer was used as a last step of purification. Protein was concentrated, and stock solutions were frozen in a dry ice/ethanol bath and kept at −80 °C. The same procedure was used for RPA70NAB samples using NiNTA buffer [30 mM MES (pH 6.5), 500 mM NaCl, 10 mM BME, and 5 mM $MgCl_2$] and cleaving with H3C protease. The SEC step was performed using a Superdex S200 column and a buffer containing 30 mM MES (pH 6.5), 200 mM NaCl, 10 mM BME, 10% glycerol, and 5 mM $MgCl_2$.

Preparation of Protein−DNA Complexes

RPA70AB or RPA70NAB was incubated in the presence of a 1.2−1.5-fold molar excess of d(CCACCCCC) or d(AAAAAACCACCCCC) for 20 min on ice; 500 μL of sample was purified by SEC using S75 (RPA70AB) or S200 (RPA70NAB) resin. The samples eluted as one peak for the complex followed by a DNA-only peak.

Size Exclusion Chromatography via Multi-Angle Light Scattering

The monodispersity of each sample was verified by multi-angle light scattering connected in line with SEC (SEC-MALS). All experiments were performed using a Wyatt

Technology instrument, and data were analyzed using ASTRA version 16.25. Samples were analyzed using a 2.4 mL Superdex75 column. Only samples that exhibited monodispersity were selected for data collection.

Small Angle X-ray Scattering

SAXS data of the various RPA constructs were collected at the SIBYLS 12.3.1 beam line at the Advanced Light Source, Lawrence Berkeley National Laboratory. Scattering measurements were performed on 20 μL samples at 15 °C using a Hamilton robot for loading samples from a 96-well plate into a helium-purged sample chamber. Protein−DNA samples were further purified on a 24 mL SEC column just prior to data collection to eliminate any free protein or DNA. Data were collected on both the original gel filtration fractions and samples concentrated ~2−8-fold from individual fractions. Fractions prior to the void volume and concentrator eluates were used for buffer subtraction.

The experiments with RPA70AB used an X-ray beam from a single-crystal monochomator of 11 keV, covering the following momentum transfer range: 0.007 $Å^{-1} < q <$ 0.35 $Å^{-1}$ ($q = 4\pi \sin \Theta/\lambda$, where 2Θ is the scattering angle). Sequential exposures (6, 6, 60, 6, 200, and 6 s) were taken, and data were monitored for radiation-dependent aggregation. SAXS experiments for all RPA70NAB samples and for RPA70AB with DNA were acquired using an X-ray beam from a multilayer monochromator of 12 keV covering the following momentum transfer range: 0.012 $Å^{-1} < q <$ 0.317 $Å^{-1}$. The multilayer provides increased X-ray flux allowing stronger signals for lower protein concentrations. Sequential exposures (0.5, 0.5, 5, and 0.5 s) were taken, and data were monitored for radiation-dependent aggregation. All SAXS data were collected using the MarCCD 165 detector in fast frame transfer mode and reduced via

normalization to the incident beam intensity. Buffer scattering was subtracted from protein scattering. This was followed by azimuthal averaging to obtain the intensity $I(q)$ versus q scattering plot visualized by xmgrace. The data were analyzed using PRIMUS (Primary Analysis & Manipulations with Small Angle Scattering Data) version 3.0 from ATSAS 2.0 [115], from which Guinier, Kratky, $P(r)$, and CRYSOL plots were generated.

For each sample, multiple experiments were conducted over multiple runs. Experiments providing the highest signal-to-noise ratio that remained consistent with the relative concentrations from original gel filtration fractions were selected for further analysis. For RPA70AB alone, the final data used for analysis were merged between a 60 s exposure at 147 μM and a 200 s exposure at 331 μM. The concentrations of the samples used for analysis for RPA70AB with 8mer DNA, RPA70AB with 14mer DNA, RPA70NAB alone, RPA70NAB with 8mer DNA, and RPA70NAB with 14mer DNA were 71, 81, 163, 80, and 98 μM, respectively, each with a 5 s exposure time.

Computational Modeling

Ab initio shape envelopes were calculated with GASBOR [94]. Ten GASBOR runs were merged using the DAMAVER suite. Protein Data Bank (PDB) coordinates were overlaid using SUPCOMB. PDB coordinates from the RPA70N NMR structure (entry 1EWI) and the RPA70AB ssDNA-bound X-ray crystal structure (entry 1JMC) were used to construct a model for RPA70NAB. From these coordinates, multiple conformers were generated by rigid body molecular dynamics simulations with BILBO-MD [97]. For the DNA complexes, ssDNA coordinates were removed from the RPA70AB model. For RPA70NAB, the connecting linker between domains 70N and 70A was built using the Biopolymer module of Insight II (Accelrys, Inc., San Diego, CA), followed by refinement with Rosetta [116]. To generate

RPA70NAB−ssDNA complexes for back-calculation of scattering profiles, the DNA coordinates were added back using Molecular Operating Environment, MOE 2010.09 (Chemical Computing Group, Montreal, QC) and Chimera [117]. Molecular graphics were generated using PyMol (DeLano Scientific, Palo Alto, CA).

CHAPTER III

TOWARD STRUCTURAL CHARACTERIZATION OF THE HYPERPHOSPHORYLATED FORM OF RPA

Introduction

Studies have shown that RPA phosphorylation results in decreased binding of RPA to DNA pol-α, DNA-PK and ATM, suggesting that these interactions are modulated by RPA's phosphorylation state [60]. It can be hypothesized that this corresponds to a type of negative control mechanism to regulate RPA's functions in which the kinases that phosphorylate RPA have a higher affinity to unphosphorylated substrate than to the more electronegative product, consistent with studies showing weaker affinity of RPA kinases for the phosphorylated RPA [60]. Similarly, it has been suggested that the effect of phosphorylation on the interaction with DNA pol-α may be to aide in the release of DNA pol-α from the replication initiation site (see review [104]).

Scanning transmission electron microscopy (STEM) studies aimed at understanding RPA-DNA binding and its role in heterologous interactions suggest that RPA can occupy an elongated conformation when is bound to a 30 nt ssDNA, and a closely packed conformation when bound to an 8 nt ssDNA [118]. These structural states were detected in the presence of long ssDNA depending on the concentration of RPA used and the use of cross-linking reagent. RPA phosphorylation was found to be inversely proportional to the number of RPA molecules bound to the ssDNA [118]. Notably, RPA was bound to DNA prior to *in-vitro* phosphorylation by DNA-PK. This means that when RPA was bound to 30 nt it became more easily

phosphorylated than when it was bound to 8 nt of ssDNA. These results can be interpreted in light of what could be happening to RPA32N, the flexible region that contains the phosphorylation sites in RPA32. In ^{15}N-^{1}H HSQC NMR experiments on the full-length RPA heterotrimer performed by our group, chemical shifts that belong to the RPA32N glycine residues are perturbed and have better correspondence to the chemical shifts of free RPA32N(1-46) peptide when dT30 is added [4]. This is consistent with the flexible RPA32N becoming more free or mobile in the DNA bound form of RPA. It has been suggested that DNA binding is a pre-requisite to RPA phosphorylation [53]. Together these data suggests that RPA32N may be more accessible to the various RPA kinases when RPA is bound to ssDNA in its most elongated form.

The STEM studies do not have sufficient resolution to indicate where RPA32N and other specific domains are located and how these might differ in the non-phosphorylated and phosphorylated states. Additional insights can be drawn based on the crystal structure of the minitrimer core (RPA70C/32D/14). Although RPA32N is not present in the trimeric core, the direction of the backbone at amino acid 45 orients RPA32N in close proximity to RPA70C, roughly pointing towards Phe532 and Tyr581 (Figure 3.1). These residues form a cleft that is 20 Å away from Ile46 (Figure 3.1b). Strikingly, the DNA-binding cleft of 32D is also ~20 Å from Ile46 [12] (Figure 3.1c). Thus, structural data suggests that RPA32N might interact with RPA70C and/or RPA32D. It has been proposed that such interactions could regulate the DNA binding activity of the core [14].

Given the possibility that RPA32N may interact with domains in the trimer core, it is of interest to consider whether or not phosphorylated RPA32N interacts differently with other domains of RPA. A previous report from the Wold laboratory demonstrated a direct interaction

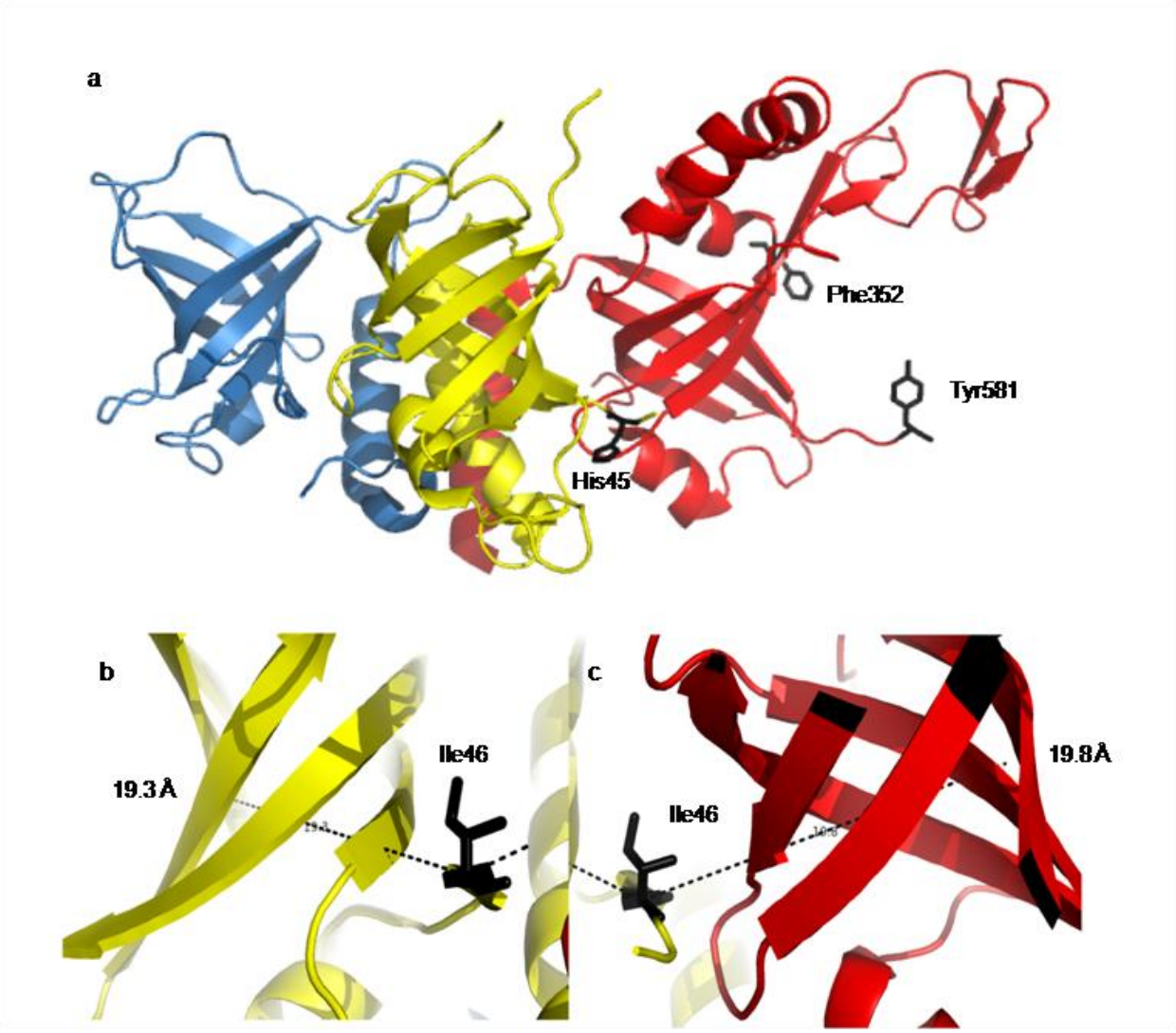


Figure 3.1. RPA minitrimer core. Direction of the backbone at RPA32 amino acid 45 orients RPA32N in close proximity to RPA70C, roughly pointing towards Phe532 and Tyr581 (a). RPA32N Ile46 is ~20 Å away from the RPA32D (b) and RPA70C (c) basic clefts.

using NMR between RPA70N and an RPA32N phospho-mimic synthetic peptide that contains all phosphorylatable residues (Ser8, Ser11, Ser12, Ser13, Thr 21, Ser23, Ser29, Ser33), except Ser4, mutated to aspartic acid [44].

Further evidence for intersubunit interactions was demonstrated in the intact RPA heterotrimer phosphorylated *in-vitro* with DNA-PK [119]. Lysine residues on the surface of the protein were biotinylated and the protein trypsin digested, then the peptide patterns analyzed by

SDS-PAGE and unique peptides identified by MALDI-TOFF and MS/MS mass spectrometry. Protection against trypsin digestion by biotin labels bound to the surface lysines in the non-phosphorylated RPA would be expected on the hyperphosphorylated RPA, unless a conformational change or rearrangement may occur upon phosphorylation, in which case the protection pattern would change as some lysine residues may no longer be exposed at the surface and be protected from biotinylation, or new lysines would be exposed to the surface. A comparison of peptide masses from non-phosphorylated and hyperphosphorylated RPA samples demonstrated a possible interaction between RPA70B and the hyperphosphorylated RPA32N, but not with 70N [119]. The authors argue that this is a more feasible scenario since RPA32N is in closer proximity to RPA70B than to 70N. However, given the long length of the RPA70N linker and direct evidence for RPA70N interaction with the phospho-mimetic peptide shown by NMR [44], this assumption may not be accurate. Moreover, this does not explain why the phosphorylated RPA32N would not interact with RPA70C or RPA32D which, according to the minitrimer core crystal structure, is even more accessible to RPA32N than RPA70B [14].

The same study confirmed RPA70B binding to the phosphorylated RPA32N [119]. Tryptophan fluorescence quenching experiments of a full-length RPA mutant construct in which all tryptophan residues, except RPA70B W361 and RPA70A W212, had been mutated to alanine was compared before and after *in-vitro* DNA-PK phosphorylation, demonstrating that hyperphosphorylation resulted in significant tryptophan fluorescence quenching. The effect was attributed only to RPA70B since biotinylation experiments indicated that RPA70B was the most likely candidate for interaction with the phosphorylated RPA32N. However this study cannot rule out the possibility that RPA70A could also contribute to the interaction with the

phosphorylated RPA32N peptide since the contribution from RPA70A W212 was not separated from RPA70B W362, and it did not address interactions with RPA70C or RPA32D.

In an effort to eliminate controversy and shed light into possible rearrangement of RPA structure induced by hyperphosphorylation of RPA32 that may lead to changes in RPA function, I have performed a systematic study of RPA32N interactions with other RPA domains. A recombinant pseudo-phosphomimetic peptide, RPA32N-D8, designed with Ser →Asp mutations on Ser8, Ser11, Ser12, Ser13, Thr21, Ser23, Ser29, and Ser33 was used to mimic the hyperphosphorylation of RPA. These mutations replicate the peptide conditions used in the previous NMR study focused solely on RPA70N [44] and those selected for *in vivo* studies showing localization of pseudo phosphorylated RPA32 to sites of DNA damage [68].

Results

Production of recombinant RPA32N wild-type and phosphomimetic peptides

In order to characterize the RPA32N domain and putative changes that occur upon hyperphosphorylation of RPA32N, wild-type and mutant peptides were subcloned using existing plasmids from full-length RPA as templates. We initially tried to build a 6X-His tagged pET15b construct of wild-type RPA32N that spanned only the 35 amino acids, $RPA32N_{1-35}$ containing the minimal region of RPA32N phosphorylation. However, although construction of the plasmid was successful, the peptide was either degraded during lysis or trapped into inclusion bodies, depending on the presence or absence of lysozyme in the lysis buffer, respectively. In order to improve the prospect of isolating the peptide, 6X-His-MBP and 6X-HIS-SUMO $RPA32N_{1-35}$ recombinant peptides were generated using in house vectors, pLM102 and pBG102, respectively.

Protein at high expression levels was successfully purified using NiNTA affinity chromatography. However, the peptide degraded immediately after cleavage of the affinity tags. Therefore, it was determined that the construct was unstable, and the fragment was extended to 46 amino acids to include the linker region that connects the RPA32N flexible domain to the RPA32D central domain of RPA32 subunit. Wild-type (RPA32Nwt) and mutant (RPA32N-D8) $RPA32N_{1-46}$ constructs were successfully generated using pLM302 and pBG102 vectors. Expression of recombinant peptides from these constructs remained as high as before, and cleavage of the affinity tags by H3C protease did not result in degradation of the peptides (Figure 3.2 and 3.3).

