Contributions to template priming from the human DNA primase 4Fe-4S cluster domain

By

Marilyn Elaine Holt

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

Chemical and Physical Biology

February 28, 2019

Nashville, Tennessee

Approved:

Charles R. Sanders, Ph.D.

Katherine L. Friedman, Ph.D.

Brandt F. Eichman, Ph.D.

Jared T. Nordman, Ph.D.

To my family,
who are endlessly supportive

ACKNOWLEDGMENTS

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

Overview

Replication of genetic information is a fundamental biochemical process, required for maintenance and propagation of genomic information. Slowing or stalling of this process can cause genomic instability, a hallmark of cancer (1-4), and the cell has developed sophisticated processes to efficiently detect and correct errors introduced during replication. In humans, DNA replication is meticulously organized by an array of DNA processing factors that form and regulate replicative machines to initiate and organize sites of DNA replication, replication forks (5). These mechanisms allow the six billion base pairs in the diploid human genome to be copied accurately and completely thousands of times per day as cells prepare to undergo cellular division.

New DNA is synthesized by splitting the double helix in two and using each strand as a template for a new complementary strand. However, DNA polymerases–which synthesize the vast majority of new DNA–are incapable of initiating DNA synthesis on the single-stranded DNA (ssDNA) template. Primases fill this functional gap by synthesizing an initial oligonucleotide "primer" that processive DNA polymerases then extend to generate the nascent complementary DNA strand. While the essential two-metal mechanism of RNA polymerase activity is well established, the molecular basis for the priming activity by primase and other components of the replisome remains unclear. Regulation of the activity of human DNA primase formed the central line of inquiry for my Ph.D. dissertation research.

DNA Replication Overview

DNA replication must occur for cells to divide and propagate their genetic legacy. In most bacteria, which have one chromosome, replication is initiated in one location, referred to as the "origin of replication" (6). Eukaryotes, which have larger genomes and slower replication machineries, initiate replication at multiple origins of replication on each chromosome to ensure that replication occurs in a timely manner (7). Archaeal replication systems differ from both bacteria and eukaryotes, as they utilize replication machinery that is reminiscent in composition of that observed at eukaryotic replication forks but operates at speeds closer to that of bacterial replication forks (8, 9). The number of origins in archaeal systems can vary widely, with some species utilizing only a single chromosome with a single origin of replication, others initiating replication from multiple origins, and at least one species that does not appear to require a defined origin of replication (10-12).

Despite differences in replication machinery components, the amount of genetic information to be copied, and where replication is initiated, the essential process of DNA replication is conserved across all three domains of life. This process starts when an origin of replication is bound by origin-binding proteins, which load two six-subunit DNA helicases onto DNA (6). The DNA helicase unwinds the double helix by enclosing one strand of the parental duplex DNA and using ATP hydrolysis to translocate across the strand, forcing the strands apart (13). This activity exposes template ssDNA, which is coated and organized by a ssDNA-binding protein (SSB). DNA primase generates an initial, short RNA oligonucleotide primer that DNA polymerases then elongate to synthesize the vast majority of nascent DNA.

Due to the 5'→3' directionality of DNA polymerases, only one strand–termed the "leading" strand–is continuously synthesized. The other, "lagging", strand is discontinuously

synthesized in segments commonly referred to as Okazaki fragments. These segments can range in length from hundreds of nucleotides in eukaryotes to thousands of nucleotides in prokaryotes. While primase only synthesizes one initial primer for the leading strand, each Okazaki fragment requires a unique primer for initiation of DNA polymerase activity. These primers are extended by a DNA polymerase to complete the fragment, after which the RNA portion is removed by a ribo-endonuclease and the resultant gap is filled in with DNA by another ploymerase (14). The completed Okazaki fragments are then joined by DNA ligase, completing the process of Okazaki fragment maturation. Other important components of the replication fork include polymerase clamps, which hold the DNA polymerases on the template strand, clamp-loading proteins, which recognize the primer-ssDNA junction and load the aforementioned clamps onto the ssDNA, and topoisomerases, which add negative supercoils to relieve strain caused by unwinding of the double helix.

To ensure cell viability and chromosomal stability, initiation of replication must be tightly coordinated to ensure the genome is copied exactly once. Bacteria regulate initiation of replication through mechanisms that modulate initiator protein expression and binding to the origin of replication; termination of bacterial DNA replication occurs when the replication fork is blocked by binding of the terminus site-binding protein to the termination site sequence (6, 15-18). Eukaryotic DNA replication occurs within the context of the cell cycle, with the G1/S checkpoint regulating initiation of DNA synthesis; timing and termination of eukaryotic replication will be discussed in more detail in the next section.

Eukaryotic DNA Replication

While bacterial, archaeal, and eukaryotic DNA utilize similar processes to replicate the genome, the mechanisms of eukaryotic DNA replication and replication timing are significantly more complex than those observed in bacterial and archaeal systems due to the increased length and complexity of the eukaryotic genome. Initiation of eukaryotic DNA replication occurs in three steps: origin licensing, activation, and firing. Origin licensing occurs during the G1 phase of the cell cycle, when the origin recognition complex (ORC), Cdc6, and CDT1 bind the origin and load the MCM2-7 helicase to assemble the pre-replication complex. In *S. cerevisiae*, origins of replication are primarily defined by autonomous replicating sequence elements (ARS) that ORC recognizes and binds (19-21). In multicellular organisms, chromatin organization and nuclear structure also play significant roles in origin recognition (22-24). Not all licensed origins are activated, and in both mammalian and yeast systems, the set of origins that undergo activation varies from cell to cell, suggesting that origin usage is flexible across Eukarya (25-27).

Origin activation, when the MCM helicase melts origin DNA and replisomes assemble around the activated helicase, is initiated at the G1/S checkpoint by phosphorylation of MCM2-7 and other key replication factors by cyclin-dependent kinase (CDK) and Dbp4-dependent kinase (DDK). During origin licensing, two MCM2-7 complexes are loaded in an inactive, head-to-head configuration. In origin activation, DDK and CDK phosphorylate MCM2-7 and a host of "firing factors" recruit Cdc45, GINS, Mcm10, and DNA polymerase ε, forming the complete Cdc45-MCM2-7-GINS (CMG) helicase (28, 29). MCM2-7 activation by DDK and formation of the CMG helicase leads to DNA melting and origin firing as the helicases begin translocating away from the origin to unwind the double helix, inducing recruitment of the other components of the replisome.
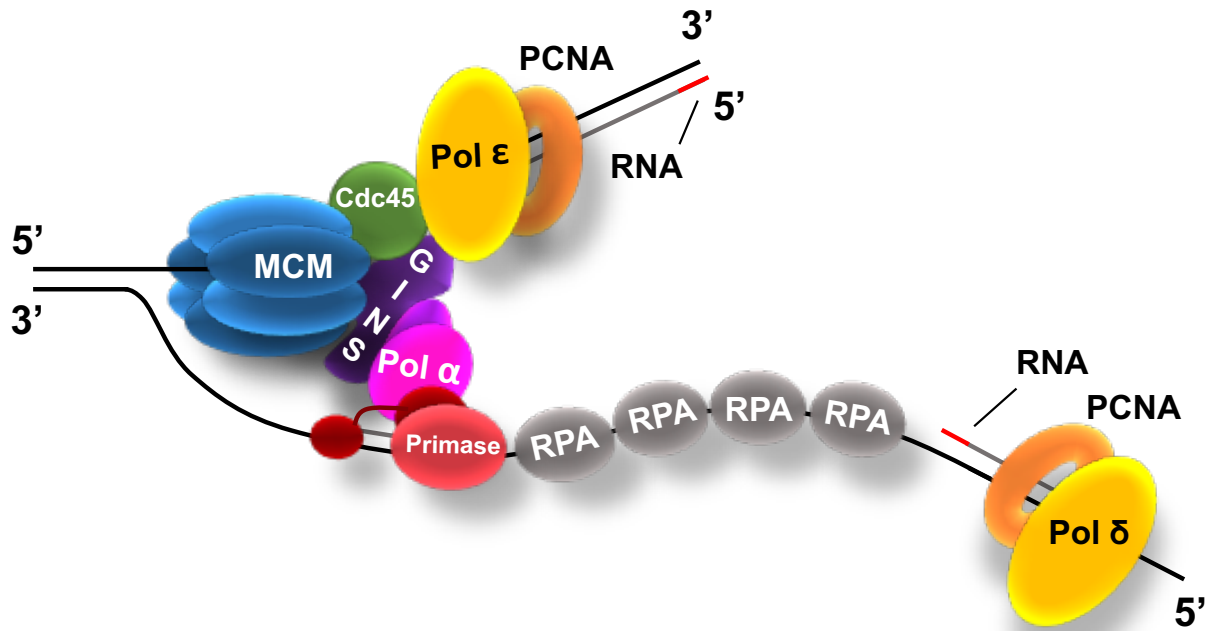
Figure 1.1. Simplified diagram of the eukaryotic replication fork. Adapted from (30).

During origin firing, the CMG helicases travel in opposite directions to form a replication "bubble" with two replisomes operating at the replication forks formed at each end of the replication bubble (Figure 1.1) (31, 32). Exposed ssDNA template is coated by replication protein A (RPA), a single-stranded DNA binding protein that protects the ssDNA and, in the case of replication fork stalling, recruits the ataxia-telangiectasia and Rad3-related kinase (ATR) and a number of other DNA damage response and repair proteins (33, 34). RPA contributes to the recruitment of polymerases, including DNA primase, which generates the initial 8-12 nt RNA primer (35, 36). In the cell, DNA primase exists as part of the polymerase α-primase (pol-prim) heterotetramer and hands off the initial primer to DNA polymerase α (pol α), which extends the primer by 20-30 DNA nts (35, 37, 38). The mixed RNA/DNA primer is in turn handed off to the processive DNA polymerases ε or δ (pols ε and δ), which synthesize the

majority of DNA on the leading and lagging strands, respectively (30). After completion of Okazaki fragment synthesis by pol δ, the Fen1 and Dna1 endonucleases remove the RNA primer and DNA ligase connects the fragments (30). During this process, fork progression is tightly regulated to prevent replication stress (39-42).

Priming by pol-prim occurs once on the leading strand and at the beginning of the generation of every Okazaki fragment on the lagging strand. Single-molecule studies with the bacteriophage T7 replication complex suggest that primase acts as a "molecular brake" during replication and causes transient pausing of the replisome during primer synthesis (43). This function ensures that the leading strand does not overtake the lagging strand despite the relatively slow rate of Okazaki fragment synthesis. Work by the Pellegrini group has revealed that the structural yeast protein Ctf4 interacts with both the GINS complex, pol α, and other proteins containing a Ctf4-interacting peptide, suggesting that Ctf4 may participate in coordination of leading and lagging strand synthesis (37). Recent work indicates that this interaction is maintained between human pol α and the human homolog of Ctf4, AND-1 (44-46).

Termination of replication occurs when two replication forks moving in opposite directions converge. Termination results in topological stress that must be resolved by either topoisomerase activity or pre-catanene formation (intertwining of sister strands behind the replication fork) (47, 48). After the converging replication forks meet, gaps are filled in with pol δ, Okazaki fragment maturation is completed, and the replisomes are unloaded from the DNA. Many details of this process are not well-established in eukaryotic DNA replication (16); however, current work suggests that converging CMG machineries continue translocating until they reach the downstream Okazaki fragment and pass onto dsDNA. At this point, MCM7 is polyubiquitinated by a Cullin-type E3 ligase, which promotes CMG unloading by p97 and,

indirectly, passive unloading of the replisome components that directly interact with the CMG helicase (49-52). The replisome components that primarily bind PCNA–including pol δ, the FEN1 endonuclease, and DNA ligase–presumably unload with PCNA, which itself is unloaded by Elg1 replication factor C-like protein (Elg1-RFC) (53).

Because the diploid human genome is approximately six billion base pairs (bp) in size and the eukaryotic replication fork travels at approximately 1-2 kb/minute, human cells utilize 30,000-50,000 origins of replication to ensure that DNA replication occurs in a timely manner (54). Coordinating nearly 100,000 replication forks to ensure the genome is copied exactly once requires robust mechanisms to regulate initiation of replication (7, 24, 55). Eukaryotic cells have evolved multiple mechanisms to ensure that licensed origins do not re-fire (56). CDK-mediated phosphorylation of MCM2-7, ORC, Cdc6, and Cdt1 both inhibits origin licensing during S phase and activates the growing replication assembly to recruit key initiation factors, including GINS, Sld2, and pol ε (57-60). The opposing roles of CDK–inhibition of origin licensing and stimulation of origin activation–temporally segregate the origin licensing and activation processes to reduce origin re-firing during S phase (61). Origin licensing in S phase is also prevented through interactions with Cullin-based E3 ligases, which target ORC, E2F1, and CTD1 for ubiquitination and proteolysis, and geminin, a coiled-coil protein expressed during the S and G2 phases that specifically binds CDT1 to inhibit origin licensing (62-65).

Unlike bacteria, not all eukaryotic origins fire simultaneously, and replication is further organized by division of the genome into replication domains during S phase. These domains consist of sets of replicons ranging from 30 to 450 kilobases (kb) in size and their associated origins, which fire as a unit (32, 66). Pulse-labeling studies have established that replication of these domains occurs in temporal and spatial patterns that correspond with chromatin

organization and nuclear structure in a cell type-specific manner (22, 31, 32, 67-71). This spatial

regulation of DNA replication has been proposed to allow the cell to adjust the number of

replication forks during replication stress and to facilitate coordination of transcription and

replication (7, 61). Regulation of replication timing at this scale may also have implications for

the fidelity of DNA replication, as research from the Sunyaev group suggests that regions that

replicate later in S phase experience higher mutation rates (72).

Replication stress, caused when the replication fork stalls or collapses, can lead to

genomic instability. Replication stress can occur for multiple reasons, including DNA damage,

protein-DNA crosslinks that block the replisome, and collisions with the transcription machinery

(73). Often, these blocks to the replication fork only cause the replisome to pause transiently:

damage on the lagging strand will be bypassed by the next Okazaki fragment, and repriming on

the leading strand can permit continued movement of the replication fork (74-76). More

persistent challenges to the replisome, such as interstrand crosslinks, may be bypassed through

replication by a fork initiated from an adjacent origin of replication.

Despite these mechanisms for restarting replication during replication stress, a replication

fork may be blocked without an adjacent functional replication fork that can bypass the damage.

Persistent stalling of the replication fork can cause replication fork collapse, or dissociation of

the replisome from the replication fork. Fork collapse results in a variety of DNA structures

vulnerable to inappropriate processing and damage. Replication fork stalling activates the G1/S

checkpoint, which ensures that the cell cycle does not progress until the damage is fixed and

replication is complete, and DNA damage tolerance pathways, which stabilize the stalled forks

and ensure they restart appropriately (76, 77). Collapsed replication forks require more invasive

interventions to restart the replication fork, such as homologous recombination (78, 79). Failure

to appropriately address replication fork stalling or collapse, as well as errors introduced by the repair pathways (80), can cause mutations, chromosomal rearrangements, cell death, and a plethora of human diseases (81-83). This highlights the fact that well-functioning replication forks are a critical component of human health.

The Problem of Priming: Initiation of DNA Synthesis

Traditionally, DNA polymerases have been defined by their requirement for a free 3'-OH group to initiate DNA synthesis, and only recently have DNA polymerases capable of initiating DNA synthesis *de novo* on ssDNA been identified (75, 84, 85). Multiple hypotheses have been proposed to explain this near-universal requirement (86-88). Because NTPs exist at concentrations in the cell that are orders of magnitude higher than dNTP concentrations, some have suggested that the greater availability of NTPs confers a catalytic advantage, resulting in the evolution of mechanisms that utilize mass action effects to overcome thermodynamic barriers to initiation of synthesis (87, 88). It has also been hypothesized that cells may use RNA as a "tag" to identify synthesis start sites and drive error-free removal and replacement of these error-prone regions of replication (86-88). Whatever the reason, this requirement has driven evolution of a diversity of solutions for the problem of priming. Perhaps the most outré of these can be found in viruses, some of which utilize host tRNAs to prime genome synthesis or employ a terminal protein that sits on the 5' end of the viral genome to provide a free 3'-OH group (89-97). However, the most common solution to this problem is an enzyme with RNA polymerase activity, primase (88, 98).

Structurally, primases are divided into those utilizing a topoisomerase-primase (toprim) fold (bacterial and most viral primases) and those in the archaeo-eukaryotic primase (AEP)

superfamily (Figure 1.2) (99, 100). Bacterial and viral primases are typically single-chain enzymes that strongly associate with the replicative helicase at the replication fork to form multimeric primosome complexes (101-107), and this helicase-primase interaction stimulates priming activity by as much as 1000-fold (105, 108). One notable exception to this rule is the herpes simplex helicase-primase complex, a member of the AEP superfamily for which knocking out helicase activity actually enhances priming activity (100, 109-111).

S. aureus
PDB ID: 4EDG

P. horikoshii
PDB ID: IV34

H. sapiens
PDB ID: 4LIL
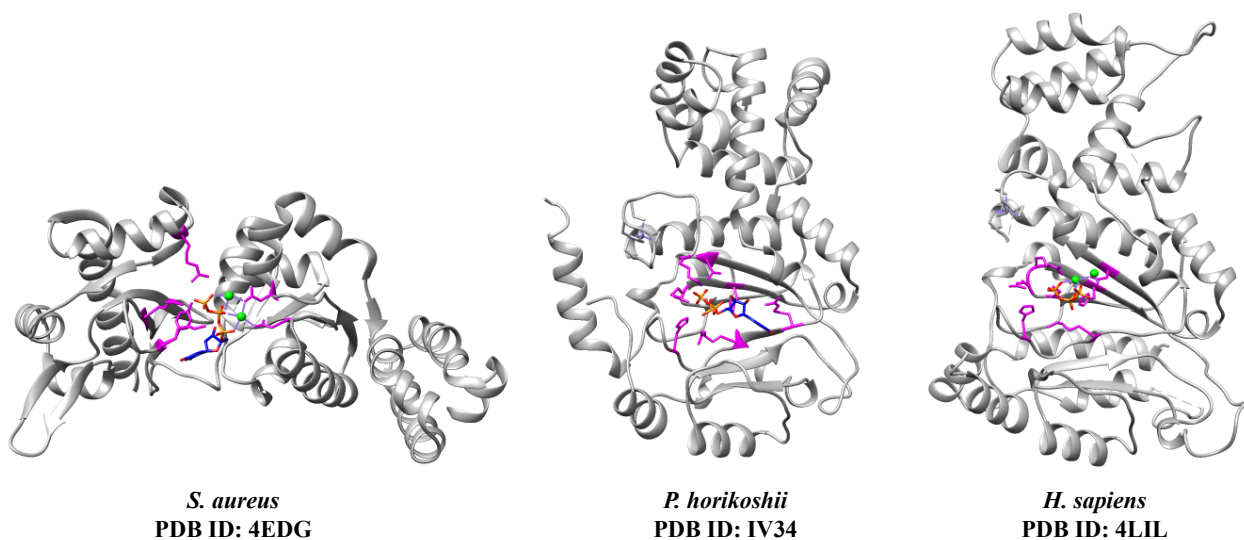
Figure 1.2. Comparison of the RNA polymerase domains of representative bacterial, eukaryotic, and archaeal primases. Catalytic subunits depicted in grey, with UTP colored in orange and blue, manganese ions in green, and key catalytic residues in pink.

In contrast to single-chain bacterial primases, AEP replicative primases are composed of at least two subunits: a small subunit, PriS, that contains the primary catalytic activity, and a large subunit, PriL, that is a critical component of the initiation active site and regulates primase processivity (112-115). The C-terminal domain of eukaryotic PriL subunits (PriL-CTD) contain an iron-sulfur (Fe-S) cluster that is required for primase activity (116-118). However, while some archaeal PriL-CTDs contain Fe-S clusters, this cluster is not required for priming activity

(119). Recent studies have identified a third subunit in some archaeal primases, PriX, that is structurally homologous to the Fe-S cluster-containing C-terminal domain of eukaryotic PriL subunits (PriL-CTD) and appears to play similar roles in forming the initiation active site and regulating primase processivity (119-121). Homologous PriX domains have not been identified in all species of archaea; however, this discovery does beg speculation that functionally homologous subunits may exist to fill the role played by the Fe-S cluster-containing PriL-CTD in eukaryotic primases (121). Unlike bacterial primases, eukaryotic primases do not directly associate with the replicative helicase; instead, they directly bind pol α to form the pol-prim heterotetramer (122). While most bacterial and viral primases exhibit sequence specificity during initiation of primer synthesis, eukaryotic primase start site selection appears to be driven more by interactions with other components of the replication fork and relative nucleotide concentrations (37, 123-129).

Regulation of eukaryotic primer synthesis

Despite the differences in sequence and structure between bacterial, archaeal, and eukaryotic primases, all primases exhibit common catalytic features, including the two-metal catalytic mechanism common to many polymerases (130). Eukaryotic primases work in three steps: initiation via synthesis of the first diribonucleotide, elongation to 7-10 ribonucleotides, and transfer of primed template to a pol α (37, 131, 132). Primases in general are quite error-prone; however, they have a remarkable ability to limit primer length, causing Lehman to dub primase the "enzyme that can count but can't read" (86, 133).

In human DNA primase, the small catalytic subunit is referred to as p48 and the large regulatory subunit is p58 (132, 134). When the human primase heterodimer (p48/p58) is

subjected to limited proteolysis, the p48/p58 heterodimer forms two stable domains, one composed of p48 and the amino terminal domain of p58 (p48/58N) and one containing the C-terminal domain of p58 (p58C) (Figure 1.3) (116). Biochemical studies have established that, while p48 contains the catalytic apparatus, p58C contributes to template binding and is required for initiation of primer synthesis and appropriate termination of primer synthesis. Regulation of termination of primer synthesis is commonly referred to as primer counting because it results in primers of defined length(112-115, 132, 135-137).
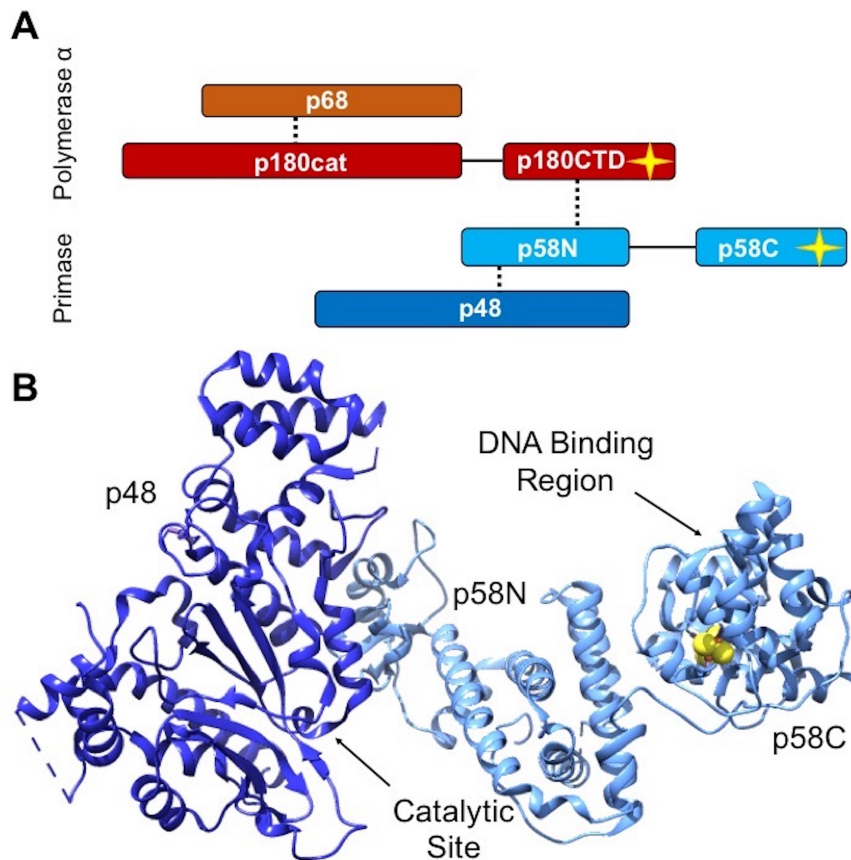
Figure 1.3. Organization of human DNA primase. A. Domain map of the DNA polymerase α -primase. Dotted lines show where domains interact. Yellow stars are shown for iron-sulfur clusters. B. The crystal structure of full-length primase (PDB ID: 4RR2). The iron-sulfur cluster in p58C is represented with yellow spheres. From Holt, M.E., et al. A polymerase with potential: the Fe-S cluster in human DNA primase. **2017**. *Methods in Enzymology*. 595:361-90.

In the context of pol-prim, human primase activity is negatively regulated by pol α after

synthesis of the initial RNA primer, preventing further primer elongation (138). This negative

regulation can be alleviated through the addition of either dNTPs or the pol α inhibitors

aphidicolin and DNA$_G$ –i.e. by allowing elongation of the RNA primer by pol α or by preventing

handoff of the primer to pol α. Negative regulation by pol α can also be alleviated by completing

the reaction at lower temperature or on a poly(dT) substrate, suggesting that this negative

regulation by pol α requires formation of a stable primer-template species (37, 112, 138, 139).

However, while pol α regulates the number of primers generated by primase, termination of

primer synthesis is regulated by primase even in the absence of pol α.

Initial experiments to identify the catalytic residues of primase mapped the essential

catalytic core to p48; however, attempts to determine the mechanism of primer counting by

mutation of these essential catalytic residues simply disrupted total primase activity, suggesting

that primer length is regulated by p58 (36, 135). Therefore, to develop a hypothesis for the

structural basis of primase counting and regulation, one must delve into the inner workings of

p58 and its interactions with pol α. Copeland and colleagues initiated this effort by creating

deletion constructs of p58 and coexpressing these with full-length constructs of p48 to map the

regions of p58 necessary for each stage of primer synthesis (see Figure 1.4) (134). They found

that, of these deletion constructs, the only construct capable of initiating primer synthesis in

complex with p48 was one in which they deleted the unstructured N-terminal tail on p58

(p48/p58$_{155-509}$).   Removing either p58N (p48/p58$_{270-509}$) or p58C (p48/p58$_{1-270}$) completely

abrogated primer initiation, as did deletion of residues 400-509 (which contain the final

coordinating cysteine for the Fe-S cluster, C424). However, while deletion of these regions

precluded initiation by the p48/Δp58 complex, all of these constructs were apparently capable of

elongating a synthetic primer. In fact, a construct without the final 109 residues (including C424)

appeared to increase p48's capacity for elongation relative to wild type. They suggested that this construct contained the regions necessary to stabilize p48 while excluding a feature involved in negative self-regulation (i.e. primer counting).
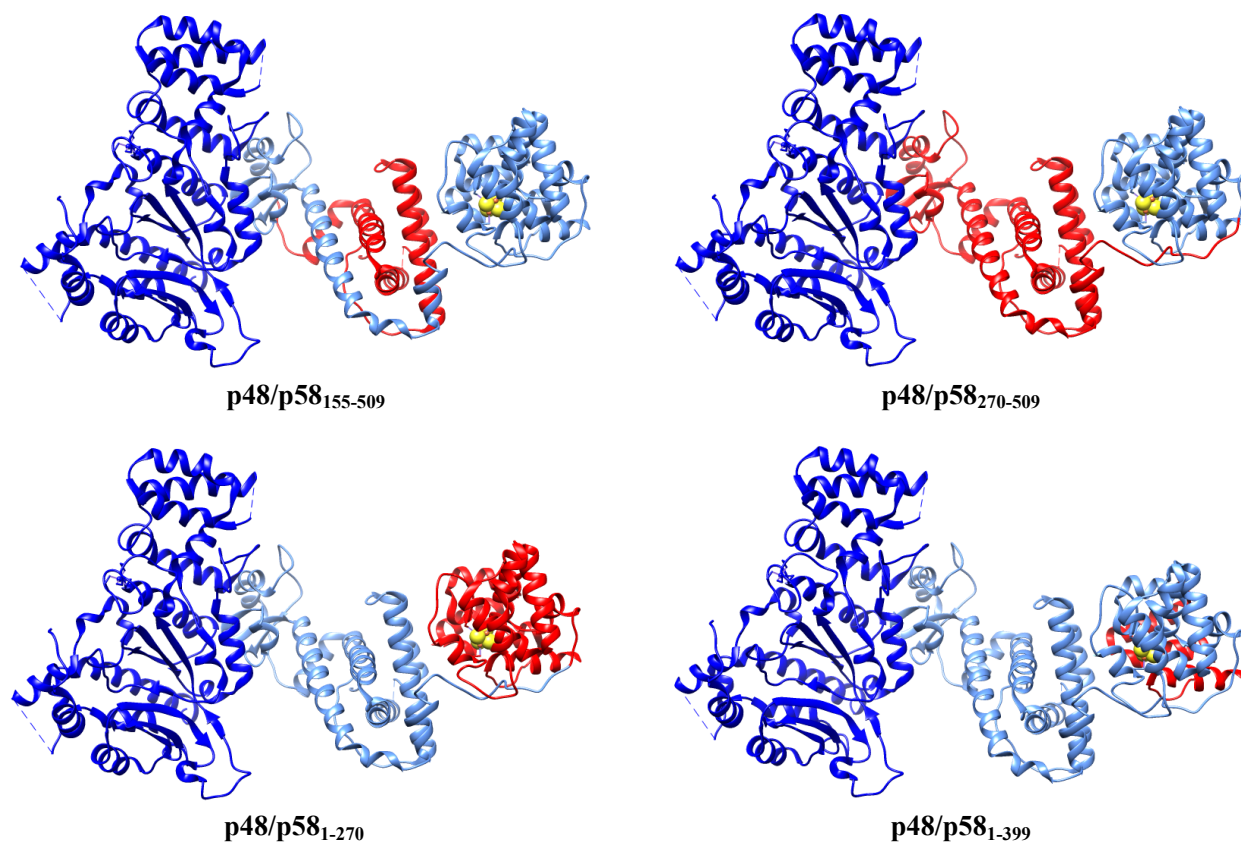
Figure 1.4. Visualization of deletion constructs assayed by Copeland et *al.* (134). The crystal structure of full-length primase (PDB ID: 4RR2) colored to show truncated residues of each construct in red. p48 is colored in dark blue, p58 in light blue, and the iron-sulfur cluster in p58C is represented with yellow spheres.

Kuchta and colleagues further studied the role of p58 in primase function by conducting a series of experiments that identified p58 as the primary DNA binder before and after primer synthesis, as well as identifying a pol β-like region they proposed to be key to primer counting activity (112, 114). Deletion of this pol β-like region, residues M288-L313, completely abrogated primase's ability to count, and alanine scanning of conserved charged residues in this

region revealed that residues R302, R306, and K314 were key to initiation of primer synthesis (114). While the fact that the structural integrity of the deletion construct was not validated complicates evaluation of these results, these charged residues were later shown to be part of a positively-charged putative DNA binding surface on p58C (36). Further investigations ultimately revealed the presence of an essential Fe-S cluster in p58C (116, 118).

As more structure-oriented labs took a healthy interest in primase, a structurally-based model for p58's interactions with p48 and pol α began to emerge. At this time, crystal structures have been determined for free full-length primase and the pol-prim heterotetramer, as well as the globular domains, including p48/p58N, p48 bound to the 3'-NTP and catalytic metals, p58C, and p58C bound to primed template (36, 137, 140-144). These studies clearly delineated the regions of p58 responsible for binding p48 and pol α, providing a structural context for further investigations into substrate binding by primase. Based on these crystal structures, it has been hypothesized that primase engages in two distinct enzymatic processes (initial dinucleotide synthesis and primer extension) and that these enzymatic processes are completed in the context of two distinct configurations of primase (Figure 1.5) (145). In this model, p58C is a critical component of the active site for initial dinucleotide synthesis but is not required for elongation activity and travels with the 5' end of the nascent primer. This disengagement from the primase active site potentially positions p58C to modulate primer termination and handoff to pol α (137, 145). Testing this hypothesis and elucidating the details interactions between members of the pol-prim heterotetramer during primer termination and handoff will require structural characterization of full-length primase and pol-prim with catalytically-relevant substrates.
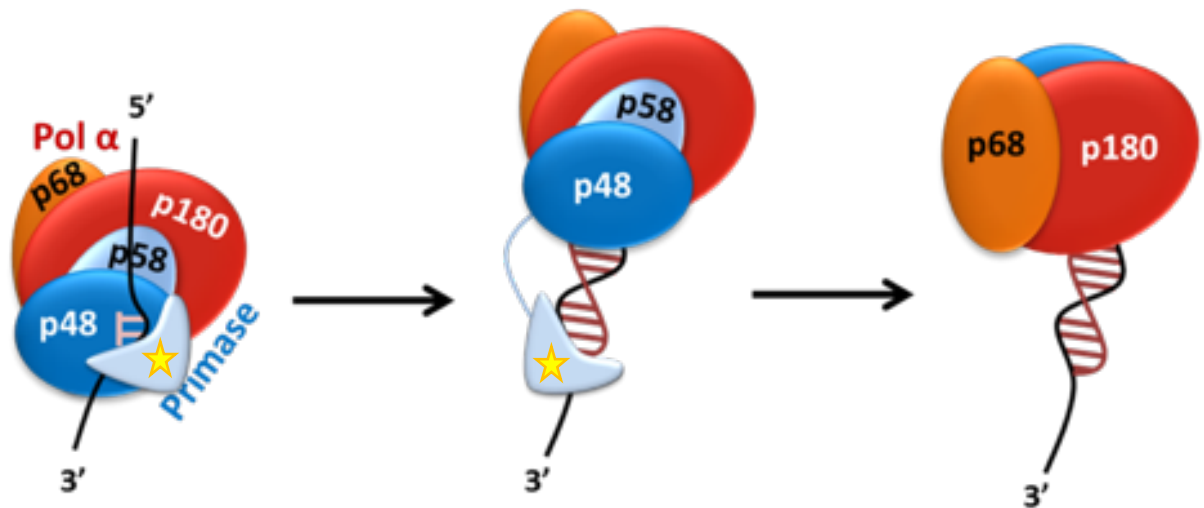
Figure 1.5. Mechanism of primer synthesis. After assembly of the primosome on the ssDNA template, p48 and p58C collaborate to synthesize the initial dinucleotide in the rate-determining step of primer synthesis. p48 rapidly elongates the nascent primer to a length of 7-10 nucleotides before the catalytic subunit of pol α (p180) displaces p48 at the 3' end of the nascent primer and extends to a final length of 20-30 nt. Star denotes Fe-S cluster in p58C.

Experimental Methods[1]

Technological advances have rapidly advanced the range of techniques for structure

determination and ways to apply that structural information to understand essential biochemical

processes and further drug development (146-149). X-ray crystallography is a powerful

technique for determining protein structures at atomic resolution, and crystallography-based

structure determination has both furthered our understanding of fundamental processes and

enabled structure-based drug development (150-153). However, crystal generation is often

technically challenging, especially with dynamic or flexible proteins, because it requires

formation of an ordered crystal lattice (154). Additionally, because diffraction quality depends

on the degree of regularity within the crystal lattice–i.e., to what degree the proteins making up

---

[1] Portions of this section were previously published in Holt, M.E., et al. A polymerase with potential: the Fe-S cluster in human DNA primase. **2017**. *Methods in Enzymology*. 595:361-90.

the lattice possess similar configurations–high-resolution data generated from well-diffracting crystals almost certainly does not reflect the dynamics or flexibility of the functional protein (154). This is a significant drawback when attempting to determine structure-function relationships for proteins for which dynamics are proposed to drive protein function, like human DNA primase (145, 155).

To overcome this challenge, a number of techniques have been developed to experimentally measure or predict protein structure and dynamics in solution (156-159). Ideally, these studies are completed in conjunction with biochemical assays, such as measurements of binding affinity and enzymatic activity, to validate proposed structure-function relationships. This section describes the application of circular dichroism (CD), nuclear magnetic resonance (NMR), and small-angle x-ray scattering (SAXS) to characterize the biochemical and structural properties of full-length primase and the isolated p58C domain. In addition, the design of affinity measurement assays will be discussed in the context of fluorescence anisotropy-based binding assays.

UV-Visible absorbance spectroscopy and circular dichroism

Absorption spectroscopy in the UV-visible (UV-Vis) region of the electromagnetic spectrum is a commonly used technique for quantitative determination of various analytes, including biological macromolecules and transition metals. In short, a solution with the analyte of interest is irradiated with light in the 190-700 nm range of wavelengths. Absorption at each wavelength is determined by the molar absorptivity of the analyte and can be used to calculate analyte concentration in solution through the use of the Beer-Lambert Law (Equation 1), in which A is absorbance, ε is molar absorptivity, b is path length, and c is concentration.

$$A = \varepsilon bc \qquad\qquad [1]$$

UV-Vis spectroscopy is of particular interest in the characterization of iron-sulfur cluster containing proteins because iron-sulfur clusters absorb in the 380-550 nm range (160, 161). This phenomenon, which results in a distinctive orange-brown color, has been used to identify iron-sulfur clusters in replicative polymerases (116, 118, 162). Because 2Fe-2S clusters primarily absorb at 400 and 456 nm with a shoulder at 510 nm, while 4Fe-4S clusters display a wide absorbance peak at approximately 410 nm, close examination of the UV-Vis spectrum can provide evidence for cluster type (160, 161, 163). Additionally, UV-Vis absorption spectroscopy has been routinely used to monitor iron-sulfur cluster formation and loading in cluster-containing proteins (161, 163-167).

Circular dichroism, which is essentially UV-Vis spectroscopy with circularly polarized light, has additional practical applications in the analysis of biological macromolecules (168). In this technique, a sample is irradiated with circularly polarized light, and the degree of right- and left-polarized light absorbed by the sample is measured by a detector fitted with a polarization filter. The CD spectrum is then generated by plotting the difference in absorption of right- and left-polarized light versus wavelength. CD is particularly useful for bioanalytical analysis; for example, near- and far-UV CD spectra of different protein preparations can be compared to determine if they are structurally similar (169-171). Secondary structures and aromatic residues result in characteristic contributions to the CD spectrum in the range of 190–260 and 250–330 nm, respectively (170, 172). These properties allow rapid analysis of secondary structure content and enables swift detection of mutation-induced structural perturbations through comparison of CD spectra collected on protein variants (168, 169, 171). Interestingly, iron–sulfur clusters also display unique spectra in the visible region of CD spectra, which can be

used to help identify the iron–sulfur cluster type (167, 173-175). However, it should be noted that while these spectroscopic methods are useful for monitoring cluster loading and condition, initial identification of cluster type requires confirmation by spectroscopic methods geared toward evaluation of paramagnetic centers, such as electron paramagnetic resonance, magnetic circular dichroism, and Mössbauer spectroscopy (175-177).

Nuclear magnetic resonance

NMR is a powerful tool for characterizing the structure and dynamics of proteins (156, 178). NMR in solution can be used for a range of applications, spanning from simple analytical biochemistry to determination of structural dynamics at atomic resolution. One very basic application of NMR is to obtain a rapid overview of the tertiary structure (or lack thereof) of a protein. The most common approach involves acquiring heteronuclear two-dimensional correlation experiments (typically ^{15}N–^{1}H HSQC) on ^{15}N-enriched samples. Such spectra provide a direct reflection of the extent of folding of a polypeptide chain and a ready means to determine whether a mutation causes perturbation of the tertiary structure (156).

Because NMR signals are highly sensitive to any changes in their electronic environment, NMR is widely used to monitor binding events(179). These experiments involve titration of a ligand into a protein solution and acquisition of the NMR spectrum at each point. The changes in the NMR signals–broadening of signals and chemical shift perturbations (CSPs)–can be used to map binding sites. CSPs can also be plotted versus the substrate concentration to generate a binding curve, from which a dissociation constant (K_D) can be extracted. The NMR-based approach is particularly valuable for characterizing weak binding in the μM to mM range.

NMR with iron–sulfur cluster-containing proteins can be challenging because the

paramagnetism of the cluster causes rapid relaxation, reflected as severe line broadening that

reduces the ability to observe signals, particularly for residues near the paramagnetic site.

Despite this, the partial NMR spectra can be used for basic biochemical analyses, such as

validating the structural integrity of p58C mutants and demonstrating binding of ligands to p58C

(116, 143). Methods for detailed NMR analysis of proteins with paramagnetic centers have been

developed (180), but these have yet to be applied to primase.

Small-angle x-ray scattering

Originally developed approximately seventy years ago, a combination of advancement in

technology and theory in the last twenty years have transformed SAXS into a powerful technique

for assessing the structural properties of biological macromolecules in solution. SAXS data

collection is similar to that of x-ray crystallography: a protein sample is exposed to a high flux,

coherent X-ray beam that scatters from electron centers within the sample, resulting in unique

intensity patterns that can be used to derive information about the distribution of these electron

centers (157, 181-183). SAXS is fundamentally different, however, in that a solution-state

sample is used for this process, rather than an ordered crystal. Isotropic tumbling of molecules in

solution precludes formation of the ordered diffraction patterns used to calculate high-resolution

x-ray crystal structures, instead resulting in a scattering curve that is radially symmetric around

the detector center. This scattering curve can be described as I(q), a function of the momentum

transfer, q, described by Equation 2; here, 2θ is the scattering angle and λ is the wavelength of

the incident beam.

$$q = \frac{4\pi \sin(\theta)}{\lambda} \qquad [2]$$

For a homogenous solution, I(q) can be described by Equation 3, using the distance

distribution function, P(r), where D_{max} is the maximum distance observed in the particle.

$$I(q) = 4\pi \int_0^{D_{max}} P(r) \frac{\sin (qr)}{qr} dr \qquad [3]$$

Practically speaking, the lowest resolution region of the scattering curve is dominated by the Rg, the average squared distance of each scatterer from particle center, as described by the Guinier approximation.

$$I(q) = I(0)e^{-\frac{q^2 Rg^2}{3}} \qquad [4]$$

Plotting the log of I(q) against q^2 yields the Guinier plot, from which the Rg and I(0) of the scattering particle can be determined. This plot should be linear over the range q*Rg < 1.3; "bending" or "curving" in this range suggests either aggregation or long-range interparticle interactions, and further analysis of the dataset be undertaken with great care (Figure 1.6) (157).
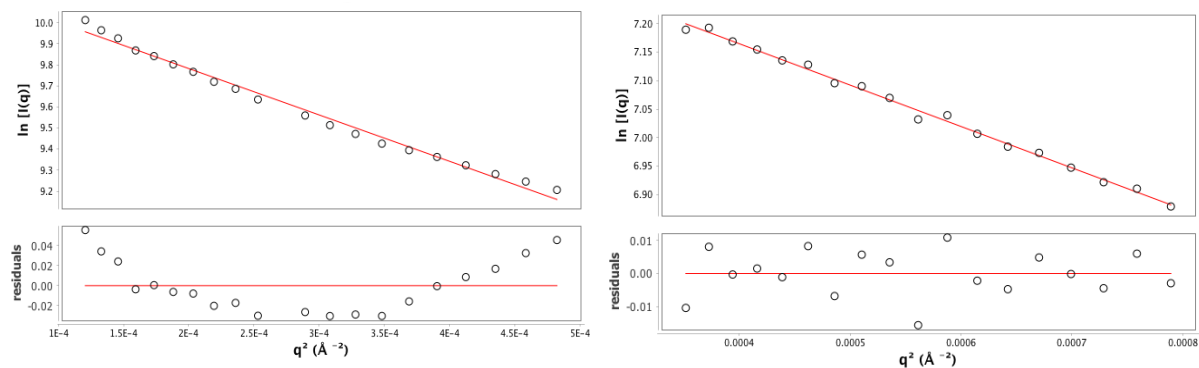
Figure 1.6. Guinier analysis reveals aggregation. Aggregation is reflected in curvature of the Guinier plot (top), and linearity in this region indicates that the sample is monodisperse (bottom).

Evaluation of molecular shape and flexibility requires analysis of scattering at higher q values. For folded molecules, scattering intensity decreases according to Porod's law, which can be described in a general manner by Equation 5, where P_x is the Porod exponent (157, 184, 185).

$$I(q) \propto q^{-P_x} \qquad [5]$$

Porod exponent values are determined by the fractal degrees of freedom, which is related

to the shape and foldedness of the molecule (185). Well-folded, globular particles will have P_x

values near 4, while the P_x of random coils will be near 5/3. The relationship described in

Equation 5 breaks down at high q values, when atomic resolution information begins to

significantly contribute (157). The region in which it applies is known as the power law regime,

and the value of P_x is determined by assessing the linearity of this region when transformed by

q^2, q^3 or q^4 (157, 184). Foldedness can also be determined in a qualitative manner from the

Kratky plot ($q^2I(q)$ vs q), which can be calculated directly from the scattering curve and yields

distinctive patterns based on the shape and foldedness of the protein (Figure 1.7A) (157, 184).
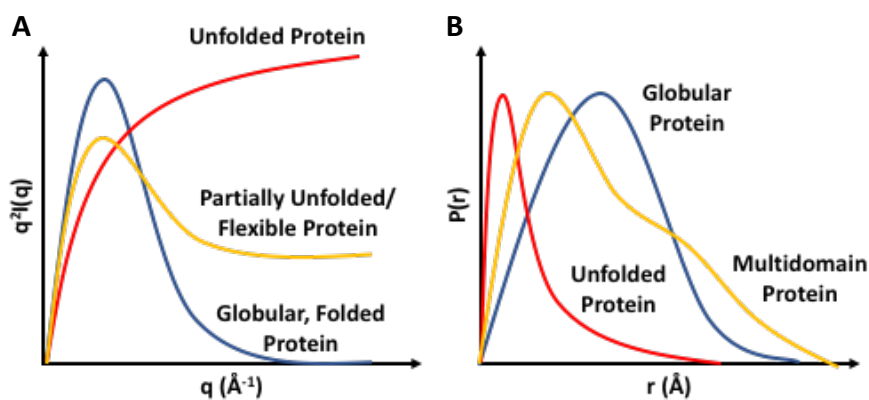
Figure 1.7. Kratky and P(r) transformations of scattering data. Kratky plots (**A**) and P(r)
distributions (**B**) can be used to evaluate protein foldedness and shape. Figures adapted from
(157).

Further details of domain or subunit orientation and distance separation can be extracted

by plotting the pairwise distance distribution function, P(r) (Figure 1.7B) (157, 184, 186). As

described in Equation 3, P(r) is related to the scattering intensity through a Fourier transform; to

reduce systemic errors caused by the discontinuous nature of experimental data collection and

low information content of SAXS scattering profiles, the P(r) distribution is calculated using an

indirect Fourier transformation (187). This yields a histogram of probability versus radius that describes the incidence of interatomic vector lengths within the particle and from which the maximum distance, D_{max}, can be determined. Importantly, "real-space" Rg and I(0) values that utilize the entire scattering curve, instead of just the Guinier region, can be calculated from the P(r) function; this serves as a useful check of the Guinier Rg and I(0) values and can be used to assess sample quality (157).

The advantages of SAXS–the solution nature of the experiment, the modest amount of sample required, and the lack of size limitation–have made SAXS an increasingly powerful tool when used in conjunction with computational methods and high-resolution structural techniques. Models developed from crystal structures can be rapidly compared to solution-state behavior by back-calculating the theoretical scattering profile of the model and comparing it to the experimental scattering profile to assess model quality (188-190). Additionally, molecular envelopes can be calculated from the P(r) distribution and crystal or NMR structure docked within the molecular envelope to gain insight into domain organization and architectural rearrangements (191-194). This integrated approach has empowered analyses of the macromolecular dynamics that link protein structure to biological outcome (194-197).

Fluorescence anisotropy-based binding assays

A variety of methods have been developed to measure binding affinity (198), which can be described by the equilibrium association constant (K_A). In the case of a specific, one-site binding interaction between two binding partners, this quantity is defined by Equation 6 (199). Here, P denotes protein, L is ligand, and PL is the protein-ligand complex.

$$K_A = \frac{[PL]}{[P][L]}$$ [6]

Due to the large numbers associated with K_A, affinity is typically reported in terms of the inverse of the K_A, the equilibrium dissociation constant (K_D).

$$K_D = \frac{[P][L]}{[PL]} \qquad [7]$$

From this, we can derive a general equation describing the behavior of protein-ligand interactions. Essentially, if

$$[P_{free}] = [P_{total}] - [P_{bound}] \qquad [8]$$

$$[L_{free}] = [L_{total}] - [L_{bound}] \qquad [9]$$

then we can rewrite Equation 7 as Equation 10.

$$K_D = \frac{([P_{total}] - [P_{bound}])([L_{total}] - [L_{bound}])}{[PL]}. \qquad [10]$$

Assuming that the protein and ligand are only binding each other (i.e. that there are no competing interactions), this simplifies to

$$K_D = \frac{([P_{total}] - [PL])([L_{total}] - [PL])}{[PL]} \qquad [11]$$

which, with a little rearrangement, can be expressed as Equation 12.

$$0 = [PL]^2 + ([P_{total}] + [L_{total}] + K_D)[PL] + [P_{total}][L_{total}] \qquad [12]$$

To describe the relationship of [PL] to the total concentrations of protein and ligand in solution, we simply apply the quadratic equation to generate Equation 13.

$$[PL] = \frac{-([P_{total}] + [L_{total}] + K_D) \pm \sqrt{([P_{total}] + [L_{total}] + K_D)^2 - 4[P_{total}][L_{total}]}}{2} \qquad [13]$$

We can then relate this to the signal produced over the course of a titration of protein into ligand by applying the following relationship, in which S_x is the signal when $[P] = x$, S_0 is the signal when $[P] = 0$, and ΔS is the maximum change in signal.

$$S_x = S_0 + \Delta S \frac{[PL]}{[L_{total}]} \qquad [14]$$

This results in Equation 15, which relates signal to the concentration of the binding

partners in solution.

$$S_x = S_0 + \Delta S \frac{-([P_{total}]+[L_{total}]+K_D)\pm\sqrt{([P_{total}]+[L_{total}]+K_D)^2-4[P_{total}][L_{total}]}}{2[L_{total}]} \qquad [15]$$

This derivation incorporates several key assumptions with practical implications. First, as stated at the beginning of this section, we assume the interaction occurs at a single site. While considerations are similar for multiple site binding, the order of Equation 7 increases with the number of binding sites, and equations relating signal to binding events at multiple sites can become considerably more complex (200). Second, in the derivation of Equation 11, we assumed that the binding event is specific–that there are no competing ligands for A or B. Again, addition of competitive binding interactions requires further consideration that is beyond the scope of this chapter. Finally, to use Equation 14, we assumed that signal varies linearly with the proportion of bound protein over the course of the titration. If this is untrue for an assay in question, a switch should be made to a different probe or assay that provides a linear response.

While Equation 15 is extremely useful for simulating binding interactions and determining initial experimental parameters (i.e. range of concentrations to use given an expected K_D), generating a K_D from experimental data using this equation is often nontrivial due to current software limitations. In standard DNA-protein binding assays, protein is typically titrated into DNA that has been fluorescently or radioactively labeled. Because the concentration of DNA in these assays is often held an order of magnitude or more below the expected K_D, we the approximation that $P_{free} = P_{total} = [P]$ is made and the following binding polynomial derived from free reactant concentrations is used (201):

$$S_x = S_0 + \Delta S \frac{\frac{[P]}{K_D}}{1+\frac{[P]}{K_D}} \qquad [16]$$

If S_0 is 0 at $[P] = 0$, this can be further simplified to Equation 17, the binding equation

used to derive the K_D values described in this dissertation.

$$S_x = S_{max} \frac{[P]}{K_D + [P]}$$ [17]

One approach for assaying the binding of DNA to proteins that utilizes this theoretical framework is to measure changes in the fluorescence anisotropy of a fluorophore covalently bonded to a DNA substrate. In this experiment, the fluorophore is irradiated with linearly polarized light, and the resulting polarized fluorescence is measured by a detector fitted with a polarization filter. The fluorescence anisotropy, r, is then calculated by determining the difference in intensity of fluorescence parallel (I_{II}) and perpendicular ($I_{\perp}$) to the polarized light source and dividing by total fluorescence emission, as described in Equation 18 (202-205).

$$r = \frac{I_{II} - I_{\perp}}{I_{II} + 2I_{\perp}}$$ [18]

Because fluorescence anisotropy correlates with the rotational relaxation time of fluorescently-labeled DNA, observed anisotropy is proportional to the extent of binding in the solution (206). This particularly appealing because it permits data acquisition under equilibrium conditions–i.e., separation of bound and free species is not required to determine the ratio of bound ligand. Fluorescence anisotropy is particularly suitable for interactions with K_D values in the nanomolar to micromolar range; however, obtaining sufficient signal may prove challenging for binding reactions with sub-nanomolar K_D values. Extremely high affinity interactions may require a radiolabel-based method, such as the electrophoretic mobility gel shift assay.

