Biochemical Model for Crossover Control

By Jenny Klein
April 17, 2013

Abstract

A DNA Double Strand Break (DSB), where both strands of the duplex are severed, is one of the most genotoxic types of DNA damage an organism can sustain. Accurate repair of DSBs is essential for genome stability and cancer avoidance. One method of repair, called Homology Directed Repair (HDR), relies on homologous recombination events that allow new DNA to be synthesized using the sister chromatid or homologous chromosome as a template. A key intermediate step in this process is the formation of a DNA structure called a D-loop, where single-stranded DNA (ssDNA) from the end of a processed DSB invades and anneals to a complementary region in the double-stranded DNA template. In this project an entirely synthetic D-loop structure has been constructed. Using the proteins from the T4 Bacteriophage model system, the ability of the D-loop to be unwound by the DNA helicase Gp41and the ability of the DNA polymerase Gp43 to extend the invasive strand have been examined. In line with a newly proposed model of crossover control, the manner of the D-loop's unwinding by Gp41 has implications for the exact mechanism of the repair process and whether it will lead to a potentially deleterious crossover or not.

Introduction

Maintaining a stable genome is crucial for an organism's survival and health. When DNA becomes damaged, it becomes increasingly difficult for the cell's metabolism to function as needed and for the cell to progress through its life cycle correctly. There are many different types of DNA damage, but one the most genotoxic types is a Double Strand Break (DSB), (Jackson and Bartek, 2009). A DSB occurs when both strands of the DNA duplex are broken, and can

result from either endogenous sources such as errors in DNA metabolism or exogenous sources such as chemical mutagens or ionizing radiation. A DSB can lead to serious genomic stability problems such as irregular gene expression and irregular DNA replication if it is not repaired properly.

Fortunately cells have evolved two methods of repairing DSBs: non-homologous end joining (NHEJ) and homology-directed repair (HDR), (Figure 1). In NHEJ, the two broken ends are ligated together. This is an error-prone repair pathway as segments of genetic information can be lost if the DNA sequence is not complete before ligation. Additionally, a type of NHEJ pathway called microhomology-mediated end-joining (MMEJ) always results in deleted bases, (Kass and Jasin, 2010). HDR prevents the possibility of sequence loss by using the sister chromatid or homologous chromosome as a template to newly synthesize the damaged section of DNA, (Kass and Jasin, 2010). Since a fully functional HDR mechanism is crucial to repair DNA correctly and prevent the accumulation of genetic damage that may eventually result in a cancerous cell, it is imperative that the repair process be well understood.

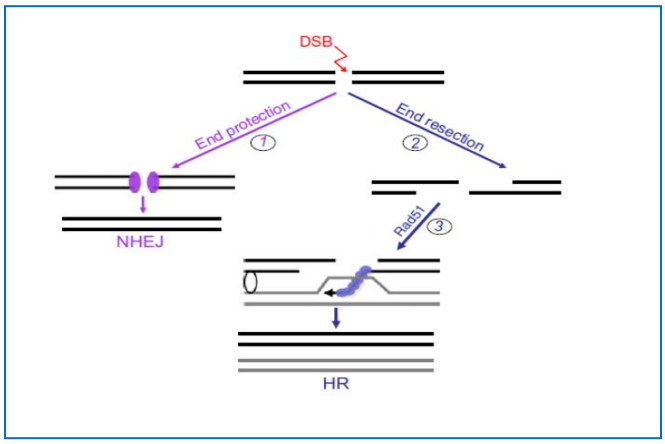

Figure 1 – Two methods of DSB repair: Non-homologous end joining (NHEJ) and Homology-directed repair (HDR). Adapted from Kass and Jasin, 2010.

To study the process of HDR the T4 bacteriophage can be used as a model. The T4 bacteriophage is a uniquely useful model organism due to the fact that during the normal course of its DNA replication it undergoes a process very similar to HDR called recombination-dependent replication (RDR). The T4 bacteriophage has a linear chromosome, and to finish

replication of its lagging strand, the 3' end must invade a region of homology to use as a template. Since this is an essential process to complete the DNA replication, the T4 bacteriophage has all the proteins needed for RDR coded for within its genome, (Kreuzer and Brister, 2010). These proteins and the role they play in RDR are shown in Figure 2 and will be discussed in more detail later in the paper. The T4 proteins have structural and functional homologs in all living organisms, including humans, and so an understanding of the role each T4 protein plays in RDR can be applied to the function of our own proteins in HDR.

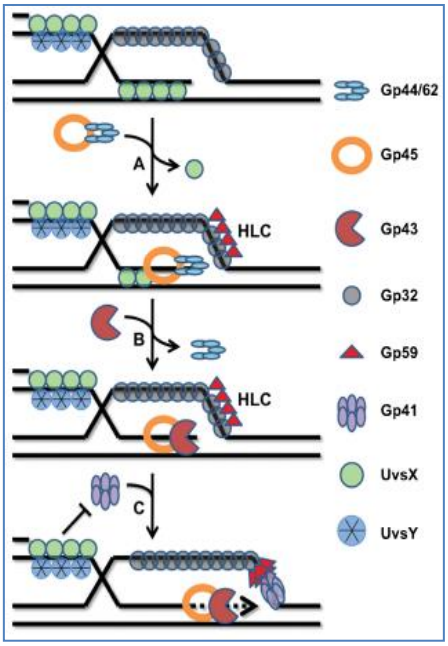

Figure 2 – The T4 Bacteriophage proteins involved in RDR. All proteins necessary for this process are encoded by the virus, making it a great model system to study HDR. The proteins and their functions are further described in the text. From Maher et al., 2011.

In order to begin HDR, a presynaptic filament needs to be formed. A presynaptic filament is a complex of ssDNA and recombinase proteins that is able to invade the dsDNA of the sister chromatid or homologous chromosome and search for homology in order to use it as a template to repair the DSB. In the T4 bacteriophage system, the recombinase protein is UvsX. UvsX catalyzes homologous pairing and directional branch migration, two crucial processes for base pairing the invasive strand with the correct complementary template in order to use it to repair the DSB, (Kodadek et al., 1988).

UvsX binds to the ssDNA formed after resection of the DSB by Gp46 and Gp47, the T4 Bacteriophage DNases, (Mickelson and Wiberg, 1981). UvsX binds in a process mediated by two accessory proteins: UvsY, the recombination mediator, and Gp32, the ssDNA-binding protein. Gp32 binds to all free ssDNA produced in a T4 infection and stabilizes it. During recombination, this stabilization helps keep the displaced strand of the duplex from rebinding as

the presynaptic filament translocates along the DNA searching for homology, (Yonesaki and Minagawa, 1989). However, UvsX has a hard time binding to ssDNA coated by Gp32, as Gp32 has a much higher binding affinity for ssDNA than UvsX, (Bleuit et al., 2001). UvsY solves this problem by mediating the exchange of Gp32 for UvsX on the ssDNA. UvsY is able to bind Gp32 coated ssDNA in a noncompetitive fashion. It is thought that the ssDNA is wrapped around the hexameric UvsY structure, (Beernink and Morrical, 1998), changing the conformation of the ssDNA. This change in conformation neutralizes the effect of the cooperativity of Gp32, weakening the Gp32-ssDNA interaction. UvsY can then mediate the direct exchange of Gp32 for UvsX via protein-protein interactions between UvsY and UvsX, (Sweezy and Morrical, 1998).

Once the presynaptic filament of ssDNA and UvsX is formed, UvsY also helps stabilize it. This enhances the homologous pairing ability of the filament since it is less likely to disassociate before finding homology, (Kodadek et al., 1989). However, because this stabilization is due to the decreased disassociation of UvsX from the ssDNA, UvsY paradoxically inhibits the branch migration ability of the UvsX coated ssDNA, (Salinas and Kodadek, 1995) since branch migration is dependent on UvsX's ability to disassociate from the filament after hydrolyzing ATP and reassemble further up on the filament (Kodadek et al, 1988). Therefore the ratio of UvsY to UvsX must be held in a delicate balance in vivo in order to promote filament assembly but not hinder the branch migration of the filament in its search for homology that is essential for forming the D-loop.

The D-loop is the DNA structure that is formed when the presynaptic filament has found the homologous region it was searching for on the template DNA and the invasive DNA strand has

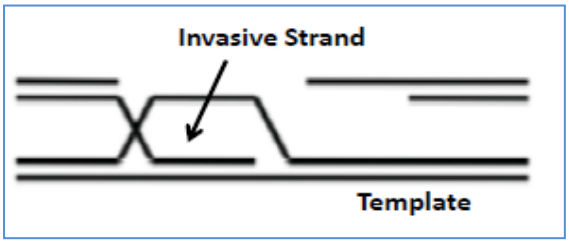

Figure 3 – The D-loop formed by the invasion of the template DNA by the presynaptic filament. The resected ssDNA of the DSB is now annealed at the homologous region in the template. Adapted from Maher et al., 2011.

annealed to it, as seen in Figure 3. The formation of the D-loop is the crucial intermediate step in HDR that links the recombination portion of the repair process and the replication portion that uses the homologous DNA as a template. It is at this point in the process that the cell can direct the repair down two separate though related pathways: the classic DSB repair model of Szostak et al. (1983), also called the double Holliday Junction (dHJ) model of repair or the synthesis-dependent strand annealing (SDSA) model of repair, (Formosa and Alberts, 1986).

