The Biochemical Characterization of *Drosophila melanogaster* RecQ4 Helicase

by

Christopher Lee Capp

Department of Biochemistry
Duke University

Date:_____

Approved:

_____
Tao-shih Hsieh, Supervisor

_____
Michael D. Been

_____
David N. Beratan

_____
Arno L. Greenleaf

_____
Kenneth N. Kreuzer

Dissertation submitted in partial fulfillment of
the requirements for the degree of Doctor of Philosophy in the Department of
Biochemistry in the Graduate School
of Duke University

2011

ABSTRACT

The Biochemical Characterization of *Drosophila melanogaster* RecQ4 Helicase

by

Christopher Lee Capp

Department of Biochemistry
Duke University

Date:_____

Approved:

_____
Tao-shih Hsieh, Supervisor

_____
Michael D. Been

_____
David N. Beratan

_____
Arno L. Greenleaf

_____
Kenneth N. Kreuzer

An abstract of a dissertation submitted in partial
fulfillment of the requirements for the degree
of Doctor of Philosophy in the Department of
Biochemistry in the Graduate School
of Duke University

2011

Copyright by
Christopher Lee Capp
2011

Abstract

RecQ4, a member of the conserved RecQ family of helicases, is involved in replication and associated with several clinical syndromes. Although biologically important, the biochemistry of RecQ4 has remained elusive. We have expressed and purified *Drosophila melanogaster* RecQ4 from a baculovirus expression system. Biochemical characterization of the helicase, ATP hydrolysis, annealing, and binding activities of the enzyme has been performed, using native and non-native gel electrophoresis and thin layer chromatography, among other techniques. These reveal that RecQ4 is a 3' to 5' helicase that is stimulated by the presence of single-stranded DNA 3' of the duplex DNA region to be unwound. The enzyme is also capable of annealing complementary DNA strands, though this is inhibited by AMPPNP, a non-hydrolyzable analog of ATP. RecQ4 also forms a stable complex with single-stranded DNA in the presence of AMPPNP. We argue that the helicase activity of RecQ4 is important to the process of DNA replication. This leads to the conclusion that two helicases, RecQ4 and the Mcm2-7 complex, are involved in replication. The manner of their simultaneous involvement is not intuitive, and so models by which the two enzymes may cooperate are discussed.

Dedication

This dissertation is dedicated to my parents, Randolph and Nancy Capp, for whose continued support and prayers I am extremely grateful.

Contents

Abstract iv

List of Tables x

List of Figures xi

Acknowledgements xiii

1. Introduction 1

1.1 The RecQ Family of Helicases 1

1.2 The Biology of RecQ4 2

1.2.1 Medical Syndromes Associated with RecQ4 2

1.2.2 RecQ4 in DNA Replication 5

1.2.3 RecQ4 in DNA Repair 9

1.2.4 Regulation of RecQ4 10

1.3 Concluding Remarks 13

2. Initial Biochemical Characterization of RecQ4 14

2.1 Introduction 14

2.2 Results and Discussion 14

2.2.1 Expression and Purification of *Drosophila* RecQ4 14

2.2.2 RecQ4 Demonstrates DNA Dependent ATP Hydrolysis 16

2.2.3 RecQ4 is a Helicase with 3'-5' Polarity 17

2.2.4 RecQ4 is an Annealase 18

2.3 Concluding Remarks 20

2.4 Methods 22
2.4.1 Cloning and Expression of the RecQ4 Construct 22
2.4.2 Purification of RecQ4 22
2.4.3 Thin Layer Chromatography 23
2.4.4 Oligonucleotide Substrate Preparation 24
2.4.5 Gel Shift Helicase Assay 25
2.4.6 Gel Shift Annealase Assay 25
3. Further Biochemical Characterization of RecQ4 Helicase 27
3.1 Introduction 27
3.2 Results and Discussion 27
3.2.1 Minimum Single-stranded Region Necessary for Helicase Activity 27
3.2.2 Helicase Activity on Bubble Substrates 29
3.2.3 Helicase Activity on Fork Substrates 30
3.2.4 Helicase Activity is Observed on Gap Substrates 32
3.3 Concluding Remarks 33
3.4 Methods 35
3.4.1 Oligonucleotide Substrate Preparation 35
3.4.2 Gel Shift Helicase Assay 36
4. The Binding Preferences of RecQ4 37
4.1 Introduction 37
4.2 Results and Discussion 37
4.2.1 Stable Complex Formation with DNA in the Presence of AMPPNP 37

4.2.2 RecQ4 Binds Single-stranded DNA 39

4.2.3 Confirmation of Binding Trends by Filter Binding Assay 41

4.2.4 RecQ4 Binding on Bubble, Fork, and Gapped Substrates 42

4.3 Concluding Remarks 44

4.4 Methods 47

4.4.1 Oligonucleotide Substrate Preparation 47

4.4.2 Electrophoretic Mobility Shift Binding Assay 47

4.4.3 Double-filter Binding Assay 48

5. Summary and Perspectives: The Biochemistry and Biology of RecQ4 49

5.1 The Comparative Biochemistry of RecQ4 49

5.2 Roles of RecQ4 in Replication 53

5.2.1 The Sld2 Domain and Initiation 53

5.2.2 The Involvement of the Helicase Domain 54

5.2.3 Cooperation Between RecQ4 and Mcm2-7 55

Appendix. Separate and Combined Biochemical Activities of the Subunits of a Naturally Split Reverse Gyrase 59

A.1 Introduction 59

A.2 Results and Discussion 62

A.2.1 Identification of the Subunits of Reverse Gyrase in *N. equitans* 62

A.2.2 Expression and Purification of the Subunits in Isolation 63

A.2.3 The Topoisomerase and Helicase-like Subunits Dimerize to Form Reverse Gyrase 66

A.2.4 Coordination Between the Topoisomerase and Helicase-like Subunits: Supercoiling....67

A.2.5 Coordination Between the Topoisomerase and Helicase-like Subunits: ATP Hydrolysis....70

A.2.6 DNA Binding Preferences of Each Subunit and the Reconstituted Holoenzyme....72

A.3 Concluding Remarks....76

A.4 Methods....80

A.4.1 Sequence Alignment....80

A.4.2 Expression and Purification....81

A.4.3 Dimerization Assays....82

A.4.4 Supercoiling Assays....83

A.4.5 Topology Shift Assays....83

A.4.6 ATP Hydrolysis....84

A.4.7 Electrophoretic Mobility Shift Assays....85

References....86

Biography....94

List of Tables

Table 1: Oligonucleotides Used In Substrate Preparation 24

Table 2: Composition of Substrates Used In Chapter 2 25

Table 3: Composition of Substrates Used In Chapter 3 35

Table 4: Composition of Substrates Used In Chapter 4 47

Table 5: Composition of Substrates Used In the Appendix 85

List of Figures

Figure 1: The symptoms of the RecQ4 syndromes. 3

Figure 2: Alignment of RecQ4 from various species with Sld2 and the other RecQ helicases in Homo sapiens and Saccharomyces cerevisiae. 8

Figure 3: RecQ4 expression in the eye imaginal disc of *Drosophila* is coincident with DNA replication. 11

Figure 4: Schematic diagram of *Drosophila melanogaster* RecQ4 construct. 15

Figure 5: Purification of RecQ4 16

Figure 6: RecQ4 exhibits DNA-stimulated ATP hydrolysis. 17

Figure 7: RecQ4 shows 3' to 5' helicase activity. 18

Figure 8: RecQ4 exhibits annealing activity. 20

Figure 9: Helicase activity with substrates of varying tail lengths. 28

Figure 10: Helicase activity with substrates having various sizes of bubble. 30

Figure 11: Helicase activity on fork substrates 31

Figure 12: Helicase activity on nicked and gapped substrates. 32

Figure 13: RecQ4 forms a stable complex in the presence of AMPPNP. 38

Figure 14: RecQ4 forms a stable complex with single-stranded DNA, not double-stranded DNA 39

Figure 15: Binding activity with substrates of varying tail lengths 40

Figure 16: Double-filter assays confirm the results from electrophoretic mobility shift assays. 41

Figure 17: Binding activity with substrates having various bubble sizes. 42

Figure 18: Binding activity with nicked and gapped substrates 43

Figure 19: The coordination of RecQ4's helicase domain with Mcm2-7 in replication.....58

Figure 20: Sequences of the subunits of *Nanoarchaeum equitans* reverse gyrase.....63

Figure 21: Purification of the Subunits of *N. equitans* reverse gyrase. 65

Figure 22: Ultracentrifugation and gel-filtration analyses of the subunits and equimolar mixtures.. 66

Figure 23: Reconstituted reverse gyrase induces positive supercoils.....70

Figure 24: The helicase-like subunit ATPase activity is stimulated by the topoisomerase subunit. 71

Figure 25: DNA substrate preference is partitioned between subunits. 73

Figure 26: DNA substrate preference is not altered by nucleotide cofactor.....74

Figure 27: Nucleotide cofactor modulates helicase-like subunit binding. 76

Acknowledgements

I am indebted to my mentor, Tao-shih Hsieh, for his support and guidance. When selecting the appropriate controls in the process of designing an experiment, or when drawing conclusions from data, it is his questions at the next lab meeting that I try to anticipate. The discussions, insights and critiques of my lab mates, Jianhong Wu, Jody Plank, Tammy Collins, Stefanie Chen, and Shane Chen, have also been particularly helpful to my project, and have also aided my development as a scientist. Yushen Qian, who I had the privilege of helping to mentor while he was an undergraduate student, was central to the work on reverse gyrase presented in the Appendix.

I would like to further thank Stefanie Chen, and my wife, Jo Anna Capp, for their part in reviewing and editing this dissertation.

1. Introduction

1.1 The RecQ Family of Helicases

The RecQ family of helicases is involved in various aspects of DNA metabolism, particularly replication and repair. Members of the family not only unwind DNA, but also anneal complementary DNA strands (reviewed in Bachrati and Hickson, 2008; Bohr, 2008; Seki et al., 2008). All prokaryotes and eukaryotes possess at least one RecQ helicase, and some plants have as many as seven. Five RecQ helicases have been identified in humans and mice, and three in *Drosophila*. Within a single species, multiple family members may exhibit a level of parallel functionality (reviewed in Seki et al., 2008). This mitigates the impact of defects in any one RecQ helicase. Even so, mutations in three of the five RecQ helicases in humans (Blm, Wrn, and RecQ4) lead to distinct clinical syndromes characterized by increased genetic instability, indicating that each of these plays a unique and important role (reviewed in Bohr, 2008). Bloom's Syndrome, caused by deficiencies in Blm, is defined by abnormally high levels of sister chromatid exchange, leading to increased occurrence of cancer. Mutations in Wrn lead to Werner's Syndrome, a severe form of progeria. Defects in RecQ4 are responsible for three different syndromes with a diverse set of symptoms. The remaining human RecQ helicases, RecQ1 and RecQ5, are not associated with any diseases, but are also implicated in suppressing inappropriate sister chromatid exchange, perhaps serving as backups to Blm (Wang et al., 2003). Though biologically important, the biochemistry of the RecQ helicases is poorly understood, and until recently the least understood was RecQ4. The recent advances in understanding the biological and biochemical roles of RecQ4 are the subject of a recent review (Capp et al., 2010b).

1.2 The Biology of RecQ4

1.2.1 Medical Syndromes Associated with RecQ4

RecQ4 was originally identified in 1998 in a human genome search for RecQ helicases (Kitao et al., 1998), and a year later was found to be associated with the rare type II Rothmund-Thomson syndrome (RTS) (Kitao et al., 1999). RTS is characterized by a wide array of defects, including cataracts, sparse hair, poikiloderma, growth deficiencies, various skeletal abnormalities and a greater propensity for osteosarcomas (reviewed in Vennos et al., 1992). The clinical impact of mutations in RecQ4 was later expanded to include the even rarer RAPADILINO (named for an acronym of the diagnostic symptoms) (Siitonen et al., 2003) and Baller-Gerold (Van Maldergem et al., 2006) syndromes. RTS, RAPADILINO and Baller-Gerold syndrome have a significant overlap of symptoms (Figure 1), all occurring in tissues with a high proliferation rate. The absence of a strong correlation between specific mutations in RecQ4 and specific phenotypes (reviewed in Larizza et al., 2006 and Siitonen et al., 2009) has led some to propose that the three syndromes be reclassified as a single one (Van Maldergem et al., 2006).

The spectrum of phenotypes associated with mutations in RecQ4 has been mirrored in studies using transgenic mice. Mouse RecQ4 knockouts were embryonic lethal (Ichikawa et al., 2002). Simply truncating the enzyme in the middle of the centrally located helicase domain showed no embryonic lethality, though only 5% survival was observed two weeks after birth. Mice that reached adulthood showed phenotypes similar to those found in RecQ4 associated syndromes, including smaller size, skin abnormalities, and hair discoloration (Hoki et al., 2003). Truncating the enzyme just before the helicase domain

(rather than within it) significantly reduced morbidity, giving a survival rate of 84%. Phenotypes associated with RecQ4 associated diseases were still present, though they were different from those found in mice with RecQ4 truncated within the helicase domain (Mann et al., 2005). Thus three different genotypic defects in murine RecQ4 result in three very different phenotypic patterns.

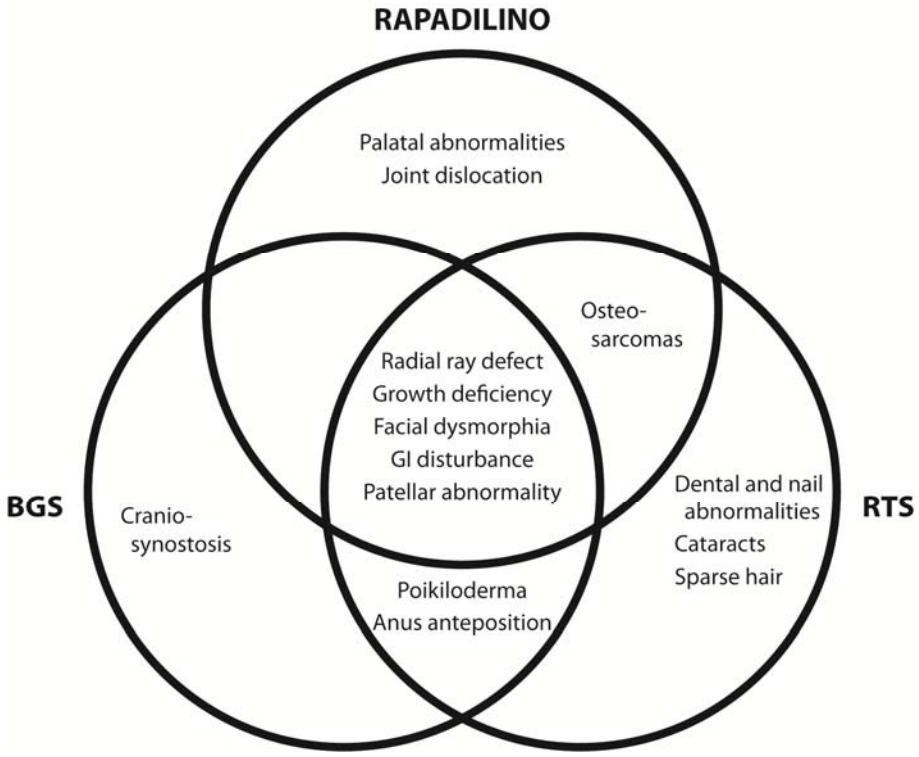

Figure 1: The symptoms of the RecQ4 associated Rothmund-Thomson (RTS), RAPADILINO, and Baller-Gerold (BGS) Syndromes. This is a Venn diagram of the symptoms of these diseases, showing which are unique and which are shared between syndromes. If these three are in fact a single syndrome, the key phenotypes would include growth deficiency, facial dysmorphia, and gastrointestinal disturbance.

The wide variation and heterozygosity of mutations in humans makes it difficult to derive enzymatic information about RecQ4 based on phenotypes, but no such difficulty exists with mice. Comparisons of human and murine RecQ4 are particularly relevant because

of the high conservation of the enzyme between the two species (63.4% overall sequence identity, and 85.7% similarity) (Ohhata et al., 2000). Such comparisons show that RecQ4 plays an essential role in development, as its complete absence leads to embryonic lethality. This may explain the extreme rarity of RecQ4 associated syndromes in humans, in that severe mutations would result in spontaneous early miscarriage and so remain unobserved. The helicase domain itself is not essential for viability, but its partial presence is more harmful than its absence. This suggests that the role played by the helicase domain is one for which there exist redundancies. In the absence of the helicase domain, these act to ensure the survival of the cell, albeit with a reduced efficiency leading to those phenotypes of RecQ4 associated syndromes that are less severe. In the presence of a partial or non-functional RecQ4 helicase domain such redundancies are unable to act. This results in the higher morbidity and more severe symptoms of RTS and other RecQ4 associated syndromes.

The above phenotypic analysis of mutant mice and RTS, RAPADILINO, and Baller-Gerold Syndrome patients demonstrates that RecQ4 plays two roles, though it does little to indicate the nature of those roles. One role is essential and does not require the helicase domain. The other is performed by the helicase domain, but either is non-essential, or can be performed by another enzyme in the absence of this domain. Evidence indicates that both roles are involved in DNA replication, and that the latter role may also be involved in DNA repair.

1.2.2 RecQ4 in DNA Replication

The first direct evidence for RecQ4's involvement in replication came in 2005, when two groups independently observed sequence homology between Sld2 in *Saccharomyces cerevisiae* and the N-terminus of xRTS, the RecQ4 homolog in *Xenopus* (Figure 2) (Matsuno et al., 2006; Sangrithi et al., 2005). Sld2 is one of two essential targets for S-phase cyclin dependent kinase (CDK), the other being Sld3. Phosphorylation of Sld2 and Sld3 by S-phase CDK leads them to interact with Dpb11 (Tanaka et al., 2007b; Zegerman and Diffley, 2007). These interactions are essential for the initiation of DNA replication in *S. cerevisiae* (reviewed in Tanaka et al., 2007a; Araki, 2010; and Sclafani and Holzen, 2007). RecQ4 is the only known metazoan homolog of Sld2 and is, by virtue of this comparison, strongly implicated in replication.

This implication is borne out by data from *Xenopus* and *Drosophila* model systems, as well as from human tissue culture cells. The N-terminus of xRTS (including the Sld2 homologous domain, but not the helicase domain) is necessary for replication initiation in *Xenopus* oocyte extract. It is also necessary for chromatin binding by DNA Polymerase α (Matsuno et al., 2006). Supplementing xRTS-depleted *Xenopus* oocyte extract with the N-terminus of human RecQ4 restores replication to ~20% of wildtype levels (Sangrithi et al., 2005). The fact that restoration is only partial suggests that the helicase domain also plays a role in replication. Similarly, *Drosophila* with null or hypomorphic expression of RecQ4 are severely defective in normal pre-mitotic replication. Hypomorphic *Drosophila* are also defective in endoreplication (genome synthesis without subsequent mitosis, leading to the polyploidy found in salivary glands) and chorion gene amplification (nested synthesis of the chorion gene, allowing for subsequent rapid production of the protein during egg shell

generation). These defects occur during initiation, not elongation (Wu et al., 2008; Xu et al., 2009b). In human tissue culture cells, RecQ4 is associated with replication origins only during late G1-phase and S-phase, and depletion of RecQ4 inhibits cell proliferation and DNA synthesis (Thangavel et al., 2010). Work in *Xenopus*, *Drosophila* and human systems thus confirms that RecQ4 plays a role in replication initiation, similar to Sld2.

In yeast, as mentioned earlier, it is known that Sld2's role in replication is mediated by Dpb11 and associated with Sld3 (reviewed in Tanaka et al., 2007a and Sclafani and Holzen, 2007). However, evidence that the interactions between Dpb11 and phosphorylated Sld2 and Sld3 are conserved in metazoans is not conclusive. It is not clear that Sld2's interaction with Dpb11 is paralleled by the interaction of RecQ4 with the known homologues of Dpb11, which include TopBP1 in humans, Cut5 in *Xenopus*, and Mus101 in *Drosophila* (reviewed in Garcia et al., 2005). In *Xenopus* oocyte extract, Cut5 and the N-terminus of xRTS co-immunoprecipitate independent of xRTS phosphorylation (Matsuno et al., 2006). Therefore, unlike Sld2 and Dpb11, RecQ4 does not need to be phosphorylated to interact with Cut5. And although xRTS is necessary for replication, it does not mediate the replication origin's interaction with either Cut5 or the CMG (Cdc45; Mcm2-7; GINS) complex, as Cut5 and the components of the CMG complex load onto chromatin even in xRTS-depleted oocyte extract (Sangrithi et al., 2005). Data from human tissue culture cells suggests that the key interaction is not between RecQ4 and TopBP1/Cut5/Mus101, but rather between RecQ4 and Mcm10. RecQ4 co-immunoprecipitates with Mcm10 (Xu et al., 2009a), and both RecQ4 and Mcm10 are necessary for the formation of the CMG complex (Im et al., 2009). In contrast, TopBP1 neither co-immunoprecipitates with RecQ4, nor is necessary for the stable formation of the CMG complex (Im et al., 2009; Xu et al., 2009a). It

is possible that the way RecQ4-Mcm10 stabilizes the CMG complex is by causing GINS to associate with Mcm2-7 and Cdc45 (Xu et al., 2009a). Thus RecQ4-Mcm10 mediates the binding of the components of the CMG complex to each other, rather than to the replication origin. The differing data from *Xenopus* and human systems with respect to the interaction of RecQ4 and TopBP1/Cut5 may reflect either species-specific differences, or subtle dynamics in the formation of the metazoan replication complex. Regardless, it is clear that RecQ4 is central to the formation of the replication complex, and that it is loaded on the origin prior to Polymerase α or RPA (Matsuno et al., 2006; Sangrithi et al., 2005).

It is clear that the N-terminal Sld2 domain of RecQ4 plays a necessary role in replication initiation, and absence of the domain leads to the early developmental lethality seen in RecQ4 null mutants. However, one cannot disregard the replicative importance of the other domains of RecQ4, particularly the helicase domain. Both the Sld2 domain and the helicase domain are required for viability in *Drosophila*, though the C-terminus is not essential (Xu et al., 2009b). More precisely, point mutants inactivating the helicase domain are unable to restore replication in either xRTS-depleted *Xenopus* oocyte extract or *Drosophila* null mutants (Capp et al., 2009; Sangrithi et al., 2005). The inability of these otherwise intact proteins to restore replication or viability indicates that RecQ4's helicase domain is also important for replication. The Sld2 and helicase domains of RecQ4 are thus both critical to replication, with the former playing a non-enzymatic role in initiation, and the latter playing a role requiring helicase activity in an undetermined stage of the process.

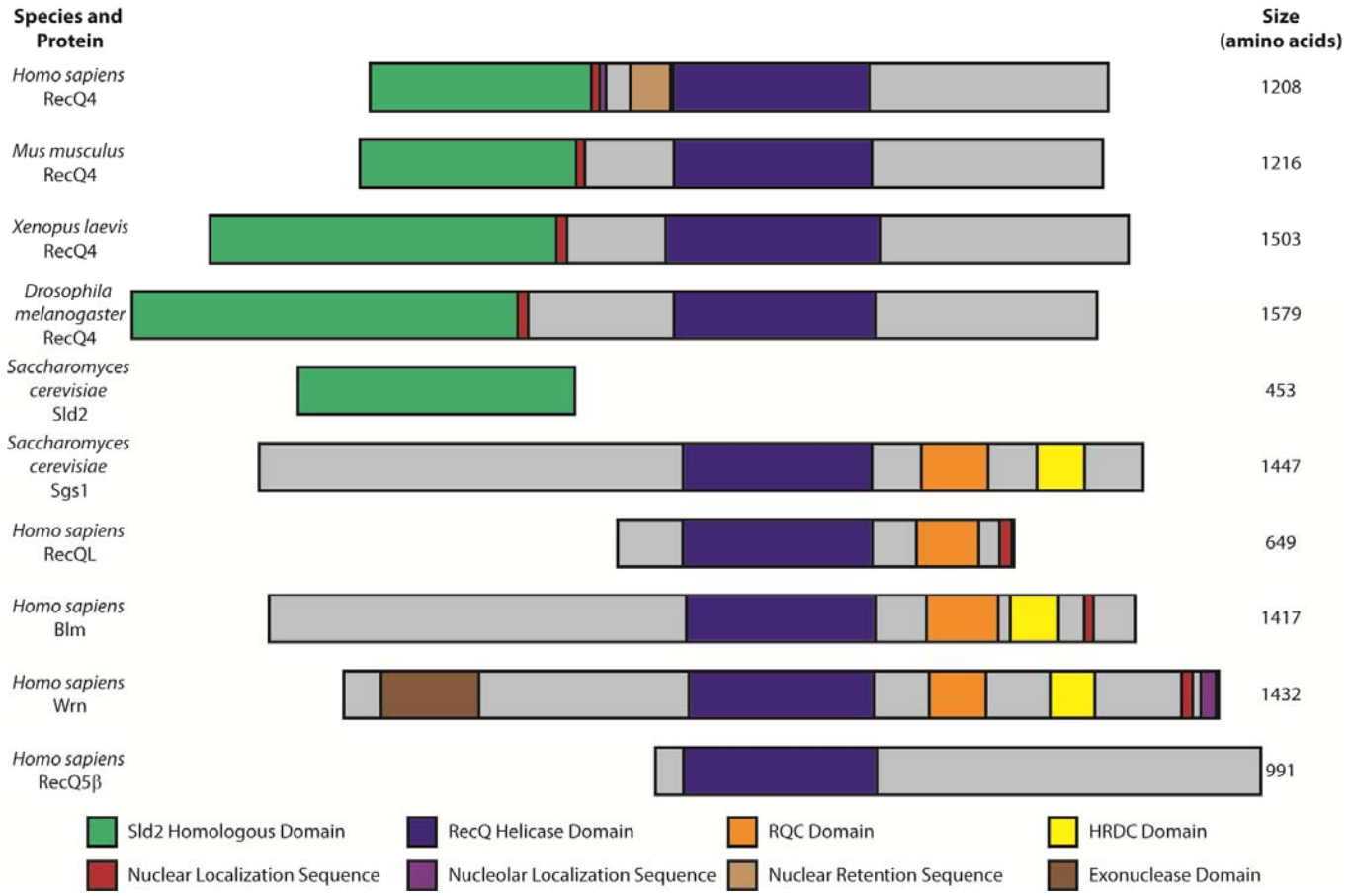

Figure 2: Alignment of RecQ4 from various species with Sld2 and the other RecQ helicases in Homo sapiens and Saccharomyces cerevisiae. The RecQ helicases are aligned based on their shared RecQ helicase domain (in blue). The N-terminus of RecQ4 is the less conserved between species than the helicase domain and C-terminus. Unlike most RecQ family members, RecQ4 does not have an RQC (RecQ C-terminal) domain (in orange) or an HRDC (Helicase and Rnase D C-terminal) domain (in yellow).