Conclusions

While the essential process of priming in eukaryotes is well-defined, mechanisms for initiation, elongation, and primer transfer in eukaryotes remain obscure. The research described in this dissertation utilized a combination of high- and low-resolution structural and biochemical

approaches to evaluate primase architecture and activity through the catalytic cycle. Chapter Two describes results leading to the conclusion that the [4Fe4S] cluster of human DNA primase functions as a redox switch using DNA charge transport. Chapter Three expands on these results by further characterizing the crystallographic and solution-state properties of two p58C constructs in response to comments on the publication described in Chapter Two. Chapter Four shifts focus to the use of small-angle x-ray scattering to investigate primase domain rearrangements during the catalytic cycle. Finally, Chapter Five describes a model for primase action based on these results and proposes future experiments to test this model.

CHAPTER II

THE [4Fe4S] CLUSTER OF HUMAN DNA PRIMASE FUNCTIONS AS A REDOX SWITCH

USING DNA CHARGE TRANSPORT[2]

Introduction

The ability of DNA to transport charge over long distances represents an intriguing

potential regulatory mechanism in biology. DNA charge transport (DNA CT) provides a rapid

means of signaling among redox moieties coupled into the DNA duplex, as well as a mechanism

to sense the integrity of DNA (207-212). Remarkably, [4Fe4S] clusters, inorganic cofactors often

associated with biological redox chemistry (213, 214), are now being identified in proteins

involved in DNA replication (36, 116, 118, 162, 215). The eukaryotic DNA primase enzyme

responsible for initiation of replication on single-stranded DNA, for example, is a [4Fe4S]

cluster enzyme. The [4Fe4S] cluster in primase has been shown to be essential for activity (36,

116, 118, 215), but the role of this cofactor was unclear.

Rapid and accurate copying of genomic DNA in humans and other higher eukaryotes is

the product of high fidelity, processive, replicative DNA polymerases (6, 88, 112, 131). While

efficient, these enzymes are unable to initiate synthesis of the new complementary strand without

a short primer on the single-stranded DNA (ssDNA) template. The task of initiating synthesis of

the new DNA strand is the responsibility of a heterotetrameric complex of two specialized

---

[2] Portions of this section were previously published in the following publications:

- O'Brien E, Holt ME, Thompson MK, et al. The [4Fe4S] Cluster of Human DNA Primase Functions as a Redox Switch using DNA Charge Transport. *Science (New York, NY)*. 2017;355(6327):eaag1789. doi:10.1126/science.aag1789
- O'Brien E, Holt ME, Thompson MK, et al. Response to Comments on "The [4Fe4S] cluster of human DNA primase functions as a redox switch using DNA charge transport." *Science (New York, NY)*. 2017;357(6348):eaan2762. doi:10.1126/science.aan2762.

polymerases: DNA primase and DNA polymerase α (pol α), both of which were discovered to contain [4Fe4S] clusters (116, 118, 162). Primase, a DNA-dependent RNA polymerase, generates an initial 8–12 nucleotide RNA primer on ssDNA before handing the nascent primer-template off to pol α, which extends the primer by ~20 DNA bases before handoff to the more processive polymerases (δ and ε). X-ray crystal structures have been determined for all globular domains of primase and pol α as well as for the primase heterodimer and the heterotetrameric DNA polymerase α-primase (pol-prim) complex (36, 137, 143, 144). However, very little structural data have been published for primase in complex with catalytically relevant ligands. In particular, although the mechanism and structure of the catalytic subunit of primase have been extensively studied (135, 143, 216, 217), the chemistry behind primer handoff to pol α is poorly understood.

Eukaryotic primases are heterodimeric, composed of a catalytic subunit (p48) and a regulatory subunit (p58) (88, 122). The regulatory subunit contains a C-terminal domain (p58C) that is unique to eukaryotes and contains the [4Fe4S] cluster cofactor required for efficient priming (116, 118). The biochemical evidence of the role for the [4Fe4S] cluster in primase, in addition to the high energetic cost paid by cells to assemble and load this cofactor into an enzyme, argues for a functional rather than structural role for this cofactor (177, 213). To better understand the function of the [4Fe4S] cluster in p58C, we collaborated with the Barton group at the California Institute of Technology to characterize the redox properties of this cluster and identify its role in primer synthesis. In this collaboration, our lab generated mutant and wild-type p58C and primase that we then structurally characterized (p58C) and assayed for DNA binding affinity (p58C and primase). Electrochemistry and primase activity and elongation assays were completed on these variants by Elizabeth O'Brien, a graduate student in the Barton group.

Ultimately, these assays revealed that the iron-sulfur cluster in p58C acts as a redox switch to modulate the DNA binding affinity of p58C.

Materials and Methods

Protein expression and purification

The Y309F, Y345F, and Y347F p58C and Y309F and Y345F primase mutants were created using a Q5 site directed mutagenesis kit from New England Biolabs. Y345C p58 and p58C mutants were created according to standard QuickChange mutagenesis protocol. Primers for mutagenesis, described in Table 2.1, were purchased from Sigma Genosys. Starting plasmids are described in (218). Recombinant WT and mutant p58C were expressed and purified as previously described (36). For crystallization of the proteins, an extra affinity purification step was added before size exclusion chromatography. Both WT and mutant p58C were buffer exchanged into 20 mM HEPES (pH 7.2), 2 mM DTT and 200 mM NaCl. Each protein sample was then loaded onto a HiTrap® Heparin column and eluted with a gradient of 20 mM HEPES (pH 7.2), 2 mM DTT and 1 M NaCl. To express primase, full-length p48 and p58 in the PBG100 and multiple cloning site 2 of the pETDuet vectors, respectively, were cotransformed into BL21 (RIL) D3 *E. coli*. The culture was grown in Terrific Broth at 37° C until it reached an O.D. of 0.5–0.8, at which point the growth was transferred to an 18° C incubator. This was allowed to continue growing until it reached an OD of 1–1.5, at which point protein expression was induced with 0.5 mM IPTG. The cells were harvested after expressing at 18° C for 20 hours.

To purify primase, cells were lysed using a sonicator and spun at 50,000 rcf. Supernatant loaded onto a nickel column equilibrated in Buffer A (20 mM Tris, 300 mM NaCl, and 20 mM imidazole (pH 8.0). After washing with five column volumes of Buffer A, the primase was

eluted with Buffer B (20 mM Tris, 300 mM NaCl, and 300 mM imidazole (pH 8.0). H3C

protease was added to primase-containing fractions, and these were dialyzed in to Buffer A at 4°

C overnight. Primase was separated from the cleaved 6xHis tag (and uncleaved protein) by

repassing over the nickel column. This sample was further purified by carefully diluting to 150

mM NaCl, loading onto a HiTrap® Heparin column, and eluting with a gradient of 20 mM

HEPES (pH 7.2), 3% glycerol, and 1 M NaCl. This was followed by size exclusion

chromatography with a Superdex S200 column from GE Healthcare. Primase was eluted into a

final storage buffer containing 20 mM HEPES, 150 mM NaCl, 3% glycerol, pH 7.2.

Table 2.1. Oligos Used in This Study		
Y309F Forward Primer	5'–CCGAATGCAGTTTGGCCTATTTC–3'	
Y309F Reverse Primer	5'–CCTCCATGACGAAGATGG–3'	
Y345F Forward Primer	5'–TGATAAAGGTTTCTCTTACAACATCC–3'	
Y345F Reverse Primer	5'–AACTTGTCTGGATCCATC–3'	
Y345L Forward Primer	5'–CAAGTTTGATAAAGGTCTCTCTTACAACATCCGTC–3'	
Y345L Reverse Primer	5'–GACGGATGTTGTAAGAGAGACCTTTATCAAACTTG–3'	
Y345C Forward Primer	5'–CAAGTTTGATAAAGGTTGCTCTTACAACATCCGTC–3'	
Y345C Reverse Primer	5'–GACGGATGTTGTAAGAGCAACCTTTATCAAACTTG–3'	
Y347F Forward Primer	5'–AGGTTACTCTTTCAACATCCGTC–3'	
Y347F Reverse Primer	5'–TTATCAAACTTGTCTGGATCC–3'	
p58C Electrochemistry Substrates	Well-matched	Abasic Site
	5'–SH–GTCGTGCAACGTGTCTGCGC–3'	5'–SH–GTCGTGCAACGTGTCTGCGC–3'
	3'–CAGCACGTTGCACAGACGCGTAC–5'	3'–CAG_ACGTTGCACAGACGCGTAC–5'
Initiation Substrate	5'–AAAAAAAAAAAAAAAAAAAAAAAAAATAAAGAGAGAGAGAGAGAGAAAAGA–5'	
Elongation Substrate	5'–(T)$_{15}$(U)$_{16}$–3'	
	3'–(A)$_{31}$TAAGAAAAGAGAGAGAGAGAGAGAGAAAAGA–5'	
p58C Fuorescence Anisotropy Substrate	3'–GAGAGTTT–5'	
	5' [Flc]–TCTCTTCTCAAA–3'	
Primase Fluorescence Anisotropy Substrate	5' [Flc]–TTTTTTTTTTTTTTTTTTTTTTTTTTTT–3'	

Table 2.1. Mutagenesis, electrochemistry, primase activity assay, and fluorescence anisotropy DNA substrates used. Electrochemistry of p58C was performed on self-assembling monolayers of a 20-mer DNA duplex substrate with a 3-nt 5'- ssDNA overhang. A 50-nt ssDNA substrate with a single thymine base complementary to the α-^{32}P radiolabeled ATP was used in the primase initiation assay comparing wild type and CT-deficient full-length enzyme. A 2'-OMea RNA-primed ss/dsDNA substrate, containing a 31-nucleotide duplex segment and a 29 nucleotide 5'-ssDNA overhang was used to assay elongation. U = 2'-OMe rU, SH = -(CH$_2$)$_6$-SH, Flc = FITC

CD spectroscopy

The WT and mutant p58C variants were concentrated to 2 mg/mL and buffer-exchanged into 10 mM potassium phosphate (pH 7.2). The far-UV CD spectrum over the range 190–260 nm was acquired at room temperature using a Jasco J-810 spectrophotometer. Each spectrum is the average of three scans acquired with a scanning rate of 0.5 nm/s.

NMR spectroscopy

Spectra were acquired using a Bruker AV-III 800 MHz spectrometer equipped with a CPTCI single-gradient cryoprobe. ^{15}N-enriched p58C constructs at a concentration of 200 μM were prepared in a buffer containing 20 mM MES (pH 6.5), 50 mM NaCl, 2 mM DTT, and 5% (v/v) D_2O. Two-dimensional ^{15}N-^1H heteronuclear single quantum coherence (HSQC) spectra were acquired at 298 K with 2,048 and 128 points in the ^1H and ^{15}N dimensions, respectively. 64 scans were recorded in the direct (^1H) dimension for each point sampled in the indirect (^{15}N) dimension. Data were processed by Topspin (Bruker) and analyzed with Sparky (University of California).

X-ray crystallography

All crystallographic data was generated by Lauren Salay and Dr. Matthew Thompson in the Chazin group. Crystals of Y347F p58C were grown by hanging drop vapor diffusion at 21 °C from a drop composed of equal volumes of 75 mg/ml protein in 20 mM TRIS (pH 7.2) and 75 mM NaCl and reservoir solution containing 100 mM Tris (pH 8.5), 300 mM Li_2SO_4 and 20% PEG 3350. Crystals of Y345F p58C were grown in 20 mM MES (pH 6.5) and 50 mM NaCl and reservoir solution containing 100 mM Tris (pH 8.5) 150 mM Li_2SO_4 and 18% PEG 3350. Crystal

growth was stimulated by streak seeding with WT. Prior to data collection, crystals were soaked in mother liquor containing 20% glycerol and flash frozen in liquid nitrogen.

X-ray data were collected at Sector 21 (Life Sciences Collaborative Access Team) of the Advanced Photon Source at Argonne National Laboratory. All data were processed by HKL2000. The Y347F crystals belong to the P3 space group and were twinned. The Y345F crystals were not twinned and belonged to the C2 space group. Prior to phasing, the twinned data set was de-twinned using the Detwin program of the CCP4 program suite (219) by applying the twin law, h,-h-k,-l. Phasing of the diffraction data was done by molecular replacement using PHASER and PDB entry 3L9Q as the search model. Manual model building for the structure was performed using *Coot* model building software, and waters were placed with the *Coot* routine, Find Waters.

The final model was obtained by iterative cycles of model building in *Coot* and structure refinement using Refmac5 in the CCP4 suite of programs. The final model of Y347F was a dimer of dimers with four Y347F p58C subunits in the asymmetric unit. The final model of Y345F was a dimer with two Y345F subunits in the asymmetric unit. All protein figures were prepared with Chimera. Data collection and refinement statistics are given in Table 2.2.

Data collection		
Variant	Y345F	Y347F
PDB Entry	5I7M	5DQO
Space Group	C2	P31
Cell dimensions		
a, b, c (Å)	109.9, 52.7, 89.2	60.40, 60.40, 246.73
α, β, γ (°)	90, 115.25, 90	90, 90, 120
Temperature (K)	100	100
Wavelength (Å)	0.987	0.999
Resolution (Å)	49.7 - 1.93	44.14-2.48
Reflections		
Total	142,467	167,361
Unique	35,604	19,561
Completeness (%)[a]	99.9 (99.8)	99.1 (89.6)
R_{merge} (%)[b]	8.7 (58.4)	5.8 (49.3)
I/σI	18.7 (2.1)	27.6 (2.1)
Redundancy	4.2 (4.0)	4.5 (2.9)
Refinement		
R_{work}/R_{free} (%)[c]	21.0/24.9	17.3/19.1
No. of residues		
Protein	314	628
4Fe-4S	2	4
Solvent	81	41
Average B-factor (Å^2)		
Protein	28.8	53.7
4Fe-4S	21.5	46.2
Solvent	27.7	55.7
RMSD bonds (Å)	0.02	0.02
RMSD angles (°)	2.11	1.91
Ramachandran[d]		
Most favored	311	530
Allowed	3	66
Disallowed	0	0

[a] Values in parentheses are for the highest resolution shell. [b] $R_{merge} = \Sigma(|I-\bar{I}|/ \Sigma I \times 100$. [c] $R_{work} = \Sigma(|F_o-F_c|/ \Sigma F_o \times 100$ where F_o is the observed structure factor amplitude and F_c is the calculated structure factor amplitude. [d] Values are numbers of residues

Figure 2.2. Crystallographic data collection and refinement statistics.

Fluorescence Anisotropy

DNA binding to wild-type and mutant p58C and primase was measured by monitoring

the increase in fluorescence anisotropy as protein was added to a solution containing 5'-FITC-

labeled DNA (see Table 2.1 for substrates). The p58C DNA substrate was annealed using a

buffer containing 20 mM MES (pH 6.5) and 75 mM NaCl. In each case, an increasing

concentration of protein was added to a solution containing DNA substrate at a concentration of

50 nM (p58C) or 20 nM (primase). Polarized fluorescence intensities were measured using

excitation and emission wavelengths of 485 nm and 520 nm using a SpectraMax M5 microplate reader (Molecular Devices).

Cyclic voltammetry

All electrochemical assays were completed by Elizabeth O'Brien. Wild type and mutant p58C samples were stored prior to experiments in p58C storage buffer. Concentrations of [4Fe4S] cluster-containing p58C or mutants were measured using UV-Visible spectroscopy, by absorbance of the [4Fe4S] cluster at 410 nm (extinction coefficient = 17000 M^{-1} cm^{-1}) (220). Aliquots of stock samples (45–90 µL) were deoxygenated using argon bubbling for 4–5 minutes. Samples were then transferred into the anaerobic chamber. Before deposition onto the gold electrode surface, p58C/mutant samples were diluted to a molar concentration of 16 µM [4Fe4S] p58C variant with previously deoxygenated p58C storage buffer. Samples were deposited onto multiplex chip quadrants in 20 µL volumes initially, with the remaining sample deposited in a well of bulk solution above the chip surface.

All electrochemistry was performed using a CHI620D potentiostat and 16-channel multiplexer (CH Instruments), in an anaerobic glove chamber. Multiplex gold electrodes were part of a three-electrode system with an external Ag/AgCl reference electrode (Bioanalytical Systems) and platinum counter electrode. Cyclic voltammetry scans were performed at 100 mV/s scan rates, over a potential range of +0.412 V to −0.288 V vs. NHE. Bulk electrolysis on DNA was performed at an applied potential of +0.412 V vs. NHE for all electrochemical oxidation reactions and −0.188 V vs. NHE for all electrochemical reduction reactions. The oxidizing potential was applied for 8.33 minutes for single oxidation reactions on a surface, and 5.83 minutes or 6.67 minutes for the iterative oxidation cycles of p58C variants. The reducing

potential was applied for 8.33 minutes in all electrochemical reduction reactions. All bulk electrolysis and cyclic voltammetry was performed in previously deoxygenated p58C storage buffer (20 mM Tris, pH 7.2, 75 mM NaCl). Charge transfer (nC) in the cathodic peak of oxidized samples' CV scans was assessed using the area under the current wave of the reduction signal.

All p58C variants were compared over three trials of oxidation at +0.412 V vs. NHE (two oxidation reactions after 500 s of applied potential, one reaction after 400 s of applied potential). The charge transfer in the bulk electrolysis curves was calculated as the area under the current curve plotted versus time for the bulk electrolysis reaction, as the current decays to a constant value. The percent recovery of bulk electrolysis charge in the CV peak after oxidation was averaged over three trials for each variant. Error bars represent the standard deviation of the average fraction of charge recovered.

Primase activity assays

All activity assays were performed by Elizabeth O'Brien. These assays were completed in an anaerobic glove chamber (Coy Products) using previously deoxygenated buffer, protein, DNA substrates, and nucleotide triphosphates. All reagents were deoxygenated by argon gas bubbling after thawing from storage temperature immediately prior to assay. Buffer was deoxygenated by bubbling argon gas for several hours (approx. 1 hour per mL buffer) and subsequent incubation in the glove chamber atmosphere for at least 1–2 days prior to the assay. Initiation assays were performed using the 50-nt Initiation Substrate listed in Table 2.1. Elongation assays were performed with the Elongation Substrate in Table 2.1, in which a 2'-OMe RNA/DNA primer is annealed to a 60-nt ssDNA complement strand. Initiation assays (18 µL total volume) contained 250 nM ssDNA Initiation Substrate, 1 µM [α-^{32}P ATP] (Perkin Elmer), 112 µM [CTP] (Sigma),

188 µM [UTP] (Sigma), and 400 nM primase or primase variant, in 50 mM Tris-HCl, pH 8.0, 3 mM $MgCl_2$.

Reactions were initiated by the addition of primase or primase variant enzyme and incubated anaerobically at 37°C. Aliquots (5.5 µL) were removed from the primase reactions after 5 minutes, 10 minutes, and 30 minutes of incubation for wild type/mutant comparison assays and after 30 minutes, 60 minutes, and 120 minutes for wild type primase well-matched and mismatched primer comparison assays. All samples were quenched by addition of an equal volume of 1% SDS (Sigma), 25 mM EDTA (Sigma) solution. Samples were then removed from the glove chamber after quenching and heat denatured at 70°C. Products were then purified by size exclusion, using mini Quick Spin Oligo columns (Roche Diagnostics) for wild type/mutant comparison assays and Micro-Bio Spin™ P-6 columns (BioRad Laboratories) for wild type well-matched/mismatched primer comparison assays. Radioactivity counts for products were measured using a LS5000TD scintillation counter (Beckman Instruments). Products were dried in vacuo, resuspended in 2 µL loading dye (Xylene cyanol, bromophenol blue), and further annealed for 1 minute at 90°C before gel separation.

Primase/mutant comparison elongation assays were performed under identical conditions, in the presence of 500 nM Well-Matched Elongation Substrate, 120 µM CTP (Sigma), 180 µM UTP (Sigma), and 1 µM α-^{32}P ATP (Perkin Elmer). The primase/variant concentrations used in elongation assays were 200 nM, 400 nM, 600 nM, or 800nM. Primase assays for elongation in the presence of a mismatch were performed with 500 µM Well-Matched Elongation Substrate or Mismatched Elongation Substrate, 200 µM CTP (Sigma), 100µM UTP (Sigma), and 1 µM α-^{32}P ATP (Perkin Elmer). Assays were performed anaerobically and quenched/purified in a manner identical to initiation assays.

All primase/primase variant assay products were separated by denaturing polyacrylamide gel electrophoresis (20% polyacrylamide). Separated products were visualized using phosphorimagery on a Typhoon FLA 9000 imager (GE Healthcare) and quantified using ImageQuant TL software. Products synthesized by different primase variants were directly compared using measured ^{32}P counts detected by software. Initiation and Elongation Substrates (Table 2.1) were designed to contain a single base complementary to the radiolabeled NTP, α-^{32}P ATP. Quantification could thus have a basis 1:1 product: radioactivity ratio, as the sole purine NTP and would be the strongly preferred site of initiation and an optimal 5'- elongation site (125, 126). Product sizes were assigned by comparison with a 10/60 Oligo Length Standard (Integrated DNA Technologies). Single-stranded oligonucleotides in standard were labeled by incubation for 1 hour at 37°C with γ-^{32}P ATP (Perkin Elmer) and T4 polynucleotide kinase enzyme (Roche Diagnostics). Labeled standard was purified using size exclusion chromatography with mini Quick Spin Oligo columns (Roche Diagnostics). All product quantifications were averaged over three trials for p48/p58Y345F and p48/p58Y345C and six trials for p48/p58. Error bars represent the standard deviation of average values obtained over these trials.

Results

p58C engages in charge transport

To study the DNA-bound redox properties of enzymes with [4Fe4S] clusters, the Barton group employs DNA-mediated electrochemistry, a robust method for directly measuring DNA CT in the ground state (Figure 2.1A/B) (210, 221, 222). An alkanethiol-terminated, annealed duplex DNA (dsDNA) substrate is deposited on a gold surface, facilitating covalent linkage of

the DNA to the gold through the thiol moiety. The gold is passivated using β-mercaptohexanol and becomes the working electrode in a three-electrode cell, with an external Ag/AgCl reference electrode and a platinum counter electrode (210, 221, 222). Charge transport through the stacked bases of dsDNA between the gold surface and a redox-active species bound at the distal end of the DNA can be measured using this platform under solution conditions. Cyclic voltammetry (CV) is employed to measure changes in current over a range of applied potentials. Earlier electrochemical studies of the base excision repair glycosylase Endonuclease III (EndoIII) using CV have shown that binding of the protein to the DNA polyanion shifts the redox potential of the [4Fe4S] cluster 200 mV negative to ~80 mV vs. NHE, into the physiological range of cellular potentials, activating the cluster for redox chemistry (207, 212). Importantly, this potential shift corresponds thermodynamically to a 1000-fold increase in DNA binding affinity for the oxidized $[4Fe4S]^{3+}$ state of EndoIII relative to the reduced $[4Fe4S]^{2+}$ state.

The p58C domain of human DNA primase, independently of the rest of the enzyme, binds a primed template DNA (Table 1) with modest affinity (Kd = 5.5 ± 0.5 µM) when initially present as isolated in the reduced, EPR-silent $[4Fe4S]^{2+}$ state (36, 116, 118). To investigate whether the p58C [4Fe4S] cluster oxidation state affects DNA binding, as observed in Endonuclease III (212), we used DNA electrochemistry. Experiments are carried out using a sixteen-electrode multiplexed chip, which allows for robust, reproducible, and internally consistent measurements of various conditions on a single surface (Figure 2.1A) (207, 211, 221, 222). Bulk electrolysis is used to convert a sample to a uniform redox state in this setup, and it is performed by passage of current through the DNA at a constant applied potential. This technique facilitates the oxidation or reduction of the DNA-bound protein with a direct transfer of electrons through the DNA, eliminating the need for exogenous chemical oxidants or reductants that could

damage the protein.
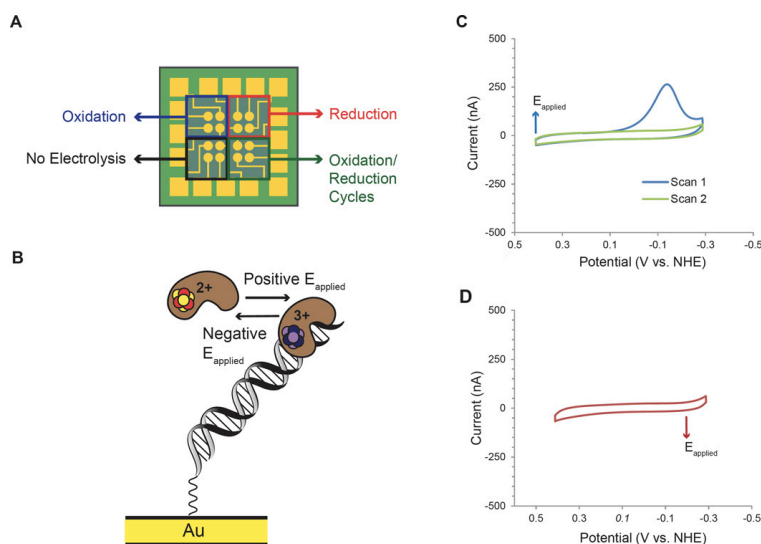
Figure 2.1. Oxidized [4Fe4S]3+ and reduced [4Fe4S]2+ p58C display different behavior on DNA-modified electrode. A) Multiplex chip with 16 DNA-modified Au electrodes (circles, center.) This platform facilitates direct comparison of oxidized, reduced, unaltered, and iteratively oxidized samples, on four separate quadrants of a single surface. B) The cartoon depicts the effects of electrochemical oxidation and reduction on p58C DNA binding and redox activity. C) CV of electrochemically oxidized p58C. After electrochemical conversion (Eapplied = 412mV vs. NHE) of the sample at the electrode/solution interface to the [4Fe4S]3+ state, CV scans display a large cathodic peak only in the initial scan to negative, reducing potentials. D) CV of electrochemically reduced p58C. After electrochemical conversion (Eapplied = −188 mV vs. NHE) of the sample to the [4Fe4S]2+ state, CV scans show no electrochemical signal on DNA. Electrochemistry was performed on 16μM p58C in 20 mM Tris, pH 7.2, 75 mM NaCl, 100mV/s for CV scans, using a Ag/AgCl reference electrode.

For DNA electrochemistry on p58C, an alkane-thiol modified primed template DNA substrate (Table 2.1) was used. All experiments are performed in an anaerobic environment with deoxygenated reagents to prevent atmospheric cluster oxidation (223) and to ensure control of the redox state of the p58C samples on the electrode. As in earlier studies (212), we confirmed that p58C electrochemistry is DNA-mediated by observing a reversible p58C CV signal in the presence of ATP, which is attenuated when an intervening abasic site perturbs the base stacking of the duplex DNA substrate (Figure 2.2). The reversible redox signal observed when p58C is

40

bound to both DNA and ATP is likely a result of the p58C [4Fe4S] cluster being better coupled

into the DNA duplex for CT when bound to DNA and a nucleotide triphosphate (NTP), both

necessary substrates for primase activity; the improved coupling likely is the result of a

conformational change. Increased coupling of the [4Fe4S] cluster to duplex DNA was previously

observed for the [4Fe4S] helicase XPD (*S. acidocaldarius*) upon binding ATP, a substrate

necessary for XPD activity (224).
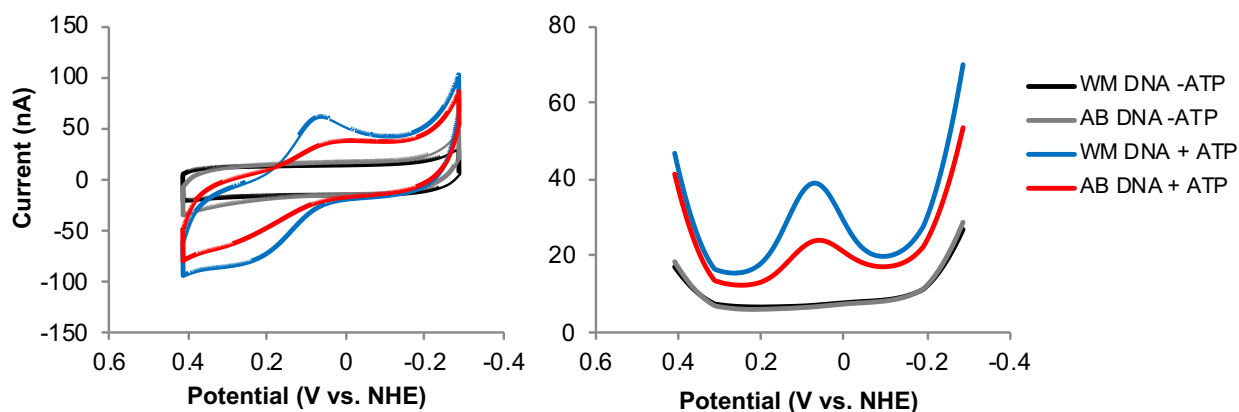
Figure 2.2. p58C electrochemistry is DNA-mediated. Cyclic voltammetry (a) and Square
Wave Voltammetry (b) signals of p58C bound to of well-matched DNA (WM DNA) or DNA
containing an abasic site (AB DNA) in the duplex segment, in the presence and absence of an
NTP pool. The p58C domain displays a reversible redox signal centered at 140-150mV vs. NHE
in the presence of ss/dsDNA and 500 µM [ATP]. This suggests that the [4Fe4S] cluster in DNA
primase is capable of undergoing reversible redox activity when the enzyme is in its active form,
bound to DNA and NTPs. All electrochemistry was performed in anaerobic conditions with a
Ag/AgCl reference, on 10 µM [p58C] in 20mM Tris, pH 7.2, 75 mM NaCl. Cyclic voltammetry
scan rate: 100 mV/s, square wave voltammetry scan frequency: 15 Hz.

To electrochemically assay p58C in the presence of DNA only, oxidized and reduced

samples are generated and subsequently compared on a single surface using bulk electrolysis

followed by CV scanning (Figure 2.1 and Figure 2.3). By passing sufficient current through the

DNA-modified electrode at an oxidizing (412 mV vs. NHE) or reducing (−188 mV vs. NHE)

potential, p58C is converted to the desired redox state (Figure 2.1B and Figure 2.3). We observe

no redox signal by CV for electrochemically unaltered p58C, indicating that the EPR-silent, [4Fe4S]$^{2+}$protein obtained upon isolation is not electrochemically active on DNA (Figure 2.4). After oxidation by bulk electrolysis, however, a large cathodic peak between −130 and −140 mV vs. NHE appears in CV (100 mV/s scan rate) during the initial scan to negative potentials (Figure 2.1C). Importantly, the electrochemical signal through DNA is lost after scanning to reducing potentials. The signal observed in the CV scan to negative potential for the oxidized p58C sample corresponds to a reduction event, which we assign to the electron transfer reaction in which the tightly bound, electrochemically active [4Fe4S]$^{3+}$ p58C is converted to a more weakly-associated, electrochemically inactive [4Fe4S]$^{2+}$ state. Consistent with this observation, CV after reduction displays no redox signal (Figure 2.1D), as observed initially for the native protein, which we attribute to the lower affinity for DNA of the [4Fe4S]$^{2+}$ state.
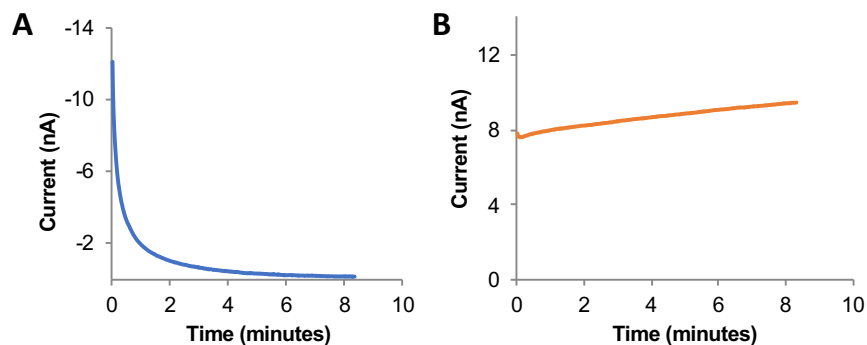
Figure 2.3 Bulk electrolysis of p58C. Oxidation (**A**) and reduction (**B**) of p58C by bulk electrolysis on a DNA-modified multiplex Au electrode. Potentials shown on oxidation/reduction plots were applied to convert the sample at the DNA/solution interface to the desired [4Fe4S] cluster oxidation state.   Bulk electrolysis performed on 16 μM p58C using a Ag/AgCl reference electrode, in 20 mM Tris buffer, pH 7.2, 75 mM NaCl, in anaerobic conditions.
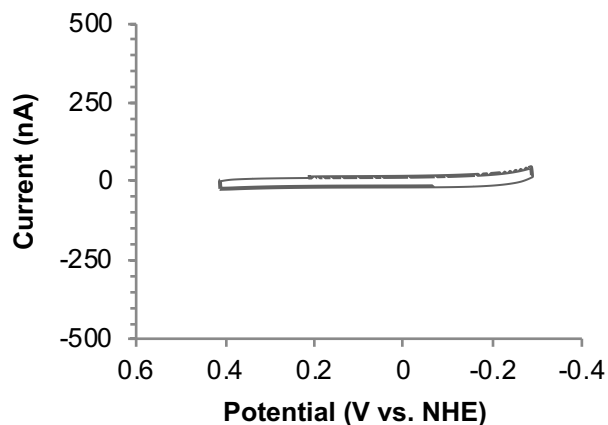
Figure 2.4. p58C does not produce a redox signal on ss/dsDNA in the absence of electrochemical alteration. CV performed in anaerobic conditions on 16μM p58C using a Ag/AgCl reference electrode, in 20mM Tris buffer, pH 7.2, 75mM NaCl. CV scans were conducted at a 100mV/s scan rate.

Progressively larger CV signals appear with each iterative oxidation performed under identical electrolysis conditions. The change in CV signal under constant oxidation conditions suggests that oxidation brings more molecules of p58C to the DNA substrate on the electrode each time that electrolysis is performed (Figure 2.5). Previously oxidized p58C molecules, together with newly oxidized p58C generated during each electrolysis at 412 mV vs. NHE, appear in the CV signal obtained after each oxidation. Moreover, it is not feasible on our experimental timescale that two serial protein diffusion events, diffusion of previously oxidized p58C molecules away from the DNA, followed by diffusion of a new sample to the DNA replacing them, can occur. This evidence therefore strongly suggests that the redox switch in the oxidation state of the [4Fe4S] cluster modulates DNA binding and coupling of the [4Fe4S] cluster into the DNA duplex for CT. The molecular basis for how oxidation controls this increase in DNA affinity is not known; progress to define this mechanism has been inhibited by the inability to produce sufficient quantities of uniformly oxidized, [4Fe4S]$^{3+}$ p58C. Part of the increase in DNA affinity is likely electrostatic in origin, associated with the oxidized protein

being bound to the DNA polyanion. It is also likely, however, that a conformational change is associated with oxidation.

Oxidation Step	Cathodic Peak Charge Transfer (nC)
1	65.4
2	112.4
3	116.4
4	151.1
5	170.9
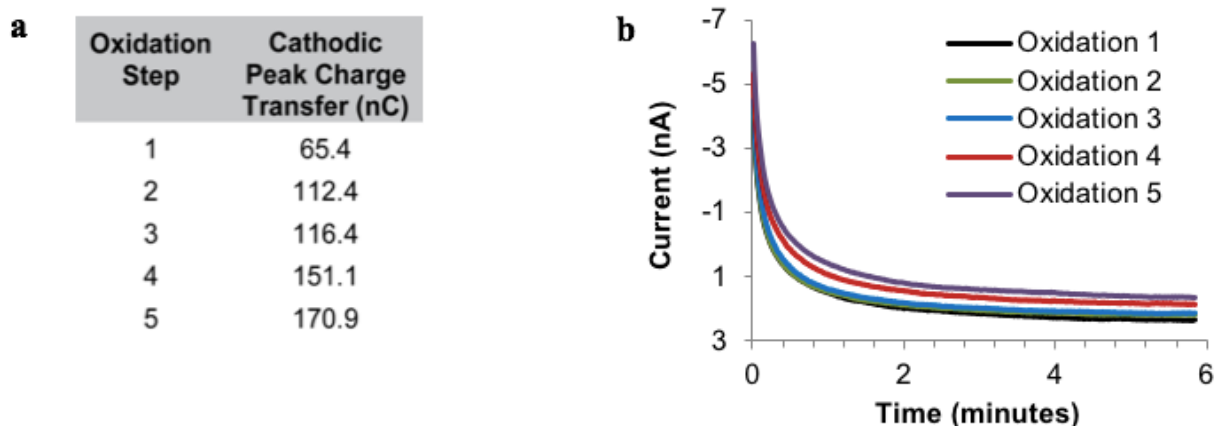
Figure 2.5. Increased charge transfer is observed with iterative oxidations. A. Charge transfer (nC) in the cathodic peak signal for each successive CV scan of iteratively oxidized p58C. CV scans performed at 50 mV/s scan rate. Charge transfer increases during successive oxidations, due to more p58C available at the DNA/solution interface during later trials. B. Iterative oxidation reactions on a single DNA-modified multiplex electrode surface. Oxidation reactions of p58C indicate that charge passes, converting a reduced species to an oxidized species at an applied potential of +412 mV vs. NHE. This shows that after reduction in CV, p58C can be oxidized again on a DNA-modified surface. All scans performed on 16 μM p58C in 20 mM Tris, pH 7.2, 75 mM NaCl.

Mutation of Y345 inhibits charge transport efficiency

Iterative, electrochemically controlled oxidation and reduction cycles demonstrate the ability of p58C to switch between weak and tight DNA binding upon a change in the oxidation state of the [4Fe4S] cluster under the control of DNA-mediated CT through the DNA electrode. For this DNA-mediated redox switch reaction to occur, however, a tunneling pathway (225) through p58C is necessary to move charge the ~25 Å distance between the primed template DNA binding site and the [4Fe4S] cluster cofactor (36, 137). Residues within p58C must therefore shuttle charge through the insulating protein matrix to mediate the DNA binding switch reaction. Tyrosine residues, whose aromatic rings can stabilize protein radicals, serve as redox

mediators in a variety of proteins (226). Moreover, conservation of tyrosine residues across p58C

domains of eukaryotic primases suggests potential participants in such a CT pathway through the

protein to the DNA interface.

In all structures of human p58C, three conserved tyrosine residues (Y309, Y345, Y347)

form a pathway from the cluster to the putative DNA binding surface (Figure 2.6). Given its low

ionization energy of 8.5 eV in solution relative to 9.4 eV for phenylalanine (226) and an average

of ~9.6 eV (225) for aliphatic residues, tyrosine is able to mediate electron/hole transfer through

proteins; the phenol side chain of tyrosine can more easily hold a radical than aliphatic residue

side chains. Tyrosine residues have additionally been shown to mediate hopping reactions

through other proteins to facilitate biological redox chemistry (225). The chain of tyrosines in

p58C is additionally within a feasible range for tunneling (225) between bound DNA and the

[4Fe4S] cluster. For these reasons, the three conserved tyrosines in p58C were viewed as a likely

conduit to facilitate electron transfer from the cluster to the DNA.
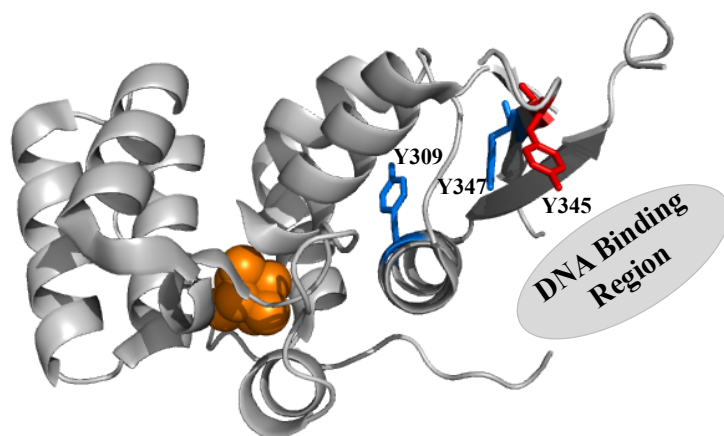
Figure 2.6. Tyrosine residues form potential charge transfer pathway from DNA binding region to [4Fe4S] cluster. Crystal structure of p58C (PDB ID: 3L9Q), with p58C in grey, iron-sulfur cluster in orange, Y309 and Y347 in blue, and Y345 in red.

To test the proposed CT pathway, we designed and isolated three human p58C Y-F

variants (Y309F, Y345F, and Y347F), which were characterized in the same manner as the wild type protein. All mutants are loaded with [4Fe4S] cluster to a similar degree as WT, as reflected in the UV-Visible absorbance 280 nm/410 nm ratio (Figure 2.7B). To validate proper folding of the tyrosine variants, circular dichroism was collected for all mutants and, as we can see from Figure 2.7A, the distribution of secondary structure in all mutants is essentially identical to wild type.
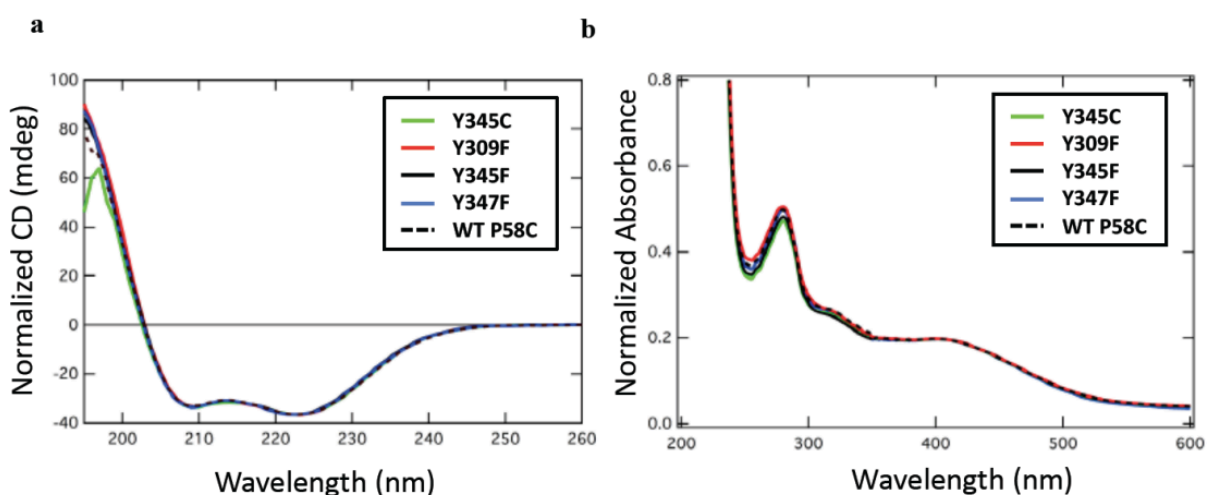
Figure 2.7. Tyrosine mutations do not affect p58C cluster loading or perturb secondary structure. A. Circular Dichroism (CD) spectroscopy of WT and mutant p58C indicate indicate the mutations do not perturbe elements of secondary structure. All spectra normalized to 222 nm. B. UV-Visible spectroscopy of WT and mutant p58C shows similar 280 nm/410 nm absorbance ratios, indicating similar degrees of [4Fe4S] cluster cofactor loading in all variants. All spectra normalized to WT at 410 nm. B.

To further investigate potential perturbations to the tertiary structure of p58 caused by mutation of Y345, ^{15}N, ^{1}H-HSQC spectra were collected for the Y345F mutant (Figure 2.8). In an overlay of the Y345F and WT p58C spectra, we observe a number of chemical shift perturbations, which is consistent with changes in the electrochemical environment caused by the tyrosine to phenylalanine mutation. Overall, however, the peak distribution is very similar to that of wild-type p58C, suggesting that these mutations do not impact the tertiary structure of p58C.
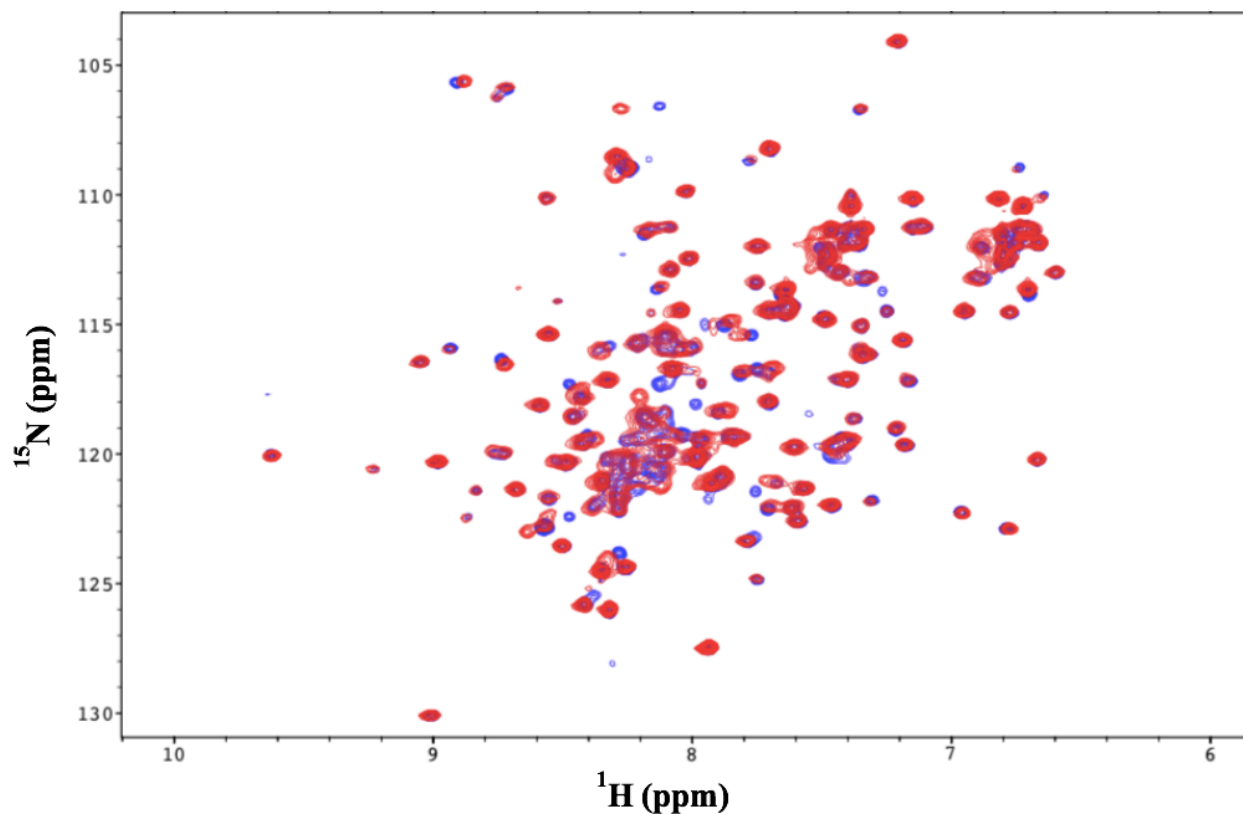
Figure 2.8. Y345F mutation does not perturb p58C tertiary structure in solution. ^{15}N-^{1}H HSQC spectra of WT (black) and Y345F p58C (red) were acquired at 25 °C on a Bruker AV-III spectrometer operating at 800 MHz. The samples contained 100 μM protein in a buffer containing 20 mM MES (pH 6.5), 50 mM NaCl, 2 mM DTT, and 5% ^{2}H$_2$O.

We also determined an x-ray crystal structure of p58C Y345F (Table 2.2, Figure 2.9).

The presence of the cluster in the mutant and the RMSD from the wild type protein of only 0.22

Å for the Y345F variant confirms that the structure is not perturbed. Critically, the crystal

structure reveals that the F345 aromatic ring in the mutant adopts the same orientation as the

Y345 ring in wild type p58C (Figure 3B). We also determined the x-ray crystal structure of the

Y347F variant, with a similarly low RMSD (0.33Å) relative to wild type p58C.
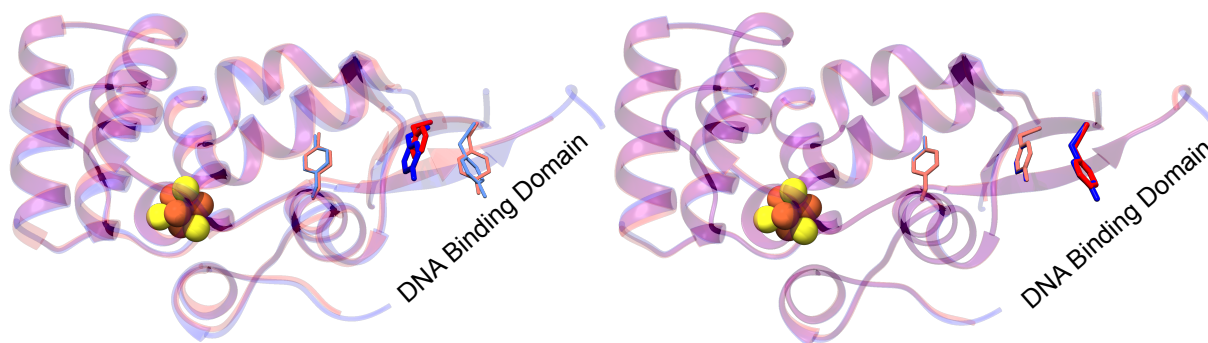
Figure 2.9. Y345F and Y347F mutations do not perturb crystal structure of p58C.
Superposition of the WT p58C (PDB 3L9Q) and Y347F p58C (PDB 5DQO) structures (left) and superposition of the WT p58C (PDB 3L9Q) and Y345F p58C (PDB 517M) structures (right). In both panels, WT p58C is colored blue and the tyrosine to phenylalanine mutant is red.

DNA binding measurements using *in vitro* fluorescence anisotropy under aerobic conditions (Figure 2.10) demonstrate that the mutations do not significantly affect the binding of p58C to a range of DNA substrates. Because the structures and biochemical properties of the mutants are the same as WT p58C, but the electron/hole transfer properties of tyrosine and phenylalanine are different, these single-atom mutations provide a powerful means to investigate the importance of the CT pathway through p58C and its effect on the DNA-mediated redox switch in primase.

Dissociation constants for p58C	
Construct	**K_D (µM)**
WT	5.6 ± 0.5
Y309F	4.0 ± 0.4
Y345C	2.7 ± 0.6
Y345F	4.6 ± 0.5
Y347F	4.1 ± 0.4
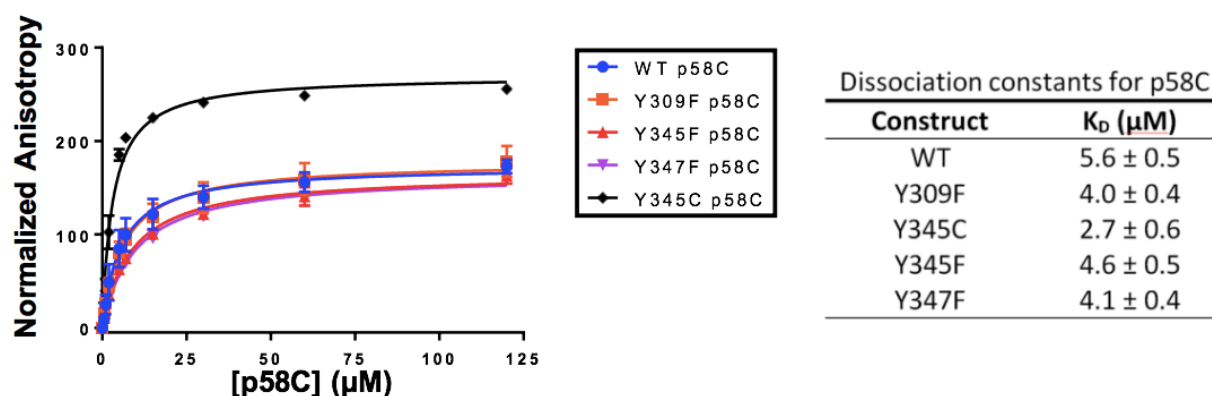
Figure 2.10. WT and mutant p58C domains have similar DNA binding affinities with low micromolar dissociation constants. Affinities were measured by fluorescence anisotropy of FITC-labeled DNA substrates in the presence of atmospheric oxygen in 20 mM MES, pH 6.5, 50 mM NaCl. Background was subtracted from anisotropy values prior to plotting. K_D values are averages over three measurements for each variant.