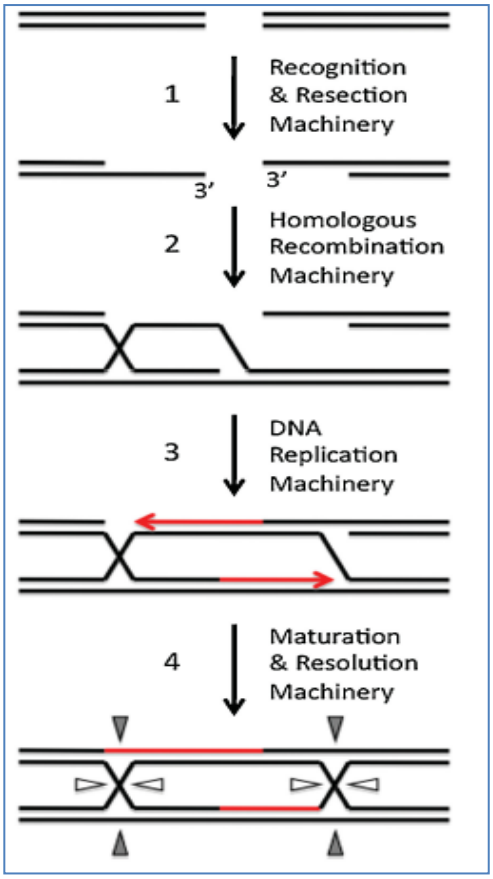

Figure 4 – The double Holliday Junction model of HDR. Maher et al., 2011.

In dHJ, once the D-loop has formed, a helicase further unwinds the dsDNA template, allowing the invasive strand to be extended in the 5' to 3' direction. As the D-loop opens up, eventually a sequence complementary to the other processed ssDNA side of the DSB becomes free allowing for second strand capture (Step 3 in Figure 4). This second ssDNA anneals to the displaced ssDNA of the D-loop and uses it as a template to synthesize new DNA. Once the broken strands have been fully filled in, the ends are ligated together forming a double Holiday Junction, (Step 4 in Figure 4). The dHJ structure is then resolved by structure specific endonucleases. Depending on which strands are cut during resolution, the process can result in either the crossover or non-crossover of markers flanking the DSB, leading to potential gene conversion. Since each newly synthesized strand of DNA is annealed to a template strand, dHJ repair is a semi-conservative method, (Szostak et al., 1983).

Similar to dHJ, in SDSA, once the D-loop has formed, a helicase opens up the D-loop allowing the invasive strand to use the bottom strand as a template and extend in the 5' to 3' direction. However, unlike in dHJ, as the D-loop opens and the invasive ssDNA extends, branch migration of the trailing end of the D-loop causes the newly synthesized DNA to be removed and allows the template to reanneal. This D-loop translocation is called bubble-migration. The ssDNA continues to extend until a sequence complementary to the other processed end of the DSB is recognized and is annealed to it. After this second end capture, the ssDNA strands can use each other as templates to fill in the gaps leaving only nicks in the backbone that a ligase can fill in. Any excess ssDNA is trimmed by nucleases, leaving the previously broken strand whole again, (Figure 5). SDSA differs from dHJ in that it is conservative replication. Only the previously broken strand ends up with newly synthesized DNA – the template strand remains as it was. This conservative replication prevents the crossing over of flanking markers on the DNA and possible gene conversion that may happen in dHJ, (Formosa and Alberts, 1986).

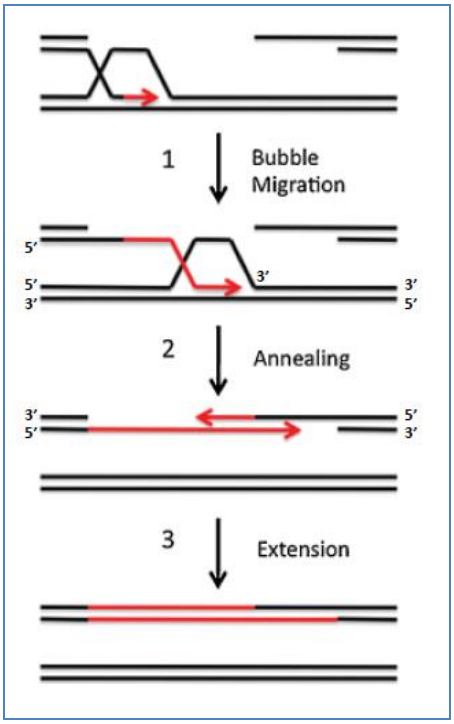

Figure 5 – The Synthesis-dependent strand annealing model of HDR. Adapted from Maher et al., 2011.

Gene conversion is defined as the nonreciprocal transfer of genetic information from one duplex to another duplex, (Szostak et al., 1983). What essentially happens is that one allele is lost and a formally heterozygous locus becomes homozygous. As mentioned above, the dHJ model of repair carries the risk of gene conversion. During repair, each newly synthesized region is paired to the older template strands and this structure has to be resolved back into two separate

dsDNAs before normal DNA metabolism can resume. Depending on how the endonucleases resolve the structure, regions flanking the junction sites may cross over from the template DNA to the newly repaired DNA (Figure 6). If these flanking regions contain a different allele than was originally on the previously broken strand, this new allele will replace the old one as mismatch repair will fix any incorrectly paired bases between the new DNA and the template. This will effectively rid the locus of heterozygousity, (Szostak et al., 1983). Gene conversion becomes a problem if a protective dominant allele is lost in the process, and the individual becomes homozygous for a recessive trait that has disease causing potential. For example the dominant copy of a tumor suppressor gene could be lost, and the cell has now lost control over its own cell cycle and become cancerous.

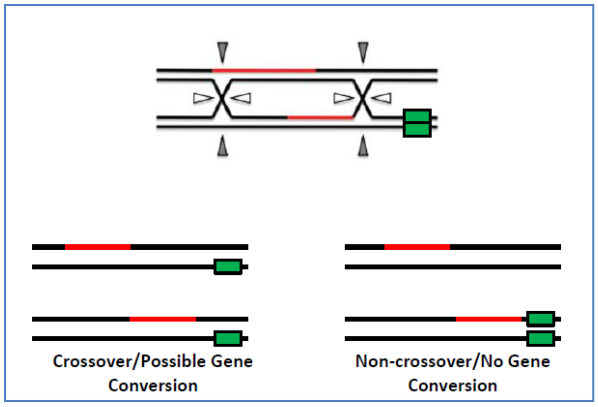

Figure 6 – Crossover of markers flanking the Holliday junctions during dHJ could lead to gene conversion if the allele sequences differ.

Because of the potential for gene conversion that exists in one of the pathways of repair, it is of great importance to understand how the cell makes the choice of which pathway to follow. This project proposes a new biochemical model of crossover control in which the helicase is the component of the reaction that directs the repair process down either the dHJ or SDSA pathways. This proposed model rests on the fact that helicases are the proteins responsible for unwinding the D-loop and how they unwind the D-loop will send it down one pathway or the other.

One of the helicases involved in unwinding the D-loop during dHJ or SDSA is Gp41, the primary replicative helicase of the T4 bacteriophage (Alberts, 1987). Gp41 exists as a dimer

physiologically but when bound to GTP or ATP, it assembles into a ring-like hexameric complex, (Dong et al., 1995). By hydrolyzing the bound ATP, Gp41 can translocate along the lagging strand in the 5' to 3' direction so that the ssDNA passes through the center of the ring structure, (Venkatesan et al., 1982). This physically excludes the leading strand allowing the duplex to be separated (Lionnet et al., 2007).

Gp41 can unwind dsDNA on its own, as long as there is an ssDNA region for it to load onto, (Venkatesan et al., 1982) but it functions much more efficiently in the presence of its accessory proteins, Gp59 the helicase loading protein (Morrical et al., 1994) and Gp32. Gp59 has the ability to bind to DNA, preferentially binding to fork structures over ssDNA over dsDNA, (Jones et al., 2000). However, even when excess free ssDNA is available, Gp59 preferentially binds to Gp32 clusters already bound to the ssDNA (Branagan et al., 2012). This preferential interaction implies that Gp41 would be loaded onto the Gp32-coated displaced strand of the D-loop, allowing the helicase to further unwind the dsDNA giving the invasive strand access to its homologous template.

In order for the invasive strand to extend and repair the DSB using the newly unwound dsDNA, a DNA polymerase is needed. In the T4 Bacteriophage, the polymerase is Gp43. Gp43 cannot initiate new strands of DNA but rather adds free deoxyribonucleotides to the 3' hydroxyl terminus end of a pre-existing primer, such as the invasive strand in the D-loop. Gp43 needs an ssDNA template in order to replicate the DNA and thus cannot replicate fully dsDNA on its own, (Goulian et al., 1968). In addition to its polymerase ability, Gp43 also has exonuclease activity in the 3' to 5' direction, (Nossal, 1969). It is only before DNA synthesis begins that the template is susceptible to hydrolysis, and even then it is only unpaired nucleotides that are hydrolyzed. This ensures that the replicated DNA is a faithful copy of the original. Though the template is only

susceptible to hydrolysis before synthesis starts, hydrolysis of newly incorporated nucleotides is continuous throughout the reaction slowing down the process of replication, (Hershfield and Nossal, 1972). To improve processivity of Gp43 and promote its polymerase activity over its exonuclease activity, it is necessary to have its accessory proteins Gp45 and Gp44/62 present in the reaction also.