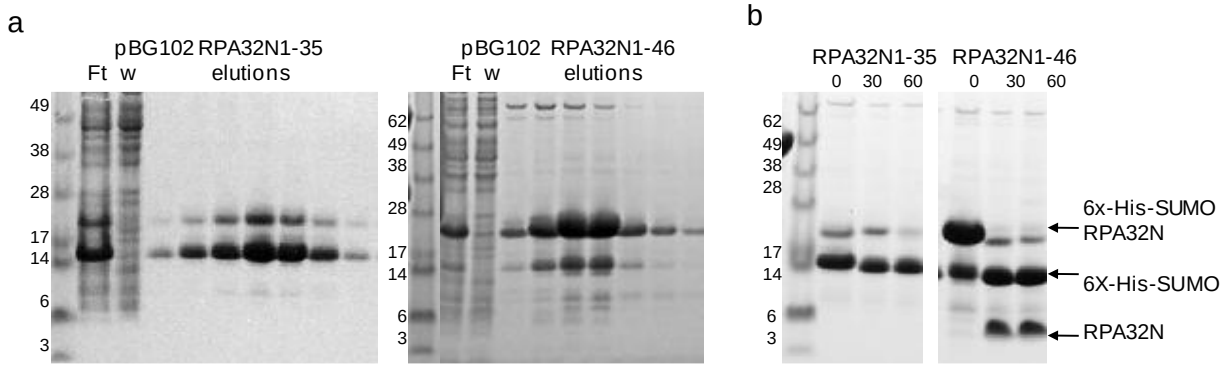


Figure 3.2. Purification of pBG102 RPA32Nwt $_{1-35}$ and $RPA32Nwt_{1-46}$ peptides. a). Nickel purification of the 35 amino acid fragment shows that degradation of the peptide has taken place during the purification as observed in the flow through and elution fractions containing mostly 6X-His sumo tag. Nickel purification of the 46 amino acid fragment shows this peptide is more stable. Elution fractions contain mostly 6X-His-Sumo-peptide. b). Cleavage reaction with H3C protease shows that only the 46 amino acid peptide can be recovered.

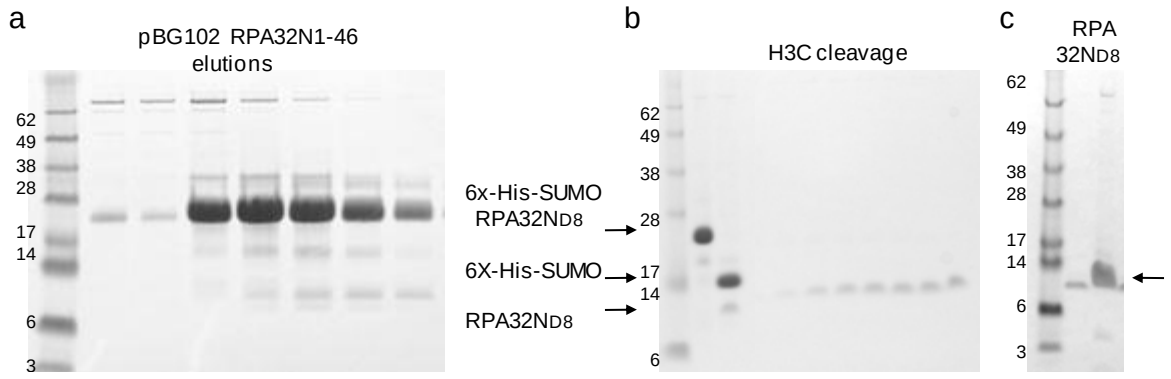


Figure 3.3. Purification of pBG102 RPA32N-D8 peptide. a). Nickel purification of the 46 amino acid 6X-His-SUMO-RPA32N peptide containing aspartic acid mutations on Ser4, 8, 11, 12, 13, 23, 29, 33 and Thr 21. Partial cleavage is seen during the purification. b). H3C protease cleavage reaction shows good separation of the peptide which runs at ~10 KDa. c). Dilute and concentrated peptide.

Wild-type and phosphomimetic RPA32N peptides are unstructured

A combination of matrix-assisted laser desorption/ionization mass spectrometry (MALDI-MS) analysis, circular dychroism (CD) and nuclear magnetic resonance (NMR) spectroscopy was used to characterize the RPA32Nwt and RPA32N-D8 peptides. Analysis of MALDI-MS spectra showed a single peak with 5087 m/z for RPA32Nwt, and a peak with 5338 m/z for RPA32N-D8. Although both peptides have the same number of residues, their difference in mass is consistent with the amino acid composition of the peptides. Each Ser → Asp mutation generates a 28 g/mol difference in molecular mass, in addition to the 13.9 g/mol difference for one Thr → Asp mutation. A single minimum observed at 200 nm in the far-UV CD spectrum for the RPA32Nwt indicated that the peptide lacks any helical or beta sheet character. In order to examine if the phosphomimic mutations elicit helical character formation, the same experiment was performed with the RPA32N-D8 peptide. However, as observed with the wt peptide, a

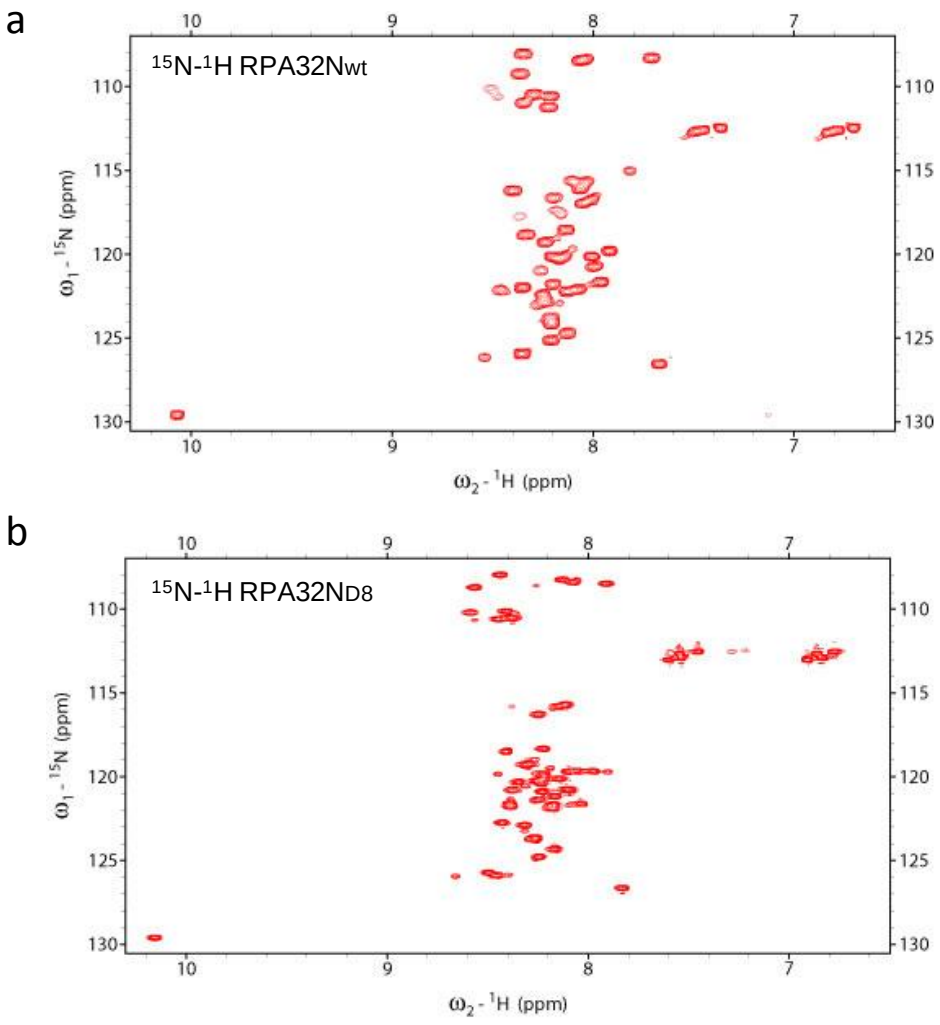


Figure 3.4. ^{15}N ^{1}H HSQC spectra of RPA32N peptides. a) ^{15}N ^{1}H RPA32Nwt spectrum is shown on the left. b) The ^{15}N ^{1}H HSQC spectrum of RPA32N-D8 is shown on the right. A gel showing the NMR sample is shown to the left of each spectrum.

single minimum at 200 mM in the far-UV spectrum indicates no secondary structure is formed. Conclusively, ^{15}N- ^{1}H HSQC NMR spectra of ^{15}N-enriched RPA32Nwt and RPA32N-D8 show a lack of peak dispersion for both peptides, which is indicative of lack of tertiary structure or unfolding (Figure 3.4). After optimization of NMR spectral conditions, which involved testing various buffers, temperature and pH conditions, a total of 30 peaks can be distinguished for the RPA32Nwt versus 43 peaks for RPA32N-D8 mutant.

Isolated RPA70N, RPA70A and RPA70B Domains Interact With Phosphomimetic RPA32N

As previously mentioned, ^{15}N-^{1}H HSQC NMR titration experiments indicate a weak interaction between RPA70N and a phosphomimetic RPA32N peptide [44]. This interaction has been mapped to the protein-protein interaction interface of RPA70N, involving residues Asn29, Ile30, Arg31, Tyr42, Leu44, Leu45, Ser55, Met57, His80, Arg92, Val93 and Val94. However, mass spectrometry footprinting assays on full-length RPA hyperphosphorylated with DNA-PK only detected a protection on RPA70B domain residues Lys343, Arg335, and Arg382 [119]. These results lead to the conclusion that interaction with RPA70N was not likely to occur in the full-length protein. However, this type of analysis presents certain restrictions as the read out mechanism involved protection of biotinylated lysine and arginine residues only, which represent only four of the identified RPA70N interacting residues. In order to fully characterize RPA32N-mediated intersubunit interactions a systematic approach was applied in an effort to accurately describe all interactions. The observations from the above mentioned studies suggest that intersubunit interactions with RPA32N occur via electrostatic forces. Given that the RPA70A domain has a more basic interacting region, it was possible that this domain also participates in interactions with RPA32N. Thus, we begin our analysis with the high-affinity ssDNA binding domains RPA70AB.

A ^{15}N-^{1}H HSQC spectrum was recorded for ^{15}N-RPA70AB in buffer containing 10 mM HEPES at pH 7.5, 100 mM NaCl and 5 mM BME, and a second spectrum was collected upon addition of a 15X molar excess of RPA32N-D8 peptide. NMR assignments for RPA70A and RPA70B were transferred to the RPA70AB spectrum, and chemical shift perturbations resulting from addition of the RPA32N-D8 peptide were mapped to the corresponding residues. Chemical

shift changes that were above the average plus one standard deviation were considered significant chemical shift perturbations. Residues Ser195, Ser213, Gly219, Arg234, Lys263, Asp301, Val334, Val341, Arg344, Arg382, Ser384, and Asp385 were perturbed in the spectrum. These results were largely reproduced upon addition of RPA32N-D8 to the isolated ^{15}N-RPA70A and ^{15}N-RPA70B domains. Curiously, many more chemical shift perturbations were observed on the isolated domains, including for RPA70A residues Lys206, Arg210, Try212, Asn214, Arg216, Gly217, Glu218, Gly219, Lys220, Leu224, Arg234, Thr236, Asn239, Ala265, Asn266, Gln268, and Phe269. Likewise, on RPA70B perturbations were observed on Asp314, Thr330, Ile332, Val334, Val341, Ala365, Gly362, Gly380, Arg382, Lys394, Val393, and Phe415 (Figure 3.5). Importantly, the residues identified are in the same vicinities of the domain suggesting that the site of interaction is specific.

Interestingly, in the RPA70AB crystal structure, loop L45 from the 70A domain curves around the DNA and is held in place by Phe238 and Phe269 aromatic ring stacking between the DNA bases forming a closed conformation upon binding DNA [45]. ^{15}N-relaxation and chemical shift perturbation experiments performed on 70A also showed that binding-induced changes in the 70A structure involved residues in L45 loop [5]. These NMR chemical shift perturbation experiments show that Phe269 is perturbed when RPA70A alone, but not in the context of RPA70AB and Phe238 is not perturbed. Other residues indirectly affected by ssDNA binding in the crystal structure include Trp212-S223 residues in the Leu12 loop and the β2 strand, Ser234-Asn239 in the β 3 strand, and Asn266-Glu277 in the Leu45 loop. Some of these residues are in the same vicinity as the perturbations observed in RPA70AB by addition of the peptide. Interestingly, polar residues Arg210, Arg216, Arg234 and Lys263 shown to be essential for high affinity to ssDNA [120] were perturbed on RPA70A. However, only Arg234 and

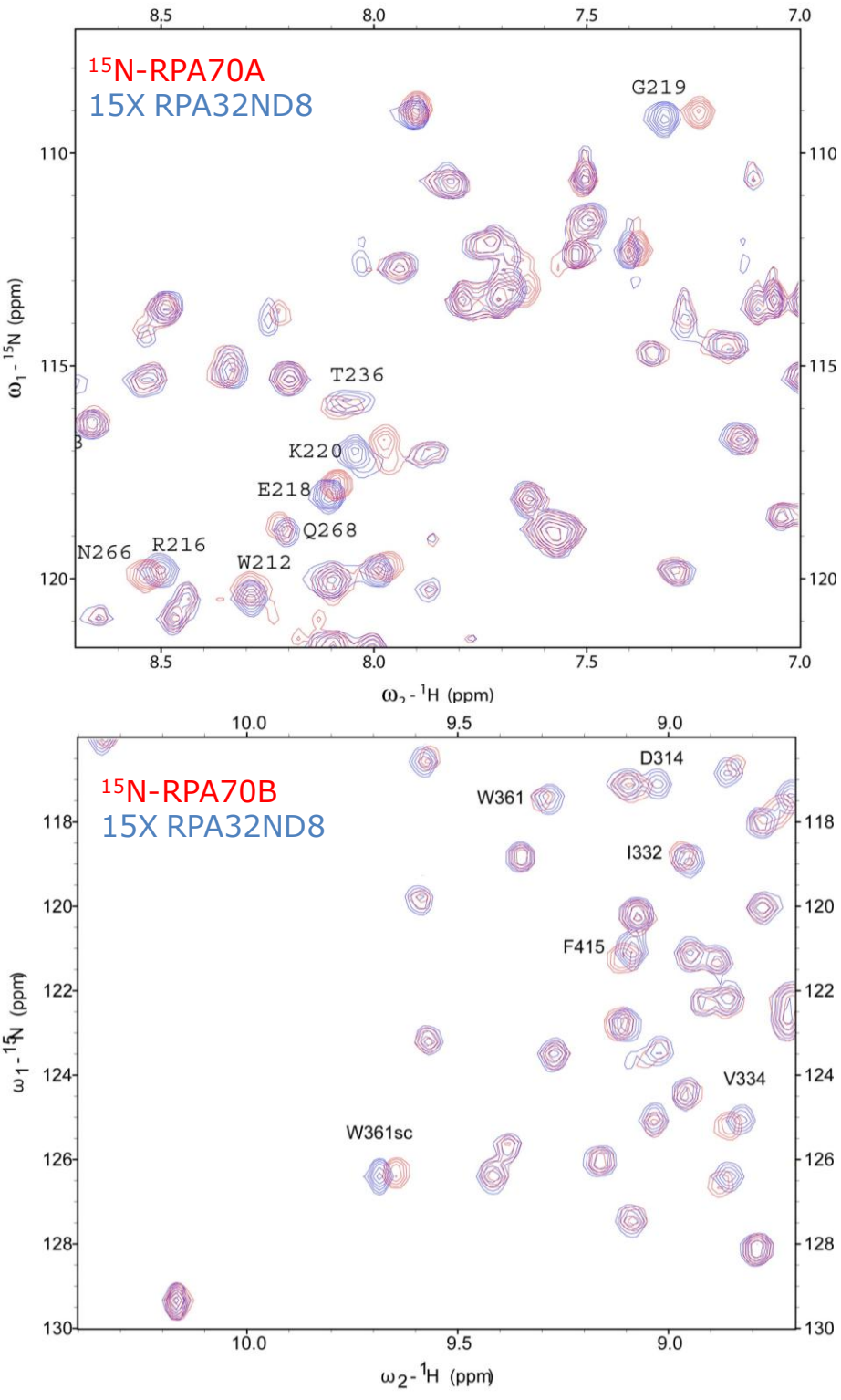


Figure 3.5. ^{15}N-^{1}H HSQC spectra RPA70A and RPA70B upon titration of RPA32N-D8.

were perturbed on RPA70AB. An RPA mutant which contained Arg234A and Lys263A mutations, in addition to Glu277A mutation, was shown to have a reduced ssDNA binding activity measured to less than 1% of the WT [120]. However, this mutant can support normal replication activity. Consistent with this, fluorescence measurements of full-length WT and mitotic phosphorylated RPA forms showed that both proteins have similar ssDNA binding activities [60]. On the contrary, a mutant with Arg210A, Try212A, Arg216A, Arg234A and Lys263A mutations, in addition to Glu277A, has reduced ssDNA binding activity and reduced replication activity while mutations on aromatic residues alone activate the cell cycle checkpoint control [120]. In addition, Try212, Arg216, Arg234 and Lys263 and other surrounding residues are involved in direct interaction with Rad51 N-terminus domain, and although this interaction had no effect on RPA70A ssDNA binding activity, it was important for RPA displacement from ssDNA. On RPA70B, perturbations on Val334, Arg344, Arg382 and Ser384 confirm the observations from MS footprinting analysis. Thus, as previously observed with RPA70N, residues surrounding the ssDNA binding region of RPA70AB participate in the interaction with the RPA32N-D8 peptide. Residues corresponding to the observed interaction have been mapped to the RPA70AB structure (Figure 3.6).