48

Oxidation and reduction using bulk electrolysis, followed by redox cycling during CV scanning were performed on all three Y-F variants using an identical multiplex chip surface as for wild type p58C (Figure 2.11). Though the general DNA-mediated electrochemical behavior of the mutants is the same as for the WT protein, it was observed that all three mutants are charge transfer-deficient compared to WT p58C. As a control that the phenomena we observed are not unique to Y-F substitution, we also assayed p58C Y345C, a somatic mutation discovered in a gastric tumor (227). Oxidation of each mutant produces a cathodic peak at potentials similar to wild type in CV scanning after electrolysis, with a much smaller signal height and a slightly positive shift in average cathodic peak potential, both of which suggest that the charge transfer pathway facilitating conversion from the native reduced, $[4Fe4S]^{2+}$ state to the oxidized, $[4Fe4S]^{3+}$ state is attenuated with mutation. The mutants of p58C are thus CT-deficient, demonstrating that all three conserved tyrosine residues aid in shuttling charge between the [4Fe4S] cluster and bound DNA.

A

B

C

	E_{peak} (mV vs. NHE)
WT	-138±3.4
Y345F	-137±4.8
Y309F	-136±16
Y347F	-137±1.7
Y345C	-128±8.7
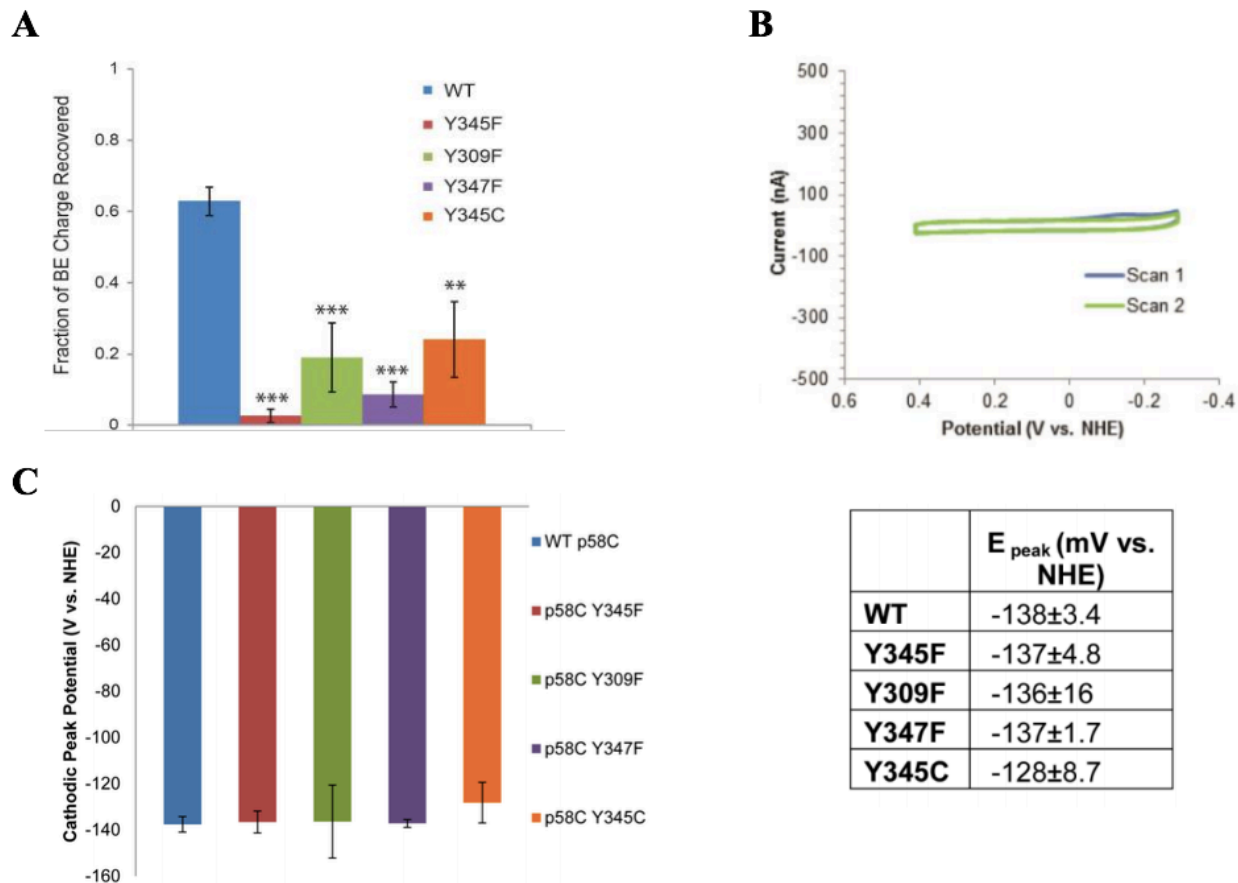
Figure 2.11. Tyrosine mutants exhibit charge transfer deficiency. A. WT p58C domain recovers significantly more (63±4%) bulk electrolysis charge than the tyrosine mutants, suggesting that perturbation of the charge transfer pathway impedes participation in redox chemistry when bound to DNA. B. Electrochemically oxidized p58C tyrosine mutants display a cathodic peak between -150mV and -160mV vs. NHE after oxidation at an applied potential of +412 mV vs. NHE; the peak signal is smaller than that observed for wild type. C. Change in potential value for cathodic peak after p58C oxidation, WT and tyrosine mutants of p58C. Potential measurements from three CV trials after p58C electrochemical oxidation show that the WT potential is more negative on average, but the difference between WT and mutant p58C cannot be determined to a significant degree outside the margin of error. This suggests that the mutants are more likely to be in the reduced $[4Fe4S]^{2+}$ state than the WT protein with an intact charge transfer pathway, as expected. Potential values are reported as mean ± SD over n= 5 trials. All bulk electrolysis reactions and CV scans were performed on 16 μM p58C/mutant in 20 mM Tris, pH 7.2, 75 mM NaCl at a 50 mV/s scan rate for CV, using a Ag/AgCl reference electrode. Mean ± SD of n=3 scans per variant, ** = 0.001<p<0.0005, *** = p<0.0005, student's t-test.

Charge transport-deficient mutants do not appropriately regulate primer counting in vitro

To investigate the connection between the DNA-dependent redox switch in p58C and

priming activity, we assayed *in vitro* initiation and elongation activity of the full-length primase

heterodimer and variants containing p58Y345F and p58Y345C mutations. Prior to testing

primase activity, both variants were found to bind ssDNA template similarly to WT primase

(Figure 2.12).
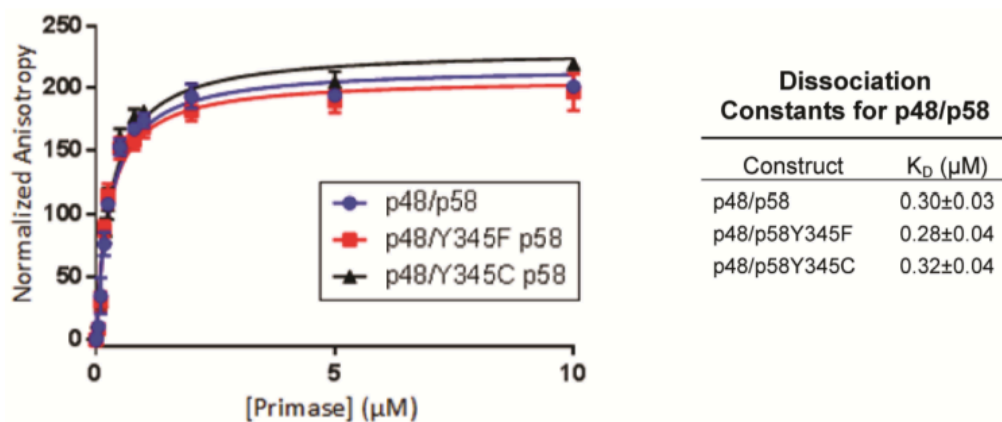
Figure 2.12. WT and mutant primase have similar DNA binding affinities with high nanomolar dissociation constants. Affinities were measured by fluorescence anisotropy of FITC-labeled DNA substrates in the presence of atmospheric oxygen in 20 mM HEPES, pH 7.5, and 75 mM NaCl. Background was subtracted from anisotropy values prior to plotting. K_D values are mean ± SD over n=3 measurements for each variant.

Anaerobic nucleotide incorporation assays were performed to test primase activity,

measuring incorporation of α-^{32}P labeled ATP on ssDNA for initiation and 2'-OMe RNA-primed

DNA for elongation. Anaerobic conditions for these assays are necessary when measuring the

effect of the redox switch on priming because atmospheric oxygen can oxidize the cluster when

the protein is not bound to DNA. In the absence of oxygen, the switch mediated by a pathway of

conserved tyrosine residues is thus the sole means through which to convert the enzyme from its

native $[4Fe4S]^{2+}$ state to its tightly bound, CT-active $[4Fe4S]^{3+}$ state. We additionally note that in

our assays, primase cannot truncate its products by handing them off to pol α; instead, we

observe a mixture of short primers and fully elongated 'primer-multimer' length products (Figure

2.13) (37, 138).
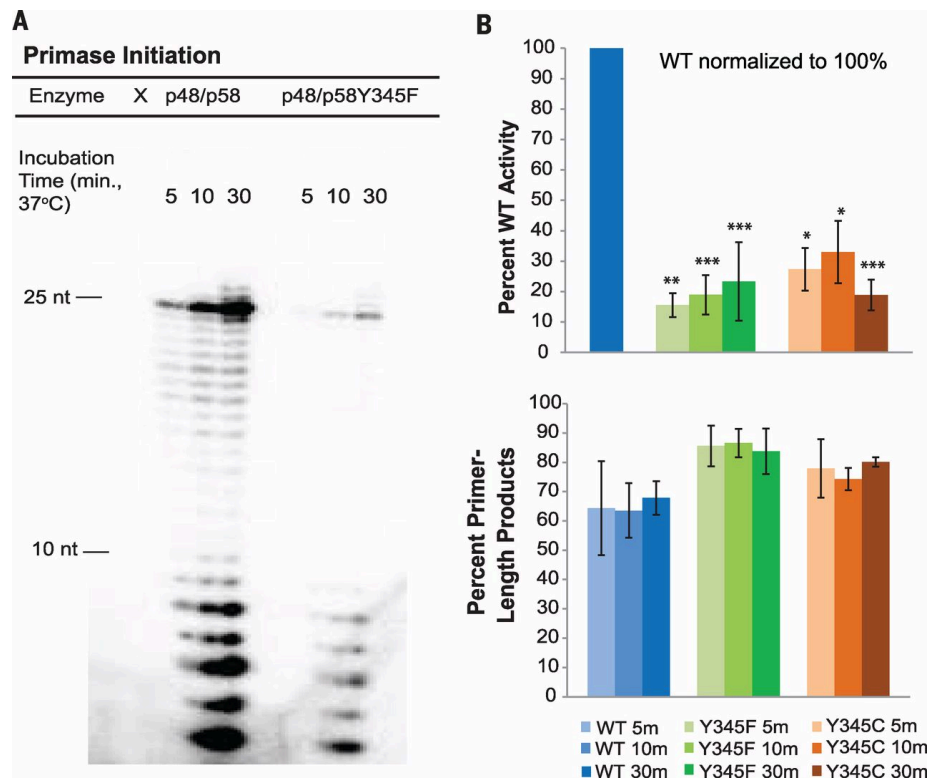
Figure 2.13. Redox switching plays a role in primase initiation. A. Gel separation of products for wild-type p48/p58 and p48/p58 Y345F reactions on ssDNA. The wild-type enzyme is significantly more active on ssDNA than either mutant. B. Quantified products for wild-type p48/p58, p48/p58 Y345F, and p48/p58 Y345C initiation assays. Top: Mutants synthesize 15 to 35% of wild-type products on average. Bottom: Mutant primase synthesizes shorter products on average. Primer-length products are defined as products 7 to 10 nt in length. Initiation assays were performed anaerobically with 250 nM ssDNA, 1 μM [α-^{32}P]ATP, 112 μM CTP, 188 μM UTP, 400 nM enzyme in 50 mM Tris (pH 8.0), and 3 mM MgCl$_2$ at 37°C. Data are means ± SD of $n \geq 3$ trials; *$P < 0.005$, **$P < 0.001$, ***$P < 0.0005$ (Student t test).

To characterize initiation, we used a 50-nt ssDNA substrate (Table 2.1) containing exactly one complementary base for the radiolabeled nucleotide triphosphate, α-^{32}P ATP. We found that WT primase has much higher overall initiation activity than both mutants tested, including the single-atom p48/p58Y345F variant (Figure 2.13). CT-deficient primase mutants are significantly less active on ssDNA, synthesizing only 15–35% of the WT initiation products over multiple trials. Compared to WT primase, a larger portion of the products synthesized by the

mutants is < 10 nucleotides in length, suggesting that the rate-determining initiation step in product synthesis is inhibited for mutants deficient in the redox switch. The significant difference in initiation activity for the WT and CT-deficient mutant primase enzymes suggests that a redox switch changing the oxidation state of the [4Fe4S] cluster in p58C is crucial for the reaction to begin primer synthesis. Additionally, the difference in activity on ssDNA caused by a very subtle mutation in the CT pathway suggests that charge migration from bound DNA to the [4Fe4S] cofactor, through the insulating protein matrix of p58C, plays a significant role in mediating the rate-determining primase activation step on genomic template DNA, prior to the presence of duplex RNA/DNA.

Discussion

Our initiation and elongation assays together suggest that DNA primase initiation and product truncation are both dependent on a redox switch changing the oxidation state of the [4Fe4S] cluster in p58C. The redox switch first activates primase to tightly bind DNA, initiating primer synthesis. DNA CT through the nascent RNA/DNA primer then mediates termination of primer synthesis. On the basis of the available structural and biochemical data (36, 112, 114, 116, 118, 137, 143, 144, 215), we propose a model for priming that incorporates this redox switch (Figure 2.14). As DNA primase elongates the RNA primer to a length of 8 to 12 nt, a second [4Fe4S] enzyme, DNA polymerase α, comes into contact with the exposed nascent RNA/DNA helix (228) and associates with the primer-template so that its cluster is able to serve as a redox donor/acceptor through primed DNA. A DNA-mediated redox event can then occur, promoting p58C dissociation from the duplex and pol α binding. Because most polymerases bind to their duplex DNA substrates with modest affinity, as we show is the case for primase, pol α is unlikely to effectively take the primed substrate from primase. DNA-mediated signaling, which

53

modulates polymerase binding affinity through a change in the redox state of the [4Fe4S] cluster, would thus provide the key driver for efficient primer handoff.
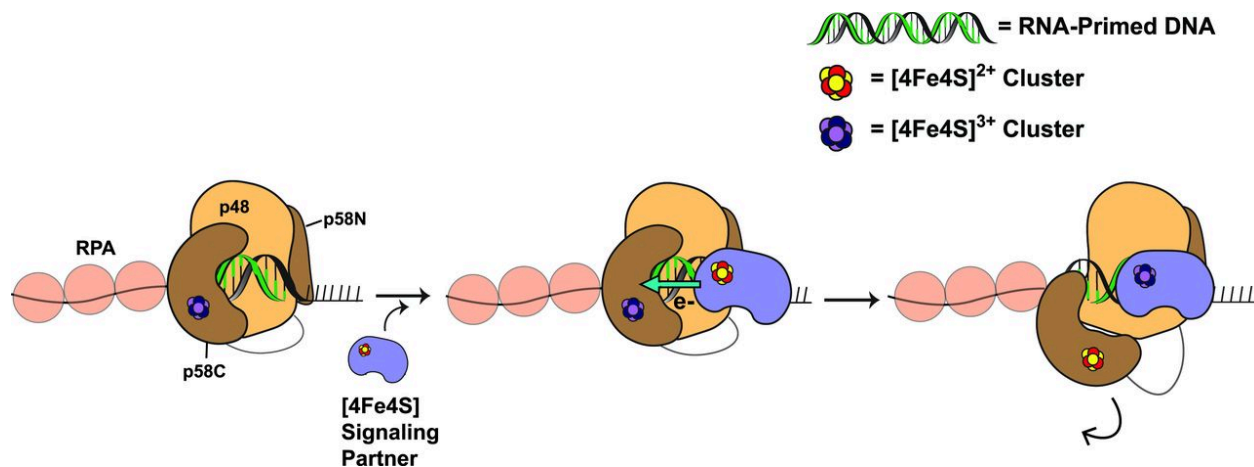
Figure 2.14. Proposed mechanism of primer handoff driven by DNA charge transport chemistry. Left: DNA primase elongates an RNA primer (green) to a length of 8 to 12 nt with both p48 and p58C contacting the nascent RNA/DNA duplex. Center: When the nascent primer is large enough, another [4Fe4S] enzyme (purple), which we hypothesize to be DNA polymerase α in vivo, participates in DNA-mediated signaling through the primer-template duplex. Right: This promotes dissociation of p58C through reduction of the cluster from the $[4Fe4S]^{3+}$ state to the $[4Fe4S]^{2+}$ state; the next [4Fe4S] enzyme is then tightly bound and can continue elongation of the primer-template.

We propose that DNA-mediated CT drives signaling between the [4Fe4S] cluster cofactors in DNA primase and DNA polymerase α, in which reduction of the primase [4Fe4S] cluster by pol α facilitates rapid handoff of the RNA/DNA primer. Generation of a short primer to initiate DNA synthesis must occur repeatedly on the lagging strand of the replication fork. DNA CT occurs on a time scale (229) that could plausibly mediate this handoff; electron hopping through the protein matrix conversely occurs approximately six orders of magnitude more slowly (230) and would not likely be the sole mediator of primase/polymerase α redox signaling between [4Fe4S] centers. A major and likely rate-limiting conformational change for DNA polymerase α binding to DNA has been modeled on a time scale of 100 to 150 ns using molecular dynamics (228). DNA-mediated redox signaling between primase and pol α to

promote primase dissociation and primer-template transfer is thus a feasible redox mechanism for the proposed critical step for the primase-to-polymerase α transfer of the primed template.

Our results support the proposal that the [4Fe4S] cluster in p58C serves as a redox switch governing binding, and therefore initiation and primer length counting, in the first step of DNA synthesis. We show that oxidation of the $[4Fe4S]^{2+}$ cluster in human DNA primase to the $[4Fe4S]^{3+}$ state facilitates DNA binding by the p58C domain. We demonstrate an electrochemically controlled redox signal for p58C on DNA, dependent on the effective concentration of p58C at the DNA/solution interface, establishing a reversible redox switch. Conserved Tyr residues in p58C on a pathway between the cluster and DNA mediate the redox reaction, affecting the [4Fe4S] cluster oxidation state and, consequently, DNA association of p58C. Moreover, primase initiation and termination activity, but notably not elongation activity or DNA binding, are compromised in these CT-deficient Tyr mutants. The p58C domain has been proposed to contact ssDNA in the rate-determining primer initiation step (44, 117, 144), and to play a role in the ability of the enzyme to count the length of the primer and signal for the arrest of further synthesis coupled with handoff of the initial primed substrate to pol α (114).

The recent discovery of [4Fe4S] cluster cofactors in each of the B-family replication-associated polymerases α, ε, δ, and ζ (162) is intriguing; combined with our results, this discovery suggests that redox reactions could be driving the DNA binding switches generally necessary for polymerase handoffs. Moreover, studies similar to the work of Liu and Huang (231) investigating primase [4Fe4S] cluster sensitivity to oxidative stress may further illuminate how this chemistry relates to eukaryotic replication activity in different cellular redox environments. Could the redox switch for primase binding to DNA be part of a larger, DNA-mediated electron transfer relay coordinating the association, transfer, and dissociation steps of

replication? Our results illustrate a chemical role for the [4Fe4S] cluster as a DNA-binding redox switch and point toward a new mechanism for coordinating activity within the dynamic replication fork. Further experiments by our groups revealed a lethal phenotype upon disruption of redox switching by yeast primase, indicating that redox switching is required for appropriate primase function at the eukaryotic replication fork (232).

Response to Comments from Tahirov and Pellegrini

After the publication of the above data, the Tahirov (Baranovskiy *et al.*) and Pellegrini groups published Technical Comments raising concerns about the experiment design and analysis (233, 234). Neither group directly disputed the electrochemical observations of charge transport through the DNA substrate to the iron-sulfur cluster of our p58 construct, nor did they dispute its inhibition by mutation of each of the three relevant tyrosine residues. Rather, they questioned the path for electron transfer through primase on the basis of differences in structures of p58C, the primase domain containing the [4Fe4S] cluster (36, 117, 215).

The structure of human p58C (36) from which we identified conserved tyrosine residues participating in a charge transfer pathway contains residues 318 to 353 folded in a β-hairpin arrangement. Other structures of human (215) and yeast (117) p58C show these residues in an α-helical arrangement and show that conserved tyrosines (Y309, Y345, and Y347 in human p58C) in the charge transport pathway through the protein are spaced and oriented differently in different structures. However, neither group appropriately considered that the requirements for charge transport mediated by tyrosines in a protein is only weakly dependent on orientation and can occur over distances up to 15 Å on our proposed time scale. Consequently, distances between tyrosine should be measured to their centroids (225). In fact, regardless of the

differences in structures, the tyrosine centroids in all human and yeast p58C structures are within feasible range for microsecond electron transfer through the protein matrix.

In addition to differences in crystallization conditions, there are differences in the two human p58C constructs used in our studies and those from the Tahirov group. Our p58C$_{272\text{-}464}$ construct contains three additional residues on the N terminus, left after cleavage of the His tag used to purify the protein. The Tahirov p58C$_{266\text{-}456}$ construct includes more residues on the N terminus but fewer on the C terminus. Notably, all electrochemistry and crystallography was performed on p58C, but all biochemical assays were performed using full-length human DNA primase (no mutations). More detailed comparative analysis of the structures of the two p58C constructs are presented in Chapter III.

The Technical Comments also raised concerns over our interpretation of data obtained for a Y345F primase mutant, which we reported has reduced electron transfer efficiency. They claim our results are ambiguous based primarily on a structural argument- that the effects observed are the result of loss of a hydrogen bond between Y345 and the triphosphate group at the 5′ end of the DNA/RNA substrate in a crystal structure (137). However, binding assays in our laboratory comparing the full, wild-type primase enzyme and the Y345F primase variant using the RNA/DNA substrate similar to the one used in the crystal structure (137) show no detectable differences. Furthermore, the placement of the Y345 side chain in the structure, which lies at the center of their claim, may be more ambiguous than previously implied (137). Retrieving the coordinate files and electron density maps for the structures (5F0Q), we observe a very poor fit for this residue. High B factors for this residue relative to most other residues call into question whether the assigned hydrogen bond has high enough occupancy to substantiate their argument.

In addition, we carried out parallel biochemical experiments comparing primer synthesis of wild-type full primase and the charge transfer–deficient Y309F variant of full p48/p58 primase on single-stranded DNA (Figure 2.15). Here, too, we observe significant inhibition of initiation in the mutant, just as we observed with the Y345F and Y345C primase variants. Thus, mutation of another tyrosine in the charge transfer pathway, one not interacting with the substrate, also inhibits initiation. Hence, inhibition of initiation observed for Y345F is the result of inhibition of the redox switch.
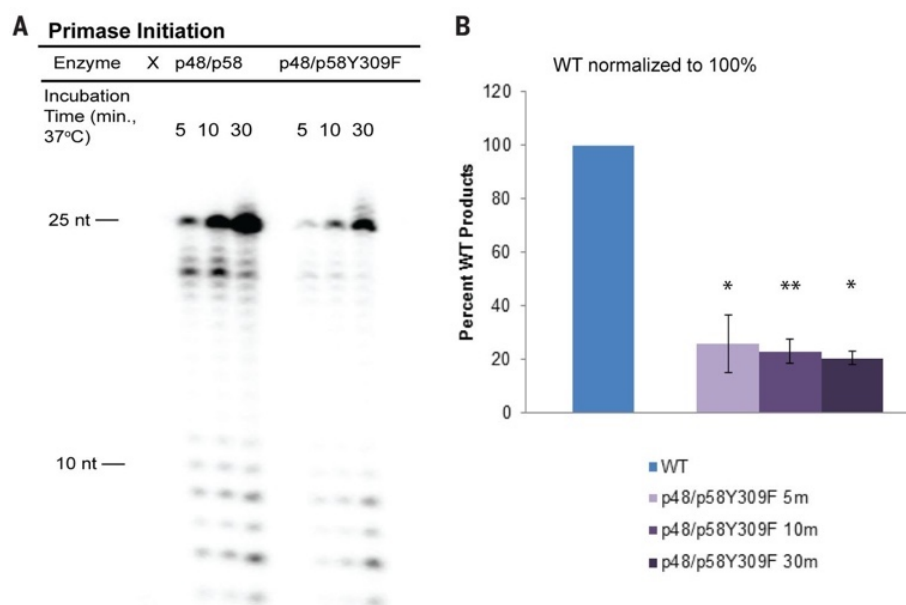
Figure 2.15. Wild-type (WT) p48/p58 versus p48/p58Y309F activity on ssDNA. A. Gel separation of products for primase initiation assay comparing WT and charge transfer–deficient primase. B. Product quantifications for p48/p58Y309F, with WT primase (p48/p58) products normalized to one. Values shown are the mean of $n = 3$ trials; error bars represent standard deviation. *, $0.001 \leq P < 0.005$; **, $P < 0.001$; Student's t test. All activity assays were performed under anaerobic conditions. Reactions contained 400 nM primase variant, 250 nM ssDNA, 188 μM uridine triphosphate (UTP), 112 μM cytidine triphosphate (CTP), 1 μM α-^{32}P adenosine triphosphate (ATP) in 50 mM Tris, pH 8.0, 3 mM MgCl$_2$.

The Tahirov group also criticized the choice of substrate for the electrochemical studies of p58C. Primase interacts with a range of DNA structures (36), enabling us to create a substrate that productively binds p58C and satisfies the technical criteria needed for DNA-mediated

electrochemistry. Our substrates were designed for the primary objectives of the experiment: (i) to assess the ability of p58C to participate in DNA charge transport and (ii) to examine the differential effects of [4Fe4S] cluster oxidation state. The binding affinity of p58C for the substrate should therefore not be particularly strong, so that any changes in affinity between oxidized $[4Fe4S]^{3+}$ and reduced $[4Fe4S]^{2+}$ protein can be detected. A substrate similar to the construct containing a 3′ single-stranded DNA (ssDNA) overhang and 5′-triphosphate, which has been shown to bind p58C very tightly, would obscure the observation of differences between the two redox states on the electrode. Moreover, as it contains a 5′-ssDNA overhang, our electrochemistry substrate is more similar to the primed ends encountered by polymerase enzymes in cells.

Tahirov and coworkers additionally suggested that a mismatch in the nascent primer inhibits initiation, but not truncation in the assay with a primed template substrate. On both well-matched and mismatched substrates, we observe a mixture of initiation and elongation products because primase binds to both ssDNA and primed DNA portions of the substrate (235). We showed that primase initiation products (7-29 nts) are synthesized on the ssDNA template used to generate a mismatched primer.

Our assignment of the oxidized $[4Fe4S]^{3+}$ species generated electrochemically was also criticized. However, we showed directly using cyclic voltammetry, as seen in Figures 2.1 and 2.3, that the oxidized $[4Fe4S]^{3+}$ product was generated. The cathodic peaks in Figures 2.1 and 2.3 furthermore show the reduction of an oxidized species on the DNA-modified electrode. Additionally, chemical oxidants with potentials similar to the potential applied to p58C can oxidize human p58C to the $[4Fe4S]^{3+}$ state, as demonstrated by electron paramagnetic resonance spectroscopy (116), supporting our assignment of the $[4Fe4S]^{3+}$ species. We were not able to

perform binding affinity measurements with the electrochemically oxidized p58C because the oxidized protein is unstable over the long periods of time required for the measurement.

Finally, our performing in vitro priming assays under anaerobic conditions was criticized because by the contention that this does not correspond to the cellular environment. Anaerobic conditions were used to ensure that we had full control over the redox state of the [4Fe4S] cluster, which can be oxidized in the presence of air (223). In fact, controls performed under aerobic conditions gave similar results overall, although with greater scatter and lower precision.

Overall, the positioning of tyrosines 309, 345, and 347 in both structures of human p58C, irrespective of local conformation, suggests a feasible pathway for electron transfer through the [4Fe4S] protein. The electrochemical experiments with p58C variants, and the biochemical activity assays with corresponding full primase variants, illuminate the electron transfer chemistry performed by the primase [4Fe4S] cluster and the effect of this chemistry on DNA binding. Primer elongation assays with well-matched and mismatched template strands, moreover, demonstrate the regulatory role of DNA charge transport in termination. Thus, the concerns raised in the two Technical Comments do not negate either our experimental observations of redox chemistry performed by the [4Fe4S] cluster or the proposal of a role for the redox chemistry, coupled to DNA charge transport, in regulating the activity of human DNA primase.

CHAPTER III

FUNCTIONAL AND STRUCTURAL SIMILARITY OF HUMAN DNA PRIMASE [4FE4S]

CLUSTER DOMAIN CONSTRUCTS[3]

Introduction

DNA synthesis at the replication fork begins with the formation of 8-12 nucleotide (nt)

RNA primers on the single-stranded DNA template (236, 237). In eukaryotes, primers are

generated by the heterotetrameric DNA polymerase α-primase (pol-prim) complex, which

possesses two enzymatic activities in two distinct active sites (37, 132, 238). Primase, a DNA-

dependent RNA polymerase, generates the initial hybrid RNA-DNA primed substrate, which is

then handed off to DNA polymerase α (pol α) to extend the initial primer by approximately

twenty DNA nts. The extended primed substrates are in turn handed off to the processive

polymerases ε and δ, which synthesize the bulk of nascent DNA on the leading and lagging

strands, respectively (145, 239, 240).

Human DNA primase is composed of catalytic (p48) and regulatory (p58) subunits. The

regulatory subunit has a C-terminal domain (p58C) that is unique to eukaryotes and contains a

[4Fe4S] cluster (116, 118, 119). This domain regulates the catalytic efficiency of primase, a

function attributed to the ability to bind nucleotides, DNA template, and primed substrate (36,

112, 114, 116-118, 137, 142, 145, 215). We have recently proposed that [4Fe4S] redox control of

DNA binding affinity may serve as a mechanism to drive handoff of the RNA primed template

---

[3] Portions of this chapter were previously published in the following publication: Holt ME, Salay LE, O'Brien E, et al. Functional and Structural Similarity of Human DNA Primase [4Fe4S] Cluster Domain Constructs. *PLoS One (San Francisco, CA)*. 2018;13(12):e0209345. doi: 10.1371/journal.pone.0209345.

from the primase to the pol α subunits of human pol-prim (235). Skepticism about some of the reported results have been expressed and debated (233, 234, 241), much of which was related to differences in crystal structures obtained from p58C constructs with different N-termini (218).

The p58C domain of human primase has been crystallized under two different conditions and these have generated structures that have the same global fold but localized differences in secondary structures (36, 215). Figure 3.1 shows a best-fit superposition of the high-resolution X-ray crystal structures of the two $p58C_{272-464}$ and $p58C_{266-456}$ constructs. The two structures are clearly very similar except for residues Leu318-His351, which are positioned near the DNA binding site. In the crystal structure of $p58C_{272-464}$, these residues occupy a β-sheet-like structure that is stabilized by cross-strand interactions with another molecule in the unit cell (36). In addition, a disulfide cross-link is formed during crystallization between the Cys449 residues of adjacent p58C molecules. In contrast, when $p58C_{266-456}$ was crystallized under a different set of conditions, this β-type interaction is not observed and instead these residues occupy a helical hairpin (215). It has been proposed that Ile271 is critical for stabilizing this helical motif in the $p58C_{272}$-464 structure and therefore the absence of this residue explains why $p58C_{266-456}$ has a different structure (233, 234).
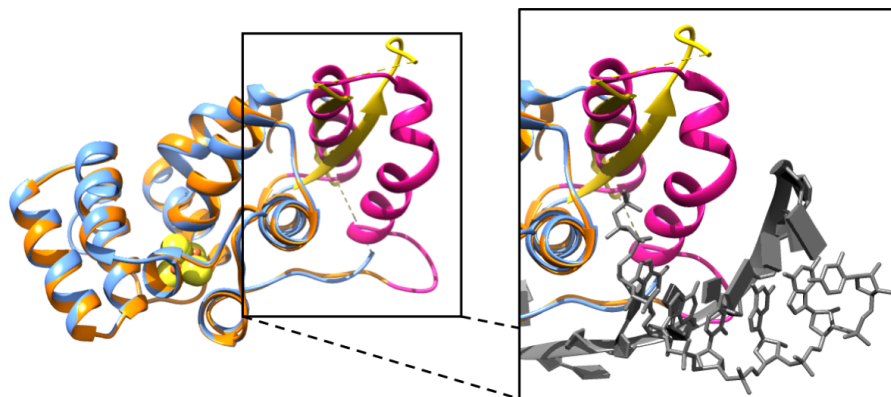
Figure 3.1. Comparison of the structures of p58C$_{272-464}$ and p58C$_{266-456}$. (A) Best-fit superposition over all backbone atoms of the crystal structures of the p58C$_{272-464}$ (PDB ID: 3L9Q) and p58C$_{266-456}$ (PDB ID: 5F0S) constructs. The substrate in the p58C$_{266-456}$ structure is removed for clarity. Inset: Best-fit superposition over all backbone atoms of the DNA binding region with the primed substrate shown in the p58C$_{266-456}$ structure. p58C$_{266-456}$ is colored orange, p58C$_{272-464}$ blue, and the RNA-primed DNA substrate in grey (RNA in sticks, DNA in slabs). In both panels, Leu318-His351 are colored yellow in the p58C$_{272-464}$ structure and pink in the p58C$_{266-456}$ structure. Connectivity between residues where electron density is missing is indicated by dashed lines.

Many of the concerns about our work were attributed to the fact that the structural differences between the two p58C constructs occurred in a region that contains residues interacting with DNA substrates. However, it is well known that differences between crystal structures arise when crystals are formed under different crystallization conditions. In order to resolve any controversy and directly address the concerns raised about our findings, we report here a comprehensive set of studies of a p58C$_{266-464}$ construct. Comparisons of the crystal structure of this construct are made to previously reported crystal structures. In addition, the structure was analyzed in solution by circular dichroism and NMR, along with assays of DNA binding affinities and electrochemical properties. These results show that the structure of Leu318-His351 varies in accord with crystallization conditions, whereas the structure in solution and biochemical properties of different p58C constructs are effectively the same.

Methods

p58C₂₆₆₋₄₆₄ construct generation

The p58C$_{266-464}$ construct was created with a Q5 site-directed mutagenesis kit from New England Biolabs, using the p58C$_{272-464}$ construct plasmid (36) as the template, 5'-GGCAAGATTTCCTTAGATCAGATTGATTTGCTTTCTACC - 3' for the forward primer, and 5' – CACATTTCCGGGCCCCTGGAACAGAAC - 3' for the reverse primer. 10 ng of p58C$_{272-464}$ plasmid was used in the exponential amplification, which was completed as described in the Q5 Site-Directed Mutagenesis Kit manual (New England Biolabs), except that the final extension time was extended to 5 minutes. The KLD reaction was also completed as described in the manual, except that 4 µL of PCR product was used in the reaction and the incubation time was increased to 20 minutes. 10 µL of the KLD reaction product was transformed into XL1-Blue cells. DNA was extracted from individual colonies with a Qiagen QIAprep Spin Miniprep Kit. Appropriate insertion of residues 266-271 was confirmed through plasmid sequencing (GENEWIZ, LLC).

Protein expression and purification

p58C constructs were expressed and purified as previously described (218, 235). In short, plasmid DNA was transformed into BL21 (DE3) cells (Novagen) and cultured in Terrific Broth media at 37 °C to an OD$_{600}$ of 0.6-0.8, when flasks were moved to a 21 °C incubator with shaking. After 30 minutes, ferric citrate and ammonium ferrous citrate were added to a final concentration of 0.1 mg/mL, and isopropyl 1-thio-β-D-galactopyranoside was added to a final concentration of 0.5 mM. Protein was expressed at 21 °C overnight. The same growth protocol was used to generate ^{15}N-labeled p58C, except that cells were cultured in M9 media

supplemented with ^{15}N-labled ammonium chloride (Cambridge Isotopes) and expressed overnight at 25 °C. The same purification scheme was used for both unlabeled and ^{15}N-labeled protein (218). In short, protein was first purified by nickel affinity chromatography (Amersham Biosciences). The 6xHis tag was cleaved with H3C protease and the protein was dialyzed into a low-imidazole buffer (218, 235). The protein was repassed over the nickel column to remove the H3C protease and uncleaved protein. A heparin column was used as the final purification step to remove residual contaminants (218, 235).

Crystallization and structure determination

The structure of p58C$_{266-464}$ was determined as previously described for p58C$_{272-464}$ (36, 218). p58C$_{266-464}$ crystals were grown by hanging drop vapor diffusion at 16 °C from a drop composed of equal volumes of 50 mg/ml protein in 20 mM MES (pH 6.5) and 75 mM NaCl and reservoir solution containing 100 mM Tris (pH 8.5), 400 mM Li$_2$SO$_4$ and 18% (v/v) PEG 3350. Prior to data collection, crystals were transferred to a drop containing 100 mM Tris (pH 8.5), 400 mM Li$_2$SO$_4$,18% (v/v) PEG 3350, and 20% (v/v) glycerol for five seconds. The crystals were looped and flash frozen in liquid nitrogen. X-ray data were collected at beamline 21ID-D (Life Sciences Collaborative Access Team) of the Advanced Photon Source at Argonne National Laboratory at 11.5 kEV. All data were processed by HKL2000 (242). The structure was determined using molecular replacement (PHASER-MR) with PDB entry 3L9Q, residues 274-316 and 360-457, as the search model. Manual model building for the structure was performed using *Coot* model building software, and waters were placed with the *Coot* routine, Find Waters (243). The final model was obtained by iterative cycles of model building in *Coot* and structure refinement using Phenix.refine in the Phenix suite of programs (244, 245). Structures were

superimposed and RMSD calculated in Chimera with the Matchmaker algorithm (246).

Programs used for structure determination and refinement were accessed through SBGrid (247).

Statistics for data collection and refinement are shown in Table 3.1.

Data collection	
Space Group	C2
Cell Dimensions	
a, b, c (Å)	110.19, 52.56, 88.77
α, β, γ (°)	90, 115.08, 90
Temperature (K)	100
Wavelength (Å)	1.08
Resolution (Å)	50.00-2.01 (2.08-2.01)
Unique Reflections	27193
R_{meas} (%)	0.14 (0.79)
I/σI	10.06 (2.12)
Completeness (%)	88.6 (84.4)
Redundancy	3.6 (3.3)
Refinement	
Resolution (Å)	50.00-2.01 (2.08-2.01)
No. reflections	27153
R_{work}/R_{free}	0.18/0.21 (0.23/0.24)
No. molecules/ASU	2
No. atoms	2872
Protein	2689
Ligand/ion	41
Water	52
B-factors	
Mean	39.1
Protein	38.7
Ligand/ion	66.9
R.m.s. deviations	
Bond lengths (Å)	0.003
Bond angles (°)	0.62
Ramachandran	
Favored	308 (99.4%)
Allowed	2 (0.6%)
Disallowed	0 (0%)

Table 3.1. Crystallographic data collection and refinement statistics. Values in parentheses are for the highest-resolution shell.

Circular dichroism (CD) spectroscopy

Samples were buffer exchanged into 20 mM K_2HPO_4 (pH 7.2) and diluted to a concentration of 0.3 mg/mL. The far-UV CD spectrum over the range 190–260 nm was acquired at room temperature using a Jasco J-810 spectrophotometer. Each spectrum is the average of three scans acquired with a scanning rate of 50 nm/min and data pitch of 1 nm. Prior to generating the overlay in Graphpad Prism 7, the p58C$_{266-464}$ spectrum was scaled to the p58C$_{272-464}$ spectrum by averaging the values of the CD_{208}(p58C$_{272-464}$)/CD_{208}(p58C$_{266-464}$) and CD_{222}(p58C$_{272-464}$)/CD_{222}(p58C$_{266-464}$) ratios to generate a scaling factor (~0.86), then multiplying the entire p58C$_{266-464}$ spectrum by this scaling factor.

NMR spectroscopy

Spectra were acquired using a Bruker AV-III 800 MHz spectrometer equipped with a CPTCI single-gradient cryoprobe. ^{15}N-enriched p58C constructs at a concentration of 200 μM were prepared in a buffer containing 20 mM MES (pH 6.5), 50 mM NaCl, 2 mM DTT, and 5% (v/v) D_2O. Two-dimensional ^{15}N-^{1}H heteronuclear single quantum coherence (HSQC) spectra were acquired at 25 °C with 2,048 and 128 points in the ^{1}H and ^{15}N dimensions, respectively. 64 scans were recorded in the direct (^{1}H) dimension for each point sampled in the indirect (^{15}N) dimension. Data were processed by Topspin (Bruker) and analyzed with Sparky (University of California).

RNA primer generation

Triphosphorylated RNA primer was transcribed with T7 RNA polymerase (248) and purified on a 25% TBE-polyacrylamide gel supplemented with 8 M urea according to standard

methods. Dried RNA pellet was resuspended in RNAse-free H_2O and aliquoted prior to further purification. RNAs used for binding assays were HPLC purified on a Luna 5 μM C18(2) 100 Å 250X4.6 mm column (Phenomenex). Buffer A: 0.1 M ammonium formate; Buffer B: acetonitrile; flow rate: 1.5 mL/min. Purification program: 1-5% Buffer B over three minutes, 5-8% Buffer B over 22 minutes, then 80% Buffer B for five minutes. RNA typically eluted around 11 minutes. HPLC-purified RNA was further validated with mass spectrometry, which confirmed that a product of the expected mass had been generated.

Substrate binding assays

The fluorescence intensity (FI) assay was performed with a Monolith NT.115 series microscale thermophoresis (MST) instrument (NanoTemper) at 25 °C. Cy5-labeled 18 nt DNA template was purchased from Sigma-Aldrich. A 1.1:1 ratio of 8 nt RNA and this DNA template was annealed in annealing buffer (20 mM MES (pH 7.0), 75 mM NaCl), resulting in a 25 μM stock of annealed, primed substrate. This stock was diluted to 100 nM with MST buffer (20 mM MES (pH 6.5), 50 mM NaCl, 2 mM DTT, 0.05% Tween). The primed substrate was mixed with p58C and allowed to incubate in the dark at room temperature for 15 min. Samples with a final substrate concentration of 50 nM were then loaded into MO-K003 Monolith NT.115 hydrophobic capillaries (NanoTemper) and fluorescence was measured at 20% LED and 40% MST power. Final K_D values were calculated using the one-site total binding equation in GraphPad Prism 7. Titrations were completed after running an SD test to ensure that the concentration-dependent changes in fluorescence intensity were not due to adsorption to the capillaries or aggregation of the fluorophore (249).

FI RNA primer: 5'-PPP-GGCUCGGA-3'
FI DNA template: 5'-Cy5-AAACTCCGAGCCAACATA-3'

Fluorescence anisotropy (FA) was measured with a SpectraMax M5 microplate reader (Molecular Devices). A 6FAM-labeled 22 nt DNA template was purchased from Sigma-Aldrich. A 1.1:1 ratio of 12 nt RNA and this DNA template was annealed in annealing buffer (20 mM MES (pH 7.0), 75 mM NaCl), resulting in a 25 μM stock of annealed, primed substrate. The stock was diluted to 800 nM with DNA binding buffer (20 mM MES (pH 6.5), 50 mM NaCl, 2 mM DTT). This primed substrate was mixed with p58C and allowed to incubate in the dark at room temperature for 15 min. Samples with a final substrate concentration of 50 nM were then loaded into a 384-well plate and polarized fluorescence intensities were measured using excitation and emission wavelengths of 485 nm and 520 nm. The fluorescein control experiments were performed with 25 nM fluorescein (Sigma Aldrich) dissolved in DNA binding buffer containing 0.016% DMSO, which was then mixed with p58C and incubated in the dark at room temperature for 15 min prior to determining fluorescence anisotropy in the same way as for the p58C-DNA titrations. Three replicates were collected for each titration. Final K_D values were calculated using the one-site specific binding equation in GraphPad Prism 7; prior to using this equation, each binding curve was normalized by subtracting the fluorescence anisotropy value of the zero point from each point on the curve. K_D values are reported as the mean ± standard deviation of three measurements for each variant.

FA RNA primer: 5'-PPP-GGACCTCCAGGA-3'
FA DNA template: 5'-6FAM-AAACTCCTGGAGGTCCAACATA-3'

Sample preparation for electrochemistry

Multiplexed chips were fabricated as described previously (235). p58C construct samples were stored prior to experiments in p58C storage buffer (20 mM Tris (pH 7.2), 75 mM NaCl). All p58C constructs were buffer exchanged into HEPES electrochemistry buffer (20 mM HEPES

(pH 7.2), 75 mM NaCl) using Amicon ultra centrifugal filters (0.5 mL, 3 kDa MWCO)

(Millipore Sigma). Protein was applied to the filter in volumes of 90-140 μL and centrifuged for

15 minutes at 14000 x g at 4 °C. After centrifugation, 400 μL of HEPES electrochemistry buffer

was applied to the filter and centrifuged at 14000 x g for 20 minutes. This was repeated four

times to exchange p58C samples into HEPES electrochemistry buffer. After buffer exchange

and recovery of sample by centrifugation (2 minutes, 1000 x g), concentrations of [4Fe4S]

cluster-containing p58C or mutants were measured by using UV-Visible spectroscopy to monitor

the absorbance of the [4Fe4S] cluster at 410 nm (extinction coefficient = 17000 M^{-1} cm^{-1}) (235,

250). Recovered samples (approx. 100-150 μL in volume) were deoxygenated for 2-3 minutes

with argon. Samples were then transferred into the anaerobic chamber (Coy Laboratory

products). Prior to deposition onto the gold electrode surface, $p58C_{266-464}$ samples were diluted

with previously deoxygenated HEPES electrochemistry buffer to a molar concentration of 40μM

[4Fe4S] p58C. Samples were initially deposited onto multiplex chip quadrants in 20 μL volumes

and the remaining sample deposited in a well of bulk solution above the chip surface, to a final

volume of 200-300 μL.

p58C construct electrochemistry

     All electrochemistry was performed using a CHI620D potentiostat and 16-channel

multiplexer (CH Instruments) in an anaerobic glove chamber. Multiplex gold electrodes were

part of a three electrode system with an external Ag/AgCl reference electrode (Bioanalytical

Systems) and platinum counter electrode. Cyclic voltammetry scans were performed at a scan

rate of 100 mV/s over a potential range of +0.412 V to -0.288 V vs. NHE or +0.512 V to -0.188

V vs NHE. Bulk electrolysis on DNA was performed at an applied potential of +0.512 V vs.

NHE for all electrochemical oxidation reactions. The oxidizing potential was applied for at least 8.33 minutes for single oxidation reactions on a surface, and 6.67 minutes for the iterative oxidation cycles of p58C variants. The reducing potential was applied for 8.33 minutes in all electrochemical reduction reactions. All bulk electrolysis and cyclic voltammetry was performed in previously deoxygenated p58C electrochemistry buffer (20 mM HEPES (pH 7.2), 75 mM NaCl). Charge transfer (nC) in the cathodic peak of oxidized samples CV scans was assessed using the area under the current wave of the reduction signal. Charge transfer was measured for oxidized samples using CHI software, assessing the area under the reductive peak in CV after electrochemical oxidation. NTP-dependence of electrochemical signals was measured by pipetting a small volume (1-3 μL) of a 0.1 M ATP stock solution into each quadrant of the multiplexed chip setup. Samples were added by quadrant, as physical barriers in the setup prevent diffusion of NTPs between electrode quadrants. After the volume of ATP stock was deposited onto the electrode quadrant, resulting in a 2.5 mM or 5 mM concentration of ATP in the quadrant, CV scans were measured (100 mV/s scan rate). Charge transfer was assessed using CHI software; charge values were determined by calculation of the area under the reductive and oxidative peak curves. Midpoint potentials of NTP-dependent redox signals were assessed using the peak selection function in CHI software.

p58C can be crystallized in different conformations

   To test if the differences between the p58C$_{266-456}$ and p58C$_{272-464}$ structures were a by-product of the differences in sequence, we produced, purified, and crystallized a p58C construct containing residues 266-464 in the conditions used to crystallize p58C$_{272-464}$ (Figure 3.2) (36, 235). These crystals diffracted to 1.6 Å and the data were phased using molecular replacement with the structure of p58C$_{272-464}$ (3L9Q). To avoid phase bias, residues 315-360 were excluded when defining the search model. With these residues omitted, density for residues in extended conformation that together formed a beta-type interaction were clearly evident in the p58C$_{266-464}$ 2Fo-Fc map (Figure 3.3).
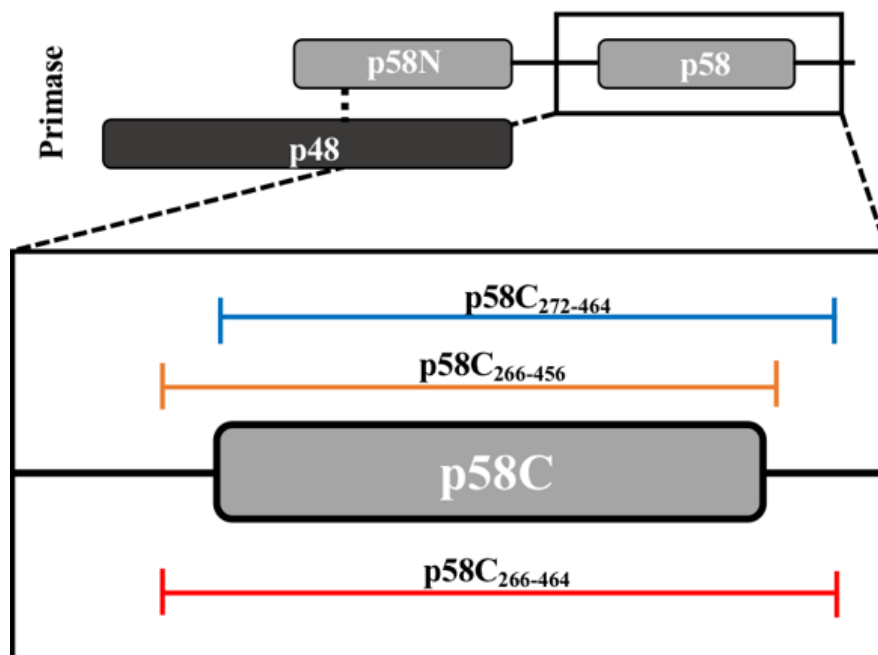
Figure 3.2. Domain map showing p58C constructs.
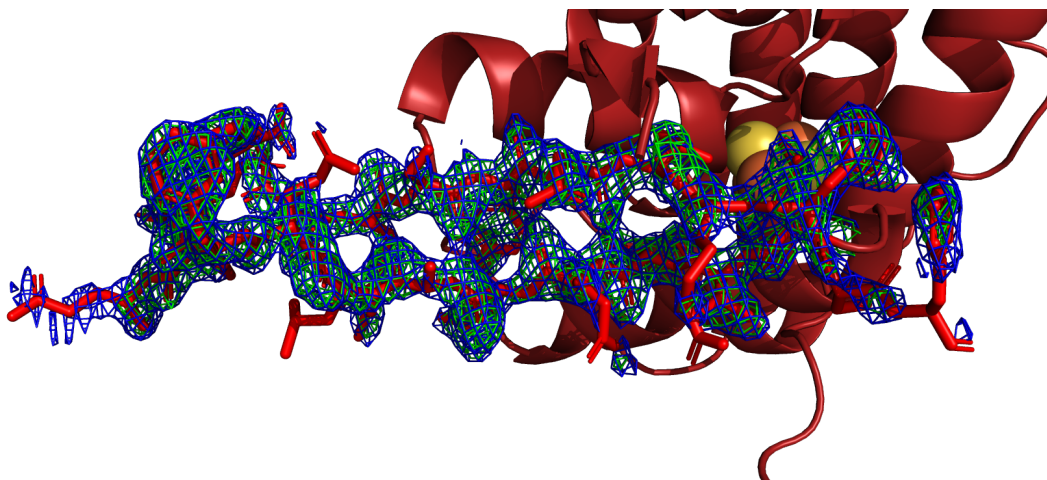
Figure 3.3. Electron density maps revealing β-sheet-like conformation for Leu318-His351.
The ribbon diagram of the structure with Leu318-His351 displayed in stick representation is docked into the starting electron density map around the β-sheet-like region. Blue mesh represents the $2F_o$-F_c map contoured at 1 σ, Green mesh represents the F_o-F_c map, contoured at 3 σ. Figure made in Pymol using the isomesh command (251).