Much like how the function of the Gp41 helicase is enhanced in the presence of its accessory proteins, Gp43 more efficiently extends the primer in the presence of its accessory proteins, Gp45 and the Gp44/62 complex. Together, these increase the processivity of the polymerase about four-fold, allowing more DNA to be replicated per polymerase loading event, (Mace and Alberts, 1984). Both Gp45 and the Gp44/62 complex must be present in a 1:1 ratio to see the full stimulatory effect, (Mace and Alberts, 1984). However, it is Gp45 that is directly responsible for the increase in processivity of the polymerase, (Reddy et al., 1993). This is because Gp45 functions directly as T4 bacteriophage's sliding clamp while the Gp44/62 complex is the sliding clamp loader (Huang et al., 1980). In other words, Gp45 helps tether the polymerase to the DNA and moves along the template with it to prevent it from disassociating rapidly. Gp45 can move along the template due to its circular structure which places the DNA in the middle of a homotrimeric complex. Encircling of the DNA by Gp45 does not naturally occur and is dependent on conformational changes of the protein mediated by conformational changes in the Gp44/62 complex during ATP hydrolysis, (Jarvis et al. 1989). The Gp44/62 complex ATPase activity is stimulated by the presence of ssDNA, thus Gp45 loading onto the template is sped up when the protein complex encounters a primer-template junction, (Pietroni et al., 1997). This helps place the Gp45 in the correct position on the DNA substrate to effectively aid the polymerase.

Previous research characterizing the function of each protein necessary for the replication component of HDR has allowed us to set up in vitro assays that strive to recreate as accurately as possible the processes that are happening in vivo. These assay systems allow us to test the proposed model that the helicase is the component of the reaction that directs the D-loop down either the dHJ or SDSA pathways of repair. In order to test this, a synthetic D-loop has been constructed for the first time to be used as the substrate in the reactions. In vivo, the D-loop is the last common structure before the repair process is directed down either dHJ or SDSA, so using a synthetic D-loop as the substrate allows the moment of choice to be examined. In this project, Gp41 is the helicase used in the reactions to illustrate the model that it is the helicase that controls pathway choice and the potential for crossover. If the Gp41 unwinds the synthetic D-loop at the template junction in order to extend it (Pathway 2 in Figure 7), it would indicate that the helicase is directing the repair down the dHJ pathway that carries the danger of gene conversion. If Gp41 unwinds the D-loop at the invasive strand to take it off the template (Pathway 1 in Figure 7) it would indicate that it is sending the substrate down the SDSA pathway.

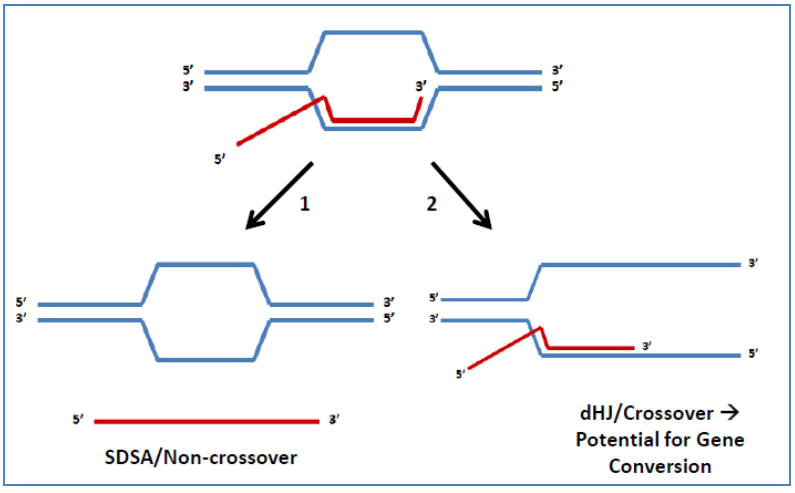

Figure 7 – The two possible ways that Gp41 could unwind the constructed D-loop, indicating that it is sending the repair process down either the dHJ pathway with the risk of gene conversion (2) or the SDSA pathway (1).

Materials and Methods

Substrate Construction

The D-loop was constructed by annealing together three HPLC purified mixed oligonucleotide sequences purchased from Operon Biotechnologies, Inc. Oligo A and Oligo B (Table 1) formed the bubble part of the D-loop and were annealed to each other first before the invasive strand (Oligo C) was added, (Figure 8). The three oligonucleotides were complementary to each other at the places shown in Figure 9.

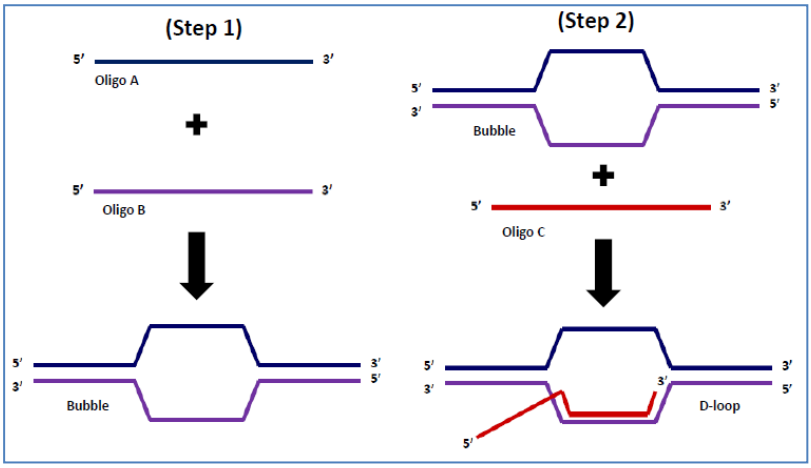

Figure 8 – Two step construction of the D-loop where the Bubble is annealed first and purified to rid the solution of any excess ssDNA before Oligo C is added to the reaction.

Oligos A and B were mixed in a 1:1 ratio in a buffer containing 20 mM Tris HCl (pH 7.5) and 50 mM NaCl. The solution was heated to 95°C for three minutes, held at 68°C for one hour, and then allowed to cool to room temperature. In order to eliminate any excess Oligo A or Oligo B that did not anneal to each other, the products of the annealing reaction were separated on a 2% agarose gel and the band corresponding to the annealed bubble cut out. The DNA was purified out of the agarose piece by using the QIAquick® Gel Extraction Kit from Quiagen® and following the protocol provided. The

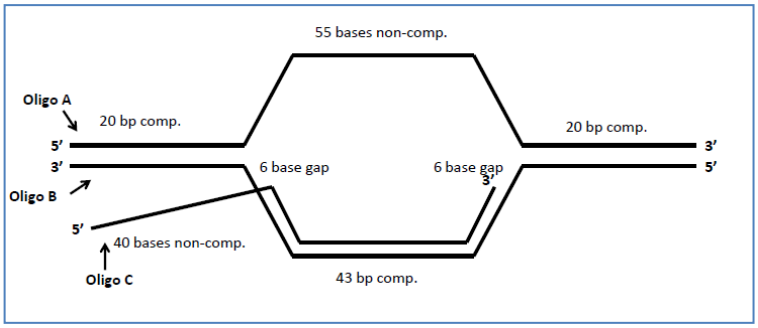

Figure 9 – Regions of complementary base pairs between Oligo A (top of the bubble) and Oligo B (bottom of the bubble) as well as between Oligo B and Oligo C (invasive strand) in the construction of the synthetic D-loop.

concentration of the final purified product was determined by measuring the absorbance at 260 nm and using a conversion factor of 41.5 μg/ml/A_{260}. By purifying the bubble before adding Oligo C to the reaction, it guaranteed that Oligo C could only anneal to the bubble instead of excess Oligo B, creating undesired products.

Oligo C was annealed to the bubble in a second step. It was mixed in a 1:1 ratio with the previously annealed bubble in a buffer containing 20 mM Tris HCl (pH 7.5) and 75 mM NaCl and incubated at 37°C for 20-24 hours.

Oligo Name	Oligo Sequence (5' → 3')	Function
Oligo A	GCCGTCGCAGGGCCATGCGCGTAGATTAGCCATCGTTACGCGCATCGAATATCGAGTACTCGTATAGG CTTCGATCACGGAGCGCTGGCCGCAGC	Top of Bubble
Oligo B	CTGCGGCCAGCGCTCCGTGGCTTCCATTCAGGTCGAGGTGGCCCGGCTCCATGCACCGCGACGCAACG CGGGGAGCGCATGGCCCTGCGACGGC	Bottom of Bubble
Oligo C	GCATAGCTAGCATAGCTACTGCATGCGATGTACTGAGCATCGTTGCGTCGCGGTGCATGGAGCCGGGC CACCTCGACCTGAAT	Invasive Strand
Oligo D	GCTGCGGCCAGCGCTCCGTG	Radiolabeled Backtrap
Oligo E	CGTTGCGTCGCGGTGCATGGAGCCGGGCCACCTCGACCTGAAT	Unlabeled Trap

Table 1- The sequences of each oligonucleotide used in the project.