The interaction of RPA32N-D8 with the RPA70N subunit was confirmed by reciprocal titration experiments on ^{15}N-^{1}H enriched RPA32N-D8 peptide. In addition, titration of unlabeled RPA70NAB resulted in a different set of perturbations that also confirms the additional interactions observed in the presence of RPA70AB. Near complete resonance assignments of RPA32N-D8 allowed for partial identification of the peptide residues involved in the interactions (Figure 3.7). Residues 10, 15-26, 28, 33 and 43 were perturbed in the interaction. These include six of the nine glycines (10, 15, 16, 18, 19, and 18), two alanines (17, 43), three of the nine

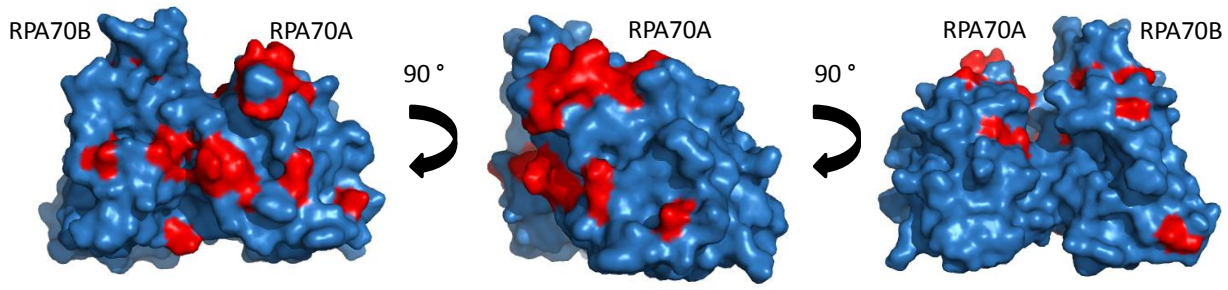


Figure 3.6. Contact points in RPA70A and RPA70B interaction with RPA32N-D8 are highlighted in red. Most interacting residues belong to RPA70A.

aspartic acids (21, 23 and 29) and Tyr20. However, due to large overlap in the RPA32N-D8 spectrum, Asp8, Asp11, Asp12, Asp13, and Asp29 have not been assigned and it is possible that these residues also experienced chemical shift perturbations. Residues Met1, Trp2, Ala31, Ser39, Arg40 and Ala41 were not perturbed.

Given that three RPA70 domains interact with the pseudo-phosphorylated peptide, and given its size and innate flexibility, we propose that in order for hyperphosphorylated RPA32N to interact with all RPA70 domains, it must do so transiently by switching between one domain and another. Therefore, we asked if any of these interactions was more favorable than the others in order to understand if some interactions are more biologically relevant. To answer this question, I used NMR reverse titration experiments. Data was acquired for two samples of known concentrations, one containing ^{15}N-labeled protein alone (RPA70N, RPA70A, RPA70B, or RPA70AB) and the other containing the matching ^{15}N-labeled protein saturated with a 15 fold excess of the RPA32N-D8. Spectra of these two samples were collected as initial and final points of titration. Then 50 µl from each sample were swapped to collect intervening titration points, and the process was repeated various times resulting in 8-10 titration points. Specific resonance assignments were transferred to each spectrum, and the chemical shift change for

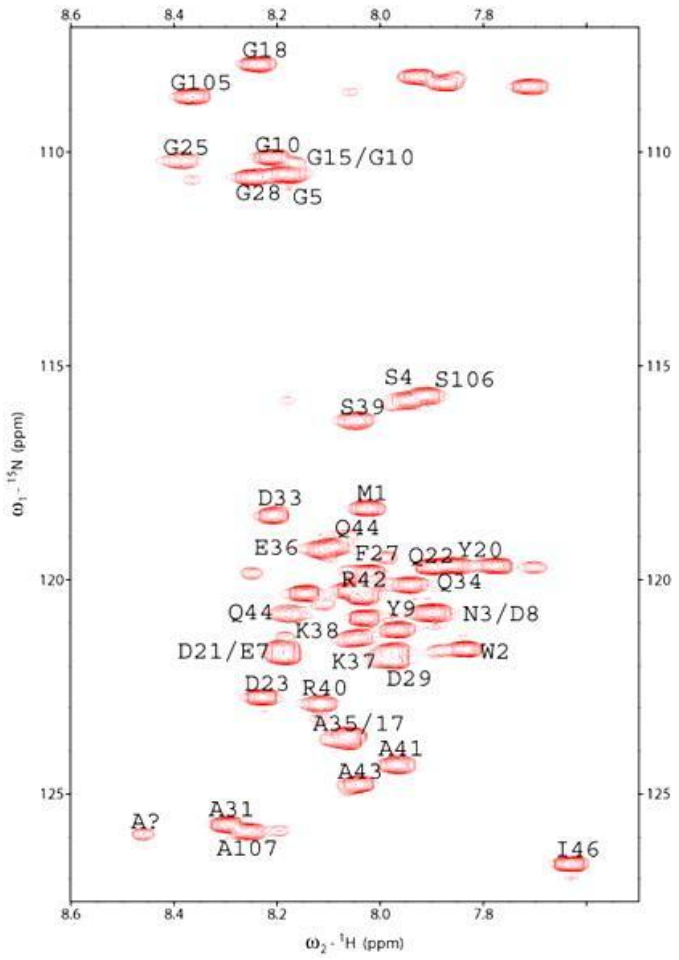


Figure 3.7. RPA32N-D8 backbone specific assignments.

every residue at each titration point was calculated. Although the number of titration points collected was not very large, it was possible to estimate a range of Kd values. An unassigned residue for RPA70N at 128, 7.712 ppm; residues Glu218 and Gly219 from RPA70A; and Arg344 and Gly362 from RPA70B were used to estimate the binding affinity using xcrvfit software. The calculated binding affinities were 4.1 $\pm$ 1.0 x 10^{-4} M, 3.9 $\pm$ 0.8 x 10^{-4} M, 7.8 $\pm$ 1.2 x 10^{-4} M, 4.6 $\pm$ 1.0 x 10^{-4} M, and 7.9 $\pm$ 2.0 x 10^{-4} M, respectively. Notably, the observed chemical shift perturbations were smaller in the context of RPA70AB. Thus, affinities obtained for RPA70A and RPA70B residues Glu218, Gly219, Arg344 and Gly362 in the isolated domains

where compared to those obtained in RPA70AB. The calculated binding affinities were 7.5 $\pm$ 1.0 x 10^{-4} M, 8.8 $\pm$ 1.1 x 10^{-4} M, 1.3 $\pm$ 0.8 x 10^{-3} M, and 1.8 $\pm$ 1.0 x 10^{-3} M, respectively.

To help envision how this is happening, electrostatic potential maps were calculated to evaluate the charge contributions to the interactions of RPA32N-D8 with RPA70 domains (Figure 3.8). A larger electrostatic potential for RPA70A may explain the greater number of residues, compared to RPA70B, that interact with RPA32N-D8.

RPA32N-D8 interaction is specific to the RPA70 subunit

Weak interactions can sometimes be observed as a product of non-specific binding. This can occur depending on the electrostatic properties of two proteins and as a consequence of high concentrations such as during saturation of one constituent in an NMR titration experiment. However, specific weak interactions are important for biological function. These result in transient associations that are often coupled to regulation of a protein's biological function or activity [121]. In order to demonstrate that phosphorylated RPA32N intersunit interactions are

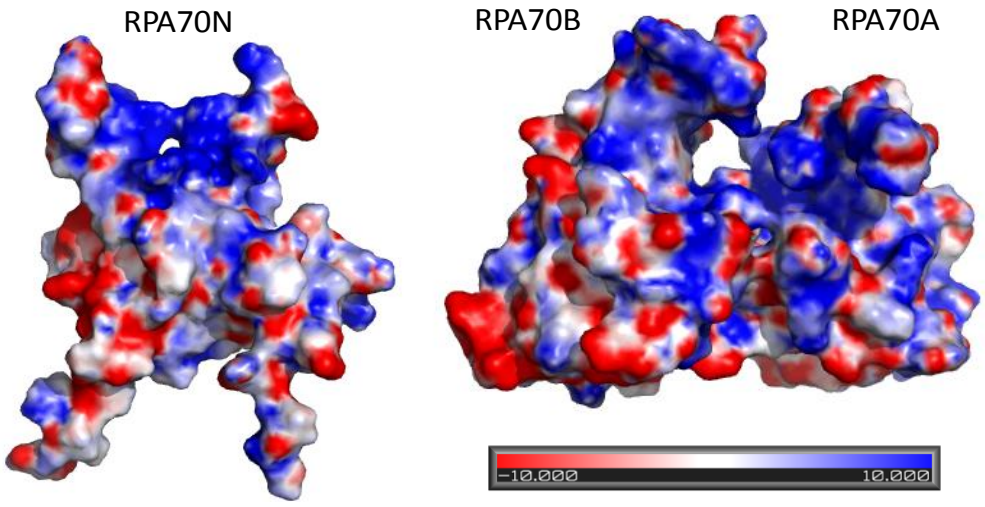


Figure 3.8. Electrostatic Potential maps. Electrostatic maps for RPA70NAB domains is shown. The blue area represents basic surfaces, while the red area represents acidic surfaces.

specific to the RPA70 subunit and do not result from macromolecular bumping due to high concentration of ligand, we asked if any other domains in RPA might interact with RPA32N-D8 peptide. Chemical shift perturbations assays were performed by titration of unlabeled RPA32/14, RPA32D/14 and RPA32C to 18 μM ^{15}N-RPA32N-D8 up to a 20-fold excess molar ratio in buffer containing 10 mM HEPES at pH 6.8, 50 mM NaCl and 5 mM BME. These experiments showed that there are no chemical shift differences in the ^{15}N-RPA32N-D8 spectrum upon addition of the proteins, and demonstrated that the pseudo-phosphorylated peptide does not interact with these domains (Figure 3.9). Likewise, chemical shift perturbation experiments with the RPA32Nwt peptide demonstrated that the unphosphosphorylated RPA32N also does not interact with RPA32D/14 or RPA32C. The same experiment was performed on ^{15}N-RPA70C/32D/14, a construct which contains the C-terminal domain of the RPA70 subunit. Although chemical shift perturbations were not detected upon addition of RPA32N-D8, only a threefold excess was achieved in this experiment. Thus, we cannot rule out the possibility that RPA70C may also have a weak interaction and is involved.

Production and Characterization of RPA$_{D9}$ phospho-mimic mutant

Previous work in our laboratory involving RPA heterotrimer has utilized a pET15b RPA construct that produces a 6X-His tagged full-length protein. However, the yield and stability of the protein are not very good with this construct. In addition, the purity of the sample after a series of purification steps is still not great. Although we have managed to perform experiments with this construct, we felt that introducing mutations could result in the same or even lower yield and/or potentially unstable protein. To optimize this construct, we decided it was best to

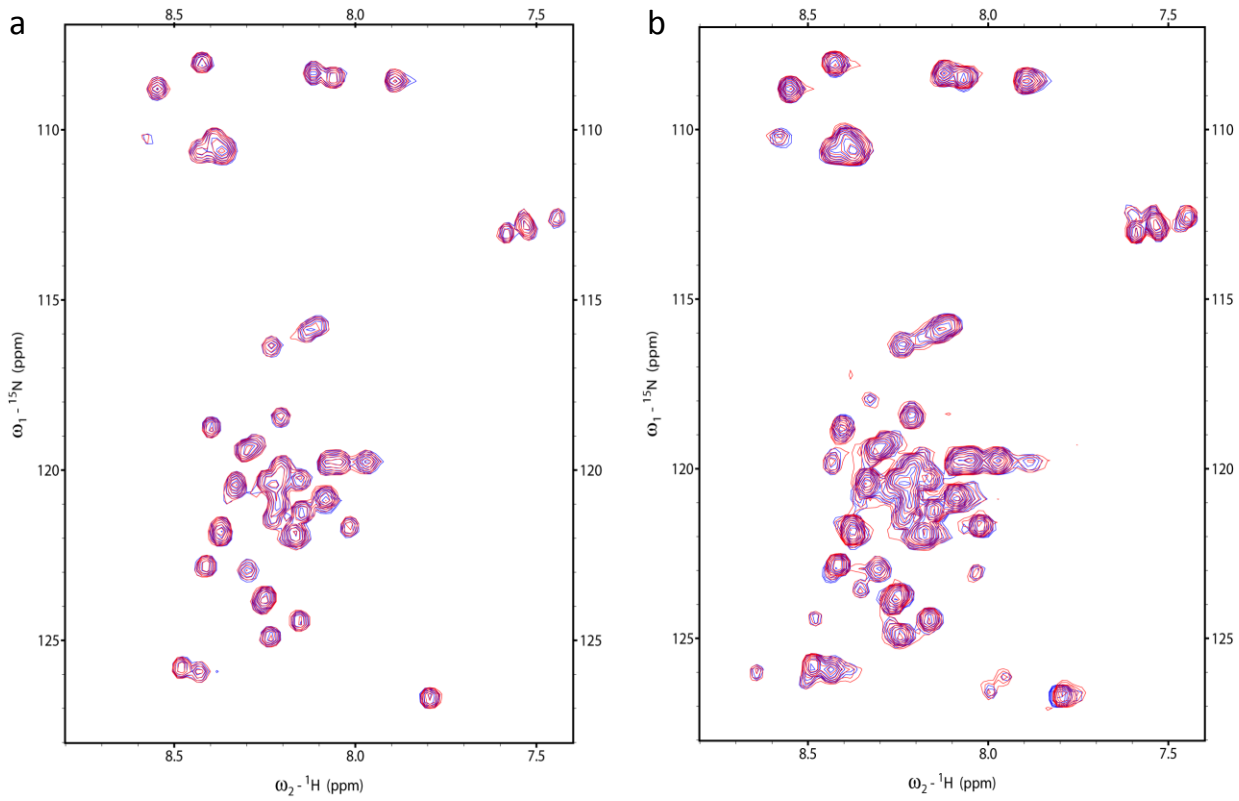


Figure 3.9. ^{15}N-^{1}H HSQC titrations. ^{15}N-enriched RPA32N-D8 (blue) was saturated with 20X excess RPA32D/14 (red, a) or RPA32C (red, b). No chemical shift perturbations are indicative of a lack of interaction.

use a different vector system. An optimized version of the full-length wild-type RPA heterotrimer and several versions of full-length phosphorylated RPA were produced through collaboration with Dr. Miaw-Sheue Tsai in the Structural Cell Biology of DNA Repair Machines (SDBR) group at Laurence Berkley National Laboratory.