This region was re-built manually in *Coot*, and the final structure was refined to 2.01 Å resolution. As previously observed for p58C$_{272-464}$ under these conditions, crystallized as a dimer, with a disulfide cross-link between the two Cys449 residues and several stabilizing interactions between symmetry-related molecules (Figure 3.4). In p58C$_{266-464}$, residues 330-340 and 353-360 in chain A and 330-345 and 353-359 in chain B are missing due to disorder. Disordered residues in the same regions are observed in the structure of p58C$_{272-464}$. A best-fit superposition of the two structures reveals they are very similar (Figure 3.5), with a backbone RMSD of only 0.23 Å. This finding shows that the differences in residues Leu318-His351 evident from comparing the previous p58C structures (36, 117, 215) are not intrinsic to the differences in the N-termini of the constructs but rather to differences in the crystallization conditions.
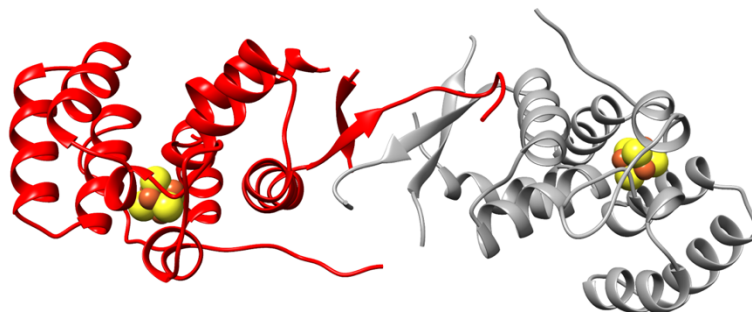
Figure 3.4. The beta sheet-like region of p58C$_{266\text{-}464}$ (red) is stabilized by interactions with the beta sheet-like region of an adjacent symmetry-related molecule (grey).
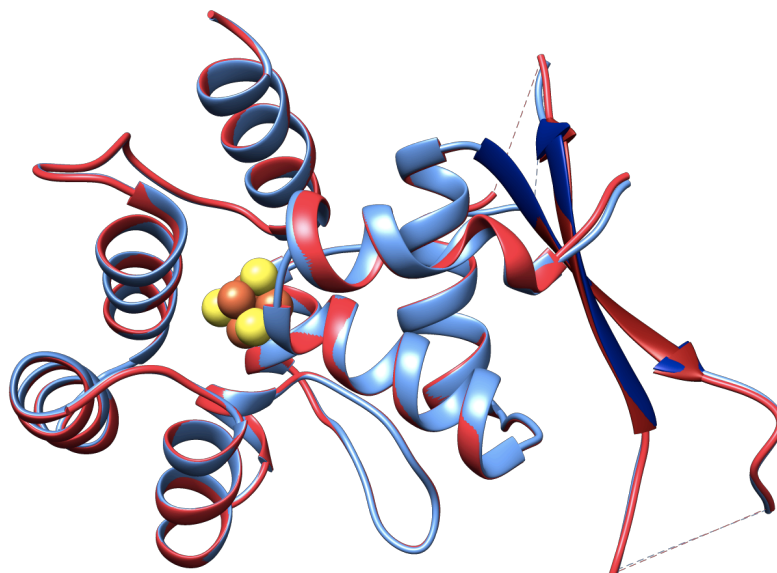
Figure 3.5. Comparison of the X-ray crystal structures of p58C$_{266\text{-}464}$ and p58C$_{272\text{-}464}$ obtained from crystals grown under identical conditions. Ribbon diagrams of the best-fit superposition over all backbone atoms of p58C$_{272\text{-}464}$ (blue) and p58C$_{266\text{-}464}$ (red) crystal structures. Residues Leu318-His351 are colored dark blue in the p58C$_{272\text{-}464}$ structure and dark red in the p58C$_{266\text{-}464}$ structure. Connectivity between residues where electron density is missing is indicated by dashed lines. The [4Fe4S] clusters are shown as the cluster of yellow and orange spheres.

CD and NMR show p58C$_{272\text{-}464}$ and p58C$_{266\text{-}464}$ are structurally similar in solution

Because the differences in the X-ray crystal structures can be attributed to differences in the crystallization conditions, we used CD and NMR spectroscopy to determine if there are significant structural differences between p58C$_{272\text{-}464}$ and p58C$_{266\text{-}464}$ in solution. The CD spectra

of p58C$_{272-464}$ and p58C$_{266-464}$ acquired under identical conditions were very similar (Figure 3.6),

showing that the two constructs contain the same distribution of secondary structure (170). To

further investigate the structures of these constructs, ^{15}N-^{1}H HSQC spectra were collected for

^{15}N-enriched p58C$_{272-464}$ and p58C$_{266-464}$. Figure 3.7 shows an overlay of HSQC spectra acquired

under identical conditions for p58C$_{272-464}$ and p58C$_{266-464}$. These match well with previously

published spectra (36, 116), indicating that both constructs are well-folded. There are a number

of chemical shift perturbations between the two spectra, consistent with the presence of six

additional N-terminal residues. Overall, the distribution of peaks is very similar in the two

spectra, indicating that the overall tertiary structure of p58C$_{272-464}$ and p58C$_{266-464}$ are the same.

Sequence-specific assignments could provide insight into the nature of the differences in the

NMR spectra. However, the presence of the paramagnetic [4Fe4S] cluster causes extreme line

broadening due to rapid relaxation of spatially proximate residues, greatly complicating efforts to

obtain resonance assignments (252). Despite the absence of assignments, because the NMR

spectra are so similar, and the CD spectra are virtually identical, we can conclude that there are

no substantial differences in the structures of p58C$_{272-464}$ and p58C$_{266-464}$ in solution.
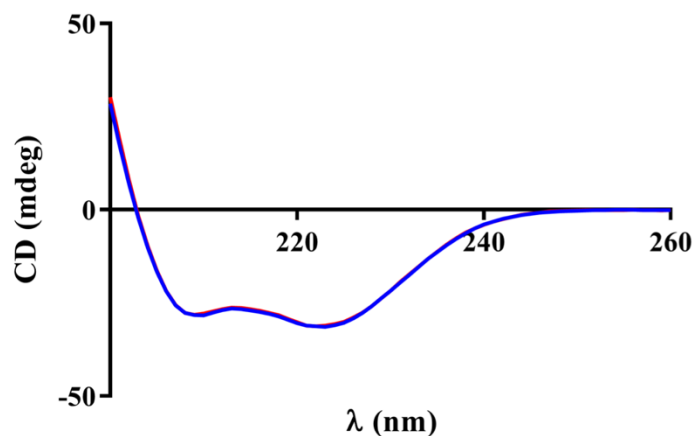
Figure 3.6. The p58C$_{272-464}$ and p58C$_{266-464}$ constructs have the same distribution of secondary structures in solution. CD spectra p58C$_{272-464}$ (blue) and p58C$_{266-464}$ (red) were collected at room temperature for a solution containing 0.3 mg/mL of protein in a buffer containing 20 mM K$_2$HPO$_4$ (pH 7.0).
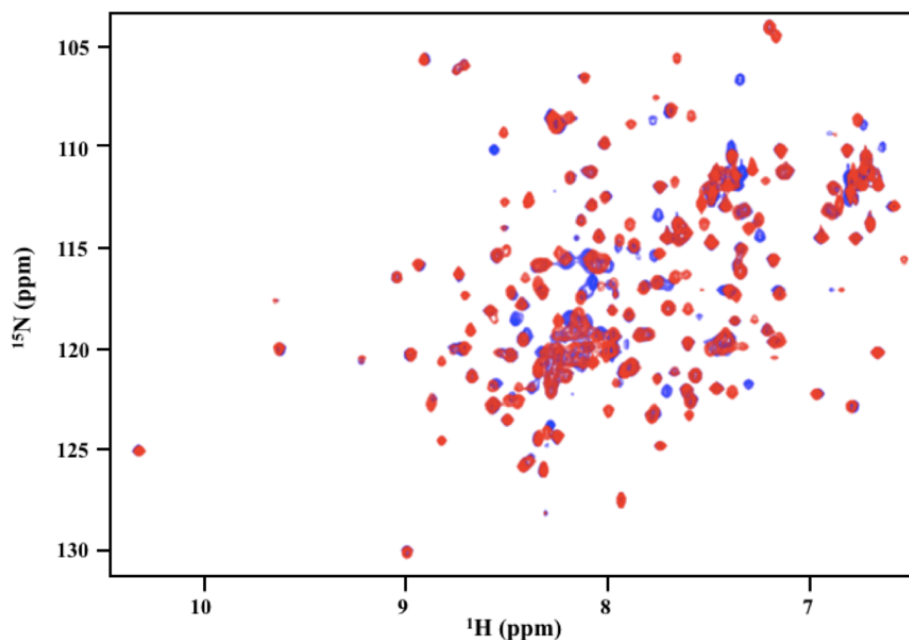
Figure 3.7. The tertiary structures of p58C$_{272\text{-}464}$ and p58C$_{266\text{-}464}$ constructs are similar in solution. ^{15}N-^{1}H HSQC NMR spectra of p58C$_{272\text{-}464}$ (blue) and p58C$_{266\text{-}464}$ (red) were acquired at 25 °C on a Bruker AV-III spectrometer operating at 800 MHz. The samples contained 200 μM protein in a buffer containing 20 mM MES (pH 6.5), 50 mM NaCl, 2 mM DTT, and 5% ^{2}H$_2$O.

The p58C$_{272\text{-}464}$ and p58C$_{266\text{-}464}$ constructs have the same affinity for a primed template

Having established that the structures of p58C$_{272\text{-}464}$ and p58C$_{266\text{-}464}$ are essentially the same, it is important to determine if the two constructs are functionally equivalent. Since DNA binding is an essential property for p58C function, we turned to fluorescence assays to compare the DNA binding properties of the two constructs. We first used a fluorescence intensity (FI)-based assay with an RNA-primed DNA substrate composed of a Cy5-labelled 18 nt DNA template strand annealed with a complementary 5'-triphosphorylated 8 nt RNA primer, corresponding to a high affinity product produced by primase (115). The affinities of p58C$_{272\text{-}464}$ and p58C$_{266\text{-}464}$ for this substrate were measured by monitoring initial fluorescence intensity (FI) changes as p58C was titrated into a fluorescently-labelled substrate (Figure 3.8). Fitting these

data to a single site binding equation returned dissociation constants (K_D) within statistical error:

2.7 ± 0.3 µM and 3.4 ± 0.5 µM for p58C$_{272\text{-}464}$ and p58C$_{266\text{-}464}$, respectively.
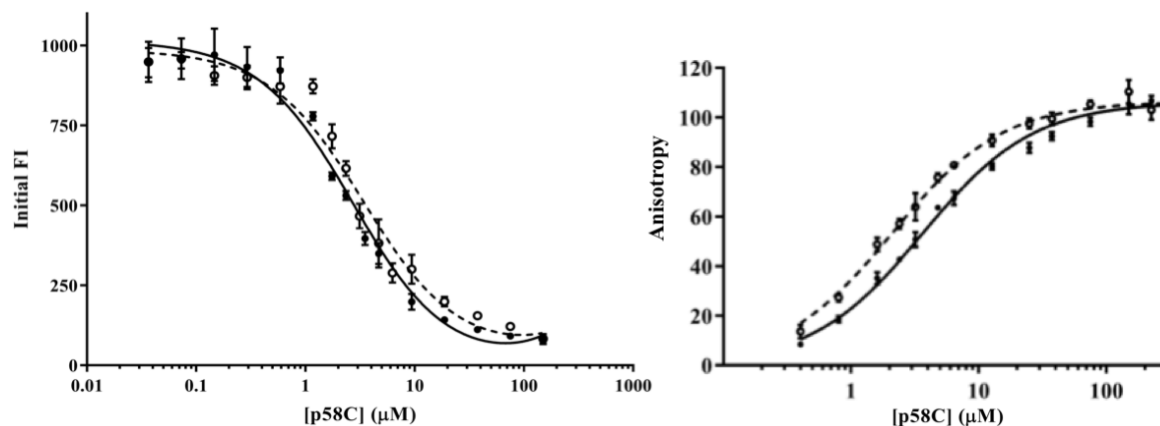
Figure 3.8. The two p58C constructs bind a RNA-primed substrate with the same affinity. Left: The p58C$_{272\text{-}464}$ (filled circle, solid line) and p58C$_{266\text{-}464}$ (open circle, dashed line) proteins were titrated into Cy5-labeled substrate in a buffer containing 20 mM MES (pH 6.5), 50 mM NaCl, 2 mM DTT, and 0.05% Tween. FI was measured at 25 °C. RNA: 5'-GGCUCGGA-3'; DNA: 5'-Cy5-AAACTCCGAGCCAACATA-3'. Right: The p58C$_{272\text{-}464}$ (filled circle, solid line) and p58C$_{266\text{-}464}$ (open circle, dashed line) proteins were titrated into 6FAM-labeled substrate in a buffer containing 20 mM MES (pH 6.5), 50 mM NaCl, and 2 mM DTT. Fluorescence anisotropy was measured at 25 °C. RNA: 5'-GGACCTCCAGGA-3'; DNA: 5'-6FAM-AAACTCCTGGAGGTCCAACATA-3'. For both panels, error bars represent the standard deviation of three replicates.
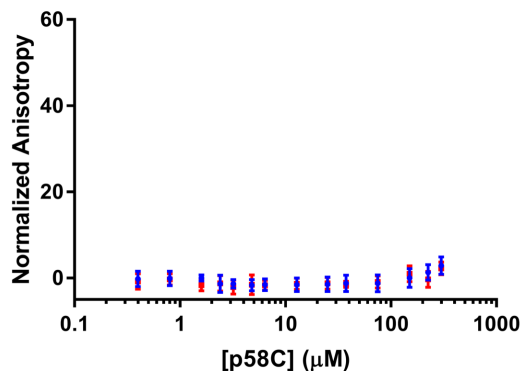
Figure 3.9. The p58C$_{266\text{-}464}$ and p58C$_{272\text{-}464}$ constructs do not bind fluorescein. Titrations of p58C$_{272\text{-}464}$ (blue circles) and p58C$_{266\text{-}464}$ (red squares) into 25 nM fluorescein in a buffer containing 20 mM MES (pH 6.5), 50 mM NaCl, 2 mM DTT, and 0.016% DMSO. Fluorescence anisotropy was measured at 25 °C. Before plotting, the data were normalized by subtracting the fluorescence anisotropy value at [p58C] = 0 µM from each titration point. Error bars represent the standard deviation of three independent measurements.

Because the fluorescence intensity-based assay is measuring fluorescence quenching, it is

conceivable that this assay could be sensing direct interaction with the probe. We turned to a fluorescence anisotropy (FA) assay to measure the binding affinity by an alternate approach and rule out the possibility of direct probe binding. For this assay, an RNA-primed DNA substrate was used composed of a 6FAM-labelled 22 nt DNA template strand annealed with a complementary 5'-triphosphorylated 12 nt RNA primer. We used the longer template to place the probe further away from the RNA-DNA hybrid region and minimize probe quenching. The K_D values of 3.6 ± 0.1 μM and 2.1 ± 0.1 μM obtained by FA for $p58C_{272-464}$ and $p58C_{266-464}$, respectively are very similar to those obtained by the fluorescence intensity assay (Figure 3.8). To test for direct binding of fluorescein to the p58C constructs, the FA assay was repeated with free fluorescein (see Figure 3.9). The absence of binding of fluorescein in these control experiments confirms that the responses in the FI and FA assays are due to interactions between p58C and the DNA substrate. The narrow range (2.1 - 3.7 μM) and average values (3.2 μM and 2.8 μM, respectively) of K_D for the two constructs measured by the two approaches demonstrate there is no functionally significant difference in the DNA binding properties of these two constructs.

The $p58C_{272-464}$ and $p58C_{266-464}$ constructs have similar electrochemical properties

The [4Fe4S] cluster in $p58C_{272-464}$ was previously shown to function as a redox switch regulating DNA binding and signaling activity (235). The oxidized, $[4Fe4S]^{3+}$ form of this protein was DNA-bound and redox-active; the reduced $[4F\text{-}e4S]^{2+}$ form was loosely associated with DNA and exhibited no DNA-mediated electrochemical signal in a cyclic voltammetry scan. To test whether the same redox switching function was present in $p58C_{266-464}$, we characterized this construct using DNA electrochemistry. Using a multiplexed, DNA-modified Au electrode

platform, we assessed the redox behavior of $p58C_{266-464}$ on a 20-nt duplex DNA substrate with a

3-nt, 5'-ssDNA overhang (235). Electrochemical experiments were performed in anaerobic

conditions to prevent nonspecific degradation of the cluster by atmospheric oxygen (223).

Electrochemically unaltered $p58C_{266-464}$, which is predominantly present in the $[4Fe4S]^{2+}$

oxidation state, displays no redox signal on DNA, as observed with $p58C_{272-464}$. However, upon

electrochemical oxidation to the $[4Fe4S]^{3+}$ state by bulk electrolysis at an applied potential of

512 mV vs. NHE, CV scans of $p58C_{266-464}$ display a large reductive peak near -140 mV vs. NHE.

As observed for $p58C_{272-464}$, this peak disappears after one scan to negative reducing potentials

(Figure 3.10). We next tested if the [4Fe4S] cluster of $p58C_{266-464}$ functions as a redox switch

and found that this construct can cycle between the tightly bound, redox-active $[4Fe4S]^{3+}$ state,

and the loosely associated $[4Fe4S]^{2+}$ state. The $p58C_{266-464}$ construct also displays a robust,

reversible, NTP-dependent signal (Figure 3.11) centered at 142 ± 12 mV vs. NHE on a DNA

electrode. This is very similar to the signal observed for $p58C_{272-464}$ in the presence of DNA and

NTPs, the two necessary substrates for primase activity. Together, these data suggest that the

two human p58C constructs, $p58C_{272-464}$ and $p58C_{266-464}$ have the same redox switching ability;

the electrochemical function of the [4Fe4S] cofactor is consistent in both proteins.
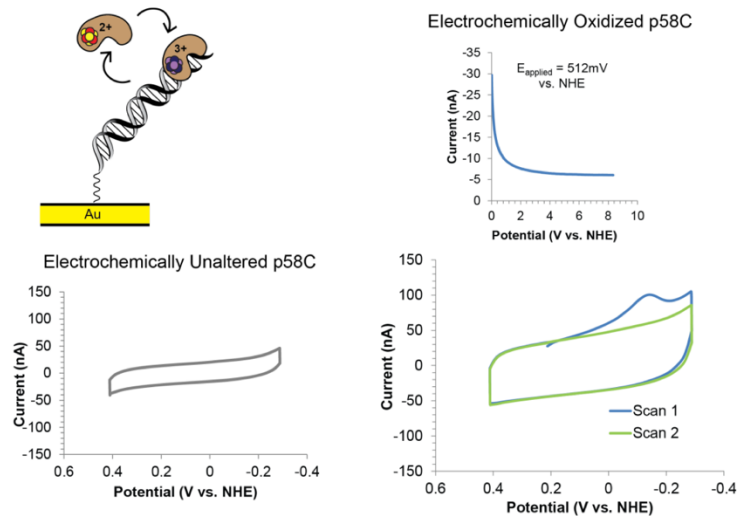
Figure 3.10. p58C$_{266-464}$ participates in redox switching on DNA. The cartoon (top left) depicts p58C DNA binding and redox switching on an Au electrode. (Bottom left) CV scan of electrochemically unaltered p58C$_{266-464}$. (Right) bulk oxidation (above) of p58C$_{266-464}$ and subsequent CV scans (below). This construct displays similar electrochemical behavior to p58C$_{272-464}$. All electrochemistry was performed in anaerobic conditions with 40 μM [4Fe4S] p58C$_{266-464}$ in 20 mM HEPES (pH 7.2), 75 mM NaCl. CV was performed at 100 mV/s scan rate.
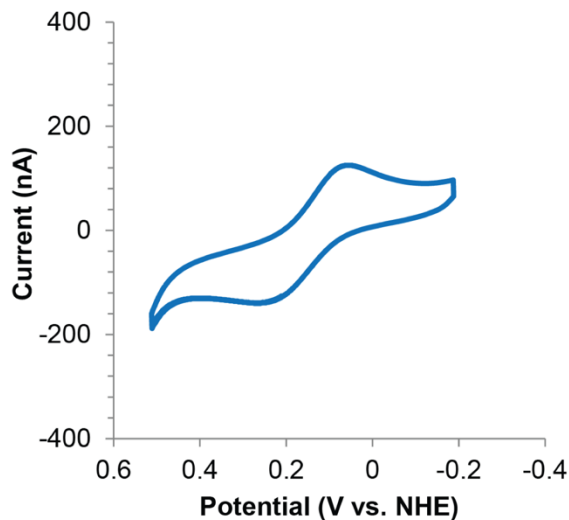
Figure 3.11. p58$_{256-464}$ Electrochemistry with 2.5 mM ATP. Upon addition of 2.5 mM ATP, a reversible redox signal generally appears in CV scans. The signal in the presence of 2.5 mM ATP was centered at an average midpoint potential measured near 142 ± 12 mV vs. NHE. The signal was observed at physiologically relevant redox potential, with a magnitude on the order of 10^2 nC charge transport. All scans were performed under anaerobic conditions on 40 μM [4Fe4S] p58C in 20 mM HEPES (pH 7.2), 75mM NaCl, at 100mV/s scan rate (CV) or 15 Hz (SQWV).

Discussion

It has been proposed that differences in the structures of residues Leu318-His351 in the crystal structures of p58C$_{266-456}$ and p58C$_{272-464}$ are due to the difference in their N-termini (233, 234). It has been hypothesized that Ile271 provides critical hydrophobic interactions stabilizing the helical hairpin in the p58C$_{266-456}$ crystal structure and that the absence of this residue leads to misfolding of Leu318-His351 in p58C$_{272-464}$ and formation of β-strands in the p58C$_{272-464}$ crystal structure. As these differences in structure are in the DNA-binding region of p58C, it is plausible to suggest that this may have significant functional repercussions. However, by determining the crystal structure of a new p58C$_{266-464}$ construct with the same extended N-terminal sequence as the p58C$_{266-456}$ construct, we demonstrated that the differences in the previously reported structures were not due to the difference in sequence, but rather to differences in the crystallization conditions.

Analysis of the p58C$_{266-464}$ and p58C$_{272-464}$ crystal structures reveals that substantial crystal lattice packing interactions stabilize this region in both structures. Specifically, Leu318-Gln329 are seen to interact extensively with the same region in the adjacent molecule in the lattice. The backbone atoms between the adjacent molecules in these regions show extensive hydrogen bonding, stabilizing the beta-like structure. In addition, on both sides of this extended interaction, Phe326 completes a hydrophobic pocket in the adjacent molecule, providing an additional anchor between the two molecules. In the structure of the p58C$_{266-456}$ crystallized under different conditions, a series of stabilizing interactions involve Leu318-His351, but none of these are formed between adjacent molecules in the crystal lattice. A number of crystal contacts in this lattice are found in loop regions and the interfaces of hydrophobic helices. The differences between the crystal contacts in these two models can be attributed to differences in

crystallization conditions, most notably the very high concentration of p58C (50 mg/mL) used originally for $p58C_{272-464}$ (218) and now for $p58C_{266-464}$. It is conceivable that at high concentration p58C forms dimers over time promoting the difference in crystal lattice and this in turn may facilitate formation of the inter-molecular Cys449 disulfide.

Crystallization contacts stabilizing alternate conformations of proteins is a well-demonstrated phenomenon (253). Detailed comparative analysis of crystal structures, structural data obtained in solution, and simulations, have revealed a number of cases in which packing interactions affect backbone and side chain conformations, and even cases of alternate folds (254-256). The addition of p58C to this group provides additional support for the importance of considering crystallization contacts between adjacent molecules in the crystal lattice when analyzing conformational changes of proteins crystallized with co-factors and ligands.

To further establish the equivalence of the p58C constructs, we turned to structural characterization in solution. Both CD and NMR (Figs 4, 5) showed that the absence or presence of residues Gly266-Ile271 has no significant impact on structure in solution. DNA binding and cyclic voltammetry assays (Figs 6 and 7) showed that $p58C_{272-464}$ and $p58C_{266-464}$ have effectively the same biochemical properties. The evidence that the DNA binding region of p58C is capable of adapting to different crystallization conditions is consistent with the need for p58C to bind substrates sequence non-specifically, a property often associated with structural plasticity in sequence non-specific DNA binding domains (e.g. (257)). This raises the question: does the structure of p58C change upon DNA binding? One structure of p58C with a DNA substrate bound is available (137). Comparison of this and the substrate-free structure reveals that only subtle changes in a limited number of side chains is required to bind the substrate (137, 215). Given the extensive similarities between the p58C structures crystallized in different conditions ,

82

there is no reason to think the alternate conformation of residues 318-351 precludes DNA

binding. Moreover, at a more fundamental level, it is the structure *in solution* with and without

substrate that is required to definitely answer this question. Hence, further solution-state

structural characterization of primase in complex with NTPs and primed substrates will be

required to establish if there is a functional role for structural plasticity in the template-binding

region of p58C.

We note that the approximately 3 µM dissociation constants we observe are similar to

those previously reported using fluorescence anisotropy (235), but significantly higher than those

reported for p58C$_{266-456}$ binding to a similar substrate measured using an electrophoretic mobility

shift assay (EMSA) (115). Molecular oxygen is known to be generated while running a

polyacrylamide gel, which may cause redox-active proteins to become oxidized over the course

of an EMSA (250). As p58C is proposed to bind substrates more tightly when oxidized, we

suggest that the EMSA measurements may report the affinity of primarily oxidized p58C,

whereas the fluorescence-based measurements report the affinity of the directly purified p58C,

which is primarily reduced (116, 235). Regardless of the differences in the K_D values measured

in solution and with EMSAs, our data show the DNA binding affinities of p58C$_{272-464}$ and

p58C$_{266-464}$ are effectively the same, indicating that the shorter construct functions normally.

All substrate binding measurements have been performed on p58C in the reduced state.

Although a direct assessment of the substrate binding affinity of oxidized p58C is highly

desirable, such experiments are precluded by the instability of the oxidized cluster. Use of

chemical oxidants causes degradation of the cluster and precipitation of the protein, as does

extended exposure to atmospheric oxygen. Moreover, although the oxidized protein can be

generated electrochemically, the lifetime of the oxidized state is not sufficiently long to make a

reliable direct measurement of DNA binding affinity by traditional methods. There is substantial interest in developing new approaches that would overcome this challenge.

In summary, we have shown that the absence or presence of residues Gly266-Ile271 does not lead to any significant differences in the structure and biochemical properties of the p58C domain of human DNA primase. Beyond addressing the concerns raised about our initial studies, these data strengthen the support for our proposal that the [4Fe4S] cluster serves as a redox switch that modulates the DNA binding affinity of p58C, setting the stage to further asses the role this switch plays in the priming function of human DNA primase. Given how readily p58C appears to switch between the two forms, it is possible that this structural change might occur during the primase catalytic cycle. However, given the lack of solution-state structures for p58C, it is difficult to state with certainty the role each structure may play in the catalytic cycle and whether the ß-structure is compatible with nucleic acid interaction. Further structural characterization of p58C and full-length primase in complex with NTPs and primed substrate may reveal a functional role for structural dynamics in the p58C template-binding region.

CHAPTER IV

FUNCTIONAL DYNAMICS OF HUMAN DNA PRIMASE

Introduction

Processive DNA polymerases synthesize the bulk of the complementary DNA strands during DNA replication. However, these enzymes cannot function without a short oligonucleotide primer with a free 3'-OH group. This functional gap is filled by primases, which generate the requisite short oligonucleotides on a single-stranded DNA (ssDNA) template (236-238). Eukaryotic replicative primase exists in complex with DNA polymerase α (pol α) (88, 122, 131, 133). Primase is an RNA polymerase and generates an initial 8-12 nt RNA primer, which is then handed off to pol α. Pol α extends this RNA primer by approximately 20 nucleotides before handing off the hybrid DNA/RNA primer to the processive polymerases ε and δ.

High-resolution crystal structures of primase domain constructs and full-length human DNA primase are available, but none provide a clear view of how the enzyme functions (36, 137, 143, 144, 215). Human DNA primase has a catalytic subunit (p48) and a regulatory subunit (p58) with two domains (p58N, p58C). To initiate primer synthesis, primase requires catalytic metal ions, two NTPs, and the ssDNA template (37, 131, 132). p58C is known to bind ssDNA and primed templates, and p48 binds the catalytic metals and NTPs, in all cases with modest affinity (36, 112, 113, 143). Some insights have been obtained from the crystal structures available for p58C in complex with an RNA-primed template and for p48 in complex with the catalytic metals and the 3'-NTP (137, 143, 144). Moreover, based on biochemical assays with p48 variants, it was proposed that the R304 residue in p48 recruits the 5'-NTP to the catalytic site (136). However, despite inclusion of the NTPs in the crystallization conditions, only the 3'-NTP is

observed; the 5'-NTP is absent in all structures. The crystal structure of p58C in complex with primed substrate revealed potential interactions of p58C with the 5' triphosphate of the RNA primer, suggesting that p58C may aid in 5'-NTP recruitment (137).

A model based on biochemical and structural evidence has been proposed for priming activity in which primase engages in two distinct enzymatic processes: initial dinucleotide synthesis and primer extension (37, 132, 145). Current structural information suggests that these enzymatic processes involve distinct configurations of primase (145). In this model, p58C is a critical component of the active site for initial dinucleotide synthesis but is not required for elongation activity and travels with the 5' end of the nascent primer. Disengagement from the p48 active site potentially positions p58C to modulate primer termination. While this model is consistent with our current structural and biophysical understanding of primase, it has not been validated with any direct experimental data on full-length primase in complex with catalytically-relevant substrates. This is due, at least in part, to the difficulties inherent in generating crystals of a flexible protein in complex with substrates and cofactors that bind with modest affinity. Nevertheless, substrates and cofactor binding is a critical step for generating the atomic resolution structures of full-length primase in catalytically competent states that would inform an atomistic understanding of the mechanisms of primer synthesis.

The proposed dual mechanism for primase relies on significant rearrangement of the p58C domain relative to p48. Small-angle x-ray scattering is uniquely suited to studying this type of large-scale inter-domain rearrangement and is especially valuable for systems like primase that are not amenable to crystallization (157, 182, 183). In this chapter, we present SAXS data acquired for primase alone, in the presence of the substrates required for initiation, and in complex with a primed substrate. Using the most state-of-the-art approach, SEC-SAXS, revealed

primase exhibits complex equilibria under all conditions. Our interpretation of this data is that free primase transiently oligomerizes and that this phenomenon increases under conditions promoting initiation, as well as in the elongation complex. Using singular value decomposition (SVD) of the complete peak extracted from the SEC (258), the SAXS data were deconvoluted into a mixture of two states, which have tentatively been assigned to monomers and dimers. In this model, the addition of substrates and cofactors for initiation has almost no effect on primase, but the conformational ensemble of the elongation complex is on average more extended than the free protein, consistent with the current model for how primase functions.

Materials and Methods

Expression and purification of human DNA primase

Full-length primase was expressed and purified as previously described (218, 235). In short, plasmid DNA was transformed into BL21RIL (DE3) cells (Novagen), and cultured at 37 °C to an OD_{600} of 0.6, when flasks were transferred to a 18 °C incubator with shaking. After 30 minutes, protein expression was induced through addition of isopropyl 1-thio-β-D-galactopyranoside to a final concentration of 5 mM. Media was also supplemented at this time with ferric citrate and ammonium ferrous citrate to a final concentration of 0.1 mg/mL. The primase subunits were expressed at 18 °C for 18 hours prior to harvesting and freezing at -80 °C. After lysis, primase was first purified by nickel affinity chromatography (Amersham Biosciences). The 6xHis tag on p48 was the cleaved with H3C protease and primase dialyzed into a low-imidazole buffer (218). Primase was then repassed over the nickel column to remove the H3C protease and un-cleaved protein. A heparin column was used to remove residual contaminants before passing over a Sephadex S200 sizing column to remove aggregates and

buffer exchange into storage buffer: 20 mM HEPES, 150 mM NaCl, 100 mM ammonium

sulfate, 1 mM TCEP, 2 % glycerol, pH 7.2 (218).

Generation and purification of triphosphorylated RNA primers

Triphosphorylated RNA primer for structural studies was synthesized *in vitro* by adding

150 µL of 5 mg/mL P266L T7 RNA polymerase (248) to a 10 mL reaction containing 5 mM

ATP, CTP, GTP, and UTP (Sigma Aldrich), 700 units RNAsin RNAse inhibitor (Promega), 10

units inorganic pyrophosphatase, 1.5 µM annealed DNA template (forward primer: 5'-

CGCCGAGCCTATAGTGAGTCGTATTA-3'; reverse primer: 5'-

AATACGACTCACTATAGG-3'), 40 mM Tris (pH 8.0), 38 mM $MgCl_2$, 1 mM spermidine (pH

8.0), 5 mM DTT, and 0.01% Triton X-100. This mixture was incubated at 37 °C for 5 hours, and

RNA products were precipitated by the addition of 0.3 M sodium acetate (pH 5.2) and 30 mL

100% ethanol at 4 °C. After incubation at -20 °C overnight, the precipitate was harvested

through centrifugation at 3700 g for 10 minutes. Precipitated RNA was resuspended in 4 °C 80%

ethanol and centrifuged at 3700 g for 10 minutes, after which the supernatant was discarded to

remove salts. A speed vacuum was used to remove any ethanol remaining in the pellet before

further purification.

Dried RNA pellet was resuspended in gel loading buffer and purified on a 25% TBE-

polyacrylamide gel according to standard methods. RNAse-free H_2O and aliquoted prior to

further purification. RNAs used for structural studies were HPLC purified on a Luna 5 µM

C18(2) 100 Å 250x4.6 mm column (Phenomenex). Buffer A: 0.1 M ammonium formate; Buffer

B: acetonitrile; flow rate: 1.5 mL/min. 1-5% Buffer B over three minutes, 5-8% Buffer B over 22

minutes, then 80% Buffer B for five minutes. RNA typically eluted around 11 minutes. HPLC-

purified RNA was further validated with mass spectrometry, which confirmed that a product of the expected mass had been generated.

Preparation of SAXS samples

Samples were prepared in two sets of SAXS buffers: a low salt buffer to limit electrostatic screening and promote binding of substrates, and a high salt buffer to promote primase stability (218). The low salt buffer contained 20 mM HEPES, 100 mM NaCl, 2 mM $MgCl_2$, 0.2 mM $MnCl_2$, 1 mM TCEP, 2% glycerol, pH 7.2. The high salt buffer contained 20 mM sodium cacodylate, 200 mM NaCl, 2 mM $MgCl_2$, 0.2 mM $MnCl_2$, 1 mM TCEP, 2% glycerol, pH 7.2. To buffer-exchange primase for SAXS, purified primase was passed over an 24 mL Superdex S200 column that had been pre-equilibrated in the SAXS buffer of choice. After elution from the column, primase-containing fractions were not concentrated to limit aggregation.

To prepare samples for high-throughput SAXS (HT-SAXS) under conditions that favor the active configuration for initiation without allowing primase to initiate primer synthesis *in vitro*, stock solutions of the ssDNA template (sequence: 5'-TCTCTCTCTCTCTCTGCTCTCTCTCTC-3'), CTP, and AMP-PNP (Sigma Aldrich) were prepared in SAXS buffer, then added to buffer-exchanged primase to final concentrations of 10 µM primase, 500 µM ssDNA, 500 µM CTP, and 500 µM AMP-PNP. To prepare the elongation complex for HT-SAXS, stock solutions of ssDNA template (sequence: ATAACTCGCCGAGCCAACATA) and triphosphorylated RNA primer (sequence: GGCUCGGCG) were mixed in SAXS buffer, then annealed at a 1.1:1 ratio of RNA:DNA by heating at 90 °C for two minutes, then slowly cooling to room temperature (115). The annealed

primed substrate was then added to buffer-exchanged primase to final concentrations of 10 μM primase and 100 μM primed substrate. HT-SAXS samples were then aliquoted into a 96-well plate and kept at 4 °C until data collection at the SIBYLS beamline. Size exclusion chromatography-coupled SAXS (SEC-SAXS) samples were prepared similarly, except that concentrated primase in SAXS buffer and the initiation and elongation master mixes were shipped separately at -80 °C to the SIBYLS beamline. There, the initiation sample was generated by adding the initiation master mix (containing ssDNA template, CTP, and AMP-PNP at a concentration of 2 mM) to primase to reach final concentrations of 30 μM primase, and 1.1 mM ssDNA, CTP, and AMP-PNP. Elongation samples were generated by adding primed substrate to primase to reach final concentrations of 30 μM primase and 220 μM primed substrate.

SAXS data collection

In SAXS data collection, a solution-state sample is exposed to high-flux x-rays; these generate unique scattering patterns as they interact with the sample, which are then recorded by a detector (see **Figure 4.1A**) (157, 181). To make this process high-throughput for HT-SAXS data collections, the SIBYLS team utilizes a Tecan Freedom Evo liquid handling robot to load samples directly into a needle containing the sample cell; this sample cell is then positioned in line with the x-ray source and shutter box and exposed to x-rays for 10 s, subdivided into 32 300 ms frames (see **Figure 4.1B**). After exposure, the robot returns the sample back to the 96-well plate and rinses and washes the sample cell. Frames that do not exhibit radiation damage are then averaged to generate an averaged scattering profile for further data analysis. This instrumentation setup, in combination with a Pilatus3 2M pixel array detector, allows collection of a 96-well plate in 2.5 hours.

In SEC-SAXS data collection, 50 µL of sample is injected onto the size exclusion (SE) column with an autosampler. Column eluent is split 2:1 between synchrotron radiation and a Wyatt multi-angle light scattering (MALS)/quasi-elastic light scattering (QUELS)/refractometer setup (see **Figure 4.1C**), enabling SAXS data collection as it comes off the column. After buffer subtraction, scattering data collected for the elution peak of the analyte of interest are averaged to generate the scattering profile for further data analysis.
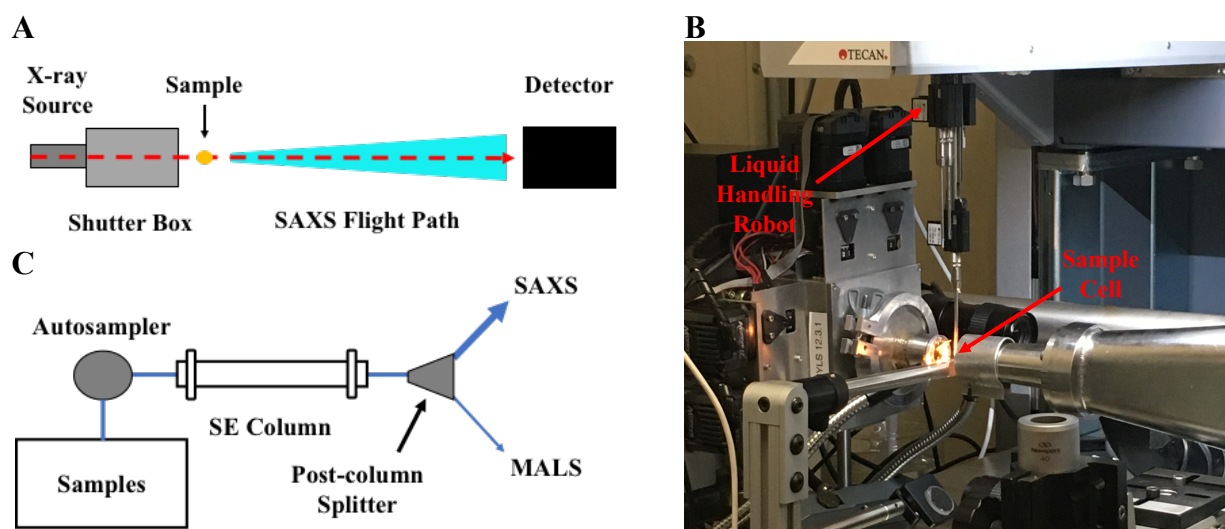
Figure 4.1. Overview of SAXS experimental setup. A. Diagram of SAXS endstation at SIBYLS. Red arrow denotes direction of x-rays. B. HT-SAXS sample collection setup. C. SEC-SAXS flowpath diagram.

SAXS data were collected at the ALS beamline 12.3.1 LBNL Berkeley, California (181, 259). For HT-SAXS experiments, the X-ray wavelength was $\lambda=1.27$ Å and the sample-to-detector distance was set to 1.5 m, resulting in scattering vectors ranging from 0.01 Å^{-1} to 0.5 Å^{-1}, where the scattering vector is defined as $q = 4\pi \sin\theta/\lambda$ and 2θ is the scattering angle. Samples were exposed for a total exposure time of 10 seconds, divided into 32 frames of 0.3 seconds each. All HT-SAXS data were collected by Dr. Greg Hura at 10 °C. Averaged scattering curves were generated in the FrameSlice web app (http://sibyls.als.lbl.gov/ran), and scattering profiles

were analyzed using the SCÅTTER software package (version 3.1r) as previously described(157, 181, 260).

For SEC-SAXS experiments, 50 µL of sample was injected onto a Shodex KW-803 column on an Agilent 1260 Infinity LC system at a flow rate of 0.5 mL/min. 3 s exposures were collected for each frame, resulting in approximately 800 3 s exposures collected over the course of a 40 min elution. The X-ray wavelength was λ=1.03 Å and the sample-to-detector distance was set to 1.5 m, resulting in scattering vectors ranging from 0.013 Å^{-1} to 0.5 Å^{-1}. All SEC-SAXS data were collected at 20 °C by Dr. Michal Hammel and Daniel Rosenberg. Data were processed in the SCÅTTER (version 3.1r) software package as previously described (181). SVD deconvolution was completed on SEC-SAXS data by Dr. Michal Hammel with the BioXTAS RAW software package (258). P(r) curves were calculated for the deconvoluted monomer curves using data in the $0.02 < q < 0.3$ region.

Results

Recent structural and biochemical data for human DNA primase indicate p58C contributes to binding of the DNA substrate in both phases of the dual mechanism (112, 115). p58C must first be positioned along with the template over the active site of p48, which binds the catalytic metals and the 3' nucleotide (135, 143, 144). In our model, some combination of interactions of p48 and p58C with the triphosphate allows recruitment of the 5' nucleotide to complete the assembly of the active complex and generate the initial dinucleotide.

The crystal structure of free human DNA primase has an extended configuration that is inconsistent with formation of the initiation complex, as the DNA binding domain, p58C, and the catalytic active site for ribonucleotide synthesis, p48, are on opposite ends of the structure (137,

141). Therefore, current models for primase function predict that substantial changes in configuration are essential for the p58C domain to reposition over the p48 active site in order to initiate primer synthesis, and that this architectural rearrangement is facilitated by the flexible linker separating the p58N and p58C domains (145). Moreover, since p58C remains bound to the 5' phosphate of the 5' nucleotide, the domain will shift away from the p48 active site as the primer is elongated (115, 137).

Here, we used SAXS analyses to investigate the configurational dynamics of free primase and characterize the global domain rearrangements that occur when primase forms the initiation and elongation complexes in solution. To promote sampling of the primase initiation complex, we designed a 27 nt ssDNA template that will base pair only with two nucleotides, a 5'-CTP and a non-hydrolyzable AMP-PNP, at a specific site (bold):

<div align="center">5'-TCTCTCTCTCTCTCTGCTCTCTCTCTC-3'</div>

To produce a primase elongation complex, we used T7 RNA polymerase to generate a short, 5'-triphosphorylated RNA oligo and annealed it to a ssDNA substrate, thereby generating a 5'-triphosphorylated primed substrate with ssDNA overhangs on both ends:

<div align="center">5'-PPP- GGCUCGGCG-3'
3'-ATACAACCGAGCCGCTCAATA-5'</div>

Dr. Matthew Thompson designed the initiation substrate and collected initial SAXS results for free primase under initiation conditions; I replicated these results prior to SEC-SAXS data collections.

SAXS analysis of free human DNA primase

To test if free primase occupies the configuration seen in the crystal structure, we obtained SAXS data for the free protein under both low salt (100 mM NaCl) and high salt (200 mM NaCl) conditions. Because primase was significantly more aggregation-prone under the low

salt conditions, data collected on primase in the high salt conditions was taken forward for further analysis. Guinier analysis of the scattering profile (Figure 4.3A) obtained under high salt conditions revealed that the Guinier plot (Figure 4.2) is linear to a q•Rg value of at least 1.3, indicating that this sample was free of aggregation (157, 181, 183). Examination of the Kratky plot (Figure 4.3B) reveals a distribution consistent with the presence of globular domains connected by one or more flexible linkers (157, 184). The Porod exponent of 3.0 derived from the data (Table 4.1) is consistent with the system containing a high degree of flexibility (157, 181, 184).
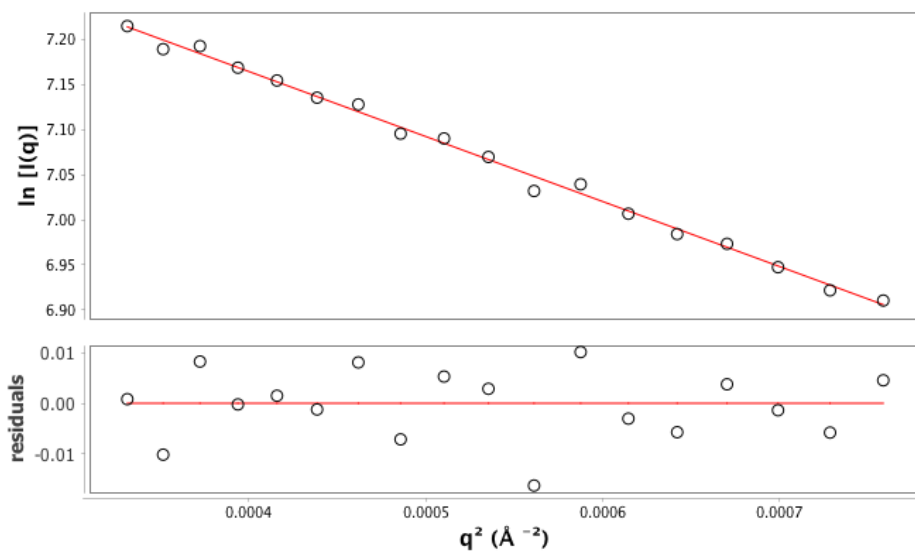
Figure 4.2. Guinier plot of free primase.
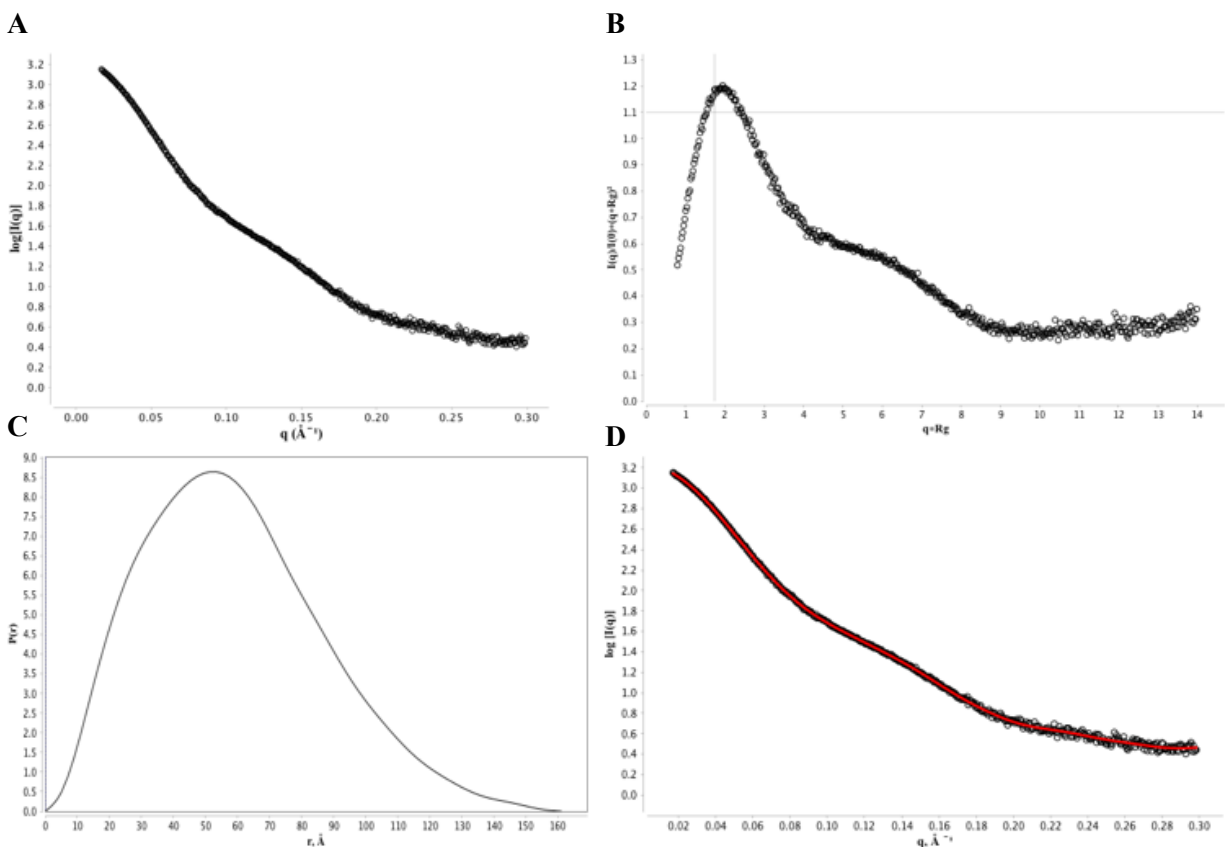
Figure 4.3. SAXS scattering data for free primase. A. Scattering intensity profile. B. Dimensionless Kratky plot (normalized using the real space Rg calculated from the P(r) distribution). C. P(r) distribution. D. Fit of the Fourier transform of the P(r) distribution to the experimental scattering curves.

The P(r) function, referred to as a distance distribution or the pair-distance distribution function, is an autocorrelation function corresponding to the Patterson function in x-ray crystallography (157, 183). This function can be determined by applying a Fourier transform to the experimental scattering data, which yields a distribution that provides information about the distance between electrons in the scattering particle. The distribution is sensitive to changes in interatomic distances, hence changes in molecular architecture will be reflected in this analysis of the data. However, in practice, the indirect Fourier transform methods used to calculate the P(r) distribution from experimental scattering data heavily depend on accurate estimation of D_{max}, more so than if a direct Fourier transform were applied. To gauge the quality of D_{max}

selection and the resulting P(r) distribution, we can apply a Fourier transform to the P(r)

distribution and assess how well it fits the experimental scattering data. Correct determination of

the D_{max} should result in a Fourier transform of the P(r) distribution that closely fits the scattering

data, which we observe for free primase (Figure 4.3D). Examination of the P(r) distribution for

free primase reveals a primary peak at ~50 Å, with a shoulder at ~30 Å, which can be attributed

to the separation of the p58C and p48/p58N domains in a manner similar to what we would

expect for the elongation complex (Figure 4.4).
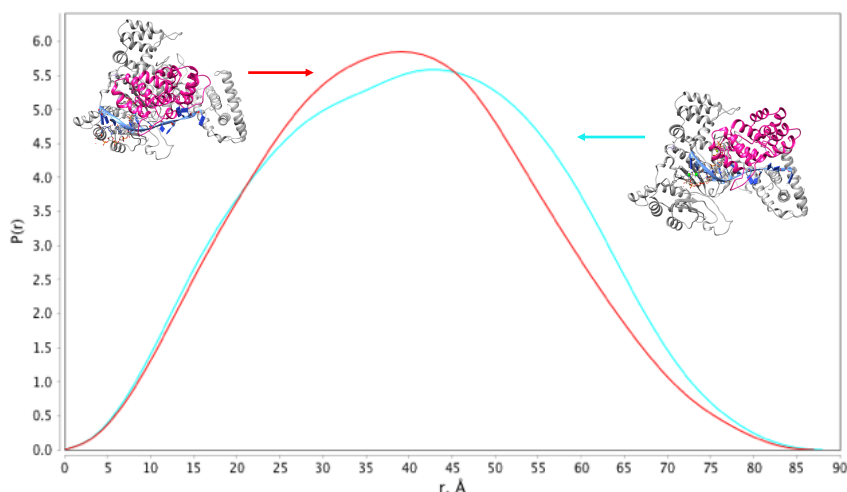
Figure 4.4. P(r) functions for models of the initiation (red) and elongation (cyan) complexes. Theoretical scattering profiles generated with FoXS. p48/p58N structural platform is depicted in grey in all panels. p58C is shown in pink for initiation/elongation models and in many colors for the free primase models. The DNA template is shown in blue, and RNA primer is in orange/red sticks. Models of primase in complex with primed substrate were generated by taking the p48 structure from PDB 4LIL, the p58N structure from PDB 4RR2, and the p58C structure from PDB 5F0S and aligning the 5' and 3' ends of the RNA primer in PDB 5F0S with the triphosphate in the p48 active site to generate the initiation and elongation complex models, respectively.
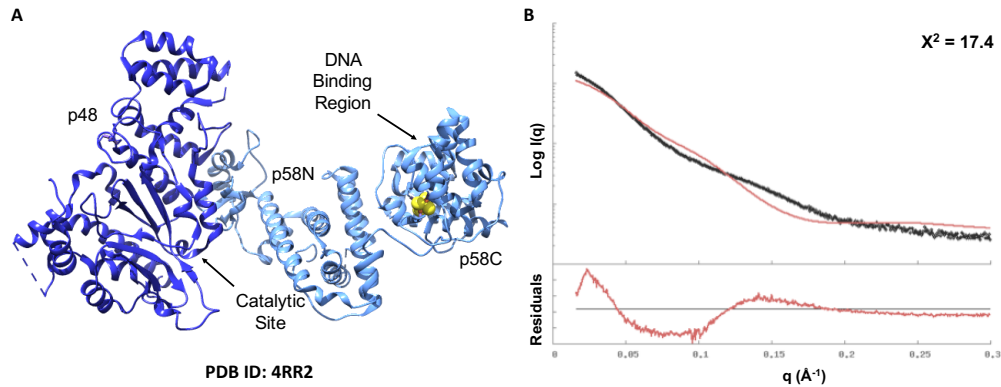
Figure 4.5. Comparison of experimental and back-calculated scattering curves for free primase. A. Crystal structure of full-length primase. p48 and p58 are colored dark blue and light blue, respectively. Iron-sulfur cluster is depicted in yellow spheres. B. Comparison of experimental scattering profile for free primase and scattering profile back-calculated from the primase crystal structure in (A) using FoXS.