Some assays required that the constructed D-loop be radiolabled. In these cases before Oligo C was added to the bubble, it was first 5'-end-labeled with γ-^{32}P ATP (Perkin Elmer) using T4 polynucleotide kinase. After the radiolabeled Oligo C was annealed to the bubble, the reaction was run on a 6% acrylamide gel in order to separate any unincorporated radiolabeled Oligo C from the complete D-loop substrate. The presence of any unincorporated radiolabeled Oligo C would make it hard to tell whether the invasive strand was coming off in the unwinding reactions and thus had to be eliminated. The band corresponding to the D-loop was cut out of the gel and electroeluted, with the substrate being trapped in a 7.5 M NH_4OAc high salt solution. The solution containing the DNA was extracted from the electroelutor and placed in an Amicon® Centrifugal Filter Unit in order to desalt the solution and concentrate the DNA. The filter cartridge was spun at 14,000 x g for 30 minutes. Deionized water was added to rinse the DNA

and get rid of any remaining salt, and the cartridge was spun for another 30 minutes at 14,000 x g. The concentrated DNA was spun out of the filter at 1,000 x g for two minutes and diluted with more deionized water. In order to ascertain the final concentration of the radiolabeled D-loop and to test for proper annealing, the purified substrate was run with a radiolabeled marker of known concentration on a 6% acrylamide gel. The gel was then exposed to a Kodak screen for 30-45 minutes and the screen scanned by a Bio-Rad phosphorimager in order to visualize the product bands. The image was viewed and the relative volumes of the products compared using Quantity One software.

T4 Bacteriophage Proteins

All proteins used in these experiments had previously been purified according to published protocols. Gp41, Gp32, and Gp59 had been purified and stored in buffers as described in Branagan et al. (2012). Gp45 and Gp44/62 were purified and stored as described by Morris et al, (1979).

The Gp43 protein used in all experiments was a mutant constructed to be deficient in exonuclease function while still capable of acting as a polymerase, (Frey et al., 1993). This mutant, Gp43 D219A was a gift from Linda Reha-Krantz.

Helicase Unwinding Assays

All unwinding reactions were carried out in a buffer containing 25 mM Tris acetate (pH 7.8), 60 mM potassium acetate, 6 mM magnesium acetate, 2 mM dithiothreitol, 200 µg/ml bovine serum albumin and 2 mM ATP. When present, Gp41, Gp59, Gp32, Gp43, Gp45, and Gp44/62 were added to the reaction for a final concentration of 600 nM, 10 nM, 600 nM, 100 nM, 250 nM, and 250 nM respectively, unless otherwise specified. When the proteins were not added to the reaction, the equivalent volume of their storage buffer was added to keep the salt

components of the reactions constant. The D-loop was present in a final concentration of 5 nM in the case of the radiolabeled substrate and 100 nM in the case of the unlabeled substrate. Reactions were carried out at room temperature, approximately 20°C.

All of the components of the reaction, except Gp41 and the DNA trap which was present at fourfold the concentration of the substrate, were added to the reaction for a final volume of 15 µl and the reaction was allowed to incubate for five minutes. The DNA trap consisted of radiolabeled Oligo D for the reactions involving the unlabeled D-loop and Oligo E for the reactions involving the radiolabeled D-loop. After the five minute incubation, the reactions were initiated by adding the premixed Gp41 and trap and allowing the reaction to proceed for twenty minutes. The reactions were stopped by the addition of 15 µl of a stopping buffer composed of 0.5 M EDTA, 1% SDS, and 2x Blue/Orange gel loading dye (Promega).

The products were separated by electrophoresis on a 6% acrylamide gel (8.5 cm x 7.5 cm) for approximately 2.5 hours. The gel was then exposed to a Kodak screen for 30-45 minutes and the screen scanned by a Bio-Rad phosphorimager in order to visualize the product bands. The images were viewed and the relative volumes of the products compared using Quantity One software.

Primer Extension Assays

All unwinding reactions were carried out in a buffer containing 25 mM Tris acetate (pH 7.8), 125 mM potassium acetate, 10 mM magnesium acetate, 2 mM dithiothreitol, 200 µg/ml bovine serum albumin and 2 mM ATP. When present, Gp41, Gp59, Gp32, Gp43, Gp45, and Gp44/62 were added to the mixture for a final concentration of 600 nM, 10 nM, 600 nM, 100 nM, 250 nM, and 250 nM respectively. When the proteins were not added to the reaction, the equivalent volume of their storage buffer was added to keep the salt components of the reactions

constant. The radiolabled D-loop was present in a final concentration of 5 nM. Each dNTP was added separately for a final concentration of 50 µM each. Reactions were carried out at room temperature, approximately 20°C.

All of the components of the reaction, except Gp41 and dGTP, were mixed together for a final volume of 30 µl and allowed to incubate for five minutes. The reactions were started by adding the premixed Gp41 and dGTP to the reaction and allowing it to proceed for 20 minutes. The reactions were quickly stopped by the addition of 11 µl of 50 mM EDTA. Twenty µg of Proteinase K was added to the reactions and they were incubated at 37°C for thirty minutes to cleave the protein components of the reaction, guaranteeing that the reaction had stopped. After incubation, Oligo E was added as a fourfold excess trap and the DNA was denatured by placing the reactions in boiling water for five minutes. The samples were cooled to room temperature and could be stored at -20°C until run on a gel.

Before loading onto the gel, the samples were boiled again for five minutes then immediately loaded onto a 12% acrylamide, 6 M Urea denaturing gel (31 cm x 38.5 cm) and run at 1500 volts for about five hours. The gel was then exposed to a Kodak screen for 30-45 minutes and the screen scanned by a Bio-Rad phosphorimager in order to visualize the product bands. The images were viewed and edited in Quantity One software.

Results

D-loop Construction

Oligos A, B, and C were successfully annealed together to form the first constructed three stranded D-loop in the two steps diagrammed in Figure 8. As seen in Figure 10, in the lane labeled Bubble + Oligo C (1:1:1), a higher molecular weight band appears compared to the lane with only the bubble (Oligo A and B) present. This would indicate that Oligo C has indeed

incorporated into the bubble during the second step of construction. Figure 10 also shows the results of an earlier construction scheme that was not pursued where Oligo E was used as the invasive strand. Not all of Oligo C was incorporated as can be seen, and a tenfold excess as seen in the lane labeled Bubble + Oligo C (10:1:1) worsened this problem while not resulting in significantly more D-loop produced. Because of this excess, there needed to be a purification scheme for the radiolabeled D-loop in order to eliminate the excess ssDNA so its unwinding could clearly be seen. The result of this purification scheme can be seen in Figure 11. On the right is the purified radiolabeled D-loop. Only the higher molecular weight band corresponding to the D-loop is visible, the lower molecular weight band that would correspond to the radiolabeled ssDNA, as seen on the left is not present.

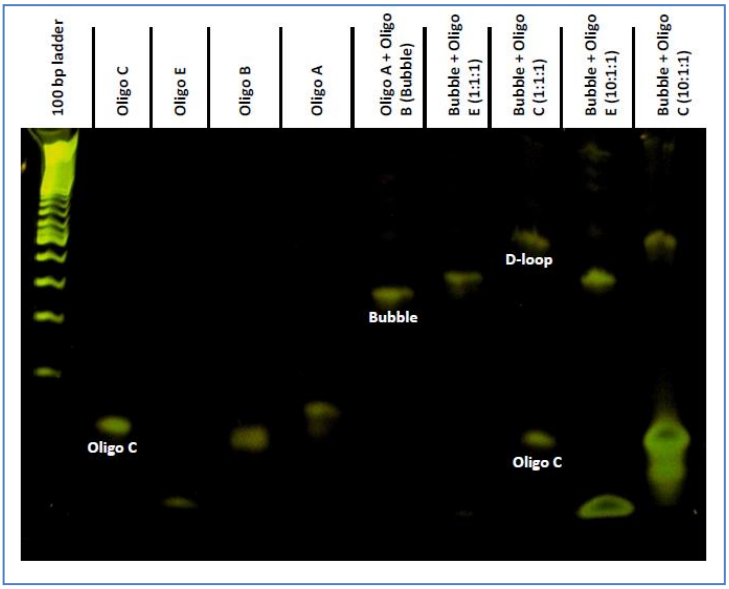

Figure 10 – Gel showing the different stages of construction of the D-loop, which can be seen appearing as the upper band in the lane labeled Bubble + Oligo C (1:1:1). 6% acrylamide in 0.5x TBE. Stained with SYBR® Gold from Invitrogen™.