The already existing pET15b RPA tri-cistronic construct was engineered by the laboratory of Dr. Alexey Bochkarev. The open reading frame for the RPA70 subunit was placed in front followed by RPA14 and RPA32, respectively. The protein contains a cleavable 6X-His tag on RPA70 and a non-cleavable RPA14 His tag. According to previous characterization of a pET11b RPA construct developed by the laboratory Dr. Mark Wold, RPA70 is the lowest copy

transcript, and was positioned in front to promote a higher yield of trimer [122]. Thus, the same subunit order was adopted. The RPA constructs were designed with cleavable 10X-His-tags at the N-terminus of RPA70 (H3C protease specific) and RPA14 (TEV protease specific) subunits and cloned into a 10X-His pBG106 vector. Pseudo-hypophosphorylated RPA (Ser23,29 → Asp) and pseudo-hyperphosphorylated RPA (Ser4,8,11,12,13,23,29,33→Asp and Thr21→Asp) mutant proteins were produced based on our knowledge of residues that become phosphorylated in response to cell cycle events and DNA damage [68, 123] (Figure 3.10). To examine the structural dynamics of RPA-D9 and investigate the impact of hyperphosphorylation on RPA conformation, purification conditions were optimized and preliminary SAXS data for RPAwt and RPA-D9 were acquired. As demonstrated in Chapter II, monodispersity of samples is critical because small amounts of aggregation or contaminants would distort the scattering leading to misinterpretation of the results. Therefore, multiple purification steps were implemented to achieve the purity of the RPA samples, albeit at the cost of reducing the final yield.

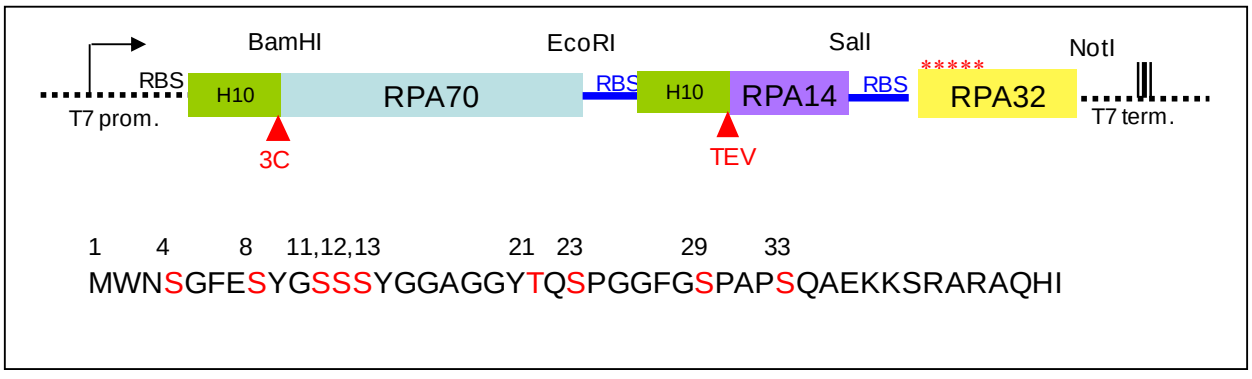


Figure 3.10. Design of pBG106 RPA constructs. Full-length RPAwt and a pseudo phosphorylated mutant were constructed in a pBG106 vector. RPA70 and RPA14 subunits contain 10X-His tags at their N-terminus cleavable with H3C and TEV proteases, respectively. RPA32N sequence is shown bellow. Phosphorylatable residues mutated to aspartic acid are highlighted in red.

For RPAwt, the data were truncated and scattering at high angles was absent in the preliminary experiments due to the small amount of protein available. RPA protein preparations are notoriously unstable, and it is possible that partial degradation of the sample may have occurred in the interim between preparation and data collection, resulting in insufficient amounts for SAXS. In contrast, scattering data for the RPA-D9 mutant were successfully recorded. Purification of RPA-D9 mutant yielded greater amounts of protein. SAXS data was acquired for the RPA-D9 sample at 14 μM (1.5 mg/ml) (Figure 3.11). An Rg of 64.7 Å was derived directly from the data based on the Guinier analysis (Figure 3.12). The Kratky analysis shows the curve plateaus at high $I(q)*q^2$ when $q > 0.15$. This pattern is characteristic for the presence of some disordered regions, consistent with the presence of RPA32N and the long inter-domain linkers. The P(r) function shows multiple peaks, reflecting the scattering of the different domains. The curve extends to a *Dmax* of 220 Å, which is consistent with the available structural information for the seven globular domains and the intrinsic flexibility of RPA. These data obtained for full-length RPA mutant demonstrate the feasibility of the SAXS approach for analyzing the effect of phosphorylation on the structural dynamics of RPA.

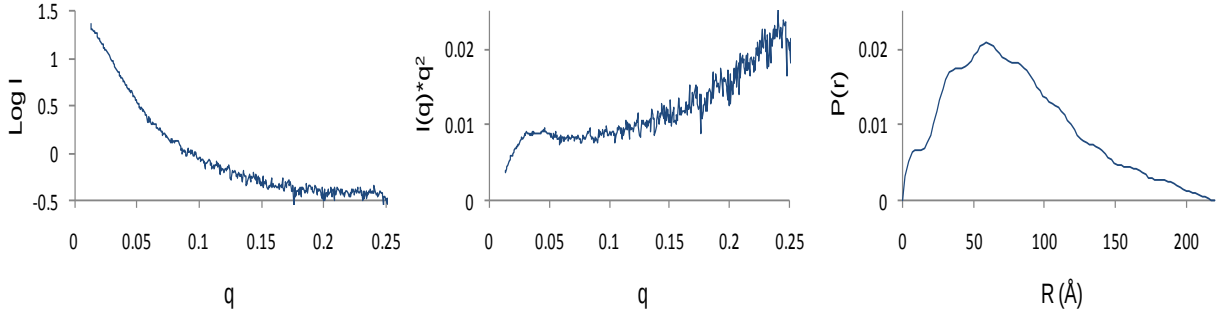


Figure 3.11. SAXS data acquired for RPA-D9: (a) Scattering curve; (b) Kratky analysis; (c) P(r) function

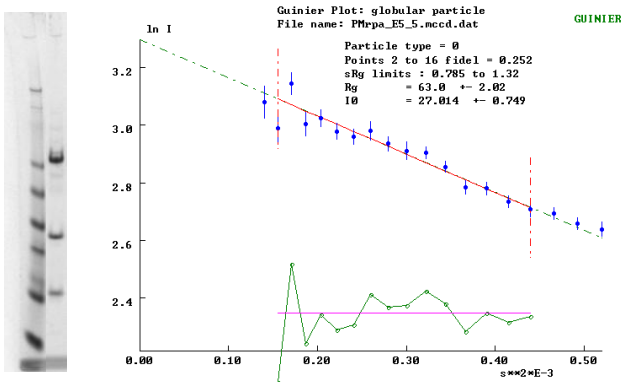


Figure 3.12. Scattering data for RPA-D9 does not show signs of aggregation. SDS-PAGE of the sample is shown in the left panel. Guinier Analysis from the SAXS data is shown at right.

Discussion

RPA mediates interactions with ssDNA and target proteins to coordinate the progression of DNA processing events. Hypophosphorylation of RPA by cdk/cyclin dependent kinases takes place as part of a regulatory mechanism during normal replication activity [56, 124]. In the presence of DNA damage, RPA becomes hyperphosphorylated by the PI3K family of kinases. This is thought to make RPA "repair competent" by means of recruitment of repair factors to sites of DNA damage [57, 58, 68, 76-79, 82]. Evidence supporting this hypothesis is very extensive. PI3K family of kinases are activated in response to DNA damage [74, 80], and the fact that RPA is a target of these kinases is the first clue to understanding that it is involved in repair activities. It has been shown that hyperphosphorylated RPA is inactive in replication, and arrest of replication activity correlates with the loss of hypophosphorylated forms of RPA [58]. Cell studies have also shown that hyperphosphorylated RPA is localized to sites of DNA damage upon UV exposure [68, 123]. Moreover, RPA association with proteins involved in replication is weakened upon hyperphosphorylation [60]. In contrast, association with proteins involved in recombinatorial repair is move avid upon hyperphosphorylation [69].

Despite the existence of atomic resolution information for its seven structured domains, the overall disposition of RPA domains and hence the structural details of its mechanism of action are unknown. Thus changes that occur in RPA structure upon hyperphosphorylation cannot easily be addressed. However, two studies have contributed insight into potential hyperphosphorylation-specific structural effects. ^{15}N-^{1}H HSQC NMR titrations of the isolated ^{15}N-RPA70N domain with a synthetic RPA32N phosphomimetic mutant peptide showed evidence for a weak interaction occurring via the basic cleft of RPA70N [44]. Further evidence for intersubunit interactions was demonstrated in intact hyperphosphorylated RPA through potential binding of RPA32N to RPA70B. However, there was no indication for RPA70N interaction [119]. It has been hypothesized that hyperphosphorylation driven intersubunit interactions may affect RPA-DNA and RPA-protein interactions [3, 44, 119].

The studies presented here demonstrate that RPA32N hyperphosphorylation-dependent intersubunit interactions occur via the RPA70N, A, and B domains and utilize SAXS to examine the structural effect of hyperphosphorylation on RPA. Although preliminary data is available for RPA$_{D9}$, several further studies are required. In order to fully interpret the effect of hyperphosphorylation on RPA, detailed comparisons to the wild-type protein are required. These studies will be challenging because comparisons will be limited to low resolution SAXS parameters such as *Dmax* and Rg values in this highly flexible molecule. To complete the analysis, our laboratory is pursuing an approach of building up models of smaller RPA fragments. These models can be used to construct a composite model for the intact RPA heterotrimer. This approach would allow us to build a comprehensive picture of the time-varying architecture of free RPA. Once these studies are complete, we will have the proper platform for further studies to addresses the effect of hyperphosphorylation on full-length RPA.

Re-acquisition of RPA-D9 SAXS data will be necessary as the experiment needs to be performed in parallel with RPAwt due the great sensitivity of SAXS to the solution conditions and the variance in the intensity of the beam line. In addition to hyperphosphorylation, delineation of the scattering contributions of the RPA domains will also facilitate studies of remodeling as RPA binds ssDNA and other DNA processing partners. With this goal in mind, I had previously optimized purification conditions and collected SAXS data for all RPA fragments available in our laboratory in a collaboration with Drs. Kevin L. Weis and Ragavan Aravinda at Oak Ridge National Laboratories (ORNL), with whom we collected data at a local source and at the Advanced Photon Source (APS) at the Argone National Laboratory. However, all SAXS data shown in this thesis were acquired through collaboration with Susan Tsutakawa in the laboratory of John Tainer at the SYBILS beam line at Berkeley National Laboratory. The implications of our findings will be further discussed in Chapter IV.

Experimental Procedures

Cloning.

Primers were designed to clone RPA32N 1-35 and 1-46 residues. The first 35 residues of RPA32 span the functional domain, and the next 10 residues (36 to 46) span the linker to the rest of the subunit. For the RPA32Nwt peptide, RPA32F forward primer (5'-CGCGGATCCGCG**ATG**TGGAACAGTGGA-3') contains a BamH1 restriction site (underlined) and ATG start codon. Reverse primers RPA32R35 (5'-CCGGAATTCCGGTTAGGCTTGAGAAGG-3') and RPA32R46 (5'-CCGGAATTCCGGCTAAATGGTCTGCTGGGCTCG-3') contain an EcoR1 restriction site (underlined) and either TTA or TAA stop codon. The same restriction sites were used for

construction of RPA32N phosphomutant peptides. RPA32FS4/8D forward primer (5'-GGATCCGCG**ATG**TGGAACGATGGATTCGAAGAC-3'), and either RPA32RS33D (5'-TGATTTCTTTTCGGCTTGATCAGGTGCTGGA-3'), or RPA32R46 reverse primer were used. Plasmids were amplified using a 20 µl polymerase chain (PCR) reaction using New England Biolabs (NEB) Vent polymerase. Full-length wild-type or mutant RPA heterotrimer were used as templates. The amplified DNA fragments were separated in 2% agarose gels, and extracted using a Quiagen gel extraction kit. Each time insert and vector, either Invitrogen pET15b (6X-His), or in-house pBG102 (6X-His-Sumo tag) and pLm102 (6X-His-MBP tag), were digested overnight at 37 °C with NEB BamH1 and EcoR1 enzymes. The next day the reactions were run on a gel and the digested plasmids isolated and eluted in 50 µl TE buffer. The ligation reaction was set up using the quick T4 ligation system from NEB, and the construct was transformed into XL1-Blue cells and plated onto ampicillin or kanamycin plates. DNA sequencing confirmed that the constructs contained the targeted sequences.

Construction of plasmids for the full-length RPA proteins was in collaboration with Dr. Miaw-Sheue Tsai at Laurence Berkley National Laboratories. To create an N-10X-His H3C tagged RPA70, RPA70-1F forward primer (5'-GCGGGATCCATGGTCGGCCAACTGAGCGAGGGGGCC-3') contained a BamH1 restriction site, and RPA70-616R reverse primer (5'-CCGGAATTCTCACATCAATGCACTTCTCCTGATGC-3') contained the Eco R1 restriction site. To create an N-10X-His-TEV tagged RPA14 subunit, a RPA14-1F forward primer(5'-CCGGAATTCTTTTTGTTTAACTTTAAGAAGGAGATATACCATGCACCATCACCATCA CCATCACCATCACCATCACCATAGCAGCGAGAATCTTTATTTTCAGGGCATGGTGGA CATGATGGACTTGCCCAGG-3') contained an EcoR1 site, RBS site, 10X-His, linker and

TEV sequences prior to start codon. The RPA14-121R reverse primer (5' AGCGGTCGACTCAATCATGTTGCACAATCCC-3') contained a Sal1 restriction site. To create an untagged RPA32 subunit, primer RPA32-1F (5'-AGCGGTCGACTTTTTGTTTAACTTTAAGAAGGAGATATACCATGTGGAACAGTGGATTCGAAAGC-3') contained a Sal1 restriction site and a RPA32-270R reverse primer (5'-ATAAGAAGCGGCCGCTTATTCTGCATCTGTGGATTTAAAATGGT) contained a NotI restriction site.

Protein Expression

pBG102 RPA32Nwt, pBG102 RPA32N-D8, pET15b RPA32D/14, pET15b RPA32C, pET15b RPA70C/32D/14 plasmids were transformed into Novagen BL21 (DE3) cells. pET28b RPA70AB, pBG100 RPA70NAB and pET15b RPA70N_{1-120} (obtained from the laboratory of Cheryl Arrowsmith) were expressed in LB using Novagen Rosetta (DE3) cells. Unlabeled proteins were prepared using LB medium containing kanamycin (pBG102 RPA32Nwt, pBG102 RPA32N-D8, pET28b RPA70AB, pBG100 RPA70NAB) or carbenicillin (pET15b RPA32D/14, pET15b RPA32C, pET15b RPA70C/32D/14, and pET15b RPA70N_{1-120}) at 37 °C. Uniform ^{15}N- and ^{15}N-^{13}C- labeled proteins were produced similarly, except for using M9 media enriched with $^{15}NH_4Cl$ and 1 ml 10% yeast extract. Proteins were purified using Sigma nickel-nitrilotriacetic acid resin (NiNTA) using an elution gradient from 10 to 300 mM imidazole, and size exclusion chromatography (SEC) using superdex 75 or superdex 200 resin.

pBG106QC7 RPA-D9 mutant was transformed into Novagen Rosetta (DE3) cells. The cells were plated after only 30 minutes of shaking rather than 1 hour. A 150 mL LB-kanamycin pre-culture was grown for 3 hours at 37 °C. Six 500 mL rich-autoinduction-kanamycin cultures

prepared in wide 2L glass flasks were inoculated with 10 ml of the 150 mL pre-culture at 37 °C overnight for a total of 13 hours. The pelleted cells were immediately processed for Nickel purification using a modified buffer to improve stability and purity of the sample. Buffers contained 30 mM Tris at pH 8.0, 1 M NaCl, 10 mM BME, 10 μM $ZnCl_2$, 5 mM $MgCl_2$, 30 → 300 mM imidazole, 0.25% NP-40 and 5% glycerol. Cleavage of the tags was only partially successful in spite of using increasing amounts of H3C and TEV proteases, therefore the tags were not cleaved in the final protein preparation.

Protein Purification

RPA32N peptides were purified in buffer containing 10 mM HEPES at pH 7.0, 50 mM NaCl, 10% glycerol. The His tag was cleaved during dialysis with H3C protease in 10 mM HEPES at pH 7.0, 100 mM NaCl, and 10% glycerol. A second NiNTA purification step was used to remove the His tag. Peptides were concentrated using a 3 kDa molecular weight cut-off concentrator in their existing buffer prior to exchanging into the NMR buffer. 1X peptide solution was mixed with 4X NMR buffer and concentrated down to 1 ml, and the process repeated twice again. After the third exchange, the concentration was re-checked and adjusted using NMR buffer.