To determine if the solution scattering data corresponds to the x-ray crystal structure of free primase, the SAXS scattering profiles and P(r) functions were back-calculated from the PDB coordinates using FoXS (188, 189). Overlays comparing the experimental to the back-calculated curves reveal a substantial disagreement (Figure 4.5). In order to better understand the range of inter-domain configurations (configurational space) accessible to primase, a set of 10,000 different primase configurations were generated by Dr. Matthew Thompson using the RANCH pool generator (Figure 4.6) (183, 261). When Rg values are plotted against Rg frequency in this ensemble of structures, one can see that, with a Rg of 37.5 Å, the crystal structure of primase represents an intermediate state (Figure 4.7). Incorporating the analysis of flexibility reflected in the Kratky plot and Porod exponent, our view of free primase is that it has substantial inter-domain dynamics, such that it can occupy the configuration observed in the crystal structure, but that it must also sample additional configurations, including some that have considerably more separation of p58C from p48/p58N than that observed in the crystal structure.
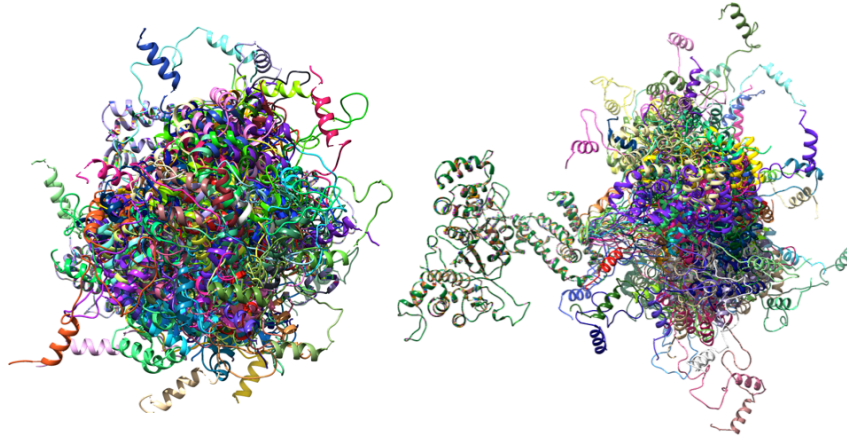
Figure 4.6. Range of configurational space sampled in pool of free primase structures.
10,000 configurations of free primase were generated using RANCH by Dr. Matthew Thompson.
Configurations with the 50 smallest (left) and 50 largest (right) Rg values are depicted here. All
structures are aligned to the p48/p58N structural domain.
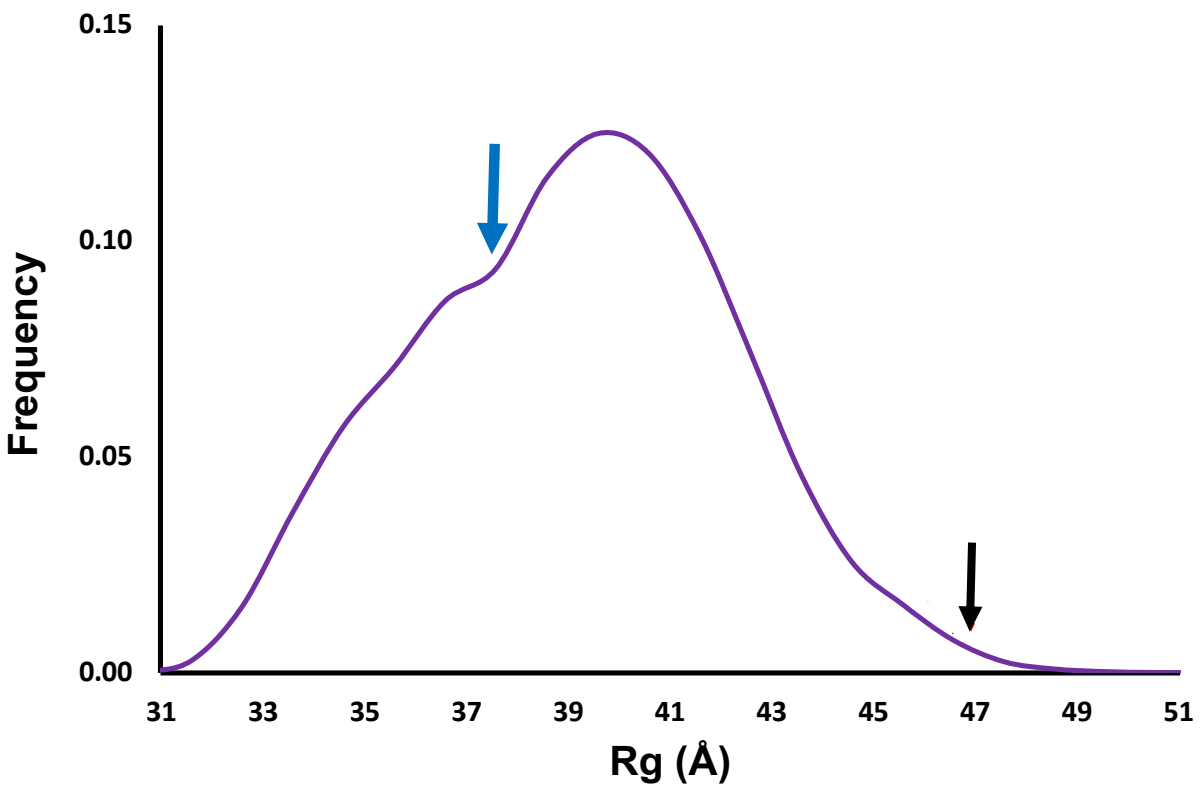
Figure 4.7. Plot of Rg frequency versus Rg value for pool of 10,000 primase structures.
Arrows indicate the Rg values calculated for the crystal structure (blue) and from Guinier
analysis of the scattering curve for free primase (black).

SAXS analysis of human DNA primase under conditions for initiation.

Guinier analyses of the scattering profile (Figure 4.9A) of primase in the presence of the template and cofactors required for initiation reveals that the Guinier plot (Figure 4.8A) is linear to a q•Rg value of at least 1.3, indicating that the sample is also free of aggregation. Examination of the Kratky plot (Figure 4.9B) reveals a slightly less hyperbolic tendency at q•Rg ~ 4 than we observe for free primase, suggesting that there is less flexibility in the system (157). This observation is supported by the Porod exponent (3.6), which is significantly higher than that determined for free primase (3.0). This assessment is also supported by the P(r) distribution, which is more symmetric than that of free primase and in which we observe near loss of the shoulder at ~30 Å, suggesting less average domain separation in the configurational ensemble.
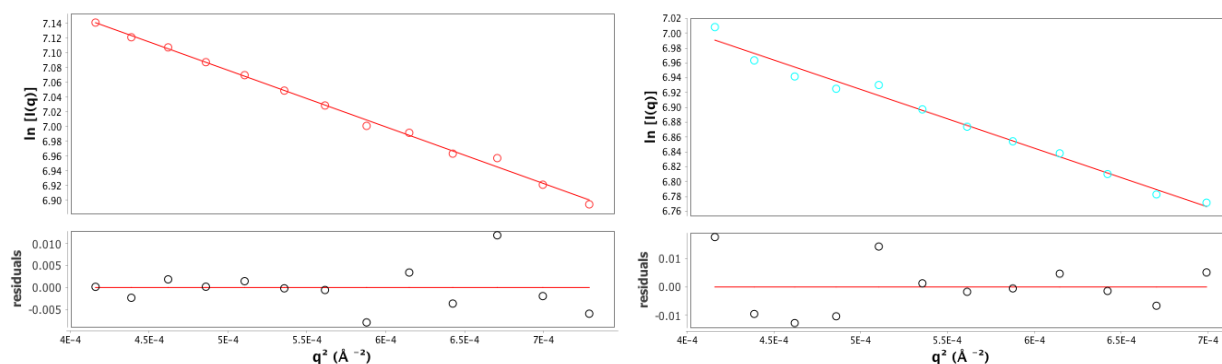
Figure 4.8. Guinier plots of primase in the presence of initiation and primed substrates. Left, red: primase in the presence of initiation substrates. Right, cyan: primase in complex with primed substrate.
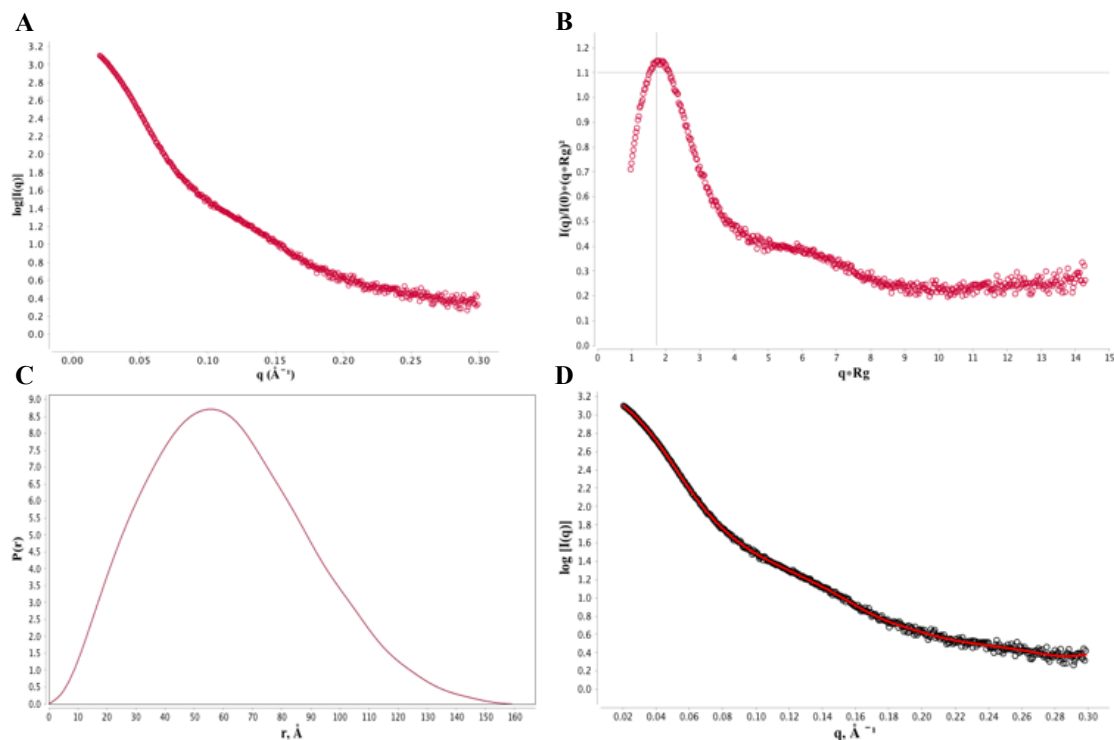
Figure 4.9. SAXS scattering data for primase in the presence of initiation substrates. A. Scattering intensity profile. B. Dimensionless Kratky plot (normalized using the real space Rg calculated from the P(r) distribution). C. P(r) distribution. D. Fit of the Fourier transform of the P(r) distribution to the experimental scattering curves.

SAXS analysis of human DNA primase in the elongation complex

Guinier analyses of the scattering profile (Figure 4.10A) of primase in the presence of primed substrate reveals that the Guinier plot (Figure 4.8B) is linear to a q•Rg value of at least 1.3, indicating that this sample is also free of aggregation. Examination of the Kratky plot for primase in the presence of primed substrate reveals a more hyperbolic tendency in q•Rg ~ 4 than is observed in the Kratky plots of either free primase or the initiation complex (see Figure 4.4B, Figure 4.9B, and Figure 4.10B), suggesting that primase may in fact be more dynamic under these conditions (157). This assessment is supported by the Porod exponent which, at 2.8, is slightly less than that determined for free primase (3.0) and significantly less than that determined for primase in the presence of initiation cofactors and template (3.6). The P(r)

100

distribution reveals a re-emergence of the shoulder at ~30 Å, suggesting increased domain

separation relative to the distribution of configurations under the conditions promoting initiation

(Figure 4.4).
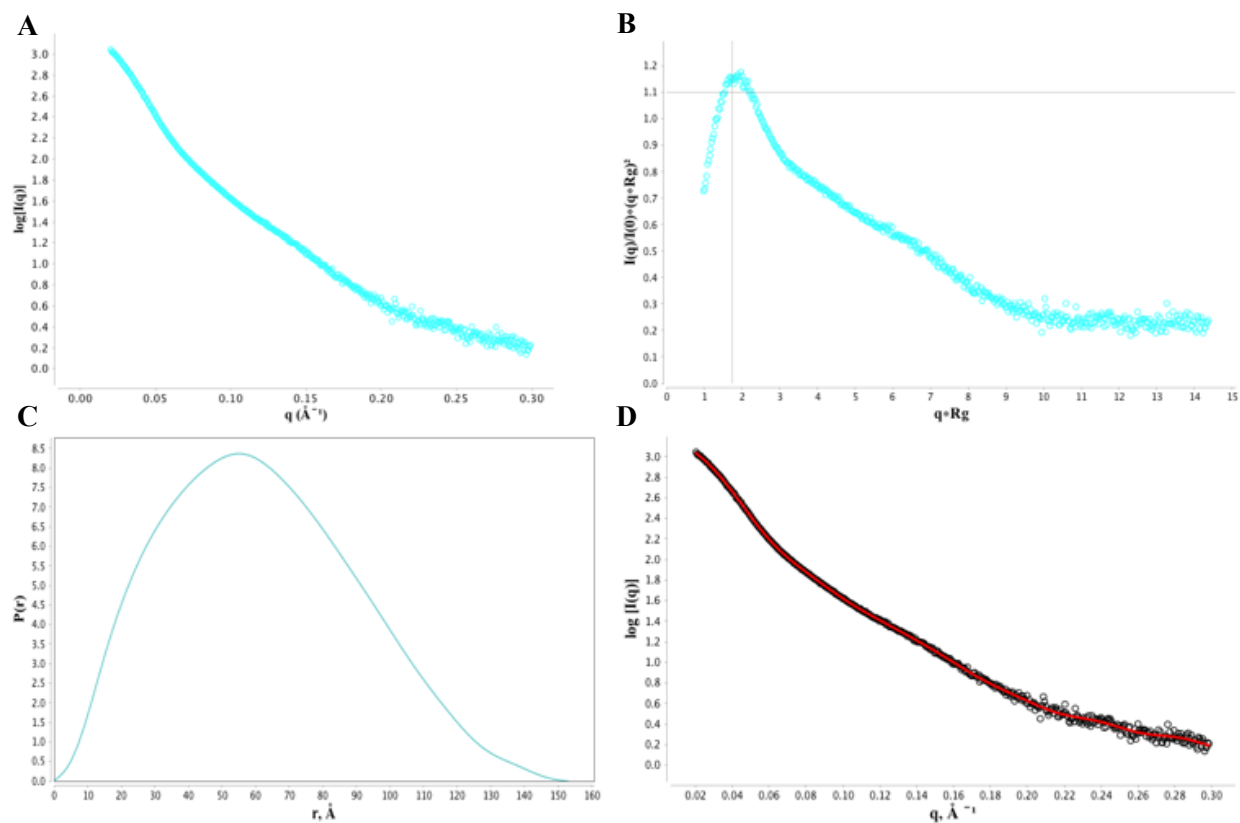
Figure 4.10. SAXS scattering data for primase in the presence of primed substrate. A. Scattering intensity profile. B. Dimensionless Kratky plot (normalized using the real space Rg calculated from the P(r) distribution). C. P(r) distribution. D. Fit of the Fourier transform of the P(r) distribution to the experimental scattering curves.
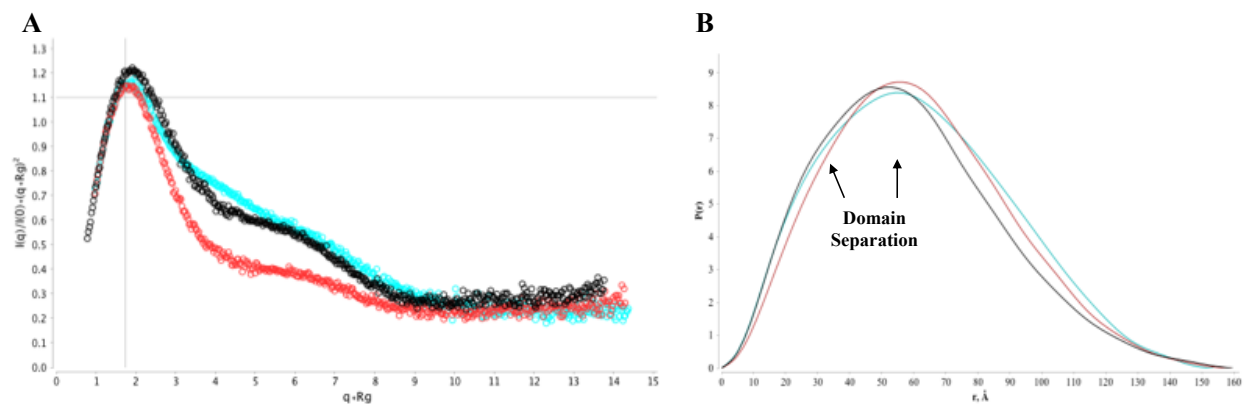
Figure 4.11. SAXS scattering data for free primase and primase in the presence of initiation and elongation substrates. Scattering data for free primase (black) and primase in the presence of initiation (red) and elongation (cyan) substrates. A. Overlay of dimensionless Kratky plots (normalized using the real space Rg calculated from the P(r) distribution). B. P(r) distributions.

Sample	I(0)	Rg (Guinier)	Rg (Real space)	Vc	Volume	MW	r	Dmax	Px
Free Primase	1.7E+03	46.8	46.0	940	430,000	150,000	58.9	161	3.0
Initiation Complex	1.7E+03	47.8	47.8	1120	540,000	210,000	61.8	159	3.6
Elongation Complex	1.5E+03	48.4	48.1	1015	620,000	170,000	61.7	153	2.8

Table 4.1. **Parameters derived from the SAXS data.**

Comparative SAXS analysis of human DNA primase free, under conditions promoting initiation, and in the elongation complex

Figure 4.11 contains direct comparisons of the Kratky plots and P(r) distributions of primase in the three conditions. Analysis of the P(r) distributions displayed in Figure 4.11B support a model in which the p48/p58N and p58C domains tumble relatively freely in free primase and form a more globular configuration in the presence of initiation substrates. The configurational space of the elongation complex ensemble is in between that of the free primase and the conditions promoting initiation, suggesting more domain separation than in the initiation complex and less than the free protein. However, this neat interpretation of the P(r) data conflicts in some ways with the flexibility analysis of the elongation complex, as reflected by the Porod

exponent and Kratky plot of the elongation complex relative to the free primase. Additionally, the Rg value for free primase (as calculated from both the Guinier region and from the P(r) distribution) was lower than that observed for either the initiation or elongation samples, which directly conflicts with the idea that the formation of the initiation complex requires a more compact configuration of primase. Additionally, the observed molecular weight values of 150,000 kDa, 210,000 kDa, and 170,000 kDa (Table 4.1) were significantly higher than the theoretical values of 109 kDa, 117 kDa, and 118 kDa for free primase, initiation, and elongation samples, respectively. Further insights were provided by considering the experimental Rg values under the three conditions in light of the distribution of Rg values of the ensemble calculated to define the accessible configurational space available to primase (Figure 4.7). It is remarkable that all three experimental values correspond to structures at the extreme high end of this distribution. Together, these observations led us to suspect that some degree of oligomerization or other perturbation of the system may be occurring in all of the primase samples. To test this hypothesis, we turned to SEC-SAXS.

SEC-SAXS analysis of human DNA primase

In SEC-SAXS, sample components are separated with SEC prior to SAXS data collection. Direct coupling of the SEC column to the SAXS flow cell allows SAXS data to be collected continuously as eluent comes off the column (Figure 4.1C). Scattering profiles corresponding to the peak of interest can then be selected and averaged for further analysis without interference from aggregates and other resolvable contaminants. This approach empowers evaluation of aggregation-prone samples and modest-affinity complexes, as the data can be analyzed to parse out contributions from aggregates and, potentially, complex from

unbound binding partners.

SEC-SAXS chromatograms were displayed as an overlay of SAXS signal (as expressed by the integral of ratio to background) and Rg values (Figure 4.12A). In a stable, monomeric species, Rg values are stable ("flat") in the region where the protein of interest is being measured. Positive or negative trends over the course of a peak indicates transient formation of multimeric complexes. This can be further assessed with an autocorrelation plot, which enables determination of which scattering profiles are similar; one expect regions in which protein is eluted to be more similar. When displayed using a heat map of the Durbin-Watson statistic in which values near 2 are blue and values near 0 and 4 are red, "patches" of blue indicate regions where the scattering profiles correlate. Areas where the Rg values are stable and the autocorrelation plot is blue can be averaged to generate an averaged spectrum of the species of interest. Knowing this, we can quickly see from Figure 4.12 that free primase is not homogenous across its SEC peak.
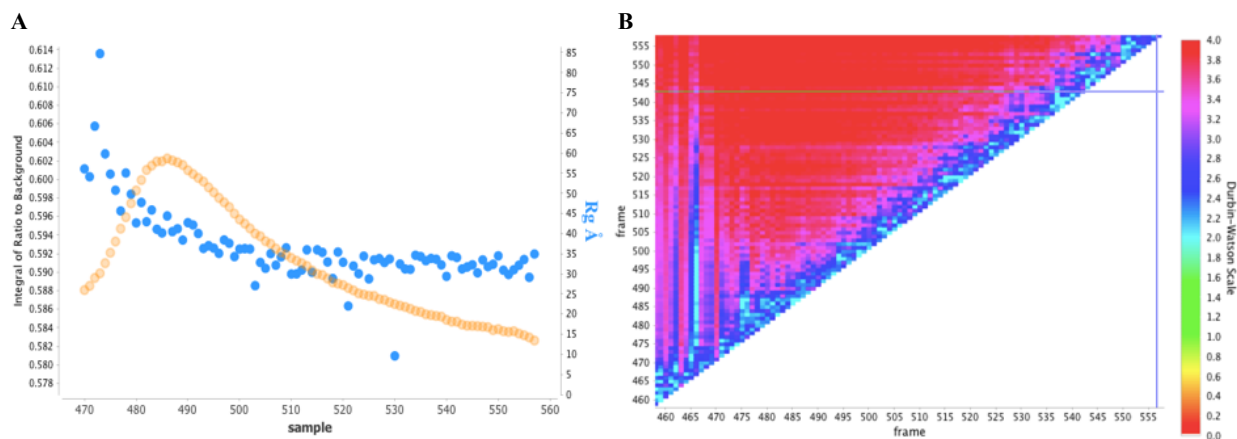
Figure 4.12. SEC-SAXS with free primase. A. Signal plot of free primase peak with signal in orange and Rg values in blue. B. Residuals similarity plot for primase peak.
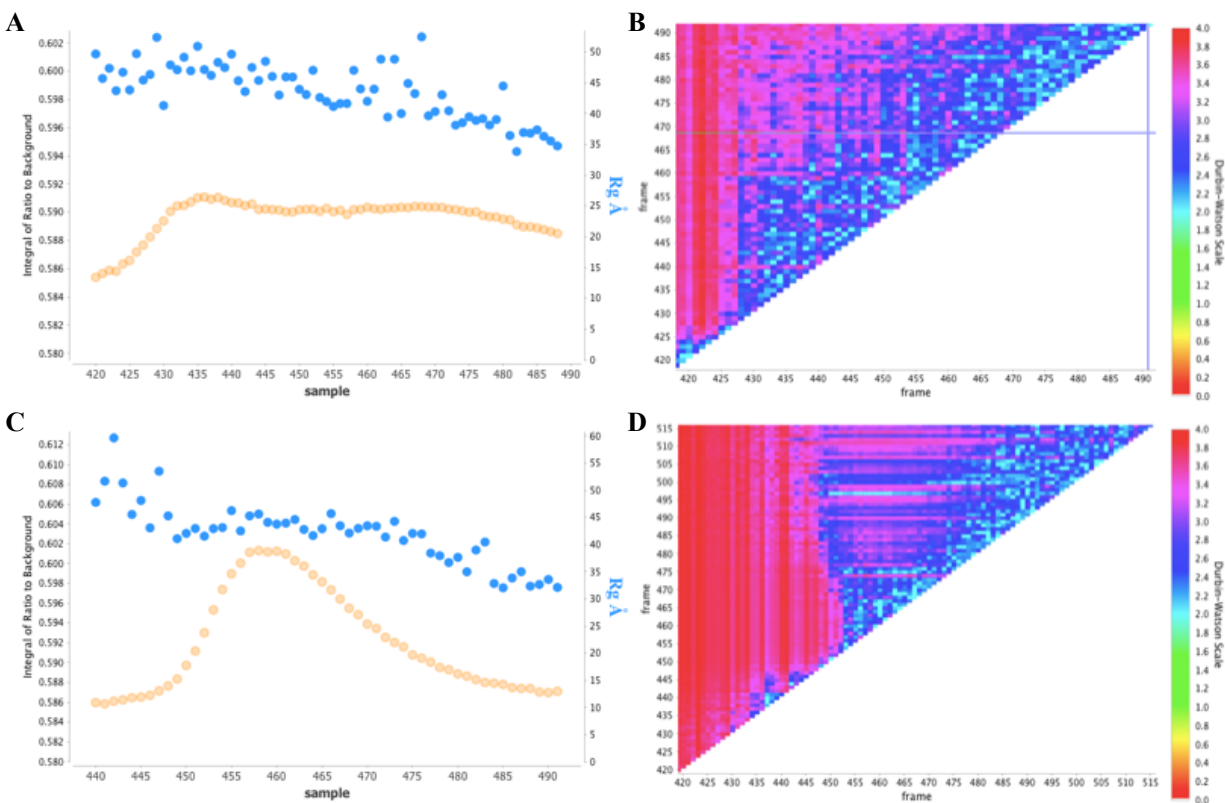
Figure 4.13. SEC-SAXS of primase under conditions promoting initiation and for the elongation complex. A. Signal plot of primase peak under conditions promoting initiation with signal in orange and Rg values in blue. B. Residuals similarity plot for scattering data shown in A.). C. Signal plot of elongation complex peak with signal in orange and Rg values in blue. D. Residuals similarity plot for elongation complex peak.

As we can see in Figure 4.13, we also observe this phenomenon in the elongation and, to an even greater extent, in the initiation SEC-SAXS runs, suggesting that addition of cofactors and a substrate induce even greater degrees of heterogeneity. With the data currently available, it is difficult to determine whether the presence of cofactors and substrate enhances the same phenomenon observed for free primase or if there is a separate, substrate-dependent mechanism at play.

Detection of transient oligomerization of primase

To assess the effect of passing primase over a size exclusion column on the scattering

profiles of free primase and primase in the presence of initiation and elongation substrates, I first

generated average scattering profiles for the primase elution peak of each sample and compared

them to the initial SAXS data, collected in high throughput mode (HT-SAXS). SEC-SAXS Rg

values are significantly lower for all samples, moving into the range that we would expect from

our examination of the primase configurational space (see Table 4.2 and Figure 4.14). However,

the overall trend is similar in that the Rg values for primase under conditions promoting

initiation and the elongation complex are larger than that observed for free primase. A similar

phenomenon is observed for the volumes, molecular weights, and D_{max} values, which are, again,

smaller but qualitatively trend in a manner similar to the values calculated from the HT-SAXS

data. As we can see from Figure 4.15B, the P(r) distributions for the averaged SEC-SAXS

profiles are qualitatively similar to those generated for the HT-SAXS profiles, and the free

primase and elongation complex curves again exhibit a "shoulder" suggesting domain separation.

The only significant difference is that the D_{max} value for the elongation complex is significantly

larger than that observed for the free primase or under conditions promoting initiation, possibly

suggesting a greater population of open configurations in the elongation complex ensemble.

Overall, these data support the idea that a small amount of aggregates or other contaminants may

have artificially inflated the apparent size of primase in the HT-SAXS data.

Sample	I(0)	Rg (Guinier)	Rg (Real space)	Vc	Volume	MW	r	Dmax	Px
Free Primase	7.5E+02	39.3	40.1	790	350,000	130,000	51.6	126	3.2
Initiation Complex	2.2E+02	43.7	42.3	875	400,000	140,000	54.8	136	3.8
Elongation Complex	7.3E+02	44.2	45.1	847	420,000	130,000	57.3	158	3.0

Table 4.2. **SEC-SAXS metrics.**
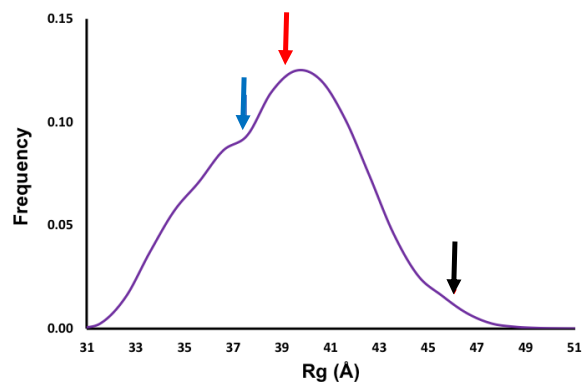
Figure 4.14. Plot of Rg frequency versus Rg value for pool of 10,000 primase structures. Arrows indicate the Rg values calculated for the crystal structure of the p48/p58 monomer (blue) and from Guinier analysis of the initial (black) and SEC-SAXS (red) scattering curves for free primase.
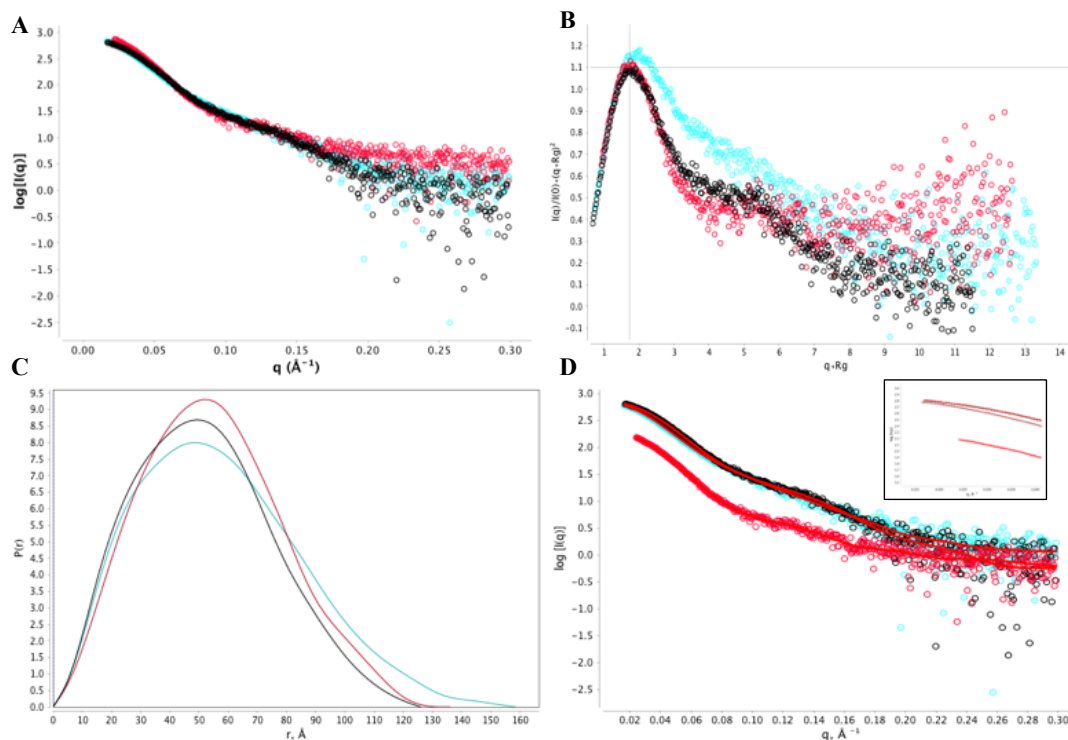
Figure 4.15. Comparison of SEC-SAXS data for primase free, under conditions promoting initiation and in the presence of the elongation complex. Scattering data for free primase (black), conditions promoting initiation (red), and the elongation complex (cyan). A. Overlay of scattering intensity profiles. B.) Overlay of dimensionless Kratky plots (normalized using the real space Rg calculated from the P(r) distribution). C. P(r) distributions. D. Fit of the Fourier transform of the P(r) distributions to the experimental scattering curves; inset is zoom into Guinier region of curves and fits.

While similar trends are observed for the Porod exponents of the HT-SAXS and averaged SEC-SAXS values, the Porod exponents are slightly higher for all SEC-SAXS samples, arguably suggesting less flexibility. However, caution should be taken when attempting to draw conclusions about relative primase flexibility and foldedness from comparing the HT-SAXS and SEC-SAXS Porod exponent values, as a difference of 0.2 is very close to the margin of error. Similar caution should be taken when attempting to analyze the SEC-SAXS Kratky plots; while we can be confident in conclusions drawn from large changes, the noisiness of the data at q > ~0.18 makes it difficult to draw strong conclusions from small perturbations to the plots. The Kratky plots generated with the averaged SEC-SAXS scattering profiles are qualitatively similar

to those generated from the HT-SAXS scattering profiles. Overall, while there may be minor changes to the apparent flexibility or foldedness as reflected in Porod exponents and analysis of the Kratky plots, the observed difference is not significant given the small change and low S/N in the SEC-SAXS data.

To evaluate how the presence of multiple species was affecting the HT-SAXS and SEC-SAXS results, I analyzed monomer scattering profiles that had been extracted from the SEC-SAXS by Dr. Michal Hammel using SVD deconvolution. Deconvolution indicated that the scattering curves for each sample (free primase, primase under initiation conditions, and elongation complex) are in fact composed of the scattering contributions of two states, currently assigned to monomer and dimer. Guinier analysis of the free primase monomer scattering profile did not reveal a significant change in Rg value or Porod exponent compared to the mixed, non-deconvoluted scattering profile (Table 4.2, Table 4.3), suggesting that scattering from the dimeric state did not significantly perturb these metrics in the averaged scattering profile of free primase (Figure 4.15A). However, we do observe a drop in the molecular weight value to within 10% of the correct value (theoretical MW: 109 kDa; experimental MW: 100 kDa).

Sample	I(0)	Rg (Guinier)	Rg (Real space)	Vc	Volume	MW	r	Dmax	Px
Free Primase Monomer	2.3E+03	39.4	40.1	712	330,000	100,000	51.4	129	2.8
Initiation Complex Monomer	1.7E+03	40.0	41.6	741	310,000	110,000	52.7	140	2.8

Table 4.3. **SEC-SAXS parameters for the monomeric state after deconvolution by SVD for free primase and primase in conditions promoting initiation.**

However, comparison of the Kratky plots and P(r) functions derived from the deconvoluted and un-deconvoluted free primase SEC-SAXS scattering spectra (Figure 4.15B/C and Figure 4.16B/C) reveals significant qualitative differences in these transformations. While it is inadvisable to attempt to derive explicit details from the plots depicted in Figure 4.16B and

Figure 4.16C (as S/N loss may introduce artefacts into the P(r) distribution and obscure details of the Kratky transformation), the overall shapes of these plots are significantly different from those derived from the non-deconvoluted and HT-SAXS scattering profiles (Figure 4.3, Figure 4.15). Most strikingly, we observe a redistribution of radii in the P(r) distribution towards a more symmetric curve and loss of the shoulder at ~30 Å which, in combination with a Porod exponent of 2.8, suggest that monomeric free primase is highly dynamic and likely occupying a combination of closed and open states.
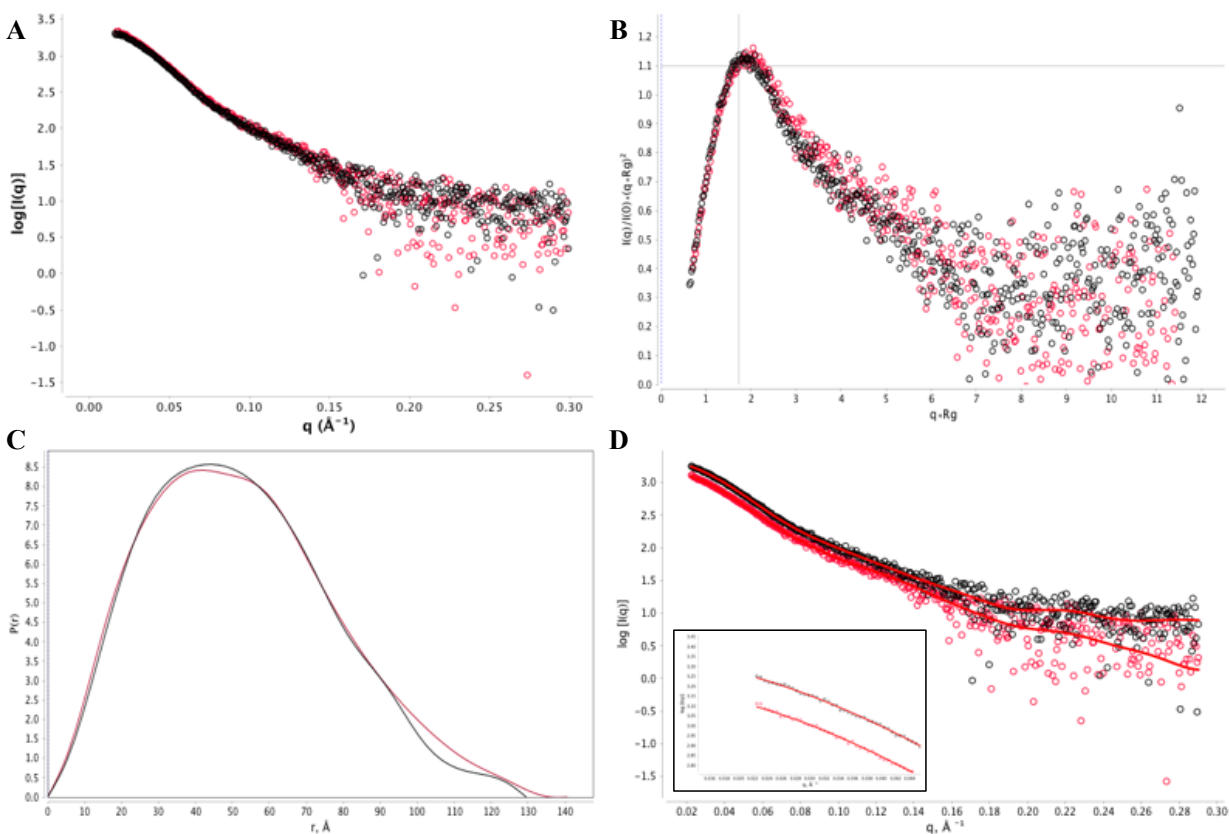
Figure 4.16. Deconvolution of SEC-SAXS data for free primase and primase in conditions promoting initiation. Scattering data after SVD deconvolution for the monomer state for free primase (black) and primase under conditions promoting initiation (red). A. Overlay of scattering intensity profiles. B. Overlay of dimensionless Kratky plots (normalized using the real space Rg calculated from the P(r) distribution). C. P(r) distributions. D. Fit of the Fourier transform of the P(r) distributions to the experimental scattering curves; inset is zoom into Guinier region of curves and fits.

Guinier analysis of the monomer "initiation" scattering profile reveals a significant decrease in Rg value from 43.7 Å to 39.4 Å (Table 4.2 and Table 4.3). Additionally, the molecular weight of the scattering complex drops to within 10% of the theoretical molecular weight for the initiation complex (theoretical MW: 118 kDa; experimental MW: 110 kDa), suggesting that the 'oligomerization' significantly affected the Rg and molecular weight values of the non-deconvoluted scattering profiles for primase in conditions promoting initiation. The Porod exponent of the deconvoluted scattering spectrum also drops from 3.8 to 2.8, the same value calculated from the scattering profile of free primase, suggesting that this scattering particle is no less flexible or dynamic than free primase.

In fact, comparing the monomer scattering profiles of free primase and under conditions promoting initiation (Figure 4.16) reveals that these scattering profiles are essentially the same. The Kratky plots also overlay very well, and the Rg, volume, and molecular weight values are not significantly different. The P(r) distributions of the monomer state for primase and the initiation species possess essentially the same shape; minor variations are likely artefacts caused by overfitting the noise in the deconvoluted scattering profile. These results suggest that the differences in free primase and under conditions promoting initiation in the HT-SAXS and averaged SEC-SAXS data may reflect cofactor/substrate-mediated oligomerization, and that the 'monomer' extracted by SVD deconvolution may correspond to the equilibrium concentration of free primase.

Guinier analysis of the monomer scattering profile for the elongation complex generates a Rg value similar to that observed in the non-deconvoluted scattering spectrum (Table 4.4). Additionally, the Porod exponent and molecular weight values are similar to those determined without deconvolution, and the Kratky and P(r) plots are qualitatively similar. This suggests that

less oligomerization occurred in this sample than in the free primase and initiation samples, resulting in a less skewed data set. In contrast to primase under conditions for initiation, we do observe significant differences when we compare the deconvoluted monomer scattering profiles for free primase and the elongation complex (Figure 4.17). The Rg value for the elongation complex is significantly larger than that calculated for free primase, and the Kratky plots and P(r) distributions exhibit the shoulder formation corresponding to domain separation observed in the HT-SAXS scattering data (Figure 4.10). These data indicate that primase in the presence of primed substrate adopts an 'open.' configuration, consistent with the model of primase action in which p58C remains with the 5' end of the nascent primer (Figure 4.4).
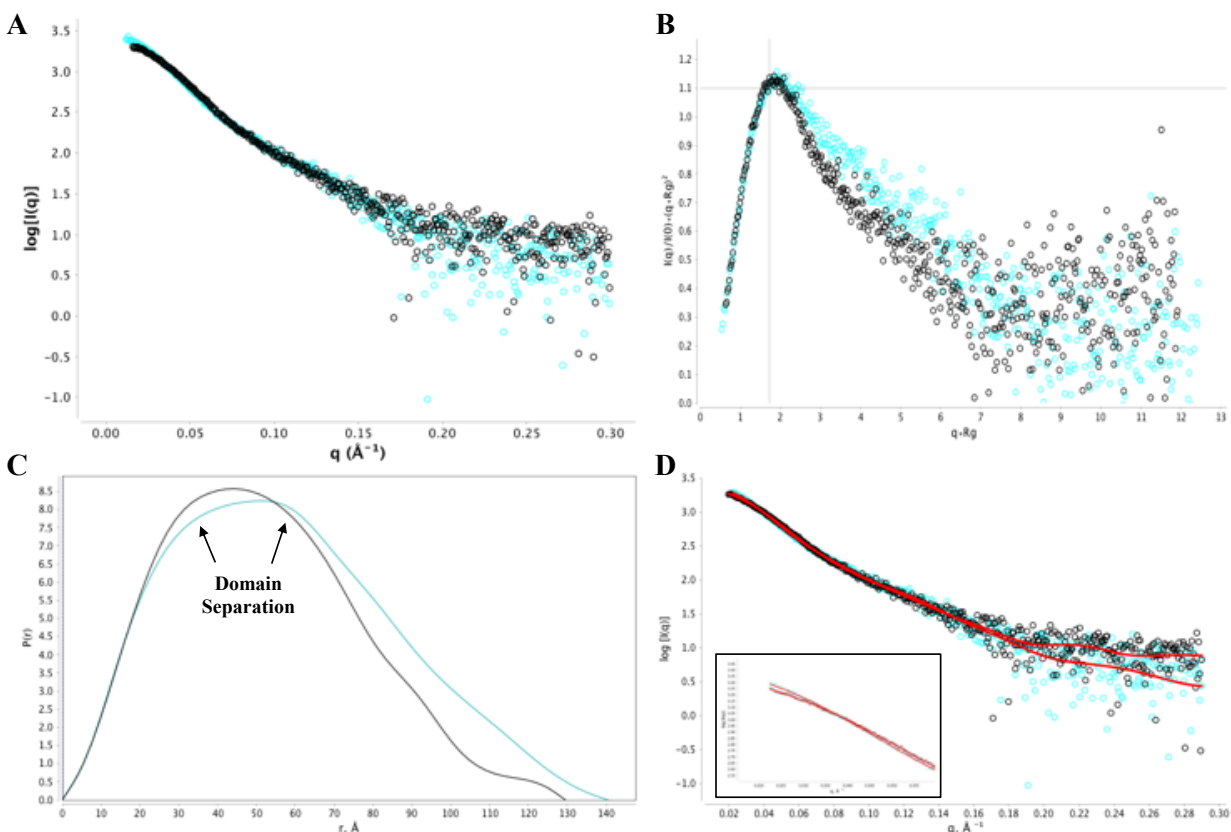
Figure 4.17. Deconvolution of SEC-SAXS data for free primase and the elongation complex. Scattering data after SVD deconvolution for the monomer state for free primase (black) and primase in the presence of primed substrate (cyan). A. Overlay of scattering intensity profiles. B. Overlay of dimensionless Kratky plots (normalized using the real space Rg calculated from the P(r) distribution). C. P(r) distributions. D. Fit of the Fourier transform of the P(r) distributions to the experimental scattering curves; inset is zoom into Guinier region of curves and fits.

Sample	I(0)	Rg (Guinier)	Rg (Real space)	Vc	Volume	MW	r	Dmax	Px
Free Primase Monomer	2.3E+03	39.4	40.1	712	330,000	100,000	51.4	129	2.8
Elongation Complex Monomer	2.8E+03	45.6	44.9	873	320,000	140,000	57.2	140	2.7

Table 4.4. **SEC-SAXS parameters for the monomeric state after deconvolution by SVD for free primase and elongation complex.**

Discussion

In our initial model of primase behavior, we postulated two things: first, that p58C forms

a bipartate initiation active site with p48. Formation of the complete initiation active site requires

NTP and ssDNA template binding to p48 and p58C and positioning of the p58C and template

over the p48 catalytic site. While formation of the active state will occur in equilibrium with free primase and transient complexes containing different combinations of catalytic metals, NTPs and template, introduction of substrates required for initiation should induce a shift in the distribution of primase configurations towards closed configurations. Such a shift should be reflected in a smaller Rg, a larger Porod exponent, and evidence for compaction in the P(r) distribution. Second, we anticipated that formation of the elongation complex would result in extension of p58C away from p48 relative to its position in the initiation complex because p58C should maintain contact with the 5' end of the nascent primer. This extension would be reflected in the SAXS scattering profile as a larger Rg, emergence of a shoulder in the P(r) distribution that would reflect domain separation, and possibly a decrease in the Porod exponent. The goal of this project was to generate SAXS data that would allow us to either confirm or refute this set of hypotheses.

Initial SAXS data collected in high throughput mode after mailing samples (HT-SAXS) were in some ways surprising. Perhaps the most notable finding was that the Rg of primase under conditions promoting initiation was larger than that of free primase (see Table 4.3), suggesting that primase adopts a more extended configuration in this state–a result that directly conflicted with our understanding of primase biochemistry. However, generation of aggregation-free initiation samples was very challenging, and we now know that even samples that were nominally free of aggregation might still possess some degree of oligomerization, which could explain our results.

To address this concern, we collaborated with Dr. Michal Hammel to further evaluate these samples with SEC-SAXS, which allows separation of primase from any contaminating aggregates or particles. The SEC-SAXS results revealed that free primase transiently

oligomerizes (see Figure 4.11), and that addition of initiation cofactors and substrate substantially enhanced this phenomenon (see Figure 4.12). Beyond providing a rationale for the surprising aspects of our HT-SAXS data, these results suggest that the differences in the HT-SAXS data for free primase and the initiation complex are likely caused by changes in primase oligomerization state, and impeded the ability to detect mechanistically relevant changes in the configurational ensemble.

One explanation for the oligomerization of free primase is that the hydrophobic ledge on p58N that would typically bind the primase-binding motif on the p180 subunit of pol α may transiently interact with a hydrophobic surface on another primase molecule (144). Additionally, one could consider the C-terminal tail of p58, which contains a helical, negatively-charged region spanning residues Leu484-Tyr504 that could conceivably transiently bind the template binding region of an adjacent primase molecule. One way to differentiate between these two hypotheses would be to collect SEC-SAXS data on primase without this C-terminal tail and observe whether or not this transient dimerization phenomenon is abolished.

While transient oligomerization was clearly evident in the free primase and multiple species were present under conditions promoting initiation, it did not appear to be present to as significant an extent in the elongation sample (see Figure 4.11 and Figure 4.12). One possible explanation for this observation is that the primed substrate organizes the primase domains, precluding the transient interactions that enable primase oligomerization. Comparison of the data extracted by SVD for the monomeric state of free primase and the elongation complex reveals perturbations to the P(r) distribution that suggest extension of the primase molecule when in complex with the primed substrate (Figure 4.17). These data support a model in which the primed substrate stabilizes primase in an open configuration, with p48 and p58C, respectively,

occupying the 3' and 5' ends of the nascent primer (Figure 4.18).
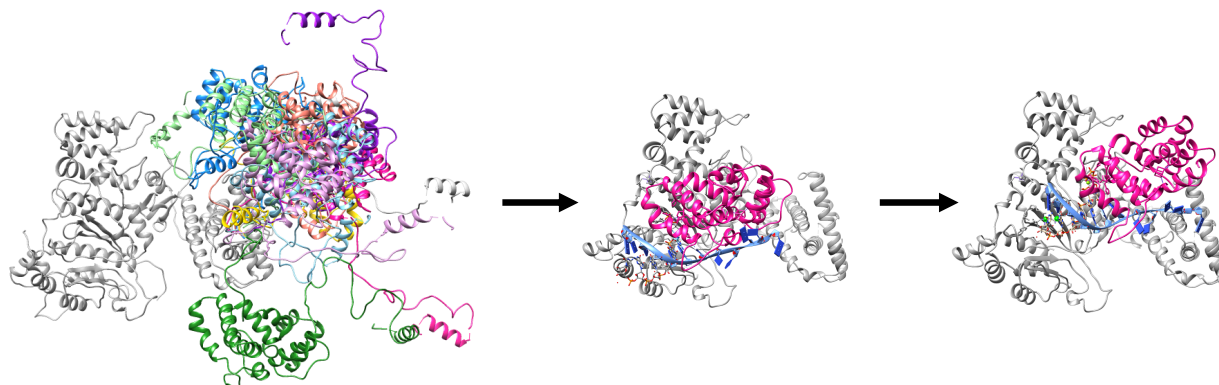
Figure 4.18. Model of primase action. p48/p58N structural platform is depicted in grey in all panels. p58C is shown in pink for initiation/elongation models and in many colors for the free primase models. The DNA template is shown in blue, and RNA primer is in orange/red sticks. To generate the models of free primase shown here, FoXS was run on the pool of 10,000 primase configurations to identify the ten best fits to the monomer scattering profile for free primase, which fit with X^2 values ranging from 1.5 to 2.4. These configurations were aligned in Chimera to generate the left panel in the figure above. Models of primase in complex with primed substrate were generated by taking the p48 structure from PDB 4LIL, the p58N structure from PDB 4RR2, and the p58C structure from PDB 5F0S and aligning the 5' and 3' ends of the RNA primer in PDB 5F0S with the triphosphate in the p48 active site to generate the initiation and elongation complex models, respectively.