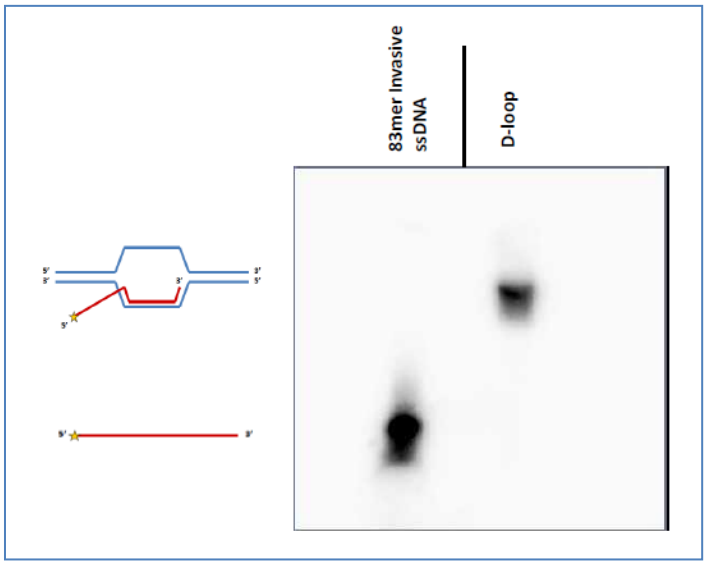

Figure 11 – The radiolabeled D-loop after the purification to remove excess radiolabeled invasive strand (Oligo C). 6% acrylamide in 0.5x TBE.

Helicase Unwinding Assays

For the helicase unwinding assays that used the D-loop with the radiolabeled invasive ssDNA, no lower molecular weight band corresponding to Oligo C alone appeared upon addition of Gp41 and other accessory proteins as seen in Figure 12. This would indicate that Gp41 did not unwind the D-loop in a manner that took the invasive ssDNA off as seen in Figure 13. It does look like there is some degradation of the D-loop that can be seen even in the control D-loop lane indicating that it is due to the conditions of the reaction, not unwinding by the helicase.

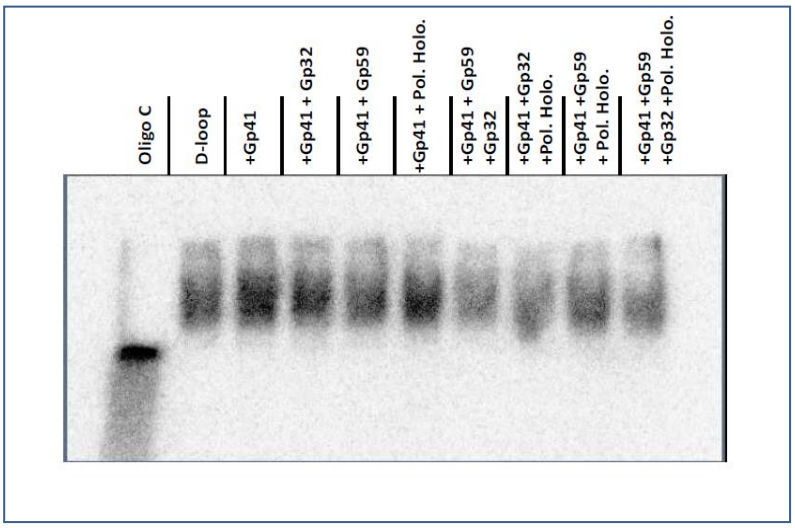

Figure 12 – No lower molecular weight band corresponding to the band in the lane labeled Oligo C appear, indicating the helicase is not taking the invasive strand off. 6% acrylamide gel in 0.5x TBE.

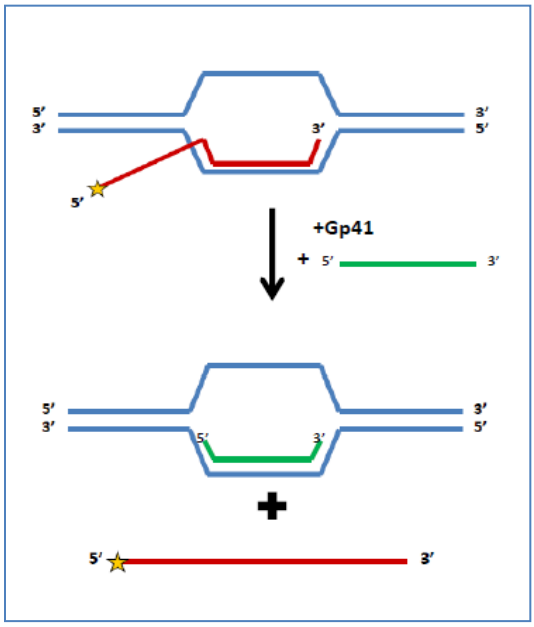

Figure 13 – Products of the reaction after initiation with Gp41 and trap if the helicase unwound the D-loop at the invasive strand.

For the helicase unwinding assays that used the unlabeled D-loop with the labeled Oligo D as a back trap to visualize the opening up of the D-loop, (Figure 14) a higher molecular weight band can be seen upon addition of the helicase (all lanes except Oligo D and D-loop) in Figure 15. This faint band

can only be produced when the back of the D-loop has been unwound, giving the radiolabeled Oligo D a chance to anneal to the higher molecular weight structure and emit a signal.

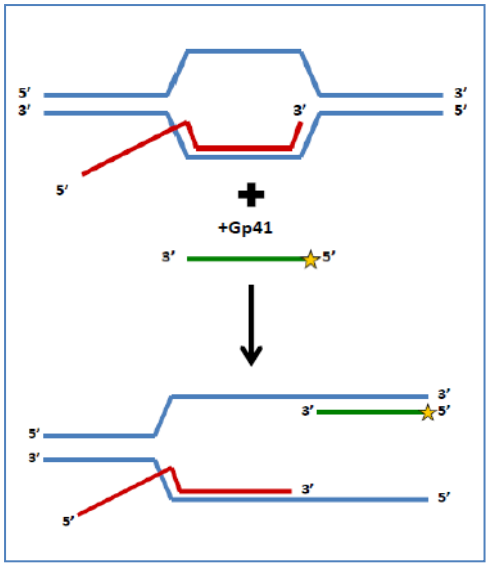

Figure 14 – Products of the reaction after initiation with Gp41 and trap if the helicase unwound the D-loop at the back.

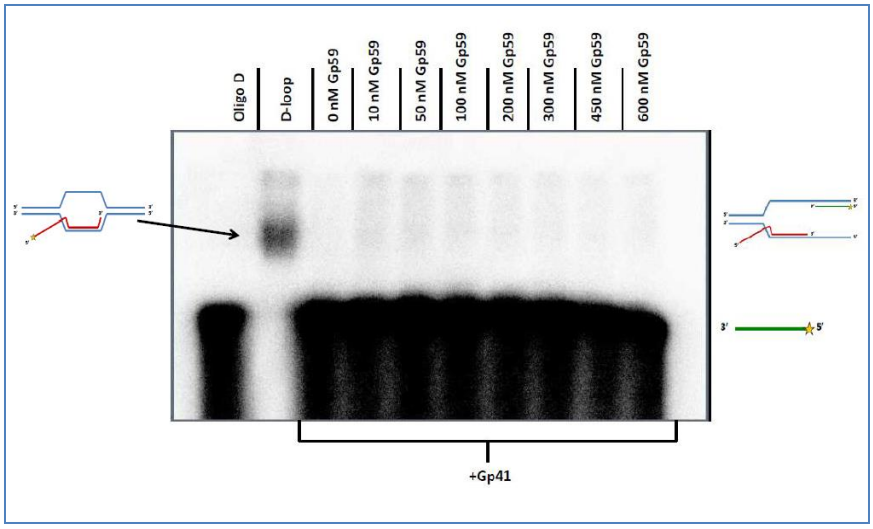

Figure 15 – Higher molecular weight bands appear in the presence of Gp41 and varying concentrations of Gp59 indicating that the helicase is unwinding the D-loop to allow radiolabled Oligo D to anneal. 6% native acrylamide gel in 0.5x TBE

As mentioned, along with Gp41, the reactions in Figure 15 all contained varying concentrations of Gp59, ranging from 0 nM to 600 nM. As can be seen in Figure 15, in the presence of low concentrations of Gp59, more high molecular weight products were produced than in the lane with Gp41 alone. This is indicated by the relative darkness, or strength of the signal, of the bands produced.

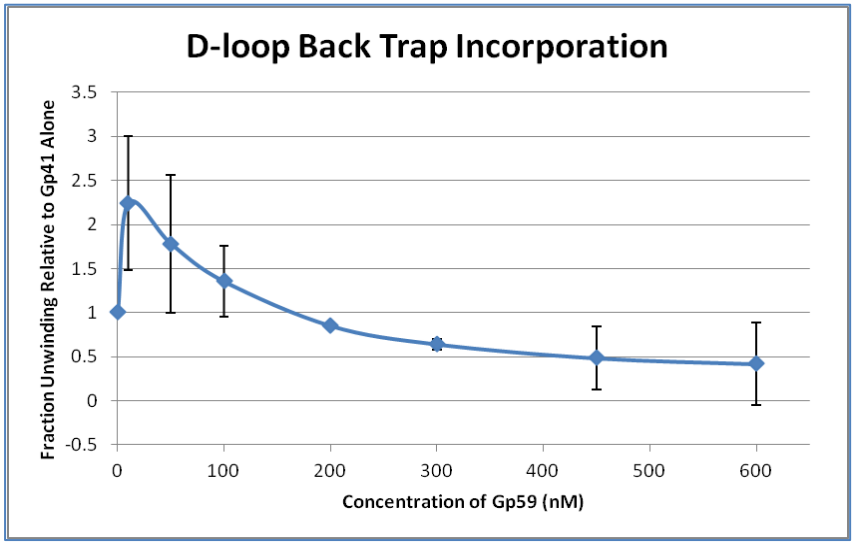

Figure 16 – Effect the presence of differing concentrations of Gp59 has on the unwinding ability of Gp41. An enhancing trend can be seen up to 100 nM Gp59 then the presence of Gp59 inhibits the reaction. The average of two trials has been graphed along with the standard deviation.