1.2.3 RecQ4 in DNA Repair

Data suggest that RecQ4 may also be involved in DNA damage repair, but this is considerably cloudier than its involvement in replication. In one study, cell lines derived from RTS patients show sensitivity similar to wildtype cell lines when subjected to a variety of DNA damaging agents, including those that induce DNA double-strand breaks (DSB), oxidative damage, inter-strand crosslinks, as well as those that interfere with replication, such as hydroxyurea (Cabral et al., 2008). However, this is the only study reporting that the loss of RecQ4 has no effect on sensitivity to such agents. Another study, using similar approaches, instead found increased sensitivity of RTS cell lines to agents that interfere with replication, and wildtype sensitivity to those that induce DSBs (Jin et al., 2008). Such data are consistent with RecQ4 being involved in the restart of stalled replication forks. While normal human cells undergo significant S-phase arrest after treatment with hydroxyurea or UV irradiation, RTS cells and T-293 cells with RecQ4 knocked down by shRNA have been shown to not enter such an arrest (Park et al., 2006). RecQ4 may thus play a role in the signaling of cell cycle arrest during DNA damage response. There is also some evidence that RecQ4 is involved in response to oxidative damage, which would suggest that it is part of the base excision repair pathway. RTS cells and those with RecQ4 transiently knocked down by siRNA have been shown by one group to be hypersensitive to oxidative damage (Schurman et al., 2009). The enzyme can also relocalize to the nucleolus in response to oxidative damage (Woo et al., 2006), indicating that the intracellular localization of RecQ4 may be regulated by DNA damage response. In human fibroblasts after induction of DSBs, the proportion of RecQ4 nuclear foci colocalizing with promyelotic leukemia protein was reduced, in favor of

associating with Rad51 (Petkovic et al., 2005). Though this shift was relatively minor, it still indicates a response to DSB induction. Taking these results together, it is difficult to form a consistent model concerning RecQ4 and DNA damage repair. There is relative consensus that RecQ4 is involved somehow in replication arrest and repair of oxidative damage. It is more likely that this involvement is due to the helicase domain than to the Sld2 domain, but one also cannot eliminate the possibility of an undiscovered activity in the C-terminus being responsible. Certainly, much remains to be done to determine the precise nature of RecQ4's involvement in DNA damage repair.

1.2.4 Regulation of RecQ4

It is generally essential that DNA be replicated once and only once during the life of a cell. Accordingly, many of the proteins involved in replication are tightly regulated in terms of expression, cellular localization, and enzymatic activity (reviewed in Remus and Diffley, 2009). RecQ4 is no exception. Both the expression and localization of RecQ4 are closely regulated, and it is possible that the activity of the enzyme is as well.

Drosophila has proved to be a useful model system for examination of the expression and localization patterns of RecQ4 (Wu et al., 2008). Cell cycle dependent expression patterns can be observed using the synchronized cell cycle progression found in the eye imaginal disc during *Drosophila* development. The eye imaginal disc has a morphogenic furrow (see arrow on Figure 3), in which the cells are all paused in G1. Anterior to the furrow (left of the arrow), cells are asynchronously cycling. Posterior to it, the cells enter S-phase at precisely the same time, as indicated by BrdU incorporation, and proceed through the following mitotic wave in lockstep before entering into a non-replicative state.

Throughout this process, RecQ4 expression is observed in asynchronously cycling cells (anterior to the furrow) and in mitotic cells (posterior to the furrow). It is notably absent within the morphogenic furrow itself. This means that RecQ4 must be destroyed prior to entrance into G1, only to be re-synthesized when replication is about to occur. Once the cells enter the extended non-replicative state (to the far right of the arrow), RecQ4 again diminishes, indicating degradation (Wu et al., 2008). It is consistent with this pattern that p53 represses expression of RecQ4 during G1 in human fibroblasts (Sengupta et al., 2005). Thus RecQ4 expression is tightly regulated by mechanisms addressing both its synthesis and degradation, confining its presence to S-phase and shortly thereafter.

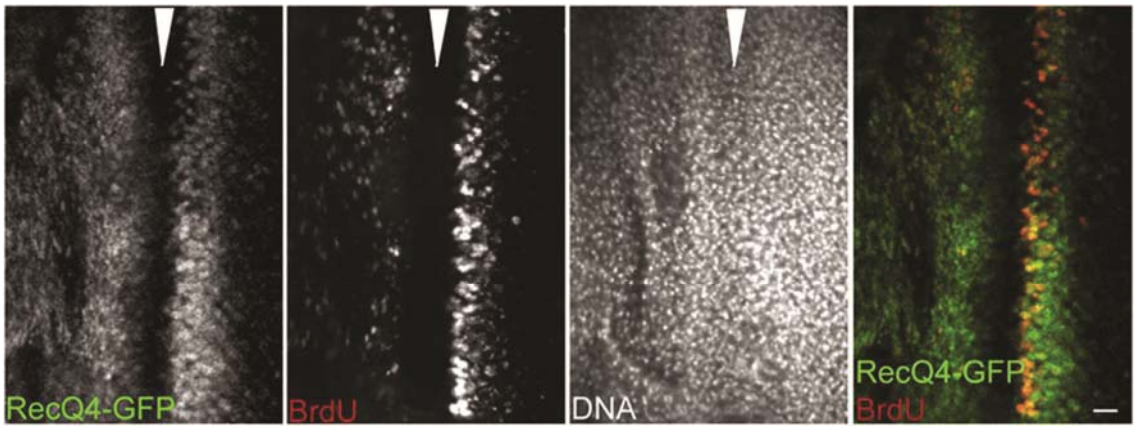

Figure 3: RecQ4 expression in the eye imaginal disc of *Drosophila* is coincident with DNA replication. Although the cell density (as seen by DAPI staining of the DNA) is constant, synthesis (as seen by BrdU incorporation) occurs only in two mitotic waves, separated by a morphogenic furrow (arrowhead) of cells in extended G1. The first, anterior to the morphogenic furrow (left of the arrowhead), is asynchronous. The second one, immediately posterior to the morphogenic furrow (right of the arrowhead), is synchronous. RecQ4 is absent upon entrance in to the extended G1 phase of the morphogenic furrow, but is again expressed with the beginning of synthesis. This indicates that RecQ4 is expressed only during S-phase, and that in other phases it is repressed and/or degraded. Scale bar: 10µm. Reprinted from *Developmental Biology.* Wu *et al.* Drosophila homologue of the Rothmund-Thomson syndrome gene: essential function in DNA replication during development. *Dev. Biol.* 2008; 323:130-142. © Elsevier.

The localization of RecQ4 within the cell is also closely regulated. Early *Drosophila* embryos do not have separate cells, though they do have distinct membrane-contained nuclei. The nuclei all proceed synchronously through the cell cycle (albeit lacking G1 and G2 phases) until cellularization occurs, after which the cell cycle lengthens and synchrony is lost (Foe et al., 1993). As expected for an enzyme involved in replication, RecQ4 is associated with the chromatin during interphase (Wu et al., 2008). During metaphase and anaphase, the nuclear membranes break down and RecQ4 is dispersed to the cytoplasm. After the membrane re-forms during telophase, RecQ4 is again found associated with the chromatin. This localization pattern is continued even after the formation of discrete cells (Wu et al., 2008). Work using human tissue culture cells has defined within the N-terminus of RecQ4 a nuclear localization sequence, a nuclear retention sequence (Burks et al., 2007), and a lysine-rich nucleolar localization sequence (Woo et al., 2006) (Figure 2). Deletion of the nuclear retention sequence results in localization to both the nucleus and the cytoplasm (Burks et al., 2007), and has been correlated to the occurrence of osteosarcomas and lymphomas among RTS and RAPADILINO patients (Siitonen et al., 2009). Acetylation by p300 of the lysines within the nucleolar localization sequence also causes RecQ4 to be excluded from the nucleus (Dietschy et al., 2009), suggesting a means of regulating the cell-cycle dependent localization of RecQ4.

Besides acetylation, there are several other post-translational modifications of RecQ4 that may also regulate the enzyme. The C-terminus of RecQ4 has been shown to co-immunoprecipitate with poly(ADP-Ribose) Polymerase-1, and to be an *in vitro* substrate for poly-ADP ribosylation (Woo et al., 2006). Although RecQ4 has also been shown to form stable interactions with UBR1 and UBR2 (ubiquitin E3 ligases of the N-end rule pathway,

which targets proteins for degradation by the proteasome), it is neither ubiquitylated nor rapidly degraded (Yin et al., 2004). As mentioned earlier, Sld2 depends on being phosphorylated for functionality. But while it is known that RecQ4 can serve as a substrate for phosphorylation *in vitro*, it has not been established that this actually occurs *in vivo*. Biochemical identification and characterization of these and other as yet to be discovered post-translational modifications will be a critical area for future investigation.

1.3 Concluding Remarks

Comparison of the data from RTS patients and mouse, *Xenopus*, and *Drosophila* model systems leads to the prediction of an important role in replication for the functionally competent helicase domain of RecQ4. This is in addition to the largely non-enzymatic role in replication initiation suggested by homology of the N-terminus of RecQ4 with Sld2. Our work characterizing the helicase activity of RecQ4 will be discussed in the following chapters. The implications this has for RecQ4's involvement in replication will also be discussed.

The work presented in this chapter was originally published in *Critical Reviews in Biochemistry and Molecular Biology*. Capp, C., Wu, J. and Hsieh, T.S. RecQ4: the second replicative helicase? *Crit. Rev. Biochem. Mol. Biol.* 2010; 45:233-242. © Informa Medical and Pharmaceutical Science - Journals.

2. Initial Biochemical Characterization of RecQ4

2.1 Introduction

The helicase activity of RecQ4 has proven difficult to biochemically assess. As such, the first two publications examining the *in vitro* activity of the enzyme were unable to detect helicase activity (Macris et al., 2006; Yin et al., 2004). This was not expected, as each of the four other human RecQ helicases had demonstrated 3′-5′ helicase activity (reviewed in Seki et al., 2008; Bohr, 2008; and Bachrati and Hickson, 2008), and these all share a well conserved helicase domain with RecQ4. In 2009 several labs independently reported the observation of helicase activity from the enzyme (Capp et al., 2009; Suzuki et al., 2009; Xu and Liu, 2009). Our results are presented below, while the comparison of the published biochemical analyses of RecQ4 is presented in Chapter 5. The research presented in this chapter, as well as in sections 3.2.3, 4.2.1, 4.2.2 and 4.2.3, was originally published in the *Journal of Biological Chemistry*. Capp, C., Wu, J. and Hsieh, T.S. *Drosophila* RecQ4 Has a 3'-5' DNA Helicase Activity That Is Essential for Viability. *J. Biol. Chem.* 2009; 284:30845-30852. © the American Society for Biochemistry and Molecular Biology.

2.2 Results and Discussion

2.2.1 Expression and Purification of *Drosophila* RecQ4

Drosophila RecQ4, with an N-terminal GST tag and a C-terminal His_{10} tag (GST-RecQ4-His_{10}, see Figure 4), was expressed in Sf9 cells using the Bac-to-Bac Baculovirus Expression System. The two tags allowed for a double-affinity purification strategy (see flowchart in Figure 5). Soluble material was first incubated with glutathione-sepharose beads, which were bound by the GST tag of GST-RecQ4-His_{10}. The GST tag was removed from

RecQ4 by PreScission Protease, resulting in RecQ4-His_{10}. The enzyme was further purified and concentrated by immobilized metal ion affinity chromatography (IMAC), using Co^{2+} as the metal. The final purification step was accomplished by ultracentrifugation across a 30%-60% glycerol density gradient. This step separated RecQ4 from a co-purifying exonuclease activity, and placed the enzyme in buffer conditions suitable for long-term storage. It should be noted that RecQ4 separated in the glycerol gradient in a manner consistent with existing as a monomer in solution (data not shown). Peak fractions were determined by ATPase activity and SDS-PAGE analysis (Figure 5), and the pooled peak fractions are more than 96% pure based on densitometric tracing.

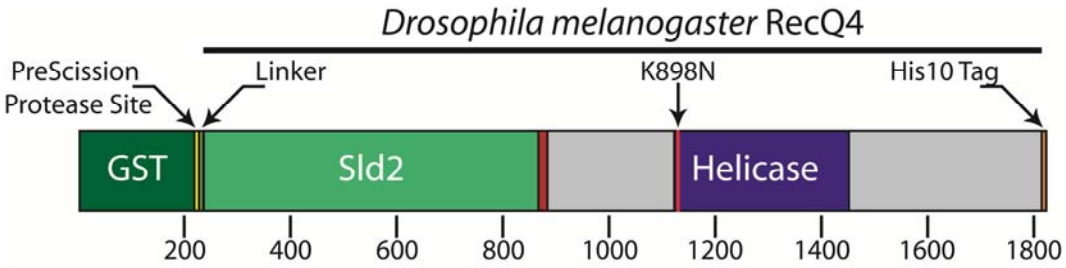

Figure 4: Schematic diagram of *Drosophila melanogaster* RecQ4 construct. An N-terminal Glutathione-S-transferase tag is attached to the enzyme by a PreScission Protease site and linker region. Also indicated are the C-terminal His_{10} tag, and lysine 898, which is mutated to asparagine in the ATPase-dead mutant (K898N).

As a negative control, lysine 898 (indicated in Figure 4) in the conserved Walker A motif (GSGKS) was mutated to asparagine, making the new motif sequence "GSGNS." This mutation is known to greatly reduce ATPase activity in Superfamily II helicases, thus preventing helicase activity (Hall and Matson, 1999). The expression and purification of the mutant protein (RecQ4-K898N) are identical to those of the wildtype, including

ultracentrifugation through a glycerol density gradient, suggesting that the mutant is neither unfolded nor aggregated.

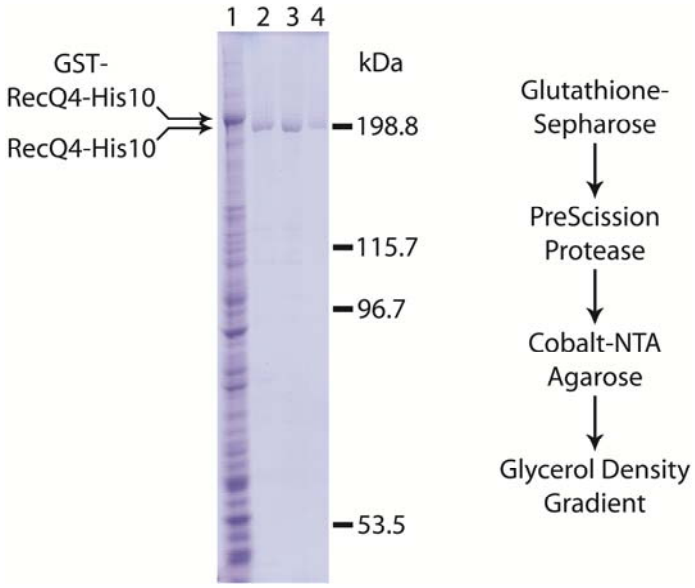

Figure 5: Purification of RecQ4. An SDS-PAGE of samples from purification stages, with a flow chart diagramming the purification strategy. *Lane 1*: Soluble fraction. *Lane 2*: RecQ4 after PreScission Protease treatment. *Lane 3*: RecQ4 after elution from Co^{2+}-NTA resin. *Lane 4*: RecQ4 peak from glycerol density gradient.

2.2.2 RecQ4 Demonstrates DNA Dependent ATP Hydrolysis

A hallmark activity for any DNA helicase is DNA-dependent ATP hydrolysis (Lohman and Bjornson, 1996). To characterize the ATPase activity of RecQ4, we incubated RecQ4 with γ-^{32}P ATP in the presence or absence of single-stranded circular DNA, and monitored the liberation of phosphate using thin layer chromatography. The presence of DNA greatly stimulated hydrolysis of ATP by RecQ4 (Figure 6A). Plasmid DNA also stimulated RecQ4's ATPase activity (data not shown). As expected, RecQ-K898N showed no detectable amount of ATP hydrolysis above background (Figure 6B).

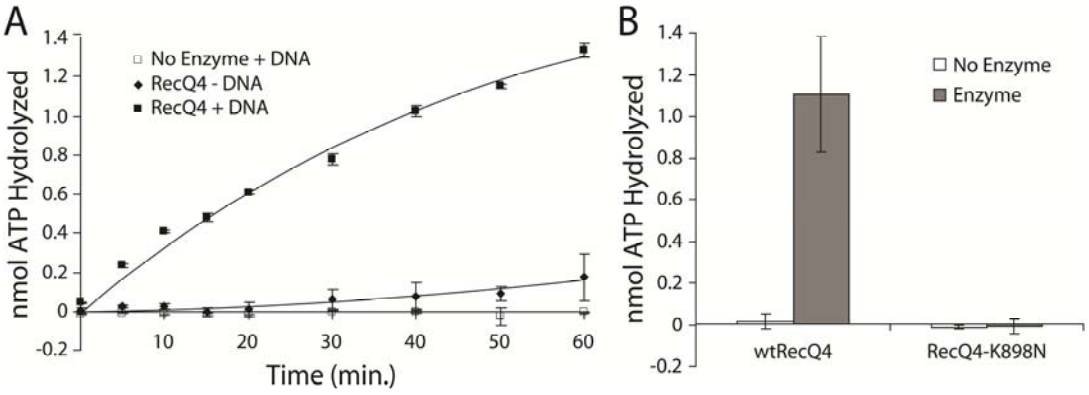

Figure 6: RecQ4 exhibits DNA-stimulated ATP hydrolysis. *(A)* ATP hydrolysis by RecQ4 is stimulated in the presence of single-stranded DNA. In the absence of RecQ4, ATP hydrolysis is not observed, independent of the presence of DNA. The reaction timecourse was performed by taking 1μL samples from a total reaction volume of 20μL at the indicated times. For each timepoint of each trial, background is removed by subtracting the amount of ATP hydrolyzed in the absence of both enzyme and DNA. Enzyme concentration was 13.4nM. *(B)* RecQ4-K898N does not hydrolyze ATP. 1μL samples were taken from 20μL reactions after 60 minutes and analyzed by TLC and phosphorimaging analysis. For each trial, a 0 minute timepoint was also taken. Background was removed by subtracting the amount ATP hydrolyzed at 0 minutes from that at 60 minutes. Enzyme concentration was 7.4nM. Error bars in these experiments indicate standard deviation (n=3).

2.2.3 RecQ4 is a Helicase with 3'-5' Polarity

To investigate the helicase activity of the purified *Drosophila* RecQ4 protein, we incubated RecQ4 with a 3' extension DNA substrate (see Figure 7A, lanes 3 and 4, and Figure 7B). Helicase activity was detected only in the presence of ATP, with no activity observed in the absence of nucleotide cofactor or in the presence of ADP or a non-hydrolyzable homolog, AMPPNP. In contrast, RecQ4-K898N did not exhibit helicase activity under any conditions (Figure 7B). This establishes that the helicase activity observed is intrinsic to purified wildtype RecQ4, and not a co-purifying contaminant.

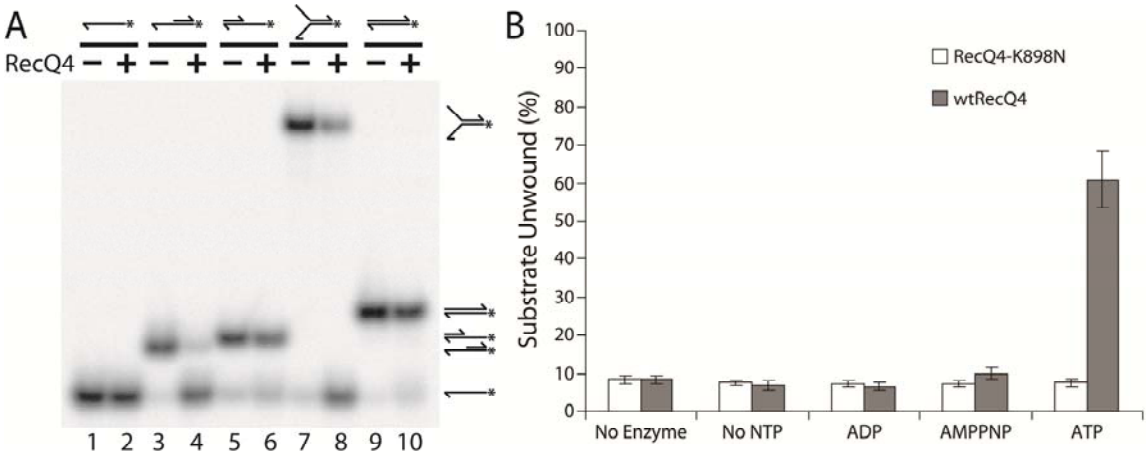

Figure 7: RecQ4 shows 3' to 5' helicase activity. *(A)* RecQ4 unwinds substrates with a 3' single-stranded region. Radiolabeled oligonucleotide substrates were incubated with RecQ4 (even numbered lanes) or an equal volume mock (odd numbered lanes), and separated by PAGE. RecQ4 has no effect on single-stranded substrate (compare lanes 1 and 2) or double-stranded substrate (lanes 9 and 10). RecQ4 can unwind substrate with a single-stranded region 3' (lanes 3 and 4, 7 and 8), but not 5' (lanes 5 and 6), of a double-stranded region. *(B)* RecQ4 (shaded bars) unwinds the 3' extension substrate (see *A*, lanes 3 and 4) in the presence of ATP. This activity is dependent on ATP hydrolysis, as it is not seen the absence of ATP, or in the presence of AMPPNP. This activity is also due to RecQ4 and not a contaminant because RecQ4-K898N (open bars) does not unwind the 3' extension substrate under any of the conditions tested. Helicase assays were conducted as seen in *A*, and were quantified by phosphorimaging analysis.

To further characterize this helicase activity, we incubated RecQ4 with several substrates constructed from oligonucleotides (see Figure 7A; substrate composition is indicated in Table 2, and oligonucleotide sequences are given in Table 1). RecQ4 does not unwind a 60 basepair duplex substrate (Figure 7A, lanes 9 and 10), but will unwind a fork substrate (lanes 7 and 8). This indicates that a single-stranded region strongly stimulates RecQ4's helicase activity, and may in fact be necessary for it. Helicases in general exhibit distinct directionality preferences, and may unwind DNA moving either from the 3' end of the DNA to 5' end, or conversely, from the 5' end to 3' end. RecQ family members have primarily been observed acting in a 3' to 5' direction. To determine the directionality of

RecQ4's helicase activity, a comparison was made between the 3' and 5' extension substrates (Figure 7A, lanes 3-6). Robust helicase activity was detected on the substrate with a 3' extension, but not the one with a 5' extension. This indicates that a single-stranded region at the 3'-end of the double-stranded region clearly stimulates RecQ4's helicase activity, and suggests the enzyme moves in a 3' to 5' direction.