Conclusions

The data reported in this chapter show that free primase transiently oligomerizes and that this phenomenon increases in the presence of initiation substrates. Extraction and analysis of the scattering of the monomeric state for the free primase and the smallest species under conditions promoting initiation revealed no significant differences between these scattering profiles. Analysis of the monomer scattering profile of the elongation complex indicates domain separation relative to free primase, suggesting that addition of primed substrate stabilizes an elongated configuration of primase. This is the first direct experimental evidence on full-length primase in complex with substrate to support the model pictured in Figure 4.17.

Because the transient oligomerization of free primase may be driven by interaction between protein binding surfaces, conducting future structural studies of full-length primase in

complex with protein interaction partners, such as RPA or pol α, may reduce this effect. Additionally, different initiation substrates that do not promote primase oligomerization may be required to elucidate the structural basis for the initiation of primer synthesis. Because the initiation substrates utilized in the set of experiments described in this chapter were specifically chosen to prevent dinucleotide polymerization, it is possible that they ultimately prevented reconfiguration of primase to form the active initiation complex. In addition to testing other nonhydrolyzable substrates, future studies may find success in pairing standard ribonucleotides with a ssDNA template engineered to prevent polymerization beyond the initial dinucleotide. Crosslinking could be used to enrich the population of primase in the active initiation configuration by trapping primase in the process of dinucleotide formation. Further parsing of the structural basis for primer elongation and termination may require utilization of other methods, such as cryo-electron microscopy or electron paramagnetic resonance, that enable structural characterization at higher resolution without requiring crystallization of the protein-substrate complex.

CHAPTER V

DISCUSSION AND FUTURE DIRECTIONS

Summary of This Work

DNA primases are required for genome replication across all kingdoms of life (88, 131). While the essential process of priming in eukaryotes is well-defined, the mechanisms driving initiation of primer synthesis, primer elongation and hand-off remain somewhat obscure. The research described in this dissertation focused on investigating the mechanisms driving primer synthesis and regulation of primase activity. Key themes from this work are summarized below.

Charge transport may play a role in polymerase dynamics at the replication fork

In 2007, our and the Pellegrini laboratories discovered that the p58C domain of human DNA primase possesses a [4Fe4S] cluster that is required for priming activity (116, 118). Biochemical evidence and the energetic cost of Fe-S cluster assembly and insertion suggest this cluster plays a functional, not structural, role in primer synthesis (116, 231). For many years, the laboratory of Jacqueline K. Barton at Cal Tech had been investigating the potential for [4Fe4S] clusters in proteins to drive charge transport through base-paired DNA (209). In 2012, the Chazin and Barton laboratories decided to join forces to determine if the [4Fe4S] cluster in human DNA primase plays a role in priming. Together, we found that p58C can act as a redox switch capable of driving charge transport through the protein to bound DNA and through the DNA wire (Chapter II). Importantly, oxidized p58C was found to engage in redox switching through charge transfer, while reduced p58C displayed no redox signal. This phenomenon was

reversible, i.e. oxidation of reduced p58C resulted in regeneration of the electrochemical signal. These results led to the conclusion that the [4Fe4S] cluster of human DNA primase functions as a redox switch that controls the affinity of the interaction with the DNA substrate (235).

For this phenomenon to occur, electrons must travel ~25 Å from the DNA substrate through the protein to the [4Fe4S] cluster. We identified three conserved tyrosine residues (Y309, Y345, and Y347) that form a possible conduit for charge transfer through the protein and generated phenylalanine mutants of these residues that should significantly diminish charge transfer. These mutants were structurally characterized using both x-ray crystallography and solution-based techniques, and DNA binding ability was confirmed with fluorescence anisotropy-based DNA binding assays. Electrochemical analysis revealed that these single-atom alterations inhibited charge transfer through the p58C protein matrix. Moreover, disruption of charge transfer reduced initiation of primer synthesis and inhibited primer counting by full-length primase *in vitro*.

The publication of our *Science* paper prompted the groups of Pellegrini and Tahirov to submit Technical Comments on our work (233, 234), which in turn led me to pursue further characterization of the crystal structures and solution-state properties of the two different p58C constructs that had been used in previous studies, as reported in Chapter III. In short, both p58C constructs had been crystallized, and the structures exhibited differences in secondary structure for 34 residues in the template binding region (36, 215). Among other things, the technical comments argued that that the results reported in the *Science* paper should be questioned because our p58C construct was mis-folded due to truncation of Ile271 (233, 234). To address this issue, I generated and crystallized a new $p58C_{266-464}$ construct that contained all residues present in both crystallized p58C constructs. When $p58C_{266-464}$ was crystallized in the same conditions used

to crystallize our previous p58C$_{272-464}$ construct, the structures proved to be remarkably similar to each other, suggesting that the differences in secondary structure observed between the previously reported p58C$_{272-464}$ and p58C$_{266-456}$ crystal structures were caused by differences in crystallization conditions, rather than differences in the intrinsic folding associated with absence or presence of Ile271.

Because p58C secondary structure appeared to be sensitive to crystallization conditions, I used CD and NMR spectroscopy to assess whether the structures of the p58C$_{266-464}$ and p58C$_{272-464}$ constructs are similar in solution. These studies revealed no significant perturbations to protein structure in solution. The p58C$_{266-464}$ and p58C$_{272-464}$ constructs also exhibited similar DNA binding affinities and charge transfer abilities, indicating that deletion of residues 266-271 does not significantly impair p58C function. These data show that the observations described in Chapter II are *not* due to mis-folding of p58C.

Investigating the role of domain rearrangements in active site remodeling

Current structural and biochemical evidence suggest that primase utilizes a biphasic mechanism, in which domain rearrangement drives formation of the initiation active site and allows elongation of the nascent primer (37, 132, 145). However, this model has not been validated with direct experimental data collected on full-length primase in complex with catalytically relevant substrates. To address this gap, I attempted to use SAXS to characterize primase domain rearrangements as primase encounters initiation and elongation substrates. Analysis of SEC-SAXS data collected on free primase, under conditions promoting initiation, and in complex with elongation substrates revealed primase exhibits complex equilibria under all conditions. For free primase, it is highly likely that the higher order states arise from transient

oligomerization. Under the conditions promoting initiation, there is an increased level of complexity, which may arise from a similar transient oligomerization phenomenon. In contrast, there is less complexity in the elongation complex. SVD deconvolution of each SEC-SAXS dataset resulted in simultaneous fitting of a mixture of two states, which have tentatively been assigned to monomers and dimers. Our interest is focused on the monomeric states, so we first compared the extracted scattering data for free primase and primase under conditions promoting initiation. Surprisingly, there were no significant differences in any of the scattering parameters. In contrast, we compared the free protein to the elongation complex and found the elongation complex to be more extended than the free protein, which supports a model in which the primed substrate stabilizes primase in an open configuration, with p48 and p58C, respectively, occupying the 3' and 5' ends of the nascent primer (Figure 5.1).

Free primase is dynamic

Substrate binding prompts formation of initiation active site

Primer elongation causes domain separation
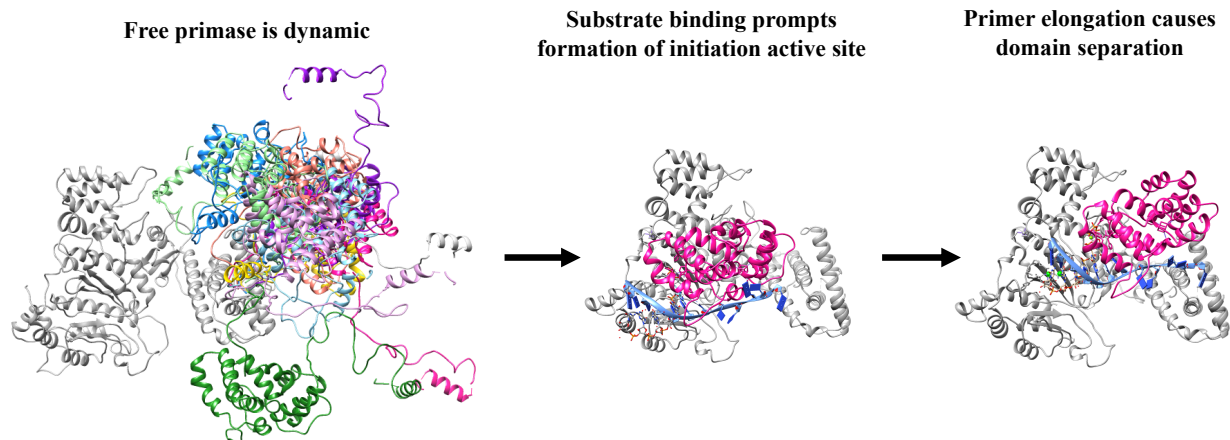
Figure 5.1. Model of primase action from SAXS and crystallographic evidence. The p48/p58N structural platform is depicted in grey in all panels. p58C is shown in pink for initiation/elongation models and in many colors for the free primase models. The DNA template is shown in blue, and RNA primer (oriented away from the reader) is in orange/red sticks. To generate the models of free primase shown here, FoXS was run on the pool of 10,000 full-length primase configurations generated by Dr. Matthew Thompson using RANCH to identify the ten best fits to the monomer scattering profile for free primase, which fit with X^2 values ranging from 1.5 to 2.4. These configurations were aligned in Chimera to generate the left panel in the figure above. Models of primase in complex with primed substrate were generated by taking the p48 structure from PDB 4LIL, the p58N structure from PDB 4RR2, and the p58C structure from PDB 5F0S and aligning the 5' and 3' ends of the RNA primer in PDB 5F0S with the triphosphate in the p48 active site to generate the initiation and elongation complex models, respectively.

Implications and Future Directions

As discussed above, the research described in this dissertation revealed that primase exhibits complex configurational equilibria in solution and that [4Fe-4S] cluster redox appears to play a role in regulating primase activity. This section discusses potential implications of these results and proposes experiments to explore the mechanism of primer synthesis and polymerase dynamics at the replication fork.

Re-envisioning the role of p58C in primer synthesis

The understanding of the role of p58C in primer synthesis was advanced in 2016, when the Tahirov group published a series of papers establishing that p58C tightly binds the 5'

triphosphate of the RNA primer and revealing the structural basis for the interaction between p58C and primed template (115, 137). Over two decades earlier, Kuchta and Copeland had discovered that primase operates via a biphasic mechanism (37, 132), and Tahirov's results helped place that discovery into a structural context (145). According to this model of eukaryotic primer synthesis, initiation requires formation of a bi-partate active site involving both p48 and p58C, in which p58C and the bound template are positioned over the p48 active site with the 3' NTP and catalytic metals bound. Dinucleotide synthesis starts when the 5' NTP is added into this constellation of primase domains, template and cofactors. In our model, both p48 and p58C contribute to the binding of the 5' NTP. This view provides an explanation for why dinucleotide synthesis is the rate limiting step in priming: the modest binding affinities for template and cofactors combined with the inability to bind the 5' NTP unless the protein is in a compact configuration with p58C properly positioned over the p48 active site results in infrequent assembly of the correct set of initiation factors. Once the initial dinucleotide is generated, only a new 3' NTP is needed to elongate, and there is no need to have p58C aligned specifically with respect to p48, so the nascent primer can be elongated much more rapidly than initial dinucleotide is formed.
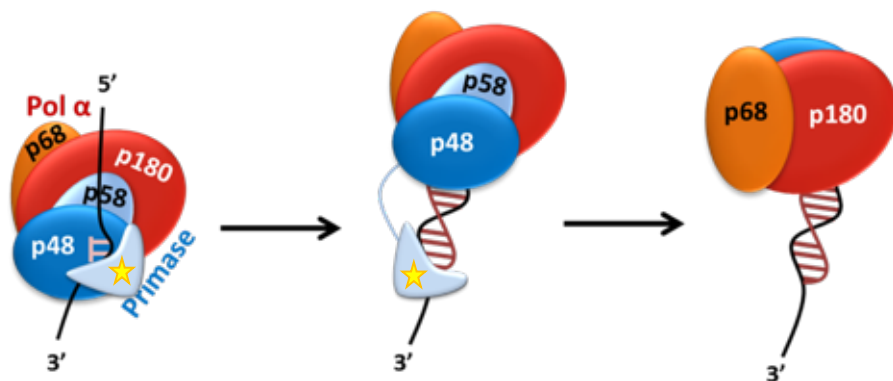
Figure 5.2. Mechanism of primer synthesis. After assembly of the primosome on the ssDNA template, p48 and p58C collaborate to synthesize the initial dinucleotide in the rate-determining step of primer synthesis. p48 rapidly elongates the nascent primer to a length of 7-10 nucleotides before the catalytic subunit of pol α (p180) displaces p48 at the 3' end of the nascent primer and extends to a final length of 20-30 nt. The star denotes the Fe-S cluster in p58C.

While this model fits the available structural and biochemical evidence, it has yet to be confirmed with structural studies of full-length primase in complex with catalytically relevant substrates. This may be due to both the dynamic nature of primase and the modest affinity with which it binds ssDNA template, catalytic metals, and nucleotides (36, 112, 115, 143). An additional important consideration was discovered during the course of this dissertation research, i.e. the propensity for transient oligomerization of isolated primase described in Chapter IV (Figure 5.3). I hypothesize that this aggregation phenomenon in primase is a result of separating primase from pol α, its binding partner *in vivo*. This separation leaves hydrophobic primase protein binding surfaces, such as the p180-binding region in p58N, exposed and therefore available to interact with each other (36, 144).
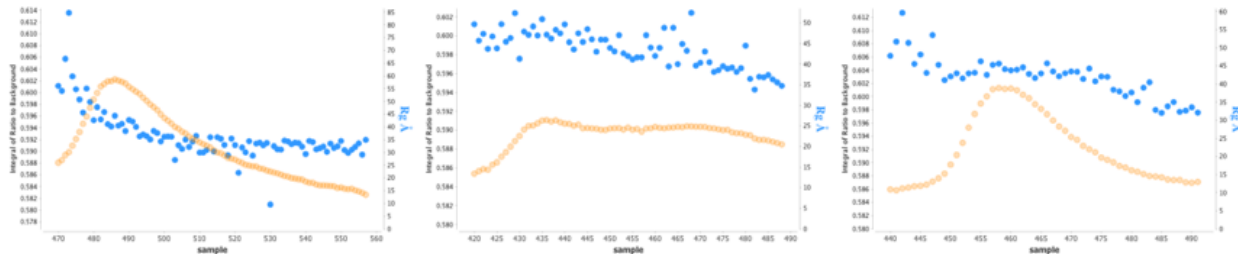
Figure 5.3. Primase transiently oligomerizes. SEC trace and corresponding SAXS Rg values for free primase (left), under conditions promoting initiation (middle), and in the elongation complex (right).

Analysis of primase constructs in which possible interacting regions have been mutated or deleted may reveal the physical basis for this phenomenon, resulting in a less aggregation-prone construct. However, since reengineering the surface is a major undertaking, it would be logical to first try a strategy to reduce transient oligomerization through modification of the buffer. This would involve screening buffer identities and considering additives that inhibit oligomerization. However, the latter introduces a risk of abolishing interactions with cofactors and substrates. Perhaps the most philosophically simple method for eliminating this phenomenon is to perform the structural characterization of these steps of priming on complexes with primase binding partners, such as pol α. However, generation of the human pol-prim complex in the quantities currently required for structural analysis is technically challenging, as the human p180 subunit is toxic to *E. coli*, and attempts to recombinantly express this protein in both *E. coli* and baculovirus-based systems have resulted in very low yields.

In contrast to free primase and primase in the presence of initiation substrates and cofactors, the complexity of the equilibria appears to be reduced in the presence of triphosphorylated primed substrate (Figure 5.3). I hypothesize that oligomerization was not observed to as significant a degree in the elongation complex because binding primed substrate organizes the p48/p58N and p58C domains, reducing their ability to transiently interact with

other primase molecules. This effect does not occur under conditions promoting initiation, possibly due to the equilibrium concentration of free primase and the low affinity with which primase binds the substrates and cofactors. The length of the ssDNA template may also play a role in the increased complexity of equilibria (possibly dimerization) we observe under conditions promoting initiation, as it is possible that two primase molecules may bind the same substrate. To assess whether the latter is the case, one could collect SEC-SAXS data on a series of primase samples in the presence of NTPs, catalytic metals, and ssDNA of varying lengths to determine whether dimer formation decreases with the length of ssDNA template. Alternatively, forcing complex formation by crosslinking primase with initiation substrates may preclude substrate-mediated dimerization.

SVD deconvolution allowed us to generate a scattering profile of the monomer peak for all three primase conditions, which revealed the complex with primed substrate is extended (Figure 5.4). While we know that scattering data in solution is caused by contributions from an ensemble of primase configurations, one can use FoXS–which calculates the theoretical scattering profile from atomic coordinates–to generate a "back of the envelope" sense for what this extension of p58C might look like. Using FoXS to determine the best fits to this scattering data from a pool of 10,000 primase configurations, we find that the average of the p58C configurational ensemble is more compact for the free primase than for the elongation complex. Figure 5.5 provides a visual guide by showing single best-fit solutions for free primase and the elongation complex.
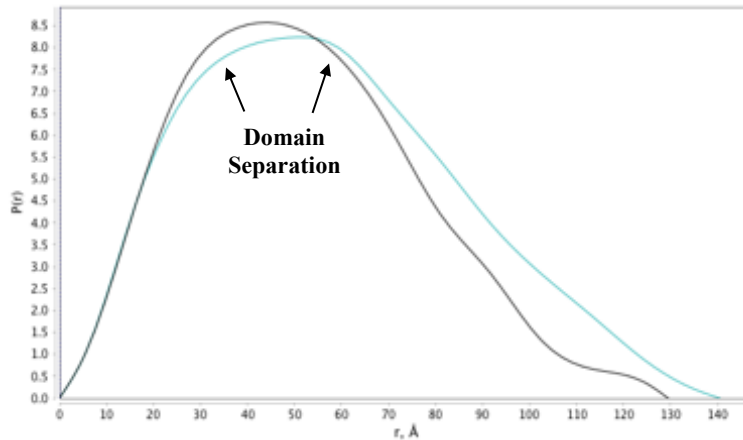
Figure 5.4. The elongation complex exhibits domain separation. P(r) distributions of free primase (black) and the elongation complex (cyan).
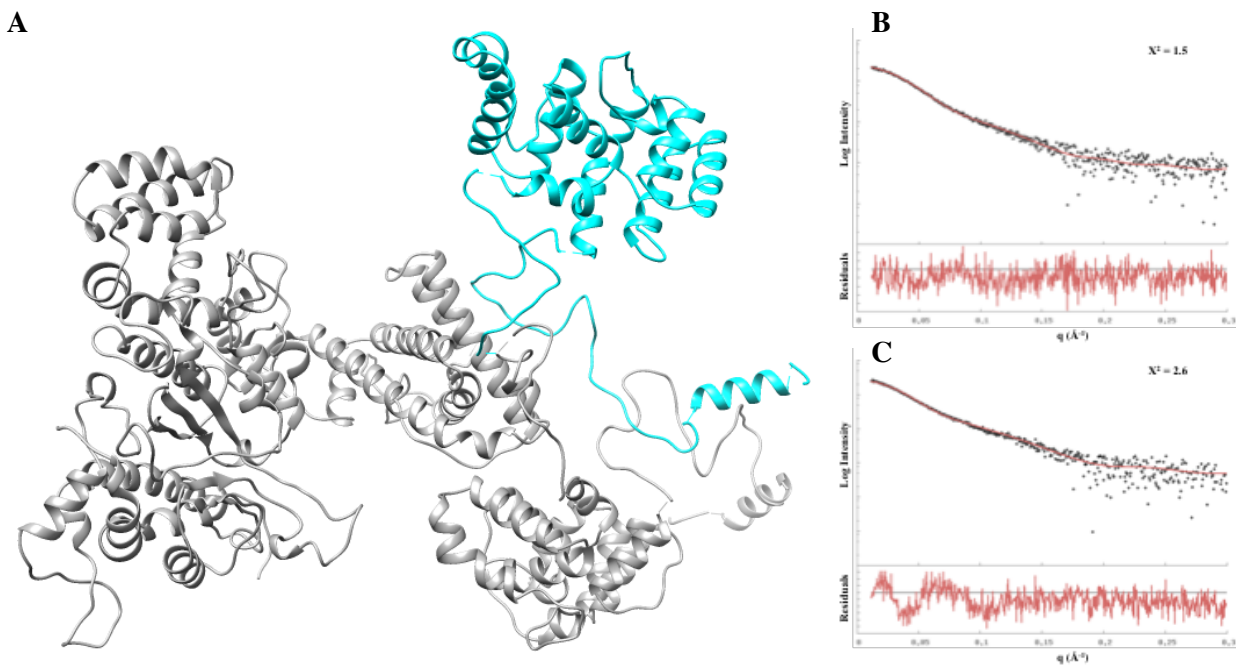
Figure 5.5. The elongation complex exhibits domain separation. A pool of 10,000 primase configurations were fit with FOXS against the deconvoluted monomer scattering profiles of free primase and primase in the presence of primed substrate. A. Best-fitting configurations for free primase (grey) and primase in the presence of primed substrate (cyan) with X^2 values of 1.5 and 2.6, respectively. B. Calculated scattering profile and residuals for best fit against free primase (red) plotted against the experimental scattering profile (black). C. Calculated scattering profile and residuals for best fit against primase in the presence of primed substrate (red) plotted against the experimental scattering profile (black).

While the data presented in Chapter IV are far from a complete structural characterization

of the elongation complex, *this is the first direct experimental evidence in the context of full-length primase that p58C travels with the 5' end of the nascent primer*. Collection of more SAXS data to support these results in combination with those from other techniques, such as EPR and electron microscopy, should enable generation of a more detailed experimentally driven structural model of the elongation complex. Ultimately, confirmation of our model of primase action may require the use of electron microscopy, in which aggregated and dimerized particles can be removed from data analysis.

In addition to the structural approaches proposed above, a combination of mutational analysis and biochemical assays may shed light on the configuration of the initiation complex. To initiate primer synthesis, primase requires catalytic metal ions, two NTPs, and DNA template (37, 132). p58C is known to bind template with modest affinity, and p48 has been crystallized with the catalytic metals and the 3'-NTP (36, 143, 144). Based on biochemical assays with p48 variants, it was proposed that the R304 residue in p48 recruits the 5'-NTP to the catalytic site (136). However, while crystal structures of p48 with catalytic metals include the triphosphate of the 3'-NTP and p58C has been crystallized in complex with primed template, the 5'-NTP has been remarkably absent in structural analyses of primase (137, 143, 144). Initially, we hypothesized that a conformational change triggered by template binding may be required for 5'-NTP binding to p48 (143). However, recent structural and biochemical studies of p58C with primed substrate suggested that p58C may interact with the 5'-NTP. This hypothesis was not confirmed by any experimental evidence such as site-directed mutagenesis, and implications for initiation of primase activity were not empirically explored beyond speculation that previous biochemical evidence indicating that p48 R304 recruits the 5'-NTP were artifacts caused by p48 misfolding (145).

I hypothesize that p58C is involved in stabilizing the 5'-NTP in the active site for initial

dinucleotide synthesis through a series of weak interactions that are only completed by additional

interactions in p48, including to R304. To test this hypothesis, I have mutated residues that

appear to interact with the 5'-NTP in the crystal structure of p58C bound to primed template

(H300, R302, H303, R306) and showed that these p58C variants are structurally intact in

solution (Figure 5.6). In addition, I used ^{31}P NMR to monitor titrations of p58C into ATP and

found that mutation to either alanine or a charge reversal to aspartic acid or glutamic acid

resulted in reduced chemical shift perturbations relative to wild-type p58C (Graph 5.1),

suggesting a deficiency in NTP binding. CD of R304A p48, which demonstrated the greatest

deficiency in 5'-NTP recruitment experiments, revealed no perturbation to the secondary

structure of the mutant compared to wild-type p48 (Figure 5.7), indicating that the previously

reported results were not caused by mutation-induced misfolding. For more details of these

experiments, see Appendix I.
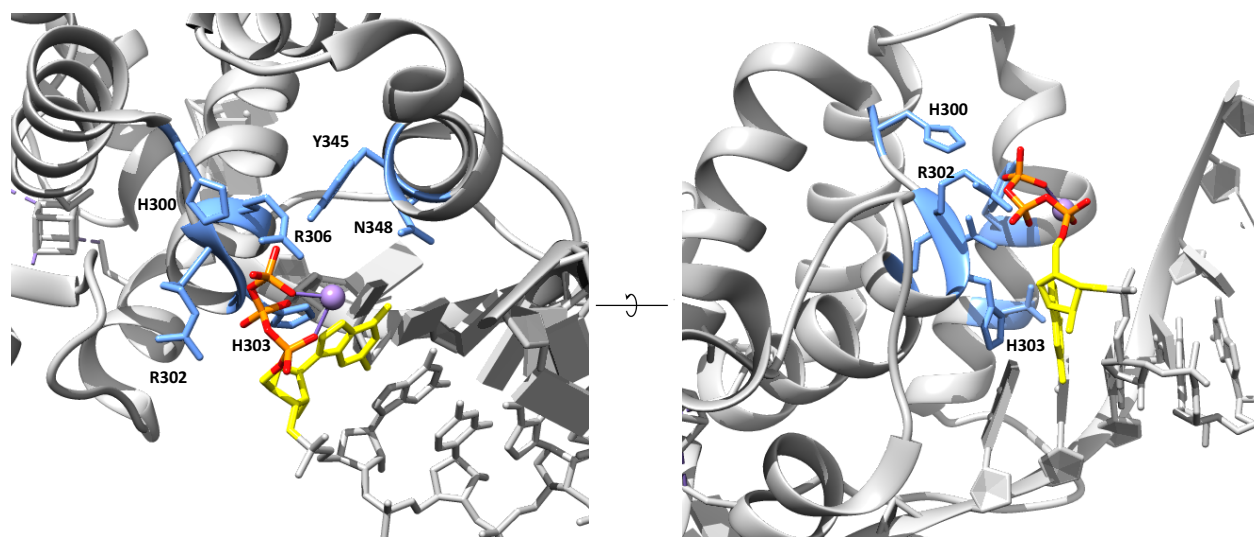
PDB ID: 5F0Q

Figure 5.6. p58C residues proposed to recruit 5'-NTP. p58C is depicted in grey, with mutated residues highlighted in blue, manganese in green, and the 5'-GTP in yellow, orange, and red.

% Chemical Shift Perturbation Relative to Free ATP

**Graph 5.1. Alanine and charge reversal mutat...
perturbations.** ^{31}P NMR chemical shift perturba...
and ATP in complex with p58C variants. Perturb...
perturbation observed with wild-type p58C.
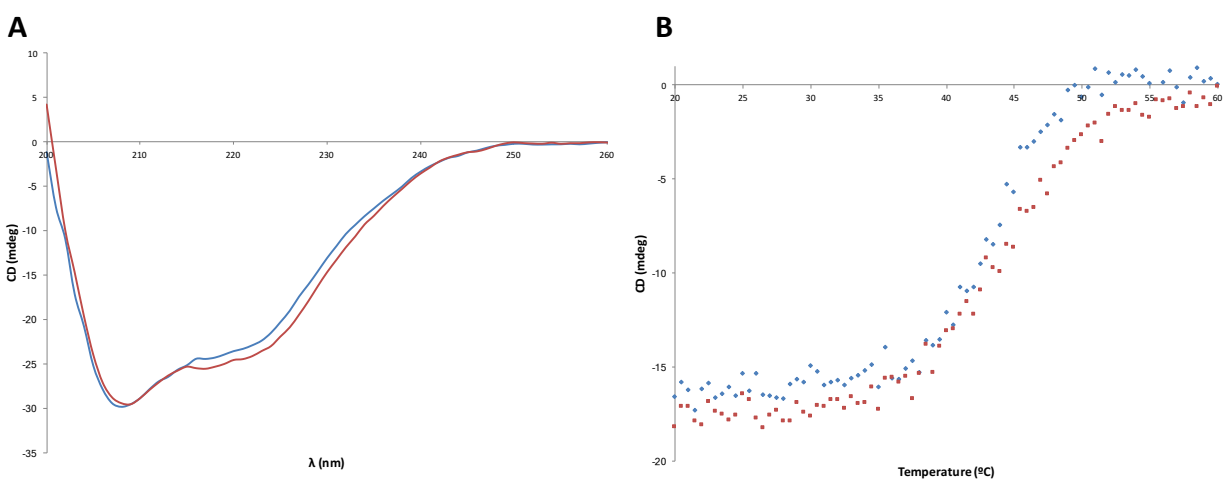
...**WT p48.** A. CD scans of WT (blue)
...ue) and R304A (red) p48, monitored at

...it would be useful to evaluate the

...TP and the ability of full-length

...substrate. I expect that these mutations

in p58C will result in significantly reduced ATP binding affinity and initiation activity, with less effect on primer extension activity. Also, as these residues appear to interact with the triphosphate of the 5'-NTP, I anticipate that wild-type primase would exhibit enhanced primer extension activity on a triphosphorylated primed substrate relative to a monophosphorylated primed substrate, and that one would not observe this difference in elongation activity with mutants that are deficient in binding ATP.

Determining the role of redox switching at the replication fork

The electrochemistry and priming assays presented in Chapter II of this dissertation suggest that initiation of primer synthesis and product truncation are dependent on a redox switch changing the oxidation state of the [4Fe4S] cluster in p58C (235). Based on this and the available structural and biochemical data, we proposed a model for priming that incorporates this redox switch. In this model, oxidation of the [4Fe4S] cluster activates primase to tightly bind DNA, which is critical to initiating primer synthesis. In priming assays such as those described in Chapter II, charge transfer between primase molecules through the nascent RNA/DNA primer can then mediate primer truncation.

Translating our data to the context of the pol-prim heterotetramer, we proposed that the putative Fe-S cluster in the catalytic subunit of DNA polymerase α may serve as a redox donor/acceptor for priming by driving charge transport through the protein and the duplex portion of the primer. Synthesis of a primer-length (>7 nt) primer and translocation of p58C exposes the duplex portion of the primer for binding by pol α. When binding occurs, the p180 subunit of pol α displaces p48 at the 3' end of the nascent primer, and charge transport through the protein and the primer promotes p58C dissociation from the 5' end. DNA-mediated

signaling, which modulates p58C DNA binding affinity through a change in the redox state of the [4Fe4S] cluster, would thus provide a key driver for efficient primer handoff.
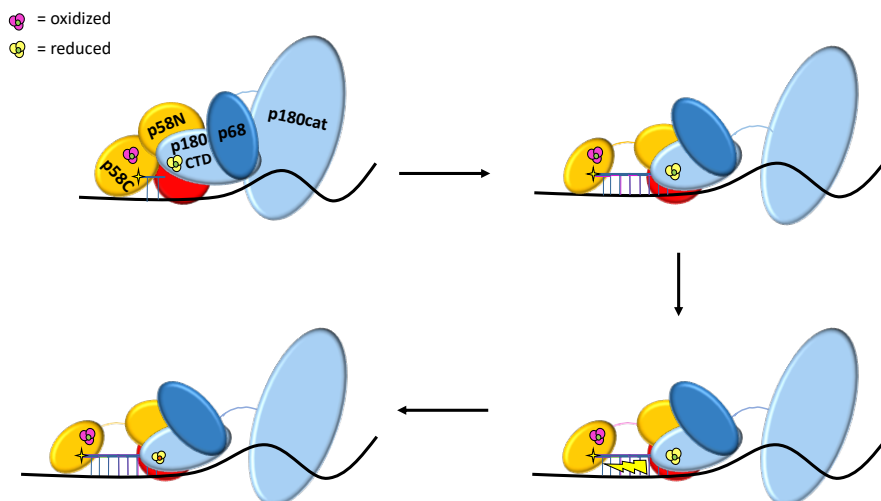
Figure 5.8. Model for redox-mediated substrate handoff between primase and pol α.

This idea of [4Fe-4S] cluster redox-driven substrate handoff is particularly interesting, but raises the question "What is on the other end?", as the change in charge of the p58C cluster requires a redox couple on the other end of the DNA wire. In this vein, in the broader context of the replication fork [4Fe4S] cluster cofactors have been reported in each of the B-family replicative polymerase polymerases α, δ, and ε. Hence, one could speculate that redox-modulated DNA binding by primase is part of a larger, DNA-mediated electron transfer cascade that coordinates polymerase activity within the dynamic replication fork. However, while this is an interesting hypothesis, the idea of primase participating in redox-modulated substrate handoff at the replication fork depends on the existence of a redox-matched Fe-S cluster in pol α.

Like primase, pol α is a heterodimer composed of catalytic (p180) and regulatory (p68) subunits. The C-terminal domain of p180 (p180-CTD) contains eight conserved cysteine residues that form two metal binding sites: a zinc-ribbon motif termed CysA and CysB, a set of four cysteine residues in an atypical pattern that form the proposed Fe-S cluster site (162). Pulldown

assays with p180 from ^{55}Fe-labeled yeast cells indicated that p180 contains an Fe-S cluster *in vivo*, and the Burgers lab was able to generate Fe-S cluster-loaded p180-CTD (162). However, the cluster loading of this sample was poor (~0.1 Fe per subunit), and zinc is bound to the CysB site in all published crystal structures of p180-CTD (137, 145, 262, 263). This has led to a controversy over whether p180-CTD does, in fact, contain an Fe-S cluster (145, 162).

Therefore, the first step towards assessing whether redox switching plays a role in primer handoff in the context of the replication fork is to determine whether pol α contains an Fe-S cluster. Published attempts to generate cluster-loaded pol α have focused on expressing the recombinant protein in *E. coli* and insect cell systems (137, 162); however, it is possible that this cluster is particularly labile and requires a careful expression scheme to remain intact. Additionally, the purification scheme may need to be optimized with the goal of cluster retention. Specifically, Fe-S cluster loading is known to be sensitive to exposure to atmospheric oxygen in some proteins, and generation of the holoenzyme may require anaerobic purification (223, 264, 265). Structural and biochemical studies of pol α have relied on aerobically purified constructs, and it is possible that this has precluded analysis of the enzyme with a potentially native Fe-S cluster. To circumvent the technical issues surrounding recombinant pol α expression, one could mutate the cysteine residues that would coordinate the proposed Fe-S cluster in yeast p180 and look for phenotypic changes (either cell toxicity or thermosensitivity). While this approach would provide limited information about the role of the Fe-S cluster in pol α activity, phenotypic changes would confirm that the cluster both exists and is necessary for pol α function.

If a pol α expression and purification scheme were developed that resulted in Fe-S cluster-loaded protein, this would open the door to a host of assays to characterize the

133

biochemistry and redox properties of the cluster loaded protein. It would be fascinating to observe whether use of intact (cluster-loaded) pol α resulted in a change in product lengths for either the primase or pol α components of the pol-prim heterotetramer, as this would support the hypothesis that redox switching regulates substrate transfer from primase to pol α. To assess the effect of Fe-S cluster loading on RNA primer termination and handoff, one could run polymerase activity assays with the pol-prim complex in the presence of rNTPs and ssDNA and evaluate changes to primase product distribution. Either adding dNTPs to this reaction or simply evaluating pol-prim activity in the presence of template bound to an exogenous primer and dNTPs would provide clues regarding the role of the pol α Fe-S cluster in primer extension. Comparing the results of these experiments with those performed with a pol α mutant in which the cysteines proposed to coordinate the Fe-S cluster were mutated to alanine would provide a clear picture of the role of an Fe-S cluster in pol α activity. Additionally, it would be interesting to assess handoff of the RNA-DNA primer from cluster-loaded pol α to pols δ and ε to determine whether the cluster has a role in polymerase switching.

Beyond the replication fork, charge transport has been also proposed as a mechanism for long-range signaling between glycosylases; however, interesting observations *in vitro* have yet to be confirmed *in vivo* in a more than a correlative manner (209). This raises the question: even if we show that pol α contains an Fe-S cluster and that redox switching plays a role in regulating substrate handoff *in vitro*, how do we show that this phenomenon occurs in the context of the cell?

In our model of charge transfer at the replication fork, redox switches modulate polymerase release from the nascent strand. Deactivating the primase redox switch by disrupting the charge transfer pathway in p58C leads to dysregulation of primer length. In the context of the

cell, this should cause increased incorporation of ribonucleotides into the genome, resulting in replication stress and a mutator phenotype (266). To determine whether this occurs in human cells, one could theoretically introduce the Y345F and Y309F mutations into p58 and use whole-genome sequencing to determine the frequency of single nucleotide polymorphisms between mutated and control primase (267, 268). In practice, however, high error rates resulting in a low S/N can complicate analysis when using this approach (269, 270). Alternatively, one could use a targeted approach–such as ribose-seq or embedded ribonucleotide sequencing (emRiboSeq)–to specifically map ribonucleotide lesions in the genome of cells with charge transfer-deficient primase (271-274). The ribose-seq approach capitalizes on RNA's sensitivity to alkaline conditions to generate nicks at ribonucleotide incorporation sites; the 2',3'-phosphate moiety at the 3' end of the nick is then targeted by an *Arabadopsis thaliana* tRNA ligase to generate a ssDNA circle, which protects from DNA degradation prior to high-throughput sequencing (271). Conversely, with emRiboSeq, RNase H2 generates nicks at sites of ribonucleotide incorporation, after which the 3' end of the nick is captured by a sequencing adaptor prior to sequencing (274). With either technique, I predict that the rate of ribonucleotide incorporation would be higher in cells with charge transfer-deficient primase than those with wild-type primase.

Additionally, dysregulation of polymerase switching could conceivably lead to replication fork stalling or collapse, which can be directly assessed through DNA fiber analysis (275-277). In this experiment, nascent DNA is labeled with halogenated nucleotides, and the labeled DNA fibers are then stretched on a slide and immunostained to visualize replication fork progression. Consecutive labeling with differently modified nucleotide analogues allows identification of fork stalling or collapse through analysis of labeling patterns. Observation of either increased ribonucleotide incorporation or replication fork stalling would open the door to

further experiments to elucidate the role of redox switching at the replication fork.

Concluding Remarks

The work presented here has provided fundamentally new insights into the role of p58C in primer synthesis. Specifically, my work with p58C contributed significantly to the discovery that the [4Fe4S] cluster acts as a redox switch that modulates priming activity (235). This has led to a new model for substrate handoff at the replication fork based on redox switching by the [4Fe4S] clusters present in the C-terminal domains of primase and pol α. Additionally, SEC-SAXS studies of primase in complex with primed substrate have provided our first experimental view of the elongation complex, confirming that p58C travels with the 5' end of the nascent primer. Continuing this work will further our understanding of redox switching at the replication fork and the role inter-domain dynamics plays in driving primase function.

REFERENCES

1.	Negrini S, Gorgoulis VG, Halazonetis TD. Genomic instability--an evolving hallmark of cancer. Nat Rev Mol Cell Biol. 2010;11(3):220-8.
2.	Zhang BN, Bueno Venegas A, Hickson ID, Chu WK. DNA replication stress and its impact on chromosome segregation and tumorigenesis. Semin Cancer Biol. 2018.
3.	Kotsantis P, Petermann E, Boulton SJ. Mechanisms of Oncogene-Induced Replication Stress: Jigsaw Falling into Place. Cancer Discov. 2018;8(5):537-55.
4.	Gaillard H, Garcia-Muse T, Aguilera A. Replication stress and cancer. Nat Rev Cancer. 2015;15(5):276-89.
5.	Branzei D, Foiani M. Maintaining genome stability at the replication fork. Nat Rev Mol Cell Biol. 2010;11(3):208-19.
6.	O'Donnell M, Langston L, Stillman B. Principles and concepts of DNA replication in bacteria, archaea, and eukarya. Cold Spring Harb Perspect Biol. 2013;5(7).
7.	Fu H, Baris A, Aladjem MI. Replication timing and nuclear structure. Curr Opin Cell Biol. 2018;52:43-50.
8.	Myllykallio H, Lopez P, Lopez-Garcia P, Heilig R, Saurin W, Zivanovic Y, et al. Bacterial mode of replication with eukaryotic-like machinery in a hyperthermophilic archaeon. Science. 2000;288(5474):2212-5.
9.	Samson RY, Bell SD. Cell cycles and cell division in the archaea. Curr Opin Microbiol. 2011;14(3):350-6.
10.	Wu Z, Liu J, Yang H, Xiang H. DNA replication origins in archaea. Front Microbiol. 2014;5:179.
11.	Hawkins M, Malla S, Blythe MJ, Nieduszynski CA, Allers T. Accelerated growth in the absence of DNA replication origins. Nature. 2013;503(7477):544-7.
12.	Gehring AM, Astling DP, Matsumi R, Burkhart BW, Kelman Z, Reeve JN, et al. Genome Replication in Thermococcus kodakarensis Independent of Cdc6 and an Origin of Replication. Front Microbiol. 2017;8:2084.
13.	Li H, O'Donnell ME. The Eukaryotic CMG Helicase at the Replication Fork: Emerging Architecture Reveals an Unexpected Mechanism. Bioessays. 2018;40(3).
14.	Zheng L, Shen B. Okazaki fragment maturation: nucleases take centre stage. J Mol Cell Biol. 2011;3(1):23-30.
15.	Duggin IG, Bell SD. Termination structures in the Escherichia coli chromosome replication fork trap. J Mol Biol. 2009;387(3):532-9.
16.	Dewar JM, Walter JC. Mechanisms of DNA replication termination. Nat Rev Mol Cell Biol. 2017;18(8):507-16.
17.	Nordstrom K, Austin SJ. Cell-cycle-specific initiation of replication. Mol Microbiol. 1993;10(3):457-63.
18.	Dame RT, Kalmykowa OJ, Grainger DC. Chromosomal macrodomains and associated proteins: implications for DNA organization and replication in gram negative bacteria. PLoS Genet. 2011;7(6):e1002123.
19.	Cocker JH, Piatti S, Santocanale C, Nasmyth K, Diffley JF. An essential role for the Cdc6 protein in forming the pre-replicative complexes of budding yeast. Nature. 1996;379(6561):180-2.

20.   Diffley JF, Cocker JH. Protein-DNA interactions at a yeast replication origin. Nature. 1992;357(6374):169-72.

21.   Fangman WL, Brewer BJ. A question of time: replication origins of eukaryotic chromosomes. Cell. 1992;71(3):363-6.

22.   O'Keefe RT, Henderson SC, Spector DL. Dynamic organization of DNA replication in mammalian cell nuclei: spatially and temporally defined replication of chromosome-specific alpha-satellite DNA sequences. J Cell Biol. 1992;116(5):1095-110.

23.   Gilbert DM. In search of the holy replicator. Nat Rev Mol Cell Biol. 2004;5(10):848-55.

24.   Fragkos M, Ganier O, Coulombe P, Mechali M. DNA replication origin activation in space and time. Nat Rev Mol Cell Biol. 2015;16(6):360-74.

25.   Cayrou C, Coulombe P, Vigneron A, Stanojcic S, Ganier O, Peiffer I, et al. Genome-scale analysis of metazoan replication origins reveals their organization in specific but flexible sites defined by conserved features. Genome Res. 2011;21(9):1438-49.

26.   Friedman KL, Brewer BJ, Fangman WL. Replication profile of Saccharomyces cerevisiae chromosome VI. Genes Cells. 1997;2(11):667-78.

27.   Heichinger C, Penkett CJ, Bahler J, Nurse P. Genome-wide characterization of fission yeast DNA replication origins. EMBO J. 2006;25(21):5171-9.

28.   Heller RC, Kang S, Lam WM, Chen S, Chan CS, Bell SP. Eukaryotic origin-dependent DNA replication in vitro reveals sequential action of DDK and S-CDK kinases. Cell. 2011;146(1):80-91.

29.   Yeeles JT, Deegan TD, Janska A, Early A, Diffley JF. Regulated eukaryotic DNA replication origin firing with purified proteins. Nature. 2015;519(7544):431-5.

30.   Burgers PM. Polymerase dynamics at the eukaryotic DNA replication fork. J Biol Chem. 2009;284(7):4041-5.

31.   Huberman JA, Riggs AD. On the mechanism of DNA replication in mammalian chromosomes. J Mol Biol. 1968;32(2):327-41.

32.   Jackson DA, Pombo A. Replicon clusters are stable units of chromosome structure: evidence that nuclear organization contributes to the efficient activation and propagation of S phase in human cells. J Cell Biol. 1998;140(6):1285-95.

33.   Fanning E, Klimovich V, Nager AR. A dynamic model for replication protein A (RPA) function in DNA processing pathways. Nucleic Acids Res. 2006;34(15):4126-37.

34.   Ball HL, Ehrhardt MR, Mordes DA, Glick GG, Chazin WJ, Cortez D. Function of a conserved checkpoint recruitment domain in ATRIP proteins. Mol Cell Biol. 2007;27(9):3367-77.

35.   Copeland WC, Wang TS. Catalytic subunit of human DNA polymerase alpha overproduced from baculovirus-infected insect cells. Structural and enzymological characterization. J Biol Chem. 1991;266(33):22739-48.

36.   Vaithiyalingam S, Warren EM, Eichman BF, Chazin WJ. Insights into eukaryotic DNA priming from the structure and functional interactions of the 4Fe-4S cluster domain of human DNA primase. Proc Natl Acad Sci U S A. 2010;107(31):13684-9.

37.   Sheaff RJ, Kuchta RD. Mechanism of calf thymus DNA primase: slow initiation, rapid polymerization, and intelligent termination. Biochemistry. 1993;32(12):3027-37.

38.   Tseng BY, Goulian M. Initiator RNA of discontinuous DNA synthesis in human lymphocytes. Cell. 1977;12(2):483-9.

39.   Nordman JT, Kozhevnikova EN, Verrijzer CP, Pindyurin AV, Andreyeva EN, Shloma VV, et al. DNA copy-number control through inhibition of replication fork progression. Cell Rep. 2014;9(3):841-9.

40.   Sirbu BM, McDonald WH, Dungrawala H, Badu-Nkansah A, Kavanaugh GM, Chen Y, et al. Identification of proteins at active, stalled, and collapsed replication forks using isolation of proteins on nascent DNA (iPOND) coupled with mass spectrometry. J Biol Chem. 2013;288(44):31458-67.

41.   Lopez-Contreras AJ, Ruppen I, Nieto-Soler M, Murga M, Rodriguez-Acebes S, Remeseiro S, et al. A proteomic characterization of factors enriched at nascent DNA molecules. Cell Rep. 2013;3(4):1105-16.

42.   Collins N, Poot RA, Kukimoto I, Garcia-Jimenez C, Dellaire G, Varga-Weisz PD. An ACF1-ISWI chromatin-remodeling complex is required for DNA replication through heterochromatin. Nat Genet. 2002;32(4):627-32.

43.   Lee JB, Hite RK, Hamdan SM, Xie XS, Richardson CC, van Oijen AM. DNA primase acts as a molecular brake in DNA replication. Nature. 2006;439(7076):621-4.

44.   Kilkenny ML, Simon AC, Mainwaring J, Wirthensohn D, Holzer S, Pellegrini L. The human CTF4-orthologue AND-1 interacts with DNA polymerase alpha/primase via its unique C-terminal HMG box. Open Biol. 2017;7(11).

45.   Simon AC, Zhou JC, Perera RL, van Deursen F, Evrin C, Ivanova ME, et al. A Ctf4 trimer couples the CMG helicase to DNA polymerase alpha in the eukaryotic replisome. Nature. 2014;510(7504):293-7.

46.   Villa F, Simon AC, Ortiz Bazan MA, Kilkenny ML, Wirthensohn D, Wightman M, et al. Ctf4 Is a Hub in the Eukaryotic Replisome that Links Multiple CIP-Box Proteins to the CMG Helicase. Mol Cell. 2016;63(3):385-96.

47.   Zechiedrich EL, Osheroff N. Eukaryotic topoisomerases recognize nucleic acid topology by preferentially interacting with DNA crossovers. EMBO J. 1990;9(13):4555-62.

48.   Postow L, Crisona NJ, Peter BJ, Hardy CD, Cozzarelli NR. Topological challenges to DNA replication: conformations at the fork. Proc Natl Acad Sci U S A. 2001;98(15):8219-26.

49.   Dewar JM, Low E, Mann M, Raschle M, Walter JC. CRL2(Lrr1) promotes unloading of the vertebrate replisome from chromatin during replication termination. Genes Dev. 2017;31(3):275-90.

50.   Maric M, Maculins T, De Piccoli G, Labib K. Cdc48 and a ubiquitin ligase drive disassembly of the CMG helicase at the end of DNA replication. Science. 2014;346(6208):1253596.

51.   Moreno SP, Bailey R, Campion N, Herron S, Gambus A. Polyubiquitylation drives replisome disassembly at the termination of DNA replication. Science. 2014;346(6208):477-81.

52.   Langston LD, Zhang D, Yurieva O, Georgescu RE, Finkelstein J, Yao NY, et al. CMG helicase and DNA polymerase epsilon form a functional 15-subunit holoenzyme for eukaryotic leading-strand DNA replication. Proc Natl Acad Sci U S A. 2014;111(43):15390-5.

53.   Kubota T, Nishimura K, Kanemaki MT, Donaldson AD. The Elg1 replication factor C-like complex functions in PCNA unloading during DNA replication. Mol Cell. 2013;50(2):273-80.

54.   Mechali M. Eukaryotic DNA replication origins: many choices for appropriate answers. Nat Rev Mol Cell Biol. 2010;11(10):728-38.

55.   Deegan TD, Diffley JF. MCM: one ring to rule them all. Curr Opin Struct Biol. 2016;37:145-51.

56.   Parker MW, Botchan MR, Berger JM. Mechanisms and regulation of DNA replication initiation in eukaryotes. Crit Rev Biochem Mol Biol. 2017;52(2):107-44.

57.   Coulombe P, Gregoire D, Tsanov N, Mechali M. A spontaneous Cdt1 mutation in 129 mouse strains reveals a regulatory domain restraining replication licensing. Nat Commun. 2013;4:2065.

58.   Nguyen VQ, Co C, Irie K, Li JJ. Clb/Cdc28 kinases promote nuclear export of the replication initiator proteins Mcm2-7. Curr Biol. 2000;10(4):195-205.

59.   Petersen BO, Lukas J, Sorensen CS, Bartek J, Helin K. Phosphorylation of mammalian CDC6 by cyclin A/CDK2 regulates its subcellular localization. EMBO J. 1999;18(2):396-410.

60.   Saha T, Ghosh S, Vassilev A, DePamphilis ML. Ubiquitylation, phosphorylation and Orc2 modulate the subcellular location of Orc1 and prevent it from inducing apoptosis. J Cell Sci. 2006;119(Pt 7):1371-82.

61.   Rivera-Mulia JC, Gilbert DM. Replicating Large Genomes: Divide and Conquer. Mol Cell. 2016;62(5):756-65.

62.   Tada S, Li A, Maiorano D, Mechali M, Blow JJ. Repression of origin assembly in metaphase depends on inhibition of RLF-B/Cdt1 by geminin. Nat Cell Biol. 2001;3(2):107-13.

63.   Wohlschlegel JA, Dwyer BT, Dhar SK, Cvetic C, Walter JC, Dutta A. Inhibition of eukaryotic DNA replication by geminin binding to Cdt1. Science. 2000;290(5500):2309-12.

64.   Mendez J, Zou-Yang XH, Kim SY, Hidaka M, Tansey WP, Stillman B. Human origin recognition complex large subunit is degraded by ubiquitin-mediated proteolysis after initiation of DNA replication. Mol Cell. 2002;9(3):481-91.

65.   Li X, Zhao Q, Liao R, Sun P, Wu X. The SCF(Skp2) ubiquitin ligase complex interacts with the human replication licensing factor Cdt1 and regulates Cdt1 degradation. J Biol Chem. 2003;278(33):30854-8.

66.   Berezney R, Dubey DD, Huberman JA. Heterogeneity of eukaryotic replicons, replicon clusters, and replication foci. Chromosoma. 2000;108(8):471-84.

67.   Nakayasu H, Berezney R. Mapping replicational sites in the eucaryotic cell nucleus. J Cell Biol. 1989;108(1):1-11.

68.   Nakamura H, Morita T, Sato C. Structural organizations of replicon domains during DNA synthetic phase in the mammalian nucleus. Exp Cell Res. 1986;165(2):291-7.