This would indicate that low concentrations of Gp59 had an

enhancing effect on Gp41 unwinding ability. However, at higher concentrations, the presence of Gp59 switches from enhancing to inhibiting the reaction. This differing effect can be seen more clearly in Figure 16. In Figure 16 the amount of unwound products, as measured by volume analysis of the complete higher molecular weight band in Quantity One, produced by Gp41 alone has been set at a baseline of 1.The effect of the different concentrations of Gp59 then has been made a fraction of the amount unwound by Gp41 alone. Two trials were averaged together and the average along with the standard deviation is shown. In the presence of 10 nM Gp59, an average of 2.25x more substrate is unwound than in the presence of Gp41 alone. This enhancing effect over only Gp41 continues, but not as significantly in the presence of 50 nM Gp59 and 100 nM Gp59. By 200 nM Gp59, the enhancing effect has been lost, and the presence of Gp59 actually starts to inhibit the reaction. The inhibition continues for the remaining concentrations of Gp59 as less and less unwound product is formed.

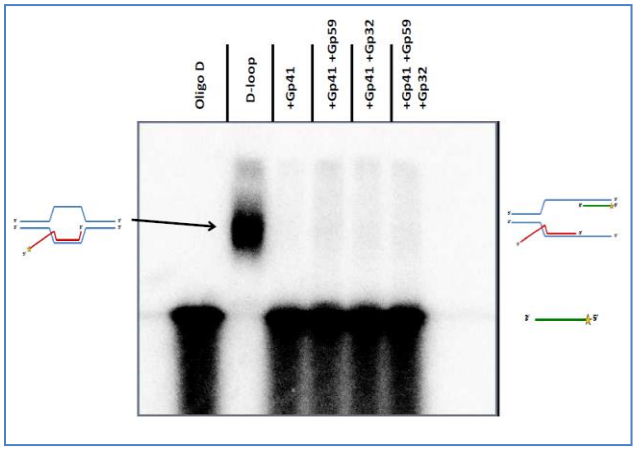

Figure 17 – Higher molecular weight bands appear in the presence of Gp41 ±Gp59 and Gp32 indicating that the helicase is unwinding the D-loop to allow radiolabled Oligo D to anneal.

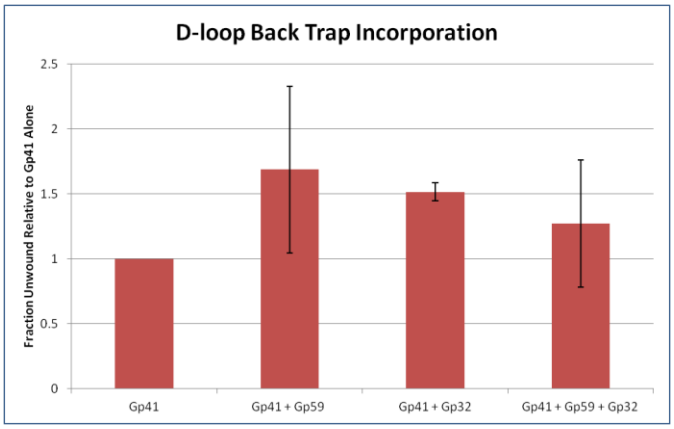

Figure 18 – Effect Gp59 and/or Gp32 has on the efficiency of the D-loop unwinding by Gp41. The average of two trials is graphed and the standard deviations are shown.

The effect of the presence of the other Gp41 accessory protein, Gp32 in the unwinding reaction was also tested. As seen in Figure 17, once again the higher molecular weight band

appears in the presence of the helicase and accessory proteins indicating that the D-loop is being unwound and the back trap is incorporating in. Like Figure 16, in Figure 18 the amount of unwinding of the D-loop by Gp41 has been normalized to a value of 1 and the amount of the substrate unwound in the presence of the other accessory proteins has been made a fraction of this value. In Figure 18 it can be seen that the presence of 10 nM Gp59 enhances the efficiency of the reaction since about 1.7x more D-loop is unwound in the presence of Gp59 than in its absence. Gp32 also plays an enhancing role though not to the extent that Gp59 does, as about 1.5x more D-loop is unwound. Finally, in the presence of both Gp32 and Gp59, there is still an enhancing effect though not as strong as with either of the accessory proteins alone, as about 1.3x more D-loop is unwound.

Primer Extension

In the presence of the polymerase holoenzyme, Gp43, Gp45, and Gp44/62, the invasive strand in the D-loop is extended the entire way to the end of the template producing a 109 base oligonucleotide as modeled in Figure 19.The fully extended 109mer appears as the higher molecular weight band in Figure 20. Not all of the invasive strand was extended as there is still a band present that corresponds to the original 83mer oligonucleotide length present in all lanes. Additionally there are faint bands of intermediate lengths present indicating that the invasive

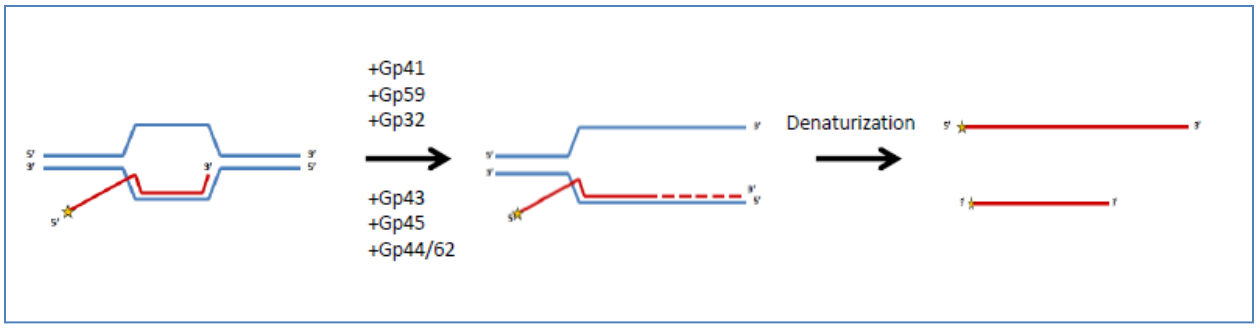

Figure 19 – Primer extension of the invasive strand of the D-loop. Oligo C was extended to a 109mer length using the D-loop as a template in the presence of the T4 Bacteriophage polymerase holoenzyme.

strand was not able to be fully extended before the reaction was interrupted for some reason. There does not seem to be an enhancing effect on the extension reaction generated by the addition of the Gp41 helicase and its accessory proteins.

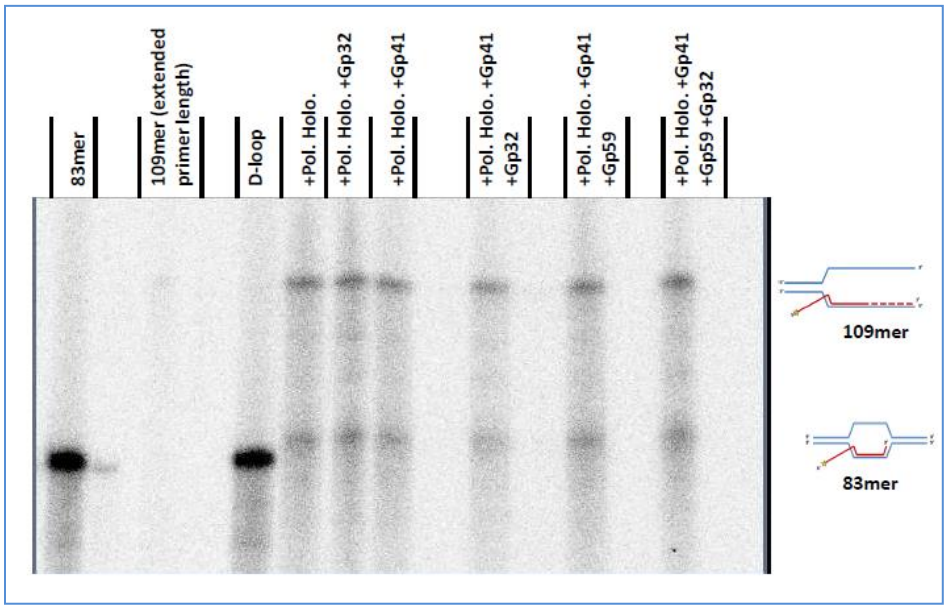

Figure 20 – Primer extension of the invasive strand of the D-loop. Oligo C was extended to a 109mer length using the D-loop as a template in the presence of the T4 Bacteriophage polymerase holoenzyme. 12% acrylamide, 6 M Urea denaturing gel in 1x TBE. Since the gel is a denaturing gel, the bands in the gel are equivalent to only the red radiolabeled strands in the corresponding DNA diagrams.