2.2.4 RecQ4 is an Annealase

Complementary DNA strand annealing activity has been exhibited by all human RecQ helicases (Cheok et al., 2005; Machwe et al., 2005; Macris et al., 2006; Muftuoglu et al., 2008). This rewinding activity may coordinate with the helicase activity to play an important role in the formation and resolution of recombination intermediates. Human RecQ4 was demonstrated to have strand annealing activity (Macris et al., 2006), and it has been suggested that this interferes with detection of the helicase activity in this enzyme (Xu and Liu, 2009). *Drosophila* RecQ4 has robust helicase activity, raising the question of whether it, like other RecQ family members, also possesses annealing activity. We incubated enzyme with complementary single-stranded DNA, which forms duplex DNA 60 basepairs long when renatured. The absence of a single-stranded region in the annealed reaction product prevents it from being a substrate for RecQ4's helicase activity. Annealing reactions were performed in the absence of nucleotide, and in the presence of ADP, ATP, or AMPPNP (Figure 8A, lanes 7-10). RecQ4 demonstrated annealing activity under all conditions tested, though in the presence of AMPPNP annealing was reduced by ~50% (Figure 8B). Interestingly, RecQ4-K898N did not demonstrate this AMPPNP-dependent reduction, but was able to anneal complementary single strands equally well under all conditions tested

(Figure 8A, lanes 2-5; quantified in Figure 8B). The fact that RecQ4-K898N shows strand annealing activity comparable to the wildtype enzyme (though lacking inhibition by AMPPNP) confirms that the mutation does not render the protein grossly misfolded or aggregated.

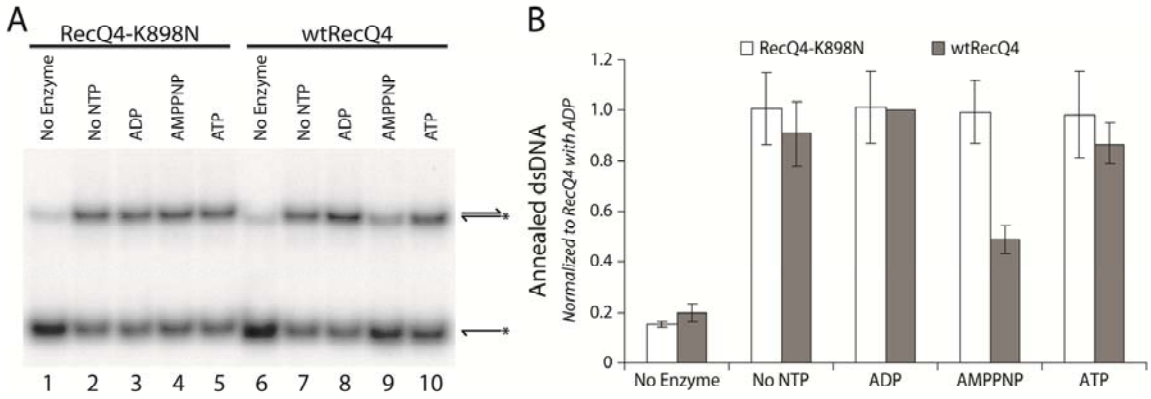

Figure 8: RecQ4 exhibits annealing activity. *(A)* AMPPNP reduces the annealing activity of RecQ4. wtRecQ4 (lanes 7-10) and RecQ4-K898N (lanes 2-5) were each incubated with two complementary single strands of DNA, in the absence of nucleotide (lanes 2 and 7), or in the presence of ADP (lanes 3 and 8), AMPPNP (lanes 4 and 9), or ATP (lanes 5 and 10). Annealing is observed in all of these conditions, with wtRecQ4 exhibiting decreased annealing in the presence of AMPPNP. *(B)* Quantification of the annealing activity of RecQ4. Phosphorimaging analyses of the annealing assays were quantified to determine the fraction of duplex product formed by wtRecQ4 (shaded bars) and RecQ4-K898N (open bars). For each trial, results were normalized to those from wtRecQ4 and ADP. Error bars indicate standard deviation (n=3).

2.3 Concluding Remarks

We have used the baculoviral expression system to express *Drosophila* RecQ4, and purified it with two affinity chromatography steps, followed by a glycerol density gradient (Capp et al., 2009). Enzyme thus obtained was more than 96% pure, and demonstrated DNA-stimulated ATP hydrolysis, annealing, and helicase activity. Helicase activity was derived exclusively from the helicase domain, because a point mutant inactivating ATP hydrolysis abolished helicase activity (Figure 7B). RecQ4 unwound duplex regions as long as

30 basepairs, and required single-stranded DNA 3′ of the region to be separated (Figure 7A). Neither a 60 basepair blunt-ended DNA substrate, nor a substrate with single-stranded DNA 5′ of a 30 basepair duplex region could be unwound. This confirms the 3′-5′ directionality of the enzyme, and is consistent with the activity of other RecQ helicases.

Interestingly, although the annealing activity of RecQ4 was not dependent on ATP hydrolysis, it was suppressed in the presence of AMPPNP. This suppression is derived from the helicase domain, as the same K898N mutant that abolishes helicase activity also abolishes the effect of AMPPNP on annealing. This does not establish that the helicase domain is itself capable of annealing activity. Recent work by Kanter et al (Kanter and Kaplan, 2010) has found Sld2 to anneal complementary DNA strands, and so it is possible that the homologous domain in RecQ4 is also able to anneal DNA. If this is the case, inhibition of annealing by AMPPNP implies communication and possible inter-regulation between the Sld2 domain and the helicase domain.

These results confirm that RecQ4, like the other RecQ family members, exhibits a robust helicase activity with a preference for 3' single-stranded ends. Several other groups have independently purified human RecQ4 and published initial biochemical characterizations of the enzyme. These articles show significant variation in approaches and in results. Comparison and contrast of our results with the other published comparisons should thus yield further insights into the activity of RecQ4, and will be addressed below in Chapter 5.

2.4 Methods

2.4.1 Cloning and Expression of the RecQ4 Construct

Drosophila RecQ4 cDNA was cloned into the pFastbac 1 vector (Invitrogen), with an N-terminal Glutathione-S-Transferase (GST) tag (separated from the enzyme by a PreScission Protease recognition site), and a C-terminal tag consisting of ten histidines (His_{10}). A diagrammatic representation of the recombinant RecQ4 construct is shown in Figure 4. The RecQ4-K898N mutant was generated by site-directed mutagenesis of the pFastbac 1 recombinant vector using Stratagene's Quikchange Lightning system.

RecQ4 was expressed using the Bac-to-Bac Baculovirus Expression System (Invitrogen). From the pFastbac 1-RecQ4 vector, the Bac-to-Bac system was used to create baculovirus encoding the RecQ4 construct under control of the *p10* promoter. Sf9 cells were grown to a density of $1x10^6$ cells/mL, and infected with a multiplicity of infection of 1. Virus-infected cells were grown for an additional 48 hours, pelleted, snap frozen and stored at -80°C.

2.4.2 Purification of RecQ4

Pelleted cells were lysed by resuspension with a Dounce homogenizer in Lysis Buffer (10mM Tris HCl, pH 7.8; 20mM KCl; 25% glycerol; 2mM EDTA); NaCl was added to a final concentration of 350mM, and the lysate was incubated on ice for 20 minutes with stirring. All subsequent steps were performed at 4°C. The soluble fraction was isolated by centrifugation at 20,000Xg for 20 minutes. Soluble material was incubated with Glutathione Sepharose 4B resin (GE Healthcare) for one hour, allowing the GST-tagged enzyme to bind. Glutathione resin was washed three times with buffers of 50mM Tris HCl, pH 7.8, 10%

glycerol, 0.02% Triton X-100, containing successively: 350mM NaCl, 2mM EDTA; 1M NaCl; and 150mM NaCl. To remove the GST tag, enzyme was incubated for two hours with 40U of PreScission Protease (GE Healthcare) per 1L culture. Cleaved enzyme was then passed over a Co^{2+} IMAC (Clontech) column, and washed with buffer containing 50mM Tris HCl, pH 7.8, 150mM NaCl, 10% glycerol, and 0.02% Triton X-100. The same buffer, with 400mM imidazole added, was used for step elution. Peak fractions from the IMAC column, as determined by Bradford assay, were combined and applied to a glycerol density gradient from 30% to 60%. Gradients were centrifuged for 20 hours at 60,000 rpm in rotor TV-865B, and collected in 0.5mL fractions. Gradient peak was determined by ATPase assay for wtRecQ4 (see below), or SDS-PAGE analysis for RecQ4-K898N. Purified enzyme was flash frozen and stored at -80°C.

2.4.3 Thin Layer Chromatography

Enzyme was incubated with 1mM ATP (3.8nM γ^{32}P-labeled ATP) for 60 minutes at 27°C in 2.5mM Tris acetate, pH 7.9, 2.5mM potassium acetate, 8mM magnesium acetate, 50μg/mL BSA, 1mM DTT, and 50μg/mL pBSK+ single-stranded circular DNA in a total volume of 20μL. At the indicated timepoints of 0, 5, 10, 15, 20, 30, 40, 50 and 60 minutes, 1μL of the reaction was spotted on Polygram Cel 300 PEI thin layer chromatography plates (Macherey-Nagel) that had been pre-run in distilled water. Plates were soaked in methanol for 5 minutes, dried, and developed in 0.5M LiCl, 1M formic acid to separate inorganic phosphate from ATP. Dried plates were subjected to phosphorimaging analysis.

2.4.4 Oligonucleotide Substrate Preparation

PAGE-purified DNA oligonucleotides were obtained from Integrated DNA Technologies (sequences shown in Table 1).

Table 1: Oligonucleotides Used In Substrate Preparation

Name	Sequence (5' to 3')
A30	CGAAGGCCAT GATTGCGCAC TGAATACATC CTGCCCTGTT ATTAATTACG TTATCTTACG
A10	CGAAGGCCAT GATTGCGCAC TGAATACATC CTGCCCTGTT
A5	CGAAGGCCAT GATTGCGCAC TGAATACATC CTGCC
A3	CGAAGGCCAT GATTGCGCAC TGAATACATC CTG
A2	CGAAGGCCAT GATTGCGCAC TGAATACATC CT
A1	CGAAGGCCAT GATTGCGCAC TGAATACATC C
A	CGAAGGCCAT GATTGCGCAC TGAATACATC
A 5Bubble	CGAAGGCCAT GATTGCGCAC TTTTTGTTAA TTACGTTATC TTACG
A 2Bubble	CGAAGGCCAT GATTGCGCAC TTGTTAATTA CGTTATCTTA CG
BC	CGTAAGATAA CGTAATTAAT AACAGGGCAG GATGTATTCA GTGCGCAATC ATGGCCTTCG
B	GATGTATTCA GTGCGCAATC ATGGCCTTCG
B 20Bubble	CGTAAGATAA CGTAATTAAT CCACTCTAGA AGCTCTGAAT GTGCGCAATC ATGGCCTTCG
B 5Bubble	CGTAAGATAA CGTAATTAAC TTTTTGTGCG CAATCATGGC CTTCG
B 2Bubble	CGTAAGATAA CGTAATTAAC TTGTGCGCAA TCATGGCCTT CG
C	CGTAAGATAA CGTAATTAAT AACAGGGCAG
C-1	CGTAAGATAA CGTAATTAAT AACAGGGCA
C-2	CGTAAGATAA CGTAATTAAT AACAGGGC
D	GTGTTGCCGT CTACATGCTT GATTATTCTC GATGTATTCA GTGCGCAATC ATGGCCTTCG
E	GAGAATAATC AAGCATGTAG ACGGCAACAC
F	GCCGTCCGGT CGACTGTGCA GACGCAGGCA GGTCGTCAGC AGCGCACACT GCCGACACTC TGGATGCTAG GACAGGCCGG TCGCGTCGCG
G	CGCGACGCGA CCGGCCTGTC CTAGCATCCA GAGTGTCGGC AGTGTGCGCT GCTGACGACC TGCCTGCGTC TGCACAGTCG ACCGGACGGC
H	TTTTTTTTTT TTTTTTTTTT TTTTTTTTTT

Oligonucleotides were resuspended in TE buffer, and 5' radiolabeled with ^{32}P, using $\gamma^{32}P$-labeled ATP (PerkinElmer) and T4 Polynucleotide Kinase (New England Biolabs). Radiolabeled oligonucleotides were annealed to non-labeled conjugate strands, as indicated in Table 2, at a ratio of at least 1:3 in 10mM Tris HCl, pH 7.9, 100mM KCl, 5mM $MgCl_2$, by

controlled temperature decrease in a thermocycler from 95°C to 25°C. Annealed substrates were separated on a polyacrylamide gel (8% or 15%) in 89mM Tris borate, pH 8.3, 2mM EDTA (TBE). Substrates were extracted from excised gel fragments by electroelution with the S&S Elutrap system (Whatman). Purified substrates were stored at -20°C in TBE with 5mM $MgCl_2$.

Table 2: Composition of Substrates Used In Chapter 2

Substrate Name	Component Oligonucleotides
Single-stranded DNA	A30
3' Extension	A30 and B
5' Extension	A30 and C
Fork	A30 and D
60 Basepair Duplex	A30 and BC

2.4.5 Gel Shift Helicase Assay

For helicase assays, enzyme was incubated with 0.2nM radiolabeled substrate for 30 minutes at 27°C in 2.5mM Tris acetate, pH 7.9, 2.5mM potassium acetate, 8mM magnesium acetate, 50μg/mL BSA, 1mM DTT, and 1mM ATP (or indicated nucleotide cofactor) in a total reaction volume of 20μL. Reactions were stopped by the addition of 4μL stop mix to a final concentration of 0.2% SDS, 5% sucrose, 10mM EDTA, 0.1mg/mL bromophenol blue, and 0.1mg/mL xylene cyanol. Reaction products were separated on an 8% TBE polyacrylamide gel, and subjected to phosphorimaging analysis.

2.4.6 Gel Shift Annealase Assay

Annealing assays were conducted as detailed for helicase assays, except that enzyme was incubated with 0.2nM single-stranded radiolabeled oligonucleotide and 0.2nM non-labeled complementary oligonucleotide. Radiolabeled oligonucleotide was purified as

described previously. Complementary non-labeled oligonucleotide was resuspended in TE upon receipt from IDT. Radiolabeled and complementary non-labeled oligonucleotides were separately heated for 5 minutes at 80°C before addition to the reaction.

3. Further Biochemical Characterization of RecQ4 Helicase

3.1 Introduction

We have established that RecQ4 possesses a robust helicase activity. Further, this activity is clearly stimulated by a single-stranded region 3' of the double-stranded region to be unwound, which indicates that it moves in a 3' to 5' direction. This is consistent with the activities of the other members of the RecQ family (reviewed in Seki et al., 2008; Bohr, 2008; and Bachrati and Hickson, 2008). To distinguish RecQ4 from other RecQ helicases, and to better understand its function in the context of DNA replication and repair, it is necessary to further characterize RecQ4's helicase activity. With this in mind, we applied the enzyme to a variety of substrates designed to dissect the precise parameters of helicase activity. We also applied the enzyme to substrates that mimic different intermediates encountered in the process of replicating and repairing DNA.

3.2 Results and Discussion

3.2.1 Minimum Single-stranded Region Necessary for Helicase Activity

We have demonstrated that RecQ4 acts as a helicase on a substrate with a 30 basepair duplex region and a 3' single-stranded region of 30 nucleotides (hereafter referred to as the "3' extension substrate"). A substrate consisting of 60 basepairs of double-stranded DNA, and one made of a 30 basepair duplex region and a 5' single-stranded tail of 30 nucleotides were not significantly acted upon. To determine the minimum 3' single-stranded region necessary for helicase activity, we applied RecQ4 to substrates with a variety of tail lengths (Figure 9). These substrates all possessed a 30 basepair duplex region and a 3' single-stranded tail of 30 nucleotides or shorter (as indicated).

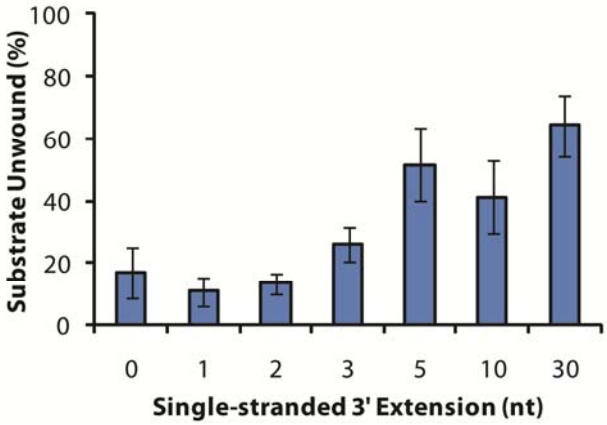

Figure 9: Helicase activity with substrates of varying tail lengths. RecQ4 maximally unwinds substrates with a 3' single-stranded region of 5 nucleotides or longer. At a lower level of activity, it also unwinds substrates lacking a single-stranded region, or with one of only 2 nucleotides or shorter. Radiolabeled oligonucleotide substrates were incubated with RecQ4 or an equal volume non-enzyme containing storage buffer, separated by PAGE, and quantified by phosphorimaging analysis. Background was removed by subtracting the average unwinding in the absence of enzyme from the average unwinding in the presence of RecQ4. Error bars indicate standard deviation (n≥3).

Substrates fell into two general categories with regards to activity. The first consists of those with a 3' tail of 5 nucleotides or longer. These were preferred substrates, and between 40% and 70% of the substrates were unwound above background. While there was not significant variance in activity on substrates with tails of 5 nucleotides or longer, the greatest activity was seen on the previously tested 3' extension substrate (with a 30 nucleotide single-stranded region). This suggests that 5 nucleotides is approximately the footprint size of RecQ4.

The second category consists of substrates with tails 2 nucleotides long or shorter. These exhibited much less activity – less than 20% unwound above background. Nonetheless, these substrates were subject to helicase activity, albeit at a low level. Notably, helicase activity was observed even in the complete absence of a 3' single-stranded region.

RecQ4 is thus able to unwind 30 basepairs of DNA in the absence of a single-stranded region, though helicase activity is greatly stimulated by the presence of a 3' tail of at least 5 nucleotides.

The activity observed on the substrate with a 3 nucleotide single-stranded region was about 26% above background. This places it between the two categories, but more consistent with the low activity on substrates with tails 2 nucleotides or shorter. The substrate with a 4 nucleotide single-stranded region has yet to be examined.

3.2.2 Helicase Activity on Bubble Substrates

RecQ4 shows maximal helicase activity on substrates with a 3' tail of 5 nucleotides or longer. In the substrates examined thus far, the single-stranded ends have been unconstrained. Unconstrained DNA ends are biologically problematic, and so they are efficiently metabolized by homologous recombination or other repair mechanisms (reviewed in Longhese et al., 2010 and Kass and Jasin, 2010). Given its role in replication, it is possible that RecQ4 may not actually encounter such substrates. In order to examine helicase activity in the context of more biologically relevant substrates, we incubated RecQ4 with substrates containing bubbles of various sizes. These substrates all possess 20 basepairs of duplex DNA on either side of the unpaired bubble region. Thus each substrate has a single-stranded region 3' to either duplex end, but there are no free 3' ends. As a note of clarification, when referring to the size of the bubble the term "nucleotide" is used, as the nucleotides in the bubble are not paired (and hence are not basepairs); however, the indicated number of nucleotides is present on each side of the bubble. Thus the 5 nucleotide bubble has 5 unpaired nucleotides on each strand.

RecQ4 was able to efficiently unwind a substrate with a 20 nucleotide bubble, at levels comparable to the 30 nucleotide 3' extension substrate (Figure 10). The 2 nucleotide bubble substrate was not a preferred helicase substrate, and was only about 10% unwound above background. The 5 nucleotide bubble substrate was clearly unwound less efficiently than the 20 nucleotide bubble, but at 44% unwound above background it was still a decent substrate for RecQ4 helicase activity. This is consistent with the data from substrates with unconstrained single-stranded regions.

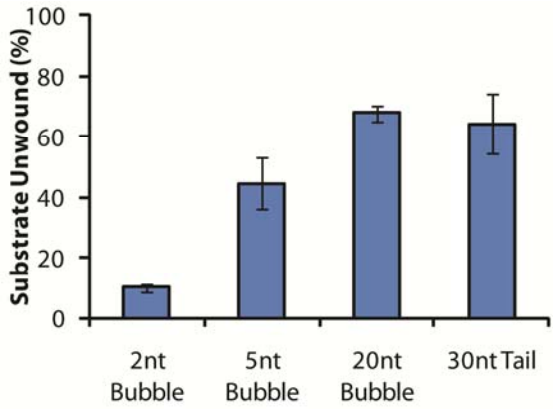

Figure 10: Helicase activity with substrates having various sizes of bubble. RecQ4 maximally unwound the substrate with a 20 nucleotide bubble. Activity was much reduced when the bubble was only 2 nucleotides. Helicase activity on these substrates was assayed and quantified as in Figure 9. Error bars indicate standard deviation (n=3).

3.2.3 Helicase Activity on Fork Substrates

We also examined the activity of RecQ4 on several variations of the fork substrate tested earlier (Figure 11). These variations are similar to the fork substrate, except that one arm (or both) is double-stranded. Thus they resemble portions of D-loops encountered as intermediates in replication and repair. The 5' flap substrate (a fork with a 5' single-stranded arm) only resulted in low helicase activity (Figure 11A, lanes 3 and 4), comparable to the 2

nucleotide bubble substrate or other non-preferred substrates. The contrasting 3' flap substrate was highly unwound at levels similar to the 3' extension substrate (lanes 5 and 6). Both single-stranded DNA and simple fork substrates were detected as products, indicating that RecQ4 unwinds the substrate in both directions available to it (down the center of the substrate, and along the 5' double-stranded branch).

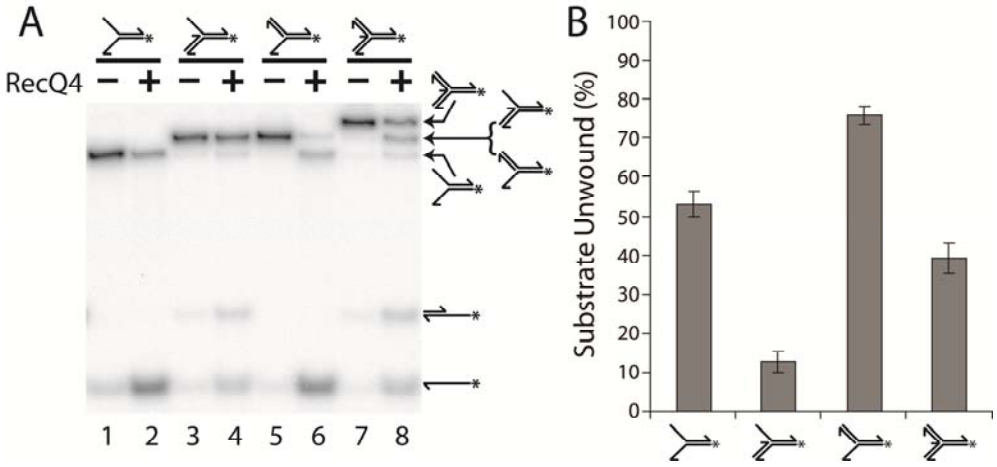

Figure 11: Helicase activity on fork substrates. *(A)* RecQ4 unwinds the 3' flap substrate (lanes 5 and 6), but not the 5' flap substrate (lanes 3 and 4). The duplex fork substrate (lanes 7 and 8) was also unwound by RecQ4. *(B)* Quantification of RecQ4's activity on substrates. Helicase assays as seen in *A* were quantified by phosphorimaging analysis. The fraction of fully annealed substrate lost after incubation without enzyme was subtracted from the fraction of fully annealed substrate lost after incubation with RecQ4. Error bars in these experiments indicate standard deviation (n=3).

Interestingly, RecQ4 unwound a duplex fork substrate (a fork with both the 3' and 5' branches double-stranded) (Figure 11A, lanes 7 and 8). Given that RecQ4 exhibits a low level of activity on a 30 basepair substrate with no 3' single-stranded region, it is possible the same basal level of activity could be observed a duplex fork substrate. However, the extent of the loss of the initial substrate is greater than observed with the simple 30 basepair duplex substrate (compare Figures 9 and 11B), implying that another process is at work. The

presence of partial reaction products, including 3' or 5' flap, fork, and 3' or 5' extension substrates, suggests that RecQ4 may instead take advantage of transiently exposed single-stranded regions at the three-way junction to enter and unwind along any of the three duplex arms.

3.2.4 Helicase Activity is Observed on Gap Substrates

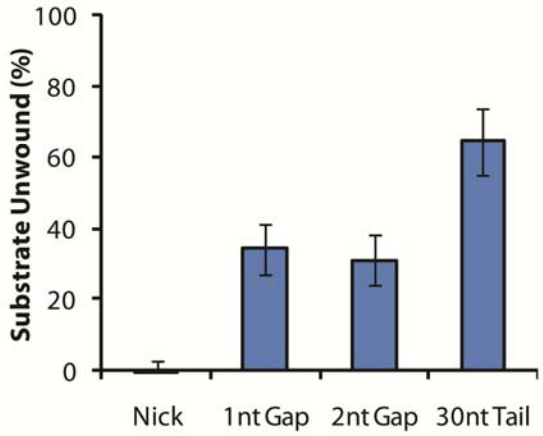

Figure 12: Helicase activity on nicked and gapped substrates. RecQ4 unwound 60 basepair duplex substrates with a central single or double nucleotide gap at an intermediate level. RecQ4 did not unwind a 60 basepair duplex substrate with a central nick. Helicase activity on these substrates was assayed and quantified as in Figure 9. Error bars indicate standard deviation (n=3).