69.   Sparvoli E, Levi M, Rossi E. Replicon clusters may form structurally stable complexes of chromatin and chromosomes. J Cell Sci. 1994;107 (Pt 11):3097-103.

70.   Hansen RS, Thomas S, Sandstrom R, Canfield TK, Thurman RE, Weaver M, et al. Sequencing newly replicated DNA reveals widespread plasticity in human replication timing. Proc Natl Acad Sci U S A. 2010;107(1):139-44.

71.   Hiratani I, Ryba T, Itoh M, Yokochi T, Schwaiger M, Chang CW, et al. Global reorganization of replication domains during embryonic stem cell differentiation. PLoS Biol. 2008;6(10):e245.

72.   Stamatoyannopoulos JA, Adzhubei I, Thurman RE, Kryukov GV, Mirkin SM, Sunyaev SR. Human mutation rate associated with DNA replication timing. Nat Genet. 2009;41(4):393-5.

73.   Lambert S, Carr AM. Replication stress and genome rearrangements: lessons from yeast models. Curr Opin Genet Dev. 2013;23(2):132-9.

74.   Mouron S, Rodriguez-Acebes S, Martinez-Jimenez MI, Garcia-Gomez S, Chocron S, Blanco L, et al. Repriming of DNA synthesis at stalled replication forks by human PrimPol. Nat Struct Mol Biol. 2013;20(12):1383-9.

75.   Garcia-Gomez S, Reyes A, Martinez-Jimenez MI, Chocron ES, Mouron S, Terrados G, et al. PrimPol, an archaic primase/polymerase operating in human cells. Mol Cell. 2013;52(4):541-53.

76.   Cortez D. Preventing replication fork collapse to maintain genome integrity. DNA Repair (Amst). 2015;32:149-57.

77.   Byun TS, Pacek M, Yee MC, Walter JC, Cimprich KA. Functional uncoupling of MCM helicase and DNA polymerase activities activates the ATR-dependent checkpoint. Genes Dev. 2005;19(9):1040-52.

78.   Branzei D, Szakal B. Building up and breaking down: mechanisms controlling recombination during replication. Crit Rev Biochem Mol Biol. 2017;52(4):381-94.

79.   Iyer DR, Rhind N. The Intra-S Checkpoint Responses to DNA Damage. Genes (Basel). 2017;8(2).

80.   Carr AM, Lambert S. Replication stress-induced genome instability: the dark side of replication maintenance by homologous recombination. J Mol Biol. 2013;425(23):4733-44.

81.   Halazonetis TD, Gorgoulis VG, Bartek J. An oncogene-induced DNA damage model for cancer development. Science. 2008;319(5868):1352-5.

82.   Zhang F, Carvalho CM, Lupski JR. Complex human chromosomal and genomic rearrangements. Trends Genet. 2009;25(7):298-307.

83.   Zeman MK, Cimprich KA. Causes and consequences of replication stress. Nat Cell Biol. 2014;16(1):2-9.

84.   Wan L, Lou J, Xia Y, Su B, Liu T, Cui J, et al. hPrimpol1/CCDC111 is a human DNA primase-polymerase required for the maintenance of genome integrity. EMBO Rep. 2013;14(12):1104-12.

85.   Redrejo-Rodriguez M, Ordonez CD, Berjon-Otero M, Moreno-Gonzalez J, Aparicio-Maldonado C, Forterre P, et al. Primer-Independent DNA Synthesis by a Family B DNA Polymerase from Self-Replicating Mobile Genetic Elements. Cell Rep. 2017;21(6):1574-87.

86.   Zhang SS, Grosse F. Accuracy of DNA primase. J Mol Biol. 1990;216(3):475-9.

87.   Hauschka PV. Analysis of nucleotide pools in animal cells. Methods Cell Biol. 1973;7:361-462.

88.   Kuchta RD, Stengel G. Mechanism and evolution of DNA primases. Biochim Biophys Acta. 2010;1804(5):1180-9.

89.   Arts EJ, Stetor SR, Li X, Rausch JW, Howard KJ, Ehresmann B, et al. Initiation of (-) strand DNA synthesis from tRNA(3Lys) on lentiviral RNAs: implications of specific HIV-1 RNA-tRNA(3Lys) interactions inhibiting primer utilization by retroviral reverse transcriptases. Proc Natl Acad Sci U S A. 1996;93(19):10063-8.

90.   Paul AV, Wimmer E. Initiation of protein-primed picornavirus RNA synthesis. Virus Res. 2015;206:12-26.

91.   Longas E, Villar L, Lazaro JM, de Vega M, Salas M. Phage phi29 and Nf terminal protein-priming domain specifies the internal template nucleotide to initiate DNA replication. Proc Natl Acad Sci U S A. 2008;105(47):18290-5.

92.   de Jong RN, van der Vliet PC, Brenkman AB. Adenovirus DNA replication: protein priming, jumping back and the role of the DNA binding protein DBP. Curr Top Microbiol Immunol. 2003;272:187-211.

93.   Kishi T, Miura K, Matsumoto K, Hirokawa H. Mechanism of protein priming DNA replication of B.subtilis phage M2. Nucleic Acids Symp Ser. 1992(27):187-8.

94.   Salas M. Protein-priming of DNA replication. Annu Rev Biochem. 1991;60:39-71.

95.   Pronk R, van der Vliet PC. The adenovirus terminal protein influences binding of replication proteins and changes the origin structure. Nucleic Acids Res. 1993;21(10):2293-300.

96.   Mak J, Kleiman L. Primer tRNAs for reverse transcription. J Virol. 1997;71(11):8087-95.

97.   Wakefield JK, Wolf AG, Morrow CD. Human immunodeficiency virus type 1 can use different tRNAs as primers for reverse transcription but selectively maintains a primer binding site complementary to tRNA(3Lys). J Virol. 1995;69(10):6021-9.

98.   Rowen L, Kornberg A. Primase, the dnaG protein of Escherichia coli. An enzyme which starts DNA chains. J Biol Chem. 1978;253(3):758-64.

99.   Aravind L, Leipe DD, Koonin EV. Toprim--a conserved catalytic domain in type IA and II topoisomerases, DnaG-type primases, OLD family nucleases and RecR proteins. Nucleic Acids Res. 1998;26(18):4205-13.

100.   Iyer LM, Koonin EV, Leipe DD, Aravind L. Origin and evolution of the archaeo-eukaryotic primase superfamily and related palm-domain proteins: structural insights and new members. Nucleic Acids Res. 2005;33(12):3875-96.

101.   Corn JE, Pelton JG, Berger JM. Identification of a DNA primase template tracking site redefines the geometry of primer synthesis. Nat Struct Mol Biol. 2008;15(2):163-9.

102.   Kato M, Ito T, Wagner G, Richardson CC, Ellenberger T. Modular architecture of the bacteriophage T7 primase couples RNA primer synthesis to DNA synthesis. Mol Cell. 2003;11(5):1349-60.

103.   Keck JL, Roche DD, Lynch AS, Berger JM. Structure of the RNA polymerase domain of E. coli primase. Science. 2000;287(5462):2482-6.

104.   Hinton DM, Nossal NG. Bacteriophage T4 DNA primase-helicase. Characterization of oligomer synthesis by T4 61 protein alone and in conjunction with T4 41 protein. J Biol Chem. 1987;262(22):10873-8.

105.   Yang J, Xi J, Zhuang Z, Benkovic SJ. The oligomeric T4 primase is the functional form during replication. J Biol Chem. 2005;280(27):25416-23.

106.   Valentine AM, Ishmael FT, Shier VK, Benkovic SJ. A zinc ribbon protein in DNA replication: primer synthesis and macromolecular interactions by the bacteriophage T4 primase. Biochemistry. 2001;40(50):15074-85.

107.   Mitkova AV, Khopde SM, Biswas SB. Mechanism and stoichiometry of interaction of DnaG primase with DnaB helicase of Escherichia coli in RNA primer synthesis. J Biol Chem. 2003;278(52):52253-61.

108.   Johnson SK, Bhattacharyya S, Griep MA. DnaB helicase stimulates primer synthesis activity on short oligonucleotide templates. Biochemistry. 2000;39(4):736-44.

109.   Biswas N, Weller SK. A mutation in the C-terminal putative Zn2+ finger motif of UL52 severely affects the biochemical activities of the HSV-1 helicase-primase subcomplex. J Biol Chem. 1999;274(12):8068-76.

110.   Biswas N, Weller SK. The UL5 and UL52 subunits of the herpes simplex virus type 1 helicase-primase subcomplex exhibit a complex interdependence for DNA binding. J Biol Chem. 2001;276(20):17610-9.

111.   Graves-Woodward KL, Gottlieb J, Challberg MD, Weller SK. Biochemical analyses of mutations in the HSV-1 helicase-primase that alter ATP hydrolysis, DNA unwinding, and coupling between hydrolysis and unwinding. J Biol Chem. 1997;272(7):4623-30.

112.   Arezi B, Kirk BW, Copeland WC, Kuchta RD. Interactions of DNA with human DNA primase monitored with photoactivatable cross-linking agents: implications for the role of the p58 subunit. Biochemistry. 1999;38(39):12899-907.

113.   Kirk BW, Kuchta RD. Human DNA primase: anion inhibition, manganese stimulation, and their effects on in vitro start-site selection. Biochemistry. 1999;38(31):10126-34.
114.   Zerbe LK, Kuchta RD. The p58 subunit of human DNA primase is important for primer initiation, elongation, and counting. Biochemistry. 2002;41(15):4891-900.
115.   Baranovskiy AG, Zhang Y, Suwa Y, Gu J, Babayeva ND, Pavlov YI, et al. Insight into the Human DNA Primase Interaction with Template-Primer. J Biol Chem. 2016;291(9):4793-802.
116.   Weiner BE, Huang H, Dattilo BM, Nilges MJ, Fanning E, Chazin WJ. An iron-sulfur cluster in the C-terminal domain of the p58 subunit of human DNA primase. J Biol Chem. 2007;282(46):33444-51.
117.   Sauguet L, Klinge S, Perera RL, Maman JD, Pellegrini L. Shared active site architecture between the large subunit of eukaryotic primase and DNA photolyase. PLoS One. 2010;5(4):e10083.
118.   Klinge S, Hirst J, Maman JD, Krude T, Pellegrini L. An iron-sulfur domain of the eukaryotic primase is essential for RNA primer synthesis. Nat Struct Mol Biol. 2007;14(9):875-7.
119.   Holzer S, Yan J, Kilkenny ML, Bell SD, Pellegrini L. Primer synthesis by a eukaryotic-like archaeal primase is independent of its Fe-S cluster. Nat Commun. 2017;8(1):1718.
120.   Yan J, Holzer S, Pellegrini L, Bell SD. An archaeal primase functions as a nanoscale caliper to define primer length. Proc Natl Acad Sci U S A. 2018.
121.   Liu B, Ouyang S, Makarova KS, Xia Q, Zhu Y, Li Z, et al. A primase subunit essential for efficient primer synthesis by an archaeal eukaryotic-type primase. Nat Commun. 2015;6:7300.
122.   Pellegrini L. The Pol alpha-primase complex. Subcell Biochem. 2012;62:157-69.
123.   Davey SK, Faust EA. Murine DNA polymerase.alpha-primase initiates RNA-primed DNA synthesis preferentially upstream of a 3'-CC(C/A)-5' motif. J Biol Chem. 1990;265(7):3611-4.
124.   Yamaguchi M, Hendrickson EA, DePamphilis ML. DNA primase-DNA polymerase alpha from simian cells: sequence specificity of initiation sites on simian virus 40 DNA. Mol Cell Biol. 1985;5(5):1170-83.
125.   Suzuki M, Savoysky E, Izuta S, Tatebe M, Okajima T, Yoshida S. RNA priming coupled with DNA synthesis on natural template by calf thymus DNA polymerase alpha-primase. Biochemistry. 1993;32(47):12782-92.
126.   Holmes AM, Cheriathundam E, Bollum FJ, Chang LM. Initiation of DNA synthesis by the calf thymus DNA polymerase-primase complex. J Biol Chem. 1985;260(19):10840-6.
127.   Koepsell SA, Larson MA, Frey CA, Hinrichs SH, Griep MA. Staphylococcus aureus primase has higher initiation specificity, interacts with single-stranded DNA stronger, but is less stimulated by its helicase than Escherichia coli primase. Mol Microbiol. 2008;68(6):1570-82.
128.   Thirlway J, Soultanas P. In the Bacillus stearothermophilus DnaB-DnaG complex, the activities of the two proteins are modulated by distinct but overlapping networks of residues. J Bacteriol. 2006;188(4):1534-9.
129.   Lao-Sirieix SH, Bell SD. The heterodimeric primase of the hyperthermophilic archaeon Sulfolobus solfataricus possesses DNA and RNA primase, polymerase and 3'-terminal nucleotidyl transferase activities. J Mol Biol. 2004;344(5):1251-63.
130.   Steitz TA. DNA polymerases: structural diversity and common mechanisms. J Biol Chem. 1999;274(25):17395-8.

131.   Frick DN, Richardson CC. DNA primases. Annu Rev Biochem. 2001;70:39-80.
132.   Copeland WC, Wang TS. Enzymatic characterization of the individual mammalian primase subunits reveals a biphasic mechanism for initiation of DNA replication. J Biol Chem. 1993;268(35):26179-89.
133.   Cotterill S, Chui G, Lehman IR. DNA polymerase-primase from embryos of Drosophila melanogaster. DNA primase subunits. J Biol Chem. 1987;262(33):16105-8.
134.   Copeland WC. Expression, purification, and characterization of the two human primase subunits and truncated complexes from Escherichia coli. Protein Expr Purif. 1997;9(1):1-9.
135.   Copeland WC, Tan X. Active site mapping of the catalytic mouse primase subunit by alanine scanning mutagenesis. J Biol Chem. 1995;270(8):3905-13.
136.   Kirk BW, Kuchta RD. Arg304 of human DNA primase is a key contributor to catalysis and NTP binding: primase and the family X polymerases share significant sequence homology. Biochemistry. 1999;38(24):7727-36.
137.   Baranovskiy AG, Babayeva ND, Zhang Y, Gu J, Suwa Y, Pavlov YI, et al. Mechanism of Concerted RNA-DNA Primer Synthesis by the Human Primosome. J Biol Chem. 2016;291(19):10006-20.
138.   Sheaff RJ, Kuchta RD, Ilsley D. Calf thymus DNA polymerase alpha-primase: "communication" and primer-template movement between the two active sites. Biochemistry. 1994;33(8):2247-54.
139.   Kuchta RD, Reid B, Chang LM. DNA primase. Processivity and the primase to polymerase alpha activity switch. J Biol Chem. 1990;265(27):16158-65.
140.   Baranovskiy AG, Gu J, Babayeva ND, Agarkar VB, Suwa Y, Tahirov TH. Crystallization and preliminary X-ray diffraction analysis of human DNA primase. Acta Crystallogr F Struct Biol Commun. 2014;70(Pt 2):206-10.
141.   Baranovskiy AG, Zhang Y, Suwa Y, Babayeva ND, Gu J, Pavlov YI, et al. Crystal structure of the human primase. J Biol Chem. 2015;290(9):5635-46.
142.   Zhang Y, Baranovskiy AG, Tahirov TH, Pavlov YI. The C-terminal domain of the DNA polymerase catalytic subunit regulates the primase and polymerase activities of the human DNA polymerase alpha-primase complex. J Biol Chem. 2014;289(32):22021-34.
143.   Vaithiyalingam S, Arnett DR, Aggarwal A, Eichman BF, Fanning E, Chazin WJ. Insights into eukaryotic primer synthesis from structures of the p48 subunit of human DNA primase. J Mol Biol. 2014;426(3):558-69.
144.   Kilkenny ML, Longo MA, Perera RL, Pellegrini L. Structures of human primase reveal design of nucleotide elongation site and mode of Pol alpha tethering. Proc Natl Acad Sci U S A. 2013;110(40):15961-6.
145.   Baranovskiy AG, Tahirov TH. Elaborated Action of the Human Primosome. Genes (Basel). 2017;8(2).
146.   Winter A, Higueruelo AP, Marsh M, Sigurdardottir A, Pitt WR, Blundell TL. Biophysical and computational fragment-based approaches to targeting protein-protein interactions: applications in structure-guided drug discovery. Q Rev Biophys. 2012;45(4):383-426.
147.   Ward AB, Sali A, Wilson IA. Biochemistry. Integrative structural biology. Science. 2013;339(6122):913-5.
148.   Liang Z, Hu G. Protein Structure Network-based Drug Design. Mini Rev Med Chem. 2016;16(16):1330-43.
149.   Chen L, Morrow JK, Tran HT, Phatak SS, Du-Cuny L, Zhang S. From laptop to benchtop to bedside: structure-based drug design on protein targets. Curr Pharm Des. 2012;18(9):1217-39.

150. Aydin H, Taylor MW, Lee JE. Structure-guided analysis of the human APOBEC3-HIV restrictome. Structure. 2014;22(5):668-84.

151. Dickey TH, Altschuler SE, Wuttke DS. Single-stranded DNA-binding proteins: multiple domains for multiple functions. Structure. 2013;21(7):1074-84.

152. Focke PJ, Wang X, Larsson HP. Neurotransmitter transporters: structure meets function. Structure. 2013;21(5):694-705.

153. Valkov E, Sharpe T, Marsh M, Greive S, Hyvonen M. Targeting protein-protein interactions and fragment-based drug discovery. Top Curr Chem. 2012;317:145-79.

154. Rupp B. Biomolecular crystallography : principles, practice, and application to structural biology. New York: Garland Science; 2010. xxi, 809 p. p.

155. Buonfiglio R, Recanatini M, Masetti M. Protein Flexibility in Drug Discovery: From Theory to Computation. ChemMedChem. 2015;10(7):1141-8.

156. Cavanagh J. Protein NMR spectroscopy : principles and practice. 2nd ed. Amsterdam ; Boston: Academic Press; 2007. xxv, 885 p. p.

157. Putnam CD, Hammel M, Hura GL, Tainer JA. 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):191-285.

158. Hagen WR. EPR spectroscopy of iron—sulfur proteins. Advances in inorganic chemistry. 38: Elsevier; 1992. p. 165-222.

159. Schuler B, Eaton WA. Protein folding studied by single-molecule FRET. Current opinion in structural biology. 2008;18(1):16-26.

160. Adinolfi S, Iannuzzi C, Prischi F, Pastore C, Iametti S, Martin SR, et al. Bacterial frataxin CyaY is the gatekeeper of iron-sulfur cluster formation catalyzed by IscS. Nat Struct Mol Biol. 2009;16(4):390-6.

161. Crack JC, den Hengst CD, Jakimowicz P, Subramanian S, Johnson MK, Buttner MJ, et al. Characterization of [4Fe-4S]-containing and cluster-free forms of Streptomyces WhiD. Biochemistry. 2009;48(51):12252-64.

162. Netz DJ, Stith CM, Stumpfig M, Kopf G, Vogel D, Genau HM, et al. Eukaryotic DNA polymerases require an iron-sulfur cluster for the formation of active complexes. Nat Chem Biol. 2011;8(1):125-32.

163. Agar JN, Krebs C, Frazzon J, Huynh BH, Dean DR, Johnson MK. IscU as a scaffold for iron-sulfur cluster biosynthesis: sequential assembly of [2Fe-2S] and [4Fe-4S] clusters in IscU. Biochemistry. 2000;39(27):7856-62.

164. Latimer MT, Painter MH, Ferry JG. Characterization of an iron-sulfur flavoprotein from Methanosarcina thermophila. J Biol Chem. 1996;271(39):24023-8.

165. Nakamaru-Ogiso E, Yano T, Ohnishi T, Yagi T. Characterization of the iron-sulfur cluster coordinated by a cysteine cluster motif (CXXCXXXCX27C) in the Nqo3 subunit in the proton-translocating NADH-quinone oxidoreductase (NDH-1) of Thermus thermophilus HB-8. J Biol Chem. 2002;277(3):1680-8.

166. Leartsakulpanich U, Antonkine ML, Ferry JG. Site-specific mutational analysis of a novel cysteine motif proposed to ligate the 4Fe-4S cluster in the iron-sulfur flavoprotein of the thermophilic methanoarchaeon Methanosarcina thermophila. J Bacteriol. 2000;182(19):5309-16.

167. Mapolelo DT, Zhang B, Naik SG, Huynh BH, Johnson MK. Spectroscopic and functional characterization of iron-sulfur cluster-bound forms of Azotobacter vinelandii (Nif)IscA. Biochemistry. 2012;51(41):8071-84.

168.   Beychok S. Circular dichroism of biological macromolecules. Science. 1966;154(3754):1288-99.

169.   Li CH, Nguyen X, Narhi L, Chemmalil L, Towers E, Muzammil S, et al. Applications of circular dichroism (CD) for structural analysis of proteins: qualification of near- and far-UV CD for protein higher order structural analysis. J Pharm Sci. 2011;100(11):4642-54.

170.   Greenfield NJ. Using circular dichroism spectra to estimate protein secondary structure. Nat Protoc. 2006;1(6):2876-90.

171.   Johnson WC. Analyzing protein circular dichroism spectra for accurate secondary structures. Proteins. 1999;35(3):307-12.

172.   Louis-Jeune C, Andrade-Navarro MA, Perez-Iratxeta C. Prediction of protein secondary structure from circular dichroism using theoretically derived spectra. Proteins: Structure, Function, and Bioinformatics. 2012;80(2):374-81.

173.   Bonomi F, Iametti S, Morleo A, Ta D, Vickery LE. Facilitated transfer of IscU-[2Fe2S] clusters by chaperone-mediated ligand exchange. Biochemistry. 2011;50(44):9641-50.

174.   Bridwell-Rabb J, Fox NG, Tsai CL, Winn AM, Barondeau DP. Human frataxin activates Fe-S cluster biosynthesis by facilitating sulfur transfer chemistry. Biochemistry. 2014;53(30):4904-13.

175.   Stephens PJ, Thomson AJ, Dunn JB, Keiderling TA, Rawlings J, Rao KK, et al. Circular dichroism and magnetic circular dichroism of iron-sulfur proteins. Biochemistry. 1978;17(22):4770-8.

176.   Stephens PJ, Thomson AJ, Keiderling TA, Rawlings J, Rao KK, Hall DO. Cluster characterization in iron-sulfur proteins by magnetic circular dichroism. Proc Natl Acad Sci U S A. 1978;75(11):5273-5.

177.   Fuss JO, Tsai CL, Ishida JP, Tainer JA. Emerging critical roles of Fe-S clusters in DNA replication and repair. Biochim Biophys Acta. 2015;1853(6):1253-71.

178.   Wüthrich K. NMR of proteins and nucleic acids. New York: Wiley; 1986. xv, 292 p.

179.   Takeuchi K, Wagner G. NMR studies of protein interactions. Curr Opin Struct Biol. 2006;16(1):109-17.

180.   Bertini I, Luchinat C, Parigi G, Pierattelli R. NMR spectroscopy of paramagnetic metalloproteins. Chembiochem. 2005;6(9):1536-49.

181.   Hura GL, Menon AL, Hammel M, Rambo RP, Poole FL, 2nd, Tsutakawa SE, et al. Robust, high-throughput solution structural analyses by small angle X-ray scattering (SAXS). Nat Methods. 2009;6(8):606-12.

182.   Rambo RP, Tainer JA. Bridging the solution divide: comprehensive structural analyses of dynamic RNA, DNA, and protein assemblies by small-angle X-ray scattering. Curr Opin Struct Biol. 2010;20(1):128-37.

183.   Bernado P, Mylonas E, Petoukhov MV, Blackledge M, Svergun DI. Structural characterization of flexible proteins using small-angle X-ray scattering. J Am Chem Soc. 2007;129(17):5656-64.

184.   Rambo RP, Tainer JA. Characterizing flexible and intrinsically unstructured biological macromolecules by SAS using the Porod-Debye law. Biopolymers. 2011;95(8):559-71.

185.   Debye P, Anderson Jr H, Brumberger H. Scattering by an inhomogeneous solid. II. The correlation function and its application. Journal of applied Physics. 1957;28(6):679-83.

186.   Tsutakawa SE, Hura GL, Frankel KA, Cooper PK, Tainer JA. Structural analysis of flexible proteins in solution by small angle X-ray scattering combined with crystallography. J Struct Biol. 2007;158(2):214-23.

187.   Moore PB. Small-angle scattering. Information content and error analysis. Journal of applied crystallography. 1980;13(2):168-75.

188.   Schneidman-Duhovny D, Hammel M, Tainer JA, Sali A. FoXS, FoXSDock and MultiFoXS: Single-state and multi-state structural modeling of proteins and their complexes based on SAXS profiles. Nucleic Acids Res. 2016;44(W1):W424-9.

189.   Schneidman-Duhovny D, Hammel M, Tainer JA, Sali A. Accurate SAXS profile computation and its assessment by contrast variation experiments. Biophys J. 2013;105(4):962-74.

190.   Svergun D, Barberato C, Koch MH. CRYSOL–a program to evaluate X-ray solution scattering of biological macromolecules from atomic coordinates. Journal of applied crystallography. 1995;28(6):768-73.

191.   Volkov VV, Svergun DI. Uniqueness of ab initio shape determination in small-angle scattering. Journal of applied crystallography. 2003;36(3-1):860-4.

192.   Franke D, Svergun DI. DAMMIF, a program for rapid ab-initio shape determination in small-angle scattering. Journal of applied crystallography. 2009;42(2):342-6.

193.   Svergun DI. Restoring low resolution structure of biological macromolecules from solution scattering using simulated annealing. Biophysical journal. 1999;76(6):2879-86.

194.   Brosey CA, Soss SE, Brooks S, Yan C, Ivanov I, Dorai K, et al. Functional dynamics in replication protein A DNA binding and protein recruitment domains. Structure. 2015;23(6):1028-38.

195.   Hashiguchi T, Fusco ML, Bornholdt ZA, Lee JE, Flyak AI, Matsuoka R, et al. Structural basis for Marburg virus neutralization by a cross-reactive human antibody. Cell. 2015;160(5):904-12.

196.   Fan L, Fuss JO, Cheng QJ, Arvai AS, Hammel M, Roberts VA, et al. XPD helicase structures and activities: insights into the cancer and aging phenotypes from XPD mutations. Cell. 2008;133(5):789-800.

197.   Sarker AH, Tsutakawa SE, Kostek S, Ng C, Shin DS, Peris M, et al. Recognition of RNA polymerase II and transcription bubbles by XPG, CSB, and TFIIH: insights for transcription-coupled repair and Cockayne Syndrome. Mol Cell. 2005;20(2):187-98.

198.   Pollard TD. A guide to simple and informative binding assays. Mol Biol Cell. 2010;21(23):4061-7.

199.   Harris DC. Quantitative chemical analysis: Macmillan; 2010.

200.   Brown A. Analysis of cooperativity by isothermal titration calorimetry. Int J Mol Sci. 2009;10(8):3457-77.

201.   Correia JJ, Detrich HW. Biophysical tools for biologists. Vol. 1, In vitro techniques. Amsterdam ;: Boston : Elsevier; 2008. xxiii, 972 p., [12] p. of plates p.

202.   Zhang H, Wu Q, Berezin MY. Fluorescence anisotropy (polarization): from drug screening to precision medicine. Expert Opin Drug Discov. 2015;10(11):1145-61.

203.   Yengo CM, Berger CL. Fluorescence anisotropy and resonance energy transfer: powerful tools for measuring real time protein dynamics in a physiological environment. Curr Opin Pharmacol. 2010;10(6):731-7.

204.   Mocz G. Information content of fluorescence polarization and anisotropy. J Fluoresc. 2006;16(4):511-24.

205.   Jameson DM, Ross JA. Fluorescence polarization/anisotropy in diagnostics and imaging. Chem Rev. 2010;110(5):2685-708.

206.   Jameson DM, Sawyer WH. Fluorescence anisotropy applied to biomolecular interactions. Methods Enzymol. 1995;246:283-300.

207.   Boal AK, Genereux JC, Sontz PA, Gralnick JA, Newman DK, Barton JK. Redox signaling between DNA repair proteins for efficient lesion detection. Proceedings of the National Academy of Sciences. 2009;106(36):15237-42.

208.   Sontz PA, Mui TP, Fuss JO, Tainer JA, Barton JK. DNA charge transport as a first step in coordinating the detection of lesions by repair proteins. Proceedings of the National Academy of Sciences. 2012;109(6):1856-61.

209.   Grodick MA, Muren NB, Barton JK. DNA charge transport within the cell. Biochemistry. 2015;54(4):962-73.

210.   Pheeney CG, Arnold AR, Grodick MA, Barton JK. Multiplexed electrochemistry of DNA-bound metalloproteins. J Am Chem Soc. 2013;135(32):11869-78.

211.   Grodick MA, Segal HM, Zwang TJ, Barton JK. DNA-mediated signaling by proteins with 4Fe-4S clusters is necessary for genomic integrity. J Am Chem Soc. 2014;136(17):6470-8.

212.   Gorodetsky AA, Boal AK, Barton JK. Direct electrochemistry of endonuclease III in the presence and absence of DNA. J Am Chem Soc. 2006;128(37):12082-3.

213.   Rouault TA. Mammalian iron–sulphur proteins: novel insights into biogenesis and function. Nature reviews Molecular cell biology. 2015;16(1):45.

214.   Rees DC, Howard JB. The interface between the biological and inorganic worlds: iron-sulfur metalloclusters. Science. 2003;300(5621):929-31.

215.   Agarkar VB, Babayeva ND, Pavlov YI, Tahirov TH. Crystal structure of the C-terminal domain of human DNA primase large subunit: implications for the mechanism of the primase-polymerase alpha switch. Cell Cycle. 2011;10(6):926-31.

216.   Lue NF, Chan J, Wright WE, Hurwitz J. The CDC13-STN1-TEN1 complex stimulates Pol alpha activity by promoting RNA priming and primase-to-polymerase switch. Nat Commun. 2014;5:5762.

217.   Georgescu RE, Schauer GD, Yao NY, Langston LD, Yurieva O, Zhang D, et al. Reconstitution of a eukaryotic replisome reveals suppression mechanisms that define leading/lagging strand operation. Elife. 2015;4:e04988.

218.   Holt ME, Salay LE, Chazin WJ. A Polymerase With Potential: The Fe-S Cluster in Human DNA Primase. Methods Enzymol. 2017;595:361-90.

219.   Collaborative Computational Project N. The CCP4 suite: programs for protein crystallography. Acta Crystallogr D Biol Crystallogr. 1994;50(Pt 5):760-3.

220.   Cunningham RP, Asahara H, Bank JF, Scholes CP, Salerno JC, Surerus K, et al. Endonuclease III is an iron-sulfur protein. Biochemistry. 1989;28(10):4450-5.

221.   Slinker JD, Muren NB, Gorodetsky AA, Barton JK. Multiplexed DNA-modified electrodes. J Am Chem Soc. 2010;132(8):2769-74.

222.   Slinker JD, Muren NB, Renfrew SE, Barton JK. DNA charge transport over 34 nm. Nat Chem. 2011;3(3):228-33.

223.   Imlay JA. Iron-sulfur clusters and the problem with oxygen. Molecular Microbiology. 2006;59(4):1073-82.

224.   Mui TP, Fuss JO, Ishida JP, Tainer JA, Barton JK. ATP-stimulated, DNA-mediated redox signaling by XPD, a DNA repair and transcription helicase. J Am Chem Soc. 2011;133(41):16378-81.

225.   Winkler JR, Gray HB. Long-range electron tunneling. J Am Chem Soc. 2014;136(8):2930-9.

226.   Plekan O, Feyer V, Richter R, Coreno M, Prince KC. Valence photoionization and photofragmentation of aromatic amino acids. Molecular Physics. 2008;106(9-10):1143-53.

227.   Forbes SA, Bindal N, Bamford S, Cole C, Kok CY, Beare D, et al. COSMIC: mining complete cancer genomes in the Catalogue of Somatic Mutations in Cancer. Nucleic Acids Res. 2011;39(Database issue):D945-50.

228.   Perera RL, Torella R, Klinge S, Kilkenny ML, Maman JD, Pellegrini L. Mechanism for priming DNA synthesis by yeast DNA polymerase α. Elife. 2013;2:e00482.

229.   O'Neill MA, Becker HC, Wan C, Barton JK, Zewail AH. Ultrafast dynamics in DNA-mediated electron transfer: base gating and the role of temperature. Angew Chem Int Ed Engl. 2003;42(47):5896-900.

230.   Gray HB, Winkler JR. Electron flow through metalloproteins. Biochim Biophys Acta. 2010;1797(9):1563-72.

231.   Liu L, Huang M. Essential role of the iron-sulfur cluster binding domain of the primase regulatory subunit Pri2 in DNA replication initiation. Protein Cell. 2015;6(3):194-210.

232.   O'Brien E, Salay LE, Epum EA, Friedman KL, Chazin WJ, Barton JK. Yeast require redox switching in DNA primase. Proc Natl Acad Sci U S A. 2018;115(52):13186-91.

233.   Baranovskiy AG, Babayeva ND, Zhang Y, Blanco L, Pavlov YI, Tahirov TH. Comment on "The [4Fe4S] cluster of human DNA primase functions as a redox switch using DNA charge transport". Science. 2017;357(6348).

234.   Pellegrini L. Comment on "The [4Fe4S] cluster of human DNA primase functions as a redox switch using DNA charge transport". Science. 2017;357(6348).

235.   O'Brien E, Holt ME, Thompson MK, Salay LE, Ehlinger AC, Chazin WJ, et al. The [4Fe4S] cluster of human DNA primase functions as a redox switch using DNA charge transport. Science. 2017;355(6327).

236.   Stodola JL, Burgers PM. Mechanism of Lagging-Strand DNA Replication in Eukaryotes. Advances in experimental medicine and biology. 2017;1042:117-33.

237.   Burgers PMJ, Kunkel TA. Eukaryotic DNA Replication Fork. Annu Rev Biochem. 2017;86:417-38.

238.   Yao N, O'Donnell M. Bacterial and Eukaryotic Replisome Machines. JSM biochemistry and molecular biology. 2016;3(1).

239.   Yeeles JTP, Janska A, Early A, Diffley JFX. How the Eukaryotic Replisome Achieves Rapid and Efficient DNA Replication. Mol Cell. 2017;65(1):105-16.

240.   Pellegrini L, Costa A. New Insights into the Mechanism of DNA Duplication by the Eukaryotic Replisome. Trends in biochemical sciences. 2016;41(10):859-71.

241.   O'Brien E, Holt ME, Thompson MK, Salay LE, Ehlinger AC, Chazin WJ, et al. Response to Comments on "The [4Fe4S] cluster of human DNA primase functions as a redox switch using DNA charge transport". Science. 2017;357(6348).

242.   Otwinowski Z, Minor W. Processing of X-ray diffraction data collected in oscillation mode. Methods Enzymol. 1997;276:307-26.

243.   Emsley P, Lohkamp B, Scott WG, Cowtan K. Features and development of Coot. Acta Crystallogr D Biol Crystallogr. 2010;66(Pt 4):486-501.

244.   Adams PD, Afonine PV, Bunkoczi G, Chen VB, Davis IW, Echols N, et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr D Biol Crystallogr. 2010;66(Pt 2):213-21.

245.   Afonine PV, Grosse-Kunstleve RW, Echols N, Headd JJ, Moriarty NW, Mustyakimov M, et al. Towards automated crystallographic structure refinement with phenix.refine. Acta Crystallogr D Biol Crystallogr. 2012;68(Pt 4):352-67.

246.   Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, et al. UCSF Chimera--a visualization system for exploratory research and analysis. Journal of computational chemistry. 2004;25(13):1605-12.

247.   Morin A, Eisenbraun B, Key J, Sanschagrin PC, Timony MA, Ottaviano M, et al. Collaboration gets the most out of software. Elife. 2013;2:e01456.

248.   Guillerez J, Lopez PJ, Proux F, Launay H, Dreyfus M. A mutation in T7 RNA polymerase that facilitates promoter clearance. Proc Natl Acad Sci U S A. 2005;102(17):5958-63.

249.   Linke P, Amaning K, Maschberger M, Vallee F, Steier V, Baaske P, et al. An Automated Microscale Thermophoresis Screening Approach for Fragment-Based Lead Discovery. Journal of Biomolecular Screening. 2016;21(4):414-21.

250.   Tse ECM, Zwang TJ, Barton JK. The Oxidation State of [4Fe4S] Clusters Modulates the DNA-Binding Affinity of DNA Repair Proteins. J Am Chem Soc. 2017;139(36):12784-92.

251.   Schrodinger, LLC. The PyMOL Molecular Graphics System, Version 1.8. 2015.

252.   Bertini I, Luchinat C, Parigi G, Pierattelli R. NMR Spectroscopy of Paramagnetic Metalloproteins. ChemBioChem. 2005;6(9):1536-49.

253.   Eyal E, Gerzon S, Potapov V, Edelman M, Sobolev V. The limit of accuracy of protein modeling: influence of crystal packing on protein structure. J Mol Biol. 2005;351(2):431-42.

254.   Diener JL. Complex conformations and crystal contacts. Nat Struct Biol. 2003;10(7):494; author reply -5.

255.   Layten M, Hornak V, Simmerling C. The open structure of a multi-drug-resistant HIV-1 protease is stabilized by crystal packing contacts. J Am Chem Soc. 2006;128(41):13360-1.

256.   Taylor P, Dornan J, Carrello A, Minchin RF, Ratajczak T, Walkinshaw MD. Two structures of cyclophilin 40: folding and fidelity in the TPR domains. Structure. 2001;9(5):431-8.

257.   Arunkumar AI, Stauffer ME, Bochkareva E, Bochkarev A, Chazin WJ. Independent and coordinated functions of replication protein A tandem high affinity single-stranded DNA binding domains. J Biol Chem. 2003;278(42):41077-82.

258.   Hopkins JB, Gillilan RE, Skou S. BioXTAS RAW: improvements to a free open-source program for small-angle X-ray scattering data reduction and analysis. J Appl Crystallogr. 2017;50(Pt 5):1545-53.

259.   Classen S, Hura GL, Holton JM, Rambo RP, Rodic I, McGuire PJ, et al. Implementation and performance of SIBYLS: a dual endstation small-angle X-ray scattering and macromolecular crystallography beamline at the Advanced Light Source. J Appl Crystallogr. 2013;46(Pt 1):1-13.

260.   Dyer KN, Hammel M, Rambo RP, Tsutakawa SE, Rodic I, Classen S, et al. High-throughput SAXS for the characterization of biomolecules in solution: a practical approach. Methods Mol Biol. 2014;1091:245-58.

261.   Tria G, Mertens HD, Kachala M, Svergun DI. Advanced ensemble modelling of flexible macromolecules using X-ray solution scattering. IUCrJ. 2015;2(Pt 2):207-17.

262.   Coloma J, Johnson RE, Prakash L, Prakash S, Aggarwal AK. Human DNA polymerase alpha in binary complex with a DNA:DNA template-primer. Sci Rep. 2016;6:23784.

263.   Baranovskiy AG, Babayeva ND, Suwa Y, Gu J, Pavlov YI, Tahirov TH. Structural basis for inhibition of DNA replication by aphidicolin. Nucleic Acids Res. 2014;42(22):14013-21.

264.   Jervis AJ, Crack JC, White G, Artymiuk PJ, Cheesman MR, Thomson AJ, et al. The O2 sensitivity of the transcription factor FNR is controlled by Ser24 modulating the kinetics of [4Fe-4S] to [2Fe-2S] conversion. Proc Natl Acad Sci U S A. 2009;106(12):4659-64.
265.   Artz JH, Zadvornyy OA, Mulder DW, King PW, Peters JW. Structural Characterization of Poised States in the Oxygen Sensitive Hydrogenases and Nitrogenases. Methods Enzymol. 2017;595:213-59.
266.   Williams JS, Kunkel TA. Ribonucleotides in DNA: origins, repair and consequences. DNA Repair (Amst). 2014;19:27-37.
267.   Foster PL, Niccum BA, Popodi E, Townes JP, Lee H, MohammedIsmail W, et al. Determinants of Base-Pair Substitution Patterns Revealed by Whole-Genome Sequencing of DNA Mismatch Repair Defective Escherichia coli. Genetics. 2018.
268.   Zhao H, Thienpont B, Yesilyurt BT, Moisse M, Reumers J, Coenegrachts L, et al. Mismatch repair deficiency endows tumors with a unique mutation signature and sensitivity to DNA double-strand breaks. Elife. 2014;3:e02725.
269.   Schirmer M, D'Amore R, Ijaz UZ, Hall N, Quince C. Illumina error profiles: resolving fine-scale variation in metagenomic sequencing data. BMC Bioinformatics. 2016;17:125.
270.   Sloan DB, Broz AK, Sharbrough J, Wu Z. Detecting Rare Mutations and DNA Damage with Sequencing-Based Methods. Trends Biotechnol. 2018;36(7):729-40.
271.   Koh KD, Balachander S, Hesselberth JR, Storici F. Ribose-seq: global mapping of ribonucleotides embedded in genomic DNA. Nat Methods. 2015;12(3):251-7, 3 p following 7.
272.   Clausen AR, Lujan SA, Burkholder AB, Orebaugh CD, Williams JS, Clausen MF, et al. Tracking replication enzymology in vivo by genome-wide mapping of ribonucleotide incorporation. Nat Struct Mol Biol. 2015;22(3):185-91.
273.   Daigaku Y, Keszthelyi A, Muller CA, Miyabe I, Brooks T, Retkute R, et al. A global profile of replicative polymerase usage. Nat Struct Mol Biol. 2015;22(3):192-8.
274.   Reijns MAM, Kemp H, Ding J, de Proce SM, Jackson AP, Taylor MS. Lagging-strand replication shapes the mutational landscape of the genome. Nature. 2015;518(7540):502-6.
275.   Kalimutho M, Bain AL, Mukherjee B, Nag P, Nanayakkara DM, Harten SK, et al. Enhanced dependency of KRAS-mutant colorectal cancer cells on RAD51-dependent homologous recombination repair identified from genetic interactions in Saccharomyces cerevisiae. Mol Oncol. 2017;11(5):470-90.
276.   Schwab RA, Niedzwiedz W. Visualization of DNA replication in the vertebrate model system DT40 using the DNA fiber technique. J Vis Exp. 2011(56):e3255.
277.   Mochizuki AL, Katanaya A, Hayashi E, Hosokawa M, Moribe E, Motegi A, et al. PARI Regulates Stalled Replication Fork Processing To Maintain Genome Stability upon Replication Stress in Mice. Mol Cell Biol. 2017;37(23).

APPENDIX

INVESTIGATION OF 5'-NTP RECRUITMENT BY p48 AND p58C

Introduction

While collecting the SAXS data discussed and analyzed in Chapter IV, we worked to

develop a clearer model for primase activity. During this process, we came to the important

insight- that primase engages in two distinct enzymatic processes (initial dinucleotide synthesis

and primer extension), which are completed in the context of two distinct configurations of

primase. In our model, p58C is a critical component of the active site for initial dinucleotide

synthesis but is not required for elongation activity and travels with the 5' end of the nascent

primer. This disengagement from the primase active site potentially positions p58C to modulate

primer termination through structural and electrochemical mechanisms. We find it fascinating

that if you define "active site" as a region where substrate molecules bind and undergo a

chemical reaction, primase arguably utilizes two distinct active sites, and believe that this may be

key to understanding why primase is the only replicative polymerase to initiate synthesis *de novo*

on ssDNA.

The idea of primase utilizing a biphasic mechanism is not new (37, 132), but the roles of

each subunit have remained obscure until relatively recently. To initiate primer synthesis,

primase requires catalytic metal ions, two NTPs, and DNA template (88). p58C is known to bind

template with modest affinity, and p48 has been crystallized with the catalytic metals and the 3'-

NTP (36, 112, 115, 143, 144). Based on biochemical assays with p48 variants, it was proposed

that the R304 residue in p48 recruits the 5'-NTP to the catalytic site (136). However, while

152

crystal structures of p48 with catalytic metals include the triphosphate of the 3'-NTP, the 5'-NTP

has been absent in structural analyses of primase (143, 144). Initially, we hypothesized that a

conformational changed triggered by template binding may be required for 5'-NTP binding to

p48 (143). However, recent structural and biochemical studies of p58C with primed substrate

suggested that p58C may interact with the 5'-NTP (Figure A.1), and it has been proposed that the

previous results with R304X p48 were caused by protein misfolding (137, 145). This hypothesis

was not confirmed by any experimental evidence, such as site-directed mutagenesis or structural

characterization of the p48 mutants, and implications for initiation of primase activity were not

explored beyond speculation (145).
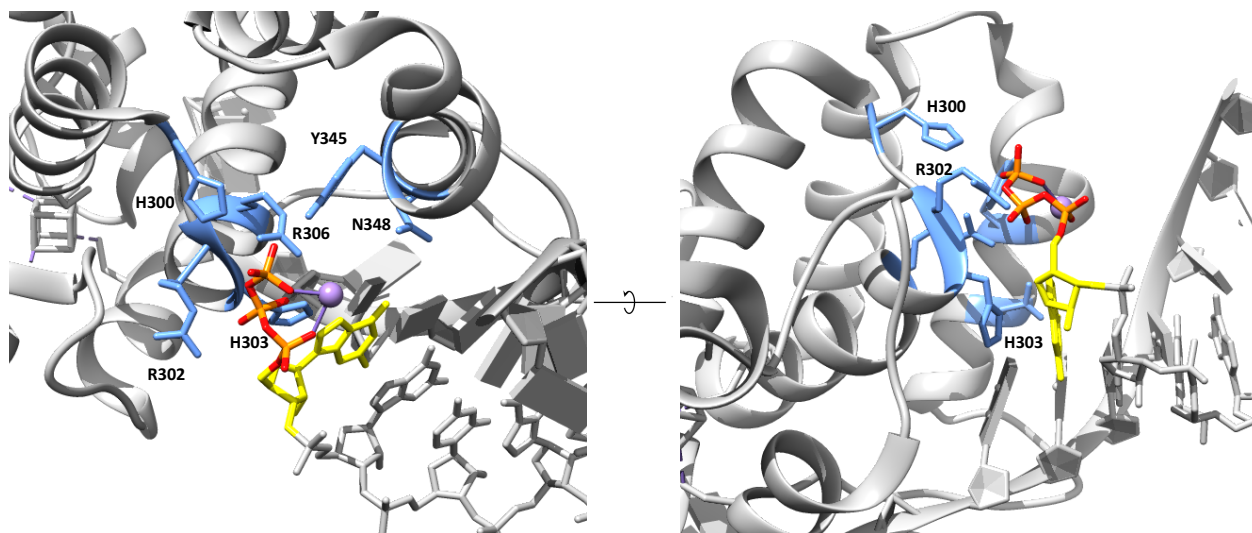
5′-NTP recruitment to initiation active site. p58C and primed substrate (PDB ID: 5F0S) are
shown in grey, with residues that are implicated in 5'-NTP recruitment colored blue and the 5'-
GTP in yellow, orange, and red.

We hypothesize that p58C is involved in stabilizing the 5'-NTP in the active site for

initial dinucleotide synthesis through a series of weak interactions that are only completed by

additional interactions in p48, including R304. To test this hypothesis, I mutated residues that

appear to interact with the 5'-NTP in the crystal structure of p58C bound to primed template and

developed a method for testing the affinity of these mutants for ATP. Additionally, I generated

R304A p48 to assess whether this mutation causes protein misfolding. Here, I present my

structural characterization of these mutants and NMR studies with mutant p58C.

Methods

p58C variant generation

      p58C variants were created with a Q5 site-directed mutagenesis kit from New England

Biolabs, using the p58C$_{266-464}$ construct described in Chapter III for template. Primers are

described in Table A.1. 10 ng of plasmid was used in the exponential amplification, which was

completed as described in the Q5 Site-Directed Mutagenesis Kit manual (New England Biolabs),

except that the final extension time was extended to 5 minutes. The KLD reaction was also

completed as described in the manual, except that 4 µL of PCR product was used in the reaction

and the incubation time was increased to 20 minutes. 10 µL of the KLD reaction product was

transformed into XL1-Blue cells. DNA was extracted from individual colonies with a Qiagen

QIAprep Spin Miniprep Kit. R304A p48 was generated using a QuikChange mutagenesis kit

(Agilent) with full-length p48 plasmid as the template and the primers described in Table A.1. 25

ng of template was used in the exponential amplification, which was completed as described in

the QuikChange manual (Agilent). DpnI digestion of the PCR reaction products was completed

as described in the manual, except that the digestion time was extended to two hours. 10 µL of

DpnI digestion reaction product was transformed into XL1-Blue cells. DNA was extracted from

individual colonies with a Qiagen QIAprep Spin Miniprep Kit. In all cases, mutagenesis was

confirmed through plasmid sequencing (GENEWIZ, LLC).

Table A.1	
R304A p48 Forward Primer	5'–GCTCCAGTACTGTTTTCCAGCGCTGGATATCAATGTCAGC–3'
R304A p48 Reverse Primer	5'–GCTGACATTGATATCCAGCGCTGGAAAACAGTACTGGAGC–3'
H300A p58C Forward Primer	5'–GCGGGAAAATCACGCTCTTCGTCATGGAGGCC–3'
H300A p58C Reverse Primer	5'–AAGGCTTTATGTAACTGACGCATGCAAGG–3'
H300D p58C Forward Primer	5'–GCGGGAAAATCACGATCTTCGTCATGGAGGCC–3'
H300D p58C Reverse Primer	5'–AAGGCTTTATGTAACTGACGCATGCAAGG–3'
R302A Forward Primer	5'–GCGGGAAAATCACCATCTTGCTCATGGAGGCCGAAT–3'
R302A Reverse Primer	5'–AAGGCTTTATGTAACTGACGCATGCAAGG–3'
R302E Forward Primer	5'–CCATCTTCGTCATGGAGGCGCAATGCAGTATGGCC–3'
R302E Reverse Primer	5'–AAGGCTTTATGTAACTGACGCATGCAAGG–3'
H303A p58C Forward Primer	5'–GCGGGAAAATCACCATCTTCGTGCTGGAGGCCGAAT–3'
H303A p58C Reverse Primer	5'–AAGGCTTTATGTAACTGACGCATGCAAGG–3'
H303D p58C Forward Primer	5'–GCGGGAAAATCACCATCTTCGTGATGGAGGCCGAAT–3'
H303D p58C Reverse Primer	5'–AAGGCTTTATGTAACTGACGCATGCAAGG–3'
R306A Forward Primer	5'–CCATCTTCGTCATGGAGGCGCAATGCAGTATGGCC–3'
R306A Reverse Primer	5'–TGATTTTCCCGCAAGGCTTTATGTAACTGACG–3'
R306E Forward Primer	5'–CCATCTTCGTCATGGAGGCGAAATGCAGTATGGCC–3'
R306E Reverse Primer	5'–TGATTTTCCCGCAAGGCTTTATGTAACTGACG–3'

Table A.1. Primers used in this study.

Protein expression and purification

p58C constructs were expressed and purified as previously described (218, 235). In short, plasmid DNA was transformed into BL21 (DE3) cells (Novagen) and cultured in Terrific Broth media at 37 °C to an OD_{600} of 0.6-0.8, when flasks were moved to a 21 °C incubator with shaking. After 30 minutes, ferric citrate and ammonium ferrous citrate were added to a final concentration of 0.1 mg/mL, and isopropyl 1-thio-β-D-galactopyranoside was added to a final concentration of 0.5 mM. Protein was expressed at 21 °C overnight. The same growth protocol was used to generate ^{15}N-labeled p58C, except that cells were cultured in M9 media supplemented with ^{15}N-labled ammonium chloride (Cambridge Isotopes) and expressed overnight at 25 °C. The same purification scheme was used for both unlabeled and ^{15}N-labeled protein (218). In short, protein was first purified by nickel affinity chromatography (Amersham Biosciences). The 6xHis tag was cleaved with H3C protease and the protein was dialyzed into a

low-imidazole buffer (218, 235). The protein was repassed over the nickel column to remove the H3C protease and uncleaved protein. A heparin column was used as the final purification step to remove residual contaminants (218, 235).