Discussion

The construction of the D-loop in this project marked the first time a whole synthetic D-loop, instead of simplified representations of it, had been made. The complete presence of all three strands involved in a natural D-loop makes this structure more physiologically relevant to perform assays on than substrates previously used as models. The synthetic D-loop directly corresponds to the structure present in the moment of repair where the cell has to make a choice between the two main mechanisms of HDR: the double Holliday Junction model and the Synthesis-dependent strand annealing model, (Maher et al., 2011). Therefore the construction of this new substrate allowed the development of a set of assays to illustrate the main hypothesis of the proposed model: that the helicase is the component of the reaction that directs the repair process down one pathway or the other and thus is in charge of crossover control. The unwinding

and extension assays performed using this substrate will better illustrate the biochemical mechanism involved in vivo HDR since the D-loop is a structure that appears naturally in the repair mechanism.

The results from both sets of helicase unwinding assays indicate that Gp41, one of the three helicases of the T4 Bacteriophage, directs the repair process down the dHJ pathway of repair and not SDSA. In the first set of reactions, the helicase did not take the invasive strand (Oligo C) off the D-loop. In the SDSA process of repair, the newly synthesized DNA is removed from the template as it is made, allowing the bubble to migrate along the template without expanding. The lack of Oligo C removal is equivalent to the invasive strand and newly synthesized DNA not being removed from the template, which it needs to be in SDSA. This can be interpreted as Gp41 not directing the repair process down the SDSA pathway. On the other hand, in the second set of reactions, Oligo D was incorporated into the D-loop. This would only be possible if the back of the D-loop was unwound by Gp41 allowing Oligo D to anneal. This corresponds to the dHJ model of repair in which the helicase extends the D-loop. This correspondence to the dHJ model of repair and variance with the SDSA model of repair both indicates that Gp41 directs the repair process down the dHJ model of repair (Figure 21).

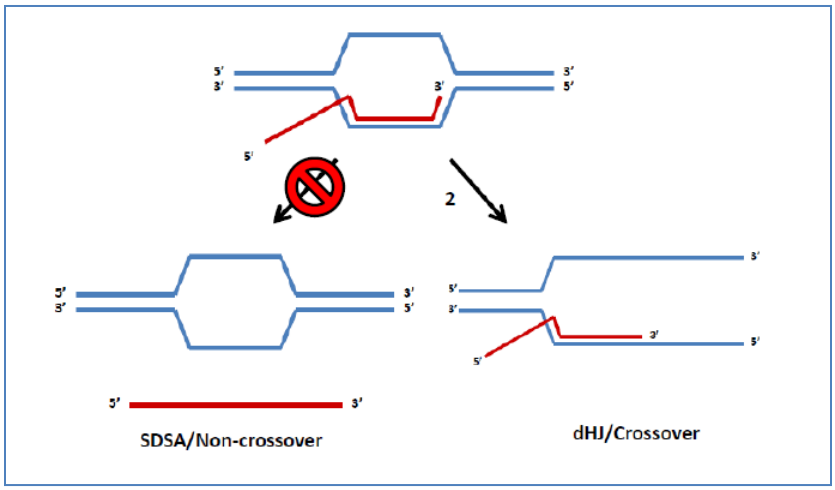

Figure 21 – Gp41 directs the substrate down the dHJ pathway of repair.

The difference in efficiency of back trap incorporation in the presence and absence of the different helicase accessory proteins indicate that protein-protein interactions are just as important as protein-DNA reactions in unwinding the substrate. Lower concentrations of Gp59 enhance the unwinding reaction perhaps due to the fact that they preferentially bind to fork DNA structures, (Jones et al., 2000). At low concentrations they will mostly be present at the opening of the D-loop, loading the helicase right where it needs to be to extend the D-loop. At higher concentrations of Gp59, it no longer just binds to fork structures, but also to less preferred structures such as ssDNA and dsDNA. Gp59 can no longer load the helicase solely at the fork where it needs to be to unwind the substrate, but loads it at less relevant locations due to the amount of protein present. Eventually at extremely high concentrations, the helicase loading protein no longer has an enhancing effect on the reaction, but rather an inhibiting one. It could be that the DNA is so saturated with Gp59 that the protein can no longer be exchanged out for the helicase. It is also possible that the Gp59 is binding to Gp41 in the solution and sequestering it away from interacting with the DNA. This serves to illustrate that the stoichiometry of the proteins present in the reaction are crucial to allow it to proceed efficiently.

On the same note, the presence or absence of Gp32 is also important in determining the efficiency of the reaction. Like Gp59, when Gp32 is added into the reaction with Gp41, there is a stimulatory effect on the reaction, although not as strong as Gp59. This may be due to the fact that Gp32 stabilizes the ssDNA of the D-loop, creating a more stable structure for Gp41 to bind to. It also coats the ssDNA newly created by the unwinding, preventing the D-loop from annealing back together. Gp59 and Gp32 added to the reaction together still have an enhancing effect, though it is closer to the enhancing effect of Gp32 alone rather than Gp59. Again the stoichiometry of these proteins is crucial as Gp59 will preferentially bind to Gp32 over ssDNA.

Therefore, the enhancing effect closer to Gp32 alone could be the result of the excess Gp32 in the solution sequestering Gp59 away from the DNA, much like the high concentration of Gp59 might be sequestering the Gp41 away from the substrate. More work will have to be done to optimize the concentrations of each of these three proteins as well as the other protein components present in the in vivo system in order to most precisely replicate biochemically what is happening naturally as well as to create the most efficient unwinding reaction.

No matter the presence or absence of the helicase accessory proteins, the unwinding of the back of the D-loop was not an efficient process. One cause of this inefficiency may be the small size of the D-loop constructed. As mentioned in the introduction, Gp41 is a large hexameric protein. It may be that it is simply too large compared to the size of the D-loop to efficiently load onto it and unwind it. On the other hand, it could also be the case that Gp41 is able to load onto and unwind the D-loop without a problem, but the back of the D-loop anneals back to itself again before the trap can get in and anneal instead, thus distorting our perception on the amount of D-loop actually unwound. The D-loop is already annealed to each other at one end so the strands are very close to each other even when they have been unwound apart, making it very easy for the unwound end to anneal back together. This problem could be mitigated by adding in more back trap, but because the trap is radiolabeled to allow for visualization of the incorporation, the assay is constrained in the amount that can be added to the assay. It may be that a combination of these two factors, difficulty loading and fast reannealing, or possibly other factors that cause the low efficiency of the unwinding reaction.

There is more to HDR than just the unwinding of the D-loop. In order for the DSB to be fully repaired, new DNA has to be synthesized using the homologous DNA as a template. The primer extension reactions model this component of the repair process. Like the unwinding

reactions, the primer extensions assay can also be used to differentiate between the two repair pathways that the helicase could send the reaction down. If the helicase only unwinds the back of the D-loop to extend it like it does in dHJ, the invasive strand will be able to be extended by the polymerase the entire way to the end of the template. The appearance of a band corresponding to this higher molecular weight product would therefore indicate that the helicase directs the substrate down the dHJ repair pathway. However if the helicase takes off the invasive strand before it could be fully extended, a ladder of different DNA lengths would appear. This result would be indicative of the SDSA repair pathway where the helicase takes the newly made DNA off the template. What was seen in the reaction was a mix between these two possible results. As seen in Figure 14, the majority of the product produced by the reaction was the 109 base length fully extended product. This would be indicative of the dHJ model of repair, with the helicase opening the D-loop allowing the invasive strand to be fully extended. However, there are also faint intermediate length products between the full 109mer and the original 83mer. These intermediate lengths would be indicative of SDSA with the helicase taking the invasive strand off of the DNA. It is interesting to note that instead of a ladder of all extended lengths possible, there are only two that appear. This might indicate that there is another factor at work in the reaction to produce these intermediate lengths rather than the helicase taking the invasive strand off before it is finished extending. One possibility is that one of the lengths corresponds to the 6 base long free ssDNA that appears after the template but before the dsDNA part of the D-loop. The template could very easily be extended to this point then stop when it hits the dsDNA portion and this scenario might account for one of the intermediate length products.

It is also interesting to note in the extension reactions that the presence of the Gp41 helicase and its accessory proteins do not seem to enhance the activity of the polymerase. Alone,

the polymerase holoenzyme extends just as much of the invasive strand primer as it does in the presence of the helicase. It would be expected that the presence of the helicase would stimulate the reaction since it opens up the dsDNA region ahead of the polymerase, giving it easier access to the template. It may be that this enhancing effect is not seen due to the small size of the D-loop once again. There is only a 20 base region of dsDNA and it may be that this is not a long enough stretch to hinder the action of the polymerase. Constructing a D-loop with a longer duplex region after the bubble may pose enough of a block to the polymerase that the presence of the helicase becomes necessary for the extension of the primer to precede.