The RPA70A and RPA70B proteins were purified using NiNTA in 10 mM HEPES at pH 7.5, 500 mM NaCl and 5 mM BME. The His tag was not removed from these proteins. ^{15}N-RPA70AB protein was purified using NiNTA in 10 mM HEPES at pH 7.5, 500 mM NaCl, 5 mM BME, and 10% glycerol using an elution gradient from 20 to 300 mM imidazole. The His tag was cleaved with TEV protease during dialysis overnight in 10 mM HEPES at pH 7.5, 200 mM NaCl, 5 mM BME, 200 mM L-arginine, and 10% glycerol. A Superdex 75 column was used as a

last step of purification. Similarly, ^{15}N-RPA70NAB samples were purified in 30 mM MES at pH 6.5, 500 mM NaCl, 10 mM BME, and 5 mM $MgCl_2$. The His tag was cleaved using H3C protease during dialysis overnight and separated with a Superdex 200 column in 30 mM MES at pH 6.5, 200 mM NaCl, 10 mM BME, 10% glycerol, and 5 mM $MgCl_2$. Prior to NMR experiments, RPA70A, RPA70B, and RPA70AB were dialyzed overnight at 4 °C in NMR buffer containing 20 mM Tris at pH 7.2, 2 mM DTT, 50 mM KCl, and 10 mM $MgCl_2$. RPA70NAB buffer contained 10 mM HEPES at pH 6.8, 50 mM NaCl and 5 mM BME. All proteins were concentrated using a 10 kDa molecular weight cut-off concentrator the day of the experiment.

RPA32C, RPA32/14 and RPA32D/14 proteins were purified using NiNTA in buffer containing 50 mM Tris at pH 8.0, 500 mM NaCl and 5 mM BME. Prior to experiments, proteins were buffer exchanged 3 times into NMR buffer containing 10 mM HEPES, pH 6.8, 50 mM NaCl, and 5 mM BME buffer.

RPA70C/32D/14 was purified using NiNTA in buffer containing 10 mM HEPES at pH 7.5, 500 mM NaCl, 5 mM BME, 10% glycerol and 10 uM $ZnCL_2$. Prior to experiments, RPA70C/32D/14 was buffer-exchanged three times into 10 mM HEPES at pH 7.5, 125 mM NaCl and 2 mM BME.

RPA-D9 mutant was purified using NiNTA in buffer containing 30 mM Tris at pH 8.0, 1M NaCl, 10 mM BME, 10 uM $ZnCl_2$, 5 mM $MgCl_2$, 30 → 300 mM imidazole, 0.25% NP-40 and 5 % glycerol. The sample was desalted into Source Q buffer A using a 54mL desalting column at 4 °C. Source Q buffers contained 50 mM Tris at pH 8.7, 100 mM L-arginine, 10% glycerol, 10 uM $ZnCl_2$, 5 mM $MgCl_2$ and 5 mM DTT. The same buffer was used for S200 gel filtration.

DNA sequencing

Sequencing was performed at the Vanderbilt University DNA sequencing core facility. DH5α or XL1bue cells expressing RPA32N peptides were grown overnight in LB medium, which contained the appropriate antibiotic. DNA was isolated using a Quiagen mini-prep kit. Sequencing was performed using a T7 promoter primer.

MALDI Mass spectrometry

Mass spectrometry analysis was performed at the Mass Spectrometry Research Center at Vanderbilt University using a MALDI Voyager Elite instrument (Applied Biosystems). A matrix was prepared that contained 10 mg of synapinic acid, 500 ul 0.1 % TFA and 500 ul acetonitrile. Calibration standards included insulin (5.7 kDa), cytochrome C (12 kDa) and albumin (66 kDa). Analysis of RPA32Nmt peptide was done using 256 scans with a 2500 laser.

NMR Chemical Shift Perturbation Assays

All NMR experiments were performed on 500 or 600 MHz Bruker spectrometers with a cryoprobe and operating at 25 °C or 15 °C. Two-dimensional ^{15}N-^{1}H-HSQC spectra were acquired on uniformly ^{15}N-enriched samples of RPA32Nwt, RPA32N-D8, RPA70N, RPA70A, RPA70B, and RPA70AB in 90% H_2O/10% D_2O, and in the presence of unlabeled proteins titrated to assess interactions. Buffers used for NMR titration experiments were as follows: 20 mM Tris-d_{11} HCl at pH 7.2 containing 50 mM KCl, 10 mM $MgCl_2$ and 2 mM DTT.

All NMR data were processed using TOPSPIN (Bruker Biospin Ltd.) and analyzed using Sparky (University of California, San Francisco). Kd values were determined using XWINNMR V 3.1 (Bruker).

NMR Resonance Assignments

In order to map the residues in RPA32N-D8 that interact with RPA70NAB, it is necessary to obtain backbone assignments. Since RPA32N-D8 is a small peptide, an initial attempt was made to determine sequential backbone assignments using 2D homonuclear NOESY and TOCSY spectra [99, 103]. An observation of the NOE from the NOESY spectrum provides information about proton-proton interactions through space, while the TOCSY spectrum provides information of the specific spin pattern of residues. Analysis of these spectra together is sometimes enough to perform sequential assignments for small peptides. However, the severe chemical shift overlap due to absence of regular secondary structure precluded the completion of assignments using the 2D data sets [99]. Therefore, we turned to standard triple resonance NMR experiments using ^{13}C, ^{15}N-enriched RPA32N-D8 peptide. The following experiments were acquired: ^{15}N-^{1}H HSQC, ^{13}C-^{1}H HSQC, HNCA, HN(CA)CB, CBCA(CO)NH, and HNCO. Although the 3D data helped to determine assignments, I was able to complete only 50% of sequence specific assignments due to the severe overlap of cross peaks.

Small Angle X-ray Scattering

SAXS data of RPA$_{D9}$ construct was collected at the SIBYLS 12.3.1 beam line at the Advanced Light Source, Lawrence Berkeley National Laboratory. Scattering measurements were performed on 20 μL samples at 15 °C using a Hamilton robot for loading samples from a 96-well plate into a helium-purged sample chamber. Protein samples were purified on a 24 mL Superdex 200 SEC column as a final step of purification and kept at 4 °C until data collection. Data were collected on both the original gel filtration fraction and samples concentrated ~2−8-fold from

individual fractions. Fractions prior to the void volume and concentrator eluates were used for buffer subtraction.

The experiments were collected using an X-ray beam from a multilayer monochromator of 12 keV covering the following momentum transfer range: 0.012 $Å^{-1} < q < 0.317$ $Å^{-1}$, where ($q = 4\pi \sin \Theta/\lambda$, where 2Θ is the scattering angle). Sequential exposures (0.5, 1, 0.5, 5, 0.5, and 10 s) were taken, and data were monitored for radiation-dependent aggregation. All SAXS data were collected using the MarCCD 165 detector in fast frame transfer mode and reduced via normalization to the incident beam intensity. Buffer scattering was subtracted from protein scattering. This was followed by azimuthally averaging to obtain the intensity $I(q)$ versus q scattering plot visualized by xmgrace. The data were analyzed using PRIMUS (Primary Analysis & Manipulations with Small Angle Scattering Data) version 3.0 from ATSAS 2.0 [115], from which Guinier, Kratky, and $P(r)$ plots were generated.

CHAPTER IV

DISCUSSION AND FUTURE DIRECTIONS

Influence of RPA70N on RPA ssDNA binding activity

Studies presented in Chapter II investigate the putative ssDNA binding activity of RPA70N and the conformational freedom of RPA70NAB. RPA70N was found to be structurally independent of RPA70AB alone and in the DNA-bound state. SAXS-derived Dmax and Rg changes in RPA70NAB upon binding of 14mer or 8mer ssDNA were similar, and the magnitude of these changes was similar to those observed for RPA70AB alone. Molecular dynamics simulations were used to model the substantial inter-domain flexibility in RPA70AB and RPA70NAB. These results support a model in which RPA70N is freely available and therefore able to act as a protein recruitment module. This model is supported by studies showing RPA70N is not essential for DNA replication [120] and that RPA70N binds multiple proteins involved in DNA repair activities [32, 39, 41, 125].

The role of flexibly linked domains in RPA function

Binding of ssDNA forms part of the larger cellular goal of conservation of the DNA. A dynamic and continuous assembly and disassembly of DNA processing proteins, like RPA, is necessary throughout every cell cycle. RPA assembles into a heterotrimeric DNA binding core with peripheral flexible domains appended to the core (Figure 4.1). Flexible linkers joining RPA domains provide a basis for its dynamic architecture, and allow RPA remodeling as a function of ssDNA and partner binding proteins [4]. Inter-domain flexibility also plays an important role in the recruitment of proteins by RPA.

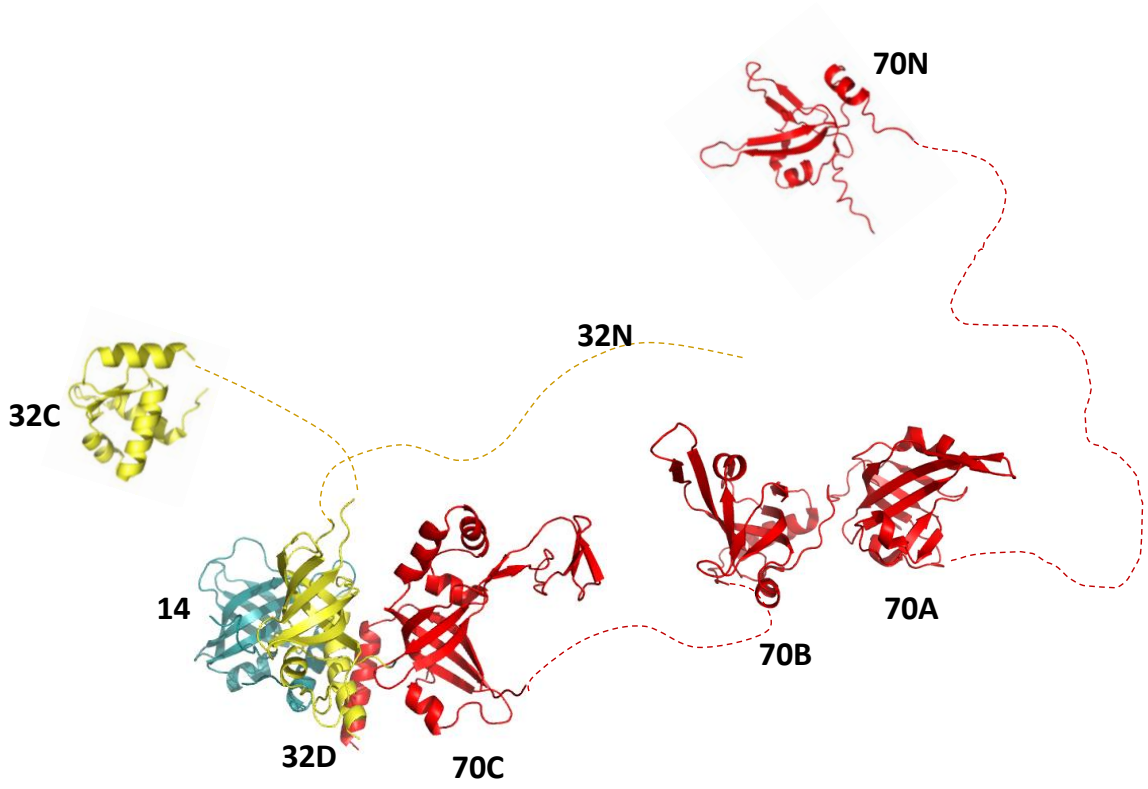


Figure 4.1. RPA is composed of globular domains combined with disordered regions. The DNA binding core is composed of RPA70ABC/32D. Peripheral domains are appended to the core through flexible linkers shown as dotted lines. The disordered RPA32N domain is also shown as a dotted line. Ribbon diagrams were generated from PDB entries (1EWI, 1JMC, 1L1O, 1DPU).

RPA binds proteins involved in all DNA metabolic pathways including replication, recombination, repair and cell cycle checkpoint control. RPA protein interactions are specific. General features include significant electrostatic contributions to binding and common interaction sites, which suggest competition between proteins for binding to RPA. A hand-off mechanism has been proposed to explain how RPA exchanges protein partners and orchestrates the dynamic assembly of DNA processes [26]. RPA70N and RPA32C protein interaction domains are appended to the heterotrimer through long linkers [4]. The RPA70 N-A linker is

especially very long, spanning 60 residues. The RPA32 D-C linker is 30 residues long. The flexible attachment of the RPA DNA and protein binding domains in RPA70NAB was directly characterized in Chapter II. We proposed that the flexibility of the connecting linkers increases the inter-domain dynamics and also allows RPA70N and RPA32C to reach out away from the DNA binding core to recruit partner proteins for effective orchestration of DNA replication and repair.

Similar flexible attachments joining protein and DNA binding domains were recently discovered for DNA-PK and polynucleotide kinase [109, 110]. In fact, inter-domain structural flexibility as experimentally defined here has been proposed to be essential to allow efficient protein handoffs and interface exchanges for FEN1-PCNA [111] and BRCA2-Rad51[112]. RPA binding protein partners, such as the Mre11−Rad50−Nbs1 complex, have similar ordered and flexible domains, as shown by Nbs1 SAXS and crystal structures [113]. Such dynamic character may be a hallmark for scaffold proteins such as RPA. However, the usefulness of the various linker lengths has never been assessed. It can be predicted that shortening of the linkers may have an effect on RPA protein-protein interactions. It would be interesting to investigate the minimum number of residues necessary to mediate functions. RPA70N for instance, interacts with a number of proteins involved in the ATR checkpoint response. It is conceivable that a mutant with a shorter RPA70N-A linker may not effectively recruit these checkpoint proteins and cause in a defective checkpoint response. Alternatively, proteins may be recruited but complexes may not form properly to mediate function. This would indicate that the length of the linker allows strategic placement of proteins by sustaining appropriate intermolecular distances required for complex formation. I believe RPA32C functions similarly, mediating interactions with proteins involved in recombinatorial repair and nucleotide excision repair.

The effectiveness of the linkers can also be assessed through deletion of the linked peripheral domains in the heterotrimer, in the presence of the same peripheral domains in isolation. In addition to loss of direct binding, such experiments could enable assessment of the effect of dilution of the high local concentration of the peripheral domains and increased entropy in the recruitment of protein binding partners. This information could provide a useful way to design RPA mutants that may potentially discriminate specific protein interactions to assay for the effect of specific interactions in replication initiation or in response to DNA damage.

Protein interactions with RPA70AB present a different scenario. These major ssDNA binding domains are presumably bound to ssDNA when protein interactions take place. At least one interaction that overlaps with the RPA70A ssDNA binding site has been identified in recombinatorial repair: Rad51 N-terminus interacts with RPA70A residues that participate in ssDNA binding [25]. Competition with ssDNA for RPA70A site was found to be important in RPA displacement assays leading to RPA disassembly and Rad51 filament formation in the initiation of homologous recombination. Similarly, the SV40 large T-antigen (Tag) helicase domain interaction with the RPA70 A-B linker is important in loading and disassembly of RPA [20]. Interestingly, preliminary studies suggest multiple contact points exist between RPA and Rad51 as has been shown for Tag [20, 21]. However, even for Tag, these networks have yet to be completely characterized. Such studies would refine our understanding of RPA displacement in replication, initiation of homologous recombination, damage response and repair.

RPA structural dynamics

Our understanding of the role of RPA structural dynamics of RPA in ssDNA binding is incomplete. Biochemical studies found that RPA binds ssDNA in multiple modes, which

correlate with different structural and functional states of the protein. The initial ssDNA binding mode involving RPA70AB (binding of 8 nt) has been characterized by crystallography and NMR [7, 45]. We have yet to understand how RPA70C (binding of 12-23 nt, [9, 126]) and RPA32D (binding of 27-30 nt, [9, 126]) engage in ssDNA binding. The crystal structure of RPA70C/32D/14 leads to the hypothesis that RPA70C may align in tandem with RPA70B, and ssDNA may wrap around the trimer core to reach the basic cleft of RPA32D [14]. Intrinsic fluorescence spectroscopy analysis of RPA using various lengths of ssDNA derived a model for ssDNA-RPA complex [15]. Fluorescent probes on the ssDNA identified that RPA70C makes contact with nucleotide 16 and RPA32D with nucleotide 24, when intact RPA is fully engaged. This information was put into context with the assumption that the RPA70C/32D/14 crystal structure model exists in solution. With this caveat, a model of a kinked ssDNA bound to RPA has been proposed [15]. However, RPA remodeling via domain reorganization must be studied to better understand the implications of ssDNA binding. Our understanding of how initiation of replication and protein recruitment takes place is dependent upon this critical aspect of RPA function. Therefore it would be important to characterize the structural basis for these last two RPA ssDNA binding modes.