We further investigated the effect of RecQ4 on substrates with minimally exposed single-stranded DNA. To that end three substrates were designed based on a 60 basepair duplex DNA, but with one strand broken in the center either by a nick, a single nucleotide gap, or a dinucleotide gap. RecQ4 failed to demonstrate any helicase activity on the nicked duplex DNA (Figure 12). However, the substrates with an actual gap, whether single- or di-nucleotide, were unwound at levels similar to that seen with the duplex fork substrate examined previously (Figure 11B). Thus the gap substrates did not exhibit the maximal activity seen on the substrates with 3' tails of 5 nucleotides or longer; neither did they exhibit

the minimal activity seen on 3' extension substrates with 2 nucleotides or less. The intermediate level of activity seen on the gap substrates has interesting implications for the mechanism of RecQ4 helicase activity.

3.3 Concluding Remarks

We have further characterized the substrate preference of RecQ4's helicase activity. We have observed that substrates with a 3' tail of 5 nucleotides or longer are unwound well, as are substrates containing an interior 20 nucleotide bubble, and a fork substrate with a 3' single-stranded branch. These substrates all possess a single-stranded region 3' of the duplex DNA to be unwound. This confirms that RecQ4 is a helicase that moves in a 3' to 5' direction. The activity on the bubble substrate also indicates that the 3' single-stranded region does not have to be unconstrained. These results are consistent with the activity of the other RecQ helicases (Bohr, 2008; Bachrati and Hickson, 2008; Seki et al., 2008).

A 3' single-stranded tail is not absolutely necessary for unwinding DNA, though one of 5 nucleotides or greater clearly stimulates RecQ4's helicase activity. Thus a low level of unwinding is observed on the 5' flap substrate, where the 3' branch is double-stranded and the 5' branch is single-stranded. Low levels of unwinding are also observed on 3' extension substrates with single-stranded tails shorter than 3 nucleotides, or even entirely lacking a 3' single-stranded region. It is possible that the action of RecQ4 on such substrates lacking a sufficient 3' tail is enabled by fraying at the ends of the duplex DNA. RecQ4 would take advantage of transiently exposed 3' single-stranded DNA to unwind the rest of the substrate. Unwinding of such substrates would be limited by the rate of DNA fraying, which accounts for the lower activity level compared to substrates which allow RecQ4 to load on a 3' tail.

RecQ4 shows more activity on the duplex fork substrate than on either the simple 30 basepair duplex, or the 5' flap substrate. This increased activity could be attributable to the greater number of double-stranded ends of the duplex fork substrate (three on the duplex fork, two on the simple duplex and 5' flap). More ends would provide more sites for fraying, producing more potential starting sites for helicase activity. Alternatively, it may be the case that the duplex fork substrate is sterically constrained at the junction of the two duplex branches. The crowding forces the branches to move apart, creating more transiently exposed single-stranded DNA in the center of the substrate, which is then accessible to RecQ4 helicase activity.

The evidence from the gapped and nick substrates is consistent with the second interpretation of action on the duplex fork substrate. RecQ4 unwinds substrates with a one or two nucleotide gap, but not a nearly identical substrate with a simple nick. The gap substrates may possess a greater flexibility, allowing the helicase to access the duplex DNA. The nick substrate would lack this flexibility, perhaps due to pi stacking by the bases (Lukin and de Los Santos, 2006; Lin et al., 2009). However, the fact that no activity at all is observed, even though the substrate possesses two 30 basepair duplex regions, suggests a reduced contribution of end fraying to the unwinding of the gap substrates. Further, activity on the one and two nucleotide gap substrates is greater than that observed on the one or two nucleotide 3' tail substrates. This suggests that the stimulative effect on helicase activity seen with substrates having longer 3' tails may be provided by interactions with the flexible backbone of the single-stranded region. Such interactions would also be present with the duplex fork and gap substrates, though the 3' single-stranded regions are clearly less than 5 nucleotides.

3.4 Methods

3.4.1 Oligonucleotide Substrate Preparation

PAGE-purified DNA oligonucleotides were obtained from Integrated DNA Technologies (sequences shown in Table 1). The oligonucleotide composition of each substrate is indicated in Table 3. The fork DNA substrates used in section 3.2.3 were prepared as indicated in section 2.4.4. Other DNA substrates were prepared as in section 2.4.4 with the following modifications. Annealed substrates were separated on a polyacrylamide gel (8% or 15%) in 89mM Tris borate, pH 8.3, 0.5mM EDTA. After extraction from the gel by means of electroelution, purified substrates were stored at 4°C in 89mM Tris borate, pH 8.3, 0.5mM EDTA, 10mM $MgCl_2$.

Table 3: Composition of Substrates Used In Chapter 3

Substrate Name	Component Oligonucleotides
3' Extension (30nt Tail)	A30 and B
10nt Tail	A10 and B
5nt Tail	A5 and B
3nt Tail	A3 and B
2nt Tail	A2 and B
1nt Tail	A1 and B
30 Basepair Duplex (0nt Tail)	A and B
20 Nucleotide Bubble	A30 and B 20Bubble
5 Nucleotide Bubble	A 5Bubble and B 5Bubble
2 Nucleotide Bubble	A 2Bubble and B 2Bubble
Fork	A30 and D
3' Flap	A30, D, and E
5' Flap	A30, C, and D
Duplex Fork	A30, C, D, and E
Nick	A30, B, and C
1nt Gap	A30, B, and C-1
2nt Gap	A30, B, and C-2

3.4.2 Gel Shift Helicase Assay

Helicase assays for the fork substrates were performed as indicated in section 2.4.5. Helicase assays on other substrates were performed as indicated in section 2.4.5, with the following modifications. Reactions were done in 25mM Tris acetate, pH 7.9, 2.5mM potassium acetate, 8mM magnesium acetate, 50μg/mL BSA, 1mM DTT, and 1mM ATP, in a total reaction volume of 20μL. After stopping the reaction, products were separated on a 8% TBE polyacrylamide gel, and subjected to phosphorimaging analysis.

4. The Binding Preferences of RecQ4

4.1 Introduction

It has been established that RecQ4 is 3' to 5' helicase. The helicase activity is stimulated by the presence of single-stranded DNA 3' relative to the double-stranded DNA to be unwound. While insight has been gained into the types of substrates RecQ4 prefers to unwind, much remains to be determined about the manner in which RecQ4 interacts with the substrates it acts upon. We used a series of binding assays to further examine the interactions of RecQ4 with DNA substrates.

4.2 Results and Discussion

4.2.1 Stable Complex Formation with DNA in the Presence of AMPPNP

In considering the mechanism of RecQ4 helicase activity, it is evident that the phosphorylated state of the nucleotide cofactor may alter the mode of interaction between the enzyme and the substrate. Differences between the enzyme-ATP-DNA complex and the enzyme-ADP-DNA complex could be critical to utilizing ATP hydrolysis to unwind DNA. To this end, we examined RecQ4's DNA binding as a function of nucleotide cofactor. Wildtype RecQ4 was incubated with the 3' extension substrate in the absence of nucleotide cofactor, or in the presence of ADP or AMPPNP, and reaction products were analyzed by electrophoresis on a native polyacrylamide gel (Figure 13). Complex formation evidenced by electrophoretic mobility shift was observed in the presence of AMPPNP (lane 4), but not in the presence of ADP (lane 3) or in the absence of nucleotide cofactor (lane 2). One may reasonably extend this result to say that RecQ4 tightly binds DNA in the presence of ATP. Hydrolysis of the ATP then yields ADP, which reduces the affinity of RecQ4 for DNA.

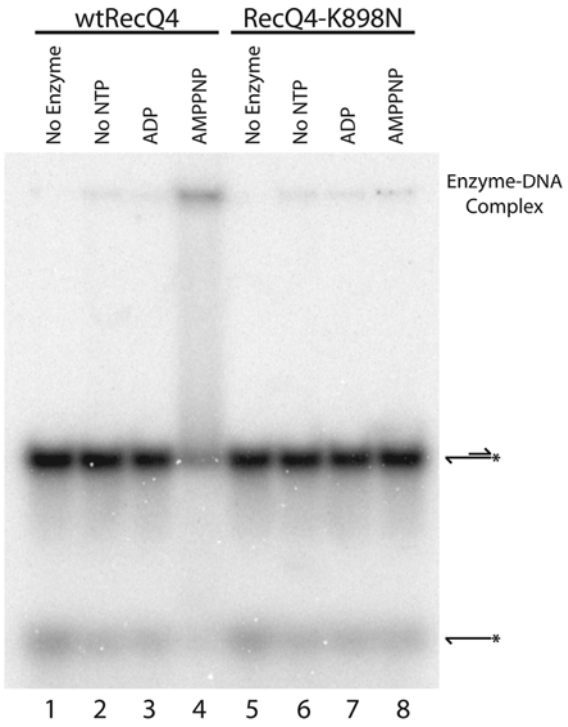

Figure 13: RecQ4 forms a stable complex in the presence of AMPPNP. wtRecQ4 (lanes 2-4) and RecQ4-K898N (lanes 6-8) were incubated with the 3' extension substrate in the absence of nucleotide (lanes 2 and 6), or in the presence of ADP (lanes 3 and 7) or AMPPNP (lanes 4 and 8). Significant binding was only seen with wtRecQ4 and AMPPNP.

Work by Liu *et al* has suggested that RecQ4 may interact with ATP outside the predicted helicase domain, through the Sld2 domain (Xu and Liu, 2009). With two potential ATP binding sites, it is relevant to determine if the helicase domain is implicated in the formation of the AMPPNP-DNA complex. To examine this, we repeated the above binding assays, this time using RecQ4-K898N, which is defective in both ATPase and helicase activities due to a point mutation in the helicase domain. The mutant did not cause a shift under any of these conditions (Figure 13, lanes 6-8). This establishes that the ATP binding site of the helicase domain is responsible for the formation of this stable enzyme-substrate complex. Further, if the RecQ4-K898N mutation were to cause a defect in ATP hydrolysis,

rather than ATP binding, one would expect to see the formation of a complex identical to that of the wildtype enzyme with the non-hydrolyzable ATP analog, AMPPNP. These results indicate that the defects seen in the ATPase and helicase activities of the mutant are based on an inability to bind ATP, rather than to hydrolyze it.

4.2.2 RecQ4 Binds Single-stranded DNA

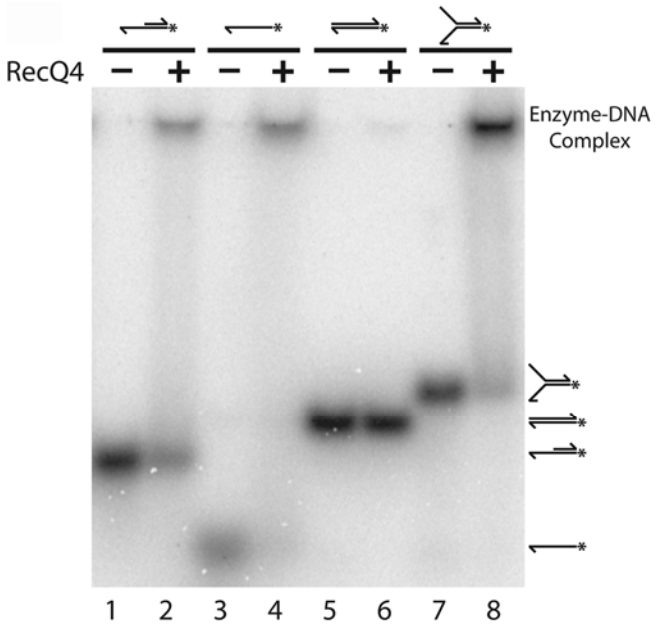

Figure 14: RecQ4 forms a stable complex with single-stranded DNA, not double-stranded DNA. Substrates were incubated with AMPPNP in the presence (even lanes) or absence (odd lanes) of RecQ4. Stable complexes were formed by RecQ4 with the 3' extension substrate (compare lanes 1 and 2), single-stranded substrate (lanes 3 and 4), and the fork substrate (lanes 7 and 8), all of which contain ssDNA regions. No stable complex was detected with RecQ4 and the duplex substrate (lanes 5 and 6).

The 3' extension substrate has both single-stranded and double-stranded regions. It is conceivable that the stable complex formed by RecQ4 in the presence of AMPPNP is formed with the either or both of these regions, or the junction between the two. To determine which portion of the 3' extension substrate is critical for stable RecQ4 binding,

enzyme was incubated with AMPPNP and single-stranded, duplex, and fork substrates (Figure 14, lanes 3-8). RecQ4 bound both the single-stranded and fork substrates. However, it was not observed to bind the duplex substrate. Thus in the presence of AMPPNP RecQ4 forms a stable complex with single-stranded DNA, but not with double-stranded DNA.

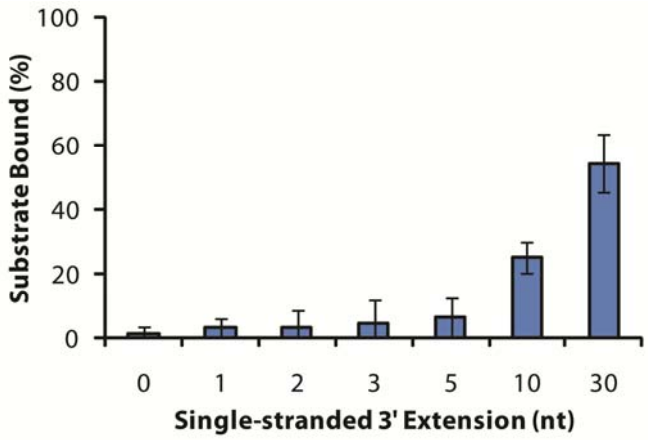

Figure 15: Binding activity with substrates of varying tail lengths. RecQ4 showed maximal stable complex formation with the substrate having a 3' single-stranded region 30 nucleotides long, and less so with the substrate having one only 10 nucleotides long. It did not significantly bind substrates with 3' single-stranded regions of 5 nucleotides or less. Radiolabeled oligonucleotide substrates were incubated with RecQ4 or an equal volume of non-enzyme containing storage buffer, separated by native PAGE, and quantified by phosphorimaging analysis. The fraction of mobility-shifted substrate after incubation without enzyme was subtracted from the fraction of mobility-shifted substrate after incubation with RecQ4. Error bars in these experiments indicate standard deviation (n=3).

In order to investigate the amount of single-stranded DNA necessary to form a stable complex with RecQ4 and AMPPNP, we incubated the enzyme with 3' extension substrates that had varying lengths of 3' single-stranded DNA (Figure 15). This is the same series of substrates used in section 3.2.1 to determine the minimum single-stranded region necessary for helicase activity. Maximal binding levels were seen with the original 3' extension substrate, having a single-stranded region 30 nucleotides long. The substrate with a

10 nucleotide tail demonstrated appreciable levels of binding, though significantly less than seen with the 30 nucleotide tail. This is in contrast with substrates that had tails 5 nucleotides long or shorter, which showed minimal complex formation. Although a single-stranded region 5 nucleotides long is sufficient to stimulate helicase activity (Figure 9), it is insufficient for the formation of a stable enzyme-DNA-AMPPNP complex.

4.2.3 Confirmation of Binding Trends by Filter Binding Assay

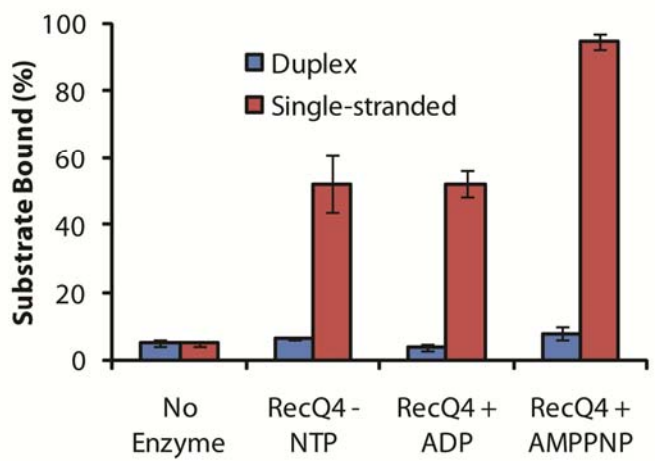

Figure 16: Double-filter assays confirm the results from electrophoretic mobility shift assays. RecQ4 demonstrated maximal binding with single-stranded DNA and AMPPNP, though lesser binding of single-stranded DNA was observed with ADP and without nucleotide cofactor. Duplex DNA was not bound under any of the conditions examined. Error bars indicate standard deviation (n=3).

Electrophoretic mobility shift assays allow observation of tight enzyme-substrate complexes. However, less stable complexes may not survive the process of electrophoretic separation, and thus may not be captured by the technique. In order to detect the formation of less stable complexes, as well as to confirm the results already stated, we performed double-filter binding assays with RecQ4, single-stranded and double-stranded DNA substrates, and a variety of nucleotide cofactors (Figure 16). RecQ4 did not bind double-

stranded DNA under any of the conditions examined. It showed almost complete binding with single-stranded DNA and AMPPNP. Binding of single-stranded DNA was observed at significantly lower levels in the absence of nucleotide cofactor, and in the presence of ADP. This confirms the general trends revealed by the electrophoretic mobility shift assays, in that RecQ4 preferentially binds single-stranded DNA in the presence of AMPPNP. It is significant to note, however, that RecQ4 also interacts with single-stranded DNA in the absence of AMPPNP, albeit with a lower affinity.

4.2.4 RecQ4 Binding on Bubble, Fork, and Gapped Substrates

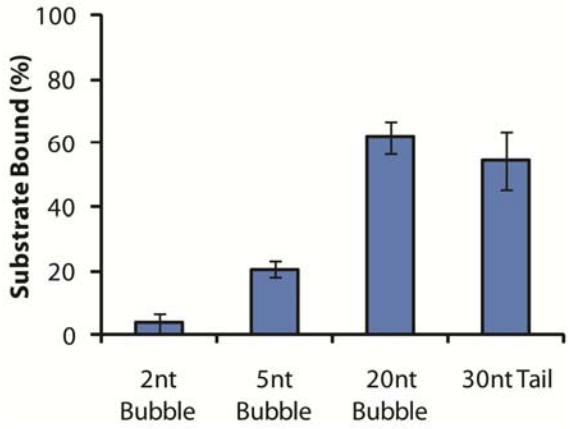

Figure 17: Binding activity with substrates having various bubble sizes. RecQ4 showed maximal stable complex formation with the 20 nucleotide bubble substrate, and did not significantly bind the 2 nucleotide bubble substrate. The 5 nucleotide bubble substrate was bound at an intermediate level. Binding activity on these substrates was assayed and quantified as in Figure 15. Error bars in these experiments indicate standard deviation (n=3).

Electrophoretic mobility shift assays were performed with RecQ4 and AMPPNP on other substrates previously investigated for helicase activity. The 20 nucleotide bubble substrate showed complex formation at levels comparable to that seen with the 30 nucleotide 3' extension substrate (Figure 17). This confirms other data showing equivalent

binding levels between 3' extension substrates with single-stranded tails 20 and 30 nucleotides long (data not shown). The two nucleotide bubble substrate did not show appreciable binding, while the 5 nucleotide bubble substrate was bound at a level intermediate to the 2 and 20 nucleotide bubble substrates. Thus the binding and helicase assays on the bubble substrates yield similar results. It is interesting that the 5 nucleotide bubble substrate was subject to binding, while the 5 nucleotide 3' extension substrate was not. This suggests either that the bubble itself enhances binding, or that 10 nucleotides need not be on a single strand to stimulate binding (in contrast to the 10 nucleotide tail substrate – see Figure 15). In either case, this result may prove significant for future analysis of RecQ4's binding preferences.

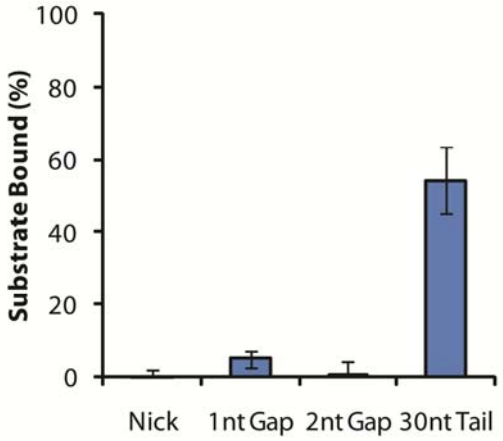

Figure 18: Binding activity with nicked and gapped substrates. RecQ4 did not demonstrate significant binding substrates with a central nick, or with a central gap of one or two nucleotides. Binding activity on these substrates was assayed and quantified as in Figure 15.Error bars in these experiments indicate standard deviation (n=3).

Binding studies were also conducted with the variations of fork substrates, as well as the single and di-nucleotide gap substrates, and the nicked substrate. RecQ4 bound the fork substrates with one single-stranded branch, whether 3' and 5', at equal levels, but did not

significantly bind the duplex fork substrate (data not shown). The 60 basepair duplex substrate with a central nick was not bound (Figure 18). Substrates that extended this nick to an actual gap one or two nucleotides long did not show binding either. This is consistent with previous results indicating single-stranded DNA is necessary for formation of a stable complex with the enzyme in the presence AMPPNP.

4.3 Concluding Remarks

It has been shown that RecQ4 binds single-stranded DNA, but not completely duplex DNA. Double-filter binding assays show that the interaction with single-stranded DNA is stimulated by AMPPNP. However, the more stringent electromobility shift assays only detect single-stranded DNA binding when AMPPNP is a cofactor. This indicates that the complex formation in the presence of AMPPNP is a qualitatively different mode of binding, which is abolished by both the absence of cofactor and the substitution of ADP for AMPPNP. Further, a point mutation inactivating the ability of the helicase domain to bind AMPPNP also abolishes complex formation with single-stranded DNA. This establishes that the helicase activity and AMPPNP-induced complex formation with single-stranded DNA are related. It is possible that these activities are not derived from the same domain, but there is at least communication between the two domains responsible for them.

While the profiles of preferred helicase substrates and preferred binding substrates overlap, they are not identical. The differences may prove instructive. Less single-stranded DNA is necessary for proficient helicase activity than for complex formation. Substrates such as the duplex fork or gapped substrates are readily unwound, but are not subject to stable complex formation at all. This suggests that the interactions necessary for helicase

activity are primarily with the DNA backbone, requiring only a single nucleotide of actual 3' single-stranded DNA for activity, provided more DNA is present on the 3' end of the single-stranded nucleotide. In contrast to this, the interactions necessary for formation of a stable enzyme-DNA-AMPPNP complex are more extensive, and require between 5 and 10 nucleotides of single-stranded DNA.

In contemplating the functional role of the enzyme-substrate-AMPPNP complex, it is tempting to think of it as a pre-unwinding step of the helicase mechanism. The enzyme, after binding ATP, but prior to hydrolyzing it, would bind single-stranded DNA tightly. The energy from ATP hydrolysis could then be used to propel the enzyme forward, driving apart adjacent double-stranded DNA. Alternatively, the hydrolysis could alter the conformation of the enzyme-DNA complex, contorting nearby duplex DNA to make strand separation more favorable. But this hypothesis is contradicted by the fact that substrates that are clearly subject to helicase activity are not tightly bound. Therefore it is unlikely that the enzyme-DNA-AMPPNP complex observed by these binding assays is a key step in the helicase mechanism.

It is significant that AMPPNP both induces the formation of a stable complex between RecQ4 and single-stranded DNA, and partially inhibits the annealing of complementary strands (compare Figures 8 and 13). Evidently, the stable association of the enzyme with single-stranded DNA interferes with the ability to rewind complementary strands. This result is at first non-intuitive, as annealing enzymes such as Rad52 and Rad59 do so by binding single-stranded DNA (Wu et al., 2006). This should reduce the available degrees of freedom, increasing the probability of forming basepairs with the appropriate complementary strand. However, it has already been established that RecQ4 is able to

interact with single-stranded DNA in the absence of AMPPNP, as seen by double-filter binding assay. This looser binding may be responsible for the enzyme's annealing activity, using the passive mechanism just mentioned. Sld2, the homolog of RecQ4's N-terminal Sld2 domain, has recently been shown to both bind single-stranded DNA and stimulate its annealing (Kanter and Kaplan, 2010). This is consistent with the data here presented, and provides insight into which domain of RecQ4 is responsible for the looser interactions associated with annealing. The enzyme is therefore capable of two apparently mutually exclusive interactions with single-stranded DNA: one in the absence of AMPPNP, which promotes annealing; and the formation of a more stable complex in the presence of AMPPNP, which is inhibitory to annealing. The biological relevance of these two interactions, as well as their precise mechanisms, remains unclear, and should be of interest to future research.

4.4 Methods

4.4.1 Oligonucleotide Substrate Preparation

Oligonucleotide substrates, composed as indicated in Table 4, were prepared as indicated in section 3.4.1. Sequences used are indicated in Table 1.