Along with modifying the stoichiometry of the proteins involved in the reaction and constructing a longer D-loop, it would be interesting in the future to examine the ability of the other two T4 Bacteriophage helicases, Dda, and UvsW to unwind the constructed D-loop. In accordance with the proposed model, how the helicases unwind the D-loop would allow them to be classified as either promoting SDSA or dHJ or perhaps a mixture of the two. Furthermore, as the newly constructed D-loop is a physiologically relevant structure in all HDR pathways, it will be a useful substrate in clarifying the mechanism of action of many proteins that play a role in repairing DSBs. This includes using it as a substrate to test whether different human helicases direct the repair process down the dHJ model of repair or SDSA. So far all the human helicases studied have been shown to direct the repair process down the SDSA pathway, (Colavito et al., 2010). As mentioned, it is the dHJ pathway rather than the SDSA pathway that carries the potential for gene conversion. Thus it is important to classify which human helicases direct HDR down this pathway in order to understand the mechanism in which protective dominant alleles are lost leaving only disease causing recessive alleles in the cell. The loss of the dominant allele of tumor suppressors such as p53 (Hollstein et al., 1991) or retinoblastoma protein (Rb) (Sherr,

1996) causes the cell to lose control of its own proliferation and become cancerous. Understanding the gene conversion process that results in this loss of control is crucial for a better understanding and treatment of cancer.

Works Cited

Alberts, B.M. (1987). Prokaryotic DNA replication mechanisms. *Philosophical Transaction of the Royal Society of London* **317(1187)**, 395-420.

Beernink, H.T.H. and Morrical, S.W. (1998) The uvsY Recombination Protein of Bacteriophage T4 Forms Hexamers in the Presence and Absence of Single-Stranded DNA. *Biochemistry* **37**, 5673-5681.

Bleuit, J.S., Xu, H., Ma, Y., Wang, T, Liu, J, and Morrical, S.W. (2001). Mediator proteins orchestrate enzyme-ssDNA assembly during T4 recombination-dependent DNA replication and repair. *PNAS* **98(15)**, 8298-8305.

Branagan, A.M., Maher, R.L., Morrical, S.W. (2012). Assembly and Dynamics of Gp59-Gp32-Single-stranded DNA (ssDNA), a DNA Helicase Loading Complex Required for Recombination-dependent Replication in Bacteriophage T4. *Journal of Biological Chemistry* **287(23)**, 19070-19081.

Colavito, S., Prakash, R., and Sung, P. (2010). Promotion and regulation of homologous recombination by DNA helicases. *Methods* **51**, 329-335.

Dong, F., Gogoal, E.P., and von Hippel, P.H. (1995). The Phage T4-coded DNA Replication Helicase (gp41) Forms a Hexamer upon Activation by Nucleoside Triphosphate. *Journal of Biological Chemistry* **270(13)**, 7462-7473.

Formosa, T., and Alberts, B.M. (1986). DNA Synthesis Dependent on Genetic Recombination: Characterization of a Reaction Catalyzed by Purified Bacteriophage T4 Proteins. *Cell*

47, 793-806.

Frey, M.W., Nossal, N.G., Capson, T.L., and Benkovic, S.J. (1993). Construction and characterization of a bacteriophage T4 DNA polymerase deficient in 3' →5' exonuclease activity. *PNAS* **90**, 2579-2583.

Goulian, M., Lucas, Z.J., and Kornberg, A. (1968). Enzymatic Synthesis of Deoxyribonucleic Acid: Purification and Properties of Deoxyribonucleic Acid Polymerase Induced by Infection with Phage T4. *Journal of Biological Chemistry* **243(3)**, 627-638.

Hershfield, M.S., and Nossal, N.G. (1972). Hydrolysis of Template and Newly Synthesized Deoxyribonucleic Acid by the 3' to 5' Exonuclease Activity of the T4 Deoxyribonucleic Acid Polymerase. *Journal of Biological Chemistry* **247(11)**, 3393-3404.

Hollstein, M., Sidransky, D., Vogelstein, B., and Harris, C.C. (1991). P53 Mutations in Human Cancers. *Science* **253**, 49-53.

Huang, C.C., Hearst, J.E., and Alberts, B.M. (1981). Two Types of Replication Proteins Increase the Rate at Which T4 DNA Polymerase Traverses the Helical Regions in a Single-stranded DNA Template. *Journal of Biological Chemistry* **256(8)**, 4087-4094.

Jackson, S.P. and Bartek, J. (2009). The DNA-damage response in human biology and disease. *Nature* **461**, 1071-1078.

Jarvis, T.C., Paul, L.S., Hockensmith, J.W., and von Hippel, P.H. (1989). Structural and Enzymatic Studies of the T4 DNA Replication System: ATPase Properties of the Polymerase Accessory Protein Complex. *Journal of Biological Chemistry* **264(21)**, 12717-12729.

Jones, C.E., Mueser, T.C., and Nossal, N.G. (2000). Interaction of the Bacteriophage T4 Gene 59 Helicase Loading Protein and Gene 41 Helicase with Each Other and with Fork, Flap,

and Cruciform DNA. *Journal of Biological Chemistry* **275(35)**, 27145-27154.

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

Kodadek, T., Gan, D-C., and Stemke-Hale, K. (1989). The Phage T4 uvsY Recombination Protein Stabilizes Presynaptic Filaments. *Journal of Biological Chemistry* **264(28)**, 16451-16457.

Kodadek, T., Wong, M.L., and Alberts, B.M. (1988). Mechanism of Homologous DNA Strand Exchange Catalyzed by Bacteriophage T4 uvsX and Gene 32 Proteins. *Journal of Biological Chemistry* **263(19)**, 9427-9436.

Kreuzer, K.N. and Brister, J.R. (2010). Initiation of bacteriophage T4 DNA replication and replication fork dynamics: a review in the Virology Journal series on bacteriophage T4 and its relatives. *Virology Journal* **7(358)**.

Lionnet, T., Splering, M.M., Benkovic, S.J., Benslmon, D., and Croquette, V. (2007). Real-time observation of bacteriophage T4 gp41 helicase reveals an unwinding mechanism. *PNAS* **104(50)**, 19790-19795.

Mace, D.C. and Alberts, B.M. (1984). Characterization of the Stimulatory Effect of T4 Gene 45 Protein and the Gene 44/62 Protein Complex on DNA Synthesis by T4 DNA Polymerase. *Journal of Molecular Biology* **177**, 313-327.

Maher R.L., Branagan A.M., and Morrical S.W. (2011). Coordination of DNA Replication and Recombination Activities in the Maintenance of Genome Stability. *Journal of Cellular Biochemistry* **112**, 2672-2682.

Mickelson, C. and Wiberg, J. S. (1981) Membrane-Associated DNase Activity Controlled by

Genes 46 and 47 of Bacteriophage T4D and Elevated DNase Activity Associated with the T4 *das* Mutation. *Journal of Virology* **40(1)**, 65-77.

Morrical, S.W., Hempstead, K., and Morrical, M.D. (1994). The Gene 59 Protein of Bacteriophage T4 Modulates the Intrinsic and Single-stranded DNA-stimulated ATPase Activities of Gene 41 Protein, the T4 Replicative DNA Helicase. *Journal of Biological Chemistry* **269(52)**, 33069-33081.

Morris, C.F., Hama-Inaba, H., Mace, D., Sinha, N.K., and Alberts, B. (1979). Purification of the Gene 43, 44, 45, and 62 Proteins of the Bacteriophage T4 DNA Replication Apparatus. *Journal of Biological Chemistry* **254(14)**, 6787-6796.

Nossal, N.G. (1969). A T4 Bacteriophage Mutant Which Lacks Deoxyribonucleic Acid Polymerase But Retains the Polymerase-associated Nuclease. *Journal of Biological Chemistry* **244(1)**, 218-222.

Pietroni, P., Young, M.C., Latham, G.J., and von Hippel, P.H. (1997). Structural Analyses of gp45 Sliding Clamp Interactions during Assembly of the Bacteriophage T4 DNA Polymerase Holoenzyme: Conformation Changes Within the gp44/62-gp45-ATP Complex During Clamp Loading. *Journal of Biological Chemistry* **272(50)**, 31666-31676.

Reddy, M.K., Weitzel, S.E., and von Hippel, P.H. (1993). Assembly of a functional replication complex without ATP hydrolysis: A direct interaction of bacteriophage T4 gp45 with T4 DNA polymerase. *PNAS* **90**, 3211-3215.

Salinas, F. and Kodadek, T. (1995). Phage T4 homologous stand exchange: A DNA helicase, not the strand transferase, drives polar branch migration. *Cell* **82(1)**, 111-119.

Sherr, C.J. (1996). Cancer Cell Cycles. *Science* **274**, 1672-1677.

Sweezy, M.A. and Morrical, S.W. (1999). Biochemical Interactions within a Ternary Complex of the Bacteriophage T4 Recombination Proteins uvsY and gp32 Bound to Single-Stranded DNA. *Biochemistry* **38**, 936-944.

Szostak, J.W., Orr-Weaver, T.L., Rothstein, R.J., and Stahl, F.W. (1983). The Double-Strand-Break Repair Model for Recombination. *Cell* **33**, 25-35.

Venkatesan, M., Silver, L.L., and Nossal, N.G. (1982). Bacteriophage T4 Gene 41 Protein, Required for the Synthesis of RNA Primers, Is Also a DNA Helicase. *Journal of Biological Chemistry* **257(20)**, 12426-12434.

Yonesaki, T. and Minagawa, T. (1989). Synergistic Action of Three Recombination Gene Products of Bacteriophage T4, uvsX, uvsY, and Gene 32 Proteins. *Journal of Biological Chemistry* **264(14)**, 7814-7820.