NMR relaxation analysis is currently in used in the Chazin laboratory as an approach to understand the motions that dominate ssDNA binding to RPA70AB. This approach will be extended to the RPA DNA binding core (RPA70ABC/32D/14). Neutron scattering contrast variation experiments in combination with SAXS will extract basic scattering functions for the RPA-DBC and ssDNA for three ssDNA substrates of eight, twenty and thirty bases long. Observations of changes in vector distribution or P(r) functions, R_g and D_{max} will be translated into structural models for RPA at every transition step of ssDNA binding. These studies will

also provide a guide for analysis of scattering data for the full length RPA heterotrimer, including the hyperphosphorylation mutant as discussed in Chapter III.

Significant architectural heterogeneity much like that reported for RPA70NAB in Chapter II, is expected in full length RPA (Figure 4.2). In fact, preliminary SAXS data on a mimic of hyperphosphorylated RPA (RPA$_{D9}$), provides evidence of RPA's intrinsic flexibility. NMR spectra collected on the full length RPA heterotrimer show that in the free and ssDNA bound forms, RPA70N and RPA32C domains tumble independently of the core in solution. Even RPA70A and RPA70B domains remain flexible enough to tumble independently from the rest of the heterotrimer when engaged in binding ssDNA [4]. Therefore, remodeling upon

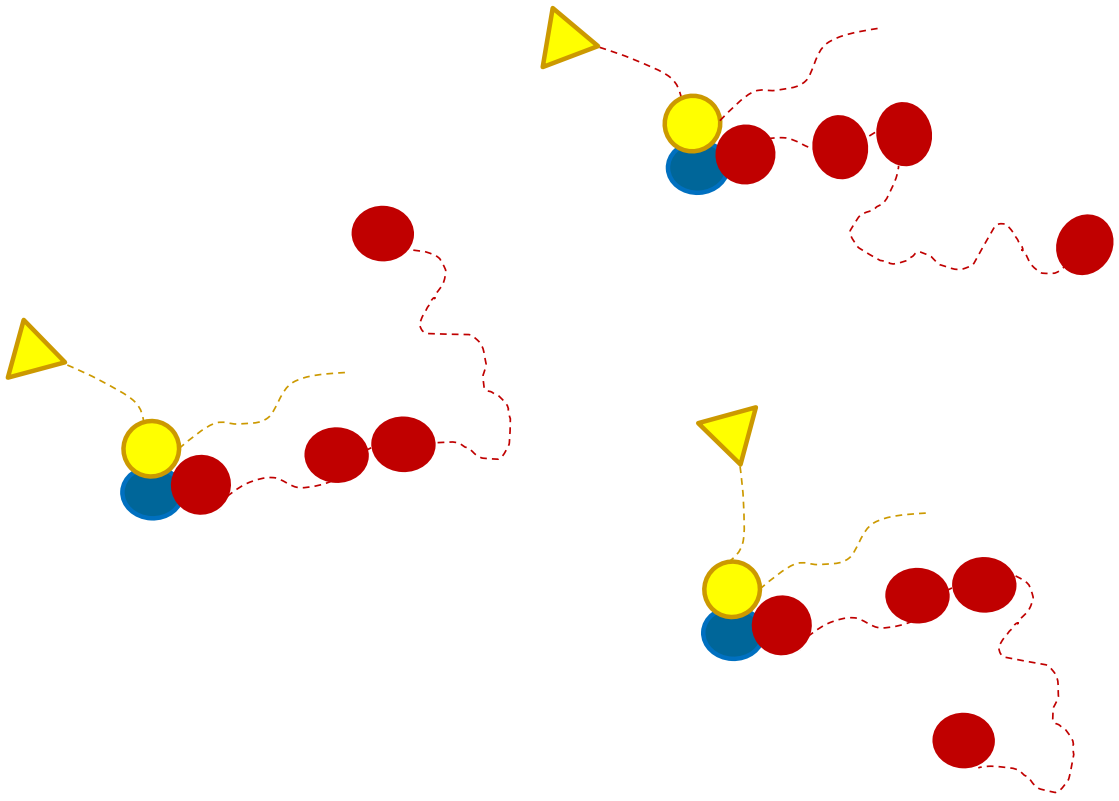


Figure 4.2. Structural dynamics of RPA. The flexible linkers in RPA allow formation of many different architectures. Subunits of RPA are color coded with RPA70 in red, RPA32 in yellow and RPA14 in blue.

ssDNA binding may directly affect protein-protein interactions with proteins involved in the initiation of replication. However, it should be noted that many critical RPA interactions presumably occur after RPA is loaded onto ssDNA.

Role of protein dynamics in DNA processing

The key role of dynamics in facilitating the organization and progression of large multiprotein machines is increasingly recognized, particularly for DNA replication and repair machinery that requires precise coordination to efficiently preserve genome integrity. The role of dynamics in facilitating recruitment, organization, and exchange of DNA processing factors has been characterized in several model systems, most notably in a recent study of homotetrameric *E. coli* SSB diffusion dynamics along ssDNA [114]. In that study, SSB diffusion was shown to be critical for resolving DNA secondary structures to enable RecA filament formation. Unlike the modular, multidomain RPA, the homotetrameric single-domain SSB does not utilize preexisting structural dynamics to facilitate organizing strands of ssDNA for DNA processing. Instead, the compact, globular SSB homotetramer is encircled by the ssDNA and is thought to "roll" along the template via a consecutive unwrapping and wrapping of ssDNA. Thus, while the structural organization of these two SSB systems remains fundamentally different, dynamic motion would appear to be integral aspects of both. Specifically, the nature of the structural dynamics of linked, ordered, and flexible RPA domains as identified here appears to be critical to the accommodation of the large-scale complex conformational changes proposed to regulate RPA-related functions, while preserving the integrity of DNA and protein partner interactions for maintaining genetic fidelity.

RPA32N interactions within RPA

RPA32N harbors all RPA phosphorylation sites shown so far to be functionally relevant [68]. NMR and CD spectra demonstrated that this region is unstructured. In addition, I showed that a pseudo-hyperphosphorylated RPA32N peptide can bind to RPA70NAB, and that this

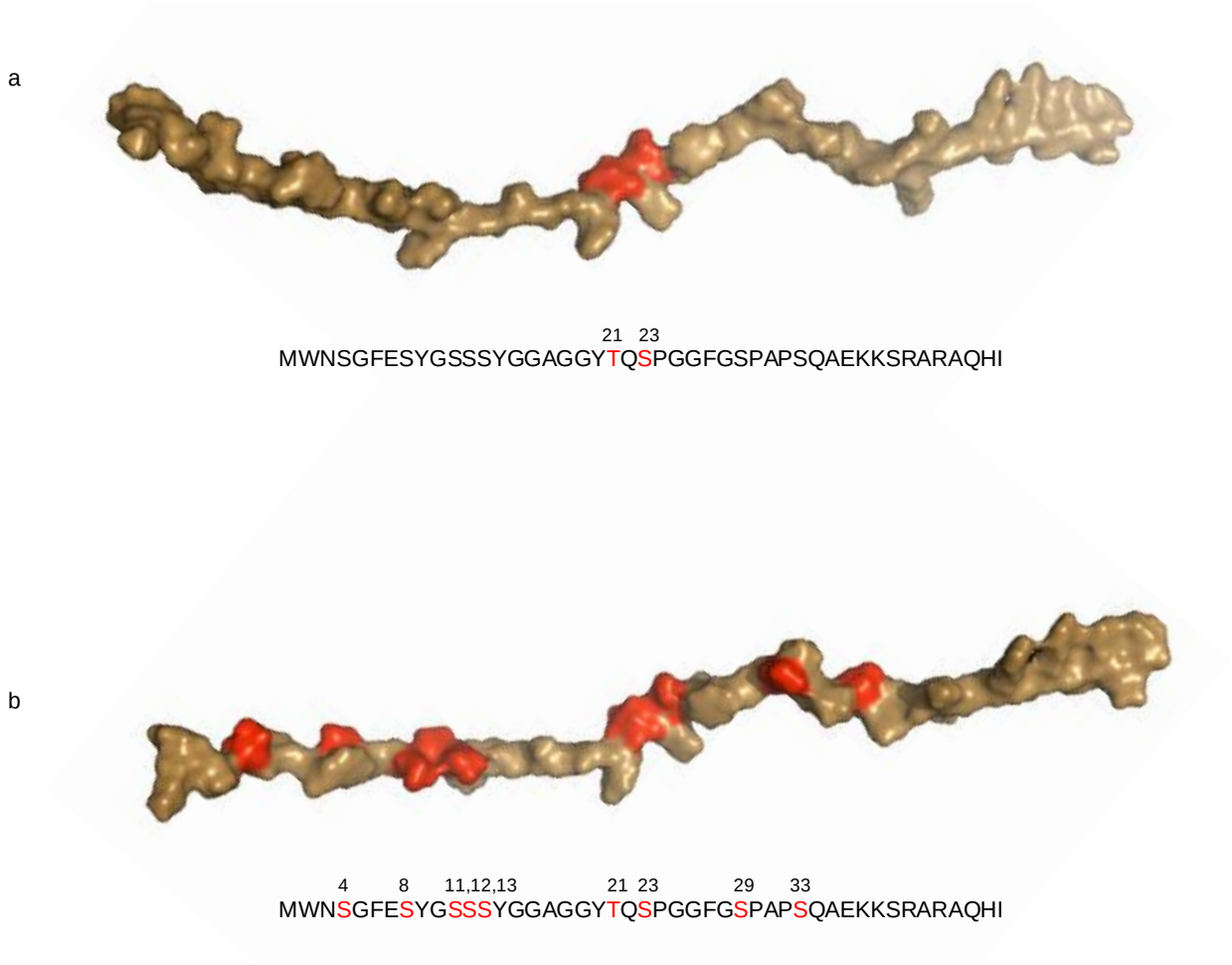


Figure 4.3. Phosphorylated RPA32N peptide. A surface representation of the hypophosphorylated (a) and hyperphosphorylated (b) RPA32N peptide is shown in an arbitrary extended conformation. Residues that are phosphorylated are highlighted in red.

interaction is mediated through residues that are normally involved in protein-protein interactions (RPA70N and RPA70A) and ssDNA binding (RPA70AB). These interactions are consistent with a role of RPA32N in regulating RPA function via phosphorylation.

In the cell, the number of phosphorylated residues at any given time varies, with increased phosphorylation of RPA32N in response to DNA damage. Phosphorylation of multiple residues dramatically changes the electrostatic properties of RPA32N (Figure 4.3) and this correlates with observations that phosphorylation of RPA32N affects RPA function. It is well established that phosphorylation activates or deactivates many proteins causing or preventing the progression of cellular processes. Proteins that become activated via phosphorylation may interact with new partners. So far, interactions with RPA32N have only been studied within RPA. Therefore, other than inter-subunit interactions, no specific protein-protein interactions have been identified involving this domain. One obvious expectation is interaction with kinases that phosphorylate RPA32N. However, it would be logical to investigate if proteins involved in DNA damage response also interact specifically with the hyperphosphorylated RPA32N.

The role of hyperphosphorylation in the function of RPA

Phosphorylation of RPA32N perturbs RPA interactions with replication proteins, and hyperphosphorylated RPA is not active in replication [60, 68]. It can be expected that inter-subunit interactions form part of a phosphorylation dependent regulation of RPA function. The standing hypothesis in the literature is that hyperphosphorylation causes a global conformational change in RPA70AB, affecting its ssDNA binding activity, and that this is regulated by RPA70N preferred binding to the hyperphosphorylated RPA32N [3]. This hypothesis was built on NMR evidence for a specific intersubunit interaction between the RPA70N and an RPA32N phosphomimetic peptide [44], and MALDI MS fingerprinting evidence of conformational changes involving RPA70B in DNA PK *in vitro* phosphorylated RPA [119]. However, the

interaction between RPA70N and RPA32N phosphomimic peptide is very weak, in the high μM range. Also, the observation that a large percentage of the RPA70 subunit in the *in vitro* phosphorylated RPA was partially proteolyzed prompted the suggestion that RPA70N was cleaved in the MS fingerprinting studies.

Intrinsic fluorescence studies in this protein showed decreased interactions with short ssDNA [119]. This information was correlated with analysis of a pseudo-phosphorylated RPA mutant harboring a deletion of the RPA70N that showed reduced binding to short ssDNA and no detectable DNA helix destabilization activity [3]. An idea was set forth that changes in RPA70AB only occur when RPA70N cannot exert a protective function.

Hyperphosphorylation and ssDNA binding activity

The data presented here shows that RPA32N intersubunit interactions occur mainly via basic residues in the clefts of RPA70NAB OB-fold domains. RPA32N-D8 interacting residues in RPA70AB participate in ssDNA binding, although estimates of the affinity indicate binding in the 100-1000 μM regime. In contrast, RPA70AB binds ssDNA with Kd ~100 nM [7]. Thus, an effect due to direct competition between a hyperphosphorylated RPA32N and ssDNA for RPA70AB binding site seems unlikely. This is consistent with evidence in the literature suggesting that hyperphosphorylation does not affect RPA ssDNA binding: surface plasmon resonance experiments indicated that hyperphosphorylation does not change RPA affinity for ssDNA [44]. Similarly, fluorescence studies demonstrate that hypophosphorylated RPA has a similar affinity for ssDNA as the wild type protein [60]. Thus, the level of phosporylation is not expected to affect ssDNA binding. I also note that hyperphosphorylation of RPA occurs in its ssDNA bound state. The persistence of foci at sites of DNA damage suggests the

hyperphosphorylated RPA remains bound to ssDNA [68]. This is consistent with ssDNA binding activity of RPA being comparable to that of wild type RPA.

Although RPA32N does not appear to directly disrupt RPA binding to ssDNA, studies investigating direct competition studies between an RPA70NAB-ssDNA complex and hyperphosphorylated RPA32N would confirm this prediction. However, one caveat is such studies examine RPA in isolation, while DNA processes occur in presence of many different factors that could be important in modulating the effect of hyperphosphorylation of RPA on binding ssDNA. Importantly, the observed weak inter-subunit interactions I observed by NMR may be relevant for RPA function in the cellular context. The high local concentration of hyperphosphorylated RPA32N resulting from its linkage to the heterotrimer should increase the frequency of interactions and could therefore affect the overall structural dynamics of RPA.

The effect of hyperphosphorylation on RPA: The compaction model

Since hyperphosphorylation does not mediate direct changes in DNA binding activity, the question of how does RPA32N hyperphosphorylation influences RPA function remains obscure. Based on the physical interaction observed between RPA32N and RPA70NAB, we favor a model of RPA compaction upon hyperphosphorylation (Figure 4.4). In this model, RPA70N is on average closer to RPA70AB than in the non-phosphorylated protein. Such a scheme does not alter the ssDNA mechanism because RPA70N is not involved in ssDNA binding. It does however provide the advantage of close localization of these domains, which may facilitate the exchange of protein interactions.

We also considered the possibility of a more extended RPA model in which RPA70N is extended away from RPA70AB. In this model, intersubunit and intermolecular protein-protein

interactions can be shared with the neighboring RPA molecule (Figure 4.5). Such a model has been proposed to explain the formation of ATR-ATRIP complex with Rad9 and TOPBP1, each of which have been shown to bind to RPA70N [41]. Insight into the validity of this model is available from analysis of wt and RPA hyperphosphorylation mutants by SEC. Estimates of the Stokes Radius of RPA species from SEC shows that the stokes radius of an RPA32N mutant that contained aspartic acid mutations to mimic hyperphosphorylation is very similar to the wild type RPA [3]. If the model in Figure 4.5 were accurate, one would anticipate RPA molecules would exhibit at least some tendency to aggregate, which would be reflected in a significant increase in Stokes Radius. Although the SEC observations do not support the extended model, they do not conclusively refute it.