Table 4: Composition of Substrates Used In Chapter 4

Substrate Name	Component Oligonucleotides
3' Extension (30nt Tail)	A30 and B
10nt Tail	A10 and B
5nt Tail	A5 and B
3nt Tail	A3 and B
2nt Tail	A2 and B
1nt Tail	A1 and B
30 Basepair Duplex (0nt Tail)	A and B
Single-stranded DNA	A30
60 Basepair Duplex	A30 and BC
20 Nucleotide Bubble	A30 and B 20Bubble
5 Nucleotide Bubble	A 5Bubble and B 5Bubble
2 Nucleotide Bubble	A 2Bubble and B 2Bubble
Fork	A30 and D
3' Flap	A30, D, and E
5' Flap	A30, C, and D
Duplex Fork	A30, C, D, and E
Nick	A30, B, and C
1nt Gap	A30, B, and C-1
2nt Gap	A30, B, and C-2

4.4.2 Electrophoretic Mobility Shift Binding Assay

Electrophoretic mobility shift assays were conducted as described for helicase assays, with the following modifications. After 30 minutes, 4μL of non-denaturing loading mix was added to each reaction to a final concentration of 5% sucrose, 33μg/mL bromophenol blue, and 33μg/mL xylene cyanol, and reactions were placed on ice. Reaction products were

separated on a 6% Tris borate polyacrylamide gel, and subjected to phosphorimaging analysis.

4.4.3 Double-filter Binding Assay

Double-filter binding assays were done with DE81 paper and nitrocellulose according to Wong and Lohman (Wong and Lohman, 1993), with binding buffer (25mM Tris acetate, pH 7.9, 25mM potassium acetate, 8mM magnesium acetate, and 1mM DTT). Filters were pre-washed with 500μL binding buffer drawn through at 30 kilopascals of vacuum. Reactions (50μL) were carried out in binding buffer with 3.4nM RecQ4, 0.25nM radiolabeled DNA substrate (as indicated), 1mM nucleotide cofactor (as indicated) and 50μg/mL BSA at 27°C for 30 minutes. Each reaction was drawn through at 30 kilopascals of vacuum, and then washed with 500μL binding buffer with 500mM potassium acetate. Nitrocellulose and DE81 sheets were dried and subjected to phosphorimaging analysis.

5. Summary and Perspectives: The Biochemistry and Biology of RecQ4

5.1 The Comparative Biochemistry of RecQ4

RecQ4 is the last of the RecQ helicase family members found in humans to demonstrate helicase activity. This is not simply a function of lack of observation, as the first two papers examining the *in vitro* biochemistry of RecQ4 looked for helicase activity and failed to detect it (Macris et al., 2006; Yin et al., 2004). In 2009, we and two other groups independently published papers having observed helicase activity for the enzyme (Capp et al., 2009; Suzuki et al., 2009; Xu and Liu, 2009). The several approaches to examining the biochemistry of RecQ4 have yielded results that do not perfectly parallel each other. Accordingly, comparative analysis of these publications can yield further insights into the enzyme's activity.

In 2004, RecQ4 was first isolated by immunoprecipitating the endogenous enzyme from HeLa cells. This enzyme was unable to act as a helicase or DNA translocase, though DNA-stimulated ATP hydrolysis was detected (Yin et al., 2004). In 2006, human RecQ4 was expressed in *E. coli*, and purified by several ion exchange and affinity chromatography steps. Helicase activity was not observed, though both DNA-stimulated ATP hydrolysis and DNA annealing activities were (Macris et al., 2006). It is not immediately evident why neither group detected helicase activity. Both detected ATP hydrolysis, indicating that the enzyme was not misfolded. In the case of the second group, the use of a heterologous expression method may be implicated. If post-translational modifications or eukaryotic-specific protein folding mechanisms are important for activation of the helicase, these would not be

provided by expression in *E. coli*. However, this does not clarify why the initial isolation of RecQ4 from HeLa cells failed to demonstrate helicase activity. It is possible that the use of immunoprecipitation in purification resulted in an inhibition of helicase activity. The potential effect of an antibody on enzymatic activity was not examined, and therefore cannot be eliminated as a cause of inactivation.

Helicase activity, in addition to annealing and ATP hydrolysis, was detected by Xu et al. (Xu and Liu, 2009), who also purified human RecQ4 using a heterologous *E. coli* expression system. The helicase activity was masked by strong annealing activity, and so required an excessive amount of complementary single strand for observation. As such, the reaction observed is more akin to DNA strand exchange than the helicase activity commonly found in other RecQ homologs. The inability of Macris et al. to detect helicase activity from similarly expressed and purified protein may nonetheless be due to failure to include such a trap (Macris et al., 2006). Helicase activity was determined to proceed in a 3′-5′ direction, like the other RecQ helicases, but also to act on blunt-ended substrates, which is uncommon in the family. This helicase/strand-exchange activity was attributed to both the helicase domain and the Sld2 domain in the N-terminus of the protein, the latter contributing most of the activity. The Sld2 domain was found to interact with ATP using both spin column assays and UV crosslinking. This is unprecedented, as neither the *S. cerevisiae* protein Sld2, nor the homologous domain in RecQ4, are predicted to have any helicase activity. An interaction between the Sld2 domain and ATP would also suggest the presence of completely new ATP binding motifs within the domain. While not inconceivable, the unexpected nature of these results renders further confirmation critical.

There is significant divergence between our results presented here and those of Xu et al. Sld2 domain-derived helicase/strand-exchange activity was not found in *Drosophila* RecQ4. Inactivation of the *Drosophila* helicase domain abolished both helicase activity and AMPPNP-based complex formation and inhibition of annealing. It may however be that this is simply reflective of differences between species. If this alternate source of helicase activity is upheld by future investigations, it could explain why humans and mice entirely lacking the helicase domain are still viable. However, recent work showing that the Sld2 protein stimulates DNA annealing (Kanter and Kaplan, 2010) provides an alternate interpretation of the helicase/strand-exchange activity. The Sld2-derived strand annealing may be a component of the observed strand-exchange activity, when combined with binding to single-stranded DNA that was either pre-existing within the substrate, or was generated as a result of thermal fraying (as observed in Figure 9 with the 30 basepair duplex substrate). RecQ4's annealing activity is inhibited by AMPPNP, and so one would predict that a strand-exchange activity would also be inhibited by AMPPNP, but not by the absence of nucleotide. Xu et al. only examined the Sld2 derived helicase/strand-exchange activity in the presence of either ATP or AMPPNP, and not in the absence of nucleotide cofactor, and so this prediction remains untested.

It is notable that Xu et al. also observed helicase activity in the context of a chimera consisting of the helicase domain of RecQ4 and the regions external to the helicase domain from RecQ5 (Xu and Liu, 2009). The helicase domain itself is thus shown to be competent for helicase activity, and there is nothing in the domain that inherently prevents it from acting as a helicase. This suggests that when RecQ4 is expressed by a significantly heterologous system the other domains of the enzyme inhibit unwinding by the helicase

domain. This inhibition may be due to misfolding or the absence of appropriate post-translational modifications, or other as-yet undetermined factors, and should be the subject of further analysis.

The second group to observe helicase activity from RecQ4 purified the human enzyme using a baculoviral expression system (Suzuki et al., 2009). They observed 3′-5′ helicase activity in the absence of a reaction product trap. RecQ4 was found to efficiently displace a 17 basepair-long oligonucleotide annealed to single-stranded circular DNA, but could not unwind a similar 37 basepair region. This gives some sense of the step-size of the enzyme, and indicates that it may not be processive. Further analysis showed maximal helicase activity between pH 8 and pH 10, at 5mM ATP, and at $MgCl_2$ concentrations of 8mM and above. All of these assays used either oligonucleotides annealed to single-stranded circular DNA, or blunt-ended linear DNA with an internal single-stranded region. As such, beyond determining the direction of helicase activity, substrate preference was not characterized (Suzuki et al., 2009).

Our results are generally consistent with those of Suzuki et al. (Suzuki et al., 2009). In both cases, robust helicase activity in a 3' to 5' direction was observed in the absence of any reaction product trap (Capp et al., 2009; Suzuki et al., 2009). Human RecQ4 was able to unwind double-stranded DNA that was 17 basepairs long, but not 37 basepairs (Suzuki et al., 2009). We have observed that *Drosophila* RecQ4 readily unwinds double-stranded DNA as long as 30 basepairs (Capp et al., 2009), but not 60 basepairs. These results do not preclude each other, though the differences between the human and *Drosophila* enzymes make direct comparison improper. It remains that helicase activity is clearly more robust when RecQ4 is

expressed in insect cells with baculovirus vectors than when it is expressed in *E. coli*. This strongly indicates that a eukaryotic expression system is necessary for activation of the helicase activity. Whether this is due to the presence of the conditions necessary for optimal folding of the enzyme, or to post-translational modifications, or both, has yet to be determined.

5.2 Roles of RecQ4 in Replication

5.2.1 The Sld2 Domain and Initiation

Consideration of both the biological and biochemical data concerning RecQ4 suggests two critical roles for the enzyme during replication. The first of these roles is in the loading and licensing of the replication machinery, and is dependent upon the N-terminal Sld2 domain. The nature of the second role is less clear, but it requires an active helicase domain, and may occur during the elongation phase of replication.

The best understood role of RecQ4 is that which is played by the Sld2 domain. This understanding is largely based on comparison with the function of Sld2 in *S. cerevisiae*. The similarity is instructive but not authoritative, due to the more complicated nature of replication in metazoans. Like Sld2 in *S. cerevisiae*, RecQ4 is involved in the construction of the replicative complex. When preparing for replication during G1, the origin recognition complex and Mcm2-7 bind the origin, forming the pre-replicative complex (pre-RC). RecQ4 and Mcm10 (along with Ctf4) interact with the pre-RC at the beginning of S-phase, causing it to form the CMG complex with Cdc45 and GINS (Im et al., 2009; Xu et al., 2009a). This complex then interacts with Polymerase α and RPA as replication is initiated. TopBP1/Cut5 plays some role in this process and can interact with RecQ4. However, unlike Dpb11 and

Sld2, it is able to bind chromatin independently of RecQ4 (Matsuno et al., 2006). If the interaction between TopBP1/Cut5 and RecQ4 is important to replication, it is not in the same manner as the interaction between Dpb11 and Sld2. In fact, it is possible that the functional homolog of the Sld2-Dpb11 interaction is that of RecQ4 and Mcm10. Phosphorylation by metazoan S-CDKs probably regulates or modulates this process, but this also remains to be determined.

The role played by RecQ4's Sld2 domain in the assembly of the replicative complex is absolutely essential, and one for which there is no "backup" redundant system. It is failure to perform at least this role that renders RecQ4 knockout mutants early embryonic lethal (the small amount of development in such mutants being attributed to maternal loading of functional RecQ4). The Sld2 domain, however, cannot be responsible for the symptoms of RecQ4 associated syndromes. Patients of these syndromes are able to replicate DNA and generally have RecQ4 mutations that occur outside the Sld2 domain. Such mutations may affect the protein sequence outside the Sld2 domain, or may occur in the promoter and only affect expression level, but either way the Sld2 domain remains intact. Therefore, the phenotypes seen in RecQ4 associated syndromes and in RecQ4-truncated mouse models are due to defects occurring in the rest of the enzyme.

5.2.2 The Involvement of the Helicase Domain

The most obvious potential source of RecQ4 defects that could result in the symptoms observed in RTS, RAPADILINO, and BGS, is the helicase domain. We suggest that RecQ4 plays a second important role in replication, one which involves the active helicase domain. This role may also be applied to DNA damage repair. Work in mouse and

fly model systems indicate that an inactivated or partially present helicase domain results in lethality. On the other hand, the complete absence of the helicase domain in transgenic mice only slightly reduces viability, though developmental abnormalities do occur. One can envision that the role played by the helicase domain may not be essential, but, when inactive, it interferes with the normal mechanics of replication, which results in defects of varying severity depending on the extent of interference. If the helicase domain is entirely absent, it provides neither benefit nor detriment, and so defects are less severe. A more specific variation of this model is that RecQ4's helicase domain would have an essential role that could also be performed by a redundant system. Such a system would not be able to compete with the helicase domain for binding, and so the presence of an inactive helicase domain would hinder its ability to act, leading to more severe defects. When employed, the redundancy would show reduced efficiency compared to RecQ4, which would result in an increased incidence of replication defects, but not annul replication altogether. It must be noted that the redundant system could come from within RecQ4 itself. If helicase activity attributed to the Sld2 domain is confirmed, this could serve in the absence of the helicase domain. In this case, the lethality of helicase-inactivated point mutants from *Xenopus* and *Drosophila* would imply that such an internal redundancy was chiefly characteristic of mammalian RecQ4.

5.2.3 Cooperation Between RecQ4 and Mcm2-7

The question remains, what is the functional role for a helicase other than the Mcm2-7 complex at the replication fork? One possibility is that Mcm2-7 and RecQ4 work in concert to unwind DNA, in an active-passive pairing. Active helicases use the energy from

ATP hydrolysis to dissociate annealed strands. This can happen either by acting as a wedge to force the strands apart, or by twisting the duplex DNA to force it to separate. Passive helicases, on the other hand, take advantage of thermal fluctuations in the DNA between annealed and unannealed states, using ATP hydrolysis to preserve single-stranded DNA from reannealing. If RecQ4 and Mcm2-7 work as an active-passive pair, one helicase actively unwinds the DNA, while the other trails behind and passively maintain the separation (Figure 19A). Mcm2-7, which is capable of passing double-stranded DNA through the center of its hexameric ring, could act to actively destabilize duplex DNA (reviewed in (Remus et al., 2009) and (Bochman and Schwacha, 2009). RecQ4 would then advance to passively maintain the separation. The converse pairing (active RecQ4, passive Mcm2-7) is difficult to conceive with Mcm2-7 remaining bound to double-stranded DNA, because this puts it ahead of the replication fork.

It has been proposed that Mcm2-7 acts as a double-stranded DNA pump (Bochman and Schwacha, 2009). This suggests another means of cooperating with RecQ4 at the replication fork (Figure 19B). Mcm2-7 may move ahead of the replication fork, pushing double-stranded DNA towards RecQ4, which then unwinds it. In this scenario RecQ4 would fulfill the role of active helicase. Mcm2-7, on the other hand, would act as a DNA translocase. In this capacity, it would serve as a motor for the replicative complex, and to clear off DNA-bound proteins ahead of the fork.

RecQ4 and Mcm2-7 may also cooperate as active helicases (Figure 19C), in a variation of the Ploughshare model (Bochman and Schwacha, 2009). The Ploughshare model proposes that Mcm2-7 unwinds DNA by pushing it towards a rigid wedge, either on Mcm2-7 itself, or on an associated protein. It is possible that rather than a simple wedge,

another helicase (RecQ4) is involved. In this situation, Mcm2-7 would push the DNA directly into RecQ4, which would actively unwind it. This coordination between the two helicases would lessen the burden borne by either individually.

It is also possible that Mcm2-7 is the primary replicative helicase, and RecQ4 does not generally function as a helicase at the replication fork. When the replicative complex encounters structures that Mcm2-7 is unable to unwind, such as certain forms of DNA damage, or topologically constrained DNA structures like Holliday junctions, RecQ4 may be required (Figure 19D). In such cases, Mcm2-7 would either pass the blockage through its central pore or otherwise allow RecQ4 access, but the replicative complex would be paused until RecQ4 unwound it. This would imply a role for RecQ4 consistent with that of the other RecQ helicases, which also act on topologically constrained DNA structures. Any role in DNA repair would also primarily involve unwinding such topological abnormalities.

Much remains to be discovered about RecQ4. A number of experiments clearly demonstrate both that the enzyme acts during replication, and that it acts as a helicase. It is not immediately obvious how this is so, but there may be an advantage to combining the non-enzymatic replication licensing function of Sld2 with the enzymatic activity of a RecQ helicase. Exactly how these two functions are coordinated will likely be a key element in DNA replication or repair, and will certainly be an area of active investigation in the future.

The work presented in this chapter was originally published in *Critical Reviews in Biochemistry and Molecular Biology*. Capp, C., Wu, J. and Hsieh, T.S. RecQ4: the second replicative helicase? *Crit. Rev. Biochem. Mol. Biol.* 2010; 45:233-242. © Informa Medical and Pharmaceutical Science - Journals.

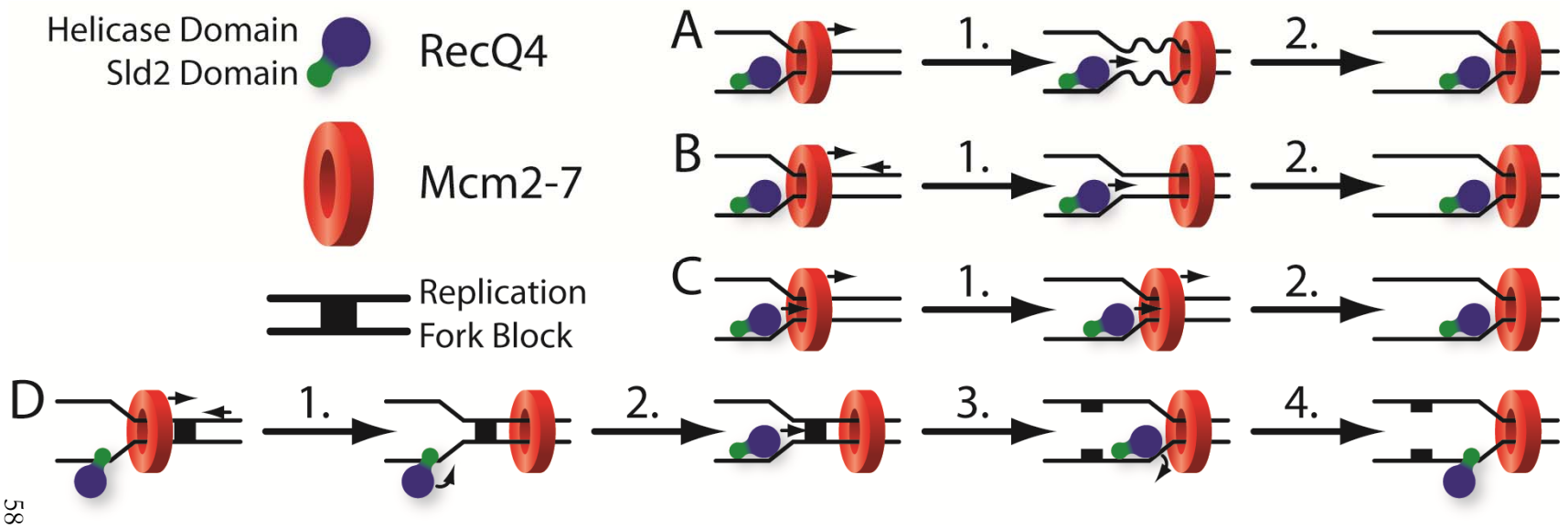

Figure 19: The coordination of RecQ4's helicase domain with Mcm2-7 in replication. In the following models RecQ4 is represented as a monomer, as indicated by initial biochemical experiments (Capp et al., 2009). The rest of the replicative complex is assumed, but not drawn. *(A)* Active Mcm2-7, passive RecQ4: Mcm2-7 actively destabilizes duplex DNA (1), causing it to transiently come apart. RecQ4 then enters the new single-stranded region (2), and prevents reannealing. *(B)* Active RecQ4, Mcm2-7 as double-stranded DNA pump: DNA is pushed by Mcm2-7 towards RecQ4 (1), but is not immediately unwound. This allows RecQ4 to actively unwind double-stranded DNA (2). *(C)* Active RecQ4, active Mcm2-7 (the Ploughshare model): Mcm2-7 pushes double-stranded DNA into RecQ4, which is simultaneously moving forward, unwinding the DNA (both 1 and 2). *(D)* RecQ4 as a specialty helicase: Mcm2-7 acts as the primary replicative helicase; RecQ4 is at the replication fork because of the licensing function of the Sld2 domain, but does not act as a helicase. Occasionally a replication block is encountered which Mcm2-7 is unable to unwind. This blocks the replicative complex and fork, but not Mcm2-7, which is around the double-stranded DNA (1). The replicative complex reconfigures to allow activity of the RecQ4 helicase domain (2). RecQ4 unwinds the blockage, which is repaired or bypassed (3). The replicative complex reconfigures again (4), and Mcm2-7 continues as the primary helicase.

Appendix. Separate and Combined Biochemical Activities of the Subunits of a Naturally Split Reverse Gyrase

The work presented in this appendix was originally published in the *Journal of Biological Chemistry*. Capp, C., Qian. Y., Sage, H. Huber, H., and Hsieh, T.S. Separate and combined biochemical activities of the subunits of a naturally split reverse gyrase. *J. Biol. Chem.* 2010; 285:39637-39645. © the American Society for Biochemistry and Molecular Biology. It is the work of several people, most notably Yushen Qian. An undergraduate research assistant that I helped mentor, Yushen purified the individual subunits of reverse gyrase, and performed the size-exclusion dimerization assays. We collaborated on the ATP hydrolysis assays and sequence alignment. I was responsible for the gel shift binding assays, organizing the data, and producing an initial draft that was worked on by the co-authors. In order to preserve the flow of the narrative, the article (Capp et al., 2010a) is presented here with minimal revisions.

A.1 Introduction

In the most apparently inhospitable environments life can survive and even thrive. A striking example of this adaptability is that of hyperthermophilic organisms, which have optimal growth conditions at temperatures of 80°C or higher (Stetter, 2006). Life at such extreme temperatures presents unique challenges, which include the propensity for double-stranded DNA to denature and a host of other potential genomic defects (reviewed in Lopez-Garcia and Forterre, 1999 and Grogan, 1998).

Hyperthermophiles, whether bacterial or archaeal, deal with the thermal instability of DNA by means of an enzyme called reverse gyrase (reviewed recently in Plank and Hsieh,

2009; Perugino et al., 2009; and Heine and Chandra, 2009). Since its discovery in a hyperthermophilic Archaeon (Kikuchi and Asai, 1984), there is growing evidence suggesting a role of reverse gyrase in stabilizing genomes in hyperthermophiles. It appears to be the only gene specific to hyperthermophiles (Forterre, 2002), and is present in all hyperthermophiles and in some thermophiles as well (Brochier-Armanet and Forterre, 2007). This unique function of reverse gyrase is further supported by genetic evidence. While reverse gyrase is dispensable for growth of *Thermococcus kodakaraensis* below 65°C, the strain without reverse gyrase showed retarded growth between 65-90°C, and no growth at 93°C (Atomi et al., 2004), demonstrating the essential role of reverse gyrase in supporting life at extreme temperatures.

Reverse gyrase is a type IA topoisomerase and has a unique enzymatic activity in utilizing ATP hydrolysis to induce positive supercoils in DNA. It has not been established precisely how this enzyme can protect against DNA thermal instability. DNA positive supercoiling per se may not be the direct cause, since supercoiling in hyperthermophiles is highly variable ranging from positively supercoiled to relaxed, or even negatively supercoiled (Lopez-Garcia and Forterre, 2000). The biochemical activity of reverse gyrase as a renaturase for single stranded DNA may have a critical role in maintaining genome stability at high temperature (Hsieh and Plank, 2006). In addition, reverse gyrase can promote DNA integrity through its role as a DNA chaperone to coat the damaged sites (Kampmann and Stock, 2004).

Reverse gyrase carries out directional strand transfer leading to an increase in linking number, rendering it capable of introducing positive supercoils to plasmid DNA and annealing complementary single stranded circles. While the biochemical mechanism for

directional strand transfer remains unclear, the structural biological and biochemical studies on the enzyme provide important insight. Reverse gyrase is composed of a Superfamily II helicase-like domain at its N-terminal half linked to a type IA topoisomerase domain (Brochier-Armanet and Forterre, 2007; Confalonieri et al., 1993). The crystal structure of *Archaeoglobus fulgidus* reverse gyrase indicates that these domains are arranged back-to-back with the active site of each domain facing away from its counterpart (Rodriguez, 2002; Rodriguez and Stock, 2002). When isolated recombinantly, the topoisomerase domain of *Sulfolobus acidocaldarius* reverse gyrase behaves like a type IA topoisomerase, and weakly relaxes negatively supercoiled DNA, but is unable to induce positive supercoils. The helicase-like domain hydrolyzes ATP, but does not demonstrate helicase activity. Combining the separately expressed recombinant topoisomerase and helicase-like domains results in positive supercoiling, indicating that the domains are able to reconstitute reverse gyrase *in vitro*. Positive supercoiling does not occur when a non-cognate topoisomerase or helicase is substituted for the respective domain (Declais et al., 2000). Thus the helicase-like and topoisomerase domains of reverse gyrase are unique, and can specifically cooperate to induce positive supercoils. One proposed mechanism for positive supercoiling is that a switch in the binding affinity of the helicase-like domain depends on its bound nucleotide (del Toro Duany et al., 2008). This switch in DNA binding is intimately coupled with the strand passage activity of the topoisomerase domain, thus resulting in reannealing of single stranded DNA and increasing the linking number (Hsieh and Plank, 2006; Plank and Hsieh, 2009).