Circular dichroism (CD) spectroscopy

p58C samples were buffer exchanged into 20 mM K_2HPO_4 (pH 7.2) and diluted to a concentration of 0.3 mg/mL. The far-UV CD spectrum over the range 200–260 nm was acquired at room temperature using a Jasco J-810 spectrophotometer. Each spectrum is the average of three scans acquired with a scanning rate of 50 nm/min and data pitch of 1 nm. Prior to generating the overlay in Graphpad Prism 7, the p58C variant CD spectra were scaled to the $p58C_{266-464}$ spectrum by averaging the values of the $CD_{208}(p58C_{266-464})/CD_{208}(p58C_{variant})$ and $CD_{222}(p58C_{266-464})/CD_{208}(p58C_{variant})$ ratios to generate a scaling factor, then multiplying the entire CD spectrum by this scaling factor. p48 samples were buffer exchanged into 10 mM K_2HPO_4 (pH 7.0) and diluted to a concentration of 10 μM. The far-UV CD spectrum over the range 200–260 nm was acquired at room temperature using a Jasco J-810 spectrophotometer. Each spectrum is the average of three scans acquired with a scanning rate of 50 nm/min and data pitch of 1 nm. Prior to generating the overlay in Graphpad Prism 7, the R304A p48 CD spectrum was scaled to the wild-type p48 spectrum using a process similar to that used for the p58C variant samples. Thermal melts were collected on the same instrument by measuring CD signal at 222 nm as the temperature was raised from 20 to 60 °C over the course of an hour.

NMR spectroscopy

Two-dimensional ^{15}N-^1H heteronuclear single quantum coherence (HSQC) spectra were

acquired using a Bruker AV-III 800 MHz spectrometer equipped with a CPTCI single-gradient cryoprobe. To collect the spectra shown in Figures A.3-6, ^{15}N-enriched p58C constructs at a concentration of 75 µM were prepared in a buffer containing 20 mM Tris (pH 7.2), 75 mM NaCl, 2 mM DTT, and 5% (v/v) D_2O. For all other HSQC data shown in this appendix, ^{15}N-enriched p58C constructs at a concentration of 60 µM were prepared in a buffer containing 20 mM MES (pH 6.5), 50 mM NaCl, 2 mM DTT, and 5% (v/v) D_2O. Spectra were acquired at 25 °C with 128 and 2,048 points in the ^{15}N and ^{1}H dimensions, respectively. 32 scans were recorded in the direct (^{1}H) dimension for each point sampled in the indirect (^{15}N) dimension. Data were processed in Topspin (Bruker) and analyzed with Sparky (University of California).

^{31}P 1D spectra of ATP in the presence of p58C variants were collected at 25 °C using a Bruker AV-III 600 MHz spectrometer equipped with a CPQCI ^{1}H-^{31}P/^{13}C/^{15}N/D Z-GRD cryoprobe. Samples were prepared a buffer containing in 20 mM MES (pH 6.5), 50 mM NaCl, 2 mM DTT, and 5% (v/v) D_2O. 1048 scans were recorded in the direct (^{31}P) dimension for each spectrum. Data were processed and analyzed in Topspin (Bruker).

Results and Discussion

In the 1999 paper of Kuchta and coworkers demonstrating that mutating R304 impaired the ability of p48 to recruit the 5' nucleotide (136), R304A was the mutant with the highest impact (out of R304A, R304K, and R304Q) and therefore of particular interest as a potential positive control for potential primase activity assays to investigate 5'-NTP recruitment by p58C. In Tahirov and coworker's review of structural characterization of the primosome, they stated that this phenomenon was likely due to R304 mutations destabilizing p48 (145). To evaluate whether these mutations are in fact destabilizing, I generated these mutants in full-length p48 and

evaluated the thermal stability of R304A p48 with CD. As we can see in Figure A.2B, R304A does not appear to be significantly less stable than WT primase. In fact, the apparent Tm of R304A p48 (44.5 °C) is one degree higher than that of WT p48 (43.5 °C). Therefore, it seems likely that the disruption of primase activity observed with the R304 p48 mutant of primase is, in fact, due to the disruption of an interaction between p48 and the 5' NTP. Crystallization of this mutant, coupled with analysis of the ^{15}N-^{1}H HSQC spectrum would provide further insights into the effect of this mutation on the tertiary structure of p48.
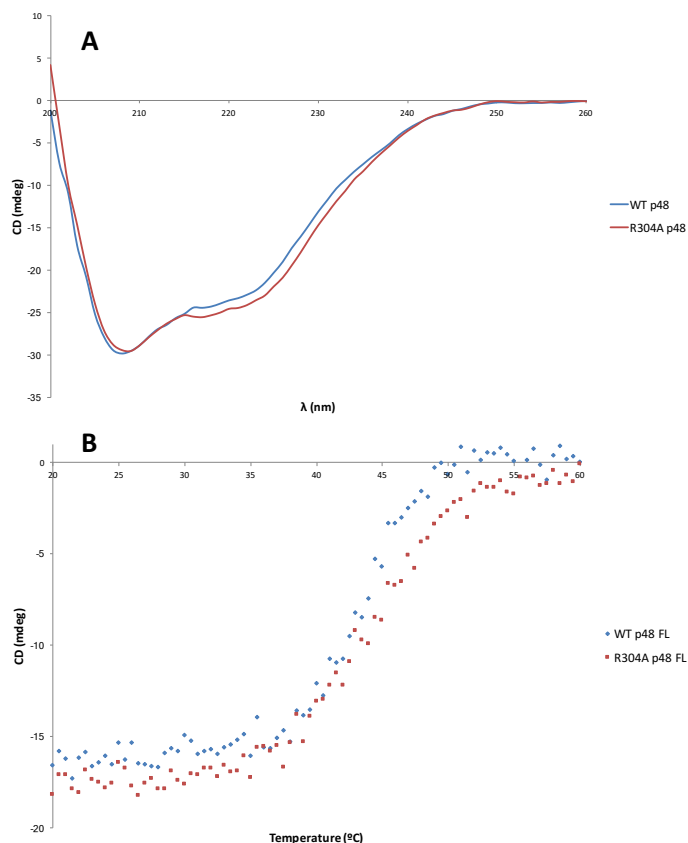
Figure A.2. R304A p48 is not less thermally stable than wild-type 48. A. CD scan of 5 μM WT (blue) and R304A (red) p48 collected in 10 mM KPO4 pH 7.0. B. Thermal melt of 5 μM WT (blue) and R304A (red) p48 collected in 10 mM KPO4 pH 7.0. To perform this melt, CD was measured at 222 nm as the temperature was raised from 20 to 60 °C over the course of an hour.

Turning to the role of p58C in 5'-NTP recruitment, I initially hypothesized that p58C

binds the 5' NTP via a DNA-mediated interaction and brings both the ssDNA template and the

5' NTP to the catalytic site. Given that we had previously used NMR to identify a weak

interaction between NTPs and p58 (Figure A.3), I attempted to determine whether DNA-bound

p58C interacted more strongly with ATP than apo p58C. Ultimately, I found that the addition of

$MgCl_2$ appeared to inhibit the p58C-ATP interaction, both with and without saturating ssDNA

template. This suggests that p58C may not appropriately bind the Mg•ATP complex (**Figure

A.4-6**).
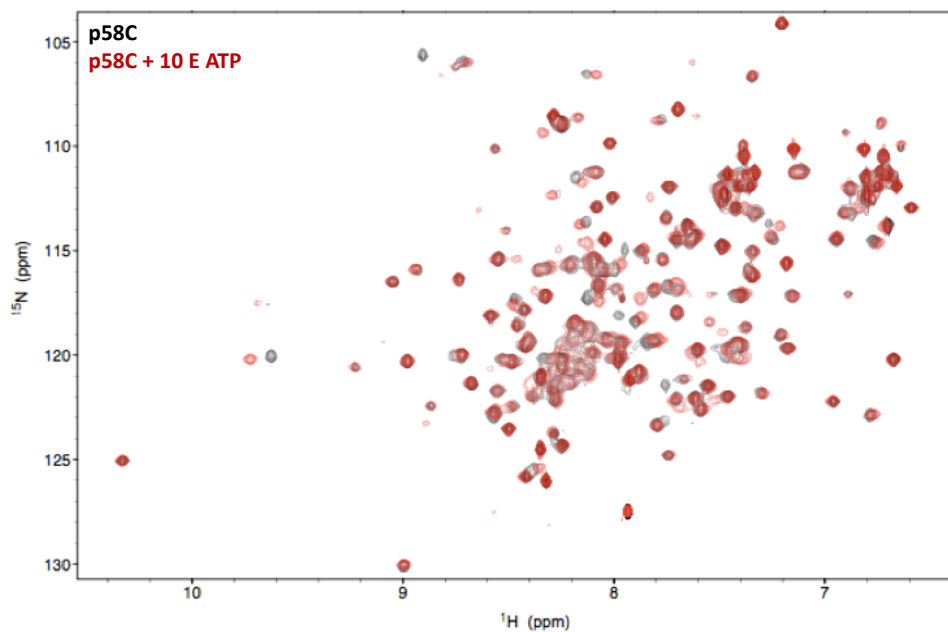
Figure A.3. p58C binds NTPs. ^{15}N-^1H HSQC spectra collected with free p58C (black) and in the presence of 10 molar equivalents of ATP (red). ATP binding is reflected by chemical shift perturbations in the red spectrum compared to the black.
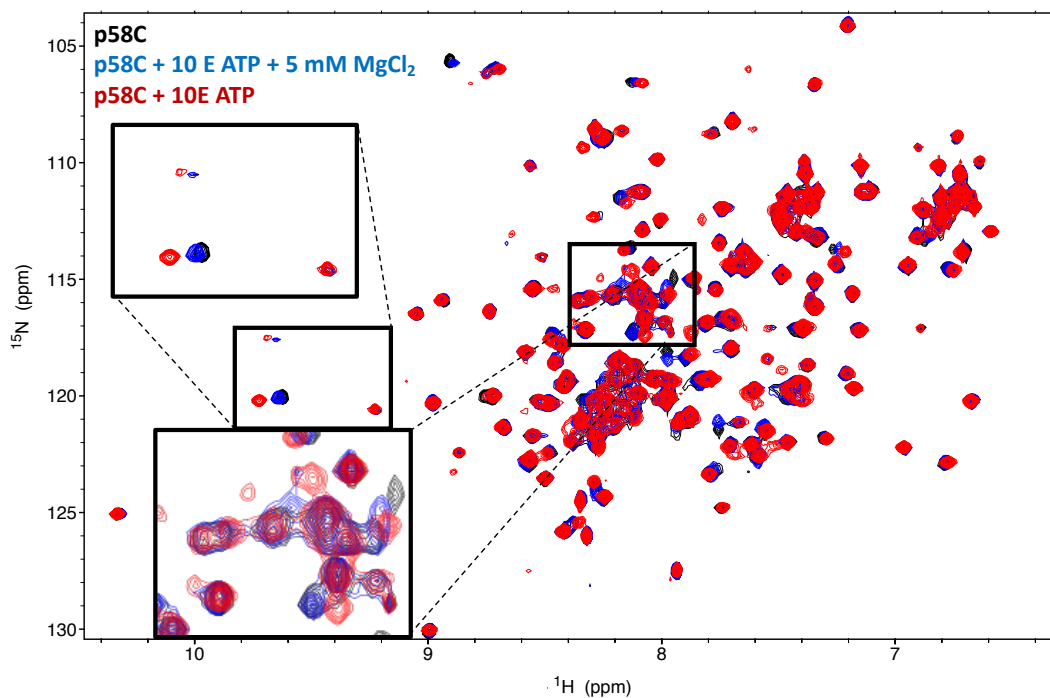
Figure A.4. Addition of Mg²⁺ inhibits ATP binding by p58C. ¹⁵N-¹H HSQC spectra collected with free p58C (black), in the presence of 10 molar equivalents of ATP (red), and in the presence of 10 molar equivalents of ATP and 5 mM MgCl₂ (blue). Reduced chemical shift perturbations indicate a weaker binding interaction.
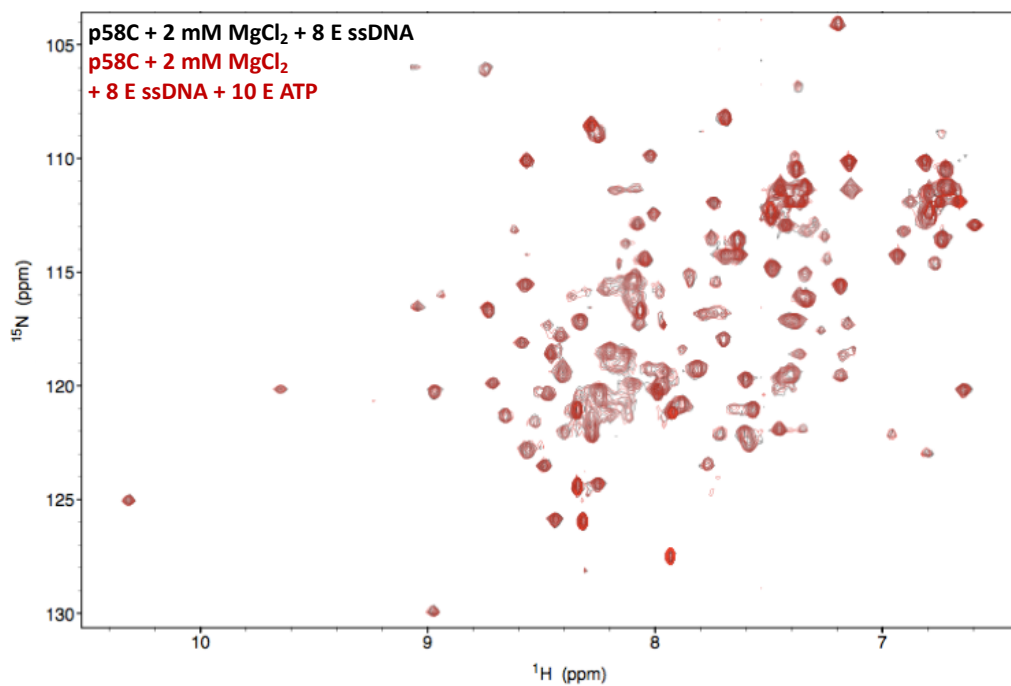
Figure A.5. Addition of ssDNA does not improve ATP binding by p58C. ^{15}N-^{1}H HSQC spectra collected with p58C in the presence of 2 mM MgCl$_2$ and 8 molar equivalents ssDNA (a (TC)$_{24}$ repeat) without (black) and with 10 molar equivalents of ATP (red). Loss of chemical shift perturbations indicates that p58C does not bind ATP under these conditions.
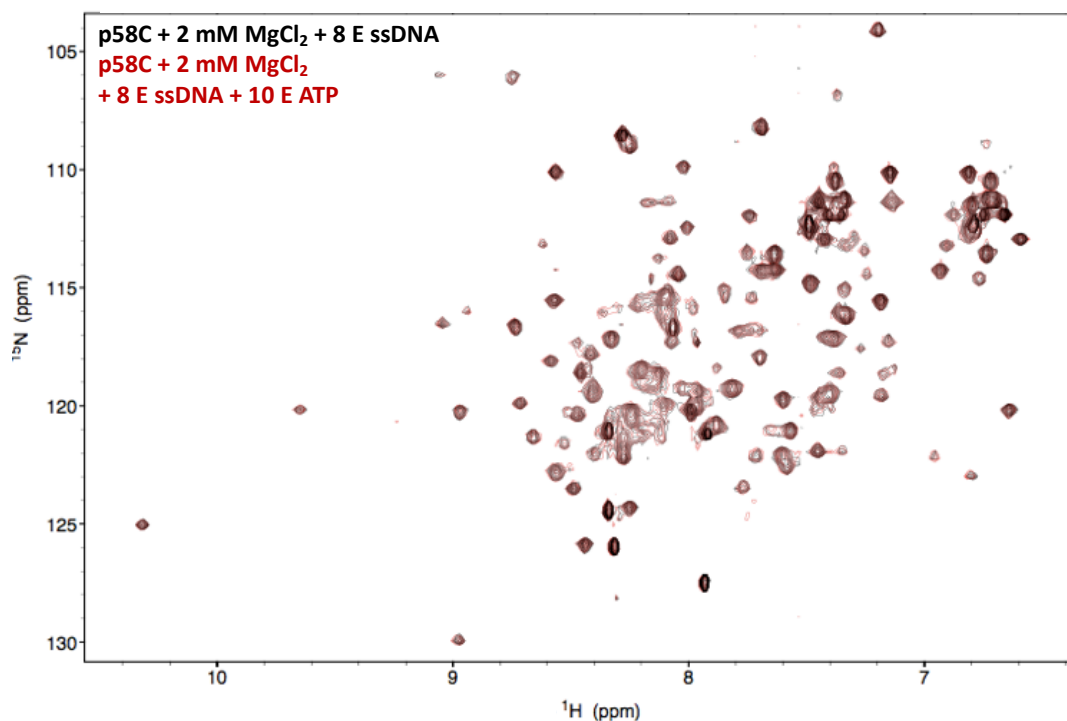
Figure A.6. Addition of ssDNA does not improve ATP binding by p58C. Flipped version of overlay shown in Figure A.4 (i.e. black spectrum is overlaid on top of the red spectrum, instead of the other way around). As in Figure A.4, loss of chemical shift perturbations indicates that p58C does not bind ATP under these conditions.

My next hypothesis was that p58C brings ssDNA template to p48 pre-loaded with catalytic metal and the 3' NTP, forming a pre-initiation complex that the 5' NTP binds to trigger the priming reaction. To begin testing this hypothesis, I generated alanine and charge reversal mutants of H300, R302, H303, and R306, the residues proposed to recruit the 5'-NTP. Structural characterization of these mutants by ^{15}N-^{1}H HSQC NMR revealed that these p58C variants are structurally intact in solution (Figures A.7-14).
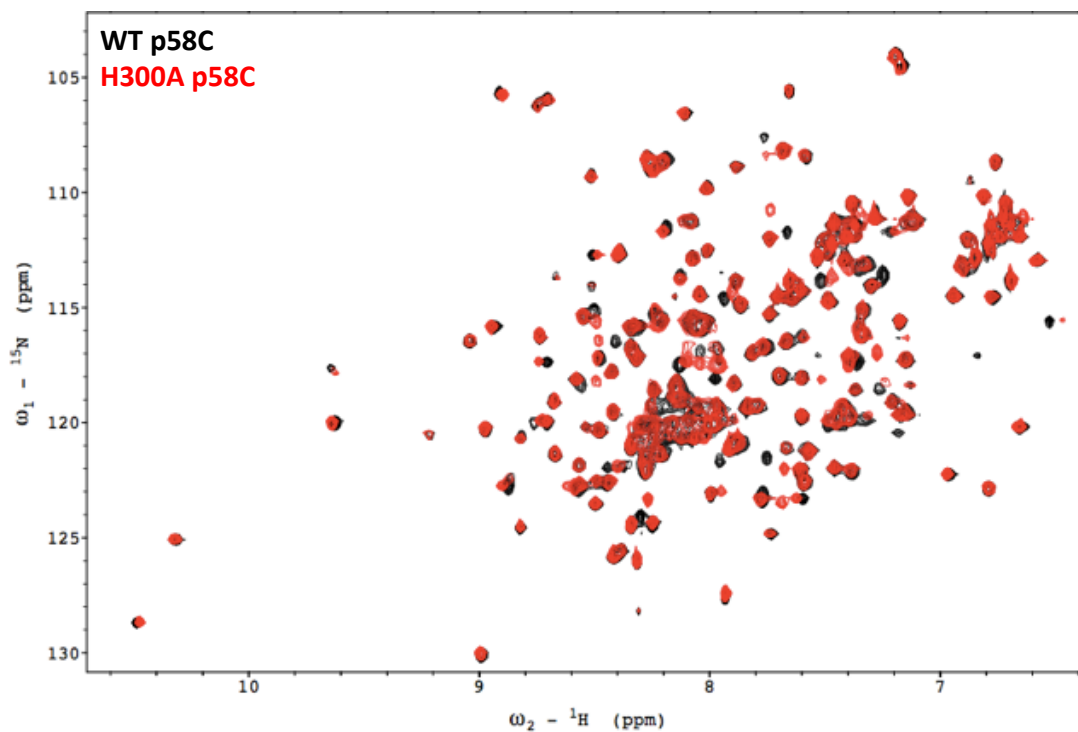
Figure A.7. Comparison of tertiary structure of wild-type and H300A p58C. The ^{15}N-^{1}H HSQC spectra of wild-type (black) and H300A (red) p58C were acquired at 25 °C on a Bruker AV-III spectrometer operating at 800 MHz. The samples contained 200 µM protein in a buffer containing 20 mM MES (pH 6.5), 50 mM NaCl, 2 mM DTT, and 5% ^{2}H$_2$O.
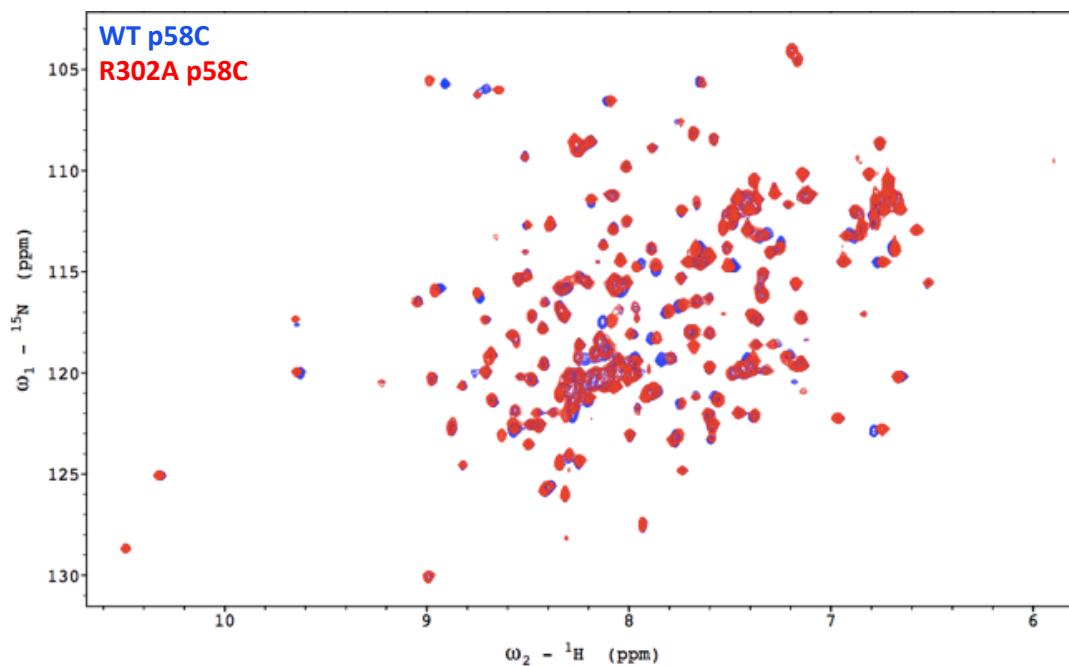
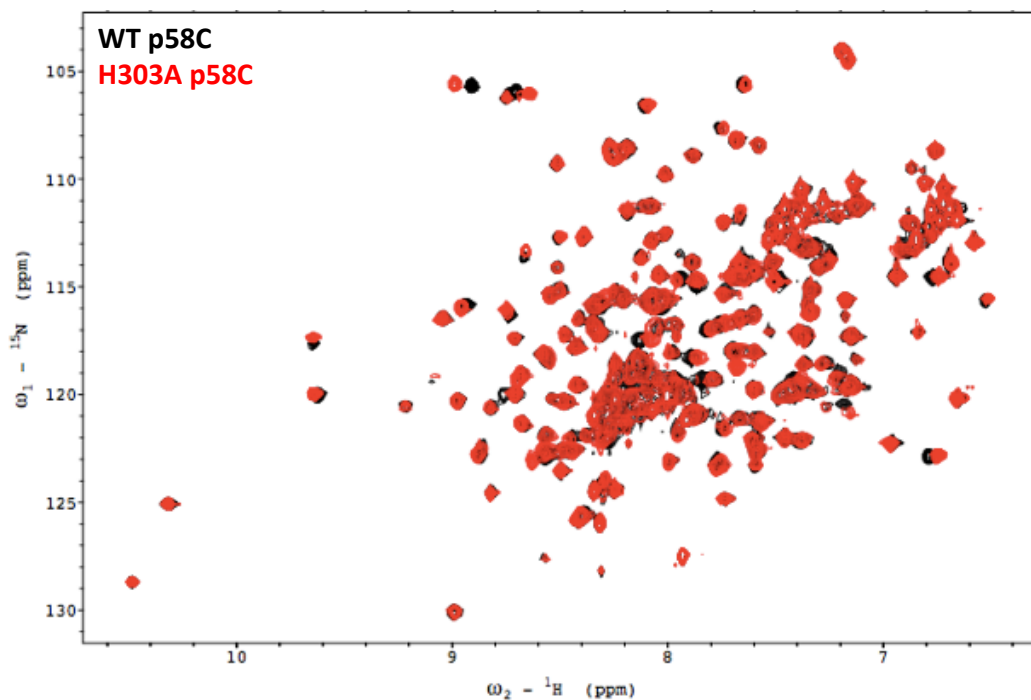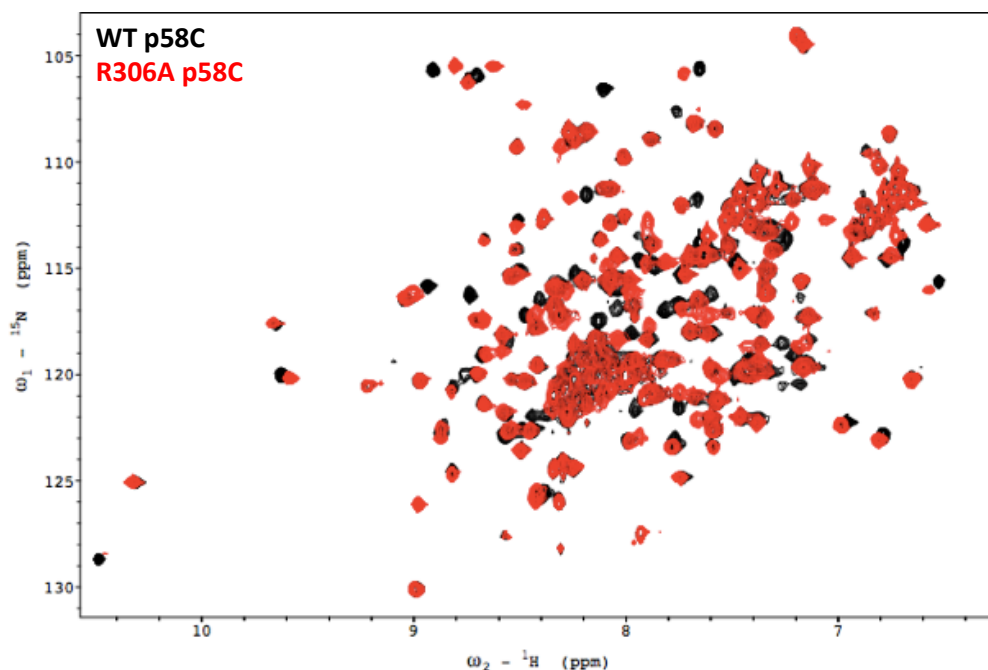**Figure A.8. Comparison of the tertiary structure of wild-type and R302A p58C.** The ^{15}N-^{1}H HSQC spectra of wild-type (blue) and H300A (red) p58C were acquired at 25 °C on a Bruker AV-III spectrometer operating at 800 MHz. The samples contained 200 μM protein in a buffer containing 20 mM MES (pH 6.5), 50 mM NaCl, 2 mM DTT, and 5% ^{2}H$_2$O.
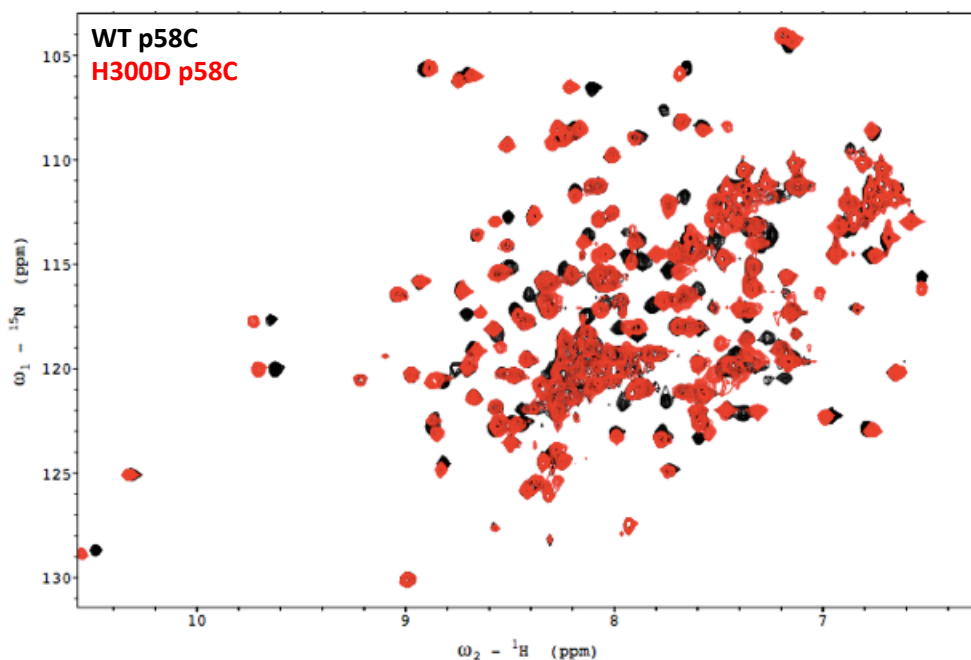
Figure A.9. Comparison of the tertiary structure of wild-type and H303A p58C. The ^{15}N-^{1}H HSQC spectra of wild-type (black) and H300A (red) p58C were acquired at 25 °C on a Bruker AV-III spectrometer operating at 800 MHz. The samples contained 200 µM protein in a buffer containing 20 mM MES (pH 6.5), 50 mM NaCl, 2 mM DTT, and 5% ^{2}H$_2$O.
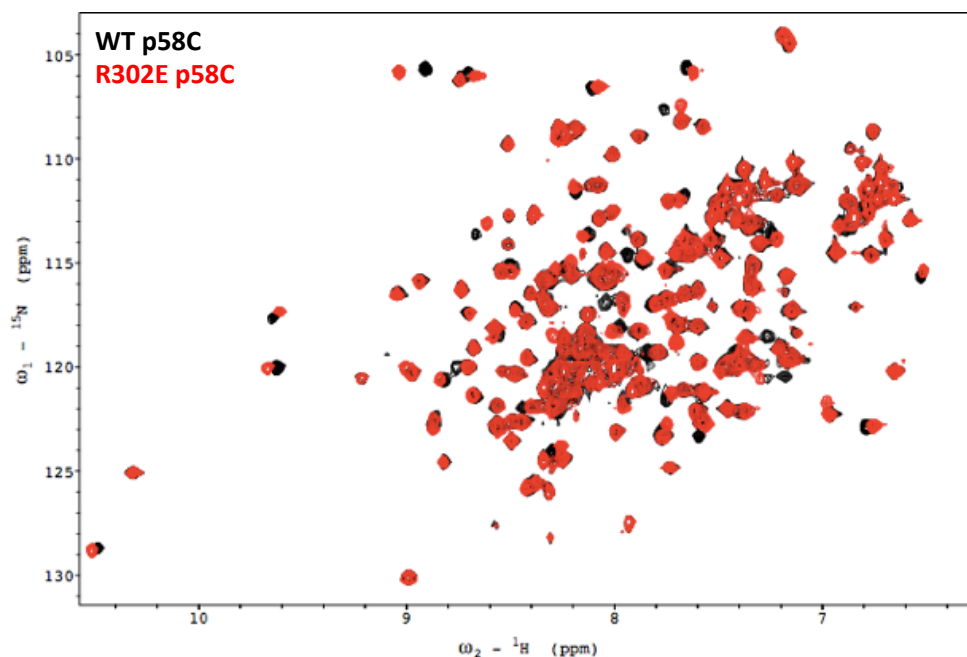
Figure A.10. Comparison of the tertiary structure of wild-type and R306A p58C. The ^{15}N-^{1}H HSQC spectra of wild-type (black) and H300A (red) p58C were acquired at 25 °C on a Bruker AV-III spectrometer operating at 800 MHz. The samples contained 200 µM protein in a buffer containing 20 mM MES (pH 6.5), 50 mM NaCl, 2 mM DTT, and 5% [2]H$_2$O.
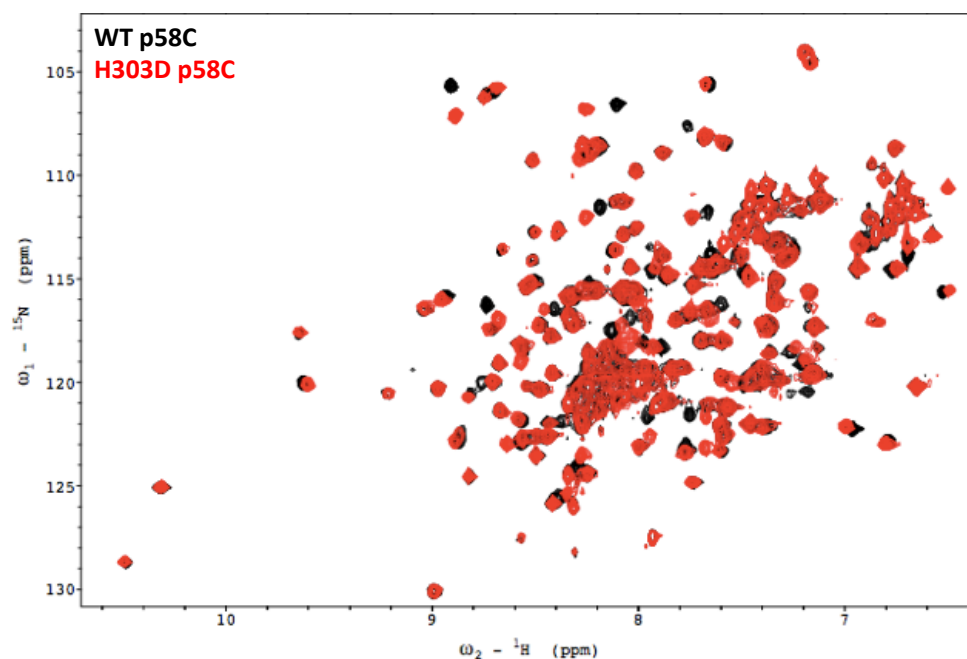

Figure A.11. Comparison of the tertiary structure of wild-type and H300D p58C. The ^{15}N-^{1}H HSQC spectra of wild-type (black) and H300A (red) p58C were acquired at 25 °C on a Bruker AV-III spectrometer operating at 800 MHz. The samples contained 200 µM protein in a buffer containing 20 mM MES (pH 6.5), 50 mM NaCl, 2 mM DTT, and 5% [2]H$_2$O.

Figure A.12. Comparison of the tertiary structure of wild-type and R302E p58C. The ^{15}N-^{1}H HSQC spectra of wild-type (black) and H300A (red) p58C were acquired at 25 °C on a Bruker AV-III spectrometer operating at 800 MHz. The samples contained 200 µM protein in a buffer containing 20 mM MES (pH 6.5), 50 mM NaCl, 2 mM DTT, and 5% ^2H_2O.
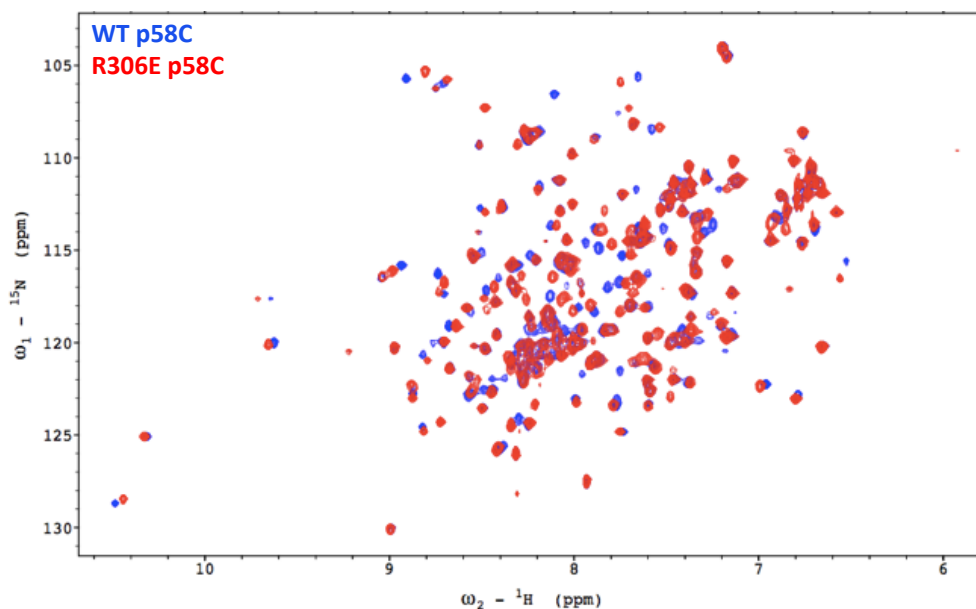

Figure A.13. Comparison of the tertiary structure of wild-type and H303D p58C. The ^{15}N-^{1}H HSQC spectra of wild-type (black) and H300A (red) p58C were acquired at 25 °C on a Bruker AV-III spectrometer operating at 800 MHz. The samples contained 200 µM protein in a buffer containing 20 mM MES (pH 6.5), 50 mM NaCl, 2 mM DTT, and 5% ^2H_2O.

Figure A.14. Comparison of the tertiary structure of wild-type and R302E p58C. The ^{15}N-^{1}H HSQC spectra of wild-type (blue) and H300A (red) p58C were acquired at 25 °C on a Bruker AV-III spectrometer operating at 800 MHz. The samples contained 200 µM protein in a buffer containing 20 mM MES (pH 6.5), 50 mM NaCl, 2 mM DTT, and 5% ^{2}H$_2$O.

Because generating sufficient ^{15}N-labeled p58C samples to conduct ATP titrations with eight p58C mutants could become expensive, I developed an assay in which, instead of monitoring chemical shift perturbations in the p58C HSQC, I evaluated chemical shift perturbations of the phosphate atoms of ATP in the presence of p58C. This was much cheaper because the magnetically susceptible ^{31}P isotope is 100% abundant and has a high gyro-magnetic ratio so it can be monitored directly via ^{31}P NMR. Additionally, because there are so few P atoms, it has a dramatically simplified NMR spectrum (Figure A.15). In ^{31}P-monitored titrations of p58C into ATP, I initially found that my p58C samples were contaminated with ATPase activity. This activity was evident by the disappearance of the γ-phosphate of ATP and emergence of the inorganic phosphate peak at ~1 ppm (Figure A.16). Adding a heparin column purification to the p58C purification scheme eliminated this activity (Figure A.17).
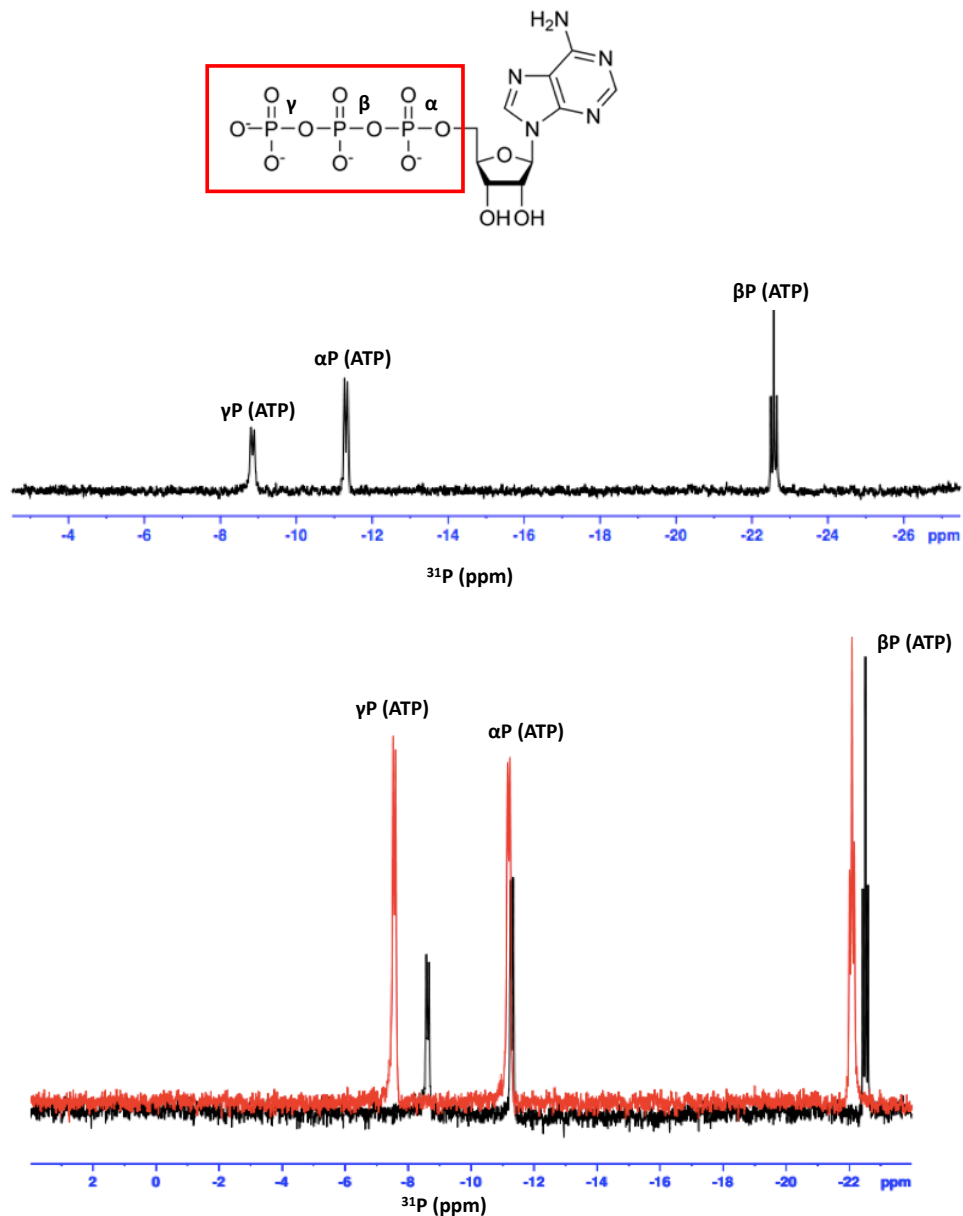
Figure A.15. ^{31}P spectrum of ATP. Top: ^{31}P spectrum of 300 µM ATP was collected on a 600 MHz Bruker spectrometer. Samples were prepared in a buffer containing 20 mM MES (pH 6.5), 50 mM NaCl, 2 mM DTT, and 5% ^{2}H$_2$O. Bottom: Overlay of 250 µM free ATP (black) and 350 µM ATP in the presence of 500 µM ATP (red), collected after removal of ATPase activity.
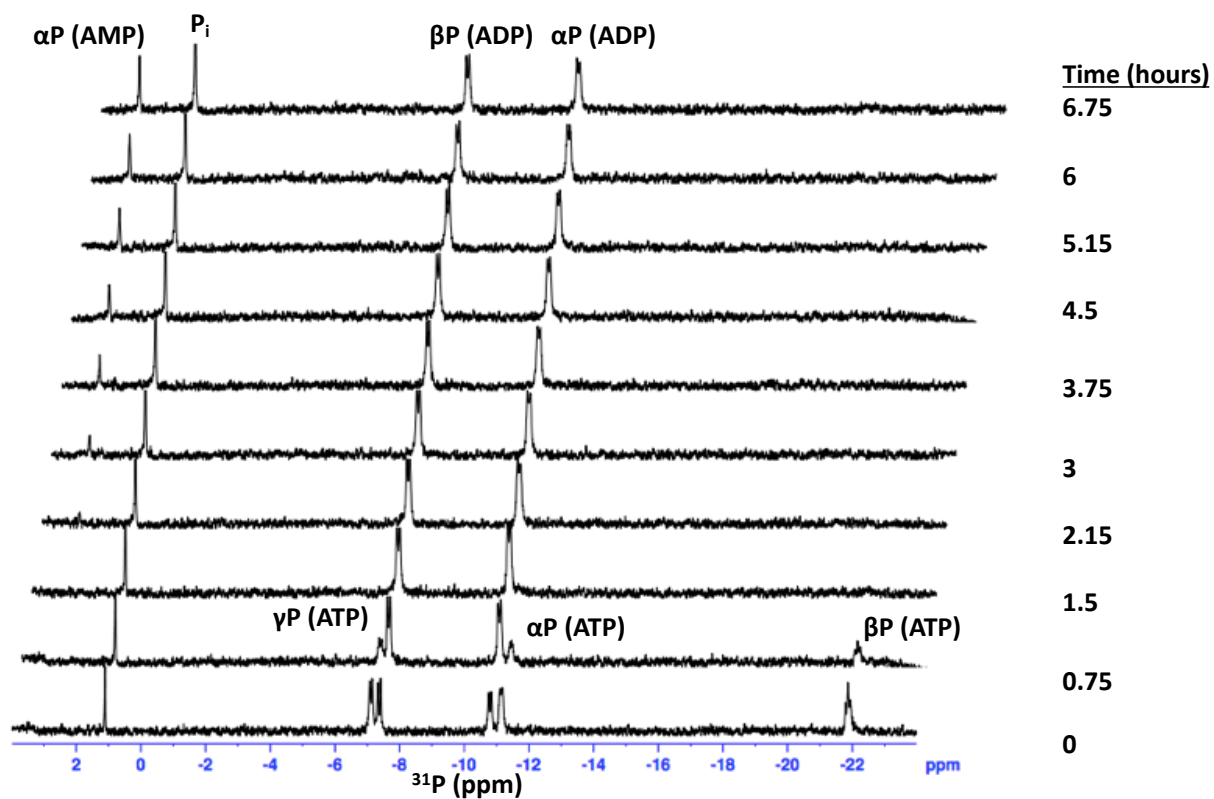
Figure A.16. Detecting contamination of p58C samples by ATPase activity. ^{31}P NMR spectra of 300 μM ATP in the presence of 450 μM p58C collected over the course of 7 hours revealed emergence of peaks corresponding with inorganic phosphate and AMP formation.
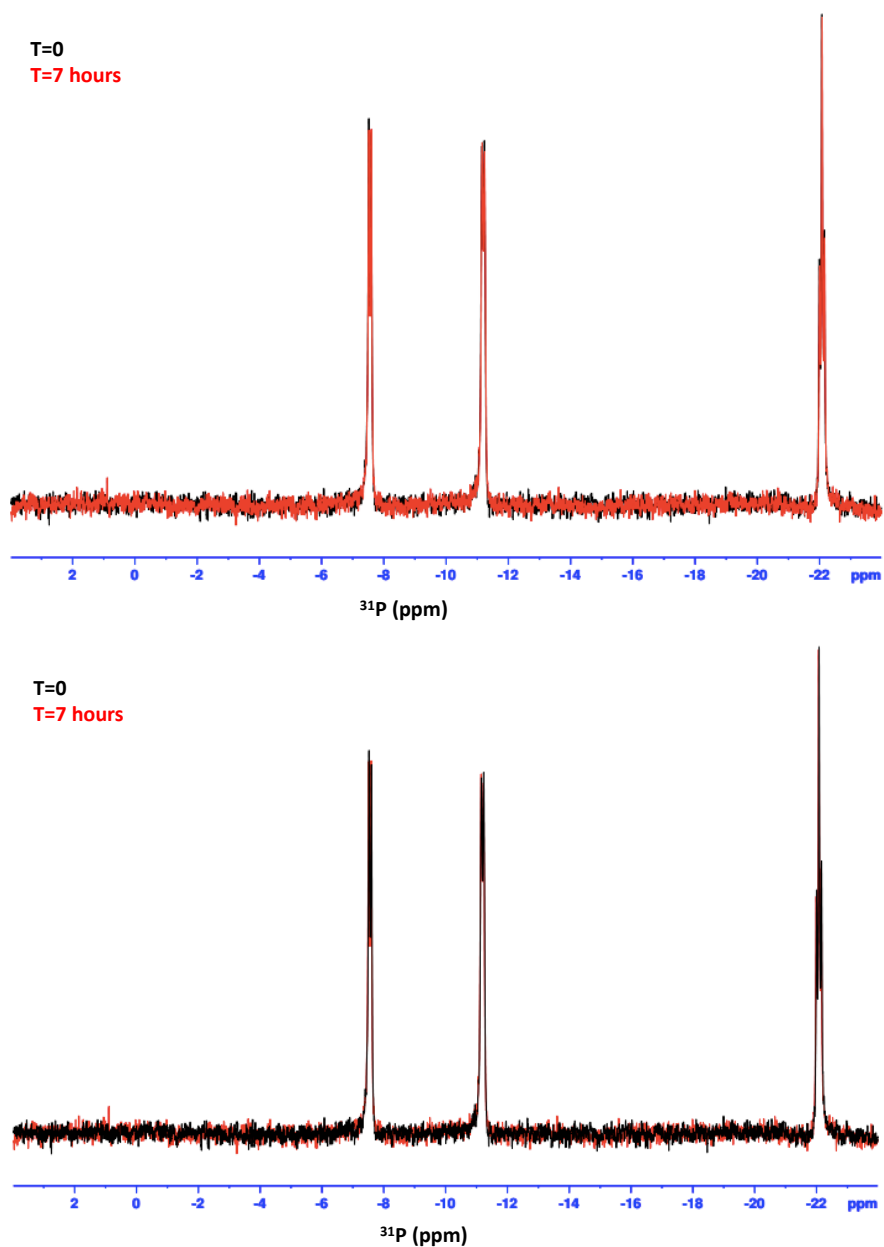
Figure A.17. Evidence that ATPase contamination was eliminated. Spectra collected at 0 (black) and 7 (red) hours after optimizing the p58C purification strategy revealed no ATPase activity.

After assay optimization was complete, I found that mutation of these residues to either alanine or a charge reversal mutant resulted in reduced chemical shift perturbations relative to wild-type p58C (see **Graph A.1**).

% Chemical Shift Perturbation Relative to Free ATP
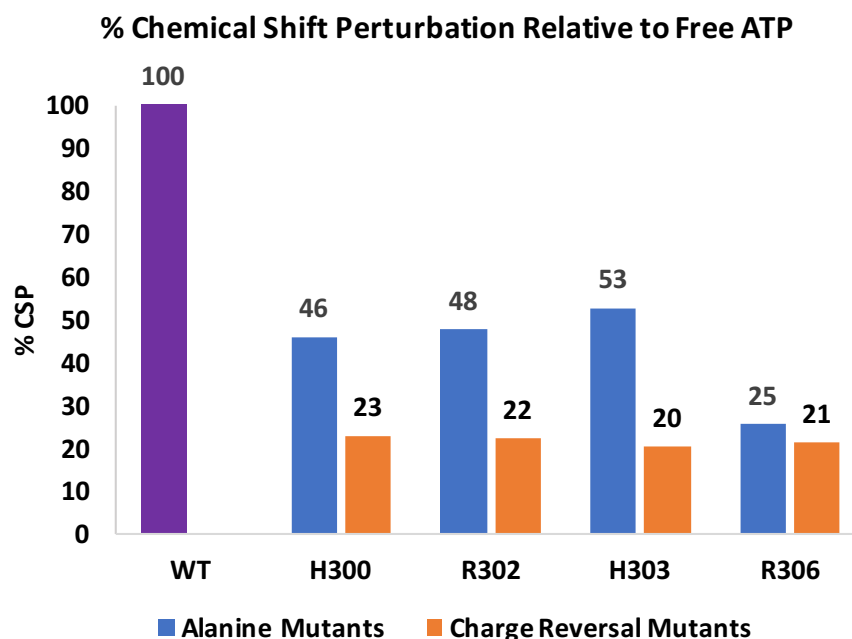
Graph A.1. Chemical shift perturbations of ATP induced by the binding of WT and p58C variants. Perturbations of the ^{31}P chemical shift of the γ-phosphate expressed as a percentage of the perturbation observed with wild-type p58C. The γ-phosphate was chosen because it displayed the greatest chemical shift differences between WT and mutants (**Figure A.15**).

Further directions for this project include completion of these binding studies with ATP and the other p58C variants combined with an investigation of the ability of the mutant primases to initiate primer synthesis and elongate a primed substrate. I expect that these mutations will result in reduced ATP binding affinity and initiation activity, with less effect on primer extension activity. Multiple mutations may be needed to see a significant effect, for example if all sites contribute equally to ATP binding. As these residues appear to also interact with the triphosphate of the 5'-NTP, wild-type primase is expected to exhibit higher primer extension activity on a triphosphorylated primed substrate relative to a monophosphorylated primed substrate, whereas mutants that are deficient in binding ATP will not. Ultimately, combination of this biochemical study with further structural characterization of the initiation active site in full-length primase will enable a more complete understanding of initiation of primer synthesis.