Electron microscopy characterization of RPA ssDNA bound forms identified RPA complexes aligned along the ssDNA in either a contracted or an extended form that was dependent on ssDNA length [118]. Importantly, phosphorylation of RPA with DNA PK appeared to stimulate formation of extended RPA-30mer complexes. However, imaging of RPA

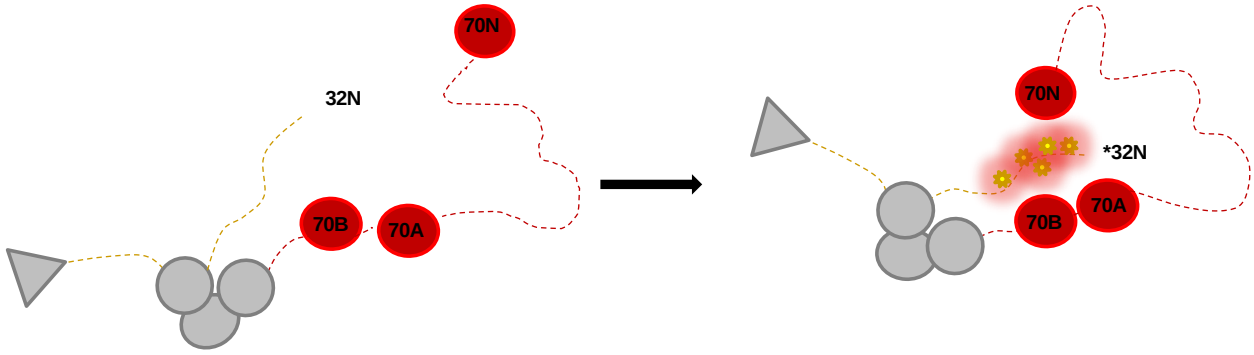


Figure 4.4. Model for compaction upon hyperphosphorylation. RPA70 domains N, A and B are shown in red. Highlighted stars represent hyperphosphorylation on RPA32N. Hyperphosphorylation causes a general compaction of RPA that might affect the overall protein structural dynamics and protein interaction network. ssDNA and potential interacting proteins have been left out of this illustration for simplicity.

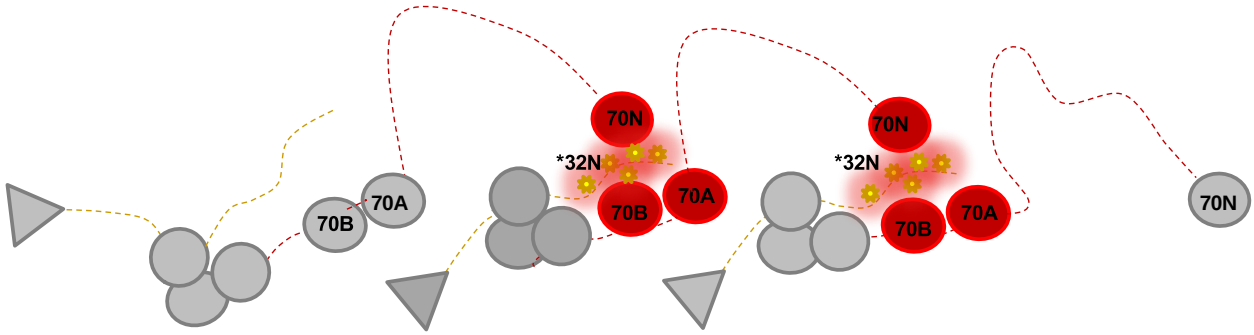


Figure 4.5. Hyperphosphorylation model of extended RPA. RPA70 domains N, A and B are shown in red. Highlighted stars represent hyperphosphorylation on RPA32N. This alternative extended model implicates interaction between neighboring RPA molecules. ssDNA and potential interacting proteins have been left out of this illustration for simplicity.

in these electro micrographs is dependent on the density of ssDNA. This suggests that only density from RPA ssDNA binding core domains is highlighted, and the density of the flexibly linked domains may not contribute to their observations. This means that phosphorylation of RPA takes place on the full assembled (30 nt bound) protein. The results from this study reflect changes in population of binding modes and is a good indication that RPA can exist in more than one conformation. However, there is no evidence that RPA oligomerizes [118]. In summary, although both models are consistent with phosphorylation not affecting ssDNA binding and are able to satisfy all interactions involving RPA32N, the global compaction model is better supported by the available data and is therefore more likely.

Given the lack of data, it will be necessary to gain a better understanding of hyperphosphorylation and the role of intersubunit interactions by directly characterizing the overall changes in the intact RPA heterotrimer. One potential approach for this is SAXS, since the distribution of interatomic distances and maximum dimensions of the wild-type versus hyperphosphorylated RPA proteins can be compared. As shown in Chapter III, a full-length hyperphosphorylation mutant constructed in collaboration with SBDR Molecular Biology Core

at Lawrence Berkley National Laboratory provided promising preliminary SAXS data. Thus, there are excellent prospects for directly testing the validity of the compaction and oligomerization models for the effect of hyperphosphorylation on RPA.

The compaction model and RPA hyperphosphorylation in replication, recombinatorial repair, and checkpoint signaling

The functional switch associated with hyperphosphorylation of RPA may result from change in the ability of RPA to unload and reload on ssDNA. The rearrangement and compaction resulting from RPA32N hyperphosphorylation would prevent RPA from being re-loaded on ssDNA for replication by simply preventing or weakening selected RPA protein interactions with RPA-loading proteins. However, I note that there is no evidence that RPA loading by the SV40 large Tag helicase is disrupted by RPA32N hyperphosphorylation.

Interestingly, RPA32C interacts with an acidic surface on Tag-OBD, analogous to the acidic hyperphosphorylated RPA32N. However, RPA32C does not bind RPA32N-D8. It would be interesting to test if hyperphosphorylated RPA32N could bind Tag OBD. If this were the case, the hyperphosphorylated peptide could compete with RPA32C for Tag OBD, affecting the stimulation of primer synthesis perhaps by interfering with Tag interaction with Pol α. This would be consistent with studies showing that phosphorylation of RPA affects its interaction with DNA pol α [60]. Tag OBD binds origin DNA with greater affinity than RPA32C. Although hyperphosphorylation of RPA32N mimics ssDNA, it is unlikely that Tag OBD-origin DNA interaction could be impaired by hyperphosphorylation since it occurs after initiation of replication. These speculations clearly require further studies.

In humans, the MCM2-7/CDC45/GINS complex is the active helicase in DNA replication. Whether MCM2-7 loads RPA analogous to what has been shown with Tag helicase and if this activity may be impaired by hyperphosphorylation of RPA is unknown. Interestingly, it was found that depletion of CDC45, a helicase cofactor, suppresses RPA hyperphosphorylation and foci formation. The same is true in CHK1 depleted cells exposed to HU or thymidine, ultimately suppressing apoptosis [89]. One explanation could be that RPA loading on ssDNA was precluded, therefore hyperphosphorylation was impaired. In *Xenopus* egg extracts, functional uncoupling of helicase and polymerase activities in response to aphidicolin is dependent on MCM helicase [127]. This uncoupling causes the helicase to continue to unwind DNA that results in increased generation of ssDNA and RPA foci formation, leading to amplification of CHK1 signaling. However, uncoupling was abrogated when MCM7 or CDC45 were inhibited [128]. Moreover, inactivation of MCM helicase and DNA pol α activity abrogates CHK1 signaling [128] . Together, this evidence also indicates that helicase complexes may be important for RPA loading, which is a pre-requisite for hyperphosphorylation. However, experiments directly addressing this question must be performed before any conclusions can be drawn.

Although the ssDNA binding capacity of RPA70AB may not be significantly affected by hyperphosphorylation, RPA binding to ssDNA may be indirectly affected. One idea is that hyperphosphorylation could indirectly promote the removal of RPA from ssDNA. A simple explanation of how this may happen is that hyperphosphorylation promotes interactions with recombination and repair proteins. For example, RPA is displaced from exposed ssDNA ends by Rad51, with help of Rad52. Rad51 N-terminal domain interaction with RPA70A occurs at the same site for RPA32N and ssDNA binding. In fact, it has been shown that ssDNA and Rad51N

compete for RPA70A binding site. Although Rad51N affinity for ssDNA is weak and it cannot displace ssDNA, mutagenesis analysis showed the importance of this contact for efficient Rad51 mediated displacement of RPA [25]. Likewise, weak RPA intersubunit interactions in hyperphosphorylation may indeed have a subtle effect in RPA activity. And it remains to be investigated why Rad51 and Rad52 efficiency of interaction with RPA is enhanced in the presence of hyperphosphorylated RPA32N [129]. Rad51 and Rad52 are DNA binding proteins and hyperphosphorylation of RPA32N mimics ssDNA. It is possible that hyperphosphorylated RPA32N mediates an additional interaction point that provides an advantage in protein recruitment, or a contribution to a synergistic binding effect through multiple weak interactions. These properties may contribute to the displacement of RPA and Rad51 filament formation. In order to explore this possibility, it would be necessary to perform direct protein-protein interaction studies and displacement assays in the presence of hyperphosphorylated RPA.

Most recently, our laboratory has found that Dss1, a 70 residue protein involved in recombinatorial repair activities through modulation of BRCA2 also interacts simultaneously with the RPA70N, RPA70A and RPA70B domains (Dr. Sivaraja Vaithiyalingham unpublished results). Because Dss1 is so small, simultaneous interaction of all three RPA domains with DSS1 must result in compaction of RPA, suggesting that an RPA compaction mechanism as proposed here may also be utilized in other contexts of RPA function.

It has been shown that RPA recruits checkpoint proteins to sites of DNA damage [41]. Although hyperphosphorylation of RPA has not been considered in the recruitment of checkpoint proteins under DNA damage conditions, a vast number of research articles have shown that treatment of cells with UV or IR causes the hyperphosphorylation of RPA, which is the predominant form found at DNA damage sites. Hyperphosphorylated RPA32N specifically

recruits checkpoint proteins in close proximity to the damaged DNA region. In support of this idea, RPA70N domain has been shown to bind to various proteins involved in cell cycle control and DNA damage response, such as ATRIP, Rad9, MRE11 and p53 [32, 39-41]. It has been shown that RPA bound to ssDNA is required for high levels of ATRIP to accumulate at sites of DNA damage [39]. ATRIP and ATR both bind RPA: ATR can bind to RPA32 independent from ATRIP and ATRIP can bind RPA70N independent of ATR. After irradiation of HCT116 cells with 50 J/m^2 of UV, or HeLa cells with 10 Gy of IR, the ATR-ATRIP complex colocalizes with RPA-covered ssDNA [90]. Thus, persistent localization of hyperphosphorylated RPA at DNA damage sites would result in accumulation of the ATR-ATRIP complex and other repair factors. One interesting observation is that ATR/ATRIP phosphorylation of CHK1 is dependent on RPA [90]. A reduction of CHK1 phosphorylation was observed upon HU or UV treatment in HeLa cells, or UV treatment in U2OS cells transfected with siRNA to RPA70 (absence of RPA70 subunit precludes the assembly of the RPA heterotrimer). A greater reduction was observed at higher doses of UV, which would normally correlate with higher levels of hyperphosphorylation of RPA. Because the kinases respond to DNA damage, it is possible that hyperphosphorylation of RPA during DNA damage may contribute to efficient checkpoint complex formation promoting CHK1 phosphorylation and proficient DNA damage signaling.

Significance

The focus of the work presented in this thesis is the influence of peripheral domains RPA32N and RPA70N on the organization and DNA binding activity of Replication Protein A. This research contributes observations on RPA domain dynamics and intersubunit interactions. Consideration of these findings must be taken into account in defining the mechanism of RPA in

protein recruitment and response to DNA damage. Although we and others have gained many insights from studying the isolated RPA protein and domain fragments, the processing of ssDNA occurs through the combined effect of multiple full-length proteins in complexes. Protein structural dynamics are influenced by the interaction with their partners and actions leading to function will not be completely understood until the complete machinery is characterized. Developing the capability to work with intact proteins and their complexes will be most important in discovering the nuances of DNA processing events.

RPA is a critical component for the progression of DNA processing events essential for eukaryotic cell survival and reproduction. Elucidating the actions of RPA in the protection of ssDNA and orchestration of dynamic multi-protein machines is critical for understanding how DNA processing is coordinated in the cell. My research has expanded the knowledge of the dynamic RPA architecture and explored the effects of phosphorylation of RPA32. Protection of the ssDNA intermediate by RPA is fundamental to all DNA processing events, and the phosphorylation of RPA plays a central role in the switch of replication activity to repair when DNA damage occurs. The work presented here contributes to basic understanding of how RPA functions and its ability to integrate the many DNA transactions essential for life.

REFERENCES

1. Bochkarev, A. and E. Bochkareva, *From RPA to BRCA2: lessons from single-stranded DNA binding by the OB-fold.* Curr Opin Struct Biol, 2004. **14**(1): p. 36-42.

2. Mer, G., et al., *Three-dimensional structure and function of replication protein A.* Cold Spring Harb Symp Quant Biol, 2000. **65**: p. 193-200.

3. Binz, S.K. and M.S. Wold, *Regulatory functions of the N-terminal domain of the 70-kDa subunit of replication protein A (RPA).* J Biol Chem, 2008. **283**(31): p. 21559-70.

4. Brosey, C.A., et al., *NMR analysis of the architecture and functional remodeling of a modular multidomain protein, RPA.* J Am Chem Soc, 2009. **131**(18): p. 6346-7.

5. Bhattacharya, S., et al., *Characterization of binding-induced changes in dynamics suggests a model for sequence-nonspecific binding of ssDNA by replication protein A.* Protein Sci, 2002. **11**(10): p. 2316-25.

6. Iftode, C. and J.A. Borowiec, *5' --> 3' molecular polarity of human replication protein A (hRPA) binding to pseudo-origin DNA substrates.* Biochemistry, 2000. **39**(39): p. 11970-81.

7. Arunkumar, A.I., et al., *Independent and coordinated functions of replication protein A tandem high affinity single-stranded DNA binding domains.* J Biol Chem, 2003. **278**(42): p. 41077-82.

8. Kim, C., R.O. Snyder, and M.S. Wold, *Binding properties of replication protein A from human and yeast cells.* Mol Cell Biol, 1992. **12**(7): p. 3050-9.

9. 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.

10. Blackwell, L.J. and J.A. Borowiec, *Human replication protein A binds single-stranded DNA in two distinct complexes.* Mol Cell Biol, 1994. **14**(6): p. 3993-4001.

11. 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.

12. Bochkarev, A., et al., *The crystal structure of the complex of replication protein A subunits RPA32 and RPA14 reveals a mechanism for single-stranded DNA binding.* Embo J, 1999. **18**(16): p. 4498-504.

13. 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.

14. 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.

15. Cai, L., et al., *Structural characterization of human RPA sequential binding to single-stranded DNA using ssDNA as a molecular ruler.* Biochemistry, 2007. **46**(28): p. 8226-33.

16. Wold, M.S. and T. Kelly, *Purification and characterization of replication protein A, a cellular protein required for in vitro replication of simian virus 40 DNA.* Proc Natl Acad Sci U S A, 1988. **85**(8): p. 2523-7.

17. Wold, M.S., et al., *Identification of cellular proteins required for simian virus 40 DNA replication.* J Biol Chem, 1989. **264**(5): p. 2801-9.

18. Li, J.J. and T.J. Kelly, *Simian virus 40 DNA replication in vitro.* Proc Natl Acad Sci U S A, 1984. **81**(22): p. 6973-7.

19. Brush, G.S., T.J. Kelly, and B. Stillman, *Identification of eukaryotic DNA replication proteins using simian virus 40 in vitro replication system.* Methods Enzymol, 1995. **262**: p. 522-48.

20. Jiang, X., et al., *Structural mechanism of RPA loading on DNA during activation of a simple pre-replication complex.* EMBO J, 2006. **25**(23): p. 5516-26.

21. Arunkumar, A.I., et al., *Insights into hRPA32 C-terminal domain--mediated assembly of the simian virus 40 replisome.* Nat Struct Mol Biol, 2005. **12**(4): p. 332-9.

22. Weisshart, K., P. Taneja, and E. Fanning, *The replication protein A binding site in simian virus 40 (SV40) T antigen and its role in the initial steps of SV40 DNA replication.* J Virol, 1998. **72**(12): p. 9771-81.

23. Loo, Y.M. and T. Melendy, *Recruitment of replication protein A by the papillomavirus E1 protein and modulation by single-stranded DNA.* J Virol, 2004. **78**(4): p. 1605-15.

24. 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.

25. Stauffer, M.E. and W.J. Chazin, *Physical interaction between replication protein A and Rad51 promotes exchange on single-stranded DNA.* J Biol Chem, 2004. **279**(24): p. 25638-45.

26. 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.

27. New, J.H., et al., *Rad52 protein stimulates DNA strand exchange by Rad51 and replication protein A.* Nature, 1998. **391**(6665): p. 407-10.

28. 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.

29. Akbari, M., et al., *Repair of U/G and U/A in DNA by UNG2-associated repair complexes takes place predominantly by short-patch repair both in proliferating and growth-arrested cells.* Nucleic Acids Res, 2004. **32**(18): p. 5486-98.

30. He, Z., et al., *The transactivator proteins VP16 and GAL4 bind replication factor A.* Cell, 1993. **73**(6): p. 1223-32.