The hyperthermophilic archaeal parasite *Nanoarchaeum equitans* (Jahn et al., 2008) does not have a single reverse gyrase gene. Instead, further analysis of its genome reveals two

separate genes encoding the apparent topoisomerase IA and helicase-like domains of reverse gyrase (Brochier-Armanet and Forterre, 2007; Waters et al., 2003). This raises the possibility that what in other hyperthermophiles are two domains of a single peptide, in *N. equitans* are distinct subunits of a multiprotein complex. In this work we separately express each protein, and biochemically characterize them in isolation and in combination with each other. We establish that they form a heterodimer in solution which is able to induce positive supercoils and is thus a functional reverse gyrase. The unique nature of *N. equitans*' reverse gyrase allows analysis of each subunit in isolation while avoiding potential complications due to artificial separation of the subunits. We exploit this to dissect the contribution of each subunit/domain to positive supercoiling.

A.2 Results and Discussion

A.2.1 Identification of the Subunits of Reverse Gyrase in *N. equitans*

Most of the reverse gyrases consist of a single peptide containing a type IA topoisomerase domain and a helicase-like domain. No such single gene is found in *N. equitans*, but genomic sequence analysis reveals two genes, NEQ 318 and NEQ 434, whose protein products show significant homology to reverse gyrase's topoisomerase and helicase-like domains, respectively (Figure 20). The protein encoded by NEQ 318 is 66% similar (47% identical) to the topoisomerase domain of *A. fulgidus* reverse gyrase. That encoded by NEQ 434 is 58% similar (35% identical) to the corresponding helicase-like domain.

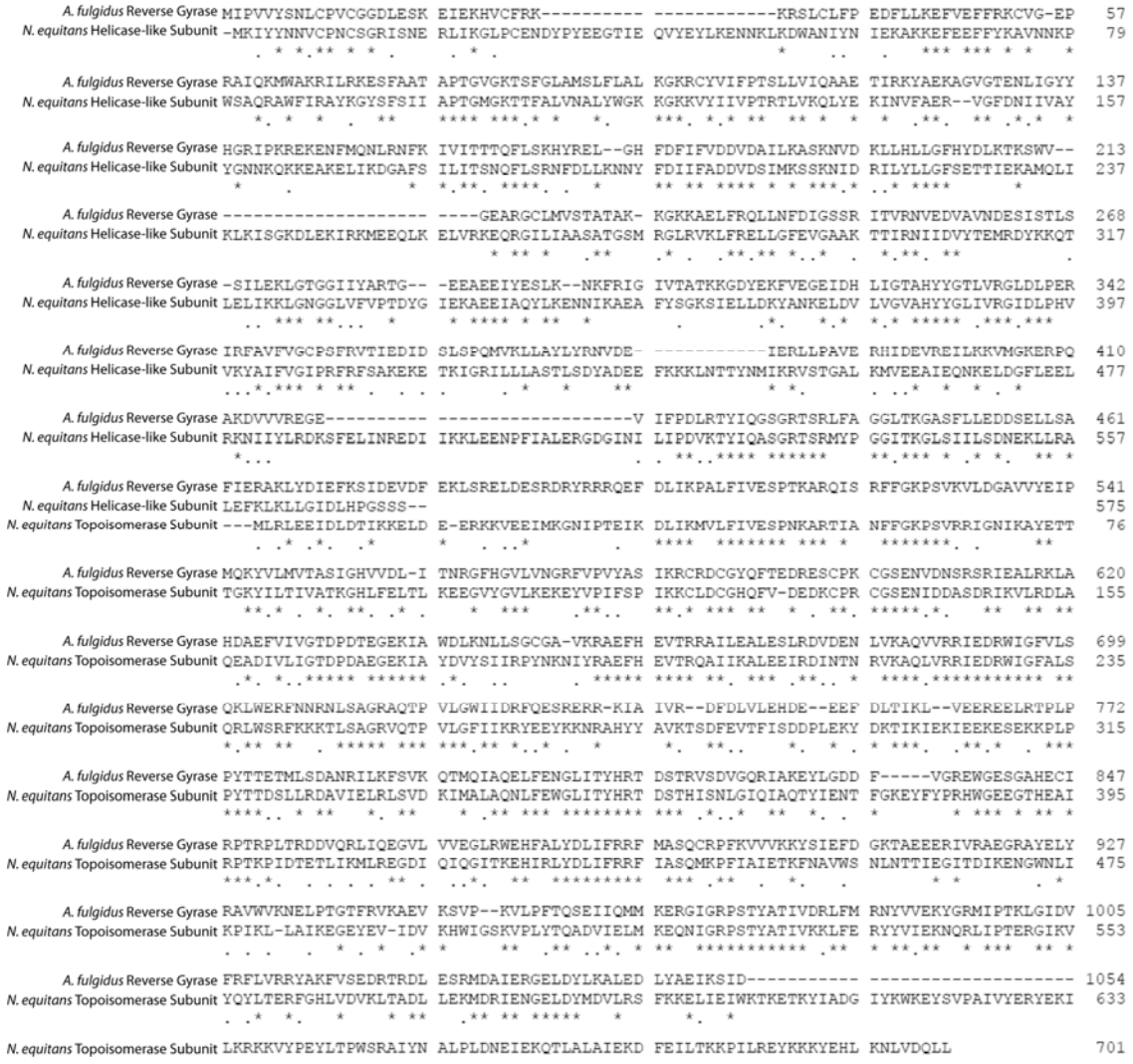

Figure 20: Sequences of the subunits of *Nanoarchaeum equitans* reverse gyrase. The amino acid sequences of the protein products of *N. equitans* genes 318 and 434 were aligned with that of *Archaeoglobus fulgidus* reverse gyrase. Identical residues are indicated by "*", and similar ones by ".". The N-terminal helicase domain of *A. fulgidus* reverse gyrase is 35% identical (58% similar) to NEQ 434. The C-terminal topoisomerase domain is 47% identical (66% similar) to NEQ 318.

It is interesting to note that splitting genes into separately expressing proteins is not unusual in the *N. equitans* genome, and the organism possesses split genes for both a type IA topoisomerase unrelated to reverse gyrase (NEQ 045 and NEQ 324) and a type II topoisomerase (NEQ 144 and NEQ 542). While it is plausible to expect that the products of

NEQ 318 and NEQ 434 dimerize to form a reverse gyrase holoenzyme, direct experimental evidence is requisite to demonstrate that they are genes for reverse gyrase. We present in the following sections biochemical evidence that the purified recombinant proteins can combine to reconstitute reverse gyrase holoenzyme. Hereafter the products of NEQ 318 and NEQ 434 will be referred to as the topoisomerase and helicase-like subunits of *N. equitans* reverse gyrase.

A.2.2 Expression and Purification of the Subunits in Isolation

The topoisomerase and helicase-like subunits of *N. equitans* reverse gyrase were expressed separately in *E. coli*. The C-terminus of the topoisomerase subunit was fused with a hexahistidine tag, and expressed in CodonPlus-RIL *E. coli* cells (Figure 21, lane 1). A three-step purification strategy was adopted. This consisted of Ni-NTA affinity chromatography (Figure 21, lane 2), Bio-Rex 70 cation exchange (Figure 21, lane 3), and hydroxyapatite chromatography (Figure 21, lane 4). The final enzyme was at >91% purity, and stored in 50% glycerol, 15mM sodium phosphate, pH 7.0, 1M NaCl, and 0.1mM DTT at -20°C.

The C-terminus of the helicase-like subunit was also fused with a hexahistidine tag. The resulting recombinant gene was expressed in Rosetta *E. coli* cells (Figure 21, lane 5). A three-step purification strategy similar to that used for the topoisomerase subunit was employed, the primary difference being the reordering of columns used. Thus the final strategy for purifying the helicase-like subunit was Bio-Rex 70 cation exchange (Figure 21, lane 6), followed by Ni-NTA affinity chromatography (Figure 21, lane 7), and lastly hydroxyapatite chromatography (Figure 21, lane 8). The final enzyme was at >92% purity,

and stored in 50% glycerol, 15mM sodium phosphate, pH 7.0, 1M NaCl, and 0.1mM DTT at -20°C.

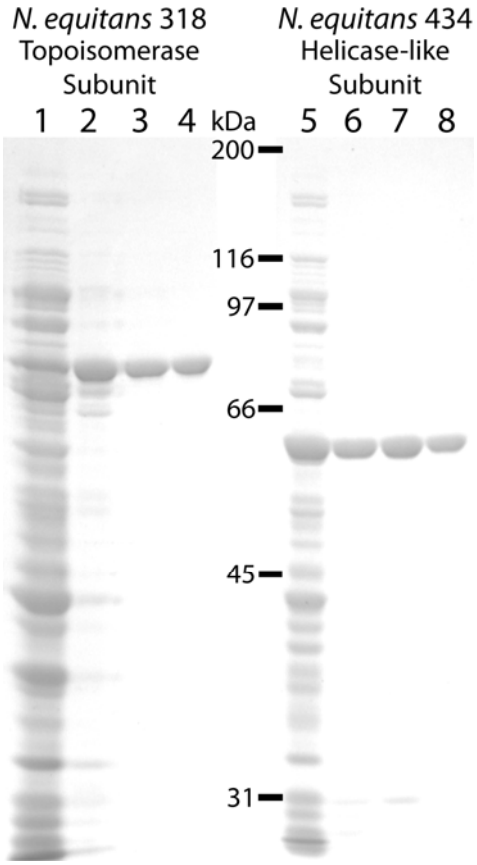

Figure 21: Purification of the Subunits of *N. equitans* reverse gyrase. Gel outlining the three-step purification of NEQ 318 and NEQ 434. Lanes 1 and 5 are the soluble portions of the respective whole cell lysates (Fraction I). Lanes 2 and 6 are the pooled peak fractions of the first columns employed (Fraction II), which are a Ni2+ IMAC column for NEQ 318 and a Bio-Rex 70 column for NEQ 434. Lanes 3 and 7 are the pooled peak fractions of the second columns employed (Fraction III), which are a Bio-Rex 70 column for NEQ 318 and a Ni2+ IMAC column for NEQ 434. Lanes 4 and 8 are the dialyzed pooled peak fractions from a hydroxyapatite column (Fraction IV).

A.2.3 The Topoisomerase and Helicase-like Subunits Dimerize to Form Reverse Gyrase

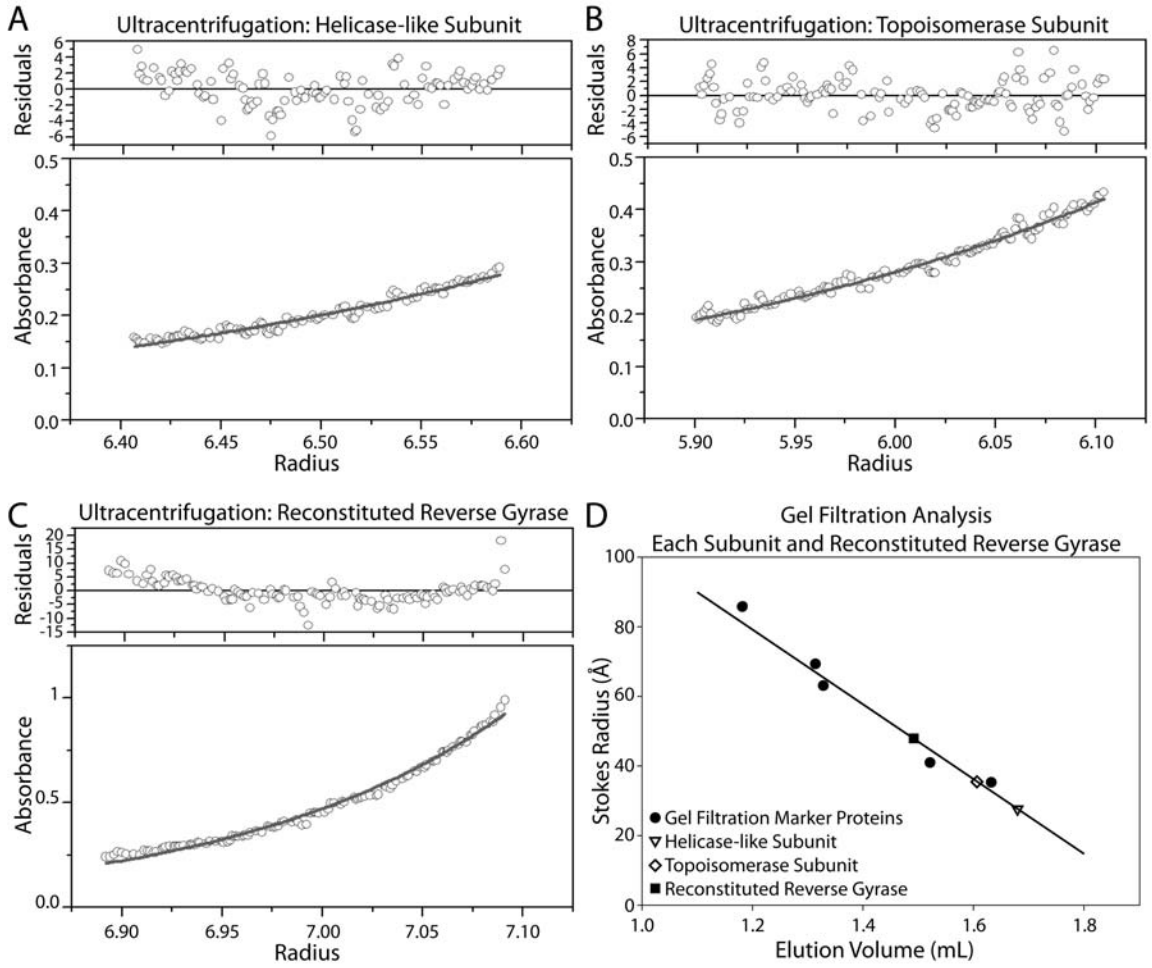

Figure 22: Ultracentrifugation and gel-filtration analyses of the subunits and equimolar mixtures. *(A, B, C)* Sedimentation-equilibrium analyses at 8,000rpm, 20°C. Plots are absorbance at 275nm versus the radius (cm), for *(A)* 4.7μM NEQ 434 (helicase-like subunit): best fit molecular weight is 74kDa (theoretical is 69kDa); *(B)* 4.3μM NEQ318 (topoisomerase subunit): best fit molecular weight is 89kDa (theoretical is 82kDa); *(C)* 1:1 mixture of NEQ 434 and NEQ 318 (4.7μM, and 4.3μM, respectively): best fit molecular weight is 155kDa (theoretical for a 1:1 complex is 151kDa). *(D)* Determination of the Stokes radii of the reconstituted reverse gyrase and its component subunits. Elution volume was measured after application to a Superdex 200 Precision Column with a bed volume of 2.4mL. Marker proteins used for calibration (see Methods) are indicated by filled circles. The elution volumes of the helicase-like and topoisomerase subunits, as well as the reconstituted reverse gyrase holoenzyme, are indicated by an open triangle, an open diamond, and a filled square, respectively.

To experimentally establish the predicted dimerization between the topoisomerase and helicase-like subunits to reconstitute reverse gyrase, size exclusion chromatography was conducted on each subunit and the in vitro reconstituted holoenzyme. Each of the samples was eluted from the column as a single protein peak. Calculating the Stokes radius of each based on elution volume with respect to marker proteins leads to the following predicted radii: for the topoisomerase subunit, 35Å; for the helicase-like subunit, 26Å; and for the reconstituted holoenzyme, 48Å (Figure 22D). If the Stokes radius is instead calculated using the partition coefficient, the following radii are predicted: for the topoisomerase subunit, 35Å; for the helicase-like subunit, 21Å; and for the reconstituted holoenzyme, 49Å (data not shown). The radii derived from each method of calculation are in reasonable agreement with each other, and are consistent with the topoisomerase and helicase-like subunits forming a single globular mass when combined in solution.

To confirm that a heterodimer is formed by the two subunits in solution, sedimentation-equilibrium analysis was conducted on each subunit and the reconstituted holoenzyme. Sedimentation equilibrium experiments are based on the principles of thermodynamics, rather than hydrodynamics, and thus the molecular weight determination is independent of any reference proteins. Analysis of such experiments resulted in a calculated molecular mass of 74kDa for the helicase-like subunit (Figure 22A), 89kDa for the topoisomerase subunit (Figure 22B), and 155kDa for the reconstituted holoenzyme (Figure 22C). These values are comparable to those predicted by the amino acid sequence (69kDa, 82kDa and 151kDa, respectively). Furthermore, the calculated mass of the reconstituted holoenzyme is consistent with the formation of a heterodimer from a topoisomerase subunit and a helicase-like subunit, and suggests that the formation of this heterodimer is rapid and

nearly complete given an equimolar concentration of the two subunits. Thus two independent assays indicate that the topoisomerase and helicase-like subunits combine to form a heterodimeric holoenzyme, confirming predictions from sequence alignments. The consistent results from both methods also suggest that the predominant species in solution is heterodimer. The concentrations of protein samples used in the sedimentation equilibrium is in the μM range or less (Figure 22, and data not shown), indicating a dimerization constant significantly lower than 1μM.

A.2.4 Coordination Between the Topoisomerase and Helicase-like Subunits: Supercoiling

Positive supercoiling by reverse gyrase requires the active participation of each domain. The energy from ATP hydrolysis, occurring in the helicase-like domain, is coordinated with the strand passage activity mediated by the topoisomerase domain, to generate an increase in the linking number and thus positive supercoiling. To confirm that the holoenzyme reconstituted from separately expressed subunits retains such an activity, we assayed the supercoiling activity of the topoisomerase domain and its complex with the helicase-like domain. Timecourses were conducted by incubating negatively supercoiled plasmid DNA with the topoisomerase subunit and varying amounts of the helicase-like subunit. The DNA products were analyzed by agarose gel electrophoresis in the presence of chloroquine (Figure 23). With such gel electrophoretic analysis, DNA species with positive/negative supercoiling or in the relaxed state can be distinguished. Densitometric analysis of these species at the 3.5 and 30 minute timepoints indicate that the initial timepoint is generally reflective of the early stages of the reaction. In the absence of helicase-like subunit, the DNA is quickly relaxed by the topoisomerase subunit, with the reaction

nearly complete at the first timepoint taken at 3.5 minutes (Figure 23, lanes 1-4). At a helicase-like to topoisomerase subunit ratio of 0.5:1, relaxed and positively supercoiled DNA species are observed (Figure 23, lanes 5-8). Some positively supercoiled species are observed at the initial timepoint, indicating rapid positive supercoiling, and confirming that the heterodimeric reconstituted holoenzyme is an active reverse gyrase. Negatively supercoiled species are also present, unlike in the absence of the helicase-like subunit (Figure 23, compare lanes 1 and 5). Thus the helicase-like subunit reduces the topoisomerase subunit's relaxation activity while allowing the positive supercoiling activity. At a helicase-like to topoisomerase subunit ratio of 1:1 (Figure 23, lanes 8-12), rapid positive supercoiling is observed; relaxed species are not observed, but a significant amount of negatively supercoiled species are. The retention of negative supercoils and absence of relaxed species confirms the trend seen at the 0.5:1 ratio. Since the relaxed DNA is an intermediate formed while positively supercoiling the negatively supercoiled substrate, the absence of relaxed species suggests that the helicase-like subunit sequesters the topoisomerase subunit, causing it to switch from mere relaxation to processive positive supercoiling activity. The two subunits undergo multiple strand passage events per DNA binding event, and this occurs even under high salt conditions that should favor subunit dissociation (data not shown). Supercoiling rates from helicase-like to topoisomerase subunit ratios greater than 1:1 (Figure 23, lanes 13-20) are not significantly different from those at 1:1. This confirms that the dimeric interaction between the two subunits is tight, and that they do not readily dissociate from each other.

Therefore, the topoisomerase subunit of *N. equitans* reverse gyrase, when isolated, acts as a topoisomerase capable of relaxing supercoiled DNA. When the topoisomerase

subunit is incubated with the helicase-like subunit, supercoil relaxation activity is diminished. The two subunits dimerize tightly and cooperate to induce positive supercoils in a processive manner. These experiments demonstrate that the reconstituted heterodimer has a clear reverse gyrase activity that neither of its subunits possesses.

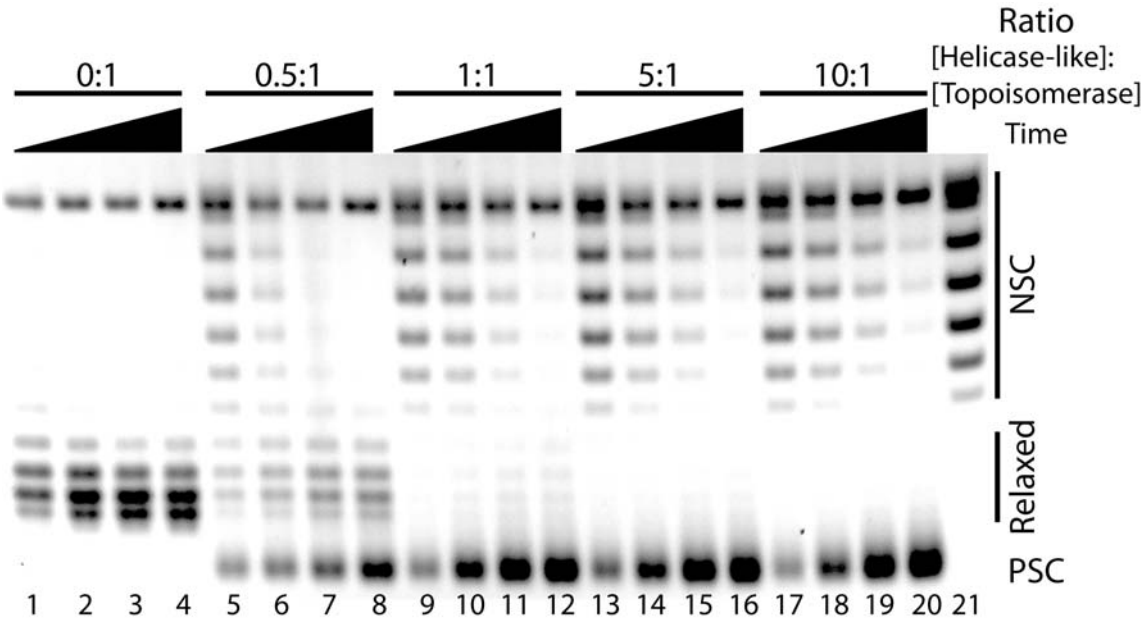

Figure 23: Reconstituted reverse gyrase induces positive supercoils. Supercoiling assays were performed with helicase-like and topoisomerase subunits at the indicated ratios, with the topoisomerase subunit concentration held constant at 29nM. For each ratio, timepoints were taken at 3.5, 7, 15 and 30 minutes. In isolation the topoisomerase subunit rapidly relaxes negatively supercoils (compare the starting material in lane 21 to the 3.5 minute timepoint with topoisomerase only in lane 1). This activity is inhibited in the presence of the helicase-like subunit, which instead activates a positive supercoiling activity (compare relative amounts of negative supercoils and relaxed DNA in lanes 1-4, lanes 5-8, and lanes 9-12).

A.2.5 Coordination Between the Topoisomerase and Helicase-like Subunits: ATP Hydrolysis

The helicase-like subunit acts to switch the function of the topoisomerase subunit from relaxing supercoiled DNA to positive supercoiling induction. Further examination of the nature of the helicase-like subunit's contribution to positive supercoiling requires analysis

of its activity in isolation and in the presence of the topoisomerase subunit. The ability to unwind DNA has never been directly observed for the reverse gyrase helicase-like domain. However, other studies have shown it to hydrolyze ATP in a DNA dependent manner. This then provides a convenient means of examining the effect of the topoisomerase subunit on the helicase-like subunit.

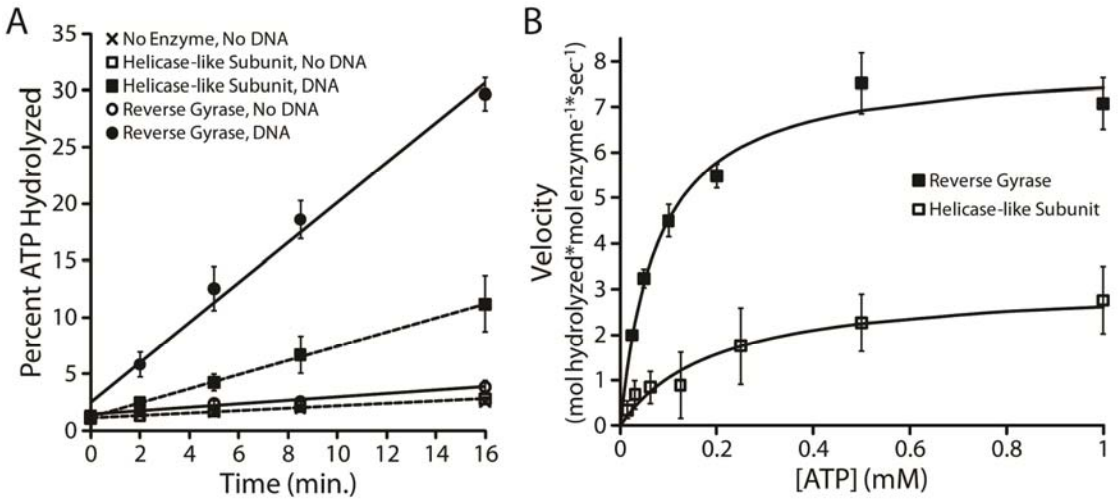

Figure 24: The helicase-like subunit ATPase activity is stimulated by the topoisomerase subunit. *(A)* Timecourses of ATP hydrolysis by the helicase-like subunit (square markers, dashed lines) and the reconstituted reverse gyrase (circle markers, solid lines) in the presence or absence of DNA (filled and open markers, respectively). Background ATP hydrolysis (in the absence of enzyme) is indicated by X markers. Hydrolysis in the absence of DNA is by either enzyme is not significantly above background levels. In the presence of DNA reconstituted reverse gyrase is more active than the helicase-like subunit alone. *(B)* Measurement of the velocity of ATP hydrolysis by the helicase-like subunit (open squares) and the reconstituted reverse gyrase (filled squares) as a function of ATP concentrations. The Michaelis-Menten parameters determined from these data are: k_{cat}= 7.99 mmol ATP hydrolyzed/(mmol enzyme * sec), K_M= 76.4μM, for the holoenzyme; and k_{cat}= 3.12 mmol ATP hydrolyzed/(mmol enzyme * sec), K_M= 197μM for the helicase-like subunit. The reconstituted holoenzyme has a higher maximal velocity and a lower K_M, indicating that it has an improved catalytic specificity over the helicase-like subunit alone. Error bars indicate standard deviation. Each data point represents at least three independent trials.

To that end, ATP hydrolysis was assayed for the helicase-like subunit in isolation and in the context of the reconstituted holoenzyme. ATP hydrolysis, regardless of context, required the presence of DNA (Figure 24A). The holoenzyme was found to hydrolyze ATP at a significantly faster rate than the helicase-like subunit alone. The enzymological basis for the rate enhancement was further dissected by determining the Michaelis-Menten parameters, k_{cat}, K_M, and catalytic specificity k_{cat}/K_M. Kinetic analysis of ATP hydrolysis found the helicase-like subunit to have a catalytic specificity of $16 mol^{-1}*sec^{-1}$ (Figure 24B). The reconstituted holoenzyme, on the other hand, has a catalytic specificity of $105 mol^{-1}*sec^{-1}$, due to both a 2.6 fold increase in k_{cat} and a 2.6 fold decrease in K_M of the enzyme for ATP. This indicates that the topoisomerase subunit significantly improves the helicase-like subunit's ability to hydrolyze ATP. Because the effect is seen in both k_{cat} and K_M, the activity enhancement is due to an increase in both the catalytic efficiency of ATP hydrolysis and binding affinity for ATP.

A.2.6 DNA Binding Preferences of Each Subunit and the Reconstituted Holoenzyme

To further dissect the contribution made by each subunit to supercoiling, the DNA binding preference of each individual subunit and the reconstituted holoenzyme were examined. Electrophoretic mobility shift assays were conducted with single-stranded and double-stranded oligonucleotide substrates at 80°C (with the double-stranded substrates being designed to have a melting temperature of 90°C). The reconstituted holoenzyme showed a slight preference for double-stranded DNA over single-stranded DNA (Figure 25A, quantified in Figure 25B). The helicase-like subunit exhibited a marked preference for double-stranded DNA, while the topoisomerase subunit preferred single-stranded DNA.

The biochemical function of reverse gyrase as a DNA renaturase suggests that it must have a dual binding affinity for both single- and double-stranded DNA. The separation of substrate preference between subunits indicates a similar segregation in the reconstituted holoenzyme. This has important implications for the mechanism of inducing positive supercoils.

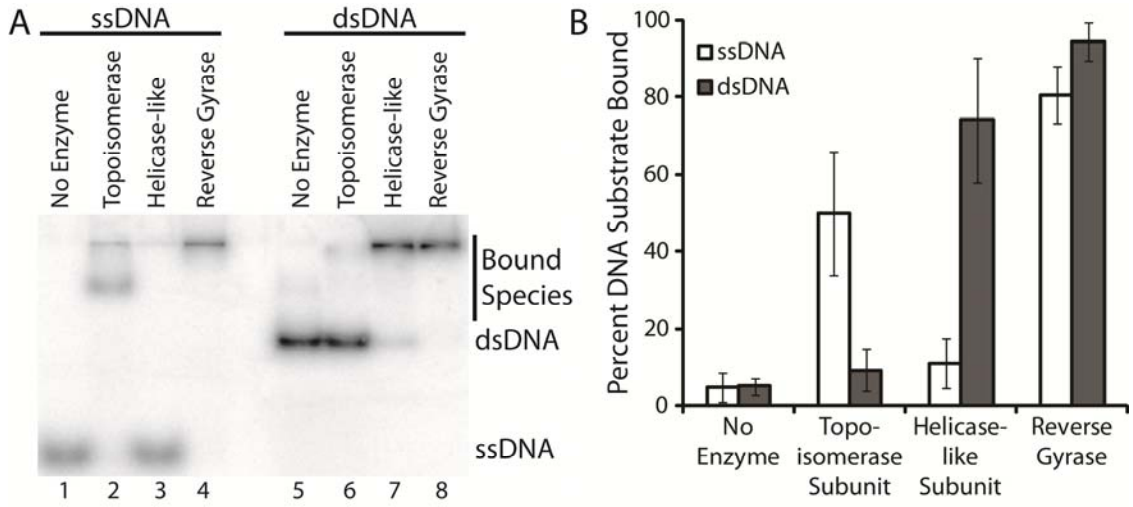

Figure 25: DNA substrate preference is partitioned between subunits. *(A)* A representative electrophoretic mobility shift assay of each subunit and reconstituted reverse gyrase with single-stranded and double-stranded oligonucleotide substrates (lanes 1-4 and lanes 5-8, respectively). Reverse gyrase binds either substrate well (lanes 4 and 8). The topoisomerase subunit strongly prefers single-stranded substrates (lanes 2 and 6), and the helicase-like subunit strongly prefers double-stranded substrates (lanes 3 and 7). *(B)* Quantification of electrophoretic mobility shift assays. Error bars indicate standard deviation (n=3).

Helicase activity has not been observed for either the reconstituted enzyme or the isolated helicase-like subunit (data not shown). This is consistent with other published results (Declais et al., 2000) and confirms that, although the helicase-like subunit demonstrates DNA-dependent ATP hydrolysis, it does not have strand translocation or separation activities. It is possible that ATP hydrolysis is employed in positive supercoiling to modulate DNA substrate affinity. To test this, electrophoretic mobility shift assays were conducted for

the helicase-like subunit and the reconstituted holoenzyme with both single-and double-stranded DNA substrates, varying the nucleotide cofactor (no cofactor; ADP; ATP; and the non-hydrolysable ATP analog, AMPPNP). Varying nucleotide cofactor did not significantly alter binding affinity for either the helicase-like subunit or the reconstituted holoenzyme (Figure 26). If the phosphorylation state of the nucleotide cofactor has any effect on substrate affinities, it is possible that they are too small or too transient to be detected by electrophoretic mobility shift assay.

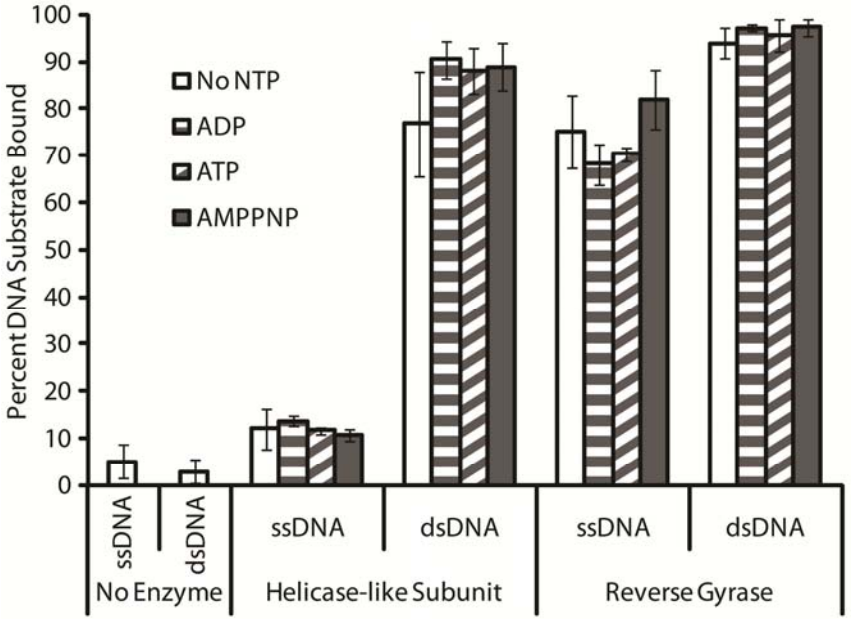

Figure 26: DNA substrate preference is not altered by nucleotide cofactor. Quantification of electrophoretic mobility shift assays of nucleotide dependence of substrate preference for the helicase-like subunit and reconstituted reverse gyrase. Enzyme was incubated with single-stranded and double stranded oligonucleotide substrates in the absence of nucleotide cofactor (open bars), and in the presence of ADP (horizontally striped bars), ATP (diagonally striped bars), or AMPPNP (filled bars). Reverse gyrase binds either substrate well under all conditions examined. The helicase-like subunit strongly prefers double-stranded substrates over single-stranded substrates, and this preference is not altered by nucleotide cofactor either. Error bars indicate standard deviation (n=3).

To detect subtle changes due to nucleotide cofactor, a topology shift assay was employed. In this assay, reactions with negatively supercoiled plasmid DNA were done in the presence of *T. maritima* topoisomerase I, which relaxes the plasmid unless an induced topological shift interferes. The extent of the induced shift correlates precisely with variation in the extent of plasmid relaxation. The effect of helicase-like subunit binding to plasmid DNA was thus assayed in the presence of varying nucleotide cofactors. Topological shift was observed in the absence of nucleotide cofactor and in the presence of AMPPNP (Figure 27, compare lanes 9 and 12 to 1 and 4), demonstrating that under these conditions the helicase-like subunit binds to plasmid DNA and reduces twist. No such shift was observed in the presence of ADP or ATP (Figure 27, lanes 10 and 11). It should be noted that the helicase-like subunit readily hydrolyzes ATP in the presence of double-stranded DNA (data not shown), making ATP and ADP functionally equivalent for this assay. Comparing with the relaxed species generated without helicase-like subunit (lanes 1-4), it is apparent from the topological shift experiments that the binding of the subunit with ADP as a cofactor does not result in DNA unwinding. The difference in topological shift between substrates bound in the presence of AMPPNP and ADP (Figure 27, lanes 12 and 10) indicates that there is a modulation in binding with the underwound DNA, resulting from ATP binding and hydrolysis.

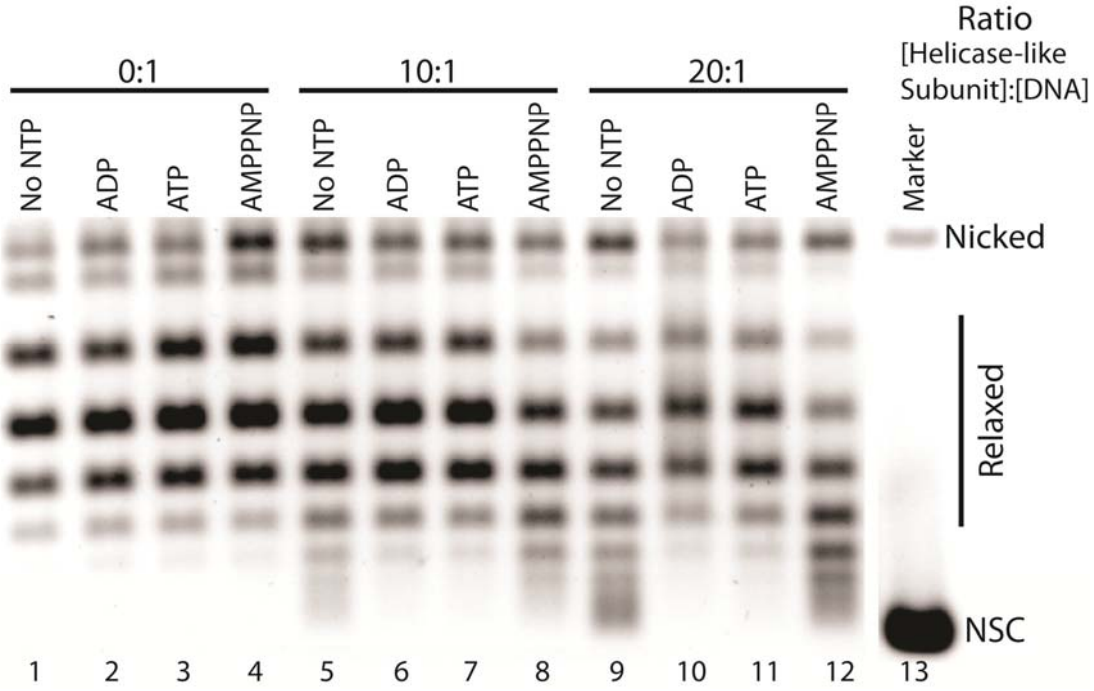

Figure 27: Nucleotide cofactor modulates helicase-like subunit binding. Topology shift assays were done with *T. maritima* topoisomerase I and the helicase-like subunit in the presence of ADP, ATP, AMPPNP, or the absence of nucleotide cofactor. Shifts were evident in the absence of cofactor, or in the presence of AMPPNP, indicating that under these conditions the helicase-like subunit alters the topology of DNA. These shifts are not seen in the presence of ADP or ATP (which is readily hydrolyzed to form ADP), indicating that ATP hydrolysis modulates the interaction of the helicase-like subunit with DNA. The DNA species generated without any helicase-like domain (lanes 1-4) mark the positions for relaxed species under these reaction conditions. Increasing amounts of helicase-like domain generate DNA species with more negative supercoils (compare lanes 9 and 12 with lanes 5 and 8). We have also run the identical samples in an agarose gel with chloroquine to confirm that the DNA species with higher mobilities shown here are more negatively supercoiled (data not shown).

A.3 Concluding Remarks

We have established that the genes NEQ 318 and NEQ 434, found in *N. equitans*, encode the topoisomerase and helicase-like subunits, respectively, of a heterodimeric reverse gyrase. Most reverse gyrases are single polypeptide chains with the helicase-like domain in the N-terminal half and topoisomerase domain in the C-terminal half. Previously, the only

known exception was the reverse gyrase from *Methanopyrus kandleri*. It is a heterodimer with part of the topoisomerase domain inserted into the helicase-like domain (Krah et al., 1996). However, the reverse gyrase from *N. equitans* is the only one encoded by separate genes that cleanly split the helicase-like domain from the topoisomerase domain. We showed here that the purified recombinant subunits associate in a one to one ratio, and work together to induce positive supercoils. The topoisomerase subunit is able to act on its own to relax negative supercoils. Similarly, the helicase-like subunit in the absence of the topoisomerase subunit acts as a DNA-dependent ATPase, though it does not actually unwind DNA. Characterizing these activities as functions of the respective isolated subunit and of the reconstituted holoenzyme has allowed insight into the mechanism of positive supercoiling by reverse gyrase.

Positive supercoiling is not a function of the simple combination of the separate activities of each subunit. Each subunit alters the activity of its counterpart when associated. In the case of the helicase-like subunit's ATP hydrolysis activity, the association of topoisomerase domain results in an enhancement in catalytic specificity by increasing substrate affinity and turnover rate. The supercoil relaxation activity of the topoisomerase subunit, on the other hand, is reduced by the presence of the helicase-like subunit. Along with this partial inhibition of supercoil relaxation activity by the topoisomerase subunit, there is also a concomitant increase in the processivity of the positive supercoiling. The holoenzyme does not rapidly relax the population of negatively supercoiled DNA, and then induce positive supercoils. Instead, it relaxes and positively supercoils a given DNA molecule in a single binding event, before dissociating and working on the next negatively supercoiled molecule. As a result, incomplete reactions contain both negatively and positively

supercoiled DNA, but little (if any) relaxed DNA. While processive positive supercoiling is an efficient means for reverse gyrase reactions, such a mode of reaction may not be universally adopted by all reverse gyrases. For example, the kinetic analysis of the supercoiling reaction by *A. fulgidus* reverse gyrase demonstrates an apparent accumulation of relaxed DNA intermediates, suggesting a much lower processivity (Hsieh and Capp, 2005).

DNA binding studies shed further light on the mechanism of positive supercoiling. DNA substrate preference is partitioned between subunits, with the topoisomerase subunit binding single-stranded DNA and the helicase-like subunit binding double-stranded DNA. In order to reanneal single strands to form a duplex and to induce positive supercoils, it is likely that reverse gyrase binds at the junction between double- and single-stranded DNA. It is well established that type IA topoisomerases prefer single stranded DNA for binding (Kirkegaard and Wang, 1985). Though the biochemical analysis of the helicase-like domain is still at its early stage, it probably has a basal affinity for the double-stranded DNA, which can be modulated by its binding to ATP/ADP. Using electrophoretic mobility shift assays to monitor DNA binding affinities by the helicase-like domain, we could not detect any effect of ATP/ADP on protein binding to double- versus single-stranded DNA. Since these binding assays are not sensitive to small or transient changes in DNA binding affinities, we used topology shift assays to detect any such changes. Our results indicated that either in the absence of any cofactor or in the presence of AMPPNP, the helicase-like domain can induce a slight DNA unwinding, implicating a relative preference of binding to underwound DNA under such conditions.

DNA gyrase and reverse gyrase are unique among topoisomerases in that they can mediate directional strand passage to induce either an increase (reverse gyrase) or decrease

(gyrase) in linking number. The mechanistic basis for gyrase to reduce linking number or to generate negative supercoils is better understood (reviewed in Schoeffler and Berger, 2008). The right-handed wrapping of the DNA gate segment by gyrase provides a clear topological basis for vectorial strand passage. In contrast, the mechanistic basis for the positive supercoiling action by reverse gyrase remains to be fully elucidated. There are a number of experimental evidences to suggest a plausible mechanism. Reverse gyrase has an important biochemical function as a DNA renaturase (Hsieh and Plank, 2006). This is supported by two lines of experiments. The first is that the enzyme can readily reanneal two single-stranded DNA circles with complementary base sequences. The second is that reverse gyrase can more efficiently generate positive supercoils in a DNA substrate containing a permanent single-stranded bubble. For the bubble substrate, the continuous (but futile) action of reannealing leads to a higher level of positive supercoiling. DNA reannealing is proposed to start with the targeting of reverse gyrase to the underwound or single stranded region (Hsieh and Capp, 2005), and this preferential binding is induced by the association of helicase-like domain with a specific nucleotide (e.g. ATP). The subsequent nucleotide switch (e.g. from ATP to ADP) can induce a binding preference of the helicase-like domain for rewound or double-stranded DNA, thus promoting the rewinding of DNA helix. Such a topological shift can be facilitated and fixed through the strand passage action of the topoisomerase domain. The key element in this mechanistic proposal, DNA binding preference shift as modulated by the nucleotide switch, has also been described by del Toro Duany et al (del Toro Duany et al., 2008). The specific details remain to be fully elucidated. The data presented here suggest that the AMPPNP (or ATP) bound form has a slight preference for underwound DNA, relative to the ADP bound form. Earlier work with the *A. fulgidus* enzyme

demonstrated that in the presence of AMPPNP, highly negatively supercoiled DNA products can be generated (Hsieh and Capp, 2005). This was interpreted to indicate that reverse gyrase when bound with AMPPNP can associate with underwound DNA and thus entrap negative supercoils. However, there could be variations to the theme. Using fluorescence anisotropy to monitor DNA binding at 37°C, the recombinantly engineered helicase-like domain from *T. maritima* reverse gyrase was shown to have a binding affinity for single-stranded DNA regardless of the cofactors (del Toro Duany et al., 2008). The relative preference of single-stranded DNA is reduced in the presence of ADP when compared to the conditions without cofactor or with ATP analog. On the other hand, the engineered recombinant helicase-like domain from the *Sulfolobus solfataricus* enzyme showed no nucleotide-modulated shift in DNA binding affinity by mobility shift assays (Valenti et al., 2008), similar to our observations reported here. The complicated difference in the binding preferences observed in these experiments could be due to methodology for monitoring binding reactions, or to the use of reverse gyrases from different species. Despite this, the essential feature of the proposed nucleotide switch mechanism through which DNA structures could be modified appears to be retained by all reverse gyrases. Further experimentation will be needed to dissect the biochemical mechanism of this unique DNA machine.

A.4 Methods

A.4.1 Sequence Alignment

Sequence alignments were performed using CLC Sequence Viewer 6.3 from CLC Bio A/S.

A.4.2 Expression and Purification

The culturing of *Nanoarcheum equitans* and isolation of its genomic DNA was described earlier (Waters et al., 2003). The topoisomerase subunit (NEQ 318) was cloned from *N. equitans* genomic DNA by PCR amplification, and placed in the pET-23b (Novagen) vector with a C-terminal hexahistidine tag. The resulting protein was expressed in *E. coli* BL21-CodonPlus-RIL (DE3, pLysS) cells (Stratagene) and grown in Luria Broth modified for a final NaCl concentration of 20g/L. Cells were induced with 1mM isopropyl β-D-1-thiogalactopyranoside for six hours at 30°C. Pelleted cells were resuspended by Dounce homogenization in lysis buffer (1mg/mL lysozyme, 20mM Tris HCl, pH 8.0, and 0.5M NaCl). The lysate was incubated at 25°C for 20 minutes and then sonicated. The soluble fraction was isolated by centrifugation at 12,000Xg for 30 minutes, and set aside. The pelleted fraction was re-extracted according to the above procedure, and the resulting soluble fraction was combined with the original soluble fraction to form Fraction I. This was passed over a Ni^{2+} IMAC column (Qiagen) and washed with buffer containing 15mM imidazole, and 0.5M NaCl. The protein was eluted with a gradient from 40mM to 400mM imidazole, maintaining constant 0.5M NaCl. Peak fractions were pooled and labeled as Fraction II. Fraction II was passed over a Bio-Rex 70column (Bio-Rad). The column was washed with buffer containing 15mM NaP_i, 10% glycerol, and 0.1M NaCl. The protein was eluted with a gradient from 0.1M NaCl to 1.1M NaCl, holding constant 15mM NaP_i and 10% glycerol. Fractions containing the topoisomerase subunit were collected and labeled as Fraction III. This was loaded on a hydroxyapatite column (Bio-Rad), which was washed with 0.1M NaP_i and 0.5M NaCl. Protein was manually eluted with buffer containing 0.5M NaP_i and 0.5M NaCl. Fractions containing the purified recombinant protein were pooled and

labeled as Fraction IV. Fraction IV was dialyzed overnight in 50% glycerol, 15mM NaP_i, 1M NaCl, and 0.1mM DTT, and stored at -20°C.

The cloning and expression of the helicase-like subunit (NEQ 434) was the same as that of the topoisomerase subunit, with the following differences. The helicase-like subunit was expressed in *E. coli* Rosetta (DE3, pLysS) cells (Novagen). Following lysis and isolation of the soluble fraction, Fraction I was diluted ten-fold to a final salt concentration of 0.1M NaCl. The column order was altered, with Bio-Rex 70 as the first column, and the Ni^{2+} IMAC column as the second. The Bio-Rex 70 column was washed with buffer containing 0.5M NaCl, and was eluted with a 0.5–1.5M NaCl gradient.

Co-expression of NEQ 434 and NEQ 318 in *E. coli* cells was accomplished by the pETDuet-1 vector (Novagen), and purification of the heterodimeric protein complex followed essentially the same protocols described for purifying NEQ 318 protein.

A.4.3 Dimerization Assays

Gel filtration assays were conducted using a pre-packed Superdex 200 Precision Column 2.3/30 (Amersham Biosciences), on an AKTA FPLC Purifier system (Amersham Biosciences). To calibrate the column, the following proteins were used as column standards: thyroglobulin (Sigma), with a Stokes radius of 86Å; β-galactosidase (Boehringer Mannhelm), 69Å; type I horse spleen ferritin (Sigma), 63Å; lactate dehydrogenase (Worthington), 41Å; and bovine serum albumin (Pentex), 35Å. All protein samples were between 10 and 15μg, and in a buffer of 15mM Tris HCl and 0.5M NaCl. The Stokes radii of the individual subunits and the reconstituted holoenzyme were calculated based on the characteristic retention patterns of the standard proteins.

Sedimentation-equilibrium analyses of NEQ 434 and NEQ 318 separately and in mixtures were performed on a Beckman XL-A ultracentrifuge by standard procedures. Samples with a range of concentrations from 1-10μM were run at 6,000rpm, 8,000rpm and 10,000rpm and followed by absorbance at 275nm. Best fit molecular weights were calculated by the Ideal I program from Beckman.