31. Dutta, A., et al., *Inhibition of DNA replication factor RPA by p53.* Nature, 1993. **365**(6441): p. 79-82.

32. Bochkareva, E., et al., *Single-stranded DNA mimicry in the p53 transactivation domain interaction with replication protein A.* Proc Natl Acad Sci U S A, 2005. **102**(43): p. 15412-7.

33. Sommers, J.A., et al., *p53 modulates RPA-dependent and RPA-independent WRN helicase activity.* Cancer Res, 2005. **65**(4): p. 1223-33.

34. Shen, J.C., et al., *The N-terminal domain of the large subunit of human replication protein A binds to Werner syndrome protein and stimulates helicase activity.* Mech Ageing Dev, 2003. **124**(8-9): p. 921-30.

35. Chen, C.Y., J. Graham, and H. Yan, *Evidence for a replication function of FFA-1, the Xenopus orthologue of Werner syndrome protein.* J Cell Biol, 2001. **152**(5): p. 985-96.

36. Yan, H. and J. Newport, *FFA-1, a protein that promotes the formation of replication centers within nuclei.* Science, 1995. **269**(5232): p. 1883-5.

37. Brosh, R.M., Jr., et al., *Replication protein A physically interacts with the Bloom's syndrome protein and stimulates its helicase activity.* J Biol Chem, 2000. **275**(31): p. 23500-8.

38. Sowd, G., et al., *Replication protein A stimulates the Werner syndrome protein branch migration activity.* J Biol Chem, 2009. **284**(50): p. 34682-91.

39. Ball, H.L., J.S. Myers, and D. Cortez, *ATRIP binding to replication protein A-single-stranded DNA promotes ATR-ATRIP localization but is dispensable for Chk1 phosphorylation.* Mol Biol Cell, 2005. **16**(5): p. 2372-81.

40. Oakley, G.G., et al., *Physical interaction between replication protein A (RPA) and MRN: involvement of RPA2 phosphorylation and the N-terminus of RPA1.* Biochemistry, 2009. **48**(31): p. 7473-81.

41. Xu, X., et al., *The basic cleft of RPA70N binds multiple checkpoint proteins, including RAD9, to regulate ATR signaling.* Mol Cell Biol, 2008. **28**(24): p. 7345-53.

42. 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.

43. Daughdrill, G.W., et al., *The weak interdomain coupling observed in the 70 kDa subunit of human replication protein A is unaffected by ssDNA binding.* Nucleic Acids Res, 2001. **29**(15): p. 3270-6.

44. 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.

45. 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.

46. Iftode, C. and J.A. Borowiec, *Unwinding of origin-specific structures by human replication protein A occurs in a two-step process.* Nucleic Acids Res, 1998. **26**(24): p. 5636-43.

47. Georgaki, A., et al., *DNA unwinding activity of replication protein A.* FEBS Lett, 1992. **308**(3): p. 240-4.

48. Georgaki, A. and U. Hubscher, *DNA unwinding by replication protein A is a property of the 70 kDa subunit and is facilitated by phosphorylation of the 32 kDa subunit.* Nucleic Acids Res, 1993. **21**(16): p. 3659-65.

49. Helmbrecht, K., E. Zeise, and L. Rensing, *Chaperones in cell cycle regulation and mitogenic signal transduction: a review.* Cell Prolif, 2000. **33**(6): p. 341-65.

50. Gabrielli, B.G., et al., *Hyperphosphorylation of the N-terminal domain of Cdc25 regulates activity toward cyclin B1/Cdc2 but not cyclin A/Cdk2.* J Biol Chem, 1997. **272**(45): p. 28607-14.

51. Roberts, J.M. and G. D'Urso, *An origin unwinding activity regulates initiation of DNA replication during mammalian cell cycle.* Science, 1988. **241**(4872): p. 1486-9.

52. 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.

53. 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.

54. 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.

55. Nuss, J.E., et al., *DNA damage induced hyperphosphorylation of replication protein A. 1. Identification of novel sites of phosphorylation in response to DNA damage.* Biochemistry, 2005. **44**(23): p. 8428-37.

56. 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.

57. 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.

58. 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.

59. Fotedar, R. and J.M. Roberts, *Cell cycle regulated phosphorylation of RPA-32 occurs within the replication initiation complex.* EMBO J, 1992. **11**(6): p. 2177-87.

60. Oakley, G.G., et al., *RPA phosphorylation in mitosis alters DNA binding and protein-protein interactions.* Biochemistry, 2003. **42**(11): p. 3255-64.

61. Patrick, S.M. and J.J. Turchi, *Replication protein A (RPA) binding to duplex cisplatin-damaged DNA is mediated through the generation of single-stranded DNA.* J Biol Chem, 1999. **274**(21): p. 14972-8.

62. Erdile, L.F., M.S. Wold, and T.J. Kelly, *The primary structure of the 32-kDa subunit of human replication protein A.* J Biol Chem, 1990. **265**(6): p. 3177-82.

63. 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.

64. Kaldis, P. and E. Aleem, *Cell cycle sibling rivalry: Cdc2 vs. Cdk2.* Cell Cycle, 2005. **4**(11): p. 1491-4.

65. Pan, Z.Q., A. Amin, and J. Hurwitz, *Characterization of the in vitro reconstituted cyclin A or B1-dependent cdk2 and cdc2 kinase activities.* J Biol Chem, 1993. **268**(27): p. 20443-51.

66. Fang, F. and J.W. Newport, *Evidence that the G1-S and G2-M transitions are controlled by different cdc2 proteins in higher eukaryotes.* Cell, 1991. **66**(4): p. 731-42.

67. 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.

68. 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.

69. Wu, X., et al., *Preferential localization of hyperphosphorylated replication protein A to double-strand break repair and checkpoint complexes upon DNA damage.* Biochem J, 2005. **391**(Pt 3): p. 473-80.

70. Feng, J., et al., *Protein phosphatase 2A-dependent dephosphorylation of replication protein A is required for the repair of DNA breaks induced by replication stress*. Mol Cell Biol, 2009. **29**(21): p. 5696-709.

71. Lee, D.H., et al., *A PP4 phosphatase complex dephosphorylates RPA2 to facilitate DNA repair via homologous recombination.* Nat Struct Mol Biol. **17**(3): p. 365-72.

72. Pommier, Y., et al., *Repair of topoisomerase I-mediated DNA damage*. Prog Nucleic Acid Res Mol Biol, 2006. **81**: p. 179-229.

73. Cortez, D., *Unwind and slow down: checkpoint activation by helicase and polymerase uncoupling*. Genes Dev, 2005. **19**(9): p. 1007-12.

74. Matsuoka, S., et al., *ATM and ATR substrate analysis reveals extensive protein networks responsive to DNA damage.* Science, 2007. **316**(5828): p. 1160-6.

75. Harper, J.W., et al., *The p21 Cdk-interacting protein Cip1 is a potent inhibitor of G1 cyclin-dependent kinases.* Cell, 1993. **75**(4): p. 805-16.

76. Brush, G.S., C.W. Anderson, and T.J. Kelly, *The DNA-activated protein kinase is required for the phosphorylation of replication protein A during simian virus 40 DNA replication.* Proc Natl Acad Sci U S A, 1994. **91**(26): p. 12520-4.

77. 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.

78. 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.

79. Park, J.S., et al., *Involvement of DNA-dependent protein kinase in UV-induced replication arrest.* J Biol Chem, 1999. **274**(45): p. 32520-7.

80. Pandita, T.K., et al., *Ionizing radiation activates the ATM kinase throughout the cell cycle.* Oncogene, 2000. **19**(11): p. 1386-91.

81. Gately, D.P., et al., *Characterization of ATM expression, localization, and associated DNA-dependent protein kinase activity.* Mol Biol Cell, 1998. **9**(9): p. 2361-74.

82. Oakley, G.G., et al., *UV-induced hyperphosphorylation of replication protein a depends on DNA replication and expression of ATM protein.* Mol Biol Cell, 2001. **12**(5): p. 1199-213.

83. Chan, D.W., et al., *Purification and characterization of ATM from human placenta. A manganese-dependent, wortmannin-sensitive serine/threonine protein kinase.* J Biol Chem, 2000. **275**(11): p. 7803-10.

84. Shao, R.G., et al., *Replication-mediated DNA damage by camptothecin induces phosphorylation of RPA by DNA-dependent protein kinase and dissociates RPA:DNA-PK complexes.* Embo J, 1999. **18**(5): p. 1397-406.

85. Brush, G.S., et al., *The ATM homologue MEC1 is required for phosphorylation of replication protein A in yeast.* Proc Natl Acad Sci U S A, 1996. **93**(26): p. 15075-80.

86. Brush, G.S. and T.J. Kelly, *Phosphorylation of the replication protein A large subunit in the Saccharomyces cerevisiae checkpoint response.* Nucleic Acids Res, 2000. **28**(19): p. 3725-32.

87. Olson, E., et al., *RPA2 is a direct downstream target for ATR to regulate the S-phase checkpoint.* J Biol Chem, 2006. **281**(51): p. 39517-33.

88. Syljuasen, R.G., et al., *Inhibition of human Chk1 causes increased initiation of DNA replication, phosphorylation of ATR targets, and DNA breakage.* Mol Cell Biol, 2005. **25**(9): p. 3553-62.

89. Rodriguez, R., M.E. Gagou, and M. Meuth, *Apoptosis induced by replication inhibitors in Chk1-depleted cells is dependent upon the helicase cofactor Cdc45.* Cell Death Differ, 2008. **15**(5): p. 889-98.

90. Zou, L. and S.J. Elledge, *Sensing DNA damage through ATRIP recognition of RPA-ssDNA complexes.* Science, 2003. **300**(5625): p. 1542-8.

91. Sleeth, K.M., et al., *RPA mediates recombination repair during replication stress and is displaced from DNA by checkpoint signalling in human cells.* J Mol Biol, 2007. **373**(1): p. 38-47.

92. Putnam, C.D., et al., *X-ray solution scattering (SAXS) combined with crystallography and computation: defining accurate macromolecular structures, conformations and assemblies in solution.* Q Rev Biophys, 2007. **40**(3): p. 191-285.

93. Rambo, R.P. and J.A. Tainer, *Bridging the solution divide: comprehensive structural analyses of dynamic RNA, DNA, and protein assemblies by small-angle X-ray scattering.* Curr Opin Struct Biol. **20**(1): p. 128-37.

94. Svergun, D.I., M.V. Petoukhov, and M.H. Koch, *Determination of domain structure of proteins from X-ray solution scattering.* Biophys J, 2001. **80**(6): p. 2946-53.

95. Pretto, D.I., et al., *Structural dynamics and ssDNA binding activity of the three N-terminal domains of the large subunit of Replication Protein A from small angle X-ray scattering.* Biochemistry. 2010 49(13) p. :2880-9

96. Svergun, D.I., *Restoring low resolution structure of biological macromolecules from solution scattering using simulated annealing.* Biophys J, 1999. **76**(6): p. 2879-86.

97. Pelikan, M., G.L. Hura, and M. Hammel, *Structure and flexibility within proteins as identified through small angle X-ray scattering.* Gen Physiol Biophys, 2009. **28**(2): p. 174-89.

98. Wuthrich, K., *Protein structure determination in solution by NMR spectroscopy.* J Biol Chem, 1990. **265**(36): p. 22059-62.

99. Cavanagh, J., Fairbrother, W. J., Palmer, A. G., Skelton, H. J., *Protein NMR Spectroscopy Principles and Practice*. Second ed, ed. A. Press. 2006: Academic Press. 912.

100. Evans, J.N.S., *Biomolecular NMR Spectroscopy*. 1995: Oxford University Press.

101. Rule, G., Hitchens, T. K., *Fundamentals of NRM Spectroscopy*. 2006, The Netherlands: Springer.

102. Teng, Q., *Handbook of Structural Biology: Practical NMR Applications*. 2004: luwer Academic/Plenum Publishers.

103. Wuthrich, K., *NMR of Proteins and Nucleic Acids*. 1986: Wiley-Interscience. 320.

104. 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.

105. Pfuetzner, R.A., et al., *Replication protein A. Characterization and crystallization of the DNA binding domain.* J Biol Chem, 1997. **272**(1): p. 430-4.

106. de Laat, W.L., et al., *DNA-binding polarity of human replication protein A positions nucleases in nucleotide excision repair.* Genes Dev, 1998. **12**(16): p. 2598-609.

107. Hura, G.L., et al., *Robust, high-throughput solution structural analyses by small angle X-ray scattering (SAXS).* Nat Methods, 2009. **6**(8): p. 606-12.

108. Kolpashchikov, D.M., et al., *Polarity of human replication protein A binding to DNA.* Nucleic Acids Res, 2001. **29**(2): p. 373-9.

109. Bernstein, N.K., et al., *Mechanism of DNA substrate recognition by the mammalian DNA repair enzyme, Polynucleotide Kinase.* Nucleic Acids Res, 2009. **37**(18): p. 6161-73.

110. Hammel, M., et al., *Ku and DNA-dependent protein kinase dynamic conformations and assembly regulate DNA binding and the initial non-homologous end joining complex.* J Biol Chem. **285**(2): p. 1414-23.

111. Chapados, B.R., et al., *Structural basis for FEN-1 substrate specificity and PCNA-mediated activation in DNA replication and repair.* Cell, 2004. **116**(1): p. 39-50.

112. Shin, D.S., et al., *Full-length archaeal Rad51 structure and mutants: mechanisms for RAD51 assembly and control by BRCA2.* EMBO J, 2003. **22**(17): p. 4566-76.

113. Williams, R.S., et al., *Nbs1 flexibly tethers Ctp1 and Mre11-Rad50 to coordinate DNA double-strand break processing and repair.* Cell, 2009. **139**(1): p. 87-99.

114. Roy, R., et al., *SSB protein diffusion on single-stranded DNA stimulates RecA filament formation.* Nature, 2009. **461**(7267): p. 1092-7.

115. Konarev P.V., V.V.V., Sokolova A., Koch M.H. J. andSvergun D. I. , *PRIMUS: a Windows PC-based system for small-angle scattering data analysis.* Journal of Applied Crystallography, 2003. **36**(Part 5): p. 1277-1282.

116. Rohl, C.A., et al., *Modeling structurally variable regions in homologous proteins with rosetta.* Proteins, 2004. **55**(3): p. 656-77.

117. Pettersen, E.F., et al., *UCSF Chimera--a visualization system for exploratory research and analysis.* J Comput Chem, 2004. **25**(13): p. 1605-12.

118. Blackwell, L.J., J.A. Borowiec, and I.A. Mastrangelo, *Single-stranded-DNA binding alters human replication protein A structure and facilitates interaction with DNA-dependent protein kinase.* Mol Cell Biol, 1996. **16**(9): p. 4798-807.

119. Liu, Y., et al., *Modulation of replication protein A function by its hyperphosphorylation-induced conformational change involving DNA binding domain B.* J Biol Chem, 2005. **280**(38): p. 32775-83.

120. Haring, S.J., et al., *Cellular functions of human RPA1. Multiple roles of domains in replication, repair, and checkpoints.* J Biol Chem, 2008. **283**(27): p. 19095-111.

121. Nooren, I.M. and J.M. Thornton, *Diversity of protein-protein interactions.* EMBO J, 2003. **22**(14): p. 3486-92.

122. 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.

123. Anantha, R.W., V.M. Vassin, and J.A. Borowiec, *Sequential and synergistic modification of human RPA stimulates chromosomal DNA repair.* J Biol Chem, 2007. **282**(49): p. 35910-23.

124. Francon, P., et al., *A hypophosphorylated form of RPA34 is a specific component of pre-replication centers.* J Cell Sci, 2004. **117**(Pt 21): p. 4909-20.

125. Kaustov, L., et al., *p53 transcriptional activation domain: a molecular chameleon?* Cell Cycle, 2006. **5**(5): p. 489-94.

126. Bastin-Shanower, S.A. and S.J. Brill, *Functional analysis of the four DNA binding domains of replication protein A. The role of RPA2 in ssDNA binding.* J Biol Chem, 2001. **276**(39): p. 36446-53.

127. Pacek, M. and J.C. Walter, *A requirement for MCM7 and Cdc45 in chromosome unwinding during eukaryotic DNA replication.* EMBO J, 2004. **23**(18): p. 3667-76.

128. Byun, T.S., et al., *Functional uncoupling of MCM helicase and DNA polymerase activities activates the ATR-dependent checkpoint.* Genes Dev, 2005. **19**(9): p. 1040-52.

129. Wu, X., S.M. Shell, and Y. Zou, *Interaction and colocalization of Rad9/Rad1/Hus1 checkpoint complex with replication protein A in human cells.* Oncogene, 2005. **24**(29): p. 4728-35.