A.4.4 Supercoiling Assays

Reactions contained 0.45μg of pUC19 DNA, 10mM Tris HCl, pH 8.0, 50mM KCl, 10mM $MgCl_2$, 0.1mM EDTA, 50μg/ml gelatin, 1mM ATP, and the indicated concentrations of topoisomerase and helicase-like subunits in a total volume of 30μl. A drop of mineral oil was added to cover the surface of the solution, to prevent evaporation. Reactions were initiated by immersing the reaction tube in a water bath set at 80°C, and terminated by adding EDTA and SDS to a final concentration of 10mM and 0.1%, respectively. Sucrose and tracking dyes (bromophenol blue and xylene cyanol) were added to a concentration of 5% and 100μg/ml respectively. Timecourse reactions were carried out in separate aliquots for each time point, due to the difficulty of manipulating the reaction mixture at high temperature. Time points used were 3.5, 7, 15, and 30 minutes. DNA products were analyzed by agarose gel electrophoresis in the presence of 30μM chloroquine as outlined previously (Hsieh and Capp, 2005).

A.4.5 Topology Shift Assays

Topology shift assays were carried out under conditions similar to supercoiling assays, with the following modifications. *Thermatoga maritima* topoisomerase I was purified and relaxation activity assayed as described before (Hsieh and Capp, 2005), and was included

at a concentration of 7 units per reaction. The helicase-like subunit only was used, to allow observation of the activity of that subunit, rather than the supercoiling of the reconstituted holoenzyme. Adenine cofactor was included at a concentration of 1mM, as indicated.

A.4.6 ATP Hydrolysis

Reconstituted holoenzyme or helicase-like subunit was incubated at a concentration of 25nM with 0.5mM ATP (2.1nM γ^{32}P-labeled ATP (Perkin Elmer)) at 80°C in 10mM Tris HCl, pH 7.9, 50mM KCl, 0.1mM EDTA, 10mM $MgCl_2$, 50μg/mL BSA, and 1mM DTT, in a total volume of 20μL. To determine DNA dependence, reactions were done in the presence or absence of 56μg/mL pBS/SK$^+$ single-stranded circular DNA. Reactions were covered with a drop of mineral oil to prevent evaporation. Timepoints were taken between 0 and 16 minutes, as indicated. Timecourses were conducted and analyzed by thin layer chromatography, as described previously (Capp et al., 2009).

To determine reaction velocity, reactions were done with 12nM reconstituted holoenzyme or 427nM helicase-like subunit and the indicated concentration of ATP (maintaining a constant concentration of 4.2nM γ^{32}P-labeled ATP) in the presence of 50μg/mL pBS/SK$^+$ single-stranded circular DNA. The timepoints used were 0, 5, 10 and 15 minutes or 0, 2.5, 5 and 10 minutes for the reconstituted holoenzyme, and 2.5, 5, 7.5 and 10 minutes for the helicase-like subunit. Velocity for each timecourse was determined by calculating the line of best fit, and at least three independent velocities were averaged together for each ATP concentration. Averaged velocities determined in the absence of enzyme were subtracted from enzymatic velocities as a background control.

A.4.7 Electrophoretic Mobility Shift Assays

For gel electrophoretic mobility shift assays, 0.1μM enzyme as specified, or an equal volume of storage buffer, was incubated with 0.25nM radiolabeled oligonucleotide substrate for 30 minutes at 80°C in 10mM Tris HCl, pH 7.9, 50mM KCl, 10mM $MgCl_2$, and 50μg/mL BSA in a total reaction volume of 20μL. For assays examining nucleotide cofactor modulation of binding preference 0.5mM nucleotide cofactor (ADP, ATP, or AMPPNP, as indicated) was included in the reaction mixture. After 30 minutes, 4μL of non-denaturing loading mix was added to each reaction to a final concentration of 5% sucrose, 33μg/mL bromophenol blue, and 33μg/mL xylene cyanol. Reaction products were separated on an 89mM Tris borate, 1mM $MgCl_2$, 8% polyacrylamide gel by electrophoresis similar to the conditions described earlier (Capp et al., 2009), and subjected to phosphorimaging analysis.

Oligonucleotide substrates were prepared as described earlier (Capp et al., 2009). Substrate compositions are indicated in Table 5, and oligonucleotide sequences are included in Table 1.

Table 5: Composition of Substrates Used In the Appendix

Substrate Name	Component Oligonucleotides
Duplex	F and G
Single-stranded DNA	H

References

Araki, H. (2010). Regulatory mechanism of the initiation step of DNA replication by CDK in budding yeast. Biochim Biophys Acta *1804*, 520-523.

Atomi, H., Matsumi, R., and Imanaka, T. (2004). Reverse gyrase is not a prerequisite for hyperthermophilic life. J Bacteriol *186*, 4829-4833.

Bachrati, C.Z., and Hickson, I.D. (2008). RecQ helicases: guardian angels of the DNA replication fork. Chromosoma *117*, 219-233.

Bochman, M.L., and Schwacha, A. (2009). The Mcm Complex: Unwinding the Mechanism of a Replicative Helicase. Microbiol Mol Biol Rev *73*, 652-683.

Bohr, V.A. (2008). Rising from the RecQ-age: the role of human RecQ helicases in genome maintenance. Trends in biochemical sciences *33*, 609-620.

Brochier-Armanet, C., and Forterre, P. (2007). Widespread distribution of archaeal reverse gyrase in thermophilic bacteria suggests a complex history of vertical inheritance and lateral gene transfers. Archaea *2*, 83-93.

Burks, L.M., Yin, J., and Plon, S.E. (2007). Nuclear import and retention domains in the amino terminus of RECQL4. Gene *391*, 26-38.

Cabral, R.E.C., Queille, S., Bodemer, C., Prost, Y., Neto, J.B.C., Sarasin, A., and Daya-Grosjean, L. (2008). Identification of new *RECQL4* mutations in Caucasian Rothmund-Thomson patients and analysis of sensitivity to a wide range of genotoxic agents. Mutat Res *643*, 41-47.

Capp, C., Qian, Y., Sage, H., Huber, H., and Hsieh, T.S. (2010a). Separate and combined biochemical activities of the subunits of a naturally split reverse gyrase. J Biol Chem *285*, 39637-39645.

Capp, C., Wu, J., and Hsieh, T.S. (2009). Drosophila RecQ4 has a 3'-5' DNA helicase activity that is essential for viability. J Biol Chem *284*, 30845-30852.

Capp, C., Wu, J., and Hsieh, T.S. (2010b). RecQ4: the second replicative helicase? Crit Rev Biochem Mol Biol *45*, 233-242.

Cheok, C.F., Wu, L., Garcia, P.L., Janscak, P., and Hickson, I.D. (2005). The Bloom's syndrome helicase promotes the annealing of complementary single-stranded DNA. Nucleic Acids Res *33*, 3932-3941.

Confalonieri, F., Elie, C., Nadal, M., de La Tour, C., Forterre, P., and Duguet, M. (1993). Reverse gyrase: a helicase-like domain and a type I topoisomerase in the same polypeptide. Proc Natl Acad Sci U S A *90*, 4753-4757.

Declais, A.C., Marsault, J., Confalonieri, F., de La Tour, C.B., and Duguet, M. (2000). Reverse gyrase, the two domains intimately cooperate to promote positive supercoiling. J Biol Chem *275*, 19498-19504.

del Toro Duany, Y., Jungblut, S.P., Schmidt, A.S., and Klostermeier, D. (2008). The reverse gyrase helicase-like domain is a nucleotide-dependent switch that is attenuated by the topoisomerase domain. Nucleic Acids Res *36*, 5882-5895.

Dietschy, T., Shevelev, I., Pena-Diaz, J., Huhn, D., Kuenzle, S., Mak, R., Miah, M.F., Hess, D., Fey, M., Hottiger, M.O.*, et al.* (2009). p300-mediated acetylation of the Rothmund-Thomson-syndrome gene product RECQL4 regulates its subcellular localization. J Cell Sci *122*, 1258-1267.

Foe, V.E., Odell, G.M., and Edgar, B.A. (1993). Mitosis and morphogenesis in the *Drosophila* embryo: point and counterpoint. In The Development of *Drosophila melanogaster*, M. Bate, and A.M. Arias, eds. (Cold Spring Harbor, New York, Cold Spring Harbor Laboratory Press), pp. 149-300.

Forterre, P. (2002). A hot story from comparative genomics: reverse gyrase is the only hyperthermophile-specific protein. Trends Genet *18*, 236-237.

Garcia, V., Furuya, K., and Carr, A.M. (2005). Identification and functional analysis of TopBP1 and its homologs. DNA Repair *4*, 1227-1239.

Grogan, D.W. (1998). Hyperthermophiles and the problem of DNA instability. Molecular Microbiology *28*, 1043-1049.

Hall, M.C., and Matson, S.W. (1999). Helicase motifs: the engine that powers DNA unwinding. Mol Microbiol *34*, 867-877.

Heine, M., and Chandra, S.B. (2009). The linkage between reverse gyrase and hyperthermophiles: a review of their invariable association. J Microbiol *47*, 229-234.

Hoki, Y., Araki, R., Fujimori, A., Ohhata, T., Koseki, H., Fukumura, R., Nakamura, M., Takahashi, H., Noda, Y., Kito, S.*, et al.* (2003). Growth retardation and skin abnormalities of the Recql4-deficient mouse. Hum Mol Genet *12*, 2293-2299.

Hsieh, T.S., and Capp, C. (2005). Nucleotide- and stoichiometry-dependent DNA supercoiling by reverse gyrase. J Biol Chem *280*, 20467-20475.

Hsieh, T.S., and Plank, J.L. (2006). Reverse gyrase functions as a DNA renaturase: annealing of complementary single-stranded circles and positive supercoiling of a bubble substrate. J Biol Chem *281*, 5640-5647.

Ichikawa, K., Noda, T., and Furuichi, Y. (2002). [Preparation of the gene targeted knockout mice for human premature aging diseases, Werner syndrome, and Rothmund-Thomson syndrome caused by the mutation of DNA helicases]. Nippon Yakurigaku Zasshi *119*, 219-226.

Im, J.S., Ki, S.H., Farina, A., Jung, D.S., Hurwitz, J., and Lee, J.K. (2009). Assembly of the Cdc45-Mcm2-7-GINS complex in human cells requires the Ctf4/And-1, RecQL4, and Mcm10 proteins. Proc Natl Acad Sci USA *106*, 15628-15632.

Jahn, U., Gallenberger, M., Paper, W., Junglas, B., Eisenreich, W., Stetter, K.O., Rachel, R., and Huber, H. (2008). Nanoarchaeum equitans and Ignicoccus hospitalis: new insights into a unique, intimate association of two archaea. J Bacteriol *190*, 1743-1750.

Jin, W., Liu, H., Zhang, Y., Otta, S.K., Plon, S.E., and Wang, L.L. (2008). Sensitivity of RECQL4-deficient fibroblasts from Rothmund-Thomson syndrome patients to genotoxic agents. Hum Genet *123*, 643-653.

Kampmann, M., and Stock, D. (2004). Reverse gyrase has heat-protective DNA chaperone activity independent of supercoiling. Nucleic Acids Res *32*, 3537-3545.

Kanter, D.M., and Kaplan, D.L. (2010). Sld2 binds to origin single-stranded DNA and stimulates DNA annealing. Nucleic Acids Res.

Kass, E.M., and Jasin, M. (2010). Collaboration and competition between DNA double-strand break repair pathways. Febs Letters *584*, 3703-3708.

Kikuchi, A., and Asai, K. (1984). Reverse gyrase--a topoisomerase which introduces positive superhelical turns into DNA. Nature *309*, 677-681.

Kirkegaard, K., and Wang, J.C. (1985). Bacterial DNA topoisomerase I can relax positively supercoiled DNA containing a single-stranded loop. Journal of Molecular Biology *185*, 625-637.

Kitao, S., Ohsugi, I., Ichikawa, K., Goto, M., Furuichi, Y., and Shimamoto, A. (1998). Cloning of two new human helicase genes of the RecQ family: biological significance of multiple species in higher eukaryotes. Genomics *54*, 443-452.

Kitao, S., Shimamoto, A., Goto, M., Miller, R.W., Smithson, W.A., Lindor, N.M., and Furuichi, Y. (1999). Mutations in RECQL4 cause a subset of cases of Rothmund-Thomson syndrome. Nat Genet *22*, 82-84.

Krah, R., Kozyavkin, S.A., Slesarev, A.I., and Gellert, M. (1996). A two-subunit type I DNA topoisomerase (reverse gyrase) from an extreme hyperthermophile. Proc Natl Acad Sci U S A *93*, 106-110.

Larizza, L., Magnani, I., and Roversi, G. (2006). Rothmund-Thomson syndrome and RECQL4 defect: splitting and lumping. Cancer Lett *232*, 107-120.

Lin, S., Horning, D.P., Szostak, J.W., and Chaput, J.C. (2009). Conformational analysis of DNA repair intermediates by time-resolved fluorescence spectroscopy. J Phys Chem A *113*, 9585-9587.

Lohman, T.M., and Bjornson, K.P. (1996). Mechanisms of helicase-catalyzed DNA unwinding. Annual review of biochemistry *65*, 169-214.

Longhese, M.P., Bonetti, D., Manfrini, N., and Clerici, M. (2010). Mechanisms and regulation of DNA end resection. Embo Journal *29*, 2864-2874.

Lopez-Garcia, P., and Forterre, P. (1999). Control of DNA topology during thermal stress in hyperthermophilic archaea: DNA topoisomerase levels, activities and induced thermotolerance during heat and cold shock in Sulfolobus. Mol Microbiol *33*, 766-777.

Lopez-Garcia, P., and Forterre, P. (2000). DNA topology and the thermal stress response, a tale from mesophiles and hyperthermophiles. BioEssays *22*, 738-746.

Lukin, M., and de Los Santos, C. (2006). NMR structures of damaged DNA. Chem Rev *106*, 607-686.

Machwe, A., Xiao, L., Groden, J., Matson, S.W., and Orren, D.K. (2005). RecQ family members combine strand pairing and unwinding activities to catalyze strand exchange. J Biol Chem *280*, 23397-23407.

Macris, M.A., Krejci, L., Bussen, W., Shimamoto, A., and Sung, P. (2006). Biochemical characterization of the RECQ4 protein, mutated in Rothmund-Thomson syndrome. DNA Repair *5*, 172-180.

Mann, M.B., Hodges, C.A., Barnes, E., Vogel, H., Hassold, T.J., and Luo, G. (2005). Defective sister-chromatid cohesion, aneuploidy and cancer predisposition in a mouse model of type II Rothmund-Thomson syndrome. Hum Mol Genet *14*, 813-825.

Matsuno, K., Kumano, M., Kubota, Y., Hashimoto, Y., and Takisawa, H. (2006). The N-terminal noncatalytic region of Xenopus RecQ4 is required for chromatin binding of DNA polymerase alpha in the initiation of DNA replication. Mol Cell Biol *26*, 4843-4852.

Muftuoglu, M., Kulikowicz, T., Beck, G., Lee, J.W., Piotrowski, J., and Bohr, V.A. (2008). Intrinsic ssDNA annealing activity in the C-terminal region of WRN. Biochemistry *47*, 10247-10254.

Ohhata, T., Araki, R., Fukumura, R., Kuroiwa, A., Matsuda, Y., Tatsumi, K., and Abe, M. (2000). Cloning, genomic structure and chromosomal localization of the gene encoding mouse DNA helicase RecQ helicase protein-like 4. Gene *261*, 251-258.

Park, S.J., Lee, Y.J., Beck, B.D., and Lee, S.H. (2006). A positive involvement of RecQL4 in UV-induced S-phase arrest. DNA and cell biology *25*, 696-703.

Perugino, G., Valenti, A., D'Amaro, A., Rossi, M., and Ciaramella, M. (2009). Reverse gyrase and genome stability in hyperthermophilic organisms. Biochem Soc Trans *37*, 69-73.

Petkovic, M., Dietschy, T., Freire, R., Jiao, R., and Stagljar, I. (2005). The human Rothmund-Thomson syndrome gene product, RECQL4, localizes to distinct nuclear foci that coincide with proteins involved in the maintenance of genome stability. J Cell Sci *118*, 4261-4269.

Plank, J., and Hsieh, T.S. (2009). Helicase-appended topoisomerases: new insight into the mechanism of directional strand transfer. J Biol Chem *284*, 30737-30741.

Remus, D., Beuron, F., Tolun, G., Griffith, J.D., Morris, E.P., and Diffley, J.F.X. (2009). Concerted Loading of Mcm2-7 Double Hexamers around DNA during DNA Replication Origin Licensing. Cell *139*, 719-730.

Remus, D., and Diffley, J.F.X. (2009). Eukaryotic DNA replication control: Lock and load, then fire. Curr Opin Cell Biol *21*, 771-777.

Rodriguez, A.C. (2002). Studies of a positive supercoiling machine. Nucleotide hydrolysis and a multifunctional "latch" in the mechanism of reverse gyrase. J Biol Chem *277*, 29865-29873.

Rodriguez, A.C., and Stock, D. (2002). Crystal structure of reverse gyrase: insights into the positive supercoiling of DNA. EMBO J *21*, 418-426.

Sangrithi, M.N., Bernal, J.A., Madine, M., Philpott, A., Lee, J., Dunphy, W.G., and Venkitaraman, A.R. (2005). Initiation of DNA replication requires the RECQL4 protein mutated in Rothmund-Thomson syndrome. Cell *121*, 887-898.

Schoeffler, A.J., and Berger, J.M. (2008). DNA topoisomerases: harnessing and constraining energy to govern chromosome topology. Q Rev Biophys *41*, 41-101.

Schurman, S.H., Hedayati, M., Wang, Z.M., Singh, D.K., Speina, E., Zhang, Y., Becker, K., Macris, M., Sung, P., Wilson III, D.M.*, et al.* (2009). Direct and indirect roles of

RECQL4 in modulating base excision repair capacity. Hum Mol Genet *18*, 3470-3483.

Sclafani, R.A., and Holzen, T.M. (2007). Cell Cycle Regulation of DNA Replication. Annu Rev Genet *41*, 237-280.

Seki, M., Otsuki, M., Ishii, Y., Tada, S., and Enomoto, T. (2008). RecQ family helicases in genome stability: lessons from gene disruption studies in DT40 cells. Cell Cycle *7*, 2472-2478.

Sengupta, S., Shimamoto, A., Koshiji, M., Pedeux, R., Rusin, M., Spillare, E.A., Shen, J.C., Huang, L.E., Lindor, N.M., Furuichi, Y.*, et al.* (2005). Tumor suppressor p53 represses transcription of RECQ4 helicase. Oncogene *24*, 1738-1748.

Siitonen, H.A., Kopra, O., Kaariainen, H., Haravuori, H., Winter, R.M., Saamanen, A.M., Peltonen, L., and Kestila, M. (2003). Molecular defect of RAPADILINO syndrome expands the phenotype spectrum of RECQL diseases. Hum Mol Genet *12*, 2837-2844.

Siitonen, H.A., Sotkasiira, J., Biervliet, M., Benmansour, A., Capri, Y., Cormier-Daire, V., Crandall, B., Hannula-Jouppi, K., Hennekam, R., Herzog, D.*, et al.* (2009). The mutation spectrum in RECQL4 diseases. Eur J Hum Genet *17*, 151-158.

Stetter, K.O. (2006). Hyperthermophiles in the history of life. Philos Trans R Soc Lond B Biol Sci *361*, 1837-1842; discussion 1842-1833.

Suzuki, T., Kohno, T., and Ishimi, Y. (2009). DNA Helicase Activity in Purified Human RECQL4 Protein. J Biochem *146*, 327-335.

Tanaka, S., Tak, Y.S., and Araki, H. (2007a). The role of CDK in the initiation step of DNA replication in eukaryotes. Cell Division *2*.

Tanaka, S., Umemori, T., Hirai, K., Muramatsu, S., Kamimura, Y., and Araki, H. (2007b). CDK-dependent phosphorylation of Sld2 and Sld3 initiates DNA replication in budding yeast. Nature *445*, 328-332.

Thangavel, S., Mendoza-Maldonado, R., Tissino, E., Sidorova, J.M., Yin, J., Wang, W., Monnat, R.J., Falaschi, A., and Vindigni, A. (2010). The human RECQ1 and RECQ4 helicases play distinct roles in DNA replication initiation. Mol Cell Biol.

Valenti, A., Perugino, G., D'Amaro, A., Cacace, A., Napoli, A., Rossi, M., and Ciaramella, M. (2008). Dissection of reverse gyrase activities: insight into the evolution of a thermostable molecular machine. Nucleic Acids Res *36*, 4587-4597.

Van Maldergem, L., Siitonen, H.A., Jalkh, N., Chouery, E., De Roy, M., Delague, V., Muenke, M., Jabs, E.W., Cai, J., Wang, L.L.*, et al.* (2006). Revisiting the

craniosynostosis-radial ray hypoplasia association: Baller-Gerold syndrome caused by mutations in the RECQL4 gene. J Med Genet *43*, 148-152.

Vennos, E.M., Collins, M., and James, W.D. (1992). Rothmund-Thomson syndrome: Review of the world literature. J Am Acad Dermatol *27*, 750-762.

Wang, W., Seki, M., Narita, Y., Nakagawa, T., Yoshimura, A., Otsuki, M., Kawabe, Y., Tada, S., Yagi, H., Ishii, Y.*, et al.* (2003). Functional relation among RecQ family helicases RecQL1, RecQL5, and BLM in cell growth and sister chromatid exchange formation. Mol Cell Biol *23*, 3527-3535.

Waters, E., Hohn, M.J., Ahel, I., Graham, D.E., Adams, M.D., Barnstead, M., Beeson, K.Y., Bibbs, L., Bolanos, R., Keller, M.*, et al.* (2003). The genome of Nanoarchaeum equitans: insights into early archaeal evolution and derived parasitism. Proc Natl Acad Sci U S A *100*, 12984-12988.

Wong, I., and Lohman, T.M. (1993). A double-filter method for nitrocellulos-filter binding: Application to protein-nucleic acid interactions. Proc Natl Acad Sci USA *90*, 5428-5432.

Woo, L.L., Futami, K., Shimamoto, A., Furuichi, Y., and Frank, K.M. (2006). The Rothmund-Thomson gene product RECQL4 localizes to the nucleolus in response to oxidative stress. Exp Cell Res *312*, 3443-3457.

Wu, J., Capp, C., Feng, L., and Hsieh, T.S. (2008). Drosophila homologue of the Rothmund-Thomson syndrome gene: essential function in DNA replication during development. Dev Biol *323*, 130-142.

Wu, Y., Sugiyama, T., and Kowalczykowski, S.C. (2006). DNA annealing mediated by Rad52 and Rad59 proteins. Journal of Biological Chemistry *281*, 15441-15449.

Xu, X., and Liu, Y. (2009). Dual DNA unwinding activities of the Rothmund-Thomson syndrome protein, RECQ4. EMBO J *28*, 568-577.

Xu, X., Rochette, P.J., Feyissa, E.A., Su, T.V., and Liu, Y. (2009a). MCM10 mediates RECQ4 association with MCM2-7 helicase complex during DNA replication. EMBO J *28*, 3005-3014.

Xu, Y., Lei, Z., Huang, H., Dui, W., Liang, X., Ma, J., and Jiao, R. (2009b). *dRecQ4* Is Required for DNA Synthesis and Essential for Cell Proliferation in *Drosophila*. PLoS One *4*.

Yin, J., Kwon, Y.T., Varshavsky, A., and Wang, W. (2004). RECQL4, mutated in the Rothmund-Thomson and RAPADILINO syndromes, interacts with ubiquitin ligases UBR1 and UBR2 of the N-end rule pathway. Hum Mol Genet *13*, 2421-2430.

Zegerman, P., and Diffley, J.F. (2007). Phosphorylation of Sld2 and Sld3 by cyclin-dependent kinases promotes DNA replication in budding yeast. Nature *445*, 281-285.

Biography

Christopher Lee Capp was born in Cairo, Egypt, on June 13th, 1981. He received a B.S. in Chemistry from Creighton University in 2003, where he did research in various biological contexts in the labs of Marc Rendell and Mark Reedy. He has since attended Duke University, where he has pursued a Ph.D. in Biochemistry in the lab of Tao-shih Hsieh. He has authored several peer-reviewed articles. The first, a biochemical characterization of *Drosophila* RecQ4, was published in the *Journal of Biological Chemistry*, "*Drosophila* RecQ4 Has a 3'-5' DNA Helicase Activity That Is Essential for Viability." He has also published a review on the enzyme RecQ4, in *Critical Reviews in Biochemistry and Molecular Biology*, entitled "RecQ4: the second replicative helicase?" As co-author with Yushen Qian, he published a biochemical characterization of *Nanoarchaeum equitans* reverse gyrase, "Separate and combined biochemical activities of the subunits of a naturally split reverse gyrase," also in the *Journal of Biological Chemistry*.